Enhanced Hydrofabric Layers Feature Table

POD_Points

Feature	Description
WDID	Native ID of diversion reported from the source
Latitude	Latitude of POD reported from the source
Longitude	Longitude of POD reported from the source
type	1. Physical 2. Aggregated Diversion 3. Multi-structure
SiteName	Site name of POD reported from the source
BeneficialUseCategory	Designated use for the water right license associated with the POD reported in WaDE
Water Source	Name of the stream source of the POD reported from the source
Source GNIS ID	GNIS ID of the stream source of the POD derived from the “Water Source”
Source_comid	COMID of the closest segment in reference fabric with the same GNIS ID
POI_NativeID	Native ID of diversion reported from the source

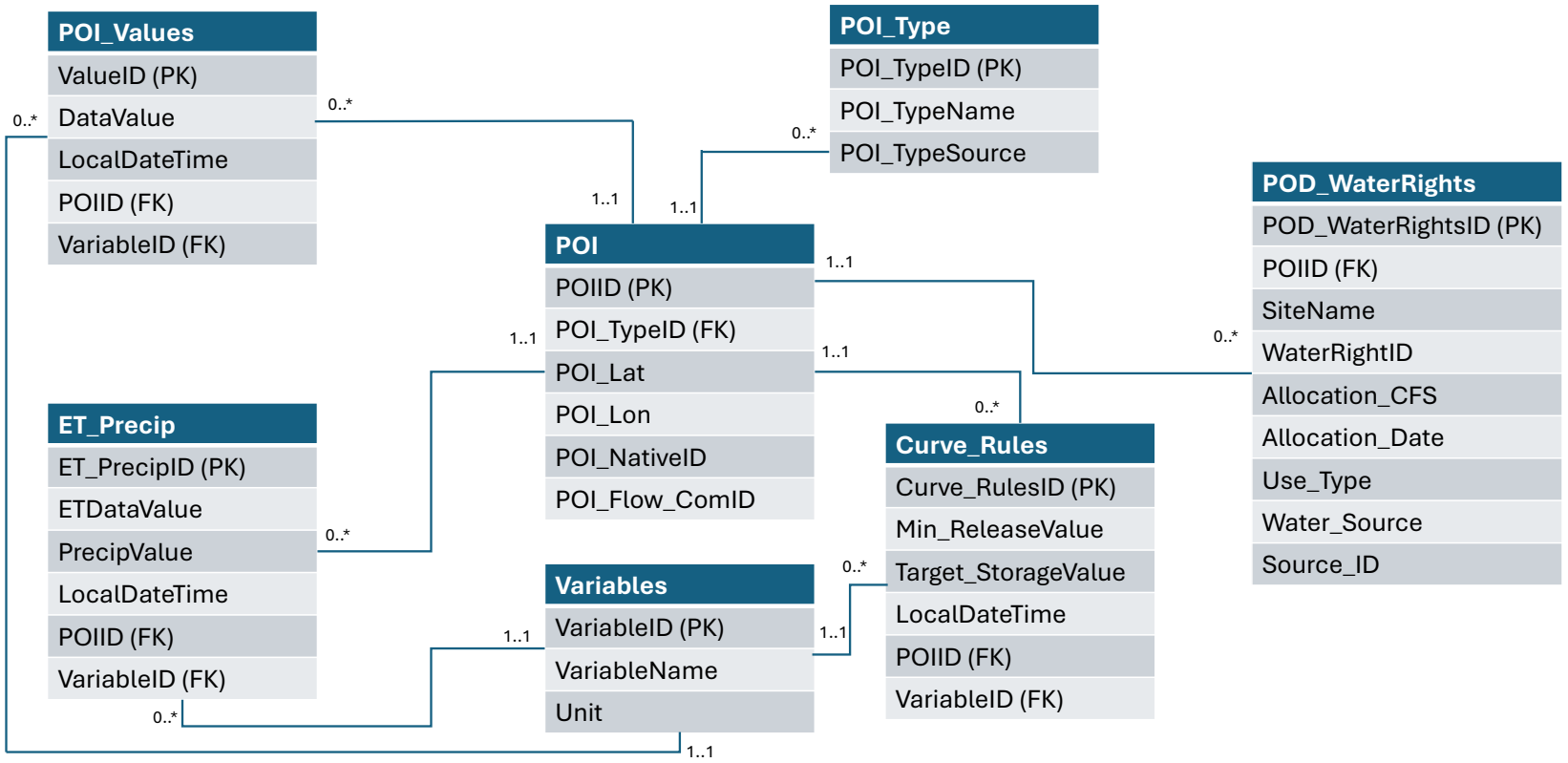
ResOps_points

Feature	Description
DAM_ID	Unique DAM_ID for each dam in ResOpsUS database
DAM_NAME	Name of the dam
State	State where the dam is located
AGENCY_CODE	Unique code for agency reporting data
LONG	The longitude of the dam's location
LAT	The latitude of the dam's location
TIME_SERIES_START	The initial starting date of the time series data for this dam in ResOpsUS database
TIME_SERIES_END	The final date of the time series data for this dam in ResOpsUS database
INCONSISTENCIES_NOTED	Any inconsistency in data reported in ResOpsUS database
NID ID	Native ID of the dam in National Inventory of Dams
RIVER	Name of impounded river
CAP_MCM	Maximum storage capacity of reservoir in million cubic meters
DAM_HGT_M	Height of dam in meters
AREA_SKM	Surface area of reservoir in square kilometers
MAIN_USE	Main purpose of reservoir: Irrigation; Hydroelectricity; Water supply; Flood control; Recreation; Navigation; Fisheries; Pollution control; Livestock; or Other
Source_comid	COMID of the closest segment in reference fabric
POI_NativeID	Native ID of dam reported from the source

Simplification of Hydrofabric

The hydrofabric simplification process involves filtering a reference flowline dataset to retain only the essential features necessary for water management modeling. The filtering operation iteratively removes flowline segments that do not contribute to downstream connectivity or are not linked to critical water management features such as PODs, reservoirs, gages, or NHM endpoints. This is achieved by checking whether each segment's `comid` appears in the `tocomid` (downstream connection) list or is associated with significant