LETTER • OPEN ACCESS

A predictive model for seasonal pond counts in the United States Prairie Pothole Region using large-scale climate connections

To cite this article: Benjamin D Abel *et al* 2020 *Environ. Res. Lett.* **15** 044019

View the article online for updates and enhancements.

Environmental Research Letters

A predictive model for seasonal pond counts in the United States Prairie Pothole Region using large-scale climate connections

Benjamin D Abel[1,4] , Balaji Rajagopalan[1,2] and Andrea J Ray[3]

[1] Department of Civil, Environmental, and Architectural Engineering, University of Colorado Boulder, Boulder, CO, United States of America
[2] Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, United States of America
[3] National Oceanic and Atmospheric Administration, Earth System Research Laboratory, Physical Sciences Division, Boulder, CO, United States of America
[4] Author to whom any correspondence should be addressed.

E-mail: benjamin.abel@colorado.edu

Abstract

The Prairie Pothole Region (PPR), located in central North America, is an important region hydrologically and ecologically. Millions of wetlands, many containing ponds, are located here, and they serve as habitats for various biota and breeding grounds for waterfowl. They also provide carbon sequestration, sediment and nutrient attenuation, and floodwater storage. Land modification and climate change are threatening the PPR, and water and wildlife managers face important conservation decisions due to these threats. We developed predictive, multisite forecasting models using canonical correlation analysis (CCA) for pond counts in the southeast PPR, the portion located within the United States, to aid in these important decisions. These forecast models predict spring (May) and summer (July) pond counts for each region (stratum) of the United States Fish and Wildlife Service's pond and waterfowl surveys using a suite of antecedent, large-scale climate variables and indices including 500 millibar heights, sea surface temperatures (SSTs), and Palmer Drought Severity Index (PDSI). Models were developed to issue forecasts at the start of all preceding months beginning on March 1st. The models were evaluated for their performance in a predictive mode by leave-one-out cross-validation. The models exhibited good performance (*R* values above 0.6 for May forecasts and 0.4 for July forecasts), with performance increasing as lead time decreased. This simple and versatile modeling approach offers a robust tool for efficient management and sustainability of ecology and natural resources. It demonstrates the ability to use large-scale climate variables to predict a local variable in a skilful way and could serve as an example to develop similar models for use in management and conservation decisions in other regions and sectors of the environment.

1. Introduction

Encompassing nearly 800 000 km^2 in North America [1], the Prairie Pothole Region (PPR) is important hydrologically and ecologically. It is the location of over 2.5 million wetlands [2] which are located in depressions, or potholes, left from the recession of the Laurentide ice sheet in the Pleistocene epoch [3]. It is the largest wetland complex in North America [4]. The depressions sometimes contain ponds or lakes, ranging in permanence from ephemeral to persistent, that

serve numerous environmental purposes. Colloquially called the 'duck factory', it is a favorable habitat for the yearly breeding of over 50% of the North American duck population during spring and early summer [5]. Other biota, such as insects and riparian vegetation, also call these potholes their home [1, 5, 6]. In addition to being a very important habitat for various forms of life, the wetlands also provide carbon sequestration, sediment and nutrient attenuation, and floodwater storage [7].

Wildlife managers in the PPR must develop and implement hunting and conservation rules and regulations and actions to preserve or maintain habitat. These decisions become more complicated in the face of threats to the ecosystems they manage. Near-term threats include land modification due to agriculture or oil/gas well production and climate variability, and longer-term threats include climate change. Managers must respond to both, though near-term threats typically take priority for addressing within agency resource constraints [8]. Complicating matters further, much of the land in the United States portion of the PPR is privately owned [9]. Thus, managers must work closely with the private land owners and convince them to enter conservation easements, typically allocated by the Conservation Reserve Program. Additionally, the program is losing federal funding, further hampering the process of allocating easements.

Agricultural modifications (e.g. tiling) and energy development (e.g. oil/gas wells) are affecting the PPR and have reduced the ecological productivity of the region. These anthropogenic modifications can diminish or drain wetlands entirely, alter their landscape, or introduce contaminants. Between 16% and 18% of the PPR was historically (circa 1850) covered by wetlands, but up to 65% had been drained by the mid-1980s, primarily due to agriculture [10]. Total wetland area declined between 1997 and 2009 by another 1.1%, or approximately 30 100 ha [2]. Wright and Wimberly [11] found that there was a net loss of 528 ha of grassland from 2006 to 2011. Agricultural modifications change the landscape, altering soil permeability, runoff pathways, and evapotranspiration [12]. A United States Geological Survey evaluation found that 34 of 48 water samples from the PPR were contaminated with brine [13]. These modifications decrease pond and wetland numbers and introduce contaminants, therefore decreasing the area for viable habitat of dependent species.

Average temperature and precipitation are increasing in the region due to climate change [14], and the region is experiencing much climate variability on both spatial and temporal scales [14–17]. The region is highly sensitive to annual variation in weather conditions, with these playing an integral role in the hydrologic productivity of the region [18–20]. Summer rainfall in the southeast PPR is extremely important because it can contribute approximately half of the annual precipitation [16, 21]. Summer rainfall in the southeast PPR is projected to decrease [22] therefore increasing the danger this region faces.

In the face of these multiple threats, managers in the region must make informed choices on actions to protect and manage wetlands and implement recreational regulations. Several conservation strategies are undertaken to maintain and improve habitat, both ponds and surrounding grasslands, in order to maximize waterfowl nesting success, and are timed within a season to ensure that treatments do not adversely affect migratory birds [23]. For example, light to moderate grazing treatments are performed to ensure proper nesting habitat for migratory waterfowl [23, 24], and are conducted after 1 June to ensure waterfowl are not disturbed by livestock [23, 25]. Similarly, noxious weed control is administered to prevent further degradation of grasslands and wetlands [23], and haying/prescribed burns are executed after 1 August each season [23]. These decisions are made while considering near-term threats, longer-term threats, and agency resource constraints.

These management responses are on two distinct time scales. First, there are month-to-month/year-to-year actions that may include habitat restoration/revegetation or management of invasive plants that often respond to seasonal variability in temperature and precipitation. Some decisions, such as determining hunting quotas, use information on prior year hunting totals in addition to upcoming season forecasts. Managers often make decisions early in the year on whether and where to implement these and other actions, making the best use of their resources. Second, there are longer-range planning decisions with effects over multiple years, such as easements, intended to protect wetlands in the face of land use conversion or agreements with landowners to apply treatments such as those described above.

The focus of this work is on predictive models that may help managers make decisions earlier in the year regarding what actions to choose by providing longer lead times on a sub-PPR scale in regions managers are accustomed to working with. These decisions will therefore be based on a better model of how observable, predictable atmospheric and hydroclimate metrics relate to pond habitats.

