

Science

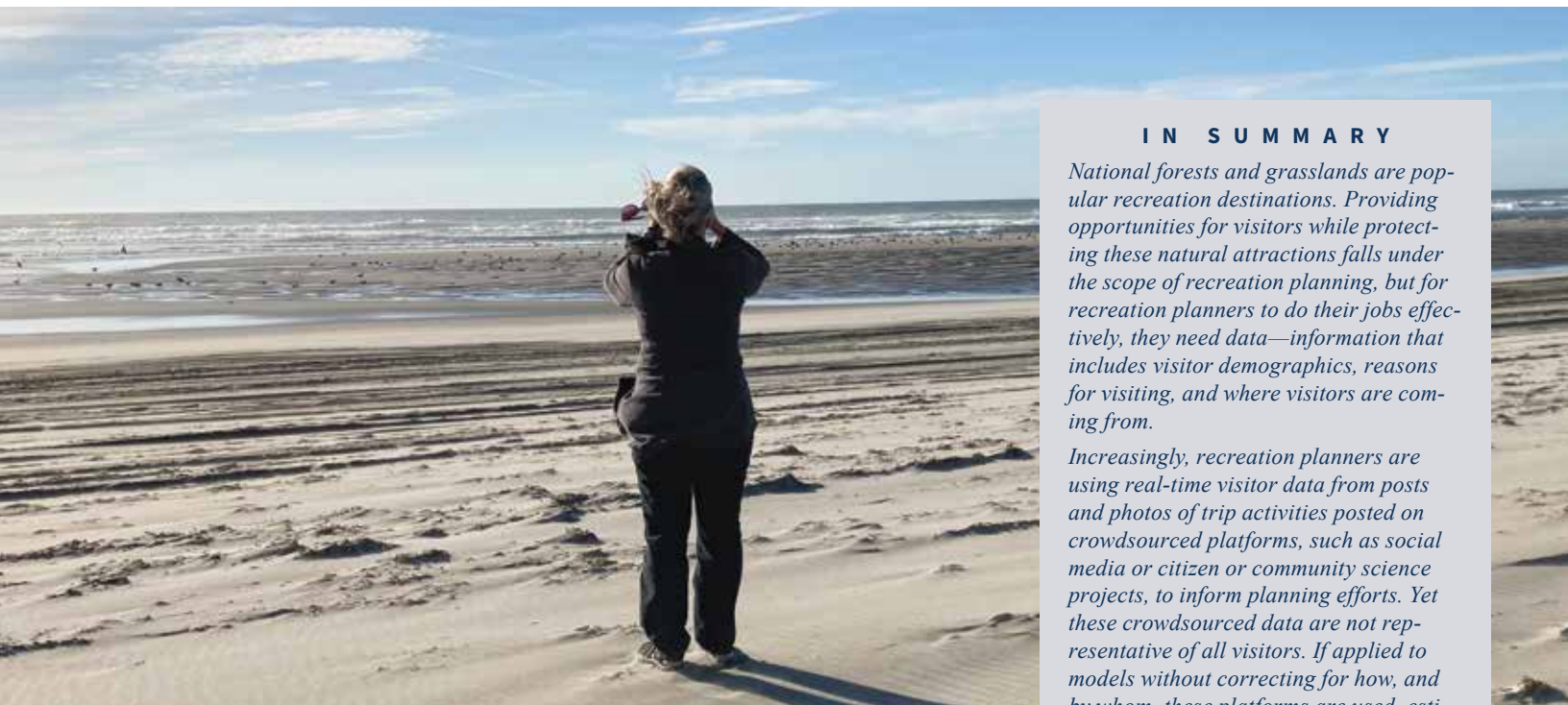
FINDINGS

INSIDE

Calculating the Value of Recreation	2
The Value of Birding	3
Viewing Rare Birds and Protecting Habitat	4
The Value of Salmon Viewing.	5

"Science affects the way we think together."
Lewis Thomas

Who's in the Crowd? Addressing Bias in Crowdsourced Data



A birdwatcher on the Washington coast. Through community science projects such as eBird, researchers can learn more about where birders visit and their willingness to pay to view birds. Photo by Trudy Cameron.

