Pacific Northwest
NATIONAL LABORATORY

PNNL-33871

Accelerated Data Analytics for Power-System Time-Series (ADAPT)

Final Project Report

January 2023

JD Follum

C Lackner

H Wang

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
Under contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-33871

Accelerated Data Analytics for Power-System Time-Series (ADAPT)

Final Project Report

January 2023

JD Follum
H Wang
C Lackner

Prepared for
the U.S. Department of Energy
Under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Abstract

Over the past decade, electric utilities have made significant progress in deploying networks of phasor measurement units (PMUs). The high-speed measurements provided by PMUs are valuable for offline analysis, but the sheer volume of data presents a challenge for utilities. The Accelerated Data Analytics for Power-System Time-Series (ADAPT) project was created to address this challenge by developing SciSync, an open-source software tool that enables utilities to rapidly read, process, analyze, and review grid measurements. SciSync was designed to provide the capabilities of Archive Walker, a research tool developed by PNNL, along with the reliability and performance of commercial-grade software. Some of the key benefits of SciSync over Archive Walker include:

- SciSync is professionally coded in a programming language well-suited to application development
- Interfaces with the PI System and openHistorian databases commonly used by utilities are seamless in SciSync
- Analysis configurations in SciSync can readily be applied to different sets of input signals
- SciSync achieves faster performance, both in processing data and displaying results
- When displaying results, SciSync seamlessly transitions from summaries to full-fidelity measurements.

This report discusses the key outcomes of the ADAPT project. It details the design of SciSync, tests that were performed to evaluate its performance, and use cases that motivate its deployment. These discussions demonstrate that SciSync is a capable tool with benefits to electric utilities and grid researchers that will grow over time.

Acronyms and Abbreviations

ADAPT	Accelerated Data Analytics for Power-System Time-Series
BPA	Bonneville Power Administration
DLL	Dynamic Link Library
DOE	Department of Energy
GMCL	Grid Modernization Laboratory Consortium
GPS	Global Positioning System
GUI	Graphical User Interface
OBAT	Oscillation Baselining and Analysis Tool
PMU	Phasor Measurement Unit
PNNL	Pacific Northwest National Laboratory
PPMV	Power Plant Model Validation
RMS	Root Mean Square
ROCOF	Rate-of-Change-of-Frequency
SCADA	Supervisory Control and Data Acquisition
TVA	Tennessee Valley Authority

Acknowledgments

The project team would like to express their gratitude to Dr. Ali Ghassemian, manager of the Department of Energy Office of Electricity Advanced Grid Modeling Program, for funding this work. The authors are grateful to Lindsey Sheets and the rest of the AVEVA team for providing guidance and licenses for the PI System to facilitate development and testing. We also gratefully acknowledge Bonneville Power Administration, particularly Tony Faris and Steve Yang, for providing access to the field-measured data used in testing. The project would not have been a success without these contributions.

Contents

Figures

Tables

1.0 Introduction

Over the past decade, electric utilities have made great strides in deploying advanced wide-area measurement systems. These systems enhance the situational awareness of system operators and reliability coordinators by enabling a variety of online monitoring and control applications. After being used in the online environment, measurements are commonly stored in archives that can span years. Utilities vary greatly in the degree to which they use these archives to support offline applications such as post-mortem analysis and model validation. For many, the sheer volume of data presents a challenge in identifying periods of interest that warrant further analysis. The Accelerated Data Analytics for Power-System Time-Series (ADAPT) project discussed in this report was created to address this challenge.

Specifically, the goal of the ADAPT project was to increase the value of archived measurements by better enabling utilities to rapidly read, process, analyze, and review the data. To accomplish this goal, PNNL partnered with Grid Protection Alliance (GPA) to produce SciSync, an open-source software tool for the analysis of power grid measurements. SciSync was designed to provide the capabilities of Archive Walker (J. D. Follum et al. 2018), a research tool developed by PNNL, along with the reliability and performance of commercial-grade software. Archive Walker has been very valuable in supporting PNNL's research of advanced power system measurements by allowing researchers to quickly prototype new analysis methods and test them on archived measurements. However, electric utilities are reluctant to utilize research-grade tools that are not developed and supported by a software vendor. The ADAPT project was established to make the benefits of Archive Walker available to utilities in a package that they can readily deploy.

With this objective, GPA was ideal to lead the development of SciSync. GPA specializes in the development and deployment of open-source tools for the management of power system measurements. Many major utilities already use GPA tools as part of their measurement infrastructure, making SciSync's integration seamless. For example, SciSync can natively read data from GPA's open-source archiving software, openHistorian. To further SciSync's broad appeal to the power industry, it was designed to also read data from AVEVA's PI System and the PDAT file format developed by Bonneville Power Administration (BPA).

AVEVA and BPA were both project partners, and their participation was key to developing SciSync. AVEVA supported testing and development by providing licenses for their PI System, which is widely used in the electric power industry to archive measurements. Along with the initial concept and funding for Archive Walker in previous projects, BPA provided input on SciSync's design and field-measured data for testing under the ADAPT project. The contributions from GPA, AVEVA, and BPA were all critical in making the ADAPT project a success.

Along with a tool useful to industry, the project's successful outcome has implications for future research. SciSync is free to download and can therefore support researchers that may not have funding to purchase other commercial-grade tools, such as university students. As an open-source tool, SciSync can even be modified to test researchers' own analysis algorithms. SciSync also provides a natural path for future PNNL research to be delivered to utilities. With the core software developed, algorithms developed by PNNL can now be integrated relatively easily for efficient transfer to utilities. Based on these use cases, SciSync is expected to have a positive impact on the industry's ability to analyze advanced power system measurements in both the short and long term.

The remainder of this report is structured as follows. A review of advanced power system measurements is provided in Chapter 2.0. A summary of the Archive Walker tool is then provided in Chapter 3.0 to establish a reference point for SciSync's development, which is described in Chapter 4.0. The performance tests that compared Archive Walker with SciSync, along with the environment they were performed in, are then described in Chapter 5.0. A set of example use cases for SciSync are provided in Chapter 6.0 before concluding remarks in Chapter 7.0.

2.0 Advanced Power System Measurements

Traditional power system measurements are produced by supervisory control and data acquisition (SCADA) systems. In SCADA systems, measurements of voltage and current magnitude are commonly reported every one to six seconds. Time tags are assigned to the measurements when they reach the control center, meaning the exact time that the measurement was taken is unknown. These measurements are critical to the safe and reliable operation of power systems, but they cannot capture some important grid behaviors. For example, Figure 1 shows that SCADA measurements cannot accurately represent some power system oscillations due to their low reporting rate. In contrast, the higher speed measurements from a phasor measurement unit (PMU) are able to show the oscillation in detail. SciSync is primarily intended to analyze the measurements collected from PMUs, which are termed *synchrophasors*.

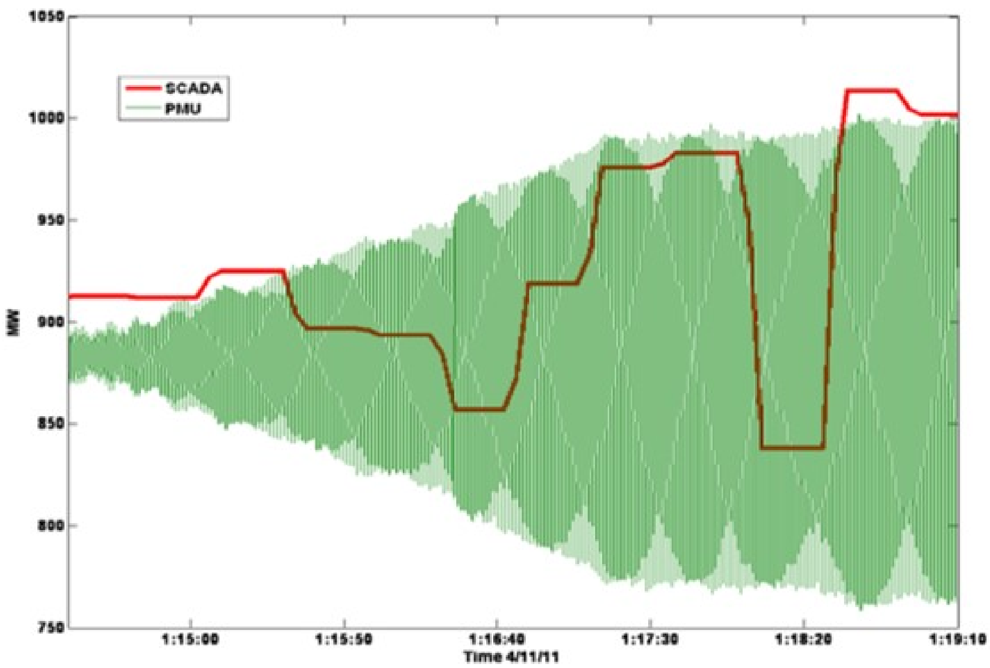

Figure 1. Comparison between SCADA measurements and more advanced PMU measurements during an oscillation. This figure is reproduced from (Silverstein 2015) and was originally created by Dominion Virginia Power.

The reason the synchrophasor measurements in Figure 1 are able to capture the oscillation is because they are reported at a much higher rate than SCADA measurements, typically 30 or 60 times per second in 60 Hz systems. Along with the higher reporting rate, synchrophasors are timestamped at the point of measurement, typically using the global positioning system (GPS). This means that measurements generated by geographically dispersed devices are synchronized, a characteristic that is crucial for several wide-area analyses. The synchronization also allows voltage and current angles to be measured. In addition to voltage and current phasors (a magnitude with an associated angle), PMUs report measurements of frequency and rate-of-change-of-frequency (ROCOF). For further reading on the history, applications, and widespread adoption of PMUs throughout the world, see (Phadke and Bi 2018).

As the use of PMU data has increased, the need for higher resolution synchronized measurements has been identified. Measurements of voltage and current waveforms, referred to as point-on-wave (POW) measurements, are necessary to analyze high-speed events, such as the behavior of inverters used to connect renewable resources to the grid. While synchrophasor measurements are typically collected 30 or 60 times per second, POW data is typically collected thousands of times per second. This presents a major challenge for the electric utility industry, but one that SciSync can help address. For many applications of POW measurements, only a short snapshot is required. However, there are practical applications that require long record lengths, and instruments that continuously record POW data are commercially available. Long record lengths will also be an important aspect of POW-based research moving forward (J. D. Follum et al. 2021). For this reason, SciSync was also designed to allow for the ingestion of POW measurements. As described later in the report, BPA provided a large archive of POW data collected from a commercial building to support the testing of this capability.

For both PMU and POW measurements, high reporting rates are an important feature that enable a much broader set of analyses than traditional SCADA measurements. However, it also introduces significant challenges when attempting to analyze archived data. The SciSync tool can significantly reduce this burden for electric utilities. Before describing SciSync, the research software it was based on will be described in the following chapter.

3.0 Archive Walker

The Archive Walker software was initially conceived by BPA as a tool to identify periods of interest in large volumes of PMU data. Along with helping BPA analyze their own data, they hoped that an open-source tool would support utilities that were in the early stages of adopting synchrophasor technology and encourage further growth. The tool was also intended to accelerate the testing and implementation of new analysis methods by avoiding the development of custom code to perform data reading, cleaning, and preprocessing for every new concept. The SciSync software, which is the focus of this report, will go even further to achieve these outcomes by providing the industry with a commercial-grade version of Archive Walker. To establish the background for SciSync's development, this chapter will explain the development history, capabilities, and limitations of Archive Walker.

3.1 Development History

The initial development of Archive Walker began in 2015 and was performed by PNNL with funding from BPA's Technology Innovation Program. The project concluded in 2018 and resulted in a useful tool for identifying and displaying grid disturbances in PMU data (J. D. Follum et al. 2018). The tool was posted publicly on GitHub, where newer releases continue to be published (PNNL 2018).

