Automatically Discerning Power System Dynamics in Synchrophasor Measurements Data Spectra

Chetan Mishra
Dominion Energy, Richmond, VA
chetan31@vt.edu

Luigi Vanfretti,
Rensselaer Polytechnic Institute,
ECSE, Troy, NY
vanfrl@rpi.edu

Jaime De La Ree Jr., Kevin D. Jones
Dominion Energy,
Richmond, VA

***Abstract*— Techniques for time-frequency analysis can be used to identify, isolate, and track the emergence of numerous underlying power system dynamic processes just from measurements. These are especially useful for drawing meaningful conclusions from data obtained under ambient conditions, when power system phenomena can be less obvious in the time domain. However, because spectrograms have typically been evaluated by visual inspection, it is impossible to routinely perform assessment on large amounts of data. More crucially, relying just on visual inspection could result in missing crucial dynamics due to human error, e.g., setting incorrect color scaling. To address these issues, this work proposes an approach to automatically identify regions of interest on spectrograms over large data sets, which reflect the underlying power system's dynamic behavior. Results for actual synchrophasor measurements from the Dominion Energy system are obtained to show the effectiveness of the proposed method.**

***Index Terms*--Spectrogram, Synchrophasors, Oscillations**

I. INTRODUCTION

With the growth of renewable generation, utilities are faced with major challenges, one of them being a limited ability to anticipate dynamic performance issues, particularly due to the performance of local inverter-based resources (IBRs) controllers and other power electronic-based devices. These issues arise from the lack of access to transparent and accurate models and changing grid conditions that would require a more frequent performance monitoring and control tuning. What makes this even more complex is the inherently different nature of these devices when compared to well-understood synchronous machines. As a result, it is frequent to encounter power dynamic behavior that are challenging to understand. Luckily, with the proliferation of synchrophasor technology [1], it has been possible in theory to discover the various dynamic phenomena at play.

Large disturbance transients/events that allow to assess the dynamic performance of power system components are rare, even though they are excellent for providing insights into the behavior of the underlying system. Furthermore, it is frequently too late when a specific dynamic process erupts into an unstable oscillation event, which in turn calls for an immediate and costly fix. In our experience, before significant disturbances occur, there are several clues left from the preceding "normal" operating conditions and therefore, there is a need to be able to automatically extract relevant information from ambient conditions [2]. In this regard, spectral analysis techniques have proven valuable [3] for ambient data analysis, owing to the stochastic linear nature of the system in ambient conditions.

Power system transient/event analysis has benefited greatly from the use of time-frequency analysis techniques like the short term Fourier transform (STFT), wavelets, and others [4] by providing insights into the evolution of event characteristics and also by distinguishing them from the regular behavior. Surprisingly, these are underutilized in synchrophasor ambient data analysis, where the major focus of this field has been in electromechanical mode estimation. Since wide area modes, have largely stationary properties such as frequency and damping, only a sub-set of time-frequency analysis techniques are exploited. Unfortunately, this does not hold true for local control dynamics, which can be nonstationary. As a result, time-frequency analysis techniques are more suited for identifying and characterizing dynamic behaviors from measurements because they allow us to characterize power system dynamics (modes) using time-frequency traces over a broader frequency range rather than for single or a few specific frequencies. Additionally, some dynamics are only possible to understand by integrating the temporal component in modal analysis. As an example, Fig. 1 spectrogram is generated from the voltage magnitude at a STATCOM in the Dominion Energy system [5]. One might believe that the underlying system has numerous oscillatory dynamics ranging from 1-4 Hz if it were processed independently frame by frame. A time-frequency representation, however, exposes connections between traces that maybe product of harmonic relationships. This example's mode becomes progressively less damped as the network's current imbalances increase (i.e., the three-phase network currents are not fully balanced). The absence of labeled historical data in the form of dominant modes and their attributes (frequency, damping, etc.) at each epoch is the sole significant obstacle preventing machine learning from automatically understanding such relationships [6].

This brings us to the main contribution in this paper, which is to develop a simple approach to automatically capture frequencies in real world measurements corresponding to dominant dynamic behaviors that evolve over time. This will

allow utilities to discover and characterize new dynamic processes, track previously discovered ones, to detect ones tending to instability and finally, generate labels in historical data to enable machine learning applications for possibly learning preventive operational solutions to unmodeled dynamic issues. To the best of our knowledge, this problem has not been addressed in the power system literature.

