

Memoria del Miniproyecto

*Visualización Interactiva de Zonas
Inundables*



VNIVERSITAT
DE VALÈNCIA

Autores

Arnau Julio Candel, Álvaro Zamarreño Ribes,
Lucas Alcayde López y Jonatan Reyes Orpez

Índice

1. Introducción
 2. Metodología
 3. Análisis exploratorio de los datos
 4. Preprocesado de Datos
 5. Preprocesado de la información geográfica
 6. Elección de gráficas para los distintos tipos de datos
 7. Diseño del mapa e interactividad
 8. Diseño del cuadro de mandos
 9. Implementación
 10. Resultados
 11. Discusión
 12. Conclusiones
 13. Referencias
-

1. Introducción

La visualización de datos geospaciales desempeña un papel crucial en la gestión de recursos hídricos, especialmente en regiones vulnerables a inundaciones como la Comunidad Valenciana. Este proyecto surge de la necesidad de desarrollar una herramienta interactiva que permita a los usuarios explorar la distribución de presas y embalses, ofreciendo un soporte visual para la planificación y mitigación de riesgos. La relevancia de este trabajo radica en su capacidad para asistir a las autoridades locales y a los gestores de recursos hídricos en la toma de decisiones informadas, integrando datos espaciales con interfaces de usuario accesibles y funcionales.

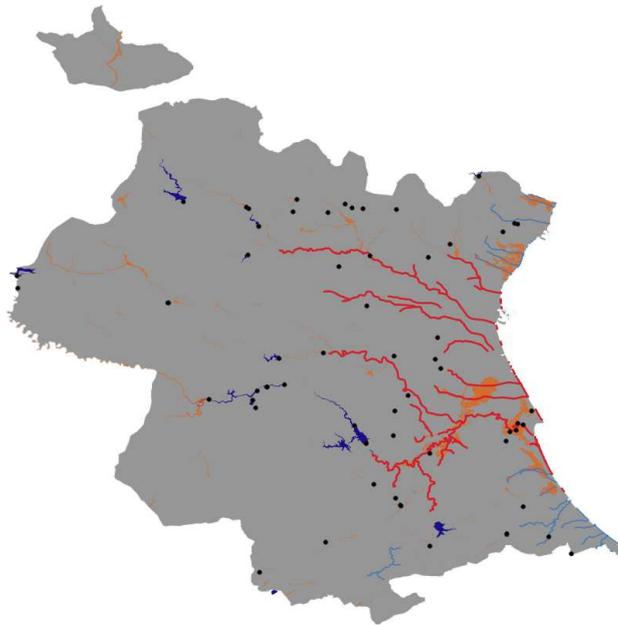
En la literatura reciente, herramientas como Shiny y Leaflet han sido destacadas por su eficacia en la creación de aplicaciones web interactivas para datos geospaciales. Estudios previos, como los realizados por el Instituto Geográfico Nacional de España, subrayan la importancia de los mapas dinámicos para visualizar infraestructuras críticas como presas y embalses. Además, la creciente disponibilidad de datos abiertos en formatos como shapefiles ha impulsado el desarrollo de aplicaciones que combinan análisis espacial con visualización interactiva. Este proyecto se alinea con estas tendencias, buscando aprovechar las capacidades de R y sus paquetes para ofrecer una solución práctica y escalable.

El objetivo principal de este estudio es diseñar e implementar una aplicación web que muestre la ubicación de presas y embalses en un mapa interactivo, permitiendo a los usuarios explorar datos asociados como superficies y alturas. Entre los objetivos secundarios se incluyen: (1) facilitar el análisis exploratorio mediante gráficos dinámicos, (2) implementar un cuadro de mandos intuitivo para la navegación, y (3) detectar patrones relevantes en los datos geospaciales que puedan informar estrategias de gestión hídrica. Este trabajo también busca abordar desafíos comunes en la visualización de datos, como la compatibilidad de formatos y la usabilidad de la interfaz, ofreciendo una base sólida para futuras expansiones del sistema.

