Deep Learning Projects Jurisdiction of New Clean Water Act Regulation

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

Affiliations:

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

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

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

[4] 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

Abstract: Policymakers and firms struggle to interpret changes to regulation. We introduce a relabeling methodology that uses deep learning to provide ex ante projections of new regulations' effects. We compare the performance of ex ante deep learning and geophysical models against ex post algorithms, all analyzing judicial and executive branch Clean Water Act rules. Geophysical projections modestly beat a naïve benchmark; an ex ante deep learning model performs much better, though ex post analysis has the strongest performance. Several months of implementation provide sufficient data to train an accurate ex post geo-foundation model, which indicates that the Supreme Court's 2023 *Sackett* decision and a March 2025 White House proposed rule deregulate more waters than most previous rules, particularly agricultural areas and floodplains.

Please do not circulate or post online. Until submitted for publication, references to the "test set" use the validation set. The model in this draft excludes the geo-foundation components, currently under implementation.

Policymakers and firms struggle to interpret changing regulations. Complexity creates one challenge—the Code of Federal Regulations fills 200,000 pages, and the Federal Register adds 90,000 pages annually (Crews 2023; Regulatory Studies Center 2024). Ambiguity creates another—many regulations qualitatively describe general requirements but provide imprecise guidance for specific applications. Ambiguity particularly complicates the analysis of potential future regulations, which inherently lack ex post implementation data.

Researchers, governments, and firms use deep learning to interpret existing regulations, with applications including energy, environmental, financial, health, judicial, and labor market policies (Coglianese and Lehr 2016; Engstrom et al. 2020; Kleinberg et al. 2018; Greenhill et al. 2024; Blumen 2025; Norm Ai 2025). This nascent agenda has the potential to decrease compliance costs and increase policy effectiveness by helping policymakers and regulated firms rapidly and inexpensively interpret regulation. Deep learning can complement domain experts like engineers, lawyers, and scientists, who have long interpreted existing regulations (Meyer and Robertson 2019).

Deep learning models rarely, if ever, evaluate the effects of potential future policies. Such models rely on ex post data for training and assessing model performance. Instead, domain experts typically project the effects of potential policies, and many policymakers use such projections when designing regulations (Cubasch et al. 2001; Owen 2024).

We address these challenges to provide among the first ex ante deep learning projections of new regulations' effects. To address the absence of training data ("labels" in machine learning lexicon) we "relabel" historical regulatory decisions to reflect decisions under a new regulation. This relabeling methodology potentially has applicability in settings where analysts can use data from prior policies to characterize potential new policies. We evaluate this methodology by exploiting a setting permitting ex ante and ex post analysis. To understand the potential value added of this relabeling methodology, we compare its performance against published projections from domain experts.

We apply the relabeling methodology to study recent and proposed reforms to the 1972 US Clean Water Act (CWA), the cornerstone of federal water pollution regulation. The CWA protects the "Waters of the United States" but does not enumerate which streams and wetlands this phrase covers. To determine whether the CWA protects a site, a developer asks the Army Corps of Engineers (ACE) to evaluate the site's geophysical conditions and issue a legally binding Approved Jurisdictional Determination (AJD), indicating whether the CWA regulates the site under the CWA rule in force. CWA jurisdiction has changed repeatedly under recent White House and Supreme Court decisions. We develop and train the Clean Water Act Analysis of Regulation (CLEAR) model using 190,000 AJDs. We use it to provide the first ex post analysis of what the Supreme Court's *Sackett* decision regulates. Compared to machine learning-based analysis of earlier regulation (Greenhill et al. 2024), the CLEAR model uses geo-foundation methods, synthetic training data, fine-tuning on individual rules, and analyzes *Sackett,* the current CWA rule.

To evaluate our relabeling projection methodology and demonstrate its applicability, we project regulation under *Sackett* and two potential future reforms: the February 2025 White House National Emergency Executive Order, and the plaintiffs' argument in *White v. EPA*, an ongoing case that a recent proposed rule echoes (EPA and ACE 2025).

Examples illustrate the challenges of the "regulatory ping pong" (Cramer 2022; Phillis 2025) that frequent regulatory reforms create, the value of accurate projections, and the different model performance metrics relevant across stakeholders. Federal elected officials, judges, and regulatory agency staff value evidence on how potential new rules affect jurisdiction; thus, these groups may especially care about model performance on broad sets of sites (e.g., the share of wetland or agricultural areas jurisdictional nationally). State regulatory agencies plan around federal regulation, so they may value model performance on the share of sites jurisdictional in each state. Some state laws rely on federal agencies like the ACE to interpret the jurisdiction of state laws; rules restricting federal regulation may lead to state investment in capacity to interpret the jurisdiction of state laws (e.g., provide state AJDs), particularly for states wishing to maintain broad regulation during federal deregulation (Keiser et al. 2021). Developers may delay investment in land due to ambiguity about its future regulation, so they care about model performance on whether individual sites are jurisdictional. Environmental restoration firms, which earn several billion dollars annually by making investments to comply with federal regulations requiring wetland offsets within watersheds (Martin and Madsen 2023), value projections of where regulatory changes affect their investments, so may care most about performance on percent of sites jurisdictional in each watershed.

CWA Background

CWA interpretation has changed frequently (Gardner 2024). The CWA initially regulated traditional navigable waters due to the Constitution's interstate commerce clause. In the late 1970s, the Environmental Protection Agency (EPA) further regulated tributaries. In 1986 and 1988, ACE clarified the jurisdiction of nearby wetlands. In the Supreme Court's 2006 *Rapanos* case, Justice Kennedy's concurring opinion found that jurisdictional waters required a significant nexus involving biological, physical, or chemical connections to traditional navigable waters.

Six different rules have applied in the last decade. *Rapanos* applied between 2006 and 2015. The Clean Water Rule (CWR) primarily sought to clarify jurisdiction of *Rapanos;* it was issued in 2015, litigated, and repealed in 2019. In 2020, the Navigable Waters Protection Rule (NWPR) followed Justice Scalia's plurality in *Rapanos* to restrict jurisdiction to relatively permanent waters with a continuous surface water connection to traditional navigable waters, excluding ephemeral streams and isolated wetlands. Biden's 2023 Rule is essentially *Rapanos*. The Supreme Court's May 2023 *Sackett* decision echoed Justice Scalia's *Rapanos* opinion, removed the significant nexus standard, required a continuous surface water connection, less clearly embraced the relatively permanent standard, and excluded certain wetlands separated from navigable waters by barriers. Due to litigation, in September 2023, the EPA implemented two versions of *Sackett* across states. Our primary analysis pools the *Sackett* rules given their similarity, though we investigate differences (SM).

Model Description

We compare four types of models. First, domain experts use geophysical data to characterize CWA jurisdiction (Meyer and Robertson 2019; Gold 2024). These models use maps of streams and wetlands and assume that features with predefined characteristics describe jurisdiction under new rules. We analyze a geophysical rule from Gold (2024) that assumes wetlands from a vector dataset inundated a certain share of the year lose jurisdiction. Gold (2024) analyzes several inundation frequency scenarios. We report results from all scenarios, which have similar accuracy (Table S3).

To give geophysical rules the best opportunity in comparison to other methods, we focus on the best-performing scenario from Gold (2024). We also analyze a geophysical rule assuming that wetlands from a vector dataset within 50 meters of a connected stream in a different vector dataset are jurisdictional. We also consider a series of geophysical models that use individual layers to predict jurisdiction.

