

AI on the Grid
An NI4AI Workshop
October 28, 2020

Course Overview

- Day 2 of 2 day course
 - Day 1: PMU Fundamentals
 - Day 2: Intro to AI
- Logistics
 - Day 2 will use the NI4AI platform
 - Sign up for a login at ni4ai.org
 - Q&A in webex
 - Raise your hand to be unmuted

Agenda

Day 1

Understanding PMU Data

*Alexandra "Sascha" von Meier
(UC Berkeley)*

Lessons & Case Studies

*Kevin Jones
(Dominion Energy)*

Get Practice and Learn More

*Laurel Dunn
(NI4AI)*

Day 2

The Power of Data

*Sean Patrick Murphy
(PingThings)*

Interfacing with Sensor Data

*Chris Ryan
(PingThings)*

Use Cases & Analytics

*Mohini Bariya and Miles Rusch
(UC Berkeley)*

Panelists



Sean Murphy

CEO

PingThings



Chris Ryan

Data Scientist

PingThings



Mohini Bariya

PhD Candidate

UC Berkeley



Miles Rusch

PhD Candidate

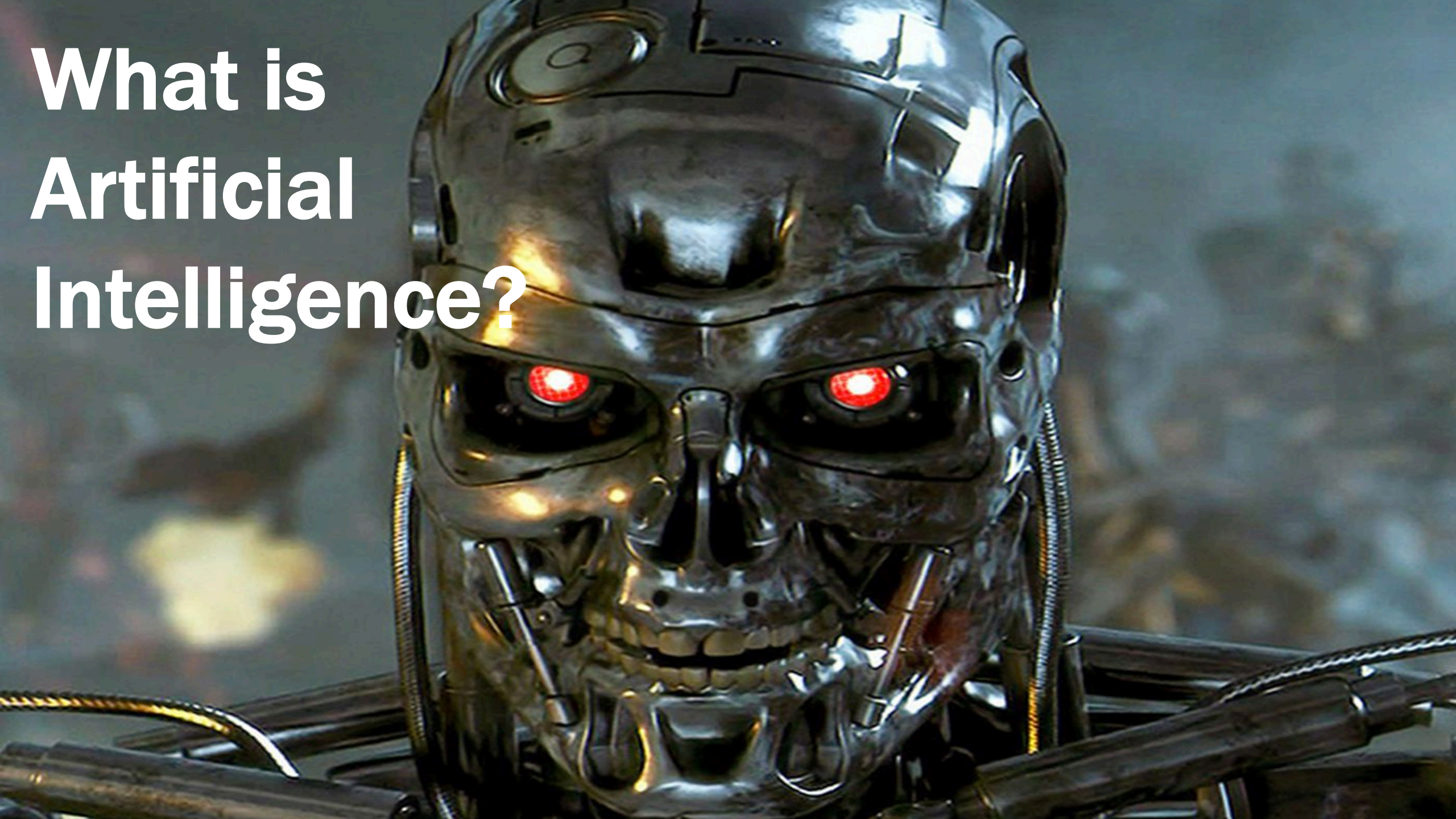
UC Berkeley

The Power of Data

Sean Patrick Murphy
PingThings



What is Artificial Intelligence?

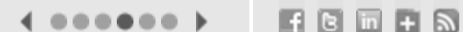


“The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

- The Oxford English Dictionary, 2017

BREAKING NEWS

NEWS & ANALYSIS: Coventor Touts Unified MEMS Platform



designlines INDUSTRIAL CONTROL

News & Analysis

Microsoft, Google Beat Humans at Image Recognition

Deep learning algorithms compete at ImageNet challenge

R. Colin Johnson

2/18/2015 08:15 AM EST

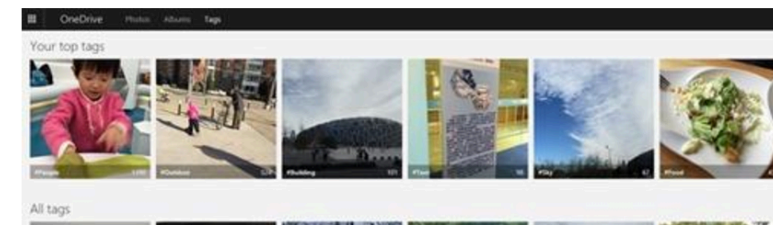
14 comments

NO RATINGS
1 saves
LOGIN TO RATE

Like 99 | Tweet | in Share 51 | G+1 44

PORTLAND, Ore. -- First computers beat the best of us at [chess](#), then [poker](#), and finally [Jeopardy](#). The next hurdle is image recognition — surely a computer can't do that as well as a human. Check that one off the list, too. Now Microsoft has programmed the first computer to beat the humans at image recognition.

The competition is fierce, with the [ImageNet Large Scale Visual Recognition Challenge](#) doing the judging for the 2015 championship on December 17. Between now and then expect to see a stream of papers claiming they have one-upped humans too. For instance, only 5 days after Microsoft announced it had beat the human benchmark of 5.1% errors with a 4.94% error grabbing neural network, Google announced it had one-upped Microsoft by 0.04%.



Most Recent Comments



resistion Samsung System LSI will become fabless customer of Samsung Foundry, competing with Qualcomm and Apple.

5/16/2017
9:49:31 PM

Navigate to Related Links

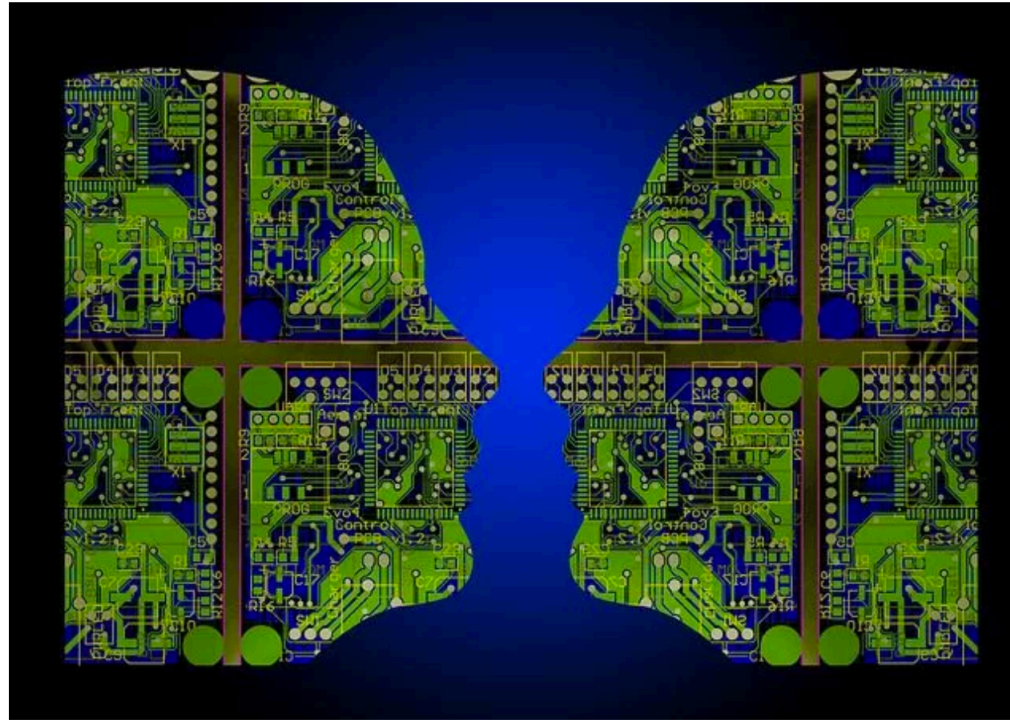
- [Home Alarms Ring Up Sales](#)
- [Businesses Set High Expectations For AI](#)
- [Q&A: BLDC Motor Control Class Questions](#)
- [Trinamic Surfing 'Automation' Wave](#)
- [Day 3: Field Oriented Control](#)



Artificially Intelligent Computer Outperforms Humans on IQ Test

The deep learning machine can reach the intelligence level between people with bachelor degrees and people with master degrees

By Sage Lazzaro • 06/15/15 11:46am



(Photo: Pixabay.com) Pixabay

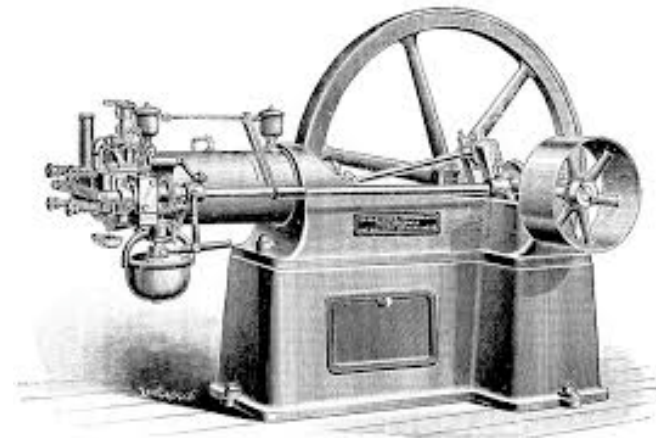
SEE ALSO: [This Website Will Rate Your Attractiveness Using Artificial Intelligence](#)

For decades, the Intelligence Quotient (IQ) has had numerous uses for humans but little importance when it comes to computers. With the focus and importance of AI research increasing, it was only a matter of time before we used this measure of intelligence to compare humans and machines.

<https://www.youtube.com/watch?v=V1eYniJ0Rnk>



Per Economists,
AI is a
**Fundamental Technological
Advancement**

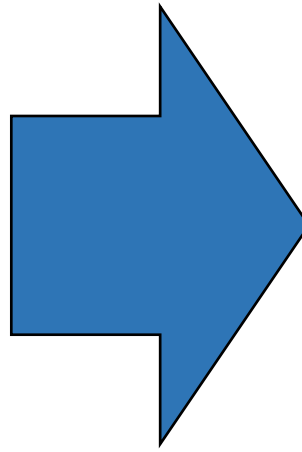




The Cost of Calculations Fell

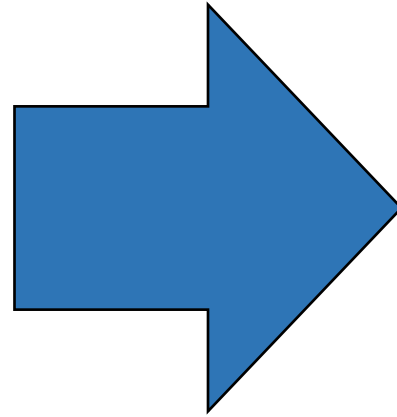
Trigonometric Functions

	sin	cos	tan	cot	sec	csc	
0°	0.0000	1.0000	0.0000	...	1.000	...	90°
1°	0.0175	0.9998	0.0175	57.29	1.000	57.30	89°
2°	0.0349	0.9994	0.0349	28.64	1.001	28.65	88°
3°	0.0523	0.9986	0.0524	19.08	1.001	19.11	87°
4°	0.0698	0.9976	0.0699	14.30	1.002	14.34	86°
5°	0.0872	0.9962	0.0875	11.43	1.004	11.47	85°
6°	0.1045	0.9945	0.1051	9.514	1.006	9.567	84°
7°	0.1219	0.9925	0.1228	8.144	1.008	8.206	83°
8°	0.1392	0.9903	0.1405	7.115	1.010	7.185	82°
9°	0.1564	0.9877	0.1584	6.314	1.012	6.392	81°
10°	0.1736	0.9848	0.1763	5.671	1.015	5.759	80°
11°	0.1908	0.9816	0.1944	5.145	1.019	5.241	79°
12°	0.2079	0.9781	0.2126	4.705	1.022	4.810	78°
13°	0.2250	0.9744	0.2309	4.331	1.026	4.445	77°
14°	0.2419	0.9703	0.2493	4.011	1.031	4.134	76°
15°	0.2588	0.9659	0.2679	3.732	1.035	3.864	75°
16°	0.2756	0.9613	0.2867	3.487	1.040	3.628	74°
17°	0.2924	0.9563	0.3057	3.271	1.046	3.420	73°
18°	0.3090	0.9511	0.3249	3.078	1.051	3.236	72°
19°	0.3256	0.9455	0.3443	2.904	1.058	3.072	71°
20°	0.3420	0.9397	0.3640	2.747	1.064	2.924	70°
21°	0.3584	0.9336	0.3839	2.605	1.071	2.790	69°
22°	0.3746	0.9272	0.4040	2.475	1.079	2.669	68°
23°	0.3907	0.9205	0.4245	2.356	1.086	2.559	67°
24°	0.4067	0.9135	0.4452	2.246	1.095	2.459	66°
25°	0.4226	0.9063	0.4663	2.145	1.103	2.366	65°
26°	0.4384	0.8988	0.4877	2.050	1.113	2.281	64°
27°	0.4540	0.8910	0.5095	1.963	1.122	2.203	63°
28°	0.4695	0.8829	0.5317	1.881	1.133	2.130	62°
29°	0.4848	0.8746	0.5543	1.804	1.143	2.063	61°
30°	0.5000	0.8660	0.5774	1.732	1.155	2.000	60°
31°	0.5150	0.8572	0.6009	1.664	1.167	1.942	59°
32°	0.5299	0.8480	0.6249	1.600	1.179	1.887	58°
33°	0.5446	0.8387	0.6494	1.540	1.192	1.836	57°
34°	0.5592	0.8290	0.6745	1.483	1.206	1.788	56°
35°	0.5736	0.8192	0.7002	1.428	1.221	1.743	55°
36°	0.5878	0.8090	0.7265	1.376	1.236	1.701	54°
37°	0.6018	0.7986	0.7536	1.327	1.252	1.662	53°
38°	0.6157	0.7880	0.7813	1.280	1.269	1.624	52°
39°	0.6293	0.7771	0.8098	1.235	1.287	1.589	51°
40°	0.6428	0.7660	0.8391	1.192	1.305	1.556	50°
41°	0.6561	0.7547	0.8693	1.150	1.325	1.524	49°
42°	0.6691	0.7431	0.9004	1.111	1.346	1.494	48°
43°	0.6820	0.7314	0.9325	1.072	1.367	1.466	47°
44°	0.6947	0.7193	0.9657	1.036	1.390	1.440	46°
45°	0.7071	0.7071	1.000	1.000	1.414	1.414	45°
	cos	sin	cot	tan	csc	sec	





Problems Recast as Calculations



What does AI make cheaper?

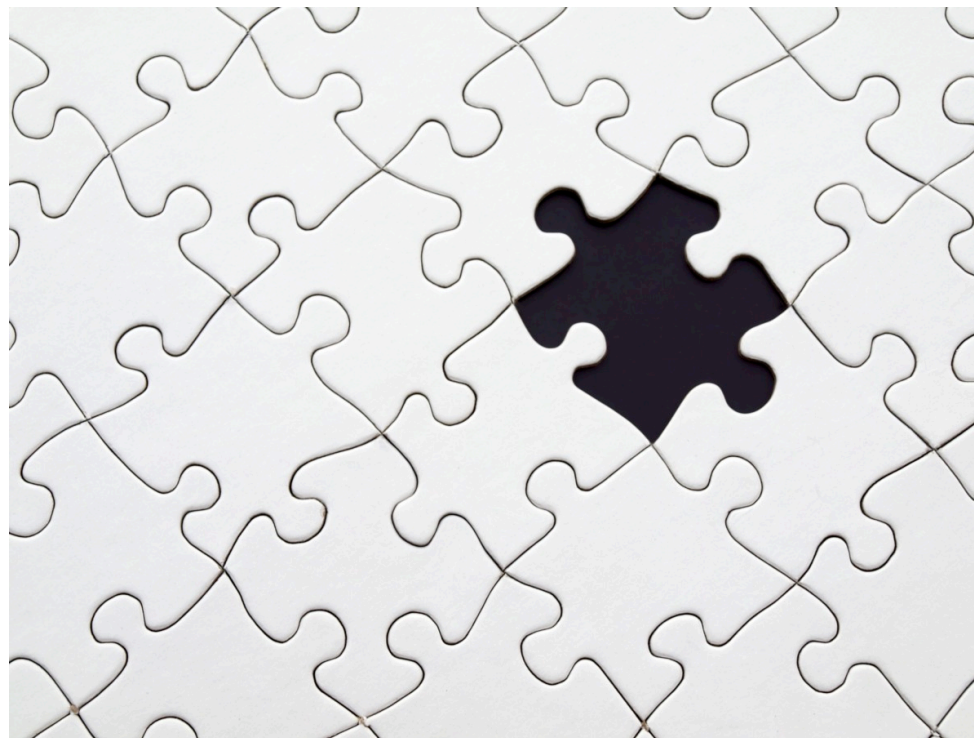
“The task that AI makes abundant and inexpensive is prediction — in other words, the ability to take information you have and generate information you didn’t previously have.”



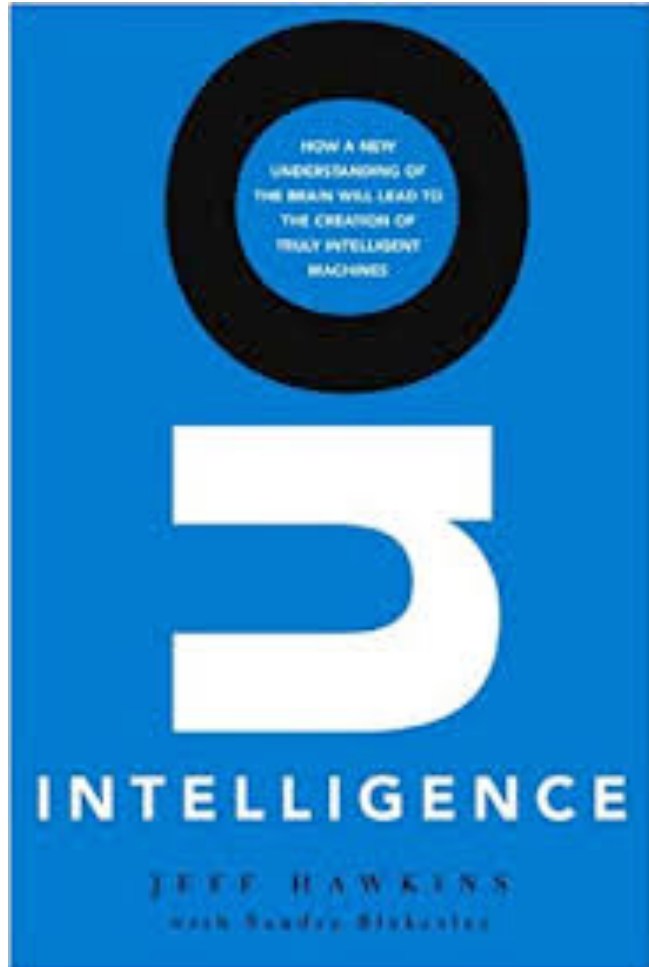
Forecasting



Resolving Ambiguity



Miss ng Data




“We are making continuous low-level predictions in parallel across all our senses. But that’s not all. I am arguing a much stronger proposition. Prediction is not just one of the things your brain does. It is the primary function of the neo cortex, and the foundation of intelligence”

**1.Improve quality of
goods/services that depend on
prediction**


**2.Problems get recast as
prediction problems**


**3.As prediction gets better,
unforeseen possibilities arise**


 Compose


Back Archive Spam Delete Mark as unread Snooze Move to Labels More

 **Inbox** 2


 Snoozed


 Chats


 Sent


 **Drafts** 10

 All Mail

 **Spam** 32


 Trash

▾  Categories

 Social

 **Updates** 3

 **Forums** 1

 Sean ▾ +

(no subject) Inbox x Life/family x



jandcmurphy@verizon.net via yahoo.com 12:05 AM (9 hours ago) ☆ ↶ Reply ⋮

to me ▾



Tonight, we had 3 raccoons, 2 foxes, and 1 ugly animal that I can't remember it's name.

All over the yard.

So cute!

Wow!

Very cool!

↶ Reply

➦ Forward



Computer
History
Museum

Google

6WRE5

Google

self-driving car

Google™



SEE SOMETHING NEW, EVERY DAY.

TAKE A LOOK



All

Prime Video Stream movies & TV show

Deliver to Sean
Silver Sp... 20910

Buy Again Your Pickup Location

EN

Hello, Sean
Account & Lists

Orders

Prime



"Alexa, play my '90s pop playlist on Amazon Music"

Bose QuietComfort 35 II



Historical Data Display



Hi, Sean

Customer since 2010

Top links for you



Your Orders



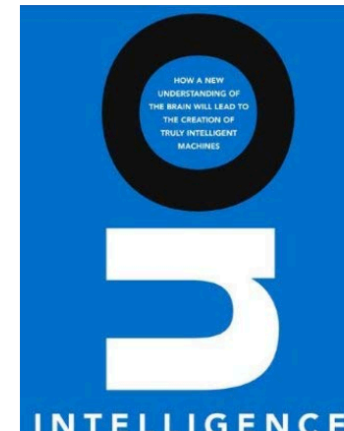
Tools & Home Improvement

<https://www.amazon.com/#>

Continue watching



Recently viewed



Real Time Use Case

Predictive Use Cases

The screenshot displays the Amazon.com homepage with a focus on predictive recommendations. At the top, the browser address bar shows the URL http://www.amazon.com/ref=nav_log. The main content is divided into three horizontal sections:

- Video: Recommended for you:** Features a large banner for **Deadpool 2** with the title in large red letters and a "See more from Prime Video" link at the bottom.
- The Widow:** A promotional card for the Prime Video series, featuring a portrait of Hailee Steinfeld and a "Watch now" link.
- Audiobook for you:** A card for the audiobook **Thinking, Fast and Slow** by Daniel Kahneman, with a "See more FREE from Audible" link.

Below these sections is a category titled **Inspired by your shopping trends**, which displays a carousel of household products including:

- Cottonelle toilet paper
- Bounty paper towels
- Unstoppable disinfectant spray
- Cascade Platinum Plus dishwasher detergent
- Q-tips swabs
- Viva paper towels

The final section is **New items to stock up on**, featuring a carousel of health and wellness products:

- Emergen-C Immune+ dietary supplement
- Optimum Nutrition Pro Whey Protein
- Quest Protein Bars
- Oral-B electric toothbrushes
- Gold Standard 100% Whey Protein
- Nutricost Vitamin B2 supplement

A Pragmatic Definition of AI

Artificial Intelligence is

the cost-effective

use of data

at **scale**

to use prediction

to solve real world problems.

