

Optimal population prediction of sandhill crane recruitment based on climate-mediated habitat limitations

Brian D. Gerber^{1*}, William L. Kendall², Mevin B. Hooten², James A. Dubovsky³ and Roderick C. Drewien^{4,†}

¹Colorado Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife and Conservation Biology, Colorado State University, Fort Collins, CO 80523-1403, USA; ²U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife and Conservation Biology, Colorado State University, Fort Collins, CO 80523-1403, USA; ³Division of Migratory Bird Management, U.S. Fish and Wildlife Service, Lakewood, CO 80215-5573, USA; and ⁴University of Idaho, Moscow, ID 83844, USA

Summary

1. Prediction is fundamental to scientific enquiry and application; however, ecologists tend to favour explanatory modelling. We discuss a predictive modelling framework to evaluate ecological hypotheses and to explore novel/unobserved environmental scenarios to assist conservation and management decision-makers. We apply this framework to develop an optimal predictive model for juvenile (<1 year old) sandhill crane *Grus canadensis* recruitment of the Rocky Mountain Population (RMP). We consider spatial climate predictors motivated by hypotheses of how drought across multiple time-scales and spring/summer weather affects recruitment.

2. Our predictive modelling framework focuses on developing a single model that includes all relevant predictor variables, regardless of collinearity. This model is then optimized for prediction by controlling model complexity using a data-driven approach that marginalizes or removes irrelevant predictors from the model. Specifically, we highlight two approaches of statistical regularization, Bayesian least absolute shrinkage and selection operator (LASSO) and ridge regression.

3. Our optimal predictive Bayesian LASSO and ridge regression models were similar and on average 37% superior in predictive accuracy to an explanatory modelling approach. Our predictive models confirmed *a priori* hypotheses that drought and cold summers negatively affect juvenile recruitment in the RMP. The effects of long-term drought can be alleviated by short-term wet spring–summer months; however, the alleviation of long-term drought has a much greater positive effect on juvenile recruitment. The number of freezing days and snowpack during the summer months can also negatively affect recruitment, while spring snowpack has a positive effect.

4. Breeding habitat, mediated through climate, is a limiting factor on population growth of sandhill cranes in the RMP, which could become more limiting with a changing climate (i.e. increased drought). These effects are likely not unique to cranes. The alteration of hydrological patterns and water levels by drought may impact many migratory, wetland nesting birds in the Rocky Mountains and beyond.

5. Generalizable predictive models (trained by out-of-sample fit and based on ecological hypotheses) are needed by conservation and management decision-makers. Statistical regularization improves predictions and provides a general framework for fitting models with a large number of predictors, even those with collinearity, to simultaneously identify an optimal predictive model while conducting rigorous Bayesian model selection. Our framework is impor-

*Corresponding author. E-mail: bgerber@colostate.edu

†Retired.

tant for understanding population dynamics under a changing climate and has direct applications for making harvest and habitat management decisions.

Key-words: *Grus canadensis*, least absolute shrinkage and selection operator, modelling, multicollinearity, Palmer drought index, Palmer drought severity index, predictive, ridge regression, SPEI, standardized precipitation–evapotranspiration index

Introduction

It remains true that if we can predict successfully on the basis of a certain explanation, we have good reason, and perhaps the best sort of reason, for accepting the explanation (Kaplan 1964; Pg. 350).

Prediction is a fundamental part of science and paramount in science-based decision-making. There are both a need and a demand for predictive models to assist decision-makers (Kendall 2001; Sutherland 2006). Generalizable predictive models are needed to explore how novel conditions, such as future environmental changes and anthropogenic disturbances, will affect ecosystems, species distributions and populations of managed species (e.g. harvested or endangered). These models are also crucial to evaluating alternative conservation and management actions (e.g. hunting, habitat manipulation or reintroduction scenarios; Nichols *et al.* 2007). As such, predictive modelling has practical importance to conservation decision-making and is an important approach to provide insights about ecological processes (Sutherland 2006).

Ecologists tend to use explanatory modelling to understand ecological processes, rather than predictive modelling, perhaps even when prediction is the goal (Shmueli 2010; Evans, Norris & Benton 2012; Hooten & Hobbs 2014). An explanatory modelling approach often focuses on a small set of discrete models to evaluate competing hypotheses about mechanistic processes motivated from theory and past observations to infer processes from data (Burnham & Anderson 2002). In contrast, a predictive modelling approach emphasizes the forecasting of new or future observations, possibly under novel scenarios. A continuous model set may be evaluated, which can lead to better predictive performance than from a discrete set (Hastie, Tibshirani & Friedman 2009). From a statistical perspective, predictive modelling is concerned with the generalizability of the model to predict beyond the sample space, and thus, it is preferred to measure model fit based on out-of-sample performance, rather than within-sample as often used for explanatory modelling (Shmueli 2010; Hooten & Hobbs 2014). Predictive and explanatory approaches are not mutually exclusive and thus not inconsistent or incompatible (Kaplan 1964; Bickel & Li 2006); however, there can be advantages in model performance when the intended application of the model is consistent with the modelling approach (Hastie, Tibshirani & Friedman 2009).

Ecologists may not be readily applying predictive modelling (i.e. out-of-sample model fitting) for a number of reasons. First, statistical tools may not be known and accessible (Mac Nally 2000; Hooten & Hobbs 2014). Alternatively, ecologists may consider prediction to be hopeless; species, population and ecosystem responses to environmental change are often highly complex, nonlinear and perhaps computationally irreducible (Beckage, Gross & Kauffman 2011). A more insidious reason may be that ecologists conflate prediction and explanation (Shmueli 2010); it is often assumed that a highly explanatory model will also be a good predictive model. This is not strictly the case. Even when the causal mechanisms of a process are identified and estimated well enough by an approximating model (\mathcal{M}), a model other than \mathcal{M} might be preferable for prediction (Shmueli 2010; Gelman, Hwang & Vehtari 2014). Lastly, ecologists may be concerned that predictive modelling lacks the ability to provide explanation of underlying processes. This is also not strictly true. While predictive modelling can be done without predictors based on clear motivating hypotheses (due to Simpson's Paradox, Simpson 1951; Pearl 2014), doing so will likely lead to a more robust generalizable model, as well as provide explanatory power (Ashley, Granger & Schmalensee 1980). Comparing models by their out-of-sample predictive performance is often considered the very best way for investigating support of hypotheses (Kaplan 1964; Ashley, Granger & Schmalensee 1980; Hooten & Hobbs 2014).

We are interested in developing a predictive modelling framework that links correlated climate variables to migratory animal population vital rates. Specifically, we focus on developing a generalizable model that links drought across multiple temporal scales to measures of juvenile (<1 year old) sandhill crane (*Grus canadensis*) recruitment in the Rocky Mountains. Exploring how climate affects crane recruitment is important in understanding future crane population dynamics, especially when considering climate change. Secondly, by taking a predictive approach, novel climate scenarios can easily be explored to help inform managers about habitat or harvest decisions or be integrated into a formal decision framework. Lastly, knowledge of how climate affects crane breeding productivity in the Rocky Mountains could also be informative to other migratory, ground-nesting birds that breed in similar wetland habitats, including shorebirds, wading birds and waterfowl (Austin & Pyle 2004).

Sandhill cranes are a long-lived North American bird and have the lowest annual recruitment of any harvested game bird in North America (Drewien, Brown & Kendall 1995). Typically, cranes do not breed until several years of age, after which they produce two eggs in a single clutch per year, but generally only one chick survives to fledging (see Gerber *et al.*, 2014). Re-nesting can occur if the first nesting attempt fails early in the incubation period, but multiple successful clutches per year have not been observed. We are specifically focused on the Rocky Mountain Population (RMP), consisting of a single subspecies (*G. c. tabida*), which migrate annually (Fig. 1). Cranes breed during the summer within palustrine and riparian wetlands throughout the central Rocky Mountains, along with other species, including the Canada Goose (*Branta canadensis*), Eared Grebe (*Podiceps nigricollis*), Mallard (*Anas platyrhynchos*), American Avocet (*Recurvirostra americana*), Long-billed Curlew (*Numenius americanus*) and Virginia Rail (*Rallus limicola*; Austin & Pyle 2004). In the fall and spring during migration, birds stop over in the San Luis Valley (SLV) in central southern Colorado, while wintering takes place primarily in the Rio Grande Valley of New Mexico (Drewien & Bizeau 1974). A survival analysis of RMP cranes has revealed low temporal variation in adult survival (Drewien, R.C., unpublished data), suggesting recruitment may be a driver of RMP dynamics (Sæther & Bakke 2000).

