Spatio-temporal transferability of environmentally-dependent population models: Insights from the intrinsic predictabilities of Adélie penguin abundance time series

Bilgecan Şen [a,b,*,1], Christian Che-Castaldo [a,1], Kristen M. Krumhardt [c], Laura Landrum [c], Marika M. Holland [c], Michelle A. LaRue [d,e], Matthew C. Long [c], Stéphanie Jenouvrier [f,2], Heather J. Lynch [a,2]

[a] Institute for Advanced Computational Science, Stony Brook University, Stony Brook, NY, United States
[b] Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY, USA
[c] Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, United States
[d] School of Earth and Environment, University of Canterbury, Christchurch, New Zealand
[e] Department of Earth and Environmental Sciences, University of Minnesota, Minneapolis, MN, United States
[f] Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, United States

ARTICLE INFO

ABSTRACT

Ecological predictions are necessary for testing whether processes hypothesized to regulate species population dynamics are generalizable across time and space. In order to demonstrate generalizability, model predictions should be transferable in one or more dimensions, where transferability is the successful prediction of responses outside of the model data bounds. While much is known as to what makes spatially-oriented models transferable, there is no general consensus as to the spatio-temporal transferability of ecological time series models. Here, we examine whether the intrinsic predictability of a time series, as measured by its complexity, could limit such transferability using an exceptional long-term dataset of Adélie penguin breeding abundance time series collected at 24 colonies around Antarctica. For each colony, we select a suite of environmental variables from the Community Earth System Model, version 2 to predict population growth rates, before assessing how well these environmentally-dependent population models transfer temporally and how reliably temporal signals replicate through space. We show that weighted permutation entropy (WPE), a model-free measure of intrinsic predictability recently introduced to ecology, varies spatially across Adélie penguin populations, perhaps in response to stochastic environmental events. We demonstrate that WPE can strongly limit temporal predictive performance, although this relationship could be weakened if intrinsic predictability is not constant over time. Lastly, we show that WPE can also limit spatial forecast horizon, which we define as the decay in spatial predictive performance with respect to the physical distance between focal colony and predicted colony. Irrespective of intrinsic predictability, spatial forecast horizons for all Adélie penguin breeding colonies included in this study are surprisingly short and our population models often have similar temporal and spatial predictive performance compared to null models based on long-term average growth rates. For cases where time series are complex, as measured by WPE, and the transferability of biologically-motivated mechanistic models are poor, we advise that null models should instead be used for prediction. These models are likely better at capturing more generalizable relationships between average growth rates and long-term environmental conditions. Lastly, we recommend that WPE can provide valuable insights when evaluating model performance, designing sampling or monitoring programs, or assessing the appropriateness of preexisting datasets for making conservation management decisions in response to environmental change.

* Corresponding author.
  E-mail address: bilgecan.sen@gmail.com (B. Şen).
[1] These authors contributed equally to this work.
[2] These authors also contributed equally to this work.

1. Introduction

Predicting the abundance, dynamics, and distributions of species are fundamental objectives in ecological studies and are essential for informing conservation management decisions under global change (Guisan and Thuiller, 2005; Ehrlén and Morris, 2015). Consequently, efforts in the last decade to make ecology a more predictive science have garnered considerable multi-disciplinary interest (Dietze et al., 2018; Lewis et al., 2022). Predictive ecological models are often built by relating an ecological response to one or more environmental factors. Depending on the modeling goals and data availability, environmental variability through space and/or time is used to explain variation in some ecological response in the same dimension(s) (e.g. Keith et al., 2008; Fordham et al., 2013; Gorzo et al., 2016; Pearce-Higgins et al., 2015; Bateman et al., 2016; Wilson et al., 2018). This relationship can then be transferred to predict ecological responses elsewhere in time (a temporal transfer) or space (a spatial transfer) (e.g. Osborne and Suárez-Seoane., 2002; Wenger and Olden, 2012; Sequeira et al., 2018; Yates et al., 2018). Evaluating how well predictive models transfer across time and space is crucial to gauging their overall predictive performance and the reliability of ecological inferences (Houlahan et al., 2017). In this paper, we present an approach based on a recently developed metric for measuring time series complexity to assess both the temporal and spatial transferability of site-specific population models used to predict abundances in the future, the past, and at new locations.

Temporal transfers can indicate the degree to which predictive models accurately reflect the temporal dynamics structuring natural systems (Houlahan et al., 2017). While many factors may affect the temporal transferability of ecological models, time series complexity is of particular importance for assessing the transferability of a model given its negative correlation with predictability (Riedl et al., 2013; Garland et al., 2014). More complex time series have less temporal signals carried forward from past states and more diversity in their temporal patterns. Time series complexity places an upper limit, called intrinsic predictability, on the temporal predictive performance of models fit to these time series (Beckage et al., 2011; Petchey et al., 2015; Pennekamp et al., 2019).

In addition to temporal transfers, time series models fit at one or more locations can be used to predict responses at new locations (e.g. Jenouvrier et al., 2014; Ryu et al., 2016; Smith and Edwards, 2021). The success of this kind of spatial transfer is conditional on the relationship found between response and predictor(s) at the initial location(s) being broadly applicable across space. This assumption is a frequent necessity in ecology and conservation management, especially for species of concern whose population dynamics are the subject of long-term ecological monitoring programs that are often restricted in spatial scope due to practical necessity (e.g. Menges, 2000; Morris et al., 2002; US Fish and Wildlife Service, 2020). Despite their ubiquity, general guidelines that are already well developed for determining when spatial transfers are appropriate for spatially-oriented models (e.g. Roberts and Hamann, 2012; Rapacciuolo et al., 2014; Soininen and Luoto, 2014; Sequeira et al., 2018; Yates et al., 2018; Iturbide et al., 2018) are lacking for assessing the spatial transferability of time series models. This is problematic if the time series underlying a temporal relationship modeled at a focal site, that is then used to predict time series at other sites, is highly complex. Given this, intrinsic predictability may also place an upper limit on spatial predictive performance, as well as the degree to which spatial transferability decays as a function of distance from the focal site.

Weighted permutation entropy (WPE) is a model-free metric recently introduced to ecology for approximating time series complexity and intrinsic predictability (Pennekamp et al., 2019). WPE can be useful as an indicator for the spatio-temporal transferability of time series models. For example, determining whether forecast performance is lower than expected when compared to WPE can help decide whether the model or the predictability of the data is the cause of low forecast performance (Beckage et al., 2011; Pennekamp et al., 2019). Such a forecasting assessment framework has been demonstrated through simulations and across time series of different species (Pennekamp et al., 2019). However, no studies have investigated the relationship between WPE and the spatio-temporal transferability of time series of different populations of the same species or used this framework diagnostically for wildlife population modeling.

Here, we use long-term (1980–2018) time series of Adélie penguin (*Pygoscelis adeliae*) breeding population abundances at 24 breeding sites across Antarctica to test the spatio-temporal transferability of predictive models that link Adélie population growth rates to environmental conditions derived from Earth System Models (Fig. 1). We use intrinsic predictability (as measured by WPE) to diagnose how well these predictive models transfer temporally and also how reliably temporal signals replicate through space. This case study is broadly applicable to ecological forecasting in general, as the large aperiodic fluctuations in abundance observed in Adélie penguin breeding populations, which play an outsized role in affecting the success of spatial and temporal transfers, are a feature of many natural systems for which ecological predictions are sought (Doak et al., 2008; Anderson et al., 2017; Clark and Luis, 2020).

2. Methods

2.1. Adélie colony growth rates and abundances

Adélie penguins have long been used as a sentinel species to measure Antarctic ecosystem health; their sensitivity to environmental variability is often reflected by changes in breeding population abundance and success, which can be monitored effectively given their gregarious nature (Ainley, 2002; Boersma, 2008). Adélie population dynamics have been studied intensively for nearly four decades, and while the Antarctic Peninsula region has seen significant declines in abundance, populations in eastern Antarctica and the Ross Sea have shown periods of increase (Lynch et al., 2012; Lyver et al., 2014; Che-Castaldo et al., 2017). Using Adélie nest count data from the Antarctic Penguin Biogeography Project, a database containing all publicly available Adélie penguin abundance and distribution data since 1979 (Che-Castaldo et al., 2023), we restrict our analyses to those breeding colonies whose time series span at least 24 years and do not have a larger than 5 year gap between any two counts. This criteria resulted in 24 colony-level time series of nest counts, that range from 24 to 39 years in length and contain between 15 to 36 nest counts.

We modify the Adélie penguin global population model introduced by Che-Castaldo et al. (2017) and expanded by Iles et al. (2020) to estimate annual nest abundance, for the 24 selected Adélie breeding colonies using a Bayesian framework (Fig. 1). This approach allows us to model complete time series of true nest abundance, which is otherwise unobservable due to observation error, from incomplete time series of nest counts. After fitting, we restrict our modeled time series of true nest abundance for each colony to begin and end with its first and last year of nest counts (see Appendix S1 for details regarding this model). We keep the colony growth rate posterior distribution means, hereafter called colony growth rates, which serve as the response variable in the colony-covariate models. Colony growth rates can be categorized as those with corresponding nest counts for both transition years (68%), those with nest counts for one of two transition years (21%), and those with no nest counts (11%).

