**Submission Documents:**

Pacific Northwest National Laboratory

**Abstract**:

This dataset includes an aggregated and event-correlated analysis of power outages in the United States, synthesized by integrating three data sources: the Environment for the Analysis of Geo-Located Energy Information (EAGLE-I), the Electric Emergency Incident Disturbance Report (DOE-417), and Annual Estimates of the Resident Population for Counties 2024 (CO-EST2024-POP). The EAGLE-I dataset, spanning from 2014 to 2023, encompasses over 146 million customers and offers county-level outage information at 15-minute intervals. The data has been processed, filtered, and aggregated to deliver an enhanced perspective on power outages, which are then correlated with DOE-417 data based on geographic location as well as the start and end times of events. For each major disturbance documented in DOE-417, essential metrics are defined to quantify the outages associated with the event. This dataset supports researchers in examining outages triggered by major disturbances like extreme weather and physical disruptions, thereby aiding studies on power system resilience.

Link to the raw data:

EAGL-E: <https://figshare.com/articles/dataset/The_Environment_for_Analysis_of_Geo-Located_Energy_Information_s_Recorded_Electricity_Outages_2014-2022/24237376>

DOE-417: [Electricity - Annual Disturbance Events Archive](https://www.eia.gov/electricity/data/disturbance/disturb_events_archive.html)

CO-EST2024-POP: <https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-total.html>

**Documentation & Instructions:**

The shared dataset is spitted into three groups, including grouped statistical outage data, merged outage dataset, and event-correlated (lagged) data. They are introduced as follows.

1. Aggregated Outage Data

The grouped outage data files summarize key metrics of power outages at multiple temporal and spatial resolutions, such as state, year, and month. These metrics help quantify the overall impact of power outages, such as the longest severe outage duration and the customer-weighted hours of disruption.

File Format: CSV

Filename: eaglei\_outages\_year\_group

Columns:

state: The state where the outage occurred.

year: The year when the outage occurred.

               month: The month when the outage occurred (with 0 representing yearly summary).

outage\_count: The total number of outages in the specified period.

              max\_outage\_duration: The maximum duration (in hours) of any single outage during the specified period.

              customer\_weighted\_hours: The total customer-weighted hours, calculated by multiplying the number of affected customers by the outage duration.

2. Merged Outage Data

The merged outage data offers a deeper view of individual county-level power outage scenario. This dataset is a result of processing the raw outage data by combining continuous outages into cohesive scenarios. It also calculates the minimum, maximum, and mean number of affected customers per outage.

File Format: CSV

Filename: eaglei\_outages\_year\_merged

Columns:

        fips: FIPS code identifying the county where the outage occurred.

        state: The state where the outage occurred.

        county: The county where the outage occurred.

        start\_time: The time when the outage started.

        duration: The duration of the outage in hours.

        min\_customers: The minimum number of affected customers during the outage.

        max\_customers: The maximum number of affected customers during the outage.

        mean\_customers: The average number of affected customers during the outage.

3. Event-Correlated Data

The event-correlated data links outage events with major disturbance incidents reported through the DOE-417. By matching geographic locations (state, county) and outage times with reported incidents, this dataset provides crucial insights into how major events affect the power grid. To ensure robust correlation, we incorporated both 8-hour and 24-hour lag periods between DOE-417 events and reported outages, and provided the results in separate CSV files.

File Format: CSV

Filename: eaglei\_outages\_with\_events\_year

    eaglei\_with\_events\_year\_8\_hour\_lag;

    eaglei\_with\_events\_year\_24\_hour\_lag.

Columns:

        fips: FIPS code representing the county where the outage occurred.

        state: The state where the outage occurred.

        county: The county where the outage occurred.

        start\_time: The start time of the outage.

       duration: The duration of the outage in hours.

       A unique identifier for the associated event (e.g., major storm)

       event\_type: The type of disturbance (e.g., hurricane) causing the outage.

Each of these files is stored in CSV format, ensuring compatibility with various analytical tools, and facilitating easy access and processing by researchers and industry professionals. The datasets provide detailed, high-resolution insights into power outages across the United States and offer essential data for studying power system resilience and response to major events.