We, therefore, develop predictive models that forecast pond counts on a sub-PPR scale using a novel, reduced-dimension approach based on principal component analysis (PCA) and canonical correlation analysis (CCA). These models have varying lead times back to 1 March. Other available predictive models make a single, regional pond count forecast, do not include large-scale climate variables, make current-year forecasts with only weeks of lead time, or some combination thereof [26–28]. Multiple linear regression was used by Larson [26] to determine the amount of pond count variability accounted for by local climate variables. The model used a suite of 10 predictors: previous year's pond counts plus seasonal and yearly averages of minimum/maximum temperature and precipitation. However, most of these variables were not available until May meaning forecasts would be issued at that time. Sorenson *et al* [28] used simple linear regression on pond and duck counts with concurrent May Palmer Drought Severity Index (PDSI). Forecasts with this model can be issued after May PDSI becomes available—approximately early June. In a more recent study, multiple linear regression was used on May pond count numbers from sub-regions
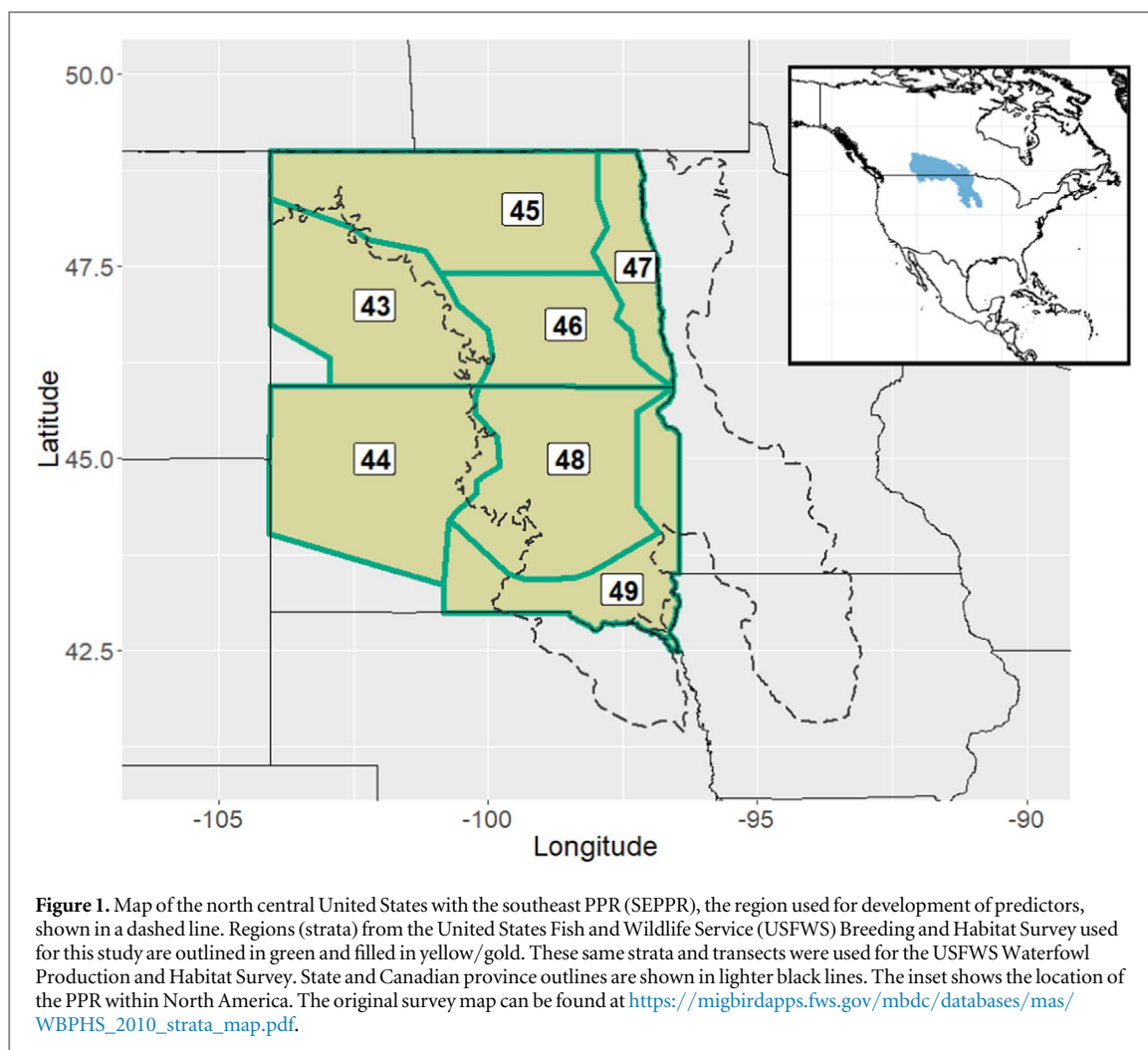
Figure 1. Map of the north central United States with the southeast PPR (SEPPR), the region used for development of predictors, shown in a dashed line. Regions (strata) from the United States Fish and Wildlife Service (USFWS) Breeding and Habitat Survey used for this study are outlined in green and filled in yellow/gold. These same strata and transects were used for the USFWS Waterfowl Production and Habitat Survey. State and Canadian province outlines are shown in lighter black lines. The inset shows the location of the PPR within North America. The original survey map can be found at https://migbirdapps.fws.gov/mbdc/databases/mas/ WBPHS_2010_strata_map.pdf.

Table 1. Datasets utilized in this study, the variables they provided, and the years they covered. See text for acronym definitions.

Dataset/source	Variable(s)	Temporal domain
USFWS Waterfowl Breeding and Habitat Survey (May)	Pond count, waterfowl population	1955–2015
USFWS Waterfowl Production and Habitat Survey (July)	Pond count, waterfowl population	1957–2003
IRI Data Library	Sea surface temperature (SST) anomalies	1896–2017
IRI Data Library	500 millibar height (500 mb) anomalies	1949–2017
NOAA NCEI nClimDiv	Palmer Drought Severity Index (PDSI)	1895–2017

of the PPR [27]. Variables included the preceding year's temperature and precipitation (May–April), the previous year's pond counts, a temperature and precipitation interaction term, and location (longitude and latitude). Forecasts with this model can be issued in early May. We seek to improve on these forecast models by developing predictive models using a novel approach not used previously for this purpose; the models are informed by a suite of predictors including large-scale climate indices, have lead times beginning 1 March (therefore providing longer lead time), and forecast pond counts on a sub-PPR scale.
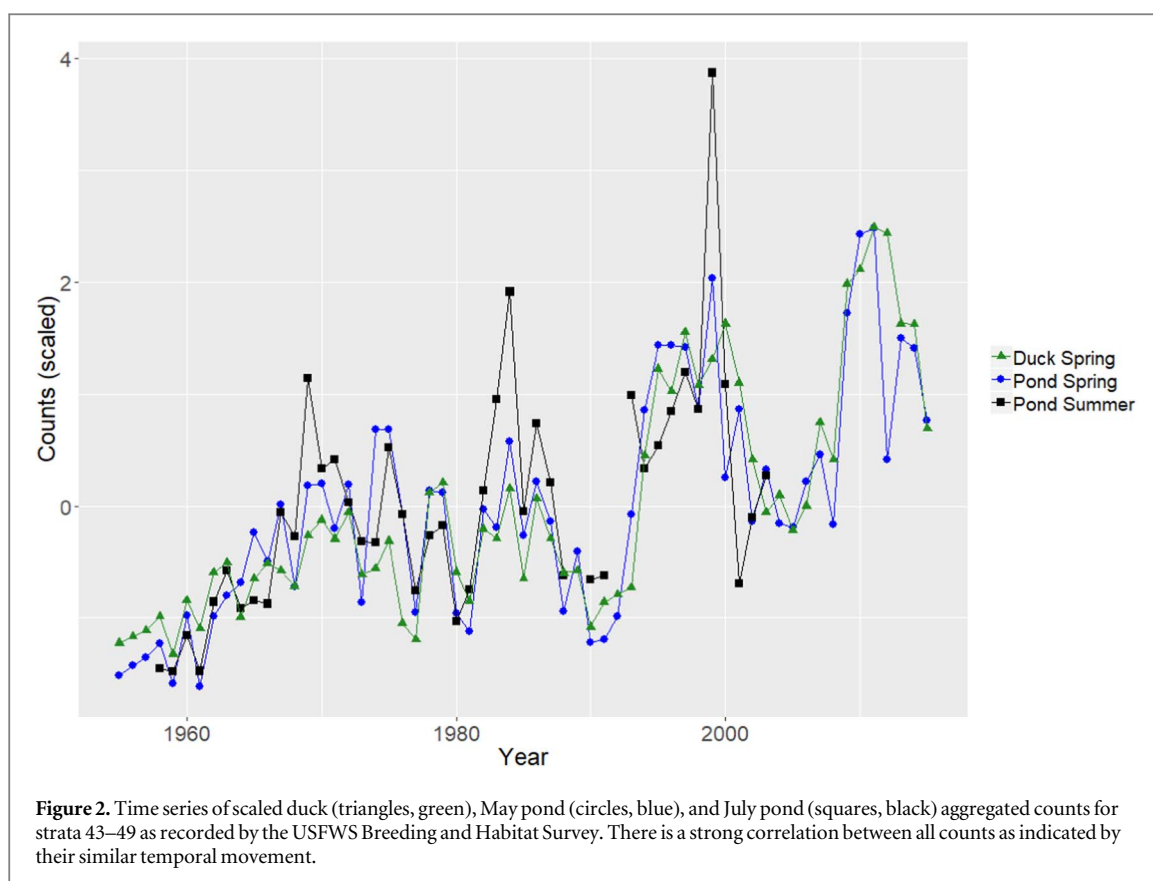
We begin by outlining the study region and data incorporated in the work. Following that, we discuss the methods involved in creating predictors, developing our predictive models, and evaluating model performance. We conclude by providing results, comparing them to other similar models, and discussing their relevance for decision-making.

