ECOLOGICAL
APPLICATIONS
ECOLOGICAL SOCIETY OF AMERICA

Both real-time and long-term environmental data perform well in predicting shorebird distributions in managed habitat

Erin E. Conlisk[1] | Gregory H. Golet[2] | Mark D. Reynolds[2] |
Blake A. Barbaree[1] | Kristin A. Sesser[1] | Kristin B. Byrd[3] |
Sam Veloz[1] | Matthew E. Reiter[1]

[1]Point Blue Conservation Science, Petaluma, California, USA

[2]The Nature Conservancy, Sacramento, California, USA

[3]U.S. Geological Survey, Western Geographic Science Center, Moffett Field, California, USA

Correspondence
Erin E. Conlisk
Email: econlisk@pointblue.org

Funding information
NASA Ecological Forecasting, Grant/Award Number: NNX17AG81G; Nature Conservancy; S. D. Bechtel Jr. Foundation

Handling Editor: Beth Gardner

Abstract

Highly mobile species, such as migratory birds, respond to seasonal and inter-annual variability in resource availability by moving to better habitats. Despite the recognized importance of resource thresholds, species-distribution models typically rely on long-term average habitat conditions, mostly because large-extent, temporally resolved, environmental data are difficult to obtain. Recent advances in remote sensing make it possible to incorporate more frequent measurements of changing landscapes; however, there is often a cost in terms of model building and processing and the added value of such efforts is unknown. Our study tests whether incorporating real-time environmental data increases the predictive ability of distribution models, relative to using long-term average data. We developed and compared distribution models for shorebirds in California's Central Valley based on high temporal resolution (every 16 days), and 17-year long-term average surface water data. Using abundance-weighted boosted regression trees, we modeled monthly shorebird occurrence as a function of surface water availability, crop type, wetland type, road density, temperature, and bird data source. Although modeling with both real-time and long-term average data provided good fit to withheld validation data (the area under the receiver operating characteristic curve, or AUC, averaged between 0.79 and 0.89 for all taxa), there were small differences in model performance. The best models incorporated long-term average conditions and spatial pattern information for real-time flooding (e.g., perimeter-area ratio of real-time water bodies). There was not a substantial difference in the performance of real-time and long-term average data models within time periods when real-time surface water differed substantially from the long-term average (specifically during drought years 2013–2016) and in intermittently flooded months or locations. Spatial

predictions resulting from the models differed most in the southern region of the study area where there is lower water availability, fewer birds, and lower sampling density. Prediction uncertainty in the southern region of the study area highlights the need for increased sampling in this area. Because both sets of data performed similarly, the choice of which data to use may depend on the management context. Real-time data may ultimately be best for guiding dynamic, adaptive conservation actions, whereas models based on long-term averages may be more helpful for guiding permanent wetland protection and restoration.

INTRODUCTION

Species distribution models are commonly employed to investigate species occurrence across the landscape (Franklin, 2010), where species' ranges are a fundamental concern in conservation. Although broadly useful, species distribution models make a number of implicit, simplifying assumptions about the relationship between target taxa and environmental conditions. Of particular interest, species' ranges are often predicted based on the long-term average of rapidly changing environmental conditions, such as 30-year averaged temperature and precipitation. Long-term average environmental data are employed because frequent, temporally resolved measurements of environmental conditions across large spatial extents can be difficult to obtain and analyze (Young et al., 2017). In addition to sparse environmental data, there may not be enough species occurrence data to build models of species' response to fluctuating environmental conditions. Further, because average environmental conditions are correlated to the real-time observations making up that average, species occurrence can be described by average conditions, especially for species with high site fidelity (Bradshaw et al., 2004). Conservation practitioners often use models built on long-term data because they are seeking to protect the best overall habitat, regardless of year-to-year variability in suitability.

The frequency of environmental measurements can be divided into three broad categories (Mannocci et al., 2017): instantaneous, real-time, and long-term. Although instantaneous measurements occurring at the finest temporal (e.g., daily or multiple times per day) resolution can be used in physiological species distribution models (e.g., Buckley et al., 2010), they are not widely employed (Dormann et al., 2012). Real-time data describe the state of the environment in a time window of days to weeks around an observation, and long-term data represent the "typical" state of the environment, often obtained by averaging across years.

To date, many types of data have been used to describe species distributions: climate (Stralberg et al., 2009), time since disturbance (such as fire, as in Connell et al., 2017), forest canopy structure (Burns et al., 2020), and land-cover characteristics (Fink et al., 2020; Shirley et al., 2013). Monitoring disturbance frequency, forest canopy, and land cover with data taken infrequently or averaged across years—that is, long-term data—is sensible because these characteristics typically change on seasonal, annual, or multiyear time frames. The use of temporally resolved environmental data is most likely to be important for highly mobile species responding to extreme or ephemeral environmental conditions, conditions such as changing weather and flooding. In these situations, distribution models built with frequently collected, real-time environmental data (Pettorelli et al., 2014) and complementarily robust observations of species through time and space could provide temporally resolved habitat associations useful for dynamic conservation decision making (D'Aloia et al., 2019). Satellite and remotely sensed real-time data have been used extensively to forecast movement of marine species (Scales et al., 2017, Welch et al., 2018) and model animal migration following spring green-up (Bauer et al., 2008; Middleton et al., 2018). Although the use of real-time measurements has proceeded much more quickly for marine than terrestrial systems, weather data have been used instead of climate data to improve distribution predictions for vagile, Australian birds (Reside et al., 2010) and bettongs (Bateman et al., 2012). Habitat predictions of wetlands birds have benefited from multitemporal characteristics, finding that predictors of temporary water

availability were best (Pickens & King, 2014). Further, real-time conditions allow for the quantification of the amount of a resource as well as the shifting patterns of resource availability, such as landscape connectivity (MacGarigal et al., 2012) of temporary water bodies. Despite the potential benefit of real-time variables, few studies have tested the assumption that real-time covariates provide critical, additional information about species distributions in rapidly changing environments (but see Reside et al., 2010). Nor have studies tested how maps of predicted habitat differ under models built with real-time versus long-term average data (but see Reside et al., 2010).

Despite their potential for dynamic habitat predictions, real-time measurements can be of limited value in species-distribution models if there are not enough observations of the target taxa in extreme environmental conditions across space or time. For example, a species-distribution model trained on data from a subset of its range provides an incomplete description of species' environmental tolerances because only part of the species' climate tolerances was used to train the models (Sánchez-Fernández et al., 2011). Similarly, climatic variation in space can be less extreme than variation across time (Veloz et al., 2012); thus, extreme climate conditions cannot be effectively introduced into a model trained on average data (Pérez Navarro et al., 2019). Although rigorous data sets across time and space are rare, one study in the US Prairie Pothole region demonstrated that shorebirds preferentially occupied reliably flooded, northerly wetlands in dry years and more ephemeral, southerly ponds in wet years (Steen et al., 2018). The relative value of intermittently flooded areas in wet versus dry years was identified because there were enough observations, across years, within highly variable sites. Further, there had been prior work demonstrating hierarchical habitat choice— from coarse land cover at large spatial extent to local habitat features—across dynamic landscapes (Albanese & Davis, 2013).

Beyond requiring more robust environmental and ecological sampling through space and time, there are additional disadvantages to using real-time data. Translating satellite measurements to meaningful ecological predictors can be nontrivial, requiring preprocessing for sensor, solar, atmospheric, and topographic effects (Young et al., 2017). Cloud cover can lead to gaps in satellite-derived, real-time data (Mannocci et al., 2017), reducing the number of observations available for distribution modeling. Finally, real-time data can make conservation decisions more complex because habitat forecasts across varying environmental conditions require that each set of conditions be given a prioritization weighting.

With highly variable water availability annually and seasonally, the Central Valley of California offers an opportunity to test the potential benefit of real-time data used in species distribution models for highly mobile shorebirds. The Central Valley lies along a large, flat latitudinal gradient (720 km northwest to southeast) between the Coast Range to the west and the Sierra Nevada to the east. Heavy water infrastructure is designed to moderate seasonally and annually varying water availability in this Mediterranean-type climate with El Niño fluctuations. The Central Valley has lost >90% of its once extensive natural wetlands (Frayer et al., 1989). Despite this habitat loss, the Central Valley remains a critically important region for resident and migratory shorebirds of the Pacific Flyway (Page et al., 1999). Nonbreeding migratory shorebirds (Order: Charadriiformes; Families: Scolopacidae and Recurvirostridae) in dynamic wetland landscapes have been shown to change distribution in response to changing environmental conditions over relatively short time periods (~2 weeks; Barbaree et al., 2015, 2018; Stenzel & Page, 2018). The dynamic nature of freshwater wetland habitat availability in the American west and the ability of the birds to respond has argued for more dynamic approaches to conservation that provide habitat only when and where it is most needed (Reynolds et al., 2017). To aid such conservation efforts, we hypothesized that, compared to long-term average conditions, real-time covariates used in distribution models will (1) better incorporate year-to-year shorebird response to climate variability and important landscape patterns, resulting in higher overall model performance; (2) better predict habitat suitability in years that deviate from the long-term average, yielding the highest relative performance on validation data withheld from drought years; (3) better predict habitat suitability in months (such as October; Reiter et al., 2018) and spatial locations with high year-to-year variability in surface water availability; (4) allow for flooding pattern metrics (e.g., perimeter area relationship) that improve prediction; (5) result in a smaller data set because of potential loss of measurements (e.g., because of cloud cover), which will result in decreased performance; and (6) yield similar spatial predictions to models built with long-term average conditions when model performance is high.

We tested these hypotheses on species with varying occurrence during the migration and over-wintering periods (Shuford et al., 1998): dunlin (*Calidris alpina*), black-necked stilt (*Himantopus mexicanus*), American avocet (*Recurvirostra americana*), and long- and short-billed dowitchers combined (*Limnodromus griseus* and *Limnodromus scolopaceus*). Avocets and stilts breed in the region (Shuford et al., 2004) and dunlin and dowitchers

tend to move in large flocks, breeding in the arctic and subarctic (Shuford et al., 1998). Although we focus on the Central Valley of California, our results are relevant to midcontinental regional conservation management of these species in the Rainwater Basin (Gillespie & Fontaine, 2017), the southern Great Plains (Albanese & Davis, 2013), and beyond (Skagen & Knopf, 1993).

METHODS

Bird data

We combined multiple data sets on the occurrence of five species of shorebirds (treating long- and short-billed dowitchers as a single taxa), including structured data sets and complete checklists in the opportunistically collected eBird reference data set (ERD) through 2016 (Sullivan et al., 2009). The structured data sets (Barbaree et al., 2020; Golet et al., 2018; Reiter et al., 2011; Sesser et al., 2016, 2018; Shuford et al., 2019; Strum et al., 2013) were largely conducted during times of year and in rice and other agricultural habitats where, given flooding, shorebirds were expected. To supplement the spatial and temporal extent of the structured data, we included eBird observations that used sampling protocols that closely matched the structured data: namely, observations collected during daylight hours and by stationary observers. Structured and eBird data have recently been shown to complement one another in modeling shorebird distribution in the Central Valley (Robinson et al., 2020) because eBird data represent a broader suit of habitat types compared to structured data in wetlands and rice.

