Frequency Variation AI Monitoring and Prediction for Preventing Power Failure Using the Stochastic Model and Adaline

Amanda G. Neuenfeldt, M.S.E.E., EIT *amanda.neuenfeldt@students.tamuk.edu*
EECS Department, MSC 192
Texas A&M University-Kingsville
Kingsville, Texas, 78363-8202, U.S.A.

Rajab Challoo, Ph.D., P.E. *rajab.challoo@tamuk.edu*
EECS Department, MSC 192
Texas A&M University-Kingsville
Kingsville, Texas, 78363-8202, U.S.A

Abstract

Power reliability is becoming problematic nationwide. Power outages are becoming increasingly common for various reasons which include climate change, the increase in distributed energy resources (DERs), and old infrastructure. Climate change has promoted a societal push toward carbon neutrality. Energy demands and the correlated load requirements on the grid are increasing and are expected to rise. Additionally, consideration is needed for the predicted large-scale use of electric vehicles (EV), along with proliferation of data centers for things such as bitcoin and generative artificial intelligence (AI) types of industries with large computation requirements. Currently, data centers are contributors to the largest growth of power consumption. The capacity requirements to meet future predicted load demand along with anticipated variable energy resources with limited inertia necessitate a quicker method for handling the dynamic impacts to grid stability. Current infrastructure, processes and procedures are not capable of meeting future requirements and a new methodology is necessary. Reliable, stable, resilient forecasting is needed. State estimation is foundational for monitoring real time grid conditions. However, today there are not enough data points, and scalability is needed. AI is critical for monitoring and solving real time issues. There is a growing need for AI integration into the power grid, due to an increase in complexity, demand, and a reduction in overall grid inertia.

In this research work we use a proactive model versus the currently used reactive model in the research community and the available literature where changes are not made until after an event has occurred. Our proposed method utilizes Common Information Model (CIM) connectivity and integration to implement Univariant Linear Local Trend (ULLT) to produce predictive grid state values and Adaptive Linear Neural Network (Adaline) to provide an optimized control signal value. Generation source frequency sensor data is input to ULLT as a time sequence trend and predictive frequency values are generated for each generation source. The predictive frequency values are processed by Adaline to obtain an optimized control value. The system is tuned to utilize predictive future frequency values to correlate to the time the optimized control values signal is implemented.

Keywords: Univariant Linear Local Trent ULLT, Adaptive Linear Neural Network Adaline, Common Information Model CIM, Artificial Intelligence AI, Machine learning ML, Frequency Synchronization.

1. BACKGROUND AND LITERATURE REVIEW

1.1 Standards

The International Electro-technical Commission (IEC) developed and adopted IEC 61970-301, which defines the CIM and provides the guidelines for interfaces of Energy Management Systems

(EMSs) and Distribution Management Systems (DMS). IEC 61850 addresses control and protection and are useful for operations and monitoring. This CIM model is already being used in parts of Europe and is currently being implemented and projected to be operational in parts of North America through the AEP T-Nexus project. (B. Lee et al.,2015; G.A. Taylor et al., 2013; X. Cheng et al.,2017).

The T-Nexus project is an implementation of the IEC common information model and is compliant with IEC 61970-301. The CIM provides a model of how the entities, data, processes, and relationships are represented in the models of the grid. This uniform model will allow systems to share data that cannot be directly shared now. It will also centralize the information, eliminating conflicting data in different systems. The project's long-term goals include laying the foundation for artificial intelligence integration into the grid, resulting in the implementation of the IEC Smart Grid standard 61850, as well as the greater goals of providing the foundation to allow communication between internal AEP systems and external entities which is IEC standard 61968 (B. Lee et al.,2015; X. Cheng et al.,2017).

1.2 Related Work

Currently, utilities are already moving to a one-platform communications and modeling network, such as the T-Nexus project, which creates an integrated information environment for the grid. Within AEP, Transmission Planning PSSE, Protection and Control Engineering ASPEN and Transmission Operations EMS models are misaligned and have conflicting information. This increases O&M costs and disrupts grid reliability since multiple systems require subsets of the same data, interrupting synchronization between the systems. The CIM provides a model of how the entities, data, processes, and relationships are represented in the models of the grid. This uniform model will allow systems to share data that cannot be directly shared now. It will also centralize the information, eliminating conflicting data in different systems. The CIM is currently implemented as a Unified Modeling Language (UML) model. The CIM UML provides a visual model of the CIM standards' entities, processes, relationships, and data. Compliance with the CIM UML is requisite for the sharing and utilizing of the data. IEC 61970-301 is built off UML addressing electric power energy and provides standardized semantics and syntactic interoperability. This implementation of the CIM is also the first large-scale integration of these standards in North America.(AEP, 2017; G.A. Taylor et al., 2013, Siemens AG, 2017). Model credibility will be enhanced through automation. A CIM-based model management procedure is introduced to target recurrent EMS model update that is error-prohibiting and maintenance friendly. This will enhance smart grid interoperability by fulfilling recommendations from regulating authorities. It allows an effective alternative for cross-company information sharing, representing data management across different entities using a unified modeling language. (Siemens AG, 2017, AEP, 2017)Once data integration has been attained through projects like T-Nexus, the data can be processed in an ML model such as a State Space Model (SSM) with a stochastic component. The data can be used to look at the impacts of grid reliability and predict the results of corrective measures taken. Those measures could include corrective maintenance, future power generation and transmission planning, predictive failure analysis, efficiency analysis, and response to system outages and instabilities.

1.3 Research Objectives

This paper compares current manual methods related to frequency deviation correction and synchronization with a new approach using ML/AI methods. Grid frequency influences power grid stability and functionality as well as equipment maintenance. Low frequency results in lowered reactance and increased current flow, which may exceed current ratings on equipment. Large frequency differentials between generation units create harmonic distortion also resulting in damaged equipment, potentially leading to maintenance issues and shorter life spans on equipment which already have significant supply chain issues and large costs. More importantly, extended outage duration can inadvertently induce loss of life.

Additional challenges are introduced with variable energy sources which can change dynamically because of sudden changes in load or capacity. The frequency of photovoltaics can change rapidly, while the frequency of a gas turbine will not change quickly comparatively since there is more inertia. The use of DERs creates more dynamics since generation fluctuates more rapidly. DERs are often used as a supplemental generation source, however, this is shifting as society pushes for more renewable energy on the grid. This shift has already begun in Texas with 42 % of current generation sources being renewable. This often creates a need for additional generator source capacity such as storage or non-renewable sources such as coal to stabilize the frequency to prevent load shedding and power outages. Currently this is managed with renewable curtailment or gas. Renewable curtailment intentionally restricts output of renewable generation sources to prevent grid overloads, balance energy generations supply or due to transmission constraints. Grid balancing is necessary due to varying load requirements. As a protective measure, generation sources may be taken offline, or load shedding may take place resulting in power outages to ensure a stable grid frequency within acceptable limits is maintained. To reduce power outages the frequency can be predicted and managed during events to minimize power outages and protect equipment, reducing costs.