hydrological features. If a segment does not meet these criteria, it is removed in each iteration until no further deletions are necessary. Once the filtering is complete, the script creates a simplified version of the original geopackage, preserving all other datasets but replacing the `reference_flowline` layer with the reduced dataset. This ensures that the essential hydrological structure remains intact while eliminating unnecessary flowlines, reducing computational complexity, and improving the efficiency of the subsequent water management model.

Database Schema



Database Tables Documentation

This document contains an explanation of the tables created in the SQLite database along with their structures.

1. POI_Type Table

This table stores information about different types of Points of Interest (POIs).

Column	Data Type	Description
POI_TypeID	INTEGER (Primary Key)	A unique identifier for each type of POI.
POI_TypeName	TEXT	The name of the POI type.
POI_TypeSource	TEXT	The source or origin of the POI type data.

2. POI Table

This table stores individual Points of Interest (POIs), which represent geographical locations of interest.

Column	Data Type	Description
POIID	TEXT (Primary Key)	A unique identifier for each POI.
POI_TypeID	INTEGER (Foreign Key)	References POI_Type table for type of POI.
POI_Lat	REAL	The latitude of the POI.
POI_Lon	REAL	The longitude of the POI.
POI_NativeID	TEXT	An optional native identifier for the POI.
POI_Flow_ComID	INTEGER	NHD segment ID associated with the POI.

3. Variables Table

This table stores information about different variables related to the POIs

Column	Data Type	Description
VariableID	INTEGER (Primary Key)	A unique identifier for each variable.
VariableName	TEXT	The name of the variable.
Unit	TEXT	The unit of measurement for the variable.

4. POI_Values Table

This table stores the actual values of different variables measured or recorded at specific POIs at different times.

Column	Data Type	Description
ValueID	INTEGER (Primary Key)	A unique identifier for each value record.
DataValue	REAL	The actual recorded data value.
LocalDateTime	TEXT	The timestamp of the data recording.
POIID	INTEGER (Foreign Key)	References POI table for the POI where data was recorded.
VariableID	INTEGER (Foreign Key)	References Variables table for the variable measured.

5. ET_Precip Table

This table stores evapotranspiration (ET) and precipitation data values for reservoir POIs.