Our second model type, Relabeled *Sackett*, implements an ex ante methodology using deep learning. AJD data include rule-specific legal classifications ("resource types") for each waterbody. Under NWPR, for example, ACE classifies some sites as "adjacent wetlands" and others as "non-adjacent wetlands" (Table S4). The ex ante model uses all AJDs preceding *Sackett*, and projects *Sackett* as NWPR but with jurisdictional labels changed for several resource types. Table S4 presents our relabeling scheme; we formalized a version of relabeling in June 2023, before ACE implemented *Sackett* (Greenhill and Shapiro 2023), though we believe the scheme we implement here follows directly from *Sackett*. We train a deep learning model to predict the relabeled AJDs. Our ex ante models only use data and knowledge available before *Sackett* implementation. For example, Relabeled *Sackett* uses the *Sackett* text but no *Sackett* AJDs.

Third, we train the CLEAR model to predict jurisdiction under each rule, using all AJDs on all rules. CLEAR is the most accurate model we estimate. Of course, an ex post model was infeasible to train before *Sackett* implementation.

Fourth, we use relabeling to project jurisdiction under two possible future rules. The National Emergency Executive Order of January 2025 eliminates CWA jurisdiction for fossil fuel and hydroelectric energy sites. To project this rule, we relabel prediction points using spatial data on energy infrastructure from the Energy Information Agency (EIA 2025) and the locations of high-potential fossil fuel resources (Bartik et al. 2019). Our *White v. EPA* projection relabels all *Sackett* AJDs classified by ACE as having a "wetland" resource type as non-jurisdictional (Table S5). This relabeling strategy reflects the prosecution's argument that if a water body can be delineated as a wetland, it is not jurisdictional. The argument follows a prominent *Rapanos* line that *Sackett* embraces: "(waters should) include only those wetlands that are as a practical matter indistinguishable from waters of the United States" (*Rapanos* 547 U.S. plurality opinion, 2006).

All ex ante models, including geophysical models and relabeled models, are policy projections. As in climate science research, a "projection" reflects an assumed policy scenario. In contrast, ex post models reflect predictions of what is regulated when a policy is implemented.

Model Architecture

The Relabeled *Sackett*, CLEAR, and *White* deep learning models have the following architecture. We divide the AJDs into disjoint training, validation, and test data sets, avoiding footprint overlap to prevent leakage across folds. We train the algorithm using 80% of the data, then tune hyperparameters and choose model design using the validation set. All performance statistics here represent a held-out test set.

We use 34 geophysical input layers analyzed in prior work (Greenhill et al. 2024), plus land cover data from the Coastal-Change Analysis Program (CCAP) (NOAA, 2025) and 64-dimensional embedding fields from a Google geospatial foundation model. The embeddings encode information from a wide variety of data sources and types. The 34 layers include aerial imagery,

spatial data on wetlands, streams, elevation, land cover, soil types, and climate, as well as data on political and physical geography, including US states, ACE districts, and ecoregions. We include data from CCAP in addition to these 34 layers because of its higher resolution and accuracy relative to the National Land Cover Dataset (NLCD) (Zou et al. 2022). Because CCAP covers only coastal areas, we retain NLCD, which covers the entire contiguous US. Appropriate input layers are one-hot encoded. These inputs provide a detailed snapshot of ground conditions affecting the likelihood a CWA rule regulates a site.

We add synthetic labels to AJDs. AJDs focus on ambiguous cases and contain few points corresponding to unambiguously regulated or unregulated locations. We randomly select points in open waters, which all WOTUS rules regulate, and points from isolated wetlands, which *Sackett* does not regulate. This hand-labeling procedure is a case of human-in-the-loop machine learning, a growing area where algorithms and human feedback together improve performance (Mosqueira-Rey). Including synthetic data improves performance both for traditional navigable waters and more ambiguous cases in the AJD test set.

We report model performance on groups of sites. We randomly select 4 million points around the US and report the share that are jurisdictional. We separately report performance for streams, wetlands, agricultural sites, and floodplains.

Model Performance

Table 1 reports performance statistics for our four types of models. The appropriate performance metric depends on the user's objective. Given relevance to stakeholders, we emphasize mean absolute error at state and federal levels. We also emphasize accuracy given its common and simple interpretation. Depending on the objective, avoiding false positives versus false negatives may have different importance; we thus also report precision, recall, and area under the receiver operating characteristic curve (AUC).

ACE completes AJD forms and then manually enters summaries into a database. We manually compare a sample of AJD forms against corresponding entries in this database and found 3% mislabeled due to typing errors or confounding of multiple JDs for the same site. Others have modest error in reported coordinates (SM). Maximum possible accuracy is thus 97%, and true performance may exceed Table 1 values given error in ground-truth data. Human error in decision-making adds further noise.

A naïve benchmark assuming no sites are jurisdictional achieves 75.4% accuracy for *Sackett*, since a fourth of test set AJDs are jurisdictional (Table 1, Panel A). Because no sites in the benchmark are jurisdictional, the has model undefined precision, a recall of zero, and an AUC of one-half. Adopting the perspective of a state regulator or a wetland remediation firm, the benchmark has mean absolute error for of 25% for the average state, 22% for the average district, and 17% for the average watershed.

The geophysical model of wetness from Gold (2024) (Table 1, Panel B) marginally improves on this benchmark. Using the most accurate wetness criterion from Gold (2024) gives accuracy of 75.8%, implying a 0.4 percentage point improvement relative to the benchmark. Other Gold scenarios have worse accuracy, including some with accuracy below the benchmark (Table S3). Unlike the benchmark, this geophysical model predicts some sites as jurisdictional and thus has

positive precision and recall, though recall of 12.8% implies only correctly identifying one in seven true positives.

We also consider simpler geophysical models, which take one input layer at a time to project *Sackett* (Table S1). Individual layers have classification errors; this table provides a numerical test of the claim that they are unsuitable for projection (EPA and ACE 2020). Row 4 of Table S1, for example, assumes that *Sackett* only regulates sites in National Wetlands Inventory (NWI) wetlands. This projection performs poorly, with an accuracy 7% below the benchmark, partly because NWI has inclusion and exclusion errors and lacks information on relatively permanent surface water connections. Using other geophysical layers one-at-a-time gives similarly poor performance. These statistics echo quantitative criticism of inaccuracies in these layers (EPA and ACE 2020).

The ex ante deep learning model, Relabeled *Sackett* (Table 1, Panel C), substantially outperforms both geophysical models. Relabeled *Sackett* has an accuracy of 77.5%, implying four times more learning than the best geophysical model, where learning equals [accuracy - max(share regulated, share unregulated)]. Relabeled *Sackett* has twice the recall and 3.5% lower state mean absolute error than the best geophysical model. Its AUC also exceeds the best geophysical model by 0.16, an enormous margin.

CLEAR outperforms all other models on all metrics by wide margins (Table 1, Panel D). It accurately predicts jurisdiction for 83.3% of sites, implying four times the learning of the ex ante deep learning model and 20 times the learning of geophysical models. The accuracy of its positive predictions (precision) exceeds other models by about 20%. Comparing accuracy and precision indicates that its negative predictions have greater accuracy than its positive predictions, perhaps because most *Sackett* AJDs are negative, providing more learning opportunity. Recall is 18.2 percentage points higher than the ex ante relabeled model and 33.3 percentage higher than the Gold (2024) geophysical model.

