UC Berkeley
Energy Institute Working Papers

Title

Machine Learning Predicts Which Rivers, Streams, and Wetlands the Clean Water Act Regulates

Permalink

https://escholarship.org/uc/item/31c855v4

Authors

Greenhill, Simon
Druckenmiller, Hannah
Wang, Sherrie
et al.

Publication Date

2024-02-01

eScholarship.org Powered by the California Digital Library
University of California

Energy Institute WP 346

Machine Learning Predicts Which Rivers, Streams, and Wetlands the Clean Water Act Regulates

Simon Greenhill, Hannah Druckenmiller, Sherrie Wang, David A. Keiser, Manuela Girotto, Jason K. Moore, Nobuhiro Yamaguchi, Alberto Todeschini, and Joseph S. Shapiro

February 2024

Revised version published in *Science*, 383(6681), 406-412, 2024

Energy Institute at Haas working papers are circulated for discussion and comment purposes. The Energy Institute acknowledges the generous support it has received from the organizations and individuals listed at https://haas.berkeley.edu/energy-institute/about/funders/.

2024 by Simon Greenhill, Hannah Druckenmiller, Sherrie Wang, David A. Keiser, Manuela Girotto, Jason K. Moore, Nobuhiro Yamaguchi, Alberto Todeschini, and Joseph S. Shapiro. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit is given to the source.

Machine Learning Predicts Which Rivers, Streams, and Wetlands the Clean Water Act Regulates

Simon Greenhill*[1,2], Hannah Druckenmiller[3,4]†, Sherrie Wang[2,5,6]†, David A. Keiser[7,8,9], Manuela Girotto[10], Jason K. Moore[11], Nobuhiro Yamaguchi[12], Alberto Todeschini[12], Joseph S. Shapiro*[1,9,13]

Affiliations:

[1] Department of Agricultural and Resource Economics, University of California, Berkeley; Berkeley, 94720, USA

[2] Goldman School of Public Policy, University of California, Berkeley; Berkeley, 94720, USA

[3] Resources for the Future; Washington, DC, 20036, USA

[4] Division of Humanities and Social Sciences, California Institute of Technology; Pasadena, 91125, USA

[5] Department of Mechanical Engineering, Massachusetts Institute of Technology; Cambridge, 02139, USA

[6] Institute for Data, Systems, and Society, Massachusetts Institute of Technology; Cambridge, 02139, USA

[7] Department of Resource Economics, University of Massachusetts, Amherst; Amherst, 010013, USA

[8] Center for Agricultural and Rural Development, Iowa State University; Ames, 50011, USA

[9] National Bureau of Economic Research; Cambridge, 02139, USA

[10] Department of Environmental Science, Policy, and Management, University of California, Berkeley; Berkeley, 94720, USA

[11] Department of Energy, US Government; Washington, 20585, USA

[12] School of Information, University of California, Berkeley; Berkeley, 94720, USA

[13] Department of Economics, University of California, Berkeley; Berkeley, 94720, USA

*Corresponding authors. Email: sgreenhill@berkeley.edu, joseph.shapiro@berkeley.edu

† These authors contributed equally to this work

Abstract: We assess which waters the Clean Water Act protects and how Supreme Court and White House rules change this regulation. We train a deep learning model using aerial imagery and geophysical data to predict 150,000 jurisdictional determinations from the Army Corps of Engineers, each deciding regulation for one water resource. Under a 2006 Supreme Court ruling, the Clean Water Act protects two-thirds of US streams and over half of wetlands; under a 2020 White House rule, it protects under half of streams and a fourth of wetlands, implying deregulation of 690,000 stream miles, 35 million wetland acres, and 30% of waters around drinking water sources. Our framework can support permitting, policy design, and use of machine learning in regulatory implementation problems.

One-Sentence Summary: A deep learning model finds that Clean Water Act regulation differs dramatically across Supreme Court and White House rules.

This is the author's version of the work. It is posted here by permission of the AAAS for personal use, not for redistribution. The definitive version was published in *Science* on January 25, 2024 383(6681), DOI: https://www.science.org/doi/10.1126/science.adi3794.

The 1972 Clean Water Act (CWA), a critically important US environmental policy, represents the cornerstone of federal water quality regulation. Given the importance of healthy waterways for flood protection, clean drinking water, ecosystem health, and economic activity (*1–3*), the CWA and reforms to it have enormous potential ecological and economic consequences.

Four recent judicial and executive CWA rules have substantially rewritten CWA coverage—*Rapanos*, the Clean Water Rule (CWR), the Navigable Waters Protection Rule (NWPR), and *Sackett*. A third of the US Supreme Court's environmental cases since 1972 have addressed the CWA, far more than any other environmental policy (*4*). Supreme Court Justice Kennedy's 2006 *Rapanos* opinion found that the CWA protects water resources with a "significant nexus" to navigable waters, meaning a biological, chemical, or physical connection. Justice Scalia's *Rapanos* plurality opinion requiring a surface water connection was the basis for the Trump Administration's NWPR, which excluded isolated wetlands and ephemeral streams, and required a surface water connection to navigable waters. The Obama Administration's CWR clarified jurisdiction under *Rapanos*, though did not seek to change jurisdiction dramatically. The Biden Administration implemented *Rapanos* with modest modifications (*5, 6*). The Supreme Court's 2023 *Sackett* decision limits regulation, especially the significant nexus standard.

The CWA protects the "Waters of the United States," but does not define which waters this phrase describes. This makes it difficult to understand precisely which waters gain and lose protection under recent rules. The CWA originally protected navigable waters and their tributaries, under the Constitution's interstate commerce clause. The CWA targets pollution, though CWA Section 404's regulation of the discharge of dredged or fill material into jurisdictional waters affects land use development.

The US Environmental Protection Agency (EPA) and Army Corps of Engineers (ACE, *7*) summarize, "**EXISTING TOOLS CANNOT ACCURATELY MAP THE SCOPE OF CLEAN WATER ACT JURISDICTION**" (formatting in original). Accurate mapping has been infeasible because the CWA and rules interpreting it give sufficiently general guidance that ACE must evaluate the geophysical conditions of a water resource to determine whether the CWA regulates it.

Media reports and an *Amicus* brief by the American Water Works Association, National Association of Wetland Managers, and others, for example, assert that NWPR eliminates CWA protection for at least 18% of streams and 51% of wetlands (*8*, *9*). Such statistics identify waters sharing specific characteristics that loosely approximate criteria for regulation in a CWA rule, then assume those waters are regulated (*10*). The EPA and ACE (*7*) call these statistics "highly unreliable" due to their lack of data on which waters are regulated.

This paper provides the first national, geographically resolved estimate of legally binding CWA regulation. Waters of the United States-Machine Learning (WOTUS-ML), a deep learning model we build, predicts CWA jurisdictional determinations under *Rapanos*, CWR, and NWPR. WOTUS-ML also classifies water resources into regulatory and hydrological categories. The Biden Administration's CWA Rule (*2*) called for "machine learning and artificial intelligence methods to develop a jurisdictional status predictive model," which this paper provides. Our training data come from ACE records of 150,680 Approved Jurisdictional Determinations (AJDs), legally-binding case-by-case decisions ACE engineers make, which represent possible water resources (though 15% of AJDs are uplands). Existing research has not analyzed all these AJDs.

Jurisdictional determinations work as follows (Fig. S1). A developer (e.g., a factory builder) where jurisdictional waters may be present can ask ACE to provide an AJD, which can take months, due to backlogs or ACE's desire to observe a site in multiple seasons. A developer wishing to minimize delay and uncertainty or believing the water is jurisdictional may alternately provide a Preliminary Jurisdictional Determination (PJD). If the water has a PJD or if an AJD concludes the water is jurisdictional, the CWA requires the developer to obtain a Section 404 Permit, which may mandate compensatory investments, change development plans, or involve interactions with the Endangered Species Act (*11*). Non-jurisdictional waters face no CWA regulation. Development of jurisdictional waters without a permit can incur penalties and require the site be returned to its original state. Since we observe much of the data that ACE engineers use, our setup somewhat re-creates the ACE engineer's decision problem.

The WOTUS-ML Model

We use the AJDs to train WOTUS-ML (Fig. S2). The WOTUS-ML architecture is the widely-used ResNet-18 convolutional neural network (Fig. 1) (*12*, *13*). We predict whether a site is regulated and which of nine hydrological (*14*) and nine legal classifications of water types the site represents, e.g., whether it is an isolated wetland or ephemeral stream, according to either the leading scientific classification of wetland and stream types, or according to a CWA rule's language (*13*; Tables S1, S2). We pool data on the *Rapanos*, CWR, and NWPR rules and include an input layer identifying which rule each AJD used. Since *Sackett* AJDs begin in late 2023, after this study's timeframe, we do not train or predict on *Sackett*, though we discuss methodologies relevant to it (*13*). We divide the ground-truth data into disjoint test, training, and validation sets (*13*; Fig. S3). Because ACE requests AJDs to list water resource centroids, we interpret WOTUS-ML as better suited to classify whether a resource is regulated than at delineating wetland boundaries.

The model receives as input an image of 34 layers, which include red, green, and blue (RGB) and near-infrared bands from aerial photographs from the National Agricultural Imagery Program (NAIP) (*15*); soil, groundwater, and elevation data from the Gridded National Soil Survey Geographic Database (*16*); hydrological information from the National Hydrography Dataset (NHD) (*17, 18);* wetland coverage from the National Wetland Inventory (NWI) **(***19*); ACE regulatory district and state boundaries (*20*); and related records (*13* and Table S3). Fig. S4 maps several layers. Fig. S5 shows inputs for one site. We use images from at least several months before each AJD to avoid temporal leakage (*13*). Political boundaries might reflect political forces within a CWA rule that could change, though ACE engineers are not political appointees and we find stable patterns within ACE districts over time.

WOTUS-ML outputs scores ranging from zero to one that a given latitude and longitude pair is regulated, separately for *Rapanos*, NWPR, and CWR. Zero represents confidence the point is unregulated; one represents confidence the point is regulated (Fig. S6). When evaluating predictive accuracy, we round scores to binary predictions—"regulated" versus "not regulated." WOTUS-ML also outputs scores for each of nine Cowardin codes and nine resource types.

Model Accuracy and Bias

WOTUS-ML correctly predicts outcomes for 79% of AJDs in a held-out test set (Table S4). The area under the model's receiver operating characteristic curve (AUC) is 0.85 (Fig. S7). Among test

set AJDs, 35% are regulated, so model learning accounts for a 14 percentage point improvement in accuracy above a naive baseline. We measure model learning as the difference between accuracy and max(share regulated, share unregulated). Type I and II errors vary by rule (Table S5). WOTUS-ML has accuracy below 100% for several reasons (*13*).

WOTUS-ML accuracy varies across settings (Table S4, S6-S7). Test set accuracy is 82% for AJDs without an ACE field visit and 74% for AJDs with a field visit. Field visits provide information unavailable to WOTUS-ML (*21*). WOTUS-ML has similar accuracy on wetlands and streams but greater accuracy for estuaries, which are always regulated. The model is extremely accurate in ACE districts such as St. Paul, which covers Minnesota and Wisconsin, and where regulation rates are low. WOTUS-ML has accuracy around 80% for sites with characteristics typical of AJDs and for sites very different from typical AJDs, which supports its external validity nationally (*13*, Fig. S8). WOTUS-ML has 75% accuracy for identifying ephemeral resources, which is useful given the absence of national maps of such resources (Table S8). We focus on WOTUS-ML's binary analysis of jurisdiction, which has greater accuracy than its resource type and Cowardin code predictions.

WOTUS-ML predicts a subset of sites with high accuracy (Figs. 2A, S9). In 27% of sites with scores below 0.07 or above 0.95, WOTUS-ML has 95% accuracy. If a developer used WOTUS-ML for these sites, ACE would agree for 95% of sites. For 52% of sites, where the score is below 0.17 or above 0.83, WOTUS-ML has 90% accuracy. In such sites, WOTUS-ML could save resources. The mean Section 404 permit costs $5,000 to $39,000 (*2*). While we are unaware of cost estimates for AJDs, if WOTUS-ML saved this amount in delay and uncertainty for each AJD where WOTUS-ML has 95% accuracy, it would save $209 million to $1.6 billion over our sample. If WOTUS-ML let developers avoid Section 404 permit costs for the 20% of PJDs where WOTUS-ML has 95% confidence the PJD is not regulated, it could save $150 million to $1.2 billion annually. These illustrative numbers demonstrate potentially high returns to efficient adjudication (*21*).

WOTUS-ML scores are well-calibrated to probabilities, because they provide an unbiased estimate of the probability of regulation (Fig. 2B). For example, when WOTUS-ML outputs a score between 0.3 and 0.4 for test set AJDs, 34% of those AJDs are regulated. Hence, we refer to WOTUS-ML scores as predicted regulatory probabilities. Unbiasedness supports WOTUS-ML's use as a decision support tool.

Required accuracy may vary by purpose (*13*). For example, developers might most value a signal with 95 percent accuracy of whether a site is regulated, to decide whether to provide a PJD or request an AJD. By contrast, ACE might value knowledge of non-extremal WOTUS-ML scores, which could help focus ACE resources on ambiguous cases.

Opening the Black Box

Feature importance analysis clarifies how a complex model like WOTUS-ML functions, though has limitations (*13*). We use permutation tests to elucidate which input layers WOTUS-ML relies on for predictions (*22*). We randomly permute groups of layers across samples, breaking the link between that feature and its label. A feature's permutation importance equals the difference between the accuracy with features intact and the accuracy with that feature permuted.

When permuted across all samples nationally, climate data most affect model accuracy (Fig. S10A), perhaps because precipitation and temperature predict streamflow and wetland prevalence. Stream and wetland vector data are second most important, which is intuitive since they are the regulated entities. Next are state, district, and rule measures, reflecting the dependence of jurisdiction on non-geophysical features. Elevation, soil and groundwater characteristics also matter. NAIP imagery is less important, perhaps because other layers are derived from remote sensing data.

We also perform within-state permutation tests by shuffling layers solely within state samples (Fig. S10B). This indicates which features help WOTUS-ML replicate AJDs within a state, which more resembles the work of ACE engineers. Within a state, stream and wetland data most account for model accuracy. ACE engineers consult these datasets in deciding AJDs (*21*).

WOTUS-ML Predictions of Regulatory Probabilities and Changes

Using WOTUS-ML, we predict regulatory probabilities for four million random prediction points across the US, plus a random sample of PJDs and traditional navigable waterways (*13, 23*). Because WOTUS-ML accuracy is independent of the probability that a point is an AJD, our model evaluation using the AJD test set provides useful information about the model's performance on the prediction points (SM Section B.4; Fig. S8).

Rapanos regulates 22% of points in the US; NWPR regulates 8% (Table 1). Areas where a land use model predicts that development will occur in the 2020s (*24*) have slightly higher levels of regulation than random prediction points. Because AJDs represent potential water resources, the mean AJD has 15 percentage points greater jurisdictional probability under *Rapanos* than the mean random point. Nearly half of PJDs are jurisdictional under *Rapanos*; 28% are jurisdictional under NWPR. Hence, developers may incorrectly assume sites are jurisdictional and request PJDs rather than AJDs; or may request PJDs for other reasons, including expediting processing under other federal regulations. These possibilities demonstrate potential value of WOTUS-ML as a decision support tool. Two-thirds of NHD streams are regulated under *Rapanos*, but only 46% under NWPR (Table 1B). NWPR thus deregulates 686,000 stream miles (Table S9 provides state level calculations), more than every river and stream in California, Florida, Illinois, New York, Ohio, Pennsylvania, and Texas, combined. Compared to *Rapanos*, NWPR deregulates 30 percent of stream and wetland areas in subwatersheds that provide drinking water for the average American (Table S10).

We also analyze differences across stream types (Table 1B). Our estimate that 100% of traditional navigable rivers are regulated under either rule gives confidence in WOTUS-ML since all rules regulate traditional navigable waterways (*23*). *Rapanos* regulates 55% of intermittent and ephemeral streams but NWPR regulates 30%. Thus, NWPR deregulates 45% of regulated intermittent and ephemeral streams, though NWPR only deregulates 21% of all national streams.

Rapanos regulates 52% of wetlands, while NWPR regulates 27% of all wetlands (Table 1C). Thus, NWPR removes jurisdiction for just under half of regulated wetlands, or 25% of all wetlands. This 25% statistic is far below the *Amicus* and media assertions that NWPR deregulates 51% of all wetlands (*9*). NWPR deregulates over a third of wetlands that are adjacent to or abutting a stream or river, and two-thirds of isolated wetlands. NWPR deregulates 35 million wetland acres (Table S9). This represents 15% of wetland area in the continental US at the time of European settlement,

or over a fourth of the wetlands that disappeared between the time of European settlement and today (*25*). This deregulated wetland area represents $12 to $23 billion in annual flood mitigation benefits, or $250 to $458 billion in present value flood mitigation benefits, discounted at 5%. The deregulated wetland area represents $249 billion to $381 billion of land value. Additionally, this large wetland area provides additional important species habitat protection, recreational opportunity, and other ecosystem values.

Rapanos jurisdiction reflects geophysical and political patterns. *Rapanos* regulates wetlands in the coastal South and mid-Atlantic and coastal streams and wetlands near the Pacific. *Rapanos* regulates less of the arid West, though some ephemeral streams. Parts of the Fall Line separating the Coastal Plain in the mid-Atlantic and South have discrete changes in jurisdiction. Major waterways are visible because their jurisdiction contrasts with lower jurisdiction for surrounding areas. ACE and state boundaries reveal differences (Fig. 3A). For example, the New England ACE district concludes that most AJDs are jurisdictional, while the St. Paul district concludes that few are.

Under NWPR, most predicted jurisdictional probabilities are below 20% (Fig. 3B, S6E,F). Major waterways remain regulated, though with narrower channels, potentially due to decreased jurisdiction of nearby wetlands. The least jurisdiction is in the arid West, where ephemeral streams are common.

Between *Rapanos* and NWPR, regulation decreases most around isolated wetlands in the mid-Atlantic and Gulf coast (Fig. 3C). WOTUS-ML shows no substantial areas of increased regulation under NWPR (*13*). The arid west shows limited areas of decreased regulation.

CWR has broadly similar jurisdiction patterns to *Rapanos* (Figs 3D, S6), with less coverage of ephemeral streams in the arid West. CWR, like *Rapanos*, has higher coverage than NWPR across the country (Fig. 3E).

