GIS for SCIENCE

APPLYING MAPPING & SPATIAL ANALYTICS

Book Chapter Excerpt: Identifying the Natural Efficient Frontier

DAWN WRIGHT & CHRISTIAN HARDER, EDITORS

GIS for SCIENCE

APPLYING MAPPING AND SPATIAL ANALYTICS

DAWN J. WRIGHT AND CHRISTIAN HARDER, EDITORS

Esri Press | Redlands, California

Esri Press, 380 New York Street, Redlands, California 92373-8100

Printed in the United States of America
23 22 21 20 19 1 2 3 4 5 6 7 8 9 10

Christian Harder and Dawn J. Wright, eds.; *GIS for Science: Applying Mapping and Spatial Analytics*; DOI: https://doi.org/10.17128/9781589485303

Library of Congress Control Number: 2019936340

Ask for Esri Press titles at your local bookstore or order by calling 800-447-9778, or shop online at esri.com/esripress. Outside the United States, contact your local Esri distributor or shop online at eurospanbookstore.com/esri.

Esri Press titles are distributed to the trade by the following:

In North America:
Ingram Publisher Services
Toll-free telephone: 800-648-3104
Toll-free fax: 800-838-1149
E-mail: customerservice@ingrampublisherservices.com

In the United Kingdom, Europe, Middle East and Africa, Asia, and Australia:
Eurospan Group
3 Henrietta Street
London WC2E 8LU
United Kingdom
Telephone: 44(0) 1767 604972
Fax: 44(0) 1767 601640
E-mail: eurospan@turpin-distribution.com

All images courtesy of Esri except as noted. Developmental editing by Mark Henry; Technology Showcase editing by Keith Mann; Cover layout and book design by Steve Pablo; Cover concept by John Nelson.

The cover image shows a lidar-derived colored hillshade of glacial ice and snowfields on Mount Rainier in Washington State, United States, compared to the glacial extent of the same area from 1924. The lidar data is from a survey completed in 2008 by the National Park Service. Image by Daniel Coe, Washington Geological Survey. The lidar portion of the image was modified from Robinson, Joel E., Thomas W. Sisson, and Darin D. Swinney. 2010. Digital Topographic Map Showing the Extents of Glacial Ice and Perennial Snowfields at Mount Rainier, Washington, based on the LIDAR survey of September 2007 to October 2008: US Geological Survey Data Series 549. The 1924 extent of glacial ice was derived from: US Geological Survey, 1924, Mount Rainier Quadrangle, Washington: US Geological Survey, 1 sheet, map scale 1:125,000.

See this book come alive at

GISforScience.com

PRAISE FOR *GIS FOR SCIENCE*

"This book is beautiful as well as illuminating, and it dramatizes the ways in which the new science of geospatial information is enriching and empowering all other scientific disciplines."

—James Fallows, staff writer, *The Atlantic*; former chief speechwriter for President Jimmy Carter

"If you love maps like I do, you'll be drawn to this book. But you'll quickly discover so much more: the power of harnessing multiple perspectives and data types that infuse maps with even more meaning and catalyze new insights. A veritable treasure trove of ideas. "

—Jane Lubchenco, environmental scientist, marine ecologist, former Administrator of the National Oceanic and Atmospheric Administration (2009-2013); former U.S. Science Envoy for the Ocean (2014-2016); university distinguished professor

"GIS has become *the* foundational tool for all things environmental—from conservation to climate change to environmental justice. This astonishing book beautifully displays GIS in all its scientific, artistic, and creative splendor."

—Peter Kareiva, director, UCLA Institute of the Environment and Sustainability

"Dawn Wright and Christian Harder have given us a geoscience book for the twenty-first century! Cutting-edge research examples and gloriously illustrated state-of-the-art, GIS-enabled techniques come together to show us how to understand our planet in ways not possible even a few years ago."

—Margaret Leinen, Director of Scripps Institution of Oceanography and UC San Diego Vice Chancellor for Marine Sciences

"The only thing changing faster than Earth's environment and our species' imprint on it, for better and worse, is the information environment. In that noisy realm, trolls and other troublemakers get the headlines. But this essential and beautiful book illuminates how a host of innovators are gleaning meaning from data and helping shape a sustainable human journey."

—Andrew Revkin, strategic adviser, National Geographic Society, and coauthor of *Weather: An Illustrated History, from Cloud Atlases to Climate Change*

"The Science of Where® comes alive in *GIS for Science*. The book is, yes, informative, helping us understand how the world works, how it looks, and how we see it through images, maps, and more. Above all, it is dazzling, combining knowledge with a sense of wonder, bringing a desire to press for more discovery, and invoking a deep appreciation for why smart decisions spring from taking science to action."

—Lynn Scarlett, vice president, The Nature Conservancy; chair, Science Advisory Board, NOAA

"A textbook and a work of art."

—Len Kne, U-Spatial associate director,
University of Minnesota Research Computing, Office of the Vice President for Research

"Illustrating the power of geospatial analytics to address pressing challenges facing our planet, *GIS for Science* is a *tour de force*. The editors and contributors have produced a visual delight that will inspire and enlighten researchers, citizen scientists, and the public about the contribution of the geographic perspective to the scientific process."

—Sergio Rey, director, Center for Geospatial Sciences, UC Riverside

"The editors and authors of this unique graphical science book, published by Esri Press, show the criticality of asking the 'where' question when looking for answers to the 'why' question. It is visually stunning and will certainly lead to an expanded cohort of citizen scientists."

—Noel Cressie, distinguished professor, University of Wollongong, Australia

"Given the relevance of this geospatial perspective for all aspects of society, I hope this beautiful book will inspire a wide range of people to embrace The Science of Where®."

—Thomas Crowther, director, Global Forest Biodiversity Initiative, ETH-Zurich

"With vivid imagery, lucid writing, interactive learning, and compelling, relevant examples from Earth's past, present, and future, *GIS for Science* is a modern manual for understanding that integrative spatial analysis and visualization is the big data revolution most vital to the quality of all life on Earth."

—Healy Hamilton, chief scientist, NatureServe

"There is no better tool to understand our place in the world than GIS, and this book puts its power on beautiful display. It's a book for scientists and all of Earth's stewards."

—Jessica Hellmann, director, Institute on the Environment, University of Minnesota

"This beautifully illustrated and inspiring book brings home the power of today's technology with unique effectiveness, telling and illustrating stories from the earth sciences in novel and powerful ways. A must-have book for anyone concerned about the planet's future."

—Mike Goodchild, distinguished emeritus professor and research professor of geography, UC Santa Barbara

CONTENTS

GIS FOR SCIENCE: A FRAMEWORK AND A PROCESS

by Jack Dangermond, Founder and President, Esri
and Dawn J. Wright, Chief Scientist, Esri

Science—that wonderful endeavor in which someone investigates a question or a problem using reliable, verifiable methods and then broadly shares the result, has always been about increasing our understanding of the world. In the beginning, we applied geographic information systems (GIS) to science—to biology, ecology, economics, or any of the other social sciences. It wasn't until around 1993, when Professor Michael Goodchild coined the term *GIScience,* that the world began to realize that GIS is a science in its own right. Today, we call this The Science of Where®. GIS incorporates sciences such as geology, data science, computer science, statistics, humanities, medicine, decision-support science, and much more. It integrates all these disciplines into a kind of metascience, providing a framework for applying science to almost everything, merging the rigor of the scientific method with the technologies of GIS. The study of where things happen, it turns out, has great relevance.

So why is this work all so important right now? We live in a world that faces more and more challenges. We see, we hear, and we read daily about such issues as growing population (some would say overpopulation), climate change, loss of nature, loss of biodiversity, social conflicts, urbanization, natural disasters, pollution, and political polarization. We also confront the realities of food, water, and energy shortages, and general overconsumption of resources. These concerns are not trivial for the individuals and organizations working in these fields. We must do everything we can to better understand these crucial issues and form better collaborations to address the challenges.

Our world at the same time is undergoing a massive digital transformation. Science always has been about increasing our understanding of the world. But it is also about using that understanding to enable innovation and transformation. It is about what we can measure, how we analyze things, what predictions we make, how we plan, how we design, how we evaluate, and ultimately, how we weave it all together in a kind of fabric across the planet.

What GIS provides is a language to help us understand and manage inside, between, and among organizations, to positively affect the future of the planet. It is also a framework in which we can compile and organize maps, data, and applications. We can visualize and analyze the relationships and patterns among our datasets, perform predictive analytics, design and plan with the data, and ultimately transform our thinking into action to create a more sustainable future. This technology also delivers a new way to empower people to easily use spatial information. As Richard Saul Wurman has said, "Understanding precedes action." Esri is driven by the idea that GIS as a technology is the best way to address the challenges of today and the future.

> Science itself is driven by the organic human instinct to dream, to discover, to understand, to create.

This book is full of examples that show how GIS advances rigorous scientific research. It shows how many science-based organizations use ArcGIS as a comprehensive geospatial platform to support spatial analysis and visualization, open data distribution, and communicate. In some cases, we use this research to preserve and restore iconic pieces of nature—revered and sacred places worthy of being set aside for future generations. These places belong to nature, and they also belong to science.

As scientists, the discipline of the scientific process is the central organizing principle of our work. But science itself is also driven by the organic human instinct to dream, to discover, to understand, to create. The Science of Where is a concept that brings these impulses together as we seek to transform the world through maps and analytics, connecting everyone, everywhere, every day through science. At Esri, we can't wait to see what you and your colleagues will achieve with geospatial technology.

Jack Dangermond   Dawn J. Wright

INTRODUCTION BY THE EDITORS

This book is about science and the scientists who use GIS technology in their work. This contributed volume is for professional scientists, the swelling ranks of citizen scientists, and anyone interested in science and geography. Our world, now two decades into the twenty-first century, seems to be entering a crucial time in history in which humanity still can create a sustainable future and a livable environment for all life on the planet. But if we look critically at the facts, no informed observer can refute the reality that the current downward trajectory does not bode well.

Our first objective in assembling this volume was to select relevant and interesting stories about the state of the planet in 2019. We looked for a cross section of sciences and scientists studying a wide range of problems.