Batteries can alleviate some issues but may also create a loss of visibility on the system. Inertia correlates to system strength and stability. As system inertia decreases, the rate of change of frequency (ROCOF) increases. This means with low inertia systems such as renewable the rate of frequency change is faster, creating a need for a quicker response time to prevent frequency from dropping too low. Reactive models are not able to identify issues until after there is a problem. This results in time delays which can create catastrophic repercussions for the grid. By initiating a proactive grid control system, not only can delays be minimized by predictive models but as a byproduct can result in a reduction in equipment damage and outages, increase in efficiency and reduction in cost. Several items should be stated. Measurement data needs to look ahead. There is a need to access the whole power system inertia. Better data is needed to make optimized decisions.

1.4 Proposed Proactive Methods

This thesis proposes a proactive model versus the reactive model currently used. Current methods use a reactive model, where changes are not made until after an event has occurred. Frequency is regulated at the source, but the control room dictates what frequency to synchronize to. Currently, a go-low method is being used. There are frequency sensors at the generation output points that send information to SCADA. There is some variation in latency; different parts of the grid have different latency. The network latency can be measured for each generation station since there are different transmission methods such as fiber, LORA and satellite with various protocols such as ICCP. Without empirical data from the grid, actual latency can’t be determined and is not within the scope of this paper. This proactive method replaces the control rooms’ decisions with an automated system. The load is not uniform across the system. The load is non-linear and there is a differential that needs to be controlled. An automated control room will use AI to provide proactive optimized predictive control in place of the manual control room process.

The proposed method utilizes CIM connectivity and integration to implement ULLT to produce predictive grid state values and Adaline to provide an optimized control signal value.

2. MACHINE LEARNING AND AI TOOL

2.1 Univariant Linear Local Trend

The ULLT uses State Space Model for Maximum Likelihood Estimation (MLE) with Kalman filtering. The machine learning ULLT algorithm also incorporates Gaussian noise as a stochastic component. The following equations are from reference (Commandeur et al., 2007; Durbin et al. 2012).

A general state space model is given by Eq. 1.

$$\begin{aligned} y_t &= Z_t\alpha_t + d_t + \varepsilon_t \\ \alpha_{t+1} &= T_t\alpha_t + c_t + R_t\eta_t \end{aligned} \qquad \text{Equation (1)}$$

The first equation y_t denotes the observation vector at time t, α_t denotes the (unobserved) state vector at time t. The Z_t, T_t, R_t, d_t, c_t, and Q_t describe the process design, transition, selection, obs_intercept, state_intercept, state_cov, respectively and the irregular components are given by:

$$\varepsilon_t \sim N(0, H_t)$$

$$\eta_t \sim N(0, Q_t)$$

A more efficient method than the current frequency control methods will be required and will incorporate predictions of the frequency changes due to the increased renewable energy sources (RES) that have low inertia. The stochastic model will provide a way for the system to predict changes in frequency and adjust ahead of time to prevent outages. These predictions can be utilized by the artificial intelligence tool to determine an optimized frequency synchronization control value. The evolution of AI integration into the grid is well predicted and further research is needed. (U.S. Department of Energy, 2024).

Local Linear Trend model equation (trend component) given by reference (Durbin et al. 2012).

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\ \mu_{t+1} &= \mu_t + \nu_t + \xi_t & \xi_t &\sim N\left(0, \sigma_\xi^2\right) \\ \nu_{t+1} &= \nu_t + \zeta_t & \zeta_t &\sim N\left(0, \sigma_\zeta^2\right) \end{aligned}$$

The term y_t (observations) is univariate. The slope term is v_{t}, where ε_{t} and ξ_{t} are serially independent (uncorrelated) random variables with zero means and constant variances. The trend is linear if $\xi_{\mathrm{t}} = \zeta_{\mathrm{t}} = 0$ then $v_t + 1 = v_t = v$, and $\mu_t + 1 = \mu_t + v$. This reduces to the deterministic linear trend plus noise model. The trend level and slope vary over time if $\sigma^2\xi > 0$ and $\sigma^2\zeta > 0$.

The univariate structural time series model may also be written in state space form.

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \nu_t \end{pmatrix} + \varepsilon_t$$

$$\begin{pmatrix} \mu_{t+1} \\ \nu_{t+1} \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \mu_t \\ \nu_t \end{pmatrix} + \begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix}$$

State vector $(\mu_t \nu_t)'$

State error vector $(\xi_t \zeta_t)'$

For the local linear trend model, see the observation equation.

$\mathrm{y_t} = \mu_t + \gamma_t + \varepsilon\mathrm{t}, \mathrm{t} = 1, \ldots, \mathrm{n}.$

To represent in state space form the state vector is

$$\alpha_t = \begin{pmatrix} \mu_t & \nu_t & \gamma_t & \gamma_{t-1} & \cdots & \gamma_{t-s+2} \end{pmatrix}'$$

and the system matrices are

$$Z_t = (Z_{[\mu]}, Z_{[\gamma]}), \qquad T_t = \operatorname{diag}(T_{[\mu]}, T_{[\gamma]}),$$
$$R_t = \operatorname{diag}(R_{[\mu]}, R_{[\gamma]}), \quad Q_t = \operatorname{diag}(Q_{[\mu]}, Q_{[\gamma]}),$$

Where

$$Z_{[\mu]} = (1,0), \qquad Z_{[\gamma]} = (1,0,\dots,0),$$

$$T_{[\mu]} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad T_{[\gamma]} = \begin{bmatrix} -1 & -1 & \cdots & -1 & -1 \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & & 1 & 0 \end{bmatrix}$$

$$R_{[\mu]} = I_2, \qquad R_{[\gamma]} = (1,0,\dots,0)',$$

$$Q_{[\mu]} = \begin{bmatrix} \sigma_\xi^2 & 0 \\ 0 & \sigma_\zeta^2 \end{bmatrix}, \qquad Q_{[\gamma]} = \sigma_\omega^2.$$

Parameters are estimated in the variance and covariance matrices.

$$H_t = [\sigma_\varepsilon^2]$$
$$Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 \\ 0 & \sigma_\zeta^2 \end{bmatrix}$$

Observation disturbance covariance matrix H_t

Disturbance covariance matrix Q_t

Variance parameters ($\sigma^2\varepsilon\sigma^2\xi\sigma^2\zeta$)

In its simplest form, the state space model represents a mathematics model of a physical system. In the context of machine leaning the SSMs represent a dynamic system. This makes the SSM model useful in time series analysis such as the ULLT model. This is a custom model. The Kalman filter adds functionality for maximum likelihood estimation.