In alignment with BPA's initial vision, DOE continued funding support for Archive Walker's development as part of a Grid Modernization Lab Consortium (GMLC) project to create an open-source suite of tools for synchrophasor data analysis (Etingov et al. 2020). As part of the suite, Archive Walker served as a data hub for the other tools by identifying events suitable for analysis with, for example, the Power Plant Model Validation (PPMV) tool or Oscillation Baselining and Analysis Tool (OBAT). These potential use cases are outlined in (J. Follum et al. 2020). By the time the GMLC project concluded in 2020, Archive Walker had become an integral part of PNNL's measurement-based grid research. It has since been used to support several projects and the testing of new algorithms for synchrophasor data analysis (Follum, Agrawal, and Etingov 2021; Follum, Becejac, and Huang 2021; Follum, Becejac, and Etingov 2021). The following section provides details on the capabilities that have made Archive Walker such a useful tool for research, along with the limitations that SciSync addresses to better meet the needs of utilities.

3.2 Capabilities and Limitations

Many of Archive Walker's characteristics, capabilities, and limitations are derived from the early design choice to develop the software in two distinct pieces, as depicted in Figure 2. The analysis engine was written in MATLAB, a scripting language widely used by power system researchers, while the graphical user interface (GUI) was written in .NET. The use of .NET for the GUI allowed for greater flexibility in design choices that streamline the process of setting up analyses and reviewing results. Keeping the analysis engine in a scripting language familiar to power system researchers makes the integration of proof-of-concept code for new algorithms into Archive Walker fast and straightforward. With this architecture, the user begins by setting up an analysis in the GUI. Once the setup is complete, the GUI calls the analysis engine, which is compiled as a .NET assembly. The engine reads the configuration file generated by the GUI and analyzes the data accordingly. When the analysis is complete, results returned by the analysis engine are displayed in the GUI. These three steps of setting up, running, and reviewing an analysis will be discussed in additional detail and used to highlight the capabilities and limitations of Archive Walker.

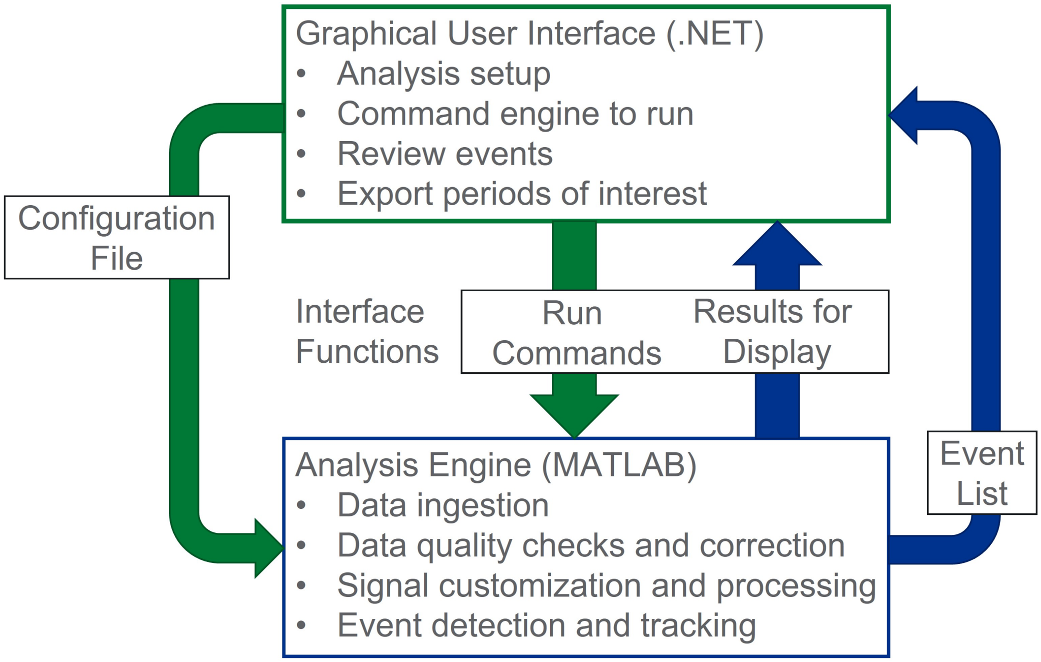

Figure 2. Summary of Archive Walker's architecture, which includes a MATLAB-based analysis engine and a graphical user interface developed in .NET.

3.2.1 Analysis Setup

Setting up an analysis in Archive Walker involves connecting to a data source and specifying the functions that should be applied to each measurement channel it contains. To do this, the user begins with the *Data Source* tab depicted in Figure 3 and then proceeds through the remaining tabs to apply data quality checks and other functions.

Archive Walker was initially designed to read file-based measurement archives but was later modified to ingest data from AVEVA's PI System and GPA's openHistorian. Though this capability is functional, it is not optimal. For MATLAB to retrieve data from PI and openHistorian, dynamic link libraries (DLLs) had to first be written in a lower-level programming language. The analysis engine calls these DLLs to retrieve data, a process that is relatively slow and requires intermediary GUIs to configure. Another shortcoming of Archive Walker's interface with PI is that it utilizes an outdated method of retrieving data that does not allow measurement channels to be grouped by measurement location. In contrast, the SciSync software utilizes the modern PI System's Asset Framework to understand which signals were collected from the same PMU. SciSync interfaces seamlessly with the PI System and openHistorian to make data access easier for the user while achieving better performance.

Once a data source is established, the user can apply a variety of functions to the measurements. As listed in Figure 4, these functions include data quality checks, mathematical and engineering operations, signal processing functions, and event detectors. Detailed descriptions of these capabilities are provided in (PNNL 2018; J. D. Follum et al. 2018). Figure 3 shows what it looks like to add available signals from the *Signal Selection Panel* on the right side of the screen to signal processing functions. Note that the signals must be specified at every step of the analysis. The GUI was designed to streamline this process, but changing the signals that an analysis is set up for still requires the change to be made carefully and for every function. As described in the next chapter, a different approach was used in SciSync to make it simple to apply an analysis setup to different sets of signals.

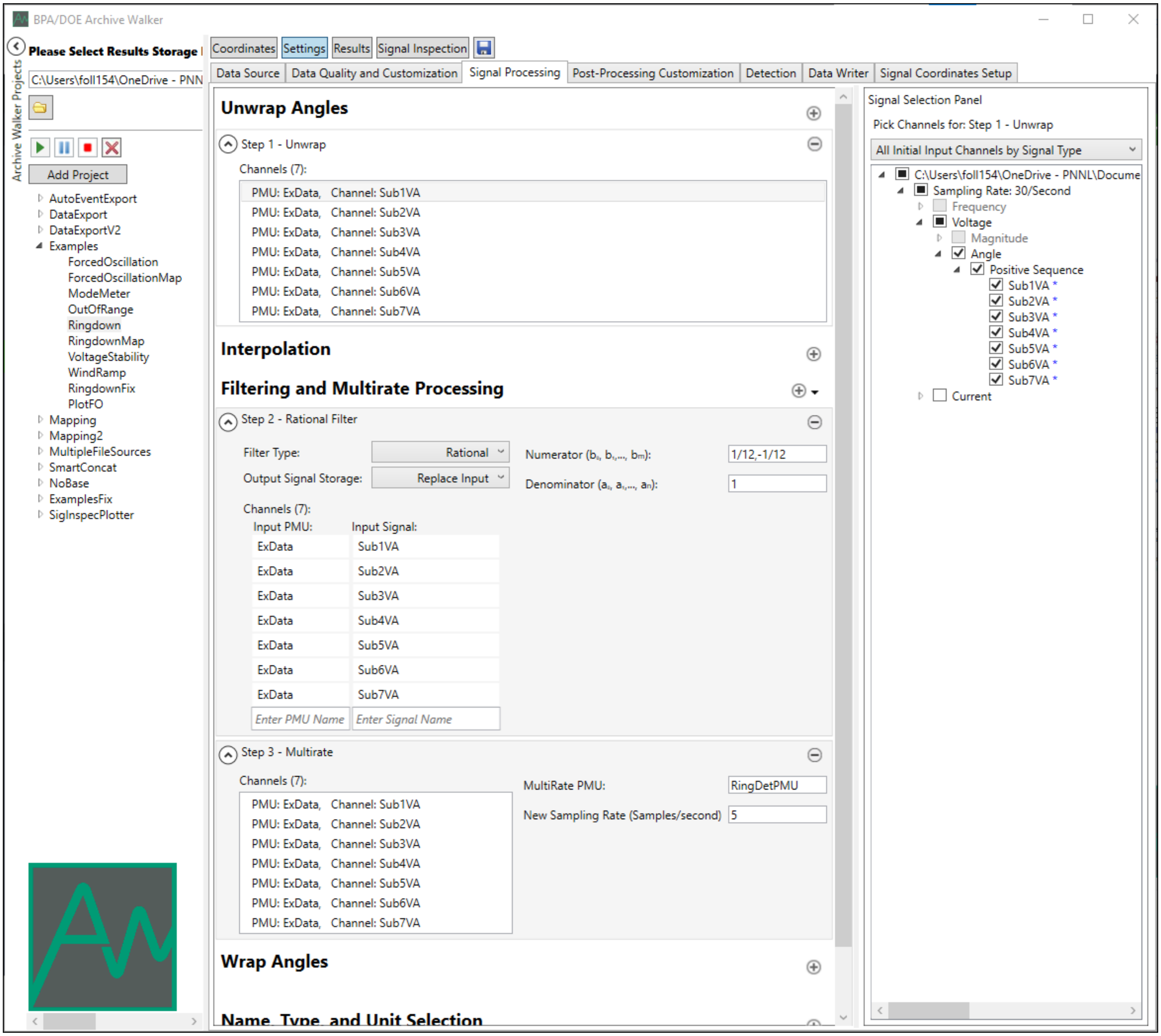

Figure 3. Screenshot of the Archive Walker user interface on a tab used to set up analyses.

3.2.2 Running Analyses

Once the analysis is configured in Archive Walker, the user prompts the GUI to run the analysis. In turn, the GUI calls a function included in the analysis engine's compiled MATLAB

Data Quality Filters

- Status Flags
- Zeros
- Missing
- Nominal Voltage
- Nominal Frequency
- Outliers
- Stale Data
- Data Frame
- Channel
- Entire PMU
- Angle Wrapping

Signal Customizations

- Scalar Repetition
- Arithmetic
  - Addition
  - Subtraction
  - Multiplication
  - Division
  - Exponential
  - Sign Reversal
  - Absolute Value
- Complex Numbers
  - Real Component
  - Imaginary Component
  - Angle Calculation
  - Complex Conjugate
  - Phasor Creation
- Engineering
  - Power Calculation
  - Signal Type/Unit
  - Metric Prefix
  - Angle Conversion
- Duplicate Signals

Signal Processing

- Wrap/Unwrap Angles
- Interpolation
  - Linear
  - Constant
  - Cubic
- Filtering
  - General
    - Rational
    - High/Low-Pass
  - Specialized
    - Frequency Derivation
    - Running Average
- Multirate Processing
- Name, Type, Unit Selection

Event Detectors

- Forced Oscillations
- Ringdown Oscillations
- Out-of-Range Events
- Wind Ramping

Figure 4. List of analysis capabilities available within Archive Walker.

code to perform the analysis. The software processes data fast enough to make analysis of even very long record lengths feasible for research purposes. For example, the analyses reported in (Ahmad et al. 2021) and (Follum, Becejac, and Huang 2021) each required the analysis of over a year of PMU data and took several weeks to run. Though sufficient for research, Archive Walker's processing performance is not optimal. As a scripting language, MATLAB is less efficient than alternative programming languages. Perhaps even more significant, the analysis engine was written to be easily understood and modified, without computational efficiency as a primary goal. As will be demonstrated in Chapter 5.0, SciSync achieves much better performance that makes it a more useful tool for industry applications.