GIS has found its way into virtually all the sciences, but the reader will notice that earth and atmospheric sciences are especially well represented. Web GIS patterns and a simultaneous explosion of earth-observation sensors fuel this growth. Between all the satellites, aircraft, drones, and myriad ground-based and tracking sensors, the science community is now awash in data. Well-integrated GIS solutions integrate all this big data into a common operating platform—a digital, high-resolution, multiscale, multispectral model of our world.

Despite all these advances, science is under attack on many fronts. From fake news to political pressure, science is too often being used as a political tool at a time when level-headed, objective scientific thinking is required. We are convinced that GIS offers a unique platform for scientists to elevate their work above the fray. We invite you to read these stories in any order; the common thread is that all this work happens at the intersection of GIS and science. As you read through these stories, you'll see that GIS is a cross-cutting, enabling technology, whose use is limited only by our imaginations.

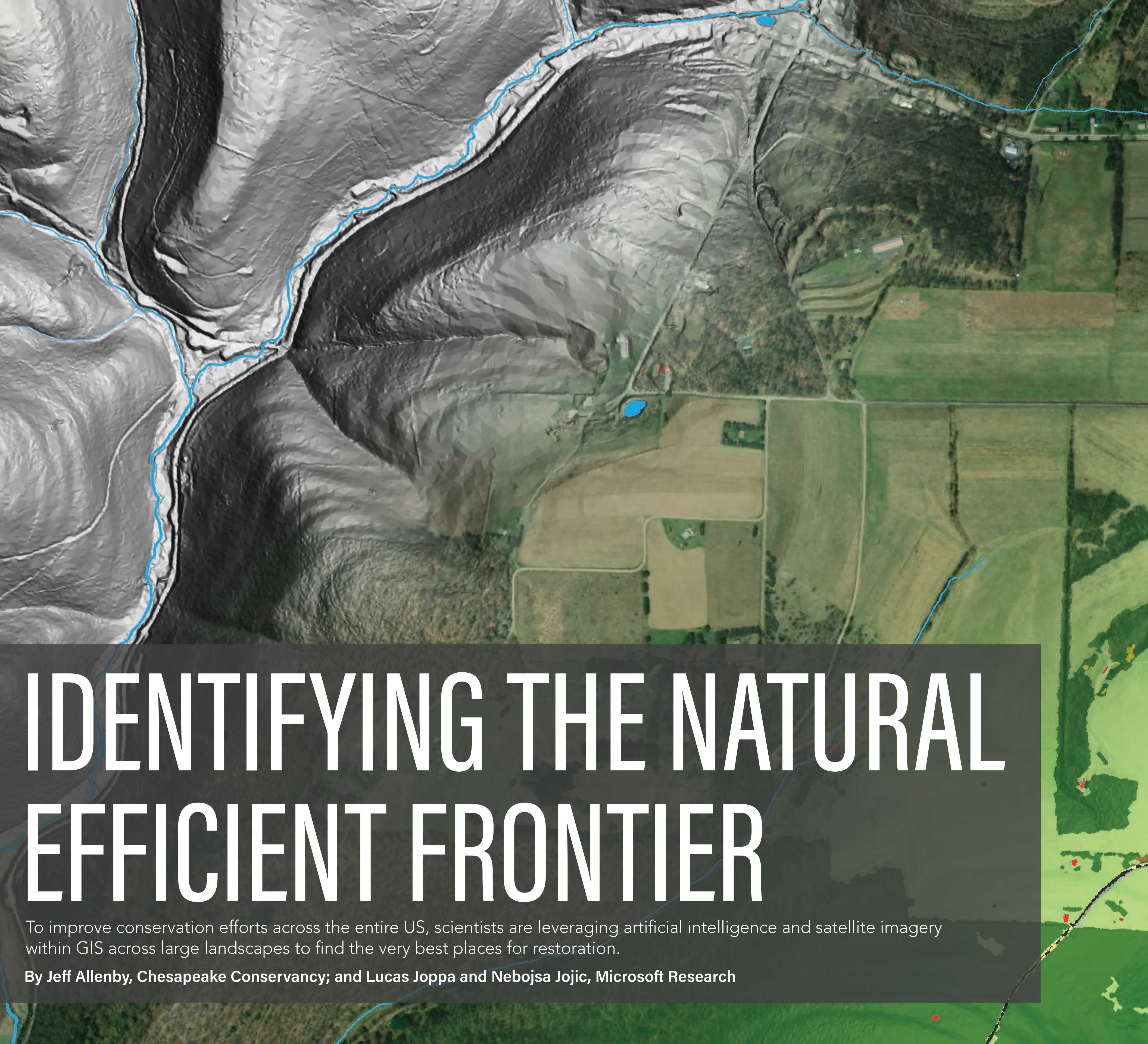
In some cases, like the fascinating work of the US Geological Survey in developing global ecosystem characterizations of the land and ocean, GIS and spatial analysis are at the core of the science. These innovations in science could only happen in the context of an advanced GIS. In other cases, like the story of glaciologists using ground-penetrating radar to measure ice loss in the high-country glaciers of Wyoming, GIS embeds itself in the science but is still mission-critical in terms of expedition planning, backcountry navigation, and analysis. GIS also serves as a vital storytelling platform that brings critical details of important research to stakeholders in the local community.

How the book and website work together

It's impossible to describe the full breadth and scope of what GIS means for science and scientists without showing digital examples. So we have created a companion and complement to this book online. You can access it here:

GISforScience.com

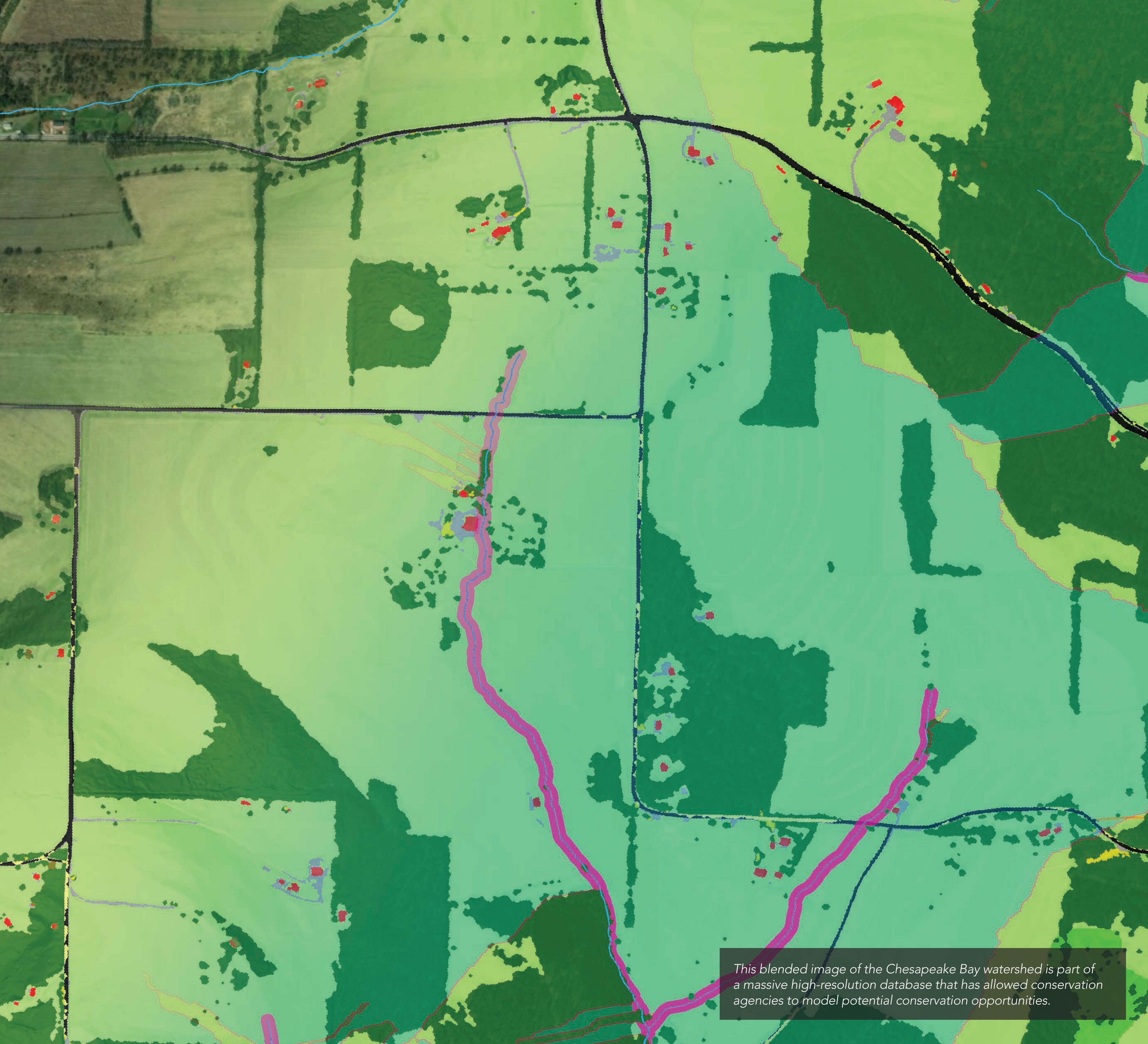
This unique website, comprising collections of ArcGIS® StoryMaps℠ stories, apps, and digital maps, brings the real-world examples to life and demonstrates the storytelling power of the ArcGIS® platform. The website also includes links to learning pathways from the Learn ArcGIS site (Learn.ArcGIS.com) and blogs related to the practical use of ArcGIS in each of the case studies.



IDENTIFYING THE NATURAL EFFICIENT FRONTIER

To improve conservation efforts across the entire US, scientists are leveraging artificial intelligence and satellite imagery within GIS across large landscapes to find the very best places for restoration.