In addition, the stochastic component Gaussian distributed noise is added to the dataset to help improve the performance of the machine learning model. This is necessary because models can learn to recognize and filter out the noise, making them more resilient to new, unseen data. Gaussian noise follows a normal distribution centered around the mean. In this scenario it is used for adding randomness. This is applicable due to the varying generation sources requirements on the grid caused by unpredictable climate changes, outages and so forth, as well as the fluctuations in load.

Normal distribution or Gaussian distribution is given by Eq. 2.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \text{Equation (2)}$$

Where, x is the variable, μ is the mean, and σ is the standard deviation.

The maximum likelihood estimation is a method of estimating the parameters of an assumed probability distribution, given some observed data, such as nominal frequency.

Series forecasting by ULLT generates a forecast for 10 sequential forecast values. This window can be altered. This provides 10 predicted frequencies. The last frequency is populated into the forecasted frequencies table for each event magnitude to be used by ADALINE to determine the optimal frequency for the generator sources. The power computation and samples needed for accuracy are not addressed in this paper.

2.2 Artificial Intelligence

According to (A.J. Abjouganrair, 2023) ADALINE is one of the AI neural networks (NN) fundamental models for data prediction. Some drawbacks include slow convergence, which is why this method has not been used much to forecast the dynamics of nonlinear systems. This slow convergence for ADALINE is well known (W. Zhang, 2007).The slow convergence mostly affects multivariate applications of ADALINE. The univariate implementation used in this process is minimal.

In (A.J. Abjouganrair, 2023), ADALINE was used to adapt control actions to address the disturbance and parameter variation issues for a cart pole system (CPS). The ADALINE controller was capable of canceling out the effect of any disturbance and accurately predicting desired values, indicating its effectiveness in controlling the system. In addition, the ADALINE controller quickly corrected the nonlinear plant to the new targets when the targets were adjusted. Simulation results proved the system could handle uncertainties by introducing variances when ADALINE is applied to the Nonlinear CPS model(A.J. Abjouganrair, 2023).In (G.S Chawda et. Al, 2022) ADALINE-LMS is characterized by adaptiveness, low computational burden and fast response. This is a popular training scheme for ADALINE which can be used online to significantly reduce the computation and storage requirements (M. Qasim, 2014). Also, GADALLINE work (W. Zhang, 2007) shows faster convergence speed, better tracking of time varying parameters and low computational complexity.

Adaline is a neural network that uses a supervised machine learning algorithm. A Stochastic Gradient Descent (SGD) Regressor is being applied. The models' parameters are updated after each training sample. This allows more frequent updates and efficient adaptation for large datasets.

Adaline is being used as a predictive optimization tool to determine the frequency for generation sources to synchronize with minimum frequency differentials to maintain grid balance. Additional considerations are being taken to calculate iterations to step frequency up or down to 60.0 Hz standard.

We are using the Scikit-Learn library for our implementation of Adaline (Scikit-learn, 2024)ADALINE is a linear regression model that uses supervised learning and is a type of machine-learning algorithm used in prediction or forecasting. It maps the data points to the most optimized linear functions. The regression model estimates the linear relationship between a dependent and one or more independent variables. The SGD is an optimization technique (Pedregosa et al, 2011; Scikit-learn, 2024).This implementation works with floating point values. The linear model is fitted by minimizing the normalized empirical loss. The predict function enables us to predict the optimal frequency to synchonize the varying inertial generating units to based on the trained model. This trained model will become more effective through each iteration providing a more optimized training data set (Pedregosa et al, Python Software Foundation, 2024).

Post estimation prediction derived from learning models, in this case raw frequency values, and forecasting output from ULLT allows the Adaline model to optimize the frequencies with a single value based on the future state. The output from ULLT is simply the input for Adaline. Adaline uses the last forecasted value in the forecasted trend for each generation source and the related inertial value for that specific generation source and Adaline outputs a single frequency value rather than the control room manual go-low method. Post estimation prediction uses the differential errors from the variance and covariance matrix and creates a forecasted output based on these matrices and the addition of Gaussian noise to provide a stochastic forecasted output to be read to Adaline's input. See Table 1.

The SGD Regressor algorithm was used in Eq. 3.

$$E(w,b) = \frac{1}{n}\sum_{i=1}^{n} L\big(y_i, f(x_i)\big) + \alpha R(w) \qquad \text{Equation (3)}$$

Training examples represent $x_1, y_1, \ldots, x_n, y_n$ where $\mathrm{x_i} \in \mathrm{R_m}$ and $y_i \in \mathrm{R}$. The goal is to learn a linear scoring function $\mathrm{f(x) = w^T x + b}$ with model parameters $\mathrm{w} \in \mathrm{R_m}$ and intercept $b \in R$. R is a regularization term or penalty where $\alpha > 0$ controls regularization strength. ' L ' is a loss function (Scikit-learn, 2024)

3. GRID SIMULATION

3.1 Preliminaries

Multiple tools and components were used for the simulation. Python version 3.12.2 64-bit is used as the primary programming language. The MariaDB Relational Database Management System (RDBMS) version 11.3.2 64-bit is used as a MySQL open source alternative to achieve software compatibility and is the database management system used to simulate the CIM. CIM Tool version 1.12.0 is used to determine schema and naming to make CIM-compliant representation of data. This is an analogous representation of a CIM implementation, e.g. the column names in the MariaDB include Generating Unit. This is not a complete representation of a CIM-compliant database. The goal of this thesis is to build pieces of a CIM-compliant database to determine how the system responds with respect to the predicted grid frequencies using the ULLT and ADALINE algorithms with Gaussian noise as a stochastic component, for optimizing grid source frequencies and determine the next steps for implementation. Gaussian noise was scaled to provide consistent reliable results for the ULLT output for the range of the initial frequency trend generated for the simulation. Univariate Local Linear Trend ULLT model is a machine learning model using the Python Stats model library. This time series model is used for capturing trends. Several statistical calculations are considered including the maximum likelihood estimation and the Kalman filter.

The generating capacity percentages were based on (Electric Reliability Council of Texas, 2024). Using 100 generator sources: 44 are natural gas, 25 are wind, 10 are coal, 13 are solar, 4 are nuclear, 3 are storage and 1 is hydro. See Table 2.The number of each type of generation source is representative by percentage of the current Texas grid. Inertia varies based on the generation source as well as other factors. The average mean inertia is provided for each generation source for the purpose of the simulation. The mean inertia in seconds [s] for each of the generation sources is CIM model based and the specific inertial constants for each type of generation source were determined (D. Kraljic et al., 2022; Electric Reliability Council of Texas, 2024; National Renewable Energy Laboratory, 2020). See Table 1.