Although the structured data sets were taken by different observers for different experimental questions (Barbaree et al., 2020, Golet et al., 2018, Reiter et al., 2011, Sesser et al., 2016, 2018, Shuford et al., 2019, Strum et al., 2013), they all employed similar search and data collection protocols. Stationary observers identified all shorebirds to species and counted all individuals within a 160–200-m fixed radius or until the intersection of levees or natural barriers, whichever was closer. Surveys lasted at least 2 min with no maximum time limit, although observers were encouraged to complete counts as rapidly as possible to guard against bird movement into and out of the unit. Surveys were not conducted in inclement weather or when visibility was low, and thus are assumed to have high detectability (Sesser et al., 2018). Surveys were typically repeated in the same location every 5–15 days during the survey season. Survey seasons ranged from a single month (15 November–15 December) to most of the wet season (July–March; see Appendix S1: Table S1).

Compared to structured surveys, eBird data covered a much broader suite of cover types, including locations where shorebirds were unlikely and locations near urban areas (see Appendix S1: Section S1). However, eBird observers in wetlands were likely to find shorebirds, leading to a higher fraction of presence locations than the structured data. Across 2000–2016, we included data for July through April—the months we modeled in this study—for all species except dowitchers. For dowitchers, we only modeled October–April because of a potentially biased eBird sampling of dowitcher presence in May–September (see Appendix S1: Section S1). We used only complete checklists (eBird observers indicated they attempted to record all species present), which can be used to infer absences. As discussed in real-time versus long-term data section, we included additional eBird traveling observations moving fewer than 250 m in a data set that included additional bird observations (called the high-n data).

All structured data and >99% of the eBird data (a small fraction of eBird data only recorded presence or absence) recorded the total number of individuals for each of the four target taxa: dunlin, black-necked stilt, American avocet, and all dowitchers combined. We modeled species presence without explicitly including detectability and encounter rates separately because our study had high detectability and low occurrence, conditions that favor exclusion of detectability (Welsh et al., 2013). To reduce spatial and temporal autocorrelation (Veloz, 2009), we thinned the combined eBird and structured data using the spThin package (Aiello-Lammens et al., 2015). We did not thin by land cover type or characteristics, but rather thinned such that no two observations occurred within the time window defined by successive Landsat mosaics and no two observations occurred within 1 km of one another (see Appendix S1: Section S1 for methodological details and autocorrelation analyses).

Covariate data

Given shorebird associations with wetlands and flooded rice and corn fields (Dybala et al., 2017; Shuford et al., 1998), we used the US Department of Agriculture–National Agricultural Statistics Service Cropland Data Layer (USDA-NASS, 2014) to define consistent cover types (The Nature Conservancy, Rodd Kelsey, unpublished data set): rice, corn, row crops, or grain; or managed wetlands (which included two categories: seasonal and semipermanent) and wastewater treatment, that is, "treated," wetlands (Petrik et al., 2014). To estimate surface water, we used remotely sensed spatial data layers built from 30-m resolution Landsat 5 (2000–2011)

T A B L E 1   Environmental covariates used in bird presence/absence models, as described in the real-time versus long-term data section

Name	Description and source	Scale	Real-time
Flooded seasonal and permanent wetlands	Overlay flooding on areas identified as seasonal or semi-permanent wetlands. In real-time layers, if both conditions (flooded and wetlands) were satisfied, the resulting layer produced a 1; if either or both conditions were not satisfied (i.e. unflooded or not wetlands), a 0 value was produced. Values 0/1 were averaged across the moving windows specified in the next column. Sources: Point Blue Water Tracker (Reiter et al., 2015), Ducks Unlimited (Petrik et al., 2014).	250 m, 5 km	Yes
Flooded rice	Same methodology as above but substituting rice for wetlands. Sources: Point Blue Water Tracker (Reiter et al., 2015), TNC classification of National Cropscape data (USDA, 2014) that were rice for 7–8 years within 2008–2014.	250 m, 5 km	Yes
Flooded corn	Same methodology as above but substituting corn for rice. Sources: Same as above row.	250 m	Yes
Flooded row, field, and grain crops	Same methodology as above but using row, field, and grain crops. Sources: Same as above row.	250 m	Yes
Flooded nonrice crops	Same methodology as above but using all corn, row, field, grain, and alternating crops. Sources: Same as above row.	5 km	Yes
Flooded treated wetlands	Same methodology as above but using treated wetlands. Treated wetlands are artificial wetlands created to treat industrial wastewater, gray water, and run-off. They are not managed for birds. Sources: Same as top row.	250 m, 5 km	Yes
Cohesion of real-time flooding	Quantifies the connectivity of all water regardless of cover type (MacGarigal et al., 2012). Sources: Point Blue Water Tracker (Reiter et al., 2015)	5 km	Yes
Mean fractal density and perimeter-area ratio of real-time flooding	Two metrics quantifying the degree to which flooding is "space-filling" and has more edge habitat as opposed to interior habitat (MacGarigal et al., 2012). Sources: Point Blue Water Tracker (Reiter et al., 2015)	250 m	Yes
Road density	Density of roads (km road/km^2) within a 5 km moving window from the bird observation. Source: OpenStreetMap, downloaded in July 2017	5 km	No
Maximum temperature	Calculated as a moving window average for the month and year when the bird observation was collected. Source: Flint and Flint (2014)	250 m	Yes
Observation type	Defined as eBird versus structured data following the protocols in (Reiter et al., 2011).	NA	No

Note: The scales used are listed in the third column. Covariates used as real-time variables in addition to long-term averages are delineated in the last column.

and Landsat 8 (2013–2016) satellite data (Reiter et al., 2015, 2018) in the four Landsat tiles covering most of the Central Valley (p44r33, p44r34, p43r34, p42r35). Because we lacked surface-water data for 2012 (Landsat 7 had a scan line error in the satellite that led to >25% loss of image data), we did not include any bird data from 2012 and did not predict species distributions for that year. For each month modeled (July–April), we overlaid surface water layers on land-cover data to determine the fractional surface area that was flooded habitat, resulting in the 15 covariates in Table 1. These covariates were used either "as is," for real-time predictions, or averaged over 2001–2017 for each month, for long-term average surface water (described in the next section). Thus, real-time spatial layers differed from long-term layers because real-time pixels were binary—the landscape is flooded or not flooded—whereas long-term average pixels were continuous—from never flooded, 0, to always flooded, 1.

For each observation location, we calculated landscape characteristics at two scales: within the search area of a given bird observation and within a 5-km radius around a bird observation. Search areas were explicitly defined in structured data (usually a circle or semicircle with an up to 200-m radius), and we used a 250-m radius around reported coordinates for eBird observations (eBird does not report search areas). We included a 5-km-radius measurement for covariates because shorebirds have been shown to select habitat moving from broad landscape characteristics to fine-resolution local characteristics hierarchically (Albanese & Davis, 2013; Gillespie & Fontaine, 2017; Skagen & Knopf, 1994).

We included additional non-water-related variables and water-pattern covariates. Specifically, we included road density, observation type (eBird or structured data), and maximum temperature observed for that month and year (Flint & Flint, 2014). We did not include month as a factor because the surface water and temperature variables sufficiently accounted for seasonality. We did not include the time observers spent sampling because it did not emerge as an influential predictor, in part because sampling times were very different between eBird and structured data. For two sets of models using real-time data, we calculated pattern metrics for all flooded areas (regardless of cover type); these variables are described in greater detail in the next section.

Real-time versus long-term data

To explore the predictive benefit of real-time versus long-term average data, we considered five sets of models built with different data sets and covariates (Table 2): (1) long term with additional checklists (hereafter "long-high-n"); (2) long-term average, only observations with matching real-time observation (hereafter "long-low-n"); (3) real time with no pattern metrics (hereafter "real-no-pattern"); (4) real-time with pattern metrics (hereafter "real-pattern"); and (5) long-term, real-time pattern (hereafter "real-long"). With the exception of models built with long-high-n covariates, all models used the low-n subset of data.

Models built with the long-high-n covariate data are most similar to typical species distribution models, where bird observations are related to long-term average environmental conditions. We matched the bird observations to the average flooding within the month of the bird observation (e.g., for a bird observed in December, we took the average flooding for the 17 Decembers from 2001 to 2017) and calculated the covariates in Table 1. Models built with long-low-n data covariate set used similar covariates, but differed from the high-n data in two crucial ways. First, in the low-n data set, only bird observations that had real-time covariate data were included. Including data points with average information but no real-time information resulted in an additional 1,652 eBird checklists and 1,215 structured surveys in the high-n data set (a roughly 25% increase in both types of data). The second difference in the two data sets was that long-high-n data included 1,239 traveling count checklists where observers moved less than 250 m, all other data sets included only stationary counts. Comparing models built with long-high-n versus long-low-n data sets tested the influence of a reduced sample size caused by cloud cover and a potentially stronger mismatch between the long-term average and the bird observation. Including traveling counts, we explored the benefit of more data as compared to a potentially stronger mismatch between the measured environmental conditions and the location of sampling.

To compare to long-term data, we used two different sets of real-time covariate data. The real-no-pattern set included only real-time variables that were analogous to the long-term data. The real-pattern covariates included data from the 16-day surface water image prior to the bird observation, to account for any potential lags in shorebird response (Barbaree et al., 2018), and pattern metrics, which could serve as a proxy for important habitat characteristics. For example, the shallow wetlands preferred by shorebirds are likely to have a high perimeter-area ratio (Barbaree et al., 2018), but also higher danger of predation (Pomeroy, 2006). Specifically, we calculated the perimeter-area ratio and mean fractal dimension within 250 m of a bird observation and landscape cohesion within a 5-km radius of a bird observation using the SDMTools package (VanDerWal et al., 2014) in R 3.5.1 (R Core Team, 2020) and the FRAGSTATS 4.2