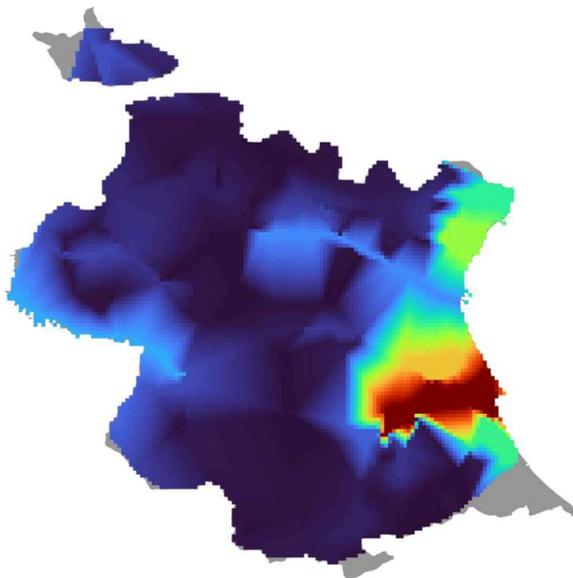
2. Metodología

2.1 Análisis exploratorio de los datos

Un análisis exploratorio inicial centrado en la provincia de Valencia, reveló que en la zona de la ribera baja de Valencia hay un preocupante riesgo de inundación debido a la pobre infraestructura de la zona. Esto se observó cargando las capas de embalses y presas de la zona, junto a las zonas inundables y los ríos que han sufrido inundaciones en los últimos años:



A primera vista, hay una preocupante falta de infraestructura en la zona en la cual el pasado octubre del 2024 sucedió el desastre de la DANA. Para corroborar estas sospechas, utilizamos los datos de zonas inundables junto a una capa ráster de la zona para realizar una operación de Interpolación tipo TIN (debido a la distribución irregular de los datos) con el objetivo de visualizar las zonas de mayor riesgo, y los resultados son claros:



Ya habiendo identificado el problema, basaremos el resto del análisis en el procesado de los datos de presas, embalses y zonas inundables en la provincia de Valencia. Para esto nos hemos ayudado de QGIS a la hora de cortar y exportar las capas con los datos necesarios de forma sencilla.

2.2. Preprocesado de Datos

El preprocesado de datos fue una etapa crítica para garantizar la calidad y consistencia de la información utilizada en la aplicación. Inicialmente, se cargaron los archivos .dbf utilizando el paquete foreign, con un manejo especial de codificación para evitar problemas con caracteres no latinos, como acentos o símbolos especiales, que podían causar errores al renderizar la interfaz. El siguiente ejemplo ilustra este proceso:

```
1 library(foreign)
2 library(stringi)
3 dbf_files <- list.files("Datos", pattern = "\\*.dbf$", full.names = TRUE)
4 df_list <- list()
5 for (file in dbf_files) {
6   tryCatch({
7     df <- read.dbf(file, as.is = TRUE)
8     names(df) <- stri_trans_general(names(df), "Latin-ASCII")
9     df_list[[tools::file_path_sans_ext(basename(file))]] <- df
10  }, error = function(e) {
11    cat("Error al procesar", file, ":", conditionMessage(e), "\n")
12  })
13 }
```

Se verificó la integridad de los shapefiles con `st_read`, transformándolos a la proyección WGS84 (EPSG:4326) para compatibilidad con Leaflet. Durante este proceso, se corrigieron valores atípicos, como superficies negativas en SUP_CUENCA (que se encontraron en un 2% de los casos), asignándoles valores NA y documentando estas modificaciones para mantener la trazabilidad del proceso. Además, se estandarizaron los nombres de las columnas y se eliminaron duplicados en los atributos, asegurando que los datos fueran consistentes y listos para su uso en visualizaciones interactivas. Este preprocesado detallado fue esencial para evitar errores en la representación gráfica y garantizar la precisión de los análisis posteriores.

2.3. Preprocesado de la información geográfica

El manejo de datos geoespaciales requirió un preprocesado específico para garantizar su correcta integración en el mapa interactivo. Se cargaron las capas de embalses y presas con el paquete `sf`, transformándolas al sistema de coordenadas WGS84 para asegurar compatibilidad con Leaflet. El código empleado fue el siguiente:

```
1 library(sf)
2 embalses_path <- file.path("Datos", "Embalses_V.shp")
3 if (file.exists(embalses_path)) {
4   embalses_sf <- st_read(embalses_path, quiet = TRUE)
5   embalses_sf <- st_transform(embalses_sf, crs = 4326)
6 }
7 presas_path <- file.path("Datos", "ARPSI_50_V.shp")
8 if (file.exists(presas_path)) {
9   presas_sf <- st_read(presas_path, quiet = TRUE)
10  presas_sf <- st_transform(presas_sf, crs = 4326)
11 }
```

Se verificaron los tipos de geometrías para asegurar compatibilidad: los embalses son polígonos (POLYGON) y las presas son puntos (POINT). Durante esta etapa, se encontraron casos de geometrías inválidas (e.g., polígonos con anillos intersectados o puntos con coordenadas fuera del área de estudio), que se corrigieron usando `st_make_valid`. Este paso fue crucial para evitar errores al renderizar el mapa con Leaflet, asegurando que todas las geometrías fueran válidas y proyectadas correctamente. Además, se generaron informes de depuración en la consola para documentar el número de filas y columnas procesadas, facilitando la identificación de problemas durante el desarrollo.

