Oscillation Analysis with SCADA using Inferential Statistics (OASIS)

Final Project Report

S-103G

Power Systems Engineering Research Center
Empowering Minds to Engineer
the Future Electric Energy System

Oscillation Analysis with SCADA using Inferential Statistics (OASIS)

Final Project Report

Project Team
Vaithianathan (Mani) Venkatasubramanian, Project Leader
Washington State University

Krishna Jandhyala
Washington State University

Graduate Student
Salman Siddique Shiuab
Washington State University

PSERC Publication 24-07

December 2024

For information about this project, contact:

Vaithianathan (Mani) Venkatasubramanian
Professor, School of Electrical Engineering and Computer Science
Washington State University
mani@wsu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-727-2052

Notice Concerning Copyright Material

Acknowledgments

We wish to thank the industry advisors, Patrick Panciatici (RTE), Gilles Torresan (RTE), and Marie-Sophie Debry (RTE) for their support and advice throughout the project.

Executive Summary

Interarea and forced oscillations are emerging as operational concerns in modern power systems because of the changing intermittent generation patterns, unusual transmission power flows, and the integration of inverter-based resources with traditional power grid equipment. The project's objective is to develop rigorous statistical methods that can use widely available SCADA measurements to detect and analyze power system oscillations. Even though millions of SCADA measurements are available in a typical power utility company, they are used chiefly for quasi-steady-state analysis in the present-day power system owing to their slow sampling rate. In this project, we propose using SCADA to detect the source of problematic dynamic oscillations in the power grid. This is accomplished by exploiting the asynchronous sampling inherent in SCADA technology. While the oscillation period cannot be determined from SCADA data as implied by the sampling theorem, the inferential statistical formulation enables estimating the oscillation amplitude from the generation outputs and transmission line flows at a specified confidence level. The proposed methodology has been tested using the archived SCADA data from RTE France and has successfully identified the source of oscillations in those events.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Background

Recently, many algorithms have been proposed that utilize synchrophasors to identify potential forced oscillation sources. However, due to the lack of sufficient Phasor Measurement Units (PMUs) installments throughout the system, sometimes it might be challenging to pinpoint the exact generators or loads from where the oscillation originated initially. In this regard, SCADA data is helpful due to its abundance in the system. Although SCADA data has a much lower sampling rate than PMU data, the asynchronous sampling nature of the SCADA data can be capitalized on in identifying the problematic source of forced oscillations. Forced oscillations in power systems are typically caused by external disturbances, such as cyclic outputs from generator prime movers or persistent cyclic fluctuations in loads ([1]-[2]). Forced oscillations can interact with natural electromechanical modes ([3]-[4]), and the resonant oscillations can be observed in broad regions of the interconnections (e.g., January 11, 2019, eastern system event [5]). Considerable research works have been done in the power system domain to investigate the identification of the oscillation source through the analysis of measurements obtained from the PMUs installed in the system [6]-[10]. Reference [11] also comprehensively reviews the existing source location methods.

The low sampling nature of SCADA can preserve some of the features of the data when it is sampled with PMU. The plot of MW output from a generating unit, measured by PMU and SCADA, is given below in Figure 1 for comparison purposes. The SCADA sampling feature preserves some characteristics, such as the amplitude of the PMU data, which can be utilized to develop the statistical method to analyze the oscillation event. This sampling nature remains the same even if the data are detrended, as seen in Figure 2. Different data-detrending methods can be applied to remove a trend from the time series data for analysis.

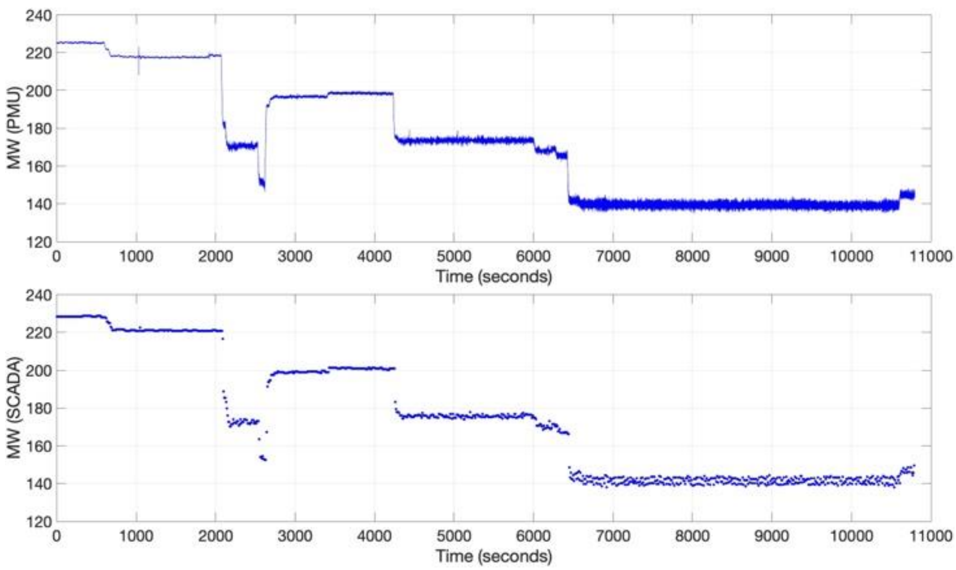


Figure 1: Comparison of raw PMU and SCADA data sampling

The duration of the measurements plotted in Figure 1 is 3 hours. The PMU data sampling rate was 30 Hz, whereas SCADA data had a sampling rate of 0.1 Hz (1 sample every 10 seconds). Although PMU data are sampled synchronously, SCADA data are sampled asynchronously and randomly, and this inherent quality of the SCADA samples preserves all the information needed to develop a rigorous statistical technique for analyzing and detecting forced oscillation sources.

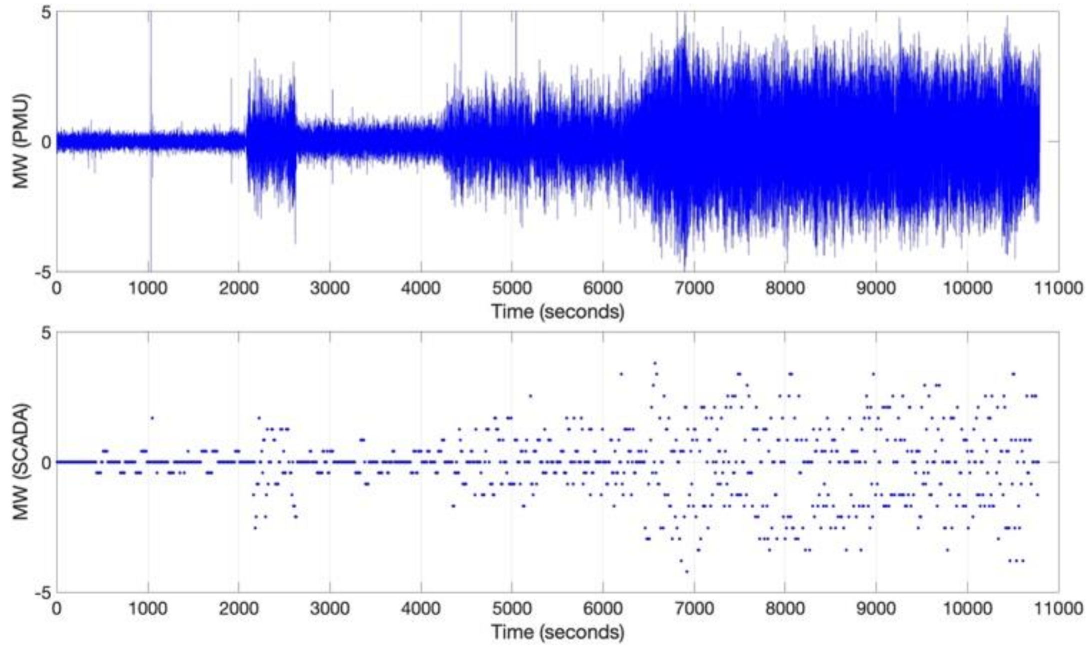


Figure 2: Comparison of PMU and SCADA data sampling after detrending

1.2 Overview of the Problem

Partial synchrophasor coverage of power generation limits the scope of oscillations that can be observed and analyzed for the source location of forced oscillations using these synchrophasor measurements. On the other hand, when combined with intelligent analysis methods, the wide availability of SCADA measurements can solve the oscillation monitoring and source location problems by easy detection and analysis of these oscillations. By deriving a novel application based on the SCADA infrastructure, the methodology provides additional value for power companies on their existing investment. Currently, oscillations are monitored using synchrophasors, which have a high sampling rate, typically 30 Hz in the North American power systems and 50 Hz in the European interconnections. While the technology for oscillation monitoring using synchrophasors is maturing, the installation cost and communication requirements have implied that the observability of the power grid using synchrophasors remains limited. On the other hand, SCADA technology has been in implementation since the 1970s, and these measurements are available at almost every synchronous generator and at the bulk interconnection interface for most of the renewable generators in the power grid due to the ample number of installations.

1.3 Technical Overview

In this project, we organized the research problem into the following sections:

(a) Statistical method for oscillation detection and analysis using SCADA data:

The key feature of the SCADA data that motivates our project is that the SCADA data is sampled asynchronously and in a somewhat random fashion in the sense of sampling time. That is, even when the SCADA sampling rate is one per second, the sampling is not done every second apart concerning any accurate timing signal. Instead, the sampling is done according to some polling schedule, and the actual sampling may occur at any time within a quarter-second window. This is significant for a signal exhibiting an oscillation because this random sampling schedule implies that the sampling may "hit" any part of the oscillating signal. This forms the basis for applying inferential statistical methods to derive probabilities for when the sampled value will be above a prespecified magnitude in the sense of an oscillation. In other words, statistical tests can be developed to ascertain whether a SCADA signal exhibits an oscillation of a certain amplitude for a given significance level. In summary, SCADA measurements can be shown to be random samples of the analog signals being monitored, and this forms the basis for applying inferential statistical methods. We can use statistical theory to derive probabilities for when the sampled value will exceed a prespecified magnitude in an oscillation. In other words, we can develop statistical tests to ascertain whether a SCADA signal exhibits signatures of an oscillation above a certain amplitude for a given confidence level.

