California Energy Commission
DRAFT Report

LONG-TERM RISK PROJECTIONS DATASET

Project ID#: CEC EPC-18-026

June 26, 2020

Prepared for California Energy Commission
Prepared by UC Merced and Spatial Informatics Group

Primary Author(s):
Leroy Westerling and Jonathan Baldwin
University of California, Merced

Contract Number: EPC-18-026

Prepared for: California Energy Commission

David Saah, PhD
Principal Investigator

Shane Romsos
Project Manager

Alex Horangic
Commission Agreement Manager

Laurie ten Hope
Deputy Director - Energy Research and Development Division

Drew Bohan
Executive Director

Gavin Newsom
Governor, State of California

DISCLAIMER

INTRODUCTION

The purpose of the Long-term Fire Risk Projections Dataset (hereafter, "dataset") is to establish a repository for disseminating and documenting data for use in model development activities across multiple labs comprising the Long-term Risk Projections workgroup (Workgroup #4). Data used in model development include (but are not limited to):

- MTBS - for historical fire data between 1984 and 2017. Includes fire perimeters and estimated burn severity. At 30m grid, Albers projection.
- LANDFIRE – vegetation classification, elevation, aspect, slope. At 30m grid, Albers projection
- gridMET – gridded historical climatological data
- A variety of derived land surface characteristics datasets (such as: biomass, housing density, structure footprints, soil moisture, landcover type, forest type, protected areas, ecoregions)

Additional data are anticipated to be used in later iterations of model development and will be added as they become available or are finalized, including:

- Revised MTBS data - Westerling's lab has generated an experimental dataset recreating fire severity maps for all large fires in California since 1984, testing algorithms that provide a consistent approach to categorizing fire severity across the different vegetation types contained within the state. These data will be added to the dataset for dissemination to modeling teams after testing and formatting are complete.
- Reference climate history dataset including temperature, precipitation, wind speed and direction, and relative humidity hourly on a 3 km grid. Preliminary data sample is being developed by Scripps Institution of Oceanography but does not cover an adequate historical period to support our modeling efforts. Once this dataset or another like it is completed—covering an adequate time period and having been tested and validated—it will be added to the Workgroup #4 dataset.
- Downscaled climate projections will be formatted and included in the dataset as they become available.
- Finalized modeled historical biomass, historical and projected land use and development scenarios, fuels management scenarios.

Scripts have been developed to facilitate merging datasets from these diverse sources into common geospatial data formats and vector data formats for use in modeling efforts across all of the teams in the working group.

DATASET ACCESS

The data listed are all input data used for various models. Most of the data is available publicly through the listed website. Each website has its own usage and referencing requirements, along with metadata. See each one for more specific information about the data provided.

1. DATA DICTIONARY
   A. MTBS - Monitoring Trends in Burn Severity
      i. Root Directory: https://www.mtbs.gov/
      ii. Description: MTBS maps fire information for fires greater than a specified size (>1000 acres for the Western US). Fires are available for 1984-2017.
      This information includes (but is not limited to) metadata about each fire (geospatial as well as thresholds used to determine severity). There are also raster (GeoTiff) files of pre- and/or post-fire Landsat images, an NBR or dNBR image (depending on one or two scenes used), a burn severity map (values 0-6), and a mask image of the fire area with clouds, shadows, snow, etc. masked. The final product we use is the ESRI Shapefile of the fire perimeter. Each of these products uses a 30m pixel. For full product descriptions see https://www.mtbs.gov/product-descriptions.
      iii. Directory Structure: Each fire is designated by a unique ID. Once a selection of fires is chosen (on the MTBS website). The downloaded files are unzipped into
      <year>/fire_level_tar_files/<fireID>.zip
      Each <fireID>.zip file contains the products described above. Please see https://www.mtbs.gov/product-descriptions for a complete list of the contents and naming conventions of this zip file.
   B. Landfire
      i. Root Directory: https://www.landfire.gov
      ii. Description: The Landfire data contains many data products (https://www.landfire.gov/version_comparison.php). The topographic and existing vegetation type data are of use to this project.The topographic information includes elevation, slope, and aspect. All are downloaded as zip files. Each zip file contains (among other data) an ArcGIS raster pyramid of the specific topographic information. The existing vegetation type data is also downloaded as a zip file. This contains the metadata for the pixel values (as a lookup table in CSV format) as well as the ArcGIS raster pyramid for the vegetation type. Each pixel is assigned a specific value that can be looked up to see all the categories for that pixel (including fuels mapping value, lifeform, physiognomy, etc.). Please see https://www.landfire.gov/evt.php for a full listing of the data dictionary.
      iii. Directory Structure: Each product is downloaded as a zip file. Each zip contains metadata in directories called CSV_Data, General_Metadata, and Spatial_Metadata. The gridded raster pyramid can be found in the Grid directory.
   C. GridMET
      i. Root Directory: http://www.climatologylab.org/gridmet.html
      ii. Description: GridMET provides daily historical climate data for the continental US from 1979-present. The data is gridded across a 1/24-degree grid. There are two types of climate data included in this dataset: primary variables (min/max temperature, precipitation, wind velocity, etc.) and derived variables (reference evapotranspiration, 100- and 1000-hour dead fuel moisture, etc.)
      iii. Directory Structure: Each variable for a given year is stored in its own netCDF file. The naming convention is <var>_<year>.nc. The files are downloaded individually or using a batch generated wget script from the gridMET website.
   D. Land Surface Hydrology Research Group – Drought Monitors
      i. Root Directory: http://www.hydro.ucla.edu/SurfaceWaterGroup/data.php
      ii. Description: The drought monitoring system provides modeled soil moisture estimates (three levels) as well as precipitation and air temperature. The historical data is available for 1920-2015. The grid is 1/16 degree.