By Jeff Allenby, Chesapeake Conservancy; and Lucas Joppa and Nebojsa Jojic, Microsoft Research



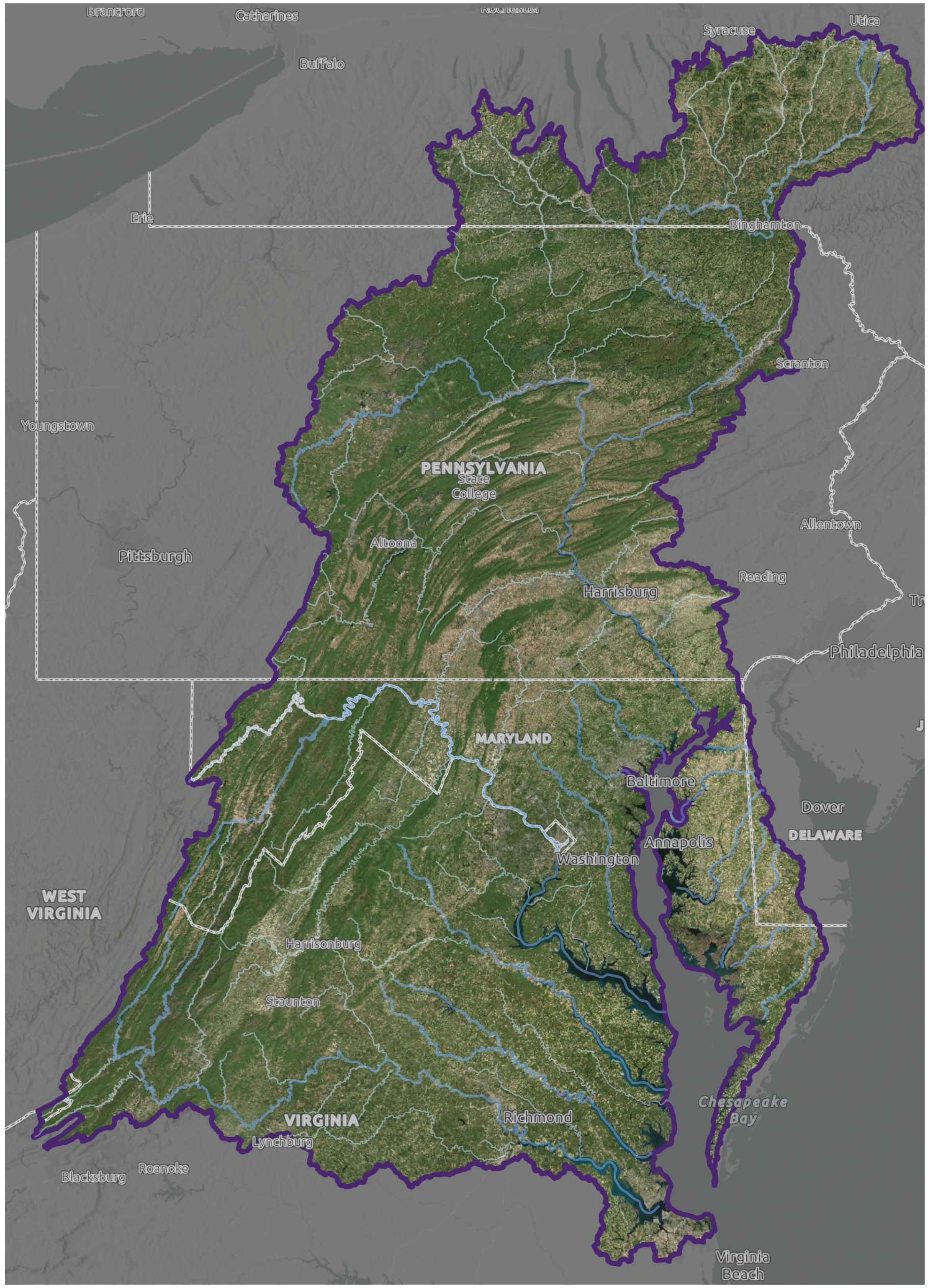
This blended image of the Chesapeake Bay watershed is part of a massive high-resolution database that has allowed conservation agencies to model potential conservation opportunities.

THE CHESAPEAKE BAY WATERSHED: A COMPLEX ENVIRONMENTAL SYSTEM

The Chesapeake Bay is the largest estuary in the United States and the third largest in the world. Its 64,000-square-mile (165-square-kilometer) watershed is home to more than 18 million people and more than 3,600 species of plants and animals. Primarily because of land runoff issues, the health of the bay declined steadily during the twentieth century, resulting in the need for a focused and coordinated effort across the six watershed states and Washington, DC, to improve habitat and water quality.

The Chesapeake Bay Program (CBP), a state-federal partnership led by the US Environmental Protection Agency (EPA), has guided restoration efforts in the Chesapeake during the last three decades, setting progessively more sophisticated restoration goals for each state based on a series of models representing the natural and physical processes of the watershed.

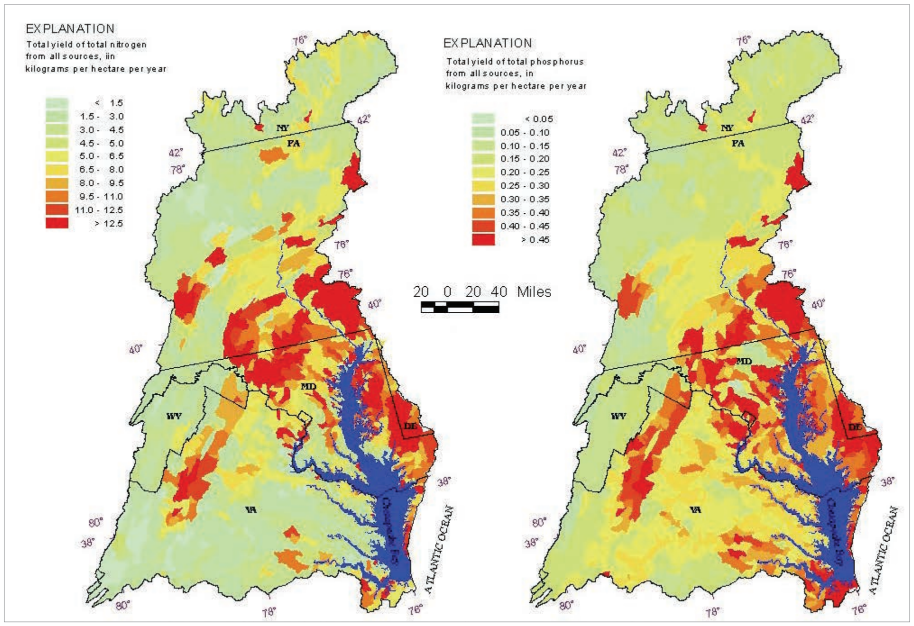
In 2010, the program partnership established a set of brand-new water quality goals—Chesapeake Bay total maximum daily load (TMDL)—a comprehensive "pollution diet" designed to restore clean water in the Chesapeake and the region's streams, creeks, and rivers by 2025. By dividing the Chesapeake watershed into thousands of subwatersheds, the team of state and federal scientists and researchers applied the TMDL thresholds to assess each subwatershed's water quality goals. Given the large scale and complexity of the problem, GIS plays a crucial role in the work.



From a jurisdictional standpoint, managing pollution flows into the Chesapeake Bay watershed is a massive challenge. The watershed drains parts of six states—New York, Pennsylvania, Maryland, Virginia, West Virginia, and Delaware—plus Washington, DC. Its total area exceeds that of 27 US states.

A PARTNERSHIP APPROACH TO RESTORATION

In 2017—roughly halfway into the 15-year project—the CBP conducted a midpoint assessment to determine how much progress had been made and where additional efforts needed to be focused. This assessment found that the watershed states met the interim goals of achieving 60 percent of the necessary reductions, compared with 2009 levels, for sediment and phosphorus reductions, but did not meet the nitrogen reduction goal. In particular, the CBP found that Pennsylvania lagged in meeting its commitments to reduce loads from its urban and agricultural runoff, as well as suburban stormwater sources. Each state is revising its goals and strategies to meet the water quality standards of the TMDL by 2025.



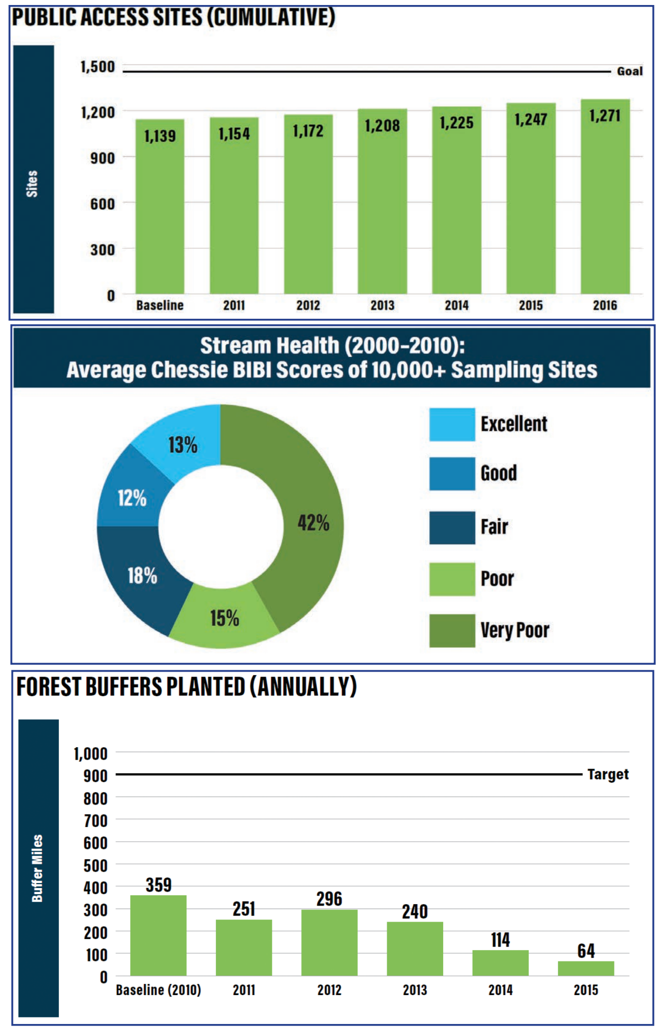
Existing modeling frameworks (like these early GIS maps from a 1992 assessment) have allowed the CBP to identify priority watersheds needing the highest pollution reductions, but local partners did not have the information to identify specific projects that would make the most progress toward these goals.

	Agriculture	Urban/Suburban	Wastewater	Trading/Offsets
Delaware	Enhanced Oversight	Ongoing Oversight	Ongoing Oversight	Ongoing Oversight
District of Columbia	Not Applicable	Ongoing Oversight	Ongoing Oversight	Ongoing Oversight
Maryland	Ongoing Oversight	Enhanced Oversight	Ongoing Oversight	Ongoing Oversight
New York	Ongoing Oversight	Ongoing Oversight	Enhanced Oversight	Ongoing Oversight
Pennsylvania	Backstop Action Levels	Backstop Action Levels	Ongoing Oversight	Enhanced Oversight
Virginia	Ongoing Oversight	Ongoing Oversight	Ongoing Oversight	Ongoing Oversight
West Virginia	Ongoing Oversight	Ongoing Oversight	Ongoing Oversight	Ongoing Oversight

The 2017 Midpoint Assessment found that most watershed states are in need of additional oversight to stay on track to meet their 2025 water-quality goals.