This produces test data that is reasonable for the types of generation sources in a ratio representative of the real grid. The machine learning model ULLT model observes 100 reasonable frequency sensor values from a single source at consistent time intervals. which are in a table in a MySQL database. The ULLT model does not read inertia values. It only reads the 100 sequential frequencies for each generator source based on the event magnitude in our simulation.

The ULLT model estimates the parameters from a local linear trend model. Based on the time-varying trend, ULLT predicts the values or frequencies for a specified time. The model produces a time sequence of predicted values. The closer the prediction is to the last observed value the more accurate the predicted value. However, sufficient time is needed to calculate the correct control signal value and deliver it to the generation sources. The selected time sequence prediction should correlate with the time required to deliver the control signal. The tenth predicted value was arbitrarily selected to provide for a reasonable time of 10 seconds to calculate and deliver the control signal.

The only way to effectively calculate this in the real grid would be through empirical data collection, as each grid or micro grid will have varying values. The consideration for which time sequence estimated value to use is outside the scope of this paper and would be contextual to each application. Consideration needs to be taken when setting this parameter, which will not be a part of this simulation. Predictions become less reliable the further out into the future the predicted values are. However, the predicted value needs to be sufficiently in the future in time to correlate to the time that an optimized control value would be received at the generation source. The predicted values are generated in a time sequence by the ULLT model, and the values of insufficient future time allowance will be discarded. This model could be adjusted to fit each implementation context by selecting a different time sequence prediction point. The effectiveness of this approach would be affected by the time needed to calculate the control signal value and the latency of the control signal value.

Adaline has access to inertial values for all generator sources for each event. This (unit inertia) table is static. See Table 1.The inertia values are taken into consideration when determining how much of a frequency change each source can move in a single step. The table can be interpreted as follows, for example, in 5.9 Hz a nuclear generating unit on average shift 1 Hz in 5.9 seconds. With this understanding, a variable generation source such as solar can shift 1 Hz in effectively zero seconds. This is a highly dynamic low inertial source.

Calculation for each generation source is required for the system to reach an optimal frequency. Sources with higher inertia values will not be able to adjust their frequency as quickly as low or non-inertia sources. This simulates our current grid as more renewable and variable generator sources are being integrated into the grid. This will help reduce large frequency differentials and reduce total harmonic distortion, prevent power outages due to large differentials and reduce the need for other stabilizing methods. This is a proactive solution. Several tables are considered in the database. ULLT outputs a certain number of future values. For the purposes of this simulation, we are using a tenth value which is equivalent to 10 seconds in the future prediction. The last value is populated into another table to be read as the input for Adaline. Adaline reads the last predicted value as well as the inertia value related to each generation source which is already stored in the database. The training set from ULLT uses reasonable values based on inertia but not the inertia values themselves. This simulates the unknown inertia of our current grid (D. Kraljic et al., 2022) The target frequency is determined based on this training set and preset reasonable inertia values to simulate the grid.

The Adaline model is supervised. It evaluates a prediction on a given data set and returns a single numerical score or frequency. In this case, Adaline will determine a single target frequency for all the generator sources to synchronize to, based on their 'weighted' inertia values. The output predicts frequencies for all generator sources and determines the optimal frequency the generators should adjust to with minimal frequency deviation considerations. This will be compared to current go-low methods. The purpose of this demonstration is to determine if AI/ML prediction and optimization methods are a viable solution for frequency stability to maintain grid reliability with hybrid generations sources by comparing it with the current go-low method.

3.2 Demonstration

The Machine learning engine reads from the MySQL ‘initial trends’ table for each of the four event magnitudes: small, medium large and serious, which is populated with 100 generation sources that resembles the Texas grid. This correlation was maintained for the generation sources within the scope of this paper related to their inertial values provided in Table 1.

For this demonstration the output of ULLT (a single predicted frequency for all generator sources - in this case 10 seconds out) is input into the Adaline model. These 100 'future' frequencies from 100 varying generator sources with weighted inertia are input into Adaline. This is an optimization technique used to train a model. This is being used to predict new values. Adaline determines the optimal frequency for each of the generator sources target value based on their inertia. See Table 1 and Table 2. This is run four times for different event magnitudes. The event magnitudes were empirically tested. For simulation purposes none of the generation units went offline. The maximum deviation event simulated was described as a serious event. A serious event resulted in the generation units with the lowest inertia falling to 59.5 Hz. For any given event the lowest inertia generation sources would have the greatest frequency deviation. A value of 59.5 Hz deviation would be the maximum permissible deviation from 60 Hz standard while remaining within standard ranges for frequency deviation. For 100 samples per second the generation units with the lowest inertia which were set at 0.01 would need to be decremented 0.005 for an event magnitude to decrement from 60 Hz to 59.5 Hz in 100 samples. The large event magnitude was 0.00375 and represented 75% of the serious event magnitude. The medium event magnitude represented 50% of the serious event magnitude at 0.0025. The small event magnitude represented 25% of the serious event magnitude at 0.00125. This easily provided for four event magnitudes evenly spaced. These event magnitudes evenly spaced were chosen to evaluate the efficiency of the proposed methods across a reasonable range of magnitudes.

4. PROCESS OVERVIEW

4.1 Process Overview

This solution operates as a closed feedback loop system. Fig 6.illustrates the simulation process. Initial SCADA data acquisition of sensor data is necessary. The sensor data is produced via equipment in the field. For the purpose of the simulation the MySQL database is used as an analog of the CIM. The sensor data is recorded in the CIM. The machine learning algorithm ULLT will read sensor data from the CIM and produce predictive states for each generation source. The ULLT algorithm writes the prediction to the CIM. The artificial intelligence algorithm Adaline reads the predictive states from the CIM and produces the optimized control value. The artificial intelligence algorithm writes the optimized control value to the CIM which is read by SCADA. The SCADA system is mathematically simulated in the simulation Python code. CIM compliant information is necessary for all systems to operate cohesively on the grid. SCADA uses the optimized control value to implement the most efficient frequency synchronization value. This is the proactive implementation of grid control allowing a quick response for dynamic low inertia or hybrid generation sources impacts on the grid systems.

The following process is for proof of concept. Maintaining stable 60 Hz frequency is a fundamental requirement for grid operations.

4.2 Results

The scale to create consistency across varying generation sources were empirically derived. The max event amplitude is defined as, E/0.01 Hz/second = 0.5 Hz.

The ULLT needs a minimum of 100 samples as a time sequence of frequency values to produce the predictive values required. Since this is representative of 100 seconds for each generation source, a max differential from beginning to end of sequence is 0.5 Hz. The 0.5 Hz represents the maximum frequency drop before the generation is taken offline. Therefore, this represents the largest event that this simulation would be able to provide control signals for.

