Predictive Analytics for Hydropower Fleet Intelligence

Yigit Yucesan[1], Pradeep Ramuhalli[2], Yang Chen[3], Jim Miller[4], Edward Hanson[5], and Stephen Signore[6]

[1,2] *Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37830, USA*
yucesanya@ornl.gov
ramuhallip@ornl.gov
[3] *North Carolina Agricultural and Technical State University, Greensboro, North Carolina, 27411, USA*
ychen1@ncat.edu
[4] *Signal Hydropower Consultants LLC, Signal Mountain, Tennessee, 37377, USA*
ljmiller-sec@epbfi.com
[5] *Pacific Gas & Electric Company, Auberry, California, 93602, USA*
eahf@pge.com
[6] *Tennessee Valley Authority: Power Operations, Knoxville, Tennessee, 37902, USA*
ssignore@tva.gov

ABSTRACT

A primary challenge in hydropower industry is the ability to maintain cost-competitiveness, reliability, and security of hydropower assets through evolving power system contexts and aging of the fleet. Maintaining cost-effective and reliable operations under these conditions is expected to require new modernization and maintenance paradigms for changing contexts. Changes in existing practices for O&M will require an understanding of the current state and health of hydropower assets, and the impact of changing paradigms on asset health and reliability. The Hydropower Fleet Intelligence project is developing and evaluating standardized methodologies and analysis tools for data-driven asset reliability and management technologies for hydropower, leading to eventual predictive maintenance planning, repair/replacement decision making, and asset-reliability and cost-optimized operations. A key question is the feasibility of using existing data sets at hydropower facilities to perform assessments of asset reliability. This document uses data from hydropower facilities to assess the potential for using available analytics methods for asset reliability estimates. In addition to reliability assessments, the feasibility of using existing analytics techniques for several other potential applications is discussed. Finally, a case study that a data-driven model is trained to learn nominal operations via vibration data from an asset of a certain plant, and then utilized to identify anomalies on a similar asset from a different plant, highlighting the generic use of proposed Prognostics and Health Management (PHM) approaches.

Notice: This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

1. INTRODUCTION

A challenge in hydropower operations is the ability to maintain cost-competitiveness and reliability of hydropower assets in evolving power system contexts and aging of the fleet. Recent analyses indicate that operations and maintenance (O&M) costs are higher for older units, and there is an increasing trend in O&M costs for older hydropower facilities over the last decade or more (Martinez et al., 2021). At the same time, O&M costs for competing sources of generation (i.e. wind and solar energy) are expected to trend down over the next several years (Stehly & Duffy, 2022; Ramasamy et al., 2022). This challenge with an aging fleet comes at the same time as the level of variability in operations is on the increase, with more intermittent generation and increased hydropower participation in the ancillary services market. Collectively, these factors can have a negative impact on hydropower asset condition and availability and require new approaches for O&M.

Maintaining cost-effective and reliable operations under these conditions is expected to require modernization of maintenance paradigms (“Water Power Technologies Office: Multi-Year Program Plan”, 2022). One possible approach to modernization of maintenance paradigms is to leverage available data sources and standardize processes to make data-driven O&M decisions. Data for this purpose might be at the unit, plant, or national scale, with the changes in contexts being driven by external or internal factors (for instance, an increase in the intensity of variable dispatch, participation in ancillary services, or aging of a powertrain component, etc.).

The proposed process for data-driven O&M decision making is shown in Figure 1. Each facility is expected to be interested in answers to a set of questions specific to that facility. These facility-specific questions (for instance, remaining life of an asset, optimal replacement times for assets, or other O&M decision making) drive a data sufficiency assessment to identify whether the data available from the facility are sufficient for this purpose. The data is then used as part of a standard process to support decision-making by providing the necessary insights on asset condition, predicted reliability, risk, and cost.

The analysis methodology being developed and evaluated focuses on hydropower assets within the powerhouse, especially powertrain components. One challenge in developing such a data-driven approach is the need to integrate multiple sources of data. Figure 2 summarizes the data workflow, using a variety of data sources to perform asset reliability, dispatch variability, consequence/risk, and cost analyses. These data sets are correlated and integrated to allow different analyses methods to be applied for extracting information on hydro asset performance, health and reliability, and quantified with metrics such as asset mileage, failure rates, health state, etc. The analyses methods account for evolving power system and hydrology contexts, for instance due to variable dispatch, and provide an estimate of the metrics in context (for instance, changes in asset reliability and availability due to increases in flexible operation). The results of these analyses may be applied to risk and cost models to quantify the impact of these evolving contexts on hydropower availability, reliability, and O&M costs, thereby providing the necessary input to stakeholders for use in O&M and investment decision making.

In the recent literature researchers mostly utilize Machine Learning methodologies to optimize the energy generation. Tubeuf et al. (2023) study the potentials of Reinforcement Learning algorithm to control the blow-out process of a hydraulic machine during pump start-up and when operating in synchronous condenser mode. Sapitang et al. (2020) investigates multiple Machine Learning models, such as Boosted Decision Tree Regression, Decision Forest Regression, Bayesian Linear Regression and Neural Network Regression, to forecast reservoir water level for the purpose of enabling optimized operations. Review paper by Bordin et al. (2020) also discusses the current and future role of Machine Learning within the hydropower sector, and provides a big picture for the potential benefits for hydropower scheduling.

In this paper, we focus on two elements of this general methodology - reliability modeling and prognostic health management (PHM). These two elements are two sides of the same coin, helping determine the current state of a component based on historical maintenance and outage records, and the future condition of an asset based on condition monitoring data. Note that these two aspects of assessing asset health rely on different types of data and data with different levels of granularity.

2. Data Sources

Preferred data sets for these analyses include data generally collected by hydropower operators and include design and capability data, unit and facility level availability and outage data (such as the North American Electric Reliability Corporation (NERC) Generating Data Availability System (GADS) data), component, asset and unit level time-series data (Supervisory control and data acquisition (SCADA) data), maintenance work orders and asset maintenance/repair histories, and operations and maintenance cost data (in the event that cost analysis is necessary). Other information, such as hydrological and power system data, are useful to place the results of the analysis in context and provide insights into how these results may change with operations practices that adapt to changing contexts. The methodology itself is applicable to other assets (such as those in the balance of plant/switchyard or those comprising the dam and related structural components), though the data sets needed for assessing the condition and reliability of these assets will be different.