CLEAR also has strong performance for the earlier NWPR and *Rapanos* rules (Table S1, Panel A). *Rapanos* lets the model observe more true positives, and thus CLEAR identifies positives more effectively, producing higher recall. Rules with less jurisdiction, NWPR and *Sackett*, have fewer true positives, so the model is less able to learn to identify positives.

Overall, Relabeled *Sackett* outperforms geophysical models but does not match the CLEAR performance. While stakeholders can use deep learning projections to add substantial insight beyond geophysical projections, ex post learning still matters.

How quickly can policymakers learn what a new rule covers? Learning curves describe the months of policy implementation data needed to achieve given accuracy (Fig. 1). Ex ante models, by definition, do not improve with additional months of ex post data. We train several versions of the CLEAR model, each using a different number of months of data, then plot accuracy. Six months of data are needed for CLEAR accuracy to exceed the Relabeled *Sackett* projections.

How quickly do agencies learn to interpret a rule? Do early phases of implementation convey the same information as later phases? Consider three hypotheses. First, in the early phase of a new rule, an agency implements a rule idiosyncratically, but converges over time. Second, rule

implementation varies idiosyncratically but steadily, challenging extrapolation across time periods. Third, a rule's geographic coverage may expand with implementation.

To distinguish among hypotheses, Fig. 1B plots learning curves varying whether we add data chronologically, randomly ordered, or reverse chronologically. The lines end with equal accuracy since data equalize in the last period. Chronological data generate the slowest learning, random ordering generates the fastest learning, and inverse chronological order is intermediate.

These learning curves do not support the first hypothesis that ACE improves the reliability of rulings over time; otherwise, the inverse ordering would learn fastest. These curves are most consistent with the third hypothesis that the algorithm learns from seeing a wide range of sites and the chronological data can be clustered (e.g., one project may have dozens of similar simultaneous AJDs for nearby wetlands).

Describing Jurisdiction

The geophysical models generate binary predictions, whereas the deep learning models produce continuous measures of the probability that a site is regulated.

Geophysical models predict low jurisdiction for *Sackett* (Table 2, columns (2) and (3)). Gold (2024) finds that 3.6% of the US is jurisdictional, covering 23% of wetlands and 7.7% of streams. The connected geophysical model concludes that 0.0% of the US, 0.2% of streams, and 0.2% of wetlands are jurisdictional. These models predict low jurisdiction rates because NHD and NWI miss many jurisdictional *Sackett* locations. Additionally, the NWI wetness and connectedness criteria exclude more jurisdictional *Sackett* AJDs (Fig. S1).

Relabeled *Sackett* in column (4) of Table 2 projects broader jurisdiction. Again, 3.5% of the US is jurisdictional. Relative to Gold, Relabeled *Sackett* predicts more regulation of streams and less of wetlands.

The CLEAR model in column (5) of Table 2, our most accurate estimate, calculates that *Sackett* regulates 7.0% of US area. This includes 35.8% of stream miles and 30.5% of wetland acres. *Sackett* is twice as likely to regulate perennial streams to regulate intermittent and ephemeral streams, using NHD's classification of these water types. About 6% of areas not in NHD and 3% of areas not in NWI are jurisdictional, underscoring limitations of these data.

Sackett regulates 2.7% of cropland as measured by the National Land Cover Dataset (NLCD). It is initially surprising that this share exceeds zero because the CWA excludes prior converted cropland. Manually investigating a sample of these AJDs indicates that most cases involve sites adjacent to fields but on roads, homes, undeveloped wetlands bordering cropland, or other NLCD misclassification. These cases help explain why agriculture has opposed stricter CWA rules.

Table S2 shows regulation under two prior rules, *Rapanos* and NWPR. *Rapanos* regulated half of the stream miles and wetland acres and 15% of US area. NWPR deregulated more than a fourth of regulated streams and 40% of regulated wetlands, ultimately deregulating half of areas that *Rapanos* regulated. *Sackett* regulates marginally more than NWPR. Compared to *Rapanos*, *Sackett* deregulates around a third of regulated streams and wetlands, two-thirds of regulated NLCD cropland area, and 40% of regulated floodplains. Deregulation of cropland may respond to

agricultural criticism of CWA regulation. Deregulation of floodplains may increase development and population exposure to floods, which sea level rise may make more common.

Column (6) of Table 2 considers the January 2025 National Energy Emergency Executive Order. We calculate that it marginally decreases jurisdiction relative to *Sackett* by a fifth of a percentage point nationally and for around 1% of streams and wetlands.

In Column (7) of Table 2, *White* dramatically restricts *Sackett* coverage. Here, 15.5% of streams, 11.6% of wetlands, and 11.4% of floodplains lose protection. Although *White* relabeling exclusively affected wetlands, the projection also predicts regulation declines for streams.

Fig. 2 visualizes the "regulatory ping pong" of recent CWA regulation. The share of US land that the CWA regulates fell by close to half in 2020, returned in 2021, fell again in 2023, and we project will fall by half again under *White.* Stream miles and wetlands acres see similarly drastic regulatory changes.

Maps visualize *Sackett* jurisdiction and contrast ex ante and ex post predictions (Fig. 3A). CLEAR predicts that *Sackett* regulates traditional navigable waters like the Mississippi and Missouri Rivers, adjacent wetlands, and wetland-rich areas of Louisiana, Mississippi, South Florida, the Northeast, and the Pacific Northwest. *Sackett* regulates few other areas.

Figs. 3B through 3E map the differences between ex ante projections of *Sackett* and the proposed rules we analyze. Fig. 3B compares Gold (2024)'s ex ante geophysical model to ex post deep learning. Fig 3C compares ex ante deep learning to ex post deep learning.

The National Energy Emergency Executive Order particularly deregulates the Marcellus and Bakken shales, including portions of the upper Missouri River in North Dakota, as well as in shale and oil fields in Texas. *White* removes regulation from wetlands across the Atlantic seaboard, currently protected under *Sackett.*

Discussion

Machine learning complements expert decision-making in a growing range of settings. In regulatory analysis, it helps interpret existing regulations. Domain experts project what proposed future regulations cover; but so far, deep learning has not. Is deep learning useful only for backward-looking regulatory settings, or can it help stakeholders understand prospective regulatory reform?

Our straightforward relabeling methodology lets deep learning project regulations, and our setting allows assessment of the method's value-added over geophysical domain expert projections. Ex ante deep learning strongly outperforms geophysical projections. At the same time, ex post deep learning still outperforms ex ante projections. About six months of ex post training data are needed to realize a large share of the benefits of ex post deep learning and to outperform ex ante projections.

The CLEAR model can support many stakeholders. State and federal policymakers, landowners, developers, environmental organizations, and environmental restoration companies can all use both ex ante and ex post projections of CWA jurisdiction. Our analysis template may help guide future CWA reforms.

More broadly, environmental regulations and other domains are undergoing dramatic changes due to US deregulation and global focus on climate change. Deep learning may have some role to contribute in interpreting these rapid changes.

References

Bartik, Alexander W., Janet Currie, Michael Greenstone, and Christopher R. Knittel. 2019. "The Local Economic and Welfare Consequences of Hydraulic Fracturing." *American Economic Journal: Applied Economics* 11 (4): 105–55. https://doi.org/10.1257/app.20170487.

Blumen. 2025. "Blumen - Automated Energy Project Diligence and Reports." 2025. https://www.blumensystems.com/.