"After a day's walk, everything has twice its usual value."

—George M. Trevelyan, English historian

In Oregon during August 2024, a member of the eBird online community observed a common tern (*Sterna hirundo*) at the Bully Creek Campground. Another eBirder noted a great egret (*Ardea alba*) in Malheur County. And at the North Jetty Beach, at the mouth of the Siuslaw River, another eBirder reported seeing a horned puffin (*Fratercula corniculata*).

Managed by the Cornell Lab of Ornithology, the eBird website describes the project as "among the world's largest biodiversity-

related science projects, with more than 100 million bird sightings contributed annually by eBirders around the world." Members opt to report to eBird where they birded and what they observed and heard. It's an example of a crowdsourced platform that leverages the observational power of a crowd to gather data about bird populations.

As a doctoral student at University of Oregon, Sonja Kolstoe saw the value of these bird-sighting data, but for reasons unrelated to ornithology. Viewing the data through a lens of economics and recreation, eBird provided an opportunity to know where people chose to go birdwatching, an outdoor activity that was previously hard to study since permits or licenses are

IN SUMMARY

National forests and grasslands are popular recreation destinations. Providing opportunities for visitors while protecting these natural attractions falls under the scope of recreation planning, but for recreation planners to do their jobs effectively, they need data—information that includes visitor demographics, reasons for visiting, and where visitors are coming from.

Increasingly, recreation planners are using real-time visitor data from posts and photos of trip activities posted on crowdsourced platforms, such as social media or citizen or community science projects, to inform planning efforts. Yet these crowdsourced data are not representative of all visitors. If applied to models without correcting for how, and by whom, these platforms are used, estimates will skew toward more frequent visitors that provide the data, instead of the general population of visitors that planners aim to survey.

Sonja Kolstoe, a research economist with the U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, has spent nearly 10 years exploring and refining techniques to use crowdsourced data and developing data correction methods to address selection bias. Her work was the basis for The Nature Conservancy's study of visitors' willingness to pay to visit the Waikamoi Preserve on the island of Maui, Hawai'i. The Washington Department of Fish and Wildlife is also developing a pilot study to analyze the economic value of salmon viewing and the project's survey methods will be based on Kolstoe's research.



A birdwatcher near Meadowbrook Pond in Seattle, Washington. Nearly 15 percent of the U.S. population age 16 and older participates in bird watching. Photo by Sonja Kolstoe, USDA Forest Service.

typically not required. When paired with information about travel time and cost, the eBird data offered an opportunity to study people's tradeoffs between travel cost and birdwatching opportunities to estimate the value of birdwatching to eBirders.

Now a research economist with the U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Kolstoe's research includes the economic value of ecosystem services as well as their management. "A lot of my work is predominately focused on valuation of ecosystem services, particularly recreation," she explains. "I focus on the valuation of nonmarket ecosystem services, which are benefits from nature that we do not pay for but that we do derive benefit from and may thus be willing to bear costs to protect."

Calculating the Value of Recreation

Harold Hotelling, an economist with the University of North Carolina, is credited with developing the first mathematical method for quantifying recreation's nonmarket value. In 1947, Hotelling, along with nine other economists, received a letter from the National Park Service Associate Director A. E. Demaray asking for their opinion on how to "evaluate the benefits of the national parks

to the national economy, to the states, and to local communities."