3.2.3 Reviewing Results

One of the challenges in analyzing large volumes of measurement data is managing the storage requirements for results. Archive Walker uses a tiered approach to allow a detailed review of processed measurements while minimizing storage requirements.

The first tier of results includes text descriptions of detected events. These event records are displayed in the tables on the right side of Figure 5, which specify the timing of the event and which signals were involved. The second tier is composed of maximum and minimum values for each minute of processed data. This information is used to provide an overview of all input signals to a detector, as in the topmost plot in Figure 5. The two events listed in the tables can be seen as large positive and negative deviations in the plot. In this case, the second event is highlighted in red because it is selected in the table. Once the user has used this high-level information to select a period for detailed review, they click the GUI's *Retrieve Detail* button. This action brings up the remaining plots in Figure 5, which show the fully processed data (second plot from the top) along with details of the detector's performance (bottom plot). Storing all this displayed information would require a great deal of disk space. Instead, the third tier of results includes only enough information to restart the analysis from the beginning of any given minute, such as initial filter conditions.

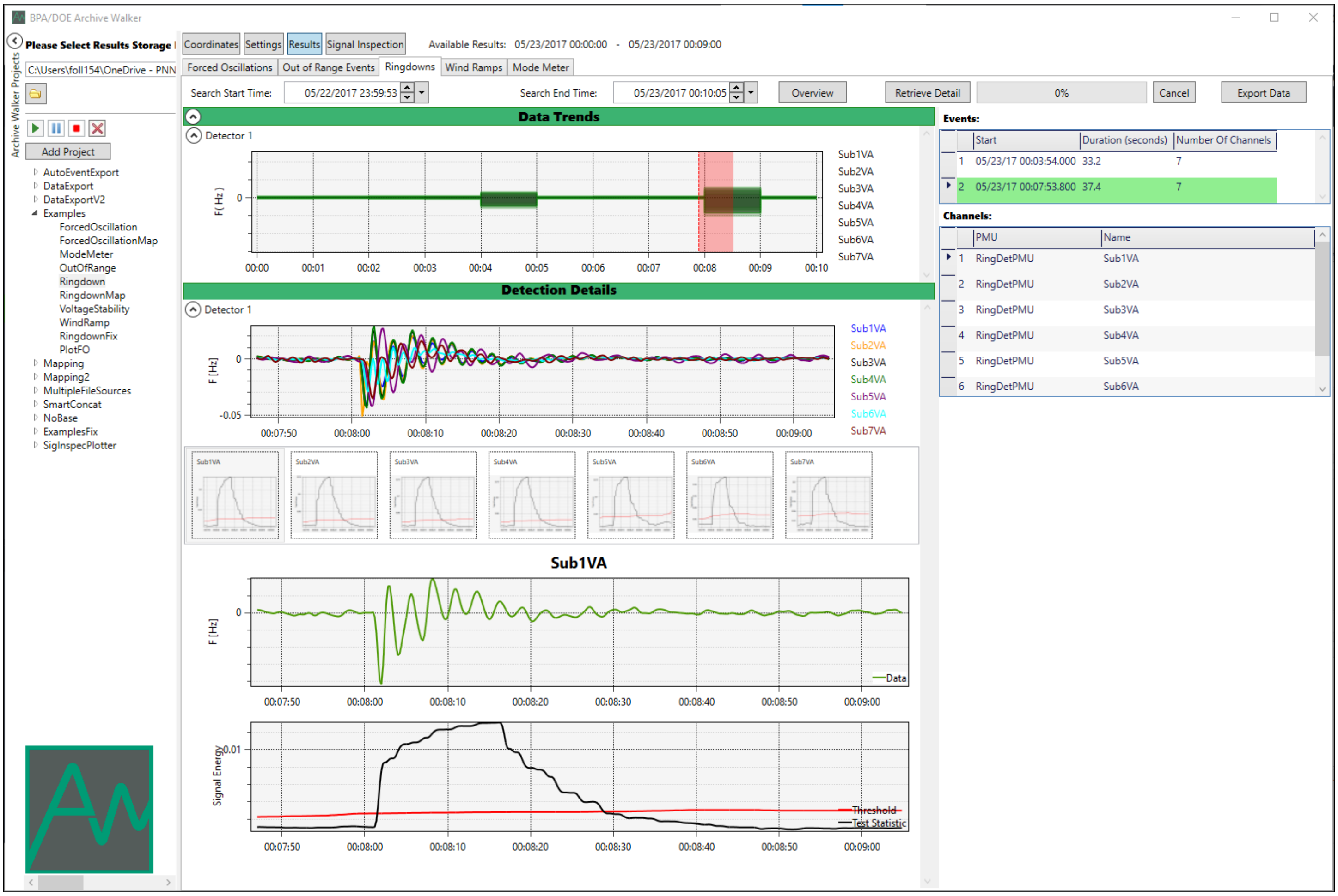

Figure 5. Screenshot of the Archive Walker user interface on a tab used to review analysis results.

This approach to result storage keeps requirements on disk space manageable while also allowing detailed results to be observed. The limitation of this approach is that results are only ever displayed at a very coarse resolution (maximum/minimum over a minute) or at full resolution. As will be discussed in the next chapter, SciSync automatically adjusts the resolution of displayed results based on the amount of data to be displayed. SciSync also stores all results in an efficient structure to avoid the multi-step process required in Archive Walker.

3.3 Summary

The Archive Walker software provides excellent functionality to support research with advanced power system measurements. The tool's architecture allows rapid modification of the analysis engine while providing an effective user interface. It's wide array of analysis capabilities gives it great flexibility to meet a broad set of research needs. As a result, it has been used by PNNL to analyze thousands of hours of data in support of several different projects. Still, research-grade tools like Archive Walker come with limitations that make them difficult for electric utilities to adopt. The ADAPT project was established to make Archive Walker’s capabilities more broadly available to the industry through the development of SciSync. The design and capabilities of SciSync are described in the next chapter.

4.0 SciSync

SciSync was designed to replicate the core functionality of Archive Walker while achieving improved performance. It's development, marketing, and support by GPA, a trusted software vendor in the electric power industry, make it readily available to utilities. The tool is available to download for free ("Version 1.0.0 of SciSync" 2023), which is also attractive to utilities. Its open-source release under the permissive MIT license also means that third parties, including PNNL, can extend its capabilities to meet their needs. The software was specifically designed to support such modifications. This and other design principles were determined after identifying the set of use cases outlined in Section 4.1. The architecture used to achieve the design principles is then outlined in Section 4.2. This chapter concludes with a comparison between SciSync and Archive Walker in Section 4.3.

4.1 Use Cases

The use cases summarized in this section were identified based on GPA's prior experience with utilities and specific discussions with BPA and the Tennessee Valley Authority (TVA). They span everything from the analysis of a few minutes of data by a utility engineer to the review of years of data by a researcher. Identifying these use cases guided the design of SciSync and resulted in a tool that can meet a broad set of needs. The five use cases are reviewed in the following sections.

4.1.1 Utility - Simple

This case focuses on a system engineer at an electric utility interested in a known system disturbance associated with a specific asset. Because the event is already known, only a small time window, perhaps less than 10 minutes, must be analyzed. The asset is also known, so measurements are only needed from a few locations.

In this use case, SciSync is simply providing access to the measurements for quick review and export so that a specialized analysis can be performed. Event detection, signal manipulation, and summarizing visualizations are unnecessary. This use case was not part of PNNL's initial conception and was instead added based on solicited feedback from a utility. It requires seamless integration with data sources, as described in Section 4.2.1.

4.1.2 Utility - Complex

A system engineer is also the user in this case, but here they are interested in a more complex analysis involving a longer time window, say, less than an hour, and/or a wider set of measurement locations. To meet the user's needs in this case, the tool must be able to manage many input signals and provide useful visualizations to browse results. To meet these needs, SciSync was designed with a simple interface for quickly assigning multiple signals to an analysis (see Figure 9).

4.1.3 Utility - Summary

The final case focused on a utility engineer addresses the need to analyze operation patterns over the course of multiple months. Engineering functions are needed to calculate values such as power flow or voltage angles across a transmission path. Rather than identifying a specific period of interest, the results need to provide statistics and visual summaries for the created signals. This use case was also developed based on specific feedback from a utility and motivated the inclusion of the widgets described in Section 4.2.4 that summarize created signals.

4.1.4 Industry Research

The user in this case is a research engineer at a utility, national laboratory, or industry organization. The researcher is envisioned to have a close tie to a utility, allowing for access to a large archive of measurements from several locations. Event detection capabilities are important in this use case so that periods of interest can be identified for further study. The researcher also needs to be able to implement custom analytics and review analysis results in detail to evaluate performance and make adjustments. With this use case in mind, SciSync was designed to store analyzed signals in full resolution for detailed review, as described in Section 4.2.4.

4.1.5 Academic Research

The final use case focused on research performed at an academic institution. These users typically have access to example datasets from a utility consisting of measurements from many locations collected over the course of a short time window of, say, 10 minutes. The simple integration of custom analytics is particularly important in this case, as is the need for visualizations that are easily interpreted while allowing results to be reviewed in detail. SciSync was specifically designed with a modular architecture to allow custom analytics to be integrated easily in support of academic research.

4.2 Architecture

SciSync's architecture was designed to achieve the broad set of use cases outlined in the previous section. To this end, it was structured around the four components listed in Figure 6: Data Source, Template, Task, and Results. These components are described in the following subsections.

4.2.1 Data Source

As the name implies, the *Data Source* establishes the connection between SciSync and measurement archives. At present, SciSync can retrieve data from three sources. The first is the PDAT file format developed by BPA to store PMU data. The others are AVEVA's PI System ("AVEVA PI System" 2022) and GPA's openHistorian ("openHistorian" 2022), both of which are in wide use by electric utilities. SciSync can read either PMU or POW data from openHistorian. If POW data is selected, SciSync automatically estimates phasors from the measurements as they are read to facilitate further analysis with the tool. Future upgrades to SciSync could include analytics for direct analysis of POW data and the addition of new data sources.

Once a *Data Source* is established, measurement values can be displayed in simple plots to help the user design the rest of the analysis. For example, the plot in Figure 7 could be used to determine the nominal voltage for the substation bus, which could impact thresholds for data

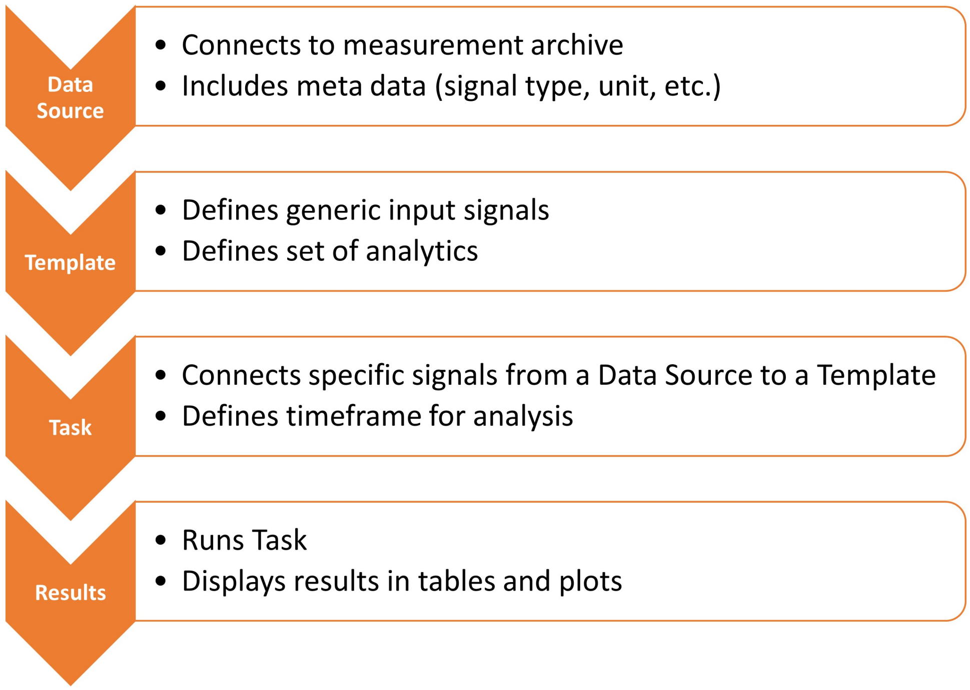

Figure 6. Summary of the four components of SciSync's architecture.

quality checks set up by the user. SciSync also retrieves meta data that specifies signal types, units, and associations. A particularly important association is which measurements belong to which PMUs. For the PI System, this meta-data is made available to SciSync via the Asset Framework capability.