Coglianese, Cary, and David Lehr. 2016. "Regulating by Robot: Administrative Decision Making in the Machine-Learning Era." *Georgetown Law Journal* 105:1147.

Cramer, Senator Kevin. 2022. "RESTORING STATES' RIGHTS & ADHERING TO COOPERATIVE FEDERALISM IN ENVIRONMENTAL POLICY." *Harvard Journal of Law & Public Policy* 45 (2).

Crews, Clyde Wayne. 2023. "The 90,402-Page 2023 Federal Register Contains 3,018 Final Rules and Regulations." *Forbes*, December 29, 2023. https://www.forbes.com/sites/waynecrews/2023/12/29/bidens-2023-federal-register-page-count-is-the-second-highest-ever/.

Cubasch, U., G. Meehl, G. J. Boer, Ron Stouffer, M. Dix, A. Noda, C. A. Senior, et al. 2001. "Projections of Future Climate Change." PNNL-SA-39649. J.T. Houghton, et al; Cambridge University Press, New York, NY, United States(US). https://www.osti.gov/biblio/901481.

EIA. 2025. "Maps - U.S. Energy Information Administration (EIA)." https://www.eia.gov/maps/index.php.

Engstrom, David Freeman, Daniel E. Ho, Catherine M. Sharkey, and Mariano-Florentino Cuéllar. 2020. "Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3551505.

EPA and ACE. 2020. "Economic Analysis for the Navigable Waters Protection Rule: Definition of "Waters of the United States'."

———. 2025. "Memorandum to the Field Between the U.S. Department of the Army, U.S. Army Corps of Engineers and the U.S. Environmental Protection Agency Concerning the Proper Implementation of "Continuous Surface Connection" Under the Definition of 'Waters of the United States' Under the Clean Water Act."

Gardner, Royal C. 2024. *Waters of the United States: POTUS, SCOTUS, WOTUS, and the Politics of a National Resource*. 1st ed. Washington, D. C: Island Press.

Gold, Adam C. 2024. "How Wet Must a Wetland Be to Have Federal Protections in Post-Sackett US?" *Science* 385 (6716): 1450–53. https://doi.org/10.1126/science.adp3222.

Greenhill, Simon, Hannah Druckenmiller, Sherrie Wang, David A. Keiser, Manuela Girotto, Jason K. Moore, Nobuhiro Yamaguchi, Alberto Todeschini, and Joseph S. Shapiro. 2024. "Machine Learning Predicts Which Rivers, Streams, and Wetlands the Clean Water Act Regulates." *Science* 383 (6681): 406–12. https://doi.org/10.1126/science.adi3794.

Greenhill, Simon, and Joseph S. Shapiro. 2023. "What Does the Clean Water Act Regulate?" Presented at the Environmental Protection Agency.

Keiser, David A., Sheila M. Olmstead, Kevin J. Boyle, Victor B. Flatt, Bonnie L. Keeler, Catherine L. Kling, Daniel J. Phaneuf, Joseph S. Shapiro, and Jay P. Shimshack. 2021. "A Water Rule That Turns a Blind Eye to Transboundary Pollution." *Science* 372 (6539): 241–43. https://doi.org/10.1126/science.abf8885.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. “Human Decisions and Machine Predictions*.” *The Quarterly Journal of Economics* 133 (1): 237–93. https://doi.org/10.1093/qje/qjx032.

Martin, Steve, and Becca Madsen. 2023. “The Time It Takes for Restoration: An Analysis of Mitigation Bank Instrument Timelines.” Washington, D.C.: Environmental Policy Innovation Center and Ecological Restoration Business Association.

Meyer, Roger, and Andrew Robertson. 2019. “Clean Water Rule Spatial Analysis.” St. Mary’s University of Minnesota.

Mulligan, Stephen P. 2018. “Wading Into the ‘Waters of the United States.’” Congressional Research Service.

Norm Ai. 2025. “Norm Ai.” 2025. https://www.norm.ai/.

Owen, Dave. 2024. “Mapping the New Clean Water Act.” *Science* 385 (6716): 1414–16. https://doi.org/10.1126/science.adr6670.

Phillis, Michael. 2025. “EPA Likely to Move to Further Limit Federal Protections for Wetlands.” *AP News*, March 12, 2025, sec. Climate. https://apnews.com/article/epa-trump-clean-water-act-wetlands-protection-021ff0aacd77b91c4b0e70bc5caedd06.

Regulatory Studies Center. 2024. “Total Pages Published in the Code of Federal Regulations (1950-2021).” The George Washington University.

Zou, Zhenhua, Chengquan Huang, Megan Lang, and Rusty Griffin. 2022. “Derivation of Wetland Difference Products by Comparing the NWI Geospatial Dataset with C-CAP (10-m) and NLCD (2019) Data.”

Tables

Table 1: Geophysical models modestly improve on naïve benchmark, ex ante *Sackett* model does better, ex post CLEAR model has strongest performance. MAE: mean absolute error in predicted share jurisdictional in US or state. For example, column (1) equals $|\mathrm{mean}(J_i) - \mathrm{mean}(C_i)|$, where J_i represents jurisdiction of validation set AJDs and C_i represents jurisdiction of CLEAR predictions for validation set AJDs. Similarly, column (2) equals $\frac{1}{S}\ \sum_{\mathrm{s}} |\mathrm{mean_s(J_{is})} - \mathrm{mean_s(C_{is})}|$, i.e., the mean across states of the mean jurisdiction rate of validation AJDs within each state minus the mean jurisdiction rate of CLEAR model predictions. Accuracy: percent correct. Precision: TP / (TP + FP), where TP is the count of true positive predictions and FP is the count of false positive predictions. Note that recall is not defined if a model makes no positive predictions. Recall: (TP / (TP + FN), where FN is the count of false negative predictions. AUC: Area under the receiver operating curve. Row 1 describes forecasting based on a naïve benchmark that predicts no location is jurisdictional. Row 2 describes forecasting based on the "seasonally flooded" model from Gold (2024). Row 3 defines points as jurisdictional if they fall within an NWI polygon that has a surface water connection with a NHD flowline that terminates in a navigable water. Row 4 describes the forecast of Sackett ex-ante, by relabeling NWPR resource types. All models describe *Sackett*. We report performance statistics only for models where there exists ex post AJDs to evaluate the model's performance, which excludes the Energy Emergency Executive Order and *White*.

	Mean absolute error								
	US (1)	**State** (2)	**District** (3)	**Water-shed** (4)	**Accuracy** (5)	**Precision** (6)	**Recall** (7)	**AUC** (8)	**N** (9)
A Naïve benchmark									
1. No jurisdiction	0.246	0.254	0.222	0.171	0.754	—	0.000	0.500	892
B Ex ante geophysical models									
2. Gold (2024)	0.186	0.215	0.185	0.178	0.758	0.528	0.128	0.545	892
3. Connected	0.246	0.254	0.222	0.171	0.754	—	0.000	0.500	892
C Ex ante deep learning									
4. Relabeling	0.129	0.181	0.135	0.175	0.775	0.587	0.279	0.704	892
D Ex post deep learning									
5. CLEAR	0.098	0.125	0.125	0.131	0.833	0.765	0.461	0.803	892

Table 2: CLEAR and ex ante models project Clean Water Act regulation. Table shows the share of points each model estimates are regulated. Column (1) describes a naïve model that predicts all points as non-jurisdictional. Column (2) describes forecasting based on the "seasonally flooded" model from Gold (2024). Column (3) defines points as jurisdictional if they fall within an NWI polygon that has a surface water connection with a NHD flowline that terminates in a navigable water. Column (4) describes the ex ante forecast of *Sackett* obtained by relabeling NWPR resource types. Column (5) defines a point as jurisdictional if it has CLEAR model score above 0.50. Column (6) describes the forecast of the 2025 Energy Emergency Executive Order, which deregulates all CLEAR-*Sackett* predictions near energy infrastructure, projects, and high-potential fossil fuel resources. Column (7) describes the forecast of regulation under the *White* prosecution argument, by relabeling *Sackett* resource types. Panels B through D describe subsets of the four million prediction points. In NHD, perennial is fcode 46006, intermittent or ephemeral are fcodes 46003 or 46007.