To help meet the TMDL water-quality goals, each of the watershed states signed on to the Chesapeake Bay Watershed Agreement, which defines 10 goals that, if achieved by 2025, indicate a healthy and thriving Chesapeake Bay watershed. Each goal links to a set of management strategies and measurable outcomes with deadlines that are used to assess how each state is achieving success.

Partners throughout the bay's watershed have focused restoration efforts on implementing the 30 strategies outlined in the 2014 Chesapeake Bay Watershed Agreement. While each strategy is unique, they aim to create quantifiable metrics that allow CBP to track progress toward the goal of a healthy Chesapeake. Many of these metrics are simple to measure, but a number of strategies rely solely on quantifying the amount of effort that has gone into a goal, such as totaling the number of stream miles of riparian forest restoration that occurred in the last year.

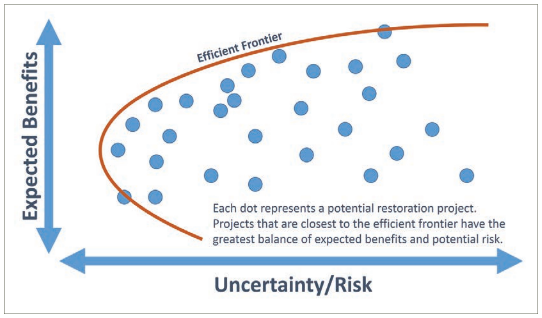


Tracking outcomes for each strategy varies and might detail cumulative metrics or the results of field assessments. In many cases, progress is calculated based on the amount of effort expended each year because appropriate data did not exist to measure implementation in a more comprehensive way (Chesapeake Conservancy 2017).

INVESTING ALONG THE EFFICIENT FRONTIER

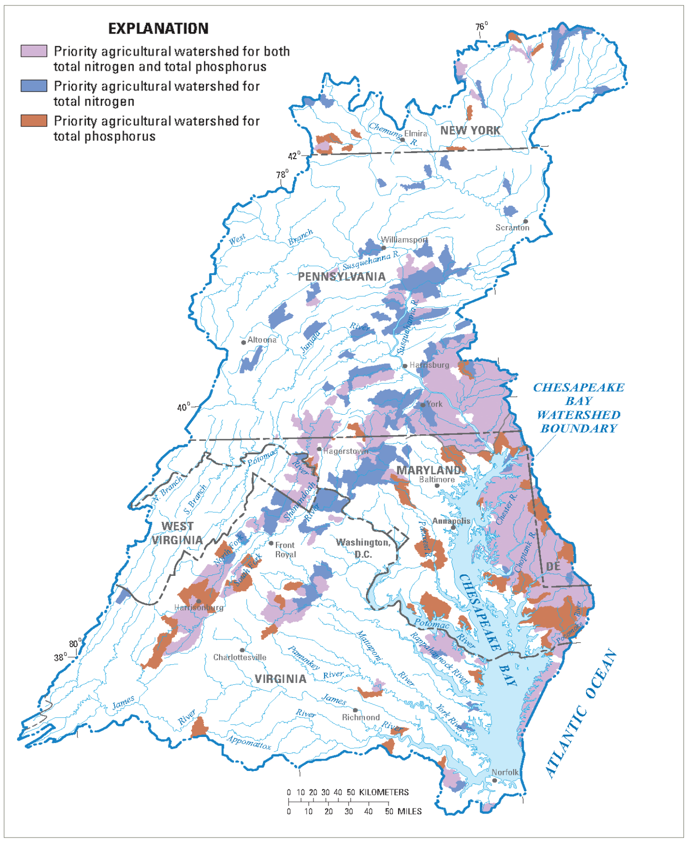
Achieving an optimal balance between risk and reward—a priority across many fields and industries—also applies to environmental management. Frameworks and strategies from finance and investing, for example, are useful in complex environmental scenarios like the Chesapeake watershed problem.

In finance, investors have historically evaluated companies and the markets in which they operate to better understand the likely performance of a stock and gauge its likely risks and returns. In recent decades, because of advancements in computing and data availability, these evaluations increasingly use large datasets to quantify a market's behavior to gain deeper insight into a stock's expected performance. Originally proposed by Harry Markowitz in 1952, the "efficient frontier" represents a hypothetical set of investments that offers the highest expected return for a defined level of risk or the lowest risk for a given level of expected return. By keeping investments near the efficient frontier, investors can maximize their returns while managing the associated risks. Integrating the concept of the efficient frontier into environmental decision making can dramatically improve our ability to invest in the highest performing, most cost-effective projects.

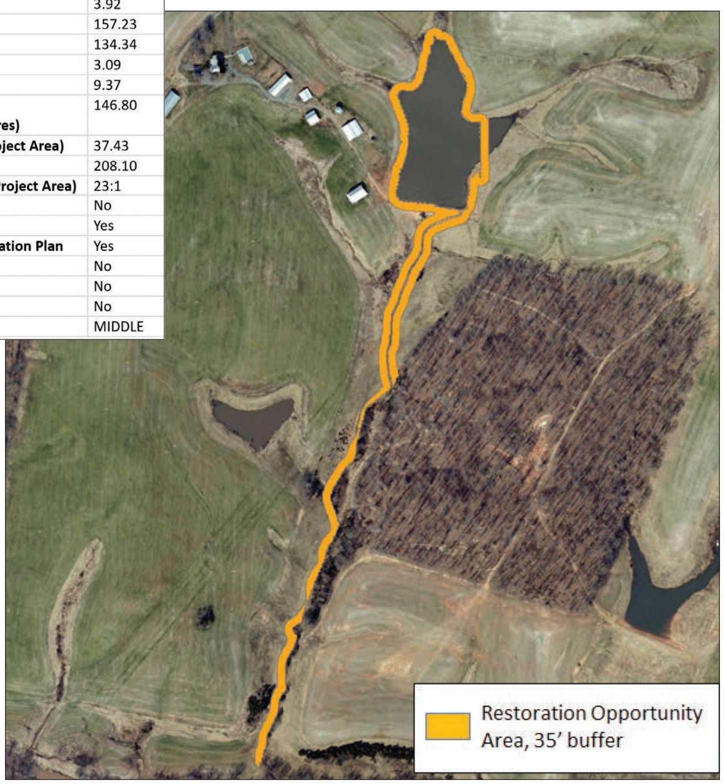


For example, environmental modeling has historically focused on identifying the general landscapes where action is most needed to improve conditions. This approach has helped direct efforts where they will be more effective, improving a program's "return," or success in reaching its goals. Treating all projects within a geography as equally important will result in a wide variability in performance and suboptimal outcomes.

New high-resolution geospatial datasets acquired through precision fieldwork and remote sensing provide a much more detailed understanding of the benefits that a competing proposed project might generate. These new data sources help reduce the risk that a project will deliver lower than expected results. Additionally, these analyses can be completed across entire landscapes, not only demonstrating an individual project's benefits but also providing context for how it compares with all other potential projects in the surrounding area.



Tier	1
Restoration Opportunity Area (ROA) ID	38483
ROA Project Area (Acres)	3.92
Drainage Area (DA) (Acres)	157.23
Agriculture in DA (Acres)	134.34
Impervious in DA (Acres)	3.09
Turf in DA (Acres)	9.37
Total Land Cover of Concern (Agriculture/Impervious/Turf) in DA (Acres)	146.80
Ratio (Land Cover of Concern) : (ROA Project Area)	37.43
Soil Loss in DA (tons/year)	208.10
Ratio (Soil Loss in Drainage Area):(ROA Project Area)	23:1
Tidal	No
On Impaired Waters	Yes
In Sub-Watershed with Local Implementation Plan	Yes
In Area of Potential Effects	No
In Conservation Lands	No
Projected Urban Growth	No
Upper/Middle/Lower James	MIDDLE



High-resolution data and advances in modeling allow natural resource managers to move from evaluating restoration priorities at the watershed scale to identifying the potential benefits of individual projects within a larger landscape context. This level of detail reduces the variability of project performance and facilitates targeted investments to maximize programmatic outcomes.

Maintaining forward momentum through innovation

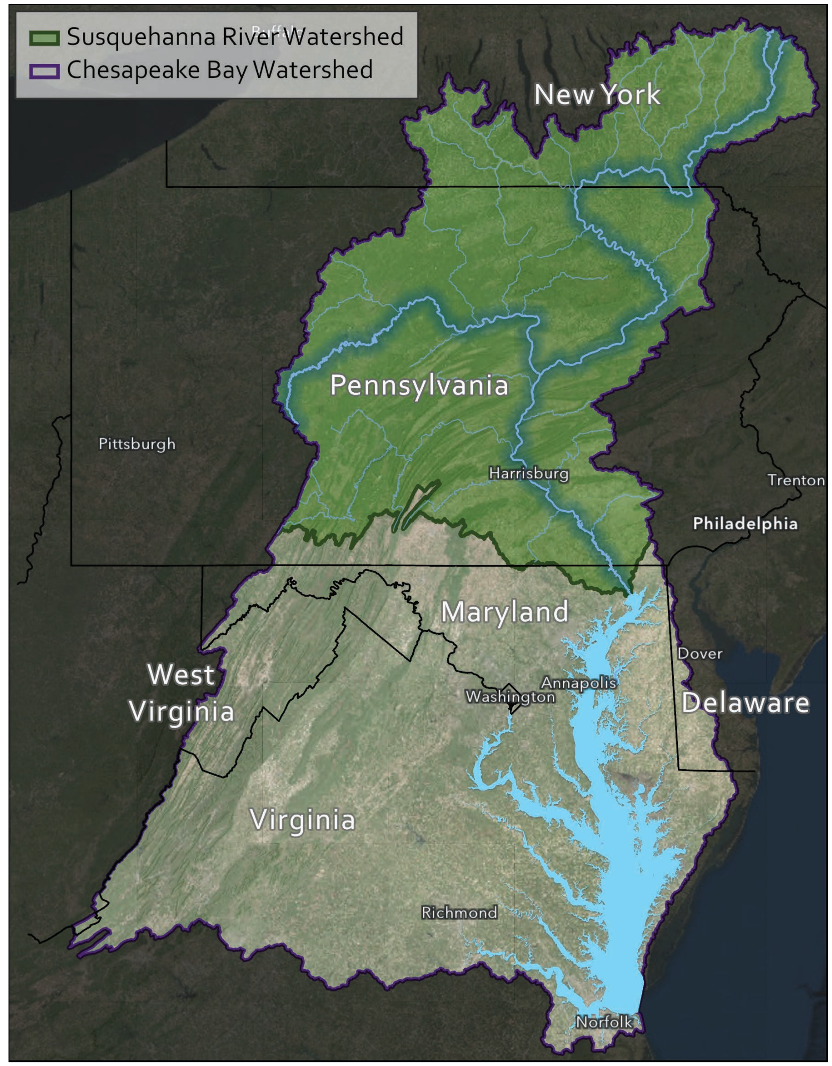
The CBP has effectively coordinated and managed Chesapeake Bay restoration for the last three decades based on strategies that use relatively coarse regional averages. Its ability to model the natural world at a finer scale was imperfect given previously available data and technology. Despite this limitation, the CBP made significant progress toward Watershed Agreement goals. Strategies for improving water quality must evolve going forward because much of the progress to date resulted from implementing the largest and most obvious ways to reduce pollution, such as upgrading wastewater treatment plants.