T A B L E 2 The five data types used to build bird presence/absence models and five validation data subsets used to test models

| Name	Description	Rationale
Long-high-n	Models using stationary and traveling bird counts moving <250 m. Bird observations collected when cloud cover masked a Landsat image were included in the model if there was information on the long-term average data for that location during the same time of year in other years.	To test the benefit of including additional observations compared to the data set described in the row below and to test a potentially stronger mismatch between a bird observation with missing real-time conditions.
Long-low-n	Same covariates as long-high-n models, but a smaller data set, where only stationary counts and observations with Landsat real-time data (i.e., no cloud cover) were included in the model.	To test the benefit of less but higher quality data compared to the data set described in the row above; to test whether real-time data provide better predictions compared to the data sets described in the rows below.
Real-no-pattern	Real-time covariates analogous to the long-low-n. Real-time spatial layers differed from long-term layers because real-time pixels were binary—flooded or not flooded within a specific 16-day window—whereas long-term average pixels were continuous—never flooded, 0, to always flooded, 1, within 2001–2017.	Simplest real-time model for comparison with the long, low-n models.
Real-pattern	Used real-time covariates, including additional covariates that only make sense in a real-time context, namely: surface water availability in the previous Landsat image and pattern metrics to quantify perimeter area ratio, fractal dimension, and landscape cohesion.	Determine the benefit of pattern metrics and previous environmental conditions, variables that are most relevant with frequent, real-time environmental measurements.
Real-long	Real-time covariates, including water pattern metrics and surface-water availability in the previous Landsat image, were added to long-term covariates.	Quantify the performance benefit of a combination of data types.
Validation data	We compared the use of withheld validation data from a random 20% subset of data and four subsets of data based on year of observation (2008–2009, 2010–2011, 2013–2014, 2015–2016). The 80% training data (with concomitant 20% validation data) were used to create spatial predictions (Figures 4–6).	Comparing drought years—the 2013–2014 data, and to a lesser extent the 2015–2016 data—to the other years, explored the relative performance of real-time and long-term data in years that deviated from the long-term average. Comparing the randomly withheld data to the 2-year subsets, we analyzed the degree to which years differed from one another, potentially inflating performance measured with a random subset of data.
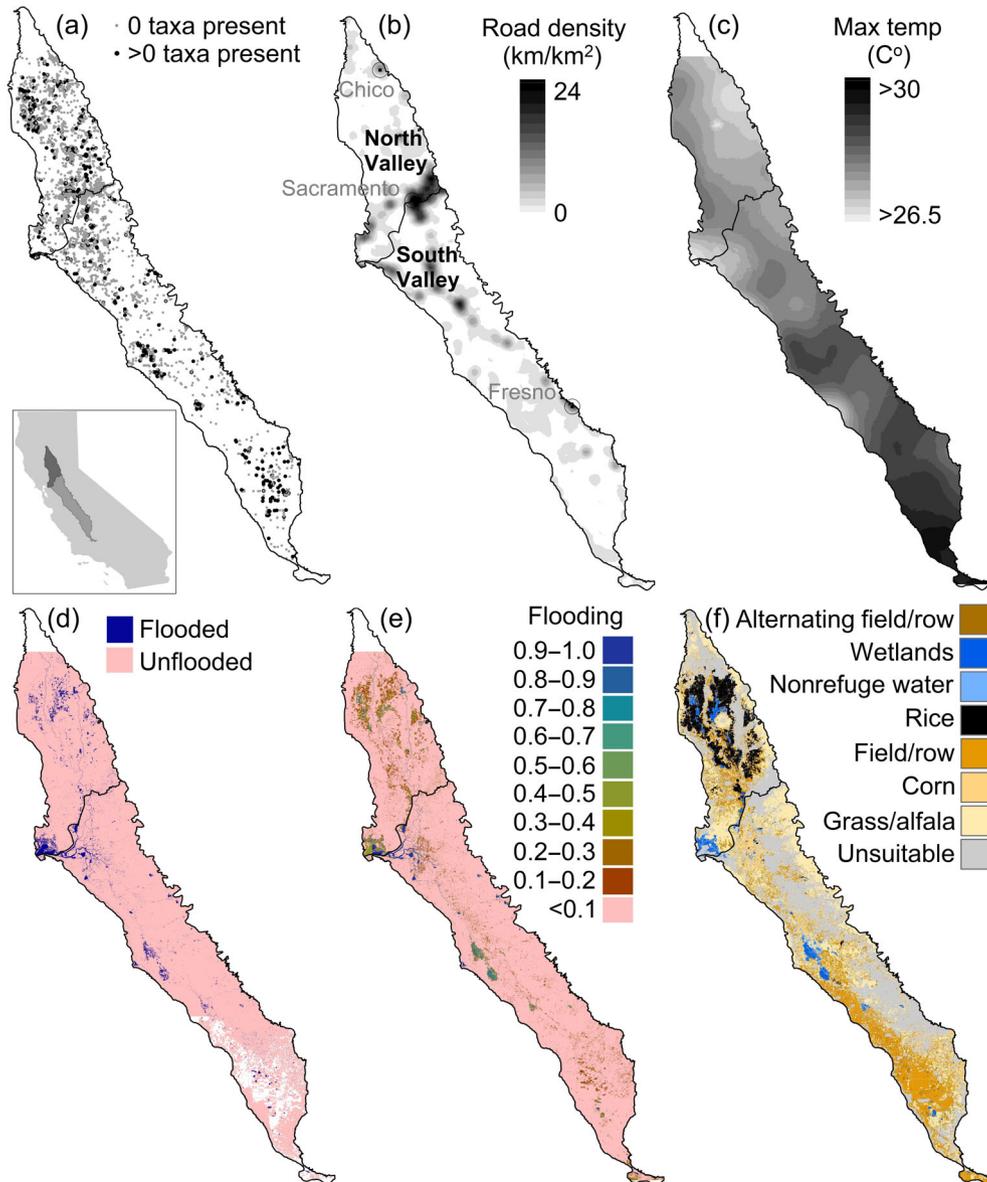
FIGURE 1 Covariates used in the shorebird models: (a) locations of bird observations across 2000–2016 and colored based on whether at least one of the four taxa was present; (b) road density, within a 5-km moving window, and three cities—Chico, Sacramento, and Fresno–within the study region; (c) October 2014 maximum temperature; (d) a nearly complete surface water mosaic from 12–21 October 2014, during the California drought, with some cloud cover in the south valley; (e) percentage of years with surface water in October 2001–2017; and (f) land-cover types. All legends are to the upper right of their respective maps. The map extent within California can be seen in the inset and represents the Central Valley Joint Venture boundary. The northernmost region of the Joint Venture boundary is not included in the four Landsat tiles used in this analysis. The black line through the center of the valley delineates where the data were split to create a North Valley and South Valley for the model ensemble [also labeled in (b) along with three cities in the Central Valley]

software described by MacGarigal et al. (2012). We compared real-pattern to real-no-pattern, and tested the importance of real-pattern covariates that are meaningful because the data are real time. For example, the perimeter area ratio is defined on categorical data (flooded or unflooded), and the long-term average is continuous; thus a threshold would have to be applied before calculating the perimeter area ratio.

The fifth and final data set, the real-long data, were calculated with both long-term and real-time covariates. Only observations for the low-n data set were used. Although we expect environmental covariates for real-time and long-term flooded cover types to be correlated with one another, we do not expect this to affect prediction accuracy with boosted regression trees (described in the next section), although correlated variables can be

problematic for determining variable performance (Elith et al., 2008).

Statistical approach and performance metrics

We used boosted regression trees (Elith et al., 2008) to model habitat suitability for each shorebird taxon. We used the presence/absence of each taxon as the response with a binomial distribution, where site observations were weighted by abundance (log[abundance +2]) to account for large variation in the number of birds present in each observation. Weighting improves model performance compared to boosted regression trees where all observations receive equal weighting (Yu et al., 2020). All models were implemented using the gbm.step function in the dismo package (Hijmans et al., 2017). Although boosted regression trees can handle highly correlated explanatory variables (Buston & Elith, 2011), we did not include any two variables that had correlation coefficients higher than 0.65 even across spatial scales, with the exception of the models that combined long-term and real-time data or used real-time, pattern data. To identify parsimonious models, we removed variables with <2% relative influence because variables with <2% importance were typically removed with the gbm.simplify function. Finally, we created an ensemble of two models: the south model was trained on data from the Tulare, San Joaquin,

and Delta Basins of the Central Valley (see Figure 1); and the north model was trained on all basins north of the Delta (American, Butte, Colusa, Suisun, Sutter, and Yolo Basins). Following other spatial ensemble-based techniques (Fink et al., 2010), we created two models because the two regions had very different habitat availabilities and data availabilities. (See Appendix S1: Section S1.) In the North and South Valleys, respectively, we used 4,536 and 6,641 bird observations, roughly 46% and 36% from structured surveys, across 1.08 million hectares and 2.42 million hectares. The ensemble was weighted according to the true skill statistic (TSS) calculated before the ensemble: The northern prediction was weighted twice as high as the southern prediction when creating the ensemble for the northern region and similarly for the south. All reported performance metrics are calculated after the ensemble.

Validation data to explore variability in flooding

To test whether real-time covariates would better forecast bird observations when surface water was highly variable, we explored model performance across month, year, and spatial location. To test across years, we used five different training and validation data sets (Table 2). Four validation subsets were defined by year, where withheld data were restricted to 2008–2009, 2010–2011, 2013–2014,
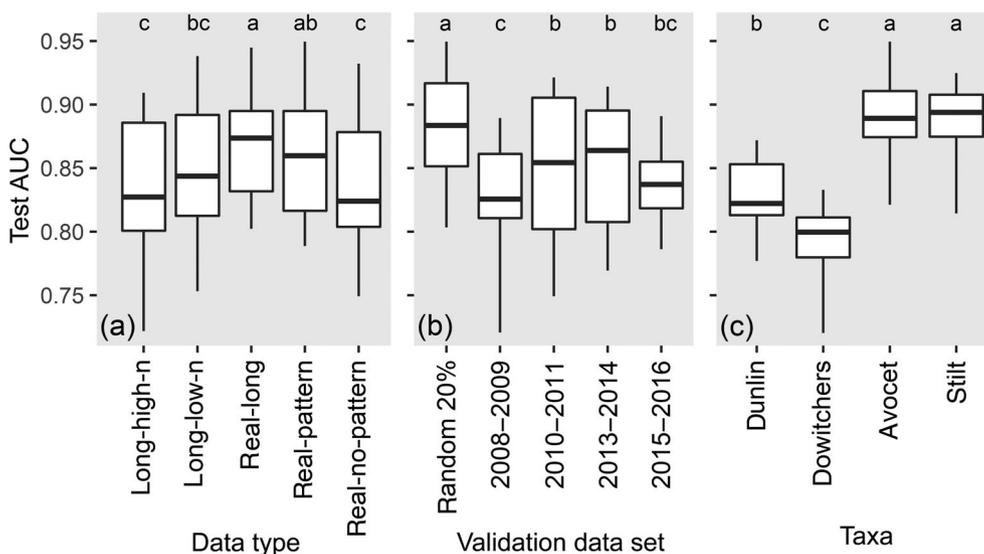

FIGURE 2 Area under the curve values (or AUCs) across (a) data type: long-high-n, long-low-n, real-long, real-pattern, and real-no-pattern; (b) validation data set; and (c) taxa. Validation data sets and data types are described in Table 2. Box plots show the median, interquartile range, and minimum and maximum AUCs. Letters (a, b, c) designate which data types lead to significantly different performance in a Tukey test. There were 20 AUC values contributing to the median and spread in boxplots across taxa-validation data sets and species-data types in (a) and (b), respectively (see Appendix S1: Figure S3 to see all points specifically). There were 25 values contributing to the median and spread in boxplots in (c)

and 2015–2016, using 2-year time periods to incorporate at least one full water year. Validation data sets based on 2-year increments were also chosen because there was enough training data in both wet (2005, 2006, and 2011) and dry years (2001, 2002, 2007, 2008, 2009, and 2013–2015; Reiter et al., 2018) and there were enough observations for testing models in the validation data set. Across these validation subsets, we were most interested in whether real-time data increased model performance in 2013–2014 and, to a slightly lesser degree the 2015–2016 validation subset, because they overlapped the California drought (2013–2016). Validation subsets chosen by year were compared to a validation data subset created by randomly selecting 20% of the data from across years; a standard approach for distribution modeling (Franklin, 2010). Across months, we were most interested in October, when the start of rice flooding typically depends on that year's water availability (see Appendix S1: Section S2 and Figure S4), leading to higher interannual surface water variability. We compared AUC (area under the receiver operating characteristic curve) on withheld subsets for all months in Appendix S1: Section S2 (Figure S5). Finally, we wanted to look at performance in different cover types with different irrigation regimes, comparing, for example, consistently flooded wetlands to intermittently flooded locations or crop types. Intermittently flooded grid cells were defined as those with 30%–70% fractional flooding across years in long-term average surface water layers (see Appendix S1: Sections S3 and Figure S6).