In regard to data sources, the key difference from Archive Walker is how seamlessly the PI System and openHistorian data sources are integrated in SciSync. Archive Walker is not configured to use the PI Asset Framework, and therefore cannot organize measurement channels by PMU. Further, Archive Walker requires intermediate tools to configure both the PI System and openHistorian as data sources, whereas everything is contained within SciSync. SciSync is also able to retrieve data from these sources much faster than Archive Walker, as will be demonstrated in Chapter 5.0.

4.2.2 Template

Unlike Archive Walker, analyses in SciSync are not configured for specific measurements. Instead, SciSync *Templates* specify which functions to apply to a generically defined input signal. For example, the *Template* in Figure 8 defines a frequency input signal simply named *Signal 1*. Once input signals are defined, a wide array of analytics can be applied to them. The set of available analytics is nearly identical to the set listed in Figure 4 for Archive Walker. Each time a function is applied, it creates a new version of the input signal. The last step in setting up a *Template* is to specify which signals should be made available on the *Results* tab. As discussed in the *Results* section, this is an important feature because SciSync handles result storage very differently than Archive Walker.

4.2.3 Task

The *Task* portion of SciSync is where a *Template* is applied to a specific set of measurements. This is accomplished by mapping signals from a configured *Data Source* to defined *Template* inputs, as in Figure 9. This approach makes it easier to change the signals an analysis should be applied to than Archive Walker, which requires the signals to be updated at every step during configuration. The final step in the *Task* portion of SciSync is to specify the time range of interest. Once done, the user clicks the *Run Task* button to begin the analysis.

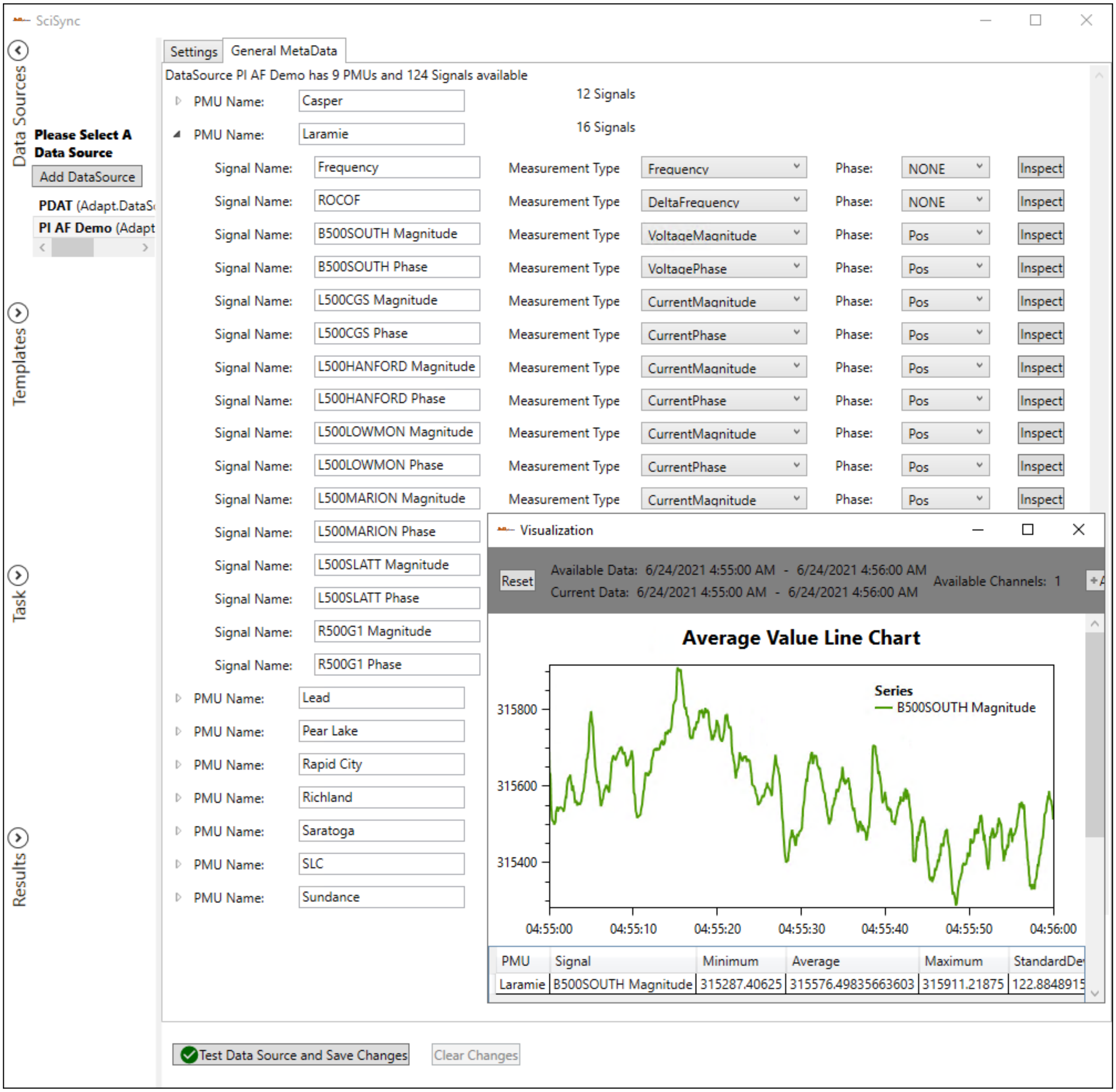

Figure 7. Screenshot of SciSync's Data Sources tab showing measurements from a configured source.

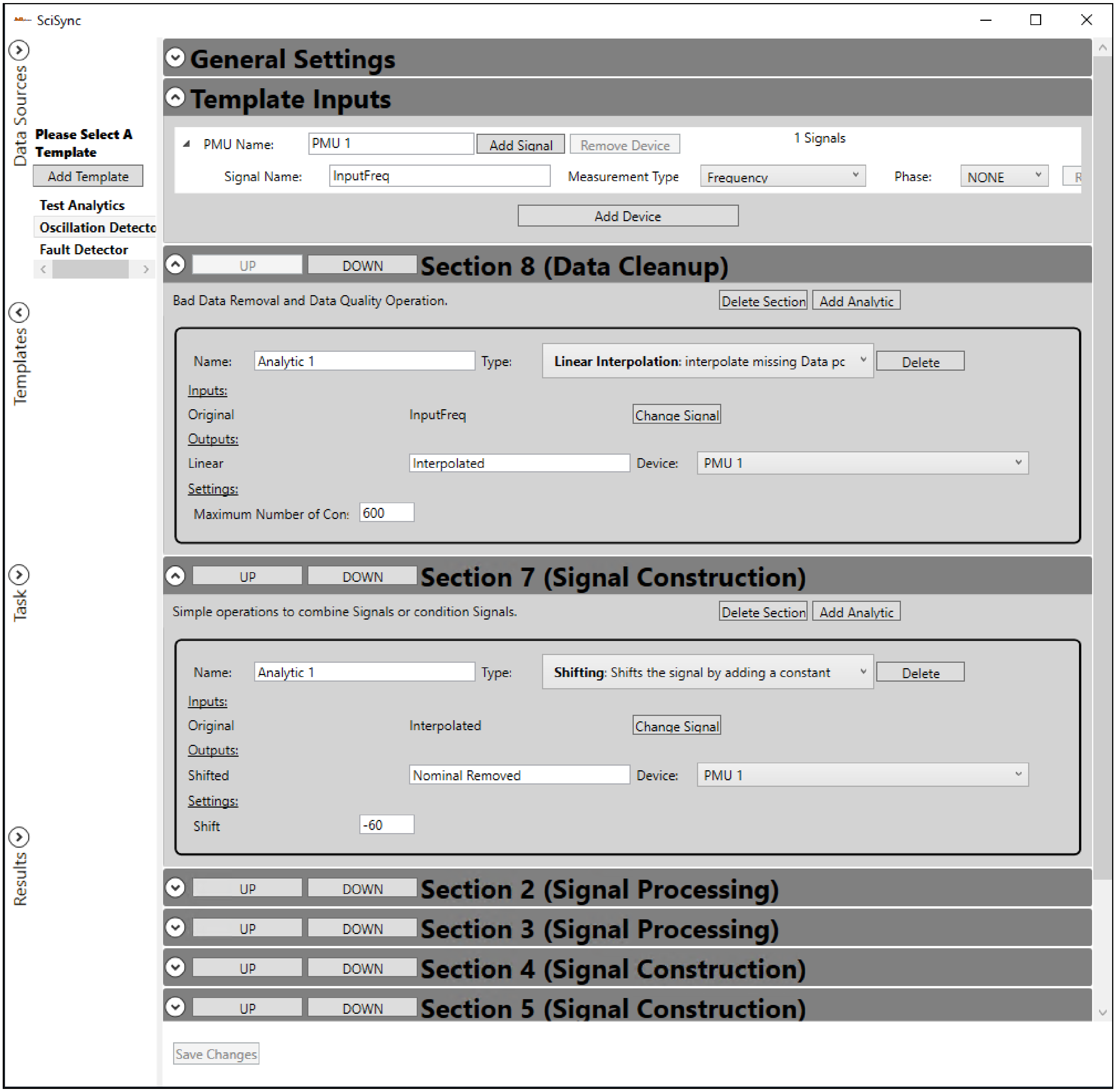

Figure 8. Screenshot of SciSync's Templates tab showing the configuration of an input signal and two analytics.

