UNIVERSITY OF CALIFORNIA
Santa Barbara

Assessing Fire in Groundwater-Dependent Ecosystems

A Capstone Project submitted in partial satisfaction of the requirements for the degree of
Master of Environmental Data Science
for the
Bren School of Environmental Science & Management

by

Jillian Allison
Meagan Brown
Andre Dextre
Wade Sedgwick

Committee in charge:
Adjunct Professor Max Moritz
Naomi Tague
Ruth Oliver

June 2023

Assessing Wildfires in Groundwater Dependent Ecosystem (GDEs)

As developers of this Capstone Project documentation, we archive this documentation on the Bren School's website such that the results of our research are available for all to read. Our signatures on the document signify our joint responsibility to fulfill the archiving standards set by the Bren School of Environmental Science & Management.

_____

Jillian Allison

_____

Meagan Brown

_____

Andre Dextre

_____

Wade Sedgwick

The Capstone Project is required of all students in the Master of Environmental Data Science (MEDS) Program. The Project is a six-month-long activity in which small groups of students contribute to data science practices, products, or analyses that address a challenge or need related to a specific environmental issue.

This Capstone Project Technical Documentation is authored by Master of Environmental Data Science students and has been reviewed and approved by:

_____

Max Moritz

_____

Ruth Oliver

Table of Contents

1. Abstract

Climatic events such as drought and wildfires have become more frequent and severe in California, causing structural and ecological damage (Goss et al., 2020, Li and Banerjee, 2021). While there is an understanding of both freshwater resource management (Lubell et al., 2020) and wildfires (McLauchlan et al., 2020, Zigner et al., 2022, Hoover and Hanson, 2023), there has been little work directed towards assessing the relationship between groundwater-dependent ecosystems (GDEs), which are riparian or wetland ecosystems dependent on groundwater, and wildfires (Seker, 2017). To fill this knowledge gap, we compiled existing wildfire datasets to spatial wildfire layers of Mode Severity, Fire Count, Modeled Fire Threat, and Time Since Last Fire (TSLF) , to conduct exploratory statistical analyses to explore if there is a relationship between wildfires and GDEs in California.. These wildfire layers and exploratory analyses are incorporated into an interactive R-Shiny dashboard to serve as an exploratory tool that allows users to view spatial relationships between groundwater variables, such as GDEs, and wildfire metrics. Moreover, exploratory statistical analyses allow users to quantify the relationship between groundwater and wildfire metrics. These results will lay the foundations for future integration of groundwater data into wildfire risk models (Nolan et al., 2020).

2. Executive Summary

Extensive research has been done to develop wildfire models (Lubell et al., 2020, McLauchlan et al., 2020, Zigner et al., 2022, Hoover and Hanson, 2023) to protect human well-being and manage ecological resources. Currently, wildfire vulnerability and risk models are assessed via field measurements and remotely sensed metrics of above ground processes such as climate, vegetation, and soil moisture content. However it only accounts for a fraction of the processes driving wildfire risk and vulnerability. Subsurface processes, such as groundwater inputs, have not yet been incorporated into wildfire risk models. To address this knowledge gap, we designed an interactive R-Shiny dashboard that serves as an exploratory tool to assess, visualize, and quantify the relationship of GDEs and wildfires. The dashboard includes compiled and generated spatial wildfire layers, integrates GDE data, and enables download of data layers for further research. Additionally, the dashboard displays results of the statistical analysis, exploring the relationship between GDEs and fire frequency and severity. Through use of the dashboard, researchers and land managers are able to determine if investing in GDE maintenance can reduce the impacts of wildfire. By quantifying the impacts GDEs have on wildfire severity, TNC and other organizations can better plan for prescribed burns through implementation of strategic fuel removal for GDEs (Lambert et al., 2010). Additionally, through better understanding between the dynamics of wildfires and GDEs, TNC and fire agencies can improve on their wildfire community planning to mitigate wildfire damage through better resource allocations.. Looking ahead, we're hopeful that this interactive dashboard can serve as an initial step for further research in including groundwater inputs in wildfire risk models. This would improve accuracy of wildfire models, aiding TNC and other organizations in conducting prescribed burning and integrating community wildfire plans.

3. Problem Statement

California's groundwater levels are being depleted due to longer periods of drought and groundwater over-pumping (CalMatters, 2023), negatively affecting the health of groundwater-dependent ecosystems (Lubell et al., 2020). It has been shown that higher moisture levels and more water content in vegetation leads to lower flammability (Park et al., 2022). However, the overall relationship between these high-moisture-content GDEs and wildfire severity has not been analyzed. Wildfire risk models do not incorporate this into their analysis, which may limit the effectiveness of these models and hinder decision-making (FHSZ, 2023). Therefore, the objective of this project is to explore how wildfire metrics, such as fire count and fire severity, differ between GDEs and non-GDEs. We compiled and produced wildfire raster layers and conducted exploratory statistical analysis that can be used to observe and quantify this relationship. Although we are not predicting wildfire severity, exploring spatial and statistical relationships between GDEs and wildfires will be useful for fire management implications. For example, agencies such as TNC,) can better assess the role of GDEs in how they perform prescribed burns and improve on wildfire community planning to protect communities and the environment.