In response to Demaray's inquiry, Hotelling replied that he was convinced it was possible to evaluate "with a reasonable degree of accuracy, the service of national parks to the public." He proposed a mathematical formula that estimated the cost of traveling to the park based on the distance traveled and the time spent doing so by a sample of visitors from the general population. Travel cost provides evidence of the price of each trip to the park, and the number of trips taken can be used to estimate the park's benefits to the public. "This approach through travel costs is one of several possible modes of attack on this problem," Hotelling wrote. "There are also others, which should be examined, though I think the method outlined above looks most promising."

Hotelling's proposed formula became the foundation for the travel cost method. Subsequent model advances in the recreation demand literature, including discrete choice modeling, led to the recreation site-choice model that Kolstoe uses.

"Recreation site-choice models allow us to understand what people would have been willing to pay to enjoy a recreation experience," she explains. "We can estimate the benefit based on the tradeoff of travel costs, which include money and time, that are willingly incurred to enjoy a site's features. Site attributes may vary over space and, in some cases, over time, to include season, like with the migration of bird species."

Recreation demand models, such as recreation site-choice, require detailed visitor data, such as visitor counts, as well as information on distances traveled to allow the calculation of travel costs. For some recreation activities, Kolstoe says the "breadcrumbs" people leave behind, such as paying for hunting licenses, permits for accessing wilderness areas, or

PNW Science Findings

Providing scientific information to people who make and influence decisions about managing land.

PNW Science Findings is published monthly by:

Pacific Northwest Research Station
USDA Forest Service
1220 SW 3rd Ave., Suite 1400
Portland, OR 97204

Send new subscription and change of address information to:

SM.FS.pnw_pnwpubs@usda.gov

Rhonda Mazza, editor; rhonda.mazza@usda.gov

Jason Blake, layout; jason.p.blake@usda.gov

To find Science Findings online, visit
<https://research.fs.usda.gov/pnw>
and click on **Products**.

To become a digital subscriber visit:
<https://research.fs.usda.gov/pnw/products/publications/subscribe>



Forest
Service

*Tundra swans (Cygnus columbianus) near Sauvie Island, Oregon.
Photo courtesy of Kathy Munsel, Oregon Department of Fish and Wildlife.*



KEY FINDINGS

- Crowdsourced data offer a rich source of information on outdoor recreation behavior and an opportunity to engage platform users about topics of interest to land managers. However, crowdsourced data only represents a portion of all recreationists, and participation also differs across platforms. This difference in use may contribute to biases in modeling estimates.
- Methods that account for sample selection issues in crowdsourced samples will allow researchers to estimate values from crowdsourced data, whether from reported trips or surveys, that may be representative of all recreationists, not just those who use crowdsourced platforms.
- Researchers used a correction method to analyze survey data collected to understand respondents' willingness to pay for a guided hike through the Waikamoi Preserve on Maui. Without the correction method, the survey results suggested differences for some hike features. Once the bias was corrected, the willingness-to-pay distributions did not show a statistically significant difference across the two samples.

reserving campsites, can be used to generate visitor counts. If home address or at least ZIP code information is also available, it is also possible to calculate travel cost and estimate a recreation demand model. The National Visitor Use Monitoring survey is another source of visitor data that the Forest Service uses.

With the advent of social media and citizen or community science projects, such as eBird, scientists have been exploring the feasibility of gleaning data from visitors' online posts. It's these unofficial breadcrumbs that, as a Ph.D. student, Kolstoe explored to determine whether they could be used in recreation site-choice models. Her initial published studies came with the caveat that the values were valid for eBird members, and not the general population of birders. This motivated her to pursue the development of a method that would account for sample selection when using crowdsourced platforms.

"Not everyone interacts with a social media platform or [community science] project in the same way," Kolstoe explains. "For example, some people will choose to report all their bird sightings to eBird, whereas others report only a small fraction."