Where E is the event magnitude 0.005, and I is the inertial value for the generation source. To ensure consistency across the generation sources with varying inertia in range of 0.01 – 5.9 Hz/second, the max event amplitude was calculated and empirically tested. The max event amplitude was arbitrarily calculated using the lowest inertial value for wind, photovoltaic and hydro generation sources, which was effectively zero. The zero value could not be used for mathematical calculations since dividing by zero would effectively be dividing by infinity and create unusable results. The significant digits were maintained. This constrained the results to useable data to prevent all low inertia generations sources from cascading. Cascading loads is feasible, especially for serious events as experienced in the Texas Winter Storm of 2022, however, is not within the scope of this paper.

Subsequently, a decrement value was needed to offset the 100 nominal frequency values to be used for the initial table fed into ULLT.

0.005/0.1 = 0.5 /100 = 0.005 Hz /second decrement steps for max event with a minimal inertia.

The maximum inertia 5.9 Hz/second represents the nuclear generation units.

0.0005/5.9 = 0.0000847/100 Hz /second = 0.000000847 decrement steps for minimum event with maximum inertia.

To ensure model accuracy and true stochastic grid nature and efficacy, several tests were performed. Grid forming and grid following methods are evaluated and compared to the ML/AI methods. Both methods have some latency. The current approach sets all generation sources low to reduce differential and then walks the sources to 60 Hz. There are several reasons to improve this. Smaller inertia systems move too quickly for the control room to respond to, and the grid is unpredictable. The traditional approach is challenging for highly dynamic systems. Currently the dynamism of the grid is increasing at a significant rate. The time it takes each method to convergence is evaluated by calculating the time differentials. Comparison is accomplished by counting the number of cycles until convergence for both methods.

Four tests were run, each representing four different event magnitudes. These magnitudes include small, medium, large, and serious. These event magnitudes are represented mathematically above. For illustrative purposes, the Texas Winter Storm 2021 would be categorized as serious.

The algorithm will simulate and determine the time to reach frequency synchronization of all generation sources. Steering is introduced mathematically. The simulation produced 812 files. The State Space Model Results were produced for each generating unit for each magnitude producing four-hundred images. See Fig3. The observation, one step-ahead prediction and forecast was also produced for each generation unit for each magnitude resulting in four-hundred additional images. See Fig4. and Fig5. In addition, the optimum control frequency for the generation units to synchronize to is provided for each four event magnitudes. Lastly, there were four text files produced which provide the efficiency gain percent for each event magnitude when comparing the current go-low method to AI/ML optimization.

The algorithm was the most efficient for serious event magnitudes, producing roughly 45.76 percent more efficiency than current go-low methods. See Table 3. The large event magnitude shows an average of 42.78 percent greater efficiency. The medium event magnitude shows an average 36.82 percent greater efficiency. The small event magnitude shows an average of 18.9 percent greater efficiency than current go-low methods. These percentages were determined with the following process and subsequent equation. See Table 3.

The go-low value was determined by taking the minimum value from the last row of the 100 generation sources frequency values. This go-low differential was subtracted from the optimal 60.0 Hz value. The optimal differential determined by the ADALINE algorithm was also subtracted by the optimal 60.0 Hz value.

golowdiff = 60.0 – golowvalue

optimaldiff = 60.0 – optimalvalue

To compare both methods the optimal differential was divided by the go-low differential.

efficiencygainpercent = 100 - (efficiencygainraw * 100).

This provides a reasonable value for simulating in a static system. This value would not be static in the real world. The grid operates as a multivariate stochastic system. The Gaussian noise was adjusted empirically.

This will allow frequency to not only optimize to a single value but walk back to 60 Hz. in the most optimized path. Through each iteration of ADALINE this will become more efficient and therefore more reliable. Since the algorithm determines the optimal frequency to synchronize to for minimum frequency deviation, the frequency will not typically synchronize at 60 Hz due to inertia, load demand and other factors. See Fig1. and Fig2.

The RDBMS, MySQL simulates the CIM in the data base using the SQL language. IEC 61970 CIM schema utilizing, CIM Tool to simulate our emerging grid which will allow all components to meet and communicate in the same language and semantics has been integrated to mimic the future grid implementations currently in process.

The purpose of this thesis is to determine the most reliable method for grid stability in relation to frequency variation. The real-world grid would be dynamic and require empirical testing for the optimized timing parameter. To be fully implemented in the most effective way would require a multivariant versus a univariant approach. However, each component must first be approached and dealt with in a univariant manner. Validation of each univariant component would be required before integrating multivariate components.

While the current go-low method produces similar optimization for small events, climate change and other fore mentioned shifts such as EV charging will increase the likelihood of bigger events which may necessitate AI/ML in the future grid. A smart grid may be the most efficient method for handling the stochastic nature of the grid in the future.

5. CONCLUSION AND FUTURE WORK

5.1 Conclusions

The grid is a stochastic system as a result of fluctuating load as well as temperature and other factors. Introducing more instability to the system creates more dynamism. One result is more frequent large and serious events. These events can be difficult to manage with our current go-low methods and can benefit from a smart grid.

The CIM will allow different Regional Transmission Organizations (RTOs) to communicate with the same data semantics and nomenclature and share data. CIM translates it regardless of the semantics and nomenclature of individual systems. All systems will be capable of utilizing the same ontology even if the systems' native ontology is different. T-Nexus is a centralized system for all systems that all RTOs can connect to it. Then, RTOs such as ERCOT can connect to it, and all will adopt all CIM nomenclature and topology. T-NEXUS matches all the systems such as SCADA, EMS, and Planning. T-Nexus is the authoritative data for all their systems. Everything will match T-Nexus. It will translate, for example, from EMS to SCADA. A change in SCADA will be read to EMS, and so forth, reducing staff and increasing profit with more reliability since so many people won't need to make the translations but reliability from data inconsistencies will be resolved. Data normalization and a shared ontology are required for full automation.

With a focus on grid resiliency due to extreme weather events such as wildfires and flooding, load growth, and affordability, AI and Automation are tools, but sufficient digital data and global data that AI can process is needed. Another huge concern is allowing low inertia, highly dynamic generation sources connection to distribution without understanding the impact to the grid. Software needs to aggregate data to be used for AI.

5.2 Future Grid

Multiple opportunities for tuning are available. Additional consideration can be incorporated for future work. Simulated states can be configured by creating excessive demand and activating load shedding and/or taking generators spontaneously offline to simulate RES limitations and real-world applications. This will allow methods to mimic changes to the grid and will include both random and selected generators to simulate various environments. During the 2021 Texas Winter Storm the frequency fell to 59.302 Hz. According to (M. Lozano, 2021), if the grid had dropped below the 59.4 Hz threshold for more than 9 minutes it would have triggered a cascading failure of the grid.