Data levels used in this work are those maintained at a facility and containing information from O&M activities. While other publicly available data sets (such as the HydroSource data (B. T. Smith et al., 2019) or USGS Streamflow data (Lins, 2012)) are relevant and can be integrated, at present these types of data are only used to provide context to the O&M findings and are not the focus of the research discussed in this paper.

While the focus of the research described in this paper is the application of PHM to a single facility, the methods themselves are likely useful to most facilities and may be used for performance benchmarking and comparative analyses across similar units and multiple facilities. Benchmarking activities will require access to aggregated O&M-related data from multiple facilities with similar units and may be able to leverage industry-led fleet-level data consolidation efforts.

The specific data used in the research discussed in this paper includes data from two utilities covering conventional and pumped storage hydropower facilites as well as data from a consolidated hydropower data repository. Facility data from utilities included data covering a reasonably long duration ($\sim$ 15 years or more for most data types). Variations in data verification practices at the different facilities meant that not

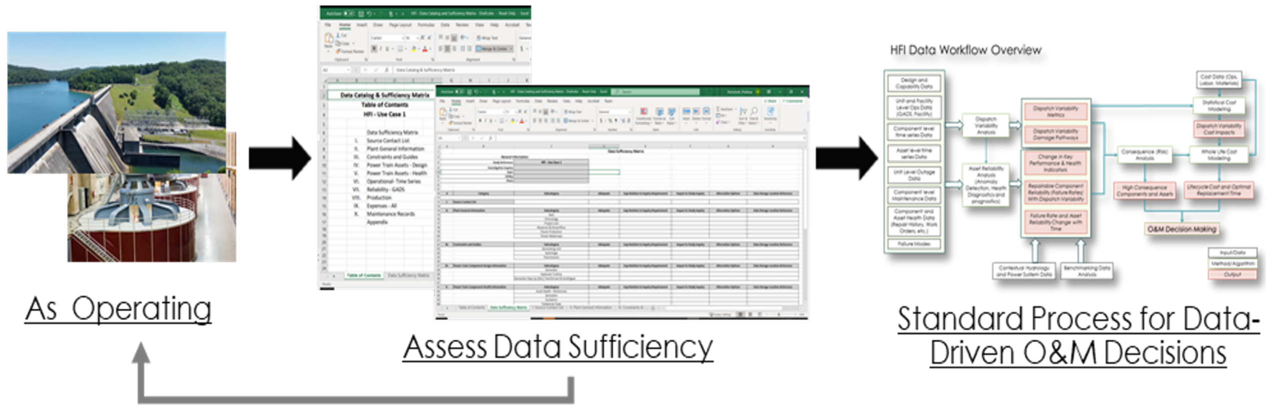


Figure 1. Overall process for a standardized approach to data-driven hydropower O&M decision making.

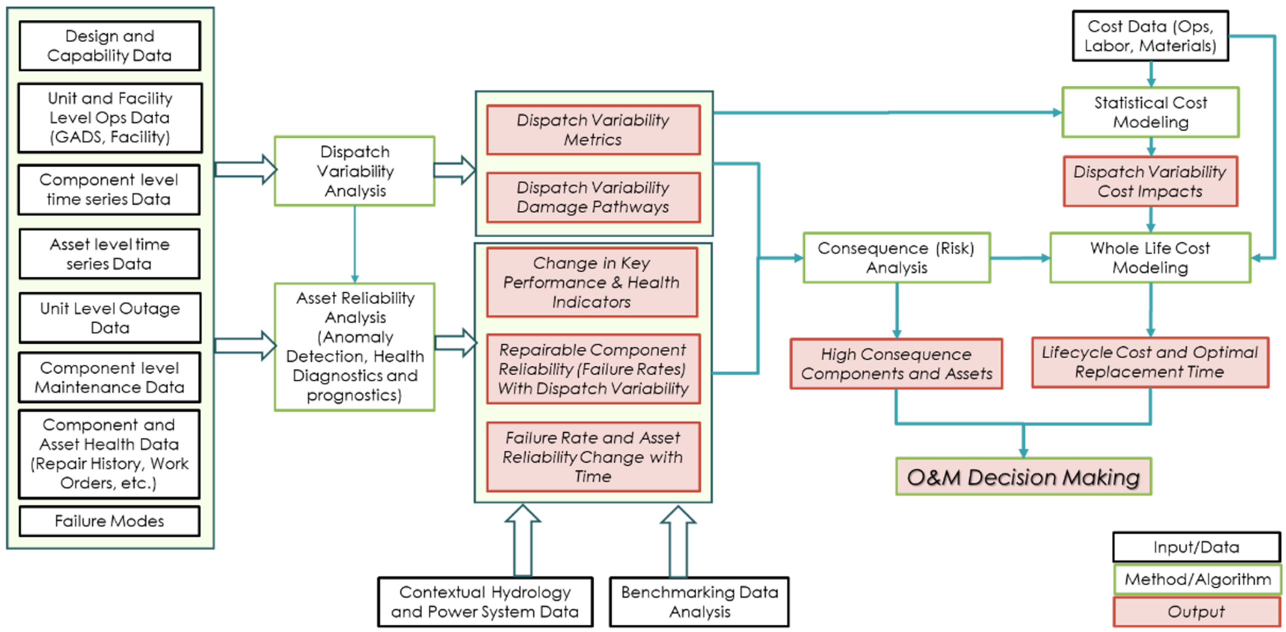


Figure 2. Workflow for data-driven analysis methodology for O&M decision making.

all data sets available to the research team covered the entire duration. While some types of information were missing from one or more facilities (for instance, complete condition reports, or maintenance histories), some of this information was able to be inferred from other data elements, and the rest were determined to not be essential for this research. Fleet-level aggregated time-series data from a consolidated repository was also evaluated used for a portion of this research.

