Deep Learning Projects Jurisdiction of
New and Proposed Clean Water Act Regulation

Simon Greenhill[1,2]*† Brant J. Walker[3]*† Joseph S. Shapiro*[1,4,5]

Affiliations

[1] Department of Agricultural and Resource Economics, University of California, Berkeley, Berkeley, CA, USA.

[2]Global Policy Laboratory, Stanford Doerr School of Sustainability, Stanford University, Stanford, CA, USA.

[3] School of the Environment, Yale University, New Haven, CT, USA.

[4] Department of Economics, University of California, Berkeley, Berkeley, CA, USA.

[5] National Bureau of Economic Research, Cambridge, MA, USA.

*Corresponding authors. Emails: sgreenhill@berkeley.edu, brant.walker@yale.edu, joseph.shapiro@berkeley.edu

† These authors contributed equally to this work

Author information: S.G., B.W., and J.S.S. conceived the project. S.G and B.W. analyzed data. S.G., B.W., and J.S.S. interpreted results and wrote the paper.

Ethics declarations: S.G. and J.S.S. are inventors on a patent pending with serial number 63/513,464, submitted by UC Berkeley, which covers machine learning systems and methods for predicting Clean Water Act jurisdiction, and allows free use for academic, government, and other non-commercial users.

Classifications: Social Sciences; Economic Sciences

Keywords: environmental policy, Clean Water Act, machine learning

This PDF file includes:
Main Text
Figs. 1-4

Abstract

Projecting the effects of proposed policy reforms is challenging because no outcome data exist for regulations that governments have not yet implemented. We propose an ex ante deep learning framework that can project effects of proposed reforms by mapping outcomes observed under past regulations onto the legal criteria of proposed future policies (i.e., by "relabeling"). We apply this framework to study changes in jurisdiction of the US Clean Water Act (CWA). We compare our ex ante deep learning projection of jurisdiction under the Supreme Court's *Sackett* decision against widely used projections from domain experts. Ex ante machine learning generates exceptional performance improvements over the leading domain expert model that the US Environmental Protection Agency currently uses, with 65 times more accurate identification of jurisdictional sites. We also develop an ex post deep learning model trained with data after policy implementation. Ex post deep learning performs best. *Sackett* deregulates one-third of all previously regulated US waters, particularly floodplains and pristine fish habitats, totaling 700,000 deregulated stream miles and 17 million deregulated wetland acres. Deep learning can effectively project consequences of far-reaching regulatory reforms before they are implemented, when projections are both most uncertain and most useful.

Significance Statement

Evaluating proposed regulations before implementation is essential for effective policymaking. Analysts, however, cannot observe how untested policies affect outcomes, which makes it challenging to produce accurate evaluations of such policies. This paper introduces a deep learning framework that addresses this challenge by modifying outcomes observed under past regulations in ways that reflect proposed regulations. This strategy allows projections of the real-world impacts of a potential policy change before it is implemented. We use this approach to predict the effects of the Supreme Court's 2023 *Sackett* decision, which restricts jurisdiction of the US Clean Water Act, using only information available before policy implementation. Ex ante deep learning dramatically outperforms widely used models created by geophysical scientists in identifying both regulated and unregulated waters. Separately, we also provide the first ex post national assessment of *Sackett*, "one of the most impactful environmental decisions in the Court's history," and the current Clean Water Act rule. *Sackett* removes federal protection from roughly one-third of previously regulated streams and wetlands, including areas with important ecological functions. These results demonstrate that deep learning can dramatically improve analysis of policies that are proposed but not yet implemented.

Evaluating proposed policy reforms is a critical task. Such evaluations can shape how policymakers choose between alternative policy proposals and how firms, citizen groups, and other stakeholders adapt to policy change. Government, academic, and private sector analysts generate numerous such evaluations annually. The stakes are high—regulatory reforms can generate hundreds of billions of dollars in annual benefits, though also enormous costs (1).

Projecting effects of proposed policy reforms is challenging because such forecasts are made before a policy is implemented, when a policy's impact is most uncertain. Because no outcome data exist for proposed policies, forecasting their effects typically relies on domain experts like scientists, engineers, and economists. The challenge of evaluating proposed policies when the outcomes are not yet observed has led to the concern that the existing evaluation system is "broken, … largely based on faith, rather than evidence" (2).

We develop a methodology that provides one of the first deep learning projections of a proposed regulation's effects. Analysts increasingly use deep learning to interpret existing energy, environmental, financial, health, judicial, and labor market regulations ex post (3–6), though largely not to evaluate proposed reforms. To address the absence of data on outcomes under proposed policies, we take data on outcomes under past policies and change ("relabel") these outcomes in ways that characterize proposed rules. We then train a deep learning algorithm that predicts regulation under the proposed rules, as captured by the relabeled outcomes. We compare performance of this ex ante deep learning projection against published ex ante projections from domain experts that rely on geophysical models. Separately, we develop an ex post deep learning model to describe a policy's effects after it is implemented and outcomes are observed, so that we can compare ex ante against ex post analysis.

We apply this methodology to study recent and ongoing reforms to the 1972 US Clean Water Act (CWA), the cornerstone of federal water pollution control. The CWA restricts water pollution discharged to the "Waters of the United States" (WOTUS) but does not enumerate which streams and wetlands this phrase covers. To determine whether the CWA protects a site (e.g., a parcel where a developer hopes to build a factory), a developer can ask the Army Corps of Engineers (USACE) to evaluate the site and issue an Approved Jurisdictional Determination (AJD), indicating whether the CWA regulates the site. AJDs are the only legally binding decisions describing CWA jurisdiction at the site level and therefore provide the natural outcome for our empirical analysis. For jurisdictional sites, a developer may request a USACE wetland permit describing conditions required to comply with the CWA.

Many stakeholders argue that CWA jurisdiction and its recent reforms are costly and uncertain. Microsoft's President summarized these wetland permits in congressional testimony as the "number 1 challenge" in data center development (7). A legal expert described courts modifying the CWA as sometimes "flying blind" (8). Media describe the regulatory landscape as "hazy" and "chaos" (9). Our analysis highlights the potential of ex ante deep learning projections to reduce the uncertainty associated with large policy reforms in this setting. Our analysis also responds to a call by USEPA and USACE (10) for evidence on whether machine learning could provide an "appropriate alternative" to geophysical models as a tool for governments to project effects of proposed regulations.

In addition to implementing ex ante deep learning, we use 200,000 AJDs to develop and train the ex post Clean Water Act Analysis of Regulation (CLEAR) deep learning model. This provides the first ex post national quantitative analysis of jurisdictional coverage under the Supreme Court's *Sackett* ruling, "one of the most impactful environmental decisions in the Court's history" (11).

Compared to algorithmic analysis of earlier CWA regulation (6), our deep learning models study new questions including projecting effects of proposed regulations, analyzing *Sackett*, and studying floodplains, fish habitat quality, and other ecosystem services. Our deep learning models also implement methodological advances including the generation and use of synthetic training data, fine tuning models on each CWA rule, fusing image and tabular data, calibrating model scores, and choosing optimal decision thresholds (SI Appendix, Sections A.1, A.2, A.3, and A.7).

CWA Background

Over the past decade, CWA jurisdiction has changed repeatedly due to alternating administrative rules and Supreme Court decisions. "Regulatory ping pong" (12, 13) under the CWA—frequent and large changes in rules between administrations and courts—includes six rules in the last decade, plus other rules under discussion or implementation (14). In the Supreme Court's 2006 *Rapanos* case, Justice Kennedy's concurring opinion found that to be jurisdictional under the CWA, a stream, wetland, or other water body required a "significant nexus," i.e., a biological, physical, or chemical connection to traditional navigable waters. The 2016 Clean Water Rule (CWR) primarily clarified *Rapanos;* USEPA and USACE repealed the CWR in 2019. The 2020 Navigable Waters Protection Rule (NWPR) restricted jurisdiction to relatively permanent waters with a continuous surface water connection to traditional navigable waters. NWPR effectively excluded ephemeral streams and isolated wetlands. The 2023 Rule, litigated then enjoined in some areas, closely resembled *Rapanos*. *Sackett* required jurisdictional waters to have a continuous surface water connection to traditional navigable waters and excluded certain wetlands separated from navigable waters by barriers. Due to litigation, in September 2023, USEPA implemented two versions of *Sackett* in different states, which our analysis combines given their similarity. In March 2025, USEPA and USACE issued revised *Sackett* guidance, prompting extensive debate, including 46,042 public comments (15). In November 2025, USEPA and USACE proposed a rule to further limit CWA jurisdiction. The PERMIT Act, which the US House passed in December 2025 with bipartisan support, rewrites the CWA to resemble *Sackett*, though further excludes groundwater (16).

Predictive Models of CWA Jurisdiction

We consider a series of models predicting which water resources each CWA rule regulates. As a benchmark, we compare all models' performance against the naïve prediction that no sites are jurisdictional.

As in climate science, a "projection" considers an assumed future policy scenario, such as a proposed CWA rule, and quantitatively describes its effects. As in machine learning, a "prediction" reflects a model's assessment of what a rule regulates (17).

Geophysical Models. We consider two geophysical models that are widely used by domain experts, which both assume that water resources with certain attributes in existing stream and wetland maps define CWA jurisdiction. Domain experts choose which characteristics in the maps define jurisdiction. USEPA and USACE once described this type of geophysical model as "highly unreliable … based on stream and wetland datasets that were not created for regulatory purposes and have significant limitations…" (18). Nonetheless, such geophysical models are prominent in research (19–21), underpin prominent Supreme Court briefs (22), guide current USEPA and USACE planning (23), and receive extensive media attention (24–26). SI Appendix Section A.1 provides details.

Model 1 (Wetness Geophysical Model). The "Wetness" model (19) assumes that non-tidal wetlands that the National Wetlands Inventory (NWI) lists as not inundated a certain share of the year lose jurisdiction under *Sackett*.

Model 2 (Connected Geophysical Model). The "Connected" model (20, 21) assumes that wetlands in the NWI that intersect a perennial or intermittent stream in the National Hydrography Dataset (NHD) are jurisdictional.

NWI and NHD are leading national maps of wetlands and streams, though both have well-documented errors of inclusion and exclusion (18, 27–29).

Deep Learning Models. We also consider approaches that take AJDs from past CWA rules and train a deep learning model to predict their jurisdictional status. We then use the trained algorithm to predict jurisdiction at any US location under each CWA rule.

Model 3 (Ex Ante Deep Learning Model). Our ex ante deep learning model predicts jurisdictional outcomes under *Sackett* using only data and knowledge available before *Sackett* implementation. For model training, we take AJDs from NWPR, a rule preceding *Sackett*, and change (i.e., "relabel") the outcomes from jurisdictional to non-jurisdictional for the two categories of waters which lost protection between NWPR and *Sackett*—wetlands separated from jurisdictional waters by artificial structures or natural features (SI Appendix, Table S2 and Section A.1). We identify these two categories of waters by reading the *Sackett* majority opinion. We formalized the relabeling in a June 2023 external email and presentation, before USEPA announced its *Sackett* rule or USACE began implementing it.

Model 4 (Ex Post Deep Learning Model). Our ex post deep learning model predicts jurisdictional outcomes under each CWA rule – *Sackett*, NWPR, *Rapanos*, and CWR – by training on AJDs from all rules.

Both the ex ante and ex post deep learning models begin from a common deep learning framework—a ResNet-18 backbone (30) pre-trained on ImageNet (31). The input layers we use to predict the AJDs include color and near infrared aerial imagery, water resource maps, elevation data, local climate and weather information, soil characteristics, land cover maps, and ecoregions (SI Appendix, Section A.5 and Table S15). We also include tabular data on location's state, USACE district, distance to USACE headquarters, and on the CWA rule under which the location is evaluated (SI Appendix, Section A.2). For each model, we pre-train on AJDs describing many CWA rules and then fine-tune the algorithm on only the CWA rule of interest (SI Appendix, Section A.1).

Measuring Model Performance

Deep learning models output a raw model score for each site in [0,1]. We use isotonic regression to translate this raw score into a calibrated jurisdictional probability, which represents the model's estimate of the probability that the site is jurisdictional (SI Appendix, Section A.7).

Generating binary classifications ("jurisdictional" versus "not jurisdictional") from deep learning models requires a decision threshold; deep learning models predict that sites with calibrated probabilities above this threshold are jurisdictional and sites with probabilities below this threshold are not jurisdictional. Analysis could default to a decision threshold of 0.5, which would imply that any site with a calibrated probability score above 0.5 is predicted to be jurisdictional. However, a benefit of utilizing a probabilistic model such as deep learning to create projections is the flexibility to choose the threshold that maximizes model performance on a given performance metric. For example, one threshold may minimize mean absolute error across the US, while another threshold may maximize accuracy.

We divide the AJDs into three spatially disjoint groups for model training, development, and evaluation. The deep learning models use the training set (80% of AJDs) to learn patterns in the data. We use the validation set (10% of AJDs) to tune model parameters. For all models, we use a held-out test set (10% of sample) to calculate the model performance statistics this paper reports. Our use of the test set helps avoid overfitting the model to the validation set and thereby inflating performance metrics (SI Appendix, Section A.2). The data have class imbalance, since 80.3% of *Sackett* AJDs are non-jurisdictional. A naïve benchmark that predicts zero jurisdiction anywhere therefore achieves accuracy of 80.3%.

While predictive performance can be measured using a range, we focus on the widely-used area under the receiver operating characteristic curve (AUC) (32). The AUC is robust to class imbalance because it evaluates a model's ability to rank positive cases above negative cases, thereby using the full ranking of predicted jurisdictional scores rather than measuring performance at one chosen decision threshold. This matters because the use of a decision threshold treats sites with the same binary prediction identically, even if the sites have different calibrated jurisdictional probabilities (e.g., if one site has 65% probability of jurisdiction and the other has 99%, and both exceed the binary decision threshold), though a user (e.g., a developer or regulator) may see these predictions differently.

We also report other model performance metrics besides the AUC that rely on binary decision thresholds. We report precision and recall, given concern with false positives and false negatives, as well as their harmonic mean (the F1 Score); accuracy, given its simple interpretation and common use; and mean absolute error (MAE) nationally, given usefulness for stakeholders. Fig. 1 and SI Appendix, Table S1, define several of these metrics. As with other model parameters, we choose binary decision thresholds using the validation set (SI Appendix, Section A.7).

The original analysis developing the Wetness geophysical model (19) has two characteristics worth discussing. First, it projects effects of eight scenarios based on different assumptions about the share of the year a wetland must be inundated to be jurisdictional, but it does not distinguish which of the eight scenarios will be enacted. The Wetness model predicts that between 19% and 91% of non-tidal wetlands lose protection under *Sackett*, a range wide enough to be "bogged down in mystery" (33). The wetness scenarios range widely because the results depend on assumptions about how USACE interprets *Sackett*. Our analysis of the Wetness model focuses on the median scenario for simplicity, though does report results for all wetness scenarios. The median scenario has the best performance in the validation set among all wetness scenarios.

Second, the original Wetness model (19) generates predictions for a narrow subset of US waters—non-vegetated, non-anthropogenically influenced, shallow water non-tidal wetlands connected to jurisdictional streams and rivers. We find that these areas only account for 1.2% of *Sackett* AJDs. This restricted availability of the Wetness geophysical model limits its applicability nationally. We therefore also report results for three separate samples of *Sackett* AJDs (SI Appendix, Section A.1).

Model Results

A naïve benchmark, which assumes that no sites are jurisdictional, has poor model performance, with AUC of 0.500 and F1 Score of 0.000 (Fig. 1 and SI Appendix, Table S1A).

The Wetness geophysical model does not uniformly improve model performance over this naïve benchmark (Fig. 1 and SI Appendix, Table S1B). For example, the Wetness model has an AUC of 0.498, just below the naïve benchmark. The wetness model only correctly identifies 1 in 250 sites that USACE classifies as jurisdictional (i.e., it has recall of 0.004). In part this happens because wetness categories have a noisy relationship to jurisdiction and only focus on non-tidal wetlands (SI Appendix, Fig. S4).

The Connected geophysical model improves slightly, with an AUC of 0.512 (Fig. 1 and SI Appendix, Table S1B). The Connected model improves performance for the sites it predicts as jurisdictional (i.e., it has high precision). It misses many waters that USACE identifies as jurisdictional (i.e., it has low recall). The Connected model performs somewhat poorly because many jurisdictional AJDs are not in national maps of streams or wetlands (NHD or NWI), and because the Connected model's geophysical criteria incorrectly exclude many jurisdictional streams and wetlands.

Ex ante deep learning substantially outperforms the geophysical models on most performance metrics (Fig. 1 and SI Appendix, Table S1C). Ex ante deep learning has an AUC of 0.693, 0.181 higher than either geophysical model. This represents an enormous performance improvement by standards common in applied machine learning, where even AUC improvements of 0.05 are considered to be substantial (34). In all eight scenarios the Wetness model examines, ex ante and ex post deep learning outperform the Wetness model in AUC and most other performance metrics (SI Appendix, Table S4). Compared to the Wetness model, ex ante deep learning is sixty-five times more likely to identify jurisdictional sites (higher recall) and has forty-seven times better performance on jurisdictional sites (F1 score).

Ex post deep learning has the strongest performance of all models (Fig. 1 and SI Appendix, Table S1D). It substantially outperforms both geophysical models on all metrics. It also outperforms ex ante deep learning on some but not all metrics, and by smaller margins. Ex ante and ex post deep learning have similar AUC. By this important metric, access to post-implementation data does not materially improve performance relative to the ex ante model. The national MAE of 0.001 from ex post deep learning means it almost perfectly projects the mean national jurisdiction of *Sackett*, while other models have MAE of 0.07 to 0.19, indicating they have some bias in projecting overall regulatory stringency of *Sackett*. Ex post deep learning achieves AUC above 0.80 on the other

CWA rules (NWPR, CWR, and *Rapanos*), exceeding its levels for *Sackett* (SI Appendix, Table S5A).

Describing Jurisdiction: *Sackett*

To understand patterns of CWA jurisdiction, we calculate each model's prediction at 4 million randomly chosen points across the contiguous US. This subsection focuses on predictions from the ex post deep learning model, since it has the strongest performance, for these 4 million points and subsets of interest. We compare against the ex ante geophysical and deep learning models to clarify their differences in substantive conclusions.

The ex post deep learning model calculates that *Sackett* regulates 11.5% of the contiguous US area, including 25% of stream miles and 28% of wetland acres (Fig. 2 and SI Appendix, Table S3). *Sackett* deregulates floodplains and other areas with important ecosystem services, many of which prior rules regulated (SI Appendix, Fig. S2 and Tables S7E and S9).

Geophysical models rely on stream and wetland maps like NHD and NWI to make predictions. Ex post deep learning indicates that 11.3% of areas not in NHD and 8.7% of areas not in NWI's palustrine wetlands are jurisdictional (SI Appendix, Table S3B and C). This further underscores the limits of geophysical models, which substantially rely on one or two stream and wetland maps like NHD and NWI without directly using AJDs.

Compared to ex post deep learning, geophysical models substantially underestimate CWA jurisdiction, while ex ante deep learning is much closer to the ex post estimates (Fig. 2, SI Appendix, Table S3). Geophysical models predict that *Sackett* regulates less than 3 percent of the US, though alternative wetness scenarios range widely (SI Appendix, Table S6). The Connected geophysical model predicts that the CWA regulates only 0.1% of the parts of the US that are not in NWI or NHD, so it performs especially poorly in these areas. The ex ante deep learning model projects that *Sackett* regulates 13.4% of the contiguous US, much closer to the ex post 11.5%. Similarly, ex post deep learning projects that the CWA protects 27.9% of wetlands, which is in the ballpark of the ex ante deep learning projection of 31.4%, though nowhere near the geophysical model projections of 10.9% to 16.5% (SI Appendix, Table S3C).

Describing Jurisdiction: All CWA Rules

Ex post deep learning shows that *Sackett* regulates fewer waters than any previous rule (Fig. 3 and SI Appendix, Table S7). *Rapanos* regulates 46% of stream miles, 41% of wetland acres, and 18% of contiguous US area. Compared to *Rapanos*, *Sackett* deregulates one-third of regulated water resources and 28% of regulated floodplains.* This amounts to over 700,000 deregulated stream miles and 19 million deregulated wetland acres. *Sackett* deregulates the most wetland acres in Florida and Michigan (SI Appendix, Table S8). Sackett also deregulates 28% of regulated floodplains, potentially encouraging development in these areas, which is important given rising national flood damages and growing extreme weather risk due to climate change.

---

* Our estimates of regulation under previous rules such as NWPR exceed prior algorithmic estimates (6), partly since we average calibrated probabilities while prior work averages binary jurisdictional predictions (SI Appendix, Section B.2). Calibrated probabilities represent the probability of jurisdiction. Thus, averaging these probabilities as we do here, rather than their rounded values as in (6), best describes the share of area that is jurisdictional.

NWPR and *Sackett* both have a basis in Justice Scalia's *Rapanos* opinion, but the differences are so far largely unquantified. Ex post deep learning finds that *Sackett* regulates systematically less than NWPR, including deregulating a fifth of wetland acres protected under NWPR (Fig. 3 and SI Appendix, Table S7).

Fig. 3 graphs the "regulatory ping pong" of recent CWA regulation. Jurisdiction fluctuated between 2018 and 2020 due to differences between CWR and *Rapanos*. The share of streams and wetlands regulated fell by about 15% in 2020 under NWPR and returned to broader jurisdiction in late 2021. Jurisdiction declined by around a third in late 2023, under *Sackett*.

Maps reveal enormous spatial differences across rules (Fig. 4 and SI Appendix, Fig. S5). Ex post deep learning finds that compared to *Rapanos*, *Sackett* deregulates isolated wetlands in coastal and inland areas, ephemeral streams across the arid West, and streams and wetlands in almost every state (SI Appendix, Table S8). Compared to NWPR, *Sackett* primarily deregulates wetlands along the East Coast and in some areas of the Pacific Northwest, but changes jurisdiction little across the Arid West (SI Appendix, Fig. S5C). Ex ante and ex post deep learning models predict qualitatively similar spatial patterns.

Case studies highlight local differences across rules and predictions (SI Appendix Fig. S1, S6). In wetland-abundant regions like Michigan's Upper Peninsula and the North Carolina coast, *Sackett* regulates fewer isolated wetlands and small water bodies than *Rapanos*. In drier regions, *Sackett* and NWPR deregulate ephemeral streams. The Wetness geophysical model has no predictions in most of these areas, given its restriction to a narrow set of non-tidal wetlands. Ex ante deep learning captures spatial patterns in the ex post data much more effectively than the geophysical models do.

Wetlands support ecosystem services including flood mitigation and water filtration, and support CWA goals of decreasing water pollution and improving water-based recreation, including fishing. *Sackett* deregulates areas important to all of these ecological functions (SI Appendix, Fig. S2, Table S9). For example, in areas used for drinking water sources, *Sackett* has 10% lower probability of regulation than *Rapanos*. In "impaired" areas where a large share of waters is too polluted to support intended uses, *Sackett* also has 10% lower probability of regulation than *Rapanos*.

Discussion

Many groups may value accurate projections of the effects of proposed environmental regulations. Developers and industrial firms can use such projections to improve regulatory compliance and guide site and investment decisions. Staff at government agencies like USEPA and USACE, the Congressional Budget Office, and state Wetland Boards can use such projections to help evaluate proposed regulations and provide a decision support tool for implementing existing regulations. Judges can use such projections to understand consequences of alternative interpretations of statutes. Environmental restoration firms can use such projections to evaluate where investment in restoring natural resources (e.g., wetland mitigation banks) is most needed. Environmental organizations can use such projections to guide public discussion of prospective environmental reforms.