**Expected
value must
exceed
cost.**

An Example

4 measurements per hour

24 hours per day

365.25 days per year

3 measurements per meter (V, I, P)

X

10M customers

1,051,920,000,000 points or 1 terapoint

An Example

60 measurements per hour

24 hours per day

365.25 days per year

3 measurements per meter (V, I, P)

X

10M customers

15,778,800,000,000 points or 15.8 terapoints

An Example

60 measurements per minute

60 minutes per hour

24 hours per day

365.25 days per year

3 measurements per meter (V, I, P)

10M customers

X

946,728,000,000,000 points or 947 terapoints

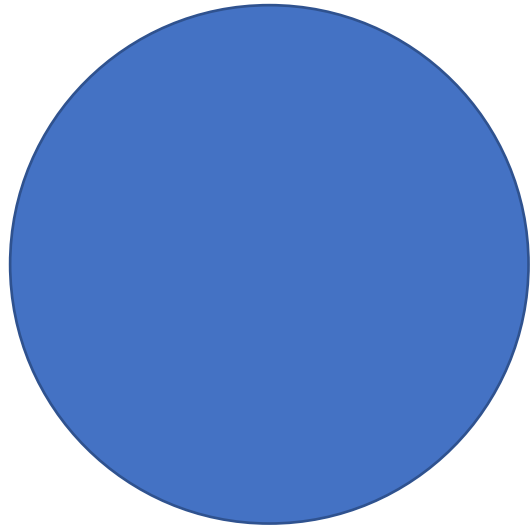
Total Annual Cost of Storing Data per Customer

Every 15 minutes - \$0.000465

Every minute - \$0.0070

Every second - \$0.43

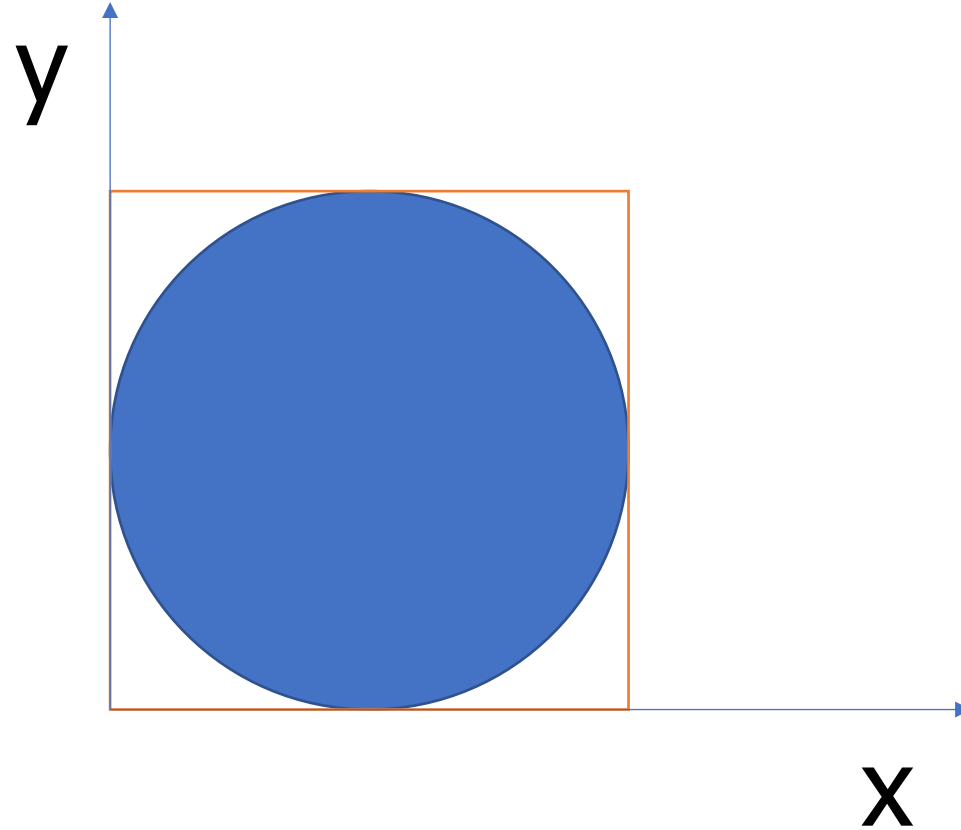
What is the Area of a Circle?



What is the Area of a Circle?

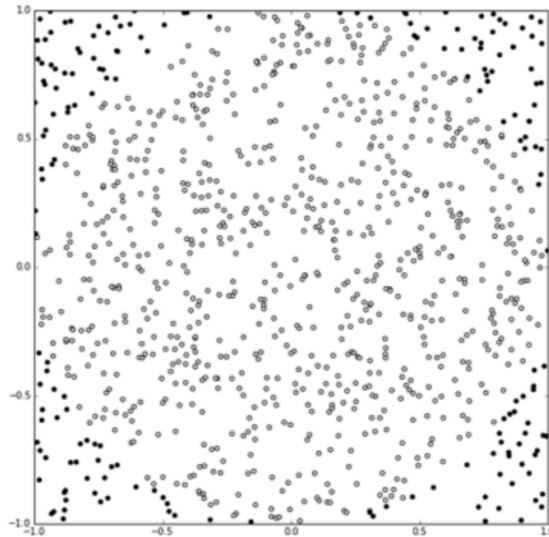
$$A = \pi r^2$$

What is the Area of a Circle?



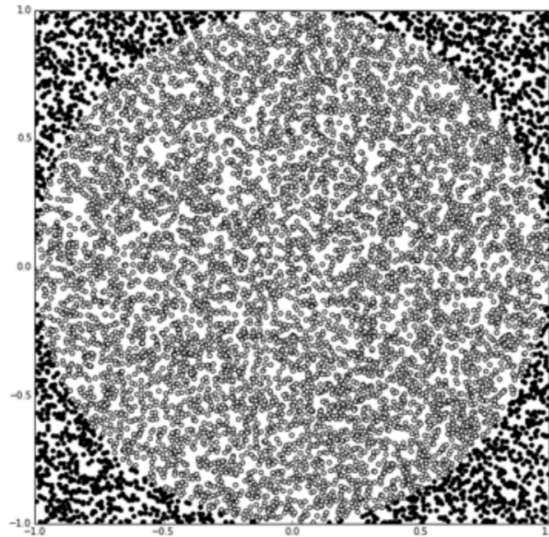
What is the Area of a Circle?

n = 1,000
Error = 0.118



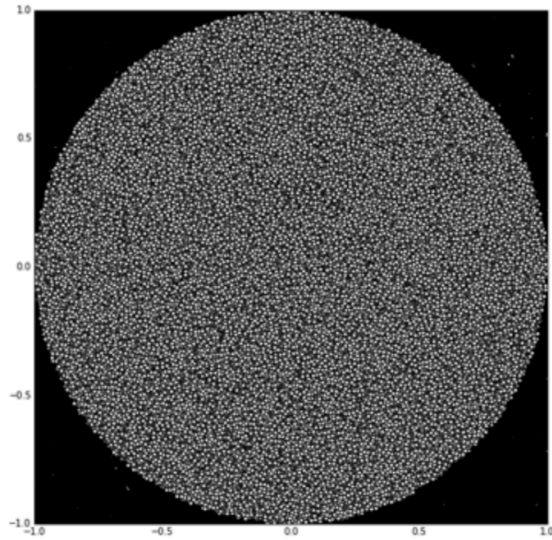
What is the Area of a Circle?

n = 10,000
Error = 0.021

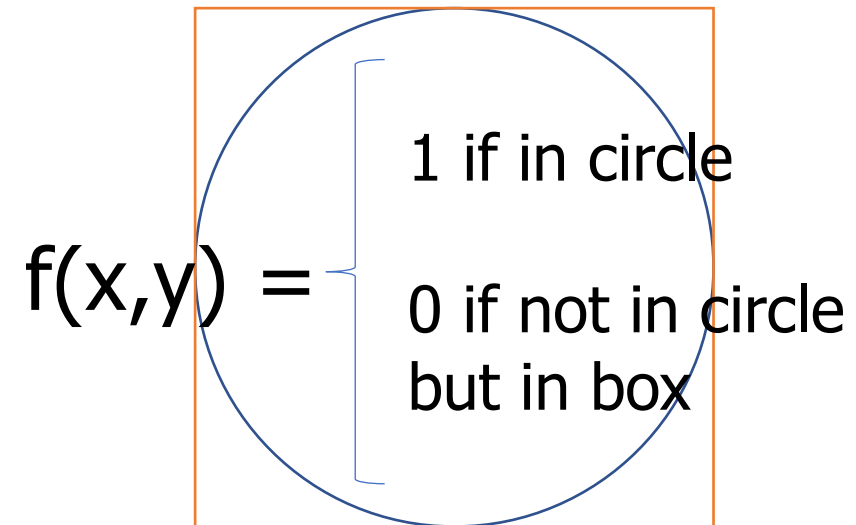


What is the Area of a Circle?

$n = 100,000$
Error = 0.008



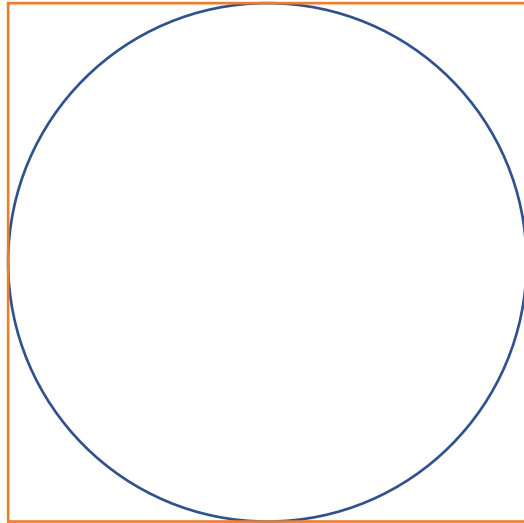
What is the Area of a Circle?



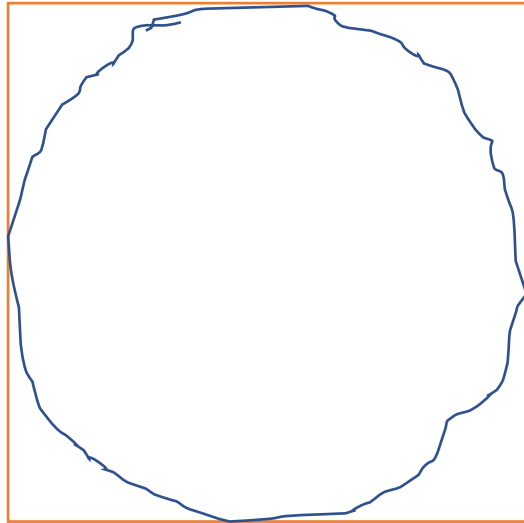
The diagram shows a blue circle centered within an orange square. To the left of the circle, the text $f(x,y) =$ is followed by a large right-facing curly bracket. The top part of the bracket is aligned with the top edge of the circle and contains the text "1 if in circle". The bottom part of the bracket is aligned with the bottom edge of the circle and contains the text "0 if not in circle but in box".

$$f(x,y) = \begin{cases} 1 & \text{if in circle} \\ 0 & \text{if not in circle} \\ & \text{but in box} \end{cases}$$

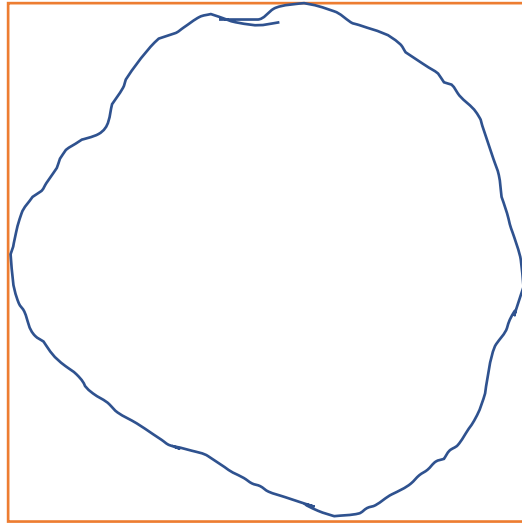
What is the Area of a Circle?



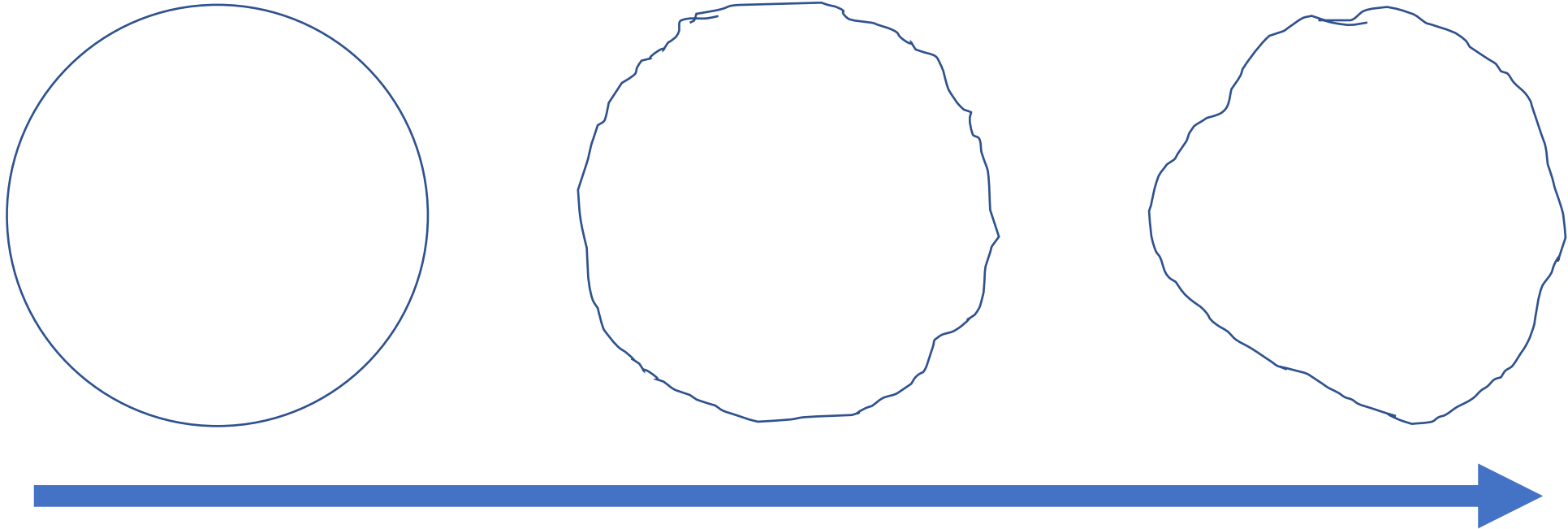
What is the Area of a Circle?



What is the Area of a Circle?

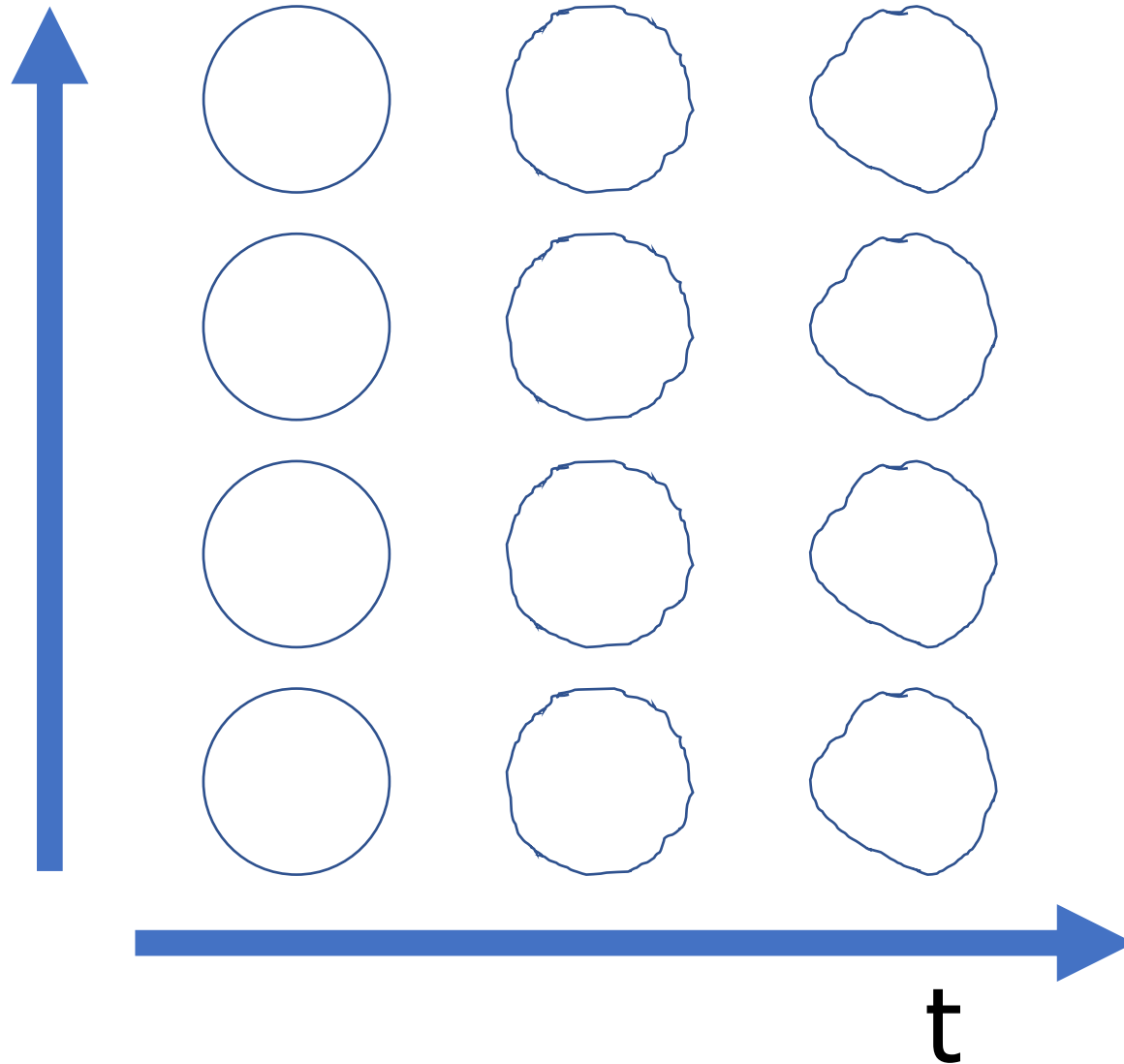


What is the Area of a Circle?



$g(t) = A$ of the Shape

What is the Area of a Circle?



Interfacing with sensor data in Python

Chris Ryan, PhD

PingThings



PingThings: Platform Summary

A number of stand alone tools have been created which make up the PredictiveGrid platform and leverage the performance of the BTrDB database.

These tools and applications provide additional value by offering pre-built services rather than having to build from the ground up.

.

Examples

- Plotter visualization website
- **On demand Jupyter notebook servers**
- File based data ingestion tools (parquet, csv, etc.)_
- DISTIL real time data processing
- And more

Data Model: Collections

Collections are used to **organize streams**.

They are **analogous to directories** in a file system. In the same way directories contain files, collections contain streams.

Collections can be **arbitrarily deep** (though the collection path is treated as a unique string).

 **ameren/METER_64301** V Stream I Stream P Stream

Data Model: Streams

Streams contain time series data. They hold data points which are time/value pairs

Each stream must contain:


- unique identifier (UUID)
- collection
- name
- unit of measure

You can query for raw values or you can query for windows of data which contain aggregate information.

 ameren/METER_64301

 V Stream

 I Stream

 P Stream

Data Model: Stream Metadata

Each stream is able to contain **rich metadata**.

Annotations are the primary means for users to manage metadata values.

Tags are used internally by the system.

Data and metadata are **separately versioned** allowing you to view historical data.

```
{
  "annotations": {
    "dataIntervalType": "Average 1-Second",
    "dataType": "variation",
    "digitalSampling": "100 Hz",
    "element": "H",
    "elevation": "424",
    "latitude": "18.113",
    "longitude": "293.849",
    "orientation": "HDZF",
    "reported": "HDZF",
    "source": "US Geological Survey",
    "station": "San Juan"
  },
  "collection": "USGS/GEOMAG/San_Juan",
  "tags": {
    "distiller": "",
    "ingress": "",
    "name": "SJG_H",
    "unit": "nanotesla"
  },
  "uuid": "84960047-a03a-40b3-b795-8c8826e1ec0b"
}
```


Data Model: Points

Raw time series values are retrieved as objects known as **RawPoints**. A stream may typically consist of many billions of RawPoint instances.

These objects contain a **time** and **value** property.

The platform treats all times in nanoseconds (since epoch).

```
>>> point  
RawPoint(1542953940000000000, -200.0009844)
```

```
>>> point.time  
1542953940000000000
```

```
>>> point.value  
-200.0009844
```

Data Model: Points

Windows of data are retrieved as **StatPoints**. The objects contain aggregate information over the time window specified.

You can view: **time**, **min**, **mean**, **max**, **count**, and **stddev**.

Aggregate (windowed) queries are exceptionally fast - **Example: you don't have to read 100 billion points to find the mean of 100 billion points**

```
>>> point
StatPoint(1542953940000000000, -910.5307974,
          85.01138333624999, 1110.5317818,
          960, 684.2895437798793)
```

```
>>> point.time
1542953940000000000
```

```
>>> point.count
960
```

```
>>> point.min
-910.5307974
```

```
>>> point.mean
85.01138333624999
```

```
>>> point.stddev
684.2895437798793
```

Python API

What do we want in an API abstraction?

Query a Range

(uuid, start, end, version) → (version, [<time, value>, ...])

Statistical Windows

(uuid, start, end, version, window_size) → (version, [<time, stats>, ...])

Insert Values

(uuid, [<time, value>, ...]) → version

Delete a Range

(uuid, start, end, version) → (version, [<time, value>, ...])

Compute Diff

(uuid, from_version, to_version, version, resolution) → <time range>

*PingThings provides all this
at high speed and scale*

Ecosystem: Plotter

The plotter is our primary **visualization website** allowing human exploration of the data.

You can seamlessly scroll from viewing data at the nanosecond scale up to decades of data.

Researchers can quickly hone in on trouble spots.

The screenshot shows the PingThings Plotter interface. At the top, there are tabs for "Stream Selection" and "Stream Visualization". Below the tabs, there are two tables: "Selected Streams" and "All Streams". The "Selected Streams" table has one row with the following data:

COLLECTION	NAME	UNIT	ACTION
ameren/METER_10072967	V	V	+

The "All Streams" table has multiple rows, with the last row highlighted in yellow:

COLLECTION	NAME	UNIT	ACTION
ameren/METER_10072965	I	I	+
ameren/METER_10072965	V	V	+
ameren/METER_10072966	I	I	+
ameren/METER_10072966	V	V	+
ameren/METER_10072967	I	I	+
ameren/METER_10072967	V	V	+

Below the "All Streams" table, there is a pagination control showing "Page 1 of 164762" and a "20 rows" dropdown menu. To the right of the tables is a line chart showing data over time. The chart has a y-axis ranging from 243 to 251 and an x-axis showing dates from Friday, June 26, 2020, to Wednesday, June 30, 2020. The chart shows a fluctuating blue line representing data points. Below the chart, there is a JSON metadata block:

```
{
  "annotations": {
    "sample_rate": "1"
  },
  "collection": "ameren/METER_10072967",
  "property_version": 3,
  "tags": {
    "distiller": "",
    "ingress": "",
    "name": "v",
    "unit": "v"
  },
  "uuid": "f624e7f1-54a3-4b3b-968d-22e0698445a1"
}
```

At the bottom right corner of the interface, the version number "v5.10.4 (1325380d)" is displayed.

Typical Analytics Workflow

Most users tend to fall into a specific workflow to perform analytical tasks.