Following previous findings of a negative trend in recruitment during a period of drought (Drewien, Brown & Kendall 1995), we hypothesize that drought within the RMP breeding range limits the availability of suitable wetlands for nesting, which negatively affects annual

recruitment. We define drought as a combination of meteorological and hydrological processes that lead to a 'period of abnormally dry weather sufficiently prolonged for the lack of precipitation to cause a serious hydrological imbalance' (Heim 2002). As with many species of waterfowl, sandhill cranes almost exclusively nest in wetlands and have high site fidelity that may limit the availability of potential wetlands for nesting (see Gerber *et al.*, 2014). Regional drying of wetlands due to drought is likely to affect recruitment in two ways: (1) reduce the proportion of the population that attempts nesting and (2) reduce nest success for those that do nest. Water depth around nests has been shown to be an important factor in nesting success, where deeper water is thought to inhibit terrestrial predators from accessing nests (Austin, Henry & Ball 2007; McWethy & Austin 2009).

We also hypothesize that cold temperatures and snowpack during the post-hatching and colt stage (June–August) will reduce survival between hatching and arrival at the SLV. Sandhill cranes are nidifugous, begin feeding within 24 h of hatching, and are brooded for up to 3 weeks (see Gerber *et al.*, 2014). Cold weather during the post-brooding stage could be fatal or may limit developmental growth due to increased thermoregulatory demands. Secondly, cold weather is likely to limit the availability of insect prey, which are an important resource for chicks/colts in obtaining necessary protein and to increase weight. Cranes that grow slowly during the colt stage are known to have a lower probability of surviving until 6 months of age (Drewien, R.C., unpublished data). Lastly, a deeper snowpack in summer may inhibit nesting or indicate a mismatch in timing of run-off and recharging of wetlands during the crane nesting season.

To consider our hypotheses in a modelling framework, we need predictor variables, which commonly in ecological and climate studies are collinear. This creates certain challenges in fitting many types of models and can prevent accurate measures of parameters and their uncertainty. We take an approach that can accommodate these challenges. We highlight a predictive modelling approach using statistical regularization to achieve a generalizable model using a relatively large set of possibly collinear predictors, motivated from ecological hypotheses. Statistical regularization is a broad concept that includes model selection and the idea of parsimony in model performance (Bickel & Li 2006; Hooten & Hobbs 2014). We specifically are interested in Bayesian ridge regression (Hoerl & Kennard 1970) and least absolute shrinkage and selection operator (LASSO; Tibshirani 1996). Ridge regression and LASSO provide a rigorous approach to Bayesian model selection and prediction (Hooten & Hobbs 2014). Each approach regularizes parameters based on out-of-sample performance, but they penalize model complexity differently. Most importantly, LASSO has variable selection properties and can remove effects of predictors by allowing them to be zero, which gives

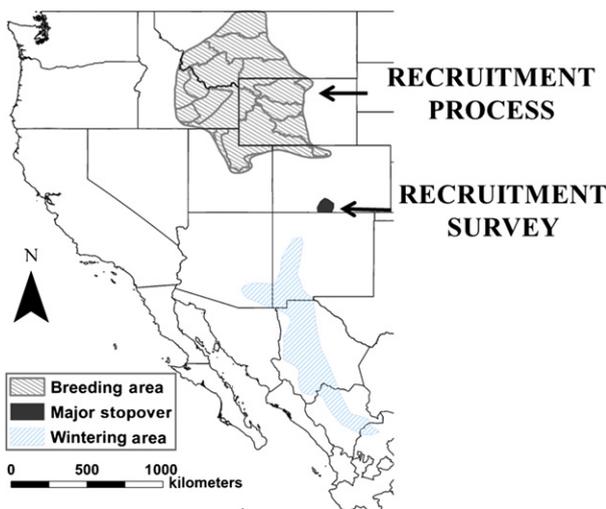


Fig. 1. Seasonal distribution of the sandhill crane Rocky Mountain population. The major stopover for spring and fall migrants is the San Luis Valley of southern Colorado, where crane biologists have conducted juvenile recruitment surveys since 1972 (Recruitment Survey). Juveniles are reared in the central Rocky Mountains (Recruitment Process). The regions within the breeding area indicate NOAA climate divisions.

the optimal model an additional amount of interpretability (Hastie, Tibshirani & Friedman 2009). Our objectives are to (1) develop a predictive modelling framework for linking climate effects to migratory animal populations using the Bayesian LASSO and ridge regression techniques, (2) identify an optimal out-of-sample (via cross-validation) predictive model of RMP juvenile recruitment (i.e. proportion of total population) and compare it to an explanatory model, (3) formally evaluate ecological hypotheses concerning climate effects on RMP juvenile recruitment and (4) provide predictions of recruitment across a range of climate scenarios to inform future management decisions.

Materials and methods

RECRUITMENT MONITORING

Starting in 1972, an annual survey of juvenile recruitment has been conducted in the RMP during the fall stopover at the SLV (Brown 2013); the survey has been conducted consistently by the same crane biologists (RCD or W. Brown) in mid-late October when most of the population is present (Drewien, Brown & Kendall 1995). The survey covers the entire distribution of cranes within the SLV where birds are assessed to subspecies and age group based on morphological characteristics. During this time, juvenile cranes can be distinguished from adults, but these differences typically disappear by the following summer and thus only allow two age classes to be identified (Drewien, Brown & Kendall 1995).

Biologists sampled cranes by observing flocks at roost, diurnal feeding and midday loafing sites. Sites were only visited once during the 3- to 6-day survey. A flock was defined as an aggregate of cranes separated from others by >50 m. Flocks of all sizes were sampled to minimize bias associated with the disproportional number of juveniles that occur in smaller-sized flocks (Drewien, Brown & Kendall 1995). Smaller flocks are generally not difficult to observe in the SLV, and thus, under-detection is not a concern (see Supporting Information). Large flocks (>300 individuals) were spot-sampled (10–25 birds) at intervals along random zigzag lines in both the perimeter and interior of the group to avoid bias due to family groups concentrating on the edges of large flocks. Between 1972 and 2013, an average of 5518 cranes were sampled each year (range, 1997–8318, Table S1, Supporting information). The recruitment estimate is a function of several biological and ecological processes, including the proportion of the population that attempts breeding, clutch size, nesting success and survival from chick to the fall migration stopover at the SLV.

PREDICTORS FOR HYPOTHESES

Drought

Measuring drought is difficult. No consensus exists on operational definitions and the complexity of the spatial and temporal climate mechanisms that make up a drought are challenging to capture. This lack of consensus led to the development of a multitude of drought indices (Heim 2002). Two popular indices that are consistent with our interpretation of drought are included in separate models: the Palmer drought indices (\mathcal{M}_P , Mishra &

Singh 2010) and the standardized precipitation–evapotranspiration indices (SPEI; \mathcal{M}_{SP} ; Vicente-Serrano *et al.* 2010a). The Palmer indices are the most widely used and readily available indices of drought (Mishra & Singh 2010). They account for precipitation, evapotranspiration and capacity for soils to hold moisture. The SPEI is a difference equation of precipitation and potential evapotranspiration that is measured at sequential temporal scales, making it applicable to multiple hydrological processes that influence drought (i.e. river discharge from headwaters, reservoir storage and groundwater storage; Vicente-Serrano *et al.* 2010a).