The Value of Birding

In December 2012, Kolstoe and her then Ph.D. advisor, Trudy Ann Cameron (now a professor emeritus at the University of Oregon) contacted Steve Kelling, the director of information sciences at the Cornell Lab of Ornithology and lead manager of the eBird project at the time. They explained the purpose of Kolstoe's research and requested access to the eBird data and home address information that members provide when registering for eBird. To build the recreation site-choice model, she organized the data to estimate the distance and time required to travel to their birdwatching destination. The unique

features of a birdwatching site might influence a birder's choice to visit, so Kolstoe gathered data about key attributes of birding sites. These features included the number of expected bird species that may be observed at the site, how the site is protected for biodiversity, its ecoregion, whether it's within an urban or rural area, and a measure for popularity and potential congestion at each site as well as controls for the time of year of the trip.

After running the model, she determined that eBirders' willingness to pay for a birding outing depended systematically upon site features as well as time of year. Based on the results of the model, she determined that eBirders were willing to pay marginally more in June, the breeding season in the region, than other months of the year.

Kolstoe demonstrated the feasibility of using eBird data in recreation site-choice models but with two crucial caveats: the eBird data represented only eBirders who submitted bird observations, and eBirders may not be representative of the greater U.S. birder community.

"This non-representativeness is a concern if you want to make policy decisions," she explains. "To leverage the valuable information coming from [community science] or social media data, we need to understand how representative these estimates may be relative to similar data from representative populations. If we want to use benefit-cost analyses to evaluate policies and programs, that's going to be important."

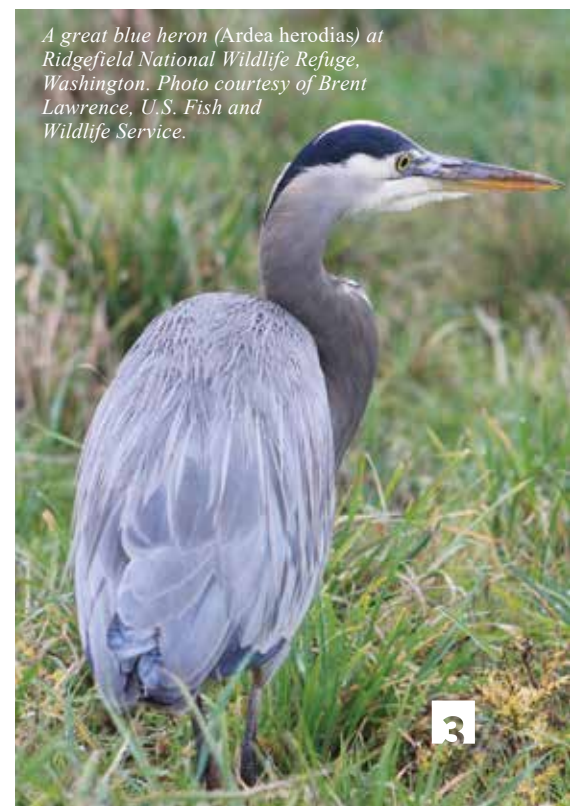
Kolstoe worked with Cameron to conduct subsequent research to develop a method to address these two caveats using survey data to facilitate testing of the method. They conducted a survey of eBird members in the Pacific Northwest and a separate survey with an online panel to obtain a

representative sample of the general population of the United States. Respondents of both surveys were asked questions about how far they would consider traveling if they were planning a daytrip to go birdwatching, whether they participated in community science projects such as eBird or the annual Audubon Christmas Bird Count, along with basic sociodemographic information. The researchers then modeled how far people would go for a birdwatching day trip and tested their prototype sample-section correction methods with the model and these two samples.

Their correction methods minimized the selection bias within the data and made their results more representative. They are currently working on incorporating these corrections into a recreation site-choice model with eBird data to demonstrate how it could be employed when using community science or volunteered geographic information data.



Birders in the Willapa National Wildlife Refuge, Washington. Photo by Sonja Kolstoe, USDA Forest Service.



A great blue heron (Ardea herodias) at Ridgefield National Wildlife Refuge, Washington. Photo courtesy of Brent Lawrence, U.S. Fish and Wildlife Service.

