Microsoft

AI for Earth Grantee Profile

NatureServe

Building a map of biodiversity importance

Summary

NatureServe is developing an unparalleled tool for identifying the places most critical for conserving at-risk species in the contiguous United States. With support from Esri, The Nature Conservancy, and Microsoft, NatureServe and its network of state natural heritage programs are applying machine learning techniques to their comprehensive biodiversity inventory data to model habitat for more than 2,600 at-risk, taxonomically and ecologically diverse species. These spatial models will be synthesized into a map that identifies high-priority biodiversity conservation areas—a dynamic, transparent, and repeatable base layer to help guide effective conservation decision-making.

Building a unique tool to map high-priority conservation areas

NatureServe's unique modeling and mapping effort will identify the most important places for the protection of threatened species, a first step in determining where to focus conservation efforts. The project will help answer important questions like: Where should we invest tax dollars or private funding to conserve land or create new nature preserves? Where should we direct mitigation and restoration activities to benefit at-risk species? And where are human activities *unlikely* to have negative impacts?

All these questions require one basic but considerable information challenge: identifying where at-risk species are found—and where are they not. In far too many cases, the answer to that deceptively simple question is inadequate in ways that make it difficult to effectively manage and conserve species. The sheer number of species to be identified and tracked presents its own challenge. Geopolitical boundaries and the disparate data collection methods used by scientists from universities, museums, and state and federal wildlife agencies add to the difficulty of deriving meaning from individual datasets. NatureServe is working to make it easier for everyone—conservation practitioners, land managers, and industries—to get better answers about where at-risk species and their habitats are found, and how best to protect them.

The task of identifying, cataloging, and monitoring where the planet's incredible diversity of species and habitats occur, and then measuring their risk of extinction, requires enormous effort. NatureServe, the umbrella organization for a network of biodiversity information centers spanning the Western Hemisphere, has been working to address this challenge since 1974, when the first state natural heritage programs were established

in South Carolina and Mississippi. For almost 50 years, these programs have been collecting standardized data on species and ecosystems using consistent methods to meaningfully answer fundamental questions about the location, status, and trends of biodiversity. The NatureServe network now encompasses all 50 US states, the provinces and territories of Canada, and much of Latin America. Individually, these programs are experts on biodiversity in their respective geographies; working together as the NatureServe network, they provide significant and seamless information about biodiversity at national and global scales.

> "Over decades, we have built an internationally standardized dataset on at-risk species that is both taxonomically rigorous and geographically precise, which we are now leveraging in new ways to inform conservation actions."—Dr. Healy Hamilton, Chief Scientist, NatureServe

Core elements of NatureServe's work are to assess the conservation status, or extinction risk, of species and to define and map the hemisphere's wide variety of terrestrial ecosystems, organizing them into hierarchies from local to broad categories of vegetation types. Expert scientists in each individual natural heritage program collect detailed, high-value data on the ecosystems and at-risk species that occur in their jurisdictions through field inventory and monitoring. By using a standardized methodology for data collection and management, NatureServe can aggregate these complex data to create analyses for conservation and management of our living resources across local, national, and international borders. Through this work over many decades, NatureServe has built, in the words of Chief Scientist Healy Hamilton, "an internationally standardized dataset on at-risk species that is both taxonomically rigorous and geographically precise, which we are now leveraging in new ways to inform conservation actions." Without these standard data, decision-makers would be without the best available data to identify and effectively protect threatened species and their habitats.

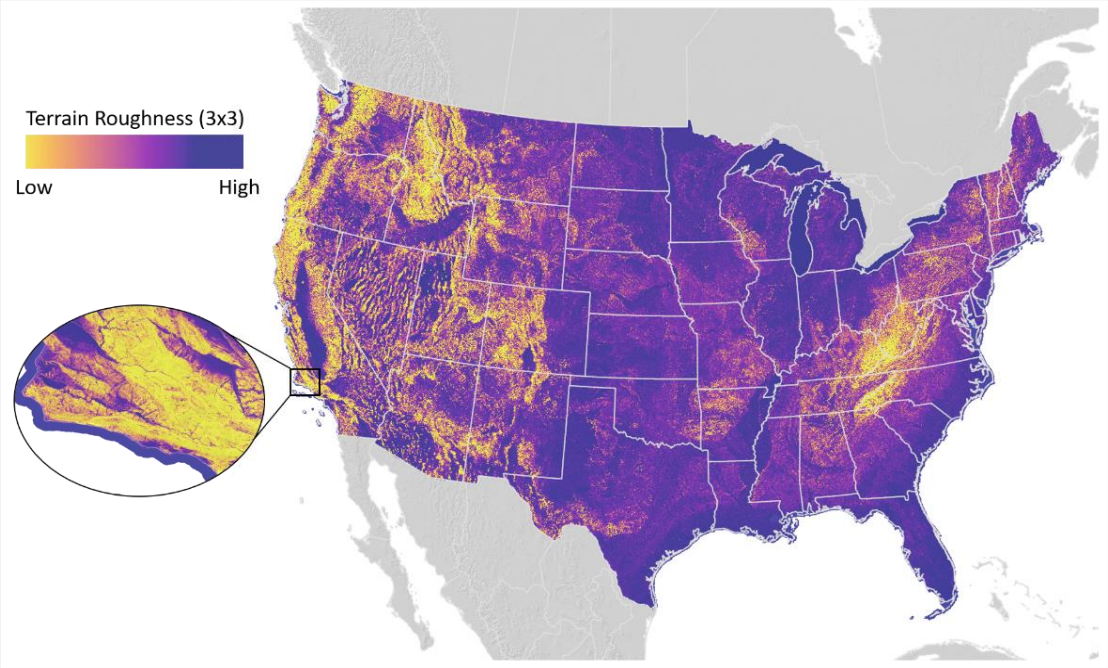
Modeling the distribution of at-risk species