1. Review data in plotter
2. Perform analysis in notebooks
3. Store / Report on results

PMU

PMU

PMU

PMU

PMU



Stream/Time Range Selection via Plotter



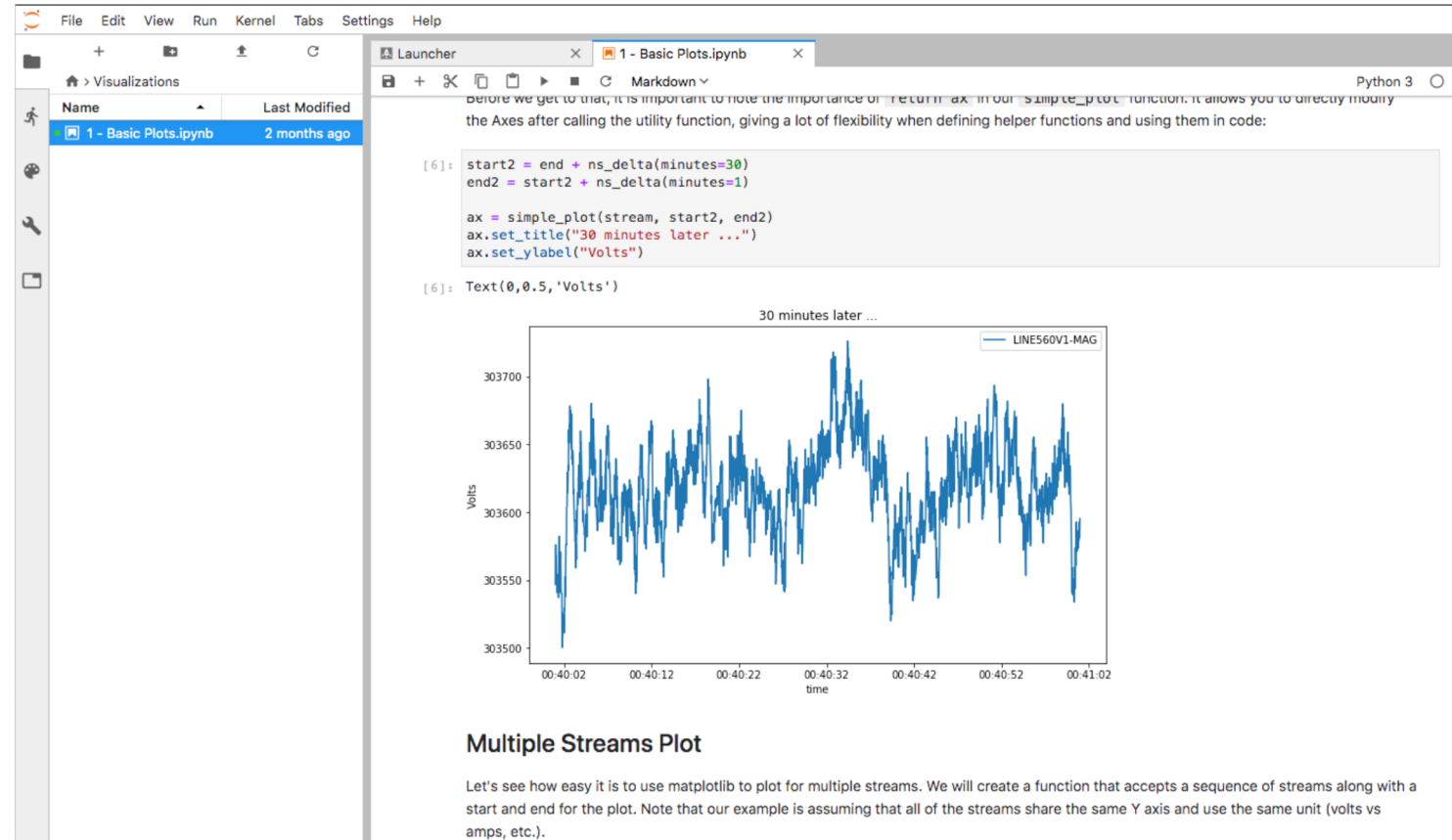
Execute Analytics and Store Results

Export Identified Fields to Notebook

Ecosystem: Jupyterhub

Jupyter notebook servers allow for **data locality** and **individual notebooks** for research.

Each user can choose from multiple server size options if they need more or less RAM, CPU, etc.



Analysis using smart meter data

simple example

“Example stream selection” demo

Analysis using smart meter data

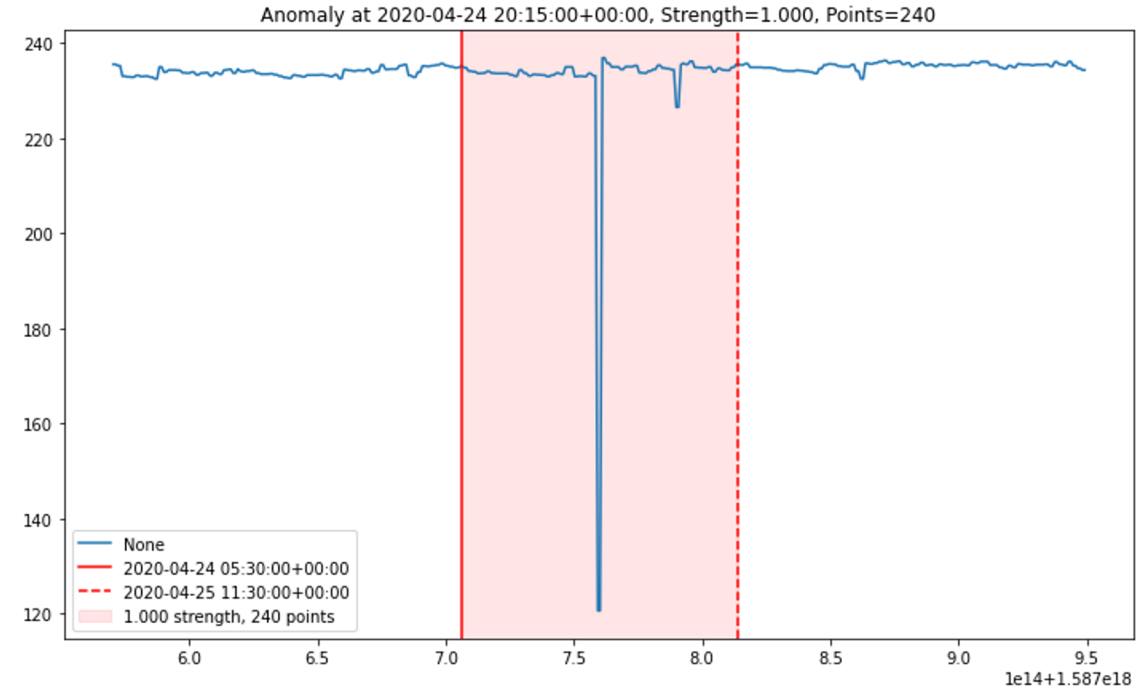
realistic example

“Multiple stream selection” demo

Other examples

df_gaps_serial

	stream	start	end	duration (seconds)	expected	actual	stream_name
0	069abd3b7783414bbb2fdd9689770bee	2020-05-30 09:45	2020-05-31 06:00	72900.0	81	0	ameren/METER_5736423/P15
1	17016815e31d4c36a654be506cc334f9	2020-05-30 10:15	2020-05-31 14:30	101700.0	113	0	ameren/METER_5568228/P15
2	1eba4971902746f28a4455492017b16e	2020-05-30 09:45	2020-05-31 14:00	101700.0	113	0	ameren/METER_5524171/P15
3	206c10dfa68e4f129c1252c67ad8c4a6	2020-05-30 10:30	2020-05-31 14:45	101700.0	113	0	ameren/METER_5523250/I
4	670b0c9c67144c5c94bf8838d89cccee	2020-05-30 11:45	2020-05-31 12:00	87300.0	97	0	ameren/METER_5500806/P15
5	686d9a6cd4d6417582f8cff4992ce3c	2020-05-30 10:45	2020-05-31 15:00	101700.0	113	0	ameren/METER_5523125/I
6	6ab1230457bc43c98ed83ec07f042a99	2020-05-30 10:45	2020-05-31 15:00	101700.0	113	0	ameren/METER_5523125/P15
7	79fe27465da4478cbadbe04346c59c4f	2020-05-30 10:15	2020-05-31 14:30	101700.0	113	0	ameren/METER_5729333/P15
8	84332093cf9447b2aef65d38481c5cf9	2020-05-30 09:45	2020-05-31 14:00	101700.0	113	0	ameren/METER_5524171/I



Thanks!

Any questions?

chris@pingthings.io

Voltage Sag Detection & Analysis

Mohini Bariya
UC Berkeley



The NI4AI platform makes it quick and easy to:

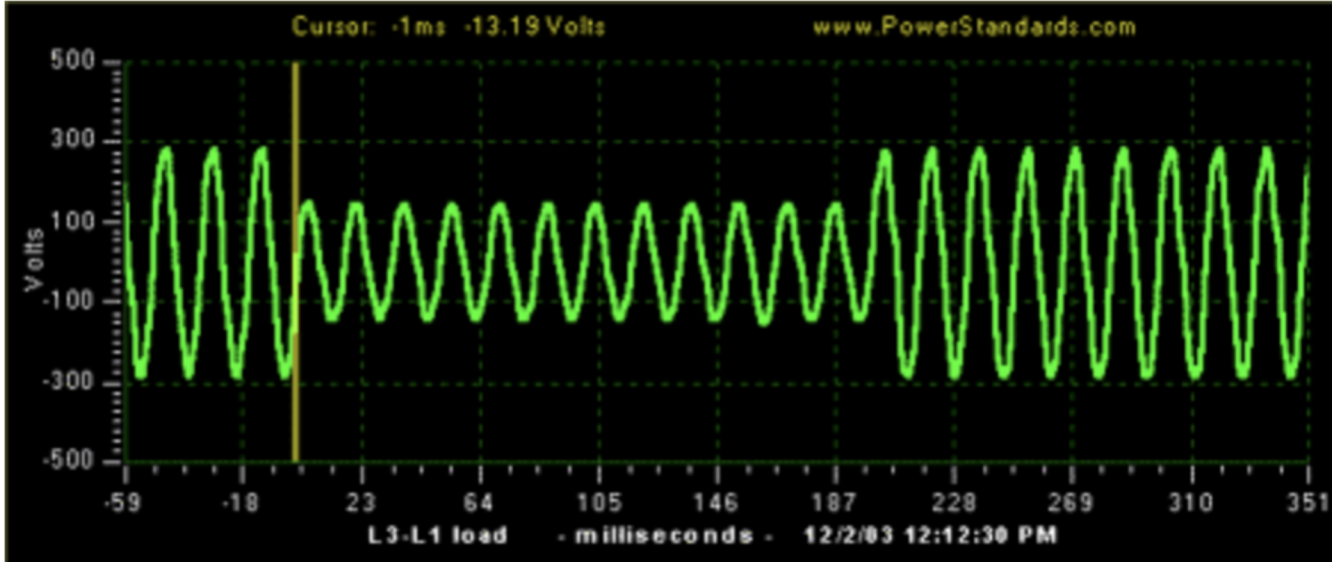
- Explore data
- Ask & answer questions
- Share results

I will demo the typical development process for one application: voltage sag detection and analysis.

Step 1: Motivation & Questions

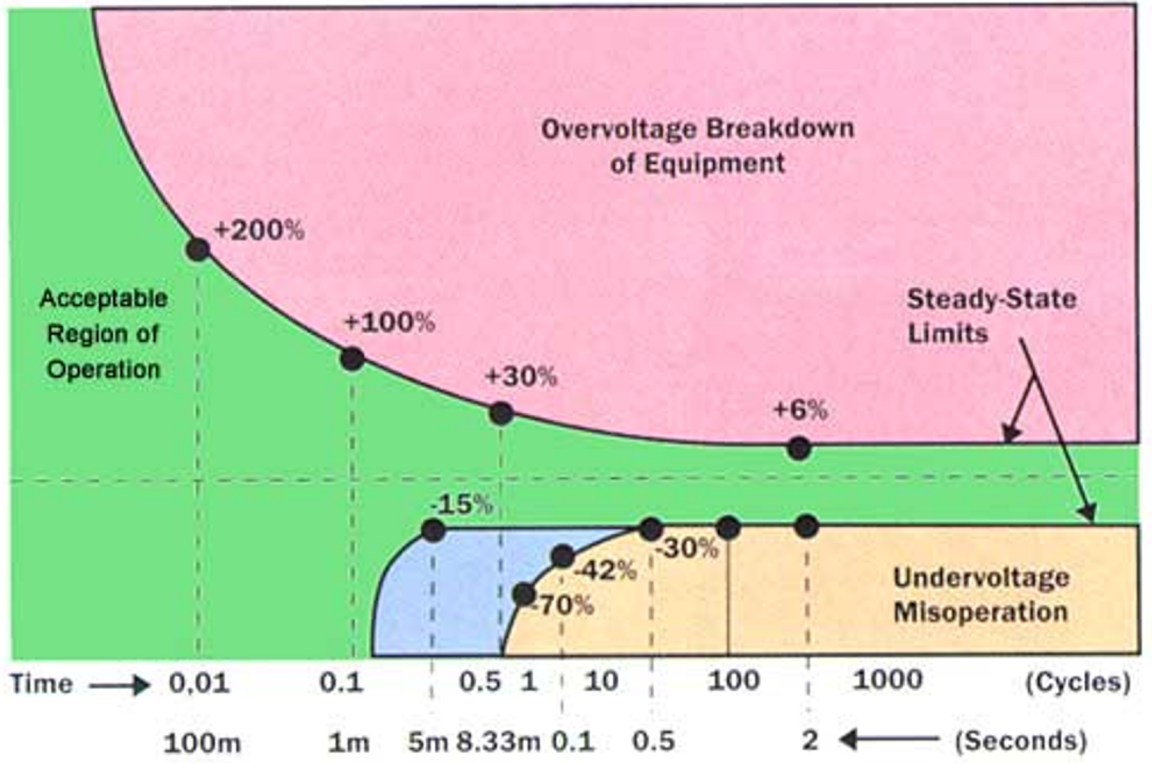
What are we interested in understanding? **Why** is this important?

- Voltage sags are brief, significant dips in system voltage.
- We want to understand the frequency & magnitude of sags.
- They can have a detrimental impact on sensitive equipment, causing inverters to trip offline, and can indicate safety issues such as faults.



A voltage sag in point-on-wave data [\[source\]](#)

rms voltage (% over or under rated)

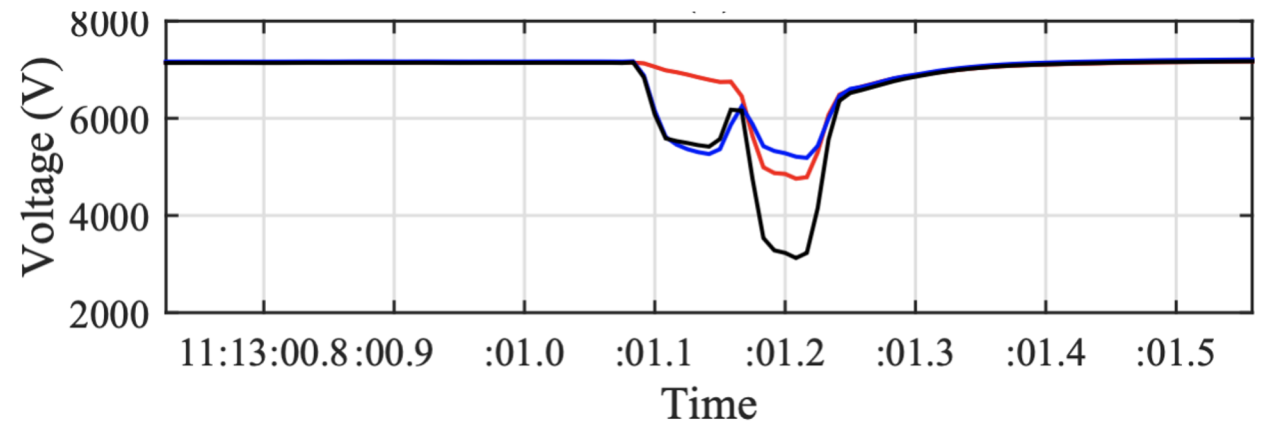
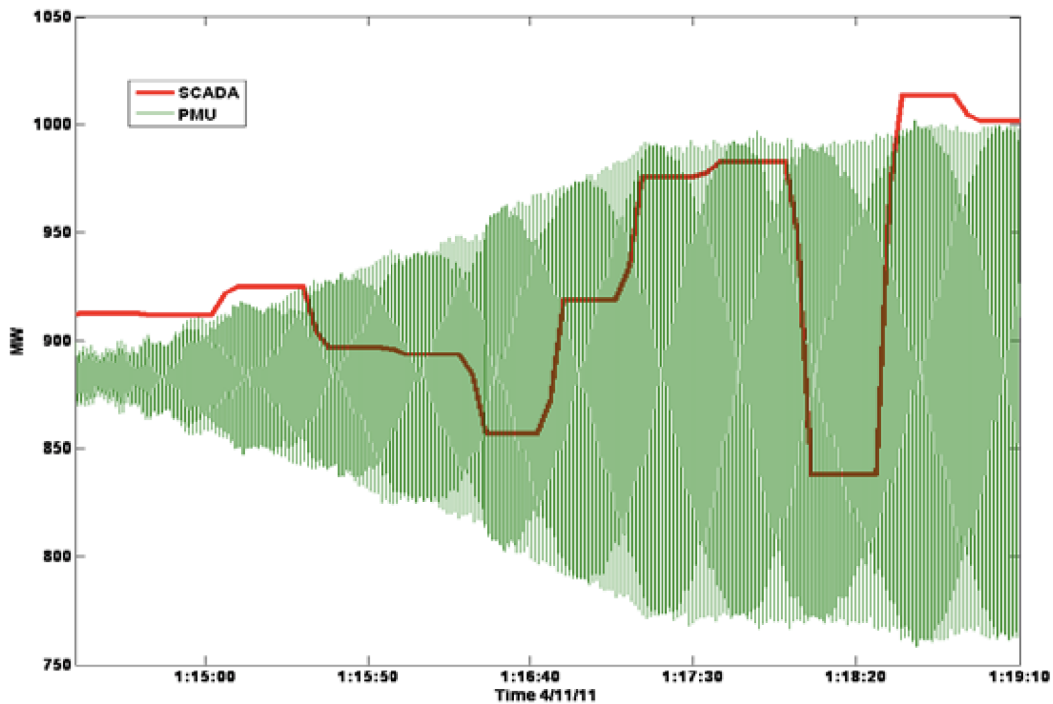


Acceptable voltage deviations for equipment

Step 2: Explore the data

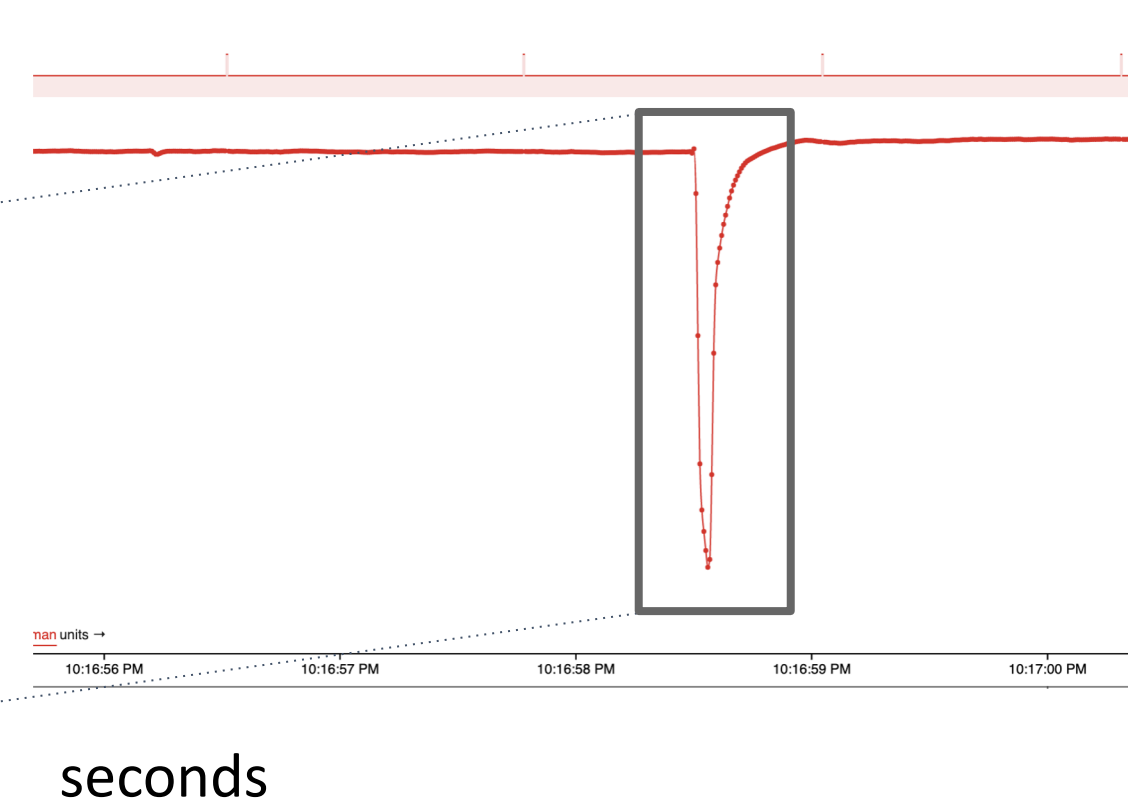
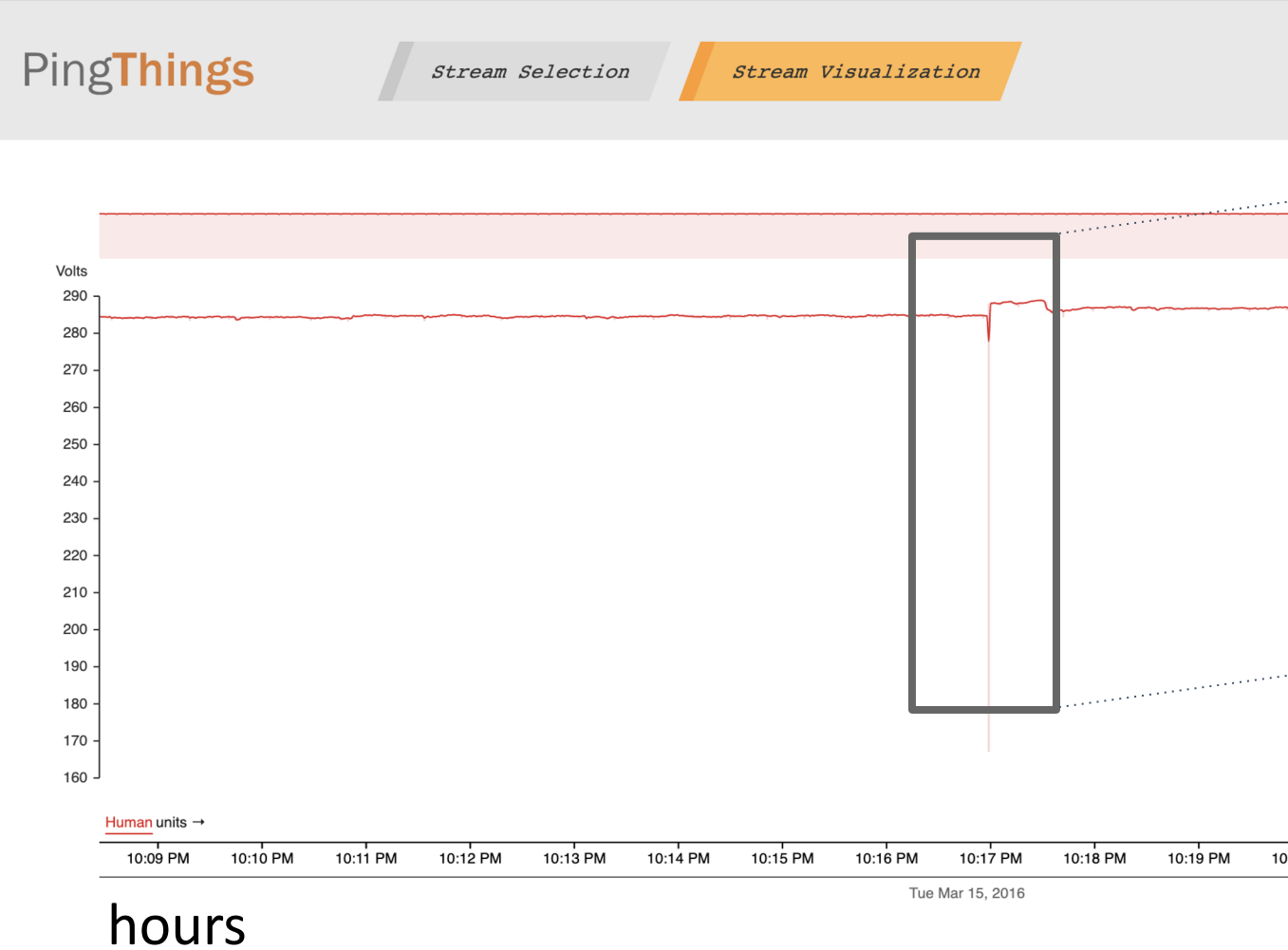
Exploration is critical for developing novel analytics.

- New measurement types
- New event types



Step 2: Explore the data

What do voltage sags look like? How can we automatically find them in the data?



Step 3: Automate sag detection

I can now write a script in a **Jupyter notebook** to automatically find voltage sags.