Ex ante deep learning can provide high-quality projections to support such needs. Ex ante deep learning helps address a critical problem of policy analysis—projecting effects of proposed policies before implementation—a time period when analysis is both most uncertain and most useful.

Our analysis of recent CWA reforms finds that ex ante deep learning far outperforms expert geophysical projections on most measures of model performance. Expert geophysical projections provide marginal improvements over a naïve benchmark. Ex post deep learning has the strongest performance and documents enormous decreases in wetland and stream jurisdiction under *Sackett* compared to all previous CWA rules.

Future work can further clarify the potential contribution of deep learning to projecting effects of other reforms. Recent or ongoing reforms to wetland protection in Chile, China, the EU, Japan, and elsewhere may provide opportunities for related analysis (35–40). Proposed reforms to the National Environmental Policy Act, Clean Air Act, Safe Drinking Water Act, and other landmark environmental US statutes may also benefit from deep learning projections. The frequency of regulatory reforms in financial, labor market, and other non-environmental domains provides many opportunities to explore related approaches.

In any setting, the relative strength of ex ante deep learning versus domain experts may depend on the extent to which relabeling effectively characterizes the policy reform. More precise descriptions of proposed reforms, and descriptions of reforms which overlap with characteristics of prior policies, may improve the performance of ex ante deep learning. A regulation's impacts ultimately depend on agencies' capacity, agencies' evolving interpretations of statutes, agencies' willingness to enforce policy changes, and regulated entities' compliance. One interpretation is that in our setting, regulatory agencies implement policy reforms in a way that can be effectively projected using a flexible interpretation of past policies.

Another intriguing question for future work involves potential combinations of ex ante deep learning with geophysical frameworks. Some prospective reforms could benefit from taking observed outcomes under past rules, using domain expertise or geophysical models to determine which outcomes within certain categories change under a proposed rule, then training an ex ante deep learning model on the resulting relabeled data. For example, USEPA and USACE released a draft CWA rule in November 2025. USEPA and USACE propose using geophysical models to project effects of this rule and dismiss the use of AJDs. Our results raise the possibility that using geophysical models to relabel AJDs from past rules and training an ex ante deep learning model to describe relabeled AJDs could substantially outperform the use of geophysical models alone.

While we project effects of regulation as implemented, a related and important question asks whether agency interpretation of a regulation fits with the intent of a law as written. This represents another area almost exclusively analyzed by domain experts, and where the potential contribution of deep learning remains unknown. "Human-in-the-loop" frameworks, where domain experts and algorithms collaboratively improve an evaluation system's capabilities, may also provide a useful avenue to compare the intent of a law as written against an agency's interpretation of it.

Acknowledgements

Participants in USEPA Regions 3 and 9, USEPA NCEE, and Environmental Law Institute presentations, D. Beaudette, J. Blumenstock, S. Burley, N. Cabrera, T. Carleton, C. Costello, J. Devine, H. Druckenmiller, P. Hann, S. Hsiang, C. Kling, M. Kotchen, D. Lobell, J. Murphy, D. Owen, B. Payne, and S. Wang provided comments. We thank A. Barker, H. Menth, M. Olander, and S. Shoemaker for research assistance.

This material is based upon work supported by the following organizations and grants: the Bakar Fellows Program; the Bakar Ignite Scholars Program; the Giannini Foundation for Agricultural Economics; NIH 1R01AG079914-01; USDA 2024-67023-42616; NSF GRFP DGE-2139841.

Data availability: Data to reproduce all results in the paper will be posted on a public internet repository before publication.

Code availability: Code to reproduce all results in the paper will be posted on a public internet repository before publication.

References

1. Office of Management and Budget, “2013 Draft report to Congress on the benefits and costs of federal regulation and unfunded mandates on state, local, and tribal entities,” OMB, Washington, DC (2013), pp. 14, 18, 19. Available at: https://obamawhitehouse.archives.gov/sites/default/files/omb/inforeg/2013_cb/draft_2013_cost_benefit_report.pdf [Accessed 20 January 2026].

2. M. Greenstone, “Toward a Culture of Persistent Regulatory Experimentation and Evaluation” in *New Perspectives on Regulation*, (The Tobin Project, Inc., 2009), pp. 111–125.

3. C. Coglianese, D. Lehr, Regulating by Robot: Administrative Decision Making in the Machine-Learning Era. *Geo. L.J.* **105**, 1147 (2016).

4. D. F. Engstrom, D. E. Ho, C. M. Sharkey, M.-F. Cuéllar, Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies. [Preprint] (2020). Available at: https://papers.ssrn.com/abstract=3551505 [Accessed 16 December 2025].

5. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, Human Decisions and Machine Predictions. *The Quarterly Journal of Economics* **133**, 237–293 (2018).

6. S. Greenhill, *et al.*, Machine learning predicts which rivers, streams, and wetlands the Clean Water Act regulates. *Science* **383**, 406–412 (2024).

7. A. Halverson, Microsoft’s Brad Smith, other tech execs give Congress AI wish list. *The Seattle Times* (2025). Available at: https://www.seattletimes.com/business/microsofts-brad-smith-other-tech-execs-give-congress-ai-wish-list/ [Accessed 22 August 2025].

8. D. Owen, Mapping the new Clean Water Act. *Science* **385**, 1414–1416 (2024).

9. E. A. Crunden, Post-Sackett, chaos erupts for wetlands oversight. *E&E News by POLITICO* (2023). Available at: https://www.eenews.net/articles/post-sackett-chaos-erupts-for-wetlands-oversight/ [Accessed 10 September 2025].

10. US EPA and US ACE, Regulatory Impact Analysis for the Proposed Updated Definition of Waters of the United States Rule. (2025). Available at: https://www.epa.gov/system/files/documents/2025-11/11132.1-01-ow_wotus_nprm_ria_20251110_508.pdf [Accessed 28 January 2026].

11. D. Owen, SACKETT V. ENVIRONMENTAL PROTECTION AGENCY AND THE RULES OF STATUTORY MISINTERPRETATION. *Harvard Environmental Law Review* **48**, 333–368 (2024).

12. S. K. Cramer, RESTORING STATES' RIGHTS & ADHERING TO COOPERATIVE FEDERALISM IN ENVIRONMENTAL POLICY. *Harvard Journal of Law & Public Policy* **45**, 481–502 (2022).

13. M. Phillis, EPA likely to move to further limit federal protections for wetlands. *AP News* (2025). Available at: https://apnews.com/article/epa-trump-clean-water-act-wetlands-protection-021ff0aacd77b91c4b0e70bc5caedd06 [Accessed 20 March 2025].

14. R. C. Gardner, *Waters of the United States: POTUS, SCOTUS, WOTUS, and the Politics of a National Resource*, 1st ed (Island Press, 2024).

15. US EPA, Implementation of the Definition of Waters of the United States. (2025). Available at: https://www.regulations.gov/document/EPA-HQ-OW-2025-0093-0001 [Accessed 25 August 2025].

16. M. Willson, Clean Water Act permitting bill clears the House. *E&E News by POLITICO* (2025). Available at: https://www.eenews.net/articles/clean-water-act-permitting-bill-clears-the-house/ [Accessed 7 January 2026].

17. D. M. Katz, M. J. Bommarito, J. Blackman, A general approach for predicting the behavior of the Supreme Court of the United States. *PLoS ONE* **12**, e0174698 (2017).

18. US EPA, Navigable Waters Protection Rule - Fact Sheet. (2021). Available at: https://19january2021snapshot.epa.gov/sites/static/files/2020-01/documents/nwpr_fact_sheet_-_mapping.pdf [Accessed 28 January 2026].

19. A. C. Gold, How wet must a wetland be to have federal protections in post-Sackett US? *Science* **385**, 1450–1453 (2024).

20. R. Meyer, A. Robertson, Clean Water Rule spatial analysis: A GIS-based scenario model for comparative analysis of the potential spatial extent of jurisdictional and non-jurisdictional wetlands. (2019). Available at: https://downloads.regulations.gov/EPA-HQ-OW-2021-0602-0752/attachment_6.pdf [Accessed 22 August 2025].

21. J. Devine, S. Lee, M. McKinzie, Mapping Destruction: Using GIS Modeling to Show the Disastrous Impacts of Sackett v. EPA On America's Wetlands. (2025). Available at: https://www.nrdc.org/sites/default/files/2025-03/Wetlands_Report_R_25-03-B_05_locked.pdf [Accessed 1 July 2025].

22. 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). Available at: https://www.supremecourt.gov/docket/docketfiles/html/public/21-454.html [Accessed 26 August 2025].

23. US EPA and US ACE, Economic Analysis of the EPA-Army Clean Water Rule. (2015). Available at: https://www.epa.gov/sites/default/files/2015-06/documents/508-final_clean_water_rule_economic_analysis_5-20-15.pdf [Accessed 7 January 2026].

24. C. Wood, C. O'Mara, D. Hall, Opinion | Trump Weakens the Nation's Clean Water Efforts. *The New York Times* (2020). Available at: https://www.nytimes.com/2020/02/10/opinion/clean-water-act-trump.html [Accessed 26 August 2025].

25. M. Joselow, A. Ellerbeck, Analysis | The Supreme Court takes on the nation's bedrock environmental laws. *The Washington Post* (2022). Available at: https://www.washingtonpost.com/politics/2022/01/25/supreme-court-takes-nation-bedrock-environmental-laws/ [Accessed 24 August 2025].

26. N. Rott, SCOTUS and wetlands. *NPR* (2024). Available at: https://www.npr.org/2024/09/26/nx-s1-5124825/a-2023-supreme-court-ruling-could-strip-u-s-wetlands-of-federal-protections [Accessed 24 August 2025].

27. M. J. Metes, *et al.*, Ephemeral Stream Network Extraction from Lidar-Derived Elevation and Topographic Attributes in Urban and Forested Landscapes. *JAWRA Journal of the American Water Resources Association* **58**, 547–565 (2022).

28. A. J. Elmore, J. P. Julian, S. M. Guinn, M. C. Fitzpatrick, Potential Stream Density in Mid-Atlantic U.S. Watersheds. *PLOS ONE* **8**, e74819 (2013).

29. J. W. Matthews, D. Skultety, B. Zercher, M. P. Ward, T. J. Benson, Field Verification of Original and Updated National Wetlands Inventory Maps in three Metropolitan Areas in Illinois, USA. *Wetlands* **36**, 1155–1165 (2016).

30. K. He, X. Zhang, S. Ren, J. Sun, Identity Mappings in Deep Residual Networks in *Computer Vision - ECCV 2016*, B. Leibe, J. Matas, N. Sebe, M. Welling, Eds. (Springer, 2016), pp. 630–645.

31. J. Deng, *et al.*, ImageNet: A large-scale hierarchical image database in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, (IEEE, 2009), pp. 248–255.

32. T. Fawcett, An introduction to ROC analysis. *Pattern Recognition Letters* **27**, 861–874 (2006).

33. D. Jackson, Wetland Protections Remain Bogged Down in Mystery. *The Equation* (2024). Available at: https://blog.ucs.org/derrick-jackson/wetland-protections-remain-bogged-down-in-mystery/ [Accessed 24 August 2025].

34. P. Baldi, P. Sadowski, D. Whiteson, Searching for exotic particles in high-energy physics with deep learning. *Nat Commun* **5**, 4308 (2014).

35. E. Hamman, Wetland Restoration in Japan: What's Law Got to Do with It? *NVJS* **11**, 47–73 (2019).

36. D. De Boer, J. Boya, New Wetland Protection Law Will Boost China's Biodiversity and Climate. *CCICED* (2022). Available at: https://cciced.eco/ecological-progress/new-wetland-protection-law-will-boost-chinas-biodiversity-and-climate-protection/ [Accessed 26 August 2025].

37. R. Nordbeck, K. Hogl, L. Schaller, The integration of peatlands into the EU Common Agricultural Policy: Recent progress and remaining challenges. *Environmental Science & Policy* **169**, 104077 (2025).

38. Province of Alberta, Water Act. (2024). Available at: https://kings-printer.alberta.ca/documents/Acts/w03.pdf [Accessed 26 August 2025].

39. P. Moschella, *et al.*, "Regulation and Protection of Urban Wetlands: A Comparative Analysis in Chile, Colombia, and Peru" in *Urban Wetlands in Latin America*, Sustainable Development Goals Series., C. Rojas Quezada, Ed. (Springer Nature Switzerland, 2024), pp. 19–31.

40. R. Quezada, "Urban Wetlands Protection Law in Chile. A Successful Tool for Urban Planning" in *Urban Wetlands in Latin America*, Sustainable Development Goals Series., (Springer Nature Switzerland, 2024), pp. 49–56.

41. S. P. Mulligan, Wading Into the "Waters of the United States." *Congressional Research Service* (2018). Available at: https://sgp.fas.org/crs/misc/LSB10236.pdf [Accessed 22 August 2025].

42. National Flood Insurance Program, Flood Insurance Data | The National Flood Insurance Program for Agents. Available at: https://agents.floodsmart.gov/flood-maps-and-data/flood-insurance-data [Accessed 27 August 2025].

43. US EPA, RPS Indicator Database. (2025). Available at: https://www.epa.gov/rps/data-downloads [Accessed 15 December 2025].

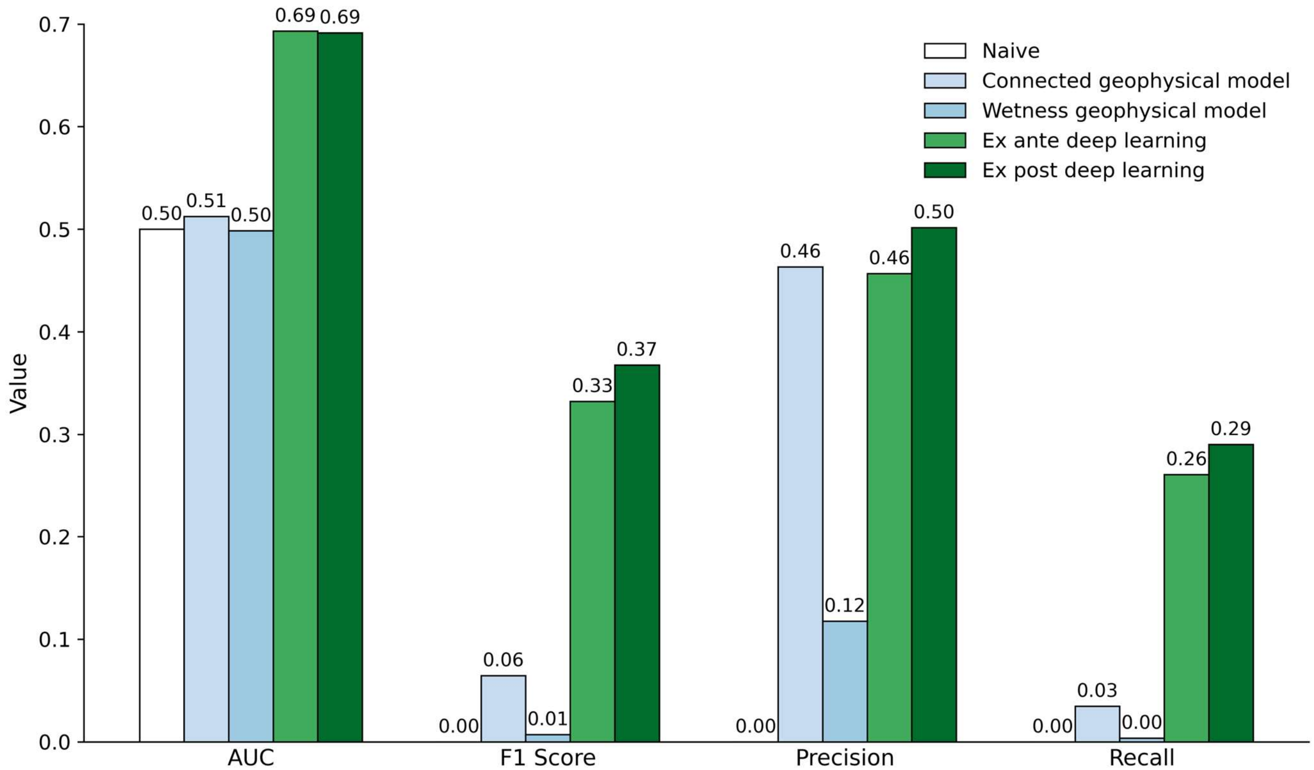


Fig. 1: Geophysical models have similar or somewhat better performance than the naïve benchmark, ex ante deep learning does much better, and ex post deep learning has the strongest model performance.

Each bar describes the performance of a separate model for *Sackett* jurisdiction, according to the performance metric listed along the x-axis. The naïve benchmark (white bar) predicts that no location is jurisdictional. The Connected geophysical model (light blue bar) defines points as jurisdictional if they fall within a potentially regulated National Wetlands Inventory (NWI) polygon that connects with a perennial or intermittent National Hydrography Dataset (NHD) flowline. The Wetness geophysical model (medium blue bar) describes the median Wetness scenario (19), "seasonally flooded." The ex ante deep learning model (medium green bar) describes a projection of *Sackett* using ex ante data. The ex post deep learning model (dark green bar) describes a deep learning model that uses ex post *Sackett* Approved Jurisdictional Determination (AJD) data. AUC is the area under the receiver operating curve. F1 Score equals the harmonic mean of precision and recall. Precision equals TP / (TP + FP), where TP is the count of true positive predictions and FP is the count of false positive predictions. Precision represents the accuracy of all jurisdictional predictions. Recall equals TP / (TP + FN), where FN is the count of false negative predictions. Recall represents the share of all true jurisdictional waters predicted as jurisdictional. Precision and recall are undefined if a model makes no positive predictions. F1 Score, precision, and recall performance use the optimal threshold for F1 performance, chosen using the validation set.

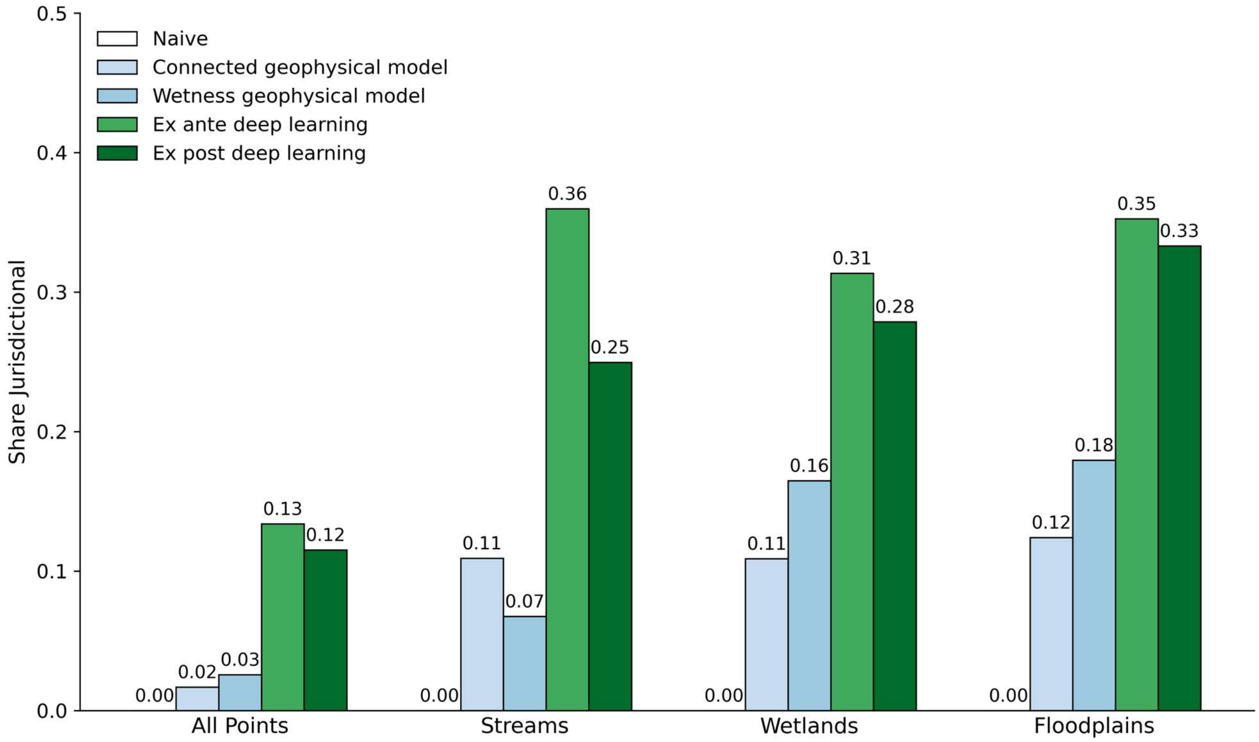


Fig. 2: Ex ante and ex post deep learning project that *Sackett* regulates a fourth to a third of water resources; geophysical models substantially under-predict *Sackett* regulation.

Each bar describes the share of points regulated under separate models for *Sackett* jurisdiction. Ex ante and ex post deep learning average calibrated probabilities. The naïve benchmark (white bar) predicts that no location is jurisdictional. The Connected geophysical model (light blue bar) defines points as jurisdictional if they fall within a potentially regulated National Wetlands Inventory (NWI) polygon that connects with a perennial or intermittent National Hydrography Dataset (NHD) flowline. The Wetness geophysical model (medium blue bar) describes the median Wetness scenario (19), "seasonally flooded." SI Appendix, Table S4 describes other wetness scenarios. The ex ante deep learning model (medium green bar) describes a projection of *Sackett* using ex ante data. The ex post deep learning model (dark green bar) describes a deep learning model that uses ex post *Sackett* AJD data. Streams include areas with 5 meters of perennial, intermittent, and ephemeral flowline feature codes (fcodes) 46006, 46003, 46007 in the NHD. Wetlands include areas with 5 meters of NWI wetlands. Floodplains are areas within floodplains from the National Flood Hazard Layer.

