CALIFORNIA ENERGY COMMISSION

Energy Research and Development Division

# Comprehensive Open Source Development of Next Generation Wildfire Models for Grid Resiliency:

## Weather Station Network Dataset

PYREGENCE™

NCAR | NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

UNIVERSITY OF SAN FRANCISCO
CHANGE THE WORLD FROM HERE

**Gavin Newsom, Governor**
**June 26, 2020**

**PREPARED BY**:

**Primary Author**:
Owen Doherty, PhD

Eagle Rock Analytics
Sacramento, CA 95820
631-766-7406
https://eaglerockanalytics.com

**Contract Number**: EPC-18-026

**PREPARED FOR**:
California Energy Commission

David Saah (Spatial Informatics Group, LLC)
**Principal Investigator**

Shane Romsos (Spatial Informatics Group, LLC)
**Project Manager**

Alex Horangic
**Commission Agreement Manager**

Laurie ten Hope
**Deputy Director**
**ENERGY RESEARCH AND DEVELOPMENT DIVISION**

Drew Bohan
**Executive Director**

**DISCLAIMER**

**This report was prepared as the result of work sponsored by the California Energy Commission. It does not necessarily represent the views of the Energy Commission, its employees or the State of California. The Energy Commission, the State of California, its employees, contractors and subcontractors make no warranty, express or implied, and assume no legal liability for the information in this report; nor does any party represent that the uses of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the California Energy Commission nor has the California Energy Commission passed upon the accuracy or adequacy of the information in this report.**

# TABLE OF CONTENTS

Page

# 1. Introduction

The purpose of the Weather Station Network Dataset is to establish a central document outlining the location of all geospatial data related to the optimization of the location of weather stations. The dataset supports the development of a routine to optimize placement of additional weather stations across the State of California, to best represent critical weather components with respect to California's electrical system. Our approach uses external, publicly available weather (and other bio-geophysical) data and derives additional material using models and algorithms, all of which the project (Pyregence) will make publicly available. For externally stored data, directions to where the data can be acquired is provided, and for material produced within the project, an online portal has been created and described herein.

We optimize weather station location by temporally aggregating relevant environmental data which are then passed to an optimization routine, from which a field that quantifies how similar or dissimilar any location is to regions where weather stations are currently located is produced. This approach ingests a large amount of geospatial data from a wide variety of sources, processes the data locally, and produces a smaller quantity of final data layers to be utilized by stakeholders. The final sources and exact variables of input data to be processed will be informed by and supplemented by layers of information from the Extreme Weather Analysis and convective-scale weather modeling research currently ongoing, along with key variables as identified by stakeholders and other subject matter experts. Below we outline our expected inputs, and the following chapters of this document explain where this data can be located.

This document describes the data and information currently being used to power the weather optimization approach. Through continual stakeholder engagement and scientific advances through the Pyregence Project our understanding of the desired outcomes for the optimization routine are continually evolving. Below we present an outline of the major potential sources of data that may be incorporated into the final product:

- Atmospheric data
    - reanalysis (i.e. ERA5)
    - climate projections (i.e. LOCA)
    - observational surface data
    - remote sensing products
- Biomass Characteristics
    - Data on fuels (LANDIS/LUCAS/LANDFIRE/California Forest Observatory)
    - Existing Vegetation Type
    - Non-burnable areas
- Abiotic Surface Characterization
    - Topography and Roughness

    - Non-Dynamic Physical characteristics (e.g. slope face, distance from ocean)
- Siting Informatics
    - Access (administrative boundaries and property rights)
    - Utility and public assets at risk (public data from Cal-Adapt and information shared by the utilities)
    - Existing weather/telecom/detection towers
    - Weather stations
    - Publicly available planned station installations
- Extreme Weather Analysis output layers
    - Layers of significant meteorological variables characterizing archetypical synoptic weather patterns associated with historical rapid-fire growth for each California fire weather zone.
- Convective-scale weather model output

Layers of analyzed CAWFE model output characterizing fine-scale weather variability pertinent to station siting associated with historical fire or weather events. The data archive will be updated over time as appropriate.

# 2. Data Archive: Inputs

A critical, necessary characteristic of all input data used to optimize the location of weather stations is high spatial resolution. The siting of electricity system infrastructure prefers information on the scale of hundreds of meters (with higher resolution preferred), as prediction information must match the scale of the infrastructure objects and components. The focus is therefore on identifying high resolution variables, with relevance for fire weather across California. Data products used to resolve large-scale, synoptic weather features in the cluster analysis may therefore be inappropriate for the weather station siting algorithm. The maximum output resolution of MaxEnt is a function of the predictor field with the lowest resolution. At present the limiting data source is Atmospheric Data, specifically wind data (2km resolution). Wind is critical in fire weather, while also being one of the most challenging variables to accurately measure and assess across large scales. We will continue to seek out higher resolution weather data to include, and the data archive will be updated over time as additional data comes online.