Future restoration efforts will increasingly focus on identifying and implementing high-performing projects across large landscapes. This transition has required the development of new models that use higher-resolution datasets that feature specific, localized priorities to deliver the greatest results.

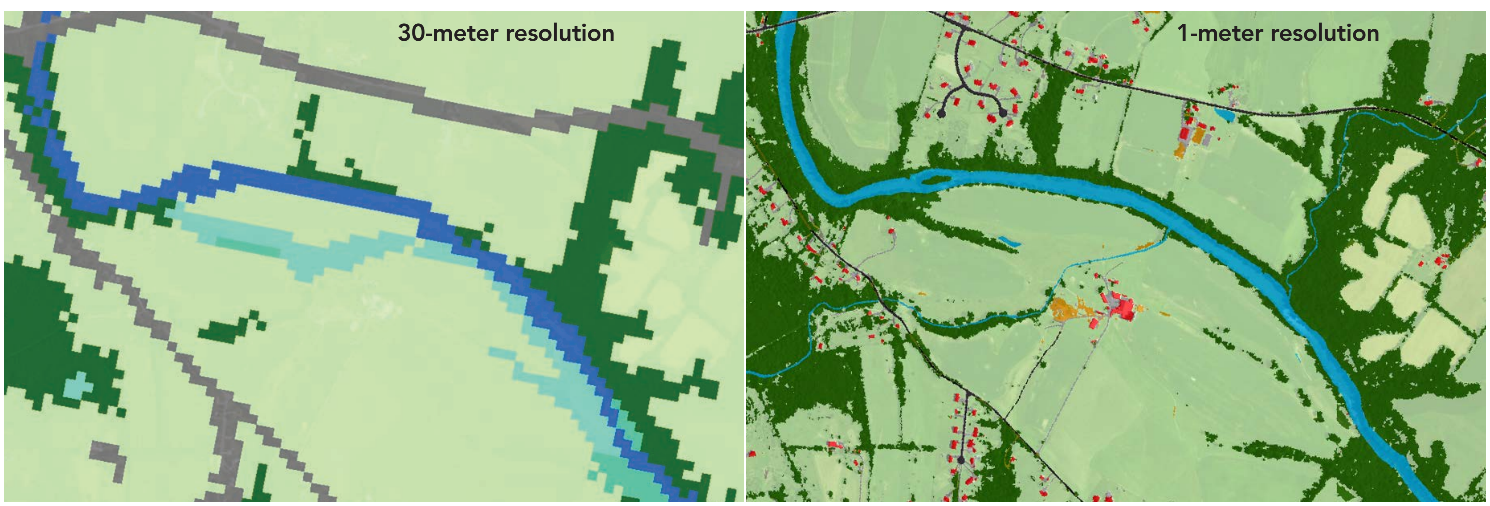
Working in the places that matter most

The Susquehanna River provides about half the freshwater entering the Chesapeake Bay, including all of Pennsylvania's portion of the watershed. The Susquehanna is also responsible for about 40 percent of the total nitrogen load entering the bay. Much of the bad runoff comes from agricultural and stormwater sources. As such, implementing best management practices (BMPs) to reduce nutrient and sediment pollution from agriculture and storm runoff has become a top priority in the restoration of the bay.

Pennsylvania has identified several strategies to reduce the amount of nitrogen, phosphorus, and sediment entering the Susquehanna. Recently, the strategy of replanting forests along streams to create riparian buffers has gained favor. These trees act as a last line of defense against runoff and have been shown to be one of the most cost-effective ways to reduce pollution. Pennsylvania state agencies estimate that more than 95,000 acres of buffers must be planted to meet the Chesapeake Bay Agreement's goal of having 70 percent of riparian areas forested. The Chesapeake Conservancy is working with local partners to identify available planting areas that will treat as much of the 27,500 square miles of watershed as efficiently as possible.



Draining over 27,500 square miles of the total watershed, the Susquehanna River spans parts of three states. Restoration work in the Susquehanna is critical to the overall health of the Chesapeake.



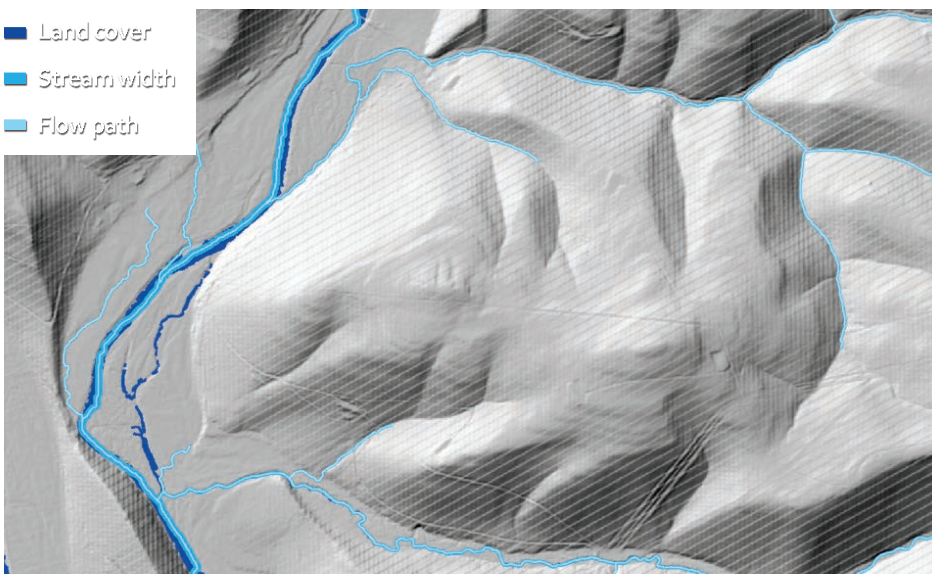
High-resolution data is transforming how managers model the landscape and identify priorities for restoration. New one-meter resolution land cover data shows a dramatically different picture of the landscape, and in particular, natural systems that have the potential to improve water quality, compared with the previously available 30-meter resolution data.

INNOVATING AT THE WATERSHED SCALE

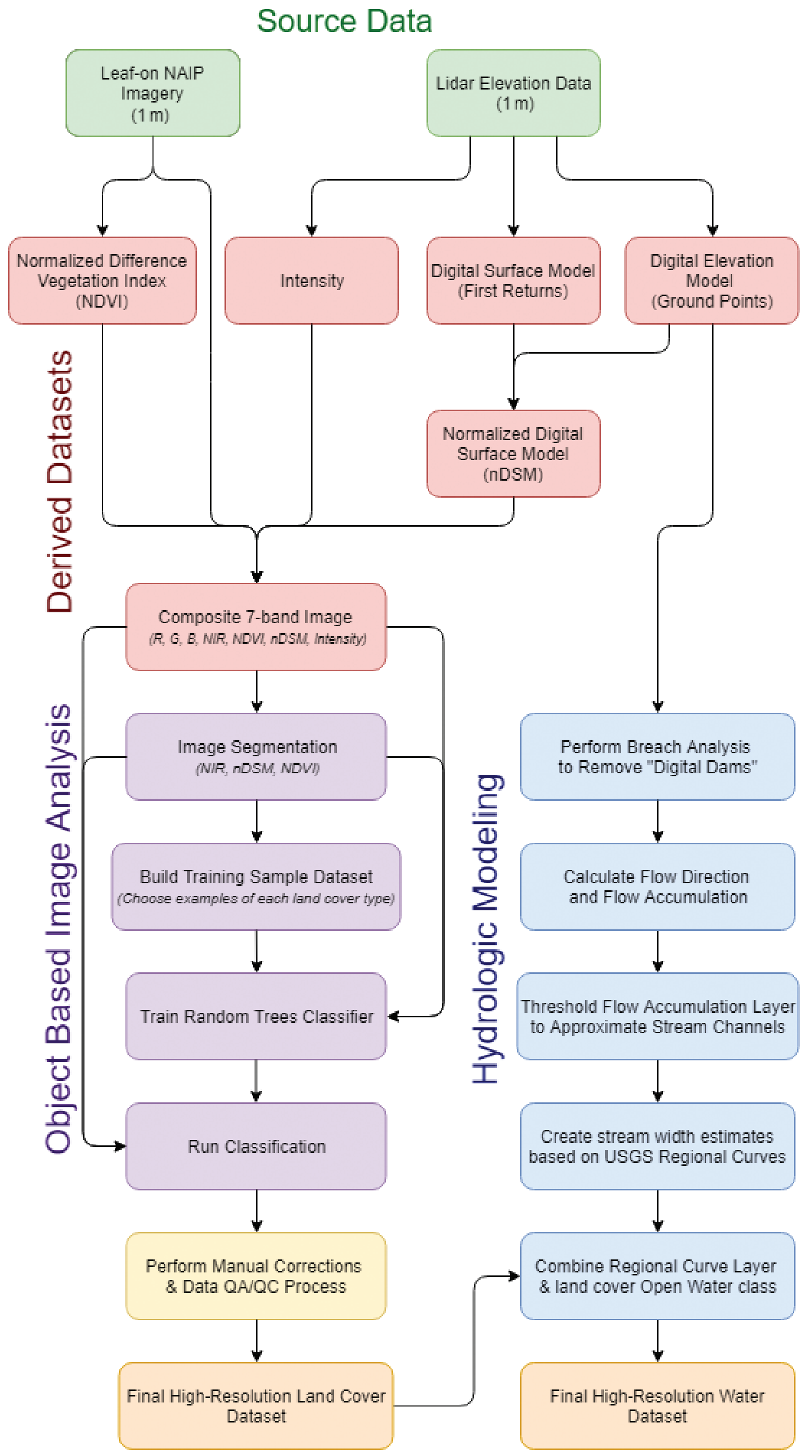
Assessing individual opportunities for restoration across a watershed requires a different mind-set, and very different data than was used previously.