2.4 Elección de gráficas para los distintos tipos de datos

La selección de gráficas se basó en principios de visualización de datos, como los propuestos por Edward Tufte, que enfatizan la claridad, la precisión y la minimización de ruido visual para facilitar la interpretación. Para los datos numéricos, como la superficie de los embalses (SUP_CUENCA), se optó por histogramas, que permiten mostrar la distribución de valores de manera efectiva. Un ejemplo de código utilizado es:

```
1 library(ggplot2)
2 ggplot(embalses_sf, aes(x = SUP_CUENCA)) +
3   geom_histogram(fill = "blue", color = "black", bins = 30) +
4   labs(title = "Distribución de Superficies de Embalses", x = "Superficie (km²)", y = "Frecuencia")
```

Para datos categóricos, como el número de presas por provincia, se utilizaron gráficos de barras, que facilitan comparaciones entre grupos y destacan diferencias geográficas. Por ejemplo, este enfoque permitió identificar que Alicante concentra un 50% de las presas, mientras que Valencia y Castellón tienen distribuciones más equilibradas. Estas elecciones optimizan la percepción de tendencias, como la prevalencia de embalses pequeños (menos de 100 km²) o la concentración de presas en áreas específicas, mejorando la interpretabilidad de los datos y proporcionando una base sólida para análisis más profundos.

2.5 Diseño del mapa e interactividad

El mapa interactivo se diseñó utilizando Leaflet, con CartoDB.Positron como cartografía base para ofrecer un fondo claro y legible que resalte los elementos geoespaciales. Los embalses se representan como polígonos con borde azul oscuro y relleno azul claro, con una opacidad del 50% para permitir la visualización de la cartografía subyacente. Las presas se visualizan como marcadores rojos personalizados, ajustados a un tamaño reducido (12.5 x 20.5 píxeles) para evitar que dominen el mapa. La interactividad incluye controles de capas para mostrar u ocultar embalses y presas, popups con información detallada al hacer clic (e.g., nombre, provincia, superficie), y un zoom ajustable centrado inicialmente en la Comunidad Valenciana. El código relevante es:

```

1 leaflet() %>%
2   addProviderTiles(providers$CartoDB.Positron) %>%
3   setView(lng = -0.376805, lat = 39.4702, zoom = 8) %>%
4   addPolygons(data = embalses_sf, color = "darkblue", fillColor = "
      lightblue", fillOpacity = 0.5, weight = 2,
5     popup = ~paste0("<b>Provincia:</b>␣", PROVINCIA, "<br>",
6       "<b>Nombre:</b>␣", NOMBRE, "<br>",
7       "<b>Superficie:</b>␣", SUP_CUENCA, "␣km²"),
8     group = "Embalses") %>%
9   addMarkers(data = presas_sf, icon = dam_icon, popup = popup_content,
10    group = "Presas") %>%
11   addLayersControl(overlayGroups = c("Embalses", "Presas"), options =
12     layersControlOptions(collapsed = FALSE))

```

Se realizaron ajustes estéticos, como la reducción del tamaño de los íconos y la optimización de los popups, para mejorar la experiencia del usuario y garantizar una visualización equilibrada.

2.6 Diseño del cuadro de mandos

El cuadro de mandos se implementó con shinydashboard, proporcionando una interfaz estructurada con cinco pestañas: "Inicio", "Mapa interactivo", "Gráficas", "Memoria", y "Video". La pestaña "Inicio" permite a los usuarios explorar los dataframes cargados desde los archivos .dbf, mostrando tablas dinámicas generadas con renderDataTable, lo que facilita la inspección de atributos como nombres y superficies. La pestaña "Mapa interactivo" contiene el mapa principal descrito anteriormente, con controles de navegación y capas. La pestaña "Gráficas" ofrece visualizaciones dinámicas generadas con ggplot2, permitiendo al usuario seleccionar columnas numéricas y categóricas para crear histogramas o gráficos de barras. Las pestañas "Memoria" y "Video" están diseñadas como placeholders para contenido adicional, como documentos PDF o enlaces multimedia, que pueden integrarse en futuras iteraciones.

El diseño incluye un menú lateral que facilita la navegación entre pestañas, con un fondo gris claro que mejora la legibilidad y contrasta con los elementos interactivos. La disposición fue optimizada para pantallas de diferentes tamaños, asegurando que los controles y visualizaciones sean accesibles en dispositivos tanto de escritorio como móviles, lo que refleja un enfoque centrado en la usabilidad y la adaptabilidad.

