DxRisk team internal data synthesis report: Phase 1 Milestone 2 deliverable

Convergence Data Analytics, Salo Sciences, Presence Product Group

Introduction

The purpose of Phase 1 of the DxRisk project is to build a predictive model, trained on PG&E internal data and external environmental data, that can assign wildfire ignitions risk scores to the Dx Grid. Phase 1 is focused on vegetation-caused outages and has been divided into 3 milestones:

Phase 1:

> **Milestone 1:** Working "backstop" model based on publicly and commercially available data (completed)
>
> **Milestone 2:** Establishing access to, performing exploratory analysis of, selecting for use in future modeling, and normalizing internal PG&E data sets relevant to modeling vegetation-caused outages and ignitions on the Dx grid (this document)
>
> **Milestone 3:** Building upon the methods of the backstop model and the information and data gained through Milestone 2 efforts to incorporate PG&E internal data and produce our official Dx grid ignition probabilities and risk scores, including modeling in support of ranking feeders by mitigation priority (upcoming)

Phase 2 of this project will focus on deploying the data and modeling pipelines developed in Phase 1 into PG&E's internal cloud IT infrastructure (as hosted by and in partnership with ARAD) and on broadening and sharpening our modeling capabilities, guided by real-world internal stakeholder needs and applications.

CDA, Presence, and Salo have defined Milestone 2 deliverables as this document which summarizes progress to date on accessing and understanding PG&E internal data, contains a collection of tables and figures that encapsulate our findings, and poses a set of deeper clarifying questions for internal subject matter experts (SMEs); as well as the delivery of the working code used to produce all of the above, suitable for reproducing and extending our exploratory work.

Most specifically, the primary objectives of this document are:

1. To synthesize the information contained in different project-relevant datasets;
2. To detail linkages (both implicit and explicit) that can be drawn between datasets;
3. To describe the relevance of each dataset in modeling wildfire risk;
4. To provide a consistent nomenclature (or field naming convention) for synthesizing information across datasets.

The datasets covered in this document include data collected for to track outages and ignitions (e.g., ignitions data, Wires Down data, ILIS, Vegetation caused outage data, SAP notifications), for tracking work that has been or needs to be done (e.g., SAP, vegetation management database), and for tracking grid assets (ED-GIS). Data is collected for a range of purposes and often with specific questions in mind. The circumstances of collection and intended use influence the nature of the information found within each dataset and the consistency and quality of the information reported.

Table of contents

Preface

Vegetation-caused wildfire ignitions are a function of asset type, location, and health in addition to the changing environmental conditions around them. A core component of the DxRisk model is therefore to determine the geographic distribution of grid assets and to couple them with spatial environmental data to ascertain which grid assets are likely to fail and/or cause an ignition. Predictive features could therefore include weather and climate data, asset characteristics and inspections, or nearby vegetation structure. Our data synthesis to-date has focused on datasets relevant to characterizing ignitions and outages, the location, type, and state of repair of equipment, and the spatial distribution of environmental conditions that affect the grid.

The DxRisk modeling team participated in an extensive requirements and information gathering effort with subject matter experts and data owners distributed throughout the company. The data sets can be loosely grouped into three categories: data about events, assets, and the environment. The sections that follow detail the DxRisk modeling team's understanding of the contents of each data set provided as a result of requests made during that process.

Event data

We intend to model wildfire ignitions as the downstream result of processes that produce asset failures and outages and are therefore interested in all data sets that contain information that helps to characterize when and where outages occur and the characteristics of the subset of outages that cause ignitions. For our current focus on vegetation-caused outages, the data sets of greatest interest are ignitions, ILIS outages, wires-down, and vegetation failure data sets.

Ignitions

Description	Records of grid-caused fires that happened between 2014 and 2019.
Temporal coverage	2014-2019[1]
Record count	● 2233 (before 2018) ● 2639 (after 2018)
Link to schema	None
Cross-references	ILIS, OIS
Key contents	● Location of fire ● Date fire occurred ● Results of investigation into what caused each ignition

Summary

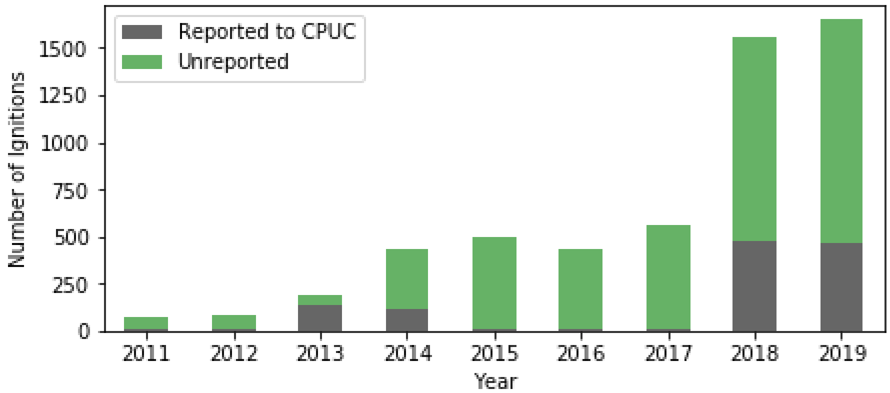
Records of ignitions that occurred in the service territory between 2013 and 2019. Figure 1 shows the number of ignitions reported each year. The severity of the 2018 wildfire season led to more fires and more comprehensive efforts to document the fires that did occur. We see an increase in the number of fires in that year among fires that met CPUC reporting requirements[2], and among fires that did not.

[1] Note: There is technically data prior to 2014, but we've been advised to expect that it is incomplete and most likely of little value to this project. We see a drastic increase in the number of ignitions recorded in 2018 -- including both reported and unreported incidents.

[2] Fires are subject to CPUC reporting if they:

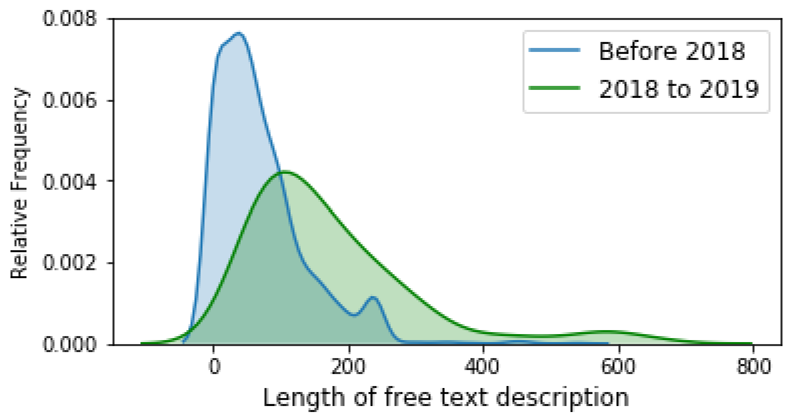
- Are self propagating
- Travel more than 1 meter from the ignition point
- Are known to the utility

Source: cpuc.ca.gov/fireincientsdata/



Beginning in 2018, the data also contain more detailed information. Notable changes include:

- More nuanced options for describing the circumstances that gave rise to each fire
- Flag indicating whether or not PG&E equipment was involved in ignition
- More extensive descriptions (in the form of free text entry) of the fire and the conditions that gave rise to ignition. Figure 2 shows the distribution of how many characters were recorded in text entry fields in ignitions data recorded before and after 2018.[3]



How the Data are Useful

- Record of when, where, and why ignitions occurred
- Provides a link between time/location of ignition and:
    - Weather data (via timestamp/location)
    - Outage data (via foreign key to ILIS)
    - Asset characteristics (via link between ILIS and asset data)
    - Inspections (via link between ILIS and asset data)
- Fields describing the circumstances that led to ignition (e.g., vegetation contact, animal contact, etc.) can help us to differentiate ignitions based on the most relevant drivers. For example, Milestone 1 results indicate that covariates such as tree height and wind speed may be strong predictors of ignitions related to vegetation contact but not animal contact. Seasonal patterns in vegetation- vs. animal-caused ignitions may also differ. Training a model on only the subset of

3

ignitions that are (or could be) related to vegetation allows us to make modeling decisions that could improve prediction accuracy.

Notable Caveats

- Pole fire is a catch all description (due to how data are collected in the field)
- Hard to link lat/lon coordinates to assets. 70% of locations are suspected accurate, others may be only within +/- 40 miles.

Questions for SMEs

- How would you go about drawing spatial links between ignitions and the grid components that failed, e.g., in EDGIS?
- In the field called “Log”, there are usually ILIS ids, but sometimes there are ids of other forms that look like: FR-2014-047, SA-2014-009, TM-2014-051, etc. Where do these other Log entries come from and can we use them to track down ILIS entries or are these ignitions without associated outages and those IDs refer to some other data set?
- ILIS tracks two separate ids. The first is the identifier labeled “Log” in this data set, but there can actually be multiple ILIS entries that share the same “Log” id. In ILIS, the more specific identifier is the

ILIS

Description	Documents the occurrence of power outages.
Temporal coverage	2007-2020
Record count	784,234 (about 10% are during storms)
Link to schema	ILIS Data Dictionary
Cross-references	SAP, ignitions, wire-down database, OIS, ED-GIS
Key contents	● Automated record generated from SCADA system describing the time and location of outages that occurred ● Records include: ○ Fault type ○ Outage cause ○ Equipment affected ○ Inputs needed to compute reliability metrics ● Location describes the equipment that triggered a notification in SCADA (e.g., substation, network, or distribution xformer).

Description

ILIS is the system of record for outages. It pulls information logged from SCADA systems in OIS into a cleaner and more interpretable format. Digesting data from OIS into ILIS involves cleansing tasks such as removing redundant reports from different devices on the system, tracking work orders involved in doing repairs, and calculating characteristics of the outage -- e.g., the duration, number of customers affected, whether momentary or sustained -- needed to calculate different reliability metrics.

ILIS also includes records of what caused the outage, based on observations made by field personnel dispatched to investigate each outage and do requisite repairs. This level of detail will not be present for self-clearing faults where no crew is dispatched to investigate what triggered an event. It is also worth noting that the causal information is collected from the perspective of the people doing a repair. A record indicating that an outage was "caused" by equipment failure will not necessarily indicate why the equipment failed. Though vegetation contact is a common cause of equipment failure, the data may not consistently report if and where this is the case.

