# Documento de Manejo y Gestión de Base de Datos

Proyecto ANILLO SACBAD

# Introducción

El repositorio de datos denominado ANILLO SACBAD ha sido creado en el marco del proyecto del mismo nombre. Este proyecto se centra en la recopilación de series temporales hidrometeorológicas desde 1990 hasta la actualidad en las cuencas de los ríos Petorca, Ligua y Quilimari. Las variables principales estudiadas son la precipitación, el caudal y los niveles de pozos. La información proviene de estaciones de monitoreo gestionadas por la Dirección General de Aguas (DGA) y la Dirección Meteorológica de Chile (DMC).

# Estructura y Parámetros Generales

## Metadatos

En el directorio raíz de la base de datos, se encuentra una carpeta de metadatos que contiene tres subcarpetas: caudal, pozos y pp. Cada subcarpeta almacena archivos que describen las estaciones de monitoreo, tanto en su totalidad como aquellas seleccionadas para el estudio SACBAD. La información incluida en estos archivos abarca el nombre de la estación, latitud, longitud, Nombre\_subcuenca (identificación asociada al proyecto SACBAD), y la columna de identificación única (Primary key), que corresponde al Codigo\_nacional.

El "código nacional" de una estación de monitoreo es un identificador único asignado a esa estación en particular. Este código es utilizado para diferenciar una estación de monitoreo de otra de manera única dentro del conjunto de datos. En el contexto del proyecto ANILLO SACBAD, el "código nacional" se utiliza como un campo que identifica de manera única cada estación de monitoreo de precipitación, caudal o niveles de pozos.

Cuando se menciona que el "código nacional" funciona como una "primary key" (clave primaria), significa que este campo se ha designado como la clave principal de la base de datos. La clave primaria es un concepto en bases de datos relacionales que se utiliza para identificar de manera única cada fila en una tabla. De esta forma se pueden utilizar las planillas de metadatos para poder filtrar las estaciones de monitoreo que fueron seleccionadas en el proyecto SACBAD.

## Series Hidrológicas

En el directorio raíz, se encuentra la carpeta de series hidrológicas, que a su vez, contiene subcarpetas para cada variable (precipitación, caudal y niveles piezométricos). Cada subcarpeta tiene sucesivamente dos subcarpetas adicionales: "crudo" y "procesado".

Precipitación: Dado que esta variable es registrada por la DGA (Dirección General de Aguas) y la DMC (Dirección Meteorológica de Chile), se presentan carpetas separadas para cada fuente y una conjunta para mostrar los datos unidos. Es importante destacar que las fuentes pueden tener distintas granularidades temporales (diaria o mensual).

En la carpeta de datos procesados, se encuentra el archivo "*pp\_mensual\_1990\_2023\_11\_rellenado.csv",* que contiene datos de precipitación con valores faltantes reemplazados por el promedio de la zona geográfica más pequeña disponible.

Caudal: Los datos de caudal se presentan desde 1990 hasta diciembre de 2023 en formato diario y resumidos por el promedio mensual en la carpeta "crudo". La carpeta de datos procesados contiene la serie mensual rellenada llamada "*q\_mensual\_1990\_2023\_12\_rellenado.csv*".

Niveles Piezométricos: La base de datos de niveles se utiliza mayoritariamente en formato mensual. Hay dos fuentes distintas: una directa de la DGA a través de API y otra obtenida mediante solicitud de transparencia. Los datos procesados incluyen el promedio del año hidrológico y el último dato mensual de cada año hidrológico.

### Diccionario de Datos

**Precipitación**

* Year: Año de medición.
* Month: Mes de medición.
* Codigo\_nacional: Identificación única de la estación de monitoreo.
* pp\_day: Acumulada diaria de preciptiación medida en milímetros.
* pp\_month: Acumulada mensual de precipitación medida en milímetros.
* Estacion: Nombre de la estación.
* Latitud: Coordenada geográfica de latitud.
* Longitud: Coordenada geográfica de longitud.
* Nombre\_Subcuenca: Nombre de la subcuenca asociada al proyecto SACBAD.
* Calidad: Proporción de datos faltantes en la data cruda. (La calidad de los datos de cada estación se calcula en función de la información cruda disponible para cada variable)
* ID: Sigla de identificación en español de la subcuenca.
* hydro\_year: Año hidrológico.
* ID\_subcuenca: Sigla de identificación en inglés de la subcuenca.
* Name\_subcuenca: Nombre de la subcuenca en inglés.

**Caudal**

* (Columnas análogas a las de la serie de tiempo de precipitación)
* q\_mean: Representa el caudal medido en metros cúbicos por segundo. Cuando se presenta agregado a escala mensual es el resultado de promediar los caudales medios diarios.

**Niveles Piezométricos**

* nivel\_pozo: Profundidad en metros del pozo desde la cota de nivel en que fue construido hasta el nivel del agua encontrado al momento de la medición.
* (Columnas análogas a las de la serie de tiempo de precipitación)

## Datos resumidos en formato SACBAD

Dentro de la carpeta de datos procesados, se encuentra una carpeta de resumen que contiene información resumida por año hidrológico y una planilla resumen en formato matriz, donde las filas representan las subcuencas de estudio y las columnas los años hidrológicos.

Este documento proporciona una visión general del manejo y estructura de la base de datos del proyecto ANILLO SACBAD. Se recomienda seguir las prácticas y convenciones establecidas para asegurar la integridad y coherencia de los datos almacenados.