

# Survey of Groundwater Wells in the United States

Alexandra L. Miller

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University  
in partial fulfillment of the requirements for the degree of

Master of Science  
In  
Civil Engineering

Landon T. Marston, Chair  
Siddharth Saksena  
Julie E. Shortridge

May 8, 2023  
Blacksburg, Virginia

Keywords: Groundwater, Wells, National, Aquifers

Copyright 2023, Alexandra L. Miller

# Survey of Groundwater Wells in the United States

Alexandra L. Miller

## ABSTRACT

Groundwater wells are critical infrastructure with significant impacts on the environment, water availability, and economy. However, comprehensive data on the purposes, locations, depths, and construction of these wells are only collected by individual states. We have compiled a nationwide dataset of groundwater wells throughout the United States. The tabular dataset consists of all groundwater well data obtained from the states, containing over nine million records. A subset of this dataset was created that excludes wells located outside of the reported county or state, with over eight million records. Our dataset represents all known groundwater well locational data that states could release. The data made available by these datasets can serve as a critical tool for refining our understanding of how groundwater is accessed and used throughout the United States, and how it impacts different industries.

# Survey of Groundwater Wells in the United States

Alexandra L. Miller

## GENERAL AUDIENCE ABSTRACT

Groundwater wells are important structures that have a big impact on the environment, water availability, and economy. However, information about these wells, like why they are built, where they are located, how deep they are, and how they are constructed, is only collected by individual states. We have put together a dataset that includes information about groundwater wells from all across the United States. This dataset is in a table format and contains over nine million records. We created a smaller version of the dataset that only includes wells located within the reported county or state, which has over eight million records. Our dataset includes all the information that states have shared about the locations of groundwater wells. This data can be a valuable tool for improving our understanding of how groundwater is accessed and used throughout the United States, and how it affects different industries.

# Table of Contents

List of Figures	v
List of Tables	vi
List of Abbreviations	vii
1 Introduction	1
2 Methods	3
2.1 Data Collection.....	4
2.2 Data Standardization.....	7
2.3 Technical Validation.....	12
3 Results	17
4 Conclusion	25
References	26

# List of Figures

1	Number of groundwater wells as reported by each state.....	2
2	A depiction that outlines the process involved in developing the groundwater well datasets. Initially, data was collected from state agencies, and then standardized using our data standard, the United States Groundwater Wells Inventory Supplementary Tables (USGWIST). The data was then visually displayed using ArcGIS Pro. Lastly, data was reviewed by members of the research team and state agencies.....	2
3	Total number of well records collected for each state with locational details. The number of the total well records that do not report coordinates are shown in orange, while records with coordinates that do not match the state-provided county or state where the well is supposed to exist are shown in gray. All other well records are shown in blue.....	14
4	Percentage of records where the county where the well is located could not be identified.....	16
5	Distribution of 8,193,428 groundwater wells in the United States after 1,342,439 records that had missing or incorrect locational details, or were duplicate entries were removed.....	17
6	Percentage of records removed after our QA/QC procedure was conducted.....	18
7	Number of wells constructed in a given year.....	19
8	Log-distribution of wells categorized by water use type.....	20
9	Number of wells with an irrigation purpose, broken into irrigation subclasses: Irrigation-Crop (C), Irrigation-Golf Courses (IR-G), Irrigation-Other (IR-O), and their combinations. The irrigation subclass of some irrigation wells is unknown.....	21
10	Distribution of well activity status - active, inactive, or unknown - by state.....	23
11	Heat map of well distribution over major aquifers in the United States (a) and the density of wells for each aquifer.....	24

# List of Tables

1	Overview of the two groundwater well data products.....	3
2	Names and source links of state agency's where we collected our data from.....	4
3	The column headers included in our dataset and their definitions.....	7
4	The water use categories and their corresponding definitions. Water use categories and definitions taken from the United States Geological Survey <sup>62</sup> .....	9

# List of Abbreviations

USGWIST United States Groundwater Wells Inventory Supplementary Tables

USGWD United States Groundwater Wells Database

QA/QC Quality assurance and quality control

FOIA Freedom of Information Act

# Chapter 1

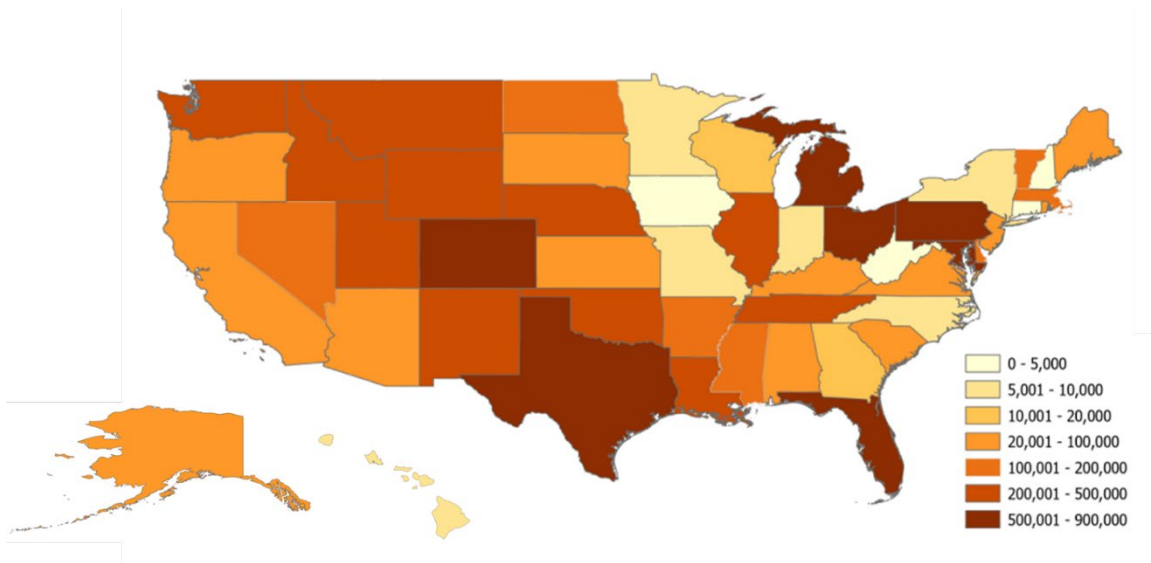
## Introduction

Millions of groundwater wells across the United States are used to extract water from aquifers for a wide variety of uses such as domestic, agricultural, commercial, industrial, and irrigation. This large quantity of wells has a considerable impact on the depletion of our groundwater storage<sup>1-3</sup>. Among the largest consumers of groundwater is the agricultural industry, whose overexploitation of aquifers threatens crop production in the United States<sup>4</sup>. Detailed data on groundwater well locations and their purpose on a national scale are needed to connect groundwater pumpage to the place of use and to quantify the volume being withdrawn.

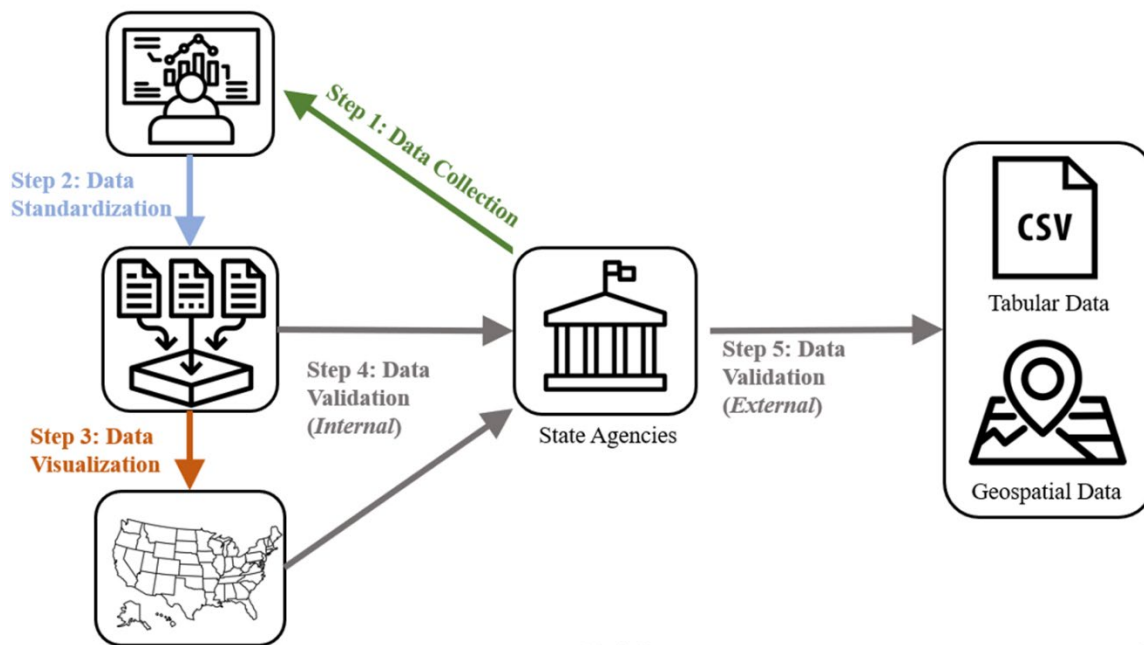
This paper presents a comprehensive dataset containing records of groundwater well locations, purposes, and other infrastructure details across the United States. While individual states maintain their own groundwater well datasets, a national collection of groundwater wells has yet to exist. Our groundwater database builds on the work done by the states to create a cohesive, standardized national dataset. The completeness and detail of our dataset are constrained by the quality of the state-level data.

In collaboration with state officials, we have identified all known groundwater well locations within the United States, for which state-level data was made available. A new data standard, named the United States Groundwater Wells Inventory Supplementary Tables (USGWIST), was developed to facilitate the standardization of the data collected. The standardized dataset, United States Groundwater Wells Database (USGWD) was sent to state officials to verify and provide additional feedback. The final data product has undergone rounds of quality assurance and quality control (QA/QC) in an attempt to confirm whether the wells were situated in the same county or state as initially reported. Some records lacked the required data needed to perform our QA/QC procedure, therefore only a portion of the data underwent analysis.