2.2. Colony-covariate models

Many studies have examined the extent to which environmental conditions, principally sea ice as measured by passive microwave satellite-based sensors, affect Adélie penguin population dynamics (e.g. Fraser et al., 1992; Wilson et al., 2001; Croxall et al., 2002; Ainley et al., 2003; Jenouvrier et al., 2006; Che-Castaldo et al., 2017; Iles et al.,

2020). However, these efforts have been unavoidably limited by the number of environmental variables that can be directly observed at broad spatio-temporal scales. Earth system models (ESMs) provide an alternate source of environmental data that overcomes several constraints associated with passive microwave and other satellite-derived products. ESMs simulate moisture, energy, momentum, and carbon fluxes within and between various components of the Earth system, including the atmosphere, ocean, land, and sea ice (Kobayashi et al., 2015). In addition to physical ocean variables, ocean biogeochemistry and lower-trophic level ecosystem dynamics are simulated within the ocean component, thus providing variables such as primary productivity and zooplankton biomass, which often lack a complete observational record, especially in remote places like the Antarctic. While ESM simulations are subject to model biases stemming from the simplification of complex processes, they provide a complete record of numerous variables that could be important to Adélie population growth rates. As such, we take advantage of ESMs to incorporate elements of the physical ocean environment and food web that are not directly observable as model predictors.

With the goal of explaining spatio-temporal variation in colony growth rates, we select 12 environmental variables simulated in a forced ocean sea ice (FOSI) configuration of the Community Earth System Model (CESM), version 2 (Danabasoglu et al., 2020) run at a nominal 1° horizontal resolution and including ocean biogeochemistry (Long et al., 2021, Table 1). We associate colony growth rates with these environmental variables in the same year (a lag of 0) and with 4 lags ranging from 3 to 6 years, because the effect of juvenile survival on colony growth rates will become apparent only after surviving juveniles first return to breed (Ainley, 2002). This results in a total of 60 environmental variables. We normalize all variables at the colony-level, causing each variable to represent anomalies from the average condition experienced at each colony over the course of its time series.

We fit colony-specific models that associate annual variability in colony growth rates with the 60 ESM environmental variables. For each colony, we do this in three steps by 1) regressing all environmental variables on colony growth rates using regularized horseshoe priors, 2)
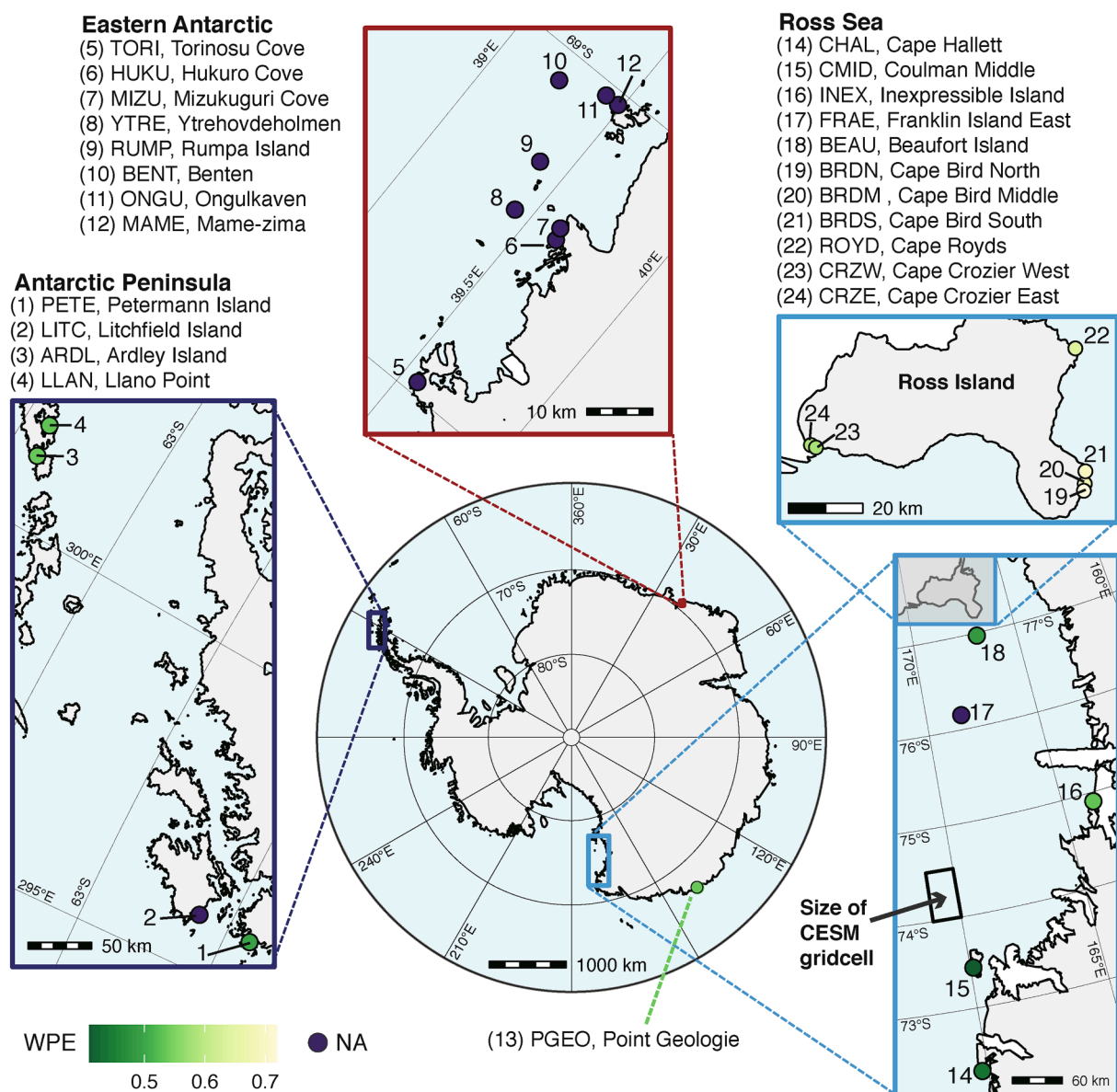

Fig. 1. Locations of the 24 Adélie penguin breeding colonies whose time series were long enough to be included in this study organized by regions in Antarctica. Weighted permutation entropy of colonies with longer than 30 year time series are shown with different shades of green. Regions are color coded, and this coloring scheme is repeated throughout the remainder of the figures to denote Antarctic region.

Table 1

Environmental variables from the Community Earth System Model version 2 (CESM2) used for variable selection in the colony covariate models. Times selected is how many times a variable was selected with regularized horseshoe regression as one of the top 3 variables across the 24 colonies.

Category	Code	Description	Times selected
sea ice	aice	area of sea ice (km^2)	12
sea ice	aicen	area of sea ice thinner than 0.6 meters (km^2)	9
sea ice	ardg	area of ridged sea ice (km^2)	7
sea ice dynamics	divu	divergence of sea ice (%/day)	6
sea ice dynamics	shear	shear of sea ice (%/day)	7
atmosphere	uatm	eastward surface wind (m/s)	5
atmosphere	vatm	northward surface wind (m/s)	0
atmosphere	rain	rainfall rate (cm/day)	4
ocean	zooC	zooplankton biomass (mmol C/m^2)	8
ocean	HMXL	mixed layer depth (m)	6
ocean	photoC	net primary production (mmol C/m^2/day)	5
ocean	temp	upper ocean temperature (top 10 m, °C)	1

selecting the three environmental variables with the largest effects, and 3) regressing these three colony-specific environmental variables on colony growth rates, replacing horseshoe priors with weakly informative ones (hereafter called colony-covariate models).

Regression modeling with horseshoe priors is a novel variable selection framework that is particularly useful when there are many possible explanatory variables but only a few with strong effects, a fairly common situation in ecology (Piironen et al., 2017). With this approach, we model the prior variances for covariate slope parameters as the product of a global shrinkage term that pulls all slopes towards 0, and a slope-specific parameter, that enables slopes to escape this pull (see Appendix S1 for details regarding the colony covariate models). After refitting these models with the three colony-specific environmental variables having the largest absolute slope estimates we compute RMSE and the traditional R^2 (square of Pearson's correlation coefficient). We also calculate RMSE for a null model for each colony, whose intercept is the mean colony growth rate. We compare the RMSE of the colony-covariate models and their respective null models to show how model fit improves when including environmental variables.

2.3. Temporal transferability

To assess the temporal transferability of the colony-covariate models, we create two training-test splits of each colony growth rate time series; a forecast split where the last 30% of the time series is the test set, and a hindcast split where the first 30% of the time series is the test set. In both splits, the remaining 70% of the data form the training sets. We refit the colony-covariate models, which were originally fit to all available colony growth rates, using the training sets, before forecasting or hindcasting as appropriate. We bypass the step where regularized horseshoe priors are used to determine which variables had the largest effect on growth rates. Instead, we refit the colony-covariate models to the training sets using the three best environmental variables originally selected for each colony.

Using model predictions of the two test sets, we calculate forecast and hindcast RMSE to score the predictive performance of these temporal transfers for each colony. We also compute the traditional R^2 for the forecast and hindcast for each colony. Using the mean colony growth rates from the training sets, we make null model forecasts and hindcasts for each colony and compute RMSE for these predictions of the test sets. Null models simply predict the average colony growth of the training set for each prediction year. We have also tested the temporal transferability of models with density dependence, and regional multi-colony

models where data for multiple colonies were pooled. In general, these models did not improve the temporal transferability of colony-covariate models (Appendix S2: Figs. S1 and S2, and Appendix S3: Fig. S1).