The Chesapeake Conservancy has worked with CBP and other partners since 2015 to create the geospatial data infrastructure for site-level planning. An endless variety of datasets can be incorporated into this type of prioritization, but the Conservancy found that two high-resolution datasets are critical for identifying potential opportunities: one-meter land cover data and lidar-based hydrology data.

Land cover data, created from one-meter resolution, four-band aerial imagery collected through the National Agricultural Imagery Program. This data is processed through an object-based image analysis workflow to segment and classify the imagery into discrete categories describing the landscape.



Hydrology and flow information, generated from lidar elevation data collected by states and the US Geological Survey (USGS). This data is processed through a series of steps to ensure hydrologic connectivity and to identify stream channels.



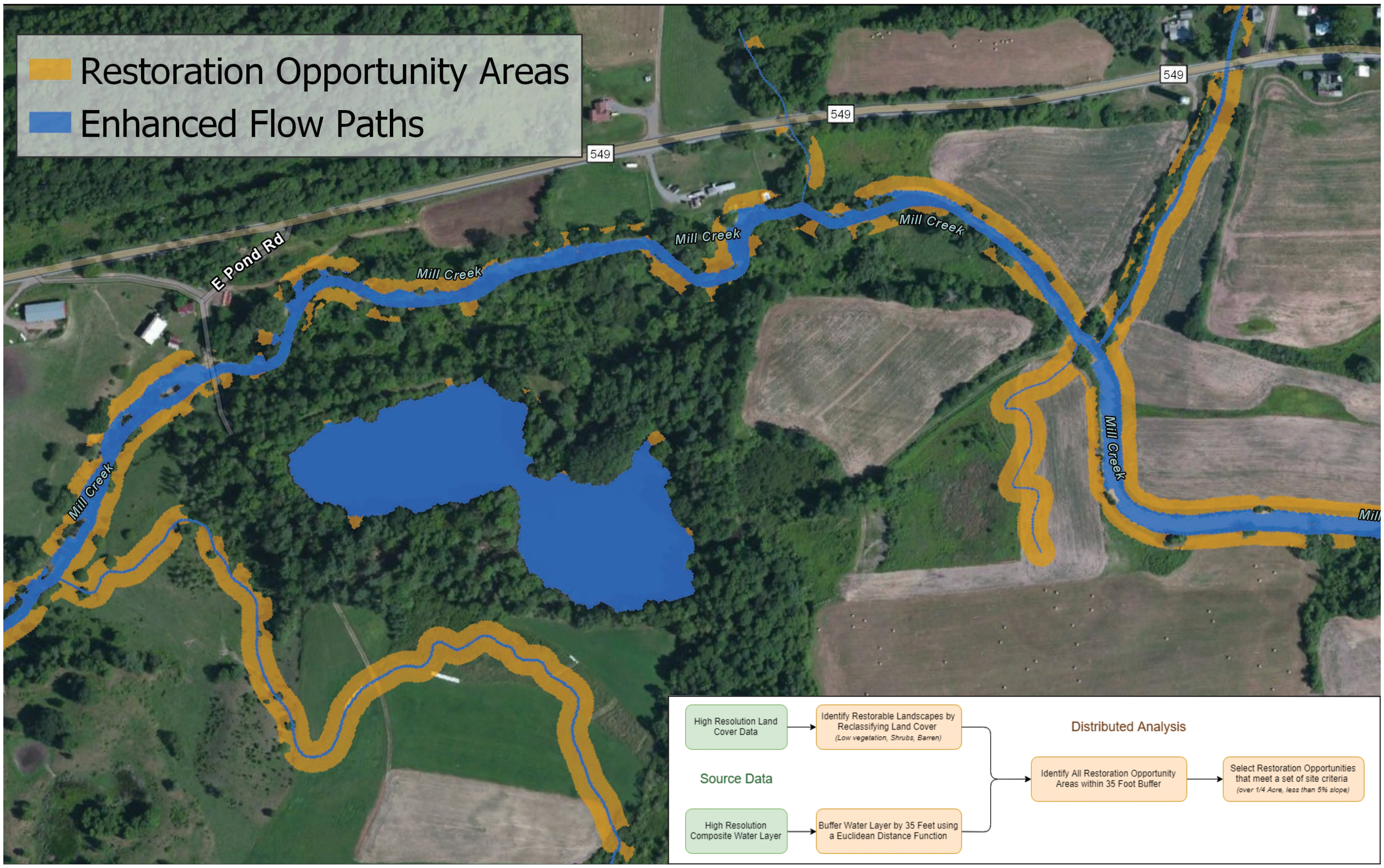
New methods of creating high-resolution datasets make it easier to create comprehensive and highly accurate layers to model the natural world. This information opens new possibilites for landscape analysis and makes it possible to take into account the highly variable nature of the landscape. Conservation and restoration activities can have dramatically different results based on their locations, and these new datasets potentially ensure the implementation of the best projects.

A NEW FRONTIER FOR ENVIRONMENTAL MODELING

In 2016, CBP released the first Chesapeake Bay High-Resolution Land Cover Dataset, the result of an 18-month collaboration between the Chesapeake Conservancy, University of Vermont, and Worldview Solutions. Detailing more than 100,000 square miles, this 12-class, one-meter resolution dataset provides a consistent and accurate understanding of every county comprising the Chesapeake Bay watershed. With this new data, planners can create models that are accurate at high resolutions across large landscapes. This breakthrough is already providing new insights as to how small changes in project placement can dramatically influence performance. This previously unavailable level of detail has also enabled a variety of new analysis opportunities.

Managers can now approach environmental decisions in a far more informed way. Increasingly, decision makers can rely on data-driven processes to compare all options across the watershed and identify landscapes that offer the most potential “bang for the buck.” Decision makers increasingly rely on data-driven processes to compare options across a watershed and identify landscapes for conservation or restoration that offer the best cost-to-benefit ratio. The Chesapeake Bay High-Resolution Project showed that watershed-wide datasets can leverage economies of scale in data generation. And advances in distributive processing showed promise in speeding up complex workflows.

To support restoration efforts in the Susquehanna River and test the capabilities of a large-scale, distributed analysis, the Chesapeake Conservancy partnered with Microsoft Research and Esri to identify all the riparian restoration opportunities in the Susquehanna River watershed. This analysis leveraged the Chesapeake Bay High-Resolution Land Cover dataset and an enhanced water layer created from lidar flow paths that accurately represent the precise location of stream banks. The team used easily scalable, cloud-based servers to process almost 125 billion pixels in a matter of hours, which represents 5 percent of the time it would have taken to run the analysis on a local server. The analysis identified tens of thousands of acres of potentially restorable land in the Susquehanna River watershed.



Chesapeake Conservancy worked with Microsoft Research and Esri to adapt a traditional raster analysis workflow to operate on a cluster of virtual servers. This analysis identified restorable areas within 35 feet (11 meters) of streams and areas of concentrated flow in the Susquehanna River watershed. This distributed process ran approximately 20 times faster than if it was excuted on a set of local machines.