This paper presents two groundwater well datasets: (i) a tabular inventory detailing locational information, water uses, operational status, and other properties of each well; (ii) geospatial data depicting the location of all collected groundwater wells. Figure 1 depicts the number of groundwater wells reported by each state that is included within USGWD. An overview of the development of USGWD, including detailed descriptions of the datasets, are located in Figure 2 and Table 1.



**Figure 1:** Number of groundwater wells as reported by each state.



**Figure 2:** A depiction that outlines the process involved in developing the groundwater well datasets. Initially, data was collected from state agencies, and then standardized using our data standard, the United States Groundwater Wells Inventory Supplementary Tables (USGWIST). The data was then visually displayed using ArcGIS Pro. Lastly, data was reviewed by members of the research team and state agencies.

# Chapter 2

## Methods

We define a groundwater well as an excavation that has been constructed with the capability to extract water from an aquifer. Our groundwater well dataset, USGWD, represents all known groundwater wells in the United States. Groundwater well data is collected independently by each state. Thus, the coverage, length and completeness, and the quality of groundwater well records vary by state. Our standardized database maps disparate state records into a common data vocabulary that includes attributes collected across the majority of states, such as well infrastructure details, locational information, operational status, water source, and primary water use. As water quality, lithology, and withdrawal rates were not uniformly reported across states, they were not included in USGWD as less than 5% of states reported this data.

The creation of USGWD involved three steps: i) data collection, ii) data standardization, and iii) data validation. These steps are described below.

**Table 1:** Overview of the two groundwater well data products.

<b>Data Products</b>	<b>Description</b>	<b>Data Format</b>
United States Groundwater Wells Database (USGWD)	Tabular data detailing all groundwater wells collected, including their locational information, water uses, operational status, and other properties of each well.	CSV
United States Groundwater Wells Geospatial Database (USGWD)	Geospatial data showing the location of each groundwater well.	Esri Shapefile (SHP)

## 2.1 Data Collection

Our efforts to compile a national groundwater well dataset began by contacting potential data sources identified by Jasechko and Perrone<sup>5</sup>. Their previous study identified various state organizations as key data collectors and repositories for groundwater well data. Access and details regarding the contents of the data collected by Jasechko and Perrone were unavailable to us at the time of this study.

Following a review of the data repositories and state websites identified by Jasechko and Perrone, we proceeded to request groundwater well data from each state by contacting their respective representatives. Our primary means of communication were emails and phone calls. In cases where these methods were unsuccessful, we submitted Open Records Act or Freedom of Information Act (FOIA) requests to obtain the necessary records. Twenty-four of the states provided private datasets whereas the remaining twenty-six states had publicly available datasets.

**Table 2:** Names and source links of state agency’s where we collected our data from

State	Organization managing the well completion reports	Link to website hosting data	Reference
Alabama	Geological Survey of Alabama	<a href="#">[Link]</a>	6
Alaska	Alaska Department of Natural Resources	<a href="#">[Link]</a>	7
Arizona	Arizona Department of Water Resources	<a href="#">[Link]</a>	8
Arkansas	United States Geological Survey-Lower Mississippi Gulf Water Science Center	<a href="#">[Link]</a>	9
California	California Department of Water Resources	<a href="#">[Link]</a>	10
Colorado	Colorado Division of Water Resources	<a href="#">[Link]</a>	11
Connecticut	Connecticut Department of Energy & Environmental Protection	<a href="#">[Link]</a>	12
Delaware	Delaware Department of Natural Resources and Environmental Control	<a href="#">[Link]</a>	13
Florida (Northwest)	Northwest Florida Water Management District	<a href="#">[Link]</a>	14
Florida (South)	South Florida Water Management District	<a href="#">[Link]</a>	15
Florida (Southwest)	Southwest Florida Water Management District Water Supply Section	<a href="#">[Link]</a>	16

Florida (St. John's River)	St. Johns River Water Management District Bureau of Water Supply Planning	<a href="#">[Link]</a>	17
Florida (Suwannee River)	Suwannee River Management District	<a href="#">[Link]</a>	18
Georgia	Georgia Department of Natural Resources	<a href="#">[Link]</a>	19
Hawaii	State of Hawaii, Department of Land and Natural Resources	<a href="#">[Link]</a>	20
Idaho	Idaho Department of Water Resources	<a href="#">[Link]</a>	21
Illinois	Illinois State Geological Survey	<a href="#">[Link]</a>	22
Indiana	Indiana Department of Natural Resources	<a href="#">[Link]</a>	23
Iowa	Iowa Department of Natural Resources	<a href="#">[Link]</a>	24
Kansas	Kansas Geological Survey	<a href="#">[Link]</a>	25
Kentucky	Kentucky Division of Water	<a href="#">[Link]</a>	26
Louisiana	Louisiana Department of Natural Resource	<a href="#">[Link]</a>	27
Maine	Maine Geological Survey	<a href="#">[Link]</a>	28
Maryland	Maryland Department of the Environment	<a href="#">[Link]</a>	29
Massachusetts	Water Management Program, MassDEP	<a href="#">[Link]</a>	30
Michigan	Michigan Department of Environment, Great Lakes, and Energy	<a href="#">[Link]</a>	31
Minnesota	Minnesota Department of Natural Resources	<a href="#">[Link]</a>	32
Mississippi	Mississippi Department of Environmental Quality	<a href="#">[Link]</a>	33
Missouri	Missouri Department of Natural Resources	<a href="#">[Link]</a>	34
Montana	Montana Bureau of Mines and Geology	<a href="#">[Link]</a>	35
Nebraska	Nebraska Department of Natural Resources	<a href="#">[Link]</a>	36
Nevada	State of Nevada Division of Water Resources	<a href="#">[Link]</a>	37
New Hampshire	New Hampshire Department of Environmental Services	<a href="#">[Link]</a>	38

New Jersey	New Jersey Department of Environmental Protection	<a href="#">[Link]</a>	39
New Mexico	New Mexico Office of the State Engineer	<a href="#">[Link]</a>	40
New York	New York Department of Environmental Conservation	<a href="#">[Link]</a>	41
North Carolina (Part 1)	North Carolina Department of Environmental Quality	<a href="#">[Link]</a>	42
North Carolina (Part 2)	North Carolina Department of Environmental Quality	<a href="#">[Link]</a>	43
North Dakota	North Dakota Department of Water Resources	<a href="#">[Link]</a>	44
Ohio	Ohio Department of Natural Resources	<a href="#">[Link]</a>	45
Oklahoma	Oklahoma Water Resources Board	<a href="#">[Link]</a>	46
Oregon	Oregon Water Resources Department	<a href="#">[Link]</a>	47
Pennsylvania	Pennsylvania Department of Conservation & Natural Resources	<a href="#">[Link]</a>	48
Rhode Island	Rhode Island Department of Health	<a href="#">[Link]</a>	49
South Carolina (Part 1)	South Carolina Department of Natural Resources	<a href="#">[Link]</a>	50
South Carolina (Part 2)	South Carolina Department of Natural Resources	<a href="#">[Link]</a>	51
South Dakota	South Dakota Department of Agriculture & Natural Resources	<a href="#">[Link]</a>	52
Tennessee	Tennessee Department of Environment & Conservation	<a href="#">[Link]</a>	53
Texas (Part 1 & 2)	Texas Water Development Board	<a href="#">[Link]</a>	54
Utah	Utah Geospatial Resource Center	<a href="#">[Link]</a>	55
Vermont	Vermont Department of Environmental Conservation	<a href="#">[Link]</a>	56
Virginia	Virginia Department of Environmental Quality	<a href="#">[Link]</a>	57
Washington	Washington State Department of Ecology	<a href="#">[Link]</a>	58
West Virginia	West Virginia Department of Environmental Protection	<a href="#">[Link]</a>	59
Wisconsin	Wisconsin Department of Natural Resources	<a href="#">[Link]</a>	60

Wyoming	Wyoming State Engineer's Office	<a href="#">[Link]</a>	61
---------	---------------------------------	------------------------	----

We established direct contact with officials in each state, who provided valuable insights into how the data was collected, defined, and the meaning behind any notations used. State officials also shared associated metadata, disclaimers, publishing restrictions, and a quantitative measure of how representative the data is for the entire state. Only twenty states were able to provide an estimate of the completeness of their well records but most noted that their well records were likely incomplete, especially for very old wells that predated the start of their datasets. Many states noted that information on new and existing wells is received on a daily basis, and as such, our study reflects only the most current data held by the primary data collectors at the time of data collection. Table S2 lists the date the data were collected from each state, as well as the primary official and organization contacted, disclaimers, and representation estimations for each state. Data collection was conducted between August 2021 and March 2023. Any wells added to a state repository after the collection date are not included in USGWD.

All states gave explicit consent for us to publish the groundwater records they provided to us. Five states requested that personal identification information, such as owner names, addresses, and phone numbers, not be made public, while only one state (Wisconsin) requested that location coordinates not be included in the dataset. We did not include the owner's name or contact information in USGWD for any state; however, we include the location of each groundwater well for all states except Wisconsin.

## 2.2 Data Standardization

The formats and type of data provided to us by states varied. USGWD contains consistent definitions and representation for all groundwater wells included in this study. Header names were developed to classify frequently reported data into common groups. To account for the variation in how each state defines their attributes, the column definitions within our dataset are broad in scope. The headers from the state datasets included in each of our headers are reported in Table S3.