2. Study region and data

This study focuses on the southeast PPR [22] located within the United States (SEPPR). We exclude the portion that extends into Manitoba, Canada. This region is drawn in a dashed line in figure 1. Data were retrieved from various sources for this study (table 1). We prioritized use of temporally concurrent and spatially conterminous data.

Data for pond counts and duck populations were retrieved from two surveys conducted by the United States Fish and Wildlife Service (USFWS). In May, the

Figure 2. Time series of scaled duck (triangles, green), May pond (circles, blue), and July pond (squares, black) aggregated counts for strata 43–49 as recorded by the USFWS Breeding and Habitat Survey. There is a strong correlation between all counts as indicated by their similar temporal movement.

Waterfowl Breeding and Habitat Survey is performed; in July, the Waterfowl Production and Habitat Survey was performed until 2003. Data for the May survey span from 1955-present; data for years up to 2015 are used for this study. Data for the July survey span from 1958 to 2003; all years of survey data are used for this study. The survey consists of multiple subregions (strata), and strata 43–49 were determined to be spatially coincident with the SEPPR (figure 1).

Sea surface temperature (SST) anomalies and 500 millibar height (500 mb) anomalies were retrieved from the International Research Institute's (IRI) Data Library which compiles raw data from various sources into a common format. The SST anomaly data are a concatenation of Kaplan *et al* [29] and Reynolds and Smith [30]. The 500 mb anomaly data are from National Oceanic and Atmospheric Administration's (NOAA) Climate Data Assimilation System I (CDAS-1) data.

We retrieved PDSI data spanning from 1895 to 2017 from the NOAA National Centers for Environmental Information (NCEI) nClimDiv dataset [31]. The data were retrieved on a monthly scale by climate division. We chose 22 climate divisions to represent the SEPPR because they were spatially coincident.

May pond counts and duck population correlate strongly (linear correlation has possible values of −1 to 1) with a linear correlation (Helsel and Hirsch [32], Montgomery and Runger [33]) of 0.87 (figure 2). We can infer duck population with pond count because of this strong correlation, a connection that has been

well-documented [22, 34, 35]. July pond counts also have high linear correlation with May pond counts (0.78, figure 2) allowing us to extend our modeling framework to predict July pond counts.

This connection between variables, however, comes with stipulations. Other factors such as wetland-cover type, emergent vegetation distribution and characteristics, and interactions among ducks can cause ducks to relocate [36]. Salinity of ponds also plays a role, because the plant communities are sensitive to the level of brine in ponds [37]. We acknowledge these confounding factors exist, but rely on the strong relationship for inferences.

Ponds and ducks have shown an increasing trend in the SEPPR (figure 2). Ponds in the SEPPR are increasing due largely to the increase in precipitation in the PPR [14] which is in part due to increases in the SEPPR in all seasons but summer [22]. Also attributable are conservation efforts over the recent decades. Duck counts have followed due to the observed relationship. Variability has increased slightly since the beginning of the observed period. The cause for this is an ongoing research area. Some potential explanations include the increase in extreme events due to climate change and effects due to land use conversion. Variability is expected to continue to increase in the region [38].

3. Methods

When forecasting one field of data with another, CCA [39] is a common approach. It has been used for multisite forecasting of precipitation (e.g. [40]), temperature (e.g. [40], [41]), streamflow (e.g. Salas *et al* 2010), SSTs (e.g. [42]), and wind (e.g. [43]). In this study, we use CCA to predict the field of pond counts using a suite of developed predictors. First, we describe how the suite of predictors was developed. This is followed by a description of the CCA process used to develop our predictive models. A more detailed description of these processes can be found in appendices A and B. We conclude this section with details on how the models were evaluated on both fit and skill.

3.1. Development of predictors

Pond count is physically linked to the amount of water available—a function of groundwater storage, direct precipitation, and snowmelt runoff [18]. This direct connection to precipitation motivated exploratory data analysis of the region's precipitation (per climate division). This analysis provided evidence that SEPPR summer precipitation variability is driven by 500 mb systems and SSTs in the Pacific and Atlantic Ocean, and it has strong connections to PDSI. We performed PCA [44] on the climate division summer precipitation data to find the leading modes of variability. These modes were then found to be correlated to SST and 500 mb anomalies thus establishing connections to these large-scale variables. It was also observed that PDSI was correlated to the leading modes. Connections were found during wet and dry years as determined by the leading mode. Informed by this, we use these three variables, developed into indices, as predictors for use in our predictive models.

Reference [28] based their models on PDSI because of its ability to serve as a single composite measure of dryness. It is a widely used index which estimates relative dryness based on current and antecedent temperature and precipitation [45]. In our models, it represents past local climate conditions. Large-scale climate features such as 500 mb and SST persist seasonally, thereby representing future conditions.

We used regions of highest magnitude linear correlation (Helsel and Hirsch [32], Montgomery and Runger [33]) between May pond counts and space-time fields of 500 mb and SST anomalies to develop eight predictors (three SST, five 500 mb). The fields were for the preceding winter (December–February), March, and April. The high correlation between May and July pond counts (figure 2) allowed us to use the same predictors along with May pond counts for July models.

Regional PDSI values from 22 climate divisions representing the SEPPR were averaged to one value per year for winter, March, and April resulting in three PDSI predictors.

3.2. Predictive model

For our predictive model, PCA [44] is performed first on the two datasets (pond counts and predictors); both are scaled (by subtracting the mean and dividing by the standard deviation) prior to PCA. It is frequently used alongside CCA as a dimension-reducing technique because, typically, datasets used are large. Though our datasets are not large, we include PCA to allow managers to add more strata or predictors making our framework versatile.

The July survey contains missing values, across all strata for some years (e.g. 1989) and for select strata during specific years (e.g. 1958). Years in which data for all strata were missing were removed for figure 2. For CCA, years with missing values in at least one stratum were removed.