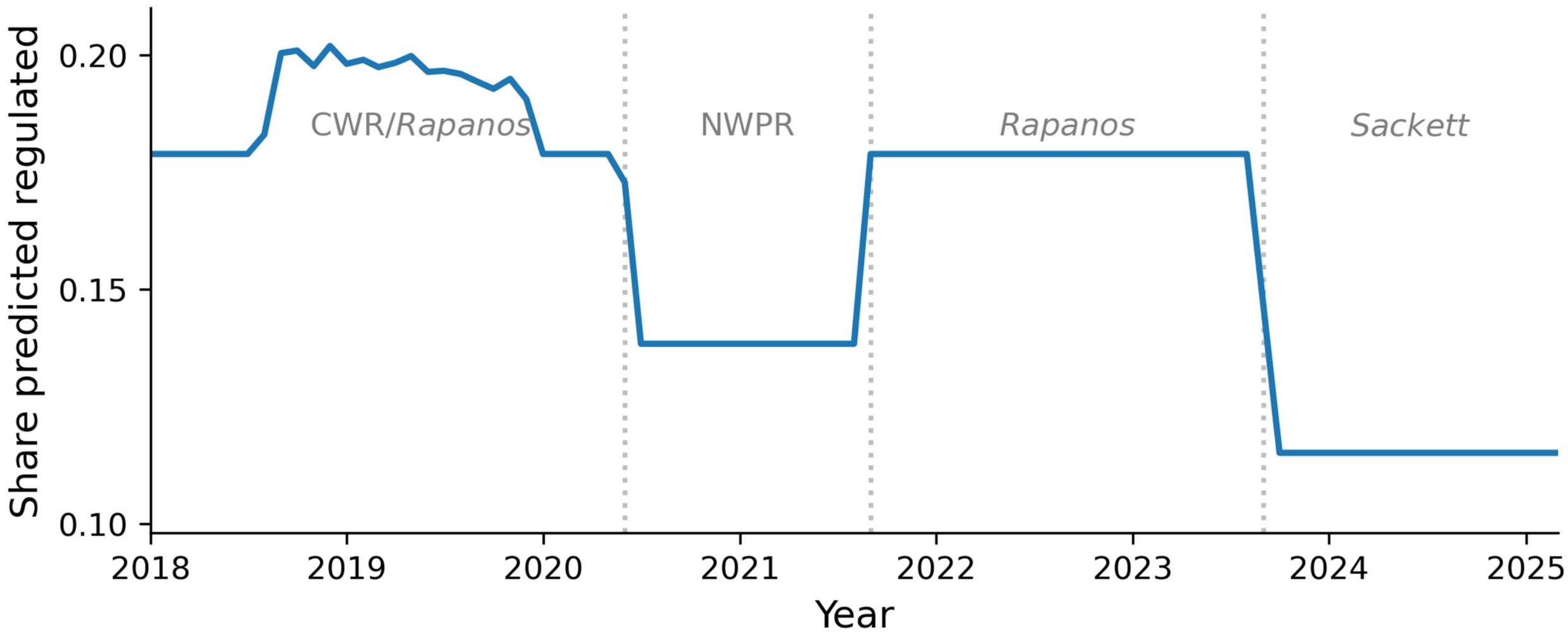


Fig. 3. Large variation in Clean Water Act jurisdiction across rules creates "regulatory ping pong." The graph shows the share of points within 5 meters of stream or wetland (National Hydrography Dataset or National Wetland Inventory) features that are predicted as jurisdictional, by month, using the ex post deep learning model. To determine which Clean Water Act rule applied in each month, we use the rule used to decide a majority of Approved Jurisdictional Determinations within each state in each month, calculate statistics by state, and average across states, weighting by the number of points in the state. Between January 2018 and August 2019, some states implemented the Clean Water Rule and others implemented *Rapanos*, due to litigation. Fluctuations in the share of locations regulated during this period reflect state-level changes in rules applied due to stays on the Clean Water Rule's implementation (41). *Rapanos* applied from September 2019 to May 2020. The Navigable Waters Protection Rule applied from June 2020 to August 2021. *Rapanos* (defined to include the 2023 rule) applied again from September 2021 to August 2023. *Sackett* applied from September 2023 onwards. The US Environmental Protection Agency and US Army Corps of Engineers are implementing two versions of *Sackett* in different states due to pending litigation, which we pool given their similarity.

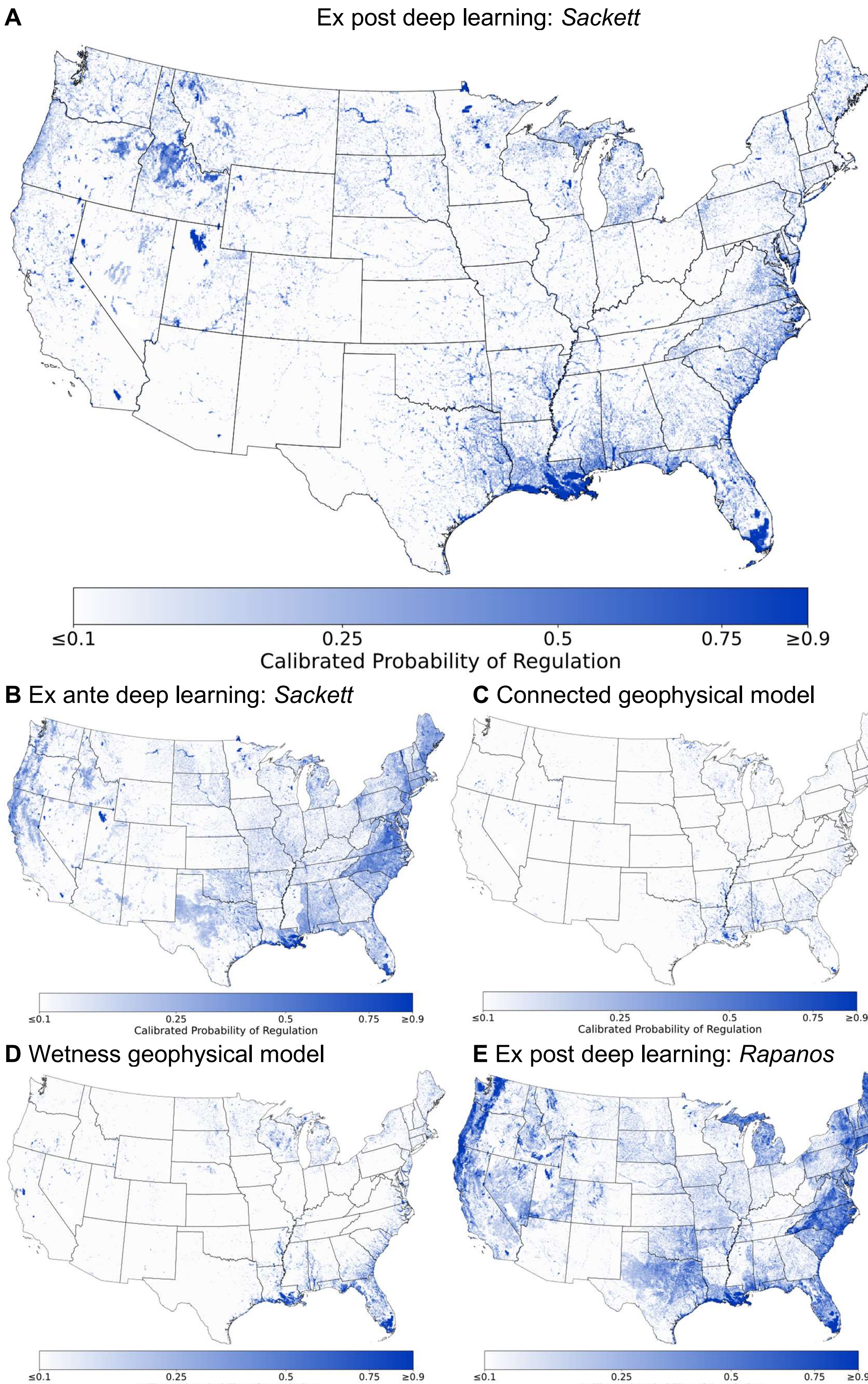

A
Ex post deep learning: *Sackett*
≤0.1
0.25
0.5
0.75
≥0.9
Calibrated Probability of Regulation
B Ex ante deep learning: *Sackett*
≤0.1
0.25
0.5
0.75
≥0.9
Calibrated Probability of Regulation
C Connected geophysical model
≤0.1
0.25
0.5
0.75
≥0.9
Calibrated Probability of Regulation
D Wetness geophysical model
≤0.1
0.25
0.5
0.75
≥0.9
Calibrated Probability of Regulation
E Ex post deep learning: *Rapanos*
≤0.1
0.25
0.5
0.75
≥0.9
Calibrated Probability of Regulation

Fig. 4. Maps show that regulation under each Clean Water Act rule varies enormously across the US. (**A**) and (**E**) show ex post deep learning projections of jurisdiction under *Sackett* and *Rapanos*. SI Appendix, Fig. S5D shows ex post deep learning projections under NWPR. (**B**) shows ex ante deep learning projection of *Sackett*. (**C**) shows Connected geophysical model projections. (**D**) shows Wetness model (19) projections. Maps aggregate the four million prediction points by taking the mean model score in 5 km by 5 km grid cells (~8 prediction points per grid cell). Extreme calibrated probabilities (0.0 – 0.1; white, 0.9 – 1.0; blue) are plotted with the same color. Color scaling uses a power transformation ($\gamma = 0.6$) to improve visual differentiation at lower probability values.

Supporting Information for

Deep Learning Projects Jurisdiction of New and Proposed Clean Water Act Regulation

Simon Greenhill[1,2]*† Brant J. Walker[3]*† Joseph S. Shapiro*[1,4,5]

Affiliations

[1] Department of Agricultural and Resource Economics, University of California, Berkeley, Berkeley, CA, USA.

[2]Global Policy Laboratory, Stanford Doerr School of Sustainability, Stanford University, Stanford, CA, USA.

[3] School of the Environment, Yale University, New Haven, CT, USA.

[4] Department of Economics, University of California, Berkeley, Berkeley, CA, USA.

[5] National Bureau of Economic Research, Cambridge, MA, USA.

*Corresponding authors. Emails: sgreenhill@berkeley.edu, brant.walker@yale.edu, joseph.shapiro@berkeley.edu

† These authors contributed equally to this work

This PDF file includes:

A. Materials and Methods

A.1: Model Details

Geophysical Models. The Wetness geophysical model (1) analyzes several scenarios for how *Sackett* could affect jurisdiction. Each scenario assumes that a water resource inundated a certain fraction of the year is jurisdictional; this fraction varies by scenario. Ex ante, it is unclear which of the wetness scenarios to consider. The main text reports the median scenario in terms of wetness (scenario 4 out of 8). Table S4 discusses all scenarios. Scenarios 3 and 4 have the best performance in the validation set. Jurisdiction is not monotonic in wetness, and we observe both jurisdictional and non-jurisdictional AJDs in six of the eight water regimes with at least one AJD (Fig. S4).

As mentioned in the main text, because the original wetness model presents results for a very specific set of non-tidal wetlands, we report results for three separate samples of *Sackett* AJDs. First, we consider AJDs within wetlands in the analysis area of the original Wetness model (1), which only includes 36 observations in the test set. Second, we consider AJDs within all NWI non-tidal wetlands (N=640). Third, we consider all AJDs. To maximize comparability across models, our main results report the performance of the Wetness model for all *Sackett* AJDs and assume that AJDs not in the analysis area of (1) are non-jurisdictional.

The Connected geophysical model closely follows the "damaging" and "very restrictive" geophysical models developed in (2, 3). We predict that any points within an NWI wetland polygon that intersects with a "navigable" NHD flowline are jurisdictional. "Navigable" in these models is defined as having an NHD feature code (fcode) of "perennial" or "intermittent". Additionally, we follow the previous models by only considering flowlines "most likely to qualify as regulatory wetlands" (2, 3). Following NRDC (3), we only consider wetlands likely to be regulated.[1] Following (2), we do not consider wetlands that have been drained, excavated, or farmed.

We also consider a series of geophysical models that use a single geophysical input layers to predict jurisdiction (see Section A.5, B.1, and Table S5). Specifically, we predict *Sackett* jurisdiction using each of the geophysical input layers used as inputs to the deep learning models. The performance of the ex post model highlights the benefits of using deep learning to parse through many input layers and learn patterns that predict jurisdiction.

Ex Ante Deep Learning Model. AJD data include classifications of water types ("resource types") that differ by rule and that classify each water body. Each resource type corresponds to a legal description of a category of waters in a rule, rather than to a geophysical classification (e.g., resource types often do not correspond to Cowardin (4) wetland types). Under NWPR, for example, AJDs classify some sites as "adjacent wetlands," others as "non-adjacent wetlands," and others as other types of sites. In machine learning terminology, each AJD provides a "label" for a location, since the AJD attaches a jurisdictional determination to the location, which is the outcome we seek to predict. Our ex ante deep learning model "relabels" AJDs by modifying the

[1] Specifically, we consider only wetlands with a vegetated component with codes EM: Emergent, SS: Scrub-Shrub, or FO: Forested and a water regime with codes A: Temporarily Flooded, C: Seasonally Flooded, D: Continuously Saturated, E: Seasonally Flooded/Saturated, F: Semi-permanently Flooded, G: Intermittently Exposed, or H: Permanently Flooded.

jurisdictional determination made under one rule to reflect the decision that we conclude would have been made under a proposed rule. Conclusions about what to relabel come from our reading of the text of the *Sackett* decision but use no ex post information about *Sackett* implementation.

Table S2 shows how we relabel AJDs to create the ex ante deep learning model. This relabeling scheme assumes that relative to NWPR, *Sackett* deregulates two categories of waters: wetlands separated from navigable waters by artificial structures and natural features. We believe these specific relabeling choices follow directly from majority opinion in *Sackett*. We chose them in June 2023, before USEPA announced a conforming rule or USACE announced associated guidance. SM B.5 of Greenhill et al. (5) discusses how changing labels in past data could describe new rules, though questions the potential promise of this approach and does not implement it.

We train the ex ante deep learning model in two steps. First, we pre-train the model using AJDs from all CWA rules besides *Sackett*. This pre-training allows the model to learn general relationships between input data and regulatory outcomes that are present across prior CWA rules. Second, we fine-tune the model using only data on the NWPR AJDs, which have already been relabeled to characterize *Sackett*. The fine tuning adapts the representation of the relationship between inputs and regulation learned in pre-training to a *Sackett*-specific interpretation of jurisdiction, without requiring the model to relearn general geophysical patterns from scratch.

Our relabeling methodology builds on past work in machine learning. Most closely related is the tradition of weak and indirect supervision, where researchers generate labels for unlabeled data using heuristics of knowledge bases (6–8). Instead of generating entirely new labels for data with no existing labels, our approach transforms data that have already been labeled to reflect what the labels would be under a different rule. Unlike ex post deep learning models (5, 9), which are trained using true labels, or simulation-based methods (10), which project outcomes using process-based environmental models, our relabeling methodology derives labels by mapping historical decisions to the criteria of a new regulation. This allows deep learning to generate projections before data from the rule under consideration exist, enabling ex ante projection.

Ex Post Deep Learning Model. As with the ex ante deep learning model, we train the ex post deep learning model in two steps. First, we pre-train a single model using pooled data from all Clean Water Act rules. Second, we fine-tune the model using only data from one rule at a time. We predict jurisdiction for each of the four main rules enforced since 2018—CWR, *Rapanos*, NWPR, and *Sackett*. We include the 2023 Rule together with *Rapanos* since they have extremely similar design. The next section describes deep learning architecture and training details.

The deep learning models can predict jurisdiction for any coordinate in the contiguous US. They therefore avoid a predetermined decision between a framework only designed to analyze streams (11) or only designed to analyze non-tidal wetlands (1); each of these categories covers a small fraction of all AJDs.

<u>A.2: Deep Learning Model Architecture</u>

The ex ante and ex post deep learning models use an architecture that takes both raster data and tabular data as inputs to predict CWA jurisdiction. The rasters are gridded spatial data such as satellite imagery and maps of stream and wetland locations. For each AJD, we assemble raster data for a 308-by-308 meter (512-by-512 pixel) area centered at the AJD's latitude and longitude. Tabular data are row-and-column data consisting of one row for each AJD. These describe characteristics of the location being evaluated – such as the USACE district deciding the AJD – that we treat as constant within the 308-by-308 meter neighborhood around the AJD. Each model outputs a raw model score for each site of interest between 0 and 1. As discussed in the main text, we then use isotonic regression to translate the raw model scores to a calibrated probability of regulation.

The model architecture has two branches: one that processes the raster data, and one that processes the tabular data. The raster branch of the model has 29 input layers: color and near infrared aerial imagery; the locations and characteristics of streams and wetlands; elevation; summary statistics of long-run average precipitation, temperature, dewpoint temperature, vapor pressure deficit, solar radiation, and cloudiness; soils data; land cover data; and Level IV Ecoregions data. Section A.5 provides additional details about input layers. Twenty-eight of these layers were used in Greenhill et al. (5); we add land cover data from the Coastal-Change Analysis Program (C-CAP) due to its resolution and quality, while recognizing that C-CAP covers only coastal areas. These inputs provide a detailed snapshot of ground conditions affecting the probability of CWA jurisdiction and include the main national layers that USACE reports using in AJDs.

The tabular branch of the model consists of 89 features. These include one-hot encoded identifiers for the state and USACE district of the location being evaluated, the distance to district headquarters, and one-hot encoded information on the WOTUS rule under which the location's jurisdictional status is being evaluated (i.e. *Sackett*, *Rapanos*, NWPR, or CWR). Section A.5 discusses these features. Greenhill et al. (4) included these features in a raster format; we include them in a tabular format to improve computational efficiency.

The branch processing the raster data is a ResNet-18 (14) convolutional neural network pre-trained on ImageNet (15). The convolutional neural network takes as inputs a stack of two-dimensional rasters and outputs a one-dimensional vector summarizing the information in those rasters that is most relevant to CWA jurisdictional status. This vector is combined with the vector of tabular features.

The combined vector of features is then passed through a small two-layer neural network (a perceptron). This step flexibly interacts all the features, allowing for non-linearities and interactions between the information in the raster data and the information in the tabular data. For example, the presence of a stream may have different implications for jurisdiction in different states or USACE districts due to regional differences in hydrology or USACE practices. This step allows the model to learn such differences if they are present in the AJD data.

Finally, the vector of fully interacted raster and tabular features is used to predict jurisdictional status. Intuitively, this last step is like running a regularized logistic regression, which penalizes model complexity, of jurisdictional status on the interacted raster and tabular features. In practice, all parts of the model are trained jointly so that the feature extraction and prediction steps are optimized together.

We experimented with a geo-foundation model in the validation set that used embedding fields (12) but found that it modestly decreased performance in the validation set, perhaps because other layers had similar information and due to the sample size. We therefore do not use these embedding fields data.

Train-Test Split. We divide the 202,295 AJDs into disjoint training, validation, and test data sets. We avoid footprint overlap between folds so as to prevent leakage across folds (Fig. S7). The deep learning models use the train, test, and validation split rules from Greenhill et al. (13) (SM, lines 33–43), with a few extensions. When assigning groups for new Approved Jurisdictional Determinations (AJDs), we first create groups of AJDs with overlapping footprints. If a new AJD's footprint group overlaps with multiple groups of AJDs used in the original model, the new AJDs take the split of the AJD it overlaps with. If the new groups connect AJDs that the original model put in separate groups, we assign or reassign all to the same split. If an AJD from the original model is in the train split, we assign all connecting AJDs in the same footprint group to training, then testing, and finally validation. We split all new AJDs that do not overlap, following the procedure in (13).

<u>A.3: Synthetic Data</u>

AJDs tend to focus on sites where jurisdiction is ambiguous. AJDs therefore describe relatively few locations that are unambiguously jurisdictional (e.g., in the middle of the Great Lakes or Mississippi River), or non-jurisdictional (e.g., on desert mountain peaks). Augmenting the AJD data with locations where prior knowledge suggests unambiguous jurisdiction may improve the model's generalizability. Adding unambiguous examples to the training data set may also improve the model's performance on the test set if the unambiguous examples provide relevant information to AJD jurisdiction, by helping the model learn features that predict both the unambiguous examples and the (typically more ambiguous) AJDs.

We therefore add synthetic AJDs to the training and validation sets. Synthetic AJDs do not represent observed USACE decisions, but instead they represent sites where we generate a data point which we can conclude with high confidence represents the jurisdictional outcome that USACE would pick for the site if it had an AJD. We generate jurisdictional synthetic AJD points in perennial streams that terminate in navigable waters and in the largest 98 inland lakes that are deep enough for boat access. We generate non-jurisdictional synthetic AJDs for *Sackett* in isolated wetlands (prairie potholes, playas, West Coast vernal pools, and salt flats) and along hydrologic region (HUC2) boundaries (Tables S12, S13). We develop separate procedures for identifying unambiguously jurisdictional and non-jurisdictional locations, as detailed below. Figs. S7A and S7B map the synthetic data that we generate.

Synthetic Data: Jurisdictional Locations. We generate jurisdictional synthetic training data within National Hydrography Dataset (NHD) (14) area stream, river, sea, and ocean polygons that connect to NHD flowlines terminating at navigable waters. All NHD flowlines list their terminal feature. We identify all NHD flowlines whose terminal feature is coastal, a large inland lake such as the Great Lakes or Humboldt lake, or at the US border; these are potentially navigable. To ensure completeness, we manually investigate the jurisdictional status of terminal features not meeting the criteria above that serve as a terminus for over 1,000 other flowlines.

We keep all NHD area polygons classified as streams/rivers (NHD fcode: 46006) or sea/ocean (NHD fcode: 44500) that spatially intersect with a flowline identified above. To ensure we select coordinates inside the water body, we exclude area within a 10 meter buffer inside the boundary of each NHD area polygon. Finally, we randomly select coordinates from these polygons. Fig. S7A shows that this procedure primarily selects traditional navigable waters.

Synthetic Data: Non-Jurisdictional Locations. We draw two sets of non-jurisdictional synthetic data: isolated wetlands and hydrologic unit code (HUC2) boundaries.

Synthetic Non-Jurisdictional Data: Isolated Wetlands. We identify wetlands that are not jurisdictional under *Sackett* or NWPR by following the classification of Tiner (15). For each isolated wetland type in Table S1 and Fig. 3 of Tiner (15), we identify the US region with that type of isolated wetlands. Tables S12 and S13 describe our mapping from Tiner (15) wetland types to geographic regions. In some cases, one wetland type spans multiple geographic regions. We were unable to link about half of the Tiner categories to specific US regions, and therefore we do not generate synthetic non-jurisdictional training data for these categories.

We then identify isolated wetlands separately for each region and isolated wetland type. We take all National Wetland Inventory (NWI) (16) polygons at least 100 meters from any navigable water, where we define navigable waters as above.

To identify wetland types for non-jurisdictional synthetic data, we tabulate all AJDs with the identified NWI polygons satisfying the criteria of the previous paragraph, separately by Cowardin (4) code (Table S13). We require that AJDs within wetlands of that Cowardin code must satisfy the following additional criteria:

1. We must observe at least 25 AJDs falling within wetlands of that Cowardin code;
2. Across all rules, no more than 10% of AJDs within these wetlands can be jurisdictional;
3. No more than 5% of Navigable Waters Protection Rule (NWPR) and *Sackett* AJDs within these wetlands can be jurisdictional.

As one test of whether this procedure effectively identifies isolated wetlands, among the *Sackett* AJDs satisfying these criteria, we find that the Army Corps of Engineers (USACE) classifies resource codes for 97% as isolated wetlands. Other CWA rules lack a distinct resource type for isolated wetlands, so we cannot report comparable statistics from AJDs for other rules. We generate synthetic non-jurisdictional training data for NWPR and *Sackett* only, since the jurisdictional status of isolated wetlands is more ambiguous under other rules.

Synthetic Non-Jurisdictional Data: HUC2 Boundaries. We generate additional synthetic non-jurisdictional training data along hydrologic region boundaries. The US Geological Survey defines a HUC as land area within which surface water drains to a point. We focus on the 21 HUC2 water resource regions, which define the drainage areas of one or multiple major rivers. HUC2 boundaries are typically uplands, since they demarcate one drainage region from another, and thus are not jurisdictional. For example, the Pacific Northwest constitutes one HUC2, bounded by several mountain ranges (Pacific Coast, Siskiyou, Absaroka, and others). The Continental Divide and Great Basin distinguish parts of other HUC2 boundaries. One could oversimplify a HUC2 boundary as a mountain ridge where one side has streams flowing to the East and the other side has streams flowing to the west, though many HUC2 boundary areas in the Midwest and South are along the highest portion of low-elevation sloped areas.