In order to conserve at-risk species, we must of course first know where they exist. The first step toward effective answers about where species are present is to gather observations in the field, but those efforts are limited by a number of factors, including the time and availability of trained observers and researchers, resources to conduct field research, as well as the breadth of lands and waters to be surveyed. Although this information can be augmented by sightings reported from the general public, called citizen science, those

reports often lack the supporting data that would make them scientifically valid for management decisions. Using only field observations collected by expert and citizen scientists, species locations can be mapped in two ways.

The first method uses confirmed observations, which provides certainty but often underrepresents a species' actual range, because a landscape can rarely be comprehensively sampled to identify every population. This method produces a range map that will not accurately indicate how widespread the species is; thus, areas that should be considered for management or protection may not be identified. The second method uses the confirmed observations (and additional sightings) to extrapolate a generalized range map, but this is subject to arbitrary decisions about where to draw boundaries—which may be influenced by artificial borders that play no part in the actual distribution of a species, such as county boundaries. This latter method leads to over-predicting where a species is found, creating potential conflicts with land use where they may not exist, and diverting resources away from the areas where they may have the greatest positive impact.

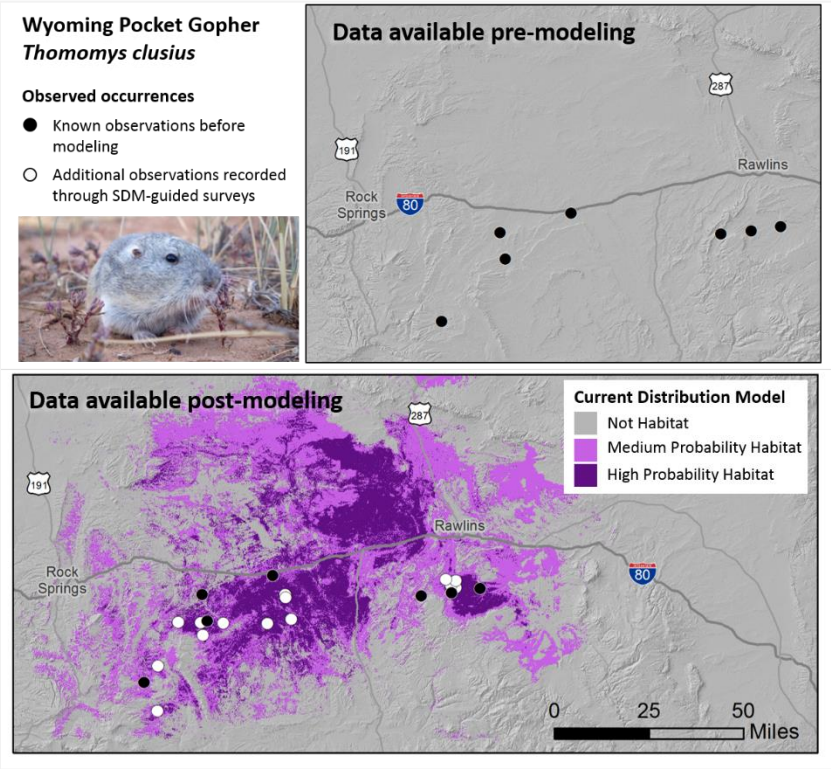
Clearly, the most effective conservation strategies require something more than these two methods. Fortunately, statistical science combined with modern computing technology offers a third option: habitat suitability modeling. Also known as species distribution modeling, this method takes into account the various environmental conditions where a species is known to occur—environmental predictors such as topography, soil type, land cover, proximity to water, and climate—and identifies other areas where that unique



Derived data for over 200 environmental predictors, such as terrain roughness, help determine where threatened species might be found. [Map Credit: NatureServe; Data Source: USGS National Elevation Dataset]

combination of conditions occurs. By mapping areas with these conditions, researchers can better determine areas of habitat where species are likely to be found, even if they have not yet been observed there.