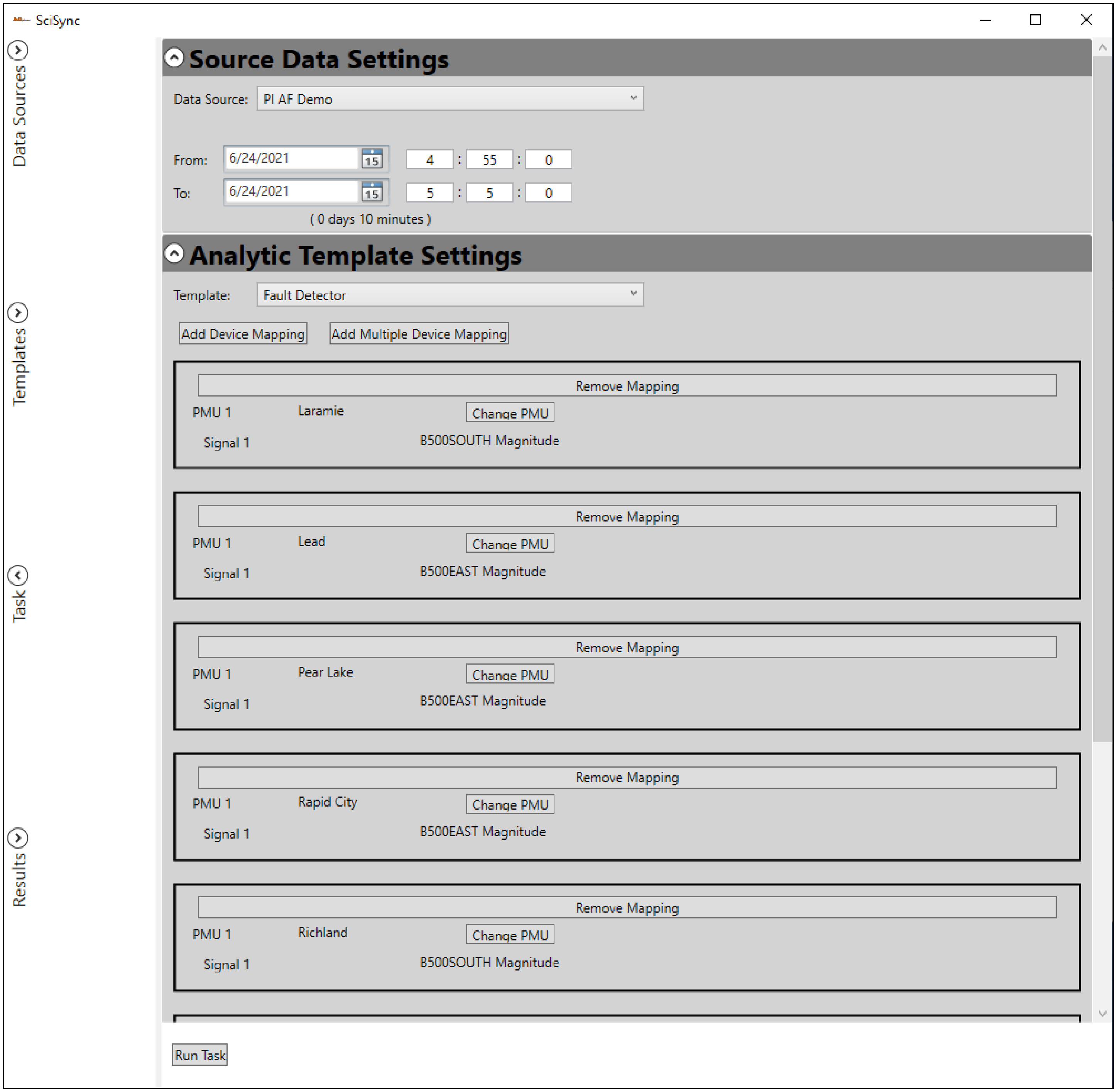

Figure 9. Screenshot of SciSync's Task tab where measurements sourced from AVEVA's PI System are assigned to a Template .

4.2.4 Results

As mentioned in the *Template* section, SciSync takes a different approach to result storage than Archive Walker. Every signal specified as a *Template* output is first stored at full resolution.

Then, summary statistics are calculated for various intervals (second, minute, hour, etc.). This information is used to seamlessly transition between signal summaries for long time ranges to full fidelity data when zooming in on a plot. In this way, the multi-step process for reviewing results in Archive Walker is avoided. As a trade-off, SciSync requires additional storage, at least temporarily. Where Archive Walker retains all records until the user deletes them manually, SciSync automatically clears the result storage every time a new analysis begins. SciSync allows users to export results and generated signals that they wish to retain.

The *Results* tab uses a variety of widgets to effectively display the output of an analysis. Specific examples with screenshots will be provided in Chapter 6.0, which describes example use cases. A summary of these widgets follows:

Statistics Overview Table Summarizes the values and availability for each signal

Summary of Detected Events Provides a summary table of detected events for each input signal, including the total time the signal was triggering an event detection

List of Detected Events Lists each detected event with summarizing information such as the input signal's deviation

Event Timeline Chart Indicates event occurrences for all input signals

Event Occurrence Chart Uses shading to indicate the period during which events occurred

Average Line Chart Displays traces for selected signals that automatically adjust measurement resolution based on the plot's stime range

Range Chart Displays the minimum, average, and maximum value of selected signals at resolutions that automatically adjust based on the plot's time range

4.3 Comparison with Archive Walker

Though SciSync was designed to replicate the core functionality of Archive Walker, there are important differences in the way this functionality is implemented. These differences, which are summarized in Table 1, span from data ingestion to the display of results.

In regard to retrieving data, SciSync interfaces seamlessly with the PI System and openHistorian databases commonly used by utilities. Archive Walker requires an additional user interface to establish a connection to these data sources, while SciSync only requires the user to enter a few pieces of information in its GUI. The actual process of retrieving measurements from these data sources is also streamlined for rapid processing in SciSync.

Once measurements are available, SciSync can rapidly change which measurement channels are fed into a set of analytics. Analyses in Archive Walker are configured for a specific set of signals. Changing the signals is possible, but requires changes at every step in the analysis. SciSync avoids this challenge by allowing *Templates* to be configured for generic input signals. The specific signals are then assigned to a *Template* when a *Task* is configured.

The tools also handle result storage differently when an analysis is performed. Archive Walker is intended to create a permanent record of results to support research. To limit storage requirements, it stores summarizing results at a very low resolution along with enough information to retrieve detailed results by rerunning a selected portion of the analysis. In contrast, SciSync temporarily stores signals of interest at full resolution along with summaries at multiple lower resolutions. This approach allows the user to quickly zoom from low resolution data for long record lengths to high resolution data for short record lengths.

These differences in design were paired with more efficient coding practices in programming languages, as opposed to the slower MATLAB scripting language used for Archive Walker. As discussed in the next chapter, SciSync's development by an experienced software development company led to marked improvement in the tool's processing speed over Archive Walker.

Table 1. A summary of differences between Archive Walker and SciSync.

Aspect	Archive Walker	SciSync
Database Interfaces	Requires Additional User Interfaces	Seamless
Data Retrieval	Inefficient	Rapid
Signal Configuration	Specific Requires Changes at Multiple Steps	Generic Rapidly Modified
Result Storage	Permanent Low Resolution with Rerun Capability	Temporary Full Resolution with Summaries
Coding	Research Grade Inefficient	Professional Efficient

5.0 Performance Evaluation

Part of the ADAPT project's goal was to better enable utilities to rapidly read, process, analyze, and review measurement data. To evaluate whether this goal was being achieved, two tests were performed to compare SciSync with Archive Walker. The specifics of these tests and their results are discussed in Section 5.2. First, the testing environment that was set up at PNNL to perform the evaluations will be described in Section 5.1.

5.1 Testing Environment

To properly evaluate SciSync, a testing environment with access to all data sources of interest was needed. The environment was built to store measurements in three archiving systems in a way that reflects actual industry practice. Both synchrophasor (PMU) and point-on-wave (POW) measurements for testing were provided by BPA. A diagram of the testing environment is provided in Figure 10.

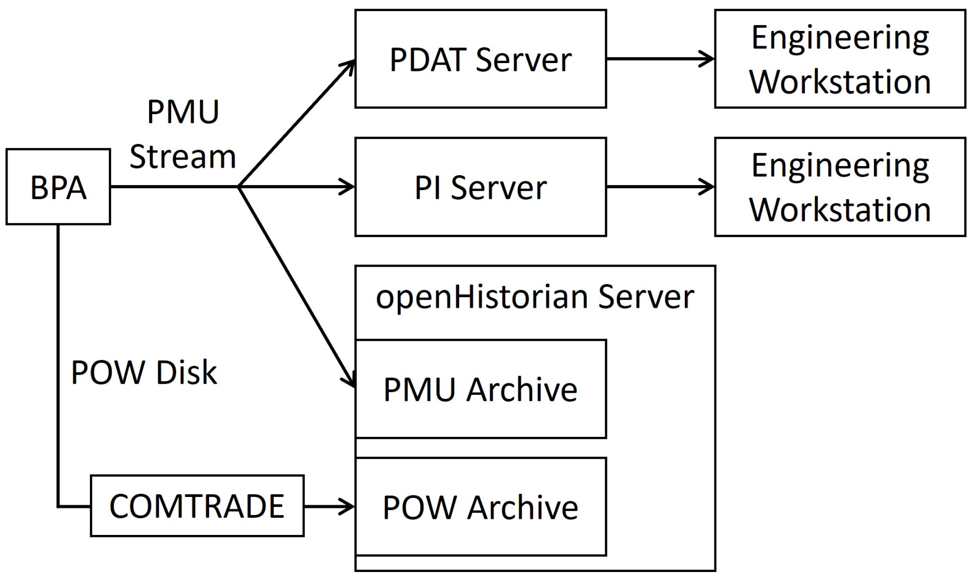

Figure 10. Block diagram of the testing environment set up to evaluate SciSync with various data sources.

A live stream of synchrophasor measurements was used to transfer the PMU data to PNNL. In total, approximately 1.2 terabytes (TB) of PMU data from 26 substations spanning March 12 through October 27, 2021 was streamed. Once at PNNL, the stream was split and passed to three servers, one for each of the data sources. The PDAT server ran BPA's custom software that generates and stores PDAT files in an archive. The second server ran AVEVA's PI System. As with the PDAT files, an engineering workstation was set up to remotely access the data on the servers to reflect the way an electric utility would manage access. The third server was initially configured to store the streaming PMU data in openHistorian.

A second openHistorian archive was later established to store POW measurements provided by BPA. The data was provided to PNNL on an external hard drive in MATLAB files with an ad hoc format. It was converted to the COMTRADE file format (IEEE/IEC 2013) to allow ingestion by openHistorian. The data provided by BPA contained measurements of one voltage and six currents within a commercial building and was split in two sets. The first set was collected at

2000 samples per second and spanned August 2-28, 2017. The seoncd set spanned November 22 through December 31, 2017 and was collected at 5000 samples per second. The volume of POW data was approximately 1.3 TB.

With multiple terabytes of field-measured PMU and POW data, the testing environment was well-suited to evaluating SciSync and comparing its performance with Archive Walker. The following subsection describes the formal tests that were performed to make this comparison.

5.2 Test Results

The performance of the SciSync tool was compared against Archive Walker in two tests. The first test focused on performance during analysis and the second focused on performance while reviewing results. As expected, SciSync outperformed Archive Walker in both cases.

The first test was based on the analysis of 51 minutes of POW voltage measurements stored in openHistorian. The tools retrieved the sets of voltage measurements, calculated three voltage phasors and a corresponding positive sequence frequency, and then applied out-of-range detectors to the calculated signals. The POW voltages were collected at 5000 samples per second, so the analysis required a large amount of data to be loaded and processed. Archive Walker required 13 minutes to process the data, while SciSync required only 3.5 minutes. The significant difference in performance was due to SciSync's seamless interface with openHistorian and its more efficient coding.

The second test compared how quickly the tools could display the results of an analysis. During the test, an out-of-range detector was applied to three frequency measurements retrieved from the PI System. After Archive Walker's analysis completed, the *Retrieve Detail* button in the GUI (see Figure 5) was used to rerun a 10-minute portion of the analysis and display results. Archive Walker required 21 seconds to rerun the analysis and display the results. Recall that in SciSync all results are stored when an analysis is executed; there's no need to rerun a portion of the analysis as in Archive Walker. To perform the test, the same 10 minutes of data was analyzed with an out-of-range detector and displayed in SciSync. This process required 16 seconds. SciSync is measurably faster than Archive Walker's 21 second performance, but a direct comparison is difficult because the tools handle result storage and display so differently. The real advantage of SciSync in this regard is its ability to automatically adjust the resolution of displayed results as the user zooms in and out.

As desired, SciSync's development by an established software vendor provided it with better performance than the research-grade Archive Walker software. These performance enhancements make it more suitable for use by electric utilities. The next chapter provides two example use cases to demonstrate how utilities might use the tool.

6.0 Example Use Cases

In this chapter, two use cases are presented to demonstrate SciSync's ease of use and flexibility. The first use case shows that the *Template* for a simple but useful detector can be set up in just a few minutes. The second use case explains how SciSync's broad set of analytics can be used to configure something as complex as an oscillation detector.