The typical oscillation frequency of power plant oscillations ranges between 0.2 Hz to 2 Hz, while subsynchronous oscillations associated with inverter-based resources typically range between 5 Hz and 50 Hz [12]. By sampling theorem, the sampling rate must be between 0.4 and 100 Hz for analyzing such oscillations, respectively. On the other hand, the sampling rate of the SCADA data varies between 0.1 Hz to 1 Hz. While this appears to be a contradiction, the random sampling schedules discussed above in the context of the SCADA data enable the sampling to preserve specific features of the oscillation, such as the oscillation amplitude in a probabilistic sense. Another essential condition for observing high amplitude samples in the SCADA data is that the transducer should not have a lower sampling frequency than the oscillation frequency present in the signal. Otherwise, the oscillation would not be observable in the SCADA samples. The MW and MVAR outputs of the generator units can measure how strongly the unit is participating in an oscillation.

(b) Oscillation analysis of SCADA data using statistical method:

After developing the theoretical background, we implemented the algorithm, using filters and data-detrending methods to preprocess the SCADA data. We analyzed the generation outputs and transmission line flows for different test cases to detect the possible source of oscillations.

(c) Validation of the methodology on offline test cases:

We used various test cases from the RTE system's past events to test the methodology and validated the results by discussing the findings with RTE.

(d) Open-source oscillation analysis toolbox OASIS:

We developed an oscillation analysis toolbox named OASIS using the Python Streamlit library with the proposed algorithm.

1.4 Discussion

1.4.1 Potential Benefits

Oscillations have emerged as one of the main stumbling blocks in reliable power grid operation, where the oscillations can appear from natural system modes or forced oscillations from control failures, rough zone operation, or incorrect setting of control parameters. Such oscillations can cause potential damage to expensive generator equipment, including rotor shafts. They can also be a problem for power quality for the consumers and a safety hazard for power plant personnel. With the increased adoption of renewable generation, the problem is primarily in integrating power electronic-based intermittent energy resources into the power grid. There is an urgent need to develop methods to detect and analyze the cause of these oscillations to quickly correct the underlying problems.

Detection and efficient analysis of oscillations can help manage and accelerate the integration of large numbers of renewable devices into the power grid. Also, timely correction of the oscillation issues in fossil fuel and nuclear power plants can prevent unscheduled outages of these energy sources and help avoid the need to dispatch power from less efficient power plants. Moreover, it can help improve the overall economy of the power grid by preventing potential damage from problematic oscillations and leading to a more operationally reliable power grid in the future.

1.4.2 Related Work

In [12], two slightly different methods named Pattern Mining Algorithm (PMA) and Maximal Variance Ratio Algorithm (MVRA) were proposed, which use SCADA data for forced oscillation source location wherein the oscillation detection was done using high-speed synchrophasor data. PMA and MVRA methods require more parameters to tune for different test cases than the proposed algorithm for accurate detection of the source. If those parameters are not set or tuned properly for various oscillation events, it might be challenging to identify the sources. On the other hand, those two methods are heuristic, whereas this proposed methodology applies inferential statistical theory for the rigorous development of the algorithm.

1.5 Report Organization

The rest of the report is organized as follows: Section 2 introduces the inferential statistical theory as the foundation for developing the algorithm. Section 3 summarizes the steps of the algorithm for locating the sources of forced oscillations using SCADA data. Section 4 illustrates and discusses the real forced oscillation events in the RTE systems. Section 5 provides an overview of the OASIS toolbox, and section 6 concludes the report.

2. Statistical Method for Oscillation Detection and Analysis Using SCADA Data

2.1 Statistical Theory

During a forced oscillation event, the amplitude of the source signals increases or decreases compared to the other signals in the system in normal or ambient conditions. To introduce the theory behind the algorithm, it is assumed that the oscillating signal behaves like a sine wave. Although this assumption might differ in real cases, it can be handled by any proper data-detrending method. For analyzing the SCADA samples, a pair of bands can be set up, which are used as an indication to check how high the amplitude of the signal is. In the case of a forced oscillation event, the oscillating signal will cross the bands, and the sampled data of the signal will show transitions across the bands. This transition behavior among the data samples can be used to formulate an algorithm to identify the system's high-amplitude signals. An illustration of data samples transitioning across the bands for both ambient and oscillating conditions is shown in the following Figure 3 with an ideal sine wave oscillation case. For example, the sine wave's amplitude is 1 per unit (pu) in ambient conditions, and the band is set to 2. As a result, the samples do not transition across the bands. On the other hand, when oscillation occurs, and the signal's amplitude can go from 1 to 4, there would be transitions among the samples across the bands.

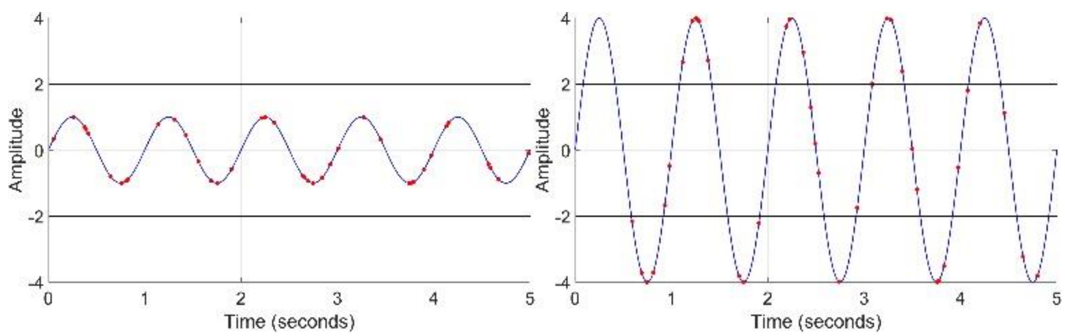


Figure 3: Illustration of data transition in ambient (left) and oscillating condition (right) across the transition bands

2.2 Analysis with One Sample

The algorithm is formulated by counting the number of transitions in the SCADA samples across the upper and lower transition bands. Let $x(k)$ be a random SCADA sample, which can lie anywhere in the upper, middle, or lower region, as shown in Figure 5. Let the transition band be represented by B. The sample space Ω consists of three outcomes.

$$\Omega = \{R_1, R_2, R_3\}$$

where, $R_1 : x(k) \in (+\infty, B)$, $R_2 : x(k) \in (B, -B)$ and $R_3 : x(k) \in (-B, -\infty)$.

The probability r can be defined as the probability of finding the sampled data $x(k)$ in the upper region which corresponds to the outcome R_1. If the highest amplitude of the oscillation is represented by A, and the transition band is set up at $+B$ in the upper region, then for outcome R_1, the probability r can be written as,

$$r = \frac{1}{2} * \left(\frac{A-B}{A-0}\right) + \frac{1}{2} * 0 = \frac{A-B}{2*A} \tag{1}$$

Similarly, the probability of finding a sampled data $x(k)$ in the lower region which corresponds to the outcome R_3 can also be defined as r. In that case, if the transition band is set up at $-B$ in the lower region, and the lowest amplitude of the oscillation is $-A$, then for outcome R_3, r can be written as,

$$r = \frac{1}{2} * 0 + \frac{1}{2} * \left(\frac{-B-(-A)}{0-(-A)}\right) = \frac{A-B}{2*A}$$

which is same as (1).

The transition happens between two consecutive samples. For example, let the two successive samples be $x_1(k)$ and $x_2(k)$. So, the transition between these two samples will occur if they correspond to the following pair of outcomes: $\{R_1, R_2\}$ or, $\{R_2, R_3\}$ or, $\{R_1, R_3\}$. So, the probability of transition, s can be written in terms of r as,

$$s = 1 - r^2 - r^2 - (1-2r)^2 = -6\left(r - \frac{1}{3}\right)^2 + \frac{2}{3} \tag{2}$$

Equation (2) represents the parabolic relation between s and r. The following claims can be made about probability r, probability of transition, s, and the transition band B.

Lemma 1: $r < 0.5$
Referring to (1),
If, $A > 0$, then, $B > 0$. As a result, $r < 0.5$.
If, $A < 0$, then, $B < 0$. As a result, $r < 0.5$.
If, $B = 0$ *and* $A \in \mathbb{R}, r = 0.5$.
Since $B \neq 0$, r must be constrained as $0 < r < 0.5$.

The relationship between r and s can be visualized from the following Figure 4.

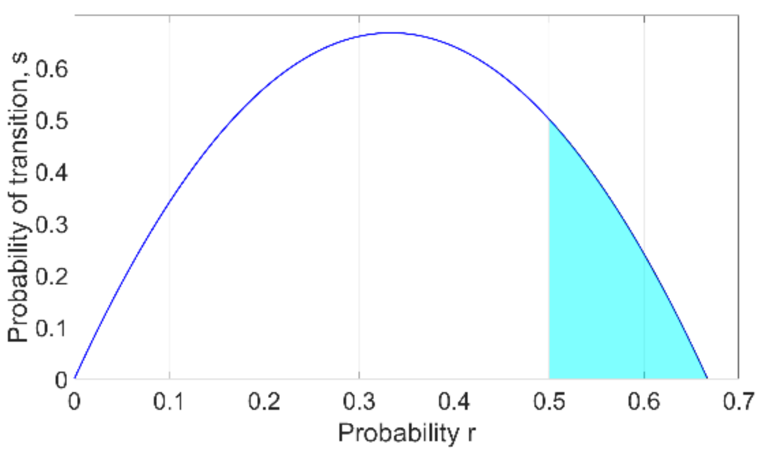


Figure 4: Relation between r and s

It can be observed that the value of s increases with the increase of r at first. When r is 0.33, s reaches its maximum value which is 0.66. After that, with the increase of r, s starts decreasing. The unshaded region of Figure 4 is the considerable region for the proposed algorithm.
The probability of transition, s can be scaled with the following formulation.

$$p = \frac{3}{2}s \qquad (3)$$

Here, p is the scaled probability of transition, and the range of this scaled probability is $0 < p < 1$ according to (3).

Lemma 2: r is an increasing function respective to the amplitude of oscillation, A.

From (1),

$$\frac{dr}{dA} = \frac{B}{2 * A^2} \qquad (4)$$

When the amplitude of oscillation increases and the band is kept fixed, the probability r will also increase. For example, let the upper transition band be at 50 MW or MVAR. Then, the lower transition band would be set at –50 MW or MVAR. If the amplitude of oscillation is 80 MW or MVAR, then r would be 0.1875. If the amplitude of oscillation is increased from 80 to 100 MW or MVAR, then r would be 0.25. It is to be noted that increasing the value of r does not mean the detection is better. Since the amplitude of oscillation is not known beforehand, the setting of the band needs to be handled intuitively for the accurate detection of the source, which is discussed in section 3.

Lemma 3: In an ideal case, $A = 3B$.

According to Figure 4, when $r = 0.33$, $s = 0.66$. Putting the value $r = 0.33$ in (1) results in $A = 3B$. So, for an ideal test case scenario, the amplitude of oscillation should be three times of the transition band.

Lemma 4: If $B > A$, $r = 0$ and the probability of source detection is 0.