4. Objectives & Deliverables

This project is centered on two main objectives to produce one main deliverable.

Objectives:

1. <u>Compile and Produce Wildfire Raster Layers</u>:
   For our first objective we processed vector/polygon wildfire data for the entire state of California to create three raster layers at a 30x30m resolution.
   The compiled spatial layers include: (1) Burn Severity, (2) Fire Perimeters, and (3) Fire Threat.
   The produced spatial layers include: (1) Mode Severity, (2) Fire Count, and (3) Time Since Last Fire (TSLF)

2. <u>Explore the Relationship Between GDEs and Wildfire</u>:
   Using the raster layers from objective 1 and the Natural Communities Commonly Associated with Groundwater v.2 to determine GDEs, we conducted the following for exploratory statistical analysis:
   1) Mann Whitney U-Test: Allows us to determine if there is a difference between the distributions of burn severity values in GDEs vs non-GDEs.
   2) Bootstrap Hypothesis Test: Allows us to determine if there is a difference between the mean fire count in GDEs vs Non-GDEs

Deliverable:

1. Produce an Interactive Dashboard:
    Integrated raster layers and exploratory statistical analysis from objectives 1 and 2, into an interactive dashboard using R Shiny, allowing users to explore and interpret the data

5. Summary of Solution Design

5.1 Approach and Methods

First, utilizing the wildfire perimeter dataset (CAL FIRE), polygon boundaries were aggregated to count the number of overlaps in order to determine the number of fire occurrences within each raster cell from years 1950 - 2022. Additionally, a raster layer was created to indicate the most recent wildfire, or Time Since Last Fire (TSLF), which was calculated for each cell by subtracting the raster layer from the year of the data publication, 2022. Both rasters were created to encompass the entire state of California in a 30x30 meter resolution.

Next, GDE data obtained from TNC's Natural Communities Dataset, representing areas in California deemed most likely to be GDEs (Klausmeyer et al., 2018), was used to overlay over the fire frequency layer and the burn severity layer for statistical analysis. Acknowledging California's size and diverse eco-regions, we decided to divide our analysis of California into 13 ecoregions to examine the relationships between GDE and wildfire within each ecoregion.

For analyzing the fire count data, we used a difference in means bootstrapping approach. We first took 1000 samples of fire counts within and outside of GDEs with a 1 kilometer buffer distance for our analysis. The sample data was resampled 1000 times with replacement, and the difference in mean fire count in GDEs vs. non-GDEs (non-GDE mean minus GDE mean) was calculated for each resample. Then, we analyzed the distribution of the 1000 differences in means and determined if the estimated population difference in means was significantly different from 0 or not at a 95% confidence level. For example, if the difference in means was 0, that would mean that there was no difference in fire counts in GDEs when compared to non-GDEs in that ecoregion. For analyzing the burn severity data, we used a Mann-Whitney U Test to compare the distributions of burn severity values in GDEs and non-GDEs for each ecoregion. We first took 30 samples of burn severities within and outside of GDEs with a 1 kilometer buffer distance for our analysis. Then we performed a Mann-Whitney U Test to determine if the distribution of burn severity values was significantly different in GDEs and non-GDEs at a 95% confidence level. Each statistical analysis is accompanied by data visualization to help users interpret and visualize results of our analysis.

Lastly, an R Shiny dashboard hosts our spatial layers and results. These are displayed on an interactive map, where users can view individual layers, see statistics and scores for specific GDEs, and download our raster layers for use in their own research. A separate tab displays

documentation and visualizations of our statistical analysis. This product was tested using the Shiny Test package and by incorporating user testing and feedback.

5.2 Data Management

All spatial layers, scripts, documentation, and instructions will be posted on Dangermond's DataOne repository.

Layers described in Table 5.2.1 and statistical analysis described in Table 5.2.2 are on an interactive dashboard created with R shiny. Dashboard users can view individual layers and explore wildfire history and statistics for specific GDEs. Additionally, on a separate tab, users can download our layers for use in their own research and view documentation on our methods.

Table 5.2.1 Data Sources and Products

No.	Raw Data	Layer Produced from Data	Layer Description
1	California Fire Perimeters (all)	tslf.tif	Raster layer of time since last fire, 30 x 30 meter resolution, NAD83 / California Albers, EPSG:3310
2	California Fire Perimeters (all)	fire_count.tif	Raster layer of number of fire occurrences, 30 x 30 meter resolution, NAD83 / California Albers, EPSG:3310
3	Burn Severity Mosaics	burn_severity.tif	Raster layer of burn severity, 30 x 30 meter resolution, NAD83 / California Albers, EPSG:3310
4	Burn Severity Mosaics	mode_severity.tif	Raster layer of burn severity, 30 x 30 meter resolution, NAD83 / California Albers, EPSG:3310
5	Fire Threat	fire_threat.tif	Raster layer of fire threat, 30 x 30 meter resolution, NAD83 / California Albers, EPSG:3310

6	Natural Communities Commonly Associated with Groundwater_v2_0	gde_boundaries.tif	Raster layer of groundwater dependent ecosystems, 30 x 30 meter resolution, NAD83 / California Albers, EPSG:3310, wetland and vegetation layers aggregated together