**Table 3:** The column headers included in our dataset and their definitions.

Header	Definition
Well ID	Unique well identifier assigned by us
Well ID (State)	Unique well identifier given by the state
Longitude	Longitude of well location in Decimal Degrees
Latitude	Latitude coordinate of well location in Decimal Degrees
County	County of well location
State	State of well location
FIPS	Unique Federal Information Processing number for each county according to coordinates

Aquifer-Specific	Subdivided broad aquifer, or aquifers, in which the well overlays based on our analysis
Aquifer-Broad	Larger aquifer, or aquifer systems, in which the well overlays based on our analysis
Subwatershed-HUC12	12-digit Hydrologic Unit Code (HUC12) subwatershed in which well is located, based on our analysis
Subwatershed-Name	Name of the subwatershed in which well is located, based on our analysis
Location Verified	The location of the well was determined by GPS or field checking (Yes, No)
Flag County	The number of wells found to be outside their reported county, based on our analysis
Flag State	The number of wells found to be outside their reported state, based on our analysis
Well Depth (Feet)	Total depth of well, measured in feet below land surface
Screen Depth (Feet)	Distance from land surface to the top of the well screen, measured in feet
Length of Screen (Feet)	Length of the well screen from the screen opening to the end of the screen, measured in feet
Well Capacity (GPM)	Estimated amount of water the well can withdrawal, measured in gallons per minute
Lithological Data	Existence of lithological records (Yes, No)
Surface Elevation (Feet)	Elevation of the ground surface at the well head, in feet above sea level
Status	Status of the well (Active, Inactive)
Year Well was Constructed	Date well was constructed
Year Reported	Date well was reported to the overseeing organization or the date in which it was entered into the system. Often times these two instances are the same
USGS Water Use Category	Primary use of the well identified by our definitions
Water Quality (Potable/Non-Potable)	Distinction between potable and non-potable water (Yes, No)
Well ID	Unique well identifier assigned by us
Well ID (State)	Unique well identifier given by the state

With guidance from state officials, we were able to identify attributes within the state records that aligned with the attribute definitions of USGWD. State officials were able to provide official documentation that reported the definitions of the headers within their dataset and labels used for each entry within those columns. When documentation was not available, officials provided us with guidance based on their expertise on how features within the dataset were defined. While our dataset includes data based on features most commonly reported across all states, some data reported in the states' datasets was not included in this study. The column headers from the state's datasets that were not included in our dataset are reported in Table S5.

The definitions for our water use categories are broad in scope to account for variations in state definitions. The naming conventions adopted for the water use categories were primarily taken from USGS water use definitions, though slight derivations or additions were required. Specifically, Irrigation-Crop (IR-C), Irrigation-Golf Courses (IR-G), Irrigation-Other (IR-O), Remediation (RM), and Other water use categories vary from the original USGS water use definition. The three irrigation subcategories (IR-C, IR-G, and IR-O) allow for greater distinction within the broader Irrigation (IR) category to which all these entries belong.

**Table 4:** The water use categories and their corresponding definitions. Water use categories and definitions taken from the United States Geological Survey<sup>62</sup>.

<b>USGS Water Use Category</b>	<b>USGS Water Use Category Definition</b>
Irrigation (IR)	“Water that is applied by an irrigation system to assist crop and pasture growth, or to maintain vegetation on recreational lands such as parks and golf courses. Irrigation includes water that is applied for pre-irrigation, frost protection, chemical application, weed control, field preparation, harvesting, dust suppression, leaching of salts from the root zone. Irrigation water use estimates also include conveyance losses. See also conveyance loss, microirrigation system, sprinkler irrigation system, and surface irrigation system.”
Irrigation-Crop (IR-C)	Water that is applied by an irrigation system to assist crop and pasture growth. Irrigation includes water that is applied for pre-irrigation, frost protection, chemical application, weed control, field preparation, harvesting, dust suppression, leaching of salts from the root zone. Irrigation water use estimates also include conveyance losses. See also conveyance loss, microirrigation system, sprinkler irrigation system, and surface irrigation system.
Irrigation-Golf Courses (IR-G)	Water that is applied by an irrigation system to recreational lands such as parks and golf courses.
Irrigation-Other (IR-O)	Irrigation not covered under IR-C or IR-G.

Livestock (LV)	“Water used for livestock watering, feedlots, dairy operations, and other on-farm needs. Types of livestock include dairy cows and heifers, beef cattle and calves, sheep and lambs, goats, hogs and pigs, horses and poultry.”
Commercial (CO)	“Water for motels, hotels, restaurants, office buildings, other commercial facilities, military and nonmilitary institutions, and (for 1990 and 1995) offstream fish hatcheries. Water may be obtained from a public-supply system or may be self-supplied. Commercial water-use estimates were included in industrial water use until 1985, then were reported as a separate category. Commercial water use estimates were last reported nationally for 1995.”
Industrial (IN)	“Water used for fabrication, processing, washing, and cooling. Includes industries such as chemical and allied products, food, mining, paper and allied products, petroleum refining, and steel.”
Aquaculture (AQ)	“Water use associated with the farming of finfish, shellfish, and other organisms that live in water, and offstream water use associated with fish hatcheries.”
Mining (MI)	“Water used for the extraction of naturally occurring minerals including solids (such as coal, sand, gravel, and other ores), liquids (such as crude petroleum), and gases (such as natural gas). Also includes uses associated with quarrying, milling and other preparations customarily done at the mine site, injection of water for secondary oil recovery or for unconventional oil and gas recovery (such as hydraulic fracturing), and other operations associated with mining activity. Does not include water associated with dewatering of the aquifer that is not put to beneficial use. Also does not include water used in processing, such as smelting, refining petroleum, or slurry pipeline operations. These processing uses are included in industrial water use.”
Hydroelectric Power Generation (PH)	“The use of water in the generation of electricity at plants where the turbine generators are driven by moving water. Hydroelectric water use is most commonly an instream use. Hydroelectric power water use was referred to as water power from 1950-1960.”
Remediation (RM)	Water use that goes back to the environment in a way to prevent or attempt to reverse environmental damage.
Water Supply / Public Supply (WS)	“Water withdrawn by public and private water suppliers that furnish water to at least 25 people or have a minimum of 15 connections. Public suppliers provide water for a variety of uses, such as domestic, commercial, industrial, thermoelectric-power, and public water use. See also commercial water use, domestic water use, industrial water use, public-supply deliveries, public water use, and thermoelectric-power water use.”

Domestic (DO)	“Water used for indoor household purposes such as drinking, food preparation, bathing, washing clothes and dishes, flushing toilets, and outdoor purposes such as watering lawns and gardens. Domestic water use includes potable and non-potable water provided to households by a public water supplier (domestic deliveries) and self-supplied water.”
Sewage Treatment (ST)	“Water returned to the hydrologic system by wastewater-treatment facilities. Wastewater-treatment return flows were referred to as sewage treatment in 1985. Wastewater treatment return flows were last reported in 1995. See also water use.”
Thermoelectric Power Generation (TE)	“Water used in the process of generating electricity with steam-driven turbine generators. Term used in previous water-use circulars to describe the combined public-supply deliveries to thermoelectric-power plants and self-supplied thermoelectric-power withdrawals.”
Other	Anything not covered by the other definitions.

With the assistance of state officials and documentation provided during the data collection phase, we categorized all of the uses of each groundwater well into our own water use categories. We matched the state water use naming conventions with our own naming conventions, as reported in Table S7. A well can be used for multiple purposes. Irrigation (IR) is the primary label to capture a well that is used for irrigation purposes when no specific details on what is being irrigated are provided. Irrigation-Crop (IR-C), Irrigation-Golf Courses (IR-G), Irrigation-Other (IR-O) are a subgroup of Irrigation (IR) and are used to capture wells that have specific irrigation uses defined. A well classified into one of the subgroups, is also grouped into the primary Irrigation (IR) classification.

Duplicate records were identified and removed from the dataset, with the assistance of state officials who provided guidance on how to handle such cases. Three states informed us of duplicate records and helped us to process those records accordingly. In Arkansas, each crop being irrigated by a well was recorded separately, potentially resulting in multiple entries for a single well. The well infrastructure data did not vary between entries which allowed us to exclude duplicate entries. In New Jersey, wells were recorded across different phases, including permit, record, and decommissioning phases. For decommissioned wells, only the decommissioning entry was included in our study, while for wells that had both permit and record entries, the permit entry was excluded. In Northwest Florida, wells were recorded when they were constructed and when they were decommissioned. In cases where both entries existed for a single well, the constructed entry was excluded, and the decommissioned entry was included in our study.

In the data collection phase, some states reported abbreviations or codes to identify the county a well was located in. These abbreviations and codes were replaced with the county names within our dataset. The corresponding county codes and abbreviations that were replaced are listed in Table S8. Each entry was assigned a Federal Information Processing System (FIPS) code for the state and county. The status of a well in terms of whether it is currently pumping water from an underground source or not

varies across states. States that did not record the status of wells provided guidance on how to classify them as active or inactive based on other data within the dataset. If there was uncertainty regarding the status of a well, it was left blank within our dataset. A list of state labels corresponding to our status designations is provided in Table S9.

States that provided screen data reported the starting and ending points of the screens in feet from the land surface or the length of the screen. Screen lengths were calculated by subtracting the depth at which the screen began from the depth at which it ended to obtain screen lengths for all states. Unit conversions were needed in some cases to maintain consistency. The well capacity for the state of Wisconsin was originally reported in gallons per day (GPD), but it was converted to gallons per minute (GPM) to maintain consistency with other states. The coordinate systems used to report data varied, so all coordinates were converted to decimal degrees for our dataset and were displayed using the same projection system.

In our study, we examined a well's potential water source and hydrologic position by identifying the aquifer and Hydrologic Unit Code (HUC12) boundary that each well was located in. We achieved this by overlying aquifer<sup>63</sup> and HUC12<sup>64</sup> shapefiles with the groundwater well locations. The aquifer geodatabase includes 440 of the most prominent aquifers in the US, but some of the smaller or less studied aquifers may be missing<sup>65</sup>. Therefore, it is possible that the well draws water from a different aquifer than the one listed in USGWD. Two separate columns were added to our dataset, one for the name of the underlying aquifer and one for the HUC12 that each well is located within.