iii. Directory Structure: The data is stored by decade with one netCDF file per month called fluxes.<year>-<month>.nc. All variables are stored within the netCDF file for each pixel in the grid,

E. LUCAS Biomass
   i. Root Directory: currently only by request (not publicly available as this is preliminary data)
   ii. Description: The biomass data includes various categories (branch, stem, fine, etc.). It also has dead organic matter broken down into additional categories (aboveground/belowground very fast/fast/medium/slow, etc.). The data is stored as a raster (GeoTiff) with a 1km resolution covering the continental US.
   iii. Directory Structure: Each variable has its own GeoTiff file, named with the variable name.

F. Ecoregions
   i. Root Directory: https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states
   ii. Description: This data includes the level III and level IV ecoregions across the continental US. The data is in shapefiles and is available with and without state boundaries.
   iii. Directory Structure: The data is downloaded in a single zip file per level which contains the ESRI shapefile components. The metadata and symbology are downloaded separately.

G. Protected Areas Database
   i. Root Directory: https://www.usgs.gov/core-science-systems/science-analytics-and-synthesis/gap/science/protected-areas
   ii. Description: This is a listing of protected areas and public lands within the continental US. The data is available as shapefiles covering the continental US.
   iii. Directory Structure: The downloaded national data is a single zip file containing one shapefile set for each of the groups (easement, fee, etc.)

H. National Forest Type Dataset
   i. Root Directory: https://data.fs.usda.gov/geodata/rastergateway/forest_type/
   ii. Description: This dataset visualizes the distribution of forest types across the continental US in 28 groups. The data is downloaded as a single zip file containing the data and metadata. The raster data is at 250m per pixel across the continental US.
   iii. Directory Structure: The zip file contains raster data and metadata for the forest group.

I. National Land Cover Database
   i. Root Directory: http://www.mrlc.gov/
   ii. Description: This dataset describes the categorical land cover for the continental US. This data is available for various years from 2001 to 2016. The data is presented as raster data with a 30m pixel resolution.
   iii. Directory Structure: The downloaded zip file contains the raster data as well as metadata.

J. Housing Density
   i. Root Directory: http://silvis.forest.wisc.edu/data/wui-change/
   ii. Description: The housing density can be extracted from the Wildland Urban Interface data provided with this dataset. The data is available for the conterminous US as a geodatabase or by state as a geodatabase or shapefile.

iii. Directory Structure: Each downloaded zip file contains the requested geodatabase or shapefile along with associated metadata.

K. Building Footprints
   i. Root Directory: https://github.com/Microsoft/USBuildingFootprints
   ii. Description: The GIS data layer of building/structure footprints for California. The data is available for the conterminous US as a geodatabase or by state as a geodatabase or shapefile.
   iii. Directory Structure: Each downloaded zip file contains the requested geodatabase or shapefile along with associated metadata.

L. Electricity Assets – Transmissions Lines
   i. Root Directory: https://cecgis-caenergy.opendata.arcgis.com/datasets/260b4513acdb4a3a8e4d64e69fc84fee_0/data
   ii. Description: The California Energy Commission (CEC) Electric Transmission Line geospatial data layer has been created to display the electric transmission grid in California. When used in association with the CEC Power Plant and CEC Electric Substation geospatial data layers, viewers can analyze the geographic relationships with the electric transmission grids across utilities, counties and state.
   iii. Directory Structure: data can be downloaded as shapefile, kml or spreadsheet or accessed via API (https://services3.arcgis.com/bWPjFyq029ChCGur/arcgis/rest/services/Transmission_Line/FeatureServer/0/query?where=1%3D1&outFields=*&outSR=4326&f=json).

M. Electricity Assets – Substations
   i. Root Directory: https://cecgis-caenergy.opendata.arcgis.com/datasets/california-electric-substation/data
   ii. Description: The California Energy Commission (CEC) Electric Substation geospatial point data layer has been created to display the locations of power substations in California.
   iii. Directory Structure: data can be downloaded as shapefile, kml or spreadsheet or accessed via API (https://services3.arcgis.com/bWPjFyq029ChCGur/arcgis/rest/services/Substation/FeatureServer/0/query?outFields=*&where=1%3D1)