3. PHM for Hydropower Assets

PHM integrates advanced monitoring technologies, data analytics, and predictive modeling to assess the health and performance of hydropower components and systems. By enabling condition-based maintenance, PHM optimizes maintenance activities, reduces costs, and can help minimize downtime. It also extends the lifespan of assets by detecting early signs of degradation or catastrophic anomalies, and implementing proactive maintenance strategies. Despite challenges, advancements in technology have paved the way for robust PHM solutions, offering improved efficiency, reliability, and sustainability in hydropower operations.

Anomaly detection is a specific element of PHM and usually focuses on identifying deviations from nominal behavior using time-series data (Ruff et al., 2021). Anomaly detection therefore requires an understanding of nominal behavior, which is usually represented using a model derived from operating data. In most applications, the model uses a variety of inputs to estimate an expected output (for instance, a measurement of temperature or vibration or other quantity) assuming nominal conditions. This estimated value may be compared against the actual measurement, with a large deviation indicative of an anomaly. The comparison may be performed by applying a simple threshold to the difference between the model estimate and the measurement, or through a statistical test.

Since anomaly detection algorithms detect changes or deviations from nominal, they can be applied to detect fault conditions in hydropower assets. For example, bearing vibration data may be used to detect the onset of excessive vibration conditions that may indicate bearing failure or other conditions (for instance, mechanical imbalance or misalignment, cracked blades or shaft, or improper lubrication) (Mohanta et al., 2017). Anomaly detection may also be useful for detecting changes in stator winding temperature (B. Smith et al., 2022) that may indicate a developing fault condition in the stator.

3.1. PHM Methods

In the context of PHM applied to hydropower assets, data-driven approaches offer great opportunities to leverage historical operational, maintenance, and failure data. In this paper we will focus on the uni-variate anomaly detection problem for specific rotating equipment of hydropower units. The main goal is to adopt a data-driven approach to model normal operation with an ability to flag deviations pass a predetermined threshold as an anomaly before catastrophic failure. Here we describe three distinct machine learning models often used for anomaly detection and justify our selection.

Variational Autoencoders (VAEs) (Sun et al., 2018) are effective machine learning tools to extract relevant features of input data, explore nonlinear relations among input variables, and decode into latent space in the form of a distribution. Encoded variables are then sampled from the latent space and decoded to reconstruct the input data and the reconstruction error is minimized for an accurate VAE model. This model is pretrained with known normal operation then used on test data, where normal operation is expected to yield to low error, and any input that inherit anomalous data would fail to reconstruct, and ultimately flagged.

Recurrent Neural Networks (RNNs) (Goodfellow et al., 2016) are special type of neural networks designed to model sequential data. State transition is carried out through a repeated cell that takes temporal data as input vector and utilizes a hidden state to learn transition from one time step to the next. Researchers tailored the RNN cell to mitigate vanishing gradient problem in long sequences (i.e. missing out long-term dependencies in the time series data), and introduced designs such as Long-short Term Memory (LSTM) cell with forget and input gates, and Gated Recurrent Unit (GRU) with update and reset gates. RNNs can be trained with nominal operational data to predict the variable in consecutive time steps, which enables monitoring and identifying any deviation.

Convolutional LSTMs (ConvLSTMs) (Petersen et al., 2019) combine convolutional layers and LSTMs to extract spatio-temporal relationships inherit in time series data and enable learning for multi-variate multi-step problems. Even though ConvLSTMs are conventionally used for image sequences such as video processing, researchers modified 1-D time series data into batches of short-term time periods through sliding window, which allows the model to explore short-term trends and seasonality effects from long-sequence time series data for accurate predictive performance.

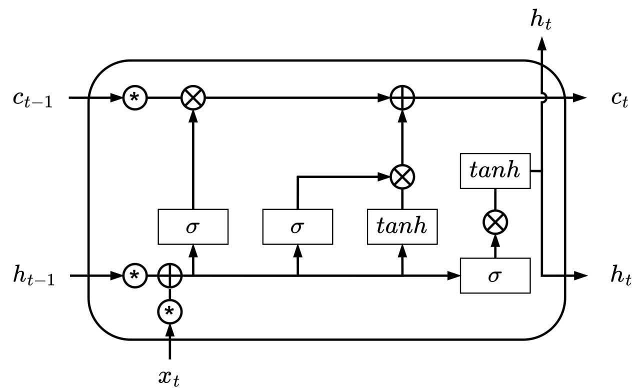


Figure 3. ConvLSTM architecture.

4. Health Management for Hydropower Assets

4.1. Asset Reliability Assessment

The previous section discussed various algorithms that may be applicable to deriving insights about hydropower asset conditions and other factors using existing data. A particular need is the analysis of these types of data to determine the reliability of assets. In this section, we consider whether available data sets at typical hydropower facilities may be used to derive estimates of reliability of assets. In this document, reliability is defined as probability that the components or units could perform its function adequately for a specified period of time without outages (Kuo & Zuo, 2003; Reid & Cox, 2018).

Key to most reliability analyses is the availability of failure data. Reliability analysis in the example discussed here uses unit availability data from NERC GADS for a specific facility. Data on outages from a sample unit are used in this study, with information on the timing and duration of maintenance and forced outages used in the analysis.

The reliability model used is at a higher-level in asset taxonomy models typically used in hydropower. Specifically, the examples in this section assume models at the level of a generator, turbine, or electrical system. Each of these asset classes includes multiple components, with a failure in any of the components within this asset class assumed to cause the asset to be in an outage. Note that reliability models of the lower level asset classes are not considered in this study, and the research demonstrates the potential for using outage data for estimating the reliability of hydropower assets.

NERC GADS outage information data over a 30 year timeframe were obtained and examined to identify maintenance outage (MO) and unplanned/forced outage (U1) events corresponding to components associated with any of the three asset classes described above. Included summary information on the activities performed and the components impacted were used to group components into higher level asset classes (for instance, the outage records of generator bearings, filters, rotors, etc., were grouped into a Generator category). Such grouping also addressed the limited data on specific components, increasing the total available data for the higher-level asset classes.