Viewing Rare Birds and Protecting Habitat

The Nature Conservancy manages the 9,000-acre Waikamoi Preserve on Maui Island, Hawai'i. Visiting the preserve "is incredibly sought-after because there are two critically endangered endemic birds, and this is the only easily accessible place [birders] can find them in all of the world," says Alison Cohan, who was the Maui Nui terrestrial program director for The Nature Conservancy at the time of the study (currently the Western U.S. and Canada Division Conservation Manager).

Cohan came across Kolstoe's name while doing some research and remembers thinking,

"Wow, there's literally a bird resource economist, and I have to talk to this person."

At the time, The Nature Conservancy offered a free monthly hike to the public, and registration would fill up within 2 hours of opening. The conservation group contemplated offering a fee-based guided hike to manage demand. However, there were concerns about increasing the number of birders in the preserve because they could inadvertently carry rapid 'Ōhi'a death, a fungal disease from other Hawaiian Islands, on their shoes or automobile tires. Once 'Ōhi'a lehua trees (*Metrosideros polymorpha*) are infected, they quickly die. Because the 'Ōhi'a tree is integral to native Hawaiian culture and

forests, The Nature Conservancy needed to prevent the spread of this disease. An equal concern was that imposing a fee would make the preserve inaccessible to some visitors.

To inform their decision, Cohan wanted to better understand the primary motivations for visitors to Waikamoi, and whether they would be willing to pay for the mitigation efforts that The Nature Conservancy would have to adopt to allow more visitors into the preserve.

Kolstoe, Cohan, and Brian Vander Naald, an associate professor of economics with Drake University, developed a stated preference survey. "If a survey option doesn't currently exist, such as paying to hike the preserve, we're able to explore what might happen if it were an option and understand people's preferences for possible programs that could be implemented," Kolstoe explains.

The Nature Conservancy posted the survey link on its social media sites, and the researchers also surveyed a representative sample of the United States and Hawai'i. To correct for the bias resulting in the social media dataset, Kolstoe applied the method she and Cameron developed earlier. After they applied the corrections, they found no significant difference between survey responses from the social media sample and the general representative sample in terms of visitors' willingness to pay for features of the guided hike. Without correcting for the bias,



*The 9,000-acre Maui Nui Preserve on the island of Maui is a sought-after destination for birders because it contains two critically endangered birds, the Maui parrotbill (kiwikiu, *Pseudonestor xanthophrys*) and the crested honeycreeper (‘ākohekohe, *Palmeria dolei*), and is the only easily accessible location to observe them. Photo by Sonja Kolstoe, USDA Forest Service.*



The Maui parrotbill is currently only found within a 12 square mile area on the island of Maui. Photo courtesy of Andrew Smith, Flickr.

however, statistically significant differences existed between the two samples in visitors' willingness to pay, which could have affected management decisions.

Behind the statistical analyses that produced usable results for The Nature Conservancy was a positive finding that stood out to Cohan. "People were willing to pay," she said. "Not just willing to pay to access the preserve, but people were willing to pay to mitigate their risk. In particular, people did care about the threats that visitors bring with them when entering the preserve and were willing to pay to offset those risks."

The Value of Salmon Viewing

Braeden Van Deynze, a natural resource economist with the Washington Department of Fish and Wildlife, regularly applies nonmarket valuation methods, including recreation demand models. These projects often use trip data, also known as revealed preference data.

Van Deynze says trip data can be expensive and difficult to obtain, which is why crowdsourced data can be very desirable to use in recreation demand models. "Sonja's really been at the forefront of exploring methods to use data people freely provide through things like eBird, or other sorts of community science or social media platforms," he explains. "She's given us, as practitioners, more options to use these sorts of data despite their drawbacks. As a result, we could use these data to make our models more powerful and make our estimates more reliable and accurate within budget."