Expected insights

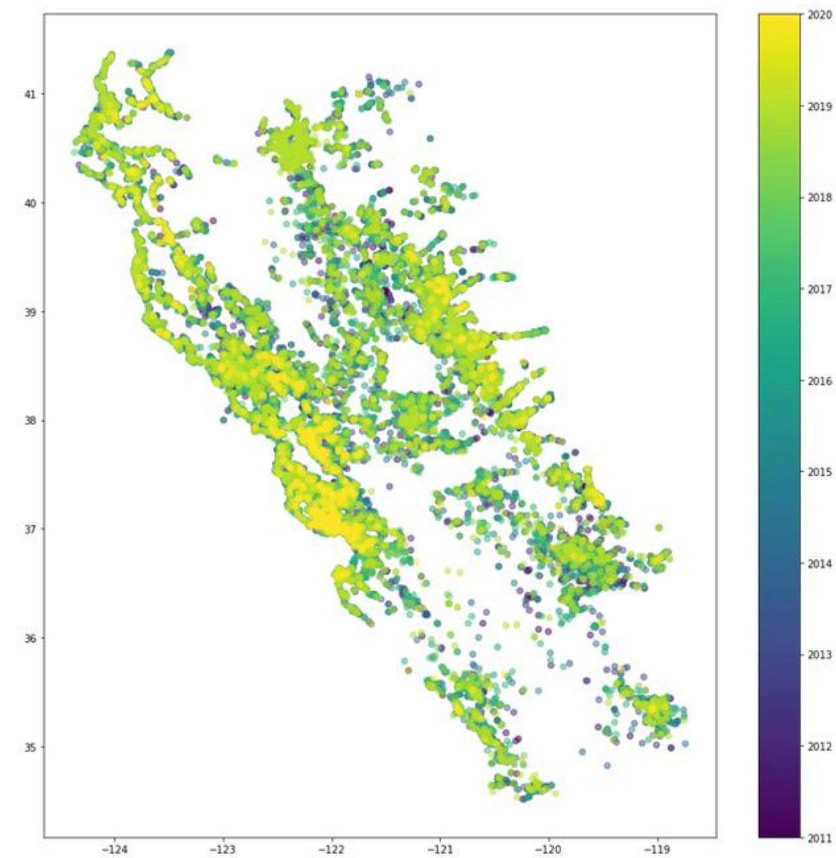
- Extensive dataset tracking outages -- many of which are the type of event that *could potentially* give rise to ignitions, but didn't. Capitalizing on these data we can observe more high-risk events than we would be able to access from working with ignitions data alone.
- Provides link between outages and mitigation measures taken in SAP.

The table below provides the counts of outages by major type since 2007. Vegetation-caused outages account for 8.4% of all outages, which is 4th highest behind Company initiated (33%), Unknown Cause (24%), and Equipment failure (21%).

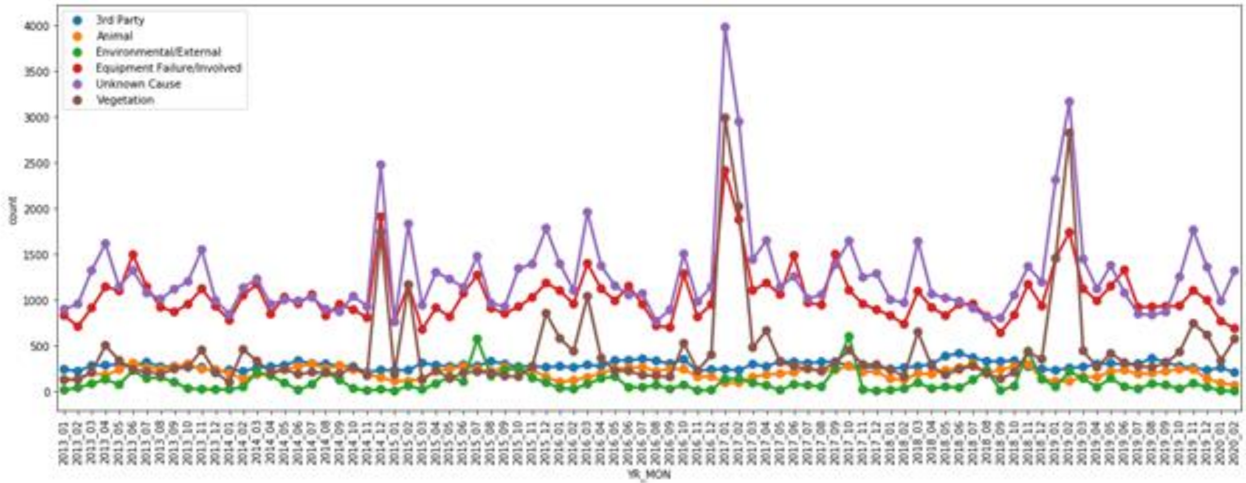
Outages by "basic cause":

Company Initiated	257528
Unknown Cause	189095
Equipment Failure/Involved	167046
Vegetation	65917
3rd Party	40890
Animal	32767
Environmental/External	15910
Wildfire Mitigation	14029
Total	784234

The map of vegetation caused outages from 2011 through 2020 plotted below give a sense of the number of vegetation-caused outages (65k+ since 2007), and the locations where they are concentrated, i.e. where the grid is in contact with trees tall enough to contact lines.



The time series of monthly total outages from 2013 to the beginning of 2020 plotted below underscores the episodic and seasonal nature of vegetation, equipment, and vegetation-caused outages, especially compared to animal, 3rd party, and environmental/external causes, which are more stable over time. For the most part, all three spike up together, with the largest spikes associated with winter storms and the smaller spikes are associated with lesser storms all year long. The correlation between unknown and vegetation caused outages, suggests that a significant number of outages of unknown cause are likely to involve vegetation as well. This suggests that our model may benefit from imputation of vegetation causes in the unknown category.



Some other useful statistics:

Veg-caused outages by the extent of the outage:

Distribution Circuit	56006
Transformer only	9002
Transmission line	878
Substation - Distribution	50
Substation - Transmission	12

Veg-caused outages leading to wires down (a little under one third):

No	45773
Yes	20144

Veg-caused outages by open equipment type - information relevant to determining the protection zone of the outage and therefore which assets were likely involved:

Fuse	38188
Transformer	9379
Line Recloser	8906
Circuit Breaker	3469
Switch	1594
Jumpers	1074
Other	549
Substation	443
Trip Saver	364
Sectionalizer	342

Sustained vs. momentary outages. The majority (85.6%) are sustained outages, which aligns well with the number of outages involving fuses:

Sus 670171
Mom 113011

Questions for SME

- We understand that the open point associated with outages is the location of the protection hardware that was triggered by the fault and that the fault was therefore within the protection zone of the piece of equipment that failed.
- We further understand that protection zones can change over time, as switches place the grid in unconventional configurations to back-feed power to restore or maintain power during outages and maintenance and repairs.
- Further we understand that circuit configurations themselves can change over time as the needs of customers change and engineers work to evolve the grid.
- Are protection zones stable enough over time to rely on the static memberships we can deduce from ED-GIS?
- At any given time, what fraction of protection zones/assets are in non-standard configurations?
- Storm caused vegetation failures have mostly to do with wind. We hypothesize that the subset of vegetation-caused outages that could occur any given day of the year (at the more or less constant "background rate") may have different root causes than the ones that occur during storm events. Are you aware of any information in this data set (or outside of it) that might be relevant to testing this hypothesis?
- Momentary outages are the result of self-correcting problems (line slap) and self-healing grid components (reclosers) whereas lonter outages are due to irreversible problems (lines down) and protection equipment (fuses). Can we determine the fraction of longer duration outages that would have been shorter with different protection hardware and what fraction are the result of irreversible conditions? We can tabulate the lines down vs. other outages, but expect that there might be more to the story there.
- Anecdotally, we were expecting to see more momentary outages than sustained. Our assumption was that there are a lot more "little" incidents on the grid than "severe" ones. However, we see the opposite. To what extent was our original expectation (i.e. that there are more minor than major problems on the grid) correct, but there are still significant numbers of sustained "minor" outages because fused protection devices can't recover quickly or automatically? Given that sustained/vs. Momentary is partially just a question of protection device type, what is the best way to assess the relative severity of the underlying condition causing an outage?
- The outage data contains information describing the type and condition of failed equipment. These are not consistently available and do not provide enough information to locate the specific hardware that failed, but do offer some insight into the type and cause of the failure. Do you share our current understanding that it will be hard to impossible to identify specific assets involved in a failure? If not, can you explain how we might reconstruct this information using ILIS and other related data sets? For example, some assets may have repair records in SAP.
- There are some outage log ids with multiple rows of data. For example event_id, aka EVENT# '19-0115935' is clearly a PSPS event with 29 rows, each with its own unique OUT_# ILIS index. We see a few different start times circuit ids within such large events but each row has a unique "open point" lat/lon value, confirming, we think, that each relates to a specific protection zone.

We also see mostly unique OIS ids. We would like to track outages with as much geographic specificity as possible, but outside of ILIS, the ILIS index value is not always found in other data sets. Specifically, the vegetation outage reports database and the wires down database have it, but the ignitions data does not. Even with vegetation outage reports and wires down, merges on the ILIS id are more sparse than merges by event id and similarly, OIS_OUTAGE# meged with the OIS id from the ignitions data also gets fewer hits than the same merge based on the EVENT#. Are we correct that the OUT_# is the unique identifier for outages at the protection zone level? What situations lead to multiple OUT_# or OIS_OUTAGE# values per EVENT#? Is it just a function of how widespread an outage is?

Wire-down database

Description	Tracks incidents of fallen conductors.
Temporal coverage	2012-2020
Record count	10,969
Link to schema	Shared as a Word document
Cross-references	ILIS; splices (no explicit link)
Key Contents	● Time and location of wire-down incidents ● Link to outage record (in ILIS) ● Characteristics of conductor that failed

Description

Wire down incidents are tracked due to the safety risks they pose to the public and to repair personnel. Recordings include observations logged by repair personnel, as well as information about the infrastructure itself—such as splices, protection, and compliance with design standards. Some of this information is obtained by querying ILIS and SAP. Most wire-down incidents include a foreign key linking to ILIS. Many incidents include a foreign key to SAP, though the field is sparsely populated.

Incidents are manually recorded when a repair is complete, and are updated to indicate if/when affected spans are replaced. Recordings may be less detailed or comprehensive when personnel are time constrained—for example during the types of severe weather events when wire-downs are most likely to happen. Most records (98.2%) can be linked to a record in ILIS. Where wire-down records are lacking in detail, refer to ILIS to determine when the incident occurred and to identify the protection zone affected.

How the Data are Useful

- ILIS data has a “WIRE_DOWN” field that is affirmative across 5.5% of all outages, but 30.5% of vegetation-caused outages. The data show that 11% of ignitions are related to wire-down incidents, and that 1.8% of wire-down incidents resulted in an ignition. Training a model to predict wire-down incidents (rather than predicting ignitions directly) can provide a much richer dataset than ignitions alone.

- Database records include information about conductors that subject matter experts have deemed relevant the occurrence of wire-down incidents. These information can provide a valuable starting point for generating hypotheses about the characteristics of the infrastructure where conductors are more to failure.