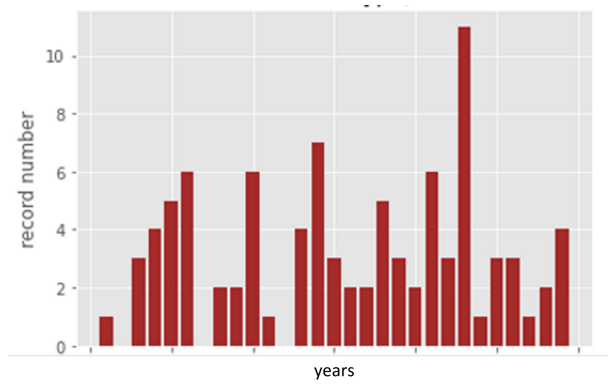
Figure 4 shows an example of the number of MO and U1 records for one unit in each year. If the assumption is that each of the U1 outages corresponds to a "failure" of the asset (i.e., asset being unable to perform its function), the timing of these outages may be used to estimate the reliability (or equivalently, the failure rate) of the assets in this initial analysis. Note that year information in these plots are omitted for proprietary confidentiality.

Note that this estimate of the reliability is not, strictly speaking, complete as data on planned outages is excluded from this analysis. As a result, repairs conducted during one or more planned outages between two maintenance/forced outages and which improve the asset condition are not accounted for. As a result, quantities such as Mean Time Between Failures (MTBF) that are computed from the reliability models need to be viewed with a grain of salt. Improved modeling through the inclusion of repair/maintenance information in the models (sometimes referred to as reliability modeling of repairable assets) can help address this issue and is planned as future work.

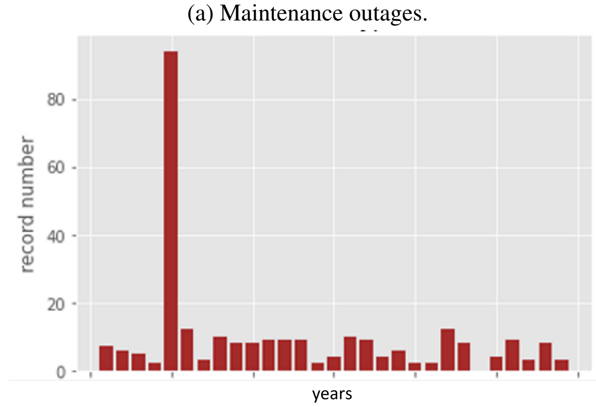
Key to reliability analysis is the assumed lifetime or failure rate distribution model. While a number of possible probability distributions can be used, the analysis discussed in this report used the Weibull distribution (Kuo & Zuo, 2003). This distribution is widely used in reliability engineering as it can represent a range of scenarios using a small set of parameters. While there are different versions of the Weibull distribution based on the number of free parameters, this study used the three-parameter version which is the most general form of the distribution. The three-parameter Weibull probability density function is given by (Martz, 2003):

$$f(T) = \frac{\beta}{\alpha}(\frac{T-\gamma}{\alpha})^{\beta-1} \cdot e^{-(\frac{T-\gamma}{\alpha})^{\beta}} \quad (1)$$

with the failure rate or hazard function defined by



(a) Maintenance outages.



(b) Forced outages.

Figure 4. Number of outage records for maintenance outages and forced outages for a sample unit in a pumped storage facility.

$$\lambda(T) = \frac{\beta}{\alpha}(\frac{T-\gamma}{\alpha})^{\beta-1} \quad (2)$$

Here, T is the time, β is the shape (Weibull slope) parameter which equal to the slope of the line in a probability plot, α is the scale parameter and γ is the location parameter.

4.2. Reliability modeling results

The available data from NERC GADS was processed to compute the time-to-failure for each occurrence of a maintenance or forced outage. These time-to-failure estimates essentially computed the time between two successive outages and included the outage duration. Figure 5 shows an example of the outage instances for the Generator asset class for one unit at this facility; the duration and occurrence time of the outage are presented. It is easy to see that most maintenance and forced outages are relatively short in duration. Once again year information is censored for confidentiality purposes.

The reliability analysis was performed using 10- and 15-year

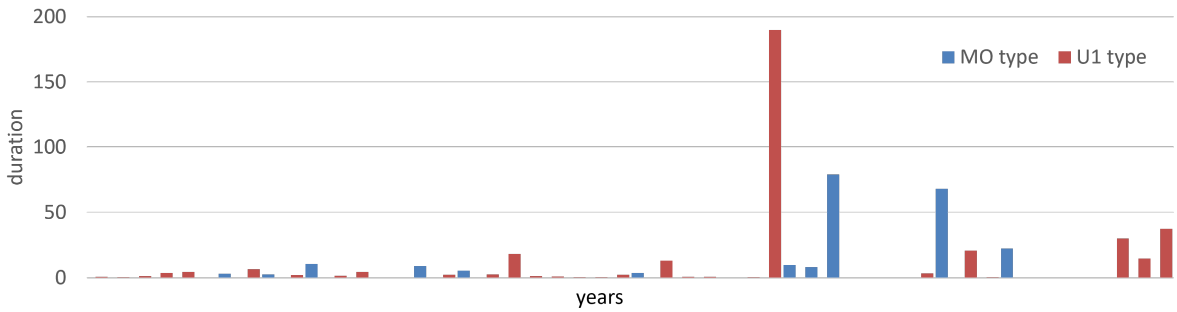


Figure 5. Maintenance outages time stamp and duration of "Generator" category in pumped storage facility for sample unit.

sliding windows, and the Weibull parameters computed. Figure 6 shows the Weibull β parameter estimated from the 10 year and 15 year windows and indicates a value that is generally low (less than 1). The low value of β indicates a failure rate that is decreasing over time in general. However, a careful examination of the data indicates an initial increase in the failure rate for the unit, followed by a slow decrease. This is likely due to two factors: the removal of planned and periodic maintenance actions that help maintain the assets in reasonable condition, significant refurbishment/replacement activities that took place around the time the failure rate flattens out, and the possibility that the asset (generators in the unit studied here) are past any initial failures and have not yet reached the wear-out failure stage of the reliability curve.