We include three types of Palmer drought index for \mathcal{M}_P : (1) the Palmer drought severity index (PDSI), (2) the Palmer hydrological drought index (PHDI) and (3) the Palmer drought Z index (PDZI). All indices are unitless, estimated monthly, and generally range from –6 to 6, where positive values indicate wet conditions and negative values dry conditions (0.00–0.49 are near normal events, 0.50–0.99 are incipient, 1.00–1.99 are mild, 2.00–2.99 are moderate, 3.00–3.99 are severe, and 4.00–6.00 are extreme). PDSI measures meteorological drought over a 12- to 18-month time period (Vicente-Serrano, Beguera & López-Moreno 2010b); PHDI is a hydrological measure of water supply that operates on time-scales longer than PDSI. PDSI and PHDI are insensitive to temporary drought relief and are thus slow to change. Therefore, for both PDSI and PHDI we include only a single annual value in \mathcal{M}_P (April). PDZI is a short-time scale measure of monthly drought with no influence from previous months. Thus, PDZI can capture temporary drought relief from precipitation, that may recharge wetlands, that is not captured in the longer time-scale drought indices. We include monthly PDZI from pre-nesting through nesting and into the colt stage (April–July).

A number of criticisms have been made of the Palmer indices (Heim 2002; Mishra & Singh 2010), which led to also investigating the SPEI. Because we are interested in drought at multiple temporal scales, we include several indices in \mathcal{M}_{SP} , with measures of SPEI on short- (1 month, April–July), medium- (12 months, April) and long-term scales (24, 36 and 48 months, April). SPEI is unitless and generally ranges from –3 to 3, where positive values indicate wet conditions and negative values dry conditions (0.00–0.39 are near normal events, 0.50–0.70 are abnormal, 0.80–1.20 are moderate, 1.30–1.50 are severe, 1.60–1.90 are extreme, and 2.00–3.00 are exceptional).

A shortcoming of all drought indices is that they ignore snowpack and subsequent spring/summer run-off, which is likely to influence wetland water levels used by RMP cranes. As such, we explicitly incorporate snowpack using measurements of maximum snow depth (MXSD) from 30 Cooperative Observer Network stations (COOP) chosen throughout the breeding range that (1) are located in areas representative of where cranes nest (e.g. elevation between 1500 and 3500 m) and (2) continuous data are available from 1972 to 2013. We include MXSD for March and April in both models, where higher values are hypothesized to be beneficial in recharging wetlands during spring runoff.

Cold weather

To incorporate our summer cold weather hypotheses into both models, predictors of the number of days below freezing in a given month (DT32; measurements from COOP) in May, June, July and August are included. MXSD is included for months with snow (May and June).

Spatially explicit predictors

Since the recruitment survey is conducted during migration, spatially explicit knowledge of recruitment by state or region within the breeding area is unavailable; thus, inference is for the entire RMP. However, regional variation in weather/drought could affect local breeding aggregations disproportionately. In addition, the RMP is not evenly distributed within its breeding range (Fig. S1, Supporting information), and thus, regional climate effects will have disproportional impact on the total recruitment depending on the proportion of the breeding population exposed to the regional climate. We investigated the within-year spatial variability in drought indices across the RMP breeding range using the climate divisions defined by NOAA and found considerable variation (see Fig. S2, Supporting information), which we attempt to accommodate.

Four weighting strategies (w_i , see Statistical Model) are applied to the predictors of both models (M_P, M_{SP}), which are compared by their predictive performance. First, M_1 predictors are weighted uniformly across all years by the long-run average proportion of the annual RMP pre-migratory fall survey count within each climate region; this count covers the known staging grounds (Thorpe, Donnelly & Collins 2013) and has been in operation since 1995 (Kruse, Dubovsky & Cooper 2014). Second, M_2 predictors are weighted by the proportional size of each climate region within the breeding area and thus assumes larger areas have more breeding cranes than smaller areas. Third, M_3 predictors are weighted equally across all climate divisions in each year. Fourth, M_4 predictors are weighted by the RMP fall count by year; for years prior to 1995, we calculate a climate region weight by the average proportion of cranes over the first 5 years of the survey.

In total, 14 (M_P) and 16 (M_{SP}) predictor variables are considered that we hypothesize are related to the realized climate mechanisms that affect sandhill crane recruitment (Table 1). However, the climate and weather variables are at least partially and potentially strongly correlated. Rather than remove collinear variables, which is often done in explanatory modelling (a form of variable selection prior to formal model selection), we see potential value in each for predicting recruitment and take a modelling approach that can accommodate the mathematical difficulties of collinear variables. Palmer indices and weather variables (MXSD and DT32) are available from NOAA (<http://www.ncdc.noaa.gov/>), and SPEI is available from CSIC (<http://sac.csic.es/spei/index.html>).

STATISTICAL REGULARIZATION AND MODEL FITTING

Statistical regularization

Statistical regularization seeks to optimize the generalizability of a model by trading off bias and variance by constraining model complexity, such that there is a decrease in variance for a smaller gain in bias. First, a criterion to optimize needs to be chosen that combines model fit (e.g. model deviance) or conversely, lack of fit (e.g. squared prediction error), with a weighted measure of model complexity (criterion = model fit + $\lambda \times$ model complexity; Reineking & Schröder 2006); the parameter λ provides the weighting or regularization that controls the relative importance (i.e. exchange rate) of model fit and complexity. Model fit can be any loss function, which measures the discrepancy between

Table 1. Ecological hypotheses about the direction of effect of climate predictors on sandhill crane recruitment. Models included either Palmer drought indices (M_P) or standardized precipitation–evapotranspiration indices (SPEI, M_{SP})

Time Scale (months)	M_{SP} Predictors*	M_P Predictors	Hypothesized effect on recruitment
Short(1)	SPEI1 (April–July)	PDZI (April–July)	Positive**
Short(1)	MXSD (March, April)	MXSD (March, April)	Positive
Short(1)	DT32 & MXSD (May)	DT32 & MXSD (May)	Negative
Short(1)	DT32 (June–August) & MXSD (June)	DT32 (June–August) & MXSD (June)	Negative
Medium(12)	SPEI12 (April)	–	Positive
Long(>12)	–	PHDI (April)	Positive
Long(12–18)	–	PDSI (April)	Positive
Long(24)	SPEI24 (April)	–	Positive
Long(36)	SPEI36 (April)	–	Positive
Long(48)	SPEI48 (April)	–	Positive

*MXSD = maximum snow depth, DT32 = number of freezing days, PDSI = Palmer drought severity index, PDZI = Palmer drought Z index, PHDI = Palmer hydrological drought index. The number following SPEI indicates the time scale and the parentheses following a predictor variable indicate the month(s) it pertains to.

**Positive values of drought indicate wetter years; thus, a positive effect on recruitment means that increasing values of a drought index leads to higher recruitment.

observations and within- or out-of-sample model predictions. Preferably, this function has the necessary characteristics (local and proper) that guarantees an honest evaluation for the chosen model and data (Gelman, Hwang & Vehtari 2014). For linear models, model complexity is often measured by the number of estimated parameters (P ; i.e. degrees of freedom) and can be generalized to include the magnitude of parameter estimates (β) as $\sum_{p=1}^P |\beta_i|^q$, when $q > 0$ (Reineking & Schröder 2006; Fig. S3, Supporting information). The value of q has a large effect on the measure of model complexity and the behaviour it imposes on estimates and model performance for different values of λ (see Discussion; Hastie, Tibshirani & Friedman 2009). Two special cases that have been studied extensively are $q = 1$ and $q = 2$, which correspond to LASSO and ridge regression, respectively. Both LASSO and ridge have likelihood and Bayesian interpretations. In a Bayesian context, choosing how model complexity is controlled can equivalently be defined via the type of prior distribution (prior) that is chosen for each β_i , where LASSO ($q = 1$) is achieved using the Laplace distribution and ridge ($q = 2$) is achieved using the normal distribution (Hastie, Tibshirani & Friedman 2009). The magnitude of the prior ‘regulating’ the parameter estimates is controlled by λ as part of the prior variance (see ‘Statistical Model’ for how λ is used in the prior). For an explicit link between constrained optimization in the likelihood paradigm ($q = 1$, LASSO and $q = 2$, ridge) and prior specification in the Bayesian paradigm (Laplace Prior, LASSO and Normal Prior, ridge), see Hoerl & Kennard (1970) and Tibshirani (1996).