This data set labels about 30% of incidents as involving an energised downed wire.

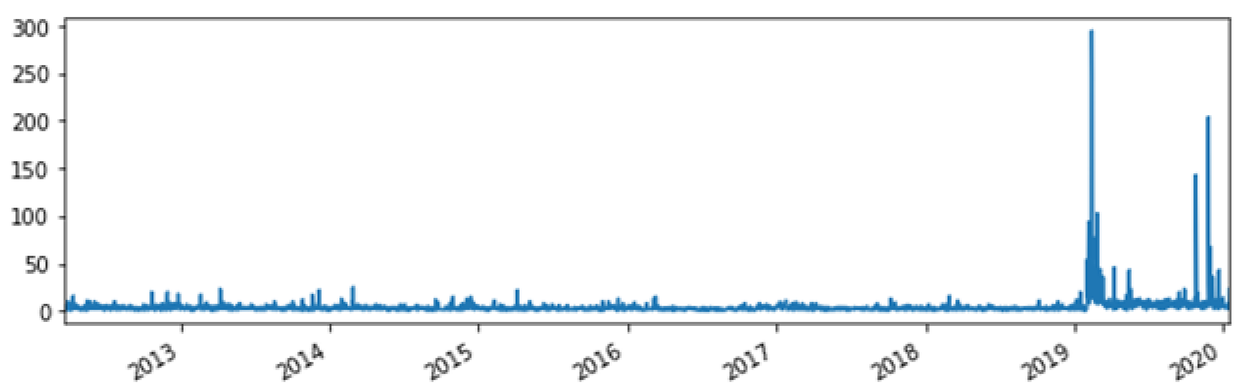
The table below summarizes the count of wire down incidents by "Basic cause". At 25% of incidents, Vegetation is the second leading cause after equipment failure (at 62%):

Equipment Failure/In	6763
Vegetation	2740
3rd Party	733
Animal	416
Third party	81
Environmental/Extern	80
Wildfire Mitigation	78
Company Initiated	48
Equipment Failure/Inv	10
Environmental/External	5
Equipment Failure/Involv	1

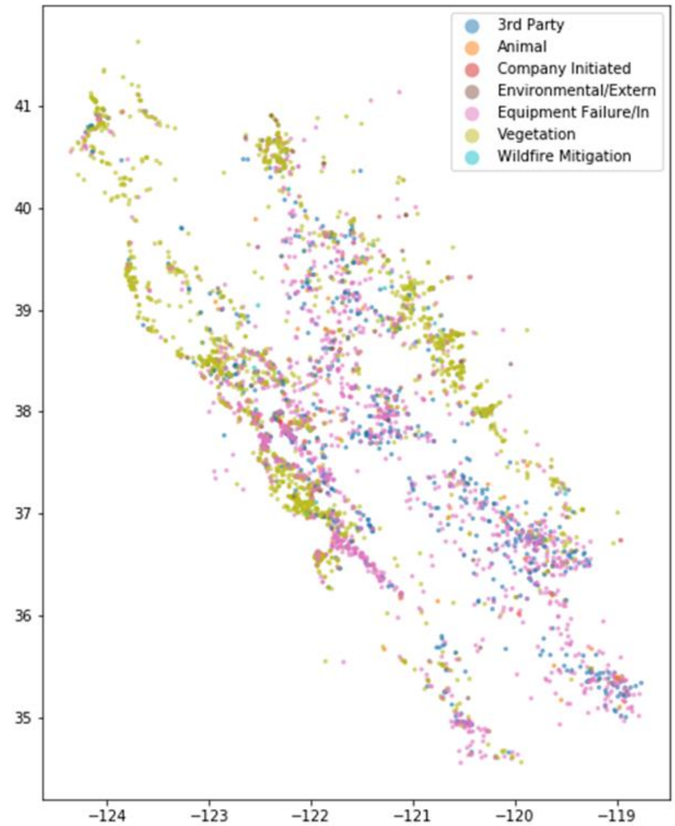
The figure below shows the number of wire-down incidents recorded each day between March, 2012 and January, 2020 (according to the field 'Outage Date'). There is a sharp uptick in events during 2019. The top 5 dates by count of incidents are:

2019-02-13	295
2019-02-14	244
2019-11-26	204
2019-10-26	143
2019-02-12	143

While all of those days are listed as ME days, meaning they were known days of widespread outages, the fact that these counts are specific to 2019 and so far above the counts from other years suggest that reporting practices have changed over time. Perhaps efforts to more thoroughly track incidents may have taken hold in 2019.



The next figure below maps wires down events since 2012 with coloration by cause. It can be readily verified that the geography of wires down caused by vegetation is substantially different from wires down from other causes.



Questions for SMEs

- We will eventually want to link these records with data reported in SAP and in EDGIS. How would you go about drawing connections between these datasets?
- Can you explain why we see such an increase in the number of incidents beginning in 2019?
- The wires down database tracks both the "OutageNumber" and "Event Log #". These correspond to the fields "OUT_#" and "EVENT#" in the ILIS data, with the OutageNumber corresponding to unique rows in ILIS and the Event Log # having the possibility of spanning multiple rows. However, when we merge ILIS data and wires down data, the "hits" based on OutageNumber are more sparse than the hits based on the Event Log #. Can you explain your understanding of these two ids and comment on why we might be seeing fewer matches when using OutageNumber? Our goal is to match each row of outages as specifically as possible, so we'd prefer to use OutageNumber if we can.

Vegetation outages investigation reports database

Contents	Investigative reports of outages caused by vegetation
Temporal coverage	Approximately 2008 to present
Rows	70,472 from 2006 through 2019
Link to schema	Microsoft SQL Server PRETAPPDBSWC019 > Projects database > OutageReports table
Cross-references	ILIS
Key contents	

Description

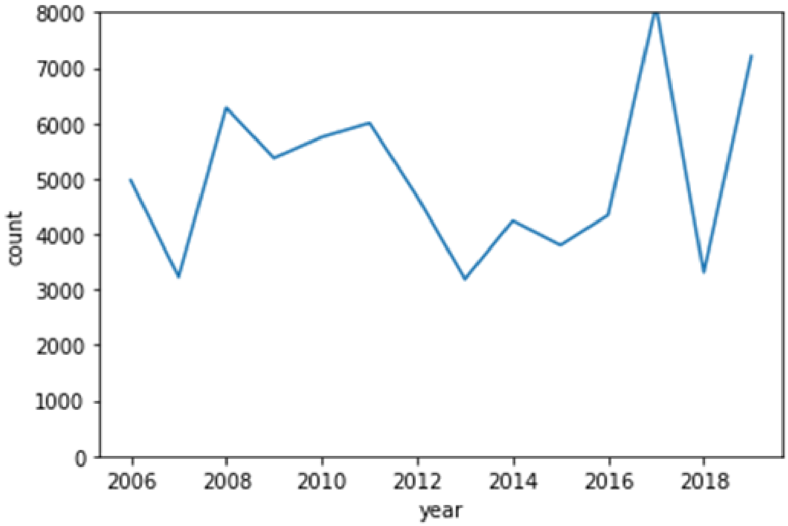
Outages believed to be caused by vegetation are investigated and reported via "Vegetation Management Distribution Outage Reports". These reports include details of the involved vegetation (species, size, proximity to conductors, health, failure mode, etc), the involved asset(s) (pole construction, splices, conductor insulation, etc), and local conditions (weather, soil, slope, snow, etc).

Additional analysis was provided that ranks tree species by general failure risk, specific failure mode risk (eg. branches fall, roots fail, etc), and fire ignition risk.

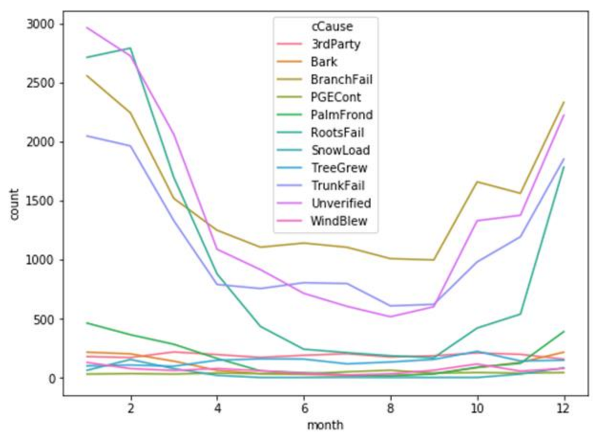
Expected insights

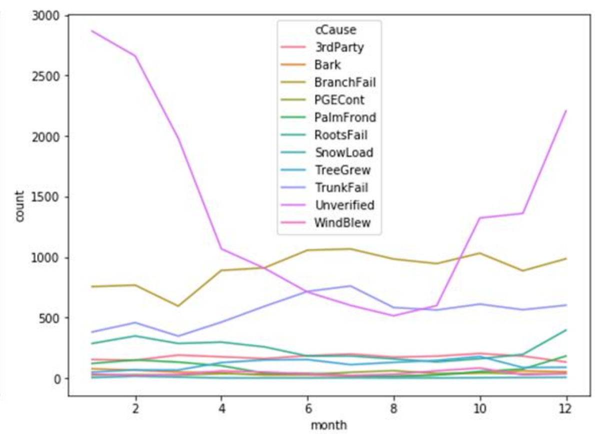
- Relative failure risks of tree species
- Relationship between tree species, local conditions, and failure modes

The figure below illustrates the count of vegetation outage reports per-year since 2006. The number is quite variable from year to year, averaging somewhere around 5,000, with 2017 and 2019 coming in as the top two years.