2.4. Spatial transferability

To assess the spatial transferability of the colony-covariate models, we use the model for each colony (focal colony) to predict colony growth rates for all other colonies (predicted colonies) within the Ross Sea. We repeat this process with the null model for the focal colony (long term average growth). We exclude colonies on the Antarctic Peninsula and from eastern Antarctica due to, respectively, the limited number of colonies in the region and the fact that colonies are located so close to one another as to preclude variation in their environmental covariates on the spatial scale used by CESM2. We do not evaluate these spatial transfers directly using RMSE, as this metric is heavily dependent on variability in colony growth rates for the predicted colonies, and is therefore not suitable for exploring the effect of distance on spatial predictive performance. For each focal colony, we instead compute the RMSE ratio of the colony covariate model and its respective null model for all predicted colonies. When the ratio is < 1, the colony-covariate model has improved spatial transferability over the null model, and vice versa. To determine how spatial transfers decay with distance from focal to predicted colony, we use linear regression to estimate the slope of the relationship between these spatial RMSE ratios and inter-colony distances.

2.5. Intrinsic predictability

Following the methodology in Pennekamp et al. (2019), we estimate the intrinsic predictabilities for the 14 colonies whose logged nest abundance time series ($lz_{i,t}$ in Appendix S1, Eq. 1) are long enough to compute weighted permutation entropy (WPE). There are several steps to calculating WPE:

1. Permutations of successive time series elements are established for a given "word" length (m). For example, if a time series consist of 7 elements (x_1 to x_7) and word length is 3, these permutations would start with x_1, x_2, x_3, followed by x_2, x_3, x_4, up until x_5, x_6, x_7. For this time series there would be 5 such time-ordered permutations.
2. Each permutation is ranked into an ordinal pattern before the relative frequencies of each of these possible permutations are calculated. For example, if $x_4 = 12, x_5 = 5$, and $x_6 = 7$, then its ordinal pattern (from lowest to highest) would be 2(x_5), 3(x_7), 1(x_4). For a given word length there are $m!$ possible permutations.
3. Unweighted and weighted permutation entropy (PE) are calculated using the Shannon entropy rate. For unweighted PE, this is computed as $\sum p(\pi)\log_2(p(\pi))$, where π is the ordinal pattern and $p(\pi)$ is its relative frequency. Weighted PE weights each time-ordered element with its variance to account for the degree of fluctuations in the time series:

$$p_w(\pi) = \frac{\sum \text{var}(X_t) \cdot \delta(\phi(X_t), \pi)}{\sum \text{var}(X_t)}, \tag{1}$$

where $\text{var}(X_t)$ is the variance of a given word and $\delta(\phi(X_t), \pi)$ is an indicator variable for cases where the ordinal pattern of $\phi(X_t)$ matches a given ordinal pattern π. Weighted PE is then computed using the Shannon entropy rate after replacing $p(\pi)$ with $p_w(\pi)$.
4. The Shannon entropy rate is normalized by $\log_2(m!)$ so that it is between 0 and 1. However, when there are ties in the ordinal patterns a tie-breaking method needs to be used. Here, we average the ranking of the tied elements, which increases the number of possible permutations, so normalization is done via $\log_2(2 \cdot m!)$.
5. A time delay may be added. For example, a time delay of 2 would mean the first word of the time series would be $x1, x3, x5$ (instead of

$x1, x2, x3$). We set the time delay to 1 as suggested by Pennekamp et al. (2019).

A WPE of 0 indicates a perfectly predictable time series, while a WPE of 1 signals that the time series is completely dominated by stochasticity. The choice of word length is limited by the length of the time series (Riedl et al., 2013). Here, we use a word length of 4 to estimate WPE, as a maximum word length of 3 or 4 is suitable for most ecological studies. We use correlation to quantify the relationship across colonies between

WPE and 1) the forecast and hindcast RMSE computed for temporal transfers described in Section 2.3 and 2) the rate of decay (slope) between RMSE ratio and physical distance computed for spatial transfers described in Section 2.4.

Fig. 2. R^2 of the 24 Adélie penguin colony covariate models computed from all data (A), the forecast test set (C), and the hindcast test set (E). Predictive performance (RMSE) for the 24 Adélie penguin colony-covariate models (orange circles) computed from all data (B), the forecast test set (D), and the hindcast test set (F) when compared to their long term average growth rate null model counterparts (white circles). Dark (light) gray lines show when a hindcast/forecast performs better (worse) than its associated null model. Sites are grouped and color coded according to their Antarctic region (Fig. 1).

3. Results

3.1. Variable selection

The set of environmental variables that best fit the data at each colony differs between the 24 Adélie colonies included in this study. Ice area was the most frequently selected variable, as it was one of top three variables with the largest slope for 12 colonies, followed by thin ice area and zooplankton biomass (Table 1, Appendix S4: Figs. S1–S21). These variables were frequently selected with different lags and they exhibited both positive and negative relationships with colony growth rates, depending on the colony (Appendix S5: Figs. S1–S72). Appendix S4 describe in detail the results of the model selection (Figs. S1–S21).

The colony-covariate models fit to all available data had an average R^2 of 0.45 (min = 0.21, max = 0.67). The R^2 values for colonies in Eastern Antarctica were generally larger than other regions, while Antarctic Peninsula colonies had the lowest R^2 values (Fig. 2A). All colony-covariate models demonstrated improvements in RMSE over their corresponding null model based on long-term average growth rate (Fig. 2B).

3.2. Temporal predictive performance

For colony-covariate models fit to the training sets, the average R^2 between temporal predictions and test data across the 24 colonies was 0.23 (min = 0, max = 0.66) for forecasts and 0.38 (min = 0, max = 0.88) for hindcasts (Fig. 2C,E). Forecasts for 45% of the colonies (11) and hindcasts for 75% of the colonies (18) had improvements in the RMSE of temporal predictions when compared to forecasts and hindcasts made with their corresponding null model based on average growth rates (Fig. 2D,F). Across sites that had lower RMSE than their null model counterparts, RMSE was reduced, on average, by 24% (SD = 15%) in forecasts and by 21% (SD = 12%) in hindcasts. Across sites that had higher RMSE than their null model counterparts, RMSE increased, on average, by 28% (SD = 29%) in forecasts, and by 14% (SD = 15%) in hindcasts. Intrinsic predictability was a strong predictor of forecast RMSE (ρ = 0.71, CI: 0.33, 0.92) and a weaker predictor of hindcast RMSE (ρ = 0.41, CI: −0.09, 0.78) (Fig. 3A,B). While colonies on Eastern Antarctica were not used for WPE and temporal transferability comparison as their time series were shorter than 30 years, their inclusion does not change the positive relationship between WPE and forecast and hindcast RMSE although it weakens it (see Appendix S6: Fig. S1).

3.3. Spatial predictive performance

Ross Sea colony-covariate models fit to all available data tended to predict colony growth rates of other colonies better when they were closer (Figs. 4 and 5). However, across Ross Island colonies, on average, only 16% (SD = 19%) of the spatial predictions of colony-covariate models showed any improvement over their null model counterparts (Figs. 4 and 5). Additionally, none of the spatial predictions from the colony covariate models for Beaufort Island (BEAU), two colonies on Cape Bird (BRDS and BRDM), Cape Royds (ROYD), or Coulman Middle (CMID) outperformed their respective null model counterparts (Figs. 4 and 5). Finally, WPE was positively associated with the slope of the relationship between spatial RMSE ratio and distance (ρ = 0.62, CI: 0.09, 0.91, Fig. 6), showing distance decay was stronger for sites with lower intrinsic predictability (higher WPE).

4. Discussion

Predictive modeling provides a robust approach to determine what (and if) we can learn from ecological systems (Houlahan et al., 2017; Lewis et al., 2023) and whether we can make useful forecasts (Clark et al., 2001). Without demonstrating transferability, models that simply explain variation in the data do not provide sufficient evidence that their modeled relationships are generalizable (Shmueli, 2010). Here, even though all colony-covariate models fit to all the data have relatively high R^2 values and perform better than models without environmental covariates, only about 45% (75%) of the site-level forecasts (hindcasts) show improvements (which are often minor) over their null model counterparts (Fig. 2D,F). In addition, differences in the temporal transferability of time series models are spatially variable. Specifically, colony covariate models from Ross Island tended to be more limited in their forecast or hindcast performances than those from other locations (Fig. 3). Also, models from Ross Island colonies transfer poorly to other colonies (have high distance decay, Figs. 4 and 5), meaning their temporal signals did not replicate reliably through space.

Interestingly, this variation in realized predictive performance across both time and space is correlated with intrinsic predictability (Figs. 3A, 6). To our knowledge, this is the first case where intrinsic predictability is shown to be: 1) spatially variable across populations for a single species and 2) associated with both the spatial and temporal transfer of time series models for populations within a species. These two findings reinforce the growing awareness that intra-specific variation in life history traits is substantial and often overlooked (e.g. Fitzsimmons, 2013; Che-Castaldo et al., 2018; Luiz et al., 2022).