## 2.3 Technical Validation

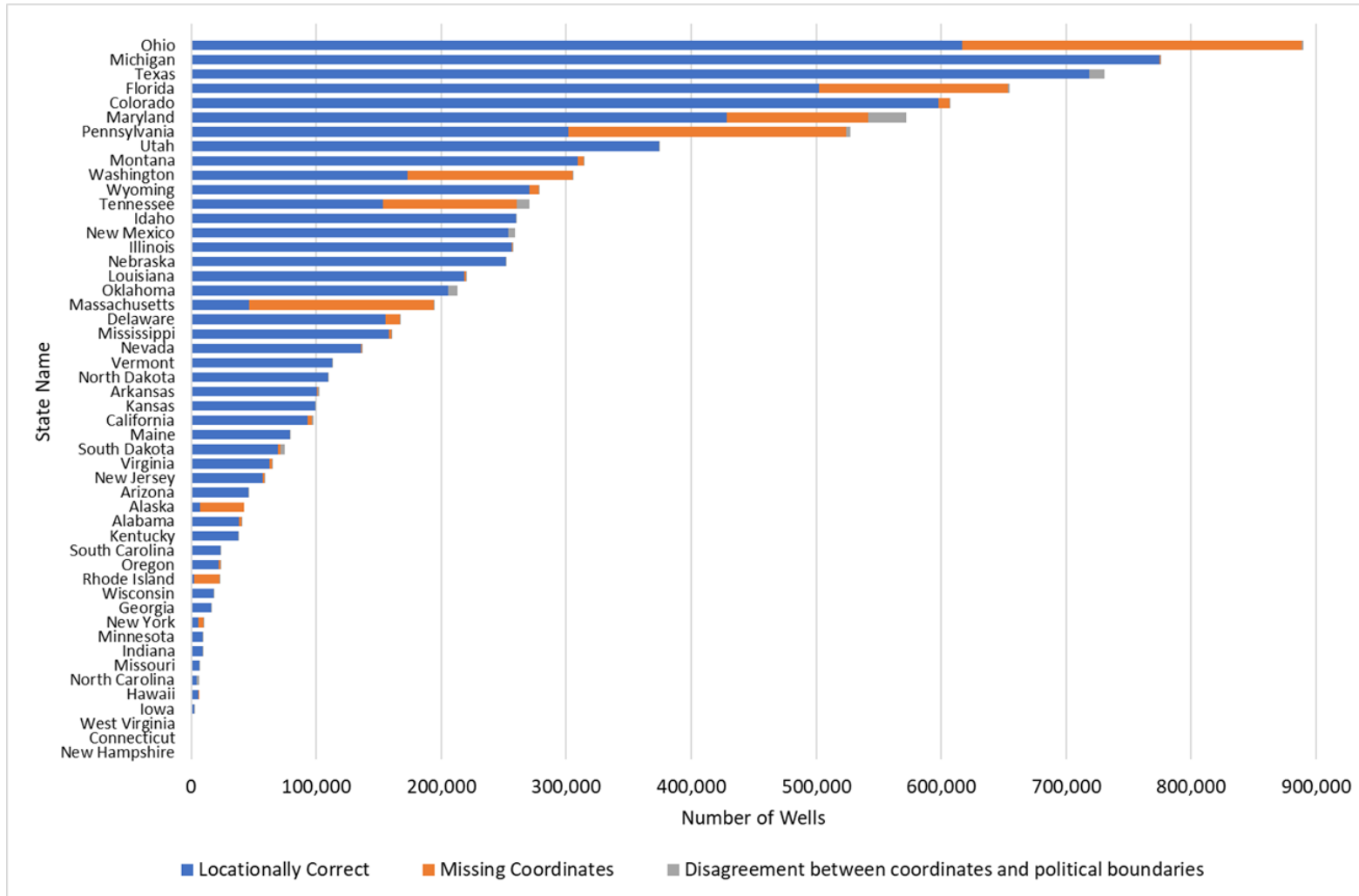
The data underwent multiple rounds of quality assurance and quality control (QA/QC) before forming our finalized dataset. During the QA/QC process, we focused on determining the accuracy of the locational data. Most groundwater well records were reported to the states by the owner or the well driller. Records that were collected using a global positioning system (GPS) device or were field verified by personnel within the collecting agency were considered to have higher locational certainty than those that were not. We have included a column in our dataset indicating the level of location verification for each record, while the corresponding state labels for these classifications are reported in Table S10. Self-reported well locations are not necessarily inaccurate, they are just not verified by the state.

There can sometimes be data entry errors, either by the state or the initial person reporting the groundwater well location to the state. To determine the agreement between the county and coordinates assigned to each well within state records, county<sup>66</sup> shapefiles were overlaid with the groundwater well locations. A new column was added to our dataset to report the county that each well resides in based on its coordinates. This allowed us to compare the state-reported counties with the counties we derived from the wells' latitude and longitude. The comparison was performed to assess the accuracy of the states' data regarding the location of the wells.

Each well's connection to different hydrologic features was assessed by comparing the aquifer and HUC12 subwatershed we associated with the well based on its given coordinates to the aquifer and watershed assigned by the state. The aquifer that is supposed to serve as the water source for a well was provided by twenty-three states. The

variability in naming conventions between states and within the literature made it difficult to compare the states' data with our own aquifer analysis. Our aquifer classification for each well is within USGWD. Four states provided information on the watershed where each well is located. We assigned wells to watersheds only at the HUC12 level; inconsistencies in the definition and classification of watersheds across the states prevented us from assessing the accuracy of the reported watershed information.

Over nine million well records were collected for the United States, but about one million of them had missing coordinates or were located outside of the state or county they were originally reported in. This resulted in only eight million records being considered locationally correct. Figure 3 depicts the distribution of these records by state, based on whether they are locationally correct, missing coordinates, or outside of the original state or county.



**Figure 3:** Total number of well records collected for each state with locational details. The number of the total well records that do not report coordinates are shown in orange, while records with coordinates that do not match the state-provided county or state where the well is supposed to exist are shown in gray. All other well records are shown in blue.

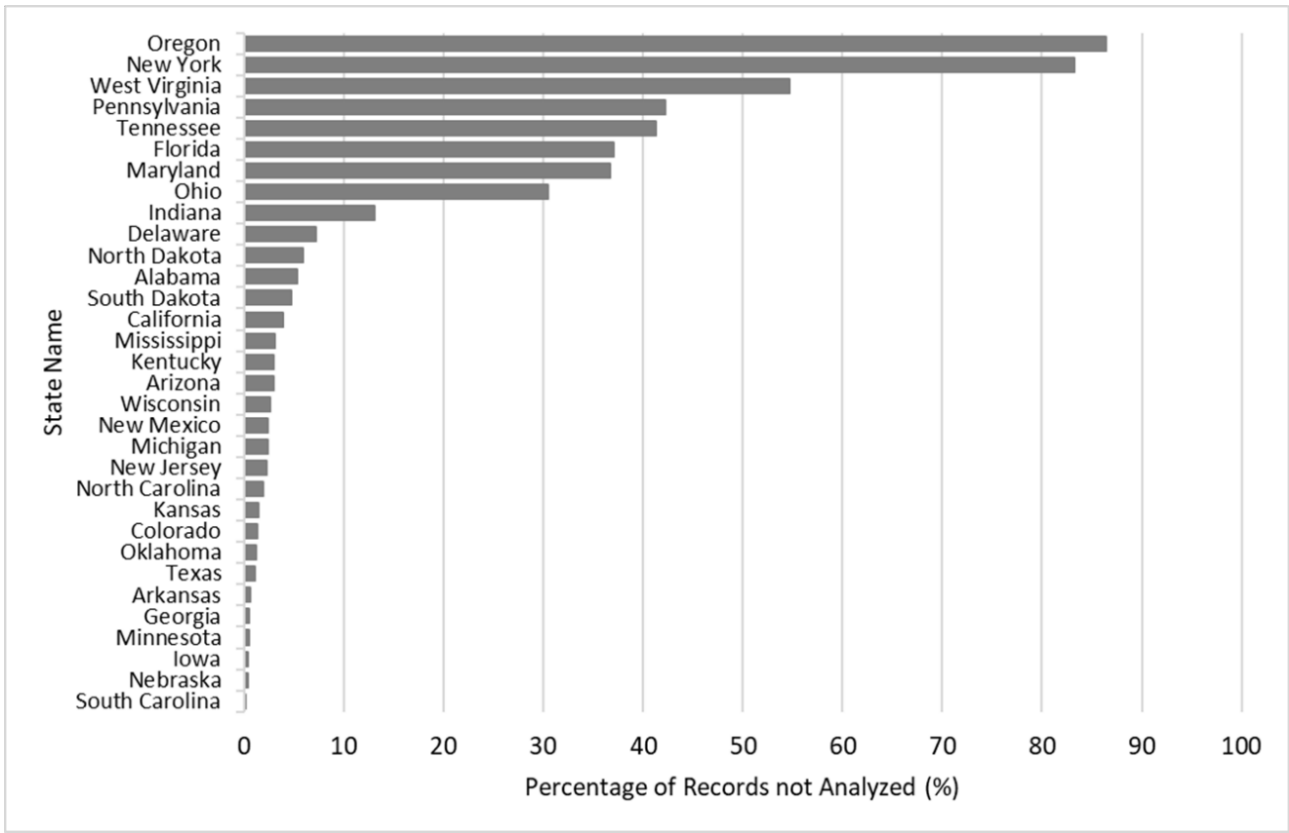
The distribution of the collected well records varies significantly across states, with a large proportion of the records concentrated in a small number of states. Approximately 49% of the well records in our dataset are located in seven states: Texas, Ohio, Florida, Michigan, Colorado, Maryland, and Pennsylvania. This uneven distribution is likely due in part to variations in data sharing restrictions and practices across states, with some states having less restrictive laws and regulations than others. It is important to note that this distribution does not necessarily reflect the actual distribution of all wells across the United States, and we caution against generalizing our findings beyond the data included in our study.

There are 7,581 wells whose coordinates are outside the reporting state. Four states, Maryland, Tennessee, North Dakota, and Massachusetts, account for approximately 70% of the total wells located outside their reported state. Maryland had the most wells located outside the state based on its locational coordinates. However, North Dakota had the highest percentage of error with approximately 1% of the well records collected from the state of North Dakota being shown as located outside of the state.