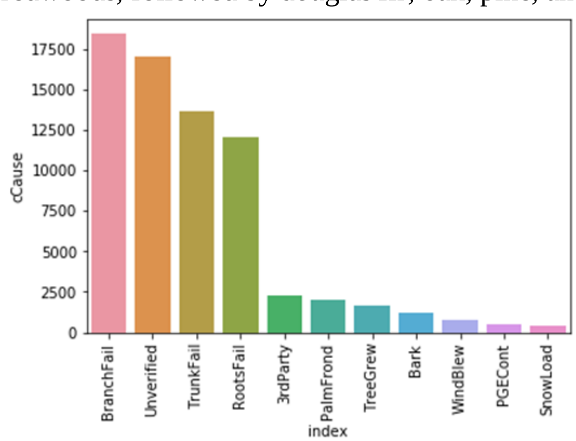


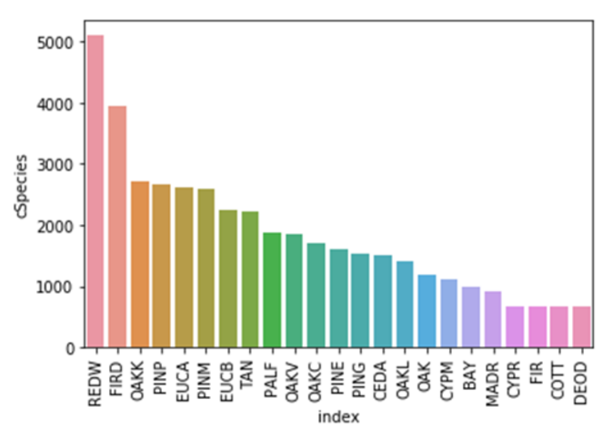
The number of Branch, Unverified, Trunk, and Root failure incidents dramatically increases in winter months (left plot). However, after removing weather with rain or snow and saturated soil, all but the unverified incidents level out (the filtering kept data missing weather or soil conditions):



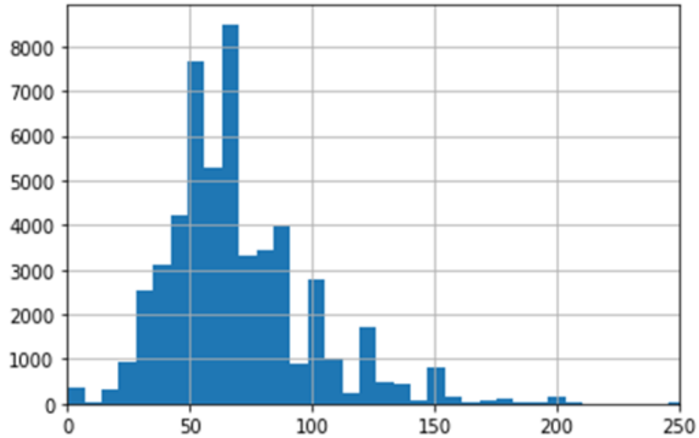


The top "cCause" of vegetation caused outages in the data set is branch failure (26.3%). Unverified causes are #2 (24.3%), followed by Trunk failure (19.5%) and Root failures (17.1%). The top species is redwoods, followed by douglas fir, oak, pine, and eucalyptis.

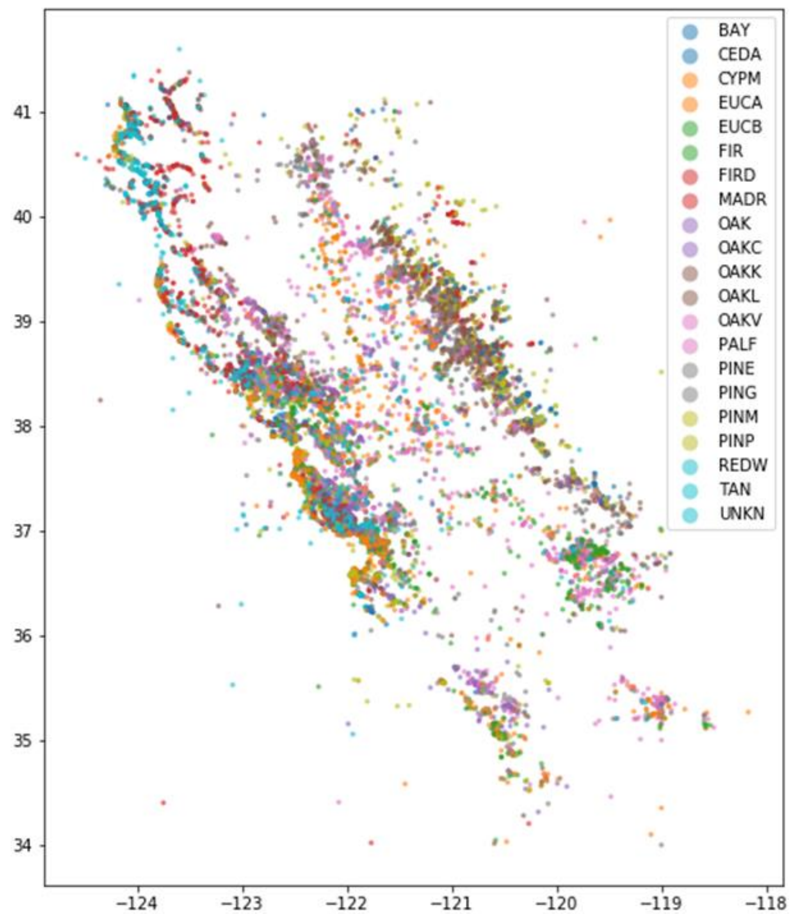




The distribution of tree heights (in feet along the x-axis) plotted below indicates that shorter trees occasionally cause outages, but those over 50 ft tall make up the bulk of the incidents:



A map of incidents by tree species under dry conditions reveals the overlap between species range and grid infrastructure.



Questions for SMEs

- How important are the context fields, including: 'cTreeRoad', 'cTreeCreek', 'cTreeSlide', 'nTreeAge', 'cTreeHealth', 'cTreeEvenAgeStand', 'cTreeRotInside', 'cTreeRotOutside', 'cTreeLean', 'cTreeNative', 'cTreeBlownOver','cWindDirection' to your understanding of what causes vegetation outages?
- Any others not mentioned in the above analysis that we should take another look at if we want to understand how to predict where/when vegetation will cause outages?
- The vegetation outage report database tracks both the "nILIS_ID" and "cRptNumber". These correspond to the fields "OUT_#" and "EVENT#" in the ILIS data, with the nILIS_ID corresponding to unique rows in ILIS and the cRptNumber having the possibility of spanning multiple rows. However, when we merge ILIS data and vegetation outage report data, the "hits" based on nILIS_ID are more sparse than the hits based on the cRptNumber. Can you explain your understanding of these two ids and comment on why we might be seeing fewer matches when using nILIS_ID? Our goal is to match each row of outages as specifically as possible, so we'd prefer to use nILIS_ID if we can.

Veg-caused ignition reports

Contents	Investigative reports of ignitions caused by vegetation
Temporal coverage	Approximately 2008 to present
Rows	“Hundreds per year”
Link to schema	Available only as part of the database or form examples, both of which contain privileged/confidential information
Cross-references	Ignitions, ILIS
Key contents	

Description

Ignitions believed to be caused by vegetation are investigated and reported via “Vegetation Management Incident Report Forms”. These reports include limited details of the involved vegetation, limited details of the involved asset(s), and limited details of vegetation conditions in the vicinity - these details are similar to, but less detailed than is captured in the Vegetation Outages reports described above.

Report forms are inconsistently transcribed into the database - some form fields are missed, with uncertain reliability. Original reports may be available in scanned PDF format.

Expected insights

- Details of how vegetation failure leads to ignition
- Note that the DxRisk team has not explored this dataset in detail

Questions for SMEs

- Is this data set sufficiently redundant with the Vegetation outages investigation reports data that it can be ignored?

Relationships between events

Synthesizing all the information uncovered to date on event data from ILIS, wires down, vegetation outage reports, and ignitions, we have developed (in Python/Geopandas) a canonical view of each data set that allows each to be joined to the others using one of outage_id, ilis_id, or ois_id, depending on circumstances. The table below summarizes the results of that work, with a sort of rosetta stone for the overlap between data sets. From this table, we can learn that 8.8% of outages are covered in the vegetation caused outage reports data and 81.6% of the reports can be linked to specific outages. Similarly, just 0.2% of outages can be linked to ignitions, but 93.8% of ignitions can be linked to outages.

These are the relationships that we expect to base our model "navigation" from predicted outage to ignition on. We expect that the percentage overlap numbers will be slightly improved with hand check for data quality of the relevant identifiers.

left	right	matches	n_left	n_right	pct_left	pct_right
outages	wires_down	10772	707312	10812	1.5	99.6
outages	veg_outage	62341	707312	76389	8.8	81.6
outages	ign	1267	707312	1351	0.2	93.8
wires_down	veg_outage	2342	10812	76389	21.7	3.1
wires_down	ign	198	10812	1351	1.8	14.7
veg_outage	ign	385	76389	1351	0.5	28.5

IPython notebooks on event data sets

We have developed notebooks as ipynb files that enabled exploration and normalization of event data.

They are all found in the PGE-Dx-model-data-science github repository and include:
notebooks/wires_down.ipynb - analysis, tables and figures found in the wires down section of this document.
notebooks/vegetation_outages.ipynb - analysis, tables and figures found in the wires down section of this document.
notebooks/ILIS_outages.ipynb - analysis, tables and figures found in the wires down section of this document.
notebooks/canonical_event_data.ipynb - tabulation of the results of different strategies for merging dataset together.
notebooks/dataset_summaries.ipynb - exploratory analysis of contents reported in event data fields and investigation of differences/similarities in reporting among data sets.
notebooks/data_synthesis.ipynb - implementation of field renaming convention across event data sets and investigation of field contents.

Physical infrastructure

Conductors

Description	Tracks locations and attributes of conductors, including splice counts, voltage and feeder IDs
Temporal coverage	Static (or periodically updated) export of system attributes
Link to schema	EDGIS_baseline_spreadsheet_052619.xlsx
Cross-references	Poles, feeders, substations
Key contents	● Multiple databases with divergent conductor data, especially as it relates to lines and spans ● GIS locations of conductors in multiple formats (lines and points) ● Conductor physical status (corrosion, splices) and environmental conditions (nearby strike trees) tracked

Description

We've identified four sources of conductor locations, 1) the publicly-accessible Integration Capacity Analysis (ICA) map, 2) the primary and secondary conductor tables in the **arad_dev/edgis** database, and 3) the conductor tables in the **aradgis/edgis** database and 4) the node/edge tables in the **arad_dev/epic320** database. These sources all have different attributes, different locations, and different levels of detail.