Kolstoe is providing technical guidance for a study Van Deynze is leading on behalf of the Washington Department of Fish and Wildlife

LAND MANAGEMENT IMPLICATIONS

- Understanding the limitations and potential biases of crowdsourced data is important. The people who share information over social media and other platforms may not represent the general population, and models that fail to correct for this problem can lead to incorrect conclusions.
- New guidance, based on tested methods, offers a way to overcome the bias inherent in crowdsourced platforms. This provides a path forward so that crowdsourced data can be incorporated into recreation planning and management.

on salmon viewers in Washington state. They want to understand people's willingness to pay to view salmon. In other words, "how much do they value the opportunity to see the salmon in the stream or at the hatchery?" Van Deynze says.

This fall, the Washington Department of Fish and Wildlife is conducting a pilot study of people who visit the Issaquah Salmon Hatchery. Going this route, using data collected opportunistically onsite, will be less costly than conducting a full study using the general population. "This was an opportunity to gather data that isn't representative of the general population, but using Sonja's methods, we can get a better idea of how we can integrate those data with other data that we might one day gather through a broader, more representative survey," he explains.

For Kolstoe, these crowdsourced data are the choices that people make about where to spend their free time. "That's really what recreation site-choice models are getting at," she explains. "Time is a scarce resource, and money is a scarce resource. When people are choosing to spend their free time to recreate, they're going to choose a site that gives them the highest expected level of satisfaction for that day's outing, that's why they choose it among all the possible options."

"By a small sample we may judge of the whole piece."

—Miguel de Cervantes, Spanish novelist

For Further Reading

Cameron, T.A.; Kolstoe, S.H. 2022. Using auxiliary population samples for sample-selection correction in models based on crowd-sourced volunteered geographic information. *Land Economics*. 98(1): 1–21. <https://doi.org/10.3368/le.98.1.040720-0050R1>.

Kolstoe, S.; Cameron, T.A. 2017. The non-market value of birding sites and the marginal value of additional species: biodiversity in a random utility model of site choice by eBird members. *Ecological Economics*. 137: 1–12. <https://doi.org/10.1016/j.ecolecon.2017.02.013>.

Kolstoe, S.; Cameron, T.A.; Wilsey, C. 2018. Climate, land cover, and bird populations: differential impacts on the future welfare of birders across the Pacific Northwest. *Agricultural and Resource Economics Review*. 47(2): 272–310. <https://doi.org/10.1017/age.2018.9>.

Kolstoe, S.; Vander Naald, B.; Cohan, A. 2022. A tale of two samples: understanding WTP differences in the age of social media. *Ecosystem Services*. 55: 101420. <https://research.fs.usda.gov/treearch/64229>.

Writer's Profile

Andrea Watts is a freelance science writer based in the Pacific Northwest. Her portfolio is available at <https://wattsinthewoods.com>, and she can be reached at andwatts@live.com.

*At the Issaquah Salmon Hatchery in Washington state, visitors can view Chinook (*Oncorhynchus tshawytscha*) and coho salmon (*O. kisutch*) that return to spawn in Issaquah Creek. In 2023, the hatchery hosted more than 200,000 visitors. Photo courtesy of Washington Department of Fish and Wildlife.*





Pacific Northwest Research Station
USDA Forest Service
1220 SW 3rd Avenue, Suite 1400
Portland, OR 97204

Official Business
Penalty for Private Use, \$300

Scientist Profile



SONJA KOLSTOE is a research economist whose research focuses on the valuation and management of ecosystem services, to include how they are affected by changes from disturbances and changes in climate.

Kolstoe can be reached at:

USDA Forest Service
Pacific Northwest Research Station
1220 SW 3rd Ave, Suite 1410
Portland, OR 97204

E-mail: sonja.kolstoe@usda.gov

Collaborators

Trudy Cameron, University of Oregon

Brian Vander Naald, Drake University

Alison Cohan, The Nature Conservancy