Table 5.2.2 Summary of Statistical Analysis

No.	Statistical Method	Layers Used
1	Mann-Whitney U Test	Burn Severity and Mode Severity
2	Bootstrapping	Fire Perimeter, Fire Count, and TSLF

5.3 Software and Tools

The objectives for this capstone project were accomplished using the R programming language along with Rstudio. R and Rstudio are free, open source software applications that can be used by anyone. R and Rstudio allowed us to access and clean our data. We used the leaflet package (Joe Cheng et al., 2021), and the R Shiny package (Winston Chang et al., 2021) to develop an interactive web application for our client. We used additional packages such as: tidyverse (Wickham et al., 2019), raster (Robert J. Hijmans, 2021), fasterize (Noam Ross et al., 2022), janitor (Sam Firke et al., 2023), and sf (Pebesma, E., 2018). We also utilized GitHub for collaboration on coding and version control. Github is an internet hosting service that offers version control and source code management functionality of Git.

6. Products and Deliverables

Table 1. Capstone deliverables and applications

No.	Deliverable Name	Description
1	Time Since Last Fire Raster (1950 - 2022)	30 x 30 m resolution. Produced from wildfire perimeter dataset
2	Fire Count Raster (1950 - 2022)	30 x 30 m resolution. Produced from wildfire perimeter dataset
3	Fire Severity Mode	30 x 30 m resolution. Produced produce by USGS MBTS dataset
4	GDE Raster	30 x 30 m resolution. Produced from NC Dataset v2

5	R-shiny App Dashboard	Holds layers and analysis on an interactive dashboard
6	Reproducible Code	In R and uploaded on DataONE

6.1 Products for the Client

Products for the client fall under two general components: a Shiny app with spatial layers and statistical analyses and a user guide designed for users unfamiliar with Shiny app. These products are freely available on the Dangermond repository (https://search.dataone.org/portals/tnc_dangermond).

The user interface displays all resampled and produced data for this project. This includes raster layers showing fire count, time since last burn (TSLF), modeled fire threat, MTBS burn severity data. The user can interact with the map with these raster layers to view the spatial relationships between GDEs and wildfires metrics. Additionally, histograms and maps are displayed summarizing the results of our statistical analysis of fire frequency and burn severity. User documentation is available in the section below, demonstrating usage of the R Shiny dashboard.

7. Summary of Testing

An initial review of each dataset had ensured that missing values and outliers did not cause inaccurate results and findings. The "testthat" library package in R had been used to test the code in our script. The code for standardizing spatial data to a common grid had included a test function that checked input spatial data formatting, and the code for our statistical analysis of fire parameters and fire threat in groundwater-dependent ecosystems had tests checking for erroneous or duplicate values. To ensure the proper functionality of the dashboard, we had used the ShinyTest package, which performed 3 main tests - unit tests, server-function tests, and snapshot-based tests. In addition, several human testers had been employed to confirm that the app was working as intended. The statistical section, which showed basic groundwater and wildfire relationships, had been tested by running all available relationships. The spatial section, which included a map of California with spatial data relating to groundwater and wildfire, had all different spatial datasets overlaid to prevent the dashboard from crashing.

8. User Documentation

This user documentation aims to provide details on the organization of our GitHub repository, including R Markdowns with our spatial layers, analyses, functions, and R Shiny Application code. It also provides details on the use of the R Shiny Application and information on results of our analysis.

Organization of our Repository

1. The Aquafire capstone project repository has multiple folders that contain the Shiny App, as well as plots created for the statistical analysis and the raw data for the GDE shapefile and other wildfire metric layers. The organization of the repository is described below.

    a. Spatial Layers: An .Rmd is in this folder that provides the user the ability to create the spatial layers used in the analysis and in the R Shiny Application. It contains code to create the National Land Cover Database raster layer, the GDE raster layer, the Ecoregion raster layer, the Fire Count raster layer, the Time Since Last Fire (TSLF) raster layer, the Fire Threat raster layer, the Mode Burn Severity raster layer, and the GDE Polygon layer with attributes for each fire metric to be displayed on the R Shiny Application.

        i. To create each raster layer, we generally used the following workflow:

            1. Load the necessary libraries: rgdal, raster, sf, fasterize, tidyverse, janitor.

            2. Load the necessary spatial data: California counties, which were used as an outline for the shape of California; California Ecoregions, which will be used to divide our rasters for statistical analysis; and the relevant dataset, ex: CalFire Fire Perimeters.

            3. Create an empty raster to use as a template. We set the extent of this to match the extent of the California counties dataset, created a valueless raster with a resolution of 30 meters, and set the CRS to EPSG:3310.

            4. Convert the relevant spatial dataset to a raster layer using Fasterize::fasterize(), which is a faster way to rasterize polygon data. In this step, use the empty raster layer as a template so that the newly created raster layer matches the extent, CRS, and resolution. This ensures that all raster layers align perfectly for analysis and comparison.

                a. If the dataset was already in raster format, we reprojected and resampled the data to align with our template raster layer.