The use of intrinsic predictability to help guide forecasting decisions and provide context for evaluating predictive outcomes is newly introduced to ecology (Beckage et al., 2011; Petchey et al., 2015; Pennekamp et al., 2019). Below we discuss the implications of this relationship between intrinsic predictability and realized predictive performance and possible ecological and statistical reasons for the spatial and temporal variability in WPE we observe across penguin colonies.
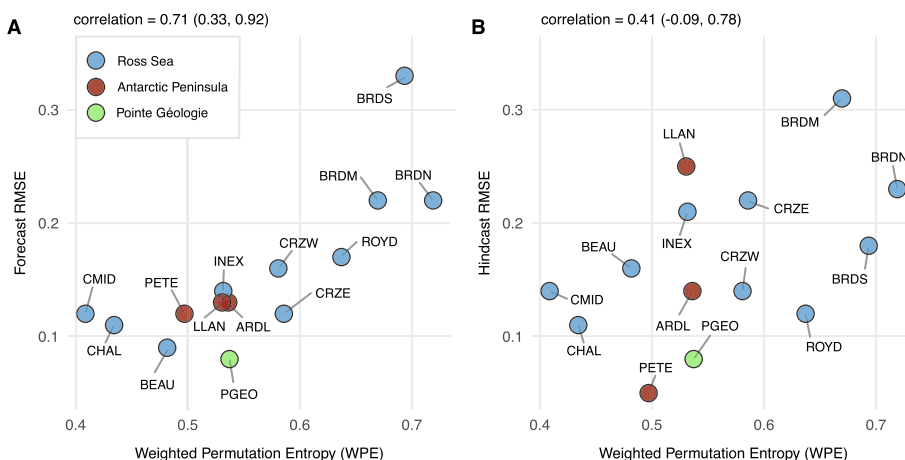

Fig. 3. The positive relationship between weighted permuation entropy (WPE) and forecast (A) and hindcast (B) RMSE for the 14 Adélie penguin colony-covariate models whose time series are long enough to calculate WPE. Colonies are color coded according to their Antarctic region (Fig. 1). Higher permutation entropy indicates lower intrinsic predictability and higher RMSE indicates lower prediction performance. We report the posterior mean and 95% equal-tailed posterior credible intervals for the forecast and hindcast correlation coefficient.
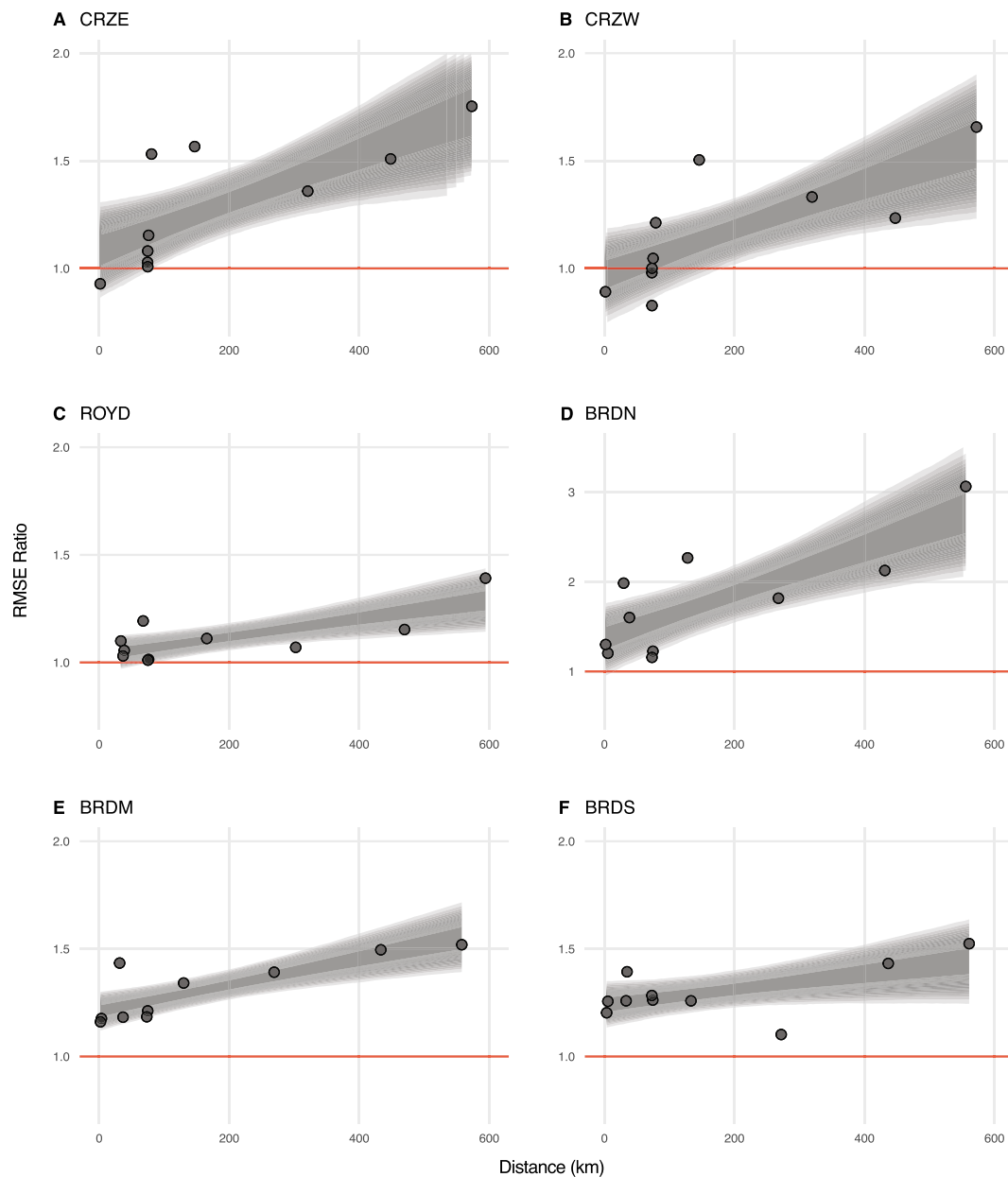
Fig. 4. Distance decay measured as the relationship between RMSE ratio and inter-colony distance for Ross Island Adélie penguin colonies at Cape Crozier East (CRZE), B) Cape Crozier West (CRZW), C) Cape Royds (ROYD), D) Cape Bird North (BRDN), E) Cape Bird Middle (BRDM), and F) Cape Bird South (BRDS). For a given focal colony, the RMSE ratio for any predicted colony is the ratio of RMSE of predictions from the colony-covariate model and its respective null model. The shaded areas show the 95% equal-tailed posterior credible intervals of the regression between RMSE ratio and inter-colony distance between focal and predicted site. The red line show when the RMSE ratio is 1, indicating that the colony covariate model and its corresponding null model have the same predictive performance. When the RMSE ratio is >1 the null model has better spatial transferability than the colony covariate model, and vice versa.

4.1. WPE as a predictor of spatial and temporal predictive performance

Intrinsic predictability (as calculated by WPE) provides a low-cost, easy to calculate measure of population time series complexity. In turn, WPE can set expectations as to the potential transferability of models before they are developed and evaluate their realized predictability after they are built. For example, when there is a positive relationship between WPE and the temporal predictive performance of site-specific models (e.g. Fig. 3A), differences in the realized predictability across sites are likely due to variation in the intrinsic predictabilities of the site-level time series themselves. This does not mean that model performances cannot be improved further. Instead, it is simply an indication that when sites are modeled individually, as we have done in this paper, those with higher WPE will tend to have lower predictive

performance compared to sites with lower WPE. A weak or no relationship (Fig. 3B) between WPE and temporal predictive performance suggests that modeling of some sites are showing lower realized predictability than expected given their WPE values relative to sites with higher WPE. This could be a sign that more model development is necessary. Model improvement can take various shapes, such as adding mechanistic processes such as density dependence, or using more relevant variables or increasing their spatial resolution. After refining, the relationship between WPE and temporal predictive performance can be checked again. Ideally, this process should continue iteratively until a positive relationship between intrinsic and realized predictability is achieved.