If the band is set up higher than the largest amplitude of the oscillation signal, there would be no transitions between samples. As a result, in the event of forced oscillation, it would not be possible to identify the source with the algorithm.

2.3 Testing for Detection of Source with Multiple Samples

Let K be the number of transitions across the $+/-B$ band and n be the length of the sample window for analysis.

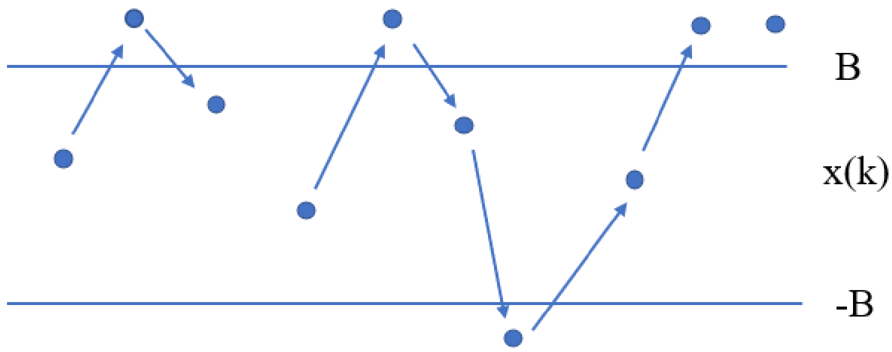


Figure 5: Sampled data transitions across the band

In the Figure 5 above, there are 10 samples in the sample window, and 7 transitions occurred between the two consecutive samples. An assumption can be made on the distribution of the number of transitions, K. Either there would be a transition between two samples, or there would be no transition. So, these two occurrences can be modeled as success and failure outcomes from the binomial distribution. Since the SCADA samples are discrete, we assume that K follows the binomial distribution.

For the detection of oscillation from a test sample, a hypothesis test can be formulated with a null hypothesis and an alternate hypothesis.

Null hypothesis (H_0): There is no oscillation (high amplitude data) present in the power system compared to the average amplitude in ambient conditions.

Alternate hypothesis (H_1): Oscillation is present in the power system, and the amplitude (absolute value) of the source signal during the oscillation period is higher than the average during the non-oscillation period.

For the samples present in the analysis window, the binomial probability mass function for K, the number of transitions among n samples can be expressed as in (5).

$$\Pr(n, k, p) = \binom{n}{k} p^k (1-p)^{n-k} \tag{5}$$

Here, Pr, the binomial probability is a function of n, k and p, where n represents the total number of samples in the oscillation window, k is the number of successes to the occurrence of transitions in n trials, and p is the scaled probability of a transition between two successive samples and $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Clearly, smaller values of transition probability p are indicative of no oscillation in the process, and when p is beyond a pre-specified value, say, p_{min}, then one concludes that oscillations are present in the process. Thus, the null and alternative hypotheses can be equivalently expressed in terms of p as:

$$H_0: p \leq p_{min}$$
$$H_1: p > p_{min}$$

The testing of H_0 against H_1 is carried out by identifying the number of transition thresholds under the binomial probability distribution for K. As is done in any hypothesis testing problem, the threshold number for transitions is obtained by assuming the null hypothesis H_0 to be true. Thus, under H_0, K follows a binomial distribution with parameters $(n,\ p_{min})$. It is well known that under the binomial distribution, the observed sum is uniformly the most powerful statistic for the one-sided hypothesis testing problem. Thus, for any given significance level α, one finds the threshold number k_α from the binomial distribution with parameters $(n,\ p_{min})$. In this manner, a Uniformly Most Powerful (UMP) test can be formulated. The time window during the oscillation event is referred to as the oscillation window, and the non-oscillation duration of the time window is referred to as the ambient or base window in this report. These hypothesis tests for both oscillation and ambient windows would give high confidence to detect the source of forced oscillation accurately and would further prevent any false detection.

Thus, p_{min} and significance level, α can be used to determine the number of transitions threshold, K_α that enables one to reject the null hypothesis H_0. So, if the actual number of transitions among the samples is denoted by K, then the test can be formulated for the oscillation window analysis as,

$$Accept\ H_0\ if\ K \leq k_\alpha$$
$$Reject\ H_0\ if\ K > k_\alpha$$

Under the above accept/reject criterion, one can expect the following. Whenever H_0 is true, the actual number of transitions K can be expected to be below k_α, and except for a small proportion α we would stay with the null hypothesis most of the time. However, if the alternative is true, then the actual number of transitions will most likely be greater than the threshold value k_α, and one will make the right decision of rejecting the null hypothesis.

Let us consider an event where the oscillation duration is 10 minutes. If the SCADA sampling rate is 0.1 Hz, there would be 60 data samples for that oscillation window. Let us also assume that the value of p is 0.30 when the band is set up. So, there would be 59 pairs to check the transition between two consecutive samples from those 60 samples, and each checking would be independent of one another. The plots of Probability Mass Function (PMF) and Cumulative Distribution Function (CDF) for this binomial distribution case are illustrated in Figure 6.

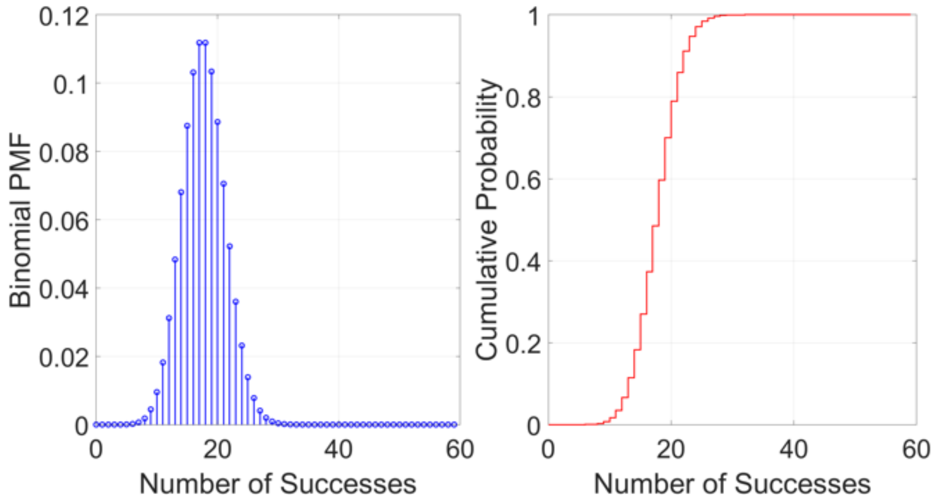


Figure 6: Plot of PMF and CDF of binomial distribution

If the oscillation duration of an event is not extremely short, there would be large enough samples where the normal approximation can also be used instead of the binomial distribution formulation as per the central limit theorem. In our analysis, it is observed that if the expected value ($E = np$) is greater than 1.5 when $p = 0.3$, the normal approximation and binomial PMF provide the same threshold count k_α for the test when the significance level is set to 5%. An illustration for the normal approximation of the binomial distribution is shown in Figure 7.

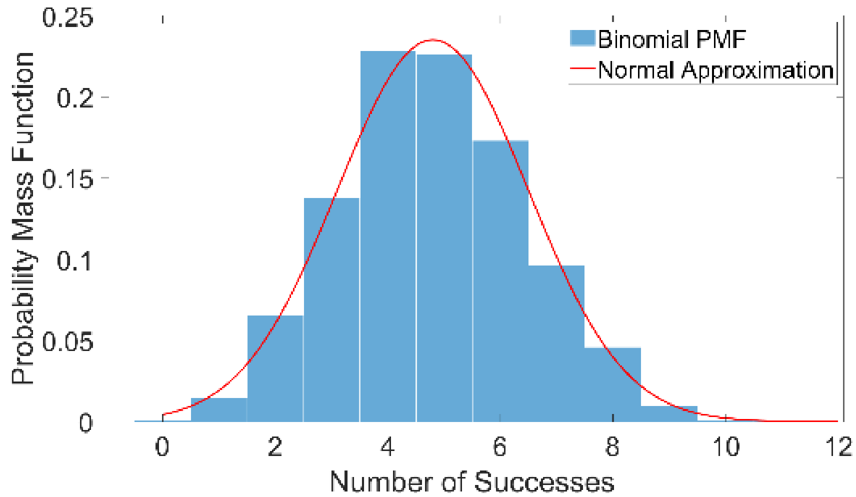


Figure 7: Normal approximation of binomial PMF for n=12 and p=0.4

2.4 Impact on High Amplitude Samples Due to Low Pass Filters

Low pass filters are commonly used in transducers within SCADA systems, particularly in the signal conditioning stages for noise reduction and signal smoothing purposes. These filters may suppress the high-amplitude data samples in the SCADA system. To illustrate this phenomenon, let us consider a signal with both ambient and oscillating conditions for 10 minutes or 600 seconds, as shown in Figure 8. The first 5 minutes of the signal is in ambient condition, and after that, the oscillation starts where the frequency of the oscillating sinusoidal signal is 0.25 Hz, the amplitude of oscillation 6 pu, and the transition bands are set at 2 and -2 pu. If the SCADA system samples 1 data in every 10 seconds, there would be 60 samples for the signal duration, and the sampling would preserve the high amplitude samples in the oscillation window. In this case, due to the transitions of the samples across the bands in the oscillation window, the proposed algorithm can be implemented successfully.