Patterns within the 38 ACE districts are somewhat persistent, which supports our methodology's medium-run validity. In all years, St. Paul is in the bottom ten districts in share of AJDs jurisdictional and Norfolk is in the top ten. New England always has among the fewest AJDs of any district. One exception is Florida, where politics led to decentralization of the AJD process in 2020.

Case studies give confidence in the model's results, illuminate new patterns, and clarify what CWA rules regulate (Fig. 4). Regulation is heterogeneous within narrow and broad geographic areas. In case studies A through C and E, some wetlands are regulated but nearby forests and farms are not. In arid streams around the Arizona, Nevada, and Utah borders, *Rapanos* regulates some ephemeral streams and NWPR deregulates most (Fig. 4D). Each case study has points near 0.50, where regulation is uncertain; and points closer to 0 and 1, where regulation is more certain.

Discussion

To inform prominent debates about which water resources should be regulated, it is important to know which resources are regulated. WOTUS-ML provides such evidence for recent CWA rules. It reveals enormous sets of natural resources that have changed jurisdiction in the last decade. Stakeholders can use WOTUS-ML to support decision-making for individual water resources or aggregate policy design and evaluation, and potentially save large costs by reducing uncertainty

and delays. This represents one set of insights that ML algorithms can provide for more general regulatory implementation problems where regulators must repeatedly interpret and apply a law. Regulators have expressed mixed views on geophysical map tools (*2, 7*). Our analysis provides a basis for caution, in that WOTUS-ML has imperfect accuracy. It also provides a basis for cautious optimism, in the ways WOTUS-ML can provide insight on one of the most complex and controversial US environmental policies.

Besides clarifying which waters the CWA regulates, WOTUS-ML can support stakeholder decision-making. A developer can use WOTUS-ML to learn the estimated probability that the CWA regulates a site. ACE can use WOTUS-ML to provide input to determining jurisdiction. The White House and EPA can use WOTUS-ML to forecast impacts of rules changing which waters are jurisdictional. State environmental agencies can use WOTUS-ML to help regulate more than federal law requires (e.g., to support enforcement of a state-level wetland rule approximating *Rapanos* under a federal *Sackett* rule). Environmental or industry associations can use WOTUS-ML to provide statistics for court briefs.

We offer a template for applying existing algorithms to regulatory implementation problems, where agencies repeatedly interpret court rulings or laws. ACE engineers interpret language from *Rapanos*, CWR, or NWPR using data and visits to determine whether water resources are regulated. Research has used related algorithms for other types of policy problems—for example, where optimal decisions depend on future events like whether a defendant will jump bail (*26*), or predicting environmental conditions like whether an aquifer has arsenic (*27*). ML has predicted court decisions (*28*), which differs from modeling regulators' implementation of such decisions. We show how ML can clarify regulatory implementation of environmental law, with potential relevance to many regulations requiring practitioners to interpret and apply textual directives.

References and Notes

1. S. Sullivan, M. Patricio, M. C. Rains, A. D. Rodewald, W. W. Buzbee, and A. D. Rosemond, "Distorting science, putting water at risk" *Science* **369**, 766-768 (2020).
2. US Environmental Protection Agency and US Army Corps of Engineers, "Economic Analysis for the Final "Revised Definition of `Waters of the United States'" Rule," (2022); https://www.epa.gov/system/files/documents/2022-12/2022_WOTUS%20EA_FinalRule_508.pdf.
3. C. Taylor, H. Druckenmiller, "Wetlands, Flooding, and the Clean Water Act" *Am. Econ. Rev.* **112**, 1334-63 (2023).
4. S. Zellmer, "Treading water while congress ignores the nation's environment" *Notre Dame L. Rev.* **88**, 2323 (2023).
5. C. Davenport, "Trump Removes Pollution Controls on Streams and Wetlands" (NYTimes, 2020).
6. S. McFetridge, "Judge temporarily blocks clean water rule in 24 states" (AP News, 2023).
7. US Environmental Protection Agency and US Army Corps of Engineers, "Mapping and the Navigable Waters Protection Rule" (2020); https://19january2021snapshot.epa.gov/sites/static/files/2020-01/documents/nwpr_fact_sheet_-_mapping.pdf.
8. A. Wittenberg, K. Bogardus, "EPA claims 'no data' on impact of weakening water rule. But the numbers exist" (E&E News, 2018) ; https://www.science.org/content/article/epa-claims-no-data-impact-weakening-water-rule-numbers-exist.
9. K. A. Fesenmeyer, S J. Wagner, D. S. Leigh, and H. M. Neville, "Large portion of USA streams lose protection with new interpretation of Clean Water Act" *Freshw. Sci.* **40** (2021).
10. *Sackett v. EPA*. "Brief amicus curiae of the National Association of Wetland Managers, the Association of Floodplain Managers, the American Planning Association, the American Water Works Association, and the New England Interstate Water Pollution Control Commission" (2022); https://www.supremecourt.gov/DocketPDF/21/21-454/228256/20220617104921002_21-454%20Brief.pdf.
11. US Army Corps of Engineers, "Compliance with the Federal Endangered Species Act," (2023); https://www.spk.usace.army.mil/Missions/Regulatory/Permitting/Endangered-Species-Agency-Consultations/.
12. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 770-778 (2016).
13. Materials and methods are available as supplementary materials.
14. L. M. Cowardin, V. Carter, F. C. Golet, and E. T. LaRoe, "Classification of wetlands and deepwater habitats of the United States." (FWS/OBS-79/31, U.S. Fish and Wildlife Service, 1979).
15. U.S. Department of Agriculture, National Agriculture Imagery Program, (January 2023); https://datagateway.nrcs.usda.gov/GDGHome_DirectDownLoad.aspx.
16. U.S. Department of Agriculture, Gridded National Survey Geographic Database, (October 2022); https://nrcs.app.box.com/v/soils/folder/191790828371.
17. U.S. Environmental Protection Agency, National Hydrography Dataset Plus V2, (October 2022); https://www.epa.gov/waterdata/nhdplus-national-data.

18. S. A. R. Colvin, S. M. P. Sullivan, P. D. Shirey, R. W. Colvin, K. O. Winemiller, R. M. Hughes, K. D. Fausch, D. M. Infante, J. D. Olden, K. R. Bestgen, R. J. Danehy, L. Eby, "Headwater Streams and Wetlands are Critical for Sustaining Fish, Fisheries, and Ecosystem Services," *Fisheries* **44**: 73-91 (2019).
19. U.S. Fish & Wildlife Service, National Wetlands Inventory, (October 2022); https://www.fws.gov/program/national-wetlands-inventory/download-state-wetlands-data
20. U.S. Army Corps of Engineers, US Army Corps of Engineers Regulatory Boundary, (January 2023); https://geospatial-usace.opendata.arcgis.com/datasets/70805e1a8fd74e42b0a9585088d6d151_0/about.
21. US Army Corps of Engineers, "Corps of Engineers Wetlands Delineation Manual," (Y-87-1, 1987); https://usace.contentdm.oclc.org/digital/collection/p266001coll1/id/4530.
22. A. Altmann, L. Toloşi, O. Sander, T. Lengauer, "Permutation importance: a corrected feature importance measure" *Bioinformatics* **26**, 1340-1347 (2010).
23. USCS § 33 1804 (2023).
24. US Environmental Protection Agency, "Land-Use Scenarios: National-Scale Housing-Density Scenarios Consistent with Climate Change Storylines." (EPA/600/R-08/076F, 2009).
25. T. E. Dahl. "Status and Trends of Wetlands in the Coterminous United States 2004 to 2009" (U.S. Fish & Wildlife Service, 2011); https://www.fws.gov/wetlands/documents/Status-and-Trends-of-Wetlands-in-the-Conterminous-United-States-2004-to-2009.pdf.
26. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan, "Human Decisions and Machine Predictions" *Q. J. Econ.* **133**, 237-293 (2018).
27. J. Podgorski and M. Berg, "Global threat of arsenic in groundwater" *Science* **368,** 845-850 (2020).
28. D. M. Katz, M. J. Bommarito II, Josh Blackman, "A general approach for predicting the behavior of the Supreme Court of the United States," *PLoS One* (2017).
29. Greenhill, S., H. Druckenmiller, S. Wang, D. A. Keiser, M. Girotto, J. K. Moore, N. Yamaguchi, A. Todeschini, and J. S. Shapiro. (2023). Code for Machine Learning Predicts Which Rivers, Streams, and Wetlands the Clean Water Act Regulates [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.10108709.
30. U.S. Geological Survey, 3-Dimensional Elevation Program, (November 2022); https://www.usgs.gov/3d-elevation-program.
31. U.S. Environmental Protection Agency, U.S. EPA Ecoregions, Level IV Ecoregion, (January 2023); https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states.
32. Prism Climate Group, Parameter-elevation Regressions on Independent Slopes Model, 30-year Normals, (January 2023); https://prism.oregonstate.edu/normals/.
33. U.S. Geological Survey, National Land Cover Database, (December 2022); https://www.usgs.gov/centers/eros/science/national-land-cover-database#overview.
34. U.S. Census Bureau, TIGER/Line Shapefiles, (December 2022); https://www2.census.gov/geo/tiger/TIGER2022/.
35. Greenhill, S., H. Druckenmiller, S. Wang, D. A. Keiser, M. Girotto, J. K. Moore, N. Yamaguchi, A. Todeschini, and J. S. Shapiro. (2023). Training data for Machine Learning Predicts Which Rivers, Streams, and Wetlands the Clean Water Act Regulates [Dataset]. Dryad. https://doi.org/10.5061/dryad.m63xsj47s.

36. Greenhill, S., H. Druckenmiller, S. Wang, D. A. Keiser, M. Girotto, J. K. Moore, N. Yamaguchi, A. Todeschini, and J. S. Shapiro. (2023). Prediction data for Machine Learning Predicts Which Rivers, Streams, and Wetlands the Clean Water Act Regulates [Dataset]. Dryad. https://doi.org/10.5061/dryad.z34tmpgm7.
37. E. Rolf, J. Proctor, T. Carleton, I. Bolliger, V. Shankar, M. Ishihara, B. Recht, and S. Hsiang. 2021. "A generalizable and accessible approach to machine learning with global satellite imagery." *Nat. Commun.* **12** 4392 (2021).
38. US Environmental Protection Agency, "Clean Water Act Approved Jurisdictional Determinations" (2022); https://watersgeo.epa.gov/cwa/CWA-JDs/.
39. U.S. Environmental Protection Agency and Department of the Army, "Economic Analysis for the Proposed "Revised Definition of 'Waters of the United States' Rule" (2021); https://www.epa.gov/system/files/documents/2021-11/revised-definition-of-wotus_nprm_economic-analysis.pdf.
40. USDA, "NAIP: Options and Challenges," (2007); https://www.fsa.usda.gov/Internet/FSA_File/2007_esri_uc_naip.pdf.
41. C. Nolte, "High-resolution land value maps reveal underestimation of conservation costs in the United States" *PNAS*, **117** 29577-83.
42. US Army Corps of Engineers, "Minimum Standards for Acceptance of Preliminary Wetlands Determinations," (2001); https://www.spk.usace.army.mil/Portals/12/documents/regulatory/pdf/Minimum%20Standards.pdf.
43. US Army Corps of Engineers, "Jurisdictional Determinations and Delineating Waters of the United States, including Wetlands," (2013); https://www.spk.usace.army.mil/Portals/12/documents/regulatory/Reg_workshop/RPW_JDs_and_Delineation.pdf.
44. Federal Register, "Revised Definition of "Waters of the United States"." (EPA-HQ-OW-2021-0602, 2023); https://www.federalregister.gov/documents/2023/01/18/2022-28595/revised-definition-of-waters-of-the-united-states.
45. *Sackett v. Environmental Protection Agency*, 598 U.S. ____ (2023) https://www.supremecourt.gov/opinions/22pdf/21-454_4g15.pdf.
46. US Environmental Protection Agency and US Army Corps of Engineers, "Amendments to the "Revised Definition of 'Waters of the United States" ", (EPA-6560-50-P, 2023); https://www.epa.gov/system/files/documents/2023-08/Pre-publication%20Version%20of%20the%20Final%20Rule%20-%20Amendments%20to%20the%20Revised%20Definition%20of%20Waters%20of%20the%20United%20States.pdf.

Acknowledgements.

Participants at AERE, the California Water Boards, the Energy Institute at Haas, EPA, Global Policy Lab, Stanford, and TWEEDS seminars, Elinor Benami, Marshall Burke, Joel Corona, Chris Costello, Todd Doley, Julie Hewitt, Sol Hsiang, Nitin Kohli, Rose Kwok, Matt Massey, Al McGartland, Sendhil Mullainathan, Kristine Swann, and Charles Taylor provided useful comments and discussions. We thank Brant Walker for outstanding research assistance.

Funding

This material is based upon work supported by the following:

Ciriacy-Wantrup Postdoctoral Fellowship (SW)

The Giannini Foundation (JSS)

Google Cloud Research Credits Program (HD, DAK, JSS)

NIH Grant 1R01AG079914-01 (DAK, JSS)

Author contributions.

Conceptualization: JSS

Data curation: SG, HD, SW, DAK, JSS

Formal analysis: SG, HD, SW, DAK, JKM, NY, AT, JSS

Funding acquisition: HD, DAK, MG, AT, JSS

Investigation: SG, HD, SW, DAK, MG, JKM, NY, AT, JSS

Methodology: SG, HD, SW

Software: SG, HD, SW, DAK, JSS

Visualization: SG, HD, SW, JSS

Writing – original draft: SG, HD, SW, DAK, JSS

Writing – review and editing: SG, HD, SW, DAK, JSS

Competing interests. DAK serves on the Scientific Advisory Board of the EPA and serves as an expert witness on water pollution litigation. During the research JSS served as an advisor on trade and environment to the European Commission's Directorate General of Trade. DAK and JSS wrote a public report on EPA's analyses of the CWR and the NWPR as part of the External Environmental Economics Advisory Committee, funded by the Sloan Foundation. All authors of this paper are inventors on a patent pending with serial number 63/513,464, submitted by UC Berkeley, which covers WOTUS-ML. The software has a Creative Commons Attribution Non Commercial No Derivatives 4.0 International license, which freely allows use for research and other non-commercial purposes. Potential commercial users should contact the Office of Technology Licensing at UC Berkeley.

Data and Materials Availability: All code is available on Zenodo (*29*). All data are publicly available online (*15-17, 19-20, 30-34*). All data used to train the models are on Dryad (*35*), as is a subset of the prediction data, as well as model weights and all data required to produce the results presented here (*36*). The model is freely available for research and non-commercial purposes (*29*, *35-36*) but its commercial use is limited (patent pending).

Supplementary Materials:
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S10
References 37-46

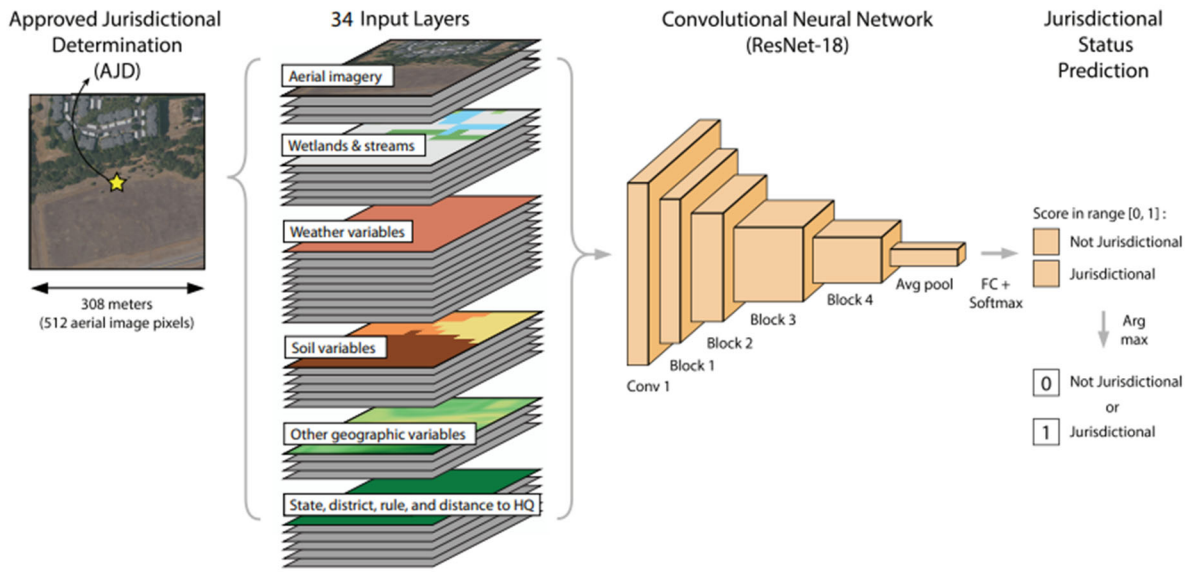


Fig. 1: WOTUS-ML Model Architecture Uses a ResNet-18 with 34 Input Layers. The training data are images centered around an AJD. The model takes 34 input layers, described in Fig. S5. The ResNet-18 architecture begins with a convolutional block (Conv 1), followed by four residual blocks (Blocks 1 through 4), an average pooling operation, and finally a fully connected layer. The outputs of the fully connected layer are passed through a softmax function, producing a score in [0, 1]. We predict that a site is jurisdictional if the score exceeds 0.5.