The relationship we find between WPE and spatial predictive performance (Fig. 6) implies that sites with more complex time series are
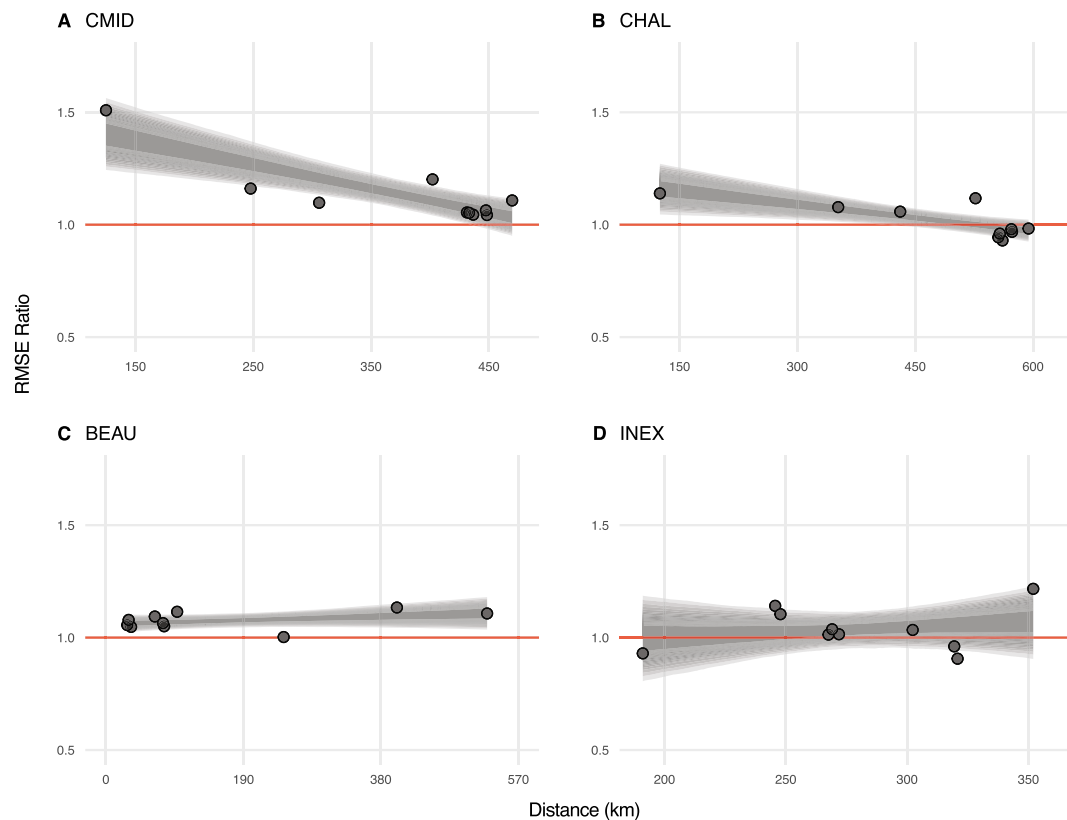
Fig. 5. Distance decay measured as the relationship between RMSE ratio and inter-colony distance for Ross Sea Adélie penguin colonies at A) Coulman Middle (CMID), B) Cape Hallet (CHAL), C) Beaufort Island (BEAU), and D) Inexpressible Island (INEX). For a given focal colony, the RMSE ratio for any predicted colony is the ratio of RMSE of predictions from the colony-covariate model and its respective null model. The shaded areas show the 95% equal-tailed posterior credible intervals of the regression between RMSE ratio and inter-colony distance between focal and predicted site. The red line show when the RMSE ratio is 1, indicating that the colony covariate model and its corresponding null model have the same predictive performance. When the RMSE ratio is >1 the null model has better spatial transferability than the colony covariate model, and vice versa.
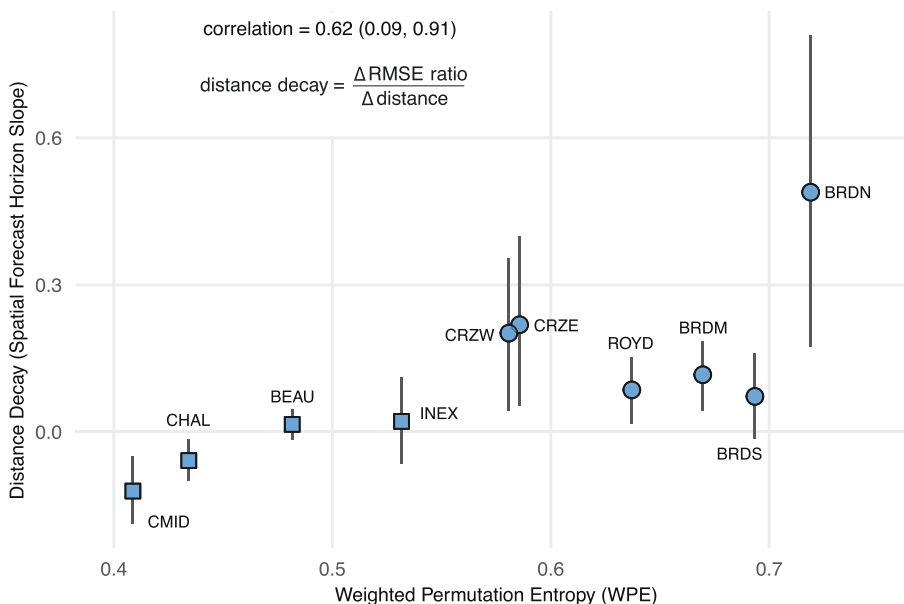

Fig. 6. The positive relationship between the distance decay of spatial transfers and weighted permutation entropy (WPE) for the 10 Adélie penguin colonies on Ross Sea whose time series are long enough to calculate WPE. We estimate distance decay as the slope of the linear regression between the RMSE ratio and the inter-colony distance between focal and predicted site (Figs. 4 and 5). We show the 95% equal-tailed posterior credible intervals for distance decay for the 6 Ross Island colonies (circles) and 4 other colonies in the Ross Sea (squares). We report the posterior mean and 95% equal-tailed posterior credible interval for the correlation coefficient between distance decay and WPE.

less spatially transferable with respect to the physical distance of the transfer. This pattern is consistent with what (Petchey et al., 2015) define as spatial forecast horizon when describing changes in community similarity and spatial transferability of species distribution models with respect to distance. Although we know of no ecological studies that

have used spatial forecast horizon to assess how well population dynamics from well-studied locations transfer across space, this definition can be expanded to quantify how the spatial transferability of a time series model at a focal site decays with distance when used to predict time series at other sites. The relationship between WPE and spatial

predictive performance raises two key concerns when developing sampling programs or gauging the appropriateness of long-term monitoring data for conservation or research purposes. The first is that generalizations to larger more policy-relevant spatial scales should not be based on a few sites with low intrinsic predictability (high WPE), as other poorly monitored or unvisited sites within the region of interest are likely to show drastically different dynamics. Second, an understanding of spatial forecast horizon as a function of WPE is critical when deciding how many sites to monitor within a region of interest and their spacing relative to one another. However, it is an unavoidable consequence of the long generation time for many species that both these determinations may require a significant investment in data collection.

When intrinsic predictability is low (high WPE), models that attempt to explain inter-annual fluctuations in growth rates may struggle to predict reliably across both space and time. In these cases, it may be better to focus on models that capture long-term average population dynamics, as these are more likely to be associated with climate or environmental conditions that change over longer time scales. For example, although the trajectories of some Adélie colonies suggest stochastic dynamics, the average growth rates of these sites still show a strong relationship with long-term winter sea ice concentrations (Iles et al., 2020). Null models can produce predictions that are as good as, or superior to, those from biologically-motivated population models, as was often the case here (Figs. 2D,F, 4, 5). Humphries et al. (2018) reach a similar conclusion when reporting on a data science competition whose purpose was to forecast Antarctic penguin abundance. In this competition, the top models included a mix of simple and sophisticated models, suggesting that well-constructed null models might be the most suitable choice for predictions of noisy systems, while also providing a yardstick by which to judge the predictive performance of more complex mechanistic models.

4.2. Geographic patterns in intrinsic predictability and the impact of black swan events

Intrinsic predictabilities across the 14 colonies included in this study whose time series were long enough to compute WPE follow a geographic pattern, where sites clustered on Ross Island have higher WPE values (lower intrinsic predictability) compared to colonies elsewhere in the Ross Sea region, Pointe Géologie (PGEO), and along the western Antarctic Peninsula (Fig. 2). Ross Island colonies also had higher levels of inter-annual variability in colony growth rates compared to other sites, and their time series were marked by large (and sometimes extreme) aperiodic fluctuations in growth (Appendix S7). Such fluctuations are not unique to Adélie penguins, but rather features of many natural systems. For example, unpredictable extreme shifts in abundance, also called black swan events, occur across a wide variety of animal taxa and usually manifest themselves as population diebacks or mass emigration events (Coulson et al., 2001; Anderson et al., 2017; Youngflesh and Lynch, 2017). Despite their rarity, these "surprise events" have been shown to be more common in ecology than previously thought and are often associated with overlooked or complex species interactions (Doak et al., 2008).

Ross Island colonies experienced what can be characterized as a black swan event, when from 2001–2005 two massive icebergs unexpectedly reduced access to these sites, dramatically altered their local environments while present, and affected Adélie metapopulation dynamics and vital rates (Robinson and Williams, 2012; Dugger et al., 2014; Lyver et al., 2014). These mega-icebergs directly impacted colony growth rates on Ross Island by causing large scale skip breeding events and/or relocation of adults from one colony to another (Dugger et al., 2010; LaRue et al., 2013). Skipped breeding, which occurs when adults forgo breeding in a given season, either due to poor body condition at the end of winter or adverse site conditions at the onset of the breeding season, is relatively common in penguins and other seabirds (Ainley, 2002; Jenouvrier et al., 2005; Massom et al., 2006). Large skip breeding

events cause extreme fluctuations in colony growth rates both in the year breeding was skipped and the following year when adults return en masse to breed (Talis et al., 2022). If the factors dictating large skip breeding events are stochastic, as is the case with historically rare but major ice calving episodes, time series of colony growth rates experiencing these events will have lower intrinsic predictability.

Skipped breeding is a specific case of demographic stochasticity (Talis et al., 2022), and any event that is related to demographic or environmental stochasticity, none of which occur in a predictable manner across time or space, will lead to variability of WPE across colonies. Ross Island colonies may be especially vulnerable to environmental stochasticity in general, as individuals from these sites undergo the longest annual migrations species-wide (Ballard et al., 2010), which limits their ability to buffer against inter-annual variation in winter and summer environmental conditions (Ainley, 2002; Dugger et al., 2014). Given this, it is not surprising that Adélie growth rates from the colonies at the leading latitudinal edge of their species range have the lowest intrinsic predictability. Understanding the population dynamics of species along range edges is a priority in ecology, as it can indicate how and why species distributions may change under climate change (MacArthur, 1972; Brown, 1984; Sexton et al., 2009). Nonetheless, our results here suggest that understanding and predicting such dynamics may be inherently limited by differences in intrinsic predictability between core and edge populations, especially since edge populations should experience higher levels of environmental and demographic stochasticity (Snyder, 2003; Sexton et al., 2009) and extreme events (Lavoie et al., 2021).