To generate synthetic training data along HUC2 boundaries, we randomly sample points satisfying the following criteria:

1. Within 50 meters of HUC2 boundaries
2. Not within 50 kilometers of international borders
3. Not within 50 meters of any NHD flowline that NHD indicates terminates in an ocean, large inland lake, or US border
4. Not within 50 meters of any NHD area polygon intersecting such NHD flowlines.

We exclude areas within 50 kilometers of international borders since some HUC2 boundaries coincide with oceans and Great Lakes.

As one test of whether this strategy accurately identifies non-jurisdictional areas, we examine the 69 true AJDs satisfying all these criteria. Among these AJDs, 37 were completed under *Rapanos*, 7 under the Clean Water Rule, 7 under NWPR, and 18 under *Sackett*. USACE concluded that none of these 69 AJDs are jurisdictional.

Synthetic Data: Model Training. In model development using the training and validation sets, we experimented with including different quantities of synthetic data, between about 500 synthetic points up to 100,000. We found that AJD validation set performance was maximized around 1,000 points of each synthetic type (i.e., 1,000 synthetic jurisdictional points, 1,000 synthetic non-jurisdictional points from HUC2 boundaries, and 1,000 of each of the synthetic non-jurisdictional isolated wetland types).

Synthetic data improve model performance both for traditional navigable waters and more ambiguous cases. Our ex post deep learning model has near-perfect accuracy on synthetic data. Because the synthetic data are not in the validation set or the held-out test set, no model performance statistics elsewhere in the paper describe the synthetic data. Additionally, including synthetic data improves model accuracy on AJDs by 2 to 3 percentage points in the validation set, as well as improving both precision and recall by 6 to 7 percentage points each. This suggests that including synthetic points helps the model distinguish between ambiguous and unambiguous decisions, and so reduces the rate at which the model produces both false positive and false negative predictions.

A.4 AJD PDF Files

We obtain labels from tabular data that USEPA and USACE provide online (17) and that we downloaded on March 24, 2025. For each AJD, USACE staff complete a document listing jurisdiction of each potential water resource in the project, and USEPA and USACE then separately hand-enter the AJD content into the tabular data we use as labels. PDFs of the AJD documents are available for a limited subset of sites, while the tabular data are available for all sites.

To assess potential classification errors in the labels, we manually compare labels in the tabular data and the AJD documents. We find that labels in the AJD PDF documents disagree with labels in the tabular data for 3.4% of AJDs, and have coordinates differing by more than 217.8 meters for 19.4% of AJDs, meaning that the input data tile does not include the location evaluated by the AJD. The percentage of differential coordinates partially reflects many project PDFs listing the project centroid, rather than the centroid of the relevant water feature. We do not use the AJD document labels or coordinates as ground truth data for a few reasons—the AJD documents are only available for 7,556 of over 40,000 projects in our sample; a single AJD document often reports labels for many potential water resources within a development project and correctly mapping each water feature's label to the water features within the project can introduce additional error; and few AJD documents list coordinates for individual water resources, while many list coordinates for the project centroid.

A.5 Input Layers

Our deep learning models take as inputs 29 raster layers and 89 tabular features. Twenty-eight of the raster layers are identical to those used in Greenhill et al. (13): three-band color and near infrared aerial imagery from the National Agricultural Imagery Program (NAIP) (18); wetland types from NWI (16); river and stream feature codes, stream order, seasonal high and low flows, and path lengths from NHD; elevation from the 3D Elevation Program; land cover data from the National Land Cover Dataset (NLCD) (19); soil taxonomic class, hydric rating, water table depth, flooding frequency, and ponding frequency from the Gridded National Soil Survey Geographic Database (gNATSGO) (20); average annual total precipitation, average daily minimum temperature, average daily maximum temperature, average daily mean temperature, average daily dew point temperature, average daily minimum vapor pressure deficit (VPD), average daily maximum VPD, average daily clear sky and total solar radiation, and average daily atmospheric transmittance (cloudiness), all for 1990–2021, from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) 30-year normal (21); and level IV Ecoregions (22). Further details about these layers are available in Table S4 of Greenhill et al. (13). We also include land cover data from the Coastal Change Analysis Program (C-CAP) (23), which has higher native resolution and is sometimes believed to be more accurate than NLCD. Because C-CAP covers only coastal areas, we also use NLCD. All raster inputs are resampled from their original resolution to match the resolution of the 0.6 meter NAIP imagery, resulting in 512 by 512 pixel rasters centered at the location being evaluated, covering an area of approximately 308 by 308 meters. Several raster input layers are available only in the contiguous US, so we restrict our analysis to this region. We selected raster input layers based on the datasets that USACE engineers most frequently cited in the PDF files accompanying AJDs (13).

The 89 tabular features consist of one-hot encoded identifiers for the state and USACE district of the location being evaluated, the distance to district headquarters, and one-hot encoded information on the WOTUS rule under which the location's jurisdictional status is being evaluated. State and USACE district boundaries have an important influence on jurisdictional rates. Similarly, distance to district headquarters may influence the likelihood that a site receives a field visit, which may also affect jurisdictional determinations (13). Including one-hot encoded rule information allows us to capture differences across rules and produce model predictions for the same locations under different rules.

Our visual review of spatial patterns in the 4 million prediction points reveals that deep learning predictions occasionally display discontinuities within a water body. Investigation indicates that discontinuities in input layers, typically the National Agricultural Imagery Program (NAIP) (18) and the Gridded National Soil Survey Geographic Database (gNATSGO) (20), drive these patterns. In all examples we investigated, the algorithm itself does not generate these discontinuous patterns except insofar as the inputs have them. The infrequent abrupt changes in NAIP inputs that we identified reflect cloud cover affecting processing of remote sensing data. gNATSGO combines the Soil Survey Geographic Database (SSURGO), State Soil Geographic Database version 2, and the Raster Soil Survey data. Analysts create SSURGO by stitching together soil survey areas. One survey area may cover one or several entire counties or parts of counties. This stitching process occasionally produces discrete spatial changes in gNATSGO inputs.

A.6: Agriculture

The CWA excludes prior converted cropland from jurisdiction, but many AJD coordinates fall within NLCD's cropland layer. To understand this contrast, an additional analysis manually investigated a sample of 88 jurisdictional AJDs from *Rapanos*, NWPR, and *Sackett* which have coordinates within NLCD's cropland layer. For each AJD where a document was available, this analysis checked the coordinate in the USEPA-USACE tabular data against the coordinate in the document. This analysis inspected Google Earth imagery from these coordinates and compared against any maps in the AJD document. This analysis found that only 12.5% of the sample of AJDs (11 AJDs) within NLCD's cropland layer represented agricultural activity. For these AJDs, the AJD documents contained insufficient information to determine why the AJD was jurisdictional and was not excluded as prior converted cropland. For example, it is possible these sites became cropland recently so were not "prior." Of the remaining AJDs, 48.9% were near agriculture but not on a field (e.g., a pond or house next to cropland), 29.5% appeared to have slight reporting error in the coordinate, and 9.1% had incorrect labels, as Section A.4 discusses.

A.7: Model Calibration and Decision Threshold Choice

Raw deep learning model scores have imperfect calibration, i.e., model scores do not reflect the probability that a point is regulated. We determine this by comparing bins of deep learning model scores against the empirical probability that AJDs in a bin are jurisdictional. To improve model calibration, we fit an isotonic regression on the training set, then use the fitted isotonic regression model to calibrate out-of-sample predictions. This procedure improves model calibration, especially for calibrated probabilities below 0.6. The Brier score, a common measure of the quality of model mis-calibration, is 0.178 on the test set before calibration and 0.148 after calibration.

Geophysical models primarily generate binary predictions of whether a site is jurisdictional. Deep learning models produce continuous model scores, which we calibrate to describe the probability that a site is regulated. Deep learning models can also generate a binary jurisdictional prediction indicating whether the site's calibrated probability exceeds a given threshold (e.g., 0.5). To use all information from the model, when we report the share of an area that is jurisdictional, we average calibrated probabilities rather than averaging binary jurisdictional predictions.

Different stakeholders may value different model performance metrics and may thus prefer different decision thresholds for binary jurisdictional classification. The AUC summarizes how well a model ranks locations, from more to less likely to be regulated. It therefore aggregates across all possible decision thresholds without requiring a binary decision cutoff. For example, an AUC of 1.0 means the model always assigns higher probability to a jurisdictional site than to a non-jurisdictional site.

For other metrics, Table S1 reports model performance for classification thresholds that differ by performance metric, with each threshold chosen to optimize the performance metric of interest. We use the validation set to choose these thresholds. To evaluate the sensitivity of threshold choice to validation set sampling variation, we implemented threshold selection using five-fold cross validation and observed minimal variation in the selected thresholds.

Fig. S8 shows how performance metrics vary across decision thresholds in the validation set. We choose optimal thresholds based on performance in the validation set, and then apply these to the test set. Table S14 reports all test set performance metrics at the optimal threshold for each metric. AUC is invariant to threshold choice. Thresholds near 0.25 optimize F1 score and state MAE. A lower threshold, 0.17, optimizes national MAE. A decision threshold near 0.50 maximizes overall accuracy. The figure shows that much of the threshold domain has flat curves, suggesting that threshold choice has only a marginal impact on overall performance. Furthermore, during model development, we implemented five-fold cross validation for threshold selection on the validation set and found that optimal thresholds and the corresponding metrics did not change substantially across folds.

Histograms show the distribution of the calibrated probabilities (Fig. S9). Ex post deep learning has high confidence——few sites have a score near 50% and most have calibrated probabilities below 20% or above 80%. Slight visual differences in probabilities seen between ex post and ex ante deep learning in Fig. 4 are explained by differences in the share of the 4 million prediction points with extreme values less than 20%. Ex post deep learning predicts 95% of the 4 million points with a probability of <20%, while ex ante deep learning only predicts 88%. Fig. S9 shows ex ante deep learning has more predictions in the 20–80% range, particularly within 20–40%. This small difference is accentuated by the gamma power transformation of the color ramp used in Fig. 4.

A.8: Prediction Points

We report model predictions for groups of sites. We randomly select 4 million points across the contiguous US, using the same set of locations analyzed by Greenhill et al. (13). These are gathered by dividing the contiguous US into approximately 80,000 0.1 by 0.1 degree grid cells, then randomly sampling 50 points in each cell. This large number of points allows us to produce high resolution maps (Fig. 4), case studies (Fig. S1, S6), and report on predicted regulation overall and at specific locations of interest (Table S3). We separately report predictions for streams, wetlands, agricultural sites, floodplains, developed urban areas, and areas likely to see urban growth in the future. We identify these areas using NHD (14), NWI (16), the National Land Cover Dataset (NLCD) (19), and the National Flood Insurance Program (NFIP) (24), and Integrated Climate and Land Use Scenarios (ICLUS) (25) .

We report the mean calibrated probability for the 4 million prediction points and for subsets of these points in important areas, including within 5 m of NWI wetlands or NHD streams (Table S3). Because we average across points, we interpret these in terms of stream miles and wetland acres.

B. Supplementary Text

B.1: Geophysical Model Projections

Table S5B presents models using one geophysical input layer at a time to determine jurisdiction. For example, the presence of hydric soils is sometimes taken as an indicator of historic wetlands (26). A prediction relying on whether a site has hydric soils has an AUC of only 0.492, which is worse than the naïve benchmark, and F1 of 0.295. Row 9 shows that a model assuming sites with water table depth less than 10 m are jurisdictional also performs poorly. The CWA excludes prior converted cropland and urban developed areas from jurisdiction, so rows 10 and 11 use crop cover and built-up area classes in the NLCD. Again, these rules perform poorly.

The Connected Wetlands model reported in all analysis assumes that only NWI wetlands that intersect with a "navigable" NHD are regulated. The Connected model performs poorly, particularly when making national predictions, in part because many jurisdictional areas under *Sackett* are not identified as wetlands in the NWI dataset.

B.2: Projections Using Probabilities Versus Binary Jurisdictional Predictions

As discussed in the main text, to estimate jurisdiction across groups of sites, Tables S3 and S7 average calibrated probabilities. These tables use the calibrated probabilities since binary jurisdictional predictions discard information by discretizing each site to an indicator for whether the calibrated probability exceeds a threshold. For example, if all sites in an area had a calibrated probability of 0.20, averaging the calibrated probabilities would indicate that 20% of sites are jurisdictional, while averaging the binary jurisdictional predictions would indicate that 0% of sites are jurisdictional.

To understand the consequences of this choice, we re-estimated Table S7 by averaging the binary jurisdictional predictions. Averaging the binary predictions would imply that 11.6% of all sites are jurisdictional under *Rapanos* and 5.8% under NWPR. These are well below the values that average calibrated probabilities. Averaging the binary predictions rather than averaging the calibrated probabilities mostly decreases the estimated share of points that are jurisdictional for NWPR, and for points without streams or wetlands. This occurs because, as in the example from the previous paragraph, binary jurisdictional predictions adjust sites with low calibrated probabilities to zero, but the calibrated probabilities retain some non-zero estimated probability of regulation for such sites.

B.3: Additional Discussion of Results

On the full sample, wetness thresholds besides the main scenario discussed in the main text all have similar performance (Table S4). In the non-tidal wetlands sample (N=36), performance varies widely across scenarios, reflecting the small sample. Wetness scenarios 3 and 4, which perform best in the validation sample, have test set AUC of 0.417, well below the benchmark.

Rapanos and CWR let deep learning observe more true positives, increasing recall. NWPR and *Sackett* have fewer true positives, decreasing opportunities to learn to predict positives for these rules. Ex post deep learning national MAE is also near zero for *Rapanos* and NWPR, though much higher for CWR, which has the smallest sample.

SI References

1. A. C. Gold, How wet must a wetland be to have federal protections in post-Sackett US? *Science* **385**, 1450–1453 (2024).

2. R. Meyer, A. Robertson, Clean Water Rule spatial analysis: A GIS-based scenario model for comparative analysis of the potential spatial extent of jurisdictional and non-jurisdictional wetlands. (2019). Available at: https://downloads.regulations.gov/EPA-HQ-OW-2021-0602-0752/attachment_6.pdf [Accessed 22 August 2025].

3. J. Devine, S. Lee, M. McKinzie, Mapping Destruction: Using GIS Modeling to Show the Disastrous Impacts of Sackett v. EPA On America's Wetlands. (2025). Available at: https://www.nrdc.org/sites/default/files/2025-03/Wetlands_Report_R_25-03-B_05_locked.pdf [Accessed 1 July 2025].

4. L. M. Cowardin, V. Carter, F. C. Golet, E. T. LaRoe, *Classification of Wetlands and Deepwater Habitats of the United States* (Fish and Wildlife Service, U.S. Department of the Interior, 1979).

5. S. Greenhill, *et al.*, Machine learning predicts which rivers, streams, and wetlands the Clean Water Act regulates. *Science* **383**, 406–412 (2024).

6. M. Mintz, S. Bills, R. Snow, D. Jurafsky, Distant supervision for relation extraction without labeled data in *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - ACL-IJCNLP '09*, Su, Keh-Yih, Su, Jian, Wiebe, Janyce, Li, Haizhou, Eds. (Association for Computational Linguistics, 2009), pp. 1003–1011.

7. A. Ratner, *et al.*, Training Complex Models with Multi-Task Weak Supervision. *AAAI* **33**, 4763–4771 (2019).

8. H. Wang, H. Poon, Deep Probabilistic Logic: A Unifying Framework for Indirect Supervision. [Preprint] (2018). Available at: https://arxiv.org/abs/1808.08485 [Accessed 22 August 2025].

9. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, Human Decisions and Machine Predictions*. *The Quarterly Journal of Economics* **133**, 237–293 (2018).

10. W. Francesconi, R. Srinivasan, E. Pérez-Miñana, S. P. Willcock, M. Quintero, Using the Soil and Water Assessment Tool (SWAT) to model ecosystem services: A systematic review. *Journal of Hydrology* **535**, 625–636 (2016).

11. C. B. Brinkerhoff, The importance of source data in river network connectivity modeling: A review. *Limnology & Oceanography* **69**, 3033–3060 (2024).

12. Google Earth Engine, Satellite Embedding V1. Available at: https://developers.google.com/earth-engine/datasets/catalog/GOOGLE_SATELLITE_EMBEDDING_V1_ANNUAL [Accessed 24 August 2025].

13. S. Greenhill, *et al.*, Machine learning predicts which rivers, streams, and wetlands the Clean Water Act regulates. *Science* **383**, 406–412 (2024).

14. US EPA, National Hydrography Dataset Plus V2. (2015). Available at: https://www.epa.gov/waterdata/nhdplus-national-data [Accessed 26 August 2025].

15. R. W. Tiner, Geographically isolated wetlands of the United States. *Wetlands* **23**, 494–516 (2003).

16. US Fish & Wildlife Service, National Wetlands Inventory. Available at: https://www.fws.gov/program/national-wetlands-inventory/download-state-wetlands-data [Accessed 26 August 2025].

17. US EPA, CWA Approved JDs. Available at: https://watersgeo.epa.gov/cwa/CWA-JDs/ [Accessed 24 March 2025].

18. US Department of Agriculture, National Agriculture Imagery Program. Available at: https://datagateway.nrcs.usda.gov/GDGHome_DirectDownLoad.aspx [Accessed 26 August 2025].

19. US Geological Survey, National Land Cover Database. Available at: https://www.usgs.gov/centers/eros/science/national-land-cover-database#overview [Accessed 26 August 2025].

20. US Department of Agriculture, Gridded National Survey Geographic Database. Available at: https://www.nrcs.usda.gov/resources/data-and-reports/gridded-national-soil-survey-geographic-database-gnatsgo [Accessed 26 August 2025].

21. Prism Climate Group, Parameter-elevation Regressions on Independent Slopes Model, 30- year Normals. Available at: https://prism.oregonstate.edu/normals/ [Accessed 26 August 2025].

22. US EPA, U.S. EPA Ecoregions, Level IV Ecoregion. (2023). Available at: https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states [Accessed 26 August 2025].

23. NOAA, Coastal Change Analysis Program (C-CAP) Regional Land Cover Data and Change Data. Available at: https://catalog.data.gov/dataset/coastal-change-analysis-program-c-cap-regional-land-cover-data-and-change-data2 [Accessed 24 August 2025].

24. National Flood Insurance Program, Flood Insurance Data | The National Flood Insurance Program for Agents. Available at: https://agents.floodsmart.gov/flood-maps-and-data/flood-insurance-data [Accessed 27 August 2025].

25. US EPA, "Land-Use Scenarios: National-Scale Housing-Density Scenarios Consistent with Climate Change Storylines." *US EPA* (2009). Available at: https://assessments.epa.gov/gcx/document/&deid%3D203458 [Accessed 27 August 2025].

26. K. Moorhead, Evaluating wetland losses with hydric soils. *Wetlands Ecol Manage* **1**, 123–129 (1991).

27. A. Gold, How wet must a wetland be to have federal protections in post-Sackett US? Dryad. https://doi.org/10.5061/DRYAD.4QRFJ6QJ1. Deposited 15 November 2024.

28. US EPA, RPS Indicator Database. (2025). Available at: https://www.epa.gov/rps/data-downloads [Accessed 15 December 2025].

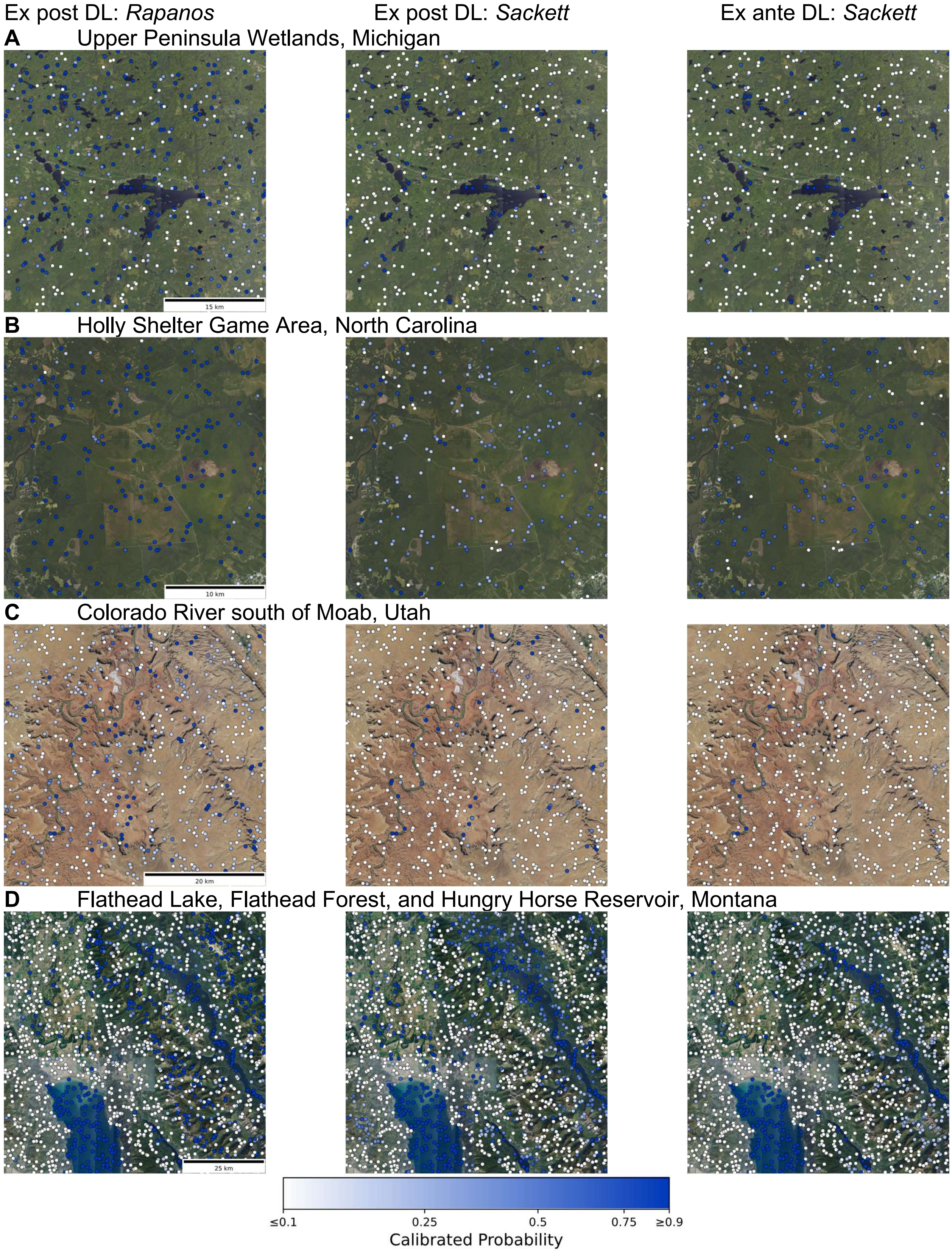
29. US Geological Survey, 3-Dimensional Elevation Program. (2022). Available at: https://www.usgs.gov/3d-elevation-program [Accessed 26 August 2025].

30. US Army Corps of Engineers, US Army Corps of Engineers Regulatory Boundary. Available at: https://geospatial-usace.opendata.arcgis.com/datasets/70805e1a8fd74e42b0a9585088d6d151_0/about [Accessed 26 August 2025].

31. US Census Bureau, TIGER/Line Shapefiles. Available at: https://www2.census.gov/geo/tiger/TIGER2022/ [Accessed 26 August 2025].