6.1 Fault Detector

A simple fault detector can be set up by applying SciSync's excursion detector, which is analogous to Archive Walker's out-of-range detector, to voltage magnitude measurements from a PMU. The first step in this process is understanding the normal range for the measurements of interest. The ability to plot signals as in Figure 7 is particularly useful for this aspect. Note that the voltage magnitude in this plot has a value near 315,600 V, which corresponds to a nominal line-to-line voltage of 500 kV. Though a more detailed review of signals and time periods would be helpful when configuring a larger scale analysis, for this demonstration we will assume that voltage magnitudes tend to remain above 305,000 V when a fault is not occurring. This value is useful when setting up the fault detector's *Template*, as depicted in Figure 11.

The top panel in Figure 11 shows that the analysis depends only on voltage magnitude measurements. The measurements, which at this stage are referred to by the generic name *Signal 1* are passed directly into SciSync's excursion detector in the middle panel. For fault detection, the lower threshold is of interest. Based on the review of data in Figure 7, a threshold value of 305,000 V was selected. If a voltage magnitude remains below this level for at least 0.01 seconds (*minDur*), an event will be detected. The bottom panel of the template uses the *DeviceName* keyword to name the *Template's* outputs so that the substation names appear when results are displayed. With the *Template* configured, a *Task* can be set up as in Figure 12 to analyze voltage magnitude measurements from multiple PMUs over a specified period, in this case 10 minutes.

After running the *Task*, the *Results* tab is used to review detected faults. Three widgets are used to summarize the results in Figure 13. The vertical bar in the *Event Occurrences* plot at the top of the figure shows that an event was detected just after 04:57. Moving down to the *Range Values Line Chart* in the middle, abnormally low voltage magnitudes are clearly visible at this time. This plot provides range information for each signal, which can be more easily seen after zooming in on a single channel, as in Figure 14. The dashed lines indicate the minimum and maximum, while the solid line indicates the average value. The amount of data used to calculate these statistics adjusts automatically so that the resolution increases as the user zooms. Eventually, the data is displayed in full fidelity, as in Figure 15. Returning to Figure 13, the *List of Detected Events* table at the bottom shows that the detection threshold was exceeded by signals from the Richland and Rapid City PMUs (these names were arbitrarily assigned to protect the data's confidentiality).

SciSync's figures and tables provide an excellent level of detail and are easy to understand. The tool's ability to automatically adjust the resolution of the displayed data allows the user to quickly identify periods of interest for further review. This use case showed how powerful even a simple *Template* can be. A more complex use case is considered in the next section.

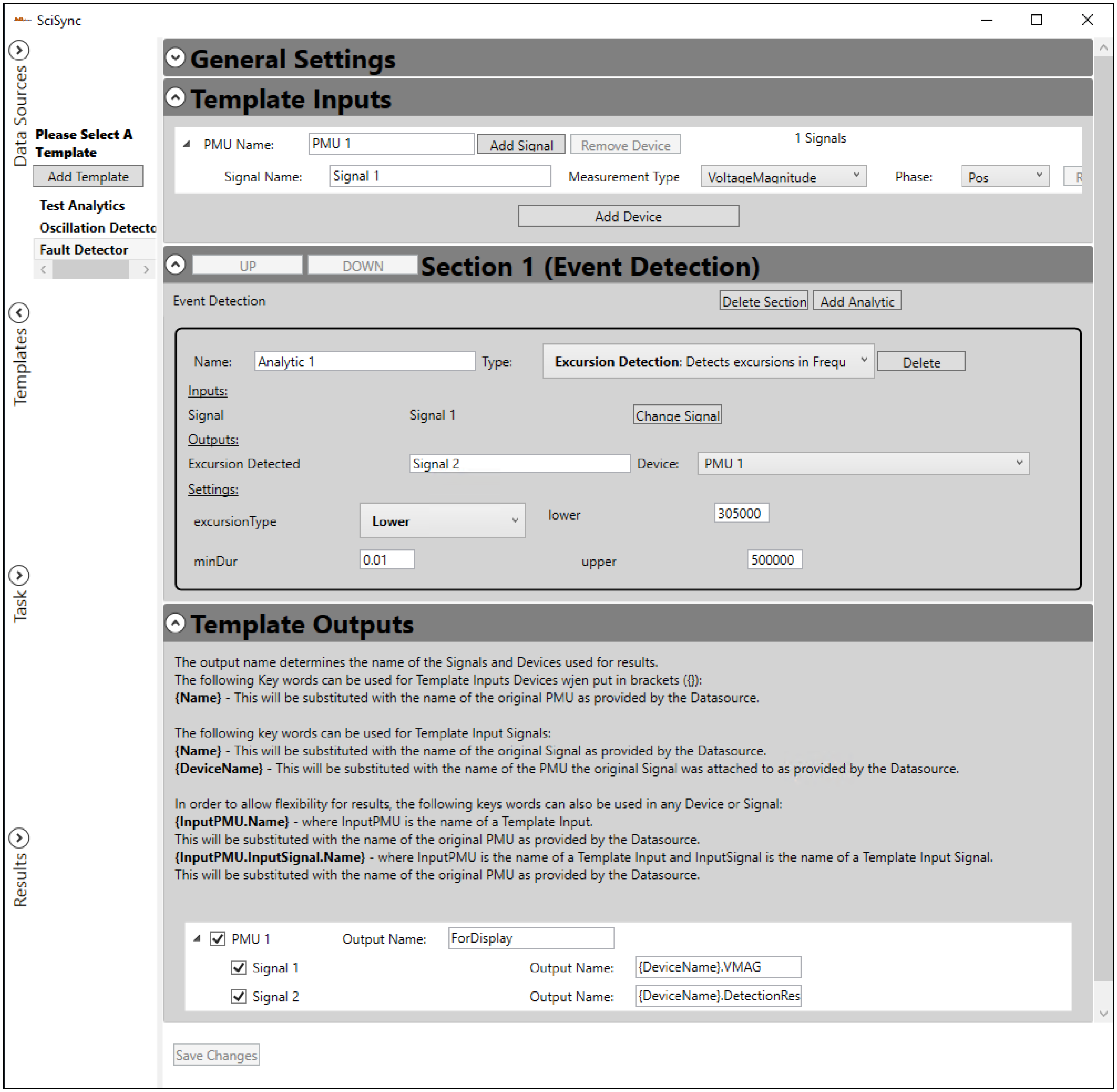

Figure 11. Screenshot of SciSync's Templates tab showing the configuration of a fault detector.

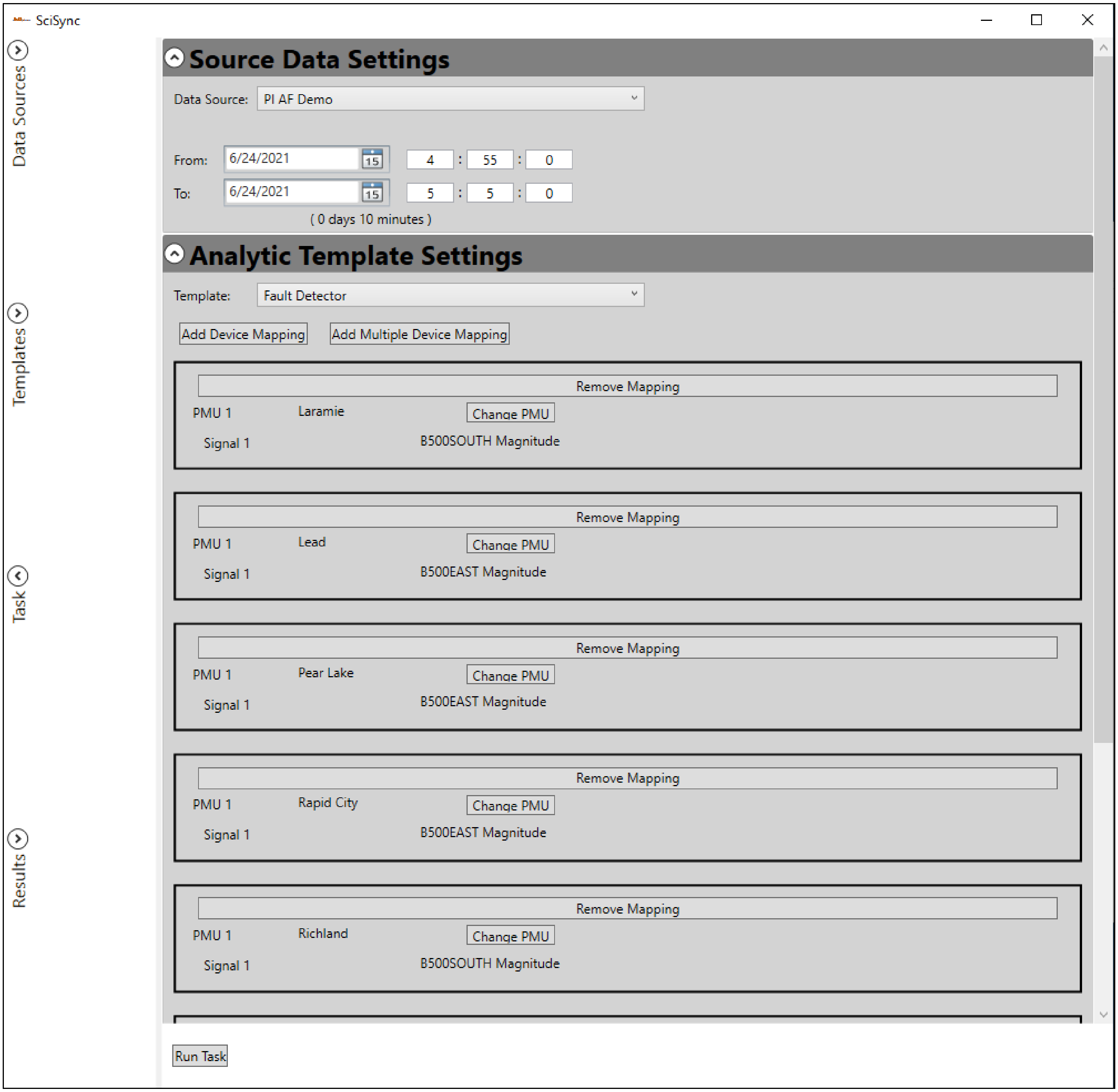

Figure 12. Screenshot of SciSync's Task tab showing the fault detector Template being applied to multiple voltage magnitude channels.