1) ICA
    - Contains detailed, multi-line format GIS data for 1,019,626 conductors
        i) Attributes: [FeederId, FeederName, Globalid, CSV_LineSection, ICA_Analysis_Date, LoadCapacity_kW, GenCapacity_kW, GenericPVCapacity_kW, GenCapacity_no_OpFlex_kW, GenericCapacity_no_OpFlex_kW]
    - Contains an additional 825,512 conductors with many attributes withheld,
        i) Attributes: [FEEDER_ID, SHOW, NUMBEROFPHASES, GLOBALID]
    - Total conductor count: 1,845,138

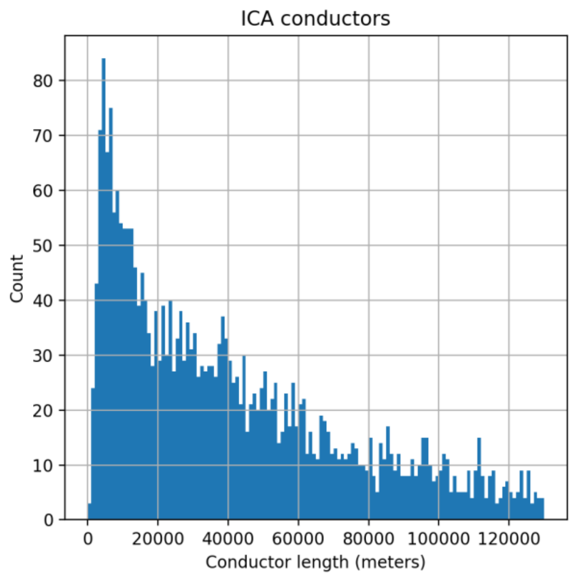
2) arad_dev/edgis
    - Compiled to support the EPIC360 project, and focuses on providing details about the attributes and specifications of assets on the grid.
    - Contains line-format GIS data for 234,974 primary conductors and 2,575,419 secondary conductors
        i) These data have been cleaned and exported on Sagemaker to **/data/vector/primary_conductors_32610.geojson** and **/data/vector/secondary conductors_32610.geojson**
        ii) Attributes: [objectid, globalid, circuitid, circuitname, convcircuitid, cedsanumberofphases, feederinfo, feedertype, jackettype, phasedesignation, operatingvoltage]

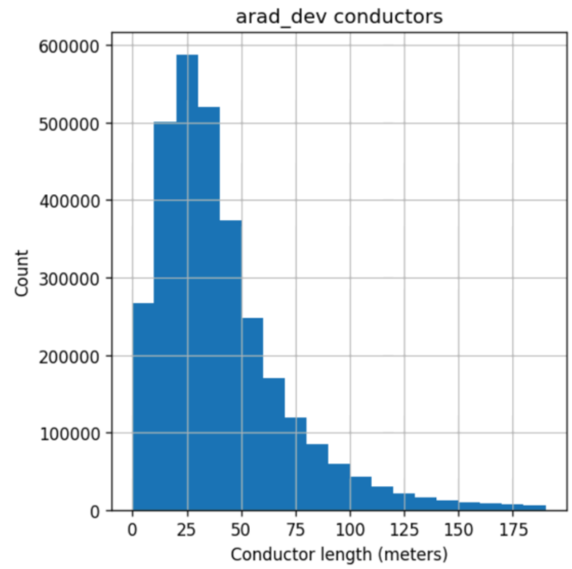
    iii) Derived from Redshift views **arad_dev/edgis/vw_snapshot_priohconductor** and **arad_dev/edgis/vw_snapshot_secohconductor**
  - Total conductor count: 2,810,393
  - UPDATE: it looks like these number are incorrect; the original data exports were incomplete (the total conductor count should contain 4,571,605 records)

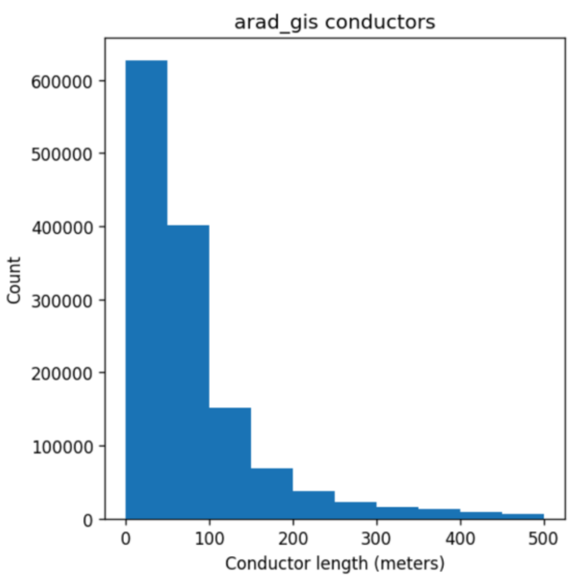
3) arad_gis/edgis
  - Compiled to support the STAR project. These data include information relevant to studying the performance characteristics of each asset. For example, the data report the number of splices and summarize inspection records for each asset.
  - Contains start/end point locations for 1,382,422 conductors
    i) These data have been cleaned and exported on Sagemaker to **/data/vector/conductors-edgis-lines.geojson** in line format and to **/data/vector/conductors-edgis-points.geojson** in point format (the centroid of the start/end location)
    ii) Attributes: [conductorguid, objectid, polecount, corrosion, totalsplice, length_m, tree_count]
    iii) Derived from Postgres multiple tables, **aradgis/edgis/conductor***
4) arad_dev/epic320
  - Compiled to support the EPIC 320 project, this database contains data derived from a snapshot of the EDGIS data.
  - See the grid topology section of this document for more details.

How the data are useful

Currently, the "best" source of conductors data is not clear, and each data set has unique advantages. ICA data appear to have the most detailed GIS geometry data, tracing multi-line features (presumably corresponding to potentially multiple spans per conductor). The arad_dev data may be similarly detailed, but with each conductor record actually corresponding to individual spans. The arad_gis data include attributes derived from airborne LiDAR, but appear incomplete. Here's a comparison of conductor lengths for each dataset:



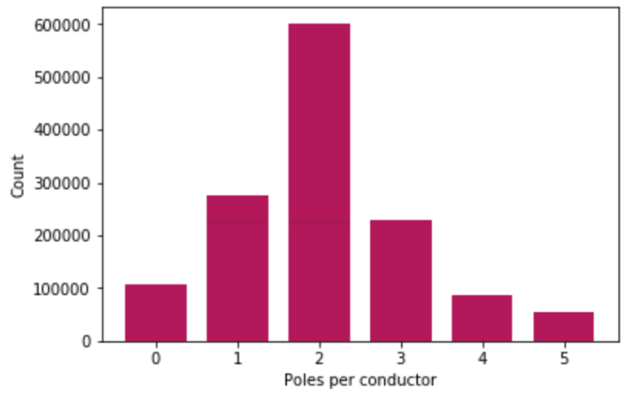




Conductor length sums:

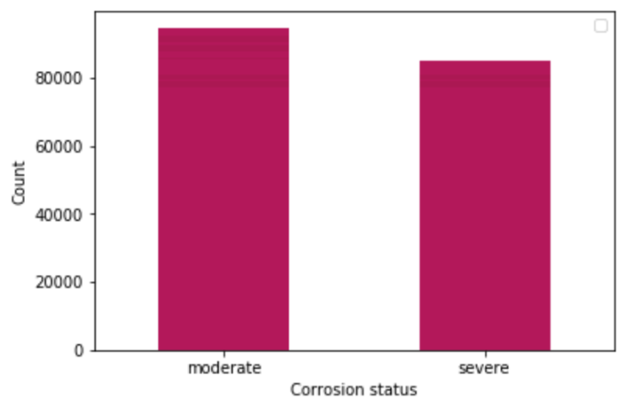
1. ICA: 171,690,894 meters
2. arad_dev: 151,449,233 meters (new revision in-process)
3. arad_gis: 128,443,810 meters

Below are some features of the **arad_gis** conductors data



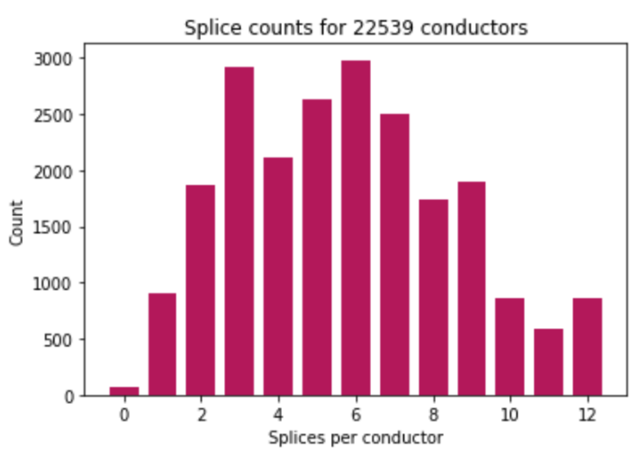
Polecount

- Number of poles per conductor ID
- Total # poles: 3,119,063
- Number of conductor IDs:
  - 1,382,422 rows
  - 1,382,422 unique



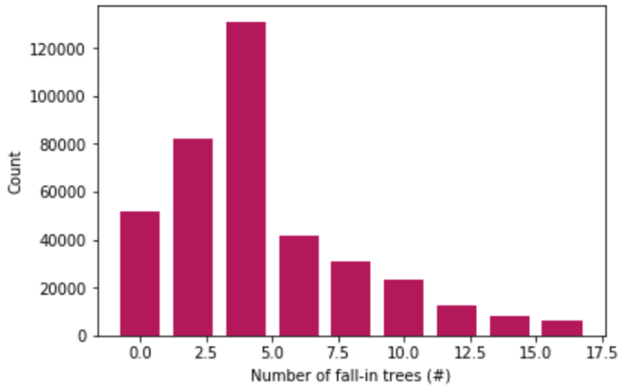
Corrosion

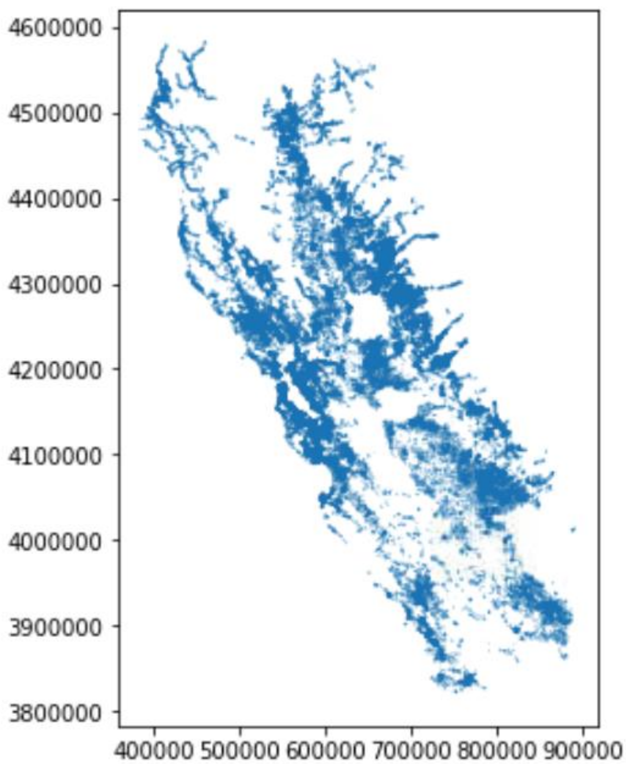
- Categorical classification of corrosion status
- Two classes: moderate and severe
- Perhaps not measured in the field--possibly based on proximity to coast
- Number of conductor IDs
  - 180,068 rows
  - 179,507 unique



Splice

- Splice count and, occasionally, year of splicing
- Number of conductor IDs
  - 22,539 rows
  - 20,829 unique IDs

	Treecount
	● Counts the number of strike trees near a conductor
	● Source of this data is unclear
	● Total treecount: 2,217,746 trees
	● Number of conductor IDs
	    ○ 396,143 rows
	    ○ 396,143 unique

	Treecount spatial coverage ● It appears that these data are more comprehensive than expected--covering more than just the T2 & T3 LiDAR flights

Questions for SME

- What do individual conductors refer to in each dataset? What does it mean to analyze a single conductor?
- Are conductors with pole counts of ‘0’ mislabeled, or a way to identify undergrounded lines?
- Where did the tree count estimate from the aradgis/edgis conductors data come from? And can conductors with no value be interpreted to say there are zero nearby trees?
- How were corrosion scores estimated?