II. BACKGROUND

A. Spectrum of Power System Ambient Dynamics

Power system ambient dynamics can be characterized by a time varying linear system driven by a white noise input [2]. The time-varying component results from the changing operating conditions that are brought on by changes in load, generation dispatch, etc, with the main assumption for modeling being that the system will respond within its linear regime. Under this assumption, it is common to model the system using a linear time-invariant for the period of analysis (~20-30 mins). However, control modes may change quickly, which may require to study ambient dynamics as a generalized time-varying linear system driven by stationary white noise $u(t) \sim N(0,1)$. The output $x(t)$ for such systems is,

$$x(t) = \int h(t,\tau)u(t-\tau)d\tau \tag{1}$$

where $h(t,\tau)$ represents the time varying impulse response and is defined as response at time t for impulse input at $t-\tau$. Cramer's representation of [7] of stationary process $u(t)$ gives,

$$x(t) = \int h(t,\tau)\int e^{j\omega(t-\tau)}dz(\omega)\,d\tau = \int e^{j\omega t}H(t,\omega)dz(\omega) \tag{2}$$

where, $z(\omega)$ is a zero mean stochastic process with orthogonal increments i.e. $E(dz(\omega_1)dz^*(\omega_2)) = \delta(\omega_1 - \omega_2)S_u(\omega_1)d\omega_1$, $S_u(\omega)$ represents $u(t)'s$ power spectral density, which is equal to 1 (constant) in this case, $H(t,\omega) = \int h(t,\tau)e^{-j\omega\tau}d\tau$ and $E(.)$ is the expectation operator. The time varying spectrum/evolutionary spectrum [7] for $x(t)$, denoted by $S(t,\omega)$ is then derived as $|H(t,\omega)|^2$. Here, the dynamic modes/oscillations of the underlying system, appear as peaks (local maxima) in $S(t,\omega)$, which we are interested in capturing. Note that the discrete system parameters such as network topology are accounted for by assuming $S(t,\omega)$ to be piecewise continuous in t.

B. Time-Frequency Analysis from Measurements

The aim of time-frequency analysis is to estimate what phenomenon happens (spectral characteristics) at what time in a signal. These operate by decomposing the measurement signal into several constituent components, which usually correspond to distinct underlying phenomena in the actual system. From computational viewpoint as well as the simplicity in implementation and analysis of the final result, of particular interest are linear class of techniques [8]. In these, the signal is characterized by inner products with a pre-assigned family of templates. Normally, these templates are skewed in time and frequency plane to distinguish between multiple phenomena at different times and time scales. In this work, we use STFT, which is defined for $x(t)$ as,

$$\boldsymbol{STFT}\{x\}(t,\omega) = X(t,\omega) = \int x(\tau)w(\tau-t)e^{-j\omega\tau}d\tau \tag{3}$$

where $w(t)$ is the window function, designed to have a compact support in time i.e., there exists a finite T s.t. $\int_{-\frac{T}{2}}^{\frac{T}{2}} w^2(t)dt \approx \int_{-\infty}^{\infty} w^2(t)dt$. Finally, spectrogram is defined as the magnitude squared of STFT coefficients i.e., $|X(t,\omega)|^2$. Substituting ambient dynamics and taking expectation,

$$\begin{aligned} &E(|X(t,\omega)|^2) \\ &= \int\int\int e^{j\omega_1(\tau_1-\tau_2)}H(\tau_1,\omega_1)H^*(\tau_2,\omega_1)w(\tau_1-t)w(\tau_2 \\ &\qquad - t)e^{-j\omega(\tau_1-\tau_2)}\,d\omega_1 d\tau_1 d\tau_2 \end{aligned} \tag{4}$$

Now, assume that T is chosen according to how fast the underlying modes/spectrum change i.e., $H(\tau,\omega) \approx H(t,\omega) = \forall\omega\ \forall\,\tau \in \left(t-\frac{T}{2}, t+\frac{T}{2}\right)$,

$$E(|X(t,\omega)|^2) \approx \int |H(t,\omega_1)|^2|W(\omega-\omega_1)|^2 d\omega_1 \tag{5}$$

Thus, spectrogram gives an estimate of the time varying spectrum $S(t,\omega) = |H(t,\omega)|^2$ convolved with the window function in frequency domain, resulting in spectral smoothing. Now, an increased time resolution (small T) comes at the expense of lowered frequency resolution according to uncertainty principles in time-frequency analysis, which will pose a challenge when dealing with very fast changing spectra. Note that while many other advanced techniques exist [8] for obtaining a better estimate, our contribution lies in detecting regions of interest/ peaks in the time-frequency plane and is independent of the specific technique for spectral estimation.

