ELECTRIC POWER RESEARCH INSTITUTE

PREDICTIVEGRID™ TIME-SERIES DATA PLATFORM

Incubatenergy Labs 2020 Pilot Project Report

PingThings

Startup

PingThings

Ameren

Host

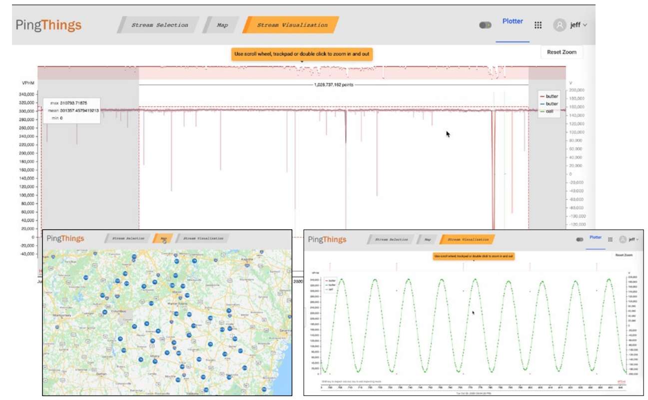
Ameren

Technology Solution

Utilities and other distribution system operators need modern, data-driven approaches for planning infrastructure improvements, diagnosing issues, and preventing outages. In the last 20 years, tens of millions of smart meters have been installed at residential and commercial endpoints to provide additional insight into distribution grid and customer behavior, and each one can take more than 100,000 measurements per year. Additional advanced metering and monitoring devices, including distribution synchrophasors, also are being widely deployed. The immense promise of the smart grid remains largely unrealized due to the challenges in handling, managing, and analyzing the vast trove of sensor data. For utility engineers and analysts, the faster they can visualize, explore, manipulate, and analyze data—transforming ideas into prototypes, testing and improving use cases, and deploying the applications to run at scale—the more value they can create.

This pilot was launched to demonstrate the PingThings PredictiveGrid™ platform for creating data-driven applications for realizing the full value of utility time-series data assets. PredictiveGrid is a data management, analytics, and artificial intelligence platform designed to work with time-series measurements from any grid sensor, from advanced metering infrastructure to synchrophasors to power quality and continuous point-on-wave monitoring. The platform, architected for speed, scalability, and ease of use, can be deployed through cloud providers—such as Amazon Web Services (AWS) and Microsoft Azure—and also as an on-premise solution.

Challenge: Customer and Community Engagement



Sample PredictiveGrid user interface

Project Overview

This project engaged PingThings, Ameren, and EPRI in evaluating the performance and capabilities of the PredictiveGrid platform when working with millions of streams of smart meter data. Ameren internally maintains a metadata table for 3.29 million smart meters across 1,387 substations in Missouri and Illinois. This table is presumed to be 80% to 90% accurate for phase and meter-secondary transformer mappings, but this has not been comprehensively tested as manual verification is prohibitively expensive.

The specific challenge was for PredictiveGrid to ingest more than 1 year of data and then implement, test, and enhance several large-scale machine learning algorithms for smart meter time-series data. The intended application was to develop predictions for correcting mis-specified phase designations and meter-transformer mappings to support the creation of hyper-accurate network models, which have high business value as foundational to an advanced grid. Since the best approaches are not known *a priori*, the pilot provided a test of the rapid prototyping and experimentation capabilities provided by the PingThings platform.

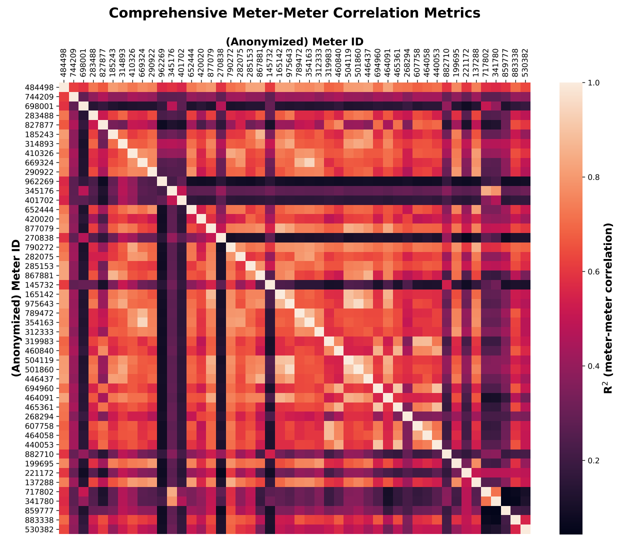
This proof-of-concept project was completed within the required 3-month timeline mainly through the efforts of a single individual. Approximately 12 months of data for 1 million smart meters were imported into the PingThings platform. Each meter reports voltage, real current, and reactive current measurements once every 15 minutes. Three "meter-meter similarity" algorithms were developed, tested, refined, and benchmarked:

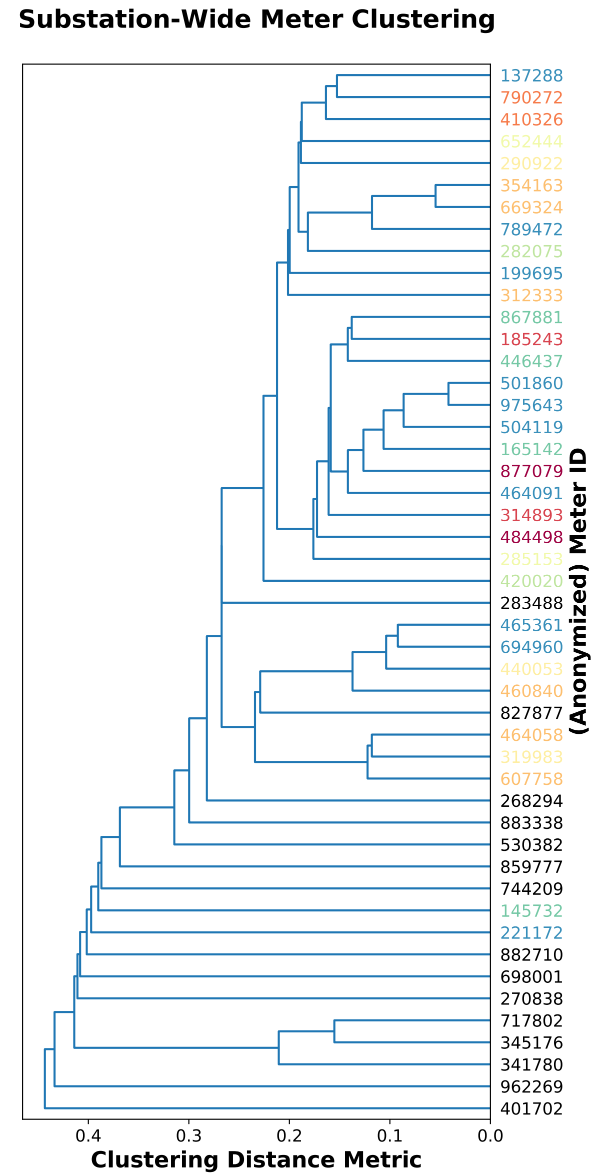
- **Pearson correlation**. A simple statistical correlation that ranges from 0 to 1.
- **Spectral clustering-based method.**[1] A machine learning method based on network theory.
- **Bottom-up grouping.**[2] A method based on statistical regression analysis that also yields estimates of physical features like resistance and reactance.

In parallel, the Ameren team assembled a solution from pre-existing AWS components (including S3 and EMR) to develop and test the same algorithms. This supported benchmarking all aspects of the PredictiveGrid in comparison to the AWS-assembled solution, including explicit cost, raw technical performance, time to value, and business process metrics.

Results & Learnings

PingThings deployed the PredictiveGrid platform, ingested and evaluated the quality of the supplied smart meter data, and prototyped and benchmarked the three algorithms. Rapid prototyping showed all algorithms to yield useful predictions, as shown in the accompanying figures. Both algorithm development





*Sample results from the methods used for inferring smart meter organization within a small substation: **(top)** heat map visualizing statistical similarity between all pairs of meters based on 1 full year of power, voltage, and current time-series data; and **(right)** hierarchical clustering results using similarity values (meter identification colors correspond to Ameren's internal grouping label)*

and large-scale testing were enabled by the platform's ability to store, load, and analyze data using best-in-class data science tools. Prototyping was performed on a small amount of ground-truth data in addition to the meter metadata table. Results suggest that the two more complex algorithms produced more accurate results, and next steps will examine this in more detail.