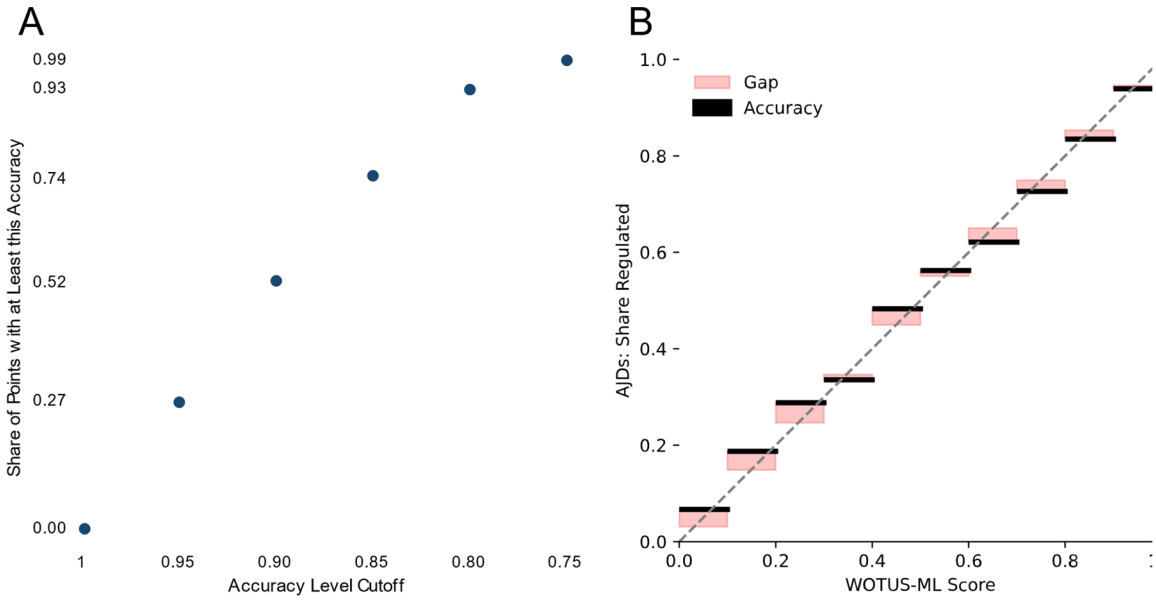


Fig. 2: WOTUS-ML Scores Allow Unbiased Estimates of Regulatory Probability. (**A**) WOTUS-ML Estimates Regulation with High Accuracy for Many Sites. This figure finds threshold WOTUS-ML scores such that the average point beyond the threshold has at least a given accuracy in the AJD test set (0.95, or 0.90, etc.). The vertical axis plots the share of points with WOTUS-ML scores beyond this cutoff. This indicates, for example, for what share of points WOTUS-ML can predict AJD outcomes with 95% accuracy, for what share with 90% accuracy, etc. Fig. S9 provides details of calculations underlying (**A**). (**B**) WOTUS-ML Scores Reflect the Probability of Regulation. AJD test set is split into ten equal-width bins containing AJDs with WOTUS-ML scores of 0.0 to 0.1, 0.1 to 0.2, etc. The black lines show the average accuracy of AJDs in each bin, i.e., the share of AJDs that are jurisdictional. The model's score is interpreted as the model's confidence that a given AJD is determined to be WOTUS. The red bar shows the gap between the model's average confidence and the accuracy of AJDs in each bin. The dashed 45-degree line is the ideal accuracy for each confidence level. If confidence and accuracy are equal, the model is calibrated and we can interpret the confidence score as a probability. If the red bar is below the diagonal, the model is too confident in its predictions, and vice-versa.

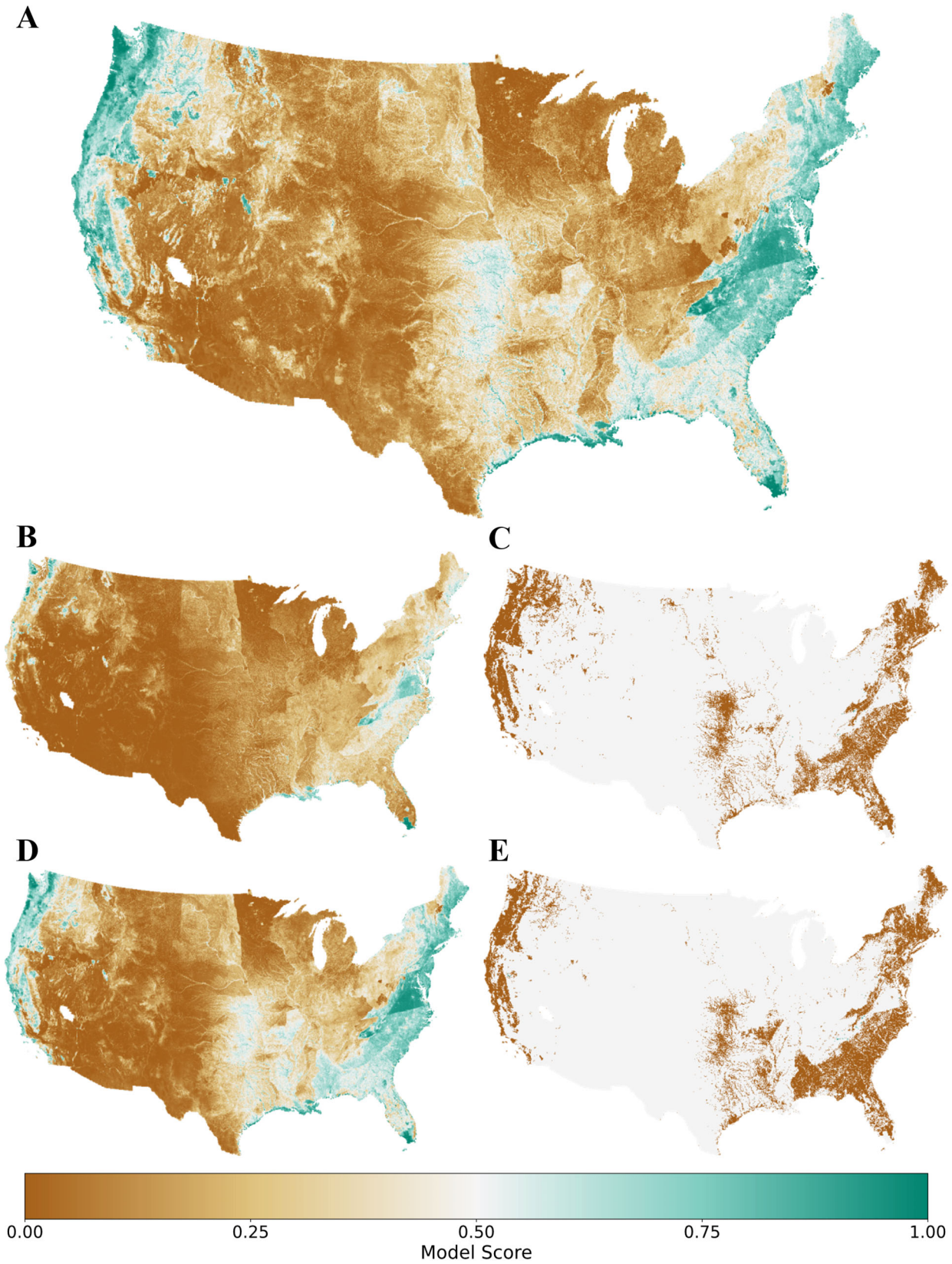


Fig. 3: Estimated probability of CWA regulation for four million prediction points across the USA. (**A**) Estimated probability of CWA regulation (WOTUS-ML score) under *Rapanos*. (**B**)

Estimated regulatory probability under NWPR. (**C**) Estimated regulation changes from *Rapanos* to NWPR. (**D**) Estimated regulatory probability under CWR. (**E**) Estimated regulation changes from CWR to NWPR. A 'regulation change' describes when the WOTUS-ML binary classification score (>50%, <50%) changed status. Brown represents deregulation, green represents new regulation. Map creates 247x576 grid and displays mean model score in each bin (~28 prediction points per bin).

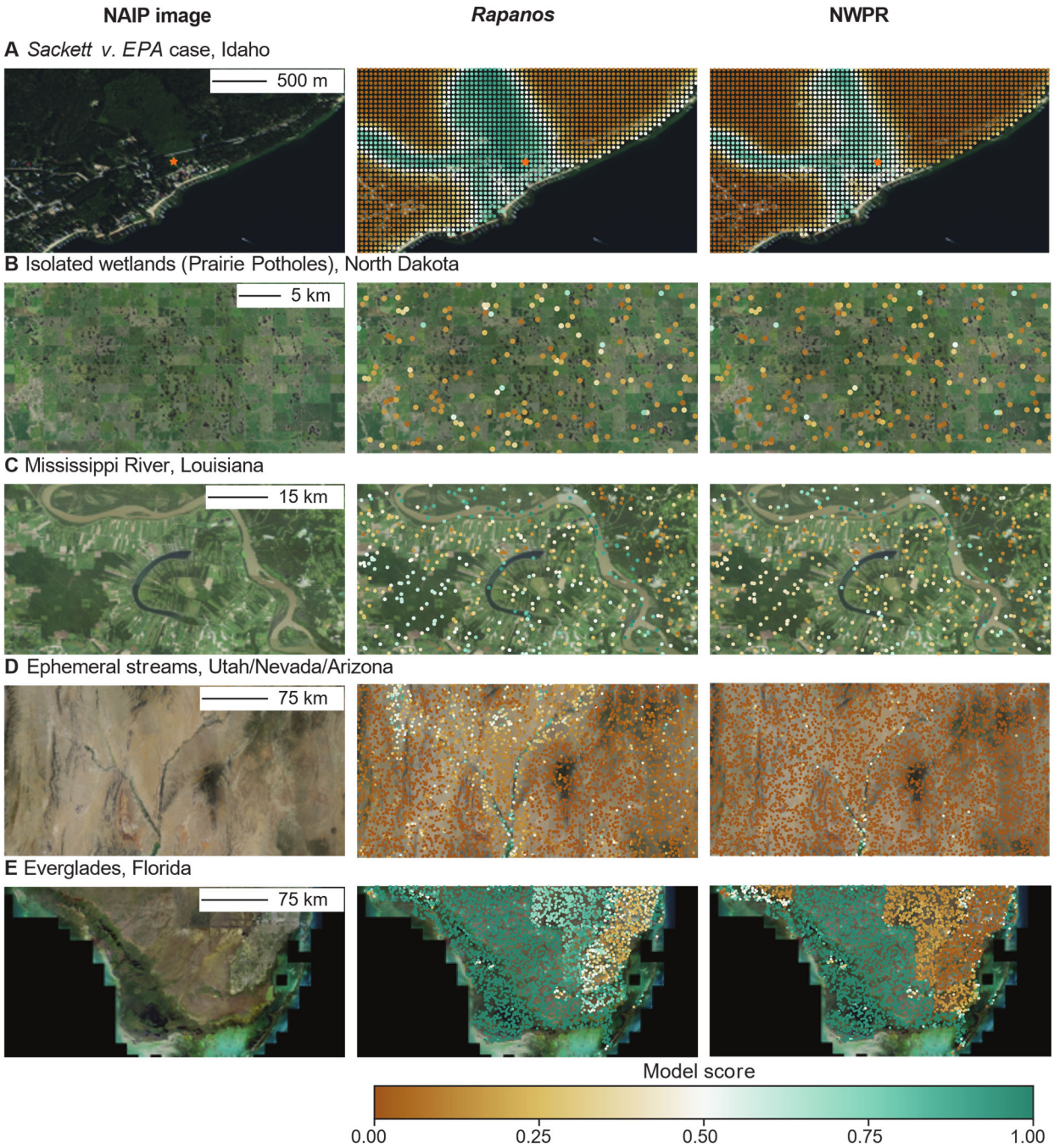


Fig. 4. Case Studies of CWA Regulation Reveal Local Heterogeneity and Differences Across Rules. (**A**) A densely sampled area around the Sackett property where jurisdiction is subject of *Sackett v. EPA*, near Priest Lake, Idaho. The property is marked with an orange star. Points on this property have mean model score of about 0.5, consistent with the ambiguity that produced litigation. Areas within and near wetlands have higher scores, though scores around the edge of the large centrally located wetland decrease under NWPR. Average model scores under *Rapanos* and NWPR are 0.34 and 0.22, respectively. (**B**) Prairie potholes (isolated wetlands), Benson County, North Dakota. Points near prairie potholes are more likely to be regulated under *Rapanos* and deregulated under NWPR. ACE categorizes 80% of AJDs in Benson County, ND, as isolated wetlands. Average model scores under *Rapanos* and NWPR are 0.22 and 0.17, respectively. (**C**) Farmland along the Mississippi River, Baton Rouge, Louisiana. WOTUS-ML predicts most points

around the Mississippi are regulated under *Rapanos* and NWPR. Farmland, which the CWA explicitly ignores, receives lower scores under both rules. Average model scores under *Rapanos* and NWPR are 0.45 and 0.35, respectively. (**D**) Ephemeral streams north of Lake Mead, near Nevada, Utah, and Arizona borders. WOTUS-ML predicts ephemeral streams in the north-central part of the image lose regulation from *Rapanos* to NWPR. Lake Mead remains regulated under both rules. The southwestern part of the image includes Solar Energy Zones near Dry Lake, Nevada, where renewable energy development is occurring and requires AJDs. ACE categorizes 69% of AJDs in Utah and Nevada under NWPR as ephemeral streams. Average model scores under *Rapanos* and NWPR are 0.09 and 0.01, respectively. (**E**) Southern Florida, including the Everglades and Miami. WOTUS-ML predicts *Rapanos* and NWPR regulate Everglades National Park and most other protected wildlife areas. Wetlands and developed areas along the northwestern and eastern image edges have much lower model scores under NWPR. Average model scores under *Rapanos* and NWPR are 0.85 and 0.64, respectively. We randomly choose foreground/background ordering of points in all panels.

Table 1: What Does the Clean Water Act Regulate? (**B**) and (**C**) describe subsets of the four million prediction points. Table shows the share of points the CWA regulates, measured as the share with WOTUS-ML score above 50 percent. For navigable waters, we evaluate jurisdiction based on the mean WOTUS-ML score among points within each named river, since CWA regulates each water resource, and report the share of named rivers predicted as jurisdictional. "WOTUS-ML resource type" refer to points where WOTUS-ML predicts that the listed resource type is the most likely under either *Rapanos* or NWPR, according to the classification listed in Table S2.

	Rapanos	**NWPR**
	(1)	(2)
A General groups of points		
All 4 million points	0.22	0.08
Urban growth areas (ICLUS)	0.30	0.09
AJD test set	0.37	0.16
Preliminary jurisdictional determinations (PJDs)	0.48	0.28
B Rivers and streams		
All (NHD all)	0.67	0.46
Traditional navigable waters (NHD+law)	1.00	1.00
Perennial (NHD 46006)	0.83	0.67
Intermittent or ephemeral (NHD 46003, 46007)	0.55	0.30
C Wetlands		
All (NWI palustrine)	0.52	0.27
Emergent (NWI)	0.47	0.20
Forested (NWI)	0.59	0.32
Adjacent or abutting (WOTUS-ML resource type)	0.88	0.57
Isolated (WOTUS-ML resource type)	0.39	0.14

Table 1: What Does the Clean Water Act Regulate? (**B**) and (**C**) describe subsets of the four million prediction points. Table shows the share of points the CWA regulates, measured as the share with WOTUS-ML score above 50 percent. For navigable waters, we evaluate jurisdiction based on the mean WOTUS-ML score among points within each named river, since CWA regulates each water resource, and report the share of named rivers predicted as jurisdictional. "WOTUS-ML resource type" refers to points where WOTUS-ML predicts that the listed resource type is the most likely under either *Rapanos* or NWPR, according to the classification listed in Table S2.

Supplementary Materials for

Machine Learning Predicts Which Rivers, Streams, and Wetlands the Clean Water Act Regulates

Simon Greenhill*, Hannah Druckenmiller†, Sherrie Wang†, David A. Keiser, Manuela Girotto, Jason K. Moore, Nobuhiro Yamaguchi, Alberto Todeschini, Joseph S. Shapiro*

*Corresponding authors. Email: sgreenhill@berkeley.edu, joseph.shapiro@berkeley.edu

† These authors contributed equally to this work

This is the author's version of the work. It is posted here by permission of the AAAS for personal use, not for redistribution. The definitive version was published in *Science* on January 25, 2024 383(6681), DOI: https://www.science.org/doi/10.1126/science.adi3794.

The PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S10
References 37-46

A. Materials and Methods

A.1 Model Choice, Design, and Assessment

We use ResNet-18 for a few reasons. Given the sample size for training data, additional blocks in the model may not improve performance. We observed some signs of overfitting in training data for ResNet-18, suggesting that the model is sufficiently complex and implying that a more complex model could exacerbate overfitting. We experimented with multi-task learning across binary regulation and the nine hydrological and the nine legal types of waters (SM section A.4) but found it did not improve accuracy in the validation data.

We also considered alternatives to ResNet but concluded they were unlikely to improve model performance. To improve model accuracy, meta-learning models may require substantially more data than we have. A foundation model would need to be pre-trained on large amounts of pre-existing data from a similar problem. In our setting, such data are not readily available. We experimented with a form of convolutional kernel ridge regression (*37*), a gradient boosting algorithm, and separate ResNet models for each rule, but found they did not have better validation accuracy than the model the main text uses.

Permutation importance tests, discussed in the main text, have limitations. A feature that is highly correlated with others may appear unimportant when removed, since the model can rely on the unpermuted features to arrive at the same prediction. We try to mitigate this by permuting features in groups, such as all PRISM or all Gridded National Soil Survey Geographic Database (gNATSGO) layers. Permutation can also produce unrealistic model inputs that generate unpredictable model behavior. Combined with the case studies and maps, however, these tests can clarify WOTUS-ML's functioning.

A.2 Train-Test Split

To prevent spatial leakage, we group data so AJDs in the same project or with geographically overlapping footprints are in the same fold. AJDs in the same project are evaluated by the same ACE engineer at the same time, so have correlated outcomes. Grouping by aerial imagery footprint ensures the test set does not partially include images and other input layers used to train the model.

We seek to assign an 80/10/10 split between train, test, and validation, stratifying the randomization by district-rule to ensure the same split in every district and every rule. Because groups span rules and districts, however, and we have limited data for some district-rules, this ideal split is not feasible in our setting.

We therefore adopt a splitting procedure that ensures we have sufficient data for testing and validation in all district-rules while getting as close as possible to the 80/10/10 target. We assign all the data from district-rules with 50 or fewer observations to test. District-rules with 100 or fewer observations are split 80/20 between test and validation. For district-rules with 500 or fewer observations, we assign 50 observations to test, then split the remaining observations 80-20 between training and validation. After the small district-rules are assigned, we split the remaining groups 80/10/10 between training, validation, and testing. We repeat the randomization until every district-rule with over 500 observations has at least 100 training observations, and at least 5% of the observations in the validation and testing sets. The entire randomization is stratified at the district level.[1]

The maximum number of points in a single image footprint group is 244. The average is 3.67, with a standard deviation of 8.53.