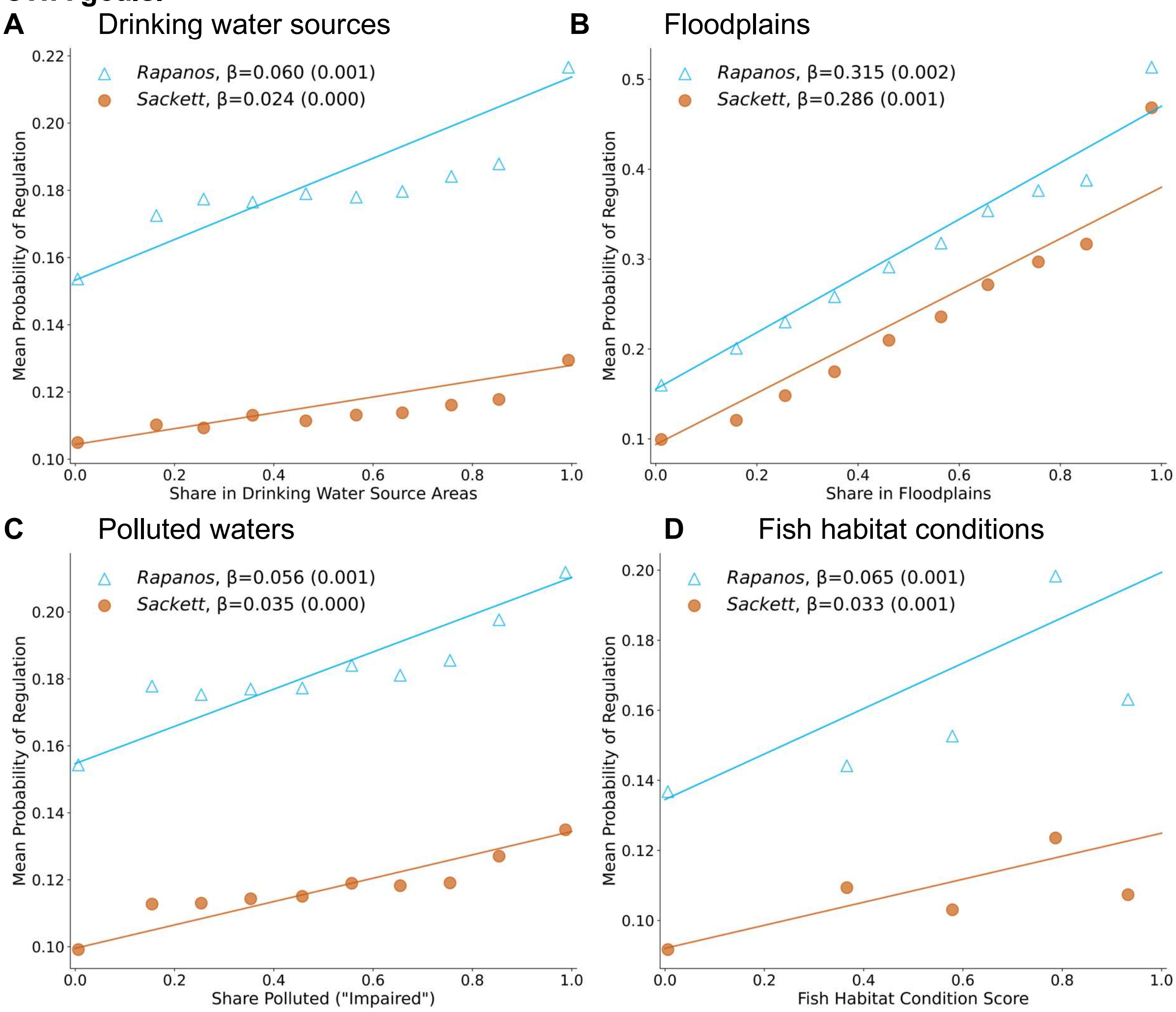
32. US Environmental Protection Agency, “Clean Water Act Approved Jurisdictional Determinations.” Available at: https://watersgeo.epa.gov/cwa/CWA-JDs/ [Accessed 26 August 2025].

Fig. S1. Case studies reveal performance of ex ante deep learning and spatial patterns of jurisdiction.



DL is deep learning. Columns show calibrated model scores for local subsets of the 4 million random prediction points under three different deep learning models. DL is deep learning. (**A**) Lakes and wetlands in the Upper Peninsula of Michigan. All models predict jurisdiction for large water bodies. Ex post deep learning predicts that Rapanos regulates most of the area, *Sackett* regulates surrounding wetlands, and ex ante deep learning closely mirrors ex post deep learning. (**B**) Holly Shelter Game Area, North Carolina. Ex post deep learning predicts that *Rapanos* regulates most of this coastal outdoor recreation area and *Sackett* predicts systematically less jurisdiction. Ex ante deep learning has predictions between these two. (**C**) Colorado River and ephemeral streams south of Moab, Utah. All models classify the Colorado River as jurisdictional. Ex post deep learning predicts that relative to *Rapanos*, *Sackett* deregulates ephemeral streams supplying the river. (**D**) Flathead Lake, Flathead Forest, and Hungry Horse Reservoir, Montana. All models classify Flathead Lake in the southwest of the image and the Hungry Horse Reservoir in the northeast corner as jurisdictional. Ex post deep learning model predicts that *Rapanos* extensively regulates areas of Flathead Forest between the water bodies, *Sackett* regulates little, and ex ante deep learning concurs. Color scaling uses a power transformation ($\gamma = 0.6$) to improve visual differentiation at lower probability values. Figure best viewed in color.

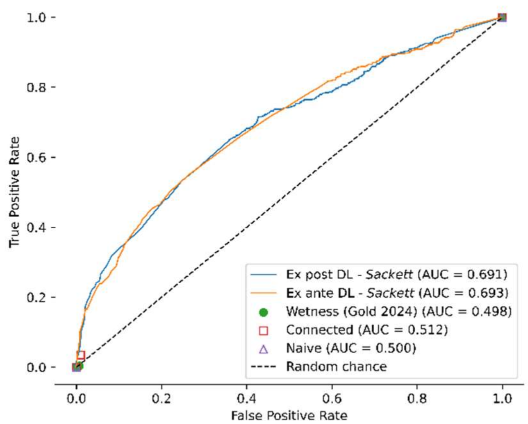
Fig. S2. *Sackett* deregulates areas that support ecosystem services and are important for CWA goals.



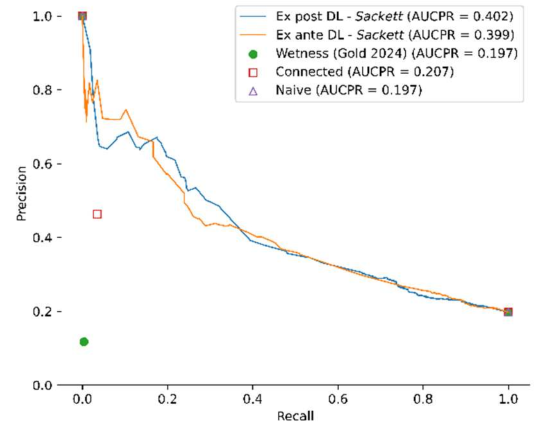
(**A**) Share of points in drinking water source areas. (**B**) Share of points in floodplains (42). (**C**) Proportion of assessed waters considered “impaired” based on pollution and intended use. (**D**) Fish habitat conditions (0=worst, 1=best). Each panel splits 4 million random points into the 5km by 5 km grid cells used to plot Fig. 4. In each graph, the y-axis shows the mean calibrated probability from ex post *Rapanos* and *Sackett* deep learning models, and the x-axis shows the mean ecosystem value within the grid cell. The x-axis divides grid cells into equal-width bins (0–1 scale) based on underlying values. The legend shows the grid-level regression coefficient, with standard errors in parentheses. In all four panels, a hypothesis test that *Rapanos* and *Sackett* have equal slopes rejects with p-value < 0.000, estimated from the interaction term in a pooled regression including both rules. Impaired waters and fish habitat conditions are measured by 12-digit hydrologic unit code (HUC12) from the EPA’s 2025 Restoration and Protection Indicator Database (43).

Fig. S3. Ex ante and ex post deep learning outperform geophysical models.

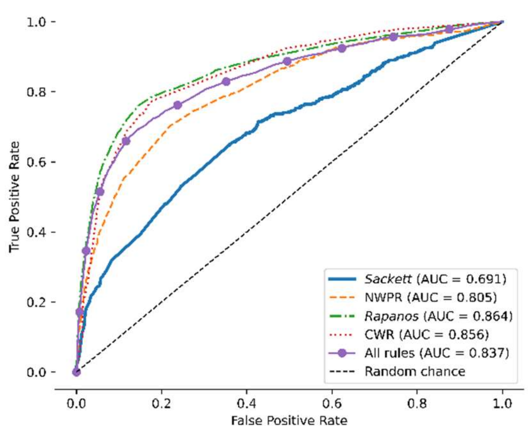
A Receiver operating curves – *Sackett*



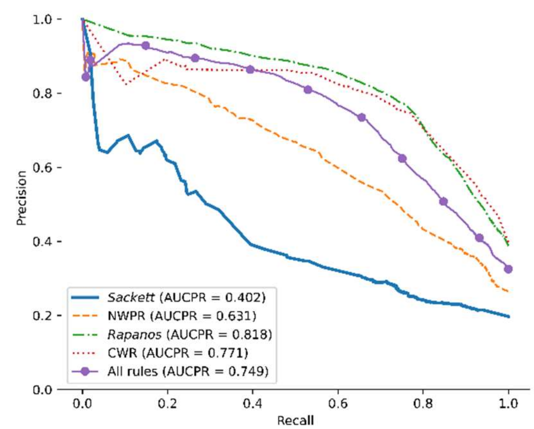
B Precision-recall curves – *Sackett*



C Receiver operating curves – comparing rules

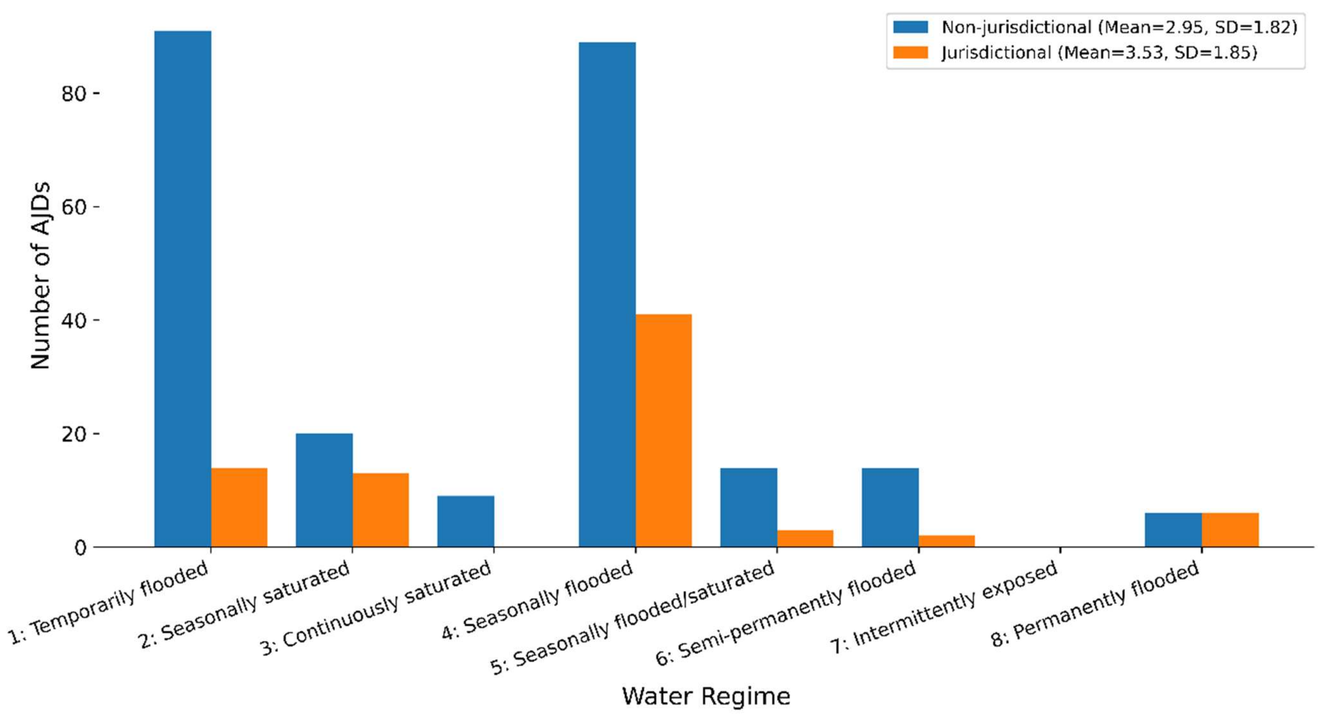


D Precision-recall curves – comparing rules



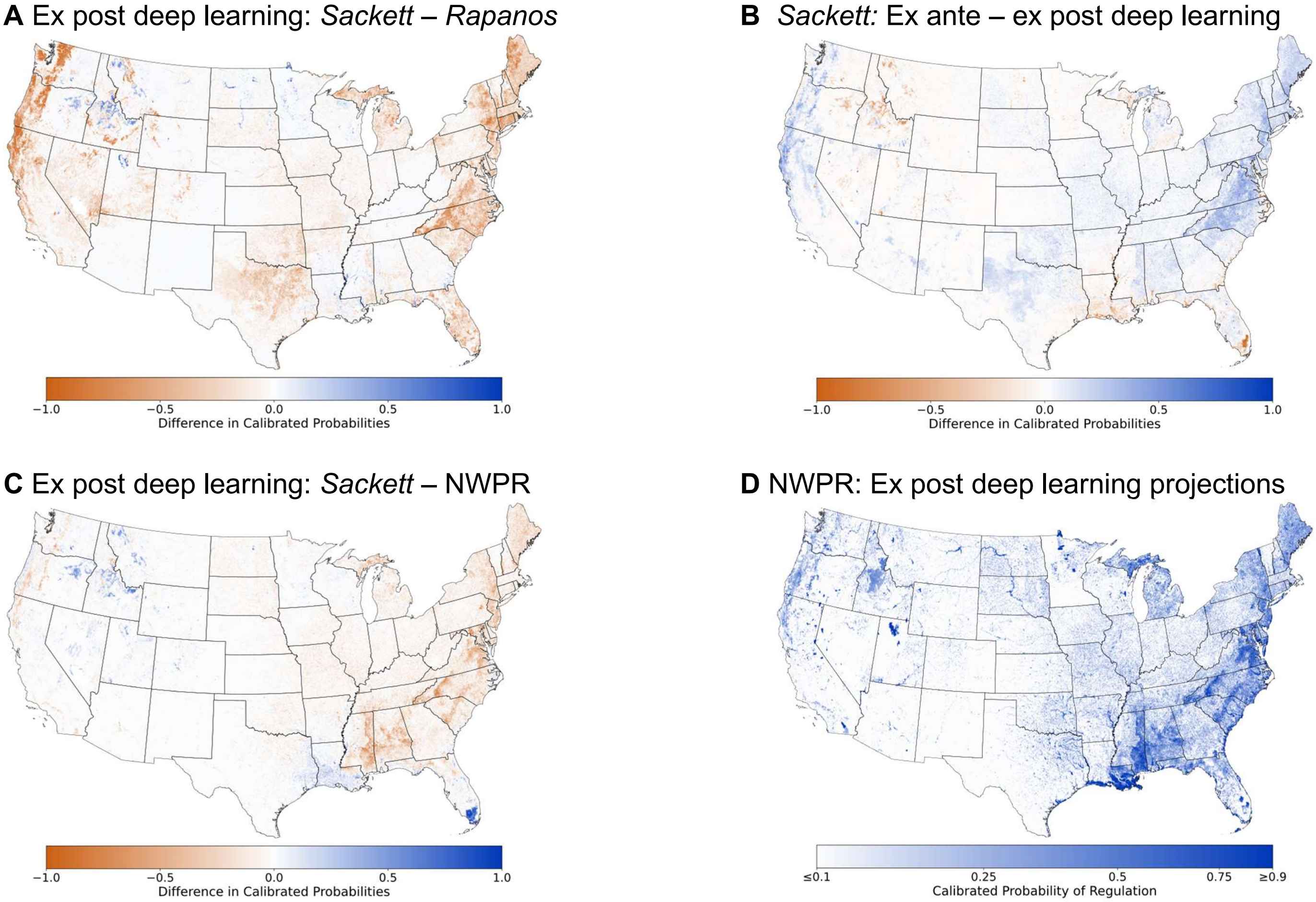
(**A**) and (**C**) show the Receiver operating curve (ROC) and the Area Under the Curve (AUC). The ROC plots the True Positive Rate (share of correctly identified positives) against the False Positive Rate (share of negatives incorrectly identified as positive) across all classification thresholds. For example, the left-most point corresponds to a threshold above one, predicting no positives. The right-most point corresponds to a threshold below zero, predicting all positives. Ex post deep learning (*Sackett*) has 69.1% probability of ranking a randomly chosen jurisdictional AJD higher than a randomly chosen non-jurisdictional AJD. AUC = 0.5 is random chance, AUC-ROC = 1 is perfect. Pooling all CWA rules, ex post deep learning has a 0.837 AUC. (**B**) and (**D**) show the Precision-Recall (PR) Curve and the Area Under the Curve (AUCPR). The PR curve plots precision (share of predicted positives that are true positives) against recall (share of true positives identified) across all classification thresholds. The AUCPR averages precision across all recall levels. A random classifier has an AUC-PR of 0.197 since 19.7% of *Sackett* AJDs are jurisdictional. Ex post deep learning (*Sackett*)'s AUCPR of 0.402 means the model identifies positive cases with about twice the precision as a random classifier, indicating strong performance in detecting jurisdictional AJDs despite class imbalance. Pooling all rules, a random classifier has an AUCPR of 0.325 since 32.5% of AJDs are jurisdictional. Also pooling all rules, ex post deep learning's AUCPR of 0.749 means the model identifies positive cases with over twice the precision as random guessing, indicating strong performance in detecting jurisdictional AJDs despite class imbalance. The PR curves focus on positive-class performance and are more informative under class imbalance. Curves are constructed by using all unique calibrated model scores as thresholds. All curves are independent of any chosen classification cutoff. Because Gold (1), Connected, and the naïve results have binary model scores, these are plotted as points rather than lines.

Fig. S4. NWI Wetness values for *Sackett* AJDs noisily measure jurisdiction.



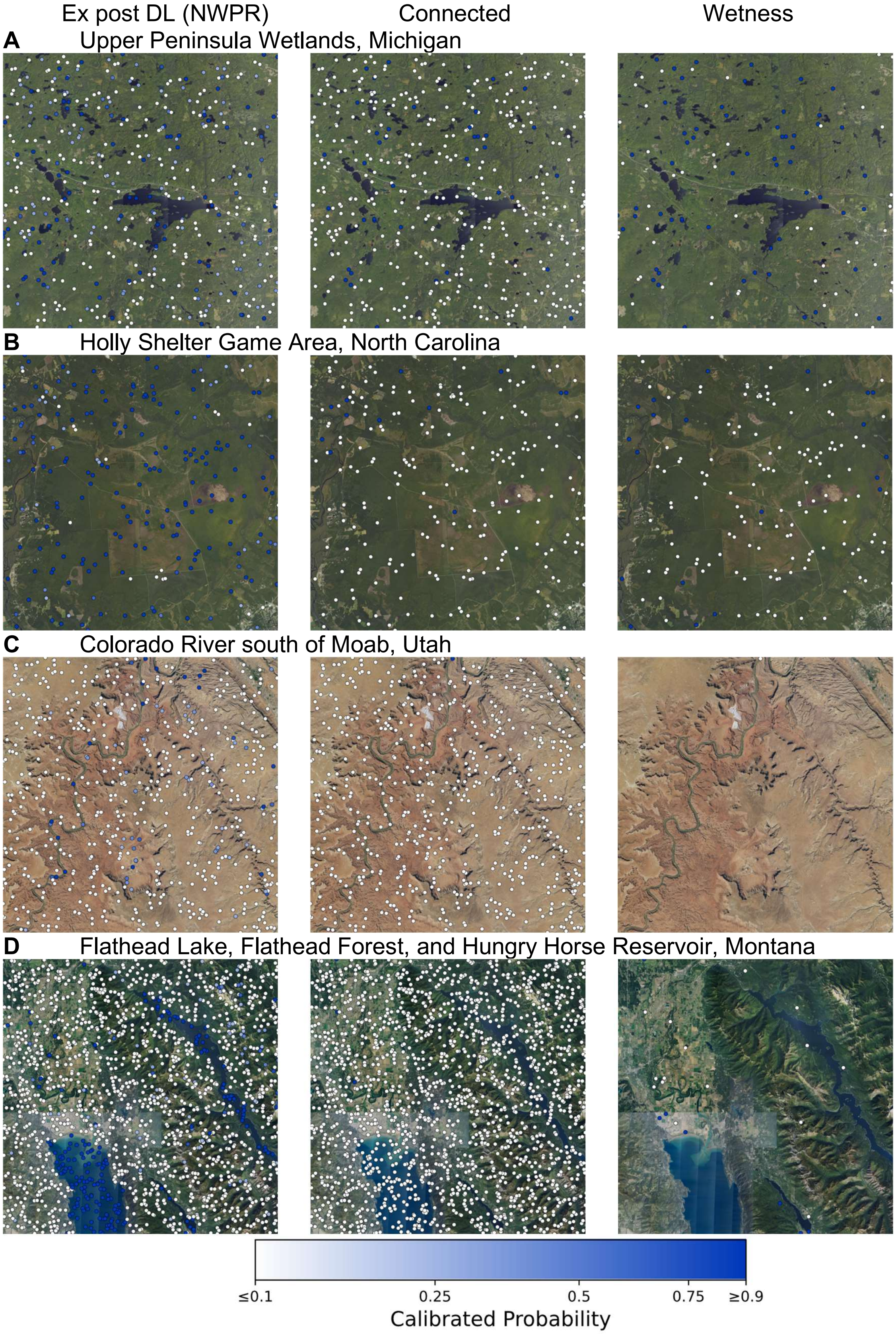
NWI "Water Regime" values differ across both non-jurisdictional and jurisdictional *Sackett* AJDs. Some jurisdictional AJDs have relatively low wetness, and some non-jurisdictional AJDs have relatively high wetness. This figure plots the water regime value, which describes "Wetness" in Gold (1) scenarios, for all 322 *Sackett* AJDs that fall within a NWI polygon in Gold (1). Dark blue bars display non-jurisdictional AJDs; light orange bars display jurisdictional AJDs. SD is standard deviation.

Fig. S5. Maps show large spatial differences in regulation across rules and models.



Brown represents newly deregulated, blue represents newly regulated. Maps show changes from (**A**) *Rapanos* to *Sackett* under ex post deep learning; (**B**) ex post to ex ante deep learning; (**C**) NWPR to *Sackett* under ex post deep learning. (**D**) shows ex post deep learning projections under NWPR. Maps aggregate the four million prediction points by taking the mean model score in 5 km by 5 km grid cells (~8 prediction points per grid cell).

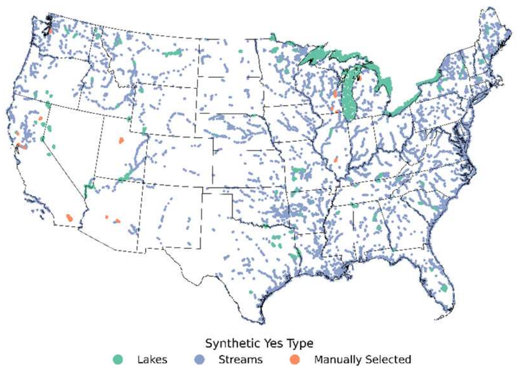
Fig. S6. Case studies show differences across rules and spatial patterns of jurisdiction.