            5. Mask this new raster layer to the shape of California using the California counties spatial data.

    b. Analysis: Two .Rmd files are in this folder that provide the user the ability to conduct analyses on both number of fires and burn severity in each of the 13 ecoregions in California. The first .Rmd file, Fire_Count_Analysis.Rmd, provides the code to read in fire count data of each region, and will provide 13 histograms displaying how often GDEs burn compared to non-GDEs. The Fire_Severity_Analysis.Rmd file will also use the most-frequent burn severity for all fires since 1950 to compare the burn severity of GDEs and non-GDEs. The

two .Rmd files will produce histograms for each ecoregion and save them as .PNG files.

c. Functions: There are a number of functions in this folder that were used to wrangle the spatial and statistical data for visualization in the Shiny App.

   i.   Remove Agricultural and Urban Areas: remove_ag_urban.R

        1. The remove_ag_urban() function is used to remove agricultural, urban, and open water areas from an input raster layer based on National Land Cover Database (NLCD) land classifications. It specifically removes the following categories: Open Water, Developed Land, Pasture/Hay, and Cultivated Crops.
        This function outputs the input raster layer with these land use types removed.
        The input raster layer should have the same extent and CRS as the nlcd_rasterlayer stored in the dataone.org repository for this project. The nlcd_rasterlayer object can be loaded by running the first code chunk in the "Objects.Rmd" file.
        To check if the input raster layer is compatible with the NLCD raster layer, use the compareRaster() function.

   ii. Divide Raster into Ecoregions: raster_to_ecoregions.R

        1. The raster_to_ecoregions() function is used to divide an input raster layer into the 13 different level 3 ecoregions of California. Inputs include:
           a. raster_layer: Raster layer users would like to divide
           b. file_name: String. Base file or object name that the user would like to use. This will be appended to the ecoregion code in the following format: file_name_#.
           c. crs: CRS code of input raster layer, default is 3310.
           d. write_to_file: TRUE or FALSE. If TRUE, files called file_name_# will be stored in the raster_output folder. If FALSE, objects called file_name_# will be stored in the environment.

   iii. Take a Stratified Sample: stratified_sample.R

        1. The stratified_sample() function can be used to take a sample from an input raster layer with n number of points in GDEs and n number of points outside of GDEs. These points are taken with a 1 kilometer buffer. Inputs include:
           a. input_raster_layer: The raster layer the user would like to take a sample from.
           b. gde_raster_layer: The raster layer containing GDE data. This raster layer can be downloaded from the dataone.org repository for this project.

c. n: The number of samples the user would like to take from GDEs and from non-GDEs.

This function returns a dataframe with the cell number, GDE status (1 = GDE, 0 = non-GDE), value from the input raster layer, x coordinate, and y coordinate.

The input raster layer should have the same extent and CRS as the gde raster layer stored in the dataone.org repository for this project. The GDE raster layer object can be loaded by running the first code chunk in the "Objects.Rmd" file.

To check if the input raster layer is compatible with the GDE raster layer, use the compareRaster() function.

iv. Perform Burn Severity Sampling: burn_sev_sampling.R

1. The burn_sev_sampling() function is similar to the stratified_sample() function, but it has extra functionality to ensure that there are an equal number of points in GDEs and non-GDEs due to a low number of available points in the MTBS data. The burn_sev_sampling() function can be used to take a sample from an input raster layer with n number of points in GDEs and n number of points outside of GDEs. n should be sufficiently low that the function can get the desired number of samples. These points are taken with a 1 kilometer buffer. Inputs include:

a. input_raster_layer: The raster layer the user would like to take a sample from.

b. gde_raster_layer: The raster layer containing GDE data. This raster layer can be downloaded from the dataone.org repository for this project.

c. n: The number of samples the user would like to take from GDEs and from non-GDEs. The default value is 30.

The input raster layer should have the same extent and CRS as the gde raster layer stored in the dataone.org repository for this project. The GDE raster layer object can be loaded by running the first code chunk in the "Objects.Rmd" file.

To check if the input raster layer is compatible with the GDE raster layer, use the compareRaster() function.

v. Perform Mann-Whitney U Test: mann_whitney_test.R

1. The mann_whitney_test.R function accepts sample data (created using the burn_sev_sampling() function) and returns a list containing the confidence interval and p-value for that ecoregion. The Mann-Whitney U Test compares the distribution of burn severity values within and outside of GDEs in each ecoregion. A p-value below 0.05 indicates that the distributions are significantly

different. This function is used within the mann_whitney_result() function described below.

 vi. Aggregate Mann-Whitney U Test Results: mann_whitney_result.R
  1. The mann_whitney_result() function accepts a list of the sample datasets (one sample for each of the 13 ecoregions) and returns a data frame indicating which sample the results came from, the lower and upper bounds of the confidence interval, and the p-value.

 vii. Perform Difference in Means Bootstrapping: diff_in_means_bootstrap.R
  1. The diff_in_means_bootstrap() function accepts sample data (taken using the stratified_sample() function) and performs difference in means bootstrapping. This means that the sample data created above is resampled 1000 times with replacement, and the difference in mean fire count in GDEs vs. non-GDEs (non-GDE mean minus GDE mean) is calculated for each resample. Then, we analyze the distribution of the 1000 difference in means and determine if the estimated population difference in means is significantly different from 0 or not. For example: If the difference in means is 0, that would mean that there is no difference in fire counts in GDEs when compared to non-GDEs in that ecoregion. This function outputs the lower and upper confidence interval bounds and the p-value.