Covariance function

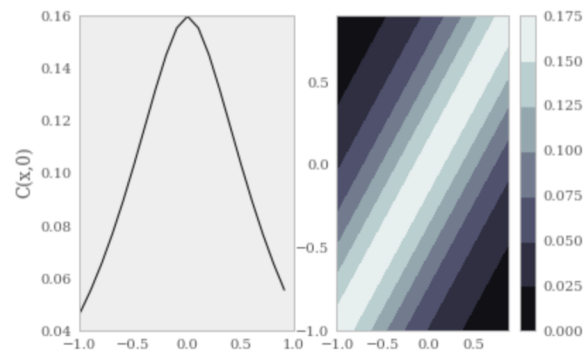
The behavior of individual realizations from the GP is governed by the covariance function. The Matérn class of functions is a flexible choice.

```
In [34]: from pymc.gp.cov_funs import matern
import numpy as np
C = Covariance(eval_fun=matern.euclidean, diff_degree=1.4, amp=0.4, scale=1, rank_limit=1000)

subplot(1,2,2)
contourf(x, x, C(x,x).view(ndarray), origin='lower', extent=(-1,1,-1,1), cmap=cm.bone)
colorbar()

subplot(1,2,1)
plot(x, C(x,0).view(ndarray), 'k-')
ylabel('C(x,0)')
```

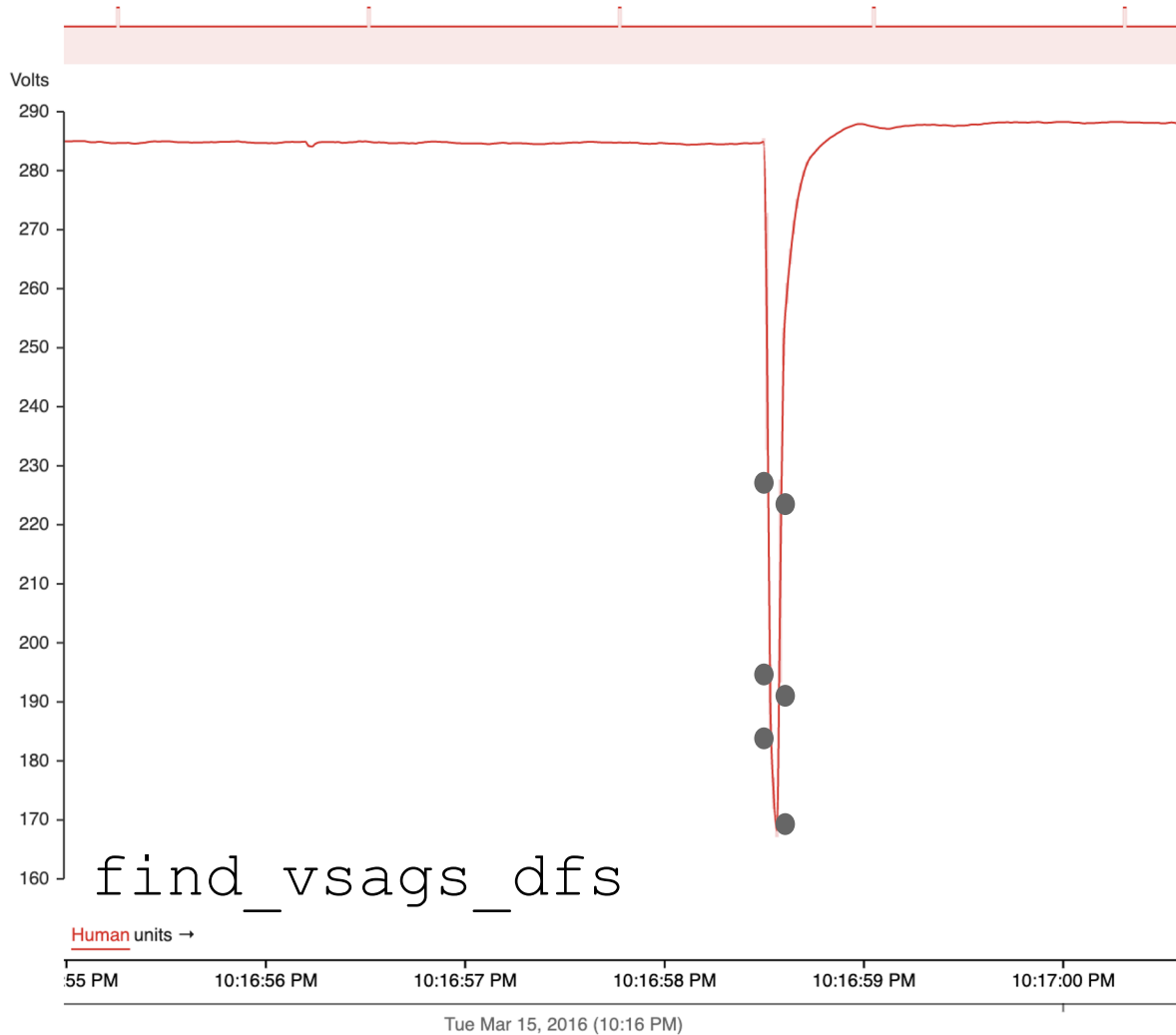
Out[34]: <matplotlib.text.Text at 0x112713290>



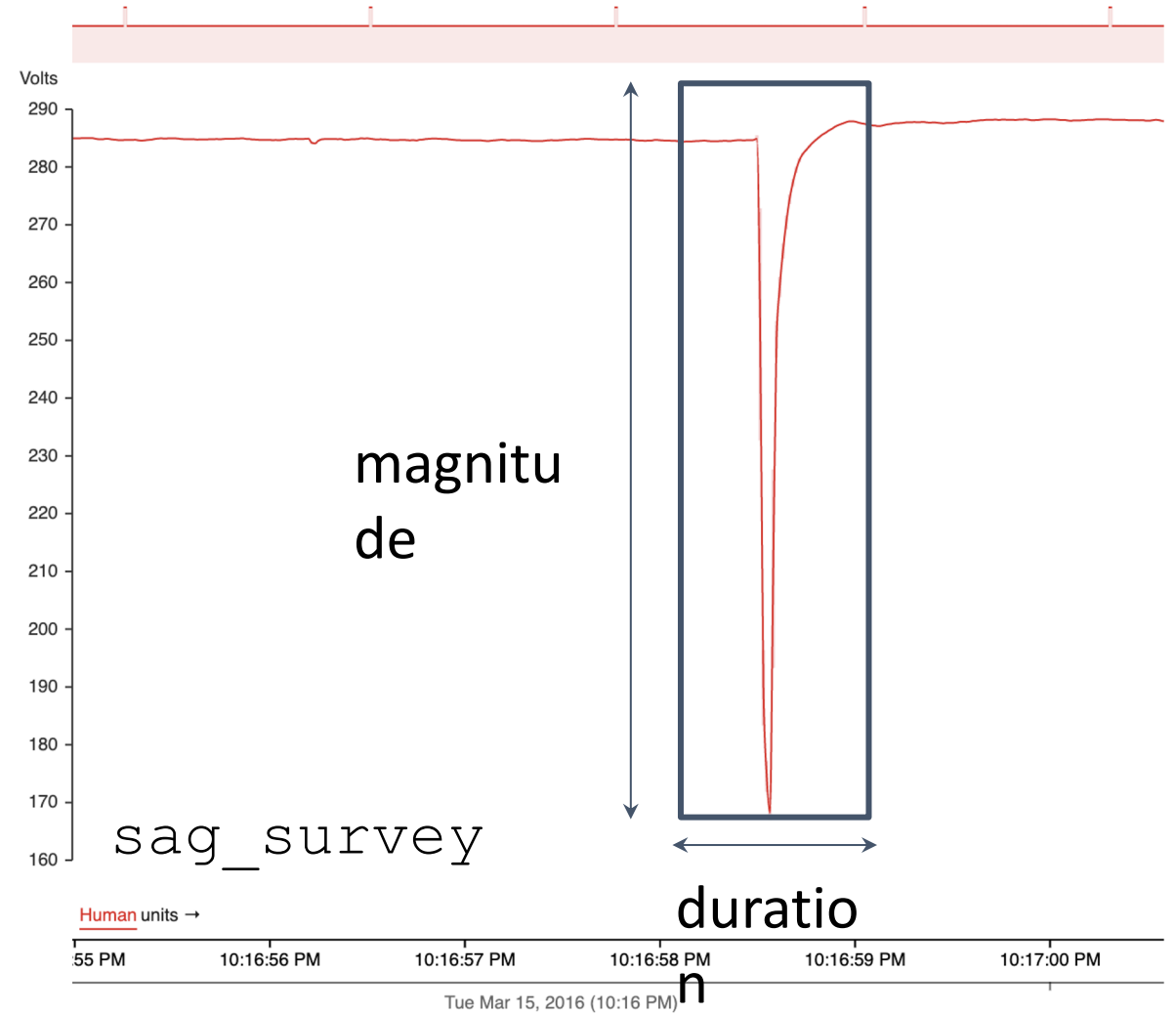
Jupyter notebooks allow for mixing of text, code, outputs including visualizations.

Enables exploration and

A sag is defined by voltage measurements below a threshold. I choose to split sag detection into two functions

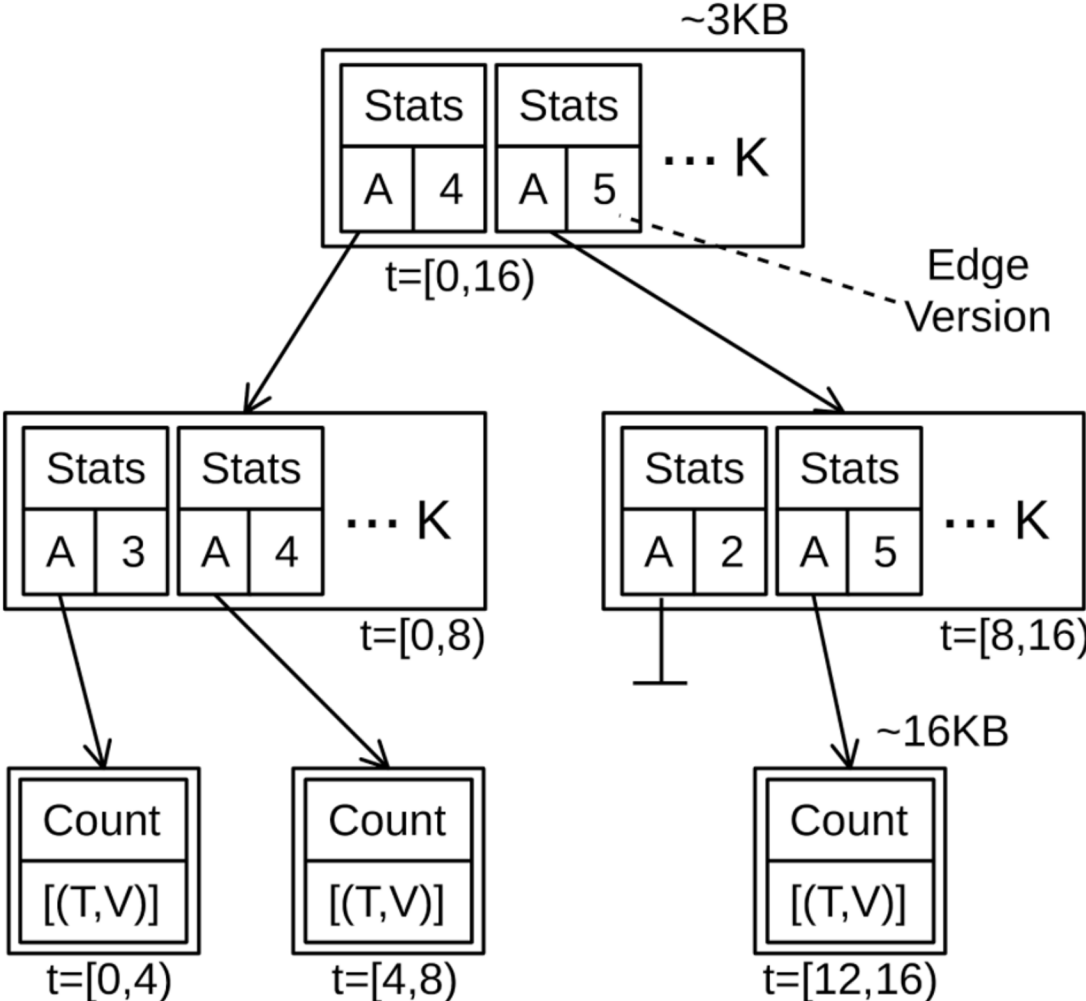


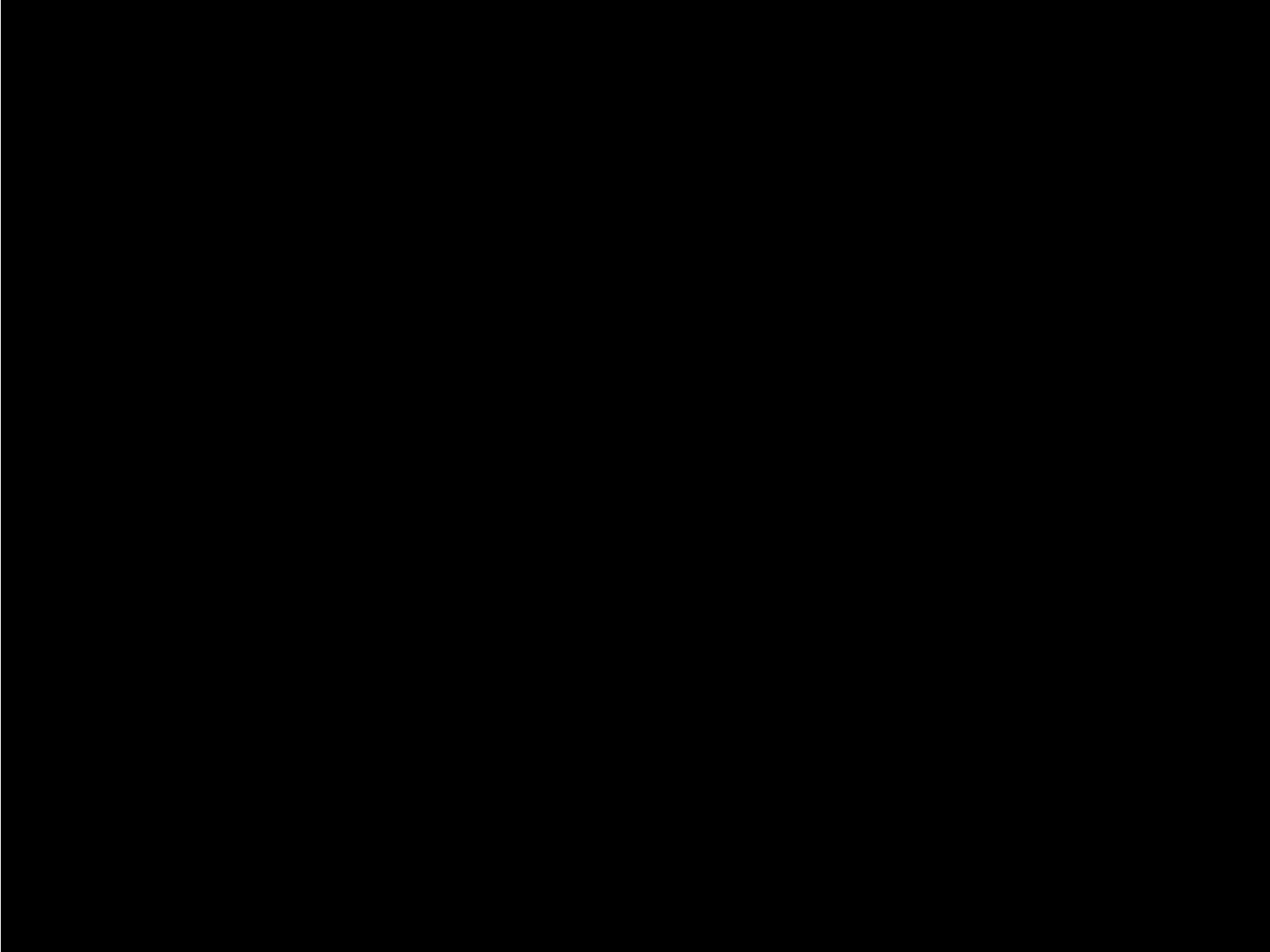
Returns all measurement points below threshold



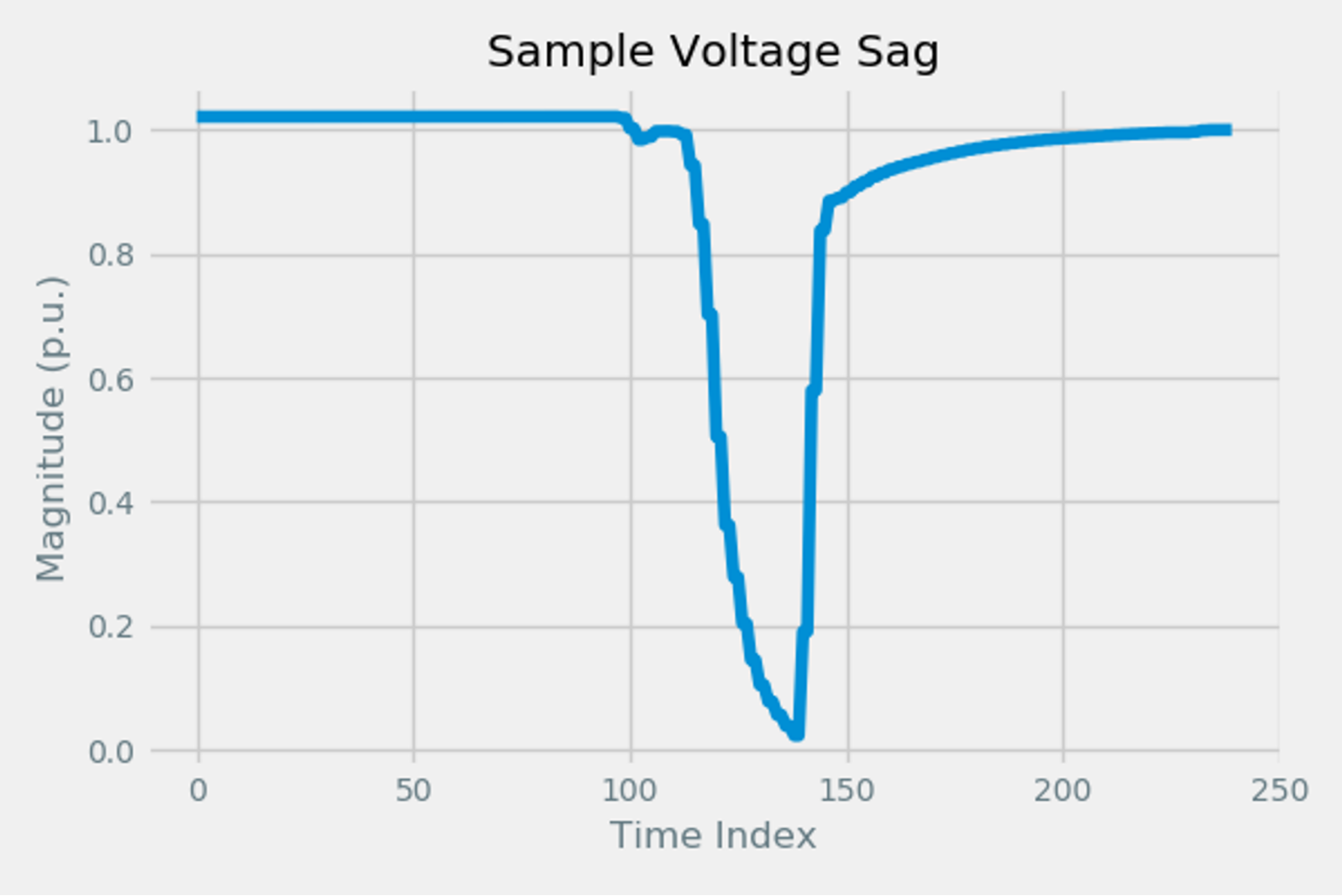
Consolidate points to get sag magnitude & duration

Database Structure





Now we can quickly find voltage sags and analyze their size and duration.

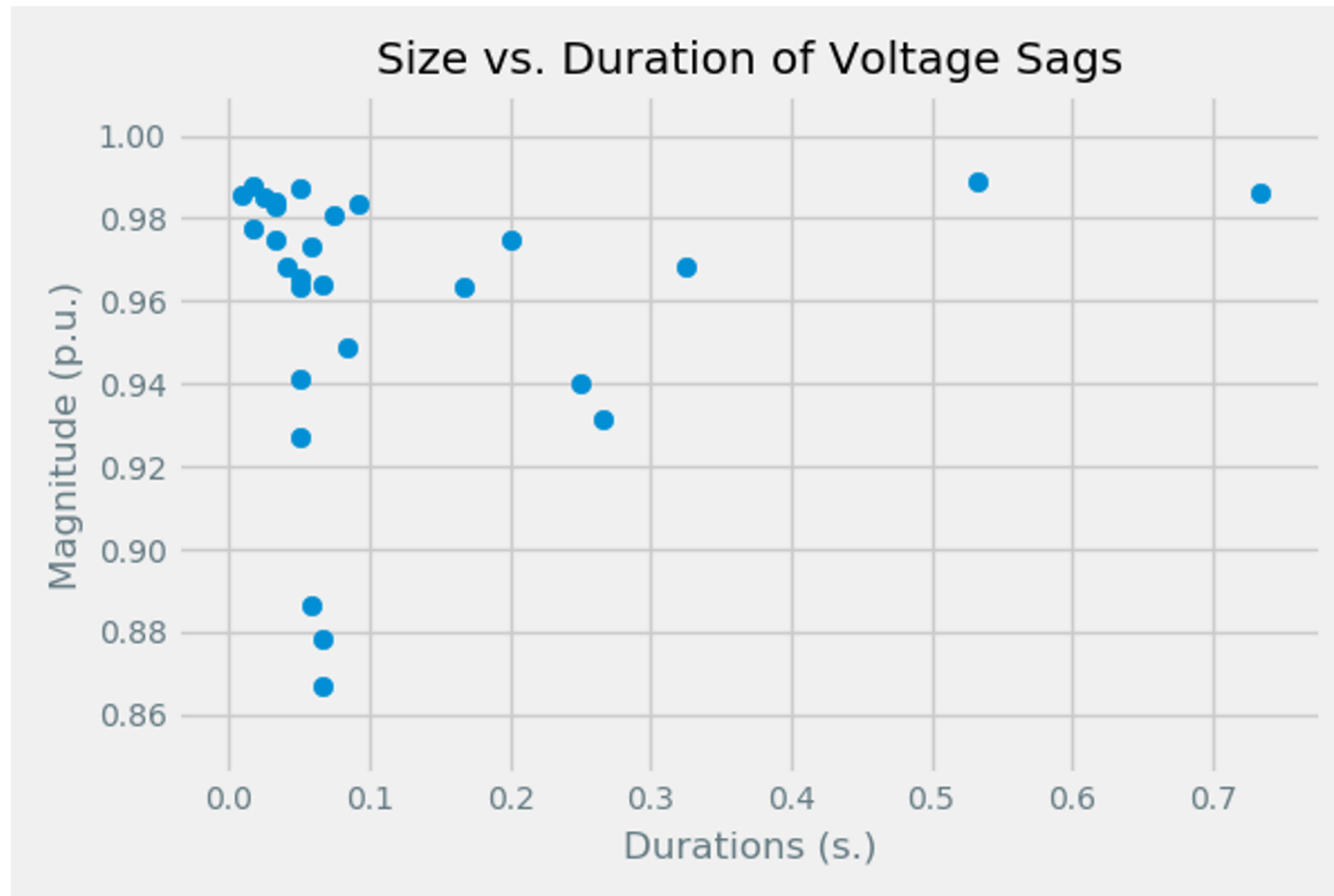


An example of a sag found by the algorithm.

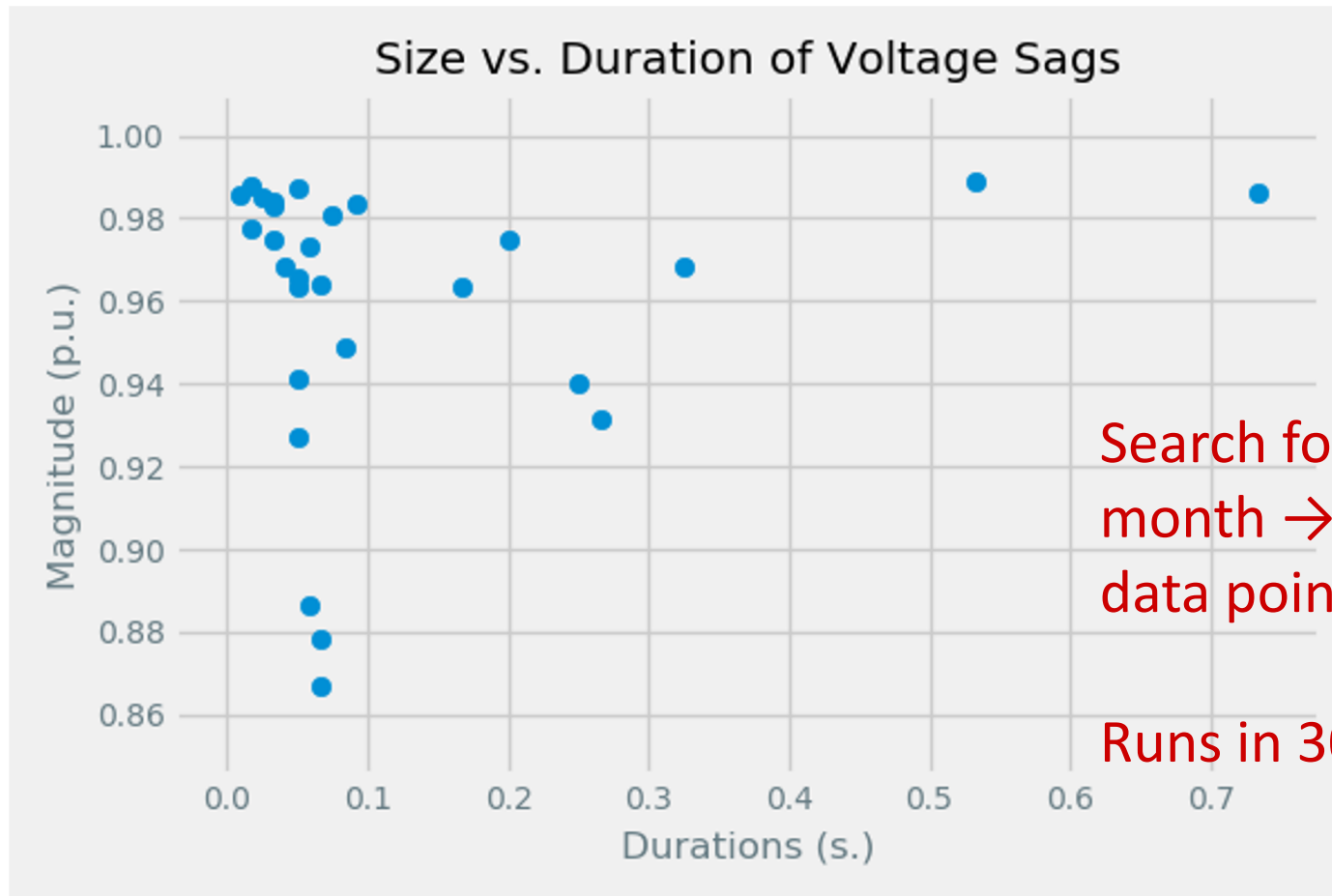
Step 4: Answer questions

- What are the magnitudes and durations of voltage sags at this location?
- How well does a model from the literature describe the frequency-vs-magnitude of the sags we observe?
- Do we see any impact of DG on sag magnitude?
- Are there weekly patterns in sag frequency?