After performing PCA, a user-defined number of PCs (N_{pc}) are retained for CCA. Then, CCA is performed on the selected N_{pc} to provide N_{pc} canonical component (CCs) pairs. Then, a regression is performed with the sets of CCs allowing prediction of N_{pc} from one dataset using the other. Finally, predicted values (scaled) can be found by multiplying the eigenvalues (from PCA) for the appropriate dataset by the predicted PCs to back transform then rescaled as desired. For plotting purposes (figures 4, 6–8), they were left scaled.

Pond count was chosen as the response variable because it is dependent on physical processes thus lowering the potential for confounding and hidden variables. Duck population, however, is more dependent on biological processes in addition to environmental conditions therefore increasing the confounding factors. The connection between the two variables (figure 2) allows managers to infer duck population based on pond count.

Models were developed forecasting both May and July pond counts for individual strata at the start of all preceding months beginning on 1 March. Predictors were included based on availability related to forecast date. For example, forecasts made on 1 April cannot include predictors developed using concurrent April SST. This resulted in a smaller suite of predictors for earlier forecasts (longer lead time).

3.3. Model performance

Model fit and forecast skill are measured by calculating the correlation coefficient, *R*, and root mean square error, RMSE (Helsel and Hirsch [32], Montgomery and Runger [33]). These two metrics, commonly used to evaluate regression models, were chosen because they provide a measure of contiguity between predicted and observed values—RMSE is a measure of difference between the two, and *R* measures how closely the two vary together. For model fit, all data are
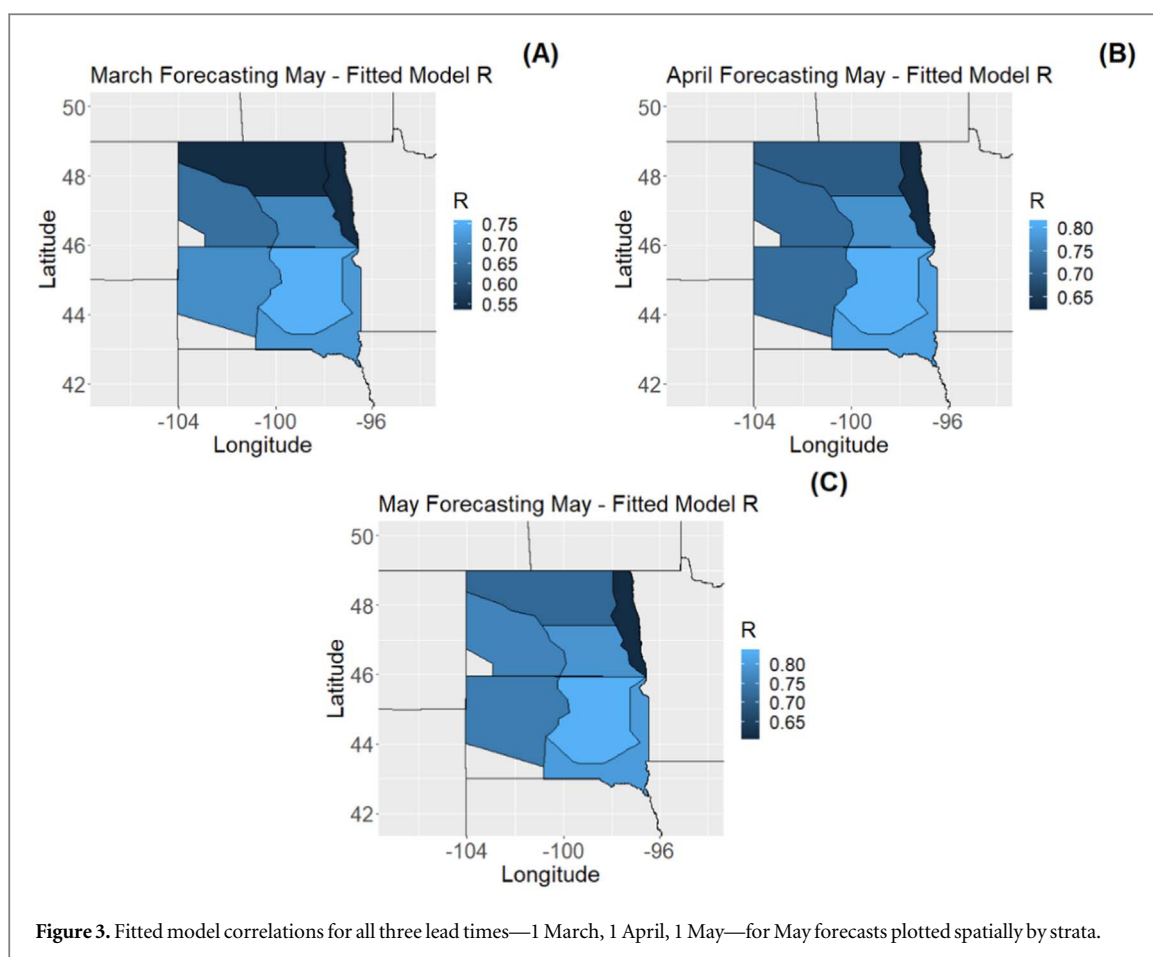
Figure 3. Fitted model correlations for all three lead times—1 March, 1 April, 1 May—for May forecasts plotted spatially by strata.

used to construct the model, predictions are made for all years, and the observed values are compared to the modeled values using R and RMSE. To test model performance in a predictive mode (i.e. model skill), leave-one-out cross-validation (LOOCV) was used.