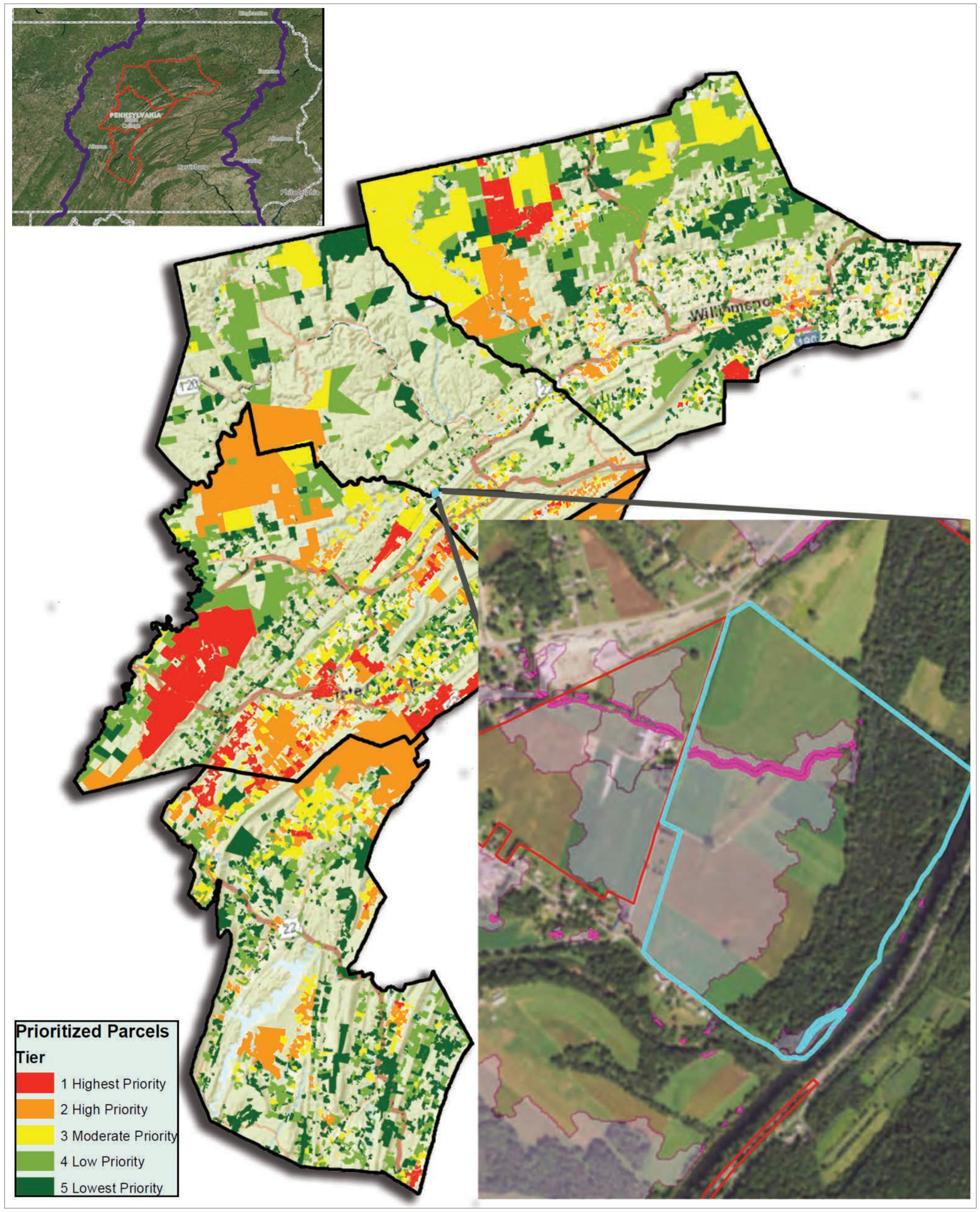
INFORMING LOCAL PRIORITIES

The Chesapeake Conservancy is working with local organizations across four counties in central Pennsylvania to further prioritize each of the 250,000 opportunities identified through the pilot study with Microsoft Research and Esri. Local partners are using this information to inform local restoration initiatives for a portion of the Susquehanna River.

This process assesses the additional benefits a buffer restoration project could provide, such as increasing habitat for threatened species, improving climate resilience, and expanding regional ecosystem connectivity. An increased understanding of the benefits of an expanded riparian buffer can help managers determine the relative importance of a potential project beyond water-quality improvements. Each project receives a “site score” detailing its benefits and a “designation score” based on whether it is located in a priority watershed. Partners can use the composite of these two values to evaluate projects and better understand how well the project meets organizational goals.

In its first year of use, ten projects were evaluated using the conservancy’s analysis prior to being implemented, including a project that was ranked highest of any potential project in all four counties.

Riparian buffer restoration projects, such as this site in Centre County, Pennsylvania, can significantly reduce the amount of nutrient and sediment pollution entering local streams and, eventually, the Chesapeake Bay.



Assessing priorities across a four-county region allowed the Conservancy and its partners to focus on a small subset of high-performing projects. Using this approach, the Conservancy was able to narrow the number of opportunities from over 200,000 to the top 278 projects, or 0.1 percent. These projects, on average, treat the runoff of 40 acres of upland for every acre of restoration, ten times what is used in CBP's modeling.

MOVING THE DIAL

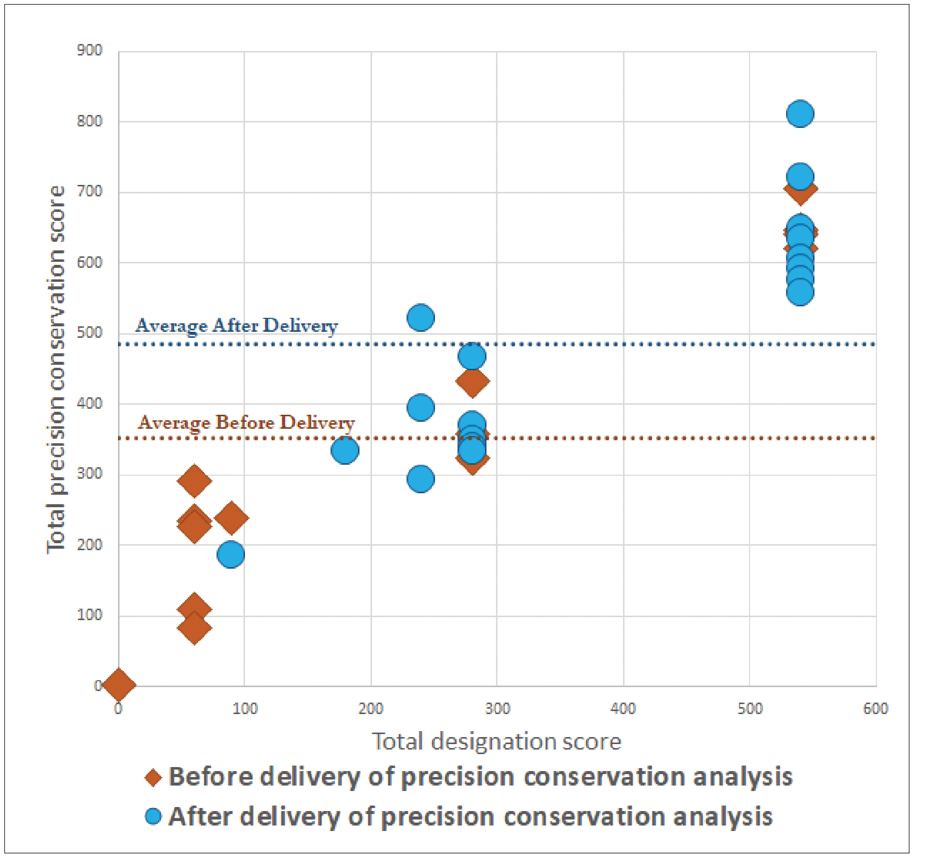
Prior to having this prioritization data, restoration projects were often selected based solely on the size of the project or, in some cases, simply because there was a willing landowner. This approach would allow partners to meet effort-based goals but has a high potential for low-performing projects to be implemented.

When the Conservancy analyzed the conservation scores from the first year's projects, generated by evaluating each project's anticipated performance and location in impaired watersheds, there was a wide variety of scores, including some extremely low-value projects. Partners reflected on how access could benefit their programs and made a commitment to use the planning data in their community outreach and recommend only higher-ranking projects for funding.

Riparian buffers act as a last line of defense for stopping pollution before it enters streams, and eventually the Chesapeake Bay. Ideally, these trees will be planted where the landscape converges so they intercept as much runoff from upslope as possible. Traditionally, models have not had the detail to measure the actual amount of area being filtered by a restoration project, so an average value was used. With the detail in our elevation data, however, we can specify the exact landscape being treated by each project and incentivize higher-performing projects.

The projects selected in our second year of funding should filter pollution from 11 times the amount of upland area that the previous models—based on a fixed upland value—would have treated. The treated area represented an 18 percent increase from the previous year's average. This four-county pilot region demonstrates the potential for new models to have an outsized impact on the performance of restoration work throughout the Chesapeake Bay and around the world.

The Chesapeake Conservancy is now working with CBP and other nonprofit and academic partners to build the datasets needed to apply models across the watershed similar to the ones developed through this project.



Conservationists now evaluate projects within a system that compares each proposal not only within the subset of projects that were submitted, but also within the larger context of the landscape. Projects implemented prior to using the prioritization system (orange diamonds) had a wide variety of performance, including one project that provided little environmental value. In the second year, projects that had previously been identified and ranked high were advanced due to their performance (blue circles), showing how a small change in management can lead to a dramatic improvement in performance.

Volunteers plant native tree species along Elk Creek in Centre County, Pennsylvania. These trees will help reduce pollution by stabilizing the stream's banks and filtering out nutrients and sediment that flow downhill from the surrounding landscape.

LEVERAGING THE POWER OF AI

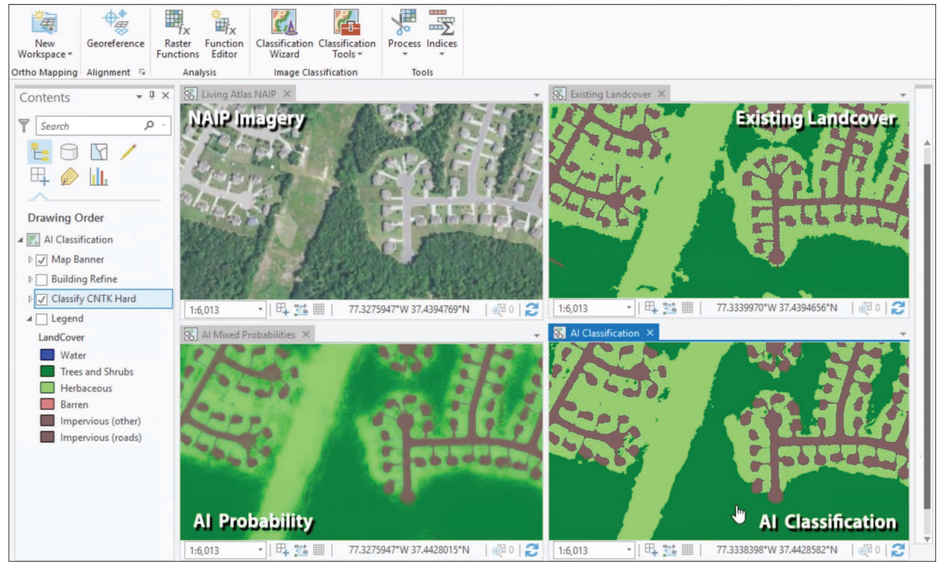
In 2017, Microsoft Research launched its AI for Earth initiative to focus the capacity of artificial intelligence on improving environmental decision making around the world. The conservancy joined the AI for Earth initiative to focus on the use of artificial intelligence in the bay's restoration. Specifically, the conservancy initiated a project to rapidly update its high-resolution land cover data as new imagery becomes available.

One of the challenges facing any approach that uses AI is the need for highly accurate training data that represents a variety of conditions. The Chesapeake Bay High-Resolution Land Cover Dataset represented an ideal training set because of its geographic scope, the variety of landscapes it includes, and its high degree of accuracy. This system uses Microsoft's Cognitive Toolkit, an architecture designed to easily train deep learning algorithms. The system also uses imagery from the National Agriculture Imagery Program (NAIP) stored on Microsoft's Azure Cloud and a new platform called Project Brainwave that integrates Intel's FPGAs—field programmable gate arrays—to increase the throughput capacity of Azure while providing increased flexibility as the system scales.