Because previous work has highlighted the trade-offs across different performance metrics (Allouche et al., 2006; Elith et al., 2008), we report AUC values here and TSS and kappa values in Appendix S1: Section S4. Differences in performance were analyzed using ANOVA (glm and Anova functions in R) with categorical explanatory variables for taxa (the three species and the combination of the two dowitcher species), validation data subset, and data-set type. For each of the five data-set types (e.g., long-low-n, real-pattern), there were four taxa for each of five data subsets, resulting in 20 observations per data-set type. We performed follow-up Tukey's tests (aov and HSD.test functions) to determine which data-set types had significantly different performance.

Spatial agreement

In addition to model performance, we determined whether the most suitable sites were similar across models built with all five data types. As with model performance, we expected that spatial agreement would be lowest in variable months and years. Thus, from the over

32 2-week, real-time surface water mosaics available for each month, we chose to calculate all real-time predictions in October across 2013–2015, the height of the drought, to compare against the long-term average prediction for October. In spatial comparisons between real-time and long-term average predictions, we averaged real-time layers across 2013–2015, preserving the unusual subset of years, but avoiding inferences based on a single, potentially idiosyncratic, 2-week period. We randomly sampled suitability at 100,000 points across the Central Valley and calculated the Pearson correlation coefficient between suitability across all pairwise combinations of data type. We also performed this calculation for the North and South Valleys separately.
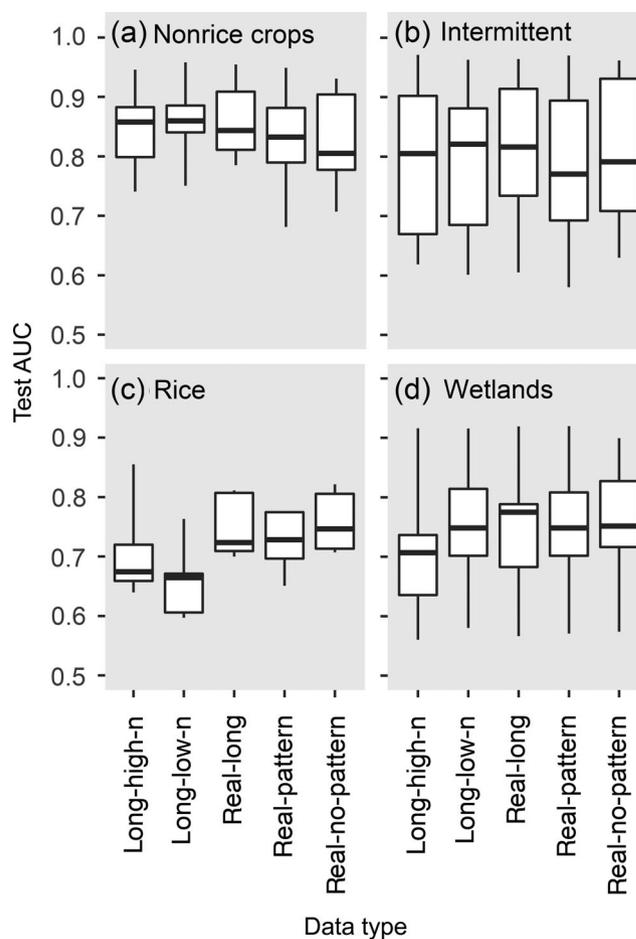

FIGURE 3 Area under the curve values (or AUCs) across different data types for different cover types: (a) nonrice crops, (b) intermittently flooded areas, (c) rice, and (d) wetlands. For (c) rice, all boxplots represent 5 AUC values, except for the high-n data which represents 6, because some species and validation datasets did not have enough data points (<10) to calculate AUCs. For the rest of the panels, there were 10 and 11 AUC values/boxplot, which is still less than the 20 values contributing to boxplots in Figure 2a. Box plots show the median, interquartile range, and minimum and maximum AUCs across species and validation subsets. The data types and validation data sets are described in Table 2

RESULTS

All models evaluated had useful discriminatory accuracy (as defined in Elith et al., 2006 as AUC > 0.75), where the average AUC across the five validation subsets of data was greater than 0.78 for the 20 taxa by data-type combinations. We found very little difference in predictive performance across the five data types (Figure 2 and Appendix S1: Figures S3, S5, and Section S4); however, small differences between models were significant ($\chi^2 = 36.7$, $df = 4$, $p = 2.1 \times 10^{-7}$). In follow-up tests (Tukey's honestly significant difference [HSD]) across taxa, models built on real-long and real-pattern data were grouped together as having the highest AUC (marked with an "a" in Figure 2a). Models built on real-pattern and long-low-n data were grouped together with the second highest AUC (b in Figure 2a). Models built on real-no-pattern, long-low-n, and long-high-n data were grouped together with the lowest AUC (c in Figure 2a). Taken together, across taxa and data subsets, there was a small benefit of using pattern metrics and using both real-time and long-term covariates. These results were similar for TSS, with some differences in follow-up tests. Kappa showed no difference across models (Appendix S1: Section S4).
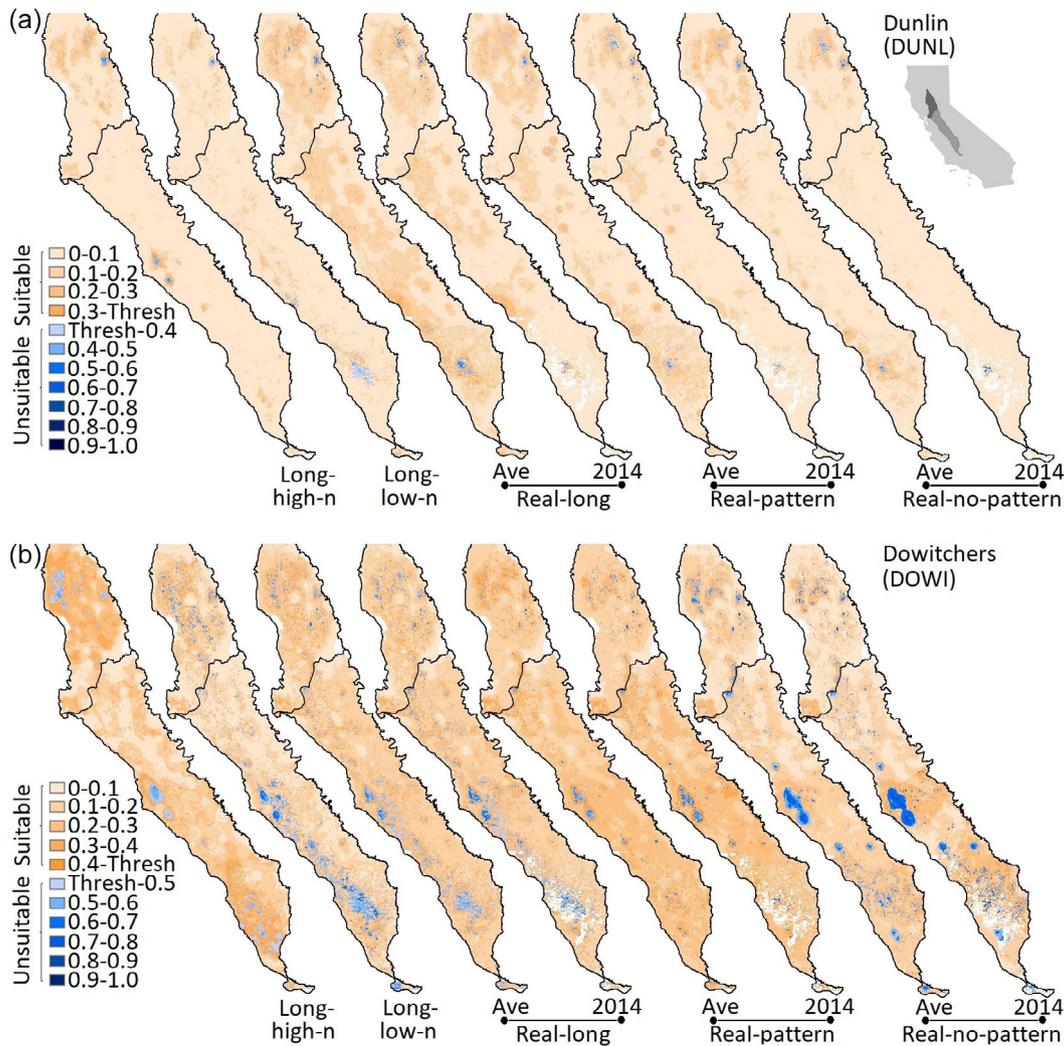

FIGURE 4   Dunlin (a) and dowitcher (b) spatial predictions for models built with long-term covariates on the larger data set (long-high-n), long-term covariates on the smaller data set (long-low-n), long-term and real-time covariates (real-long; Ave and 2014), real-time models with pattern metrics (real-pattern; Ave and 2014), and real-time models with no pattern metrics (real-no-pattern; Ave and 2014). For the real-time prediction pairs, a single mosaic for 12–21 October 2014 is shown to the right of the average of annual predictions from all real-time Landsat 8 October mosaics (Ave). White areas in 2014 predictions are due to cloud cover. The thick black lines are the boundaries of the California Valley Joint Venture and delineation of North and South Valley. Orange areas can be thought of as unsuitable based on the threshold that maximized kappa. Thresholds differed across species and data type used to create the model (see Appendix S1: Table S6 for threshold values), where the thresholds lie between 0.3 and 0.4 for dunlin and 0.4 and 0.5 for dowitchers (excepting the real-no-pattern dowitcher map, which had the same color scale as the dunlin map)

There were also significant differences in AUC across different training–testing subsets of the data ($\chi^2 = 69.0$, $df = 4$, $p < 3.7 \times 10^{-14}$; Figure 2b). Significantly higher AUCs were observed when a random 20% of the data were withheld for validation (Figure 2b), as compared to when validation data were withheld according to year. For the extreme drought 2013–2014 validation data, we did not find a significant interaction of data-type-by-subset. Thus, contrary to expectation, models built with real-time data did not perform better in drought years. Finally, there was a significant effect of taxa ($\chi^2 = 357.5$, $df = 3$, $p < 2.2 \times 10^{-16}$) with the smaller, Nearctic breeding taxa that are more likely to travel in flocks—dunlin

and dowitchers—having significantly lower AUC values than the larger-bodied, resident-breeding avocet and stilt (Figure 2c).