In LOOCV, one of the observations is dropped. The model is then fitted to the remaining observations and the dropped value is predicted. We performed this process chronologically, starting with the first year of observations then ending with the last year, similar to jackknife [46]. However, jackknife is typically used for parameter estimation. R and RMSE were then calculated using all the LOOCV predictions and the observed values.
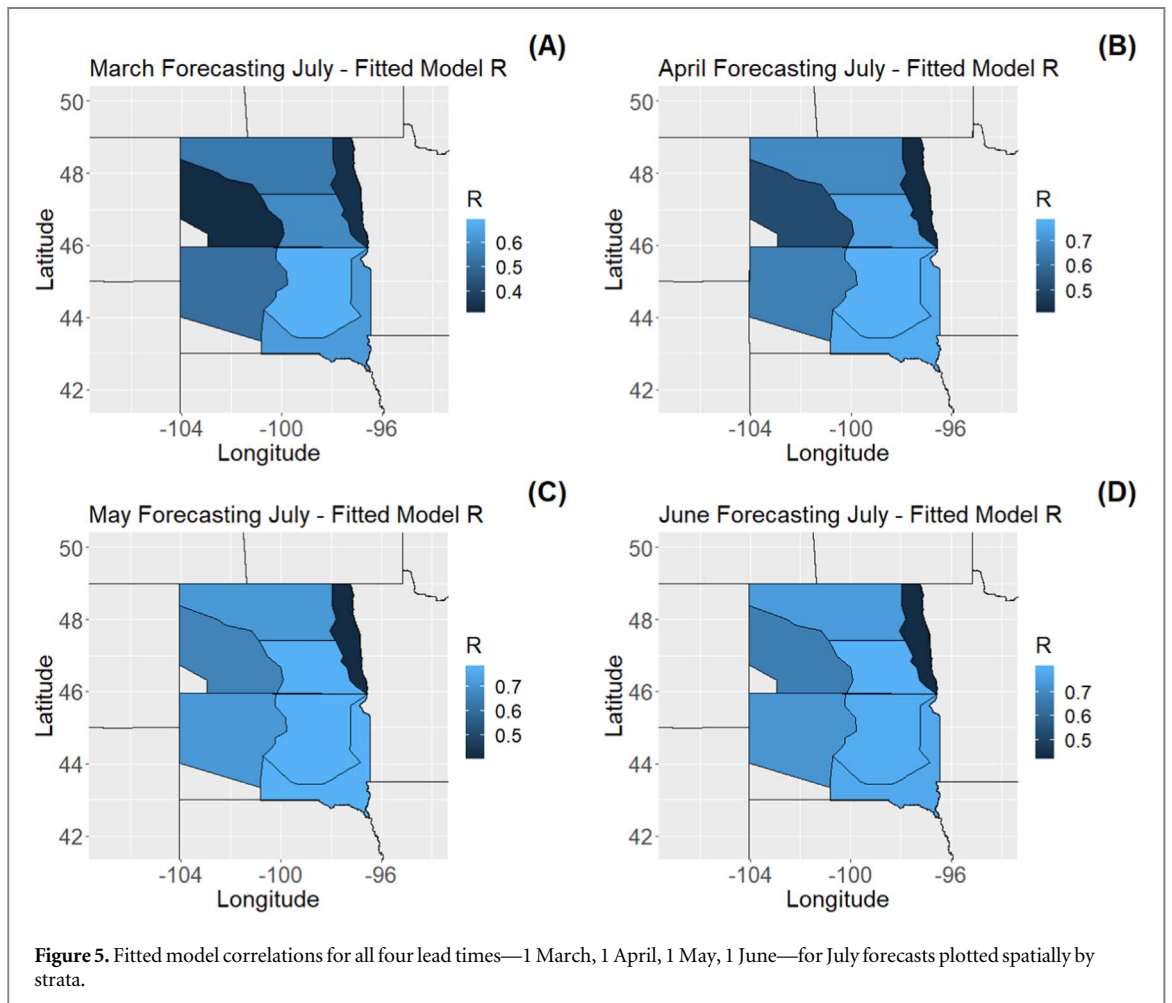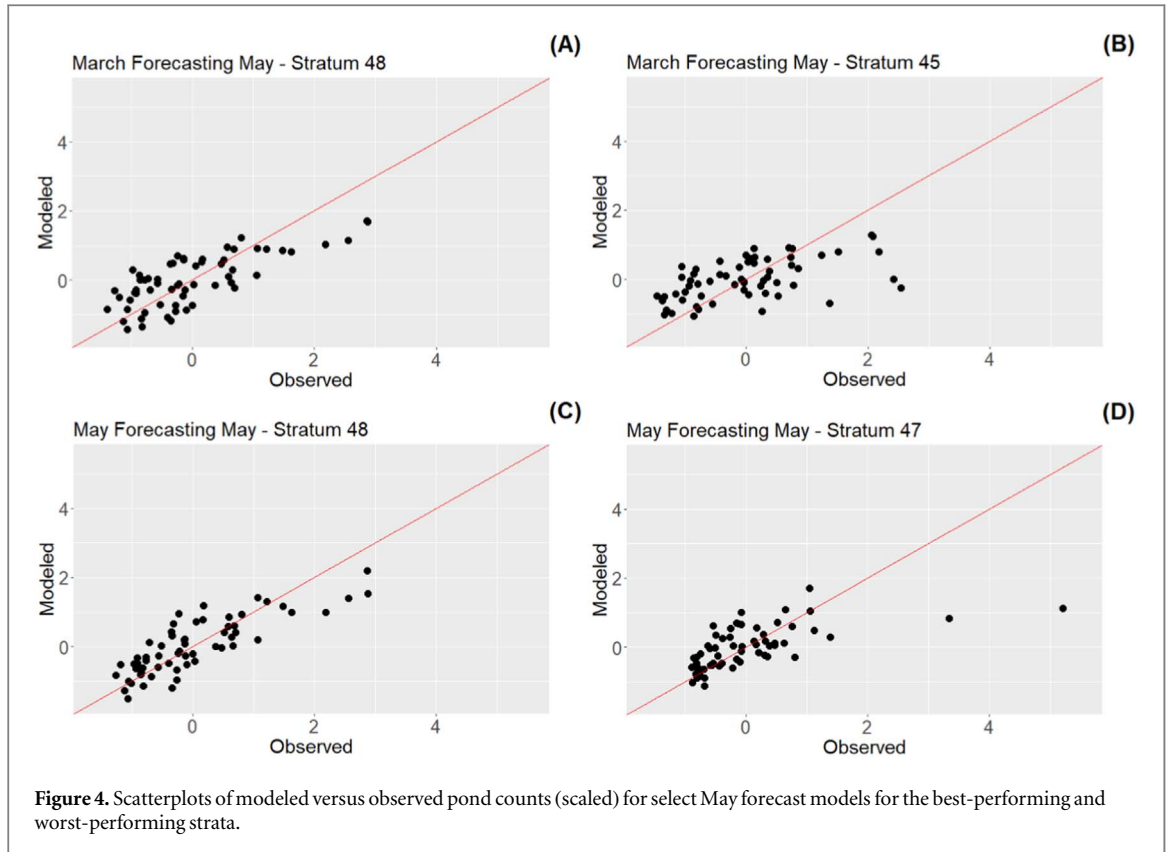
4. Results

4.1. Model fit
In general, models fit well for most strata. For the May pond count models, most R values were above 0.65, even when forecasting in March (figure 3). The R values for the July forecast models were mostly above 0.5 including models with the longest lead time (figure 5). Strata with models that fit well were close to the 1:1 line when plotting observed versus modeled values (figures 4(a), (b), 6(a) and (b)) whereas the poorly fitting strata displayed obvious outliers (figures 4(c), (d), 6(c) and (d)). Stratum 47, located in

eastern North Dakota, fit worst for most models (figures 3, 4(d), 5, 6(d)), and its fit was disparately low in comparison. Stratum 45 (figures 3(a), 4(b)) and 43 (figures 5(a) and 6(b)), located in northern and western North Dakota, fit poorly in May and July forecast models, respectively, when forecasting in March, but fit improved for the successive lead times. For most strata, R increased and RMSE (not shown) decreased with shorter lead time. The only exception was stratum 47 in May forecast models which had static fit metrics from April to May lead times.

4.2. Cross-validation forecast skill
Model performance in LOOCV mode was similar to fitted model results with a minor reduction in performance. For May pond count forecasts, most R values were above 0.6 and RMSE is below 1 (equivalent to 1 standard deviation since the data is scaled), including forecasting in March (figure 7). The R values for the July forecast models were mostly above 0.4, and RMSE was mostly below 1 (figure 8). Stratum 47 again performed weaker than other strata, and its forecast skill was low regardless of lead time. Strata 45 and 43 performed poorly in May and July forecast models, respectively, but forecast skill improved with lead time. As with model fit, R increased and RMSE decreased with shorter lead time for most strata.

Figure 4. Scatterplots of modeled versus observed pond counts (scaled) for select May forecast models for the best-performing and worst-performing strata.

Figure 5. Fitted model correlations for all four lead times—1 March, 1 April, 1 May, 1 June—for July forecasts plotted spatially by strata.

Figure 6. Scatterplots of modeled versus observed pond counts (scaled) for select July forecast models for the best-performing and worst-performing strata.

Figure 7. Correlation (*R*) and Root Mean Square Error (RMSE) values for each lead time for May forecast models in cross-validation mode separated by strata.