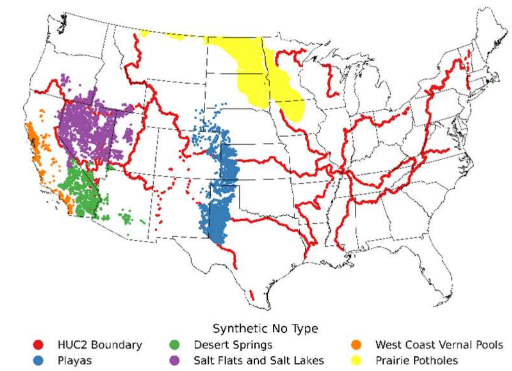
Columns show calibrated model scores for prediction points under three different models, one deep learning and two geophysical. The first column shows ex post deep learning (NWPR), the second column shows the Connected model, and the third column shows the Wetness (1) model (seasonally flooded scenario). The Wetness model only shows prediction points within wetlands used in Gold (1, 27) which lack information for most prediction points. (**A**), Lakes and wetlands in the Upper Peninsula of Michigan. Ex post deep learning (NWPR) and the Connected model predict little jurisdiction for surrounding wetlands, and the Wetness model predicts jurisdiction for different surrounding wetlands and information for many points. (**B**), Holly Shelter Game Area, North Carolina. Ex post deep learning (NWPR) classifies most points as jurisdictional in this coastal outdoor recreation area. The Connected model and the Wetness model predicts little jurisdiction. (**C**), Colorado River and ephemeral streams south of Moab, Utah. Ex post deep learning (NWPR) predicts no jurisdiction for ephemeral streams upstream of the river. The Connected model predicts no jurisdiction, and the Wetness model has no information for any points. (**D**), Flathead Lake, Flathead Forest, and Hungry Horse Reservoir, Montana. Ex post deep learning (NWPR) classifies the lake in the southwest corner and the reservoir in the northeast corner as jurisdictional, but do not regulate the Flathead Forest. The Connected model predicts no regulation, and the Wetness model has almost no information on sites in the area. Color scaling uses a power transformation ($\gamma = 0.6$) to improve visual differentiation at lower probability values. Figure best viewed in color.

Fig. S7. Synthetic and true training data span most US regions.

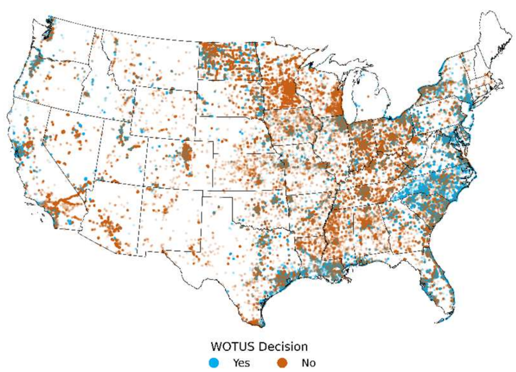
A Synthetic jurisdictional points



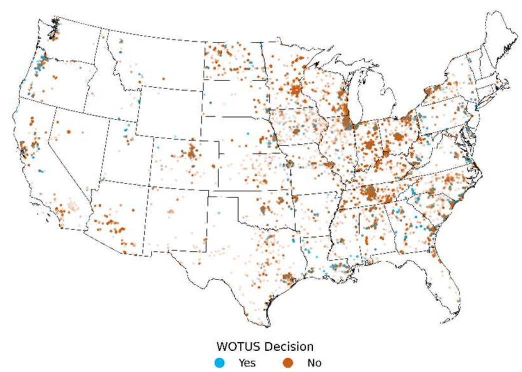
B Synthetic non-jurisdictional points



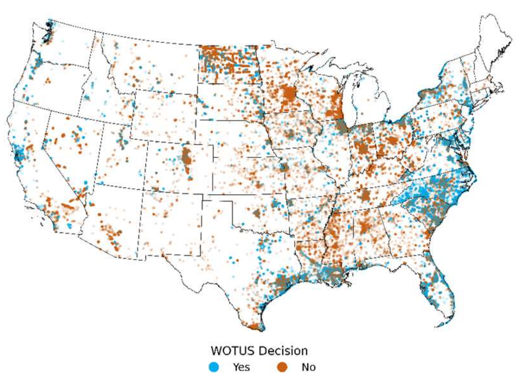
C True AJDS



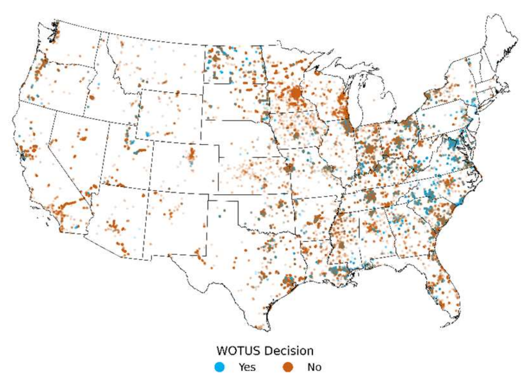
D *Sackett* AJDs



E *Rapanos* AJDs

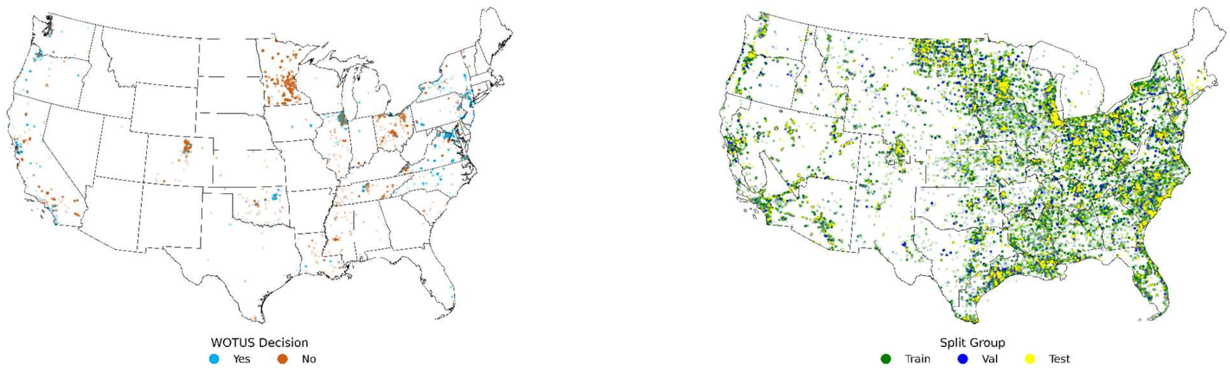


F NWPR AJDs



G CWR AJDs

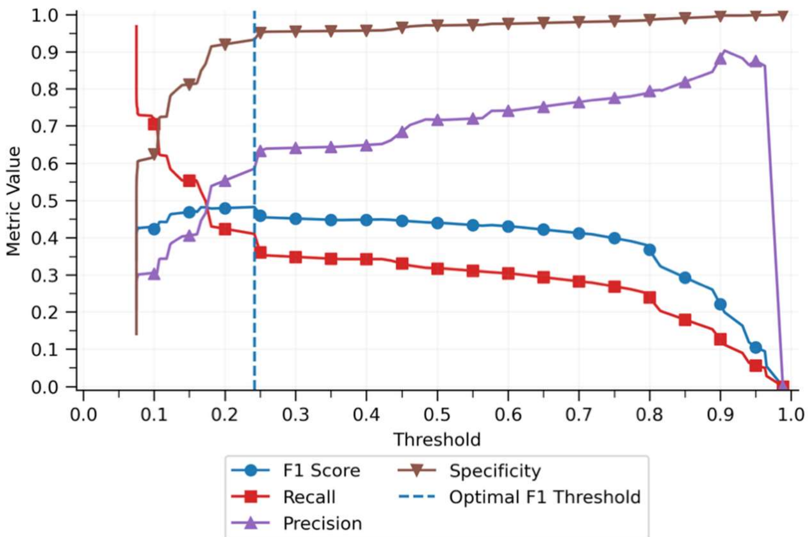
H True AJDs, by split



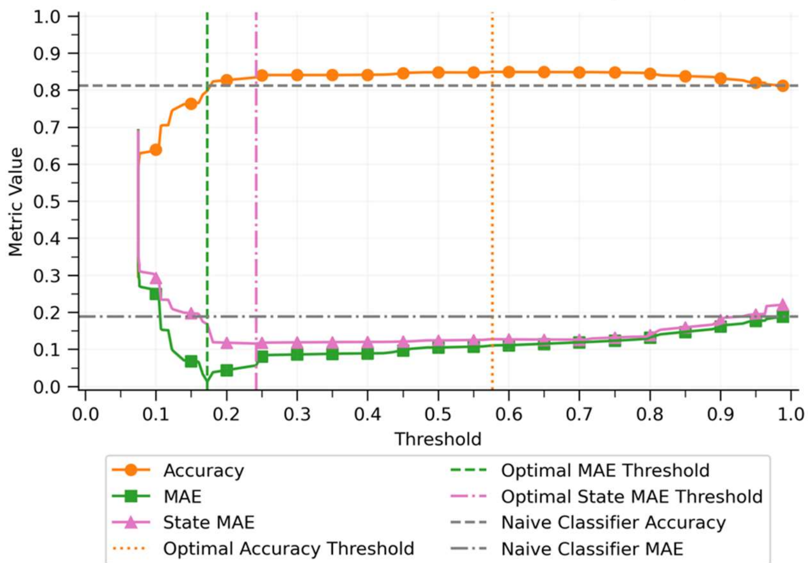
(**A**), synthetic jurisdictional AJDs and **(B**), synthetic non-jurisdictional AJDs, both colored by water resource type. (**C**)**,** true (non-synthetic) AJDs, colored by label. (**D)**–(**G**) separate true AJDs by rule. (**H**) colors true AJDs by split. Lines in (**A**)–(**F**) show states; lines in (**H**) show Army Corps (USACE) districts.

Fig. S8. Jurisdictional thresholds optimize model performance for each metric.

A F1-Score, Precision, Recall, and Specificity



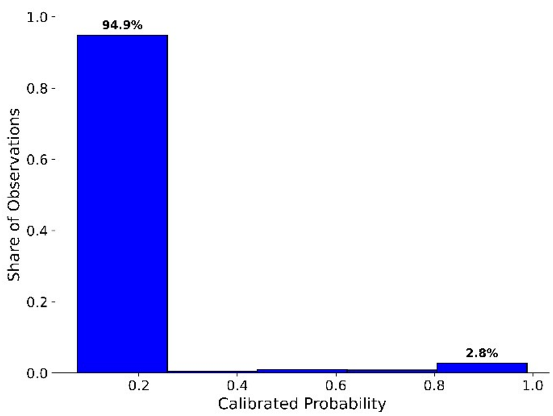
B Accuracy, Overall Mean Absolute Error (MAE), and State MAE



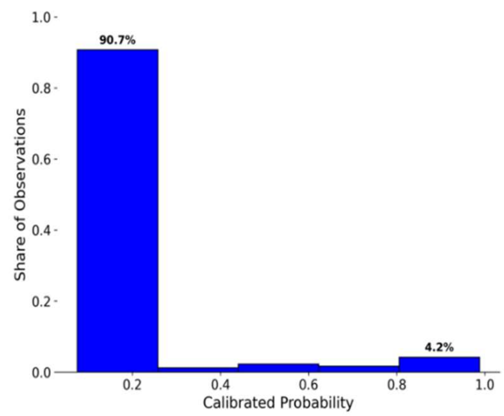
The ex post deep learning (*Sackett*) model predicts a site as jurisdictional if its calibrated probability exceeds the relevant threshold. The y-axis in each graph shows the model's performance on the metric of interest if the model uses the threshold indicated on the x-axis. Each line with markers shows a different performance metric. (**A**), the blue line with circles shows F1; the red line with squares shows recall; the purple line with triangles shows precision; and the brown line with inverted triangles shows specificity. The vertical dashed blue line shows the threshold which maximizes F1. (**B**), the orange line with circles shows accuracy, the green line with squares shows MAE, and the pink line with triangle shows state MAE. Each vertical line shows the threshold which maximize the performance metric with matching color (e.g., the dashed green line shows the threshold which maximizes MAE, which is also shown in green). The horizontal dashed lines show performance of a naïve benchmark that assumes no sites are jurisdictional.

Fig. S9. Distribution of calibrated probabilities of regulation differ by rule and sample, though concentrate below 0.2 or over 0.8.

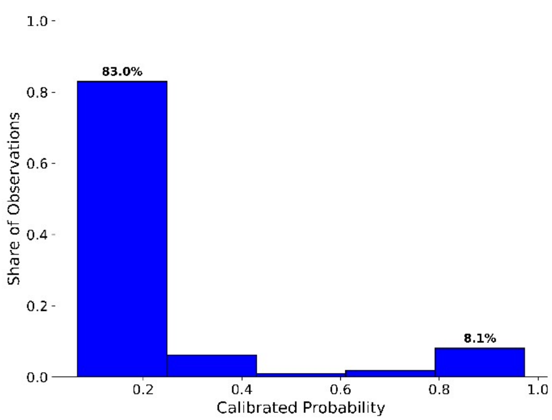
A Ex post deep learning (*Sackett*) – 4mn points



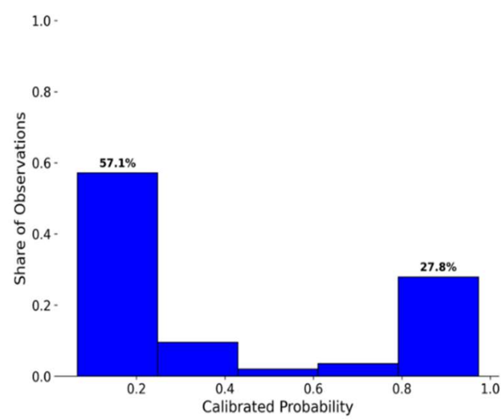
B Ex post deep learning (*Sackett*) – test set



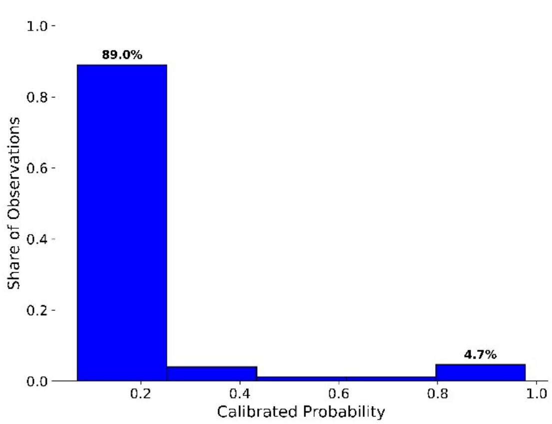
C Ex post deep learning (*Rapanos*) – 4mn points



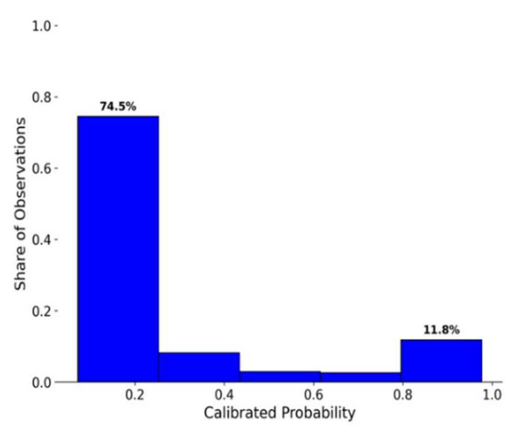
D Ex post deep learning (*Rapanos*) – test set



E Ex post deep learning (NWPR) – 4mn points



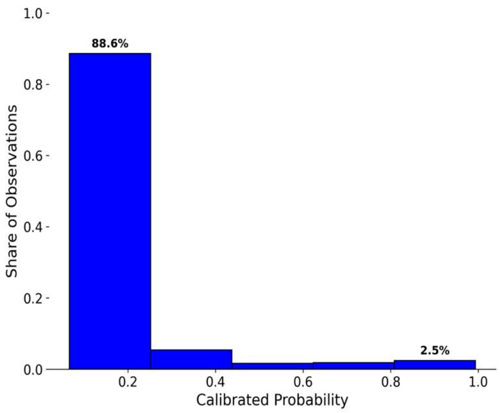
F Ex post deep learning (NWPR) – test set



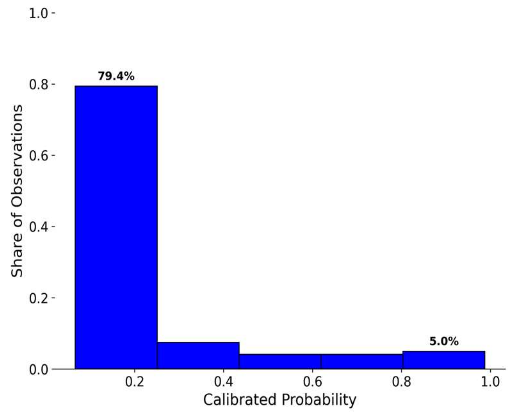
(continued next page)

(continued from previous page)

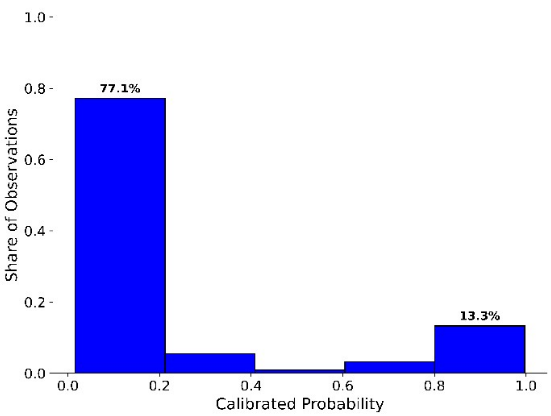
G Ex ante deep learning (*Sackett*) – 4mn points



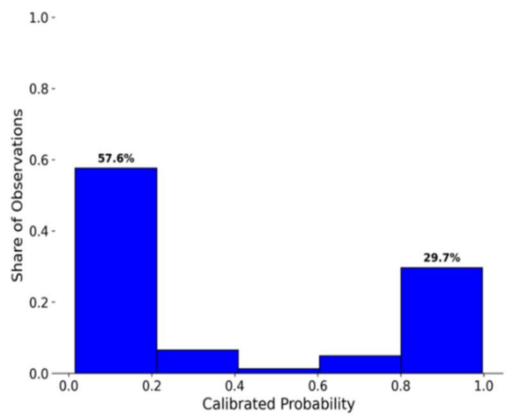
H Ex ante deep learning (*Sackett*) – test set



I Ex post deep learning (CWR) – 4mn points

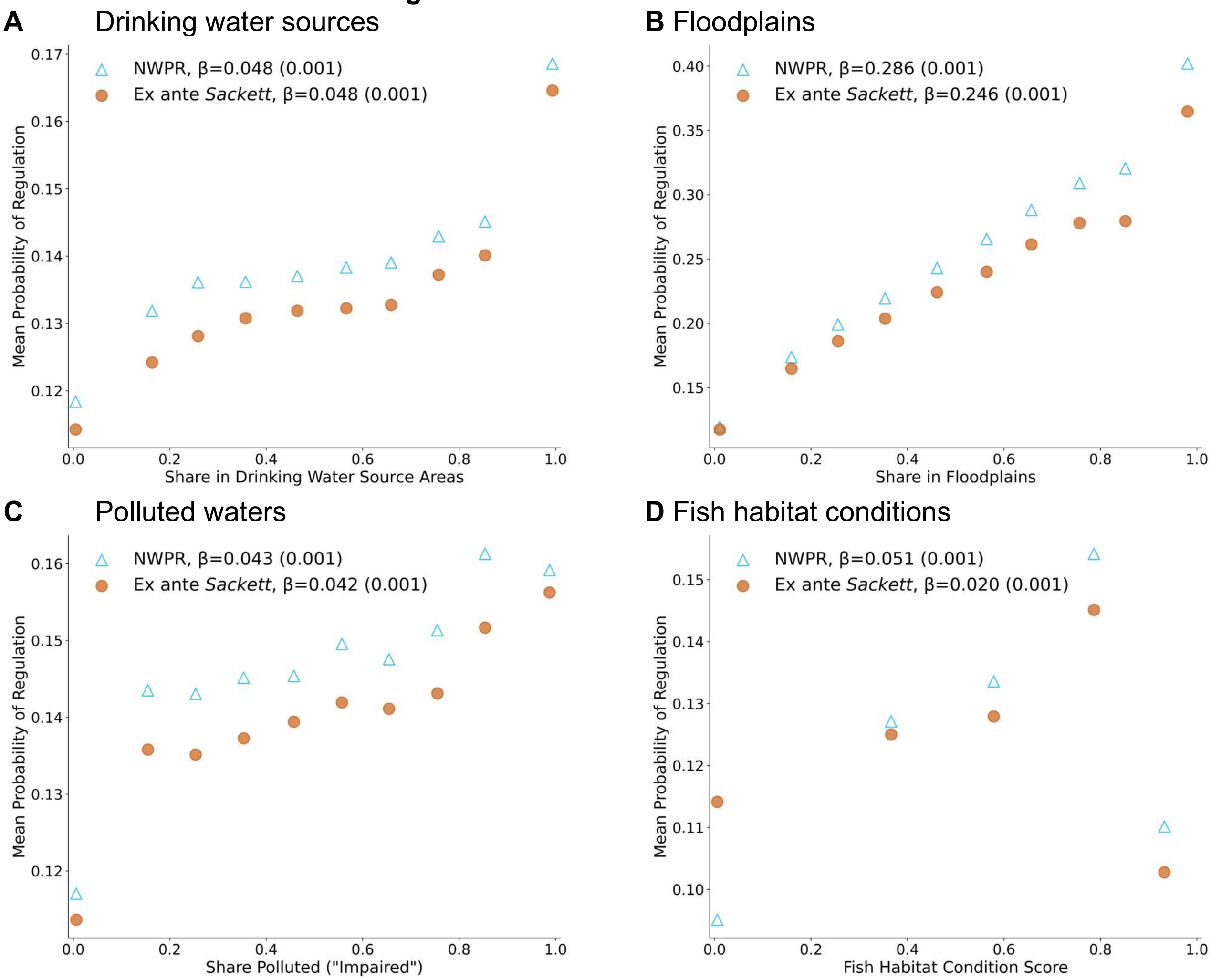


J Ex post deep learning (CWR) – test set



For each rule and for either the 4 million prediction points or the test set, each graph shows the share of points with a calibrated probability in one of five evenly sized bins spanning 0.0 to 1.0. Across all rules, and in both the 4 million random prediction points and the test set, around 90 percent of sites have calibrated probabilities below 0.2 or above 0.8, indicating that the model has high confidence. The test set has higher jurisdictional probabilities than the 4 million random prediction points because AJDs disproportionately represent sites with potential water resources.

Fig. S10. Relabeling captures *Sackett*'s deregulation of areas with concentrated ecosystem services and relevant to CWA goals.