 viii. Aggregate Bootstrapping Results: bootstrap_results.R
  1. The bootstrap_results() function accepts a list of sample datasets (one sample for each of the 13 ecoregions) and returns a dataframe indicating which sample the results came from, the lower and upper bounds of the confidence interval, and the p-value.

 ix. Get Fire Severity Summary Statistics: fire_severity_stats.R
  1. The fire_severity_stats() function accepts a sample dataset (taken using the burn_sev_sampling() function) and outputs the following values for GDEs, non-GDEs, and all data points as a dataframe: maximum, minimum, and mode.

 x. Get Fire Count Summary Statistics: fire_count_stats.R
  1. The fire_count_stats() function accepts a sample dataset (taken using the stratified_sample() function) and outputs the following values for GDEs, non-GDEs, and all data points as a dataframe: mean, maximum, minimum, and non-zero mode.

 xi. Create a Histogram for Fire Count Data: fire_count_histogram.R
  1. The fire_count_histogram() function accepts sample data (taken using the stratified_sample() function) and outputs a properly

formatted histogram showing the distribution of fire counts in
GDEs and non-GDEs side by side.

      xii.   Create a Histogram for Burn Severity: burn_severity_histogram.R
1. The burn_severity_histogram() function accepts sample data
(taken using the burn_sev_sampling() function) and outputs a
properly formatted histogram showing the distribution of burn
severity values in GDEs and non-GDEs side by side.

      xiii.   Create a Violin Plot for Fire Count Data: fire_count_violin_plot.R
1. The fire_count_violin_plot() function accepts sample data (taken
using the stratified_sample() function) and outputs a properly
formatted violin plot showing the distribution of fire counts in
GDEs and non-GDEs.

d. Raster Outputs: The raw data for the spatial and statistical analyses are
contained in this folder. This includes the GDE polygon shapefile of California, as
well as the four wildfire metric raster layers: fire count, time (in years) since the
last burn, fire threat, and most-frequent burn severity. Additional data include
the 13 level 3 ecoregion shapefiles designated by the EPA, a California counties
shapefile and a fire perimeters shapefile that was used for the fire count and
time since the last burn datasets.

e. Shiny: Contains the main code for the creation of our Shiny App. They are
separated into 3 .R files and 2 folders.

      i.   ui: contains the code required for the interface of the Shiny App. It is split
into the 4 different tabs of the shiny app, the About page, the Ecoregion
page, the Map page, and the Statistics page.

      ii.   server: contains the main code that creates the reactivity of the Shiny
App. The main sections of this .R file are: the GDE tmap interactive map
that are linked to the respective ecosystem of each raster layer, the
ecoregion map, and the statistics histograms that are reactive based on
the selected ecoregion.

      iii.   global: contains the code that only needs to be run once for the Shiny
App to work. This includes the ecoregion data, land cover type data, as
well as for loops that will make a list of each dataset to easily utilize
reactivity. Due to the large size of the GDE shapefiles, the ecoregion
GDE shapefiles are simplified with st_simplify and smaller GDEs below a
certain threshold are removed. The thresholds can easily be changed in
this section. Finally, this file contains the data table that is displayed in
the About page about the data that was used, and code to use reactivity
with the statistics histograms.

      iv.   data: the data folder contains all necessary data to run the app, and each
dataset is split into the 13 California ecoregions. These include the GDE
shapefiles (gde_ecoregions), as well as the four wildfire metric raster
layers (fire_count, tslf, fire_threat, and burn_severity). The format of these

data are [ecoregion]_[wildfire_metric] for the wildfire layers and gde_[ecoregion] for the GDE shapefiles. The folder 'ca_eco_l3' contains the California level 3 ecoregions from the EPA, and the CA_Counties has a shapefile of the 58 counties of california, and is used to outline California in the main interactive map. Finally, the two data frames used to plot the fire count and burn severity are in this folder, and are named fire_count_shiny_histogram_df.txt and burnsev_shiny_histogram_df.txt, respectively.

v.   www: contains the main TNC logo, the GDE and wildfire images on the About page, and the 2 static images of california that show whether fire count or burn severity are statistically significant in an ecoregion.

How to use the Shiny App

1.   Upon opening the Shiny App, the About page is displayed that shows the background, significance of the app, a short description on how to use the app, metadata information, and a section about the creators. Users are able to access the data sources in the metadata tab.