- What are the magnitudes and durations of sags that occur at this measurement point?



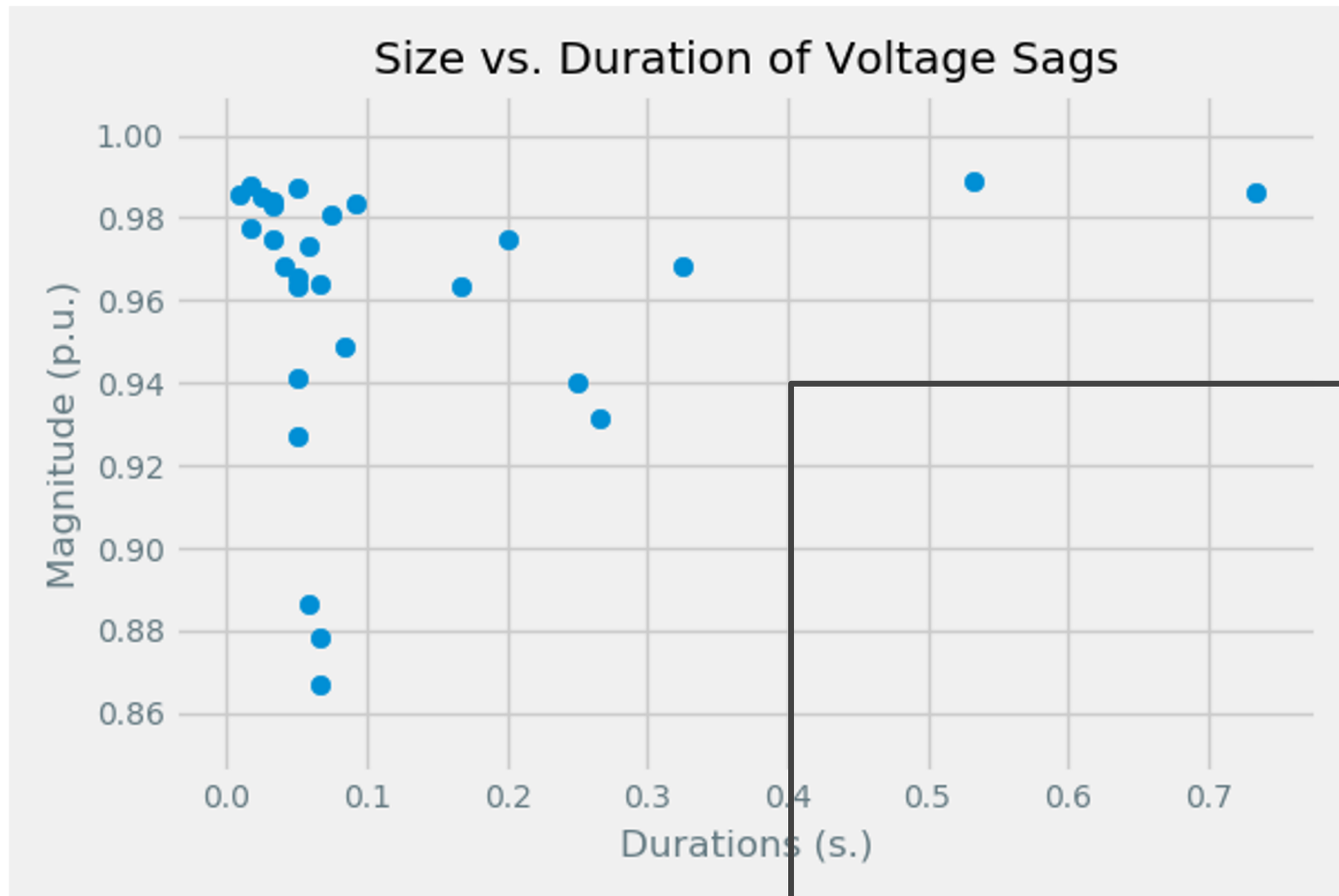
- What are the magnitudes and durations of sags that occur at this measurement point?



Search for sags across 1 month → 300 million data points.

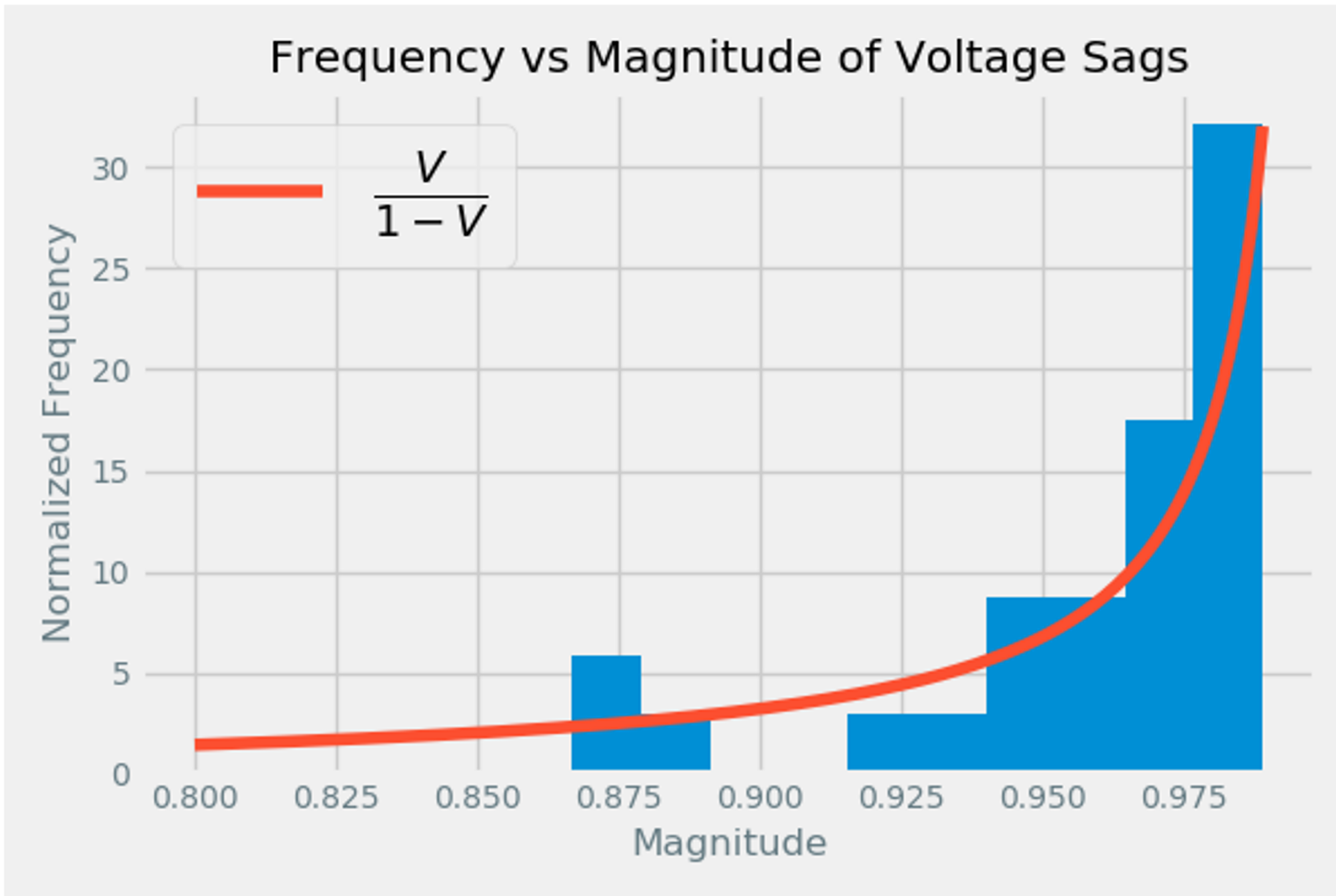
Runs in 30 seconds!

- What are the magnitudes and durations of sags that occur at this measurement point?



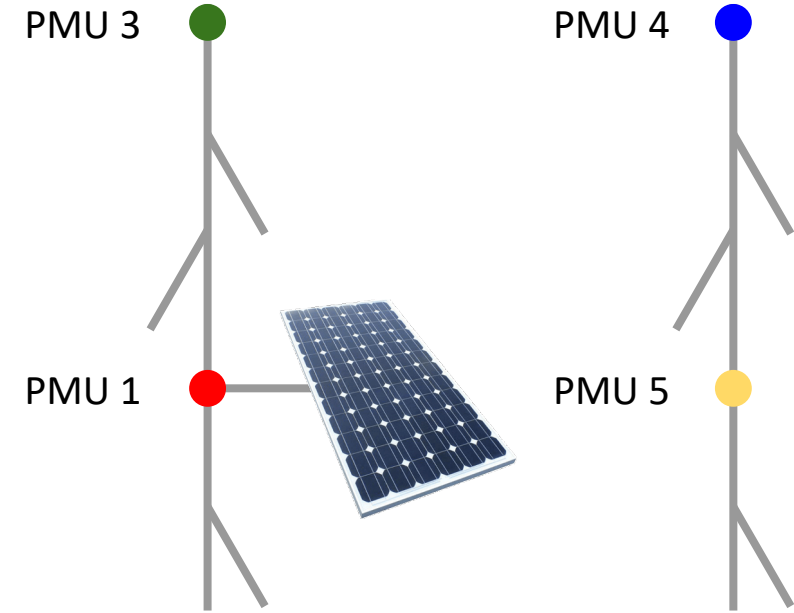
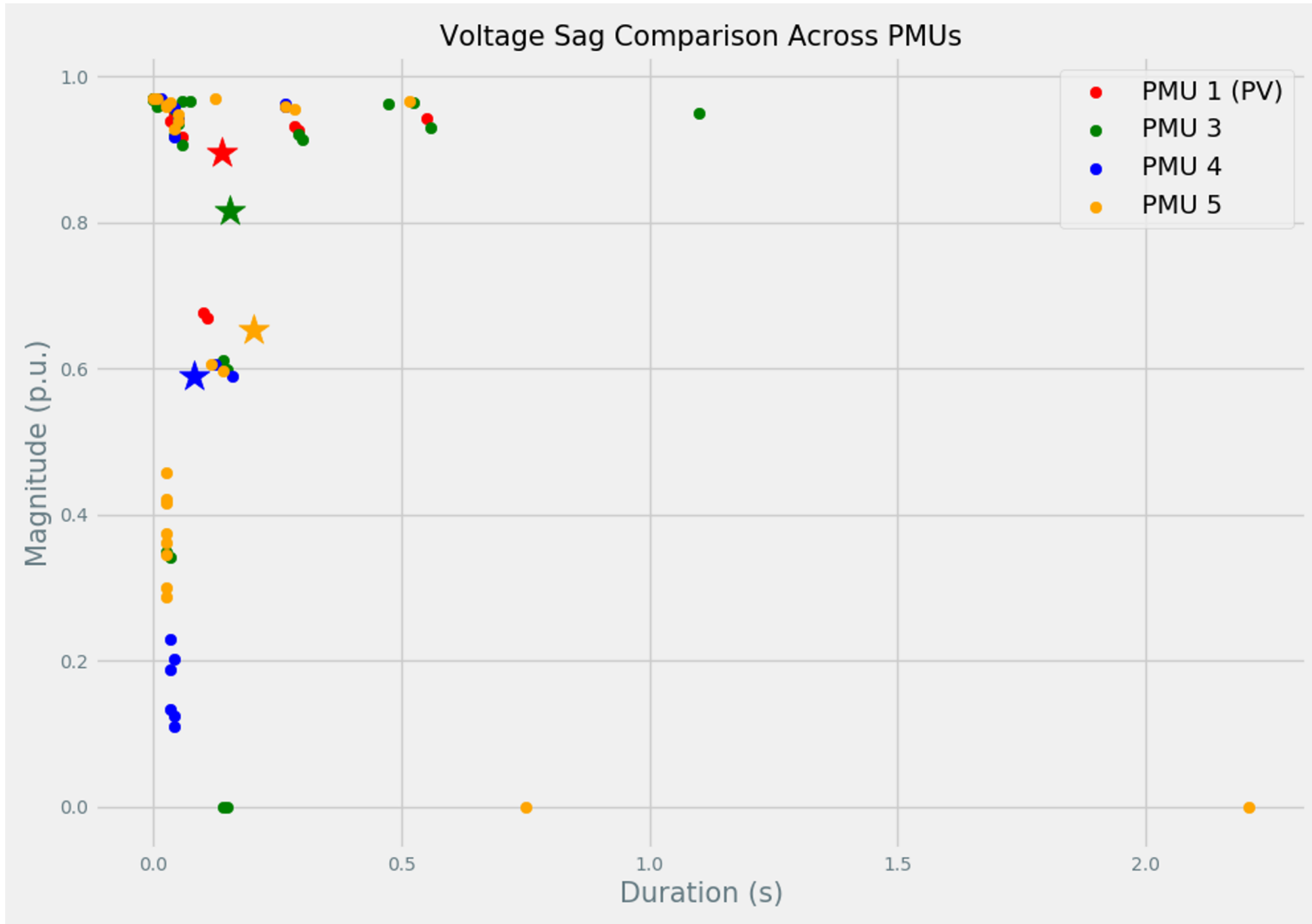
None in the danger zone!

- How well does a model from the literature describe the frequency-vs-magnitude of the sags we observe?



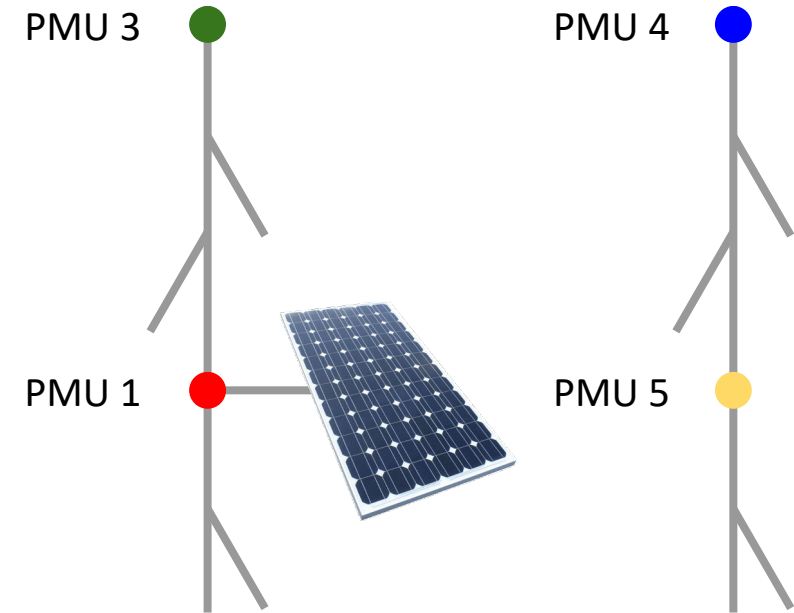
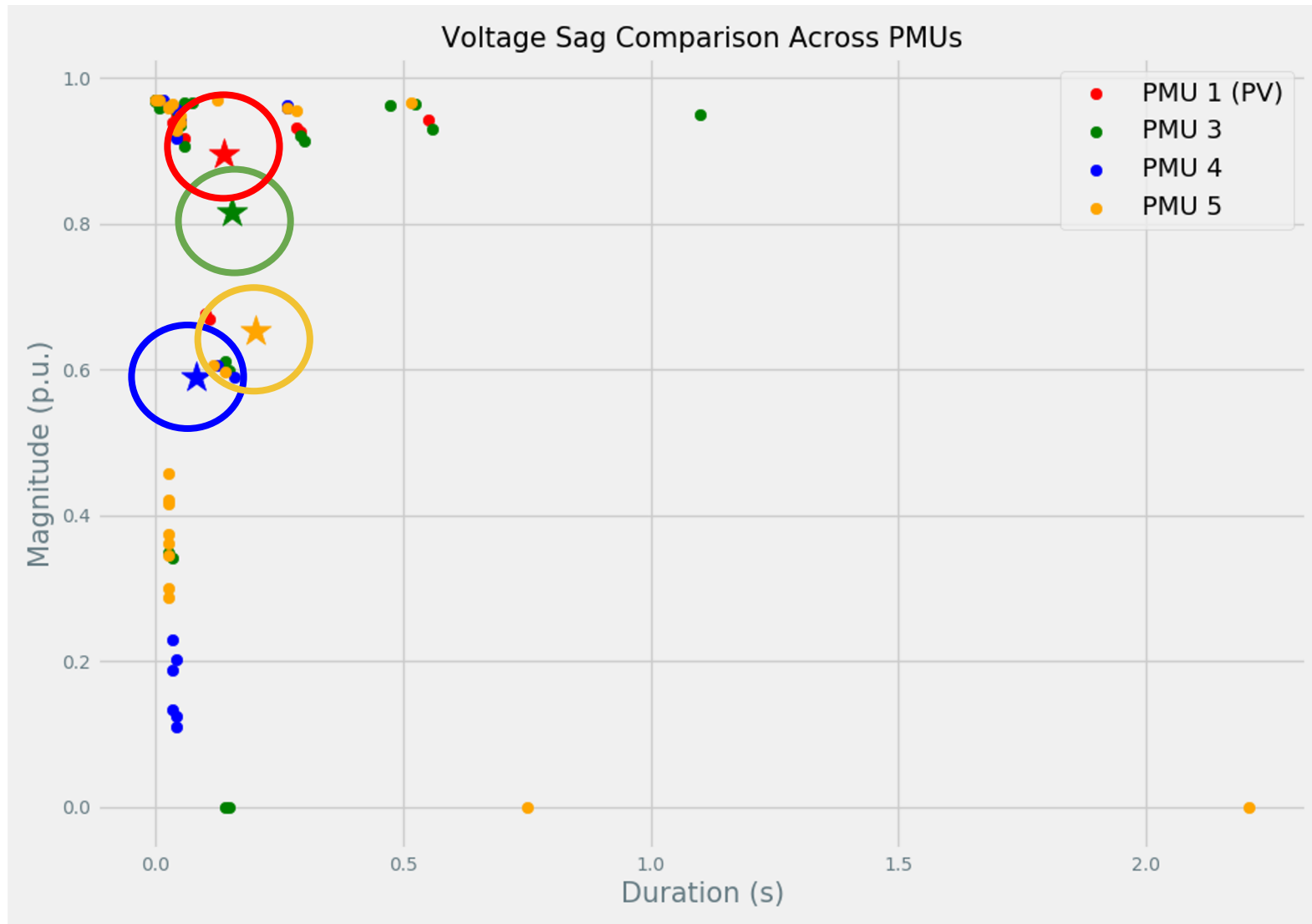
In the literature, a highly simplified model of fault-induced voltage sags predicts that the frequency of a sag with magnitude V will be proportional to $V / (1 - V)$.

- Do we see any impact of DG on sag magnitude?



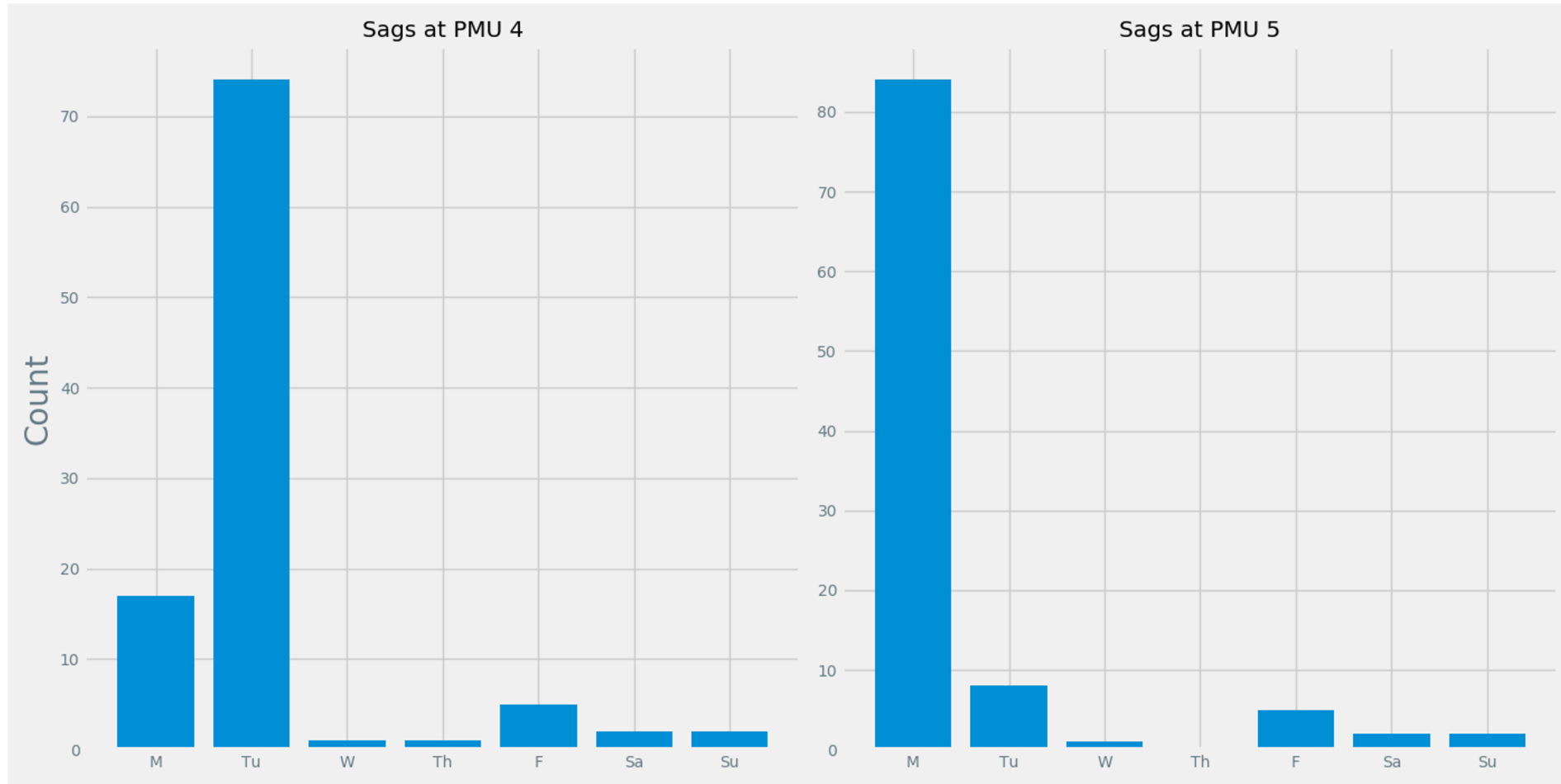
We can compare sag magnitude & frequency at multiple PMUs including one with a PV injection

- Do we see any impact of DG on sag magnitude?



We can compare sag magnitude & frequency at multiple PMUs including one with a PV injection

- Are there weekly patterns in sag frequency?



We can count sag occurrences per day to see if there are weekly patterns.

Step 5: Share Results

Creating use cases in a Jupyter notebooks allows for easy sharing of results.