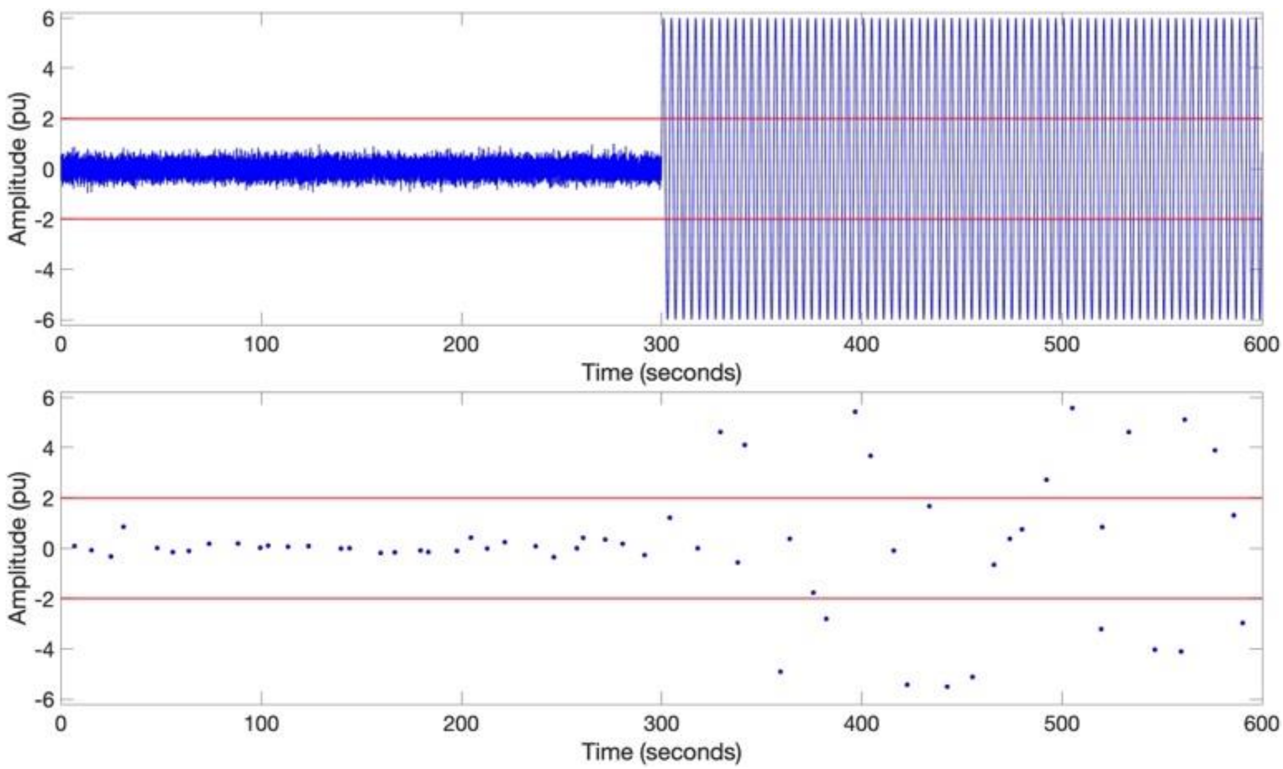


Figure 8: Emulated signal and the resulting samples from SCADA with transition bands

On the other hand, if the signal is passed through a low pass filter before the SCADA sampling, the oscillating signal would be attenuated, and the resulting sampled data would not show any transition. This is depicted in Figure 9, where the emulated signal is passed through a low pass filter first with a cutoff frequency of 0.1 Hz. This is a necessary condition for the proposed algorithm to work as intended; that is, the SCADA samples should not be attenuated due to the implementation of any low pass filter design with a lower cutoff frequency.

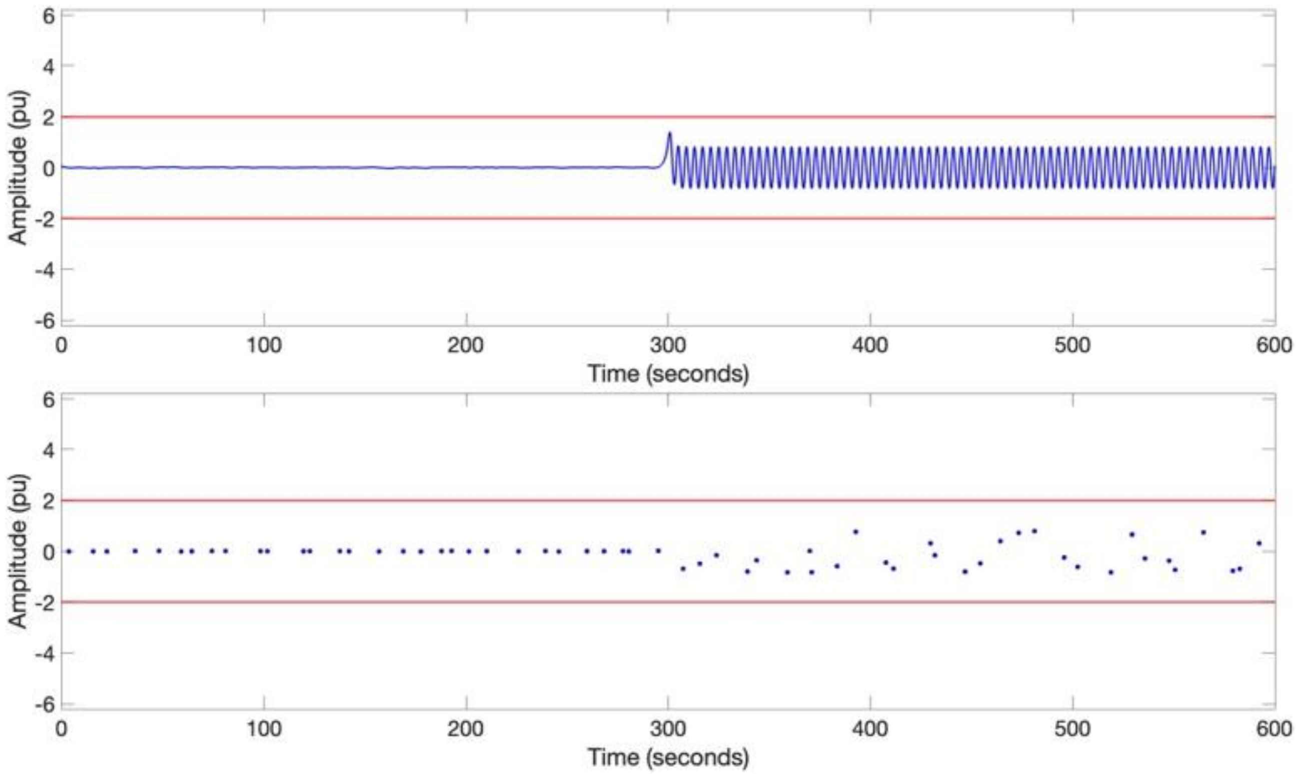


Figure 9: The resulting signal and samples after passing through a low-pass filter

2.5 Maximum Observable Oscillation Frequency

The maximum detectable oscillation frequency will depend on the transducer characteristics. Some of the characteristics of the transducer used in RTE are provided below:

Table 1. Transducer characteristics

Device Name	Sampling frequency (Hz)	Number of samples
Vendor 1	3200	192
Vendor 2	3000	300
Vendor 3	1000	500

The output of the transducer is based on a moving average on the sampled measures.

A moving average is a kind of low-pass filter for which the following formula approximates the -3 dB cutoff frequency:

$$Fcutoff = \frac{0.442947}{\sqrt{N^2 - 1}} \cdot Fsampling$$

N is the number of samples.

Then, we can calculate the cutoff frequency for the devices above:

Table 2. Transducer cut-off frequencies

Device Name	Cutoff frequency (Hz)
Vendor 1	7.38
Vendor 2	4.42
Vendor 3	0.88

From Table 2, we can conclude that, depending on the sampling frequency and number of samples the transducer uses to elaborate the measurement, the observable frequency could be up to several Hz. However, if the sampling frequency is insufficient and the number of samples used is too high, the measurement obtained would be unable to detect modes above 1 Hz, common on synchronous machines.

3. Oscillation Analysis of SCADA Data Using Statistical Method

For the algorithm formulation, the start time and end time of the oscillation event, that means the oscillation duration needs to be known beforehand. The knowledge of the start and end times is used to distinguish between the presence versus absence of the oscillations in the SCADA data. This information can be gathered from any oscillation detection engines, such as FFDD [13], which use PMU-based synchrophasor data to find the start and end times of the oscillation. Once the start and end time information of the oscillation are known, the algorithm uses the SCADA data for detecting the source.

3.1 Processing the Data

The data set is processed first before applying the algorithm. Some generators might have empty columns or contain "NaN" values, and a filter is applied to remove these generators. Again, some generators might have very low outputs. These generators can be ignored by checking them against the minimum output threshold, which can be set as an input parameter in the initial stage. For example, a filter can be designed that removes these types of generators whose maximum output is less than 10 MW or MVAR. Moreover, some generators might have minimal variation in amplitude from the samples. Another filter can also be applied that removes the generators whose maximal difference for the whole data set is less than, say, 5 MW or MVAR.

3.2 Detrending the Data

In the algorithm proposed in this report, two different data-detrending methods have been applied. For the first detrending method, a median filter is applied to the raw data. Then, the difference between the raw data and filtered data is measured to check any oscillating activity in the SCADA signals. For the second detrending method, the difference between two consecutive samples is measured for detrending the whole data set. It is observed that both detrending methods work well with the proposed algorithm in locating the accurate forced oscillation source.

3.3 Setting Up the Transition Band

The lower and upper transition band values are difficult to set up since the SCADA signals that contain oscillating samples might have different amplitudes in different cases. To resolve this issue, an iterative process is introduced in the algorithm where the p value of the respective oscillation window and the ambient window does not change in the iteration process. The transition band starts with a lower value and is increased by a fixed number in each iteration, resulting in a larger transition band in the end. From the first iteration, many generators can be flagged as oscillating sources. As the number of iterations goes up and the transition band increases, the generators that are flagged initially with the lowest transition band will be removed, and only a few generators with high amplitude of data may remain, which will point towards the actual sources of oscillation. In this regard, a ranking of the generators is also done in the algorithm based on their participation in the oscillating event. The formula used to rank the generators is as follows.

$$Ranking\ Factor = \frac{Num_{osc}}{Length_{osc}} \tag{6}$$

Here, Num_{osc} denotes the number of actual transitions among the samples in the oscillation window of the detrended measurements, and $Length_{osc}$ represents the lengths of the oscillation duration or in this case, the total number of samples in the oscillation window. In a similar fashion, Num_{amb} will represent the number of transitions among the samples in the ambient or base window of the detrended measurements, and $Length_{amb}$ will represent the total number of samples in the ambient window.

3.4 Number of Samples and Transition Threshold for the Analysis

The transition threshold k_α for both ambient and oscillation windows is calculated using the binomial distribution formulation until the cumulative probability reaches the pre-specified confidence level. For example, if the confidence level is set to be 95% and the scaled probability of transition is chosen as $p = 0.3$ for the analysis window, the threshold for the number of transitions for the analysis window needs to be at least 3 to find the number of samples n using (7). In this case, the number of samples in the analysis window would be 5. If the SCADA sampling rate is 0.1 Hz or 1 sample in every 10 seconds, the duration of the analysis window needs to be at least 50 seconds for the proposed algorithm.

$$0.95 \le \sum_{i=0}^{k} \binom{n}{i} 0.3^i (1 - 0.3)^{n-i} \tag{7}$$

Table I shows the number of samples against the transition threshold to reach the 95% confidence level, with p being 0.3. For example, if the analysis window has 30 samples, the number of transition threshold would be 13 when calculated with the stated parameters. Referring to Fig. 8, there are 20 transitions in the oscillation window between 300 and 600 seconds. So, this signal would pass the oscillation window test with respect to the transition bands set there. It is also to be noted that the condition for the hypothesis is strict regarding the transition threshold when the number of samples is small, or the duration of the analysis window is short. This condition becomes a little relaxed when the sample number increases as evident from the sample numbers 25 and 30 in Table I.