We did not see increased performance of real-time data across any months (see Appendix S1: Section S2 and Figure S5). Overall, across all models, July–September had higher AUCs, especially for stilt and avocet, with median AUC > 0.80 in all cases (except the avocet long-low-n case). Because some month–taxa combinations did not have enough presence observations (which we defined to be 10 or more) to calculate performance metrics, we did not run statistical tests across months.
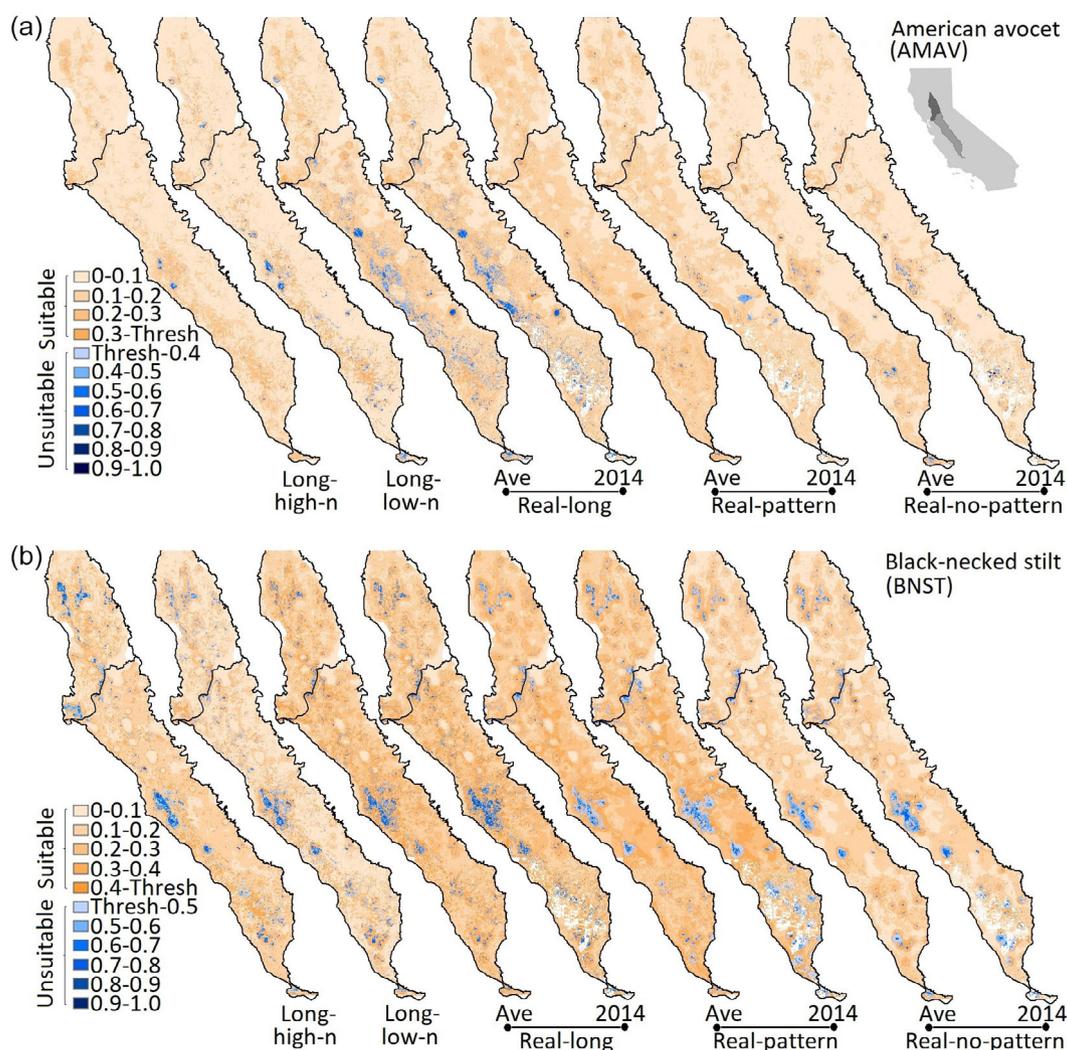

FIGURE 5 American avocet (a) and black-necked stilt (b) spatial predictions for models built with long-term covariates on the larger data set (long-high-n), long-term covariates on the smaller data set (long-low-n), long-term and real-time covariates (real-long; Ave and 2014), real-time models with pattern metrics (real-pattern; Ave and 2014), and real-time models with no pattern metrics (real-no-pattern; Ave and 2014). For the real-time prediction pairs, a single mosaic for 12–21 October 2014 is shown to the right of the average of annual predictions from all real-time Landsat 8 October mosaics (Ave). White areas in 2014 data are due to cloud cover. The thick black lines are the boundaries of the California Valley Joint Venture and delineation of North and South Valley. Orange areas can be thought of as unsuitable based on the threshold that maximized kappa. Thresholds differed across species and data type used to create the model (see Appendix S1: Table S6 for threshold values), where the thresholds lie between 0.3 and 0.4 for avocet and 0.4 and 0.5 for black-necked stilt

We found limited evidence that separating the data by cover type improves the relative performance of real-time data (Figure 3). In more variably flooded rice (Figure 3c), the three models using real-time data had median AUC > 0.70 and models with only long-term data had median AUC < 0.70. However, the performance advantage of the real-time data was not present in intermittently flooded areas (Figure 3b), where median AUC > 0.80 occurred for models with long-low-n and real-long data and AUC < 0.80 for all other data types. In flooded nonrice, nonorchard crops (Figure 3a)—a cover type concentrated in the South Valley—all models with long-term data had median AUC > 0.82 and models with real-time only data had AUC < 0.82. Compared to the
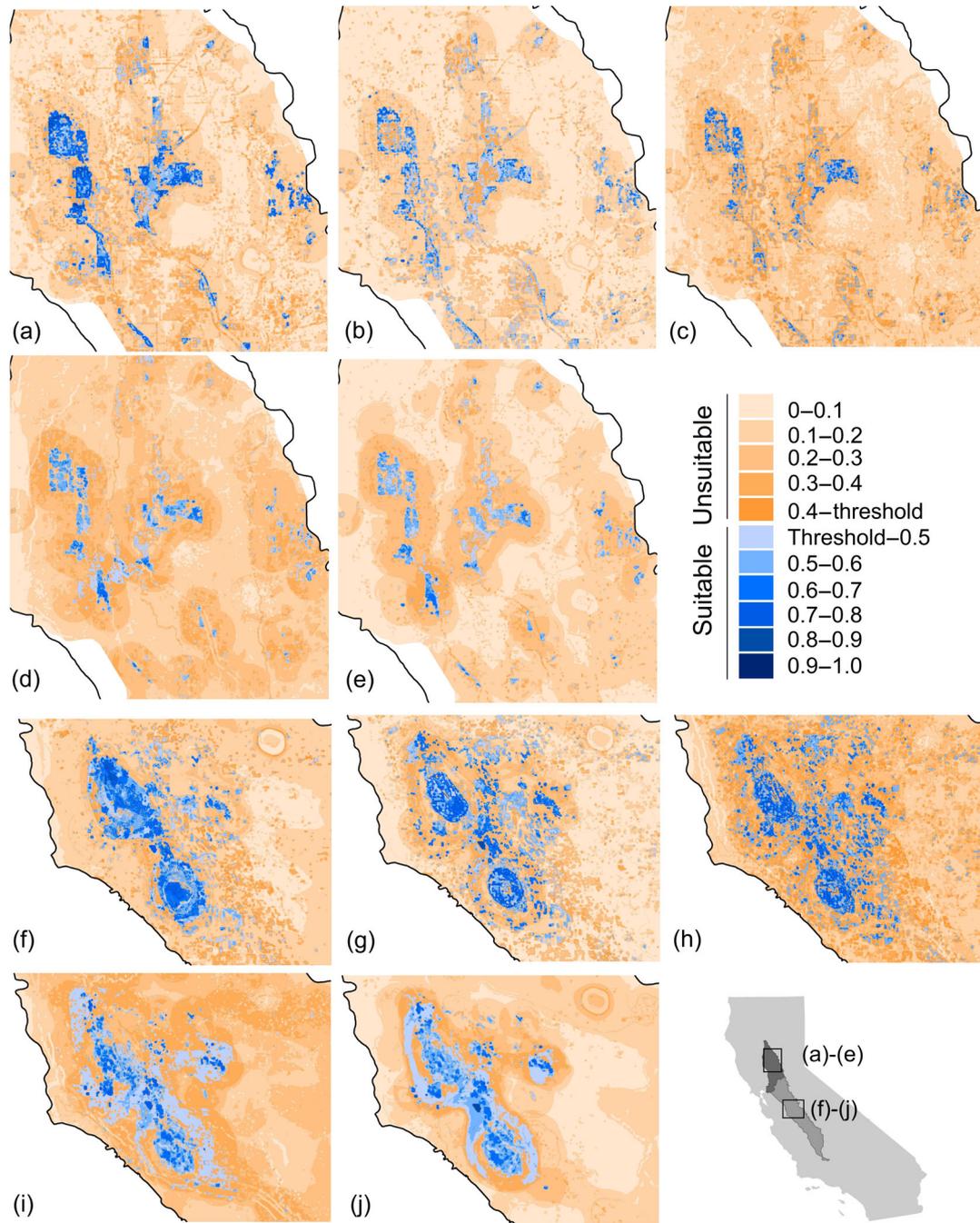

FIGURE 6 Black-necked stilt suitability in (a)–(e) the northern California Central Valley, and the Grasslands Ecological Area in the southern Central Valley (f)–(j) (see inset with bounding boxes) under the four types of data (described in Table 2), (a and f: long-high-n; b and g: long-low-n models; c and h: real-long; d and i: real-pattern; and d and i: real-no-pattern). Thresholds, defined by the suitability value that maximized kappa, differed across data type (see Appendix S1: Table S6 for threshold values), but all of them lay between 0.4 and 0.5. The spatial scale (1:1,000,000) is the same across (a)–(j)

AUC values across the entire landscape (Figure 2), AUC values for specific cover types (Figure 3) were lower because fewer observations can be discriminated into presences and absences when a smaller portion of landscape variability is considered. We did not perform significance tests on performance-by-cover analyses because there were not enough observations in each cover type (e.g., there were too few observations [<10] for avocet in rice to calculate AUC; see points in Appendix S1: Figure S8 for taxa-validation-data combinations with enough data to calculate AUC).

Maps of suitable habitat across the Central Valley (Figures 4 and 5) created from models built with the 80% training-data subset (or 20% validation subset) revealed that all taxa except dunlin had the highest predicted occurrence in flooded wetlands (also see variable importance in Appendix S1: Section S5 and Figures S9 and S10). Much of the nonrefuge suitable habitat in the north valley is on flooded rice, and much of the suitable habitat in the south valley is on flooded row, field, or grain crops. Avocet differed from the other taxa in having high prevalence in wastewater treatment sites (Appendix S1: Section S5 and Figure S10).

Correlations of suitability from predictions based on each of the five types of data showed similar patterns to Figure 6, which displays high-resolution stilt predictions in two regions of interest—North Valley refuges and Grasslands Ecological Area in the South Valley. Stilt predictions in Figure 6 and Table 3 had more agreement and higher correlations ($0.52 \leq r \leq 0.86$; Table 3) than other taxa ($0.40 \leq r \leq 0.83$). Across taxa, when comparing the two real-time predictions to each other or the two long-term predictions to each other, correlations were higher (0.47–0.86) than when comparing real-time predictions to long-term predictions (0.44–0.71). In correlations across layers for the North and South Valley in

isolation (Tables 4 and 5), the South Valley showed less agreement (0.19–0.86) than the North Valley (0.46–0.92), a result also apparent in comparisons of Figures 6a–e to Figures 6f–j. Differences in spatial predictions were not reflective of substantially different performance ranking across models built with real-time versus long-term data in either the North or South Valley (Appendix S1: Figure S11). Not surprisingly, when predictions were compared across locations where birds were observed, there was greater correlation between the different data types (0.66–0.96; Appendix S1: Section S7 and Table S5).