		Ex ante geophysical		Ex ante deep learning	Ex post	Other projections	
	Naïve benchmark	Gold (2024)	Connected	Relabeled *Sackett*	CLEAR-*Sackett*	Energy emergency	*White v. EPA*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A General groups of points							
All 4 million points	0.000	0.036	0.000	0.035	0.070	0.068	0.037
B Rivers and streams							
All (NHD all)	0.000	0.077	0.002	0.261	0.358	0.343	0.203
Perennial	0.000	0.139	0.002	0.429	0.540	0.520	0.320
Intermittent or ephemeral	0.000	0.039	0.001	0.130	0.209	0.199	0.076
None (not in NHD)	0.000	0.036	0.000	0.030	0.064	0.062	0.034
C Wetlands							
All (NWI palustrine)	0.000	0.231	0.002	0.165	0.305	0.297	0.189
Emergent (NWI)	0.000	0.401	0.001	0.122	0.199	0.194	0.084
Forested (NWI)	0.000	0.534	0.002	0.203	0.308	0.304	0.100
None (not in NWI palustrine)	0.000	0.002	0.000	0.012	0.028	0.028	0.011

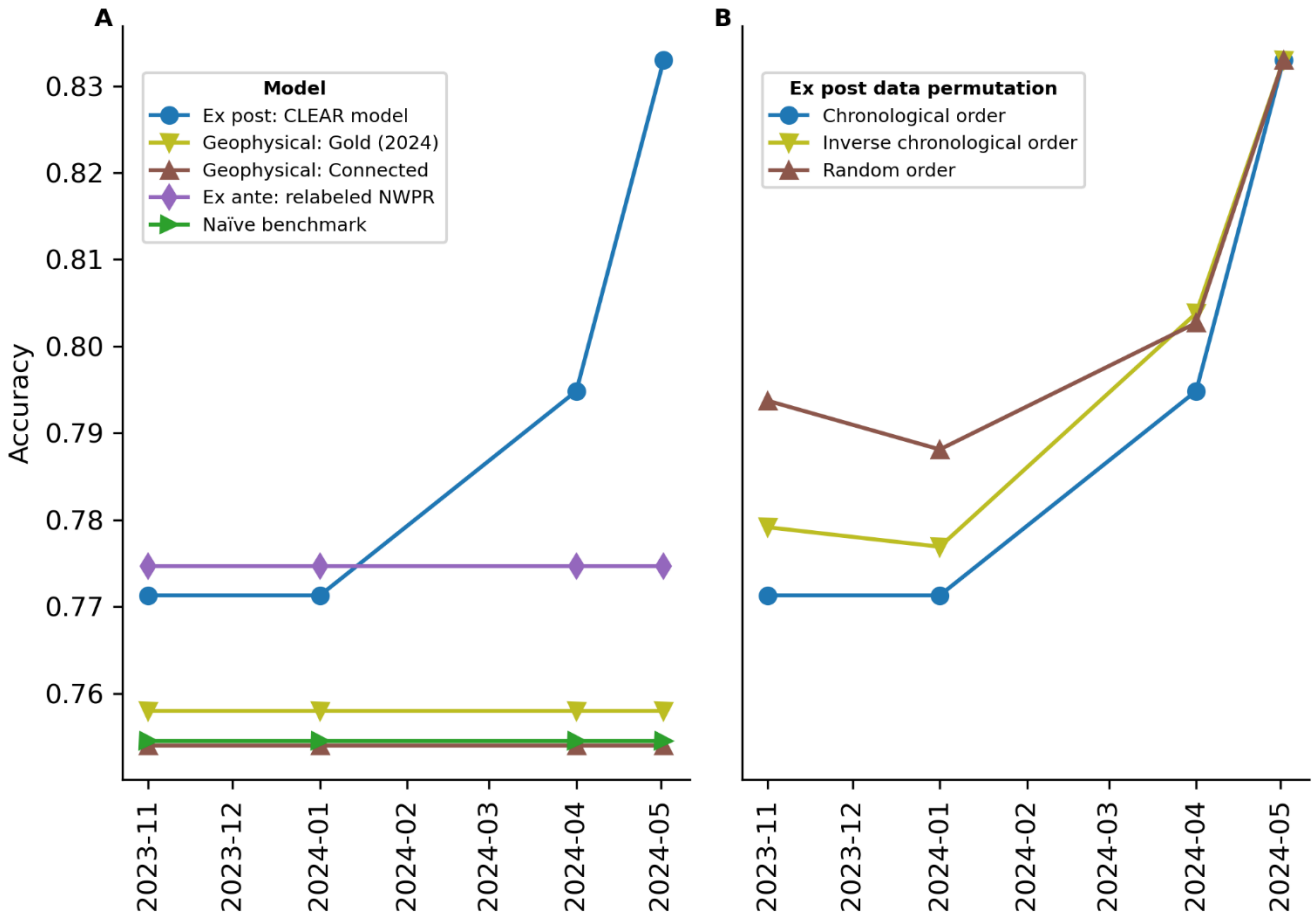
continued on next page

Table 2: CLEAR and ex ante models project Clean Water Act regulation. Table shows the share of points each model estimates are regulated. Column (1) describes a naïve model that predicts all points as non-jurisdictional. Column (2) describes forecasting based on the "seasonally flooded" model from Gold (2024). Column (3) defines points as jurisdictional if they fall within an NWI polygon that has a surface water connection with a NHD flowline that terminates in a navigable water. Column (4) describes the ex ante forecast of *Sackett* obtained by relabeling NWPR resource types. Column (5) defines a point as jurisdictional if it has CLEAR model score above 0.50. Column (6) describes the forecast of the 2025 Energy Emergency Executive Order, which deregulates all CLEAR-*Sackett* predictions near energy infrastructure, projects, and high-potential fossil fuel resources. Column (7) describes the forecast of regulation under the *White* prosecution argument, by relabeling *Sackett* resource types. Panels B through D describe subsets of the four million prediction points. In NHD, perennial is fcode 46006, intermittent or ephemeral are fcodes 46003 or 46007.