- Visualizations
- Explanations
- Replicable code

Jupyter Voltage Sag Exploration (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted | Python 3.0

```
In [10]: # Choose the stream
# PMU 1 from sunshine dataset
stream = streams["35bdb8dc-bf18-4523-85ca-8ebe384bd9b5"];
# Get nominal voltage of stream
vnom = get_mean_value(stream);

# Start and end times of period to study
start = "2016-11-19T00:00:00.000Z"
end = "2017-02-04T00:00:00.000Z"

# Threshold below which data is considered a voltage sag
thresh = 0.99 * vnom;

# Find voltage sag data points
sags = find_vsags_dfs(stream, thresh, start=start, end=end);
# Get features of voltage sags
starts, durs, mags = sag_survey(sags, verbose=False);
```

Sag Features

Now we can explore patterns and features of the voltage sags found across several months of data.

```
In [251]: plt.scatter(durs / 1e9, mags / vnom);
plt.title('Size vs. Duration of Voltage Sags');
plt.xlabel('Durations (s.)'); plt.ylabel('Magnitude (p.u.)');

plt.tight_layout();
plt.savefig('size_vs_durs', dpi=100);
```



```
In [252]: # Plot a histogram of the normalized magnitude
# of the voltage sags
pmags = mags / vnom;
plt.hist(pmags, density=True);

# A result in the literature says that the frequency of
# a voltage sag with minimum value V will be V / (1-V)
# Plot this fit
x, y = freq_vs_size(0.8, 0.99);
plt.plot(x, y, label = r'$\frac{V}{1-V}$');

# Add keys
plt.legend(fontsize=20);
plt.title('Frequency vs Magnitude of Voltage Sags');
plt.xlabel('Magnitude'); plt.ylabel('Normalized Frequency');

plt.tight_layout();
plt.savefig('freq_vs_voltage', dpi=100);
```



Compare voltage sag occurrences at multiple locations.

Exploring Phase and Frequency

Miles Rusch
UC Berkeley



What is frequency?

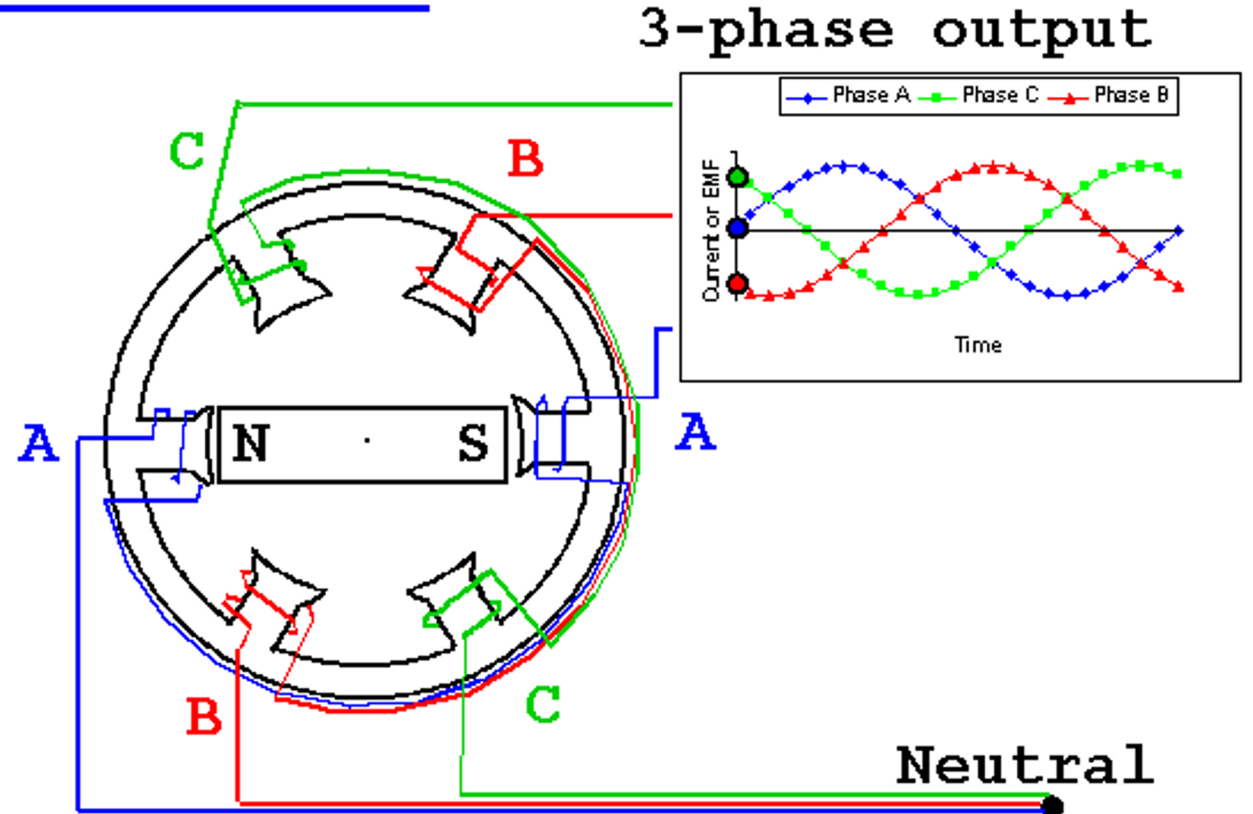
Frequency corresponds to the number of cycles per second

$$V(t) = A \cdot \cos(\omega_0 \cdot t + \phi)$$

The phase (ϕ) is the angular offset relative to a simple cosine

Definition assumes voltage and current are sinusoidal

The Generator



T. Davies 2002

What is phase?

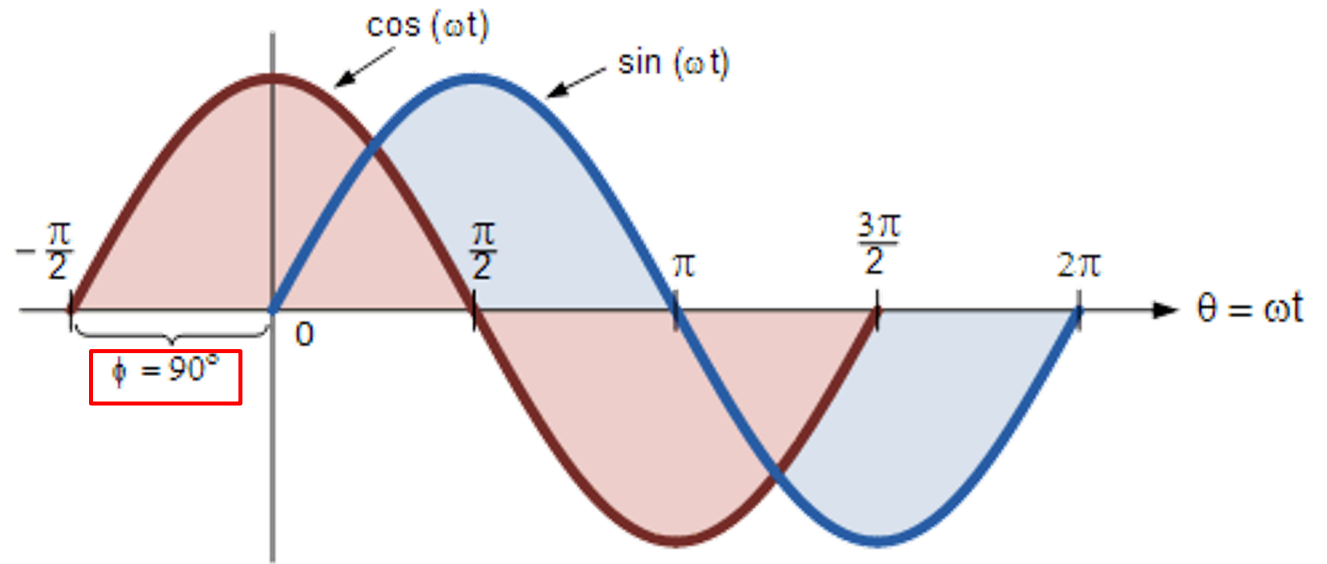
$$V(t) = A \cos(\omega_0 t + \phi)$$

Both amplitude and phase can vary with time

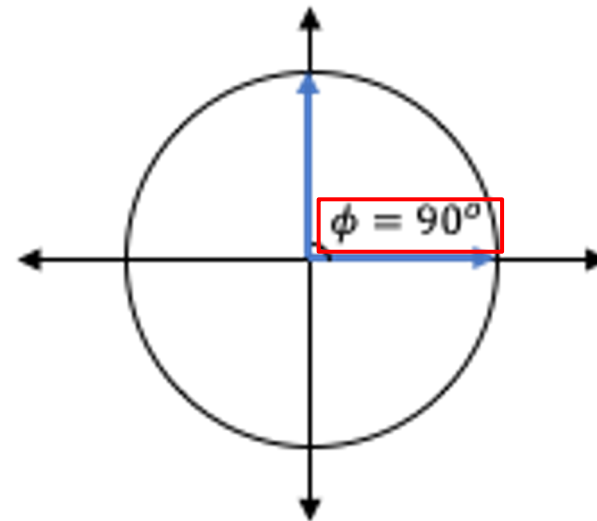
$$V(t) = A(t) \cos[\omega_0 t + \phi(t)]$$

We define frequency as the time derivative of the cosine argument:

$$\omega(t) = \omega_0 + d\phi/dt$$



<https://www.electronics-tutorials.ws/wp-content/uploads/2018/05/accircuits-acp33a.gif>



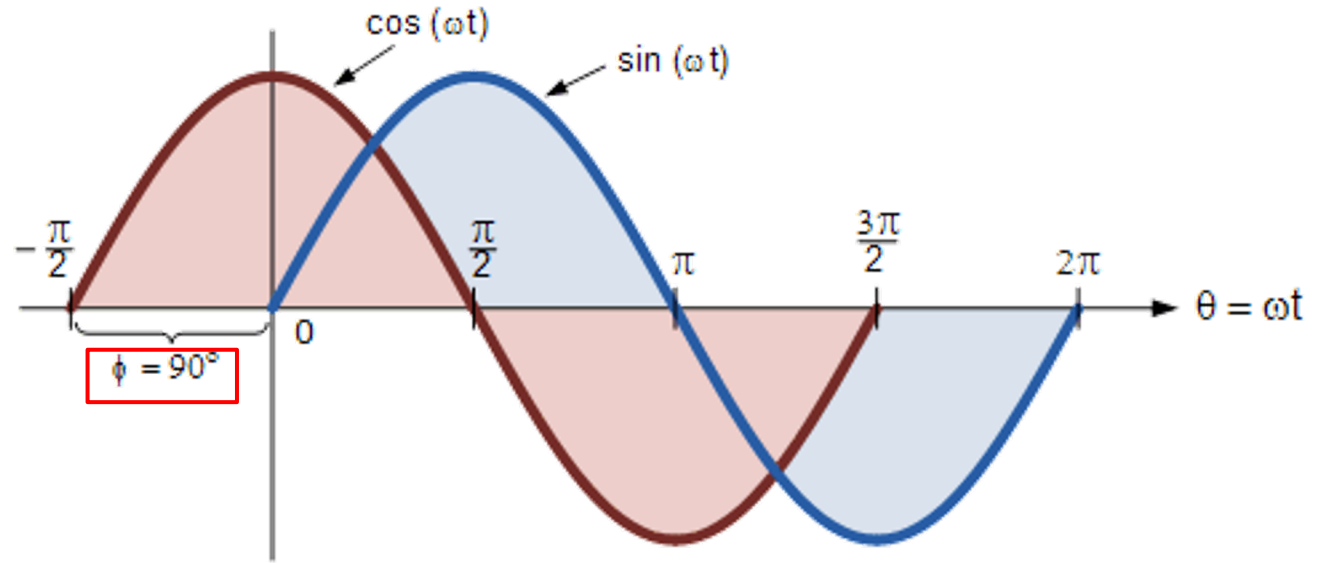
More on Phasors - Nominal Frequency

$$\omega(t) = \omega_0 + d\phi/dt$$

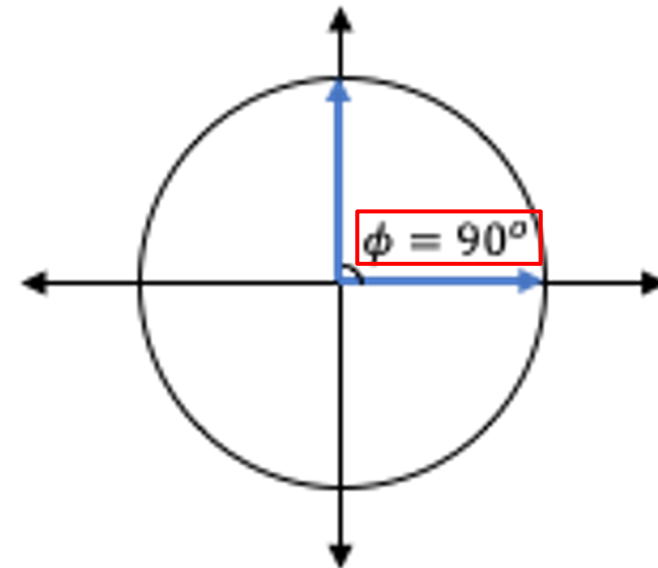
We simplify notation by subtracting the nominal 60Hz frequency (ω_0)

$$V = A(t) \angle \phi(t)$$

$$\omega(t) = d\phi/dt$$



<https://www.electronics-tutorials.ws/wp-content/uploads/2018/05/accircuits-acp33a.gif>

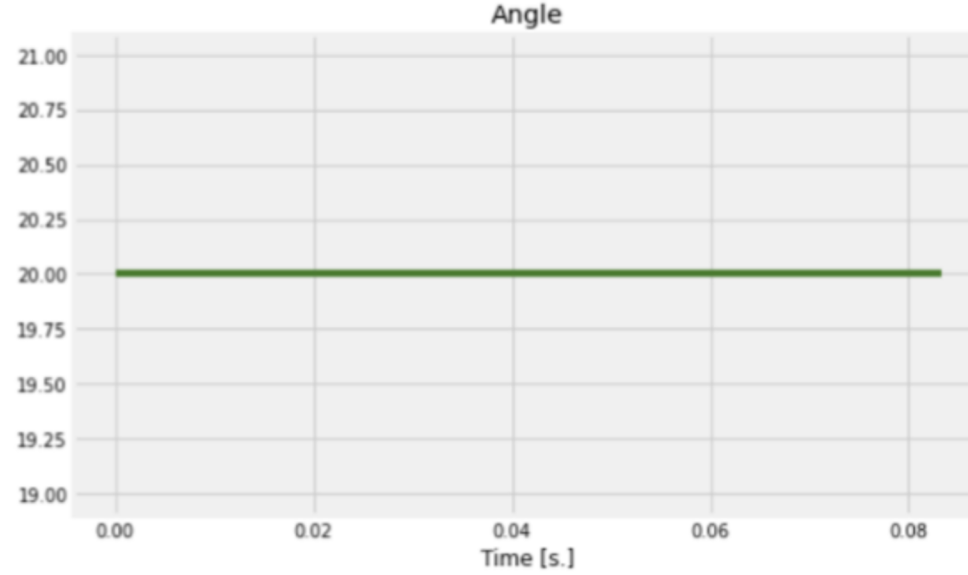
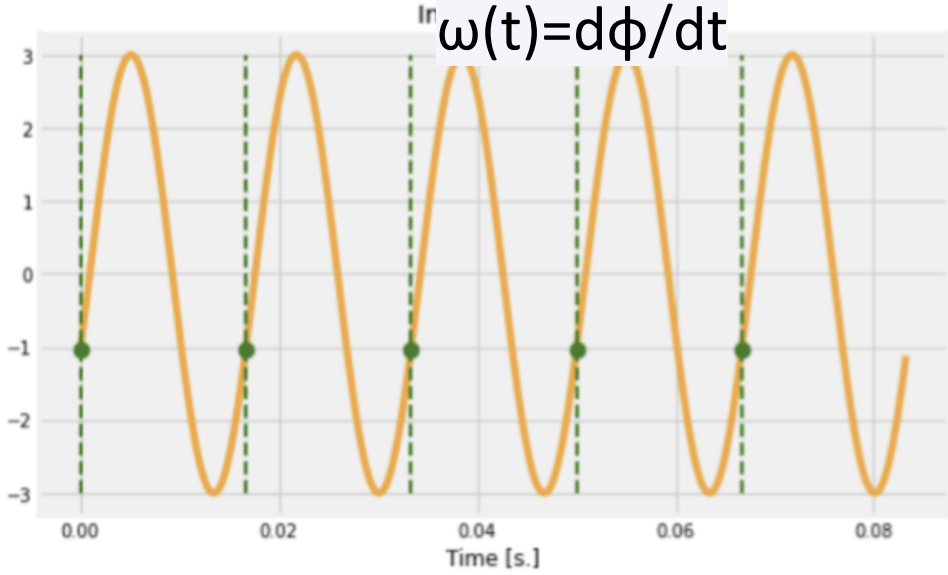


What happens to phase when the frequency isn't 60 Hz?

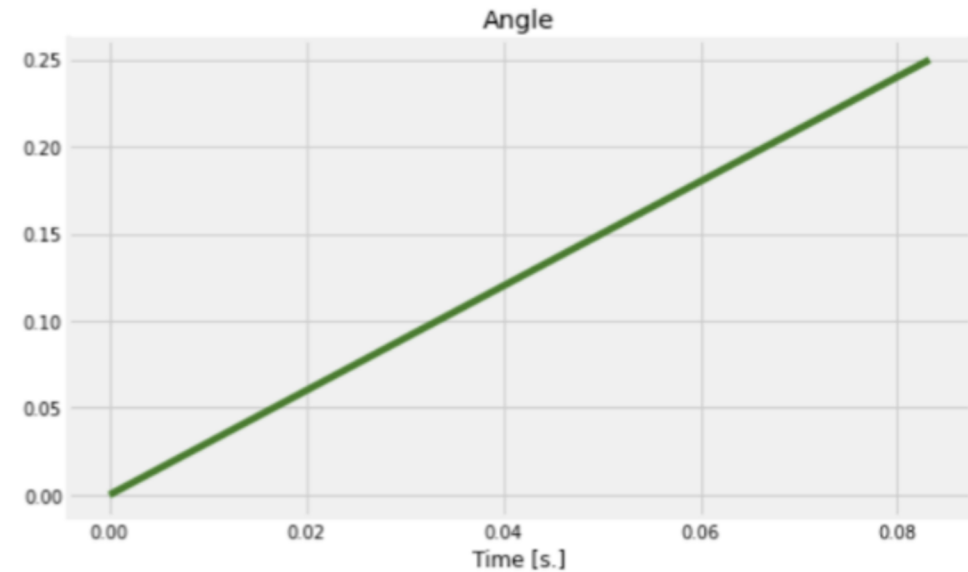
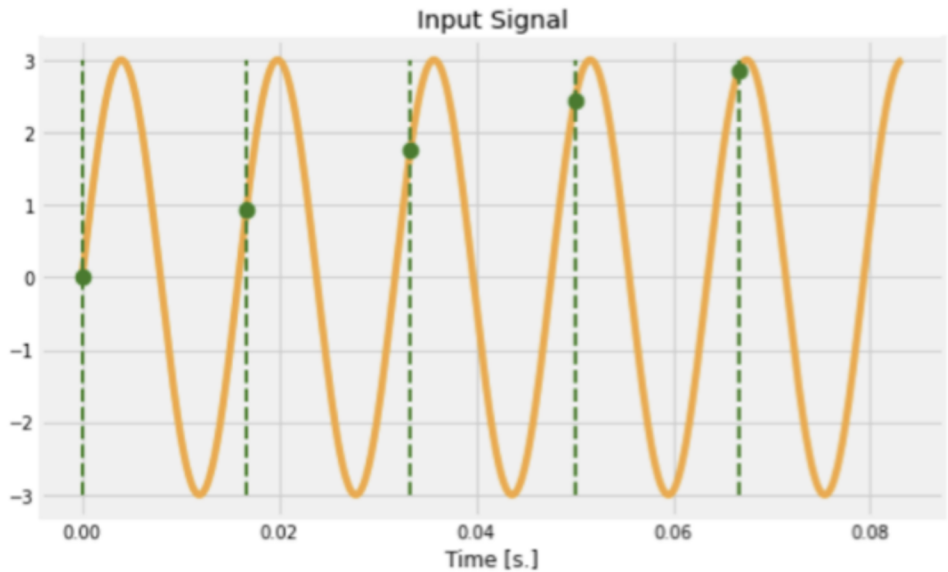
$$V(t) = A \cos(\omega_0 t + \phi)$$

$$\phi(t)$$

$$\omega(t) = d\phi/dt$$



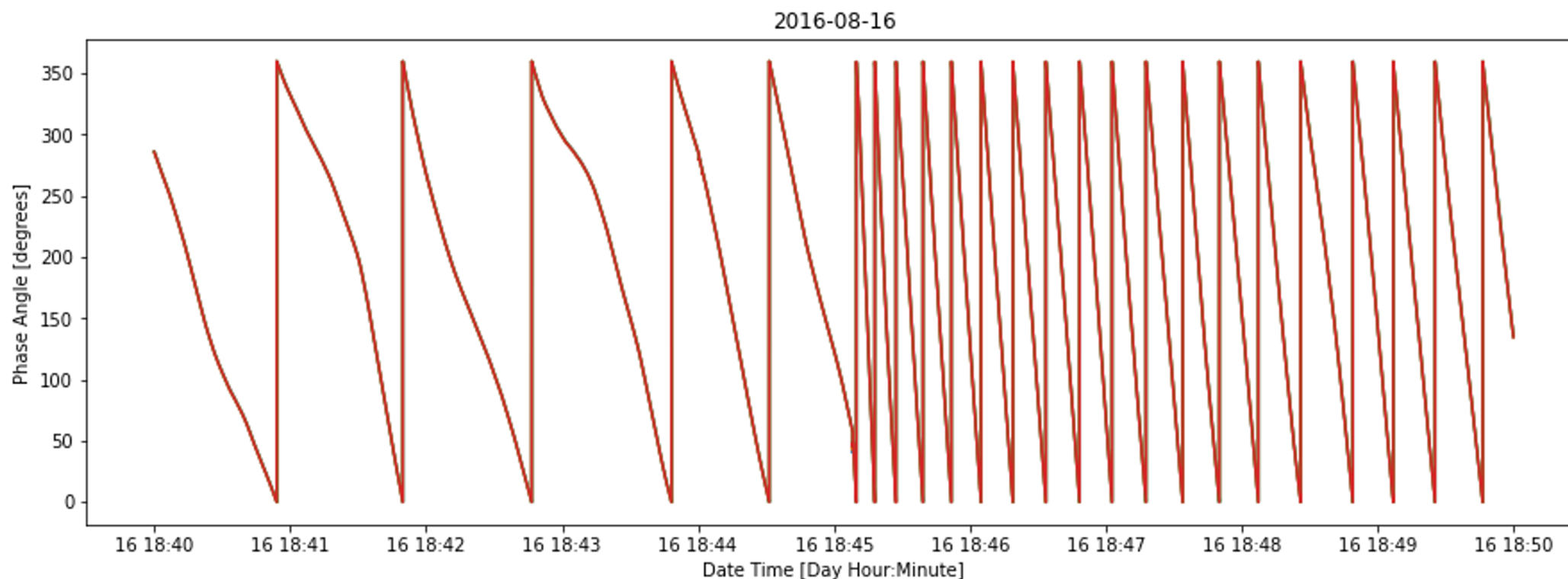
$$\omega = 0$$



$$\omega > 0$$

A first glance at the phase angle data

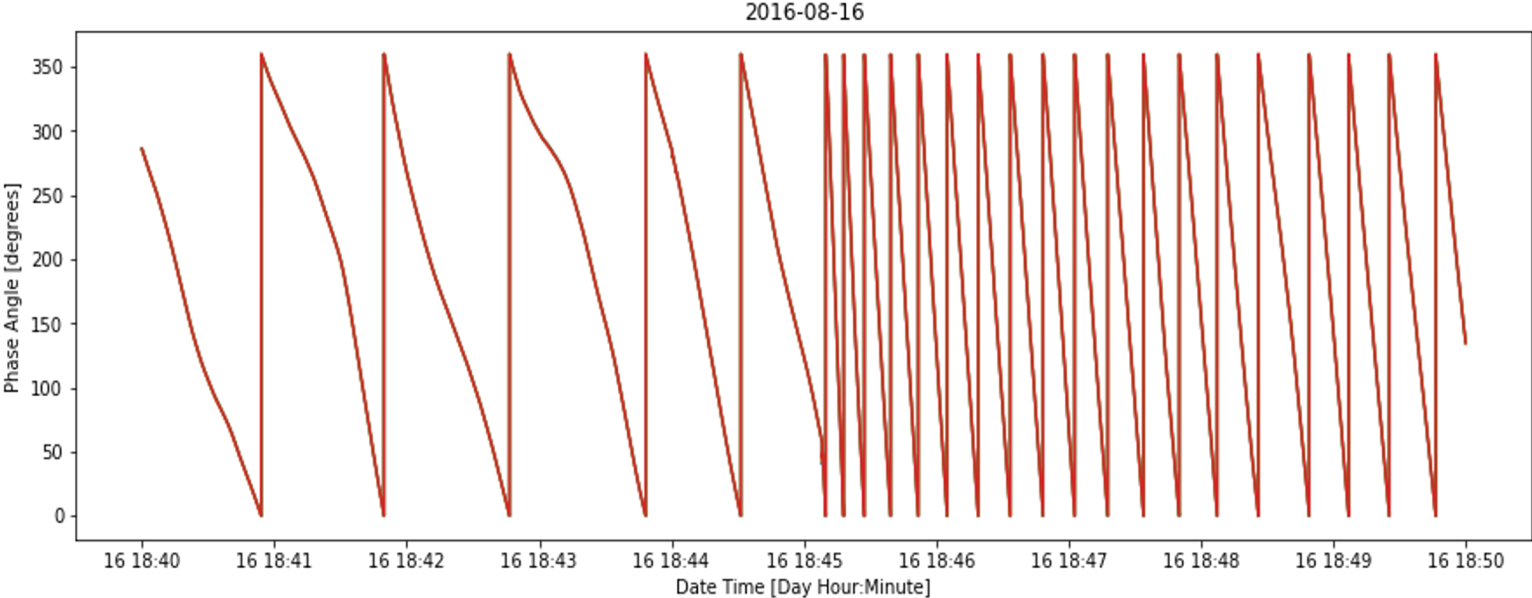
$\phi(t)$



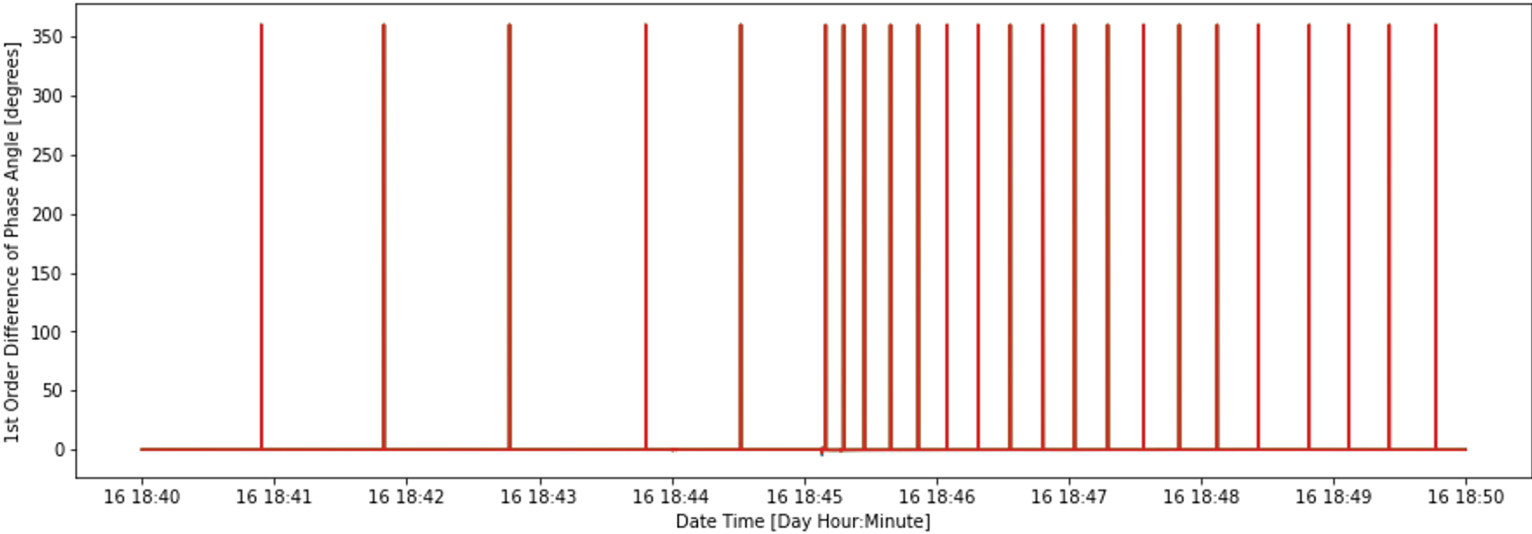
Notice the phase angles are always between 0 and 360 degrees