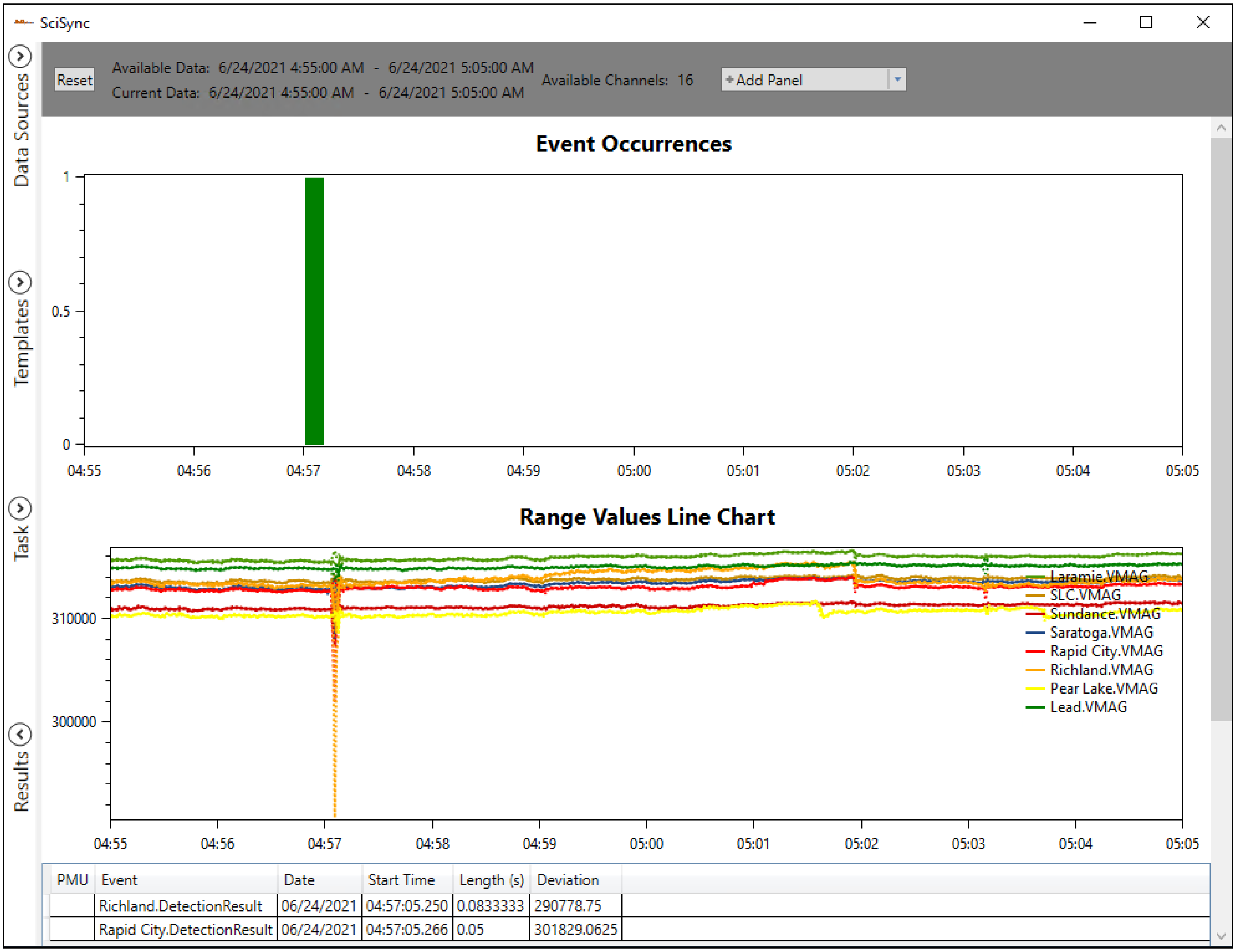

Figure 13. Screenshot of SciSync's Results tab showing a summary of results after applying the fault detector Template to 10 minutes of data.

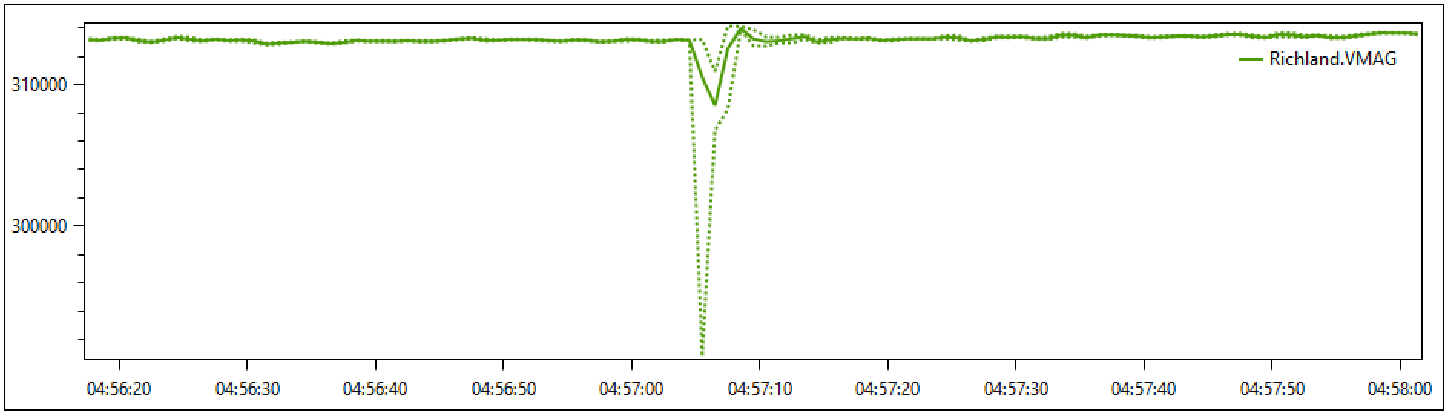

Figure 14. Screenshot of SciSync's Range Values Line Chart with traces indicating a signal's minimum, average, and maximum.

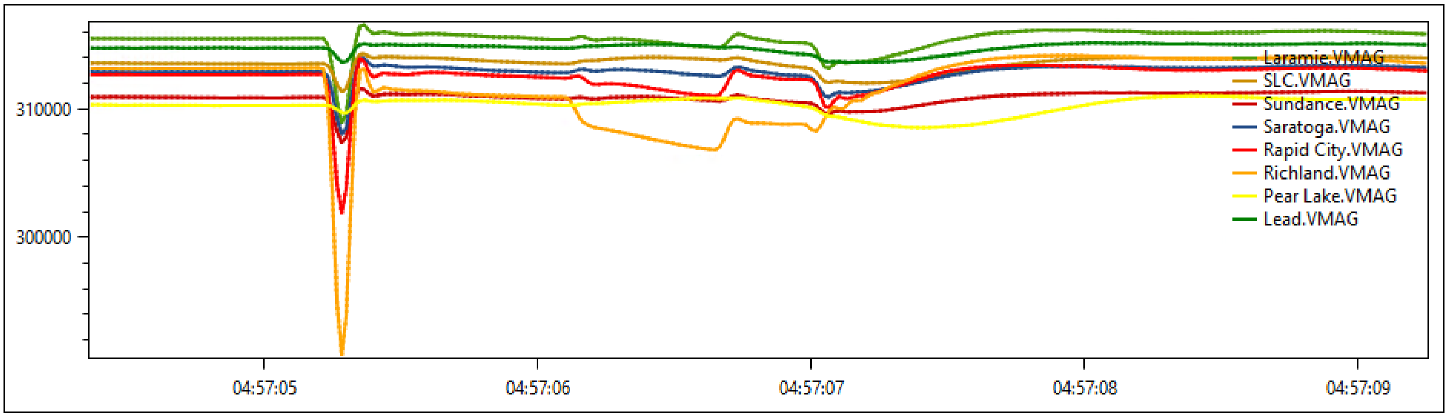

Figure 15. Screenshot of SciSync's Range Values Line Chart zoomed in to show full-fidelity data for multiple signals.

6.2 Oscillation Detector

SciSync has built in oscillation detectors that can be configured quickly, but this use case highlights how SciSync's flexibility can support a custom oscillation detector. The design is based on the root mean squared (RMS)-energy detector proposed in (Donnelly et al. 2015) and summarized in Figure 16. The output of this diagram provides a measure of the input signal's energy over a frequency band of interest. A sudden increase in this energy indicates that an oscillation is present within the specified frequency range. Thus, the detector operates by comparing the calculated RMS-energy with a threshold.

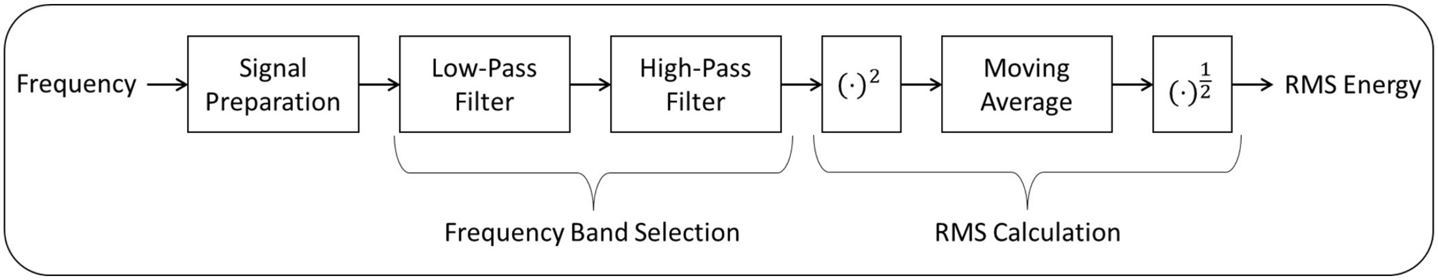

Figure 16. Diagram of the calculation of RMS-energy from frequency measurements provided by a PMU.

The steps necessary to implement the detector in Figure 16 as a SciSync *Template* are:

1. Signal Preparation
   a. The *Linear Interpolation* capability is used to fill in any missing data.
   b. Using the *Shifting* capability, the signal is centered around zero by subtracting the nominal 60 Hz frequency. This prevents large startup transients in the filters that follow.

2. The frequency band of interest is selected using two custom filters. The first is designed to retain signal content below 2 Hz, and the second retains content above 0.1 Hz. The filters are designed outside of SciSync and implemented by entering their coefficients.

3. RMS Calculation
   a. SciSync's *Exponential* capability is used to raise the signal to the power of two.
   b. The *Running Average* function is then used as a simple way to implement a moving average filter. Alternatively, a custom filter could be utilized.
   c. The *Exponential* capability is once again used to calculate the square root of the signal by raising it to a power of 0.5.

4. The resulting RMS-energy is compared against a threshold using SciSync's *Excursion Detector*, just as in the fault detector case.

To demonstrate this use case, the *Template* was applied to frequency measurements from two PMUs over a 15-minute period. The results are displayed in Figure 17. The top plot displays the RMS-energy for each of the PMUs. Note that the energy for the Sundance PMU increases significantly beginning at about 02:38:30. The bottom plot is a range chart for the Sundance frequency measurement. Note that the signal's range increases along with the RMS-energy. After zooming in further, SciSync automatically displays the data in full fidelity, as in Figure 18. Here, the presence of the oscillation is clearly visible in the frequency measurements from Sundance. The oscillation is not apparent in the SLC frequency measurements, explaining why the RMS-energy for SLC remained flat throughout the analysis.