Ecoregion Information
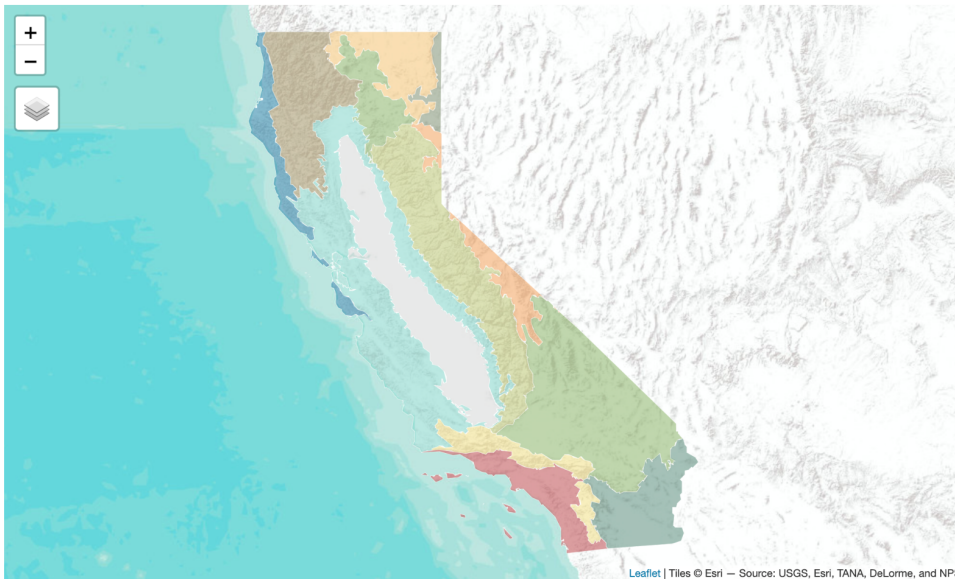
Cascades (4)

This mountainous ecoregion stretches from the central portion of western Washington, through the spine of Oregon, and includes a disjunct area in northern California. It is underlain by Cenozoic volcanics and much of the region has been affected by alpine glaciation. In Oregon and Washington, the western Cascades are older, lower, and dissected by numerous, steep-sided stream valleys. A high plateau occurs to the east, with both active and dormant volcanoes. Some peaks reach over 14,000 feet. Soils are mostly of cryic and frigid temperature regimes, with some mesic soils at low elevations and in the south. Andisols and Inceptisols are common.

a.   Ecoregion Information: This page helps the user decide which ecoregion they are interested in and want to explore in the Map page. The ecoregion key numbers are also included. Upon scrolling down, the user is able to read a short description of each ecoregion. This information contains the geology and geography of each ecoregion, as well as a description of the dominant vegetation within the different areas of the ecoregion.