Calculating Frequency Using Discrete Difference

$$\phi[t]$$



$$\phi[t+1] - \phi[t]$$



Discontinuities in the phase angle create large artifacts in the derivative calculation.

Unwrapping the Phase

A reminder ...

Angles are relative

$$360^\circ = 0^\circ$$

$$370^\circ = 10^\circ$$

$$350^\circ = -10^\circ$$

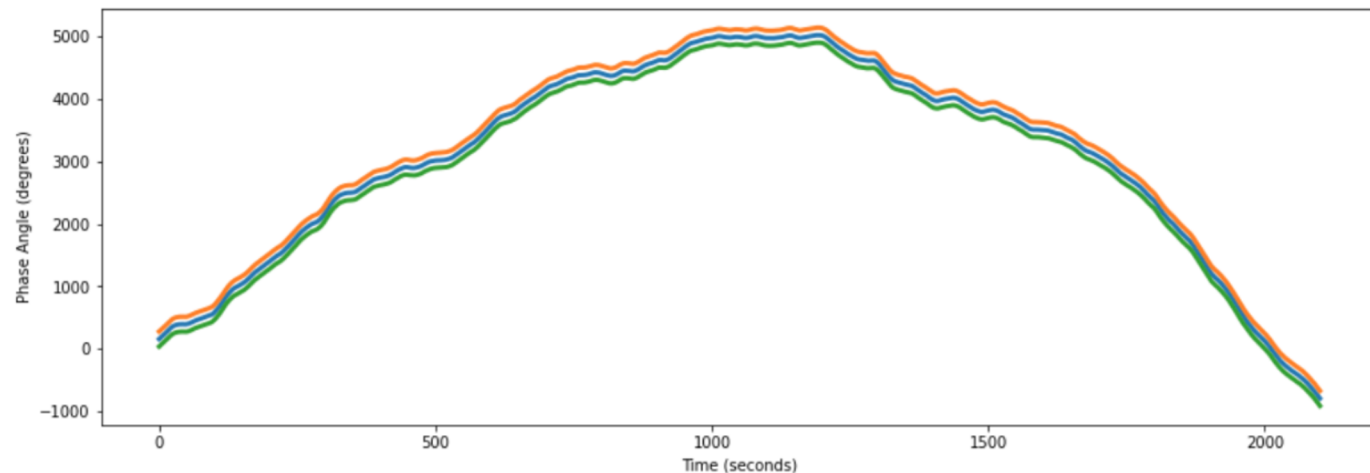
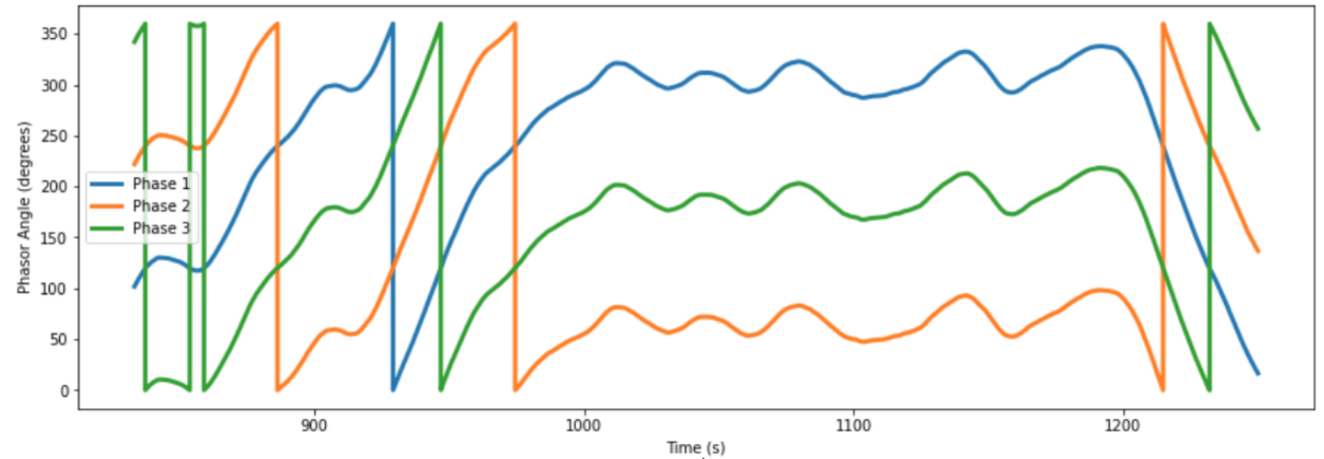
etc.

We can use this to express the angle such that we don't see these step changes in frequency.

This is called unwrapping

Unwrapping the phase

```
def unwrap_phase(points, threshold):  
    difference = []  
    for i in range(len(points)-1):  
        diff = points[i+1]-points[i]  
        if diff > threshold:  
            difference.append(diff-360)  
        elif diff < -threshold:  
            difference.append(diff+360)  
        else:  
            difference.append(diff)  
    return np.array(difference)
```



Case Study: Blue Cut Fire Solar Faults

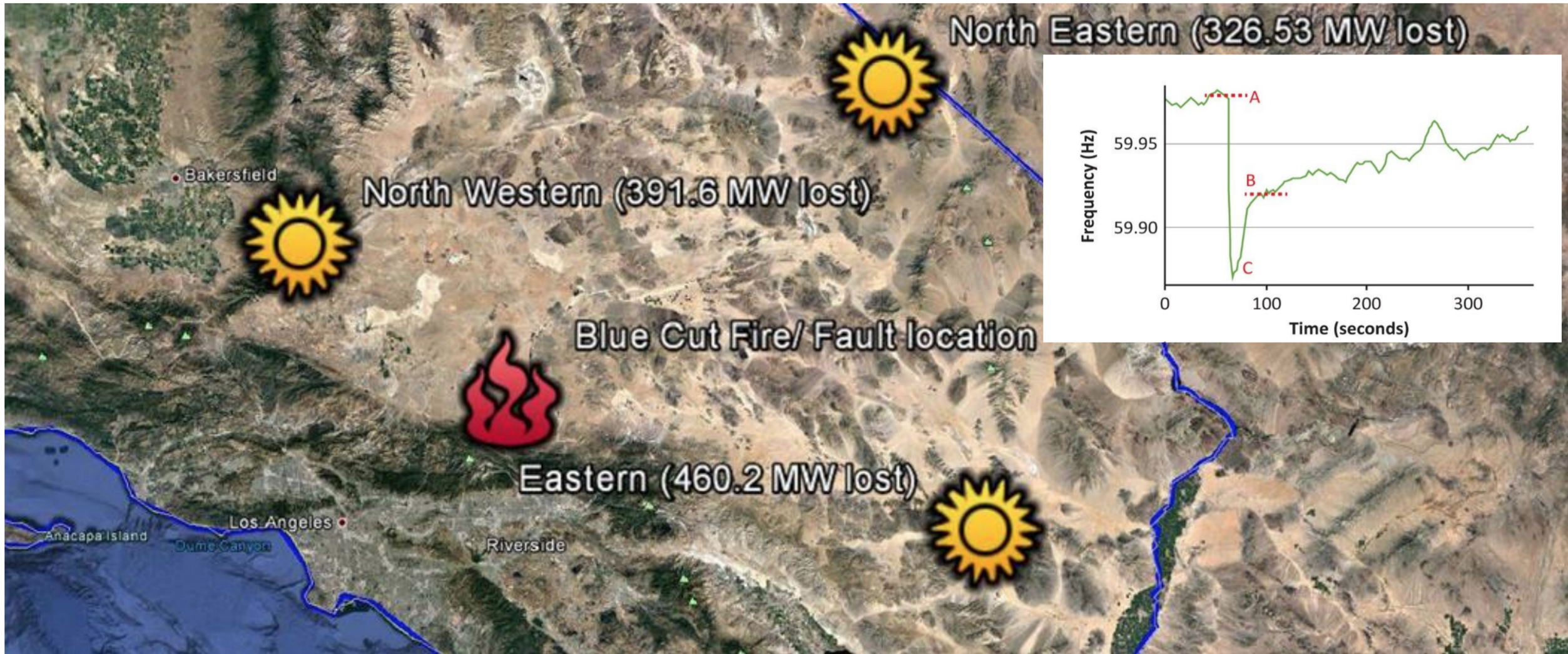


Figure 1.1: Map of the Affected Area and Blue Cut Fire Location

Blue Cut Fire in the Sunshine Dataset

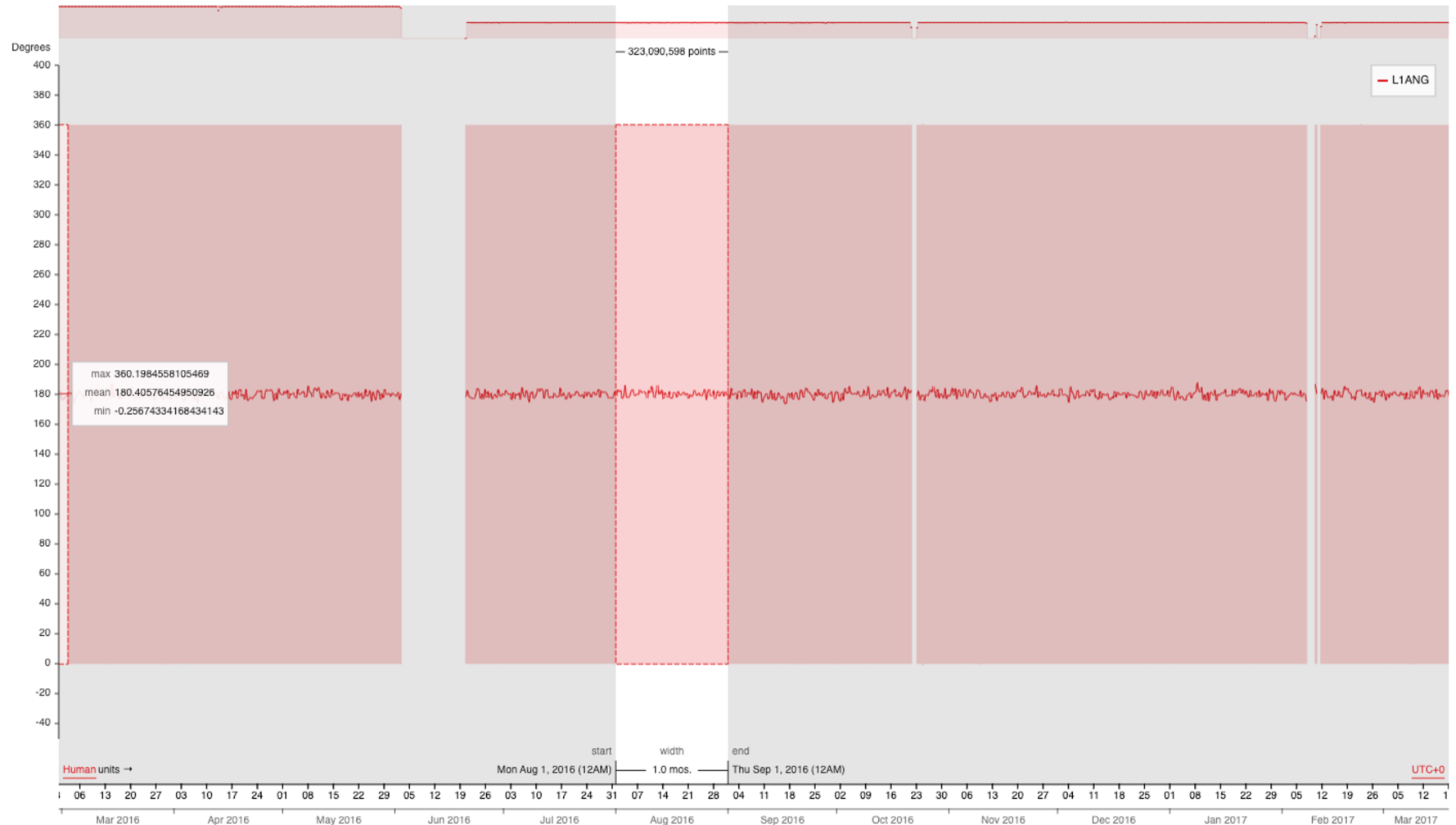
PingThings

Stream Selection

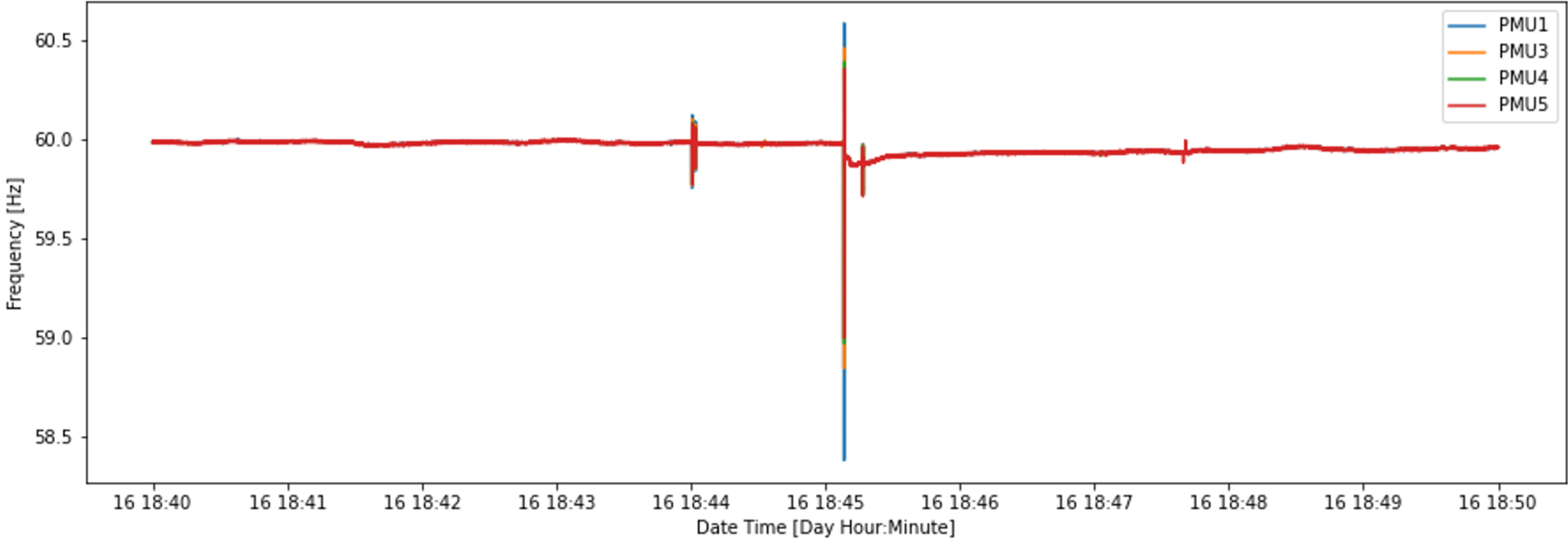
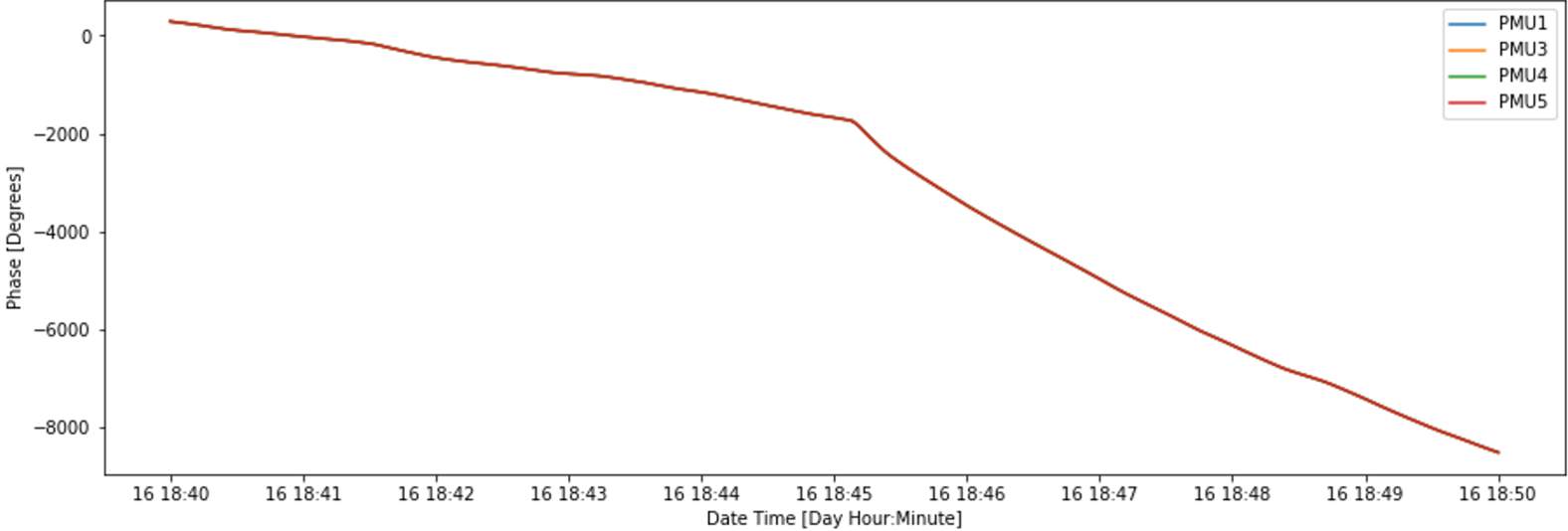
Stream Visualization

The transmission line fault cause by the Blue Cut fire occurred at 6:45pm UTC-0.

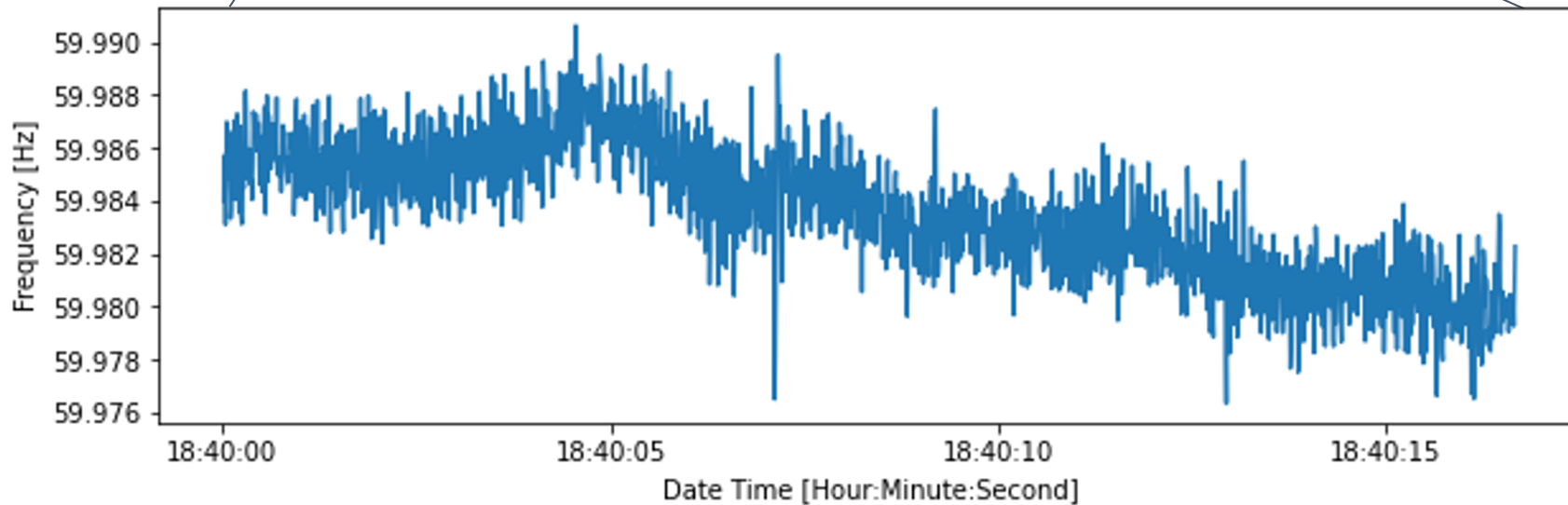
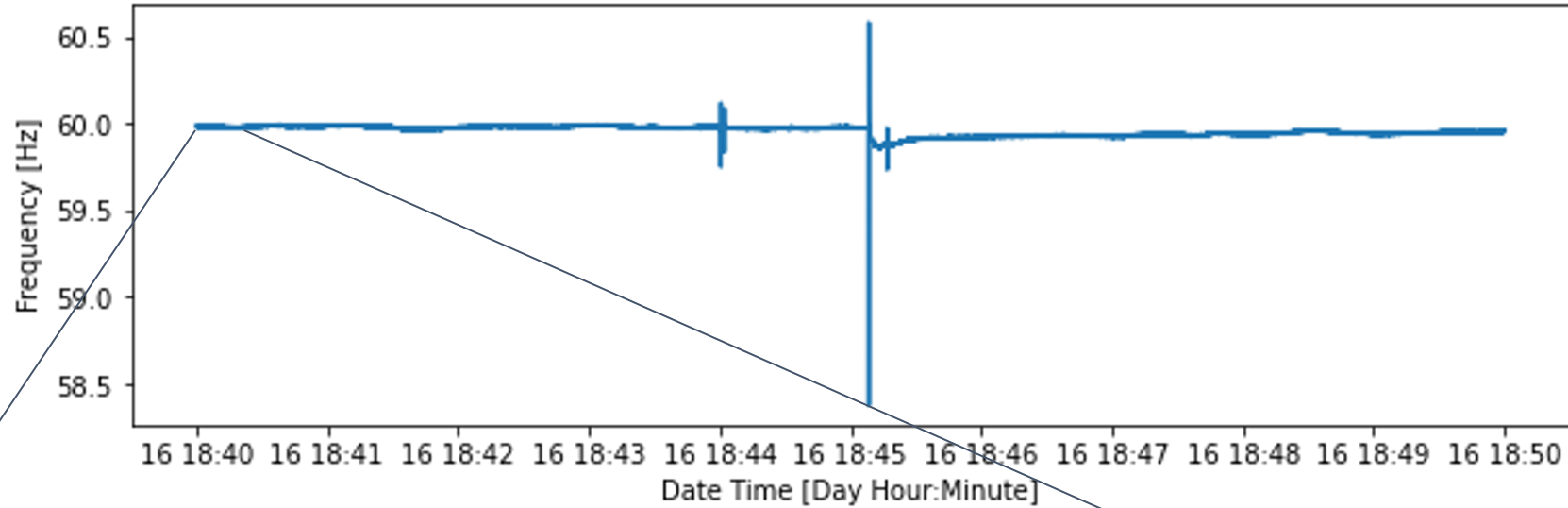
The sunshine dataset has PMUs reporting data during this event.