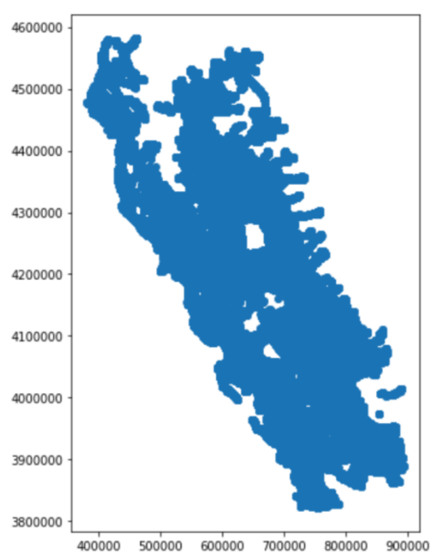
Poles

Contents	Location and IDs for dx poles
Temporal coverage	Static export of system attributes
Link to schema	[EDGIS_baseline_spreadsheet_052619.xlsx]()
Cross-references	Conductors (aradgis/edgis)
Key contents	Location and cross-walk of poles-to-conductors in aradgis

Description

Point format GIS data on the location of each pole, with associated ID codes to match each pole with an associated conductor. This cross-reference is done via the **aradgis/edgis/conductorpolerelation_updated** table, which contains information on pole installation date and conductor 'job year.' The cross-reference is done via the 'conductorguid' and 'supportstructureguid' labels, which are unique identifiers for each conductor/pole, respectively. This cross-reference table contains 4,270,367 records, just under the sum of the total number of recorded conductors (1,382,422) and poles (3,070,173) in aradgis (4,452,595 total).

Expected Insights



- Pole locations appear to correspond to aradgis/edgis condutors locations
- Total pole count in poles database: 3,070,173
- Total pole count in conductors (sum of 'polecount'): 3,119,063

Questions for SME

- The total number of poles in the poles database (3,070,173) closely matches, but is not the same as the sum of the 'polecount' attribute in the aradgis/edgis conductors database (3,119,063). Is this discrepancy due to some duplicate poles that are the joint between conductors?

Grid topology - nodes and edges

Contents	Topological information about grid structure with nodes (various types of equipment) and edges (conductors)
Temporal coverage	Recent snapshot (exact timing not know to this team)
Link to schema	araddev.epic320
Cross-references	ED-GIS, ILIS open points
Key contents	Hierarchical data on the topological relationships between grid assets

Description

Every outage in ILIS has a record of the "open point" associated with the outage. We understand these points to be the locations of protection hardware that "opened" to protect the section of the grid that suffered the failure that was the root cause of the outage. Because many outages are short lived and others are self-correcting, not all outages are actively investigated. Even when there is follow-up, there is no guarantee that it will be possible to identify the specific asset that triggered the outage. The result is that the most reliably available information about what asset(s) failure caused an outage is which protection device was triggered, and therefore which protection zone experienced the failure. To establish the protection zone, i.e. list of protected assets, associated with each protection device, requires a topological model of the grid. Such information, derived from relationships in source ED-GIS tables, exists as the result of work on EPIC320.

Expected insights

List of assets in the protection zone of protective devices provide asset associations with ILIS outages characterized by an open point at a protection device.
List of assets on a given feeder for roll-up modeling and predictions at a feeder level.

Questions for SMEs

(1) The nodes and edges describe the normal configuration of the grid, but there are switches all over the grid that allow "back feeding" from one circuit to another to provide redundancy when parts of the grid have been damaged or during maintenance and repairs. How often and for how long are switches set to non-standard positions? What percentage of protection zones are therefore configured in a manner different from the relationships found in these nodes and edges tables? What fraction of assets can be expected to be mis-classified?

(2) Circuits on the grid are subject to expansion, upgrades, and other revisions. At what rate of annual change (i.e. in terms of the percentage of misclassified assets) would a snapshot of grid "nodes" and "edges" fall out of date? Through what process might one reconstruct the time history of grid topology to recreate the relationships during an outage from, e.g. 2010?

(3) What is the most reliable way to key from the assets found in the nodes and edges to our other data sets? ED-GIS GUIDs?

Inspections and maintenance

Work records (SAP)

Contents	SAP is an enterprise tool for logging all work that is done on the system.
Temporal coverage	
Link to schema	Accessible only through user interface
Cross-references	ILIS, inspections, asset datasets
Key contents	Database includes the time, location, and nature of the work that's been done. Work includes maintenance, repairs, inspections, and replacements.

Description

SAP is a software platform that both personnel and automated systems interface with to generate records of work on the system -- including work that is anticipated, and work that has already been done. These include documentation of the assets that are in place, as well as the assets with which they are replaced. Records may be generated from outage notifications, or inspection records indicating that a component is in need of replacement.

Expected insights

SAP records can tell us about the state of repair of the grid at a particular point in space and time. Tracking repairs and replacements that have been done over time can help us to understand what condition grid assets were in when an event (e.g., an outage or ignition) occurred, and to examine how the characteristics (e.g., age, state of repair, design characteristics) of grid assets change the probability that different types of events could occur.

Given the need to geo-locate grid assets, the DxRisk team began our inquiry of asset data with GIS datasets described above. We have not yet done a thorough review of the information reported in SAP.

Questions for SMEs

- None yet! We haven't worked with this data in enough detail to have questions here, but we understand that there is a lot of valuable data in SAP and that much of it is quite nuanced.

PRONTO - asset inspections

Contents	Asset inspection reports
Temporal coverage	Unknown
Rows	~718,000
Link to schema	
Cross-references	SAP, Equipment ID
Key contents	

Description

Asset inspections are recorded on forms which are saved as PDF documents. The contents of these documents are recorded in the PRONTO database. Inspection data is also copied to SAP.

Expected insights

- Note that the DxRisk team has not yet explored this dataset in detail

Questions for SMEs

- None yet

O-Calc Pro - pole loading model

Contents	Digital twin model of a fraction of Dx network pole assets
Temporal coverage	Uncertain - believed to be only a few years
Rows	Understood to be ~700,000
Link to schema	Available by browsing the O-Calc Pro database
Cross-references	Pole Test and Treat (PTT), Equipment ID (links to SAP asset data)
Key contents	

Description

O-Calc Pro is a physics-based modeling tool that provides a "digital twin" of distribution network pole assets, third party attachments, conductors, guy wires, etc. O-Calc Pro is used to model the "safety factor" of assets under specific wind conditions - to estimate the wind speed and direction that results in a given pole's failure probability. Note that O-Calc Pro does *not* model risk probabilities due to vegetation failure (eg. it does not model the risk that a given pole will fail in high winds because the winds cause nearby vegetation to impact the pole or adjacent conductors).

O-Calc Pro relies on LiDAR mapped asset data, which is currently incomplete. Of ~2.4 million distribution network poles, ~700,000 have been LiDAR surveyed, and ~300,000 have had their lat/long locations corrected from LiDAR survey data. All remaining poles will be LiDAR mapped in the next ~3 years.

PRONTO asset inspections (described in another section of this report) are inconsistently incorporated into O-Calc Pro data inputs.

Some assets have had maintenance work planned or performed on the basis of O-Calc Pro's predictions. The history of these assets is recorded in O-Calc Pro's database at multiple stages of the review, planning, and construction process.

Expected insights

- For each pole, predict the failure probability at a given wind speed and direction(s)

Questions for SMEs

- Are the pre-computed failure probabilities present in the O-Calc Pro database potentially useful in predicting veg-caused asset failures?
- When the Dx Risk team expands our modeling to incorporate ignition causes other than vegetation-related, does the O-Calc Pro database contents become more useful?

- Can the O-Calc Pro application be utilized to predict per-asset failure probabilities under arbitrary weather conditions?

Splices

Description	Tracks incidents of broken conductors. Reports the cause, safety concerns, and details of the repair process.
Temporal coverage	2013-2019
Link to schema	None provided
Cross-references	ILIS, Circuit, SAP
Key contents	• Report number of splices on spans where inspections show more than 3 splices on a span. • Includes counts per phase, and by splice type (automatic vs. compression) • Provides link to ILIS (perhaps to corresponding wiredown?) and to reconductoring job (presumably in SAP).

Description

Dataset documents spans of conductors with >3 splices. Describes the type of splice, number of splices per phase, and reconductoring jobs that have mitigated the issues. SME estimates that 95% of splices are automatic and 5% are compression splices. Automatic splices pose a greater wildfire risk, though they are no longer allowed in the HFTD.

Splices occur due to wire-down incidents. SME estimates the breakdown of causes to be:

- 15% age
- 20% size and type
- 20% pre-existing splices
- 15% corrosion zone + conductor type (copper resists corrosion, ASME does not)
- 10% snow loading
- 15% I^2T limits exceeded (i.e., too much heat dissipation I^2 for too much time T)
- 5% load ampacity exceeded

Expected insights

- Describes current state of repair for conductor spans
- Indicative of how vulnerable a span is to failure upon veg contact

Questions for SME

- How comprehensive is reporting? Are there potentially spans with lots of splices that wouldn't be recorded here?
- What does the inspection process look like? How are splices detected? What other information is recorded?

- There's a notification type in SAP to indicate corrosion. Have these notifications been linked to splice data? What would be involved in doing so?
- There's a column indicating that a span of conductor was repaired. What generally triggers a repair? Is it a wire-down? And inspection? Or something else?