		Ex ante geophysical		**Ex ante deep learning**	**Ex post**	**Other projections**	
	Naïve benchmark	**Gold (2024)**	**Connected**	**Relabeled *Sackett***	**CLEAR-*Sackett***	**Energy emergency**	***White v. EPA***
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
D Rivers, streams, and wetlands							
All (NWI all, NHD all)	0.000	0.226	0.002	0.164	0.303	0.295	0.186
None (not in NWI or NHD)	0.000	0.002	0.000	0.011	0.028	0.027	0.010
E Other important groups of points							
Cropland and pasture (NLCD)	0.000	0.009	0.000	0.017	0.027	0.027	0.010
Floodplains (NFIP)	0.000	0.198	0.002	0.234	0.332	0.324	0.218
Urban growth areas (ICLUS)	0.000	0.023	0.000	0.036	0.062	0.061	0.026
Urban developed (NLCD)	0.000	0.008	0.000	0.022	0.056	0.055	0.025

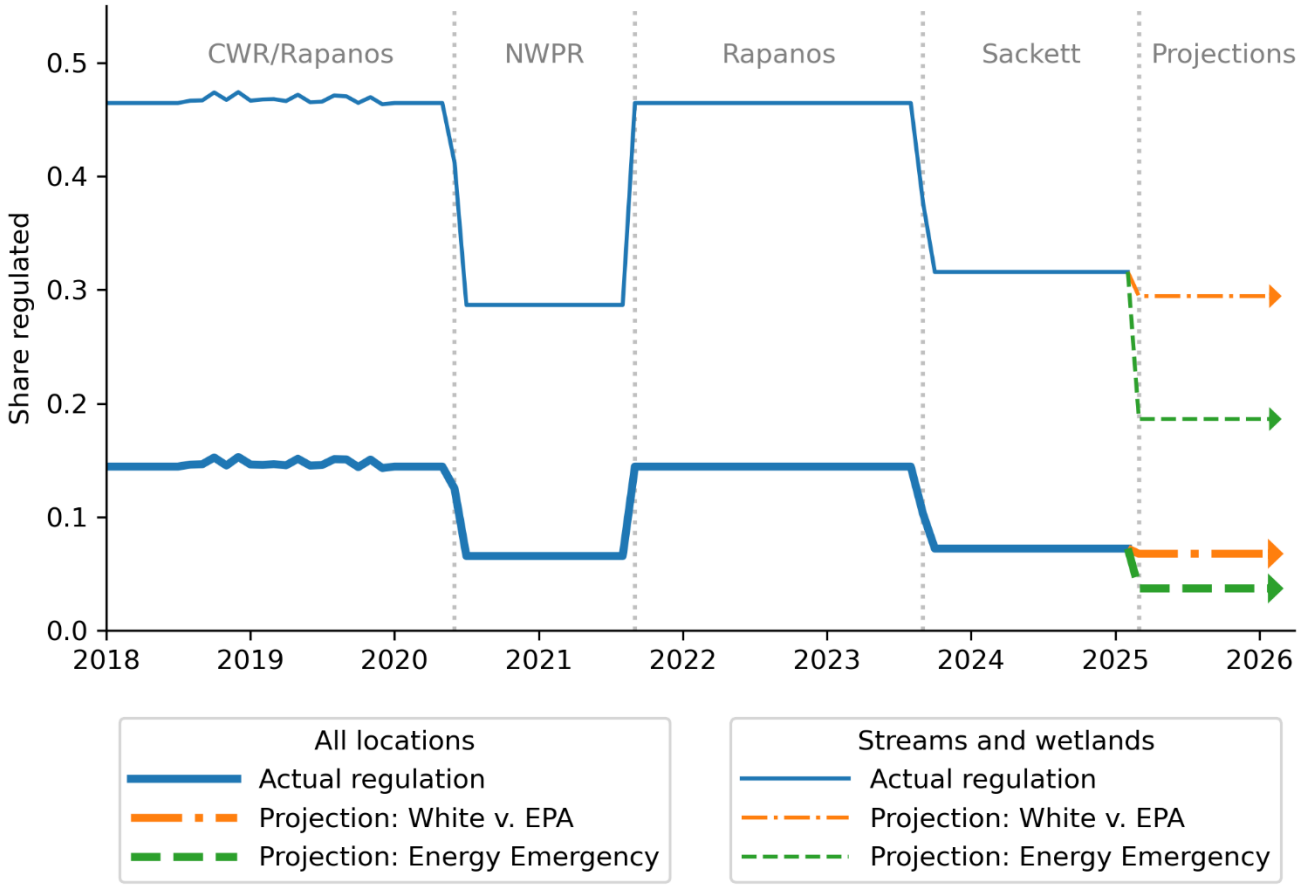
Figures

Figure 1: CLEAR learns to outperform ex ante projections with several months of data, and learns fastest with randomly-ordered data



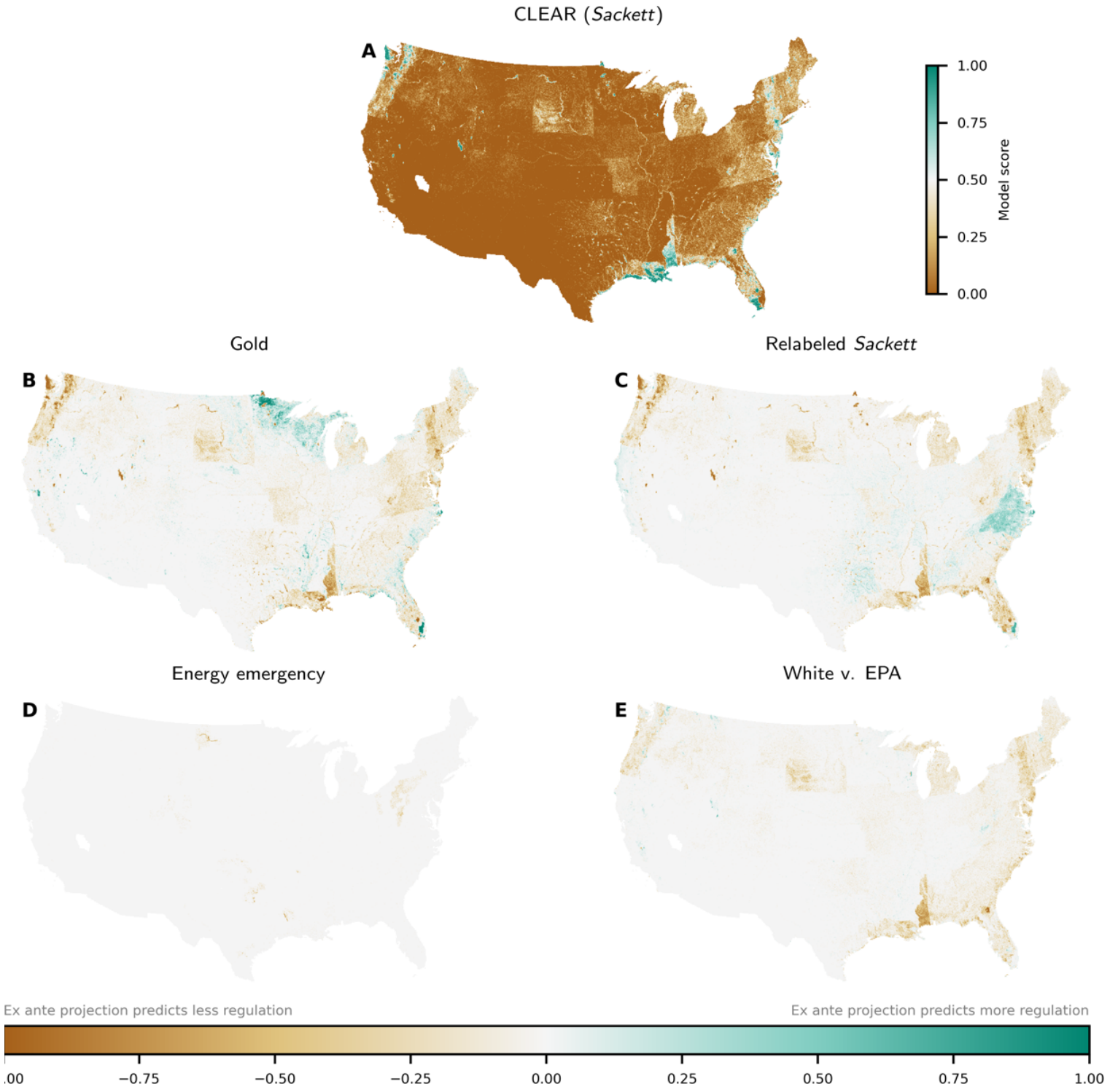
Notes: Panel A describes performance of the CLEAR deep learning model and ex ante projections on the *Sackett* AJDs. Points show the first nine months of post-*Sackett* AJD data, and the full sample. The blue line with circular markers plots the accuracy of the CLEAR model at each time step, which increases monotonically as the model is trained on more ex post data. Ex ante projections depend only on pre-implementation data so do not improve as ex post data become available. The purple line with diamonds describes relabeled *Sackett*; the olive line with downward-pointing triangles corresponds to the Gold (2024) geophysical model; the brown line with upward-pointing triangles describes the connected geophysical model; the green line with right-pointing triangles describes the naïve benchmark predicting no jurisdiction anywhere. Panel B describes CLEAR performance as in Panel A, for permutations of AJDs. The blue line with circular markers is identical to the same line in Panel A. The olive line with downward-pointing triangle markers reports performance of a CLEAR model trained on data in inverse chronological order. From left to right, the accuracies reported correspond to models trained on the last 3 months, the last 6 months, the last 9 months, and all months of ex post data. The brown line with upward-pointing triangle markers reports performance of a CLEAR model trained on subsets of the AJD data in randomly shuffled order.