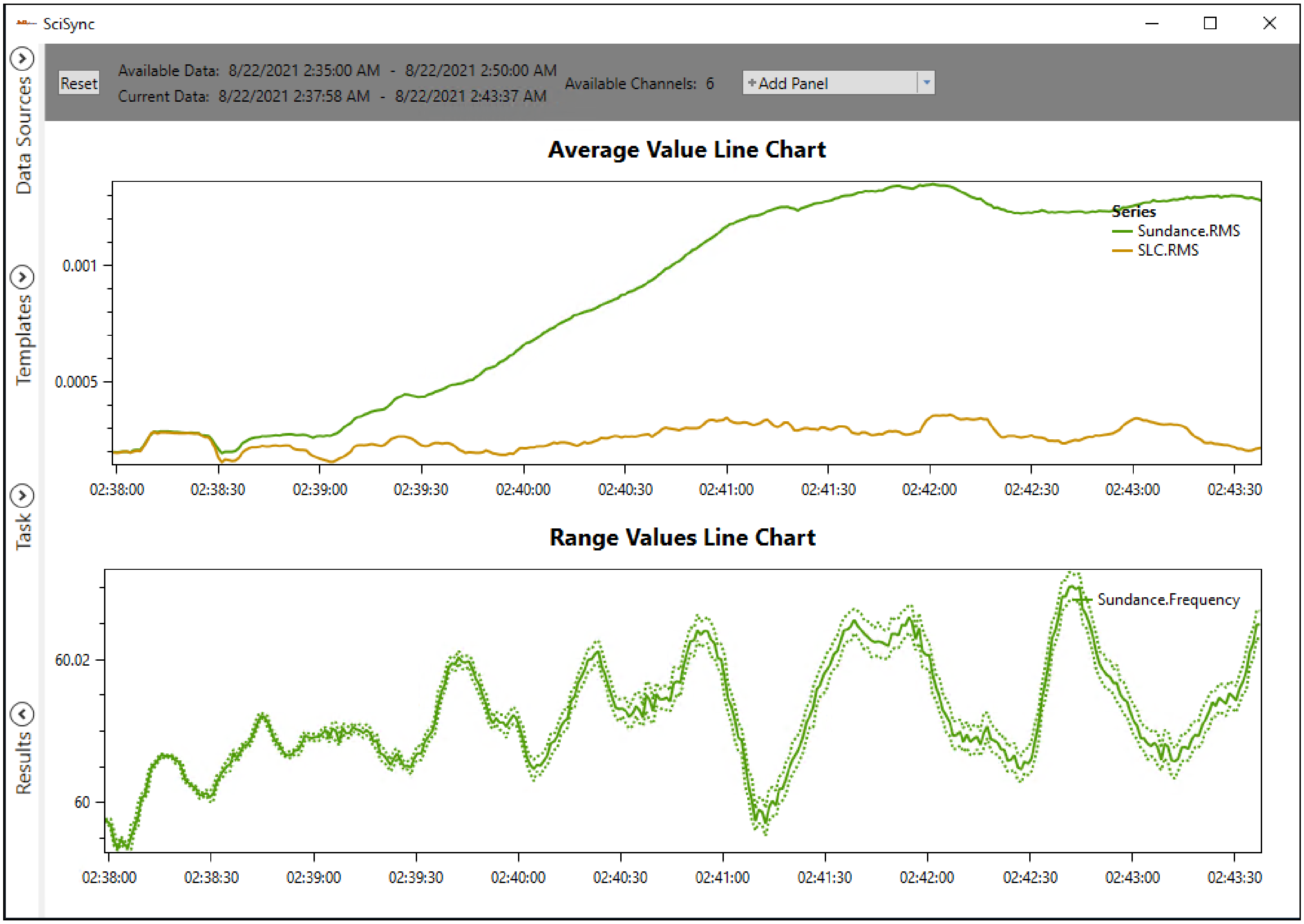

Figure 17. Screenshot of SciSync's Results tab showing the RMS-energy for two signals (top) and the range of the frequency measurement from the Sundance PMU (bottom).

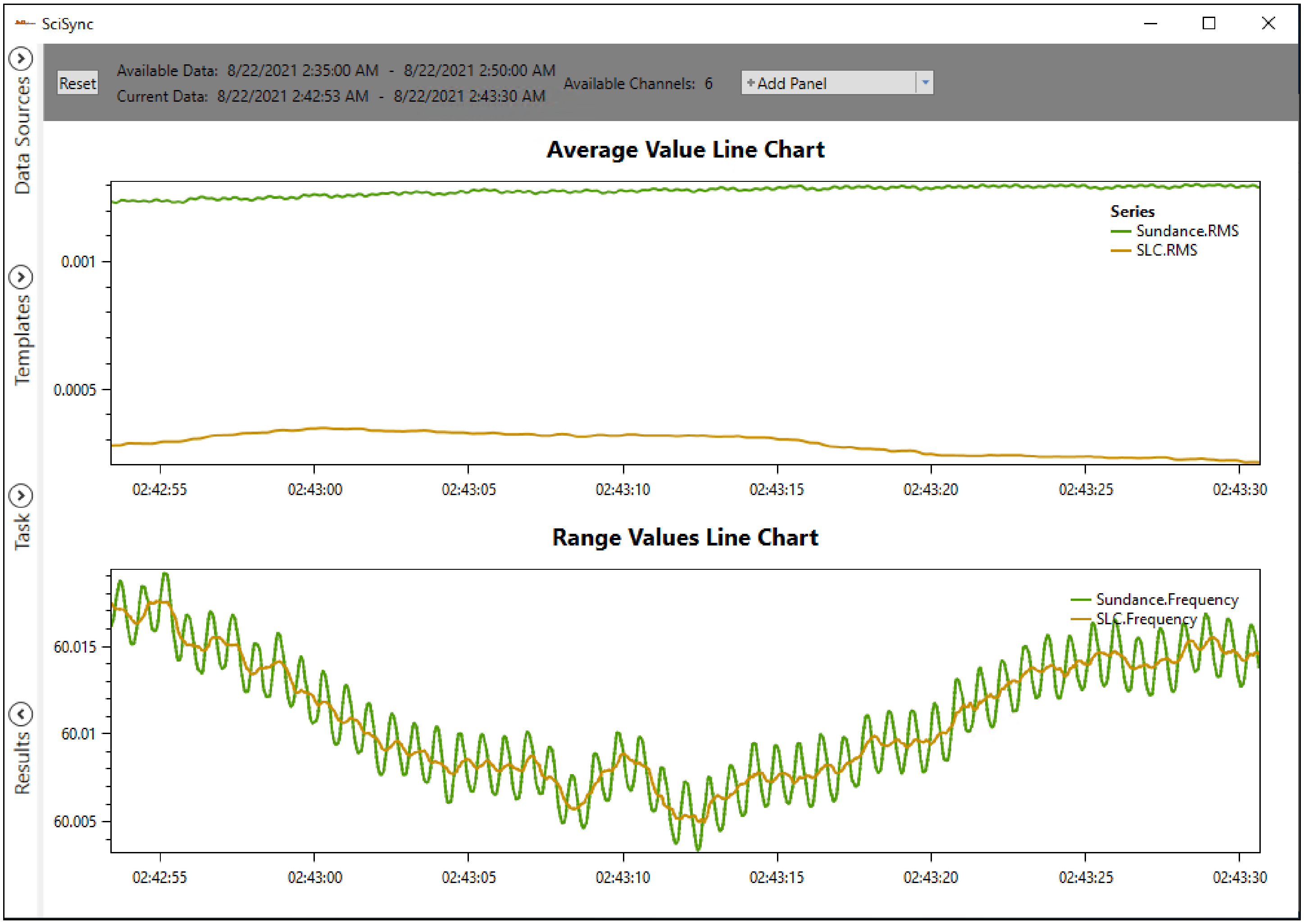

Figure 18. Screenshot of SciSync's Results tab zoomed in from Figure 17 to show the oscillation in full fidelity.

7.0 Conclusion

As demonstrated in the previous two chapters, SciSync is well equipped to help utilities rapidly read, process, analyze, and review power system measurements. The tool replicates much of Archive Walker's functionality, but in a commercial package that utilities can readily deploy. Some of the key benefits of SciSync over Archive Walker include:

- SciSync is coded in a programming language well-suited to application development, whereas Archive Walker utilizes an analysis engine developed in the MATLAB scripting language
- Interfaces with the PI System and openHistorian data sources are much more seamless in SciSync
- *Templates* are configured for generic input signals, making it simple to change the measurement channels an analysis is applied to when setting up a *Task*
- SciSync achieves faster performance, both in processing data and displaying results
- The resolution of displayed data is automatically adjusted based on the selected time range to provide a smooth transition from summaries to full-fidelity measurements.

While Archive Walker will remain a useful tool for early-stage research at PNNL, there is significant opportunity for SciSync to support research across the industry moving forward. SciSync was intentionally designed to support the integration of additional analytics. As an open-source tool, researchers have access to the code to make these updates themselves, making SciSync an ideal tool for demonstrating the value of new algorithms. SciSync also provides a new path for reseachers to transfer their algorithms to industry. With the base SciSync tool in place, PNNL can now quickly and efficiently partner with GPA to deploy proven algorithms in an industry-grade tool used by utilities. In these ways, SciSync's benefits to the electric power industry are expected to grow over time.

References

Ahmad, Tawsif, Ning Zhou, Jim Follum, Renke Huang, Shaobu Wang, Urmila Agrawal, Pavel Etingov, and Zhenyu Huang. 2021. "Estimation of the Correlation Between Oscillation Modes and Operating Conditions using Quantile Regression: A Measurement-Based Approach." In *2020 52nd North American Power Symposium (NAPS),* 1–6. https://doi.org/10.1109/NAPS50074.2021.9449647.

"AVEVA PI System." 2022. AVEVA. https://www.aveva.com/en/products/aveva-pi-system/.

Donnelly, Matt, Dan Trudnowski, James Colwell, John Pierre, and Luke Dosiek. 2015. "RMS-energy filter design for real-time oscillation detection." In *2015 IEEE Power & Energy Society General Meeting,* 1–5. https://doi.org/10.1109/PESGM.2015.7286192.

Etingov, Pavel V., James D. Follum, Urmila Agrawal, Heng Wang, Francis K. Tuffner, Lisa L. N. Newburn, Renke Huang, Tamara Becejac, and Malini Ghosal. 2020. "Open Source Suite for Advanced Synchrophasor Analysis" (September). https://doi.org/10.2172/1673609. https://www.osti.gov/biblio/1673609.

Follum, J., P. Etingov, F. Tuffner, H. Wang, U. Agrawal, D. Kosterev, S. Yang, and A. Faris. 2020. "Detecting and Analyzing Power System Disturbances in PMU Data with the Open-Source Archive Walker Tool." In *IEEE/PES Transmission and Distribution Conference and Exhibition.*

Follum, James D., Urmila Agrawal, and Pavel V. Etingov. 2021. "Evaluation of Mode Meters Robust to Forced Oscillations using Field-Measured Data" (January). https://www.osti.gov/biblio/1817687.

Follum, James D., Harold Kirkham, Artis Riepnieks, Pavel V. Etingov, Laurie E. Miller, Xiaoyuan Fan, and Emily J. Ellwein. 2021. "Roadmap for Advanced Power System Measurements" (April). https://doi.org/10.2172/1871292. https://www.osti.gov/biblio/1871292.

Follum, James D., Heng Wang, Pavel V. Etingov, Francis K. Tuffner, Urmila Agrawal, Zhuanyi Huang, Dmitry Kosterev, et al. 2018. "Archive Walker Software: Setting Up and Reviewing Analyses with the Archive Walker GUI" (September). https://doi.org/10.2172/1592685. https://www.osti.gov/biblio/1592685.

Follum, Jim, Tamara Becejac, and Pavel Etingov. 2021. "A Robust Yule-Walker Method for Online Monitoring of Power System Electromechanical Modes of Oscillation." In *2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT),* 01–05. https://doi.org/10.1109/ISGT49243.2021.9372152.

Follum, Jim, Tamara Becejac, and Renke Huang. 2021. "Estimation of Electromechanical Modes of Oscillation in the Eastern Interconnection from Ambient PMU Data." In *2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT),* 1–5. https://doi.org/10.1109/ISGT49243.2021.9372216.

IEEE/IEC. 2013. "IEEE/IEC Measuring relays and protection equipment – Part 24: Common format for transient data exchange (COMTRADE) for power systems." *IEEE Std C37.111-2013 (IEC 60255-24 Edition 2.0 2013-04),* 1–73. https://doi.org/10.1109/IEEESTD.2013.6512503.

“openHistorian.” 2022. Grid Protection Alliance. https://www.gridprotectionalliance.org/phasor-Historian.html.

Phadke, A. G., and T. Bi. 2018. “Phasor measurement units, WAMS, and their applications in protection and control of power systems.” *Journal of Modern Power Systems and Clean Energy* 6 (4): 619–629. https://doi.org/10.1007/s40565-018-0423-3.

PNNL. 2018. *Setting Up and Reviewing Analyses with the Archive Walker GUI.* [Online]. Available: https://github.com/pnnl/archive_walker/wiki.

“Version 1.0.0 of SciSync.” 2023. Grid Protection Alliance. https://github.com/GridProtectionAlliance/Adapt/releases/tag/v1.0.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99352
1-888-375-PNNL (7675)

www.pnnl.gov