Column	Data Type	Description
ET_PrecipID	INTEGER (Primary Key)	A unique identifier for each record.
ETDataValue	REAL	The recorded evapotranspiration value.
PrecipValue	REAL	The recorded precipitation value.
LocalDateTime	TEXT	The timestamp of the data recording.
POIID	INTEGER (Foreign Key)	References POI table for the POI where data was recorded.
VariableID	INTEGER (Foreign Key)	References Variables table for the variable measured.

6. Curve_Rules Table

This table stores rules and parameters related to release and storage values for reservoir POIs.

Column	Data Type	Description
Curve_RulesID	INTEGER (Primary Key)	A unique identifier for each rule.
Min_ReleaseValue	REAL	The minimum release value from the reservoir or water body.
Target_StorageValue	REAL	The target storage value for the water body.
LocalDateTime	TEXT	The timestamp for the rule.
POIID	INTEGER (Foreign Key)	References POI table for the POI the rule applies to.
VariableID	INTEGER (Foreign Key)	References Variables table for the variable measured.

7. POD_WaterRights Table

This table stores information about water rights for Points of Diversion (POD) associated with specific POIs.

Column	Data Type	Description
POD_WaterRightsID	INTEGER (Primary Key)	A unique identifier for each water rights record.
POIID	TEXT (Foreign Key)	References POI table for the POI associated with the water right.
SiteName	TEXT	The name of the water right site.
WaterRightID	TEXT	The native ID for the water allocation right
Allocation_CFS	REAL	The water allocation in cubic feet per second (CFS).
Allocation_Date	TEXT	Appropriation date for the water right holder
Use_Type	TEXT	The type of water use.
Water_Source	TEXT	The source of the water.
Source_ID	REAL	An identifier for the water source.

Database Query Examples:

1. To retrieve time-series storage levels for a specific reservoir by POI_NativeID. This SQLite query example fetches the historical storage levels for a given reservoir.

```
SELECT POI_Values.LocalDateTime, POI_Values.DataValue
FROM POI_Values
JOIN POI ON POI_Values.POIID = POI.POIID
JOIN Variables ON POI_Values.VariableID = Variables.VariableID
WHERE POI.POI_NativeID = 'RESERVOIR_ID'
AND Variables.VariableName = 'storage'
ORDER BY POI_Values.LocalDateTime;
```

2. To retrieve time-series demands for a diversion by POI_NativeID. This SQLite query fetches the historical water demand for a specific diversion.

```
SELECT POI_Values.LocalDateTime, POI_Values.DataValue
FROM POI_Values
JOIN POI ON POI_Values.POIID = POI.POIID
JOIN Variables ON POI_Values.VariableID = Variables.VariableID
WHERE POI.POI_NativeID = 'DIVERSION_ID'
AND Variables.VariableName = 'demand'
ORDER BY POI_Values.LocalDateTime;
```

3. To retrieve Allocation_Date for a specific water right by POI_NativeID. This query fetches the date when a particular water right was allocated.

```
SELECT Allocation_Date
FROM POD_WaterRights
JOIN POI ON POD_WaterRights.POIID = POI.POIID
WHERE POI.POI_NativeID = 'WATER_RIGHT_ID';
```

NHM simulated output

`nhm_data_with_endpoints.shp` contains geospatial information about the NHM segments, including their spatial locations, flowline connectivity, and computed endpoint coordinates to attach them to the relevant segment. The NHM data is merged with the reference flowline layer (`clipped_enhanced_reference_14.gpkg`) based on the `comid` column. Endpoint coordinates (longitude, latitude) are extracted from flowline geometries and stored as `end_point_lon` and `end_point_lat`. A new shapefile `nhm_data_with_endpoints.shp` is created containing the NHM data along with extracted endpoint coordinates. The resulting dataset is also saved as a CSV file `nhm_data_with_endpoints.csv`.

Headwater segments that have a NHM output simulation are identified as those that do not have an upstream segment. Segments that are in middle of streams (non-headwaters) are also extracted to take into account the local contribution of watersheds within the area of study. Time-series data from NetCDF files (`seg_outflow.nc` and `seg_sroff.nc`) is extracted for the identified headwater and non-headwater segments. The final combined dataset is saved as `NHM_outflow_time_series.csv`. The table has time-series with relevant `comid` for that segment as different columns.