A similar result for turbines and electrical equipment (not presented) seems to indicate a generally healthy set of assets despite a slight uptick towards the end of the time period being studied.

It should be noted that Weibull analysis for hydropower reliability assessment has a long history, though much of the analyses to date have been focused on failure data from one or more asset classes across the fleet. Most of these analyses are unpublished in the public literature, and have generally been used by hydropower owners/operators for internal use. The preliminary analyses presented above are intended to determine the applicability of standard reliability analysis methods using failure data from a single unit.

Given the results, it appears that with long term failure data, as is typically maintained at a facility, Weibull analyses can be applied to determine potential changes in asset reliability parameters over time. However, as indicated above, changes in asset reliability due to planned maintenance are difficult to quantify and therefore the windowed analysis approach needs further evaluation before it can be deployed for routine use.

4.3. PHM Case Study: Anomaly Detection on Generator Guide Bearing

While reliability modeling is helpful and provides an estimate of the failure rates, the information is largely based on past operational histories and difficult to use for near term maintenance scheduling. The use of PHM methods provides an alternate approach to quantifying the condition of assets and provides early warning of impending failures, allowing for near term scheduling. One example of such an approach is shown here, which utilizes vibration data. While the use of vibration to predict anomalies in rotating machinery is not novel, there appear to be few studies applying these techniques to hydropower equipment. Such a study is therefore worthwhile and helps identify specific challenges associated with these machines. In addition, in this study, we focus on the generalizability of vibration models, and show that, at least for the data sets studied, data from fleet-wide repositories may be helpful in developing robust models that are applicable to other similar units.

4.4. Site and Equipment

In this paper, we will use single axis vibration data obtained from generator guide bearing of a unit located in one site (Site A) to train our anomaly detection model. The main goal is to teach the model the trends and variations of nominal operational conditions. After a successful training session, model is expected to predict normal operation, and ideally unable to predict the variable accurately during any anomalous event affecting the bearing vibration, thus can be flagged using the deviation between prediction and true data prior to any catastrophic failure.

In the second part of the analysis, we will use the trained model to predict the vibration of a generator guide bearing of a different unit from a different site (Site B). Goal of this study is to present an efficient way to translate ML-based anomaly detection models across different sites.

Smith et al. (B. Smith et al., 2022) categorizes the comparison of multiple hydropower assets in three categories concept

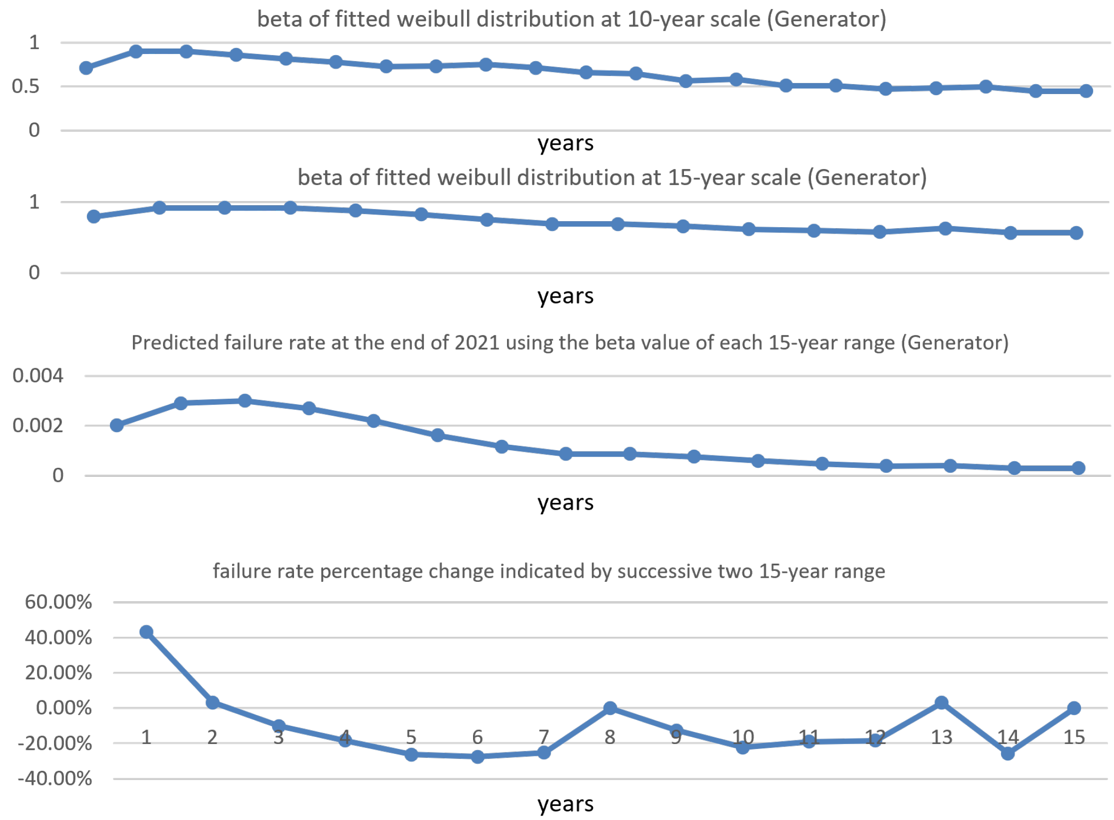


Figure 6. Estimated Weibull parameter and failure rate changes for assets assigned to the "Generator" category. Top to bottom: β using 10 year time windows for aggregating major maintenance activities, β using 15 year time windows, failure rate estimates using 15 year windows, failure rate percentage change between two 15-year range.

of similarity. When two or more hydropower assets share the exact same specifications and characteristics, these assets can be considered as *identical*. In the case where one asset is the scaled version of another asset, but otherwise identical, in terms of context or design, we can call these assets are *homologous*. Finally, when two or more assets have certain proximity in some of their characteristics, these assets are deemed to be *similar*. According to this taxonomy, we classify two assets from Site A and B as *similar*.