A.3 Input data

WOTUS-ML takes as inputs 34 layers from 11 distinct datasets (Table S3). These include imagery from the National Agriculture Imagery Program (NAIP); geographic data on the locations and characteristics of wetlands and

[1] Groups of observations occasionally overlap districts. A total of 61 out of the 14,000 image footprints groups span more than one district, accounting for about 6% of our observations. In cases where a group is assigned to multiple splits, we re-assign it to a single split, with train taking precedence over test and validation and test taking precedence over validation.

streams from the National Wetlands Inventory (NWI) and National Hydrography Dataset (NHD); soil characteristics from the Gridded National Soil Survey Geographic Database; information on local climate from the PRISM Climate Group's 30-year Normals; land cover from the National Land Cover Database (NLCD); elevation from USGS's 3DEP model; ecoregion classification (level IV) from the US EPA; distance to U.S. Army Corps of Engineers (ACE) district headquarters; and encodings for states, ACE regulatory districts, and WOTUS rules. Several input layers are available only in the contiguous US (CONUS), thus restricting our analysis to this region.

While state and ACE district layers represent political rather than geophysical inputs, they have an important influence on what the CWA regulates. For example, if the St. Paul ACE district concludes that no wetlands are regulated, and the Wilmington ACE district concludes that all wetlands are regulated, the resulting AJDs have legally binding force and describe what the CWA regulates. Similarly, we include a layer for the distance of a site to the closest ACE headquarters because the probability of in-person visits decreases with this distance, and sites with in-person visits have higher probability of being regulated. One potential explanation is that field visits are more likely for sites nearer to an ACE office, and field visits may discover water resources (e.g., swales under tree cover) that a desk analysis is less likely to observe. Although CWA represents federal regulation, its management through ACE district offices implies that its implementation is to some extent sub-national.

We rasterize all vector layers into 512x512 pixel tiles, where each pixel is 0.6 to 1.1 meters, to match the resolution of the NAIP imagery. For the input layers that are not natively in a raster or vector format (i.e., distance to ACE district headquarters and encodings for ACE districts, states, and WOTUS rules), we fill a 512x512 pixel raster with one value to make these layers conform to our convolutional neural network (CNN) architecture. For categorical data, such as the NLCD land cover classes or ACE districts, we use ordinal encoding. We assign numeric values such that similar classes or adjacent geographic areas are close together in numeric space (e.g., for states, we do not use federal information processing standards (FIPS) codes, but instead assign sequential numbers to neighboring states). After experimenting with one-hot encoding and observing that it had no impact on accuracy in the validation set, we elect to use ordinal encoding to avoid inputting an excessive number of layers into the CNN given the limited amount of training data (~150,000 observations) and computational cost of training a model with more input layers. We use one-hot encoding for the rules, so each rule has its own input layer.

For the training data layers that contain information on built-up land (NAIP and NLCD), we use images from the most recent survey *prior* to each AJD to ensure that the model's prediction is based on pre-AJD conditions rather than post-AJD development activity. We use the most recent survey available for the input layers corresponding to the four million prediction points across the country for which we predict regulatory probabilities. We exclude all AJD and prediction points that are missing input data, such as NAIP imagery or state information. All these excluded AJDs are in the oceans, Great Lakes, or military bases (e.g., Area 51).

We highlight a few other notes on NHD and NWI. NHD describes 3.15 million stream miles in the CONUS. We define a sample point as in NHD or NWI if it is within 10 meters of those resources. Documentation for NHD's high resolution version states that it distinguishes ephemeral from intermittent streams. We do not use the high-resolution NHD files given its incompleteness and measurement error. NHD high resolution's distinction between ephemeral and intermittent streams appears in only a few western states; elsewhere, NHD classifies all streams as intermittent. Additionally, we found substantial differences between NHD's classification of streams as intermittent or ephemeral and AJDs' classification of the same streams as intermittent or ephemeral.

NWI and NHD seek to identify water types of locations across the US. While this has similar outputs as the model, NWI and NHD alone provide insufficient information to determine whether a site is jurisdictional (*7*). For example, NWI identifies wetlands, but does not identify whether a wetland is abutting navigable waters or isolated, and thus does not identify if the wetlands are jurisdictional. Similarly, while the high resolution NHD identifies certain water resources as ephemeral (which would not be jurisdictional under NWPR), ACE may determine that some of those water resources are in fact intermittent and thus jurisdictional under NWPR. Additionally, NWI and NHD have imperfect measures, particularly for certain water types.

These considerations lead us to use NWI and NHD as model inputs, since they provide useful but imperfect and incomplete information about regulatory probabilities. They may also affect the model's accuracy. For example, if some future version of NHD and NWI had no errors, then those error-free inputs would provide more useful information for a model, and thus a model trained on perfect versions of NHD and NWI could have greater accuracy than WOTUS-ML.

At the same time, these considerations do not bear on our conclusions about model accuracy, reliability, or bias. Our estimates of accuracy and bias are conditional on the information available in the model inputs. Although using an NWI or NHD point alone may provide a biased estimate of regulatory probabilities, by using NWI, NHD, and many other inputs to flexibly predict legally-binding AJDs, the model obtains unbiased estimates of regulatory probabilities (Fig 2A).

Because PJDs are not certain to be jurisdictional, we use the model to assess their probability of regulation, but do not train the model on PJDs. In other words, our data on PJDs do not affect model performance since the model is trained only on AJDs. After training WOTUS-ML, we compare predicted regulatory probabilities for PJDs versus other types of points. Sites with the highest probability of regulation may be more likely to request PJDs than AJDs, so AJDs may over-represent cases with ambiguous jurisdiction. We obtained details on PJDs from a Freedom of Information Act request filed with the EPA.

A.4 Other data

A few additional details on AJDs are useful to highlight. The ground-truth data include AJDs between August 2015 and May 2022, obtained from the EPA (*38*, *39*). Of these, 81,316 were made under *Rapanos*, 61,582 under NWPR, and 7,755 under CWR. *Rapanos* applied through 2018; CWR applied through late 2020, though only in about half of states due to litigation; NWPR applied through mid-2021; and *Rapanos* applied thereafter. For a small subset of AJDs, we can manually obtain a PDF document for each AJD with modest additional detail on other variables from ACE's ORM2 database. A further subset of these PDF files includes idiosyncratic information from ACE engineers on what information the engineer gathered.

The main text mentions that ACE engineers visit only some sites. An example of information gleaned from a site visit comes from a sand and gravel plant expansion in Northeast Ohio. An ACE engineer conducting an AJD field visit discovered "small drainage swales" potentially under tree cover connecting to jurisdictional waters and concluded the site was jurisdictional. This example has project identifier LRB-2015-01193.

An AJD reports the water resource's coordinates, whether the resource is jurisdictional, the rule under which the AJD was decided, the Cowardin code, the resource type, and a unique project identification code (e.g., different wetlands on a single housing development may have separate AJDs but the same project identifier), which we use to separate AJDs into test and training sets (SM section A.2). AJDs are valid for five years. About half of AJDs also incorporate site visits.

ACE requests AJDs to list centroids of water resources. We scrutinized high-resolution satellite imagery around 20 randomly selected AJDs on lakes and found that coordinates for about half of these water resources listed the centroids, while coordinates for the other half listed other locations around the water resource. Due to this focus and the model architecture used, we interpret WOTUS-ML as classifying individual points as regulated or not. Although WOTUS-ML can output different regulatory probabilities inside versus outside water resources, it is better suited at identifying jurisdiction for a water resource overall than delineating wetland boundaries (e.g., ACE sometimes conducts wetland delineation, which seeks to identify the exact boundaries of a wetland feature).

During the AJD process, ACE engineers also assign the water resource a Cowardin class and resource type. Cowardin classes are a hierarchical system for classifying wetland and water resources that capture the aquatic system, type of substrate, and water regime of the habitat (*14*). Resource types correspond to legal groupings of aquatic habitats, such as adjacent wetlands (those abutting a stream/river) or isolated wetlands (those that do not have a surface water connection) specified in each rule. Broad categories of Cowardin classes and resource types relate to different interpretations of WOTUS. For example, *Rapanos* regulates isolated wetlands and NWPR does not.

The main text refers to areas where a land use model predicts development. We identify these areas using ICLUS, which provides a raster identifying areas with increasing urban development in the period 2020-2030 (*24*), which may be more likely to have future AJDs. Using ICLUS, we categorize each point into one of three levels of development. We define a point as undeveloped if it has any of the following ICLUS categories: natural water; reservoirs; canals; wetlands; recreation; conservation; timber; grazing; pasture; cropland; mining; or barren land. We define a point as semi-developed if it has any of the following ICLUS categories: exurban, low; suburban; urban, low; parks; or golf courses. We define a point as developed if it has any of the following ICLUS categories: exurban, high; urban high; commercial; industrial; institutional; or transportation. Finally, we define areas with increasing urban development as

those that ICLUS predicts move from undeveloped to semi-developed, semi-developed to developed, or undeveloped to developed, using the preceding definitions.

Several parts of the paper and several datasets discuss isolated wetlands and ephemeral streams. These water features have different names and sub-types in different regions of the country, including prairie potholes, vernal pools, Carolina bays, playa lakes, alpine ponds, desert depressions, swale ponds, and others.

WOTUS-ML observes differences across ACE districts and state boundaries. NAIP has some differences in collection methods across states. For example, the flying season for imagery collection can differ modestly across states, and photographic methods can also differ (*40).* The WOTUS-ML inputs include a state boundary layer, which helps account for differences in NAIP methodology across states. Differences in regulatory probabilities across ACE districts are salient in the unprocessed AJD data.

The main text reports estimates of the difference in flood mitigation benefits and land values between *Rapanos* and NWPR. The range we report reflects homogeneous versus heterogeneous estimates of flood mitigation benefits and land values from other sources. Our homogeneous estimate uses the finding that the mean US wetland provides $1,840 to $1,900 per hectare in annual flood mitigation benefits (*3*). Aggregated, this represents $25 to $26 billion in annual US flood mitigation benefits, or $500 to $520 billion in present value flood mitigation benefits, discounted at 5%. Heterogeneous estimates of flood mitigation values by population density (*3*), applied to our 4 million prediction points, imply that the mean wetland deregulated between *Rapanos* and NWPR has annual flood mitigation benefits of $1030/hectare. Aggregated, this heterogeneous estimate implies that these deregulated points represent $12 billion in annual flood mitigation benefits, or $250 billion in present value. The mean value of US wetlands is $12,700/acre (*41*). Using heterogeneous estimates of land values for each deregulated point (*41*), the mean wetland deregulated between *Rapanos* and NWPR has fair market value of $8,313/acre. Thus, the heterogeneous estimates imply the national land value of wetlands deregulated between *Rapanos* and NWPR is $249 billion. The heterogeneous estimates of flood mitigation benefits and land values for deregulated wetlands are smaller than the homogeneous estimates for the mean US wetland because deregulated points are in less populated areas than the mean US point.

A.5 Prediction points

For prediction, we collect input data on three distinct samples of points in the CONUS. First, we collect 4.1 million points at random locations across the entire CONUS. To do this, we first divide the country into approximately 80,000 0.1 degree by 0.1 degree grid cells. We then randomly sample 50 points in each cell. We choose sample sizes to support detailed maps and case studies. In practice, input data are not available for about 4% of the points we sample, so we generate predictions for a total of 3.94 million points. The large number of prediction points allows us to produce high-resolution maps, as in Fig. 3, as well as detailed case studies, as in Fig. 4. We additionally use the intersections of these points with areas of particular interest to produce the values in Table 1. In reporting aggregate statistics, the share of the four million prediction points in streams and wetlands that are jurisdictional approximates the share of wetland acres and stream miles that are jurisdictional.

The second CONUS sample we collect covers about 3,000 points in a dense sample of approximately 2.7 km sq centered on the Sackett property near the shore of Priest Lake, Idaho. These points enable us to analyze in detail an area of particular interest as it lies as the center of the 2023 Supreme Court *Sackett* case. In addition, these points clarify how spatially precise our predictions are.

Finally, we collect approximately 6,200 points along traditional navigable waters legally identified in Title 33 of the US code (*23*) and link them to corresponding flowline names in NHD. In line with EPA and ACE practice, we use "traditional navigable waters" to denote major waterways primarily used for navigation, sometimes called, "navigable-in-fact waters." We select these points because they are regulated under every WOTUS rule, and thus are less likely to appear in our AJD training set. These points allow us to test our model's out of sample performance on locations virtually certain to be jurisdictional.

B Supplementary Text

B.1 Model Accuracy and Potential Uses

To evaluate model performance, one metric is whether WOTUS-ML predictions are useful enough for different possible applications. The main text provides a few statistics that help assess the model's potential usefulness, which we summarize here. For all AJDs, WOTUS-ML test set accuracy is 79%, reflecting learning of 14 percentage points. The corresponding Area Under the Receiver Operating Characteristic Curve (AUC) is 0.85. The ROC curve is plotted in Figure S7. AUC varies little by rule at 0.82, 0.86, and 0.87 for *Rapanos*, CWR, and NWPR respectively. Finally, WOTUS-ML scores provide an unbiased estimate of the probability that a site is regulated. Correspondingly, grid search for the optimal classification threshold returned 0.51. Relative to a threshold of 0.50, using a threshold of 0.51 would increase accuracy on the validation set by 0.013 percentage points (a hundredth of a percentage point); thus, we use a threshold of 0.50.

For a developer, learning that a site's WOTUS-ML score is below 0.07 or above 0.95 provides a strong signal, with over 95% accuracy, whether a site is jurisdictional. Many developers may find this information useful for investment and permitting decisions. For example, observing the WOTUS-ML prediction that a site is 95% likely to be regulated may lead a developer to request a PJD rather than an AJD, which can increase permitting efficiency. Sites with WOTUS-ML scores closer to 0.5 provide weaker signals for an individual site, since accuracy is lower for those score ranges, but may still provide useful inputs for some developer investment and permitting decisions. For ACE, having WOTUS-ML scores available could help to prioritize between sites where WOTUS-ML indicates jurisdiction is more certain, so the return to additional ACE effort is lower (or a field visit may be less needed), versus sites where WOTUS-ML scores are closer to 0.5, and where the return to ACE effort may be greater. WOTUS-ML could flag cases where an ACE decision differs substantially from the WOTUS-ML prediction, and which ACE could use to select cases for review. Additionally, WOTUS-ML scores could allow benchmarking or provide a training tool for new ACE engineers or provide a way of comparing practices across ACE districts. Under current practice, algorithms like WOTUS-ML alone cannot provide a legally-binding jurisdictional determination, but could provide a useful input to ACE's decision process.

For EPA, environmental, or industry associations, WOTUS-ML could help estimate the share of waters that are regulated under different rules, geographic areas, or water types. The unbiasedness described in Figure 2A is useful here, since it shows that while WOTUS-ML has imperfect accuracy for individual sites, these errors average out in large samples.

Another metric to evaluate model performance is whether WOTUS-ML exceeds performance of available alternative tools. We are unaware, however, of any available alternative tools. We believe no other analyst in government, academia, or the private sector has constructed a model that predicts AJDs and has then evaluated the accuracy of this model. In this sense, a possible use of WOTUS-ML is to highlight the importance of this prediction problem and to provide a set of accuracy statistics upon which other, potential models may attempt to improve.

We briefly discuss how WOTUS-ML differs between sites with and without a field visit. Under COVID-19 lockdowns, some sites had a "visit, conducted virtually using remote tools," primarily under NWPR, which we count as not having a site visit. The main text notes that test set accuracy is 82% for AJDs without a field visit and 74% for AJDs with a field visit. The share regulated is 26% for AJDs without a field visit and 46% for AJDs with a field visit. Thus, learning is only 8 percentage points for sites without a field visit, but 20 percentage points for sites with a field visit. At the same time, the low regulatory probability for AJDs without a field visit overall (26%) gives less scope for model learning.

Regulation and model learning differ for AJDs with a field visit for several reasons. On a field visit, ACE engineers can discover evidence of regulated waters that computer analysis misses, such as drainage swales under forest cover, plant species indicative of wetlands, or the location of the ordinary high-water mark. Additionally, much of the time when NWPR was operating occurred in 2020 and 2021 when travel was limited due to COVID-19. Hence field visits occurred more under *Rapanos* and CWR than under NWPR, and *Rapanos* and CWR regulate a larger share of waters. Finally, ACE engineers choose field visits purposefully, for more complex cases (*39*).

WOTUS-ML has imperfect accuracy in part due to characteristics of the training data. AJDs (the training data) can have human error or reflect brief and uncommon environmental conditions (*42, 43*). Implementation varies; for example, ground-truth data reveal sharp regulatory differences around ACE district boundaries (Fig. S2), and individual engineers and field offices could vary further. After an AJD is finalized, landowners receive a form to request an appeal (*44*); the existence of this form implicitly acknowledges that AJDs can be incorrect. In-person visits for roughly half of AJDs provide information not available to train WOTUS-ML. In addition, most AJD PDFs note

that the decision reflects maps, data, photos, or other information provided by the applicant or an associated environmental consultant about the site, information which is unavailable to train WOTUS-ML.

The main text explains how the permutation tests highlight which features most influence the model's predicted probabilities. It is also informative to observe which features play less of a role. ACE documentation notes that soils, hydrology, and vegetation determination can play an important role in identifying wetlands (*21*). Our gNATSGO extract includes five types of categorical variables: soil taxonomic class; soil hydric rating; flooding frequency; ponding frequency; and water table depth variables, in addition to others (Table S3). Fig. S10 shows that gNATSGO is moderately important nationally and is the third-most important feature to determine which sites are jurisdictional within a state. For reasons discussed in SM section A.1, we permute features in groups, so we identify the importance of gNATSGO as a whole, but not the individual features of gNATSGO. Nonetheless, this would suggest that soil and hydrology variables play an important role in guiding WOTUS-ML's predictions across sites within a state. One can similarly look at other permutation importance levels for Fig. S10, compare them to the individual variables listed in Table S3, and determine their relative importance levels.

B.2 Data Sources

To compare our data against the data sources that ACE engineers consult, we selected a random sample of 20 AJDs, obtained the PDF decisions for these AJDs from ACE's ORM2 online database, and categorized the data sources used in these decisions.[2] The five most commonly-used data sources were maps and data sheets submitted by the applicant and associated consultant (100% of sampled AJDs), aerial imagery (100%), topographic maps (90%), NWI (75%), and soil surveys (65%). The AJDs also report that other data sources are less widely used, including ACE's Antecedent Precipitation Tool (30%). HUC maps (25%), NHD (20%), LiDAR (20%), state/local flood maps (15%), FERA/FIRM maps (10%), ACE data sheets (5%), and ACE navigable waters studies (5%). Apart from the applicant materials, which have unspecified data sources, our data cover aerial imagery via NAIP, topographic maps via the DEM, NWI, and soil surveys via gNATSGO.