6. FIGURES AND TABLES

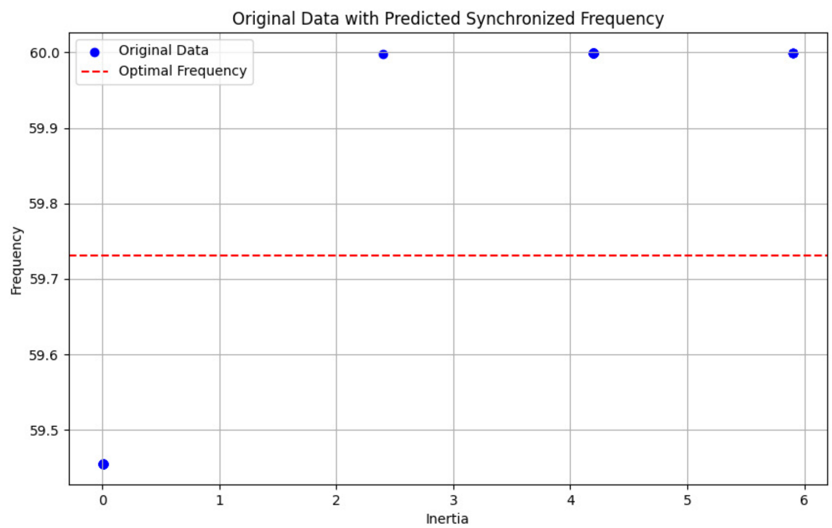

FIGURE 1: Optimum Control Frequency_serious. The optimal frequency for a Serous event magnitude is approximately 59.732. The optimal frequency float is written to a text file. The grid is a stochastic system because of fluctuating load as well as temperature and other factors. Introducing more instability or low inertia generation sources to the system creates more dynamism.