Table 1: Number of samples and transitions to reach 95% confidence level when p=0.3

Number of Samples	Transition Threshold
5	3
10	5
15	8
20	9
25	11
30	13

3.5 Summarizing the Steps of the Algorithm

The main steps of the algorithm are summarized below.

1) Input SCADA data and the start and end time of oscillation event.
2) Find the oscillation and ambient window length.
3) Detrend the data either using the median filtering method or the differencing detrending method.
4) Count the thresholds of transitions for both oscillation and ambient windows separately using the pre-specified p values and significance levels for both windows.
5) Process the SCADA data based on the empty column or "NaN" values, minimum output threshold, and minimum variations of the generator data.
6) Set up the minimum transition bands value and count the actual number of transitions for both oscillation and ambient window respective of that transition bands.
7) Compare the actual transition count against the thresholds for each signal for both windows separately. If the number of transitions found in the oscillation window is lower than the oscillation window threshold count, reject the signal and move on to the next signal. If not, then check whether the number of transitions found in the ambient window is higher than the ambient window threshold count. If yes, reject the signal and move on to the next signal. In summary, for a signal to qualify for the ranking, the number of transitions in the oscillation window must be higher than the oscillation window threshold, and the number of transitions in the ambient window must be lower than the ambient window threshold.
8) Rank the signals according to (6) and store the results in a result index.
9) Increase the transition band for the next iteration and repeat steps 6 to 8.
10) The generating units flagged where the transition band is largest at the end of the iteration process will point toward the accurate source of oscillation.

4. Validation of the Methodology on Offline Test Cases

4.1 Analysis with Archived SCADA Data

In this section, various real-world oscillation events with different scenarios are analyzed with the proposed algorithm to test the effectiveness of the proposed algorithm by finding the oscillating sources. The data for the first three test cases presented here are provided by RTE from the European interconnection system. The findings were discussed with RTE, and it was verified that the proposed algorithm was successful in identifying the accurate sources of oscillation for all the cases. The fourth case is from the Western system with a low amplitude oscillating source. For each of the cases presented below, the starting point of the transition band is set at 3 after detrending the data, and the maximum transition band is set at 100 with a step increase of 10 in each iteration. The transition band values, as well as the step increase of the iteration process, can be modified in the algorithm according to user preference.

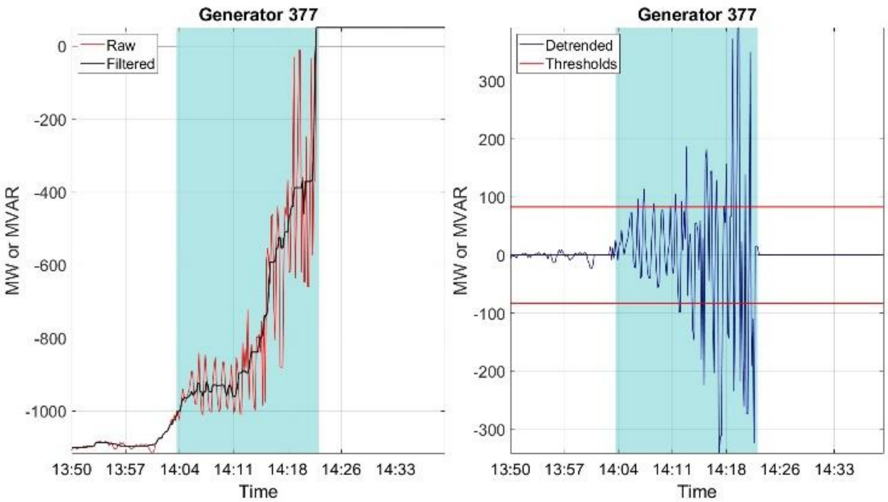
4.2 Test Case 1

This event occurred on August 10, 2022. The event duration was approximately 19 minutes. The oscillation starts at approximately 2:04 PM and ends at 2:23 PM. It occurred in a nuclear power plant during the ramping down of active power of a generation unit. The SCADA data set, which was used for finding the source generator, contains approximately 50 minutes of data and about 2250 signals consisting of real and reactive power measurements. The following table shows the ranking of the generators in ascending order with selected iterations where the highest ranked generator is shown at the top and so on. The transition band values increase in each iteration step. The generators found in which the transition band is largest (in this case, 83) are the sources of forced oscillation.

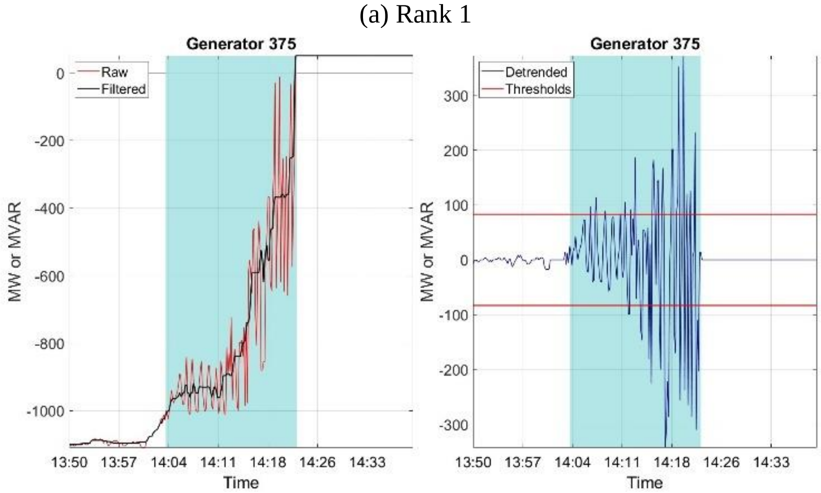
Table 2: Ranked sources for test case 1

Transition Band	±43	±53	±63	±73	±83	±93
Generator	378	378	378	377	377	NA
	1512	1512	1512	375	375	NA
	379	379	379	376	1511	NA
	380	380	376	378	376	NA
	375	376	377	1511	378	NA
	1511	375	375	1512	1512	NA
	377	377	1511	379	NA	NA
	376	1511	380	NA	NA	NA

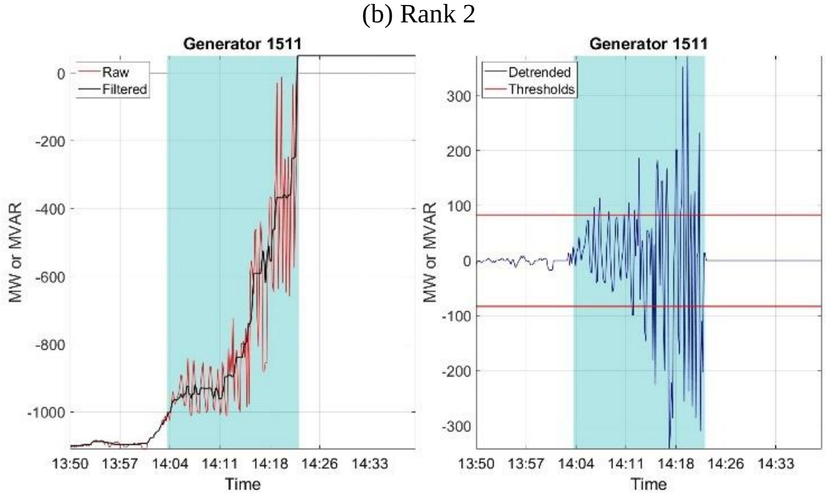
It is to be noted that when the transition band is small at the starting point, many generators can be flagged as possible sources of oscillation. But these initially flagged generators will go away with the increase of the band, and only those with the potential of being possible sources will remain. That is why the above table enlists the generators when the transition band is of a substantial large value. In some test cases, it is found that when the oscillation amplitude is low, the lower transition band can provide the correct possible sources from the beginning of the iteration process. Such examples are given in test case 3, where the sources are located with a few number of iterations.