Observing that many AJDs use a dataset does not indicate that the dataset most influences decisions. Nonetheless, we note that NWI and gNATSGO are the most important for WOTUS-ML decisions within a state in Fig. S10, and are among the most commonly used data sources in these AJDs. We also note the contrast that aerial imagery is used in 100% of these AJDs, but NAIP is much less important in Fig. S10. One potential reason for this is that much of the information in NAIP may already be in our other datasets, some of which ACE does not directly consult (e.g., ecoregions, PRISM, and NLCD).

B.3 Decision Model and Cost Savings

WOTUS-ML can support several parts of the CWA Section 404 process. At decision (a) in Fig. S1, WOTUS-ML scores can give the developer information that helps decide whether to hire a consultant, develop while ignoring WOTUS, or stop development. At (b), WOTUS-ML could give ACE and EPA potential signals for prioritizing potential enforcement cases. At (c), WOTUS-ML could give a developer and consultant useful information about deciding between PJDs, AJDs, development, or non-development. At (d), WOTUS-ML as a decision support tool could give ACE useful information about the estimated probability an AJD is regulated.

We can use this cost to provide an estimate of one component of WOTUS-ML's value, at decision node (c). Because this only includes one component of WOTUS-ML's informational value at this decision node, and since WOTUS-ML has potential value at other decision nodes, it may provide a lower bound on the value of WOTUS-ML to stakeholders.

Specifically, Fig. 2A shows that for 27% of AJDs, WOTUS-ML provides greater than 95% certainty about whether a site is jurisdictional. For these sites, requesting an AJD is more likely to be an error, since a PJD or development without permit are alternatives. The expected cost of the error depends on the number of decisions at each part of the tree and the costs of the other components, most of which are unknown. In the absence of this data, we can obtain one ballpark number from EPA and ACE's (*2*) estimate that each Section 404 permit costs $5,000 to $39,000, representing S in the decision tree. We are unaware of estimates for the cost of AJDs (A), but these estimates of the permit (S) cost provide a plausible starting estimate. If anything, AJD costs might exceed this amount since an

[2] This sample excluded AJDs missing PDFs or AJDs corresponding to more than one PDF in ORM2.

environmental consultant alone can exceed this amount, and then AJDs create additional costs due to delay, uncertainty, and resulting changes to project design.

B.4 External Validity

This paper asks what the CWA regulates, using data over the period 2015-2022. In this sense, predicting CWA jurisdiction in distant future years is not our primary focus. At the same time, for stakeholders, it may be useful to understand the external validity of WOTUS-ML to *Sackett* or future regulation. Given that WOTUS-ML inputs include non-geophysical layers like state boundaries, ACE district boundaries, and distance to a field office, it is unclear how district-specific policies could change over time. The state boundaries with the sharpest regulatory discontinuities in maps like Fig. 3A also lie on ACE district boundaries. Except for the few states which have deregulated AJD determination, our sense is that ACE districts rather than states provide the most important heterogeneity. Additionally, the importance of PRISM climate data and our use of 30-year climate normals could mean that climate change would affect future predictive accuracy.

We believe that five and possibly ten future years is the most relevant period for stakeholders. In part this is because AJDs are valid for five years, because a developer who requests an AJD is likely to develop promptly, and because CWA rules in the last decade have changed frequently. For policymakers, federal budget scoring often considers ten years.

Our cautious assessment is that conditional on rule changes, WOTUS-ML is likely to have good external validity for five to ten years. Given the new *Sackett* rule beginning in late 2023, one could train WOTUS-ML in 2024 or later on *Sackett* AJDs, and then we would expect similar external validity. External validity is more difficult to assess beyond that period. The WOTUS-ML training data covers the eight-year period from 2015 to 2022, which is a similar duration as a five- to ten- year prediction. The role of PRISM in prediction occurs despite the presence of climate change in this period, and our sense is that the PRISM data are highlighting cross-sectional variation in deserts, moisture, heat, and other features that will be broadly similar in the next 5-10 years. If ACE district jurisdictional practices or climate substantially changed after several future years, one could always train a new version of WOTUS-ML on newer data or update the model inputs, although that possibility does not shed light on the relevance of the current WOTUS-ML for predicting future jurisdiction.

We have asked policymakers why ACE district practices differ, since answering this question could help assess the persistence of these patterns. We do not have a single clear answer. While ACE seeks to harmonize AJD criteria across districts, ACE regulatory practices are more decentralized than some federal agencies. Some ACE districts and regions have supplements to the ACE handbook that guide specific aspects of AJDs for local terrain. We also emphasize that ACE district staff are not federal political appointees, and while ACE district leaders have influence, one might expect less variation in ACE district priorities over time than occurs among some federal agencies where presidential appointees change regularly.

We also train a separate deep learning model, AJD-ML, which helps us study WOTUS-ML's external validity. AJD-ML predicts the probability that an arbitrary location has an AJD. By contrast, WOTUS-ML predicts the probability that an AJD for a location concludes that the location has jurisdictional waters. In other words, AJD-ML and WOTUS-ML have two main differences: AJD-ML predicts the probability a point has an AJD while WOTUS-ML predicts the probability that an AJD is jurisdictional; and AJD-ML is trained on both AJDs and non-AJD locations, while WOTUS-ML is only trained on AJDs.

Apart from these differences, AJD-ML and WOTUS-ML are similar. They have identical model architecture, training, and train-validation-test split procedure. The input data are identical except that AJD-ML does not include a WOTUS rule layer. In other words, to predict which points have AJDs in AJD-ML, we use aerial imagery, maps of streams and wetlands, soil quality and groundwater information, a digital elevation model, land use data, etc. To train and evaluate AJD-ML, we use all AJD points and 152,201 of our four million prediction points. We randomly choose these prediction points, subject to not overlapping with any AJD. We train AJD-ML for 11 epochs. AJD-ML's test set accuracy is 92.7%.

Fig. S8 compares WOTUS-ML accuracy across deciles of AJD-ML probabilities. In other words, it investigates how accurately we can predict whether a given location is jurisdictional under the CWA (WOTUS-ML model), separately for points that are more versus less likely to have an AJD (AJD-ML model). In the top panel, the

horizontal axis shows bins indexing the probability that a given point has an AJD. These bins represent fitted probabilities from the AJD-ML model. The vertical axis shows WOTUS-ML's accuracy in each bin (blue triangles), measured using all AJDs; and the vertical axis also shows the share of AJDs in each bin that ACE determines to be jurisdictional (green circles).

Our main finding from Fig. S8 is that WOTUS-ML has similar accuracy of about 80% for both locations that look very much and very little like a typical AJD. We obtain this finding from the observation that the blue triangles in the top panel have flat slope horizontally, at a value near 0.8 on the y-axis. WOTUS-ML's accuracy is similar for sites that have characteristics typical of AJDs, on the graph's right side, as for sites that have characteristics atypical for AJDs, on the graph's left side.

This finding that WOTUS-ML's accuracy is similar for points that do and do not look like a typical AJD helps increase confidence in our headline conclusions about the share of all US wetlands and streams that are jurisdictional. One might have the concern that the wetlands and stream sites which have AJDs are larger, more complex, more ambiguous, or for some other reason different than the mean US wetland or stream. This concern could imply that WOTUS-ML has good accuracy on the kinds of sites on which WOTUS-ML is mostly trained – sites that look like a typical AJD – but not on other kinds of sites. This concern could therefore imply that our statistics on accuracy and jurisdiction of WOTUS-ML might not extrapolate well to all US streams and wetlands. Fig. S8 does not support this concern, and instead suggests that WOTUS-ML has similar accuracy for all types of points, and not only those that look like a typical AJD. It therefore suggests that we can reasonably use WOTUS-ML to assess the share of all streams and wetlands nationally that are jurisdictional, even including streams and wetlands with different characteristics than most AJDs.

B.5 *Sackett*

In May 2023, the Supreme Court issued a decision in *Sackett* (*45*). In August 2023, the EPA issued a final rule which conforms to *Sackett* (*46*). *Sackett* strikes down the significant nexus standard, and excludes interstate wetlands and wetlands separated from jurisdictional waters by a man-made dike or barrier, natural river berm, beach dune, or the like. Formally, the EPA's August 2023 rule modifies the January 2023 rule issued under the Biden administration, which is similar to *Rapanos*, rather than modifying NWPR. *Sackett* and the *Sackett* rule focus on changing the significant nexus standard but focus less on the relatively permanent standard.

We here discuss two potential ways to quantify what *Sackett* regulates. This paper trains a machine learning model on past AJDs issued under a given rule. This is infeasible for *Sackett* in the near term since the EPA has just issued the *Sackett* rule. It will take time for many *Sackett* AJDs to be issued. The recent history of CWA regulation also has had numerous delays and obstacles to changing CWA rules and implementing new rules. It is possible that procedural delays could slow implementation of the *Sackett* rule.

A possible alternative strategy to quantify what *Sackett* regulates, which differs from the methodology applied in the rest of the paper, would be a relabeling technique. Because the EPA's *Sackett* rule modifies the January 2023 rule, which is very similar to *Rapanos*, this methodology would identify which past AJDs already completed under *Rapanos* would lose jurisdiction under *Sackett*. This relabeling technique would then construct a synthetic dataset of *Sackett* AJDs by changing the labels (the jurisdictional outcomes) for the AJDs that would lose jurisdiction in *Sackett* relative to *Rapanos*. Applying this relabeling methodology would then train WOTUS-ML on this relabeled set of AJDs, and predict jurisdiction for the 4 million and other points. These points would describe a prediction of what *Sackett* regulates.

While we investigated the potential of this relabeling methodology in detail, we do not implement it here because we believe that without *Sackett* AJDs for training a model, it is difficult to implement it with sufficient accuracy. Specifically, the construction of the *Sackett* rule makes it difficult to conclusively identify which *Rapanos* AJDs would lose jurisdiction under *Sackett*. The *Sackett* rule removes the significant nexus standard for several legal categories of waters identified in the January 2023 rule. The legal categories of waters identified in *Rapanos* AJDs (formally, called "resource types" in the AJD data) differ from the legal categories of waters listed in the January 2023 rule. Thus, determining which *Rapanos* AJDs would lose jurisdiction under *Sackett* depends on a match in legal water categories between the two rules. More importantly, for several legal categories of waters in the January 2023 rule, *Sackett*'s removal of the significant nexus standard deregulates some but not all waters in the category. Within a category, the *Rapanos* AJDs alone have insufficient information both about what share and which AJDs would lose jurisdiction under *Rapanos*.

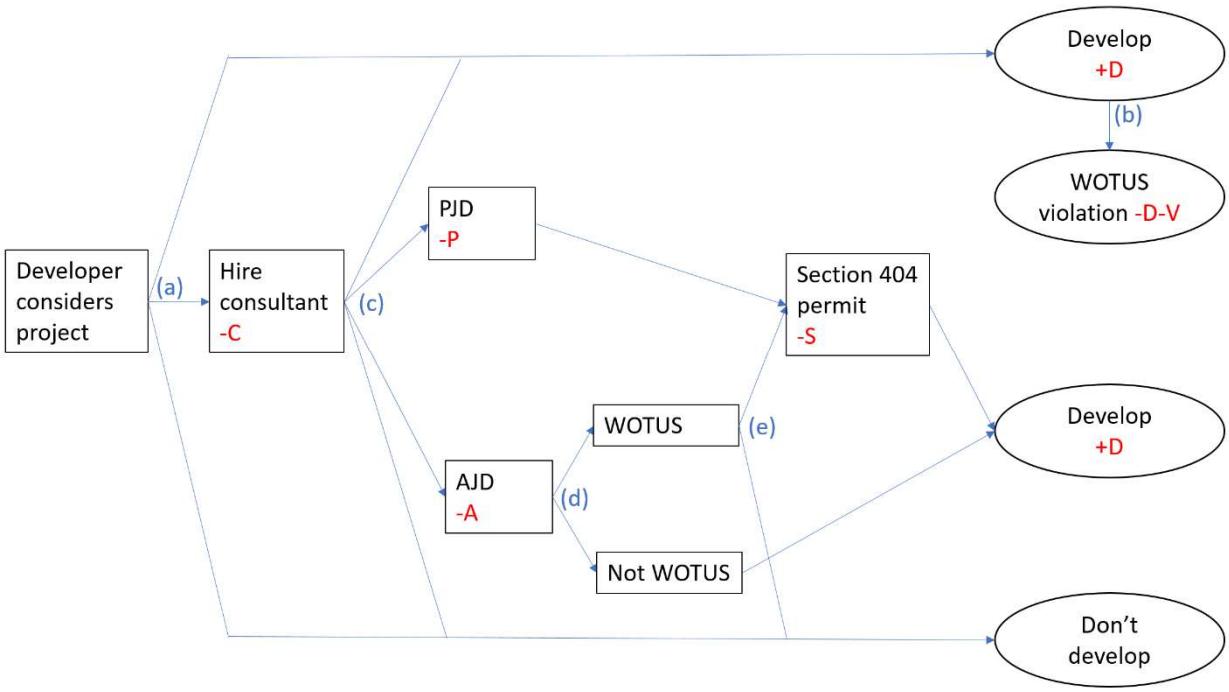
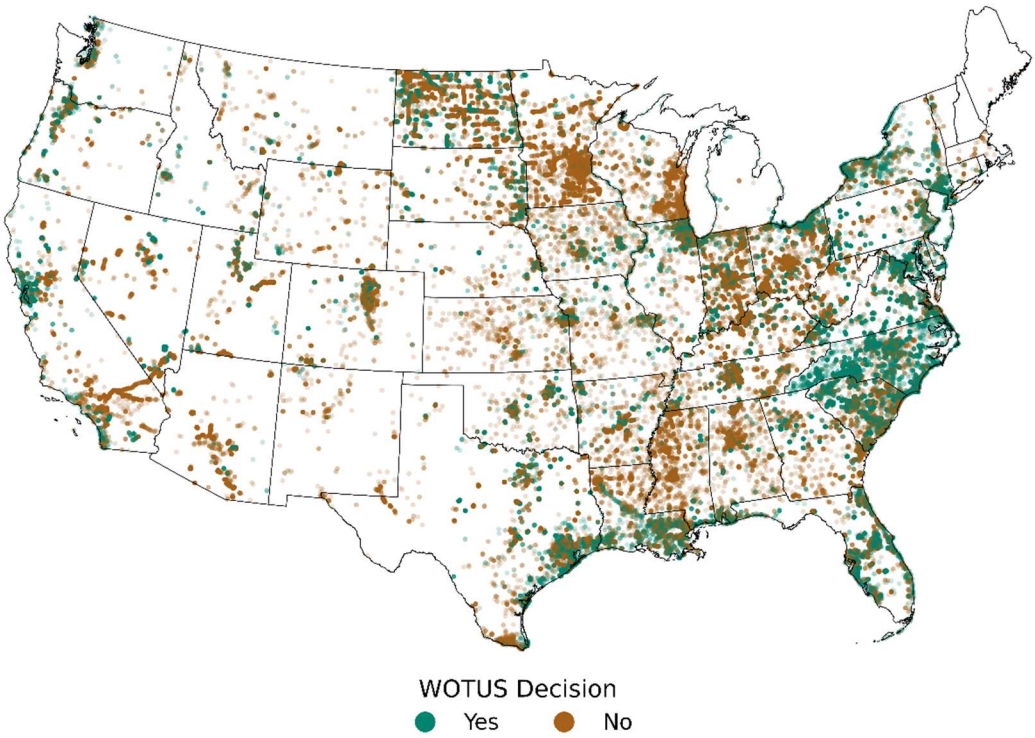
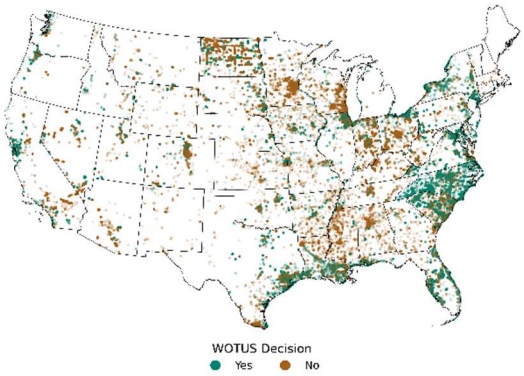


Fig. S1: A Decision Tree Clarifies the Process of Development and Jurisdictional Determinations and Shows Potential Uses of WOTUS-ML. Each box describes an intermediate decision and each oval describes a potential terminal outcome. Branches and decision nodes are in blue and lowercase, costs and payoffs are in capital red. At (a), the developer chooses whether to develop while ignoring the CWA, hire an environmental consultant to support the CWA Section 404 process, or not develop. If the project proceeds assuming jurisdictional Waters of the United States (WOTUS) are not present, at (b), ACE and EPA may discover the development, conclude WOTUS are present, and require restoring the site to its original conditions and imposing a fine. At (c), the developer and an environmental consultant may provide a PJD, request an AJD, proceed ignoring WOTUS, or not develop. At (d), ACE determines whether the AJD is WOTUS. If the AJD is not WOTUS, development occurs. At (e), if the AJD is WOTUS, the developer may obtain a Section 404 permit and develop or may stop development. AJD are approved jurisdictional determinations, PJD are preliminary jurisdictional determinations, and WOTUS are Waters of the United States. The model shows common sequences but abstracts from others (e.g., a developer could provide a PJD without a consultant, and might provide a PJD but not develop the project). Payoff variables are follows: C is cost of hiring an environmental consultant; P is cost of obtaining a PJD, including paperwork, delays, and uncertainty; A is cost of obtaining an AJD, including paperwork, delays, and uncertainty; S is cost of obtaining and complying with a Section 404 nationwide or individual permit, including mitigation, compensatory measures, or constraints on development; D is payoff from development; and V is fines and other costs due to violating CWA Section 404.