4.5. Vibration Data

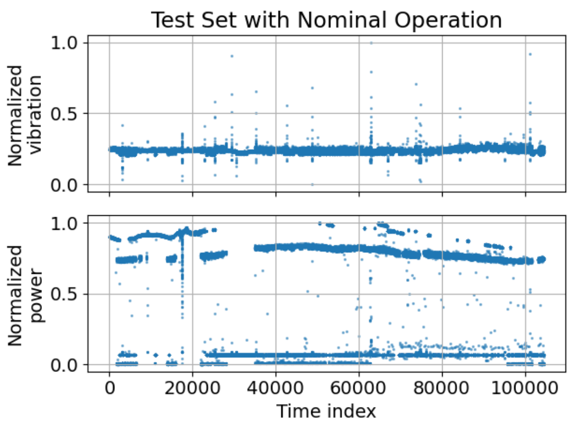
Our source data, as we cite as Site A, is obtained from a consolidated data repository. We consider vibration data with 5-minute resolution covering approximately 2,016 generating hours, where no vibration-related forced outage is recorded. We will divide this data into training and validation sets, and test the sensitivitiy of the model performance on division, by evaluating 70/30%, 80/20%, and 90/10% training and validation data points respectively.

From our target plant, cited as Site B, we take two long time periods: first one with normal operations without vibration failure (Figure 7a), second one with a forced outage caused by excessive generator bearing vibration (Figure 7b). We also illustrate the power generation during these periods to point out similar operational patterns, despite the impending anomaly in the Figure 7b. Note that all data are normalized between 0 and 1 to preserve anonymity of the sites and operators.

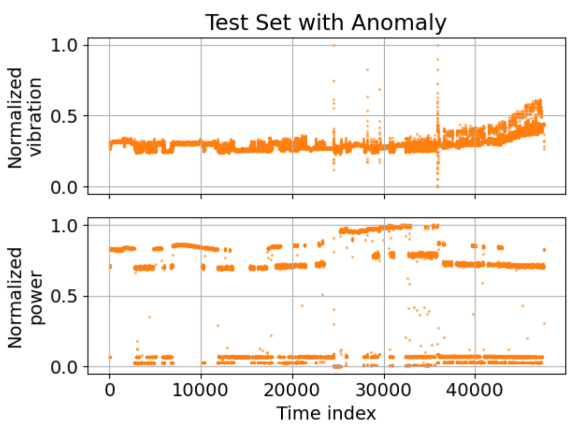
In order to feed the data into model in batches, we restructure data set by creating subsets where certain amount of data points from past is used to predict the next time step through a sliding window. We also analyze the effect of the window size on model performance by testing 60, 30, 15, and 5 data points as our sliding window of the past.

4.6. ConvLSTM Model and Training Parameters

We choose the ConvLSTM framework as the vibration model (described in Section 3.1), since our data includes both short term (start-stop) and long term (seasonality) trends and ConvLSTM models have proven track record with such datasets (Petersen et al., 2019). Another reason that we opt in for ConvLSTM is also ability to scale up the input (and output if necessary) parameter space without causing computational



(a) Vibration data with nominal operation.



(b) Vibration data with anomaly.

Figure 7. Data from Site B

overload as much as other methods that utilize covariance matrix such as Gaussian Process. There are multiple applications in the literature that use ConvLSTM for multi-variate time-series prediction (Xiao et al., 2021; Mu et al., 2019; Tayeh et al., 2022). In the future, we are aiming to include many other relevant sensor variables (temperature, current, etc.) to enhance our predictive model. Summary of the design of model architecture and number of trainable parameters is given in Table 1.

Model training is executed with 500 epochs, with a batch size of 256, and 20 steps per epoch. Adam optimizer is chosen with an initial learning rate of 0.001 to minimize Mean Squared Error (MSE) loss function. Additionally, Reduce Learning Rate on Plateau callback is utilized to ensure convergence, with a factor of 0.85 and patience of 50 epochs.

5. Results and Discussion

First, we report the results of our window size sensitivity analysis on Table 2. We evaluate our ConvLSTM model's performance on both training and validation sets for different window sizes used to structure the data and normalized root mean squared error (RMSE) metrics are reported. The results indicate that there is marginal difference in changing the window size on the model performance. Even though window size of 60 has slightly better (lower) normalized RMSE on training set, we consider the performance on the validation set (which is in the future and was not seen by the model) as our selection criterion, at which window size 30 performs marginally better.

Layer	**# parameter**
Batch Normalization	4
ConvLSTM2D	8,768
Dropout	0
Batch Normalization	64
Flatten	0
Repeat Vector	0
Reshape	0
ConvLSTM2D	10,304
Dropout	0
Time Distributed	17
Dense	2
Total parameters	**19,159**

Table 1. ConvLSTM Architecture

Window	**Norm. RMSE (Training)**	**Norm. RMSE (Validation)**
60	**0.02891**	0.04333
30	0.02897	**0.04326**
15	0.02919	0.04367
5	0.02980	0.04419

Table 2. Window size analysis. 80/20% split between training and validation sets used.

After freezing the window size, we evaluate the effect of training and validation data split on our model on Table 3. We use three different splitting ratio between training and validation sets, and train the model from scratch. Note that we always keep the chronological difference between training and validation same, i.e. train using the past data and validate using future data. Shown in Table 3, 80/20% split gives us the optimum model performance on the validation data. We also observe that larger the data we feed in to the model does not always yield to better performance.

Trg/Val Split	**Norm. RMSE (Training)**	**Norm. RMSE (Validation)**
90/10%	0.02910	0.09765
80/20%	0.02897	**0.04326**
70/30%	**0.02847**	0.04454

Table 3. Training and validation set split analysis. Window size of 30 is used.

Learning curve of the ConvLSTM model detailed in Table 1 is given in Figure 8. While we observe large fluctuations in training loss, the model is clearly converging with a stable trend. The fact that this convergence aligns with the validation loss indicates a favorable training session, where the model is able to prevent overfitting and generalize the data.