A lack of reported coordinates prevented us from verifying whether all well records were located within their respective states. Approximately 91% of the total wells missing coordinates are from Ohio, Pennsylvania, Florida, Massachusetts, Washington, Maryland, and Tennessee. Ohio has the highest number of wells missing coordinates, with 271,396 records missing latitude and longitude data. However, Rhode Island, Alaska, and Massachusetts had the highest percentage of missing coordinates compared to the total number of well records collected for each state, at approximately 89%, 83%, and 76%, respectively.

To assess whether a well was located within its reported county, we utilized both the coordinate and county data provided by each state. There were 82,870 wells whose coordinates did not match the state-reported county where the well was supposed to be located. Four states, Maryland, Texas, Tennessee, and Oklahoma account for approximately 72% of the total wells located outside their reported counties. Among all states, Maryland had the highest percentage of error with approximately 5% of the well records collected from the state of Maryland being located outside of their reported counties.

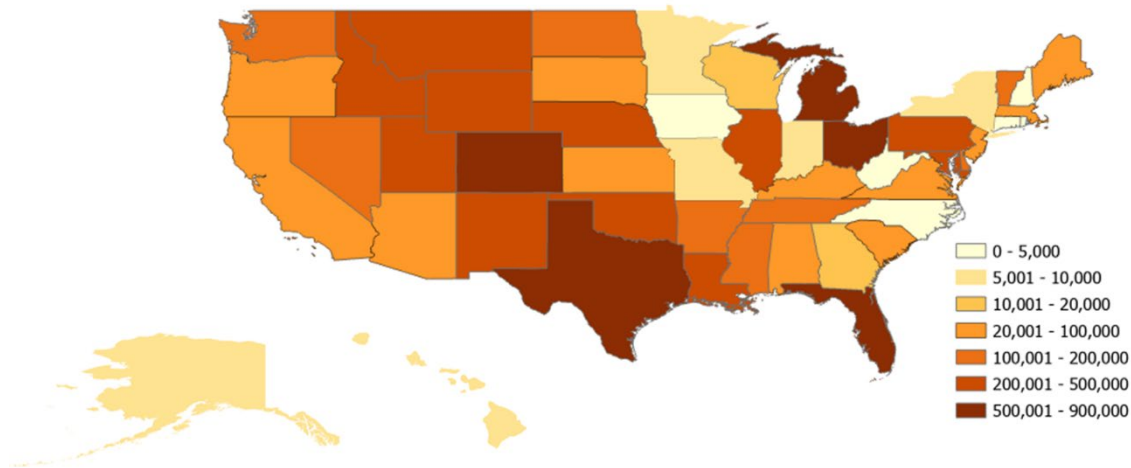
Missing coordinate or county data prevented us from determining whether all records were located in their reported counties. Oregon, New York, and West Virginia had the highest percentage of missing data compared to the total number of records collected for each state that we could perform the analysis on. The percentage of records that we could not analyze for each state is shown in Figure 4. States not included in Figure 4 did not report any county or coordinate data, so we could not perform this analysis on records from those states.



**Figure 4:** Percentage of records where the county where the well is located could not be identified.

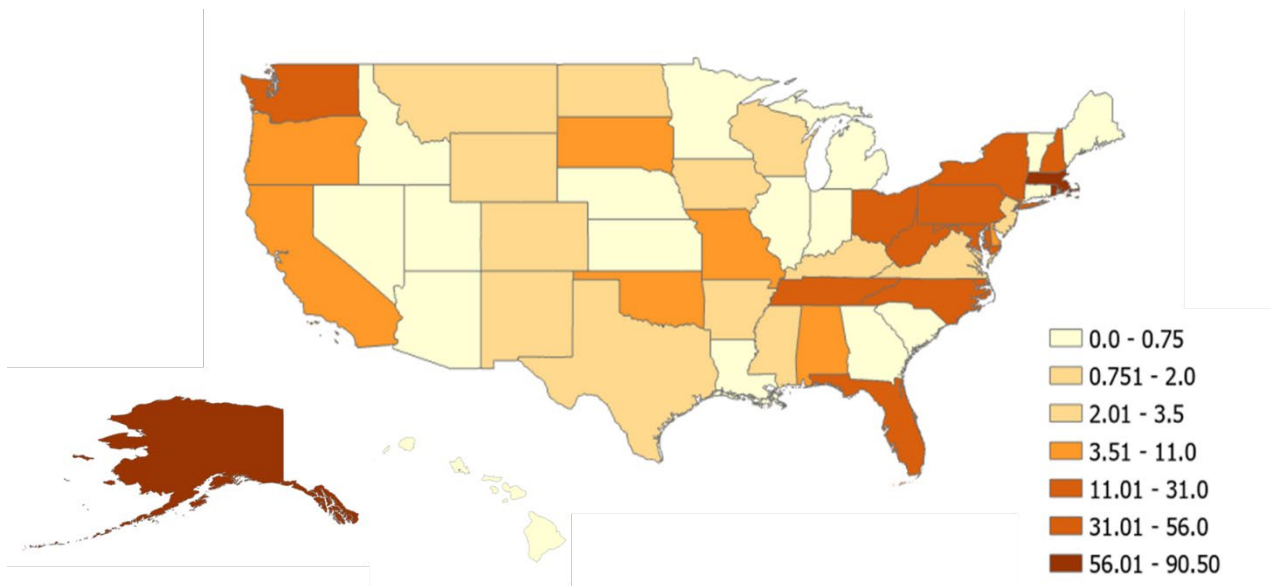
## Results

After excluding the records located outside their reported county and state, over eight million records remained for the rest of the analysis. The distribution of these records is presented in Figure 5.



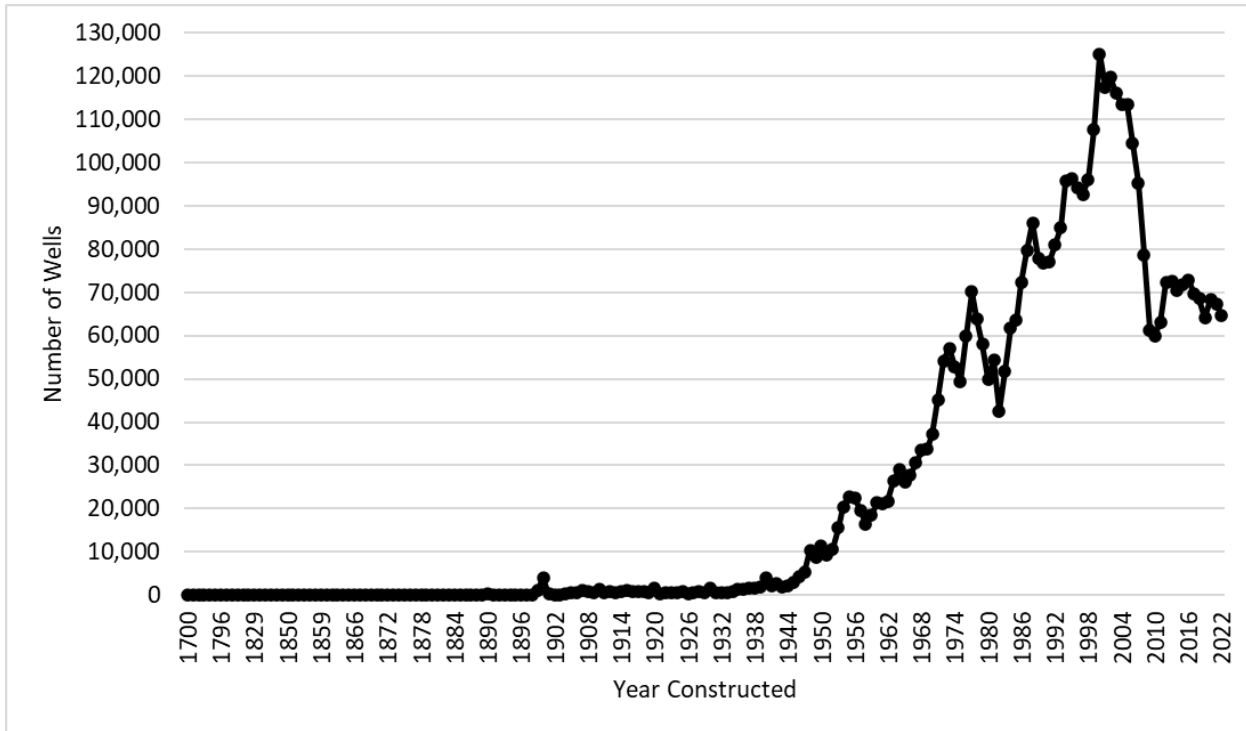
**Figure 5:** Distribution of 8,193,428 groundwater wells in the United States after 1,342,439 records that had missing or incorrect locational details, or were duplicate entries were removed.

Michigan had the highest number of wells (774,532) in the cleaned dataset, followed by Texas (718,233), Ohio (616,883), and Colorado (597,903). New Hampshire had the lowest number of wells (80), followed by West Virginia (160), Connecticut (329), and Rhode Island (2,170). Data sharing restrictions varied among the states, and it is possible that many states have more wells than the ones captured in this dataset. Similarly, states with fewer reported wells may have stricter sharing restrictions compared to states with a larger number of wells. Figure 6 depicts the distribution of records removed after our QA/QC procedure was performed in percentage.