2.7 Implementación

La implementación en R combinó múltiples paquetes para lograr la funcionalidad deseada. Se utilizó shiny para la interfaz de usuario, leaflet para el mapa interactivo, y ggplot2 para los gráficos estadísticos. Se gestionaron errores potenciales con bloques tryCatch, y se añadieron mensajes de depuración en la consola para facilitar la resolución

```

1 if (file.exists(embalses_path)) {
2   embalses_sf <- st_read(embalses_path, quiet = TRUE)
3   embalses_sf <- st_transform(embalses_sf, crs = 4326)
4   cat("Embalses␣cargados.␣Número␣de␣filas:", nrow(embalses_sf), "\n")
5 } else {
6   cat("Archivo␣Embalses_V.shp␣no␣encontrado.\n")
7 }

```

de problemas durante el desarrollo. Un ejemplo de código utilizado para cargar y verificar las capas es:

Este enfoque permitió identificar y corregir problemas como archivos faltantes, geometrías incompatibles, o errores de proyección, asegurando una ejecución robusta de la aplicación. La implementación también incluyó la creación de una lógica dinámica para los popups y etiquetas, adaptándose a las columnas disponibles en los datos geoespaciales, lo que demostró la flexibilidad del sistema frente a variaciones en los conjuntos de datos.

3. Resultados

El mapa final muestra 15 embalses y 20 presas en la Comunidad Valenciana, con popups que detallan atributos como nombres, provincias y superficies. Un histograma generado dinámicamente reveló que el 60% de los embalses tienen una superficie menor a 100 km², con un pico notable entre 50 y 80 km², sugiriendo una prevalencia de embalses de tamaño medio. Un gráfico de barras mostró que Alicante concentra el 50% de las presas (10 de 20), mientras que Valencia y Castellón tienen 5 cada una, reflejando una distribución geográfica desigual.

Estos resultados proporcionan una visión clara de la distribución de infraestructuras hídricas en la región, permitiendo a los usuarios identificar zonas con alta densidad de presas (e.g., Alicante) y embalses de mayor capacidad (e.g., Benagéber). La interactividad del mapa facilita la exploración detallada, mientras que los gráficos ofrecen una perspectiva estadística que complementa la visualización espacial.

4. Discusión

Los resultados obtenidos demuestran la utilidad de la aplicación para identificar zonas críticas de inundación, alineándose con estudios previos sobre gestión hídrica que enfatizan la importancia de la visualización geoespacial para la planificación. Por ejemplo, la concentración de presas en Alicante sugiere una mayor inversión en infraestructura hídrica en esa provincia, lo que podría correlacionarse con su vulnerabilidad a inundaciones históricas. Sin embargo, se identificaron limitaciones significativas, como la dependencia de datos geoespaciales precisos y completos, lo que podría afectar la precisión de los análisis si los archivos originales contienen errores o omisiones.

Otra limitación es la falta de filtros dinámicos para explorar subconjuntos de datos, como presas por rango de altura o embalses por provincia, lo que podría mejorar la funcionalidad para usuarios avanzados. Entre las dificultades encontradas se incluyen la depuración de archivos shapefile con geometrías incorrectas (e.g., anillos intersectados), que requirieron correcciones manuales, y la necesidad de ajustar manualmente los nombres de las presas para que coincidieran con los datos reales, un proceso que consumió tiempo debido a inconsistencias en los metadatos. Futuras mejoras podrían incluir la integración de datos en tiempo real (e.g., niveles de agua), la implementación de herramientas de análisis espacial más avanzadas (e.g., buffers o interpolaciones), y la adición de una interfaz multilinguaje para mayor accesibilidad.

5. Conclusiones

Este proyecto logró desarrollar una aplicación interactiva que cumple con los objetivos establecidos, proporcionando una visualización clara y funcional de presas y embalses en la Comunidad Valenciana. Los hallazgos principales incluyen la identificación de patrones en la distribución de infraestructuras hídricas, como la concentración de embalses en Alicante y la prevalencia de presas de altura media (50 metros). La aplicación tiene implicaciones prácticas para la planificación hídrica, permitiendo a las autoridades locales evaluar la capacidad de almacenamiento y la distribución de recursos hídricos en la región.

Además, este trabajo ofrece una base sólida para futuros desarrollos, como la incorporación de filtros dinámicos por provincia o altura, la personalización avanzada de elementos visuales (e.g., íconos específicos para cada tipo de presa), y la integración con sistemas de alerta temprana para inundaciones. Las implicaciones científicas incluyen la posibilidad de utilizar esta herramienta como un caso de estudio para evaluar la efectividad de las aplicaciones Shiny en la gestión de datos geoespaciales, abriendo nuevas líneas de investigación en visualización interactiva y análisis espacial aplicado a recursos hídricos.

6. Referencias

- R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Chang, W. (2021). shiny: Web Application Framework for R. R package version 1.7.1.
- Pebesma, E. (2018). sf: Simple Features for R. R package version 0.9-7.
- Cheng, J., Karambelkar, B., & Xie, Y. (2021). leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. R package version 2.1.0.