The spatial variability in WPE we report here also highlights how statistical methodology and survey decisions can impact intrinsic predictability estimation. Adélie colonies vary in the frequency and manner with which they are counted, leading to differences across colonies in data gaps requiring interpolation and observation uncertainty (Appendix S7). Bayesian state space models, which are widely used for wildlife population modeling, address these issues by separating process from observation error in order to estimate latent or "true" abundance for all time steps (e.g. Ahrestani and Hebblewhite, 2013). Computing WPE for growth rates derived from latent abundances, as opposed to raw counts, offers the advantages of working with time series without missing values, that have been adjusted to account for the "lost" information due to observation error as defined by Pennekamp et al. (2019), and which are ultimately of interest to population biologists. However, this approach comes at the expense of WPE no longer being a model-free evaluation of intrinsic predictability, but instead conditional on the model used to estimate abundance. In this case, these models can lead to time series of latent abundances that are smoother and, as a result, more predictable than those of raw counts on which they are based, where the level smoothing depends on the relative amounts of observation and process error and data sparseness. This smoothing occurred for colonies at Pointe Géologie (PGEO), due to observation error, and Coulman Middle (CMID) and Cape Hallett (CHAL), due to interpolation. As a result, differences in intrinsic predictability between Ross Island colonies and other sites might also be due to spatial differences in the degree of model-based smoothing, as well as ecological factors.

4.3. Intrinsic predictability of ecological time series may vary over time

The association between hindcast error and WPE, while still positive, is weaker compared to forecast error (Fig. 3B). This discrepancy could be caused by the violation of the assumption that intrinsic predictability is constant across a time series, regardless of its length. No studies of ecological predictability to date have investigated whether different parts of a time series might be more predictable than others. Instead, calculating WPE assumes that the process, or collection of processes, that generates a time series is constant. This assumption may not hold true, and different parts of time series may be more or less predictable than what a WPE value indicates for the full time series. Here, the CRZE

(Cape Crozier East) hindcast test set is more variable than the forecast test set and contains an extreme value for colony growth rate in 1987 (Fig. 7 A,B), leading to poorer predictive performance for the hindcast relative to the forecast (Fig. 3). While a single WPE calculated from the full time series is unlikely to capture differences in the population dynamics between the two test sets, neither the test or training sets are long enough for the outright calculation of their own WPE value.

Heterogeneity in the intrinsic predictability of time series may be caused by the fact that populations are usually regulated by different exogenous (e.g. environmental factors) and endogenous factors (e.g. density dependence), whose interactions can vary over time (e.g. Coulson et al., 2001; Nater et al., 2016; Nater et al., 2018). Consequently, these time-dependent interactions might alter the predictability of dynamics in different parts of a time series. For example, the effect of density dependence on the survival rates of African striped mouse (*Rhabdomys pumilio*) populations is at a minimum at moderate levels of food availability and increases when food is scarce or abundant (Nater et al., 2016). Since stronger density dependence relative to environmental stochasticity can lead to more predictable time series, as it is a deterministic factor, the intrinsic predictability of a time series of its abundance will change over time as food availability and, by extension, density dependence strength vary.

For Adélie penguins, we currently lack the ability to identify the relevant interactions that impact the intrinsic predictability of colony growth rates and how they vary, as the factors affecting Adélie population dynamics are complex and still an area of active research. Associating different ecological factors with WPE values computed from subsets of a time series requires time series longer than are currently available in this study. Nonetheless, such interactions are likely important in this system given that both atmospheric conditions, which help structure the physical environment, and competition between penguins and other marine predators for prey vary on decadal scales (e.g Ainley et al., 2005; Ainley et al., 2006; Ainley et al., 2007; Ainley et al., 2010; Warwick-Evans et al., 2022).

4.4. Conclusion

In this study, we demonstrated that intrinsic predictability, as measured by WPE, is a reliable predictor of the spatio-temporal transferability of environmentally-dependent population models. The relationship between intrinsic and realized predictability (or its lack thereof) can indicate whether the primary limiting factor affecting predictive performance is time series complexity or needed model improvements. Furthermore, spatial variability in intrinsic predictability can reveal the underlying geographical heterogeneity of population dynamics in an ecological system, as well as highlight populations that may have experienced stochastic environmental events that affect abundance. Similarly, temporal variability in intrinsic predictability across a time series for a single population can show how different regulatory mechanisms that are more predictable (e.g. density dependence) become more dominant in specific parts of the time series compared to others. We recommend that using WPE to approximate the intrinsic predictability of a time series should be standard practice when setting a priori expectations for model building and potential predictive performance across populations or sites, as well as for designing large-scale monitoring programs where spatial or temporal transfers are
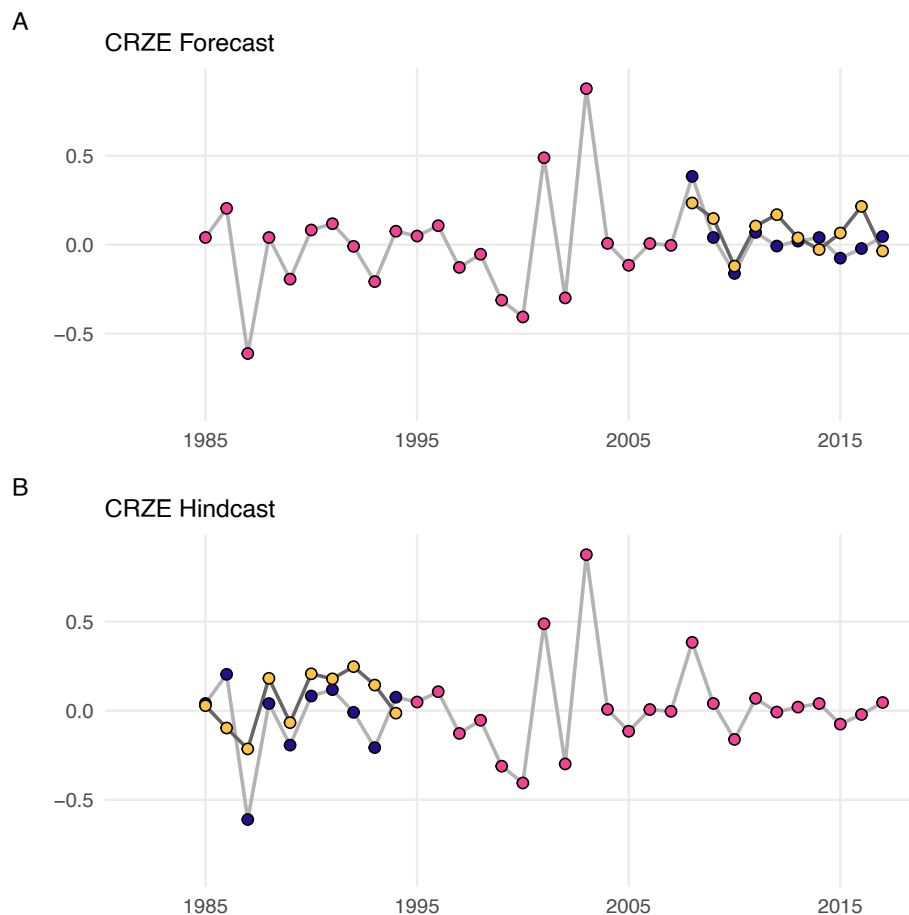

Fig. 7. Colony growth rate time series for Cape Crozier East (CRZE) showing (A) higher forecast predictive performance and (B) lower hindcast predictive performance. Training sets, test sets and forecasts/hindcasts are color coded separately. Pink is the training set, blue is the test set, and yellow is the prediction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

inevitable due to logistic constraints.

CRediT authorship contribution statement

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All code and data used in this paper is available at Github (https://github.com/bilgecansen/MAPPPD_transferability), and a Zenodo repository (https://zenodo.org/badge/latestdoi/513257374).

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ecolind.2023.110239.

References

Ahrestani, F.S., Hebblewhite, M., Post, E., 2013. The importance of observation versus process error in analyses of global ungulate populations. Sci. Rep. 3 (1), 1–10.

Ainley, D.G., 2002. The Adélie Penguin: Bellwether of Climate Change. Columbia University Press.

Ainley, D.G., Ballard, G., Emslie, S.D., Fraser, W., Wilson, P., Woehler, E., 2003. Adélie penguins and environmental change. Science 300 (5618), 429–430.

Ainley, D.G., Ballard, G., Dugger, K.M., 2006. Competition among penguins and cetaceans reveals trophic cascades in the western Ross Sea, Antarctica. Ecology 87 (8), 2080–2093.

Ainley, D.G., Clarke, E.D., Arrigo, K., Fraser, W.R., Kato, A., Barton, K.J., Wilson, P.R., 2005. Decadal-scale changes in the climate and biota of the Pacific sector of the Southern Ocean, 1950s to the 1990s. Antarct. Sci. 17 (2), 171–182.

Ainley, D., Ballard, G., Blight, L.K., Ackley, S., Emslie, S.D., Lescroël, A., Olmastroni, S., Townsend, S.E., Tynan, C.T., Wilson, P., Woehler, E., 2010. Impacts of cetaceans on the structure of Southern Ocean food webs. Marine Mammal Sci. 26 (2), 482–498.

Ainley, D., Ballard, G., Ackley, S., Blight, L.K., Eastman, J.T., Emslie, S.D., Lescroël, A., Olmastroni, S., Townsend, S.E., Tynan, C.T., Wilson, P., Woehler, E.J., 2007. Paradigm lost, or is top-down forcing no longer significant in the Antarctic marine ecosystem? Antarct. Sci. 19 (3), 283–290.