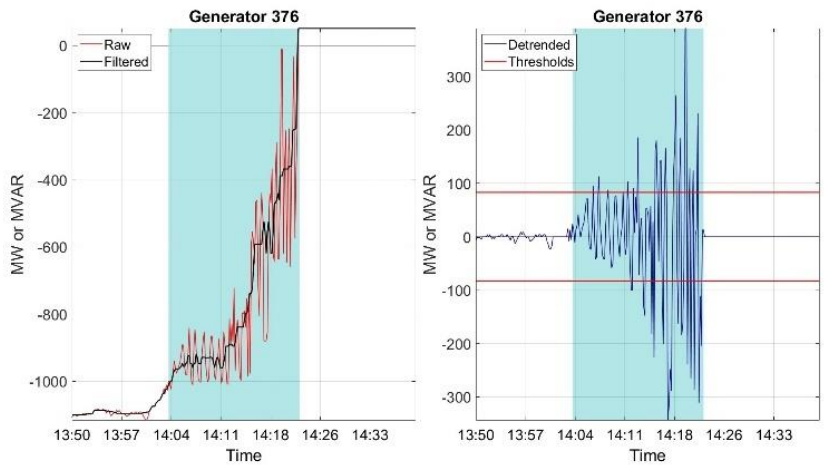
(a) Rank 1



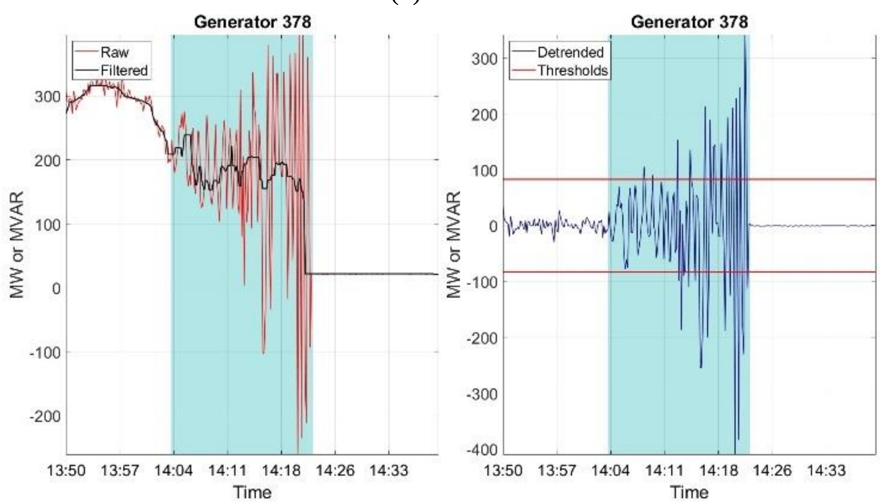
(b) Rank 2



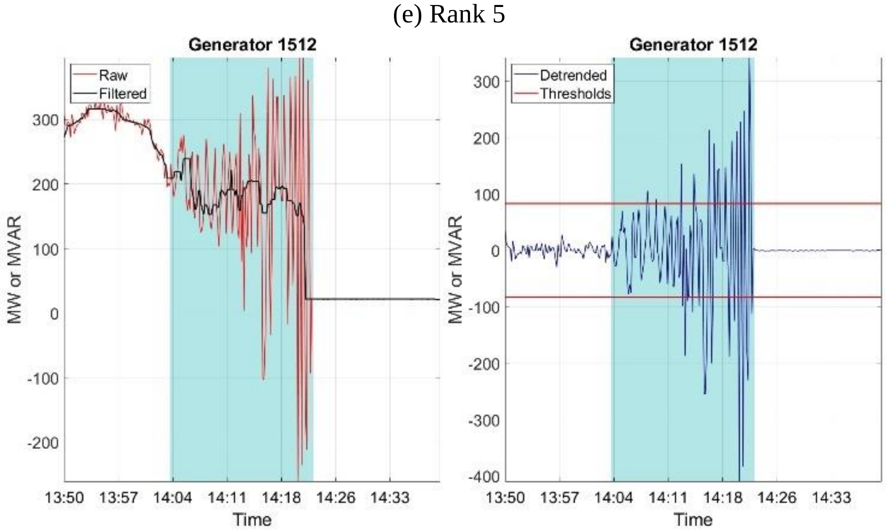
(c) Rank 3



(d) Rank 4



(e) Rank 5



(f) Rank 6

Figure 10: Outputs of the top-ranked generators for test case 1

By observing the above generator outputs in Figure 10, it can be seen that generators 377, 375, 1511, and 376 point towards the same generator. Upon discussing the result with RTE, it was found that the data of this generator was measured from multiple different SCADA units. The same goes for generators 378 and 1512. It also alternatively verifies the consistency of the result.

The same event is analyzed with the transmission line flow data from SCADA measurements to check the consistency of finding the forced oscillation source with the data acquired from the generation side. As shown in Table 3, the algorithm correctly ranks the transmission lines that are adjacent to the source generator.

Table 3: Ranked lines for test case 1

Transition Band	±53	±63	±73	±83	±93
Transmission Line	1208	1208	1204	1204	NA
	1206	1206	1203	1205	NA
	1207	1203	1205	1203	NA
	1203	1204	1208	1208	NA
	1204	1205	1206	NA	NA
	1205	1207	NA	NA	NA
	444	NA	NA	NA	NA

Line numbers 1203 to 1208 correspond to the MW and MVAR flow of the transmission lines adjacent to the oscillating generator. The line measurements were also coming from multiple SCADA units, and instead of showing all of them, line 1204 flow is given in Figure 11.

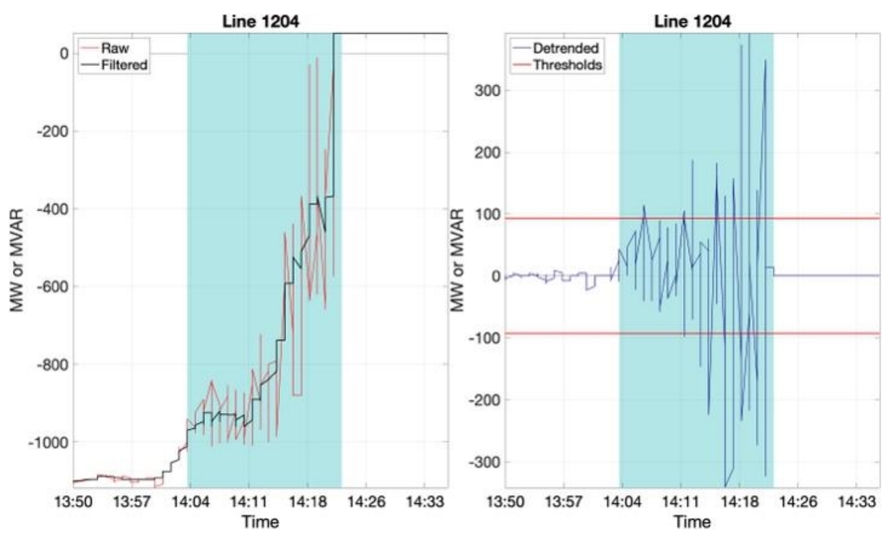


Figure 11: Transmission flow of the top-ranked line for test case 1

The line flow analysis provides another interesting insight in this case. When the transition band is at 53 for the analysis, line 444 also gets included in the ranking. Lines 1203 to 1208 correspond to region 1, and line 444 corresponds to region 2 in the transmission network map of RTE, as shown in Figure 12. Line 444 is connected between France and its neighboring country. As discussed with RTE, an interarea mode is present in this region 2, and it was observed several times from the past events. For this particular event on August 10, 2022, the mode was mildly resonating with the local mode in region 1. This type of ranking and analysis from the proposed algorithm might provide useful information from the oscillation events in the day-to-day operation of any power system network.

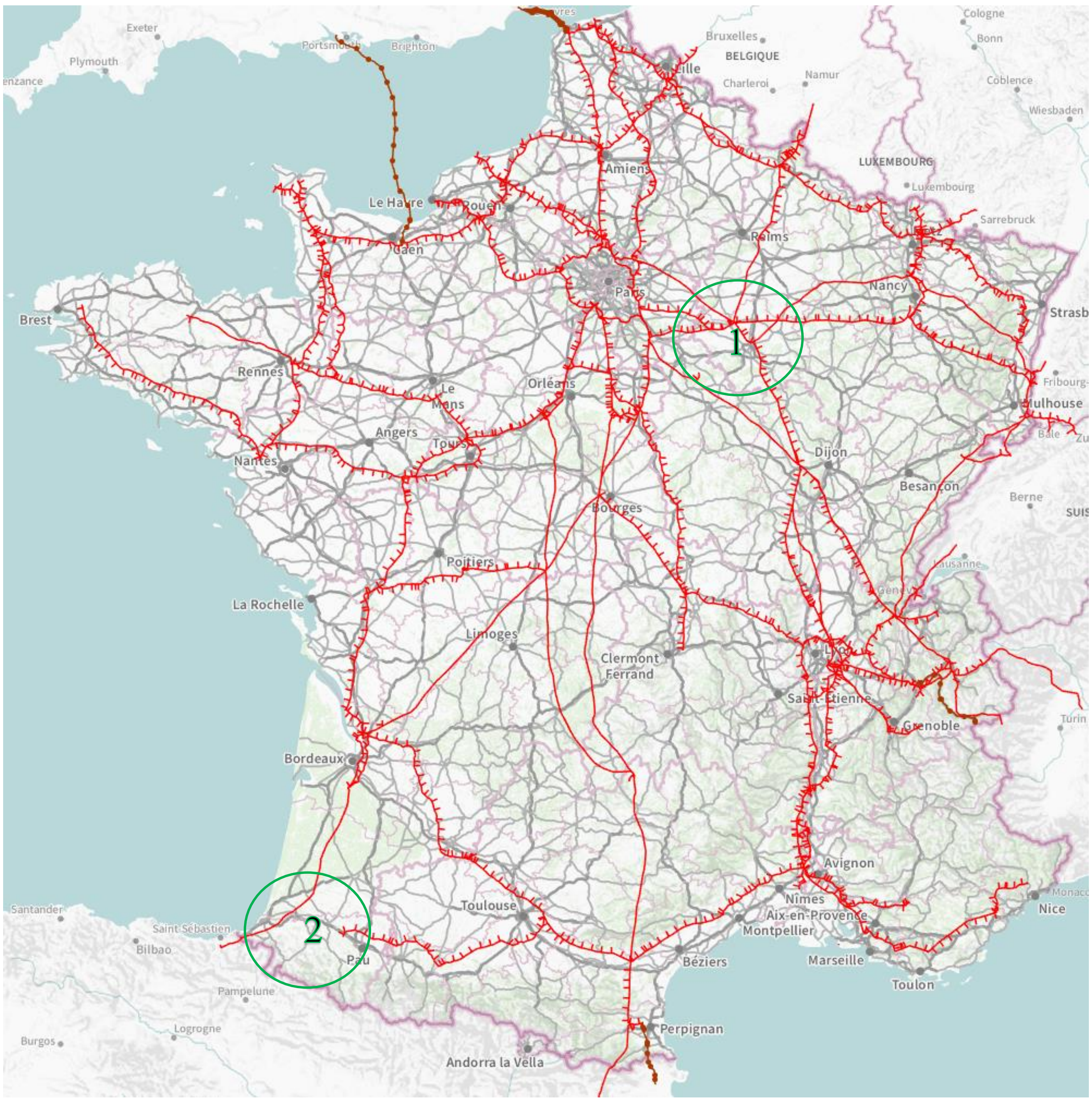


Figure 12: Transmission network of RTE, France

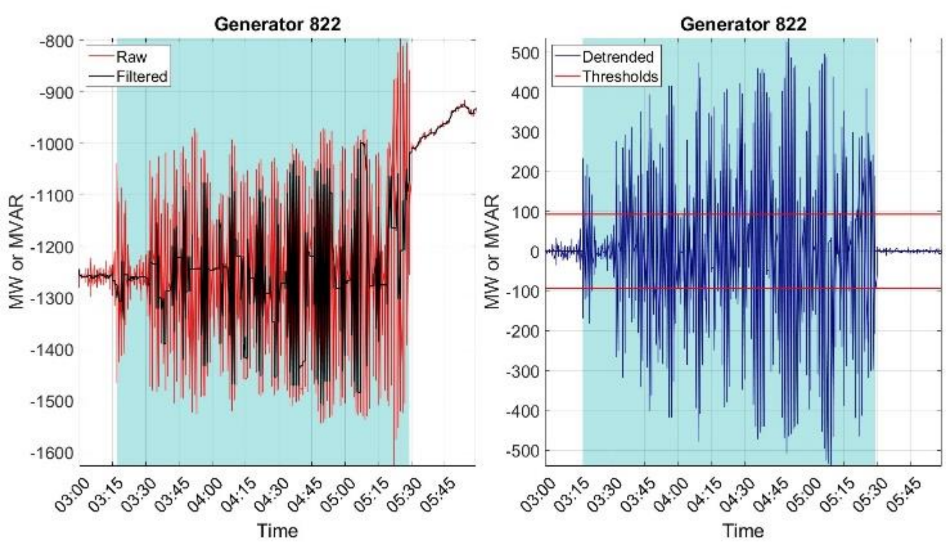
4.3 Test Case 2

This event occurred on September 30, 2017. The oscillation lasted approximately 2 hours and 12 minutes. The oscillation starts at approximately 3:17 AM and ends at 5:29 AM. The SCADA set which was used for analysis contains approximately 3 hours of data and about 1995 signals consisting of real and reactive power measurements. The following table shows the ranking of the generators in a similar fashion as shown in Table 2. The generators found in which the transition band is largest (in this case, 93) are the sources of forced oscillation.