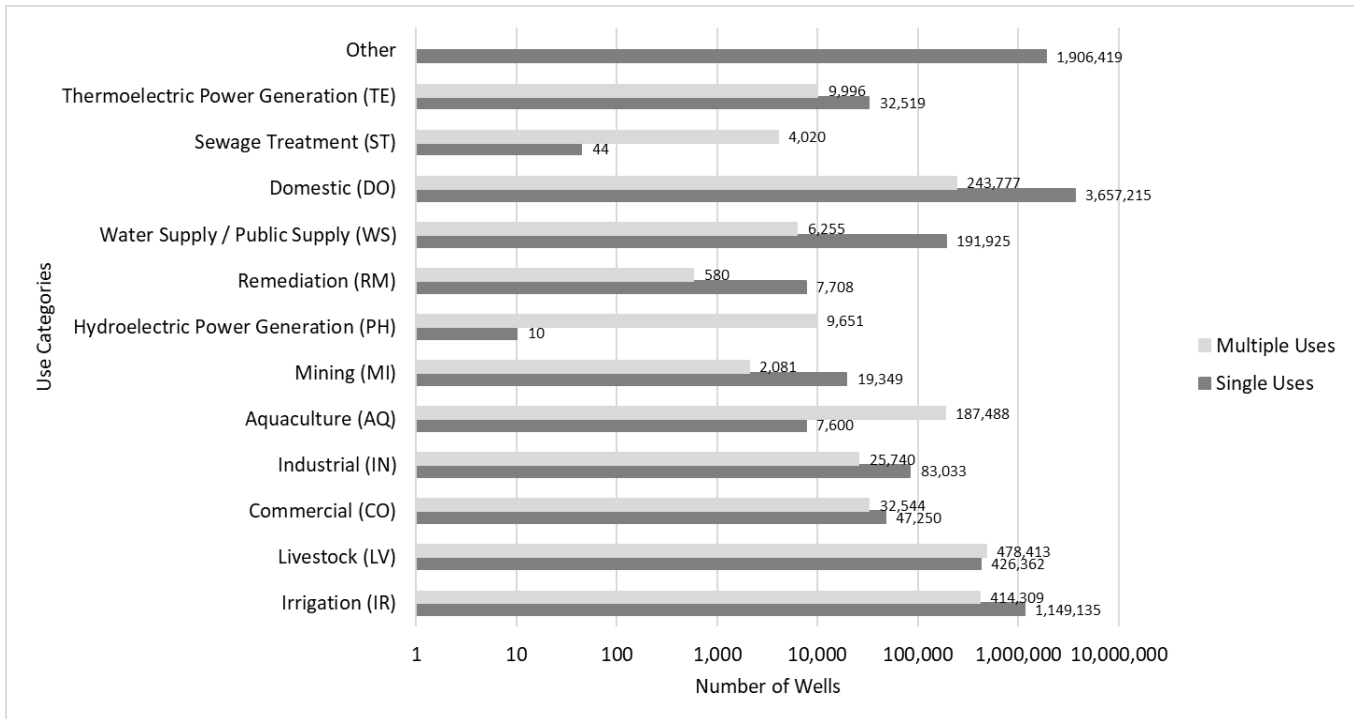
**Figure 6:** Percentage of records removed after our QA/QC procedure was conducted.

Figure 7 provides a visual representation of well construction over time. Nineteen states did not report the well construction date and were therefore excluded from our analysis of well construction over time. Of the wells included in our analysis, approximately 87% were constructed between 1972 and 2022. The year with the highest number of wells constructed was 2000. Most states noted that their current well databases did not start until the 1970s. Dates prior to the 1970s may under-represent the number of wells constructed due to lack of required reporting for that time period. As a result, caution should be exercised when interpreting the number of wells constructed during those years. After the 1970s the data is still highly variable with increases and decreases throughout the years, after further examining the data a clear indication for this was not apparent.



**Figure 7:** Number of wells constructed in a given year.

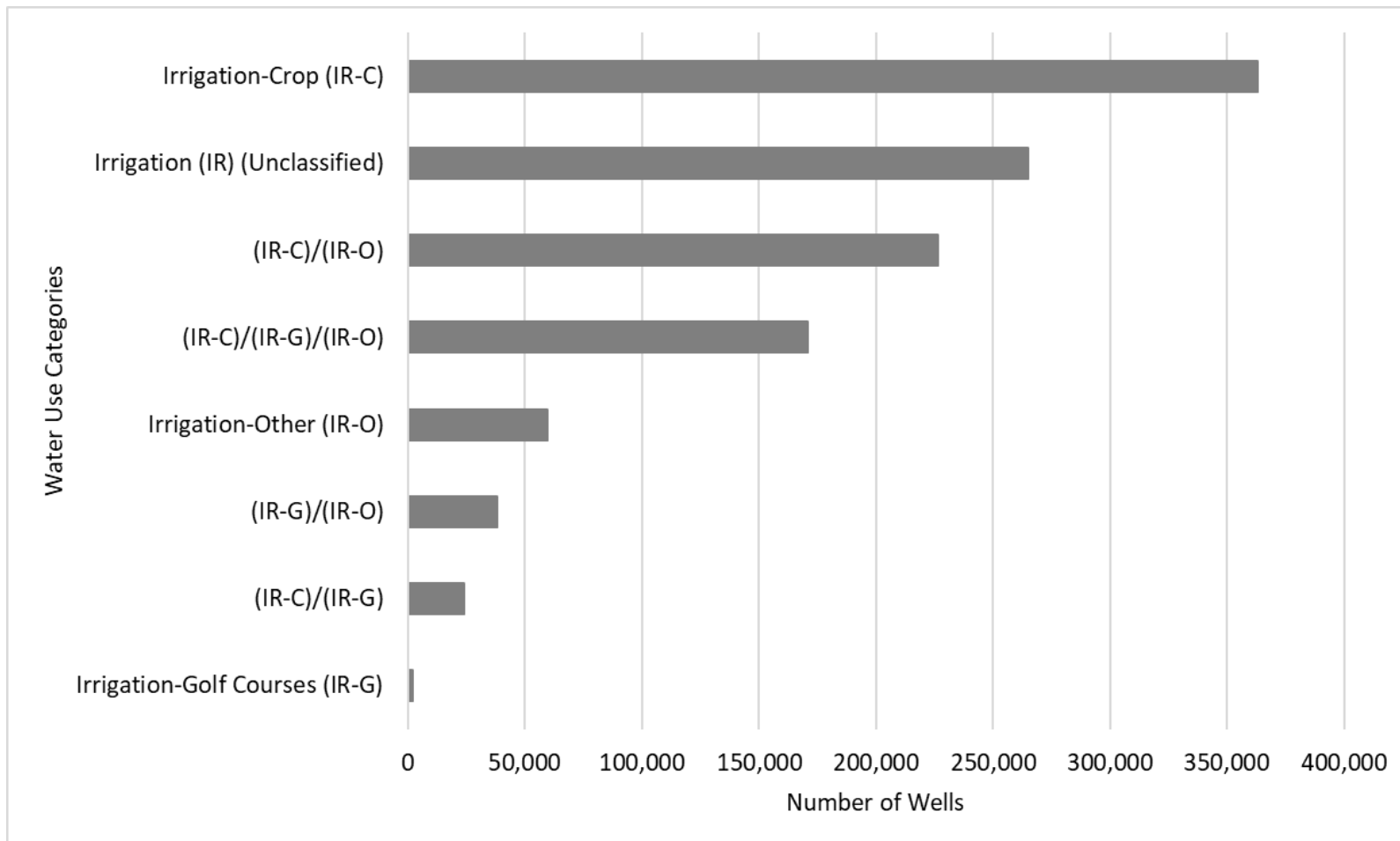
Each well in our dataset was assigned to one or more water use categories based on how the respective states categorized them. Figure 8 displays the breakdown of wells by water use category. Some wells have multiple water uses, and therefore, are included in more than one category. Of the over eight million wells analyzed, 664,859 were categorized as having multiple water uses. Wells that did not have multiple water uses were only assigned to one water use category.



**Figure 8:** Log-distribution of wells categorized by water use type.

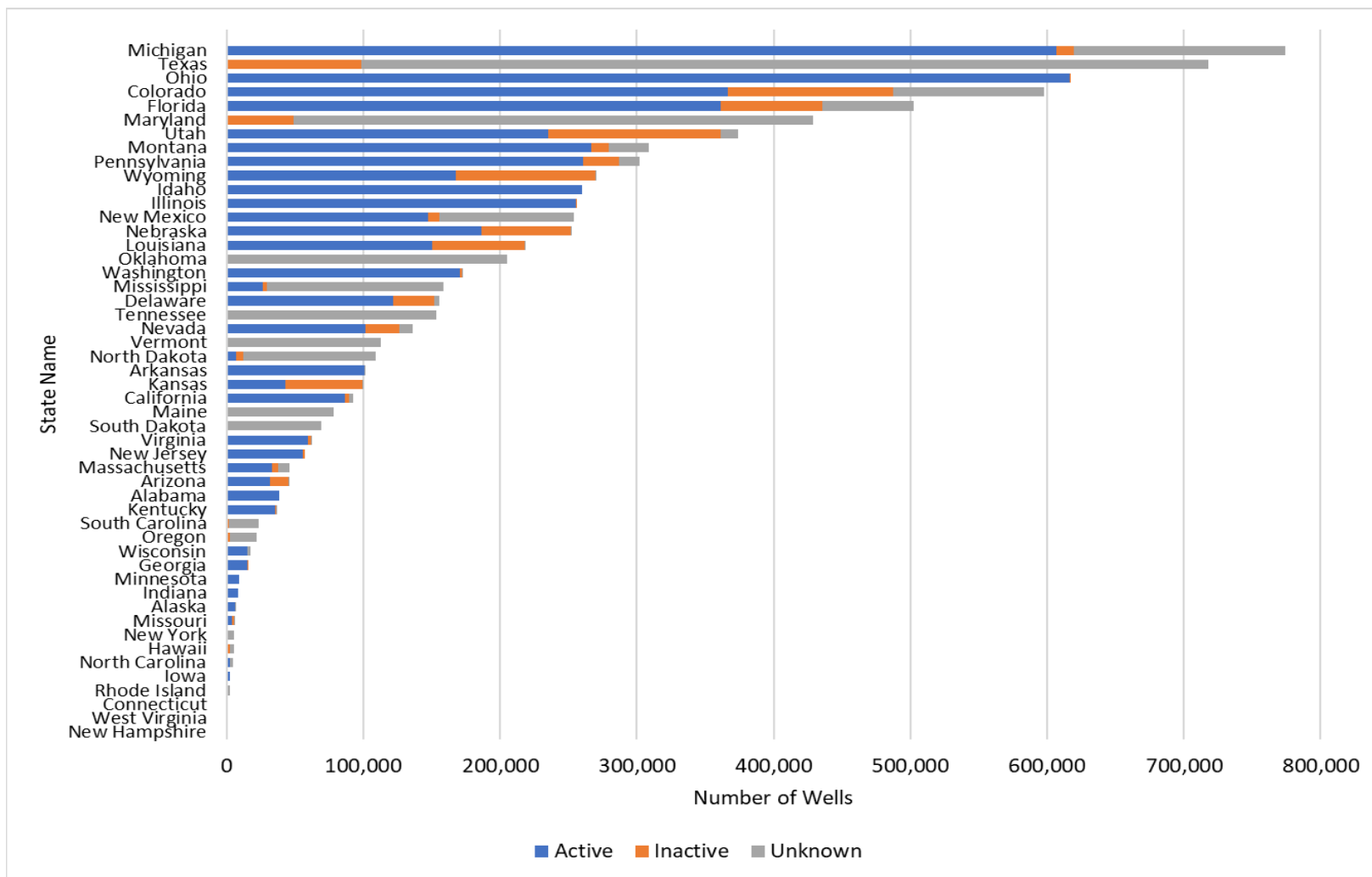
Approximately 8% of the total number of wells within this dataset have multiple water uses. The remaining 92% of the total wells have one water use assigned to them. The Livestock (LV), Irrigation (IR), and Domestic (DO) categories account for 80% of water use types of wells with multiple water uses. Among these categories, Irrigation (IR) has the highest number of wells assigned, followed by Livestock (LV) and Domestic (DO). In relation to the total number of wells that are designated as having an Irrigation (IR) water use, approximately 27% have multiple water uses, while a little under 73% are exclusively used for Irrigation (IR). Similarly, approximately 53% of the total number of wells assigned as Livestock (LV) are categorized as having multiple uses. Approximately 6% of all of the domestic wells are used for multiple purposes, with the remaining 94% being used only for domestic use. Out of the wells that are only used for a single water use, approximately 82% are designated as either Domestic (DO), Other, or Irrigation (IR). The distribution of single water use wells is shown in Figure 8.