b. Map: the main part of the Shiny App, the map page contains an outline of an ecoregion of California with GDEs displayed. To focus on the user's ecoregion of interest, select the ecoregion of interest based on the ecoregion map page. Two basemaps are available, ESRI world terrain and ESRI world street map, that can be changed with the layer icon below the zoom buttons.
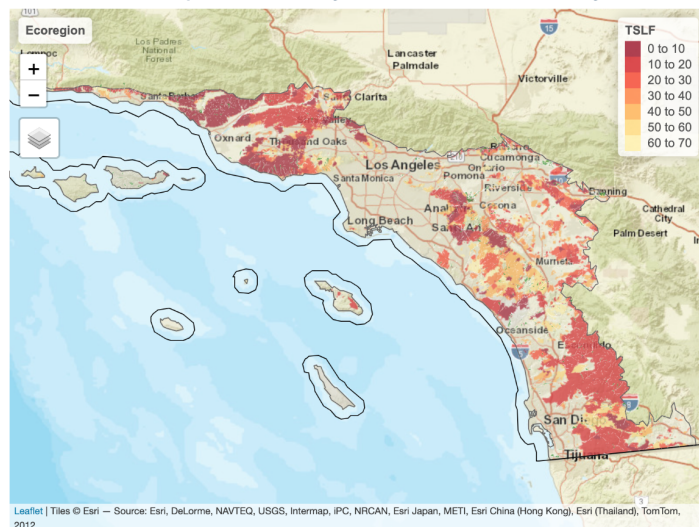
The four raster layers are displayed on the left panel: fire count, TSLF, fire threat and burn severity.
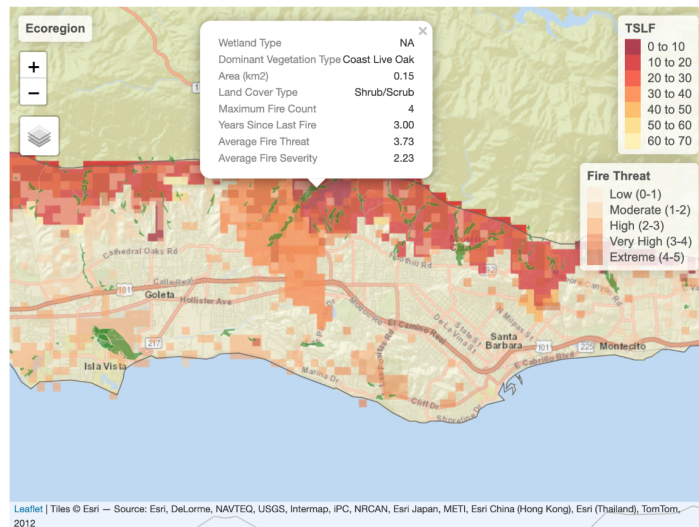
The map is interactive so the user can zoom in and select GDEs; upon clicking a GDE, a popup window will show wetland type or dominant vegetation type depending on whether it is a wetland or riparian GDE, area in square kilometers and land cover type are displayed. Additionally, four other statistics are shown: Maximum Fire Count, which is the maximum number of fires anywhere inside the GDE, Years Since Last Fire, which is the number of years since any part of the GDE burned, Average Fire Threat, which is the fire threat averaged over the entire GDE, and Average Fire Severity, which is the fire severity mode averaged throughout the GDE (fire severity is only available for fires that have occurred after 1984, the year when LandSat began recording this data). Note: Due to the large sizes of the shapefile and raster files, be sure to explore the ecoregion map to understand where and what wildfire metrics are relevant. It is also advised to first adjust the transparency of a wildfire layer before selecting the raster layer button so the raster layer does not need to be loaded twice. Additionally, if the user is viewing a smaller ecoregion and is opening a larger ecoregion, it is recommended to unselect the wildfire layers before switching to the larger ecosystem. Finally, when selecting a new raster layer, the zoom level of the map will reset to the entire ecoregion. Currently the app loads all raster files at a 30 meter resolution, but to quicken load times the user can edit the global.R file to include aggregating the larger raster layers using an if statement where the condition is being above a certain file size, and the function: fire_metric_layer[[i]] <- aggregate(fire_metric_layer[[i]], fact = 4, which is commented out but can easily be included. This will decrease resolution by a factor of 4 horizontally and 4 vertically, so will be reduced by a factor of 16.
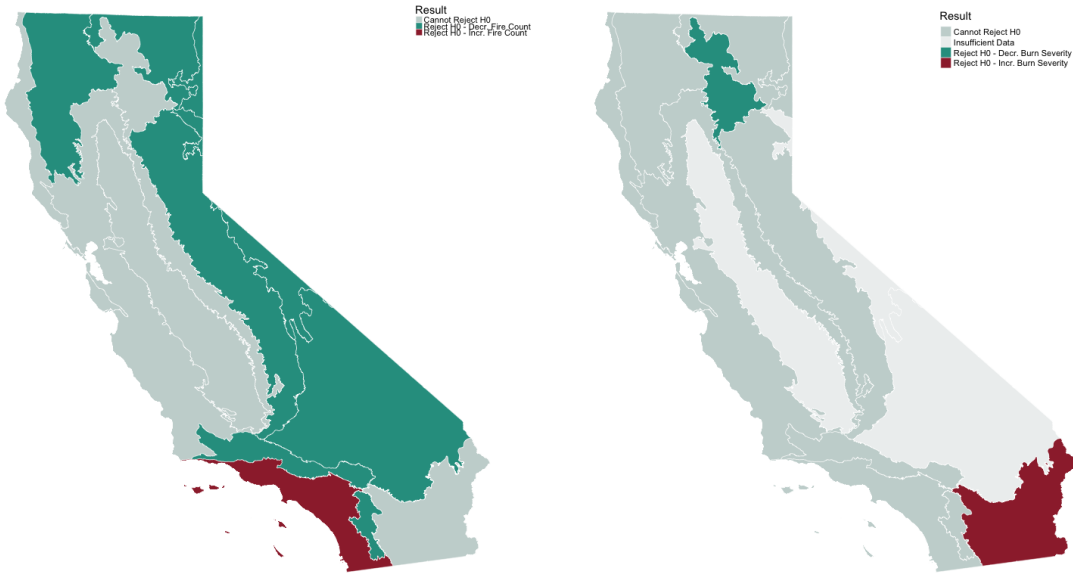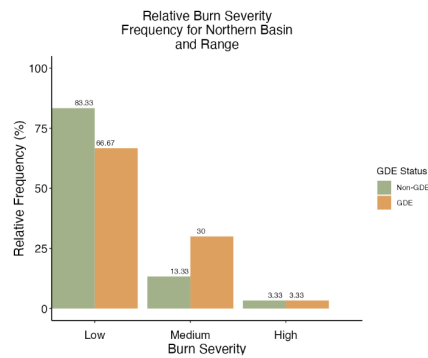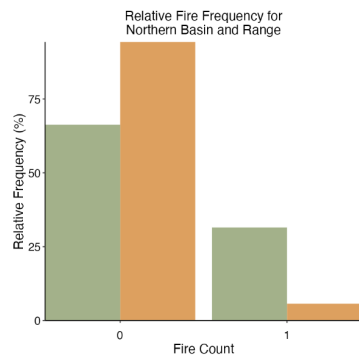
c. Statistics: The fire frequency histogram shown above displays the distribution of fire counts in the Northern Basin and Range ecoregion in GDEs (green) and non-GDEs (orange). The burn severity histogram shown above displays the distribution of burn severity values in the Northern Basin and Range ecoregion in GDEs and non-GDEs (these analyses used a sample of 1000 points from within and outside GDEs). Plots for all ecoregions are available on the R Shiny Application, along with a paragraph summarizing the statistical results for the impact of GDEs on fire frequency and burn severity. The Central Basin, Central Valley and Mojave Basin did not have enough GDE data to create meaningful results for the burn severity analysis.

9. Archive Access

All relevant data, including reproducible scripts, datasets, and the user guide,are available on the Dangermond DataOne Repository (https://search.dataone.org/portals/tnc_dangermond).

10. Acknowledgements

11. References

Bland, Alastair. "Ground Zero: Rain Brings Little Relief to California's Depleted Groundwater." *CalMatters*, 7 Feb. 2023, http://calmatters.org/environment/water/2023/02/california-depleted-groundwater-storms/.