Environmental data

Fall-in trees

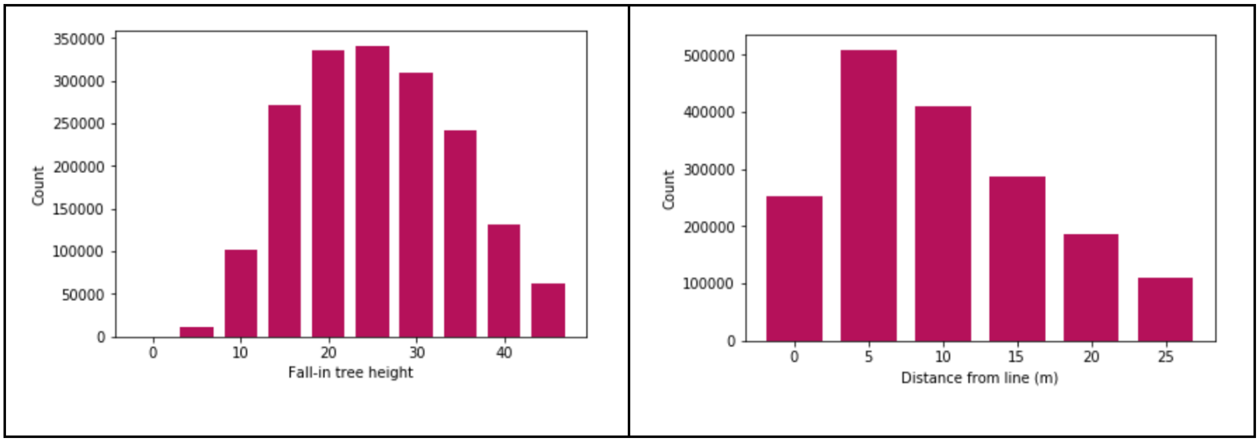
Contents	Top-of-tree lat/lon location, height, distance to line and overhang distance of all strike trees near conductors, restricted to T3 risk areas
Temporal coverage	Snapshot - 2018
Link to schema	None
Cross-references	Conductors (aradgis data)
Key contents	LiDAR-derived locations and distance to line for all strike trees in a high risk region

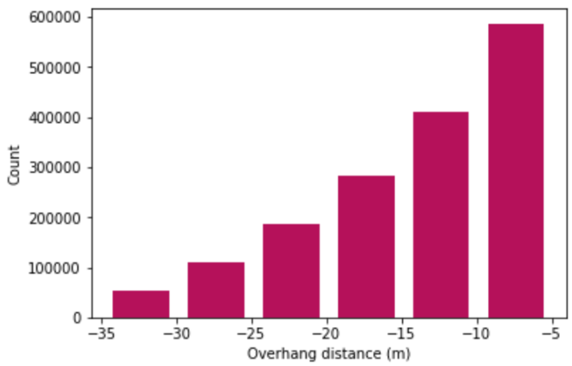
Description

These data identify the location, height, distance from line, and overhang distance for each strike tree in the T3 fire risk zone. It's our understanding that these data were generated from the airborne LiDAR data collection in 2018, and that the 2019 LiDAR data covering the Tier 2 fire risk zone has not been fully processed yet. Tier 1 does not have comprehensive LiDAR coverage.

How these data are useful

These data should help predict where vegetation contact events from fall-in trees are likely to occur.



	• This dataset includes 1,843,026 individual tree locations. • These data were likely used to compute the ‘treecount’ attribute of the aradgis conductors data

Unfortunately, these data do not cover the whole grid. However, by combining the Salo tree height data with a distance-from-grid map, we may be able to generate a similar dataset that could be validated using the fall-in trees database.

Questions for SMEs

- None

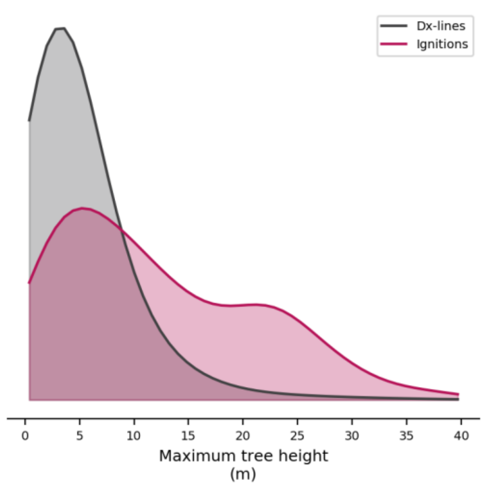
Salo Sciences tree height

Contents	Statewide map of tree height derived from satellite imagery at 10 m resolution
Temporal coverage	Derived from satellite imagery collected November, 2019
Link to schema	Data generation, description, and model performance metrics reviewed in [this spec. sheet]()
Cross-references	N/A
Key contents	Modeled prediction of the location and height of every tree across the state at high resolution

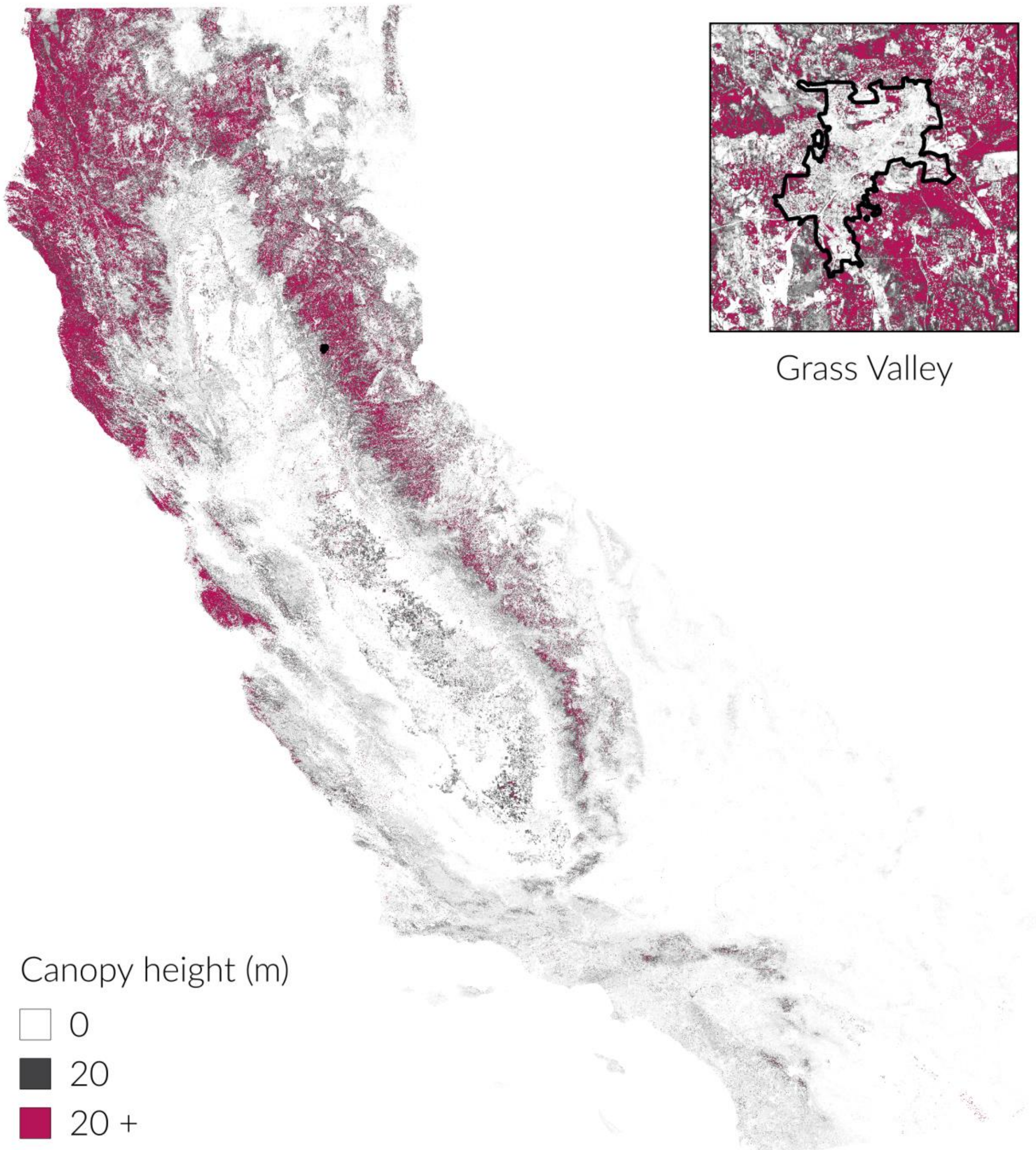
Description

Tree height maps the distance between the ground and the top of the canopy. It's a proxy for aboveground biomass and the amount of foliage that may be consumed in a canopy fire, but is also a key predictor of where vegetation contact ignitions are likely to occur across the distribution grid.

These data were generated by using publicly-accessible airborne LiDAR to map tree height patterns over small sections of the state, then upscaled by training a model to recognize height patterns in satellite data. The USGS 3D Elevation Program hosts, but these data are only available for a small fraction of California's 423,970 km² area. To overcome this, Salo trained deep learning models—a form of pattern recognition—to identify these forest structure patterns in satellite imagery, then mapped height statewide. These algorithms are of the U-net family of neural network architectures that perform pixel-wise regression and classification tasks. The satellite data includes imagery from Sentinel-1 C-band radar sensors and Sentinel-2 multispectral sensors at 10 m spatial resolution, collected in Fall 2019



- In milestone 1, the maximum tree height in a 100 meter grid cell was the top predictor of ignition location
- At full 10 meter resolution, there are 4,052,745,221 valid pixels of tree height predictions across the state



Questions for Salo

- None

Vegetation management database (VMD)

Contents	Inspections, planned work, and performed work
Temporal coverage	Approximately 2000 to present
Rows	“Millions of trees per year”
Link to schema	Unknown
Cross-references	
Key contents	

Description

The Vegetation Management Database (VMD) has been characterized to the DxRisk team as “archaic” and “limited in what data it contains”. Inspections are performed on 100% of the Dx network annually, and these inspections are recorded in this database.

Data from the Enhanced Vegetation Management (EVM) program, which began relatively recently, is in a separate database, not in the VMD.

Expected insights

- Note that the DxRisk team has not yet explored this dataset in detail

Questions for SMEs

- None

Meteorology and fuel moisture data

Description	Hourly, 3km resolution data from custom "POMMS" WRF modeling
Temporal coverage	"30" years: late-1988 to late-2018 (Weather data now through 2/2020)
Link to schema	None
Cross-references	Linked through space-time
Key contents	• Dead Fuel Moisture (1,10,100, and 1000 hour fuels), • NFDRS burn indices (ignition, spread, and energy release components, burn index), • Live Fuel Moisture • Weather (wind, temperature, relative humidity, dew point temperature, dew point depression, and precipitation)

Description

PG&E Operational Mesoscale Modeling System (POMMS) Data stored on S3 "pge-climatology/2019-30-year" as NetCDF (.nc) files in 3 folders:

1) "dfm_nfdrs": Dead Fuel Moisture (DFM) and National Fire Danger Rating System (NFDRS) fire indices {hourly data in daily files}
2) "lfm": Live Fuel Moisture
3) "weather": wind, temperature, relative humidity, dew point temperature, dew point depression, and precipitation {hourly data in monthly files}

Below are some examples of the hourly data in the daily dead fuel moisture (DFM) / NFRDS files.