Although correlations across predictions were moderate to high, prediction differences led to different footprints of suitable habitat when habitat was defined as "suitable"—namely, when its suitability value was above the suitability value that maximized Cohen's kappa (see Appendix S1: Section S8 for details and Table S6 for specification of thresholds). Some of the differences in footprint might be expected because the real-time footprint was driven by surface-water data from the 2013–2015 drought years. In the North Valley, for all taxa but dunlin, the two sets of long-term predictions had suitability footprints that were 20%–370% larger than either of the two sets of real-time predictions. Dunlin had 30%–70% less habitat with long-term than with real-time predictions. In Figure 6, the long-term predictions in the North Valley and the South Valley (Figure 6a,b,f,g) had a larger footprint of suitable habitat compared to the real-pattern and real-no-pattern (Figure 6d,e,i,j). However, higher suitable areas with long-term data were not uniform across the South Valley, where there was no consistent pattern across taxa. With more disagreement across predictions in the South Valley (Table 5), prominent features, like the extended suitability in the 12–21 October 2014 predictions around the Grasslands Ecological Area

TABLE 3 Pearson's correlation coefficients for a random sample of 100,000 points across the predictions within the Central Valley

Data Type	Long-term, low n		Real-time and long-term		Real-time, no pattern		Real-time, pattern	
Long-term, high n	*0.47*	*0.62*	0.40	0.53	*0.62*	*0.52*	0.50	0.49
	0.77	*0.86*	0.69	0.75	*0.59*	*0.71*	0.48	0.51
Long-term, low n			**0.68**	**0.85**	*0.61*	*0.62*	0.51	0.51
			0.83	**0.86**	*0.62*	*0.72*	0.44	0.53
Real-time and long-term					*0.66*	*0.61*	**0.77**	**0.76**
					0.55	*0.66*	**0.57**	**0.72**
Real-time, no pattern							**0.80**	**0.75**
							0.63	**0.75**

Note: Four numbers are listed for each grouping, with the correlation coefficient for dunlin in the upper left, dowitchers in the upper right, avocet in the lower left, and black-necked stilt in the lower right. To highlight the differences across models built with different data types, values for each group are italicized if the correlation coefficient, averaged across species, is between 0.6 and 0.7 and are in bold if the average correlation coefficient is >0.7.

TABLE 4 Pearson's correlation coefficients for a random sample of 100,000 points across predictions within the Central Valley north of the Delta basin (see Figure 1)

Data Type	Long-term, low n		Real-time and long-term		Real-time, no pattern		Real-time, pattern	
Long-term, high n	**0.77**	**0.63**	*0.69*	*0.62*	**0.81**	**0.62**	*0.74*	*0.64*
	0.67	**0.90**	*0.63*	*0.84*	**0.58**	**0.79**	*0.49*	*0.63*
Long-term, low n			**0.76**	**0.85**	*0.73*	*0.64*	0.67	0.57
			0.82	**0.91**	*0.59*	*0.77*	0.46	0.61
Real-time and long-term					**0.80**	**0.75**	**0.87**	**0.84**
					0.52	**0.75**	**0.48**	**0.77**
Real-time, no pattern							**0.92**	**0.85**
							0.67	**0.84**

Note: Four numbers are listed for each grouping, with the correlation coefficient for dunlin in the upper left, dowitchers in the upper right, avocet in the lower left, and black-necked stilt in the lower right. To highlight the differences across models built with different data types, values for each group are italicized if the correlation coefficient, averaged across species, is between 0.6 and 0.7, and are in bold if the average correlation coefficient is >0.7.

TABLE 5 Pearson's correlation coefficients for a random sample of 100,000 points across predictions within the Central Valley south of the northern border of the Delta basin (see Figure 1)

Data Type	Long-term, low n		Real-time and long-term		Real-time, no pattern		Real-time, pattern	
Long-term, high n	*0.35*	*0.65*	0.19	0.55	0.40	0.55	0.29	0.50
	0.78	*0.84*	0.70	0.72	0.59	0.68	0.46	0.47
Long-term, low n			**0.67**	**0.86**	*0.60*	*0.62*	0.45	0.50
			0.84	**0.85**	*0.62*	*0.71*	0.42	0.50
Real-time and long-term					0.58	0.56	*0.71*	*0.72*
					0.52	0.63	*0.56*	*0.70*
Real-time, no pattern							*0.70*	*0.71*
							0.62	*0.72*

Note: Four numbers are listed for each grouping, with the correlation coefficient for dunlin in the upper left, dowitchers in the upper right, avocet in the lower left, and black-necked stilt in the lower right. To highlight the differences across models built with different data types, cells of the table are italicized if the correlation coefficient, averaged across species, is between 0.6 and 0.7 and are in bold if the average correlation coefficient is >0.7.

(center of the Valley in Figure 4), can have a big impact on differences across suitable footprints.

DISCUSSION

Given the dynamic availability of surface water in the Central Valley of California, we expected that models of suitable habitat for shorebirds using frequent, real-time measurements of surface water would outperform models built with long-term average surface water data, especially in years, months, and locations that showed greater deviance from the long-term average. However, we found little evidence to support this hypothesis. Further, when we withheld validation data from the 2013–2014 drought, models built with real-time data did not outperform the long-term average despite changes in surface water availability during these unusually dry years (Reiter et al., 2018). Instead, a random subsample of validation data yielded significantly higher AUC values, suggesting that random subsamples—a standard practice—may inflate performance because of higher variability in occurrence relationships across years. Similarly, there was no clear pattern of relative model performance across months, even in October, where the timing of post-harvest flooding for rice makes this month highly variable (Reiter et al., 2018; Appendix S1: Section S2 and Figure S5). Instead, summer months had slightly higher performance for black-necked stilts and avocets, specifically avocets in July and stilts in September (median AUC ≥ 0.85; Appendix S1: Section S2 and Figure S5), possibly because shorebirds have no choice but to visit

the few regularly flooded locations. Studies of shorebirds in the Prairie Potholes region support the idea that drier conditions concentrate individuals in more permanent wetlands (Steen et al., 2018). Overall, avocet and stilts had higher performance (as measure by AUC) than dunlin and dowitchers, likely because avocet and stilts depend on predictably flooded wetlands (Barbaree et al., 2018; Shuford et al., 1998). However, the preference of dunlin and dowitchers for more variably flooded habitat did not translate into higher performance in models built with real-time data. Finally, in intermittently flooded rice there may have been a benefit of real-time data, but not in other crops and not when explicitly defining intermittently flooded areas.

One possible explanation for similar performance across long-term and real-time data is that surface water across the Central Valley is heavily managed (in part due to the Refuge Water Supply Program) and surprisingly predictable across most months (Reiter et al., 2018), despite increasing annual variability in precipitation (He & Guatam, 2016). This may be one reason why our results contradict previous studies that found weather covariates were better predictors of Australian birds and bettongs than climate averages, especially during drought conditions (Bateman et al., 2012; Reside et al., 2010). Similarly, real-time data may be more important for predicting shorebirds in rain-fed wetlands, such as the Prairie Pothole region (Steen et al., 2018), as opposed to the hydrologically managed Central Valley.

In addition to the degree of flooding at a site during a given time, year-to-year consistency in flooding may be important to predicting shorebird occurrence. Even in months like October where the timing of flooding varies year to year, that variability is mostly defined by the date of first flood and not whether the area eventually floods (Appendix S1: Section S2 and Figure S5). Assuming predictability in flooded locations, shorebirds may visit the same locations each year because locations that had high water availability in a given month in previous years would serve as a good indicator of locations with high water availability in future years. This result is consistent with previous research demonstrating Central Valley shorebird movement is correlated to long-term average water availability (Barbaree et al., 2018), suggesting the possibility of some degree of site fidelity. Further, long-term average covariate layers, by virtue of being composed of pixels varying from 0 to 1, capture some information regarding a location's previous flooding frequency.

One noteworthy exception where real-time data performed better than long-term data was in predicting habitat in rice fields (Figure 3), a less predictably flooded cover type than wetlands and the cover type most

represented in the structured data. Part of the structured data was designed to measure the impact of flooding, and flood timing on shorebird occurrence in rice fields (Golet et al., 2018). Further studies observing shorebirds in a wider variety of cover types would help better characterize the importance of real-time data. Such future studies would more closely match previous work that examined shorebird habitat choice in mid-continental wetlands (Albanese & Davis, 2013; Gillespie & Fontaine, 2017; Steen et al., 2018).

Large, unbiased data sets that have spatially representative data across time are needed to tease apart the relative performance benefit of real-time versus long-term average covariates, especially when there is disagreement in spatial predictions across models. We found high correlation coefficients (0.65–0.97) in locations where birds were observed (Appendix S1: Section S7 and Table S5) and lower correlations in a random sample of locations, especially in the South Valley, which had 65% of the sampling density as the North Valley (Tables 4 and 5). Despite differences in spatial predictions, performance was similar for models built with real-time and long-term data (Figure 2) and when the North Valley and South Valley were considered separately (Appendix S1: Figure S11). Taken together, the lack of performance differences between real-time and long-term data and increasing spatial disagreement in locations with lower sampling density suggest the need for more data to determine better which models are most accurate. Specifically, more sampling is needed in the South Valley and in locations where bird occurrence may change through time as a function of intermittent flooding.

There were still relatively few bird observations in variably flooded, moderately suitable locations despite the combination of two types of data that provide complementary information in the Central Valley (Robinson et al., 2020). Namely, the structured data provide observations in suitable shorebird cover types (e.g., rice) and eBird data in less suitable cover types (e.g., near urban areas). Thus, although more data in intermittently flooded areas would help distinguish the benefit of real-time models, the available data had more observations in intermittently flooded areas (4%–31%; see Appendix S1: Section S3) than were present on the broader landscape (1%–8% of all land). Despite the need for more bird sampling in underrepresented regions and cover types, more data did not increase performance in models built with the high-n data set, where the high-n data include eBird traveling counts (where observers moved up to 250 m) and bird observations with no real-time water measurements associated with them. This result is consistent with other studies that found no performance increase with additional, lower-quality data (Reside et al., 2011). Still,

all else being equal, larger data sets typically increase performance (Fernandes et al., 2019), and our results highlight the need for additional sampling in moderately suitable areas (e.g., flooded nonrice agriculture) and undersampled areas (the South Valley).

Although it is assumed that highly mobile species will move to new locations to take advantage of more suitable habitat, there are ecological reasons why they might not. High site fidelity may lead to higher long-term fitness (Bradshaw et al., 2004), and some species show high site fidelity in highly deviant years even when there is a fitness cost (Abrahms et al., 2018). Nonbreeding shorebirds in disturbed areas had lower survival when they were site faithful as opposed to when they emigrated to less disturbed sites (Gibson et al., 2018). Consistent with our results that show the best performing models used a combination of both real-time and long-term data, some taxa, such as whales, may respond to both historical and contemporaneous environmental conditions (Abrahms et al., 2019). Previous studies have demonstrated that shorebirds typically shift their distributions as a function of dynamic within-year water availability, both in the Central Valley (Barbaree et al., 2018, 2020) and elsewhere (Maclean et al., 2008; Steen et al., 2018). The tendency to visit the same site every year may simply be a function of available habitat. In the North Valley, dunlin and dowitchers left in late winter as flooded habitat declined. In contrast, in the Grasslands Ecological Area, dunlin and dowitchers remained throughout the winter and into the spring because the Grasslands Ecological Area provided consistently good habitat in a region with little alternative habitat (Barbaree et al., 2018). Regardless, future work might focus on the degree to which shorebirds in the heavily irrigated Central Valley display high fidelity across years and in stopover and wintering sites.