The other benefit of habitat suitability modeling is that researchers can identify areas of unsuitable habitat, where a species most likely *will not* be found, sparing the time, money, and effort to enforce management in areas that are highly unlikely to sustain a species. This modeling makes it easier to look past artificial borders and jurisdictions to focus on the locations that matter most to preventing (and even reversing) a species' decline.



An example of how modeling has been used to improve species distribution information and streamline management.

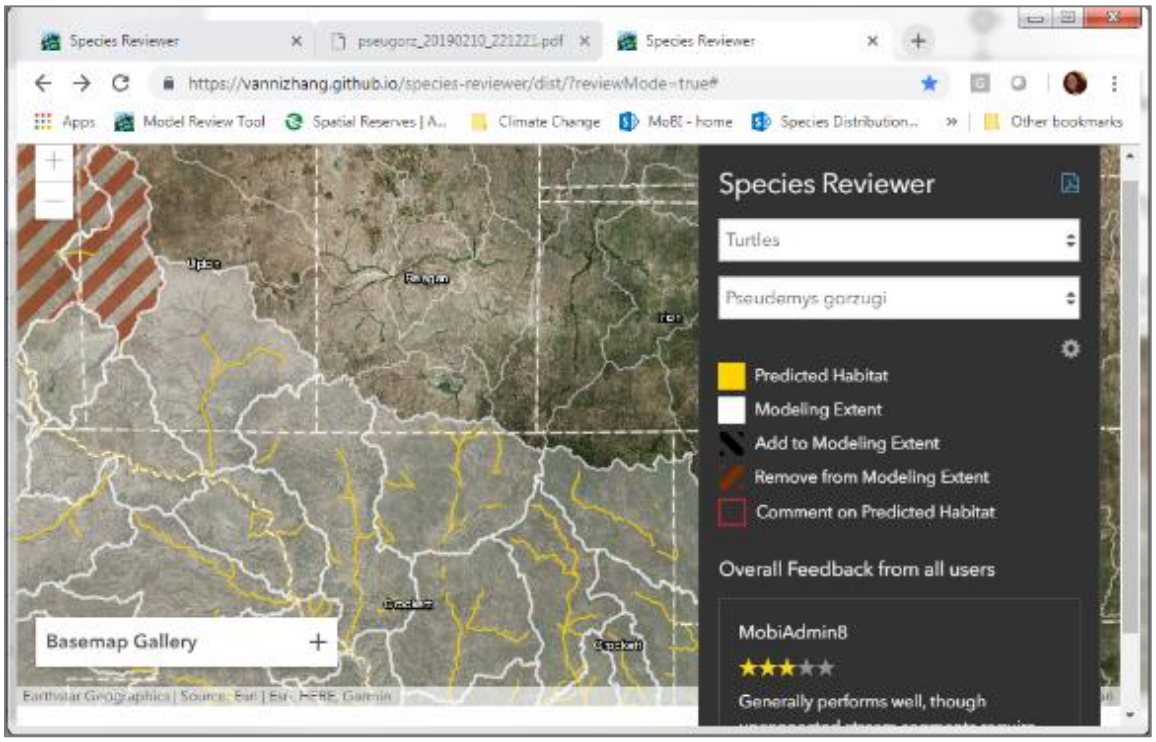
The Wyoming pocket gopher (Thomomys clusius) is a globally imperiled species found largely on BLM land in Wyoming's Red Desert. Prior to 2006, little was known about the species, with few documented occurrences (top panel). The Wyoming Natural Diversity Database, a NatureServe partner, first modeled the gopher that year, and used the model to direct field efforts. As a result, new populations were located, and the model was refined (bottom panel). Land managers now have better information on both where the species is known to occur and other areas of likely habitat. This expanded understanding of occurrence and available habitat contributed to a 2010 USFWS decision not to list the species.

Creating a guide to conservation investment

Habitat suitability modeling is not a new concept. Even in the 19th century, scientists used environmental predictors to account for the distribution of different species. However, practical application of this modeling to solve the complex issues facing biodiversity today requires diverse sets of high spatial resolution data, even for just one species. And only over the past two decades has habitat suitability modeling matured into an effective and accepted tool in the scientific community. As its natural heritage programs began making use of modeling to predict suitable habitat for at-risk species, NatureServe saw an opportunity to transform conservation decision-making by building a dynamic national library of rigorously vetted habitat suitability models for

species of conservation need. Such a library would help streamline decisions and improve conservation outcomes while saving time and money.