Figure 8. *R* and RMSE values for each lead time for July forecast models in cross-validation mode separated by strata.
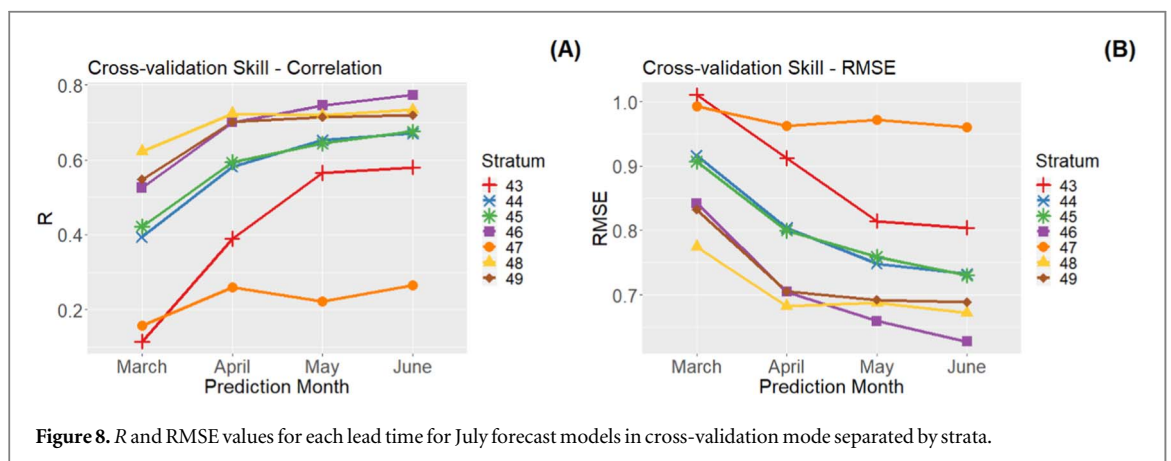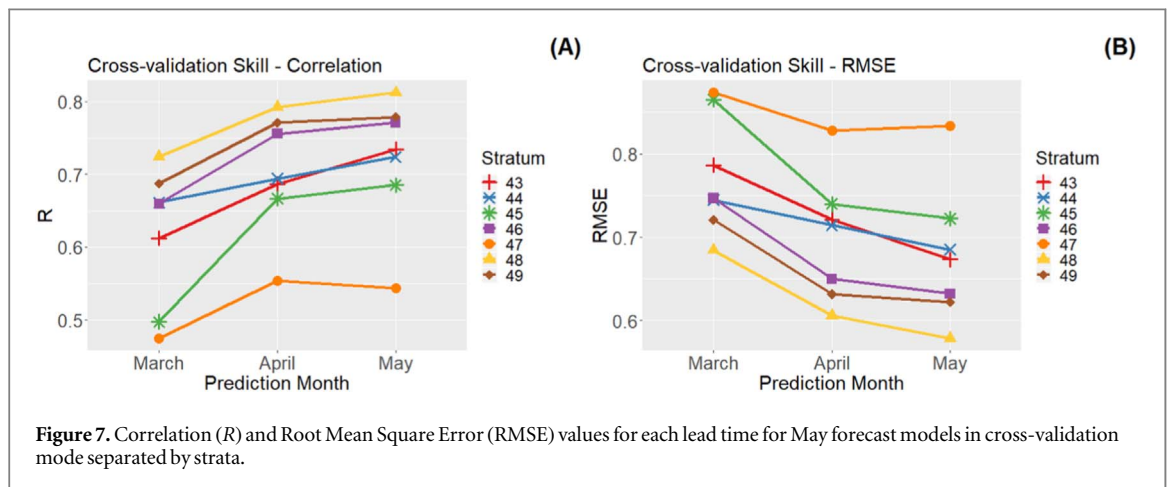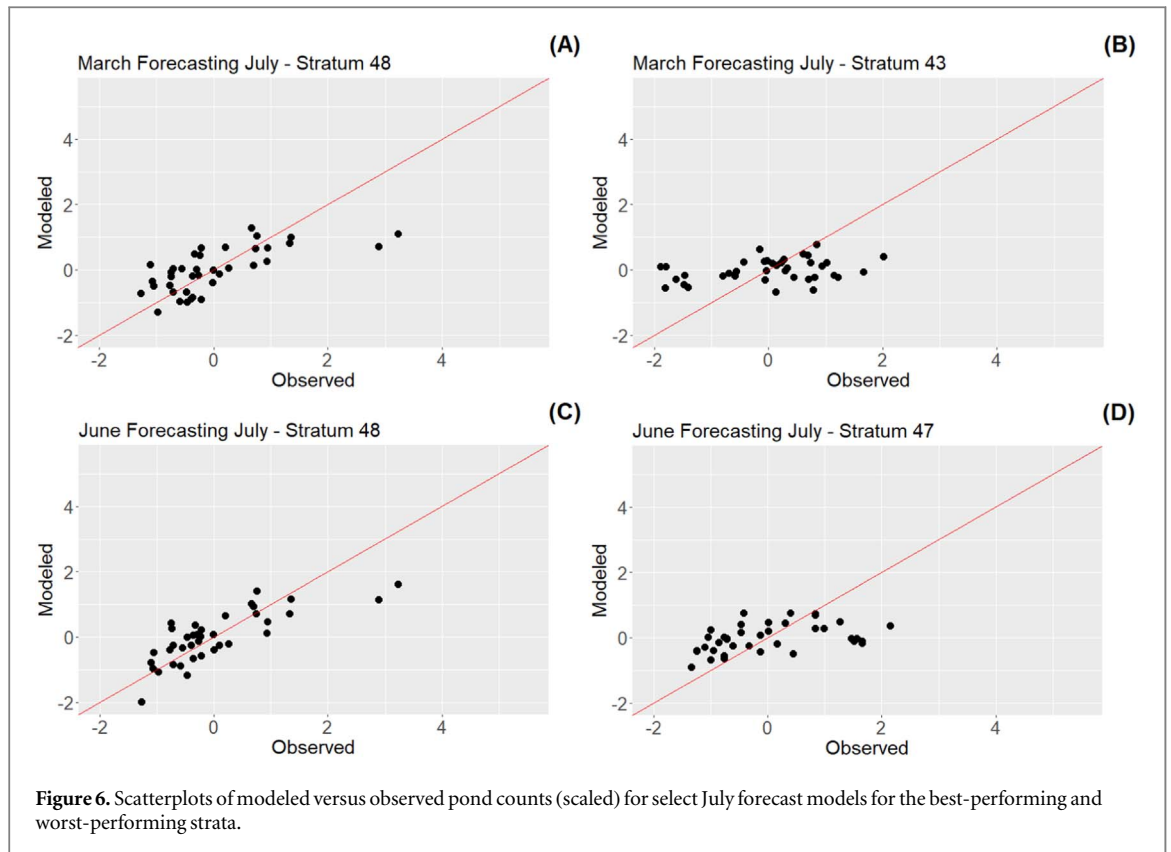
Table 2. Coefficient of determination, R^2, values for all models and all strata.

	Stratum 43	Stratum 44	Stratum 45	Stratum 46	Stratum 47	Stratum 48	Stratum 49
				May forecasts			
March	43%	46%	24%	42%	29%	52%	47%
April	52%	52%	49%	60%	39%	66%	63%
May	58%	56%	51%	62%	37%	69%	64%
				July forecasts			
March	10%	27%	29%	35%	11%	47%	40%
April	26%	43%	46%	55%	17%	60%	58%
May	44%	51%	51%	60%	17%	60%	60%
June	45%	53%	56%	64%	18%	62%	61%

Exceptions include stratum 47 in May models and strata 47 and 48 in the July models.