Goss, Michael, et al. "Climate Change Is Increasing the Likelihood of Extreme Autumn Wildfire Conditions across California." *Environmental Research Letters*, vol. 15, no. 9, Aug. 2020, p. 094016, https://doi.org/10.1088/1748-9326/ab83a7.

Seker, Sandhya. "Groundwater May Play Key Role in Forest Fires." *Mongabay Environmental News*, 16 June 2017, https://news.mongabay.com/2017/06/groundwater-may-play-key-role-in-forest-fires/.

Hoover, Katie, and Laura Hanson. *Wildfire Statistics*. 66, Congressional Research Service, 1 Mar. 2023, https://crsreports.congress.gov/product/pdf/IF/IF10244.

Keeley, Jon, and Alexandra Syphard. "Climate Change and Future Fire Regimes: Examples from California." *Geosciences*, vol. 6, no. 3, Aug. 2016, p. 37, https://doi.org/10.3390/geosciences6030037.

Klausmeyer K. et al. Mapping Indicators of Groundwater-Dependent Ecosystems in California: Methods Report. 2018. https://groundwaterresourcehub.org/public/uploads/pdfs/iGDE_data_paper_20180423.pdf

Lambert, Adam, et al. "Invasive Species and Fire in California Ecosystems." *Fremontia*, vol. 38:2, Apr. 2010, https://rivrlab.msi.ucsb.edu/sites/rivrlab.msi.ucsb.edu/files/publications/frem38.2_38.3_lambert_etal.pdf

Li, Shu, and Tirtha Banerjee. "Spatial and Temporal Pattern of Wildfires in California from 2000 to 2019." *Scientific Reports*, vol. 11, no. 1, Apr. 2021, p. 8779, https://doi.org/10.1038/s41598-021-88131-9.

Lian, Xu, et al. "Summer Soil Drying Exacerbated by Earlier Spring Greening of Northern Vegetation." *Science Advances*, vol. 6, Jan. 2020.

Lubell, Mark, et al. "Sustainable Groundwater Management in California: A Grand Experiment in Environmental Governance." *Society & Natural Resources*, vol. 33, no. 12, Dec. 2020, pp. 1447–67, https://doi.org/10.1080/08941920.2020.1833617.

Masri, Shahir, et al. "Disproportionate Impacts of Wildfires among Elderly and Low-Income Communities in California from 2000–2020." *International Journal of Environmental Research and Public Health*, vol. 18, no. 8, Jan. 2021, p. 3921, https://doi.org/10.3390/ijerph18083921.

McLauchlan, Kendra K., et al. "Fire as a Fundamental Ecological Process: Research Advances and Frontiers." *Journal of Ecology*, edited by Giselda Durigan, vol. 108, no. 5, 2020, pp. 2047–69, https://doi.org/10.1111/1365-2745.13403.

Nolan, Rachael H., et al. "Linking Forest Flammability and Plant Vulnerability to Drought." *Forests*, vol. 11, no. 7, July 2020, p. 779, https://doi.org/10.3390/f11070779.

"Cal Fire: Office of the State Fire Marshal." *Fire Hazard Severity Zones*, 12 Mar. 2023, https://osfm.fire.ca.gov/divisions/community-wildfire-preparedness-and-mitigation/wildfire-preparedness/fire-hazard-severity-zones/.

Park, Isaac, et al. "Forecasting Live Fuel Moisture of Adenostema Fasciculatum and Its Relationship to Regional Wildfire Dynamics across Southern California Shrublands." *Fire*, vol. 5, no. 110, July 2022.

Rao, Krishna, et al. "Plant-Water Sensitivity Regulates Wildfire Vulnerability." *Nature Ecology & Evolution*, vol. 6, no. 3, Feb. 2022, pp. 332–39, https://doi.org/10.1038/s41559-021-01654-2.

Stephens, Scott L., et al. "Drought, Tree Mortality, and Wildfire in Forests Adapted to Frequent Fire." *BioScience*, vol. 68, no. 2, Feb. 2018, pp. 77–88, https://doi.org/10.1093/biosci/bix146.

Taufik, Muh, et al. "Amplification of Wildfire Area Burnt by Hydrological Drought in the Humid Tropics." *Nature Climate Change*, vol. 7, no. 6, 2017, pp. 428–31, https://doi.org/10.1038/nclimate3280.

Torres-García, M. Trinidad, et al. "The Potential of Groundwater-Dependent Ecosystems to Enhance Soil Biological Activity and Soil Fertility in Drylands." *Science of The Total Environment*, vol. 826, June 2022, p. 154111, https://doi.org/10.1016/j.scitotenv.2022.154111.

Warter, Maria, et al. "Drought Onset and Propagation into Soil Moisture and Grassland Vegetation Responses during the 2012–2019 Major Drought in Southern California." *Hydrology and Earth System Sciences*, vol. 25, July 2021, pp. 3713–29.

Williams, A. Park, et al. "Growing Impact of Wildfire on Western US Water Supply." *Proceedings of the National Academy of Sciences*, vol. 119, no. 10, Mar. 2022, p. e2114069119, https://doi.org/10.1073/pnas.2114069119.

Zigner, Katelyn, et al. "Wildfire Risk in the Complex Terrain of the Santa Barbara Wildland–Urban Interface during Extreme Winds." *Fire*, vol. 5, no. 138, Sept. 2022.