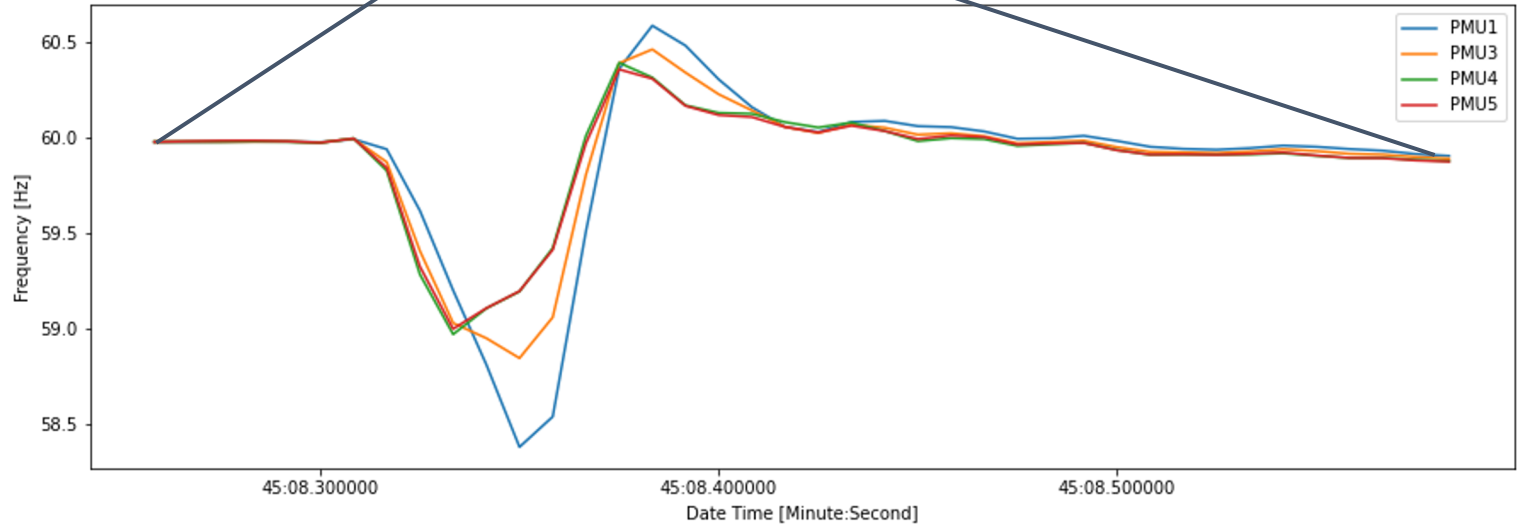
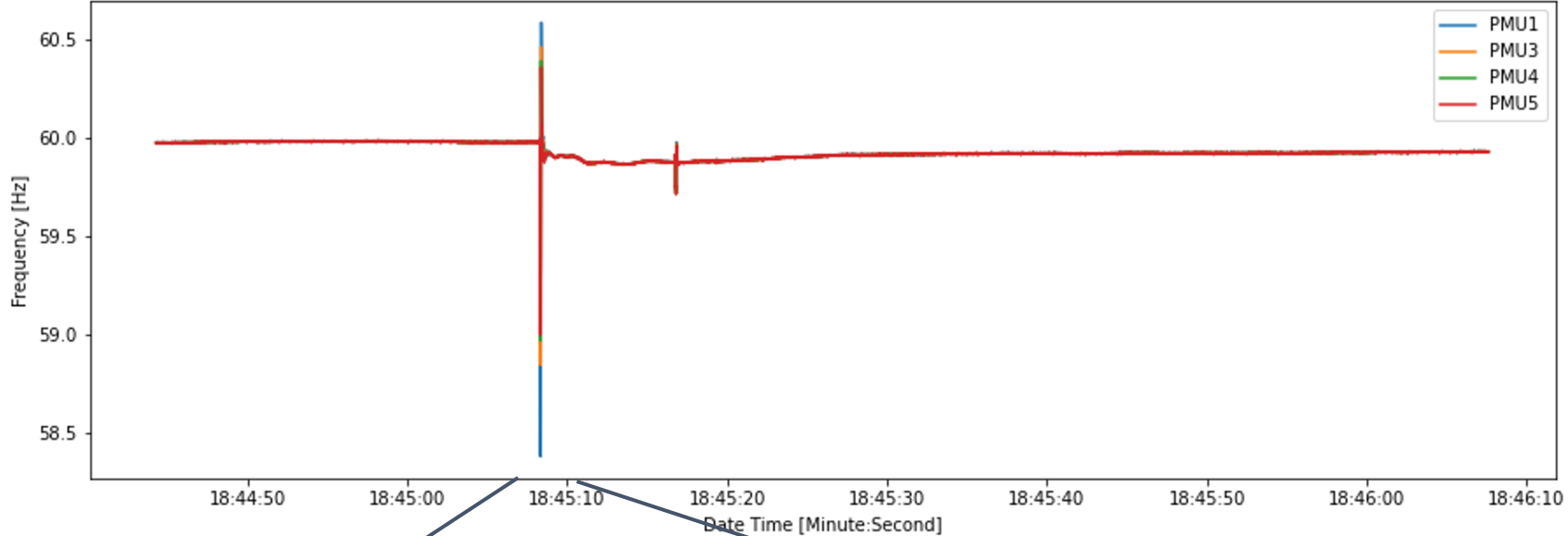
Calculating Frequency from Unwrapped Phase



Calculating Frequency from Phase

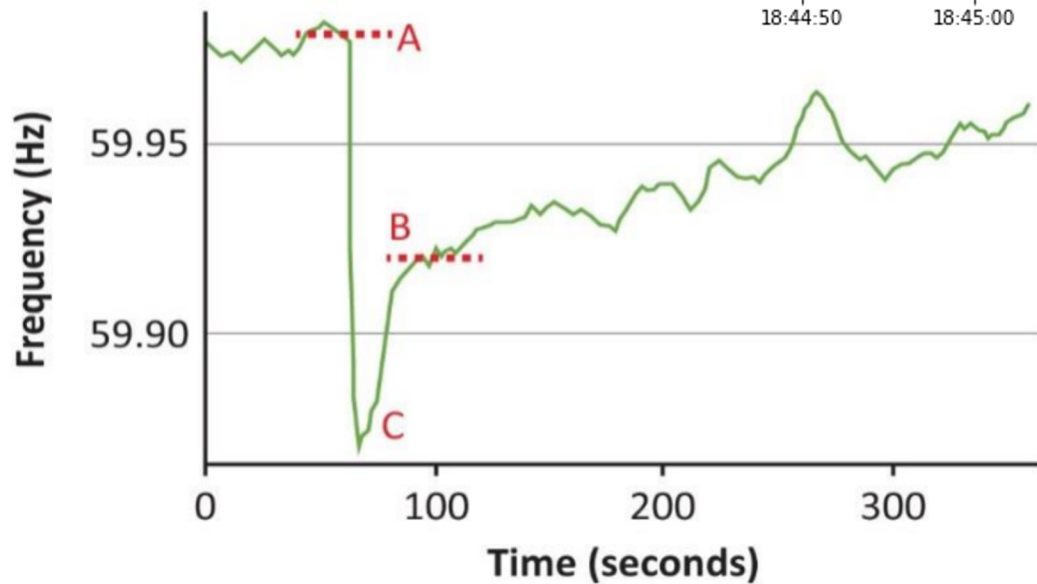
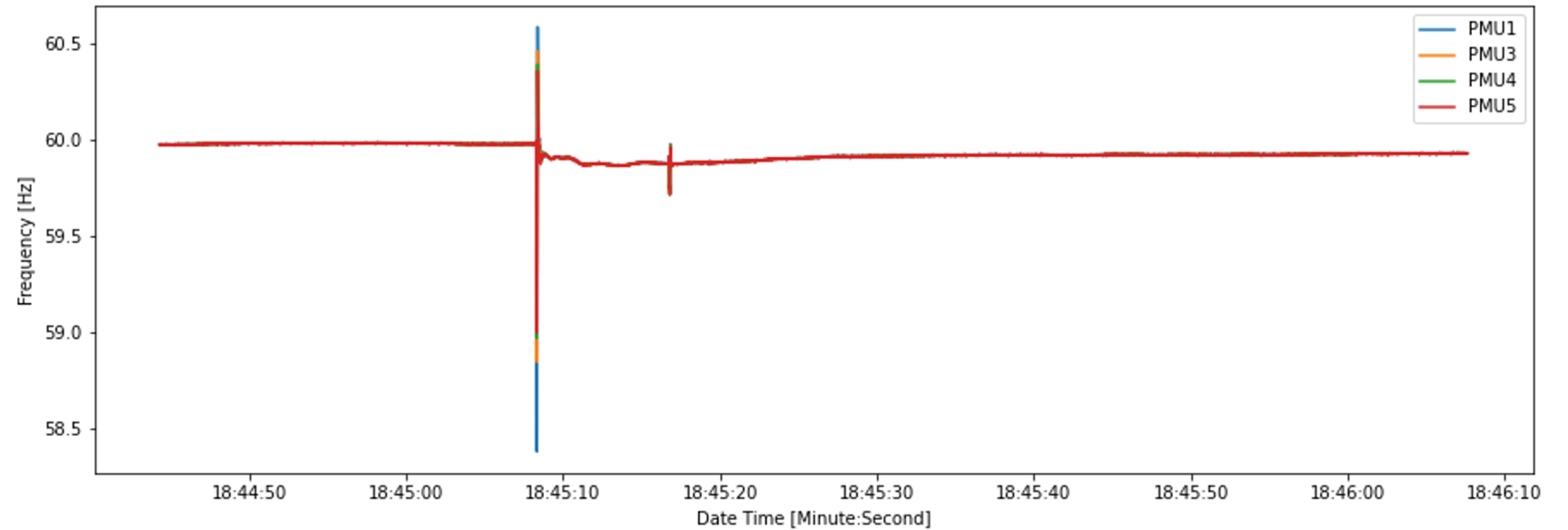


Calculating Frequency from Phase



The overshoot at such a short timescale and disagreement between PMUs indicate artifacts from the signal processing

Data Comparison

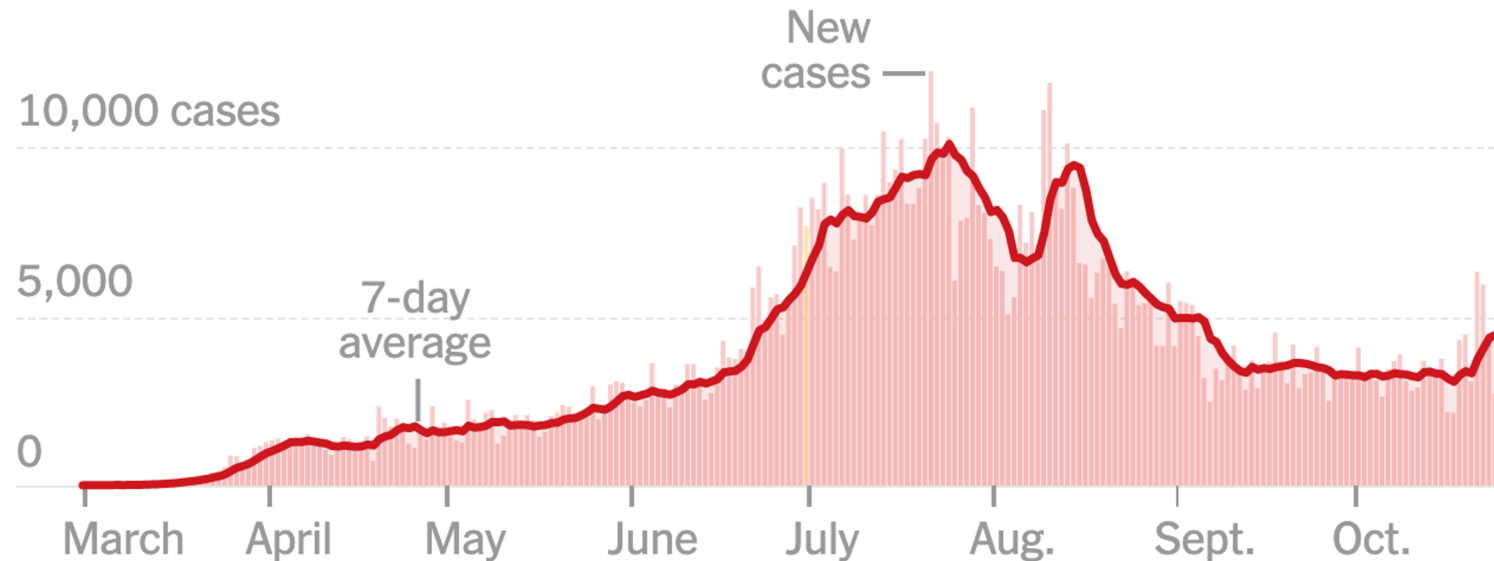


Our frequency data does not resemble the data reported from the Blue Cut Fire report.

Moving Average Filter Example

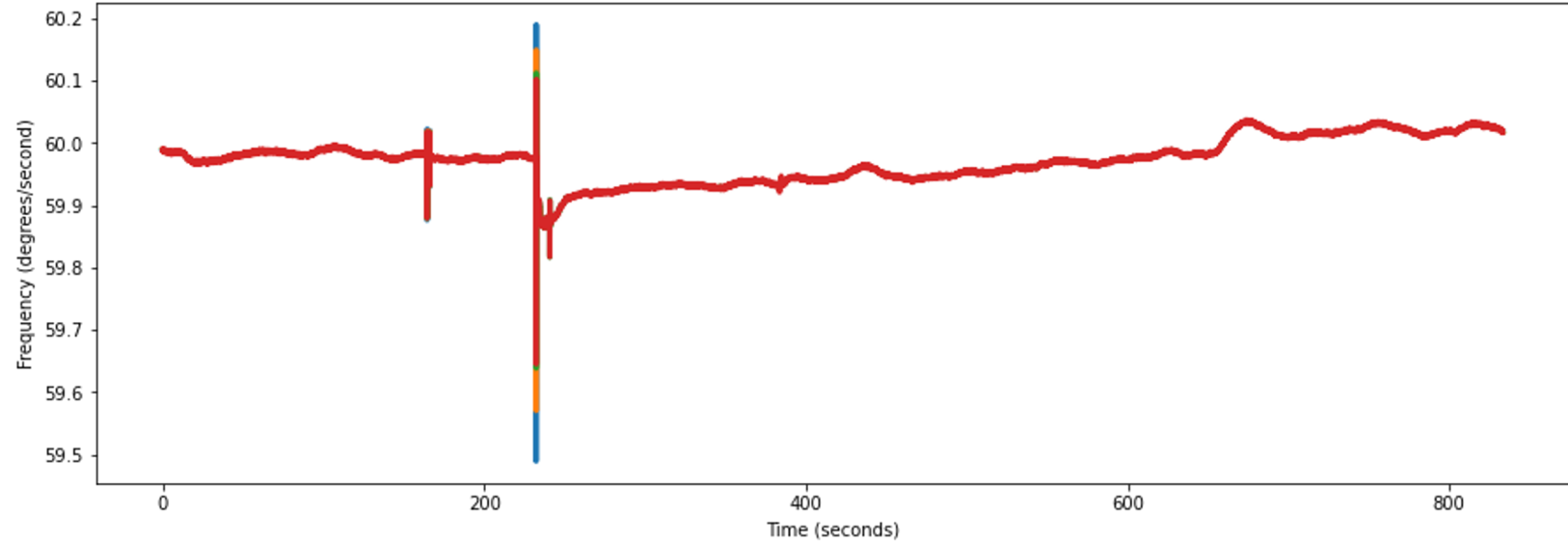
California Covid Map and Case Count

By The New York Times Updated October 27, 2020, 8:07 P.M. E.T.

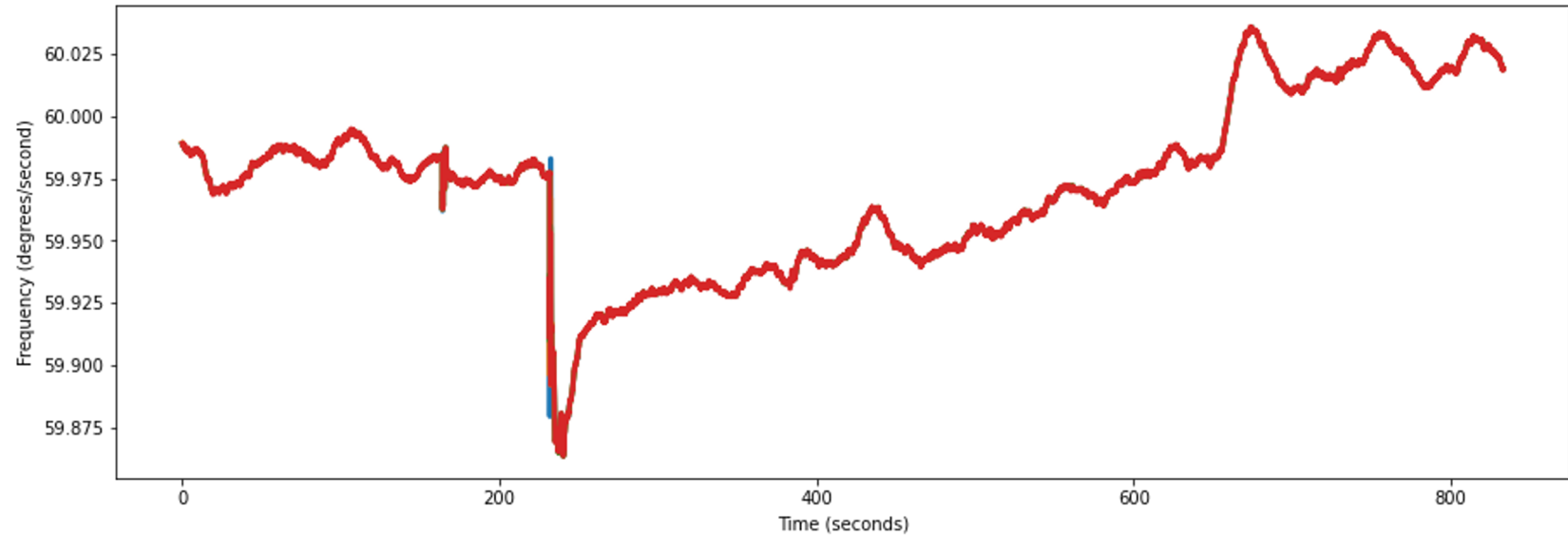


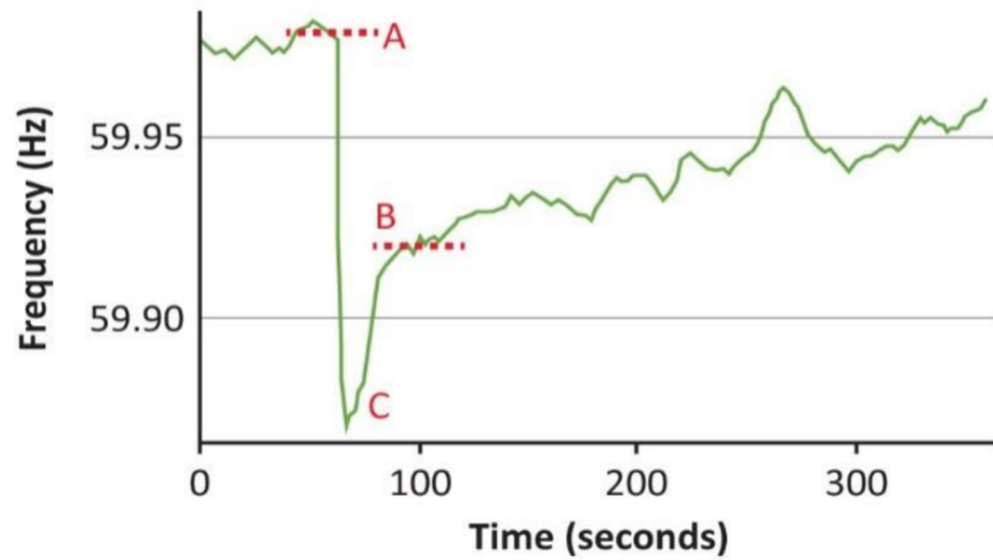
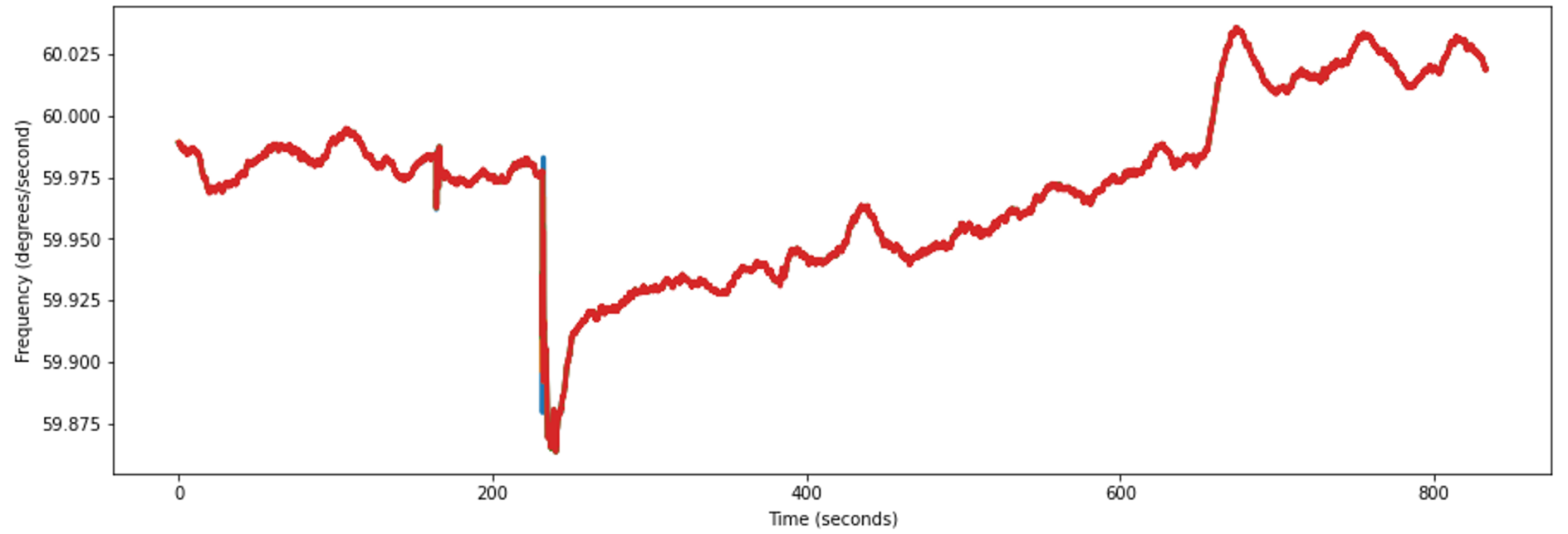
filter frequency - no overshoot

0.1 second average



0.5 second average





Conclusions from filtered frequencies

Frequency is not well defined during transient events when the waveform isn't sinusoidal.

The IEEE standard requires PMUs to have minimum step change tracking, error, etc, but it does not specify the algorithm to compute phasors from the actual point on wave data. Choosing algorithms with different fitting methods and filtering will change the reported frequency.

Point on wave data will be more accurate for analyzing transient frequency events.

Thank you!

Contact

info@ni4ai.org

Subscribe for Updates

<http://eepurl.com/he8jtj>

Upcoming Events

Nov 3-4 **NASPI working group meeting**

Nov 11-13 **IEEE SmartGridComm**

Accelerating AI on the Grid

info@ni4ai.org