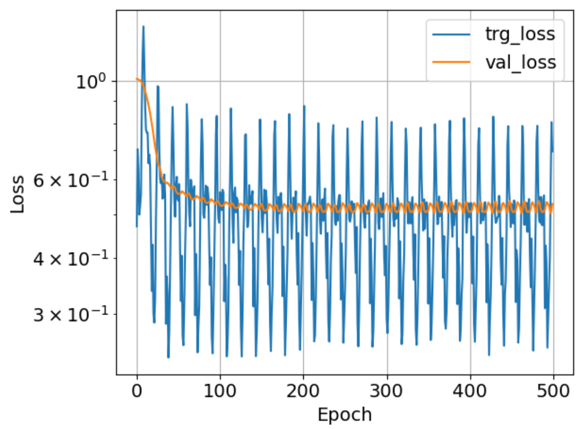
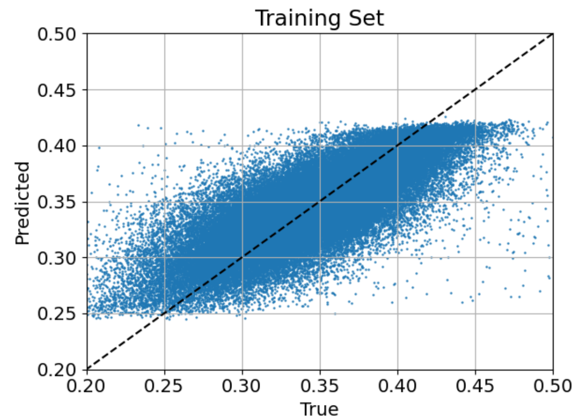


Figure 8. Learning curve.

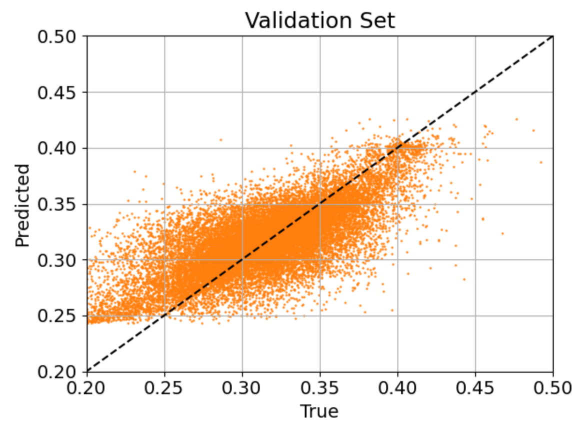
Figures 9a and 9b illustrate the model vibration prediction on training and validation data sets. In these plots, each point correspond to a vibration value at a single time step. While X-axis is the actual measured vibration, Y-axis is the model prediction for the same time step. Ideally we would like the data to be consolidated at 45 degree line. Figure 9a shows that data aligns with the 45 degree line, even though there is variation. This can be explained by noisy data, however the model predictions tend to follow general trend of the time series; hence the variation. Similar deduction can be made for the validation data from Figure 9b.

In Figure 10, we can observe the prediction error over training and validation sets extracted from Site A. Note that a moving average of 60 time steps is calculated to discard point-wise major variations and capture longer time-scale trends. This graphic confirms our deduction that model is able to follow the general time series trend and successfully predict the vibration using past data both for training and validation sets with a maximum prediction error of $\cong 5\%$. With this validated model prediction performance, now we can test the model on Site B.

Figures 11 and 12 respectively show the model vibration prediction on two distinct, the former with nominal operation and the latter with an anomaly. In Figure 11 we clearly observe that predictions follow the general trend of the actual vibration time series data. One can also easily infer that the variations seen in Figure 9 is caused by measurement variation that does not reflect the actual time series history. We can conclude that the model trained on data from different site with similar unit can be successfully used on different site to model nominal operational vibration of generator guide bear-



(a) Prediction performance on training set.



(b) Prediction performance on validation set.

Figure 9. Point-wise model performance on training and validation sets on Site A.

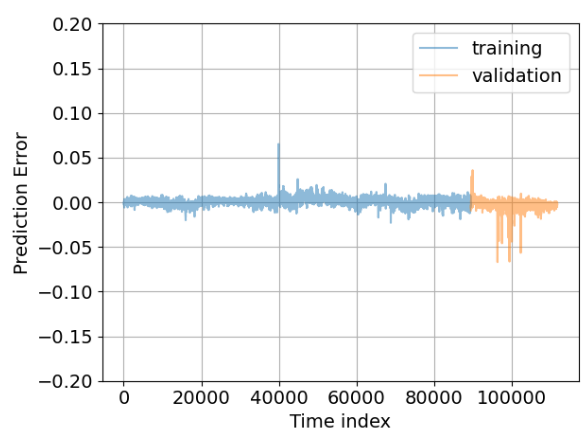
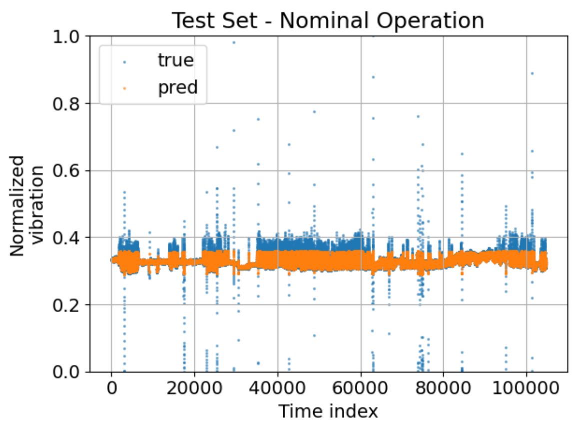


Figure 10. Model performance on training and validation sets on Site A.

ing. In Figure 12, we see similar results in the first three quarters of the time series. However, towards the end of this time period, the model is unable to pick up the unexpectedly increasing trend of vibration. This ’failure’ of the model is exactly what we were aiming for. Now we can utilize the prediction error as a tool to flag the anomaly.