(**A**) Share of points in drinking water source areas. (**B**) Share of points in floodplains. (**C**) Proportion of assessed waters considered "impaired" based on ambient pollution and relevant standards. (**D**) Fish habitat condition score. Each panel splits 4 million random points into the 251,975 5km by 5 km grid cells used to plot Fig. 4. In each graph, the y-axis shows the mean calibrated probability from ex post deep learning (NWPR) and ex ante deep learning, and the x-axis shows the mean value within the grid cell. The x-axis divides grid cells into equal-width bins (0–1 scale) based on underlying values. The legend shows the grid-level regression coefficient with standard errors in parentheses. Impaired waters and fish habitat conditions measured by 12-digit hydrologic unit code (HUC12) from the EPA's 2025 Restoration and Protection Indicator Database (28).

Table S1: Geophysical models modestly improve on naïve benchmark, ex ante deep learning does better, ex post deep learning has strongest performance.

	AUC	**F1**	**Precision**	**Recall**	**Accuracy**	**MAE**
	(1)	(2)	(3)	(4)	(5)	(6)
a Naïve benchmark						
No jurisdiction	0.500	0.000	—	0.000	0.803	0.197
b Geophysical models						
1. Wetness	0.498	0.007	0.118	0.004	0.798	0.191
2. Connected	0.512	0.065	0.463	0.035	0.802	0.183
c Ex ante deep learning model						
3. *Sackett*	0.693	0.332	0.457	0.261	0.802	0.066
d Ex post deep learning model						
4. *Sackett*	0.691	0.368	0.502	0.290	0.819	0.001

All statistics use AJD test set. AUC: Area under the receiver operating curve. All models describe *Sackett*. F1: harmonic mean of precision and recall. Precision: TP / (TP + FP), where TP is the count of true positive predictions and FP is the count of false positive predictions. Recall: TP / (TP + FN), where FN is the count of false negative predictions. Precision is undefined if a model makes no positive predictions. Accuracy: percent correct. MAE equals $|\mathrm{mean}(J_i) - \mathrm{mean}(C_i)|$, where J_i represents AJD jurisdiction and C_i represents model predictions. Row 1 describes a naïve benchmark that predicts no location is jurisdictional. Row 2 describes the median Wetness model (1), "seasonally flooded." Row 3 defines points as jurisdictional if they fall within a potential regulatory National Wetlands Inventory (NWI) polygon that connects with a perennial or intermittent National Hydrography Dataset (NHD) flowline. Row 4 describes the ex ante deep learning model projection of *Sackett* using ex ante data. Row 5 describes the ex post *Sackett* deep learning model. Rows 4 and 5 show performance of calibrated probabilities with thresholds optimized for performance for F1 in columns (2), (3), and (4), accuracy in column (5), and national mean absolute error (MAE) in column (7). Column (1) depends on model calibrated probabilities and is independent of threshold choice. SI Appendix, Table S14 and Fig. S8, show the thresholds. N = 2,777.

Table S2. Relabeling NWPR AJDs allows training of ex ante deep learning

			Jurisdictional under	
	Definition	**Share of AJDs**	**NWPR**	**Ex ante deep learning**
	(1)	(2)	(3)	(4)
A1TNW10	(a)(1) Water is also subject to Sections 9 or 10 of the Rivers and Harbors Act - RHA Tidal water is subject to the ebb and flow of the tide	0.0035	Yes	Yes
A1TNWCOMM	(a)(1) Water is currently used, was used in the past, or may be susceptible to use in interstate or foreign commerce (CWA Section 404 only)	0.00067	Yes	Yes
A1TNWFED	(a)(1) A federal court has determined the water is navigable in fact under federal law	0.00011	Yes	Yes
A1TNWSEAS	(a)(1) Territorial Seas	6.4E-05	Yes	Yes
A2TRIBINT	(a)(2) Intermittent tributary contributes surface water flow directly or indirectly to an (a)(1) water in a typical year	0.072	Yes	Yes
A2TRIBPER	(a)(2) Perennial tributary contributes surface water flow directly or indirectly to an (a)(1) water in a typical year	0.039	Yes	Yes
A3LPIFLOOD	(a)(3) Lake/pond or impoundment of a jurisdictional water inundated by flooding from an (a)(1)-(a)(3) water in a typical year	0.0013	Yes	Yes
A3LPIFLOW	(a)(3) Lake/pond or impoundment of a jurisdictional water contributes surface water flow directly or indirectly to an (a)(1) water in a typical year	0.0057	Yes	Yes
A4WETABUT	(a)(4) Wetland abuts an (a)(1)-(a)(3) water	0.11	Yes	Yes
A4WETARTSEP	**(a)(4) Wetland separated from an (a)(1)-(a)(3) water only by an artificial structure allowing a direct hydrologic surface connection between the wetland and the (a)(1)-(a)(3) water in a typical year**	**0.0076**	**Yes**	**No**
A4WETFLOOD	**(a)(4) Wetland inundated by flooding from an (a)(1)-(a)(3) water in a typical year**	**0.0081**	**Yes**	**Yes**
A4WETNATSEP	**(a)(4) Wetland separated from an (a)(1)-(a)(3) water only by a natural feature**	**0.0027**	**Yes**	**No**
B10STORM	(b)(10) Stormwater control feature constructed or excavated in upland or in a non-jurisdictional water to convey, treat, infiltrate, or store stormwater runoff	0.02	No	No
B11REUSE	(b)(11) Groundwater recharge, water reuse, or a wastewater recycling structure constructed or excavated in upland or in a non-jurisdictional water	0.00024	No	No
B12WTS	(b)(12) Waste treatment system	0.0018	No	No

(Continued next page)

Table S2. Relabeling NWPR AJDs allows training of ex ante deep learning (Continued)

	Definition	Share of AJDs	NWPR	Jurisdictional under: Ex ante deep learning
	(1)	(2)	(3)	(4)
B1EXCLUDEDOTH	(b)(1) Water or water feature that is not identified in (a)(1)-(a)(4) and does not meet the other (b)(1) sub-categories	0.011	No	No
B1LPINOSCFLD	(b)(1) Lake/pond or impoundment that does not contribute surface water flow directly or indirectly to an (a)(1) water and is not inundated by flooding from an (a)(1)-(a)(3) water in a typical year	0.014	No	No
B1SWCNOSC	(b)(1) Surface water channel that does not contribute surface water flow directly or indirectly to an (a)(1) water in a typical year	0.0078	No	No
B1WETNONADJ	(b)(1) Non-adjacent wetland	0.31	No	No
B2GRNDWATER	(b)(2) Groundwater, including groundwater drained through subsurface drainage systems	0.00011	No	No
B3EPHEMERAL	(b)(3) Ephemeral feature, including an ephemeral stream, swale, gully, rill, or pool	0.22	No	No
B4SHEETFLOW	(b)(4) Diffuse stormwater run-off over upland or directional sheet flow over upland	0.0016	No	No
B5DITCH	(b)(5) Ditch that is not an (a)(1) or (a)(2) water, and those portions of a ditch constructed in an (a)(4) water that do not satisfy the conditions of (c)(1)	0.094	No	No
B6PCC	(b)(6) Prior converted cropland	0.0053	No	No
B7ARTIRR	(b)(7) Artificially irrigated area, including fields flooded for agricultural production, that would revert to upland should application of irrigation water to that area cease	0.0013	No	No
B8LPIART	(b)(8) Artificial lake/pond constructed or excavated in upland or a non-jurisdictional water, so long as the artificial lake or pond is not an impoundment of a jurisdictional water that meets (c)(6)	0.028	No	No
B9DEPPIT	(b)(9) Water-filled depression constructed/excavated in upland/non-jurisdictional water incidental to mining/construction or pit excavated in upland/non-jurisdictional water to obtain fill/sand/gravel	0.0058	No	No
DRYLAND	The review area is comprised entirely of dry land (i.e. There are no waters or water features, including wetlands, of any kind in the entire review area)	0.034	No	No
RHA10NAV	RHA Non-tidal water is on the district's Section 10 waters list	0.00045	No	No

(Continued next page)

Table S2. Relabeling NWPR AJDs allows training of ex ante deep learning models. (Continued)

			Jurisdictional under	
	Definition	**Share of AJDs**	**NWPR**	**Ex ante deep learning**
	(1)	(2)	(3)	(4)
RHAB10STORM	Rivers and Harbors Act Section 10 water excluded from the CWA as a (b)(10) stormwater control feature constructed or excavated in upland or in a non-jurisdictional water to convey, treat, infiltrate, or store stormwater runoff	0.00010	No	No
RHAB1EXCLUDEDOTH	Rivers and Harbors Act Section 10 water excluded from the CWA as a (b)(1) water or water feature that is not identified in (a)(1)-(a)(4) and does not meet the other (b)(1) sub-categories	0.000016	No	No
RHAB1LPINOSCFLD	Rivers and Harbors Act Section 10 water excluded from the CWA as a (b)(1) lake/pond or impoundment that does not contribute surface water flow directly or indirectly to an (a)(1) water and is not inundated by flooding from an (a)(1)-(a)(3) water in a typical year	0.000016	No	No
RHAB1WETNONADJ	Rivers and Harbors Act Section 10 water excluded from the CWA as a (b)(1) non-adjacent wetland	0.0014	No	No
RHAB3EPHEMERAL	Rivers and Harbors Act Section 10 water excluded from the CWA as a (b)(3) ephemeral feature, including an ephemeral stream, swale, gully, rill, or pool	0.00032	No	No
RHAB6PCC	Rivers and Harbors Act Section 10 water excluded from the CWA as (b)(6) prior converted cropland	0.000016	No	No
RHATIDAL	RHA Tidal water is subject to the ebb and flow of the tide	0.00075	No	No

Each row describes one NWPR resource type. Ex ante deep learning relabeled resource types appear in bold. Column (2) shows non-synthetic AJDs for each resource type as a share of all NWPR AJDs.

Table S3. Ex post and ex ante deep learning models project that *Sackett* regulates relatively few water resources.

	Naïve benchmark (1)	Geophysical: Wetness (Gold) (2)	Geophysical: Connected (3)	Deep learning: Ex ante *Sackett* (4)	Deep learning: Ex post *Sackett* (5)
a General groups of points					
All 4 million points	0.000	0.026	0.017	0.134	0.115
AJD test set	0.000	0.006	0.015	0.204	0.161
b Rivers and streams					
All (NHD all)	0.000	0.067	0.109	0.360	0.250
Perennial	0.000	0.122	0.194	0.502	0.348
Intermittent or ephemeral	0.000	0.033	0.065	0.232	0.138
None (not in NHD)	0.000	0.025	0.015	0.129	0.113
c Wetlands					
All (NWI palustrine)	0.000	0.165	0.109	0.314	0.279
Non-tidal wetlands	0.000	0.524	0.336	0.286	0.319
Emergent (NWI)	0.000	0.348	0.157	0.194	0.199
Forested (NWI)	0.000	0.330	0.278	0.291	0.284
None (not in NWI palustrine)	0.000	0.001	0.001	0.102	0.087
d Rivers, streams, and wetlands					
All (NWI all, NHD all)	0.000	0.161	0.107	0.311	0.275
None (not in NWI or NHD)	0.000	0.001	0.001	0.102	0.086
e Other important groups of points					
Cropland and pasture (NLCD)	0.000	0.008	0.004	0.098	0.082
Floodplains (NFIP)	0.000	0.179	0.124	0.353	0.333
Urban growth areas (ICLUS)	0.000	0.017	0.012	0.133	0.093
Urban developed (NLCD)	0.000	0.007	0.004	0.116	0.087

Values represent share of points regulated. Columns (4)–(5) average calibrated probabilities. Column (1) describes a naïve model where no points are jurisdictional. Column (2) describes the median scenario from the original wetness model (1), "seasonally flooded." Column (3) defines points as jurisdictional in "potentially regulated" NWI polygons that intersect with perennial or intermittent NHD flowlines. Column (4) describes the ex ante deep learning projection of *Sackett*, which relabels resource types in NWPR AJDs. Column (5) describes the ex post deep learning model of *Sackett*. (**B**)–(**E**) describe subsets of the four million prediction points. NHD includes areas within 5 m of perennial, intermittent, and ephemeral flowline feature codes (fcodes) 46006, 46003, and 46007. Non-tidal wetlands include wetlands analyzed in the original wetness model (27). NHD is National Hydrography Dataset, NWI is National Wetlands Inventory, NLCD is National Land Cover Dataset, NFIP is National Insurance Program, ICLUS is Integrated Climate and Land-Use Scenarios, DL is deep learning.

Table S4. Ex ante and ex post deep learning outperform different wetness scenarios

	AUC (1)	F1 (2)	Precision (3)	Recall (4)	Accuracy (5)	MAE (6)
A All sites (N=2,777)						
1 Temporarily flooded	0.499	0.021	0.167	0.011	0.794	0.184
2 Seasonally saturated	0.501	0.021	0.214	0.011	0.797	0.187
3 Continuously saturated	0.498	0.007	0.118	0.004	0.798	0.191
4 Seasonally flooded	0.498	0.007	0.118	0.004	0.798	0.191
5 Seasonally flooded/saturated	0.498	0.004	0.083	0.002	0.799	0.193
6 Semi-permanently flooded	0.500	0.004	0.250	0.002	0.802	0.196
7 Intermittently exposed	0.501	0.004	0.500	0.002	0.803	0.197
8 Permanently flooded	0.501	0.004	0.500	0.002	0.803	0.197
9 Naïve	0.500	0.000	0.000	0.000	0.803	0.197
10 Connected	0.512	0.065	0.463	0.035	0.802	0.183
11 Ex ante DL (*Sackett*)	0.693	0.332	0.457	0.261	0.802	0.066
12 Ex post DL (*Sackett*)	0.691	0.368	0.502	0.290	0.819	0.001
B Non-tidal NWI (Emergent, Forested, Pond) (N=640)						
1 Temporarily flooded	0.502	0.030	0.250	0.016	0.800	0.181
2 Seasonally saturated	0.503	0.031	0.286	0.016	0.802	0.183
3 Continuously saturated	0.496	0.000	0.000	0.000	0.800	0.188
4 Seasonally flooded	0.496	0.000	0.000	0.000	0.800	0.188
5 Seasonally flooded/saturated	0.497	0.000	0.000	0.000	0.802	0.189
6 Semi-permanently flooded	0.500	0.000	0.000	0.000	0.806	0.194
7 Intermittently exposed	0.500	0.000	0.000	0.000	0.806	0.194
8 Permanently flooded	0.500	0.000	0.000	0.000	0.806	0.194
9 Naïve	0.500	0.000	0.000	0.000	0.806	0.194
10 Connected	0.517	0.088	0.462	0.048	0.805	0.173
11 Ex ante DL (*Sackett*)	0.703	0.370	0.487	0.298	0.820	0.056
12 Ex post DL (*Sackett*)	0.724	0.424	0.568	0.339	0.825	0.003

(Continued next page)

Table S4. Ex ante and ex post deep learning outperform different wetness scenarios (continued)

	AUC (1)	**F1** (2)	**Precision** (3)	**Recall** (4)	**Accuracy** (5)	**MAE** (6)
C Non-tidal wetlands (27) (N=36)						
1 Temporarily flooded	0.500	0.286	0.167	1.000	0.167	0.833
2 Seasonally saturated	0.633	0.353	0.214	1.000	0.389	0.611
3 Continuously saturated	0.417	0.174	0.118	0.333	0.472	0.306
4 Seasonally flooded	0.417	0.174	0.118	0.333	0.472	0.306
5 Seasonally flooded/saturated	0.400	0.111	0.083	0.167	0.556	0.167
6 Semi-permanently flooded	0.533	0.200	0.250	0.167	0.778	0.056
7 Intermittently exposed	0.567	0.250	0.500	0.167	0.833	0.111
8 Permanently flooded	0.567	0.250	0.500	0.167	0.833	0.111
9 Naïve	0.500	0.000	0.000	0.000	0.833	0.167
10 Connected	0.483	0.000	0.000	0.000	0.806	0.139
11 Ex ante DL (Sackett)	0.819	0.200	0.250	0.167	0.806	0.000
12 Ex post DL (Sackett)	0.947	0.625	0.500	0.833	0.917	0.139

MAE: mean absolute error in predicted share jurisdictional in US or state. AUC-ROC: Area under the receiver operating curve. F1: harmonic mean of precision and recall. Precision: TP / (TP + FP), where TP is the count of true positive predictions and FP is the count of false positive predictions. Recall: TP / (TP + FN), where FN is the count of false negative predictions. Accuracy: percent correct. Column (6) equals $|\mathrm{mean}(J_i) - \mathrm{mean}(C_i)|$, where J_i represents AJD jurisdiction and C_i represents model-predicted jurisdiction. Each panel describes jurisdiction predicted by scenarios analyzed in Gold (27). Each scenario indicates how "wet" a wetland must be to be protected under the Clean Water Act. In other words, in scenario 4, all AJDs within wetlands (27) at least as wet as "seasonally flooded" are predicted as WOTUS; all others are predicted as non-WOTUS. The median scenario ex-ante, Scenario 4, is used as the wetness model throughout the rest of the paper. (**A**), N=2,777. (**B**), N=640. (**C**), N=36.

Table S5. For all rules, ex post deep learning performs well but individual input layers perform poorly.

	AUC (1)	F1 (2)	Precision (3)	Recall (4)	Accuracy (5)	US (6)	N (7)
A Ex post deep learning, by rule							
1. All rules	0.837	0.665	0.787	0.575	0.811	0.088	20,844
2. *Sackett*	0.691	0.368	0.502	0.290	0.819	0.001	2,777
3. Rapanos	0.864	0.761	0.737	0.786	0.819	0.005	10,187
4. NWPR	0.805	0.603	0.561	0.652	0.801	0.010	6,373
5. CWR	0.856	0.748	0.756	0.740	0.802	0.101	1,507
B Individual input layers—*Sackett*							
5. Wetland (NWI)	0.502	0.253	0.199	0.349	0.595	0.148	2,777
6. Stream (NHD)	0.492	0.058	0.146	0.036	0.768	0.148	2,777
7. Wetland or stream	0.499	0.253	0.196	0.354	0.587	0.158	2,777
8. Hydric soil (gNATSGO)	0.492	0.295	0.194	0.624	0.413	0.439	2,777
9. Water table <10m (gNATSGO)	0.494	0.130	0.179	0.102	0.731	0.085	2,777
10. Cropland and pasture (NLCD)	0.496	0.308	0.195	0.728	0.355	0.538	2,777
11. Urban developed (NLCD)	0.491	0.303	0.193	0.701	0.364	0.518	2,777

(**A**), performance of ex post deep learning calibrated probabilities with thresholds optimized for performance for F1 in columns (2), (3), and (4), accuracy in columns (5), and national mean absolute error (MAE) in column (6). Column (1) depends on model calibrated probabilities and is independent of threshold choice. (**B**), forecasting based on individual layers on for *Sackett* AJDs. MAE: mean absolute error in predicted share jurisdictional in US or state. AUC-ROC: Area under the receiver operating curve. F1: harmonic mean of precision and recall. Precision: TP / (TP + FP), where TP is the count of true positive predictions and FP is the count of false positive predictions. Recall: TP / (TP + FN), where FN is the count of false negative predictions. Recall is not defined if a model makes no positive predictions. Accuracy: percent correct. Column (6) equals $|\mathrm{mean}(J_i) - \mathrm{mean}(C_i)|$, where J_i represents AJD jurisdiction and C_i represents model-predicted jurisdiction. Row 5 predicts regulation if within 5 m of a NWI wetland; row 6 if within 5 m of an NHD stream; row 7 if within 5 m of either an NWI wetland or NHD stream; row 8 if the area has a hydric soil according to the Gridded National Soil Survey Geographic Database (gNATSGO). Row 9 predicts regulation if the water table is less than 10 meters deep, and no regulation everywhere else. Rows 10 and 11 predict no regulation in cropland and pasture, and urban developed areas, respectively, and regulation everywhere else. NWPR is the Navigable Waters Protection Rule. CWR is the Clean Water Rule.

Table S6. Wetness models project a wide range of jurisdiction.

	Temporarily flooded	Seasonally saturated	Continuously saturated	Seasonally flooded	Seasonally flooded/saturated	Semi-permanently flooded	Intermittently exposed	Permanently flooded
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A General groups of points								
All 4 million points	0.049	0.036	0.030	0.026	0.010	0.008	0.003	0.003
AJD test set	0.013	0.01	0.006	0.006	0.004	0.001	0.001	0.001
B Rivers and streams								
All (NHD all)	0.129	0.077	0.070	0.067	0.020	0.012	0.003	0.003
Perennial	0.221	0.139	0.126	0.122	0.040	0.020	0.005	0.005
Intermittent or ephemeral	0.075	0.039	0.034	0.033	0.005	0.004	0.000	0.000
None (not in NHD)	0.047	0.036	0.030	0.025	0.010	0.008	0.003	0.003
C Wetlands								
All (NWI palustrine)	0.307	0.231	0.192	0.165	0.069	0.052	0.019	0.019
Non-tidal wetlands (27)	0.994	0.742	0.622	0.524	0.213	0.163	0.059	0.059
Emergent (NWI)	0.546	0.401	0.376	0.348	0.118	0.099	0.013	0.013
Forested (NWI)	0.711	0.534	0.412	0.330	0.122	0.077	0.007	0.007
None (not in NWI palustrine)	0.004	0.002	0.002	0.001	0.000	0.000	0.000	0.000
D Rivers, streams, and wetlands								
All (NWI all, NHD all)	0.300	0.226	0.188	0.161	0.067	0.051	0.019	0.018
None (not in NWI or NHD)	0.003	0.002	0.002	0.001	0.000	0.000	0.000	0.000
E Other important groups of points								
Cropland and pasture (NLCD)	0.016	0.009	0.008	0.008	0.003	0.002	0.002	0.002
Floodplains (NFIP)	0.276	0.198	0.181	0.179	0.085	0.078	0.033	0.033
Urban growth areas (ICLUS)	0.034	0.023	0.018	0.017	0.005	0.004	0.002	0.002
Urban developed (NLCD)	0.012	0.008	0.007	0.007	0.002	0.001	0.001	0.001

Each column shows one scenario from the Wetness model (1). Model numbers in (1) correspond to column numbers here. Table shows the share of points each framework estimates are regulated. Panels B through D describe subsets of the four million prediction points. NHD only refers to flowlines. NFIP is the National Flood Insurance Program and ICLUS is the Integrated Climate and Land Use Scenarios.

Table S7. *Sackett* regulates less than earlier CWA rules.

	CWR	*Rapanos*	NWPR	*Sackett*
	(1)	(2)	(3)	(4)
A General groups of points				
All 4 million points	0.230	0.179	0.138	0.115
AJD test set	0.402	0.383	0.246	0.161
B Rivers and streams				
All (NHD all)	0.524	0.463	0.427	0.249
Perennial	0.713	0.615	0.602	0.347
Intermittent or ephemeral	0.373	0.336	0.287	0.137
None (not in NHD)	0.225	0.173	0.133	0.112
C Wetlands				
All (NWI palustrine)	0.567	0.410	0.351	0.278
Non-tidal wetlands(27)	0.689	0.463	0.389	0.318
Emergent (NWI)	0.509	0.343	0.234	0.199
Forested (NWI)	0.702	0.455	0.400	0.284
None (not in NWI palustrine)	0.172	0.139	0.101	0.087
D Rivers, streams, and wetlands				
All (NWI all, NHD all)	0.561	0.407	0.349	0.275
None (not in NWI or NHD)	0.171	0.138	0.100	0.086
E Other important groups of points				
Cropland and pasture (NLCD)	0.147	0.124	0.098	0.082
Floodplains (NFIP)	0.611	0.459	0.386	0.333
Urban growth areas (ICLUS)	0.244	0.169	0.135	0.093
Urban developed (NLCD)	0.211	0.139	0.110	0.087

Table shows the share of points each framework estimates are regulated. Columns (1)–(4) average calibrated probabilities from ex post deep learning. Column (1) describes regulation under the Clean Water Rule (CWR). Column (2) describes regulation under *Rapanos*. Column (3) describes regulation under NWPR. Column (4) duplicates column (5) from Table S3. (**B**)–(**D**) describe subsets of the four million prediction points. NHD only refers to flowlines. NFIP is the National Flood Insurance Program and ICLUS is the Integrated Climate and Land Use Scenarios.