Figure 8 only reports the number of wells categorized as Irrigation (IR) and does not further distinguish between different types of irrigation, such as irrigation for crops, golf courses, or landscaping. When known, we further assign the irrigation subclass to each irrigation well. There are 1,149,135 single-use Irrigation (IR) wells, 31.6% of which are exclusively for Irrigation-Crop (IR-C), 0.2% are Irrigation-Golf Courses (IR-G), 5.2% are Irrigation-Other (IR-O), and 23% are unknown irrigation uses (Figure 9). The remaining irrigation wells serve multiple irrigation purposes: 14.9% are a combination of Irrigation-Crop (IR-C), Irrigation-Golf Courses (IR-G), and Irrigation-Other (IR-O), 19.7% are a combination of Irrigation-Crop (IR-C), and Irrigation-Other (IR-O), 2.1% a combination of Irrigation-Crop (IR-C) and Irrigation-Golf Courses (IR-G), and 3.3% a combination of Irrigation-Golf Courses (IR-G) and Irrigation-Other (IR-O).



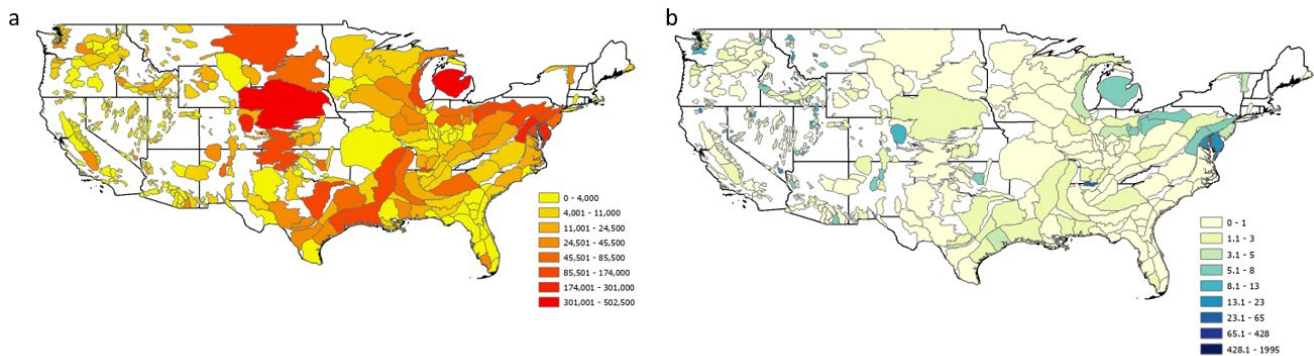
**Figure 9:** Number of wells with an irrigation purpose, broken into irrigation subclasses: Irrigation-Crop (C), Irrigation-Golf Courses (IR-G), Irrigation-Other (IR-O), and their combinations. The irrigation subclass of some irrigation wells is unknown.

Of the 8,193,428 wells in USGWD, 4,853,469 (59%) are actively being used to withdraw water, 918,395 (11%) are inactive, and 2,421,564 (30%) were classified as unknown. Each well in the dataset was assigned a status of either active, inactive, or unknown. Two-thirds (67%) of active wells are located in just nine states: Ohio, Michigan, Colorado, Florida, Montana, Pennsylvania, Idaho, Illinois, and Utah. Idaho had the highest number of active wells, with 100% of the state's reported wells being active (note that some states only reported active wells). Approximately 57% of inactive wells were located in Utah, Colorado, Wyoming, Texas, and Florida. Utah had the highest number of inactive wells, with 34% of its total wells being inactive. Approximately two-thirds of unknown wells were located in Texas, Maryland, Oklahoma, Michigan, Tennessee, and Mississippi. Texas has the highest number of wells with an unknown activity status: 86% of the wells in Texas are categorized as having an unknown status.



**Figure 10:** Distribution of well activity status - active, inactive, or unknown - by state.

Well locations were overlaid with major aquifers in the United States to identify which aquifers had the most wells overlying them. Since the aquifer database, we use (GebreEgziabher et al., 2022) only reports the 440 largest aquifers, and since some areas have overlapping aquifers, the listed aquifer may not be the one the well draws water from. With that said, the listed aquifer is the most likely water source for most wells. Figure 11 depicts a heat map showing the distribution of wells and their corresponding underlying aquifers and the density of wells for each aquifer.



**Figure 11:** Heat map of well distribution over major aquifers in the United States (a) and the density of wells for each aquifer

Of the 8,193,428 wells in USGWD, 6,459,106 wells are located in an aquifer boundary. The aquifers with the highest number of wells located within their boundaries are the Nussbaum (502,434), Michigan Basin (453,084), and Northern High Plains aquifers (301,312). The Kettle River Valley, Apalachicola Delta, Central Lake Area, Eastern Flatwoods Southshores, Eastern Flatwoods Northshores, Intermediate, and Biscayne aquifers have no wells located within their boundaries.

Additional analysis was conducted to investigate the density of wells within the aquifers, measured in number of wells per square kilometer. The Powell Park and Nussbaum aquifers had the highest density of wells per square kilometer. The Powell Park aquifer had the highest density with approximately 1,995 wells per square kilometer, which is about 1,567 more wells per square kilometer than the Nussbaum aquifer. The Kettle River Valley, Apalachicola Delta, Central Lake Area, Eastern Flatwoods Southshore, Eastern Flatwoods Northshore, Intermediate, and Biscayne aquifers had the lowest density of wells per square kilometer due to a lack of wells located within them. Despite the Michigan Basin, and Northern High Plains aquifers having the highest number of wells within their boundaries, their larger size allows for more dispersion, resulting in a lower density of wells per square kilometer.

# Conclusion

Groundwater is a critical resource that has significant impacts on the environment, water availability, and the economy. However, a publicly available national database detailing the uses and locations of groundwater wells is currently lacking. In this study, we developed a national dataset of groundwater wells, containing locations, purposes, and other infrastructure details across the United States, to address this gap. While this study attempted to lessen the existing gaps, there are limitations to our datasets. Data sharing restrictions varied across states, resulting in partial datasets for some states. The data collected by each state also varied, which affected the analysis conducted in this study. The dataset developed in this study can be utilized by the public to further analyze groundwater depletion in the United States. By overlaying land use coverage with our dataset, parcels of land can be assigned to specific water uses. Additionally, the dataset could be paired with water use data to quantify the amount of groundwater being withdrawn from each well, which could lead to identifying areas where better water management practices are needed.

## References

1. Konikow, L. F., & Kendy, E. (2005). Groundwater depletion: A global problem. *Hydrogeology Journal*, 13, 317-320.
2. Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., ... & Rodell, M. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters*, 38(3).
3. Castle, S. L., Thomas, B. F., Reager, J. T., Rodell, M., Swenson, S. C., & Famiglietti, J. S. (2014). Groundwater depletion during drought threatens future water security of the Colorado River Basin. *Geophysical research letters*, 41(16), 5904-5911.
4. Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., & McMahon, P. B. (2012). Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. *Proceedings of the national academy of sciences*, 109(24), 9320-9325.
5. Perrone, D., & Jasechko, S. (2019). Deeper well drilling an unsustainable stopgap to groundwater depletion. *Nature Sustainability*, 2(8), 773-782.
6. ARCGIS Web Application. (n.d.). Retrieved March 29, 2023, from <https://groundwater.gsa.state.al.us/>
7. Alaska Department of Natural Resources. (n.d.). Retrieved March 29, 2023, from <http://dnr.alaska.gov/>
8. GIS Data and Maps | Arizona Department of Water resources. (n.d.). Retrieved March 29, 2023, from <https://new.azwater.gov/gis>
9. Lower Mississippi-Gulf Water Science Center | U. S. Geological Survey. (n.d.). Retrieved March 29, 2023, from <https://www.usgs.gov/centers/lower-mississippi-gulf-water-science-center>
10. California Natural Resources Agency. (n.d.). Retrieved March 29, 2023, from <https://resources.ca.gov/>
11. GIS Data by Category | Colorado's Decision Support Systems. (n.d.). Retrieved March 29, 2023, from <https://cdss.colorado.gov/gis-data/gis-data-by-category>
12. Department of Energy & Environmental Protection. (n.d.). CT.Gov - Connecticut's Official State Website. Retrieved March 29, 2023, from <https://portal.ct.gov/DEEP>
13. Well Permits | Delaware Open Data Portal. (n.d.). Retrieved March 29, 2023, from <https://data.delaware.gov/Energy-and-Environment/Well-Permits/2655-qn8j>
14. Northwest Florida Water Management District. (n.d.). Retrieved March 29, 2023, from <https://nwfwmd.com/>
15. Well Construction Permits. (n.d.). Retrieved March 29, 2023, from <https://www.sfwmd.gov/doing-business-with-us/permits/well-construction>
16. Water Use Permit, Permitted Withdrawals. (n.d.). Retrieved March 29, 2023, from <https://data-sfwmd.opendata.arcgis.com/datasets/swfwmd::water-use-permit-permitted-withdrawals/explore?location=28.126469,-82.057900,8.89>.
17. Home of the St. Johns River Water Management District. (n.d.). SJRWMD. Retrieved March 29, 2023, from <https://www.sjrwmd.com/>
18. Suwannee River Water Management District | Official Website. (n.d.). Retrieved March 29, 2023, from <https://www.mysuwanneeriver.com/>