Variable Name	**Definition**
mean_wtd_moisture_1000 hr	1000 hour Dead Fuel Moisture
IC	Ignition Component
ERC	Energy Release Component
FL	Unknown at this time
FRS	Unknown at this time

SC	Spread Component
BI	Burn Index
FIL	Unknown at this time

Based on our initial walk through, the data appears to be NetCDF “Classic” format.

LFM data only contain one variable: Live Fuel Moisture.

All variables of the hourly “Weather” data in monthly files:

Variable Name	**Description**
PREC_ACC_NC	Accumulated precip over prec_acc_dt – periods of time
PSFC	SFC Pressure
Q2	QV at 2m (Specific humidity)
RAINNC	Accum. Total grid scale precip
SMOIS	Soil Moisture
SNOW_ACC_NC	Snow accumulated
SWDOWN	Downward Shortwave Radiation
T2	Temperature at 2m
TSLB	Soil Temp
U10	U component of wind at 10 m
UST	U* (Similarity Theory, Frictional Velocity)
V10	V component of wind at 10 m

ZNT	Time-varying roughness length

Expected Insights

- PG&E weather and fuel moisture data is expected to be of higher quality than the hourly 2.5 km RTMA and daily 4 km GRIDMET weather data currently incorporated into the Dx Risk

Questions for SMEs

- Is it correct that there are no wind gust speeds in POMMS data?
- Do data include wind direction, or just U-component and V-component?
- How are U and V components dealt with (individual and temporally-aggregated)
- Are "sustained winds" those derived from the hourly components?
- Were weather data processed to anything other than hourly?
- Any suggestions for temporal aggregation (e.g., for wind direction – U & V components)?
- POMMS reported to be better than RTMA for capturing high winds. Can we generally assume that same greater accuracy for other variables?
- What are the nuances of the two different precipitation/rainfall variables in the Weather data?
- Are there any known weaknesses of the POMMS data (e.g., negative precipitation values documented in Gilleran FPI document)?
- Are there any important meteorological variables we might better find in daily GRIDMET or hourly RTMA?
- LFM's are generally for a small number of plant types (all but one are shrubs) – a limitation for tree-covered areas?
- Does NFDRS' "Ignition Component" (% of ignitions that will cause a fire requiring suppression) serve as the potential connection between our ignition model and PG&E's FPI, given that FPI = P(1000 acre fire) | 40 acre fire?
- In the dead fuel moisture (DFM) / NFDRS data, what are FL, FRS, and FIL?

<u>Questions about prior modeling efforts</u>

- OUTAGE PRODUCING WIND (OPW) MODEL
  - {Trained on 10 years of ILIS outages and hourly wind speeds.}
  - Did it consider wind directions, e.g., orthogonality of winds w.r.t. lines (as discussed in a PGE powerpoint)
  - Why restrict to HFTD?
- "DIABLO STUDY" (~400 days)
  - What are the exact wind directions for "offshore" (0 through 90 degrees?)
- "600 (worst?) DAYS"
  - What exactly are the "low pass filters" for R.H. and Precip (only removing those days with pretty extreme values in these variables?)

APPENDIX

Field renaming conventions

Ignitions data

Before 2018

Source field	Dx-Risk field	Notes
Date	date	Datetime
Time	time	Datetime
Log	ilis_id	Foreign key to ilis
OIS#	ois_id	Foreign key to OIS
Div	division	
Circuit	feeder	Foreign key to feeders
Latitude	lat	Location
Longitude	lon	Location
Material at Origin	ground_cover	
Land Use at Origin	land_use	
Size	fire_size	
Other Companies attached to pole	attachments	
Voltage	voltage	
Equipment Involved With Ignition	equipment_involved	
Line	network_location	
Type	line_type	
Outage Type	outage_type	

Suspected Initiating Event	fire_cause	
Equipment Facility Failure	equipment_failed	
HTFD Area?	hftd_area	
Primary Remarks	free_entry	
Fire cause	component_cause	
BPR Category	cause_category	

After 2018

Source field	**Dx-Risk field**	**Notes**
Date	date	m/d/YYYY
OIS #	ois_id	Foreign key to OIS
Fire Zone	hftd_zone	
ILIS #	ilis_id	Foreign key to ILIS
INT # (T-line)	transmission_line	
CAP # (T-line)	tline_id	Foreign key to EDGIS
T or D	transmission_flag	
CPUC Reportable?	reportable_flag	
EIR?	eir_flag	
Suspected Initiating Event	cause	Candidate partitioning
Circuit	feeder	Foreign key to feeders
OH or UG (OH unless otherwise indicated)	line_type	
If cause = PG&E equipment, what type?	equipment_cause	
If cause = PG&E equipment, addtl notes	equipment_damaged	
Reviewer Remarks	free_text	

Notes from T-men Follow-up or Internal Records	free_text_1	
Final FireLat	lat_fire_final	
Final FireLong	lon_fire_final	
Division	division	
District	district	
Position Latitude	lat	
Position Longitude	lon	
Electric Feeder Number	feeder_id	
Primary Remarks 1	free_text_2	
Fire Latitude	lat_fire	
Fire Longitude	lon_fire	
Fire Size Description	fire_size	
Attachment Count	n_attachments	
Pole Barcode	sap_id	
Outage	outage_flag	
Est Wind	wind_category	
Fire Size	fire_size_2	
Actual Cause	fire_cause	
Equipment	equipment	
Energized Wire Down?	hotwire_flag	
PLANNED	planned_outage_flag	
Fault Type	fault_type	
Equipment Type	equipment_type	
Equipment Condition	equipment_condition	
PM	sap_id_2	
ILIS Cause Correct?	ilis_cause_correct	

bTreeFound	tree_found	
cSpecies	tree_species	
cInsulation	insulator_type	

Wire Down Database

Source fieldname	Dx-Risk fieldname	Notes
Division	division	
Outage Date	dttm	
Event Log #	ilis_id	Foreign key to ILIS
Circuit	feeder	
Weather Condition	weather_category	
ME Day?	med_flag	
Basic Cause	cause_code	
Supplemental Cause	supplemental_cause	
EquipmentType	equipment_type	
EquipmentInvolved	equipment_involved	
EquipmentCondition	equipment_condition	
OutageNumber	ois_id	Foreign key to OIS
PM #	sap_id	Foriegn key to SAP
Downed Wire Energized?	hotwire_flag	
If Energized=Yes: Wire was on ground:	onground_flag	
If Energized=Yes: Wire was Surface:	onsurface_flag	
If Energized=Yes: SGF setting installed?	sfg_protection_flag	

IF SGF installed=Yes: Lockout by SGF?	sfg_triggered_flag	
Source Side Protection Type	protection_type	
Wire Type	wire_material	
Year Installed	year_installed	
# of Phases	n_phases	
Total # splices in span (all phases)	n_splices	
Splice type	splice_type	
Span Length (ft)	span_length	
Corrosion Zone?	corrosionzone_flag	
Built to Standard	uptospec_flag	
Latitude	lat	
Longitude	lon	
CompressionSplice	compression_splice	
AutomaticSplice	automatic_splice	
Notification	sap_notification_id	
NormalPeakLoadingCyme	peak_loading	
SpliceLat	lat_splice	
SpliceLong	lon_splice	
FaultLatitude	lat_fault	
FaultLongitude	lon_fault	
HighFireThreatDistrict	hftd_dictrict	
Tree Wire	treewire_flag	

ILIS

Source fieldname	Dx-Risk fieldname	Notes

BASIC_CAUSE_DESC	cause	
CAUSE_DETAIL	free_text	
CONSTRUCTION_TYPE	line_type	
CUST_MINUTES	customer_minutes	
CUST_OUT	customers_out	
DIVISION	division	
END_DATE	dttm_end	
EVENT#	ilis_id	ID links to other datasets
EQUIP_COND_DESCRIPTION	equipment_condition	
EQUIP_COND_STATUS_DESCRIPTION	equipment	
FAULT_TYPE_DESC	fault_type	
FEEDER_NAME	feeder	
FNL	dttm_start	
MED	med_flag	
OIS_OUTAGE#	ois_id	Foriegn key to ILIS
OPEN_EQUIPMENT_TYPE	equipment_failed	
OPEN_POINT_LAT	lat	Location (protective device)
OPEN_POINT_LONG	long	Location (protective device)
OUTAGE_LEVEL_DESC	outage_equipment	
PHASES	n_phases	
PM_NUMBER	sap_id	
PREVIOUS_SWITCHING	previous_switching	
SUBSTATION_NAME	substation	
SUPPLEMENTAL_CAUSE_DESC	supplemental_cause	
SUS_MOM	momentary_flag	

UNPLAN_PLAN	planned_flag	
WIRE_DOWN	wiredown_flag	
WIRE_DOWN_HOT	hotwire_flag	

Conductors

arad_dev/edgis/vw_snapshot_priohconductor

Source fieldname	**Dx-Risk fieldname**	**Notes**
objectid	equipment_id	
installjobyear	install_year	
circuitid	circuit_id	
coastalidc	coastal_flag	
measuredlength	length	
jackettype	jacket_type	
operatingvoltage	voltage_operating	
constructiontype	construction	
windspeedcode	windspeed_rating	
windspeedrerateyear	windrating_year	
district	district	
division	division	
region	region	
installjobnumber	sap_id	
circuitname	feeder	
nominalvoltage	voltage_nominal	
shape	geometry	

arad_dev/edgis/vw_snapshot_priohconductorinfo

Source fieldname	Dx-Risk fieldname	Notes
objectid	equipment_id	
globalid	global_id	
conversionid	sap_id	
material	conductor_material	
conductorsize	conductor_size	
insulation	conductor_insulation	
conductortype	conductor_type	
installationdate	install_date	

arad_dev/edgis/vw_snapshot_secohconductor

Source fieldname	Dx-Risk fieldname	Notes
objectid	equipment_id	
installjobyear	install_year	
circuitid	circuit_id	
coastalidc	coastal_flag	
measuredlength	length	
jackettype	jacket_type	
operatingvoltage	voltage_operating	
constructiontype	construction	
windspeedcode	windspeed_rating	
windspeedrerateyear	windrating_year	
district	district	

division	division	
region	region	
installjobnumber	sap_id	
circuitname	feeder	
nominalvoltage	voltage_nominal	
shape	geometry	

arad_dev/edgis/vw_snapshot_sechconductorinfo

Source fieldname	**Dx-Risk fieldname**	**Notes**
objectid	equipment_id	
globalid	global_id	
conversionid	sap_id	
material	conductor_material	
conductorsize	conductor_size	
insulation	conductor_insulation	
conductortype	conductor_type	
installationdate	install_date	