Table S8. Regulated stream miles and wetland acres, by state.

State	Total Stream Miles	Total Wetland Acres	Stream Miles Regulated			Wetland Acres Regulated		
			Rapanos (share)	*Sackett* (share)	Difference *Sackett - Rapanos*	*Rapanos* (share)	*Sackett* (share)	Difference *Sackett - Rapanos*
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
National	3,154,478	119,825,265	-	-	-705,047	-	-	-19,321,637
Alabama	72,650	4,043,348	0.49	0.35	-10,752	0.41	0.38	-109,170
Arizona	139,281	262,281	0.16	0.09	-9,610	0.32	0.18	-36,457
Arkansas	78,496	2,558,428	0.48	0.25	-18,525	0.39	0.30	-235,375
California	173,028	2,789,804	0.40	0.15	-42,565	0.39	0.15	-694,661
Colorado	93,255	1,522,952	0.28	0.14	-13,056	0.25	0.13	-184,277
Connecticut	5,215	304,750	0.94	0.41	-2,717	0.84	0.19	-196,259
Delaware	2,234	290,940	0.79	0.30	-1,097	0.59	0.30	-84,954
Florida	22,385	12,681,770	0.76	0.60	-3,604	0.68	0.45	-2,916,807
Georgia	64,833	6,396,737	0.45	0.36	-5,381	0.28	0.27	-6,397
Idaho	94,753	1,119,249	0.45	0.38	-6,254	0.51	0.33	-194,749
Illinois	67,074	1,271,986	0.60	0.18	-27,970	0.56	0.21	-443,923
Indiana	24,066	1,008,100	0.51	0.16	-8,543	0.29	0.13	-160,288
Iowa	67,717	1,014,174	0.62	0.17	-30,473	0.45	0.14	-323,522
Kansas	118,236	1,349,856	0.30	0.11	-23,293	0.24	0.09	-206,528
Kentucky	45,616	430,781	0.17	0.09	-3,786	0.23	0.16	-27,139
Louisiana	43,096	8,092,819	0.59	0.59	-259	0.64	0.68	283,249
Maine	24,974	2,569,961	0.73	0.18	-13,961	0.63	0.14	-1,256,711
Maryland	10,263	863,198	0.88	0.43	-4,680	0.80	0.44	-308,162
Massachusetts	7,273	775,106	0.75	0.23	-3,767	0.54	0.15	-302,291
Michigan	47,861	7,712,081	0.86	0.36	-24,122	0.68	0.32	-2,814,909
Minnesota	60,103	9,973,334	0.16	0.13	-1,623	0.09	0.09	-9,973
Mississippi	77,386	4,534,181	0.40	0.32	-5,881	0.38	0.45	321,927
Missouri	95,347	1,388,966	0.63	0.16	-44,813	0.43	0.18	-352,797
Montana	166,847	1,589,844	0.28	0.23	-8,843	0.30	0.20	-163,754
Nebraska	72,506	549,755	0.33	0.14	-13,269	0.30	0.14	-87,961
Nevada	143,616	1,003,174	0.33	0.11	-30,878	0.44	0.18	-258,819
New Hampshire	9,374	384,706	0.71	0.19	-4,790	0.55	0.12	-163,115
New Jersey	7,128	1,019,092	0.90	0.40	-3,557	0.73	0.30	-440,248
New Mexico	109,260	383,873	0.11	0.10	-983	0.14	0.12	-9,213
New York	48,756	2,651,158	0.67	0.21	-22,428	0.43	0.13	-816,557

(Continued next page)

Table S8. Regulated stream miles and wetland acres, by state. (Continued)

North Carolina	56,673	4,679,517	0.92	0.50	-23,916	0.84	0.50	-1,600,395
North Dakota	59,514	2,442,160	0.45	0.24	-12,379	0.18	0.11	-180,720
Ohio	54,736	715,219	0.36	0.11	-13,465	0.27	0.13	-99,415
Oklahoma	75,615	1,274,713	0.67	0.19	-35,766	0.56	0.22	-432,128
Oregon	102,984	1,803,096	0.46	0.26	-20,185	0.50	0.22	-497,655
Pennsylvania	51,477	588,835	0.77	0.31	-23,782	0.72	0.35	-219,047
Rhode Island	978	86,061	0.88	0.19	-679	0.64	0.14	-43,203
South Carolina	29,372	4,238,935	0.82	0.48	-9,898	0.67	0.39	-1,191,141
South Dakota	96,965	3,529,693	0.54	0.24	-29,283	0.27	0.13	-465,919
Tennessee	59,244	1,148,777	0.26	0.13	-7,820	0.36	0.23	-153,936
Texas	176,194	5,551,483	0.56	0.25	-54,973	0.59	0.36	-1,276,841
Utah	82,724	624,397	0.44	0.15	-23,494	0.37	0.24	-83,045
Vermont	7,100	287,628	0.47	0.11	-2,542	0.29	0.09	-56,375
Virginia	49,280	1,682,396	0.83	0.43	-19,909	0.79	0.55	-408,822
Washington	68,964	1,297,395	0.42	0.23	-13,034	0.45	0.29	-206,286
West Virginia	30,572	81,858	0.40	0.13	-8,193	0.57	0.19	-30,942
Wisconsin	53,370	7,610,528	0.34	0.23	-5,550	0.14	0.13	-22,832
Wyoming	106,082	1,646,169	0.25	0.17	-8,699	0.25	0.16	-153,094

Total stream miles in column (2) is from NHD stream and river flowline features. Total wetland acres in column (3) is from NWI. Regulation rates in columns (3), (4), (6), and (7) display calibrated probabilities from ex post deep learning (*Sackett* and *Rapanos*), applied to the subset of four million prediction points that are within 5 meters of NHD or NWI features. The difference in column (5) is measured in stream miles, and in column (8) in wetland acres.

Table S9. Recent rules deregulate drinking water sources.

	Rapanos	**NWPR**	***Sackett***
	(1)	(2)	(3)
A Share Regulated			
1. All points	0.243	0.187	0.144
2. NHD or NWI points	0.523	0.448	0.366
3. NHD points	0.628	0.578	0.350
4. NWI points	0.525	0.449	0.369
B Pop. Served Weighted			
1. All points	0.262	0.180	0.139
2. NHD or NWI points	0.593	0.463	0.391
3. NHD points	0.637	0.565	0.355
4. NWI points	0.596	0.465	0.394

Columns show the results from ex post deep learning models fine-tuned on each CWA rule. 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). (**A**), share of prediction points within HUC12 areas that serve as drinking water inputs for an active 2019 CWS predicted as jurisdictional under each regime. (**B**), same share weighted by the population served by each CWS. NWPR is the Navigable Waters Protection Rule.

Table S10. *Sackett* divides AJDs into resource types corresponding to different legal categorizations of waters.

	Definition	Share of AJDs	Share juris-dictional
	(1)	(2)	(3)
A Pre-2015-Post-Sackett			
A1.TNW-404	(a)(1) Traditional Navigable Water (Section 404 Only)	0.00076	1.00
A1.TNW-404.10	(a)(1) Traditional Navigable Water, also subject to Sections 9 or 10 of the Rivers and Harbors Act (Section 10/404)	0.0027	1.00
A2.INTSTATE-404	(a)(2) Interstate Waters (Section 404 Only)	0.00025	1.00
A4.IMPDT-404	(a)(4) Impoundments of waters otherwise defined as "waters of the United States"	0.0052	1.00
A5.TRIB-404	(a)(5) Tributaries of waters identified in paragraph (a)(1) through (4), where the tributary is a relatively permanent, standing or continuously flowing body of water	0.094	1.00
A7-AJD.WETL-404	(a)(7) Wetland adjacent to a non-wetland water identified in (a)(1) - (a)(6)	0.095	1.00
DRY.LAND	Dry Land - The review area is comprised entirely of dry land (i.e. there are no aquatic features, including wetlands, of any kind in the entire review area)	0.055	0.00
EXCL-PCC	(a)(8) Prior converted cropland	0.0027	0.00
EXCL-WTS	(a)(8) Waste treatment systems, including treatment ponds or lagoons, designed to meet the requirements of the Clean Water Act	0.0081	0.00
NON-JD - PREAMBLE - ART.IRR	Preamble water - Artificially irrigated areas which would revert to upland if the irrigation ceased	0.0024	0.00
NON-JD - PREAMBLE - ART.LAKE.POND	Preamble water - Artificial lake/pond created by excavating/diking dry land, used exclusively for purposes such as stock watering, irrigation, settling basins, or rice growing	0.04	0.00
NON-JD - PREAMBLE - ART.REF.SWIM.ORN	Preamble water - Artificial reflecting or swimming pools or other small ornamental bodies of water created by excavating and/or diking dry land to retain water for primarily aesthetic reasons	0.0023	0.00
NON-JD - PREAMBLE - NON-TIDAL.DITCH-DRY.LAND	Preamble water -Non-tidal drainage and irrigation ditches excavated on dry land	0.0033	0.00
NON-JD - PREAMBLE - WATERFILLED.DEP-PITS	Preamble water - Waterfilled depression created in dry land and pits excavated in dry land unless and until the operation is abandoned and resulting body of water meets definition of WOTUS	0.012	0.00

(Continued next page)

Table S10. *Sackett* divides AJDs into resource types corresponding to different legal categorizations of waters. (Continued)

	Definition	Share of AJDs	Share juris-dictional
	(1)	(2)	(3)
NON-JD - RAPANOS.GUIDE - DITCH	*Rapanos* Guidance - Ditches (including roadside ditches) excavated wholly in and draining only uplands and that do not carry a relatively permanent flow of water	0.09	0.00
NON-JD - RAPANOS.GUIDE - SWALE.EROSION	*Rapanos* Guidance - Swales or erosional features (e.g., gullies, small washes, characterized by low volume, infrequent, or short duration flow)	0.073	0.00
NON-WOTUS-LAKE.POND.NEGATIVE-A5	NON-WOTUS - Intrastate Lake or Pond that is not a tributary to a water identified in paragraphs (a)(1) through (4)	0.023	0.00
NON-WOTUS-STREAM.NEGATIVE-A5	NON-WOTUS - Intrastate Stream that is not a tributary to a water identified in paragraphs (a)(1) through (4)	0.021	0.00
NON-WOTUS-TRIB.NEGATIVE-A5	NON-WOTUS: Tributary to a water identified in paragraphs (a)(1) through (4), where the tributary is not a relatively permanent, standing or continuously flowing body of water	0.19	0.00
NON-WOTUS-WETL.NEGATIVE-A7	NON-WOTUS: Wetland that is not adjacent to a water identified in paragraph (a)(1) through (6)	0.28	0.00
RHA-10NAV	RHA - Non-tidal water is on the district's Section 10 waters list (Section 10 Only)	0.00025	1.00
RHA-10TIDAL	RHA - Tidal water is subject to the ebb and flow of the tide (Section 10 Only)	0.00013	1.00
B Amended-2023-Rule			
A1-1.TNW-404	(a)(1)(i) Traditional Navigable Water (Section 404 Only)	0.0083	1.00
A1-1.TNW-404.10	(a)(1)(i) Traditional Navigable Water, also subject to Sections 9 or 10 of the Rivers and Harbors Act (Section 10/404)	0.0033	1.00
A1-2.TERSEAS-404.10	(a)(1)(ii) Territorial Seas, also subject to Sections 9 or 10 of the Rivers and Harbors Act (Section 10/404)	0.0001	1.00
A1-3.INTSTATE-404	(a)(1)(iii) Interstate Waters (Section 404 Only)	0.0002	1.00
A2.IMPDT-404	(a)(2) Jurisdictional Impoundment (Section 404 Only)	0.0027	1.00
A3.TRIB-404	(a)(3) Tributary (Section 404 Only)	0.061	1.00
A4-1.ADJ.WET.A1-INTSTATE-404	(a)(4)(i) Adjacent Wetland, adjacent to (a)(1))(iii) Interstate Water	0.0006	1.00
A4-1.ADJ.WET.A1-TERSEAS-404	(a)(4)(i) Adjacent Wetland, adjacent to (a)(1)(ii) Territorial Sea	0.0002	1.00
A4-1.ADJ.WET.A1-TNW-404	(a)(4)(i) Adjacent Wetland, adjacent to (a)(1)(i) TNW	0.013	1.00
A4-2.ADJ.WET.A2&A3-404	(a)(4)(ii) Adjacent Wetland, adjacent to a relatively permanent paragraph (a)(2) Impoundment or (a)(3) Tributary (Section 404 Only)	0.067	1.00

(Continued next page)

Table S10. *Sackett* divides AJDs into resource types corresponding to different legal categorizations of waters. (Continued)

	Definition	Share of AJDs	Share juris-dictional
	(1)	(2)	(3)
A5.INTSTATE.LKPND-404	(a)(5) Intrastate Lake or Pond not Identified in Paragraphs (a)(1) through (4), that is a relatively permanent, standing or continuously flowing body of water (Section 404 Only)	0.0023	1.00
B1-EXCL-WTS	(b)(1) Waste Treatment System (Excluded)	0.0088	0.00
B2-EXCL-PCC	(b)(2) Wetland Excluded as Prior Converted Cropland designated by USDA (Excluded)	0.0007	0.00
B3-EXCL-DITCH	(b)(3) Ditches (including roadside ditches) excavated wholly in and draining only dry land and that do not carry a relatively permanent flow of water (Excluded)	0.11	0.00
B4-EXCL-ART.IRR	(b)(4) Artificially irrigated areas that would revert to dry land if the irrigation ceased (Excluded)	0.0022	0.00
B5-EXCL-ART.LK	(b)(5) Artificial lakes or ponds created in dry land, used exclusively for specific purposes (Excluded)	0.031	0.00
B6-EXCL-ART.REF	(b)(6) Artificial reflecting/swimming/ornamental pools; created by excavating or diking dry land to retain water for primarily aesthetic reasons (Excluded)	0.0036	0.00
B7-EXCL-WTF.DEP	(b)(7) Waterfilled depressions created in dry land incidental to construction activity and pits excavated in dry land, until abandoned (Excluded)	0.012	0.00
B8-EXCL-SWAL.EROS	(b)(8) Swales and erosional features (e.g., gullies, small washes) characterized by low volume, infrequent, or short duration flow (Excluded)	0.027	0.00
DRY.LAND	Dry Land - The review area is comprised entirely of dry land (i.e. there are no aquatic features, including wetlands, of any kind in the entire review area)	0.035	0.00
NON-WOTUS-INTSTATE-LKPND.NEGATIVE.A5	NON-WOTUS - Intrastate lake or pond not identified in paragraphs (a)(1 - 4) that is not relatively permanent or does not have a continuous surface connection to (a)(1) or (3) water	0.015	0.00
NON-WOTUS-INTSTATE-STRM.NEGATIVE.A3	NON-WOTUS - Intrastate stream that does not connect to a paragraph (a)(1) or (a)(2) water	0.011	0.00
NON-WOTUS-TRIB.NEGATIVE.A3	NON-WOTUS - Tributary evaluated under (a)(3) and determined to not be a relatively permanent water with a continuous surface connection to paragraph (a)(1) or (a)(3) water	0.18	0.00
NON-WOTUS-WET.NEGATIVE.A4	NON-WOTUS - Wetland that does not have a continuous surface connection to a paragraph (a)(1) water or to a relatively permanent paragraph (a)(2) impoundment or paragraph (a)(3) tributary	0.4	0.00
RHA-10NAV	RHA - Non-tidal water is on the district's Section 10 waters list (Section 10 Only)	0.0003	1.00

Each row lists a *Sackett* resource type from the AJD data. Column (1) describes the resource type, column (2) lists the share of all *Sackett* AJDs the resource type accounts for, and column (3) shows the share of the resource type AJDs that are jurisdictional.

Table S11. *Rapanos* divides AJDs into resource types corresponding to different legal categorizations of waters.

	Definition	Share of AJDs	Share jurisdictional
	(1)	(2)	(3)
IMPNDMNT	Impoundment of Jurisdictional Waters	0.011	0.71
ISOLATE	Isolated (interstate or intrastate) waters	0.34	0.000025
NRPW	Non-relatively Permanent Water that flows directly or indirectly into Traditional Navigable Water	0.052	0.63
NRPWW	Wetland Adjacent to Non-relatively Permanent Water that flows directly or indirectly into Traditional Navigable Water	0.029	0.87
RPW	Relatively Permanent Water that flows directly or indirectly into Traditional Navigable Water	0.098	1.00
RPWWD	Wetlands Directly Abutting Relatively Permanent Water that flows directly or indirectly into Traditional Navigable Water	0.11	1.00
RPWWN	Wetlands Adjacent but not Directly Abutting Relatively Permanent Water that flows directly or indirectly into Traditional Navigable Water	0.037	0.94
TNW	Traditional Navigable Water	0.032	1.00
TNWRPW	Traditional Navigable Water - Relatively Permanent Water	0.0007	0.99
TNWW	Wetlands Adjacent to Traditional Navigable Water	0.038	1.00
UPLAND	Uplands	0.26	0.000064

Each row lists a *Rapanos* resource type from the AJD data**.** Column (1) describes the resource type, column (2) lists the share of all *Rapanos* AJDs the resource type accounts for, and column (3) shows the share of the resource type AJDs that are jurisdictional.

Table S12. Tiner (2003) categorizes many types of isolated wetlands.

	Geographic Region
	(1)
Tiner (2003) Wetland Types	
Alvar wetlands	Level IV ecoregions 50ab (Cheboygan Lake Plain)
Channeled Scablands wetlands	Level IV ecoregion 10a (Channeled Scablands)
Cypress domes	None -- area is too large/no specific agreement
Delmarva pothole wetlands	Level IV ecoregion 63f (Delmarva uplands)
Desert spring wetlands	Level III ecoregions 14 (Mojave Basin and Range)
Fens	None -- area is too large/no specific agreement
Geysers	None -- area is too large/no specific agreement
Inactive floodplain wetlands	None -- area is too large/no specific agreement
Interdunal and intradunal wetlands	None -- area is too large/no specific agreement
Kettle hole wetlands	None -- area is too large/no specific agreement
Mid- and South Atlantic Wetlands	Mid- and South Atlantic Wetlands
Natural ponds	None -- area is too large/no specific agreement
Playas	Level III ecoregion 25 (High Plains)
Prairie potholes	Mann (1974) Prairie Pothole Region
Rainwater basin wetlands	Level IV ecoregion 27f (Rainwater Basin Plains)
Rock pool wetlands	None -- area is too large/no specific agreement
Salt flats and salt lake wetlands	Level III ecoregions 13 (Central Basin and Range)
Sandhills wetlands	Level III ecoregion 44 (Nebraska Sand Hills)
Seepage slope wetlands	None -- area is too large/no specific agreement
Sinkhole wetlands	Level IV ecoregions 69c (Greenbriar Karst), 71e
Tarn wetlands	None -- area is too large/no specific agreement
Volcanic-formed wetlands	Level IV ecoregions 1d (Coast Range Volcanics)

Table shows isolated wetland types from Tiner (15). Column (1) shows mapping to geographic regions.

Table S13. We generate synthetic non-jurisdictional training data within several categories of isolated wetlands from Tiner (15).

	Geographic Region	Tiner (15) Wetland Type(s)
	(1)	(2)
Cowardin Code		
PABG	Palustrine wetland, aquatic bed, intermittently exposed	Prairie potholes
PEM1A	Palustrine emergent persistent wetland, temporarily flooded	Playas; prairie potholes
Pf	Palustrine wetland, farmed	Prairie potholes
PUBFx	Palustrine wetland, unconsolidated bottom, semi-permanently flooded, excavated	Playas; prairie potholes
PUBHx	Palustrine wetland, unconsolidated bottom, permanently flooded, excavated	West Coast vernal pools
R4SBJ	Riverine wetland, surface flooding, intermittent	Desert spring wetlands; salt flats and salt lake wetlands

Table shows Cowardin codes (4) selected for non-jurisdictional synthetic training data, by Tiner (15) wetland type. See Section A.3 under “Synthetic Non-Jurisdictional Data: Isolated Wetlands.” Column (1) describes associated geographic regions and column (2) lists associated Tiner wetland types.

Table S14. Optimal thresholds for each metric allow calculation of model performance.

	Performance Metrics								
								MAE	
Metric optimized	**Threshold** (1)	**AUC** (2)	**F1** (3)	**Precision** (4)	**Recall** (5)	**Specificity** (6)	**Accuracy** (7)	**US** (8)	**State** (9)
MAE	0.173	0.691	0.393	0.392	0.394	0.850	0.760	**0.001**	0.153
State MAE	0.242	0.691	0.368	0.502	0.290	0.929	0.803	0.083	**0.182**
Accuracy	0.577	0.691	0.297	0.635	0.193	0.973	**0.819**	0.137	0.205
F1 Score	0.242	0.691	**0.368**	0.502	0.290	0.929	0.803	0.083	0.182

Table shows ex post deep learning (*Sackett*) model performance. In each row, we choose the threshold which maximizes the performance metric indicated. AUC does not depend on threshold choice so it is identical across cases. Column (1) lists the resulting threshold. Columns (2)–(9) show all performance metrics. Values in bold show the optimized performance values. Selection of thresholds in column (1) uses the validation set. Performance metrics in columns (2)–(9) use the *Sackett* test set AJDs. MAE is mean absolute error.

Table S15. Input layers build on the inputs from Table S3 of Greenhill et al. 2025

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	(18)
	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	(16)
National Hydrography Dataset (NHD) Plus V2	FCode	Water feature type (e.g., perennial stream, intermittent stream, coastline)	Vector	1:100,000	(14)
	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	(29)
US 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	(22)

(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	(21)
	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			
US Army Corps Regulatory Boundaries	District codes	Each ACE district is assigned a unique value.	Point	1:250,000	(30)
	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	(20)
	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	(19)
Coastal Change Analysis Program (CCAP)	Landcover	C-CAP has 25 land cover classes: background, unclassified, developed (high intensity), developed (medium intensity), developed (low intensity), developed (open space), cultivated crops, pasture/hay, grassland/herbaceous, deciduous forest, evergreen forest, mixed forest, scrub/shrub, palustrine forested wetland, palustrine scrub/shrub wetland, palustrine emergent wetland (persistent), estuarine forested wetland, estuarine scrub/shrub wetland, estuarine emergent wetland, unconsolidated shore, barren land, open water, palustrine aquatic bed, estuarine aquatic bed, tundra, perennial ice/snow	Raster	30 meters	(23)

(Continued next page)

Dataset	Input layer	Variable definition	Data type	Spatial Resolution	Source
Topologically Integrated Geographic Encoding and Referencing System (TIGER)/Line State boundaries	State codes	Each state is assigned a unique value	Vector	1:250,000	(31)
CWA Approved Jurisdictional Determinations	WOTUS rule	Three WOTUS rules: *Rapanos,* CWR, NWPR	Point	--	(32)