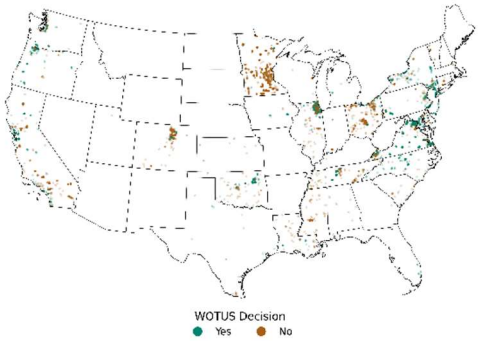
A AJDs, all rules



B AJDs, *Rapanos*



C AJDs, CWR



D AJDs, NWPR

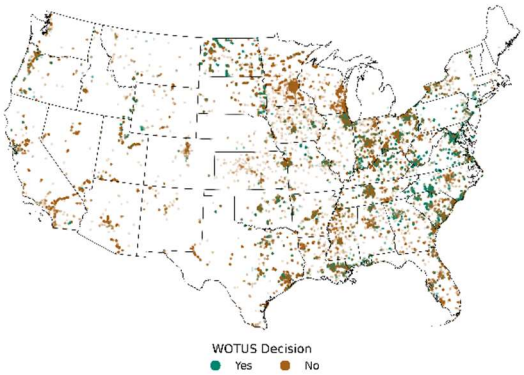


Fig. S2: AJDs (Training Data), by Rule. (**A**) All AJDs. (**B**) AJDs decided under *Rapanos*. (**C**) AJDs decided under CWR. (**D**) AJDs decided under NWPR. Green dots represent jurisdictional sites; brown dots represent non-jurisdictional sites.

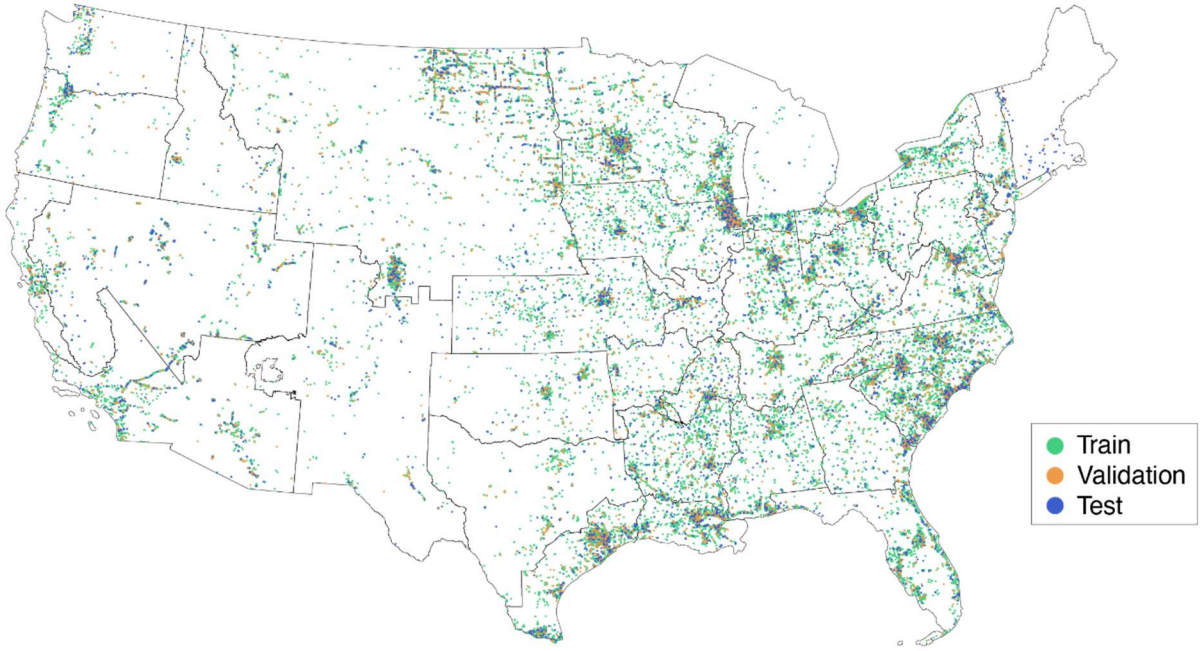
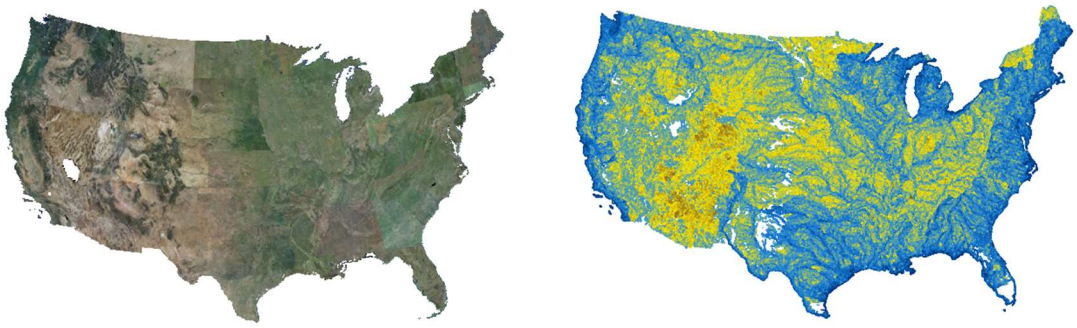
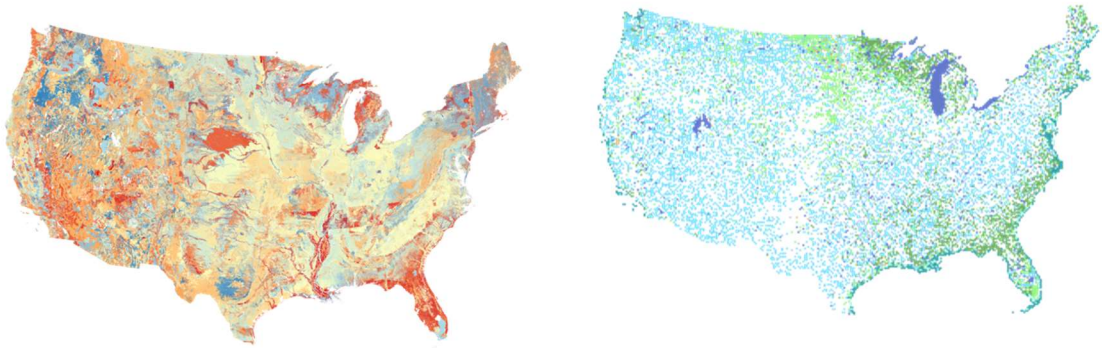


Fig. S3: Train, validation and test assignment. Train points are in green, validation points in brown, and test points in blue. Black lines show ACE district boundaries. We plot the validation and test sets on top of the train set.

A National Agricultural Imagery Program **B** National Hydrography Dataset



C Gridded National Soil Survey Geographic Database **D** National Wetlands Inventory



E National Land Cover Database **F** PRISM

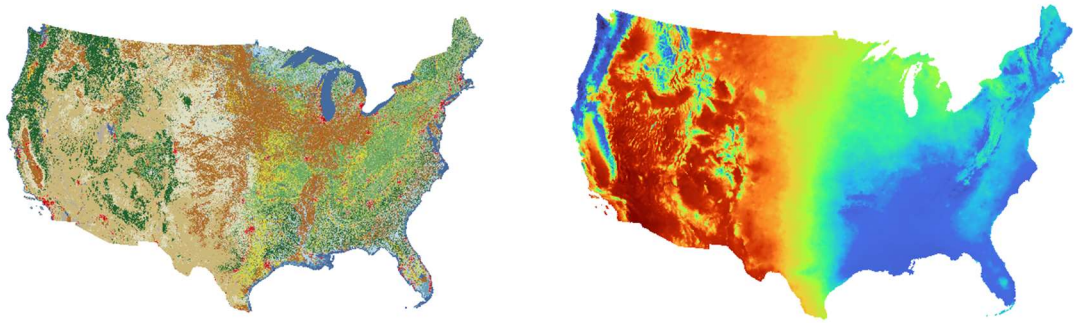


Fig. S4: Maps of six input layers. (**A**) shows NAIP imagery. (**B**) shows NHD flowlines, where darker colors represent streams that flow more directly into oceans or great lakes (stream order). (**C**) shows soil taxonomic class from gNATSGO. (**D**) shows NWI polygons. **(E)** shows land cover. **(F)** shows precipitation patterns from PRISM.

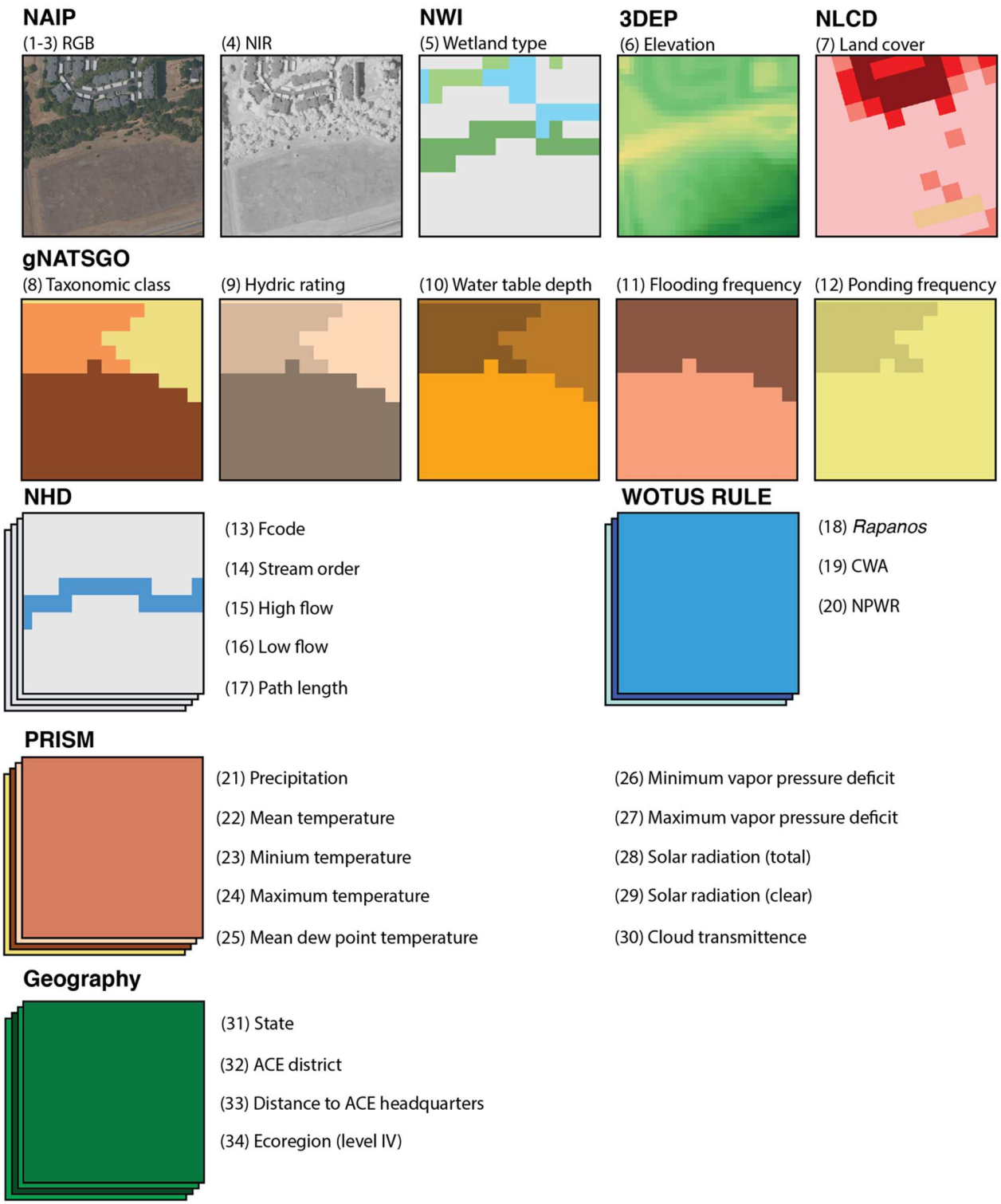
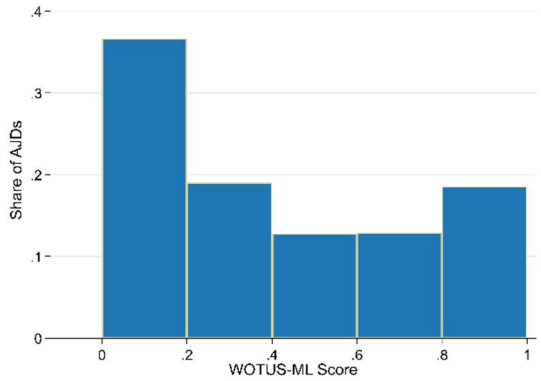


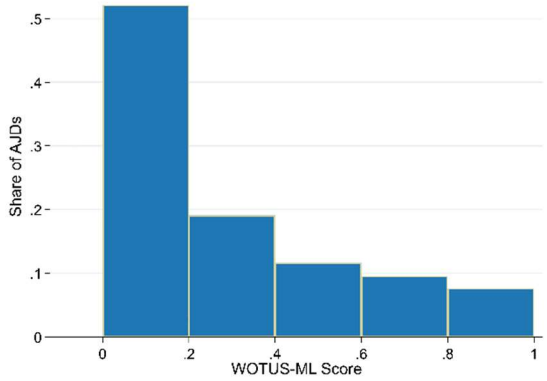
Fig. S5. Sample Inputs for One Site. Fig S5 depicts the 34 input layers for WOTUS-ML (see Table S3 for details). The location shown here is associated with an Approved Jurisdictional Determination (AJD) located at -121.795E, 39.734N in Butte County, California. The nearest water body is Dead Horse Slough, which flows into the Traditional Navigable Water (TNW) Little Chico Creek. The National Agriculture Imagery Program (NAIP) imagery shows that the site includes an undeveloped lot besides a waterbody with adjacent wetlands. The National Wetlands Inventory (NWI) shows three types of wetlands at this location: Riverine (light blue), Freshwater Emergent Wetlands

(light green), and Freshwater Forested/Shrub Wetland (dark green). The National Land Cover Database (NLCD) layer includes open space (pink), developed land (red), and grasslands (tan). 3DEP shows that elevation ranges from 73m in the streambed to 78m in the field. The water feature appears in the National Hydrography Dataset (NHD) layer as an intermittent stream (FCode 46003). WOTUS-ML also inputs soil characteristics in the gridded National Soil Survey Gridded Database (gNATSGO) layers, information on climate in the PRISM layers, and geographic variables such as state, ecoregion, and U.S. Army Corps of Engineers (ACE) district. Finally, the model includes a one-hot encoding for each WOTUS rule: *Rapanos*, the Clean Water Rule (CWR), or the Navigable Waters Protection Rule (NWPR). This site was determined to be jurisdictional under *Rapanos*. WOTUS-ML assigns this site a score (i.e., probability of being jurisdictional) of 96% under *Rapanos*.

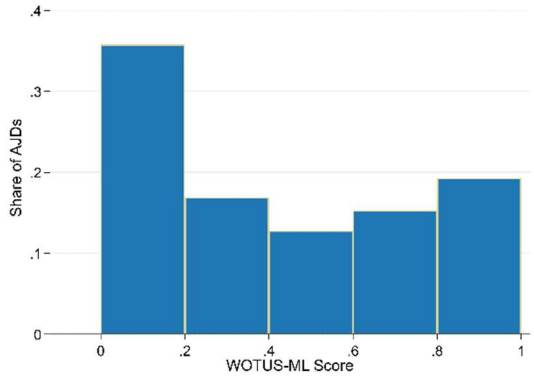
A *Rapanos*, AJD Test Set



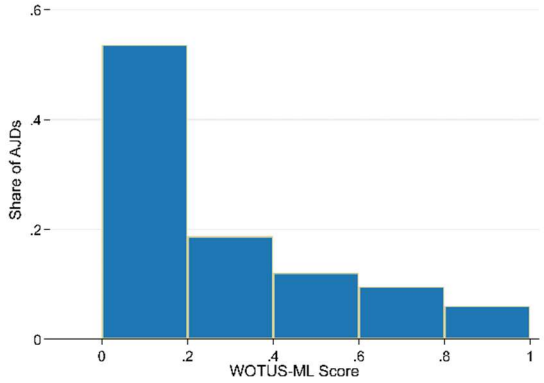
B *Rapanos*, Four Million Prediction Points



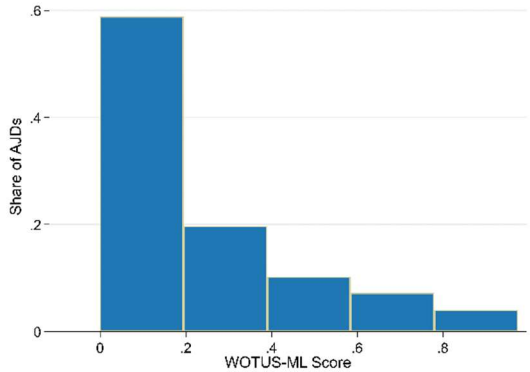
C CWR, AJD Test Set



D CWR, Four Million Prediction Points



E NWPR, AJD Test Set



F NWPR, Four Million Prediction Points

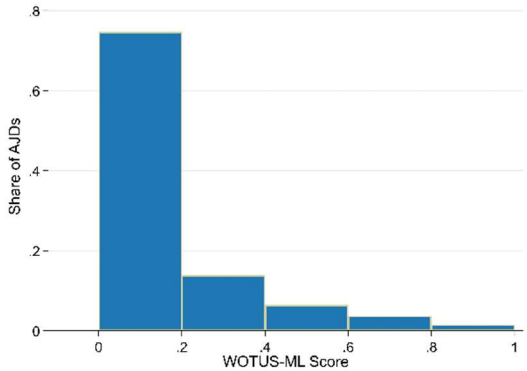


Fig. S6: Histograms of WOTUS-ML Scores for AJD Test Set and Four Million Prediction Points, By Rule.