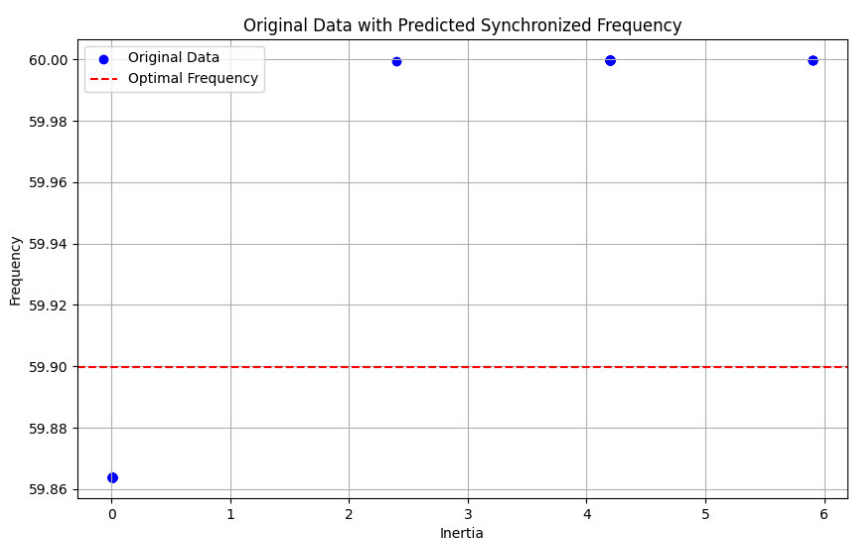

FIGURE 2: Optimum Control Frequency_small. The optimal frequency for a small event magnitude is approximately 59.900. The optimal frequency float is written to a text file. Frequency is better maintained for smaller events.

```
                           Unobserved Components Results
==============================================================================
Dep. Variable:      GeneratingUnit100   No. Observations:                  100
Model:             local linear trend   Log Likelihood                1250.967
Date:                Mon, 14 Oct 2024   AIC                          -2495.935
Time:                        11:54:37   BIC                          -2488.180
Sample:                             0   HQIC                         -2492.798
                                - 100
Covariance Type:                  opg
====================================================================================
                       coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------------
sigma2.irregular  2.667e-26   1.73e-06   1.54e-20      1.000    -3.4e-06     3.4e-06
sigma2.level      1.105e-23   3.94e-05    2.8e-19      1.000   -7.72e-05    7.72e-05
sigma2.trend          1e-12   3.88e-05   2.58e-08      1.000   -7.61e-05    7.61e-05
===================================================================================
Ljung-Box (L1) (Q):                  36.37   Jarque-Bera (JB):                10.45
Prob(Q):                              0.00   Prob(JB):                         0.01
Heteroskedasticity (H):               1.04   Skew:                             0.00
Prob(H) (two-sided):                  0.90   Kurtosis:                         1.40
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

FIGURE 3: GeneratingUnit100_serious_StateSpaceModelResults (Wind) 0.01Hz/second. This is the output of the state space model. The unobserved components are a univariate time series decomposed into a trend. It's a structural time series model. This is the output of the state space model. The unobserved components are a univariate time series decomposed into a trend. It's a structural time series model. The Unobserved Components Results include standard errors, z-statistics and prediction or forecasting. The unobserved components results are provided for each generating unit for each event magnitude. This is an example of the GeneratingUnit100 for a serious event magnitude.

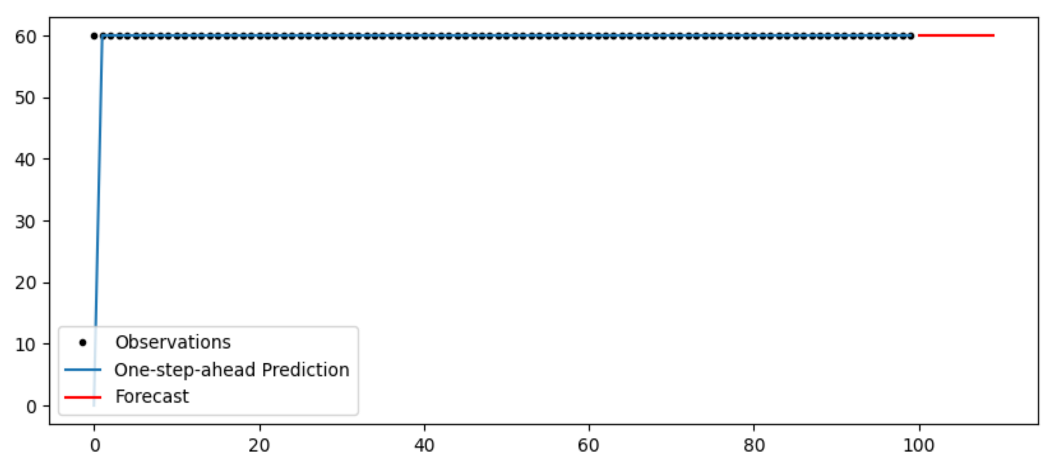

FIGURE 4: GeneratingUnit1_serious (Nuclear) 5.9 Hz/second. The 100 observations, one-step-ahead prediction and forecast are available for each frequency value for each event magnitude. Referencing the SQL table forecast_output_serious the 10th prediction value for the nuclear generation source is 59.99919. The high inertia generation source is maintaining a reasonable frequency for the serious event.

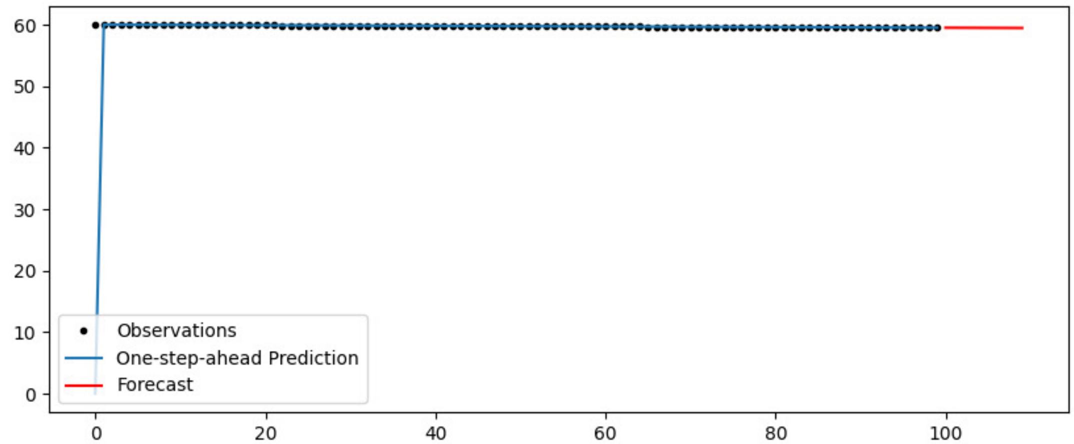

FIGURE 5: GeneratingUnit60_serious (Solar) 0.01 Hz/second.The 100 observations, one-step-ahead prediction and forecast are available for each frequency value for each event magnitude. Referencing the SQL table forecast_output_serious the 10th prediction value for the solar photovoltaic generation source is 59.455. The low inertia generation source is continuing to decrement for the serious event.

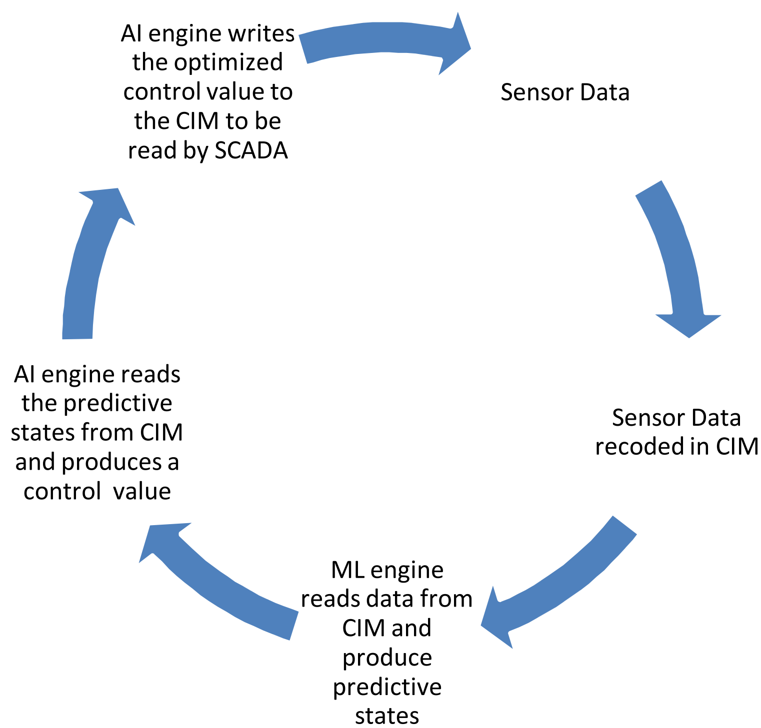

FIGURE 6: Smart Grid Cycle M/L and AI tools would require integration with multiple grid systems to implement automated control solutions. CIM implementation provides the shared data repositories and communication pathways.

Generating Unit Type	Mean Inertia
Nuclear	5.9
Natural Gas Turbine	4.2
Coal Turbine	4.2
Hydroelctric	2.4
Solar Photovoltaic	Effectively Zero (Implemented Experimentally as 0.01)
Wind Turbine	Effectively Zero (Implemented Experimentally as 0.01)
Large Scale Storage	Effectively Zero (Implemented Experimentally as 0.01)

TABLE 1: Generating Unit Types and Mean Inertia [16, 21, 23]. The generation sources are representative of the Texas grid. Nuclear has the highest inertia at approximately 5.9 Hz/second followed by Natural Gas and Coal Turbines at 4.2 Hz/seconds. Hydroelectric has minimal inertia averaging 2.4 Hz/second. The Solar Photovoltaic, Wind Turbine, and Large-Scale storage is effectively zero. A 0.01 value was used for mathematical calculation as 0.00 can interfere with computer mathematics.

Generating Unit Type	Number Of Units Proportionate to Texas Grid
Nuclear	4
Natural Gas Turbine	44
Coal Turbine	10
Hydroelctric	1
Solar Photovoltaic	13
Wind Turbine	25
Large Scale Storage	3

TABLE 2: Generating Unit Counts by Type. The generation sources are representative of the Texas grid. Currently roughly 42% of the Texas grid is utilizing renewable energy sources. (RES). Traditional energy sources include Nuclear, Natural Gas Turbine and Coal Turbine. The RES include Hydroelectric, Solar Photovoltaic, Wind Turbine, and Large Scale Storage.

Event Magnitude	Small	Medium	Large	Serious
Efficiency Gain Percent	18.85	36.82	42.78	45.76

TABLE 3: efficiency_gain_percent AI/ML and go-low comparisons. This table represents the percentage of efficiency gain of the AI/ML method over the go-low method. The more serious events have the greater efficiency gain percent. The AI/ML method will be most efficacious for serious event magnitudes. The small event magnitude is 18.85%. efficiency. The medium event magnitude of 36.82%. The large event magnitude is 42.78% and the serious event magnitude is 45.76%. efficiency.

7. REFERENCES

AEP (2017), "AEP T Nexus – Benefit Summary / White Paper," pp. 1-2, 2017. [Online] Available: AEP Transmssion T Nexus Benefit Summary_White Paper_Final.pdf (ucaiug.org).