**Weather Variables.** Weather variables passed to the optimization routine will be identified from the cluster analysis and convective-scale weather modeling being performed by this task and the Historical Extreme Weather tasks in this Agreement. Additional variables may be included at Stakeholder suggestion or request, and from recommendation by subject matter experts. As input fields are temporally aggregated, we can utilize data from reanalysis products, models or remote sensing products.

| Dataset | Resolution | Key Variables | Provider |
|---|---|---|---|
| Daymet | 1km | Precip, Water Vapor, Air Temperature | Oak Ridge National Laboratory |
| DRI WRF Runs | 2km | Wind, Precip, Water Vapor, Air Temperature | DRI and SIG |
| PRISM | 4km | Precip, Water Vapor, Air Temperature | Oregon State University |
| ERA5-Land | 9km | Wind, Precip, Water Vapor, Air Temperature | European Centre for Medium-Range Forecasts |

Project-generated layers, where the appropriate state variables and field are to be generated during the research, are given placeholders below.

| Dataset | Resolution | Possible key variables | Input data source for the algorithms |
|---|---|---|---|
| Extreme weather cluster analysis | 30 km | Pressure, Air temperature, Wind at particular atmospheric pressure level, Humidity | ERA5 (ECMWF) |
| CAWFE convective-scale weather simulations | 123m – 370m | Near-surface wind speed, wind temporal variability | NCEP North American Mesoscale (NAM) forecast or analyses |

**Biomass & Abiotic (Physiographic) Characteristics.**

The focus of this product is weather, however biological, and physical characteristics of the soil must be included to ensure sited weather stations are positioned in regions where fire risks are high. Inclusion of biotic and physical predictor field are to be determined on a region-by-region and IOU-by-IOU basis; stakeholder engagement will inform final selection. For example, Liberty Utilities has indicated that the direction of slope is important in the high Sierras and Pacific Corp noted that a set of wind regimes is most critical for their operations

Bio and physiographic variables will be assessed for utility and relevance. Our initial list of such variables (presented below) will be supplemented through further outreach. These variables are divided into their sources, the Near Term Modeling workgroup and publicly available, peer-reviewed sources

Variables coming from Near Term Modeling Workgroup:

- Fuel Availability
- Vegetative Characteristic
- Urban / Non-Burnable Regions

Information on where and how to access these variables is available in the *Near-Term Fire Forecast Dataset* report.

From Publicly Available, Peer-Reviewed Sources (i.e. PRISM)

- Slope Direction (Aspect)
- Slope Steepness
- Diurnal Anisotropic Heating Index

- Elevation
- Surface Roughness
- Valley Depth
- Effective Distance from Ocean

**Siting Information.** Security and privacy concerns, along with federal regulations, often preclude IOUs sharing infrastructure information outside their organizations. Such organizations may be hesitant to utilize tools and applications on public servers, due to data privacy concerns. Our station optimization routine will be packaged so as to run locally on IOU secure servers, therefore ameliorating the need for high resolution asset information.

For our public facing web tools that will support this software repository, we will utilize publicly available GIS assets from CEC & CPUC repositories.

https://www.energy.ca.gov/data-reports/energy-maps-california

**Processed Input Layers.** Temporally aggregated input predictors will be made publicly available on the pyregence FTP site: pyregence@ftp.pyregence.org.

The MaxEnt process requires single timepoint data layers. Even at very high (> 1km) resolution, this data is small in storage size, allowing for it to be distributed along with the code repository.

# 3. Data Archive: Outputs

The weather optimization approach will produce three types of output: code repositories; intermediate (pre-processed) data products and web-based visualization tool.

**Code Repositories**. When weather stations are installed and added to the network, the optimal location for the next set of weather stations may change. This will require iteration of the MaxEnt code. For this reason, and the fact that a majority of IOUs are sensitive to making data publicly available, or running code on secure servers, the code repository must be made available to run locally on IOU servers. This code (along with supporting documentation and metadata) will be available at: https://github.com/pyregence.

**Pre-Processed Data Products**. Temporally aggregated input predictors will be made publicly available on the pyregence FTP site: pyregence@ftp.pyregence.org, and further disseminated alongside the code repository.

**Web-Based Visualization Tools**. A web-based visualization tool displaying the similarity of locations to regions that are characterized by the current array of weather stations in a domain will be made available on https://pyregence.org/near-term-forecast. This website will be periodically updated as weather stations are deployed. The tool, as open-source code, may be ported