Choosing a loss function and measure of model complexity will depend on a study's objectives and assumptions of the structure of the data and predictors (Reineking & Schröder 2006). Because we are interested in developing a generalizable model that is appropriate for predicting outside of the sample space, we chose to find an optimal value of λ based on an out-of-sample log predictive score (Gelman, Hwang & Vehtari 2014; equation presented in 'Statistical model'). In doing so, a continuous model selection is effectively performed via regularization to optimize predictive accuracy.

Statistical model

A Bayesian binary regression model was developed to provide inference (Fig. 2). The observation process is the number of juveniles observed (y_t) from the total number of cranes assessed by age (N_t) from year $t = 1, 2, \dots, T$. The proportion of juveniles (p_t) is linked to scaled and centred predictor variables (P in total for Q climate regions) from the breeding area for each year t . Predictors are defined in a $P \times Q$ matrix (\mathbf{X}_t) that is weighted by a $Q \times 1$ vector (\mathbf{w}_t , which is different for models \mathcal{M}_{1-4}), to estimate a $P \times 1$ vector of coefficients (β). We evaluate alternative regulators (λ) by out-of-sample cross-validation, in which coefficients are shrunk towards zero using the prior as a constraint; the intercept (α_0) is not regularized and given a relatively uninformative prior (Fig. 2). For ridge regression, the exchangeable prior for each parameter is defined as the normal distribution ($\beta_p \sim N(0, \tau^2)$), where the variance $\tau^2 = \frac{1}{\lambda}$, while a Laplace or double-exponential distribution ($\beta_p \sim L(0, \tau^2)$) is used to specify the LASSO. The Laplace distribution is more pointed at the mean than the normal distribution and has heavier tails (Fig. S4, Supporting information). A random effect (η) is also included to accommodate additional process uncertainty for annual variation that is not explained by our predictors.

For each model and regularization method, a search is performed for the optimal value of λ using out-of-sample cross-validation, in which a portion of the data is left out from the model (y_{hold}), while the rest (y_{train}) is used for estimation of unknown parameters. This portion of the data that is not observed by the model (y_{hold}) and thus does not belong to the sample being used to estimate parameters is applied to evaluate the predictive accuracy of the model. We use a y_{hold} of length $n = 1, 2, 3$ and 6 (CV1, CV2, CV3, CV6, respectively) to evaluate the stability of small sample cross-validation in selecting λ . The data left out as a group are chosen randomly. We consider 100 evenly spaced regulators (ridge: $\lambda \in [exp(-5), exp(15)]$; LASSO: $\lambda \in [exp(0), exp(20)]$). For leave-one-out cross-validation (CV1), the out-of-sample predictive fit is calculated as the log pointwise predictive density ($lppd$) for a given λ across all Markov chain Monte Carlo (MCMC) samples (post-burn-in) from $s = 1, 2, \dots, S$ and all hold

outs of data from $i = 1, 2, \dots, n$, where $\theta = [\alpha_0, \beta, \sigma^2]$:

$$lppd_{CV1} = \sum_{i=1}^n \log_e \int [y_{i,hold} | \theta] [\theta | y_{train}] d\theta, \tag{eqn 1}$$

and computed as,

$$\sum_{i=1}^n \log_e \left(\frac{1}{S} \sum_{s=1}^S p(y_{i,hold} | \theta_{is}) \right). \tag{eqn 2}$$

We consider the λ value with the highest $E[lppd]$ as the regularization parameter that achieves the best prediction, and we consider the model ($\mathcal{M}_{SP1-4}, \mathcal{M}_{P1-4}$) with the highest $E[lppd]$ as our best predictive model. We fit all models using MCMC methods by sampling from full-conditional distributions using the R Programming Language (2014); 500 000 MCMC samples were used with a burn-in of 100 000 samples. Posterior convergence was primarily assessed graphically. Additional information on applying statistical regularization and CV with the $lppd$ in Bayesian modelling can be found in the Supporting Information, including R code.

PREDICTION FROM AN EXPLANATORY MODEL

We compare the $E[lppd]$ of our best model chosen using cross-validation for LASSO and ridge regression ($\mathcal{M}_{SP1-4}, \mathcal{M}_{P1-4}$) with a model developed following an explanatory modelling approach. The explanatory model includes only a subset of predictors; collinear predictors are removed when pairwise correlation coefficients between predictors ($|r|$) are >0.28 (Graham 2003) and no regularization is performed. This is the most common approach to remove collinear variables (Graham 2003). We also use the full correlation matrix (\mathbf{R}) to investigate collinearity by taking its determinant ($|\mathbf{R}|$) as well as calculate the condition number (CN). The $|\mathbf{R}|$ is a summary of all r , such that if no correlation exists, this value will be 1.0, and otherwise it will be <1.0 , with worse collinearity being present as the value approaches zero. The CN is the ratio of the largest to smallest nonzero singular value of the predictor matrix ($(\mathbf{X}_t \mathbf{w}_t)^T$), where the predictors all have a mean of zero and a standard deviation of one. Multicollinearity becomes an issue in estimating parameters when the $CN > 5.4$ (Lazaridis 2007).

ECOLOGICAL EFFECTS AND PREDICTION

We evaluate our ecological hypotheses by summarizing the posterior distribution of our coefficients of the optimal predictive model as either $P(\beta_p > 0 | \mathbf{y})$ for those effects we consider to have a positive effect on recruitment and $P(\beta_p < 0 | \mathbf{y})$ for those predic-

- Observation Process: $y_t \sim Binomial(N_t, p_t)$
- Ecological Process: $logit(p_t) = \alpha_0 + (\mathbf{X}_t \mathbf{w}_t)^T \beta + \eta_t$
- Process Uncertainty: $\eta_t \sim Normal(0, \sigma^2)$
- Intercept Prior Information: $\alpha_0 \sim Normal(0, 1000)$
- Uncertainty Prior Information: $log(\sigma^2) \sim Normal(0, 2)$
- Ridge Regularization (prior): $\beta \sim MVN(\mathbf{0}, \tau^2 \mathbf{I}_p)$
- LASSO Regularization (prior): $\beta \sim Laplace(\mathbf{0}, \tau^2 \mathbf{I}_p)$
- Joint Posterior Distribution: $[\alpha_0, \beta, \sigma^2 | \mathbf{y}, \mathbf{X}, \mathbf{w}] \propto [\mathbf{y} | \mathbf{X}, \mathbf{w}, \beta, \alpha_0, \sigma^2] [\beta] [\alpha_0] [\sigma^2]$

Fig. 2. Regularized Bayesian binary regression model used to identify a generalizable predictive model of sandhill crane recruitment based on climate effects; shown with the prior specification for ridge regression and least absolute shrinkage and selection operator, which act to control model complexity. Note, we are specifying variances, rather than precisions.

tors that have a negative effect (Table 1); predictors that have little effect on recruitment will have probabilities near 0.5. We also investigate variability in juvenile recruitment by predicting recruitment from our best model under a variety of climate scenarios.

Results

The best out-of-sample predictive model for \mathcal{M}_{1-4} was found to be similar for ridge and LASSO, but with ridge consistently performing slightly better. Overall, there was little difference in predictive accuracy between \mathcal{M}_{P2-4} and \mathcal{M}_{SP1-4} , while \mathcal{M}_{P1} was clearly superior (Fig. 3). Across all models and both regularization methods, CV1-CV3 were mostly consistent in selecting the same optimal λ , while CV6 tended to select a slightly larger regulator value (e.g. Fig. S5, Supporting information). The differences in selected regulators have a small effect on the expected predictive accuracy ($E[lppd]$), but are clearly not different when considering the complete posterior distribution of $lppd$ (Fig. S6, Supporting information). Across all weighting strategies, the $lppd$ was found to be optimal at different values of λ for ridge regression and LASSO due to their different forms of regularization. However, functionally λ occurred at similar values of the regularized coefficients (e.g. Fig. S7, Supporting information), which led to similar conclusions about our ecological hypotheses (Table 2). For \mathcal{M}_{P1} , the optimal λ using ridge regression improved the predictive accuracy by 21.3% compared to a non-regularized model (diffuse prior), while LASSO improved by 20.8%.