Anderson, S.C., Branch, T.A., Cooper, A.B., Dulvy, N.K., 2017. Black-swan events in animal populations. Proc. Nat. Acad. Sci. 114 (12), 3252–3257.

Ballard, G., Toniolo, V., Ainley, D.G., Parkinson, C.L., Arrigo, K.R., Trathan, P.N., 2010. Responding to climate change: Adélie penguins confront astronomical and ocean boundaries. Ecology 91 (7), 2056–2069.

Bateman, B.L., Pidgeon, A.M., Radeloff, V.C., Flather, C.H., VanDerWal, J., Akçakaya, H.R., Thogmartin, W.E., Albright, T.P., Vavrus, S.J., Heglund, P.J., 2016. Potential breeding distributions of US birds predicted with both short-term variability and long-term average climate data. Ecol. Appl. 26 (8), 2718–2729.

Beckage, B., Gross, L.J., Kauffman, S., 2011. The limits to prediction in ecological systems. Ecosphere 2 (11), art125.

Boersma, P.D., 2008. Penguins as marine sentinels. Bioscience 58 (7), 597–607.

Brown, J.H., 1984. On the relationship between abundance and distribution of species. Am. Nat. 124 (2), 255–279.

Che-Castaldo, C., Humphries, G.R., Lynch, H.J., 2023. Antarctic Penguin Biogeography Project: Database of abundance and distribution for the Adélie, chinstrap, gentoo, emperor, macaroni, and king penguin south of 60 S. ARPHA Preprints 4:ARPHA Preprints. doi:10.3897/arphapreprints.e101581.

Che-Castaldo, C., Jenouvrier, S., Youngflesh, C., Shoemaker, K.T., Humphries, G.R., McDowall, P., Landrum, L., Holland, M.M., Li, Y., Ji, R., Lynch, H.J., 2017. Pan-Antarctic analysis aggregating spatial estimates of Adélie penguin abundance reveals robust dynamics despite stochastic noise. Nat. Commun. 8 (1), 832.

Che-Castaldo, J., Che-Castaldo, C., Neel, M.C., 2018. Predictability of demographic rates based on phylogeny and biological similarity. Conserv. Biol. 32 (6), 1290–1300.

Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A., Lodge, D.M., Pascual, M., Pielke, R., Pizer, W., Pringle, C., Reid, W.V., Rose, K.A., Sala, O., Schlesinger, W.H., Wall, D.H., Wear, D., 2001. Ecological forecasts: An emerging imperative. Science 293 (5530), 657–660.

Clark, T., Luis, A.D., 2020. Nonlinear population dynamics are ubiquitous in animals. Nature Ecol. Evolut. 4 (1), 75–81.

Coulson, T., Mace, G.M., Hudson, E., Possingham, H., 2001. The use and abuse of population viability analysis. Trends Ecol. Evolut. 16 (5), 219–221.

Croxall, J.P., Trathan, P., Murphy, E., 2002. Environmental change and Antarctic seabird populations. Science 297 (5586), 1510–1514.

Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D.A., DuVivier, A.K., Edwards, J., Emmons, L.K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M.M., Large, W.G., Lauritzen, P.H., Lawrence, D.M., Lenaerts, J.T.M., Lindsay, K., Lipscomb, W.H., Mills, M.J., Neale, R., Oleson, K.W., Otto-Bliesner, B., Phillips, A.S., Sacks, W., Tilmes, S., Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J.E., Kinnison, D., Kushner, P.J., Larson, V.E., Long, M.C., Mickelson, S., Moore, J.K., Nienhouse, E., Polvani, L., Rasch, P.J., Strand, W.G., 2020. The Community Earth SystemModel Version 2 (CESM2). J. Adv. Model. Earth Syst. 12 (2) e2019MS001916.

Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S., Keitt, T.H., Kenney, M.A., Laney, C.M., Larsen, L.G., Loescher, H.W., Lunch, C.K., Pijanowski, B.C., Randerson, J.T., Read, E.K., Tredennick, A.T., Vargas, R., Weathers, K.C., White, E.P., 2018. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. Proc. Nat. Acad. Sci. 115 (7), 1424–1432.

Doak, D.F., Estes, J.A., Halpern, B.S., Jacob, U., Lindberg, D.R., Lovvorn, J., Monson, D.H., Tinker, M.T., Williams, T.M., Wootton, J.T., Carroll, I., Emmerson, M., Micheli, F., Novak, M., 2008. Understanding and predicting ecological dynamics: are major surprises inevitable? Ecology 89 (4), 952–961.

Dugger, K.M., Ballard, G., Ainley, D.G., Lyver, P.O., Schine, C., 2014. Adélie penguins coping with environmental change: Results from a natural experiment at the edge of their breeding range. Front. Ecol. Evolut. 2, 1–12.

Dugger, K.M., Ainley, D.G., Lyver, P.O., Barton, K., Ballard, G., 2010. Survival differences and the effect of environmental instability on breeding dispersal in an Adélie penguin meta population. Proc. Nat. Acad. Sci. 107 (27), 12375–12380.

Ehrlén, J., Morris, W.F., 2015. Predicting changes in the distribution and abundance of species under environmental change. Ecol. Lett. 18 (3), 303–314.

Fitzsimmons, J.M., 2013. How consistent are trait data between sources? A quantitative assessment. Oikos 122 (9), 1350–1356.

Fordham, D.A., Akçakaya, H.R., Brook, B.W., Rodríguez, A., Alves, P.C., Civantos, E., Triviño, M., Watts, M.J., Araújo, M.B., 2013. Adapted conservation measures are required to save the Iberian lynx in a changing climate. Nature Clim. Change 3 (10), 899–903.

Fraser, W.R., Trivelpiece, W.Z., Ainley, D.G., Trivelpiece, S.G., 1992. Increases in Antarctic penguin populations: Reduced competition with whales or a loss of sea ice due to environmental warming? Polar Biol. 11 (8), 525–531.

Garland, J., James, R., Bradley, E., 2014. Model-free quantification of time-series predictability. Phys. Rev. E 90 (5), 052910.

Gorzo, J.M., Pidgeon, A.M., Thogmartin, W.E., Allstadt, A.J., Radeloff, V.C., Heglund, P.J., Vavrus, S.J., 2016. Using the North American Breeding Bird Survey to assess broadscale response of the continent's most imperiled avian community, grassland birds, to weather variability. The Condor 118 (3), 502–512.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8 (9), 993–1009.

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

Humphries, G.R., Che-Castaldo, C., Bull, P., Lipstein, G., Ravia, A., Carrión, B., Bolton, T., Ganguly, A., Lynch, H.J., 2018. Predicting the future is hard and other lessons from a population time series data science competition. Ecolog. Inform. 48, 1–11.

Iles, D.T., Lynch, H., Ji, R., Barbraud, C., Delord, K., Jenouvrier, S., 2020. Sea ice predicts longterm trends in Adélie penguin population growth, but not annual fluctuations: Results from a range-wide multiscale analysis. Glob. Change Biol. 26 (7), 3788–3798.

Iturbide, M., Bedia, J., Gutiérrez, J.M., 2018. Background sampling and transferability of species distribution model ensembles under climate change. Global Planet. Change 166, 19–29.

Jenouvrier, S., Barbraud, C., Cazelles, B., Weimerskirch, H., 2005. Modelling population dynamics of seabirds: Importance of the effects of climate fluctuations on breeding proportions. Oikos 108 (3), 511–522.

Jenouvrier, S., Holland, M., Stroeve, J., Serreze, M., Barbraud, C., Weimerskirch, H., Caswell, H., 2014. Projected continent-wide declines of the emperor penguin under climate change. Nature Climate Change 4 (8), 715–718.

Jenouvrier, S., Barbraud, C., Weimerskirch, H., 2006. Sea ice affects the population dynamics of Adélie penguins in Terre Adélie. Polar Biol. 29, 413–423.

Keith, D.A., Akçakaya, H.R., Thuiller, W., Midgley, G.F., Pearson, R.G., Phillips, S.J., Regan, H.M., Araújo, M.B., Rebelo, T.G., 2008. Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. Biol. Lett. 4 (5), 560–563.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., Takahashi, K., 2015. The JRA-55 Reanalysis. General Specifications and Basic Characteristics. Journal of the Meteorological Society of Japan. Ser. II 93 1, 5–48.

LaRue, M.A., Ainley, D.G., Swanson, M., Dugger, K.M., Lyver, P.O., Barton, K., Ballard, G., 2013. Climate change winners: receding ice fields facilitate colony expansion and altered dynamics in an Adélie penguin metapopulation. PLoS One 8 (4), e60568.

Lavoie, M., Jenouvrier, S., Blanchette, P., Lariviere, S., Tremblay, J.-P., 2021. Extreme climate events limit northern range expansion of wild turkeys. Oecologia 197, 633–650.

Lewis, A.S.L., Woelmer, W.M., Wander, H.L., Howard, D.W., Smith, J.W., McClure, R.P., Lofton, M.E., Hammond, N.W., Corrigan, R.S., Thomas, R.Q., Carey, C.C., 2022. Increased adoption of best practices in ecological forecasting enables comparisons of forecastability. Ecol. Appl. 32 (2), e2500.

Lewis, A.S.L., Rollinson, C.R., Allyn, A.J., Ashander, J., Brodie, S., Brookson, C.B., Collins, E., Dietze, M.C., Gallinat, A.S., Juvigny-Khenafou, N., Koren, G., McGlinn, D.J., Moustahfid, H., Peters, J.A., Record, N.R., Robbins, C.J., Tonkin, J., Wardle, G.M., 2023. The power of forecasts to advance ecological theory. Methods Ecol. Evol. 14 (3), 746–756.