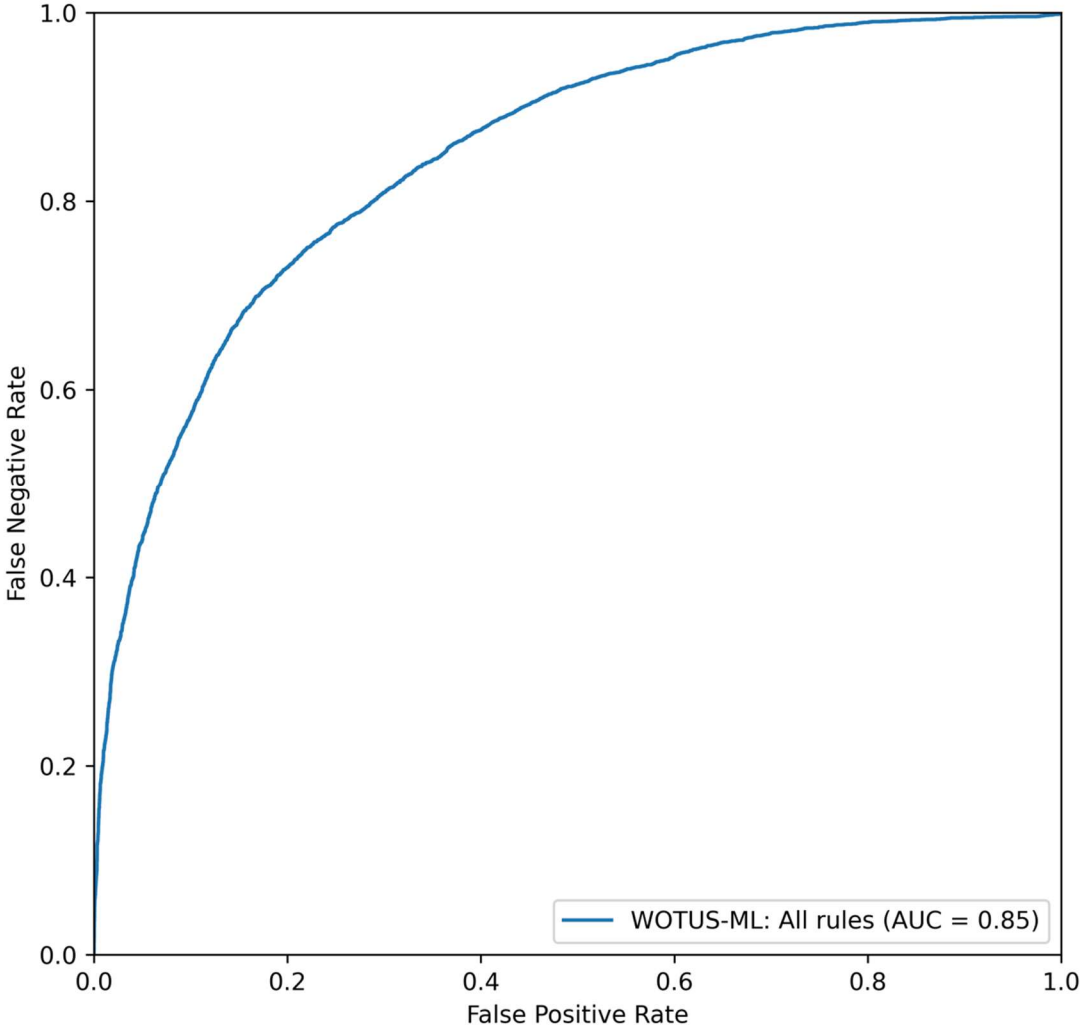


Fig S7: WOTUS-ML achieves an AUC of 0.85. The Receiver Operator Characteristic (ROC) curve is plotted by calculating false positive and false negative rates for various classification thresholds. The Area Under the ROC Curve (AUC) is the integral of the ROC curve along the x-axis.

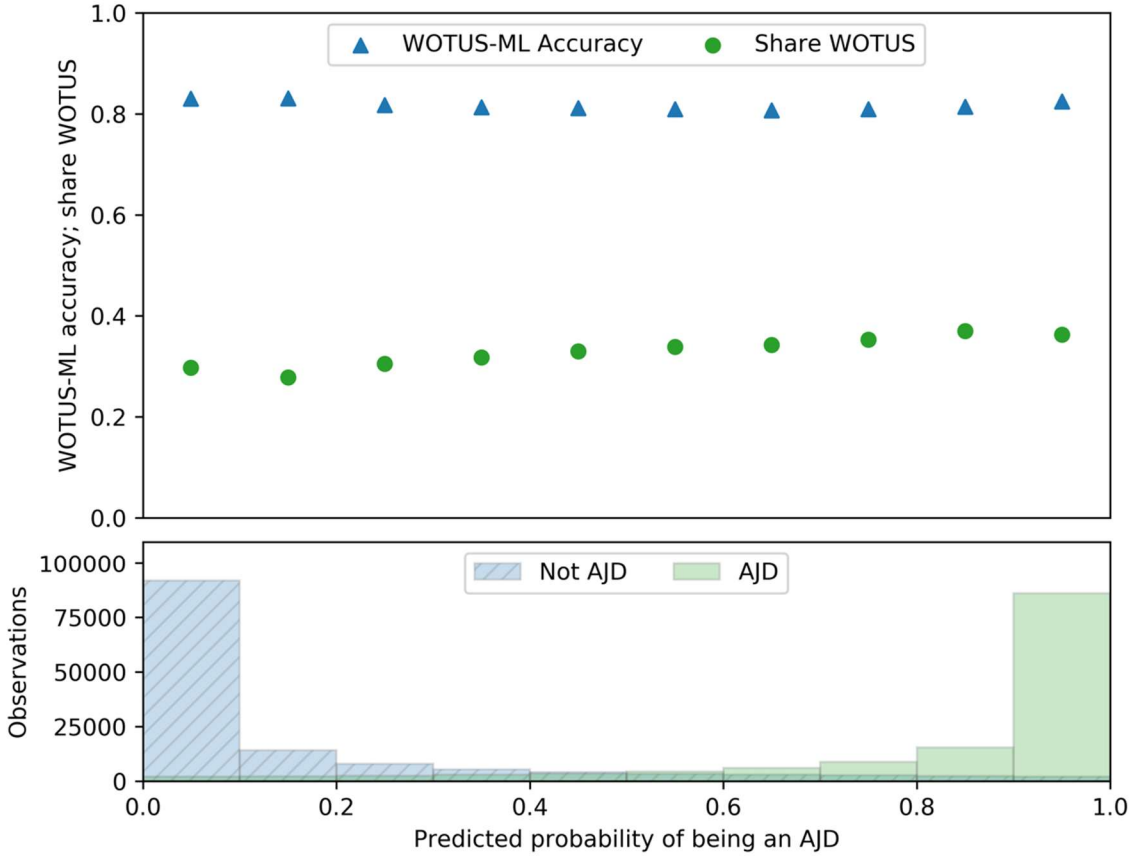
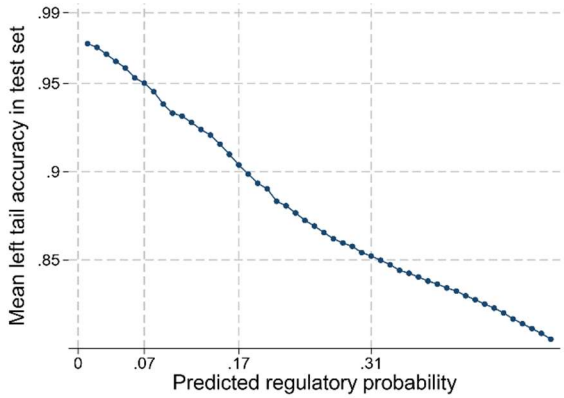
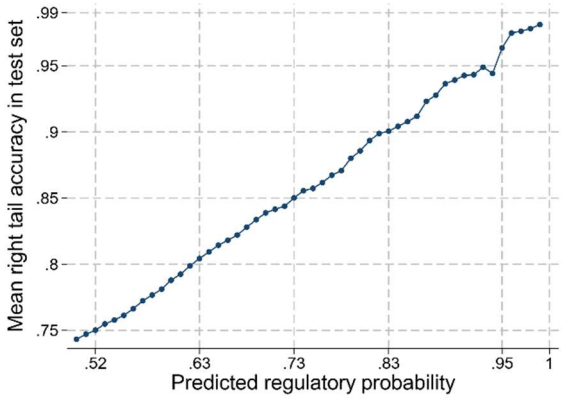


Fig. S8: WOTUS-ML Accuracy is Not Correlated with Predicted AJD Probability. Predicted AJD probabilities are merged with WOTUS-ML predictions, then WOTUS-ML accuracy is calculated for each 10 percentage point bin of predicted AJD probability. The WOTUS-ML accuracy for each bin is plotted as a blue triangle. We find no correlation between WOTUS-ML accuracy and predicted AJD probability. We also plot the share of AJDs in each bin that are WOTUS as green circles. We find a weak positive correlation between the predicted probability of being an AJD and the probability of being WOTUS. The bottom panel shows histograms of the distribution of predicted AJD probabilities for the AJDs and the sampled prediction points. Note that this figure includes points from all folds (training, validation, and test) to maximize the number of observations in each bin. The test set only (not shown) has a similar horizontal pattern, though with more variance given the smaller sample. This has little effect on interpretation because WOTUS-ML and AJD-ML both exhibit minimal overfitting.

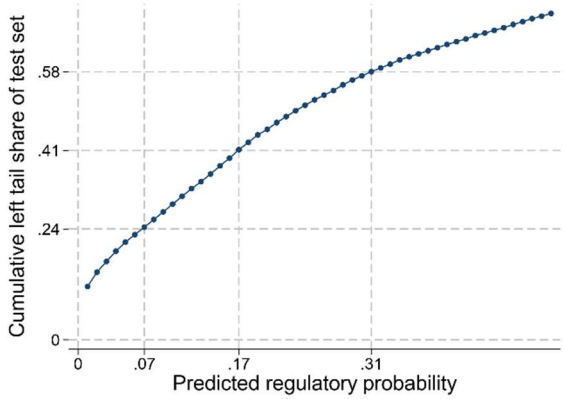
A Left Tail Accuracy



B Right Tail Accuracy



C Left Tail Share



D Right Tail Share

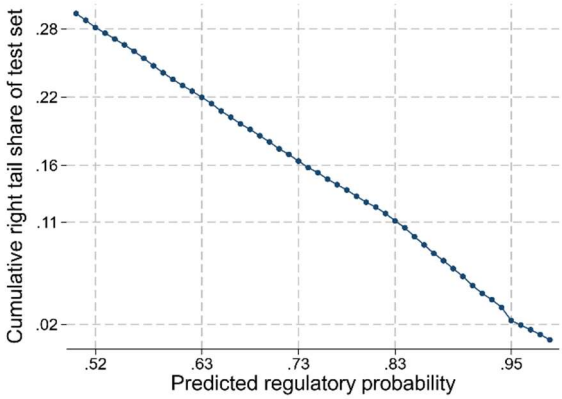


Fig. S9: Cumulative Shares and Accuracy in AJD Test Set. The x-axis in (**A**) lists threshold predicted regulatory probabilities such that for test set AJDs below each threshold, WOTUS-ML has accuracy of 99%, 95%, 90%, or 85%. (**B**) is similar but calculates accuracy for AJDs above each threshold. (**C**) calculates the cumulative share of test set AJDs in the test set with predicted regulatory probability below the thresholds identified in (**A**). (**D**) is similar but calculates the cumulative share of test set AJDs above each threshold. Summing the y-axis values in (**C**) and (**D**) gives the y-axis in Fig. 2A.

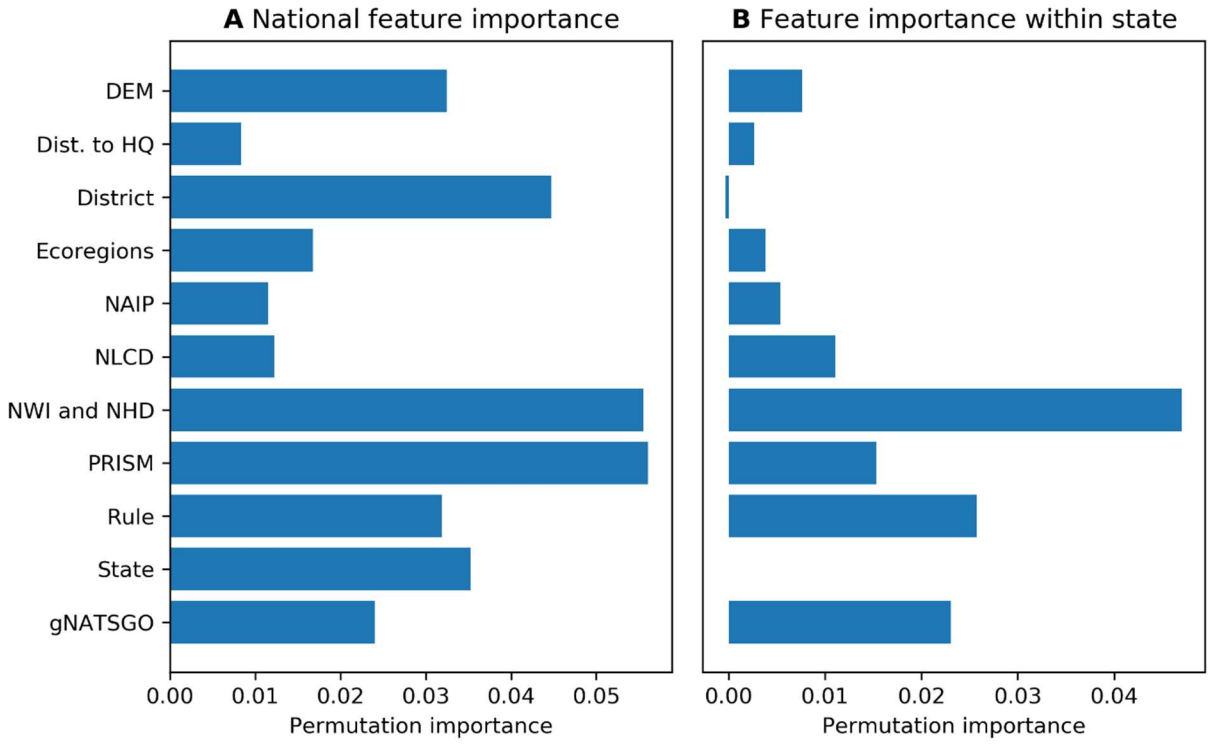


Fig. S10: Permutation Tests of WOTUS-ML Feature Importance. Features are permuted across samples in validation set (**A**) nationally and (**B**) within each state. The datasets represent the following information: NAIP is imagery, NWI and NHD are wetlands and streams, DEM is elevation, NLCD is land cover, PRISM is climate, and gNATSGO is soil and groundwater.

Table S1: Categorization of Nine Hydrological (Cowardin) Codes

Cowardin code	Description
1. Streams, ephemeral	A wetland, spring, stream, river, pond or lake that only exists for a short period
2. Streams, intermittent	Intermittent, Riverine; Streambed, Intermittent, Riverine
3. Streams, perennial and other	Upper Perennial, Riverine; Lower Perennial, Riverine; Unknown Perennial, Riverine
4. Wetland, emergent	Emergent, Palustrine; Persistent, Emergent, Palustrine; Nonpersistent, Emergent, Palustrine
5. Wetland, forested	Forested, Palustrine; Broad-Leaved Deciduous, Forested, Palustrine; Needle-Leaved Evergreen, Forested, Palustrine"; Needle-Leaved Deciduous, Forested, Palustrine; Broad-Leaved Evergreen, Forested, Palustrine; Indeterminate Deciduous, Forested, Palustrine
6. Wetland, other	All other palustrine
7. Estuaries	Estuarine
8. Uplands	Uplands
9. Other	Marine, lacustrine, riparian

Table S2: Categorization of Nine Resource Types. RPW, relatively permanent waters; TNW, territorial national waters.

	Rapanos	**CWR**	**NWPR**
1. Ephemeral			(b)(3) Ephemeral feature, including an ephemeral stream, swale, gully, rill, or pool
2. Isolated	Isolated (interstate or intrastate) waters		(b)(1)
3. Non-RPW that flows directly or indirectly into TNW	Non-RPW that flows directly or indirectly into TNW		
4. Other non-jurisdictional		(b)(1), (b)(2), (b)(3)(i), (b)(3)(ii), (b)(3)(iii), (b)(4)(v), (b)(5), (b)(7), (b)(4)(iii), (b)(4)(iv), (b)(4)(ii), (b)(4)(i), (b)(4)(vi), (b)(6), (b)(4)(vii)	(b)(2), (b)(4), (b)(5), (b)(6), (b)(7), (b)(8), (b)(9), (b)(10), (b)(11), (b)(12), the review area is comprised entirely of dry land
5. Other jurisdictional		(a)(4), (a)(7), (a)(8)	(a)(3)
6. RPW that flows directly or indirectly into TNW	Relatively Permanent Water that flows directly or indirectly into Traditional Navigable Water	(a)(5)	(a)(2)
7. TNW	Traditional Navigable Water	(a)(1), (a)(2)	(a)(1)
8. Uplands	Uplands	Uplands	
9. Wetlands and adjacent/abutting regulated waters	Wetlands Directly Abutting RPW that flows directly or indirectly into TNW; Wetland Adjacent to Non-RPW that flows directly or indirectly into TNW; Wetlands Adjacent but not Directly Abutting RPW that flows directly or indirectly intoTNW; Wetlands Adjacent to TNW	(a)(6)	(a)(4)

Table S3: Datasets to Support Predictions of Jurisdictional Waters.

Dataset	Input layer	Variable definition	Data type	Spatial Resolution	Source
National Agriculture Imagery Program (NAIP)	Red band	Red channel visible light	Raster	0.6 to 1.0 meters	(*15*)
	Blue	Blue channel visible light			
	Green	Green channel visible light			
	NIR	Near-infrared light			
National Wetlands Inventory (NWI)	Wetland type	NWI wetland types: Estuarine and Marine Deepwater, Estuarine and Marine Wetland, Freshwater Emergent Wetland, Freshwater Forested/Shrub Wetland, Freshwater Pond, Lake, Riverine, Other	Vector	1:250,000	(*19*)
National Hydrography Dataset (NHD) Plus V2	FCode	Water feature type (e.g. perennial stream, intermittent stream, coastline)	Vector	1:100,000	(*17*)
	Path length	NHD flowline distance			
	High flow	Maximum flow for this water segment over a sequential 3-month period, using NHD Value Added Attributes Enhanced Runoff Method (EROM) long-term mean flow estimates for each month.			
	Low flow	Minimum flow for this water segment over a sequential 3-month period, using EROM.			
	Stream order	Hierarchy of streams from the source (or headwaters) downstream			
USGS 3-Dimensional Elevation Program (3DEP)	Elevation	Height above sea-level	Raster	10 meters	(*30*)
U.S. EPA Ecoregions	Level IV Ecoregion	Ecoregions are areas where ecosystems (and the type, quality, and quantity of natural resources) are generally similar. There are 967 level IV ecoregions in the United States.	Vector	1:250,000	(*31*)

Continued next page

Dataset	Input layer	Variable definition	Data type	Spatial Resolution	Source
Parameter-elevation Regressions on Independent Slopes Model (PRISM) 30-year Normals	Precipitation	Average annual total precipitation	Raster	4 kilometers	(*32*)
	Minimum temperature	Daily minimum temperature, averaged over 1990-2021			
	Maximum temperature	Daily maximum temperature, averaged over 1990-2021			
	Mean temperature	Daily mean temperature, averaged over 1990-2021			
	Mean dew point temperature	Daily mean dew point temperature (the temperature to which air must be cooled to become saturated with water vapor), averaged over 1990-2021			
	Minimum vapor pressure deficit (VPD)	Minimum VPD (difference between the amount of moisture in the air and how much moisture the air can hold), averaged over 1990-2021			
	Maximum vapor pressure deficit (VPD)	Maximum VPD (difference between the amount of moisture in the air and how much moisture the air can hold), averaged over 1990-2021			
	Solar radiation (clear sky)	Total daily global shortwave solar radiation received on a horizontal surface, averaged over 1990-2021			
	Solar radiation (total)	Total solar radiation incident on a horizontal surface), averaged over 1990-2021			
	Cloudiness	Atmospheric transmittance (cloudiness), averaged over 1990-2021			
U.S. Army Corps Regulatory Boundaries	District codes	Each ACE district is assigned a unique value.	Point	1:250,000	(*20*)
	Distance to headquarters	We calculate the distance from each point to the district headquarters.			