Predictors of all our models were found to be considerably collinear, including the best performing model, (\mathcal{M}_{P1} : mean, SD and range of r was 0.26, 0.27, -0.31 to 0.96, respectively). Additionally, the $|R|$ was 0.00004, and

the CN was 33.06. After removing collinear predictors, the explanatory model $\mathcal{M}_{P1-Explanatory}$ had a total of five predictors (PDZI-May, DT32-May-July-August and MXSD-June) that were considerably less collinear ($\mathcal{M}_{P1-Explanatory}$: mean, SD and range of r was -0.01, 0.16, -0.21 to 0.22, respectively). The $|R|$ was 0.77, and the CN was 2.16. However, there was also a loss of predictive accuracy from the explanatory model, with \mathcal{M}_{P1} fit using ridge and LASSO being $\approx 37\%$ more accurate than $\mathcal{M}_{P1-Explanatory}$.

For more than four decades, the annual juvenile recruitment varied considerably while remaining consistently low (0.034–0.120 [range, Brown 2013, Fig. S8, Supporting information]). We found the majority of our ecological hypotheses regarding the effect of drought and weather on juvenile recruitment to be supported using either LASSO or ridge for the best predictive model, \mathcal{M}_{P1} (Table 2). Contrary to our hypotheses, there was strong support for a negative influence of PDZI-April and positive support of DT32-May on recruitment. There was also no significant effect of PDZI-July. Two of the drought indices (PDZI-May-July) were highly sensitive to the weighting strategy used with at least a 0.1 change in probability of the hypothesis (Table S2, Supporting information); we are therefore less confident about inferring the importance of these predictors. We found similar general support of our hypotheses using the SPEI drought indices \mathcal{M}_{SP1-4} , but more variability in support across weighting strategies (Table S3, Supporting information).

Using our best model (\mathcal{M}_{P1}), we predicted juvenile recruitment under a range of climate scenarios. We found short-term (PDZI) and long-term drought (PDSI, PHDI) to affect juvenile sandhill crane recruitment differently. Our model suggests that the alleviation of drought by the

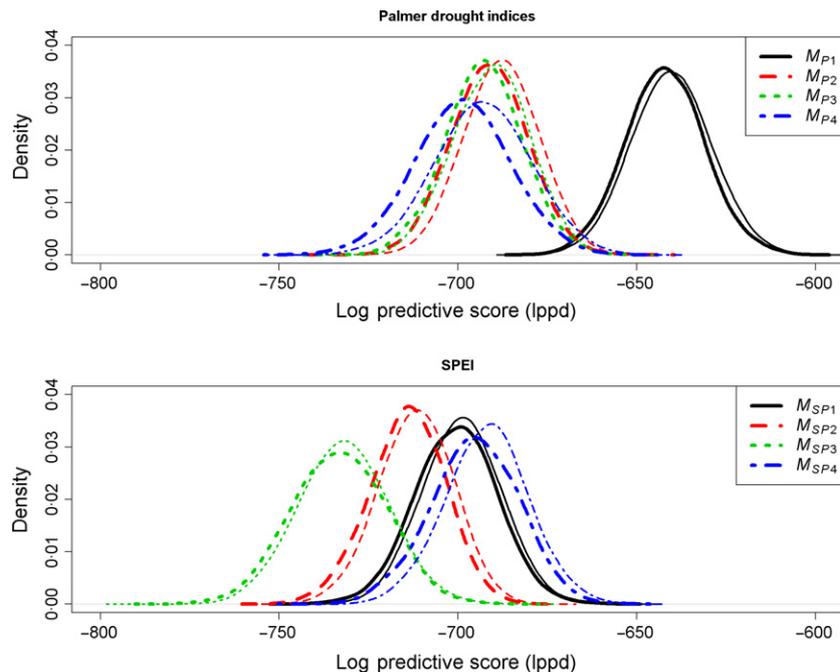


Fig. 3. Posterior distributions of the optimal log pointwise predictive density from models using Palmer drought indices and SPEI (standardized precipitation–evapotranspiration indices), spatially weighted by different strategies (\mathcal{M}_{P1-4} , \mathcal{M}_{SP1-4} , respectively); thick lines indicate the model was fit using least absolute shrinkage and selection operator and narrow lines indicate ridge.

Table 2. Ecological hypotheses, predictor variables and evidence of the relationship from the optimal out-of-sample predictive model, M_{P1}

Predictor	Hypothesized effect on recruitment	Ridge regression P (Hypothesis Data)	LASSO P (Hypothesis Data)
PDSI	Positive	1.00	1.00
- April			
PHDI	Positive	1.00	1.00
- April			
PDZI	Positive	0.01	0.01
- April			
- May	Positive	0.98	0.97
- June	Positive	1.00	1.00
- July	Positive	0.31	0.40
DT32	Negative	0.00	0.00
- May			
- June	Negative	1.00	1.00
- July	Negative	1.00	1.00
- August	Negative	1.00	1.00
MXSD	Positive	1.00	1.00
- March			
- April	Positive	1.00	1.00
- May	Negative	1.00	1.00
- June	Negative	1.00	1.00

increasing of short-term drought indices can positively affect juvenile recruitment, but not as much compared to alleviation of long-term drought (Fig. 4). We also found that the number of freezing days in the summer months (June, July and August) have a less pronounced effect on juvenile recruitment than drought, but one that is consistently negative (Fig. 5). Investigating the timing of maximum snowpack across spring and summer revealed strong effects on sandhill crane recruitment, with the best scenario for crane production being high snowpack in the spring that is drastically reduced by the summer months (Fig. 6).

Discussion

PREDICTIONS, VITAL RATES AND HARVEST DECISIONS

We demonstrated a predictive modelling framework that links ecological hypotheses of correlated climate mechanisms to a migratory population's vital rates, while overcoming common data and modelling issues. Most importantly, we did not limit our investigation of climate effects to a single value of drought, but instead allowed the model to evaluate a set of numerous temporally scaled measures of drought that were considered *a priori* to be candidate predictors of recruitment. Understanding whether climate drives variation in vital rates will often depend on evaluating temporally correlated mechanisms. By doing so, we were able to improve predictions and better understand how short- and long-term drought proportionally impact cranes.

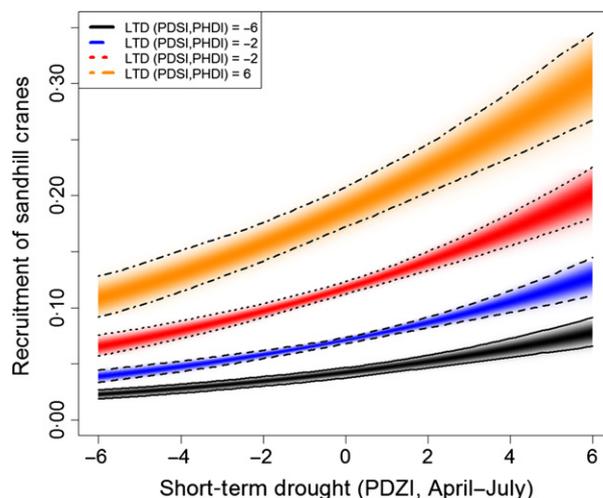


Fig. 4. Variation in predicted juvenile sandhill crane recruitment across values of long-term (LTD) and short-term drought from an optimal out-of-sample predictive model, fit by ridge (M_{P1}). Weather predictors (maximum snow depth and the number of freezing days within a month) are held constant at their mean. The uncertainty is shaded in proportion to the posterior distribution and the lines are 95% credible intervals.