III. PROPOSED METHODOLOGY

In this section, we propose a methodology to discern oscillatory dynamics/modes in the time-frequency space. Let us denote the true time varying spectrum of ambient dynamics by $S(t,f)$ and its estimate $\hat{S}(t,f)$, obtained from the spectrogram. The underlying dynamic behaviors that we wish to capture can be loosely characterized by a set P of spectral peaks $P = \{(t^*,f^*) \in R^2 | \partial_f S|_{t^*,f^*} = 0, \partial_f^2 S|_{t^*,f^*} < 0\}$.

A. Practical Challenges

Identifying P from $\hat{S}(t,f)$ is not straightforward in practice. First, the tradeoff between time and frequency resolution leads to problems such as the merging of closely spaced modes and the spreading of some modes' energy across a range of frequencies. For instance, modes with fast-varying frequencies in 0–2 Hz and 10–12 Hz regions in Fig. 8 prevents them from having a distinct spectral peak. Thus, the definition of P is modified to now represent a set of neighborhoods with locally maximum spectrum values, instead of clearly distinct peaks. Remember that we can always post-process P to further characterize its structure (e.g., we can collapse these neighborhoods wherever possible to a single frequency value.), which will be explored in future.

Secondly, the stochastic nature of ambient dynamics will result in a noisy $\hat{S}$, which could either result in spectral regions classified as relevant dynamics or dynamic behaviors corresponding to poorly observed, less excited and/or well damped modes being lost in the noise. Fig. 1 shows one such low energy mode labeled "2". Careful preprocessing is required to make the relevant dynamics stand out.

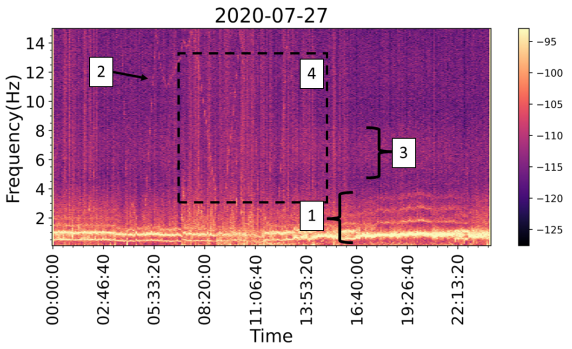

Figure 1 Test Spectrogram, key dynamics – 1. Time varying control mode and harmonics, 2. Fast varying low amplitude mode 3. 6-8 Hz dynamics 4. Arc furnace broad band effect

The proposed approach is comprised of three major steps: (i) obtaining a low variance estimate of time-frequency plot, (ii) pre-processing the spectrogram to enhance mode dynamics and finally, and (iii) detecting relatively high energy regions that represent time-varying dominant dynamics. A 24 hour voltage magnitude spectrogram from a 115 kV STATCOM shown in Fig. 1 is used to gain visual insight into each step of our approach. The STFT is computed with non-overlapping, 5-minute Hann windows. To minimize the impact of different magnitude scaling on dynamics, the spectrum values are plotted in decibels rather than real values. We will use the subscript db on spectrum variables to represent the $10\log_{10}(.)$ operation. The colormap range is set to 2-98 percentile of these values.

B. Low Noise Spectrum Estimate

The previously stated spectrogram estimate from STFT coefficients performs badly, especially in terms of variance. It should be noted that this cannot be resolved by arbitrarily smoothing the generated spectrogram using filters because doing so may cause the spectral peaks to become flat. In this work, we use Welch's method [3], which averages the periodograms obtained from shorter segments of the same data window, thus reducing variance at the expense of frequency resolution, which can be accounted for as discussed before.