19. Georgia Department of Natural Resources Division. (n.d.). Retrieved March 29, 2023, from <https://gadnr.org/>
20. Commission on Water Resource Management. (n.d.). Commission on Water Resource Management. Retrieved March 29, 2023, from <https://dlnr.hawaii.gov/cwrm>
21. Water Right Pods. (n.d.). Retrieved March 29, 2023, from <https://data-idwr.hub.arcgis.com/datasets/IDWR::water-right-pods/explore?location=45.432773%2C-114.096317%2C6.20&showTable=true>
22. Illinois State Geological Survey Illinois Water Well (Ilwater) Interactive Map | isgs. (n.d.). Retrieved March 29, 2023, from <https://isgs.illinois.edu/ilwater>
23. Water. (2021, February 4). Significant Water Withdrawal Facility Data. Water. <https://www.in.gov/dnr/water/water-availability-use-rights/significant-water-withdrawal-facility-data/>
24. Iowa Department of Natural Resources. (n.d.). Retrieved March 29, 2023, from <https://www.iowadnr.gov>
25. Welcome to the Kansas Geological Survey. (n.d.). Retrieved March 29, 2023, from <https://www.kgs.ku.edu/>
26. Division of Water—Kentucky Energy and Environment Cabinet. (n.d.). Retrieved March 29, 2023, from <https://eec.ky.gov/Environmental-Protection/Water/Pages/default.aspx>
27. Louisiana Department of Natural Resources—Interactive Map. (n.d.). Retrieved March 29, 2023, from <http://sonris-www.dnr.state.la.us/gis/agsweb/IE/JSViewer/index.html?TemplateID=181>
28. Maine Well Database—Well Depth. (n.d.). Retrieved March 29, 2023, from <https://mgs-maine.opendata.arcgis.com/datasets/maine-well-database-well-depth/explore?location=45.166189%2C-68.997750%2C8.00>
29. Permitting. (n.d.). Department of the Environment. Retrieved March 29, 2023, from [https://mde.maryland.gov/programs/water/water\\_supply/Pages/default.aspx](https://mde.maryland.gov/programs/water/water_supply/Pages/default.aspx)
30. Energy & Environmental Affairs Data Portal. (n.d.). Retrieved March 29, 2023, from <https://eeaonline.eea.state.ma.us/portal#!/search/welldrilling>
31. EGLE Maps & Data. (n.d.). Retrieved March 29, 2023, from <https://gis-egle.hub.arcgis.com/search?q=water%20well>
32. Home. (n.d.). Minnesota Department of Natural Resources. Retrieved March 29, 2023, from <https://www.dnr.state.mn.us/>
33. *MDEQ – Mississippi Department of Environmental Quality*. (n.d.). Retrieved March 29, 2023, from <https://www.mdeq.ms.gov/>
34. *MSDIS*. (n.d.). Retrieved March 29, 2023, from <https://www.msdis.missouri.edu/index.html>
35. *Montana’s Ground-Water Information Center (GWIC) | please sign in | v. 11. 2023*. (n.d.). Retrieved March 29, 2023, from <http://mbmaggwic.mtech.edu/sqlserver/v11/help/go/signin.asp>
36. *Nebraskamap*. (n.d.). Retrieved March 29, 2023, from <https://www.nebraskamap.gov/>

37. (N.d.). Retrieved March 29, 2023, from <http://water.nv.gov/mapping.aspx?mapping=Well%20Drilling%20and%20Dam%20Data>
38. *Welcome*. (n.d.). NH Department of Environmental Services. Retrieved March 29, 2023, from <https://www.des.nh.gov/>
39. NJDEP. (n.d.). New Jersey Department of Environmental Protection. Retrieved March 29, 2023, from <https://dep.nj.gov>
40. *OSE PODS*. (n.d.). Retrieved March 29, 2023, from <https://geospatialdataose.opendata.arcgis.com/datasets/OSE::ose-pods/explore?filters=eyJ1c2Vfb2Zfd2VsIjpbIklSUiJdfQ==>
41. *New York State Department of Environmental Conservation*. (n.d.). Retrieved March 29, 2023, from <https://www.dec.ny.gov/>
42. NC Dept. of Environmental Quality. (n.d.). Retrieved March 29, 2023, from <https://deq.nc.gov/>
43. *NCDEQ-DWR: Central Coastal Plain Capacity Use Area (CCPCUA) Water Withdrawal Permit*. (n.d.). Retrieved March 29, 2023, from <https://www.ncwater.org/?page=49>
44. *ND Department of Water Resources Mapservice*. (n.d.). Retrieved March 29, 2023, from <https://mapservice.dwr.nd.gov/index.phtml?active=Drillers>
45. *Division of Geological Survey*. (n.d.). Retrieved March 29, 2023, from <https://ohiodnr.gov/discover-and-learn/safety-conservation/about-ODNR/geologic-survey>
46. *OWRB Groundwater—Data & Maps*. (n.d.). Retrieved March 29, 2023, from [https://www.owrb.ok.gov/maps/PMG/owrbdata\\_GW.html](https://www.owrb.ok.gov/maps/PMG/owrbdata_GW.html)
47. *OWRD GWIS*. (n.d.). Retrieved March 29, 2023, from [https://apps.wrd.state.or.us/apps/gw/gw\\_info/gw\\_info\\_report/Default.aspx](https://apps.wrd.state.or.us/apps/gw/gw_info/gw_info_report/Default.aspx)
48. *PA Groundwater Information System*. (n.d.). Pennsylvania Department of Conservation & Natural Resources. Retrieved March 29, 2023, from <https://www.dcnr.pa.gov:443/Conservation/Water/Groundwater/PAGroundwaterInformationSystem/Pages/default.aspx>
49. *Home: Department of Health*. (n.d.). Retrieved March 29, 2023, from <https://health.ri.gov/>
50. *ARCGIS Web Application*. (n.d.). Retrieved March 29, 2023, from <https://www.arcgis.com/apps/webappviewer/index.html?id=5052e9310cb242eaa5b07340b406ab5a>
51. *South Carolina Department of Natural Resources*. (n.d.). Retrieved March 29, 2023, from <https://www.dnr.sc.gov/>
52. *SD DENR Wells Completion Reports*. (n.d.). Retrieved March 29, 2023, from <https://apps.sd.gov/nr68welllogs/default>
53. *Water Wells*. (n.d.). Retrieved March 29, 2023, from <https://tdec-division-of-water-tdec.hub.arcgis.com/datasets/3b59f1b3cf38441297fd9252243c92950/explore?location=35.748283,-86.072809,7.00>
54. *Groundwater Data | Texas Water Development Board*. (n.d.). Retrieved March 29, 2023, from <https://www.twdb.texas.gov/groundwater/data/drillersdb.asp>
55. *Utah's State Geographic Information Database(AGRC)*. (n.d.). Retrieved March 29, 2023, from <https://opendata.gis.utah.gov/search?groupIds=4de5aefb5f6540348da4c633bd57cced>

56. *Welcome to DEC | Department of Environmental Conservation.* (n.d.). Retrieved March 29, 2023, from <https://dec.vermont.gov/>
57. *Virginia DEQ | home.* (n.d.). Retrieved March 29, 2023, from <https://www.deq.virginia.gov/>
58. ARCGIS Web Application. (n.d.). Retrieved March 29, 2023, from <https://groundwater.gsa.state.al.us/>
59. WV Department of Environmental Protection. (n.d.). Retrieved March 29, 2023, from <https://dep.wv.gov/Pages/default.aspx>
60. *Wisconsin Department of Natural Resources (DNR) | Wisconsin dnr.* (n.d.). Retrieved March 29, 2023, from <https://dnr.wisconsin.gov/>
61. *E-Permit.* (n.d.). Retrieved March 29, 2023, from <http://seoweb.wyo.gov/e-Permit/common/login.aspx?ReturnUrl=%2fe-Permit%2f>
62. Water-use terminology | u. S. Geological survey. (n.d.-a). Retrieved April 20, 2023, from <https://www.usgs.gov/mission-areas/water-resources/science/water-use-terminology>
63. GebreEgziabher, M., S. Jasechko, D. Perrone (2022). United States Aquifer Database, HydroShare, <http://www.hydroshare.org/resource/d2260651b51044d0b5cb2d293d21af08>
64. (N.d.). Retrieved April 10, 2023, from <https://www.arcgis.com/home/item.html?id=0f76175ca3a4424a9ce2328b1daf931a>
65. GebreEgziabher, M., Jasechko, S., & Perrone, D. (2022). Widespread and increased drilling of wells into fossil aquifers in the USA. *Nature Communications*, 13(1), 2129.
66. Bureau, U. C. (n.d.). Cartographic boundary files—Shapefile. Census.Gov. Retrieved April 10, 2023, from <https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html>