4.3. Comparison to other similar models

Our models provide longer lead times, forecast for multiple strata, include large-scale variables, and exhibit similar performance when compared to other pond forecast models available. A common measure of model skill is the variation explained by the model, measured by the coefficient of determination, R^2 (Helsel and Hirsch [32], Montgomery and Runger [33]). Our May forecast models, from longest to shortest lead time, explain 43%, 52%, and 58% (median across strata) of pond count variation, respectively. The July forecast models, from longest to shortest lead time, explain 29%, 46%, 51%, and 56% (median across strata) of pond count variation, respectively. Individual strata R^2 values can be seen in table 2. Larson's [26] best model explained 65% of variation in percent basins holding water, and the model given by Niemuth *et al* [27] explained 62% of pond number variation Sorenson *et al* [28] were able to explain 72% of May pond count variation using only May PDSI. Generally, our shortest lead time models perform similarly to these models, and the variation explained remains strong with longer lead time aside from the March forecast for July pond counts.

Stratum 47 has lower R^2 than other strata for both models. Exploratory data analysis reveals one possible explanation. There are two extreme years present within the May pond data (figure 4(d)) and removing these improves model performance. In the July data, removing one extreme year also improves model performance (figure 6(d)). Therefore, the models are not capturing the extreme years.

5. Summary and discussion

We developed a suite of predictors for SEPPR pond count informed by previous work investigating large-scale climate links to summer precipitation in the SEPPR. Using these predictors, we developed predictive models using CCA that can provide relevant information for SEPPR managers as they plan conservation efforts such as determining hunting quotas and the timing and location for rotational grazing, prescribed burns, haying, and noxious weed control.

In providing a better predictive tool, these models can help managers make more informed choices for actions to preserve and maintain habitat in the face of near-term threats. This tool provides longer lead times than currently available models and pond forecasts in regions consistent with areas that managers use—the survey strata. Managers may use our models for spatially explicit forecasts of where pond habitat might be better or worse (i.e. where there are fewer ponds and therefore pond area), and thus inform their planning at varying times in the seasonal decision-making process. Additionally, these models use variables that influence the climate variability of the region and capture seasonal climate information and trend. Capturing this variability is important because managers are very interested in the climate variability of the region [8]. These models exhibit high skill exemplified by high linear correlation values and low RMSE in LOOCV mode. With a few exceptions, this skill increased as lead time decreased.

Our CCA framework provides versatility that can be adapted to accommodate needs and interests of managers. Variables can easily be added (or removed) to either dataset. For example, other strata could be included for predictions as desired by managers. Other variables, such as snowmelt, could be added to the predictor dataset to incorporate their information in the model.

An important utility of these models lies in their ability to provide information on ecological health based on the strong connection between pond count and duck population (figure 2). Managers must still be mindful that duck population depends on other factors such as pond salinity (LaBaugh *et al* [47]), wetland-cover type, vegetation characteristics (also dependent on salinity), and interactions among ducks [36]. They should also be careful making these inferences in strata with lower model skill, but broad inferences can still be made with confidence. Further, studies providing similar pond count models do not include large-scale climate variables, have little lead

time, do not forecast for individual strata, or a combination thereof.

This work has a broader impact because we demonstrate the use of a powerful forecasting technique, CCA, and how it might be used for management decisions. This application could serve as an example to develop similar models for use in other management and conservation decisions in other regions of the world and other sectors of environment. Further, we demonstrate the ability to use large-scale climate variables to predict a local variable (pond count) in a skillful way. This method could be applied to other variables, some of which were referenced previously. We hope the technique will reach a different audience than has been in previous works.

Acknowledgments

Data availability

The data that support the findings of this study are openly available at the following URL/DOI.

IRI Data Library: https://iridl.ldeo.columbia.edu/

Migratory Bird Data Center: https://migbirdapps.fws.gov/mbdc/databases/db_selection.asp

NOAA NCEI nClimDiv: https://doi.org/10.7289/V5M32STR

ORCID iDs

Benjamin D Abel ⓘ https://orcid.org/0000-0003-0344-8137
Balaji Rajagopalan ⓘ https://orcid.org/0000-0003-1433-7520
Andrea J Ray ⓘ https://orcid.org/0000-0001-5385-1202

References

[1] Smith A, Stoudt J and Gallop J 1964 Prairie Pothole Wetlands and Marshes *Waterfowl Tomorrow* ed J Linduska (Washington, DC: U.S. Government Printing Office) pp 39–50

[2] Dahl T E 2014 *Status and Trends of Prairie Wetlands in the United States 1997 to 2009* (Washington, DC: U.S. Department of the Interior; Fish and Wildlife Service, Ecological Services)

[3] Dyke A and Prest V 1987 Late Wisconsinan and Holocene history of the Laurentide ice sheet *Géog. Phys. Quatern.* **41** 237–63

[4] van der Valk A G 2005 Water-level fluctuations in North American prairie wetlands *Hydrobiologia* **539** 171–88

[5] Batt B D J, Anderson M G, Anderson C D and Caswell F D 1989 The use of prairie potholes by North American ducks *Northern Prairie Wetlands* ed A van der Valk (Ames, IA: Iowa State University Press) pp 204–27

[6] Kantrud H A, Krapu G L and Swanson G A 1989 *Prairie Basin Wetlands of the Dakotas: A community Profile* (Washington, DC: U.S. Fish and Wildlife Services) p 116 Biological Report85(7.28)

[7] Gleason R A *et al* (ed) 2008 *Ecosystem services derived from wetland conservation practices in the United States Prairie Pothole Region with an emphasis on the U.S. Department of Agriculture Conservation Reserve and Wetlands Reserve Programs* (Reston, VA: U.S. Geological Survey) Professional paper 1745 (https://doi.org/10.3133/pp1745)

[8] Yocum H M and Ray A J 2019 Climate information to support wildlife management in the North Central United States *Reg. Environ. Change* **19** 1187–99

[9] Doherty K E, Ryba A J, Stemler C L, Niemuth N D and Meeks W A 2013 Conservation planning in an era of change: state of the U.S. Prairie Pothole Region: state of the Prairies *Wildlife Soc. Bull.* **37** 546–63

[10] Dahl T E 1990 *Wetlands Losses in the United States 1780s to 1980s* (Washington, DC: U.S. Department of the Interior, Fish and Wildlife Service)

[11] Wright C K and Wimberly M C 2013 Recent land use change in the Western Corn Belt threatens grasslands and wetlands *Proc. Natl Acad. Sci.* **110** 4134–9

[12] Voldseth R A, Johnson W C, Gilmanov T, Guntenspergen G R and Millett B V 2007 Model estimation of land-use effects on water levels of northern prairie wetlands *Ecol. Appl.* **17** 527–40

[13] Gleason R A and Tangen B A (ed) 2014 *Brine contamination to aquatic resources from oil and gas development in the Williston Basin, United States* (Reston, VA: U.S. Geological Survey) Scientific Investigations Report 2014–5017 (https://doi.org/10.3133/sir20145017)

[14] Millett B, Johnson W C and Guntenspergen G 2009 Climate trends of the North American prairie pothole region 1906–2000 *Clim. Change* **93** 243–67

[15] Mushet D M 2016 Midcontinent prairie-pothole wetlands and climate change: an introduction to the supplemental issue *Wetlands* **36** 223–8