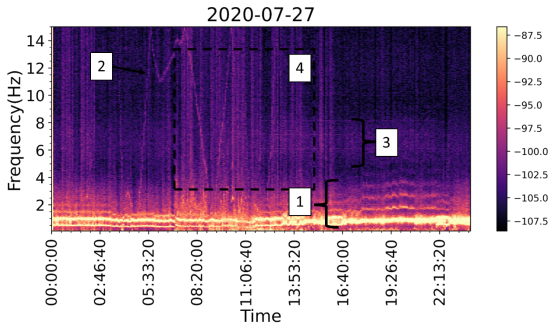

Figure 2 Welch Spectrogram

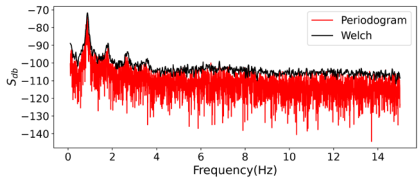

Figure 3 Welch vs Periodogram 5 Min Window at 7 PM

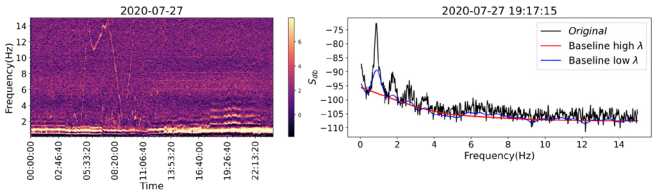

Figure 4 **Left:** Baseline removed spectrogram **Right:** non-flat baseline

In this work, the Welch's spectrogram is obtained by dividing each non-overlapping 5 min window into 1 min windows (1/5th the frequency resolution) with 50% overlap, thereby maintaining the same time-resolution as the STFT in Fig. 1. The noise suppression properties can be seen in Fig. 3 where a column representing 5 min bock of data starting at 7 PM of the two spectrograms is compared. Welch's technique reveals the higher peaks in the 2-4 Hz range, which where were barely noticeable in the original spectrum.

C. Spectrogram Preprocessing

In real-world power grids, in addition to the oscillatory ambient dynamics of interest, there are many additional undesirable behaviors that may hinder the detection of the former. One such behavior results from slow trends in the measurements caused by changing operating conditions. While a lot of approaches have been proposed for detrending synchrophasor data [2], it is difficult to completely remove them in practice because the ground truth is unknown. Furthermore, there are non-oscillatory dynamics characterized by real eigen values. Both will have the effect of a non-flat spectral baseline, especially at lower frequencies < 1 Hz as shown in Fig. 4 for the 7:17-7:22 PM part of spectrum of Fig. 2. These can occasionally interfere with the detection of spectral peaks and leak into the identified regions of interest. Step changes in operating conditions, such as changing the voltage setpoint in a generator or switching a capacitor, are another group of behaviors that may not be of relevance but exhibit high spectrum values and can be mistaken for dynamics of interest. These have a brief broadband effect, like the vertical stripes in Fig. 1 (labeled 4).

By estimating and eliminating the spectral baseline of $\hat{S}_{db}(t,f)$, it is possible to solve both issues. Here, the spectral baseline $b(t,f)$ can be defined as a smooth, under-approximation of the spectrum. Note that t here represents the 5 min blocks being processed to obtain the spectrogram. Our previously proposed approach [9], which is employed in this work at each t, estimates $b(t,f)$ by solving the following optimization problem,

$$min_{b(t,f)} \sum_f w(err(t,f)) \times err^2(t,f) + \lambda \sum_f \left(\nabla_f^2 b(t,f)\right)^2 \qquad (6)$$

where, the residual $err(t,f) = \hat{S}_{db}(t,f) - b(t,f)$, $w(err) = p(err \geq 0) + (1-p)err(< 0)$ and $\nabla_f^2(.)$ is the double difference operator w.r.t. f. The parameter p forces the baseline to be below the spectrum while λ penalizes curvature. A small value of λ runs the risk removing relevant system dynamics such as 6-8 Hz in Fig. 4 for the blue estimate. The results of baseline removal on the spectrogram of Fig. 2 are also shown in Fig. 4. Observe that the vertical lines (labeled 4 in Fig. 1) are

completely removed, thereby making the relevant dynamics much easier to discern.

D. Detecting Points of Interest Using Image Thresholding

Mode dynamics at each time are characterized by local maxima in the spectrum. Broadly speaking, there are two ways of identifying these viz. using a robust peak detection technique such as [10] or solving the problem as an image thresholding problem [11]. In this work, we explore the latter owing to the simplicity in terms of the number of hyperparameters to tune. The simplest technique for binarizing photos is thresholding. Each pixel is compared to a corresponding threshold value to determine whether to keep it or discard it. Although a variety of thresholding methods have been created over time, none of them are universally applicable to all kinds of images. Histogram-based approaches, which take advantage of the distribution of pixel intensities for the provided family of images, are of special relevance in terms of simplicity. However, understanding the nature of pixel intensity (read $S_{db}(t,\omega)$) histograms produced by our power system ambient data spectrograms is necessary to apply these techniques. To this end, we first define a cumulative probability measure $\mu(z) = \frac{\int\int I(S_{db}(t,\omega)\leq z)dtd\omega}{\int\int dtd\omega}$ where $I(.)$ is an indicator function. Simply put, $\mu(z)$ refers to the number of (t,ω) pairs for which $S_{db}(t,\omega) \leq z$. The histogram then gives the empirical $\frac{d\mu}{dz}$.