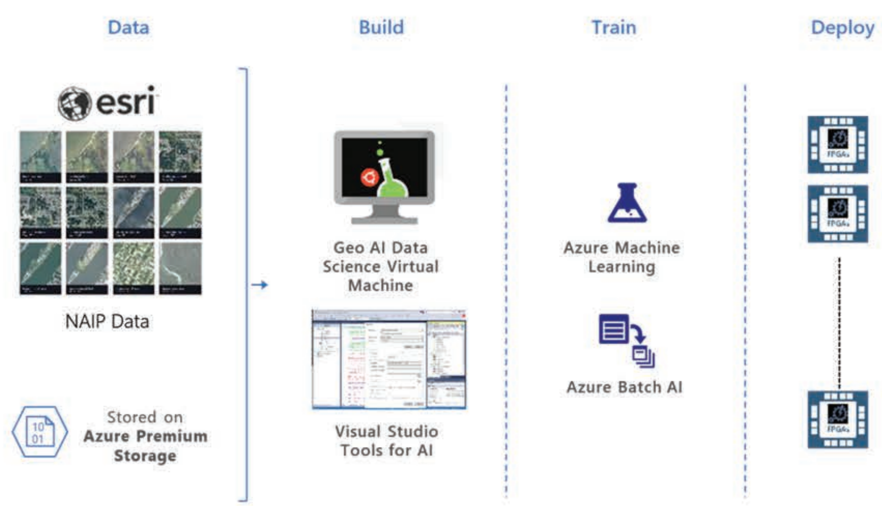
A deep learning algorithm analyzed the existing land cover data in combination with the four-band NAIP imagery to create a simplified four-class land cover dataset in the Chesapeake. With the algorithm trained, two datasets were created, one based on 2013/14 NAIP imagery (the source data for the Chesapeake Bay High-Resolution Land Cover) and one based on 2016/17 NAIP imagery. This tested the platform's ability to create a classification based on imagery that was different from the data on which the system was trained. Both datasets ran successfully, and the conservancy is comparing the two time periods to identify areas that have experienced significant levels of change over the three-year period. This analysis will inform future updates to the Chesapeake Bay High-Resolution Land Cover dataset.

The conservancy's work had moved into geographies that lacked high-resolution land cover data, including landscapes where lidar data was not available. As a result, the partners expanded the system so that it could create high-resolution land cover data anywhere in the United States.

Project Brainwave significantly accelerated the processing capabilities of this large-scale analysis and completed the classification in a little over 10 minutes at a cost of $42. The classification does a good job of detailing landscapes that look visually similar to the Chesapeake, such as the Great Lakes or Pacific Northwest, but struggled in arid landscapes. Although it doesn't completely replace our previous workflow yet, this system illustrates how AI has the potential to quickly create high-quality data.



The land cover algorithm the Conservancy developed with Microsoft and Esri leverages NAIP imagery (top left) and the Conservancy's Chesapeake High Resolution Land Cover (top right) to predict the likelihood of each pixel matching one of four classes (bottom left). Based on this probability layer, the algorithm chooses the most likely class for the final classification (bottom right).



Performing a change analysis using two time periods of land cover data in Microsoft's Azure Cloud allowed the Chesapeake Conservancy to quickly identify priority landscapes in updating the Chesapeake Bay High-Resolution Land Cover dataset.

AI IS CHANGING THE WAY WE THINK

Access to current information is critical for environmental management so that conservation and restoration partners can accurately understand the world around them. Ensuring data accuracy can be costly and can leave less funding for program implementation.

AI-generated datasets can save money and time as the conservancy and its partners update critical landscape data. As these systems improve, updated landscape data could be automatically processed each time a satellite passes over Earth.

This technology will help partners increasingly focus their efforts on the areas that are experiencing the most change. Instead of noticing the impacts to an ecosystem by observing a decline in its health years after the fact, AI-based systems will provide an early-warning system as critical landscape thresholds, such as the percentage of impervious surfaces in a watershed, are exceeded.

Updates to the Chesapeake Bay land cover are already being informed by a rapid-change analysis leveraging the two time periods of data created in Azure. This analysis is helping the conservancy identify where hotspots of change have occurred during the last three years and an updated land cover dataset is needed to better reflect the landscape.

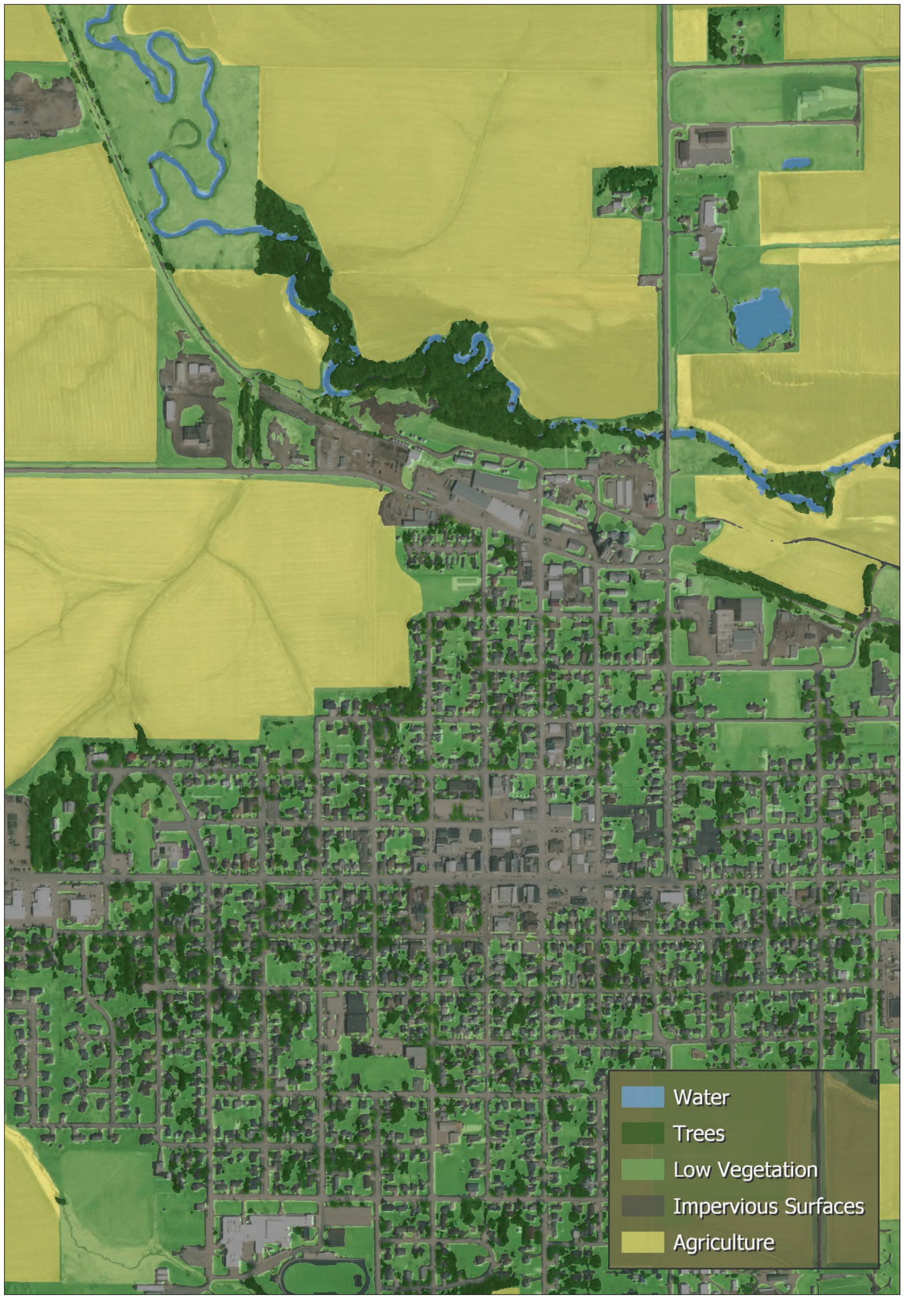
Moving to other landscapes

Managing water quality is a priority for all landscapes. Chesapeake Conservancy's approach to improving conservation and restoration design has the potential to improve decision making and ensure that projects are as efficient as possible.

One of the greatest limiting factors to expanding these analyses, however, is the time and cost of creating the base datasets needed to run more advanced modeling. Integrating AI-based workflows can remove these barriers and expand precision planning capabilities to new landscapes.

In 2017, the Iowa Agricultural Water Alliance asked the conservancy about applying the Susquehanna pilot program to the Black Hawk Creek headwaters, a 160-square-mile subwatershed in central Iowa. This project was part of a larger initiative working to identify ways to accelerate the process of developing more than 1,600 watershed management plans spanning the state.

This project demonstrated the ability of the Azure-based land cover workflow to accurately classify a landscape outside the Chesapeake. More importantly, the workflow allows projects to be completed within a tight budget and extends planning capabilities to landscapes where generating new data would be impossible otherwise.



Land cover data created in Azure allowed Chesapeake Conservancy to assess restoration opportunites on agricultural lands along Black Hawk Creek in Iowa. This project would not have been feasible given the available budget if the data had to be generated manually.

WHERE WE GO FROM HERE

Chesapeake Conservancy is expanding its portfolio of work and applying its expertise outside the Chesapeake Bay watershed to work in landscapes across the United States and, increasingly, around the world. The Conservancy's approach to precision conservation has the potential to improve environmental outcomes for both water quality and water quantity management as well as a host of other environmental issues such as habitat conservation and climate adaptation. However, moving toward a performance-based system of management requires access to high-quality, and often high-resolution, datasets to be as effective as it has been in the Chesapeake.

To help address these management challenges in varied landscapes, the Conservancy is continuing to work with Microsoft to refine the land cover classification algorithms to improve the accuracy of the output, make it easier to retrain the algorithms in new landscapes, and explore the potential to add additional classes. This work holds a tremendous amount of promise to rapidly produce usable datasets that can easily be shared with partners.

Most importantly, this system is enhancing the Conservancy's capacity to generate repeat datasets whenever new imagery becomes available. With an influx of new high-resolution satellite datasets, access to this platform will allow us to create a "living" dashboard of the watershed that reflects conditions as soon as they change instead of only incorporating changes years after it has occurred. Having this dynamic system will provide managers with the information they need to customize and tailor solutions where they will be most effective.

Clear waters in the lower Chesapeake Bay deep in Virginia.