Figure 2: Large variation in CWA jurisdiction across rules creates "regulatory ping pong"



Notes: Thick lines display the share of all 4 million prediction points predicted as jurisdictional in each month since January 2018. Thin lines display the share of jurisdictional points within NWI or NHD features. To determine which rule applied in each month, we use the rule used to decide a majority of AJDs within each state in each month, calculate statistics at the state level, and average across states, weighting by the number of points in the state. Between January 2018 and August 2019, some states implemented CWR and others implemented *Rapanos* due to contemporaneous lawsuits challenging CWR. The slight fluctuations in the share of locations regulated during this period reflect state-level changes in rules applied due to stays on CWR's implementation, which resulted from litigation (Mulligan 2018). *Rapanos* applied from September 2019 to May 2020. NWPR applied from June 2020 to August 2021. *Rapanos* (including the 2023 rule) applied again from September 2021 to August 2023. *Sackett* applied from September 2023 to the present (March 2025). The EPA is implementing two versions of *Sackett* in different states due to pending litigation. Post March 2025 plots projections. Dash-dotted orange lines projects the 2025 Energy Emergency Executive Order. Dashed green lines project *White.*

Figure 3: Maps of CWA regulation under *Sackett* and spatial differences between ex post prediction and ex ante projection



Notes: Panel A shows the predicted probability of CWA regulation under *Sackett.* Panels B through E show the difference between ex ante projections and the ex-post CLEAR model. Panel B shows the difference between the Gold (2024) geophysical projection and ex post CLEAR. Panel C shows the difference for the ex ante relabeled CLEAR model. Panel D shows the difference for the Energy Emergency Executive Order

projection. Panel E shows the difference for the *White v. EPA* projection. Map are generated from 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).

Acknowledgements

Participants in EPA Region 3 and Environmental Law Institute workshops, J. Blumenstock, S. Burley, N. Cabrera, T. Carleton, C. Costello, J. Devine, H. Druckenmiller, P. Hann, S. Hsiang, M. Kotchen, R. Kwok, D. Lobell, J. Murphy, D. Owen, B. Payne, and S. Wang provided comments. We thank A. Barker, H. Menth, and S. Shoemaker for research assistance.

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

Competing interest declaration: S. Greenhill and J. S. Shapiro 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.

Supplementary Tables and Figures

Table S1: For all rules, CLEAR performs well but individual input layers perform poorly. Accuracy: percent correct; Precision: TP / (TP + FP); Recall: (TP / (TP + FN; AUC: Area under the receiver operating curve; Jurisdictional share: rate of regulation in test set; N: number of AJDs in test set. Panel A uses the CLEAR model. Rows 4-10 describe predictions using only layers intersected with AJDs under various rules. "Wetland" predicts any AJD within a NWI polygon as jurisdictional, and non-jurisdictional otherwise. "Streams" (NHD flowline), "Hydric soil" (gNATSGO), and "Water table depth <10m" (gNATSGO) follow a similar rule. "Cropland" and "Urban Developed" predict any AJD within their respective NLCD categorized areas as non-jurisdictional, and jurisdictional otherwise.

	Mean absolute error							
	US	**State**	**Watershed**	**Accuracy**	**Precision**	**Recall**	**AUC**	**N**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A Ex post CLEAR model for other rules								
1. All rules	0.063	0.122	0.171	0.795	0.736	0.595	0.85	16,458
2. Rapanos	0.059	0.165	0.203	0.783	0.771	0.658	0.85	9186
3. NWPR	0.064	0.075	0.136	0.809	0.631	0.46	0.82	6,380
B Individual input layers—*Sackett*								
4. Wetland (NWI)	0.085	0.252	0.302	0.695	0.410	0.553	0.647	892
5. Stream (NHD)	0.168	0.214	0.174	0.760	0.536	0.169	0.561	892
6. Wetland or stream	0.096	0.256	0.311	0.693	0.410	0.571	0.652	892
7. Hydric soil (gNATSGO)	0.278	0.409	0.448	0.507	0.263	0.562	0.525	892
8. Water table <10m (gNATSGO)	0.136	0.251	0.230	0.680	0.163	0.073	0.476	892
9. Cropland and pasture (NLCD)	0.474	0.509	0.621	0.396	0.251	0.735	0.510	892
10. Urban developed (NLCD)	0.492	0.559	0.615	0.425	0.277	0.831	0.562	892

Table S2: CLEAR Model Projects Regulation Under Earlier CWA Rules. Table shows the share of points each framework estimates are regulated. Columns define a point as jurisdictional if it has CLEAR model score above 0.50. Panels **B** through **D** describe subsets of the four million prediction points. For navigable waters, we evaluate jurisdiction based on the mean model prediction among points within each named river, because the CWA regulates each water resource, and report the share of named rivers predicted as jurisdictional. Wetland permits include all types (individual, nationwide, regional, letters of permission, etc.)