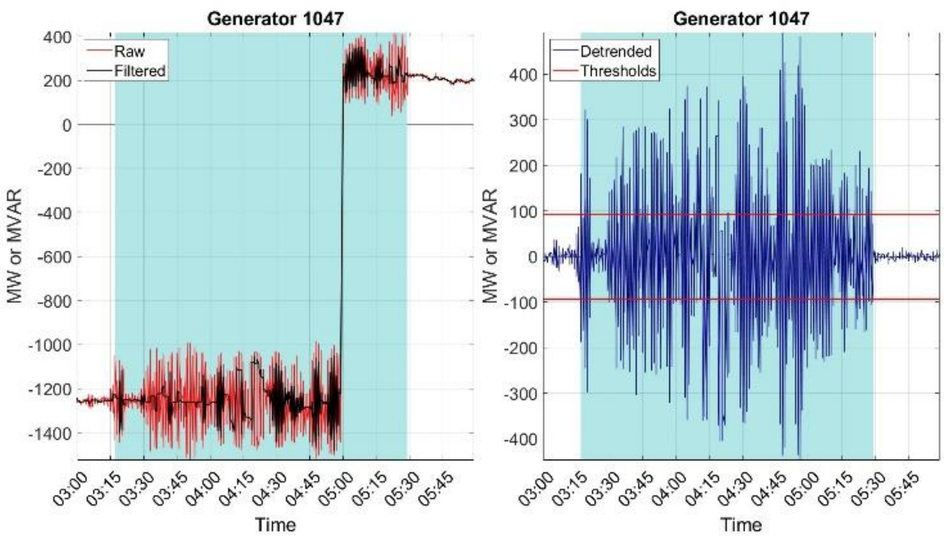
Table 4: Ranked sources for test case 2

Transition Band	±43	±53	±63	±73	±83	±93
Generator	1047	822	822	822	822	822
	822	1047	1047	823	823	823
	823	823	823	1047	1047	1047
	825	825	825	825	824	824
	824	824	824	824	825	825
	828	829	1048	1048	1048	NA
	829	828	828	828	NA	NA
	1048	1048	829	829	NA	NA

The algorithm flags multiple generators to be the possible sources of oscillation. Similar to test case 1, generators 822, 823, 824, and 825 are basically one generator where multiple SCADA units are connected to it. The output of generators 822 and 1047 are shown below in Figure 13, where the oscillation can be inspected visually.



(a) Rank 1



(b) Rank 2

Figure 13: Outputs of the top-ranked generators for test case 2

4.4 Test Case 3

This event occurred on April 11, 2021. The event duration was approximately 9 minutes. The oscillation starts at approximately 1:56 AM and ends at 2:05 AM. The SCADA set which was used for finding the source generator contains approximately 30 minutes of data and about 895 signals consisting of real and reactive power measurements. This test case is different in the sense that the oscillation amplitude is much smaller compared to the previous two test cases. Thus, it is expected that the source would be found with the smaller transition bands and with the first few iterations.

Table 5: Ranked lines for test case 3

Transition Band	±3	±13	±23
Generator	618	618	NA
	714	NA	NA
	622	NA	NA
	715	NA	NA

In this case, the source generator 618 is identified with only two iterations.

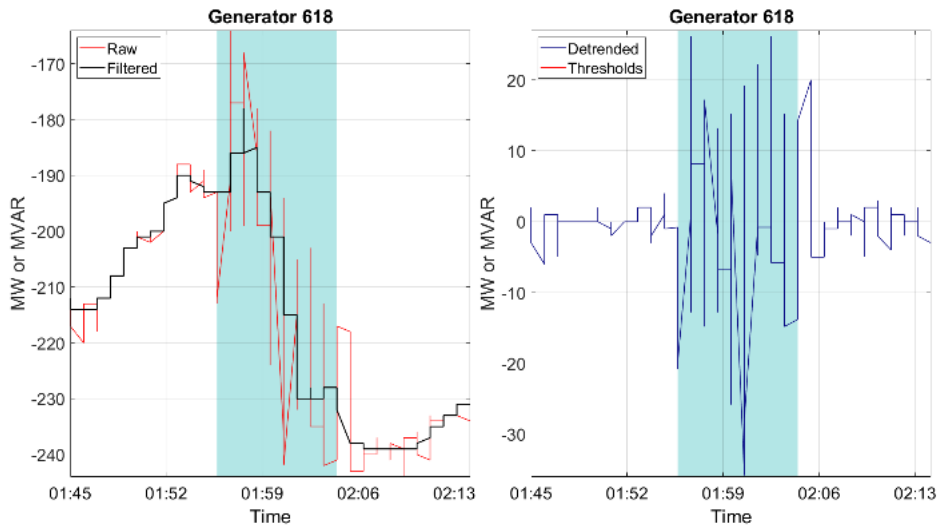


Figure 14: Output of the source generator for test case 3

4.5 Sensitivity of Data Detrending Methods

All the cases were also tested with both the median filter-based data-detrending method as well as the differencing detrending method, and it was found that the proposed algorithm identifies the accurate source of forced oscillation regardless of which data detrending method is used. For example, the algorithm was run with another real event with both median filter and differencing data-detrending methods, and it pointed to the same source both times, which are illustrated in Figure 15 and 16.

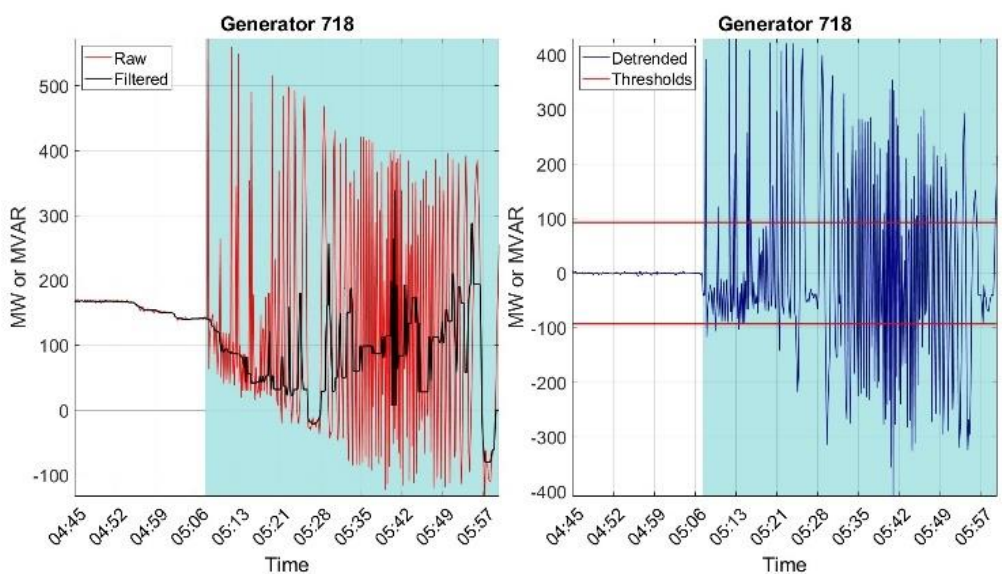


Figure 15: Source identification with the median filter-based detrending method

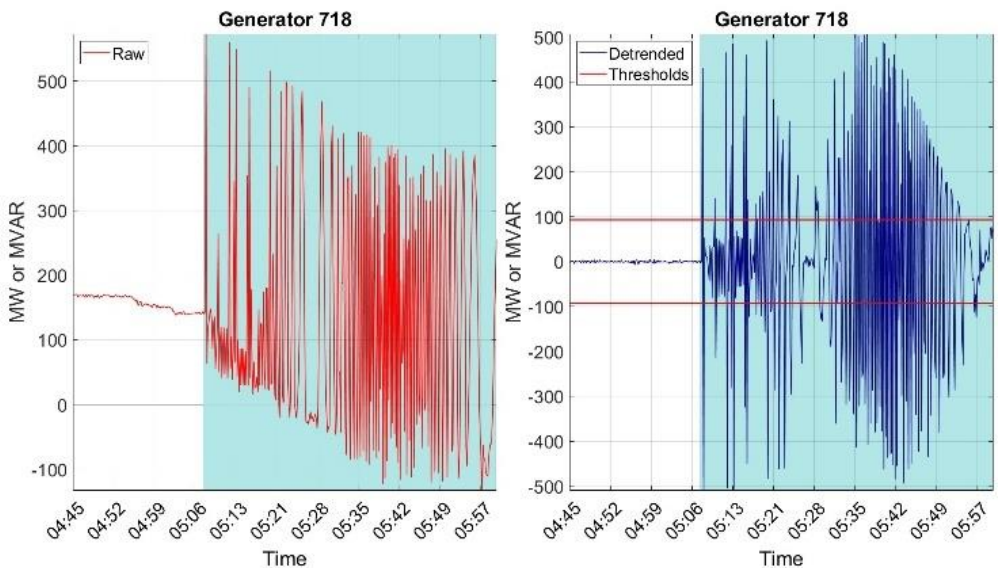
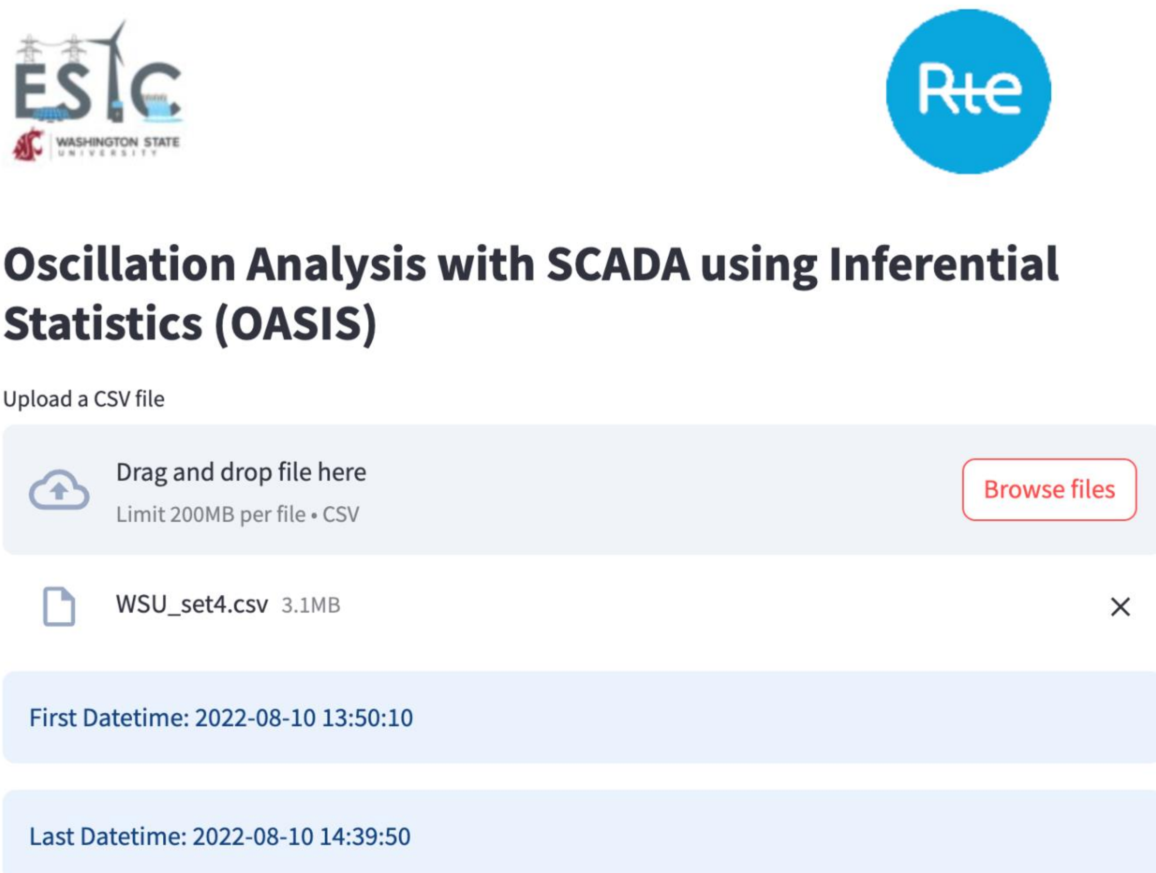