In the RMP, juvenile recruitment may be especially important in understanding crane population dynamics. While adult survival is known to have the greatest potential to affect population change (i.e. more elastic and sensitive), observed temporal variation is often low in long-lived birds (Sæther & Bakke 2000). In contrast, juvenile recruitment and survival may have a lower potential to affect population change, but greater observed influence on long-lived birds due to high temporal variation (Sæther & Bakke 2000). For the RMP, adult survival has been found to be stable over a 23-year period (Drewien, R.C., unpublished data), while recruitment has high temporal variation (Fig. S8, Supporting information, Brown 2013). Our best predictive model suggests that this variation is at least partly due to drought (short- and long-term) as well as spring/summer weather in the Rocky Mountains. Climate can be said to exert a population limiting effect on the RMP, mediated through breeding habitat constraints. Specifically, in periods of drought, breeding cranes are limited in the quality of nesting habitat and choose not to nest and/or are less successful at producing young that make it to their main migratory stopover when they do nest. These findings are likely not unique to sandhill cranes, but apply more generally to migratory, ground-nesting birds that also rely on palustrine and riparian wetlands for breeding.

Exploring variation in predictors (e.g. Figs 4–6) can lead to important general findings of how future climate scenarios may impact RMP sandhill crane recruitment, which decision-makers may find useful. Most important is that the drought in the central Rocky Mountains is generally expected to increase in severity (Dai 2011), likely due

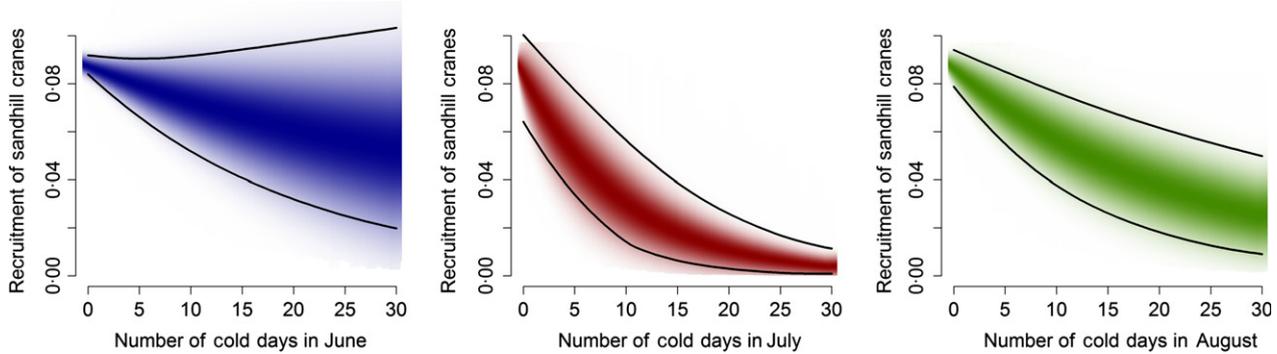


Fig. 5. Variation in predicted juvenile sandhill crane recruitment across the number of freezing days during the breeding season from an optimal out-of-sample predictive model, fit by ridge (M_{P1}). Drought indices are held constant at their mean. The uncertainty is shaded in proportion to the posterior predictive distribution and the lines are 95% credible intervals.

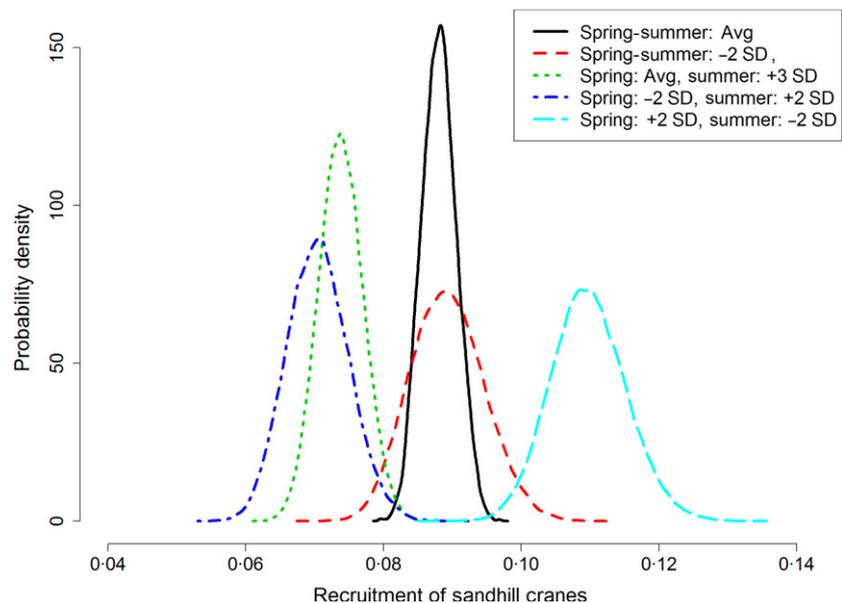


Fig. 6. Variation in predicted juvenile sandhill crane recruitment across alternative scenarios of maximum snow depth during the spring (March–April) and summer months (May–August) from an optimal out-of-sample predictive model, fit by ridge (M_{P1}). The number of freezing days within a month and drought predictors are held constant at their mean. SD indicates standard deviation such that $-2SD$ is 2 SD below the mean predictor value.

to increased temperatures and decreased spring snow cover (McKelvey *et al.* 2011). The timing of flows in the Rocky Mountains is also expected to occur earlier in the spring (Rood *et al.* 2008) with higher flows in the winter and substantially reduced flows in the summer. These changes may eventually cause a mismatch in timing of wetland recharge and water needs of nesting cranes, unless their migration chronology is sufficiently plastic to adapt. Our predictions suggest that future annual recruitment is likely to be lower than currently observed, thus limiting population growth more than in the past, unless survival rates could compensate. Limited population growth could be detrimental to the long-term stability of the RMP, as adult cranes have a natural mortality of $\approx 6\%$, which could be exceeded by harvest mortality under an increasing harvest rate (Drewien, R.C., unpublished data). Even relatively low harvest rates coupled with low recruitment could have an effect on the population's trajectory. If the population began to decline, the low expected recruitment coupled with delayed breeding would make for a very slow recovery.

MODELLING AND DATA LIMITATIONS AND CONCERNS

Some caution in using our model directly in harvest decisions is warranted. First, the juveniles we survey are those that survived through fledging and early sport hunting mortality and successfully migrated to the SLV. Thus, the recruitment we observe is likely lower than would be at the breeding area. Annual variation in the differential survival of juvenile and adults from the breeding grounds to the SLV may mask true variation in our estimates of recruitment to the population at the fall staging area in any given year. Juvenile and adult survival appear not to be highly temporally variable (Drewien, R.C., unpublished data); however, juvenile survival during this period is dependent on their body condition prior to migration (Drewien, R.C., unpublished data). Such variation across cohorts could occur as a result of changes in food availability, which could be exacerbated by environmental changes resulting from a changing climate. Climate predictors should capture the variation in survival effects on recruitment due to

climate effects, but not so for other temporally varying factors. Thus, this may not be a systemic issue but affect only a few years.

Secondly, the scale of monitoring recruitment has important implications of how we incorporate climate variables for the breeding area. Monitoring recruitment at the SLV has clear benefits in increasing the likelihood of obtaining a representative sample from the entire RMP and is less financially burdensome than a survey conducted over the entire breeding area. However, more informative measures of recruitment would result from monitoring at a finer grain than that of the entire RMP breeding area, because drought and local weather conditions vary considerably at a finer scale. Such trade-offs between sampling efficiency, scale and grain are likely common in studies of widely dispersed populations. Because fine-scale estimates of recruitment were not available, we used a modelling framework to spatially weight climate variation at the breeding area, based on relative abundance in migration staging areas, and then apply it to the population-level measure of recruitment. This is challenging because the distribution of the nesting cranes is unknown and has likely shifted over time due to habitat change, changing land practices, loss of wetlands and increasing exurban development. It gives us hope that we have generally captured the spatial variation of climate on recruitment in the population as our best predictive model used the annual pre-migratory counts (\mathcal{M}_{P1}), but we did expect the time-varying weighted model (\mathcal{M}_{P4}) to have more support. Perhaps, this was because we have no information on the distribution of cranes from 1972 to 1994.