Let us first understand $\mu(z)$ for some relevant stationary processes. For i.i.d $x(t)\sim N(0,\sigma^2)$, $S_{db}(\omega) = 20\log_{10}\sigma = \sigma_{db}$ and therefore, $\mu(z) = 1\forall z \geq \sigma_{db} \Rightarrow \frac{d\mu}{dz} = \delta(z-\sigma_{db})$ i.e. a unimodal distribution (one peak). Next, consider $N \ll \int d\omega$ number of undamped sinusoids with measurements noise $n(t)\sim N(0,\sigma^2)$ i.e., $x(t) = \sum_{i=1:N} A_i cos(\omega_i t + \phi_i) + n(t)$. $S_{db}(\omega) = 10\log_{10}(\sum_{i=1:N} A_i^2\delta(\omega-\omega_i)+\sigma^2)$. $\mu(z) \approx 1\ \forall z \geq \sigma_{db}$ i.e., approximately unimodal. Here, the histogram peak is at σ_{db} while the oscillatory portion $A_i's$ are towards the right tail of the distribution since $S_{db}(\omega_i) > \sigma_{db}$. Finally, for a linear system with a transfer function $H(\omega)$ driven by a white noise input with variance σ^2, $S_{db}(\omega) = 20\log_{10}|H(\omega)| + \sigma_{db}$. The resulting distribution tends to be unimodal for finite number of observable poles, which is true in practice. In this case, the histogram peak/mode represents the driving noise spectrum $\sigma_{db} + k$, where k is a constant. To comprehend how these findings apply to spectrograms produced by non-stationary ambient dynamics, it is important to note that the pixel distribution will be unimodal if the relevant dynamics' (high spectrum values) coverage area is much less than that of the background. This holds true since time-frequency analysis operates by choosing an appropriate template/basis to obtain a sparse representation of the signal. Note that baseline removal has an added advantage of aligning the noise floors across various timeframes, thereby further assuring unimodality.

Finally, we adopt one of the most popular techniques to threshold unimodal images as proposed by Rosin [11]. It works by drawing a straight line connecting the peak of the histogram to the first empty bin, then locating the point on the histogram that is furthest from this line to serve as the threshold, as illustrated in Fig. 5 by black dotted line. Fig. 6 shows the results for thresholding applied to Fig. 4 based on its pixel distribution shown in Fig. 5. The suggested framework effectively captures all dynamic behaviors with hardly any false positives. Despite these positive results, it's important to understand the method's underlying assumptions. First, it assumes that the dominant class of pixels (histogram peak), which in our instance represents the noise floor as previously described, has a lower intensity than the pixels we want to detect. Second, there is a discernible corner at the major histogram peak. Now, any significant dynamics that are not sparsely represented in the spectrogram, typically because of the technique's subpar frequency resolution, may potentially interfere with that. A fast varying dynamics shown in Fig. 9 results in the 0-2 Hz obfuscated region in Fig. 8 and Fig. 10, which adds the red highlighted region in the histograms in Fig. 7. These can easily be dealt with by an improved time-frequency analysis technique or through a more aggressive baseline removal (lower λ).