A. J. Abougarair (2023), "Adaptive Neural Networks Based Optimal Control for Stabilizing Nonlinear System," *2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)*, Benghazi, Libya, 2023, pp. 141-148, doi: 10.1109/MI-STA57575.2023.10169340.

B. Lee, D. -K. Kim, H. Yang, H. Jang, D. Hong and H. Falk (2015), "Unifying Data Types ofIEC 61850 and CIM," in *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 448-456, Jan. 2015, doi: 10.1109/TPWRS.2014.2326057.

Commandeur, Jacques J. F., and Siem Jan Koopmann (2007). An Introduction to State Space Time Series Analysis. Oxford; New York: Oxford University Press.

D. Kraljic, B. Sobocan, J. Katanec, M. Logar and M. Troha (2022), "Inertia constants for individual power plants," 2022 18th International Conference on the European Energy Market (EEM), Ljubljana, Slovenia, 2022, pp. 1-5, doi: 10.1109/EEM54602.2022.9921019.

Durbin, James, and Siem Jan Koopman (2012). Time Series Analysis by State Space Methods: Second Edition. Oxford University Press.

Electric Reliability Council of Texas (2024), “ERCOT Fact Sheet,” TX, April 2024. [Online]. Available: ERCOT_Fact_Sheet.pdf.

ENTSO-E (2024), "Common Information Model (CIM) - Interoperability Tests and Roadmap for Implementation of CIM/XML Data Exchange Format for System Operations and System Studies Exchanges," ENTSO-E, [Online]. Available: https://www.entsoe.eu/data/cim/. [Accessed: Nov. 4, 2024].

European Network of Transmission Systems Operators for Electricity (2020), *“Inertia and Rate of Change of Frequency (RoCoF,)”* Version 17, December 16, 2020. [Online] Available: Microsoft Word - Inertia and RoCoF_v17_clean (azureedge.net).

G. A. Taylor, N. Hargreaves, P. Ashton, M. E. Bradley, A. Carter and A. McMorran (2013), "Potential integration of Phasor Measurement Units and Wide Area Monitoring Systems based upon National Grid enterprise level CIM," *2013 IEEE Power & Energy Society General Meeting*, Vancouver, BC, Canada, 2013, pp. 1-1, doi: 10.1109/PESMG.2013.6672342.

G. S. Chawda and A. G. Shaik (2022), "Enhancement of Wind Energy Penetration Levels in Rural Grid Using ADALINE-LMS Controlled Distribution Static Compensator," in *IEEE Transactions on Sustainable Energy*, vol. 13, no. 1, pp. 135-145, Jan. 2022, doi: 10.1109/TSTE.2021.3105423.

Harris, C.R., Millman, K.J., van der Walt, S.J. et al. (2020) *Array programming with NumPy*. Nature 585, 357–362 (2020). DOI: 10.1038/s41586-020-2649-2. (Publisher link).

L. R. Gorjão *et al*. (2020), "Data-Driven Model of the Power-Grid Frequency Dynamics," in *IEEE Access*, vol. 8, pp. 43082-43097, 2020, doi: 10.1109/ACCESS.2020.2967834.

M. Lozano (2021), "Texas power grid was 4 minutes and 37 seconds away from collapsing. Here’s how it happened," Texas Standard, Feb. 2021. [Online]. Available: https://www.texasstandard.org/stories/texas-power-grid-was-4-minutes-and-37-seconds-away-from-collapsing-heres-how-it-happened/. [Accessed: 14-Oct-2024].

M. Qasim, P. Kanjiya and V. Khadkikar (2014), "Optimal Current Harmonic Extractor Based on Unified ADALINEs for Shunt Active Power Filters," in *IEEE Transactions on Power Electronics*, vol. 29, no. 12, pp. 6383-6393, Dec. 2014, doi: 10.1109/TPEL.2014.2302539.

National Renewable Energy Laboratory, University of Colorado Boulder (2020), “Inertia and the Power Grid: A Guide Without the Spin,” BCO, May 2020. [Online] Available: 73856.pdf (nrel.gov).

Pedregosa *et al.* (2011), JMLR 12, pp. 2825-2830, 2011. [Online] Available: Scikit-learn: Machine Learning in Python.

Python Software Foundation (2024). Python Language Reference, version 3.3. Available at http://www.python.org.

Siemens AG (2017), "AEP Transmission breaks new ground in network model management," Siemens, Erlangen, Germany, 2017. [Online]. Available: https://assets.new.siemens.com/siemens/assets/api/uuid:109359ad576c0533f3ba52560b18a001177b592d/aep-casestudy-intl-version.pdf. [Accessed: 14-Oct-2024].

"sklearn.linear_model.LinearRegression.predict," Scikit-learn (2024). [Online]. Available: https://scikitlearn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html#sklearn.linear_model.LinearRegression.predict. [Accessed: 14-Oct-2024].

"Stochastic Gradient Descent: Mathematical Formulation," Scikit-learn (2024). [Online]. Available: https://scikit-learn.org/stable/modules/sgd.html#sgd-mathematical-formulation. [Accessed: 14-Oct-2024].

"Stochastic Gradient Descent," Scikit-learn (2024). [Online]. Available: https://scikit-learn.org/stable/modules/sgd.html#sgd. [Accessed: 14-Oct-2024].

U.S. Department of Energy (2024), "AI for Energy: Opportunities for a Modern Grid and Clean Energy Economy," April 2024. Available: https://www.energy.gov/sites/default/files/2024-04/AI%20EO%20Report%20Section%205.2g%28i%29_043024.pdf. [Accessed: Nov. 4, 2024].

W. Zhang (2007), “A generalized ADALINE neural network for system identification,” IEEE International Conference on Control and Automation, Guangzhou, China, pp. 2705-2709, 2007.

X. Cheng, W. -J. Lee and X. Pan (2017), "Modernizing Substation Automation Systems: Adopting IEC Standard 61850 for Modeling and Communication," in *IEEE Industry Applications Magazine*, vol. 23, no. 1, pp. 42-49, Jan.-Feb. 2017, doi: 10.1109/MIAS.2016.2600732.

Y. Guo and T. H. Summers (2019), "A Performance and Stability Analysis of Low-inertia Power Grids with Stochastic System Inertia," *2019 American Control Conference (ACC)*, Philadelphia, PA, USA, 2019, pp. 1965-1970, doi: 10.23919/ACC.2019.8814402.

Y. Pradeep, P. Seshuraju, S. A. Khaparde, V. S. Warrier and S. Cherian (2009), "CIM and IEC 61850 integration issues: Application to power systems," *2009 IEEE Power & Energy Society General Meeting*, Calgary, AB, Canada, 2009, pp. 1-6, doi: 10.1109/PES.2009.5275765.