Lastly, we have attempted to link drought with crane recruitment, where we assume one effect of drought is the reduction of wetland water levels, which increases a predator's ability to locate nests. Predation risk to crane nests is likely driven by predator abundance, activity and alternative prey availability (Austin, Henry & Ball 2007). These are in turn influenced by a number of factors, including land-use change and human activity, which have varied spatially and temporally within the Rocky Mountains over the last four decades. In the future, it may be possible to incorporate land-use change into this modelling framework. Work is currently underway by the Intermountain West Joint Venture to quantify the magnitude and distribution of landscape change (Donnelly & Vest 2006). However, this will not necessarily solve the issue of spatio-temporal variation in predation risk, which may remain imprecise over such a large area. A major benefit of the land-use data could be the ability to weight climate predictors by the availability of wetland habitats. This may allow crane recruitment to be used as a general index of overall wetland breeding productivity of certain waterfowl, shorebirds and wading birds in the Rocky Mountain region. However, this application would require testing.

PREDICTION AND MODEL SELECTION

We suggest the field of ecology could benefit from increased application of formal predictive modelling. By formal, we mean models trained by out-of-sample data and measured by a truly predictive criterion that does not depend on asymptotic properties (e.g. *lppd*), which will help identify generalizable predictive models. This is often considered the gold standard in discriminating among hypotheses (Kaplan 1964; Ashley, Granger & Schmalensee 1980). In addition, this approach is ideal for predicting system response over a range of realistic scenarios that includes unobserved or even novel conditions, to help decision-makers visualize possible outcomes of environmental change. Lastly, hypotheses of ecological or climate processes will often lead to a set of correlated variables, which a predictive modelling approach can easily accommodate.

We contend that viewing model selection within the domain of regularization is advantageous, because it opens up the wide array of interesting statistical approaches to fitting and comparing models (Hastie, Tibshirani & Friedman 2009; Hooten & Hobbs 2014). Regularization helps one to think about measures of model complexity and whether within-sample explanation or out-of-sample prediction is preferred. There is no one-size-fits-all regularization method; choosing an approach should be based on the research or management issue of interest, the assumptions of the structure of the data and the operating characteristics of the method (Reineking & Schröder 2006). If predicting beyond the current sample space is a priority, as it was here, it is preferable to use an out-of-sample measure of model fit, as it generalizes the model and helps guard against overfitting (Gelman, Hwang & Vehtari 2014). This is not to suggest prediction is inappropriate when optimizing a model based on within-sample fit, but if prediction is the goal, out-of-sample predictive performance is known to be worse (i.e. less accurate and/or precise) when fitting a model optimized using within-sample fitting (Hastie, Tibshirani & Friedman 2009). When there is sufficient data, the best practice is to use true out-of-sample predictive fit where a subset of the total data is completely left out and is never used to fit the model, only to evaluate predictive accuracy; otherwise, cross-validation as used here, which randomly leaves out a portion of the data to evaluate predictive accuracy is a valid option (Hastie, Tibshirani & Friedman 2009).

Regularization is commonly discussed and applied in ecological research, but mostly narrowly and focused in the likelihood framework using variable subset selection via information theoretics (IT, Burnham & Anderson 2002). The more general notion of statistical regularization being used for model selection is more rarely discussed in the ecological literature (Hooten & Hobbs 2014). One of the widely applied IT criterion, Akaike's Information Criterion (AIC), can easily be understood in the context of regularization, as $AIC = -2 \times \text{Log-Likelihood}$

+ $\lambda \times \sum_{p=1}^P |\beta_i|^q$, where $q = 0$ and $\lambda = 2$. AIC is based on minimizing within-sample expected prediction error (Burnham & Anderson 2002; Hastie, Tibshirani & Friedman 2009). Recommendations of using AIC include comparing a small set of ecologically motivated models (Burnham & Anderson 2002), which is ideal for an explanatory modelling approach. However, because variable selection is a discrete process ($q = 0$, predictor is either in or out), it can be unstable and exhibit high variance when it comes to prediction, and thus often leads to considerable predictive loss compared to continuous regularization methods where the amount of regularization (λ) is data-driven (e.g. by cross-validation) rather than fixed (Breiman 1996; Hastie, Tibshirani & Friedman 2009).

In this study, we applied the Bayesian LASSO and ridge, as they are known to perform well under a wide range of conditions, including, binary data, and especially at small sample size, such as ours (Reineking & Schröder 2006). A key issue in modelling binary data is the known overestimation of coefficients for small sample size, which is compounded when the event of interest is rare (small p_i , low recruitment) and when interested in many predictors (Makalic & Schmidt 2010); these issues commonly lead to poor prediction, which regularization can ameliorate. For our crane data, LASSO and ridge performed equally well in predictive accuracy and led to the same ecological findings. The differences between the two in their model specification are slight, but they have different operating characteristics that were not clearly illustrative in our study. Ridge is best when there are many nonzero coefficients and when there is strong collinearity among predictors, while LASSO operates best when there are many coefficients close to or at zero with a few larger coefficients and only some predictor collinearity (Tibshirani 1996; Reineking & Schröder 2006; Friedman, Hastie & Tibshirani 2010). For our analysis, if we had considered additional variables that were less motivated by clear ecological hypotheses and thus were more likely to be unrelated to recruitment, LASSO would have likely outperformed ridge by removing these variables from the optimal model; instead, the stronger collinearity among our important predictors lead to ridge outperforming LASSO.

Ecologists should consider their study goals and examine the wide range of regularization methods and their performance for their specific situation. We took a Bayesian approach to regularization, which can easily accommodate various sources of uncertainty, including parametric uncertainty, directly in our measure of predictive accuracy. We see a high utility in regularization specifically for meeting the increasing demands from decision-makers for increased predictive modelling in ecology. Further details on applying Bayesian model fitting can be found in the Supporting Information; for those interested in likelihood statistical regularization, see the R package 'glmnet' (Friedman, Hastie & Tibshirani 2010).

Acknowledgements

Funding was provided by the U.S. Fish and Wildlife Service Webless Migratory Game Bird Program and their Mountain-Prairie Migratory Bird Office. This research utilized the CSU ISTeC Cray HPC System supported by NSF Grant CNS-0923386. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the US Government. We are very grateful to L. Cordes, P. Doherty, D. Koons and an anonymous reviewer for providing helpful edits and comments on a prior version of this manuscript.

Data accessibility

Sandhill crane RMP recruitment data are provided in Table S1 (Supporting information). R code is provided in the Supporting Information that outlines the essential elements for conducting regularization and cross-validation.

References

- Ashley, R., Granger, C.W. & Schmalensee, R. (1980) Advertising and aggregate consumption: an analysis of causality. *Econometrica*, **48**, 1149–1167.
- Austin, J.E., & Pyle, W.H. (2004) Nesting ecology of waterbirds at Grays Lake, Idaho. *Western North American Naturalist*, **64**, 277–292.
- Austin, J.E., Henry, A.R. & Ball, I.J. (2007) Sandhill crane abundance and nesting ecology at Grays Lake, Idaho. *The Journal of Wildlife Management*, **71**, 1067–1079.
- Beckage, B., Gross, L.J. & Kauffman, S. (2011) The limits to prediction in ecological systems. *Ecosphere*, **2**, 125.
- Bickel, P.J. & Li, B. (2006) Regularization in statistics. *Test*, **15**, 271–344.
- Breiman, L. (1996) Heuristics of instability and stabilization in model selection. *The Annals of Applied Statistics*, **24**, 2350–2383.
- Brown, W.M. (2013) *October 2013 Recruitment Survey of the Rocky Mountain Population of Greater Sandhill Cranes*. Administrative Report, U.S. Fish and Wildlife Service, Albuquerque, New Mexico, USA.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach*. Springer, New York.
- Dai, A. (2011) Drought under global warming: a review. *Climate Change*, **2**, 45–65.
- Donnelly, P. & Vest, J.L. (2006) *Identifying Science Priorities 2013–2018: Wetland Focal Strategies*. Intermountain West Joint Venture Technical Series 2012–3. Intermountain West Joint Venture, Missoula, Montana, USA.
- Drewien, R.C. & Bizeau, E.G. (1974) Status and distribution of greater sandhill cranes in the Rocky Mountains. *The Journal of Wildlife Management*, **38**, 720–742.
- Drewien, R.C., Brown, W.M., & Kendall, W.L. (1995) Recruitment in Rocky Mountain greater sandhill cranes and comparison with other crane populations. *The Journal of Wildlife Management*, **59**, 339–356.
- Evans, M.R., Norris, K.J. & Benton, T.G. (2012) Predictive ecology: systems approaches. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **367**, 163–169.
- Friedman, J., Hastie, T. & Tibshirani, R. (2010) Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, **33**, 1–22.
- Gelman, A., Hwang, J. & Vehtari, A. (2014) Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, **24**, 997–1016.
- Gerber, B.D., Dwyer, J.F., Nesbitt, S.A., Drewien, R.C., Littlefield, C.D., Tacha, T.C. *et al.* (2014) Sandhill crane (*Grus canadensis*). *The Birds of North America Online* (ed. A. Poole), Cornell Lab of Ornithology, Ithaca. <http://bna.birds.cornell.edu/bna/species/031>
- Graham, M.H. (2003) Confronting multicollinearity in ecological multiple regression. *Ecology*, **84**, 2809–2815.
- Hastie, T., Tibshirani, R. & Friedman, J. (2009) *The elements of statistical learning*. Springer, New York.
- Heim, R.R. (2002) A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, **83**, 1149–1165.
- Hoerl, A.E. & Kennard, R.W. (1970) Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, **12**, 55–67.