Adding pattern metrics provided moderate improvement in models built with real-time data and both real-time and long-term covariates. Fractal dimension and perimeter area metrics were used as proxies for surface water depth, an important environmental covariate for shorebirds (Skagen & Knopf, 1994). Including a metric for connected water bodies across larger spatial extents, represented by the cohesion metric, also improved predictions, which is consistent with the findings of Barbaree et al. (2018). Although these metrics did significantly improve the models, their benefits were not quantitatively large. Across landscapes, pattern metrics have been shown to be correlated with the amount of a particular cover type (Cushman et al., 2008). In our study they appeared to provide limited value to the models beyond what was already captured by including information on the prevalence of cover types.

Finally, it is important to consider the added benefit of real-time data, given that real-time data can be costly to collect, computationally intensive to analyze and process, and can lead to predictions that are more difficult to interpret. Although real-time data did not perform better than long-term data, they also did not perform markedly worse, and there was enough spatial disagreement in real-time versus long-term predictions to suggest that they provided slightly different information for habitat prioritization. Thus, if the goal of a particular conservation decision is to determine potential habitat in a dry versus wet year, using real-time data would allow for this difference to be estimated. Thus, the extra effort to use real-time data could be justified to motivate dynamic conservation actions (as in Reynolds et al., 2017). However, for more long-term conservation land acquisition, models built with long-term data performed well and could continue to be used, as they have been used historically, to determine overall suitable habitat across years for a variety of taxa efficiently. Ultimately, the collection of real-time data is worthwhile and should be pursued because real-time data are often the foundation for long-term average data.

In summary, we found a nuanced trade-off between real-time and long-term covariates in modeling suitable shorebird habitat, where some performance testing was ultimately limited by data availability, especially in the southern portion of the Central Valley. One possible explanation for the comparable performance between real-time and long-term covariates is that agricultural management in the Central Valley makes surface water more predictable, a different situation than in rain-fed, mid-continental wetlands (Skagen & Knopf, 1993). Despite the nuances, real-time and long-term data both performed well and may have distinct value for conservation and management decisions. Models built with real-time data can be thought of as providing the optimally suitable contemporaneous sites, whereas models built with long-term data can be thought of as providing a version of long-term, optimal suitability.

ACKNOWLEDGMENTS

does not imply endorsement by the U.S. Government. This is Point Blue Conservation Science contribution number 2302.

ORCID
Erin E. Conlisk https://orcid.org/0000-0002-4049-4805
Gregory H. Golet https://orcid.org/0000-0001-9805-0767
Mark D. Reynolds https://orcid.org/0000-0001-9958-6360
Blake A. Barbaree https://orcid.org/0000-0002-5912-8768
Kristin A. Sesser https://orcid.org/0000-0003-2761-7388
Kristin B. Byrd https://orcid.org/0000-0002-5725-7486
Sam Veloz https://orcid.org/0000-0003-3401-9167
Matthew E. Reiter https://orcid.org/0000-0002-0587-786X

REFERENCES
Abrahms, B., E.L. Hazen, E.O. Aikens, M.S. Savoca, J.A. Goldbogen, S.J. Bograd, M.G. Jacox, L.M. Irvine, D.M. Palacios, and B.R. Mate. 2019. "Memory and Resource Tracking Drive Blue Whale Migrations." *Proceedings of the National Academy of Sciences of the United States of America* 116: 5582–7.

Abrahms, B., E.L. Hazen, S.J. Bograd, J.S. Brashares, P.W. Robinson, K.L. Scales, D.E. Crocker, and D.P. Costa. 2018. "Climate Mediates the Success of Migration Strategies in a Marine Predator." *Ecology Letters* 21: 63–71.

Aiello-Lammens, M.E., R.A. Boria, A. Radosavljevic, B. Vilela, and R.P. Anderson. 2015. "spThin: An R Package for Spatial Thinning of Species Occurrence Records for Use in Ecological Niche Models." *Ecography* 38: 541–5.

Albanese, G., and C.A. Davis. 2013. "Broad-Scale Relationships Between Shorebirds and Landscapes in the Southern Great Plains." *The Auk* 130: 88–97.

Allouche, O., A. Tsoar, and R. Kadmon. 2006. "Assessing the Accuracy of Species Distribution Models: Prevalence, Kappa and the True Skill Statistic (TSS)." *Journal of Applied Ecology* 43: 1223–32.

Barbaree, B.A., M.E. Reiter, C.M. Hickey, N.K. Elliott, D. Schaffer-Smith, M.D. Reynolds, and G.W. Page. 2018. "Dynamic Surface Water Distributions Influence Wetland Connectivity within a Highly Modified Interior Landscape." *Landscape Ecology* 33: 829–44.

Barbaree, Blake A., Matthew E. Reiter, Catherine M. Hickey, and Gary W. Page. 2015. "Day and Night Habitat Associations of Wintering Dunlin (*Calidris alpina*) within an Agriculture–Wetland Mosaic." *Waterbirds* 38: 40–6.

Barbaree, B.A., M.E. Reiter, C.M. Hickey, K.M. Strum, J.E. Isola, S. Jennings, L.M. Tarjan, C.M. Strong, L.E. Stenzel, and W.D. Shuford. 2020. "Effects of Drought on the Abundance and Distribution of Non-Breeding Shorebirds in Central California, USA." *PLoS One* 15: e0240931.

Bateman, B.L., J. VanDerWal, and C.N. Johnson. 2012. "Nice Weather for Bettongs: Using Weather Events, Not Climate Means, in Species Distribution Models." *Ecography* 35: 306–14.

Bauer, S., P. Gienapp, and J. Madsen. 2008. "The Relevance of Environmental Conditions for Departure Decision Changes en Route in Migrating Geese." *Ecology* 89: 1953–60.

Bradshaw, C.J.A., M.A. Hindell, M.D. Sumner, and K.J. Michael. 2004. "Loyalty Pays: Potential Life History Consequences of Fidelity to Marine Foraging Regions by Southern Elephant Seals." *Animal Behaviour* 68: 1349–60.

Buckley, L.B., M.C. Urban, M.J. Angilletta, L.G. Crozier, L.J. Rissler, and M.W. Sears. 2010. "Can Mechanism Inform Species' Distribution Models?" *Ecology Letters* 13: 1041–54.

Burns, P., M. Clark, L. Salas, S. Hancock, D. Leland, P. Jantz, R. Dubayah, and S.J. Goetz. 2020. "Incorporating Canopy Structure from Simulated GEDI Lidar into Bird Species Distribution Models." *Environmental Research Letters* 15: 095002.

Buston, P.M., and J. Elith. 2011. "Determinants of Reproductive Success in Dominant Pairs of Clownfish: A Boosted Regression Tree Analysis." *Journal of Animal Ecology* 80: 528–38.

Conlisk, E., G. Golet, M. Reynolds, B. Barbaree, K. Sesser, K. Byrd, S. Veloz, and M. Reiter. 2021. "Both Real-Time and Long-Term Environmental Data Perform Well in Predicting Shorebird Distributions in Managed Habitat." *Dryad data set.* https://doi.org/10.5061/dryad.n8pk0p2w0

Connell, J., S.J. Watson, R.S. Taylor, S.C. Avitabile, R.H. Clarke, A.F. Bennett, and M.F. Clarke. 2017. "Testing the Effects of a Century of Fires: Requirements for Post-Fire Succession Predict the Distribution of Threatened Bird Species." *Diversity and Distributions* 23: 1078–89.

Cushman, S.A., K. McGarigal, and M.C. Neel. 2008. "Parsimony in Landscape Metrics: Strength, Universality, and Consistency." *Ecological Indicators* 8: 691–703.

D'Aloia, C.C., I. Naujokaitis-Lewis, C. Blackford, C. Chu, J.M.R. Curtis, E. Darling, F. Guichard, et al. 2019. "Coupled Networks of Permanent Protected Areas and Dynamic Conservation Areas for Biodiversity Conservation Under Climate Change." *Frontiers in Ecology and Evolution* 7: 27.

Dormann, C.F., S.J. Schymanski, J. Cabral, I. Chuine, C. Graham, F. Hartig, M. Kearney, et al. 2012. "Correlation and Process in Species Distribution Models: Bridging a Dichotomy." *Journal of Biogeography* 39: 2119–31.

Dybala, K., M.E. Reiter, C.M. Hickey, D.W. Shuford, and G. Yarris. 2017. "A Bioenergetics Approach to Setting Conservation Objectives for Non-Breeding Shorebirds in California's Central Valley." *San Francisco Estuary and Watershed Science* 15.

Elith, J., C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R.J. Hijmans, et al. 2006. "Novel Methods Improve Prediction of Species' Distributions from Occurrence Data." *Ecography* 29: 129–51.

Elith, J., J.R. Leathwick, and T. Hastie. 2008. "A Working Guide to Boosted Regression Trees." *Journal of Animal Ecology* 77: 802–13.

Fernandes, R.F., D. Scherrer, and A. Guisan. 2019. "Effects of Simulated Observation Errors on the Performance of Species Distribution Models." *Diversity and Distributions* 25: 400–13.

Fink, D., T. Auer, A. Johnston, V. Ruiz-Gutierrez, W.M. Hochachka, and S. Kelling. 2020. "Modeling Avian Full Annual Cycle Distribution and Population Trends with Citizen Science Data." *Ecological Applications* 30: e02056.

Fink, D., W.M. Hochachka, B. Zuckerberg, D.W. Winkler, B. Shaby, M.A. Munson, G. Hooker, M. Riedewald, D. Sheldon, and S. Kelling. 2010. "Spatiotemporal Exploratory Models for Broad-Scale Survey Data." *Ecological Applications* 20: 2131–47.

Flint, L.E., and A.L. Flint. 2014. California Basin Characterization Model: A Dataset of Historical and Future Hydrologic Response to Climate Change, (ver. 1.1, May 2017). U.S. Geological Survey Data Release. https://doi.org/10.5066/F76T0JPB (accessed August 1, 2017).

Franklin, J. 2010. *Mapping Species Distributions: Spatial Inference and Prediction*. New York: Cambridge University Press.

Frayer, W.E., D.D. Peters, and H.R. Pywell. 1989. *Wetlands of the California Central Valley: Status and Trends*. Portland, Oregon: U.S. Fish and Wildlife Service.

Gibson, D., M.K. Chaplin, K.L. Hunt, M.J. Friedrich, C.E. Weithman, L.M. Addison, V. Cavalieri, et al. 2018. "Impacts of Anthropogenic Disturbance on Body Condition, Survival, and Site Fidelity of Nonbreeding Piping Plovers." *The Condor* 120: 566–80.

Gillespie, C.R., and J.J. Fontaine. 2017. "Shorebird Stopover Habitat Decisions in a Changing Landscape." *The Journal of Wildlife Management* 81: 1051–62.