Figure 16: Source identification with the differencing detrending method

The event shown in Figure 15 and 16 occurred on January 29, 2021. The start time of the oscillation was 5:08 AM, and the end time was 6:00 AM. The source of the oscillation was generator 718, which was accurately identified with both median filtering and differencing data-detrending methods.

5. Open-Source Oscillation Analysis Toolbox OASIS

The toolbox is developed using the Python Streamlit library, and it can be run by the streamlit run command with the file location directory in the command prompt in the Windows operating system or the terminal in macOS with any Python Integrated Development Environment (IDE). This web-based application will open a new tab in the default browser of the system where the user can upload the SCADA data file. The first and last date and time information from the data set would be displayed automatically after uploading the data set. The user will then need to provide the oscillation event start and end time as inputs. The application will read the other input parameters from the "Settings" file located in the directory, and it will show the results when the "Run Algorithm" button is clicked.



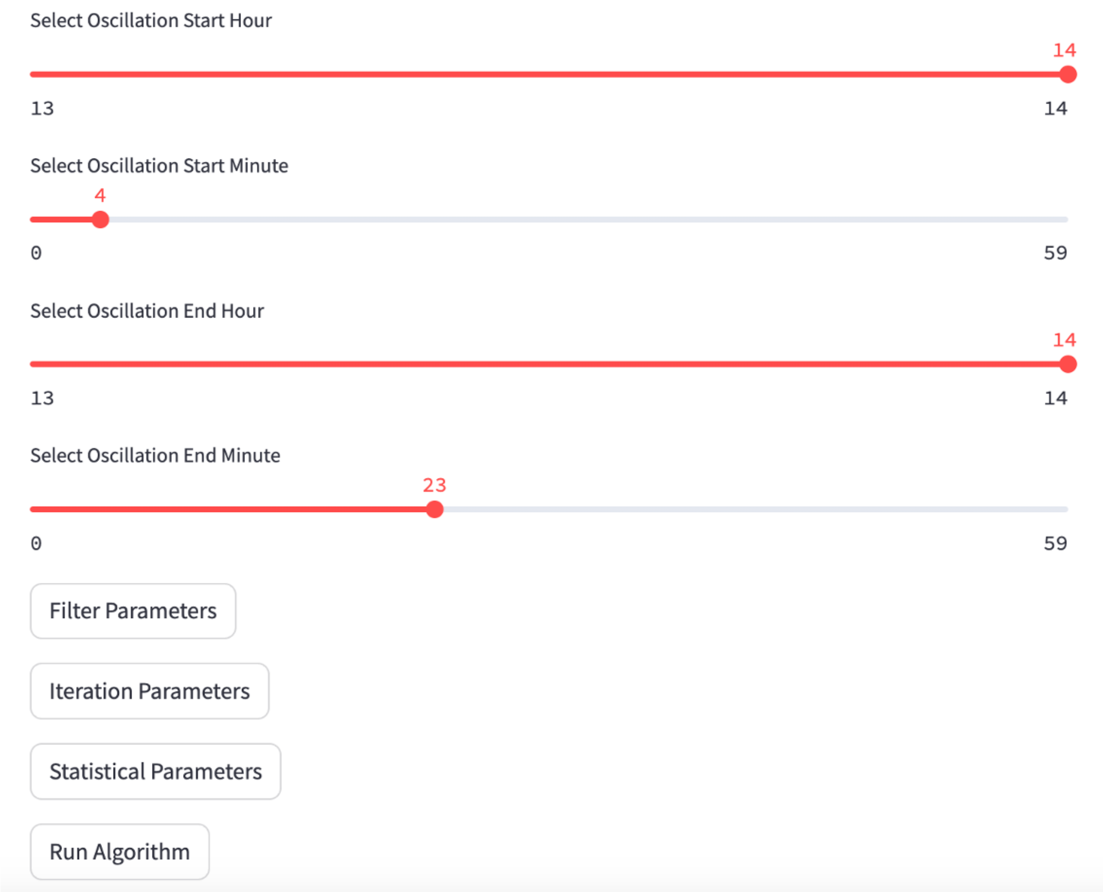


Figure 17: OASIS toolbox for oscillation analysis

6. Conclusion

Oscillations have emerged as one of the main stumbling blocks in the reliable operation of the power grid. Such oscillations can cause potential damage to expensive generator equipment, including rotor shafts, and can also be a problem for power quality for the consumers, in addition to being a safety hazard for power plant personnel. With the increased adoption of renewable generation, the problem is especially showing up in integrating power electronic-based intermittent energy resources into the power grid. Detection and efficient analysis of the oscillations can help the smooth integration of large numbers of renewable devices into the power grid. Also, timely correction of the oscillation issues in fossil fuel and nuclear power plants will prevent unscheduled outages of these energy sources and will help avoid the need for dispatching power from less efficient power plants. From the real cases presented in this report, it is observed that the proposed algorithm can detect the source regardless of the duration of oscillation, the size of the data set, and the amplitude of oscillation. In the future, there are scopes for more investigations to modify the proposed algorithm in this report or to search for statistical approaches on how SCADA data can be utilized for detecting the start of the oscillation so that the algorithm does not need to rely on synchrophasors for the oscillation timestamp information.

References

[1] C. D. Vournas, N. Krassas, and B. C. Papadias, “Analysis of forced oscillations in a multimachine power system,” in Proc. Control 1991. Int. Conf., Edinburgh, vol. 1, 1991, pp. 443–448.

[2] N. Rostamkolai, R. J. Piwko, and A. S. Matusik, “Evaluation of the impact of a large cyclic load on the LILCO power system using time simulation and frequency domain techniques,”IEEE Trans. Power Syst., vol. 9, no. 3, pp. 1411–1416, Aug. 1994.

[3] S. A. N. Sarmadi and V. Venkatasubramanian, "Inter-Area Resonance in Power Systems From Forced Oscillations," in IEEE Transactions on Power Systems, vol. 31, no. 1, pp. 378-386, Jan. 2016, doi: 10.1109/TPWRS.2015.2400133.

[4] Reliability Guideline: Forced Oscillation Monitoring & Mitigation. North American Electric Reliability Coordination (NERC): https://www.nerc.com/comm/PC_Reliability_Guidelines_DL/Reliability_Guideline_Forced_Oscillations_-_2017-07-31_-_FINAL.pdf.

[5] Eastern Interconnection Oscillation Disturbance: January 11, 2019 forced oscillation event, North American Electric Reliability Coordination (NERC), December 2019: https://www.nerc.com/pa/rrm/ea/Documents/ January_11_Oscillation_Event_Report.pdf.

[6] L. Chen, Y. Min, and W. Hu, “An energy-based method for location of power system oscillation source,” IEEE Trans. Power Syst., vol. 28, no. 2, pp. 828–836, May 2013.

[7] S. Maslennikov, B. Wang, and E. Litvinov, “Dissipating energy flow method for locating the source of sustained oscillations,” Int. J. Elect. Power Energy Syst., vol. 88, pp. 55–62, 2017.

[8] J. Follum and J. W. Pierre, “Detection of periodic forced oscillations in power systems,” IEEE Trans. Power Syst., vol. 31, no. 3, pp. 2423–2433, May 2016.

[9] N. Zhou and J. Dagle, “Initial results in using a self-coherence method for detecting sustained oscillations,” IEEE Trans. Power Syst., vol. 30, no. 1, pp. 522–530, Jan. 2015.

[10] D. Osipov, S. Konstantinopoulos and J. H. Chow, "A Cross-Power Spectral Density Method for Locating Oscillation Sources Using Synchrophasor Measurements," in IEEE Transactions on Power Systems, vol. 38, no. 6, pp. 5526-5534, Nov. 2023, doi: 10.1109/TPWRS.2022.3229255.

[11] B. Wang and K. Sun, “Location methods of oscillation sources in power systems: A survey,” J. Modern Power Syst. Clean Energy, vol. 5, no. 2, pp. 151–159, Mar. 2017.

[12] J. O’Brien, T. Wu, V. Venkatasubramanian, and H. Zhang, “Source location of forced oscillations using synchrophasor and SCADA data,” in Proc. 50th Hawaii Int. Conf. Syst. Sci., Waikoloa, HI, USA, vol. 4–7, Jan. 2017, pp. 3173–3182.

[13] H. Khalilinia, L. Zhang, and V. Venkatasubramanian, “Fast frequency-domain decomposition for ambient oscillation monitoring”, IEEE Trans. Power Del., vol.30, no. 3, pp. 1631-1633, Jun. 2015.