For the Pearson correlation method, complete data loading and analysis of all available data across 64 substations were achieved in under 40 minutes on the PingThings platform using a single compute node. This was more than twice as fast as the AWS-based solution while using 64 times fewer resources, all at significantly lower cost. The more complex algorithms, which each require millions of individual calculations, finished within 20 minutes for substation networks thousands of meters in size. All analytic methods tested were efficient enough to allow for easy parameter tuning and prediction scoring.

TESTIMONIAL: PingThings

Working with Ameren and EPRI has been great, and we have really appreciated the opportunity to show the industry just how powerful the PredictiveGrid is for all types of utility time-series data.

TESTIMONIAL: Ameren

The PredictiveGrid is a high-performance data analytics platform that can ingest and store grid data, enable easy interactions, and accelerate building artificial intelligence solutions.

The AWS-based analytics implementations by Ameren were hindered by limitations imposed by the AWS services and software but not apparent at the beginning of the process. Pilot results demonstrate that the PingThings platform provides technical advantages such as fast time-series data storage and loading and flexible and customizable analytics development to meet specific user needs. In addition, experimentation with time-series data is encouraged rather than penalized by the platform's pricing model.

This proof-of-concept project involved a one-time data export and transfer from Ameren through EPRI to PingThings. While challenges encountered resulted in some delays, project learnings will guide development of a more robust pipeline and schedule to transfer data continuously into the PingThings platform in an automated fashion.

Implications & Next Steps

Unlike many other data analytics offerings for advanced metering infrastructure, the PingThings platform assists utilities in creating their own solutions as opposed to being handed a "black box" machine that lacks observability and cannot be tuned. Importantly, utilities often underestimate the difficulty of data wrangling—getting all of the relevant data securely into one place in a form that is usable (ingestible) by analysis software. Regardless of the data analytics application or choice of platform, this portion of the process should be thoroughly assessed and tested during planning to refine the time estimate and resource commitment.

In the specific use case considered, the analytics developed and refined using the PingThings platform demonstrated high potential value as a viable alternative to field verification of phase and meter-transformer mappings, which requires each meter to be examined and measured at its location. Further, because the platform is an emerging alternative to the big data technologies of leading providers, a significant and valuable part of this project involved deploying an AWS implementation to support an objective comparison for Ameren.

For 2021, Ameren plans to continue to evaluate the PredictiveGrid platform and how it can fit into the utility's overall sensor and analytics architecture and strategy.

References

1. L. Blakely, M. J. Reno and W. Feng, "Spectral Clustering for Customer Phase Identification Using AMI Voltage Timeseries." Presented at 2019 IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, 2019, pp. 1-7, doi: 10.1109/PECI.2019.8698780.
2. T. A. Short, "Advanced Metering for Phase Identification, Transformer Identification, and Secondary Modeling." *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 651-658, June 2013, doi: 10.1109/TSG.2012.2219081.

Resources

Sean Murphy, PingThings
Chief Executive Officer
sean@pingthings.io

Scott Hixson, Ameren
Manager, Grid of the Future and Analytics Solutions
shixson@ameren.com

Ed Beroset, EPRI
Principal Technical Leader, Information and Communication Technologies
eberoset@epri.com

Resources

Erik Steeb, Incubatenergy® Lead
esteeb@epri.com; 650.680.6530

Anne Haas, Incubatenergy® Challenge Lead
ahaas@epri.com; 704.608.6314

incubatenergy labs

labs.incubatenergy.org

2020 Incubatenergy Labs Sponsors

Ameren

AMERICAN ELECTRIC POWER

conEdison, inc.

enel

Nebraska Public Power District
Always there when you need us

NEW YORK STATE OF OPPORTUNITY. NY Power Authority

PGE

SOUTHERN CALIFORNIA EDISON®
An *EDISON INTERNATIONAL®* Company

SRP

TVA

Xcel Energy®

3002020657

March 2021

Electric Power Research Institute
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA • 800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com