Long, M.C., Moore, J.K., Lindsay, K., Levy, M., Doney, S.C., Luo, J.Y., Krumhardt, K.M., Letscher, R.T., Grover, M., Sylvester, Z.T., 2021. Simulations With the Marine Biogeochemistry Library (MARBL). J. Adv. Model. Earth Syst. 13 (12).

Luiz, O.J., Olden, J.D., Kennard, M.J., Crook, D.A., Douglas, M.M., Saunders, T.M., Wedd, D., Adair, B., King, A.J., 2022. Substantial intraspecific trait variation across a hydrological gradient in northern Australian fishes. Ecosphere 13 (7), e4169.

Lynch, H.J., Naveen, R., Trathan, P.N., Fagan, W.F., 2012. Spatially integrated assessment reveals widespread changes in penguin populations on the Antarctic Peninsula. Ecology 93 (6), 1367–1377.

Lyver, P.O., Barron, M., Barton, K.J., Ainley, D.G., Pollard, A., Gordon, S., McNeill, S., Ballard, G., Wilson, P.R., 2014. Trends in the breeding population of Adélie penguins in the Ross Sea, 1981–2012: A coincidence of climate and resource extraction effects. PLoS One 9 (3), e91188.

MacArthur, R.H., 1972. Geographical ecology: Patterns in the distribution of species. Princeton University Press.

Massom, R.A., Stammerjohn, S.E., Smith, R.C., Pook, M.J., Iannuzzi, R.A., Adams, N., Martinson, D.G., Vernet, M., Fraser, W.R., Quetin, L.B., Ross, R.M., Massom, Y., Krouse, H.R., 2006. Extreme anomalous atmospheric circulation in the West Antarctic Peninsula region in austral spring and summer 2001/02, and its profound impact on sea ice and biota. J. Clim. 19 (15), 3544–3571.

Menges, E.S., 2000. Population viability analyses in plants: Challenges and opportunities. Trends Ecol. Evolut. 15 (2), 51–56.

Morris, W.F., Bloch, P.L., Hudgens, B.R., Moyle, L.C., Stinchcombe, J.R., 2002. Population viability analysis in endangered species recovery plans: past use and future improvements. Ecol. Appl. 12 (3), 708–712.

Nater, C.R., van Benthem, K.J., Canale, C.I., Schradin, C., Ozgul, A., 2018. Density feedbacks mediate effects of environmental change on population dynamics of a semidesert rodent. J. Anim. Ecol. 87 (6), 1534–1546.

Nater, C.R., Canale, C.I., Van Benthem, K.J., Yuen, C.-H., Schoepf, I., Pillay, N., Ozgul, A., Schradin, C., 2016. Interactive effects of exogenous and endogenous factors on demographic rates of an African rodent. Oikos 125 (12), 1838–1848.

Osborne, P.E., Suárez-Seoane, S., 2002. Should data be partitioned spatially before building large-scale distribution models? Ecol. Model. 157 (2), 249–259.

Pearce-Higgins, J.W., Eglington, S.M., Martay, B., Chamberlain, D.E., 2015. Drivers of climate change impacts on bird communities. J. Anim. Ecol. 84 (4), 943–954.

Pennekamp, F., Iles, A.C., Garland, J., Brennan, G., Brose, U., Gaedke, U., Jacob, U., Kratina, P., Matthews, B., Munch, S., Novak, M., Palamara, G.M., Rall, B.C., Rosenbaum, B., Tabi, A., Ward, C., Williams, R., Ye, H., Petchey, O.L., 2019. The intrinsic predictability of ecological time series and its potential to guide forecasting. Ecol. Monogr. 89 (2), e01359.

Petchey, O.L., Pontarp, M., Massie, T.M., Kéfi, S., Ozgul, A., Weilenmann, M., Palamara, G.M., Altermatt, F., Matthews, B., Levine, J.M., Childs, D.Z., McGill, B.J., Schaepman, M.E., Schmid, B., Spaak, P., Beckerman, A.P., Pennekamp, F., Pearse, I. S., 2015. The ecological forecast horizon, and examples of its uses and determinants. Ecol. Lett. 18 (7), 597–611.

Piironen, J., Vehtari, A., 2017. Sparsity information and regularization in the horseshoe and other shrinkage priors. Electron. J. Statist. 11 (2), 5018–5051.

Rapacciuolo, G., Roy, D.B., Gillings, S., Purvis, A., 2014. Temporal validation plots: Quantifying how well correlative species distribution models predict species' range changes over time. Methods Ecol. Evol. 5 (5), 407–420.

Riedl, M., Müller, A., Wessel, N., 2013. Practical considerations of permutation entropy: A tutorial review. Eur. Phys. J. Special Topics 222 (2), 249–262.

Roberts, D.R., Hamann, A., 2012. Method selection for species distribution modelling: are temporally or spatially independent evaluations necessary? Ecography 35 (9), 792–802.

Robinson, N., Williams, M., 2012. Iceberg-induced changes to polynya operation and regional oceanography in the southern Ross Sea, Antarctica, from in situ observations. Antarct. Sci. 24 (5), 514–526.

Ryu, H.Y., Shoemaker, K.T., Kneip, A., Pidgeon, A.M., Heglund, P.J., Bateman, B.L., Thogmartin, W.E., Akçakaya, H.R., 2016. Developing population models with data from marked individuals. Biol. Conserv. 197, 190–199.

Sequeira, A.M.M., Bouchet, P.J., Yates, K.L., Mengersen, K., Caley, M.J., 2018. Transferring biodiversitymodels for conservation: Opportunities and challenges. Methods Ecol. Evolut. 9 (5), 1250–1264.

Sexton, J.P., McIntyre, P.J., Angert, A.L., Rice, K.J., 2009. Evolution and ecology of species range limits. Annu. Rev. Ecol. Evol. Syst. 40 (1), 415–436.

Shmueli, G., 2010. To Explain or to Predict? Stat. Sci. 25 (3), 289–310.

Smith, A.C., Edwards, B.P.M., 2021. North American Breeding Bird Survey status and trend estimates to inform a wide range of conservation needs, using a flexible Bayesian hierarchical generalized additive model. The Condor 123 (1) duaa065.

Snyder, R.E., 2003. How demographic stochasticity can slow biological invasions. Ecology 84 (5), 1333–1339.

Soininen, J., Luoto, M., 2014. Predictability in species distributions: A global analysis across organisms and ecosystems. Glob. Ecol. Biogeogr. 23 (11), 1264–1274.

Talis, E.J., Che-Castaldo, C., Sen, B., Krumhardt, K., Lynch, H.J., 2022. Variability, skipped breeding and heavy-tailed dynamics in an Antarctic seabird. J. Anim. Ecol. 91 (12), 2437–2450.

US Fish and Wildlife Service, 2020. Species status assessment for Smith's blue butterfly (*Euphilotes enoptes smithi*) (Version 1.0). Ventura, California.

Warwick-Evans, V., Kelly, N., Dalla Rosa, L., Friedlaender, A., Hinke, J., Kim, J., Kokubun, N., Santora, J., Secchi, E., Seyboth, E., Trathan, P.N., 2022. Using seabird and whale distribution models to estimate spatial consumption of krill to inform fishery management. Ecosphere 13 (6), e4083.

Wenger, S.J., Olden, J.D., 2012. Assessing transferability of ecological models: an under appreciated aspect of statistical validation: Model transferability. Methods Ecol. Evol. 3 (2), 260–267.

Wilson, P., Ainley, D.G., Nur, N., Jacobs, S., Barton, K., Ballard, G., Comiso, J., 2001. Adélie penguin population change in the pacific sector of Antarctica: Relation to sea-ice extent and the Antarctic Circumpolar Current. Mar. Ecol. Prog. Ser. 213, 301–309.

Wilson, S., Smith, A.C., Naujokaitis-Lewis, I., 2018. Opposing responses to drought shape spatial population dynamics of declining grassland birds. Diversity Distribut. 24 (11), 1687–1698.

Yates, K.L., Bouchet, P.J., Caley, M.J., Mengersen, K., Randin, C.F., Parnell, S., Fielding, A.H., Bamford, A.J., Ban, S., Barbosa, A.M., Dormann, C.F., Elith, J., Embling, C.B., Ervin, G.N., Fisher, R., Gould, S., Graf, R.F., Gregr, E.J., Halpin, P.N., Heikkinen, R.K., Heinänen, S., Jones, A.R., Krishnakumar, P.K., Lauria, V., Lozano-Montes, H., Mannocci, L., Mellin, C., Mesgaran, M.B., Moreno-Amat, E., Mormede, S., Novaczek, E., Oppel, S., Ortuño Crespo, G., Peterson, A.T., Rapacciuolo, G., Roberts, J.J., Ross, R.E., Scales, K.L., Schoeman, D., Snelgrove, P., Sundblad, G., Thuiller, W., Torres, L.G., Verbruggen, H., Wang, L., Wenger, S., Whittingham, M.J., Zharikov, Y., Zurell, D., Sequeira, A.M., 2018. Outstanding Challenges in the Transferability of Ecological Models. Trends Ecol. Evolut. 33 (10), 790–802.

Youngflesh, C., Lynch, H.J., 2017. Black-swan events: Population crashes or temporary emigration? Proc. Nat. Acad. Sci. 114 (43), E8953–E8954.