- Hooten, M.B. & Hobbs, N.T. (2014) A guide to Bayesian model selection for ecologists. *Ecological Monographs*, **85**, 3–28.
- Kaplan, A. (1964) *The conduct of inquiry: methodology for behavioral science*. Chandler Publishing, New York.
- Kendall, W.L. (2001) Using models to facilitate complex decisions. *Modeling in Natural Resource Management* (eds T.M. Shenk & A.B. Franklin), pp. 147–170. Island Press, Washington, District of Columbia.
- Kruse, K.L., Dubovsky, J.A. & Cooper, T.R. (2014) *Status and Harvests of Sandhill Cranes: Mid-Continent, Rocky Mountain and Lower Colorado River Valley Populations*. Administrative Report, U.S. Fish and Wildlife Service, Lakewood, CO, USA.
- Lazaridis, A. (2007) A note regarding the condition number: the case of spurious and latent multicollinearity. *Quality & Quantity*, **41**, 123–135.
- Mac Nally, R. (2000) Regression and model-building in conservation biology, biogeography and ecology: the distinction between-and reconciliation of predictive and explanatory models. *Biodiversity & Conservation*, **9**, 655–671.
- Makalic, E. & Schmidt, D.F. (2010) Review of modern logistic regression methods with application to small and medium sample size problems. *Advances in Artificial Intelligence 3rd Australasian Joint Conference* (ed. L. JiuYong), pp. 213–222. Springer-Verlag, Berlin.
- McKelvey, K.S., Copeland, J.P., Schwartz, M.K., Littell, J.S., Aubry, K.B., Squires, J.R., Parks, S.A., Elsner, M.M. & Mauger, G.S. (2011) Climate change predicted to shift wolverine distributions, connectivity, and dispersal corridors. *Ecological Applications*, **21**, 2882–2897.
- McWethy, D.B. & Austin, J.E. (2009) Nesting ecology of greater sandhill cranes (*Grus canadensis tabida*) in riparian and palustrine wetlands of eastern Idaho. *Waterbirds*, **32**, 106–115.
- Mishra, A.K. & Singh, V.P. (2010) A review of drought concepts. *Journal of Hydrology*, **391**, 202–216.
- Nichols, J.D., Runge, M.C., Johnson, F.A. & Williams, B.K. (2007) Adaptive harvest management of North American waterfowl populations: a brief history and future prospects. *Journal of Ornithology*, **148**, 343–349.
- Pearl, J. (2014) Comment: Understanding Simpsons Paradox. *The American Statistician*, **68**, 8–13.
- R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R foundation for statistical computing, Vienna, Austria, <http://www.R-project.org>.
- Reineking, B. & Schröder, B. (2006) Constrain to perform: regularization of habitat models. *Ecological Modelling*, **193**, 675–690.
- Rood, S.B., Pan, J., Gill, K.M., Franks, C.G., Samuelson, G.M. & Shepherd, A. (2008) Declining summer flows of Rocky Mountain rivers: changing seasonal hydrology and probable impacts on floodplain forests. *Journal of Hydrology*, **349**, 397–410.
- Sæther, B.E. & Bakke, O. (2000) Avian life history variation and contribution of demographic traits to the population growth rate. *Ecology*, **81**, 642–653.
- Shmueli, G. (2010) To explain or to predict. *Statistical Science*, **25**, 289–310.
- Simpson, E.H. (1951) The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **25**, 238–310.
- Sutherland, W. (2006) Predicting the ecological consequences of environmental change: a review of the methods. *Journal of Applied Ecology*, **43**, 599–616.
- Thorpe, P.P., Donnelly, P., & Collins, D. (2013) *September 2013 survey of the Rocky Mountain Population of greater sandhill cranes*. Administrative report, U.S. Fish and Wildlife Service, Division of Migratory Bird Management, Lakewood, Colorado.
- Tibshirani, R. (1996) Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, **58**, 267–288.
- Vicente-Serrano, S.M., Beguera, S., Lopez-Moreno, J.I., Angulo, M. & El Kenawy, A. (2010a) A New Global 0.5 Gridded Dataset (1901–2006) of a Multiscalar drought index: comparison with current drought index datasets based on the palmer drought severity index. *Journal of Hydro-meteorology*, **11**, 1033–1043.
- Vicente-Serrano, S.M., Beguera, S. & López-Moreno, J.I. (2010b) A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate*, **23**, 1696–1718.

Received 7 January 2015; accepted 3 March 2015
 Handling Editor: Brett Sandercock

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Additional details on NOAA climate regions, sandhill crane sampling, Bayesian model fitting, statistical regularization, and cross validation.

Table S1. Annual survey data used to develop an optimal out-of-sample predictive model of juvenile sandhill crane recruitment.

Table S2. Ecological hypotheses, predictor variables, and evidence of the relationship (P(Hypothesis|Data)) of how drought (Palmer indices) and weather predictors influence sandhill crane juvenile recruitment.

Table S3. Ecological hypotheses, predictor variables, and evidence of the relationship (P(Hypothesis|Data)) of how drought (SPEI) and weather predictors influence sandhill crane juvenile recruitment.

Fig. S1. Animation of the proportion of the annual pre-migration count of the Rocky Mountain population of sandhill cranes on their staging areas (1992, 1995 to 2013); spatial units are NOAA climate regions.

Fig. S2. Drought indices from 1972 to 2013 within the sandhill crane Rocky Mountain Population breeding area (Idaho, Montana, Wyoming, Colorado, Utah); Top: the standardized precipitation evapotranspiration index (SPEI). Bottom: the Palmer Drought Severity Index (PDSI).

Fig. S3. Illustration of the continuous effect of model complexity under different forms.

Fig. S4. Comparison of the Laplace and Normal distribution, used to constrain model complexity in Bayesian regularization, where prior specification of coefficients are equivalent to LASSO and ridge regression, respectively.

Fig. S5. Out-of-sample cross validation (CV) log-predictive scores (E[lppd]) under different amounts of training data (CV1, CV2, CV3, CV6) for identifying an optimal regulator (λ).

Fig. S6. Animation of the posterior distribution of the log predictive score across a range of regulators (λ); the thick black line indicates the optimal or highest expected value.

Fig. S7. Trace plots of the standardized regression coefficients (E [λ]) when using ridge regression and LASSO.

Fig. S8. Observed variation in juvenile recruitment of the Rocky Mountain Population of sandhill cranes from 1972 to 2013 during migration at the San Luis Valley in southern Colorado.

Appendix S2. R code demonstrating the structure of how to carry out statistical regularization, cross-validation, and calculations of the lppd.