Continued next page

Dataset	Input layer	Variable definition	Data type	Spatial Resolution	Source
Gridded National Soil Survey Geographic Database (gNATSGO)	Taxonomic class	The Soil Taxonomy subgroup and family for a soil.	Raster	30 meters	(*16*)
	Hydric rating	Is the map unit "hydric soil"?			
	Flooding frequency	The annual probability of a flood event expressed as a class.			
	Ponding frequency	The number of times ponding occurs over a year			
	Water table depth	The shallowest depth to a wet soil layer			
National Land Cover Database (NLCD)	Landcover	The NLCD has 20 land cover classes: Open water, ice/snow, four classes of developed land (open, low, medium, and high), barren, three forest classes (evergreen, deciduous, mixed), two scrub classes (dwarf, shrub), four herbaceous classes (grassland, sedge, moss, lichen), two agricultural classes (pasture/hay, cultivated), and two wetland classes (woody, emergent herbaceous)	Raster	30 meters	(*33*)
Topologically Integrated Geographic Encoding and Referencing System (TIGER)/Line State boundaries	State codes	Each state is assigned a unique value	Vector	1:250,000	(*34*)
CWA Approved Jurisdictional Determinations	WOTUS rule	Three WOTUS rules: *Rapanos,* CWR, NWPR	Point	--	(*38*)

Table S4: Accuracy of WOTUS-ML Model Predictions, All Rules. Statistics describe the AJD test set. Column (1) shows the share that USACE determines to be regulated. Column (2) shows the share that WOTUS-ML predicts to have score above 0.5. Column (3) shows the share that WOTUS-ML correctly classifies as regulated. Column (4) shows the number of test set AJDs. Field visit indicators and Cowardin codes are from the AJD data. NWI and NHD include points within 5-10 meters of these vector data. (**E**) shows five ACE districts selected to represent districts with among the largest or smallest numbers of AJDs, and to show geographically heterogeneous districts.

	AJDs: regulated (1)	**WOTUS-ML: score > 0.5** (2)	**Accuracy** (3)	**N** (4)
A General groups of points				
All AJDs	0.35	0.29	0.79	15,970
USACE field visit	0.46	0.38	0.74	7,198
USACE no field visit	0.26	0.21	0.82	8,772
B By Cowardin Codes				
All rivers and streams (riverine)	0.43	0.31	0.78	4,353
Wetlands (palustrine)	0.38	0.28	0.77	8,203
Uplands	0.00	0.17	0.83	2,529
Estuaries (estuarine)	0.99	0.94	0.94	304
Lakes (lacustrine)	0.39	0.30	0.81	352
C By Vector Data Overlay				
All rivers and streams (NHD)	0.71	0.75	0.77	768
All wetlands (NWI)	0.53	0.45	0.75	2,247
D By year				
2016	0.52	0.46	0.77	1,731
2017	0.41	0.40	0.82	1,255
2018	0.34	0.32	0.80	1,580
2019	0.38	0.37	0.76	2,067
2020	0.31	0.22	0.79	3,816
2021	0.25	0.16	0.79	3,982
2022	0.43	0.39	0.75	918
E Selected USACE Districts				
Albuquerque	0.13	0.05	0.86	100
Charleston	0.51	0.44	0.73	935
Philadelphia	0.62	0.69	0.69	124
St. Paul	0.02	0.01	0.98	1,296
Wilmington	0.78	0.84	0.79	868

Table S5. Confusion Matrix. For each panel, table entries show share of test set AJDs in each cell. Top-left cell within a panel represents true positives, bottom-right represents true negatives, and top-right and bottom-left represent false positives and false negatives. Share of true positives plus true negatives equals accuracy.

		AJDs: Test set outcomes	
		Regulated (1)	**Not regulated** (2)
A All rules			
	WOTUS-ML Prediction		
	Regulated	0.22	0.08
	Not regulated	0.14	0.57
B *Rapanos*			
	WOTUS-ML Prediction		
	Regulated	0.29	0.09
	Not regulated	0.13	0.50
C CWR			
	WOTUS-ML Prediction		
	Regulated	0.29	0.12
	Not regulated	0.11	0.48
D NWPR			
	WOTUS-ML Prediction		
	Regulated	0.11	0.05
	Not regulated	0.16	0.69

Table S6: Accuracy of WOTUS-ML Model Predictions, *Rapanos*. Statistics describe the AJD test set. Column (1) shows the share that USACE determines to be regulated. Column (2) shows the share that WOTUS-ML predicts to have score above 0.5. Column (3) shows the share that WOTUS-ML correctly classifies as regulated. Column (4) shows the number of test set AJDs. Field visit indicators and Cowardin codes are from the AJD data. NWI and NHD include points within 5-10 meters of these vector data. (**E**) shows five ACE districts selected to represent districts with among the largest or smallest numbers of AJDs, and to show geographically heterogeneous districts.

	AJDs: regulated	**WOTUS-ML: score > 0.5**	**Accuracy**	**N**
	(1)	(2)	(3)	(4)
A General groups of points				
All AJDs	0.41	0.37	0.78	8,198
USACE field visit	0.53	0.49	0.76	4,011
USACE no field visit	0.30	0.26	0.80	4,187
B By Cowardin Codes				
All rivers and streams (riverine)	0.67	0.54	0.80	1,425
Wetlands (palustrine)	0.47	0.36	0.76	4,199
Uplands	0.00	0.19	0.81	2,059
Estuaries (estuarine)	1.00	0.96	0.96	236
Lakes (lacustrine)	0.59	0.49	0.77	156
C By Vector Data Overlay				
All rivers and streams (NHD)	0.79	0.87	0.84	370
All wetlands (NWI)	0.61	0.55	0.76	1,285
D By year				
2016	0.52	0.46	0.77	1,731
2017	0.41	0.40	0.82	1,255
2018	0.32	0.30	0.83	1,260
2019	0.39	0.34	0.75	1,217
2020	0.38	0.35	0.82	902
2021	0.35	0.31	0.71	597
2022	0.43	0.39	0.75	918
E Selected USACE Districts				
Albuquerque	0.26	0.11	0.72	47
Charleston	0.65	0.66	0.77	542
Philadelphia	0.48	0.79	0.69	42
St. Paul	0.04	0.01	0.97	618
Wilmington	0.85	0.93	0.84	610

Table S7: Accuracy of WOTUS-ML Model Predictions, NWPR. Statistics describe the AJD test set. Column (1) shows the share that USACE determines to be regulated. Column (2) shows the share that WOTUS-ML predicts to have score above 0.5. Column (3) shows the share that WOTUS-ML correctly classifies as regulated. Column (4) shows the number of test set AJDs. Field visit indicators and Cowardin codes are from the AJD data. NWI and NHD include points within 5-10 meters of these vector data. (**E**) shows five ACE districts selected to represent districts with among the largest or smallest numbers of AJDs, and to show geographically heterogeneous districts.

	AJDs: regulated	**WOTUS-ML: score > 0.5**	**Accuracy**	**N**
	(1)	(2)	(3)	(4)
A General groups of points				
All AJDs	0.26	0.15	0.79	6299
USACE field visit	0.33	0.16	0.71	2478
USACE no field visit	0.21	0.15	0.85	3821
B By Cowardin Codes				
Rivers and streams (riverine)	0.30	0.15	0.77	2467
Wetlands (palustrine)	0.26	0.16	0.80	3345
Uplands	0.00	0.09	0.91	240
Estuaries (estuarine)	0.88	0.79	0.73	33
Lakes (lacustrine)	0.10	0.05	0.90	129
C By Vector Data Overlay				
All rivers and streams (NHD)	0.60	0.60	0.68	329
All wetlands (NWI)	0.38	0.24	0.73	798
D By year				
2020	0.29	0.18	0.78	2914
2021	0.24	0.13	0.80	3385
E Selected USACE Districts				
Albuquerque	0.00	0.00	1.00	46
Charleston	0.33	0.14	0.67	366
Philadelphia	0.63	0.47	0.67	49
St. Paul	0.02	0.00	0.98	479
Wilmington	0.54	0.56	0.61	208

Table S8: Accuracy of Cowardin Code and Resource Type Predictions. Accuracy takes test AJDs that WOTUS-ML classifies in a particular category (e.g., Cowardin code of ephemeral streams) and reports the share of these AJDs where the ACE engineers classifies the resource as in this category.

	AJDs: share	WOTUS-ML: score > 0.5	Accuracy	N
	(1)	(2)	(3)	(4)
A. Cowardin Codes				
1. Streams, ephemeral	0.14	0.18	0.66	2,265
2. Streams, intermittent	0.07	0.03	0.11	1,128
3. Streams, perennial and other	0.06	0.02	0.16	960
4. Wetland, emergent	0.29	0.34	0.65	4,655
5. Wetland, forested	0.16	0.19	0.61	2,477
6. Wetland, other	0.07	0.01	0.05	1,071
7. Estuaries	0.02	0.03	0.85	304
8. Uplands	0.16	0.18	0.63	2,529
9. Other	0.04	0.02	0.13	581
Overall	1.00	1.00	0.52	15,970
B. Resource Types				
1. Ephemeral	0.09	0.11	0.75	1,388
2. Isolated	0.29	0.36	0.70	4,664
3. Non-RPW that flows directly or indirectly into TNW	0.03	0.01	0.26	424
4. Other non-jurisdictional	0.13	0.11	0.53	2,070
5. Other jurisdictional	0.01	0.01	0.12	164
6. RPW that flows directly or indirectly into TNW	0.11	0.06	0.20	1,698
7. TNW	0.02	0.03	0.67	353
8. Uplands	0.13	0.12	0.57	2,066
9. Wetlands adjacent/abutting regulated waters	0.20	0.18	0.44	3,143
Overall	1.00	1.00	0.54	15,970

Table S9: Regulated Stream Miles and Wetland Acres, by State. Total stream miles in (2) is from NHD stream and river flowline features. Total wetland acres in (3) is from NWI. Regulation rates in (4), (5), (6), and (7) are from WOTUS-ML, applied to the subset of four million prediction points that are within 10 meters of NHD or NWI features. The difference in column (6) is measured in stream miles, and in column (9) in wetland acres. Positive entries in (6) are associated with perennial streams, see SM section A.4.

			Stream miles regulated			Wetland acres regulated		
State (1)	**Total Stream miles** (2)	**Total Wetland Acres** (3)	***Rapanos* (%)** (4)	**NWPR (%)** (5)	**Difference NWPR - *Rapanos*** (6)	***Rapanos* (%)** (7)	**NWPR (%)** (8)	**Difference NWPR - *Rapanos*** (9)
All National	3,154,480	119,825,268	—	—	-686,257	—	—	-34,659,988
Alabama	72,650	4,043,348	0.89	0.87	-1,443	0.86	0.63	-915,215
Arizona	139,281	262,281	0.10	0.01	-12,562	0.33	0.06	-69,684
Arkansas	78,496	2,558,428	0.89	0.67	-16,885	0.77	0.39	-978,770
California	173,028	2,789,804	0.58	0.13	-78,689	0.68	0.15	-1,460,431
Colorado	93,255	1,522,952	0.45	0.17	-26,541	0.43	0.18	-387,822
Connecticut	5,215	304,750	1.00	0.99	-40	0.98	0.48	-152,838
Delaware	2,234	290,940	1.00	0.95	-120	0.99	0.78	-63,450
Florida	22,385	12,681,770	0.99	0.83	-3,636	0.88	0.24	-8,196,536
Georgia	64,833	6,396,737	0.95	0.92	-1,791	0.91	0.47	-2,772,997
Idaho	94,753	1,119,249	0.66	0.40	-24,510	0.68	0.35	-370,204
Illinois	67,074	1,271,986	0.68	0.66	-1,746	0.54	0.42	-147,421
Indiana	24,066	1,008,100	0.72	0.73	286	0.29	0.25	-37,603
Iowa	67,717	1,014,174	0.82	0.71	-7,654	0.63	0.34	-297,352
Kansas	118,236	1,349,856	0.87	0.42	-53,764	0.79	0.33	-620,966
Kentucky	45,616	430,781	0.20	0.30	4,421	0.38	0.29	-35,465
Louisiana	43,096	8,092,819	0.87	0.66	-8,932	0.80	0.55	-2,055,459
Maine	24,974	2,569,961	0.99	0.72	-6,931	0.83	0.32	-1,317,269
Maryland	10,263	863,198	0.98	0.92	-589	0.98	0.66	-278,264
Massachusetts	7,273	775,106	0.99	0.92	-477	0.96	0.41	-425,813
Michigan	47,861	7,712,081	0.55	0.29	-12,435	0.06	0.03	-214,072
Minnesota	60,103	9,973,334	0.11	0.11	-40	0.01	0.01	-58,149
Mississippi	77,386	4,534,181	0.65	0.55	-7,445	0.68	0.43	-1,092,420
Missouri	95,347	1,388,966	0.82	0.63	-17,551	0.78	0.43	-481,217
Montana	166,847	1,589,844	0.77	0.43	-55,447	0.46	0.24	-344,039
Nebraska	72,506	549,755	0.61	0.39	-16,093	0.27	0.16	-58,510
Nevada	143,616	1,003,174	0.23	0.06	-25,156	0.62	0.27	-353,148
New Hampshire	9,374	384,706	0.87	0.54	-3,090	0.75	0.20	-214,531
New Jersey	7,128	1,019,092	1.00	0.99	-81	0.99	0.83	-158,687
New Mexico	109,260	383,873	0.14	0.02	-12,810	0.23	0.05	-68,753
New York	48,756	2,651,158	0.95	0.81	-6,868	0.70	0.30	-1,079,833
North Carolina	56,673	4,679,517	1.00	0.97	-1,535	0.99	0.63	-1,688,348
North Dakota	59,514	2,442,160	0.96	0.78	-10,620	0.20	0.15	-122,294
Ohio	54,736	715,219	0.62	0.68	2,796	0.38	0.26	-88,107

Continued next page

State	Total Stream miles	Total Wetland Acres	Stream miles regulated			Wetland acres regulated		
			Rapanos (%)	NWPR (%)	Difference NWPR - *Rapanos*	*Rapanos* (%)	NWPR (%)	Difference NWPR - *Rapanos*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Oklahoma	75,615	1,274,713	0.89	0.37	-39,658	0.89	0.47	-538,402
Oregon	102,984	1,803,096	0.80	0.43	-37,975	0.88	0.37	-915,645
Pennsylvania	51,477	588,835	0.91	0.77	-7,319	0.78	0.37	-239,928
Rhode Island	978	86,061	1.00	0.94	-61	0.99	0.33	-56,368
South Carolina	29,372	4,238,935	1.00	0.96	-1,019	0.98	0.55	-1,823,810
South Dakota	96,965	3,529,693	0.83	0.57	-24,992	0.35	0.22	-482,699
Tennessee	59,244	1,148,777	0.52	0.46	-3,425	0.56	0.47	-101,837
Texas	176,194	5,551,483	0.67	0.21	-82,033	0.72	0.35	-2,051,586
Utah	82,724	624,397	0.34	0.08	-21,435	0.41	0.14	-169,600
Vermont	7,100	287,628	0.89	0.46	-3,048	0.60	0.15	-128,195
Virginia	49,280	1,682,396	1.00	0.98	-715	0.99	0.86	-219,712
Washington	68,964	1,297,395	0.98	0.60	-26,192	0.97	0.52	-588,074
West Virginia	30,572	81,858	0.50	0.66	4,900	0.65	0.49	-13,097
Wisconsin	53,370	7,610,528	0.59	0.47	-6,227	0.10	0.06	-251,016
Wyoming	106,082	1,646,169	0.49	0.21	-29,080	0.45	0.16	-474,353

Table S10: Deregulation of Drinking Water Areas after NWPR. A 12-digit hydrologic unit code (HUC12) or subwatershed is the finest polygon delineation of watershed boundaries the US Geological Survey defines, corresponding to about 80,000 HUC12s. This table considers active 2019 community water systems (CWS). Column (1) shows the number of people with a CWS served by a HUC12 with at least one point of a given classification deregulated between *Rapanos* and NWPR, divided by the number of people in a CWS served by a HUC12 with at least one point. Column (2) counts number of systems rather than number of people. Column (3) shows the share of points of a given classification that are deregulated in the sample of HUC12's that serve as a drinking water intake. With points as an analogue for water resource miles, 30% of NHD or NWI points deregulated in drinking water systems indicates a deregulation of 30% of NHD and NWI resources that serve as a drinking water source. Column (4) shows the share of points weighted by population. Including people or systems downstream of any HUC12 with at least one deregulated point does not change qualitative results.

Deregulated share of	**Population**	**PWSs**	**Points serving PWSs**	**Population Weighted Points**
	(1)	(2)	(3)	(4)
A. Same subwatershed (HUC12).				
Deregulated points	0.91	0.85	0.20	0.26
Deregulated NHD or NWI points	0.80	0.69	0.30	0.32
Deregulated NHD points	0.32	0.20	0.21	0.19
Deregulated NWI points	0.80	0.68	0.30	0.32
Deregulated NWI wetlands points	0.55	0.43	0.32	0.46