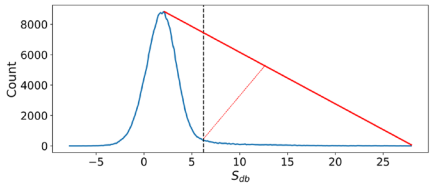

Figure 5 Rosin's Image Threshold for Spectrogram

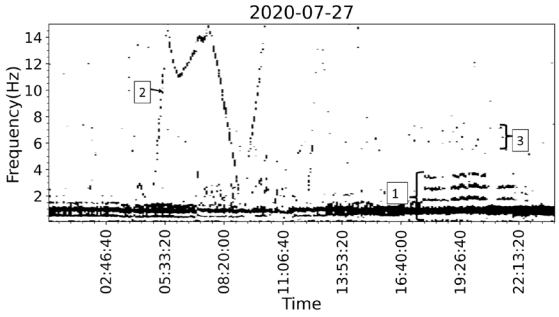

Figure 6 Binarized Spectrogram

IV. RESULTS

To contextualize the proposed methodology, we provide a two more examples from the Dominion Energy's real-world synchrophasor data in Fig. 8 and Fig. 10. First, we discuss their common dynamics. Firstly, there is an 8 Hz mode, which is only active from sunrise to sunset, found to be from a transmission level solar plant [12]. Analysis of point of wave data showed it to be at 22 Hz, showing up at 8 Hz due to aliasing. The region in 0-2 Hz does not have a clear frequency resolution as mentioned before due to low magnitude oscillations with fast varying fundamental frequency, as shown in Fig. 9. The broadband effect of arc furnace response is also common to both. Now, let us focus on their unique characteristics.

First case shown in Fig. 8 is that of a voltage magnitude of a data center load region. It has a ~ 14 Hz oscillation from 2-4 AM resulting. Besides that, there are repeating and persistent peaks at 1 Hz and its multiples. These are a result of a periodic, once a second voltage sag [13]. Another, difficult to characterize dynamics are visible as a blurred region in 10-12 Hz range. Further analysis revealed a once a second ringdown of a 10.5-11 Hz mode. Finally, starting at around 16:00 Hrs, the spectral baseline shifts up, likely due to some local device

connecting or disconnecting. The binarized spectrogram obtained through the proposed approach can capture every key feature of this spectrogram. The second case is at 115 kV side of a distribution substation. A defining feature of its spectrogram (in Fig. 10) is an oscillation with large variations in frequency. Analyzing more days showed this pattern to be random, later traced to a manufacturing plant. In our experience, such behaviors are commonly observable whenever power electronic drives are involved. Furthermore, there are very low energy, easy to miss periodic components such as the one labeled 5. The proposed approach is again able to capture all the relevant dynamics, even the lowest energy ones, while being robust to estimation noise.

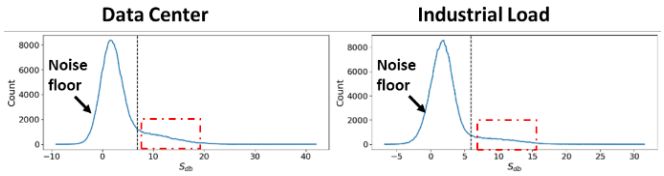

Figure 7 Spectrum Distribution

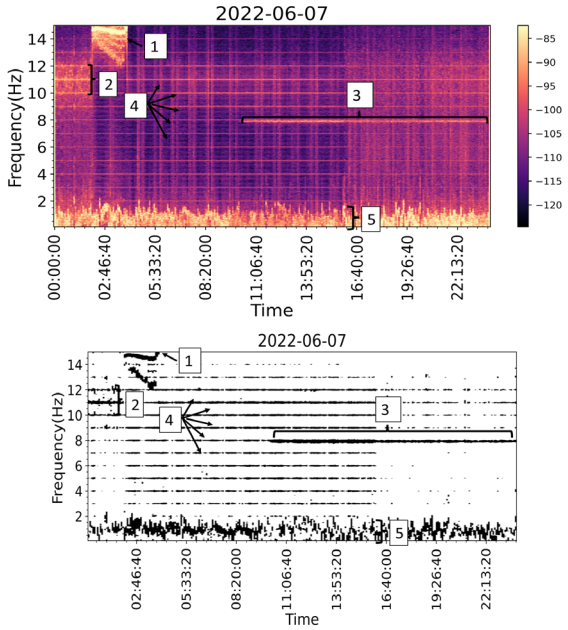

Figure 8 Data Center – 1. forced oscillation 2. intermittent ringdown dynamics 3. solar dynamics 4. periodic voltage sag 5. fast varying mode

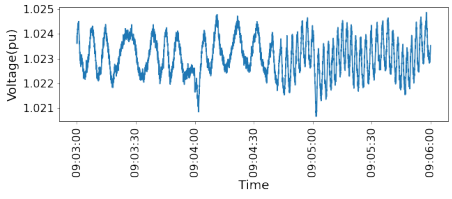

Figure 9 Time Varying 0-2 Hz Mode

V. Discussion and Future Work

A methodology for automatically capturing essential spectrogram regions that correlate to relevant power system dynamic behavior is proposed in this work. Future research will examine adding more structure to the identified regions such as time-frequency traces, harmonic coupling, etc.