	Rapanos	**NWPR**
	(1)	(2)
A General groups of points		
All 4 million points	0.146	0.064
B Rivers and streams		
All (NHD all)	0.504	0.354
Perennial	0.718	0.567
Intermittent or ephemeral	0.337	0.184
None (not in NHD)	0.138	0.058
C Wetlands		
All (NWI palustrine)	0.469	0.281
Emergent (NWI)	0.344	0.146
Forested (NWI)	0.543	0.266
None (not in NWI palustrine)	0.089	0.026
D Rivers, streams, and wetlands		
All (NWI all, NHD all)	0.466	0.279
None (not in NWI or NHD)	0.088	0.025
E Other important groups of points		
Cropland and pasture (NLCD)	0.078	0.027
Floodplains (NFIP)	0.520	0.337
Urban growth areas (ICLUS)	0.157	0.059
Urban developed (NLCD)	0.150	0.053

Table S3: Gold (2024) Scenarios have similar model performance. Accuracy: percent correct; Precision: TP / (TP + FP); Recall: (TP / (TP + FN); AUC: Area under the receiver operating curve; Jurisdictional share: rate of regulation in test set; N: number of AJDs in test set. Panel A shows scenarios analyzed in Gold (2024). 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 at least as wet as "seasonally flooded" are predicted as WOTUS; all others are predicted as non-WOTUS. See Gold (2024) for more details.

	Accuracy	Precision	Recall	AUC	Jurisdictional share	N
	(1)	(2)	(3)	(4)	(5)	(6)
A Gold (2024) Scenarios						
1. Temporarily flooded	0.752	0.500	0.193	0.565	0.248	1334
2. Seasonally saturated	0.759	0.571	0.121	0.545	0.248	1334
3. Continuously saturated	0.754	0.532	0.076	0.527	0.248	1334
4. Seasonally flooded	0.756	0.556	0.076	0.528	0.248	1334
5. Seasonally flooded/saturated	0.752	0.500	0.006	0.502	0.248	1334
6. Semi-permanently flooded	0.752	0.500	0.006	0.502	0.248	1334
7. Intermittently exposed	0.752	0.000	0.000	0.500	0.248	1334
8. Permanently flooded	0.752	0.000	0.000	0.500	0.248	1334

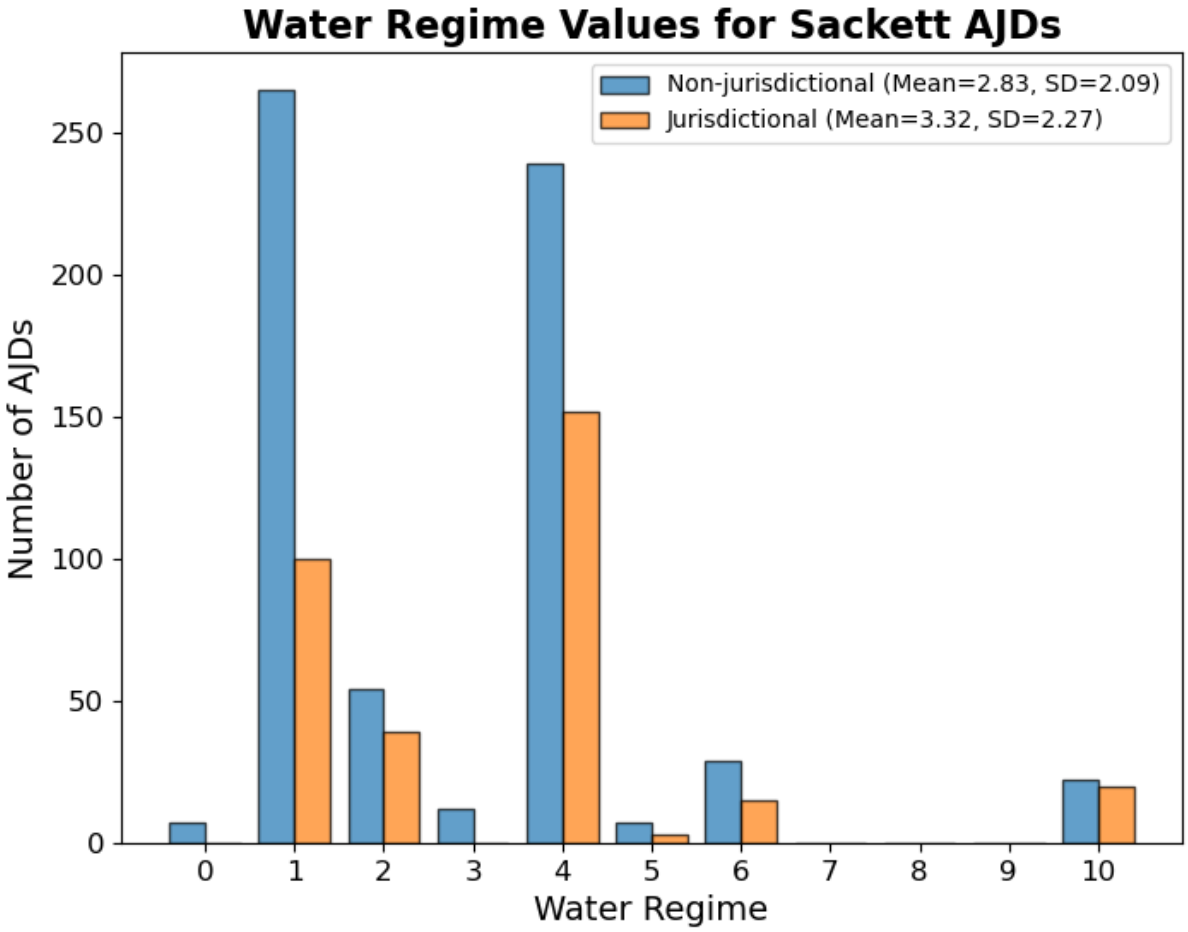
Table S4: Relabeling Scheme to train Relabeled *Sackett* Model. Each row shows details about a resource type under NWPR relabeled to match *Sackett* implementation.

	Definition	Share of AJDs	Jurisdictional under NWPR	Jurisdictional under Relabeled *Sackett*
	(1)	(2)	(3)	(4)
A NWPR Resource Types				
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
A4WETNATSEP	(a)(4) Wetland separated from an (a)(1)-(a)(3) water only by a natural feature	0.0027	Yes	No

Table S5: Relabeling Scheme to train *White v. EPA* Model. Each row shows details about a resource type under *Sackett* relabeled to match *White v. EPA.* "Pre-2015-Post-Sackett" and "Amended 2023" Rules were implemented in different states following the *Sackett* decision, though we treat them as the same rule.

	Definition	Share of AJDs	Jurisdictional under *Sackett*	Jurisdictional under *White v. EPA*
	(1)	(2)	(3)	(4)
A Pre-2015-Post-Sackett Resource Types				
A7-AJD.WETL-404	(a)(7) Wetland adjacent to a non-wetland water identified in (a)(1) - (a)(6)	0.095	Yes	No
B Amended 2023 Rule Resource Types				
A4-1.ADJ.WET.A1-INTSTATE-404	(a)(4)(i) Adjacent Wetland, adjacent to (a)(1)(iii) Interstate Water	0.0007	Yes	No
A4-1.ADJ.WET.A1-TERSEAS-404	(a)(4)(i) Adjacent Wetland, adjacent to (a)(1)(ii) Territorial Sea	0.0002	Yes	No
A4-1.ADJ.WET.A1-TNW-404	(a)(4)(i) Adjacent Wetland, adjacent to (a)(1)(i) TNW	0.014	Yes	No
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)	0.071	Yes	No

Appendix Figure 1



Notes: NWI "Water Regime" characterizations differ dramatically for both non-jurisdictional and jurisdictional Sackett AJDs. This figure plots the water regime value, which describes "wetness" in Gold (2024) scenarios, for all 964 Sackett AJDs that fall within a NWI polygon. Blue bars display AJDs determined Not-WOTUS; orange bars display Yes-WOTUS.