[16] Rosenberry D O 2003 Climate of the Cottonwood Lake area *Hydrological, Chemical, and Biological Characteristics of a Prairie Pothole Wetland Complex Under Highly Variable Climate Conditions: The Cottonwood Lake Area, East-Central North Dakota* ed T C Winter (Reston, VA: U.S. Geological Survey) Professional paper 1675 (https://doi.org/10.3133/pp1675)

[17] Winter T C and Rosenberry D O 1998 Hydrology of prairie pothole wetlands during drought and deluge: a 17 year study of the Cottonwood Lake wetland complex in North Dakota in the perspective of longer term measured and proxy hydrological records *Clim. Change* **40** 189–209

[18] Hayashi M, van der Kamp G and Rosenberry D O 2016 Hydrology of Prairie Wetlands: understanding the integrated surface-water and groundwater processes *Wetlands* **36** 237–54

[19] LaBaugh J W, Mushet D M, Rosenberry D O, Euliss N H, Goldhaber M B, Mills C T and Nelson R D 2016 Changes in pond water levels and surface extent due to climate variability

alter solute sources to closed-basin Prairie-Pothole Wetland Ponds, 1979 to 2012 *Wetlands* **36** 343–55

[20] Liu G and Schwartz F W 2012 Climate-driven variability in lake and wetland distribution across the Prairie Pothole Region: from modern observations to long-term reconstructions with space-for-time substitution *Water Resour. Res.* **48** W08526

[21] Vecchia A V 2008 *Climate Simulation and Flood Risk Analysis for 2008–40 for Devils Lake, North Dakota* (Reston, VA: U.S. Geological Survey) Scientific Investigations Report 2008–5011 (https://doi.org/10.3133/sir20085011)

[22] Ballard T, Seager R, Smerdon J E, Cook B I, Ray A J, Rajagopalan B, Kushnir Y, Nakamura J and Henderson N 2014 Hydroclimate variability and change in the prairie pothole region, the 'Duck Factory' of North America* *Earth Interact.* **18** 1–28

[23] US Fish and Wildlife Service 2015 Kulm Wetland Management District. Habitat Management Plan, Kulm, North Dakota. U.S. Department of the Interior, Fish and Wildlife Service, Mountain-Prairie Region. p 192 https://data.doi.gov/dataset/kulm-wetland-management-district-habitat-management-plan

[24] Salo E D, Higgins K F, Patton B D, Bakker K K, Barker W T, Kreft B and Nyren P E 2004 Grazing intensity effects on vegetation, livestock and non-game birds in North Dakota mixed-grass prairie *Proc. 19th N. A. Prairie Conf.* (Madison, WI) https://digitalcommons.unl.edu/napcproceedings/88

[25] Barker W T, Sedivec K K, Messmer T A, Higgins K F and Hertel D R 1990 Effects of specialized grazing systems on waterfowl production in south-central North Dakota *Trans. 55th N. A. Wildlife and Nat. Res. Conf.* (Denver, CO) (https://digitalcommons.unl.edu/usfwspubs/29)

[26] Larson D L 1995 Effects of climate on numbers of northern prairie wetlands *Clim. Change* **30** 169–80

[27] Niemuth N D, Fleming K K and Reynolds R E 2014 Waterfowl conservation in the US Prairie Pothole region: confronting the complexities of climate change *PLoS One* **9** e100034

[28] Sorenson L G, Goldberg R, Root T L and Anderson M G 1998 Potential effects of global warming on waterfowl populations breeding in the Northern Great plains *Clim. Change* **40** 343–69

[29] Kaplan A, Cane M A, Kushnir Y, Clement A C, Blumenthal M B and Rajagopalan B 1998 Analyses of global sea surface temperature 1856-1991 *J. Geophys. Res.: Oceans* **103** 18567–89

[30] Reynolds Richard W. and Smith Thomas M. 1994 Improved global sea surface temperature analyses using optimum interpolation *J. Climate* **7** 929–48

[31] Vose R S, Applequist S, Squires M, Durre I, Menne M J, Williams C N Jr, Fenimore C, Gleason K and Arndt D 2014 NOAA's Gridded Climate Divisional Dataset (CLIMDIV). [pdsidv]. NOAA National Climatic Data Center (Accessed: 3 June 2018) (https://doi.org/10.7289/V5M32STR)

[32] Helsel D R and Hirsch R M 2002 *Statistical Methods in Water Resources: Techniques of Water Resources Investigations* (Reston, VA: U.S. Geological Survey) ch A3 https://pubs.usgs.gov/twri/twri4a3/ Book 4

[33] Montgomery D C and Runger G C 2014 *Applied Statistics and Probability for Engineers* 6th edn (Hoboken, NJ: Wiley)

[34] Johnson W C, Millett B V, Gilmanov T, Voldseth R A, Guntenspergen G R and Naugle D E 2005 Vulnerability of Northern prairie wetlands to climate change *BioScience* **55** 863–72

[35] Niemuth N D and Solberg J W 2003 Response of waterbirds to number of Wetlands in the Prairie Pothole region of North Dakota U.S.A. *Waterbirds* **26** 233–8

[36] Raven G H, Armstrong L M, Howerter D W and Arnold T W 2007 Wetland selection by mallard broods in Canada's Prairie-Parklands *J. Wildl. Manage.* **71** 2527–31

[37] Stewart R E and Kantrud H A 1972 Vegetation of prairie potholes, North Dakota, in relation to quality of water and other environmental factors *Professional Paper* 585-D U.S. Geological Survey (https://doi.org/10.3133/pp585D)

[38] Conant R T *et al* 2018 Northern Great Plains *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II* ed D R Reidmiller *et al* (Washington, DC: Global Change Research Program) 941–86

[39] von Storch H and Zwiers F W 2001 *Statistical Analysis in Climate Research* (Cambridge: Cabridge University Press)

[40] Barnston A G and Smith T M 1996 Specification and prediction of global surface temperature and precipitation from global SST Using CCA *J. Clim.* **9** 2660–97

[41] Mo K C 2003 Ensemble canonical correlation prediction of surface temperature over the United States *J. Clim.* **16** 1665–83

[42] Barnston A G and Ropelewski C F 1992 Prediction of ENSO episodes using canonical correlation analysis *J. Clim.* **5** 1316–45

[43] Gill E C, Rajagopalan B, Molnar P and Marchitto T M 2016 Reduced-dimension reconstruction of the equatorial Pacific SST and zonal wind fields over the past 10,000 years using Mg/Ca and alkenone records *Paleoceanography* **31** 928–52

[44] von Storch H 1999 Spatial patterns: EOFs and CCA *Analysis of Climate Variability* ed H von Storch and A Navarra (Berlin: Springer) pp 231–63

[45] Palmer W C 1965 Meteorological drought *Research Paper No* 45 U.S. Dept. of Commerce, Washington, DC

[46] Efron B 1982 *The Jackknife, the Bootstrap and Other Resampling Plans.* (Philadelphia, Pa: SIAM)

[47] LaBaugh J. W., Winter T. C., Swanson G. A., Rosenberry D. O., Nelson R. D. and Euliss N. H. 1996 Changes in atmospheric circulation patterns affect midcontinent wetlands sensitive to climate *Limnol. Oceanogr.* **41** 864–70

[48] Salas J D, Fu C and Rajagopalan B 2011 Long-range forecasting of Colorado streamflows based on hydrologic, atmospheric, and oceanic data *J. Hydrol. Eng.* **16** 508–20