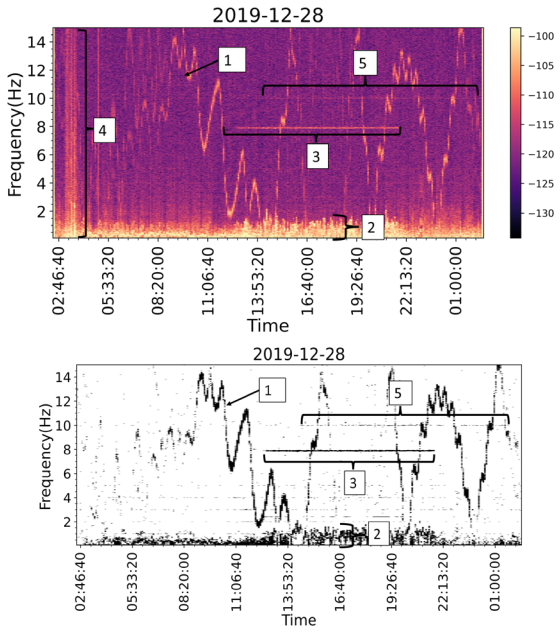

Figure 10 Industrial Load – 1. industrial motor drive 2. fast varying mode 3. solar mode 4. arc furnace 5. low energy undamped oscillation

References

[1] A. G. Phadke and J. S. Thorp, *Synchronized phasor measurements and their applications*, vol. 1. Springer, 2008.

[2] L. Vanfretti, S. Bengtsson, and J. O. Gjerde, "Preprocessing synchronized phasor measurement data for spectral analysis of electromechanical oscillations in the Nordic Grid," *Int. Trans. Electr. Energy Syst.*, vol. 25, no. 2, pp. 348–358, 2015, doi: 10.1002/etep.1847.

[3] P. Stoica and R. Moses, *Introduction to Spectral Analysis*, 1st edition. Upper Saddle River, N.J: Prentice Hall, 1997.

[4] D. C. Robertson, O. I. Camps, J. S. Mayer, and W. B. Gish, "Wavelets and electromagnetic power system transients," *IEEE Trans. Power Deliv.*, vol. 11, no. 2, pp. 1050–1058, Apr. 1996, doi: 10.1109/61.489367.

[5] C. Mishra *et al.*, "Analysis of STATCOM Oscillations using Ambient Synchrophasor Data in Dominion Energy," in *2022 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Apr. 2022, pp. 1–5. doi: 10.1109/ISGT50606.2022.9817489.

[6] X. Xu *et al.*, "Fast Oscillation Detection and Labeling via Coarse Grained Time Series Data for ML Applications," in *2022 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Apr. 2022, pp. 1–5. doi: 10.1109/ISGT50606.2022.9882712.

[7] M. B. Priestley, "Non-linear and non-stationary time series analysis," *Lond. Acad. Press*, 1988.

[8] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990, doi: 10.1109/18.57199.

[9] M. de Castro, C. Mishra, L. Vanfretti, and K. D. Jones, "A Novel Method for Despiking Spectra from Synchrophasor Measurements," in *2021 IEEE Power & Energy Society General Meeting (PESGM)*, Jul. 2021, pp. 1–5. doi: 10.1109/PESGM46819.2021.9638097.

[10] P. Du, W. A. Kibbe, and S. M. Lin, "Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching," *Bioinformatics*, vol. 22, no. 17, pp. 2059–2065, Sep. 2006, doi: 10.1093/bioinformatics/btl355.

[11] P. L. Rosin, "Unimodal thresholding," *Pattern Recognit.*, vol. 34, no. 11, pp. 2083–2096, 2001.

[12] C. Wang, C. Mishra, K. D. Jones, R. M. Gardner, and L. Vanfretti, "Identifying Oscillations Injected by Inverter-Based Solar Energy Sources," in *2022 IEEE Power & Energy Society General Meeting (PESGM)*, Jul. 2022, pp. 1–5. doi: 10.1109/PESGM48719.2022.9916830.

[13] X. Xu *et al.*, "Tracking Periodic Voltage Sags via Synchrophasor Data in a Geographically Bounded Service Territory," in *2023 IEEE PES Grid Edge Technologies*, San Diego, Apr. 2023.