A major step toward creating this national library is to produce a national-level map of the places most critical to conserving at-risk species—a map of biodiversity importance—to help all stakeholders better understand and undertake species management and conservation efforts. At the urging of Jack Dangermond, the co-founder and CEO of Esri, the global leader in geographic information system (GIS) software, NatureServe took this project on. Because of their interest in conservation, Jack and his wife Laura, who co-founded Esri, had donated $165 million to The Nature Conservancy to purchase and create a 24,000-acre nature preserve at Point Conception, California that will conserve its native biodiversity, which includes dozens of species-at-risk. Jack reached out to NatureServe to generate this national map of biodiversity importance to guide additional conservation investments, the same way he and his wife were guided to the Point Conception property.



NatureServe's Model Review Tool offer an easy and intuitive interface that allows scientists and practitioners to collaboratively validate and improve the accuracy of habitat suitability models for any species. [Credit: NatureServe]

To begin creating the map, NatureServe compiled two terabytes of data at 30-meter resolution on soils, topography, climate, land cover, and hydrologic systems for the contiguous United States from various sources. Working closely with Esri's spatial data scientists, they built a library of over 200 environmental predictor layers for both aquatic and terrestrial species. NatureServe used its conservation status ranking system to select a

wide variety of threatened species for distribution modeling on this national map—over 2,600 species, ranging beyond the "cute and fuzzy" mammals to include endangered plants, pollinators like bumblebees and butterflies, and even freshwater invertebrates like mussels and crayfish, all of which play important ecological roles in maintaining healthy functions in the communities where they occur. Esri built a web mapping application that facilitates collaboration among model reviewers— NatureServe's network of expert field scientists—allowing them to view the results of the habitat suitability models and provide feedback to improve their accuracy and quality.

Going forward

From the start of the project, NatureServe knew that creating this map of biodiversity importance to identify priority conservation areas would require the power of Microsoft Azure cloud services. "The networked computational infrastructure and the data it supports are what make this project viable. Azure is the technical foundation of our work, which would not have been possible without the Microsoft AI for Earth grant," says Dr. Healy Hamilton. "Only once we were awarded the AI for Earth grant did we have the support necessary to build the technological infrastructure required to implement this national map. Esri and The Nature Conservancy both provided financial support; from Esri, we also had significant in-kind technical support; and from AI for Earth, we've had the incredible computing power of the Azure infrastructure."

> "The networked computational infrastructure and the data it supports are what make this project viable… which would not have been possible without the Microsoft AI for Earth grant."—Dr. Healy Hamilton

"What this project has allowed us to do," explains Regan Smyth, Director of Spatial Analysis at NatureServe, "is to bring habitat suitability models to scale. The enormous computing power of Azure actually makes it feasible to solve these problems now, at an absolutely critical point for conserving biodiversity in a rapidly changing world." The project is already delivering useful results, facilitating online collaboration among natural heritage programs and reducing the time to produce and review model outputs from weeks to minutes. Production of the map of biodiversity importance will be a major accomplishment in its own right, but the potential of the project truly manifests in the strides made toward achieving NatureServe's larger goal—to protect threatened and endangered species from extinction—which requires constantly improving our understanding of where at-risk species occur. By putting this foundation on a cloud-based infrastructure, NatureServe can now build its

dynamic, rigorous modeling service, which will modernize the way we understand the distribution of habitat for species-at-risk and allow for better-informed decisions about their management and conservation.

> "The enormous computing power of Azure makes it feasible to solve these problems now, at an absolutely critical point for conserving biodiversity in a rapidly changing world."—Regan Smyth, Director, Spatial Analysis, NatureServe

About NatureServe

NatureServe empowers people to conserve biodiversity by making sure everyone has access to the knowledge needed to be better stewards of our shared lands and waters. Through its network of natural heritage programs, NatureServe collects decision-quality data about imperiled species and ecosystems, transforms that data into knowledge products and decision support tools, and delivers expert analyses that guide decision-making, implement actions, and enhance conservation outcomes. The public-private partnership that is the NatureServe network comprises more than 90 network programs throughout the Western Hemisphere, with over 1,000 conservation professionals and a collective annual budget of more than $100 million.

Resources

Websites

NatureServe home site

Publications

Helen R Sofaer, Catherine S Jarnevich, Ian S Pearse, Regan L Smyth, Stephanie Auer, Gericke L Cook, Thomas C Edwards, Gerald F Guala, Timothy G Howard, Jeffrey T Morisette, Healy Hamilton. "Development and Delivery of Species Distribution Models to Inform Decision-Making." *BioScience*, biz045, June 5, 2019. https://doi.org/10.1093/biosci/biz045