Golet, G.H., C. Low, S. Avery, K. Andrews, C.J. McColl, R. Laney, and M.D. Reynolds. 2018. "Using Ricelands to Provide Temporary Shorebird Habitat During Migration." *Ecological Applications* 28: 409–26.

He, M., and M. Gautam. 2016. "Variability and Trends in Precipitation, Temperature and Drought Indices in the State of California." *Hydrology* 3(2): 14. https://doi.org/10.3390/hydrology3020014

Hijmans, R.J., S. Phillips, J. Leathwick, and J. Elith. 2017. Dismo: Species Distribution Modeling. R Package Version 1.0-12. http://CRAN.R-project.org/package=dismo

MacGarigal, K., S.A. Cushman, and E. Ene. 2012. Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer Software Program Produced by the Authors at the University of Massachusetts, Amherst. http://www.umass.edu/landeco/research/fragstats/fragstats.html (accessed February 1, 2018).

Maclean, I.M.D., G.E. Austin, M.M. Rehfisch, J. Blew, O. Crowe, S. Delany, K. Devos, et al. 2008. "Climate Change Causes Rapid Changes in the Distribution and Site Abundance of Birds in Winter." *Global Change Biology* 14: 2489–500.

Mannocci, L., A.M. Boustany, J.J. Roberts, D.M. Palacios, D.C. Dunn, P.N. Halpin, S. Viehman, et al. 2017. "Temporal Resolutions in Species Distribution Models of Highly Mobile Marine Animals: Recommendations for Ecologists and Managers." *Diversity and Distributions* 23: 1098–109.

Middleton, A.D., J.A. Merkle, D.E. McWhirter, J.G. Cook, R.C. Cook, P.J. White, and M.J. Kauffman. 2018. "Green-Wave Surfing Increases Fat Gain in a Migratory Ungulate." *Oikos* 127: 1060–8.

Page, G.W., L.E. Stenzel, and J.E. Kjelmyr. 1999. "Overview of Shorebird Abundance and Distribution in Wetlands of the Pacific Coast of the Contiguous United States." *The Condor* 101: 461–71.

Pérez Navarro, M.Á., G. Sapes, E. Batllori, J.M. Serra-Diaz, M.A. Esteve, and F. Lloret. 2019. "Climatic Suitability Derived from Species Distribution Models Captures Community Responses to an Extreme Drought Episode." *Ecosystems* 22: 77–90.

Petrik, K., D. Fehringer, and A. Weverko. 2014. *Mapping Seasonal Managed and Semi-Permanent Wetlands in the Central Valley of California*. Rancho Cordova, California: Ducks Unlimited, Inc.

Pettorelli, N., W.F. Laurance, T.G. O'Brien, M. Wegmann, H. Nagendra, and W. Turner. 2014. "Satellite Remote Sensing for Applied Ecologists: Opportunities and Challenges." *Journal of Applied Ecology* 51: 839–48.

Pickens, B.A., and S.L. King. 2014. "Linking Multi-Temporal Satellite Imagery to Coastal Wetland Dynamics and Bird Distribution." *Ecological Modelling* 285: 1–12.

Pomeroy, A.C. 2006. "Tradeoffs Between Food Abundance and Predation Danger in Spatial Usage of a Stopover Site by Western Sandpipers, Calidris Mauri." *Oikos* 112: 629–37.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Reiter, M.E., N.K. Elliott, D. Jongsomjit, G.H. Golet, and M.D. Reynolds. 2018. "Impact of Extreme Drought and Incentive Programs on Flooded Agriculture and Wetlands in California's Central Valley." *PeerJ* 6: e5147.

Reiter, M.E., N. Elliott, S. Veloz, D. Jongsomjit, C.M. Hickey, M. Merrifield, and M.D. Reynolds. 2015. "Spatio-Temporal Patterns of Open Surface Water in the Central Valley of California 2000–2011: Drought, Land Cover, and Waterbirds." *Journal of the American Water Resources Association* 51: 1722–38.

Reiter, M.E., C.M. Hickey, G.W. Page, D.W. Shuford, and K.M. Strum. 2011. *A Monitoring Plan for Wintering Shorebirds in the Central Valley of California, Version 1.0*. Report to the California Landscape Conservation Cooperative. Petaluma, CA: PRBO Conservation Science.

Reside, A.E., J.J. VanDerWal, A.S. Kutt, and G.C. Perkins. 2010. "Weather, Not Climate, Defines Distributions of Vagile Bird Species." *PLoS One* 5: e13569.

Reside, A.E., I. Watson, J. VanDerWal, and A.S. Kutt. 2011. "Incorporating Low-Resolution Historic Species Location Data Decreases Performance of Distribution Models." *Ecological Modelling* 222: 3444–8.

Reynolds, M.D., B.L. Sullivan, E. Hallstein, S. Matsumoto, S. Kelling, M. Merrifield, D. Fink, et al. 2017. "Dynamic Conservation for Migratory Species." *Science Advances* 3: e1700707.

Robinson, O.J., Viviana. Ruiz-Gutierrez, M.D. Reynolds, G.H. Golet, M. Strimas-Mackey, and D. Fink. 2020. "Integrating Citizen Science Data with Expert Surveys Increases Accuracy and Spatial Extent of Species Distribution Models." *Diversity and Distributions* 26: 976–86.

Sánchez-Fernández, D., J.M. Lobo, and O.L. Hernández-Manrique. 2011. "Species Distribution Models that Do Not Incorporate Global Data Misrepresent Potential Distributions: A Case Study Using Iberian Diving Beetles." *Diversity and Distributions* 17: 163–71.

Scales, K.L., E.L. Scales, M.G. Jacox, C.A. Edwards, A.M. Boustany, M.J. Oliver, and S.J. Bograd. 2017. "Scale of Inference: On the Sensitivity of Habitat Models for Wide-Ranging Marine Predators to the Resolution of Environmental Data." *Ecography* 40: 210–20. https://doi.org/10.1111/ecog.02272

Sesser, K.A., M. Iglecia, M.E. Reiter, K.M. Strum, C.M. Hickey, R. Kelsey, and D.A. Skalos. 2018. "Waterbird Response to Variable-Timing of Drawdown in Rice Fields after Winter-Flooding." *PLoS One* 13: e0204800.

Sesser, K.A., M.E. Reiter, D.A. Skalos, K.M. Strum, and C.M. Hickey. 2016. "Waterbird Response to Management Practices in Rice Fields Intended to Reduce Greenhouse Gas Emissions." *Biological Conservation* 197: 69–79.

Shirley, S.M., Z. Yang, R.A. Hutchinson, J.D. Alexander, K. McGarigal, and M.G. Betts. 2013. "Species Distribution Modelling for the People: Unclassified Landsat TM Imagery Predicts Bird Occurrence at Fine Resolutions." *Diversity and Distributions* 19: 855–66.

Shuford, D.W., J.M. Humphrey, R.B. Hansen, C.M. Hickey, G.W. Page, and L.E. Stenzel. 2004. Patterns of Distribution, Abundance, and Habitat Use of Breeding Black-Necked Stilts and American Avocets in California's Central Valley in 2003. Draft Final Report for the Central Valley Shorebird Working Group.

Shuford, D.W., G.W. Page, and J.E. Kjelmyr. 1998. "Patterns and Dynamics of Shorebird Use of the California's Central Valley." *Condor* 100: 277–44.

Shuford, D.W., M.E. Reiter, K.A. Sesser, C.M. Hickey, and G.H. Golet. 2019. "The Relative Importance of Agricultural and Wetland Habitats to Waterbirds in the Sacramento–San Joaquin River Delta of California." *San Francisco Estuary Watershed Science* 17: 2.

Skagen, S.K., and F.L. Knopf. 1993. "Toward Conservation of Mid-continental Shorebird Migrations." *Conservation Biology* 7: 533–41.

Skagen, S.K., and F.L. Knopf. 1994. "Migrating Shorebirds and Habitat Dynamics at a Prairie Wetland Complex." *The Wilson Bulletin* 106: 91–105.

Steen, V., S.K. Skagen, and B.R. Noon. 2018. "Preparing for an Uncertain Future: Migrating Shorebird Response to Past Climatic Fluctuations in the Prairie Potholes." *Ecosphere* 9: e02095.

Stenzel, L., and G.W. Page. 2018. "Trends in Abundance of Wintering Waterbirds Relative to Rainfall Patterns at a Central California Estuary, 1972–2015." *Western Field Ornithologists* 3: 236–57.

Stralberg, D., D. Jongsomjit, C.A. Howell, M.A. Snyder, J.D. Alexander, J.A. Wiens, and T.L. Root. 2009. "Re-Shuffling of Species with Climate Disruption: A No-Analog Future for California Birds?" *PLoS One* 4: e6825.

Strum, K.M., M.E. Reiter, C.A. Hartman, M.N. Iglecia, T.R. Kelsey, and C.M. Hickey. 2013. "Winter Management of California's Rice Fields to Maximize Waterbird Habitat and Minimize Water Use." *Agriculture, Ecosystems & Environment* 179: 116–24.

Sullivan, B.L., M.J. Wood, R.E. Iliff, R.E. Bonney, D. Fink, and S. Kelling. 2009. "eBird: A Citizen-Based Bird Observation Network in the Biological Sciences." *Biological Conservation* 142: 2282–92.

US Department of Agriculture. 2014. "National Agricultural Statistics Service Cropland Data Layer. 2008–2014." Published Crop-Specific Data Layer. Washington, DC: USDA-NASS. https://nassgeodata.gmu.edu/CropScape/ (accessed 2014).

VanDerWal, J., L. Falconi, S. Januchowski, L. Shoo, and C. Storlie. 2014. Species Distribution Modelling Tools: Tools for Processing Data Associated with Species Distribution Modelling Exercises. R Package Version 1.1-211. https://cran.r-project.org/web/packages/SDMTools/index.html

Veloz, S.D. 2009. "Spatially Autocorrelated Sampling Falsely Inflates Measures of Accuracy for Presence-Only Niche Models." *Journal of Biogeography* 36: 2290–9.

Veloz, S.D., J.W. Williams, J.L. Blois, F. He, B. Otto-Bliesner, and Z. Liu. 2012. "No-Analog Climates and Shifting Realized Niches during the Late Quaternary: Implications for 21st-Century Predictions by Species Distribution Models." *Global Change Biology* 18: 1698–713.

Welsh, A.H., D.B. Lindenmayer, and C.F. Donnelly. 2013. "Fitting and Interpreting Occupancy Models." *PLoS One* 8: e52015.

Welch, H., R.L. Pressey, and A.E. Reside. 2018. "Using Temporally Explicit Habitat Suitability Models to Assess Threats to Mobile Species and Evaluate the Effectiveness of Marine Protected Areas." *Journal for Nature Conservation* 41: 106–15.

Young, N.E., R.S. Anderson, S.M. Chignell, A.G. Vorster, R. Lawrence, and P.H. Evangelista. 2017. "A Survival Guide to Landsat Preprocessing." *Ecology* 98: 920–32.

Yu, H., A.R. Cooper, and D.M. Infante. 2020. "Improving Species Distribution Model Predictive Accuracy Using Species Abundance: Application with Boosted Regression Trees." *Ecological Modelling* 432: 109202.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.