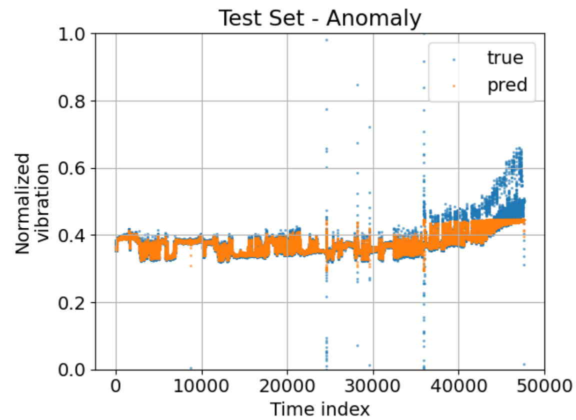
Figure 13 depicts the prediction error in both time periods. Just as the source data results, top figure proves decent predictive performance on nominal operation. The maximum prediction error is found to be ≅ 5%. However in the bottom figure, enhances the gradual increase of the prediction error, where the failure is imminent. One can simply set a threshold of 10% to this metric and raise a flag for the anomaly before the fault occurs and potentially leading to a catastrophic failure.



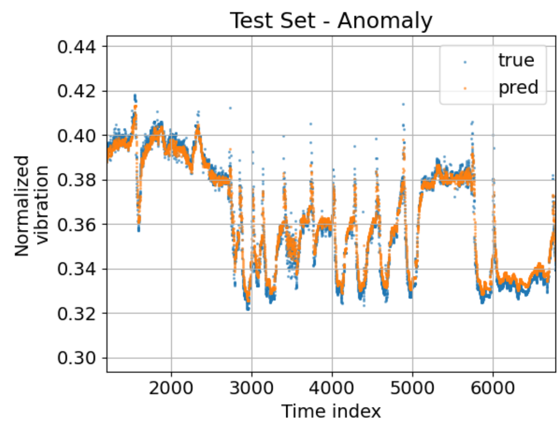
(a) Whole time series.

(b) Zoomed in to illustrate predictive performance on the beginning of the time series.

Figure 11. Prediction performance on Site B, of time period where there is no vibration related forced outages.



(a) Whole time series.



(b) Zoomed in to illustrate predictive performance on the beginning of the time series.

Figure 12. Model performance on Site B.

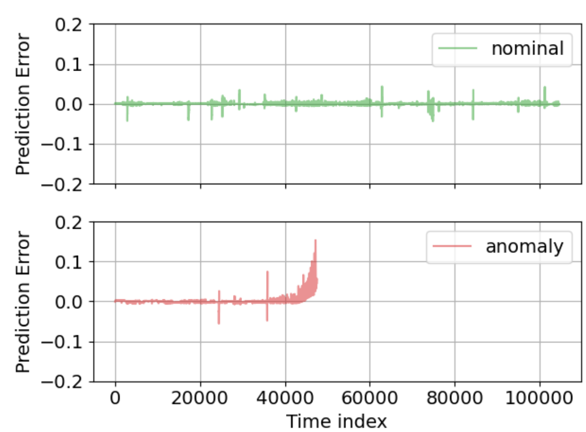


Figure 13. Prediction error over Site B for both years.

6. Conclusion

This paper discusses the application of predictive analytics and health management techniques for hydropower fleet. We lead by emphasizing the importance of data sources, including design sheets, availability and outage records, operational time-series, and historical maintenance data. We also discuss the methodologies that can be applied across different facilities for performance benchmarking and comparative analyses. Then we explore the use of PHM for hydropower assets, highlighting the benefits in optimizing maintenance activities, reducing costs, extending asset lifespan, and detecting early signs of degradation or anomalies. Additionally, we dive into asset reliability assessment and the use of existing reliability modeling techniques to estimate the reliability of hydropower assets based on failure data. We carry out an example analysis using NERC GADS outage data, demonstrating the potential for estimating asset reliability. Finally, we present a case study on anomaly detection for a generator guide bearing using vibration data. We showcase the efectiveness of machine learning-based models in predicting normal operation and identifying anomalies, across similar assets of different sites. Our results indicate the ConvLSTM model trained on a different site can predict an anomaly on a similar unit installed on another site, hence potentially enabling near-term maintenance scheduling, ultimately enhancing the efficiency, reliability, and sustainability of hydropower operations.

In the future, we aim to extend this work by exploring:

- PHM potentials on other hydropower assets, such as turbines, stators, etc.
- Uncertainty quantification components for the data-driven models, to provide prediction confidence.
- Building back-to-back predictive analytics workflow for hydropower plants, connecting asset-specific prognostics to fleet-level reliability analysis, and ultimately providing cost effect on operations.

Acknowledgment

This work was supported by the US DOE Water Power Technologies Office and conducted at Oak Ridge National Laboratory. The authors are employees of UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US DOE. Accordingly, the US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript or allow others to do so, for US Government purposes.

References

Bordin, C., Skjelbred, H. I., Kong, J., & Yang, Z. (2020). Machine learning for hydropower scheduling: State of the art and future research directions. *Procedia Computer Science*, *176*, 1659–1668. doi: 10.1016/j.procs.2020.09.190

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Kuo, W., & Zuo, M. J. (2003). *Optimal reliability modeling: principles and applications*. John Wiley & Sons.

Lins, H. F. (2012). Usgs hydro-climatic data network 2009 (hcdn-2009). *US Geological Survey Fact Sheet*, *3047*(4).

Martinez, R., Johnson, M., & Shan, R. (2021, January). *U.s. hydropower market report (january 2021 edition)* (Tech. Rep.). doi: 10.2172/1763453

Martz, H. F. (2003). *Reliability theory*. Elsevier.

Mohanta, R. K., Chelliah, T. R., Allamsetty, S., Akula, A., & Ghosh, R. (2017, April). Sources of vibration and their treatment in hydro power stations-a review. *Engineering Science and Technology, an International Journal*, *20*(2), 637–648. doi: 10.1016/j.jestch.2016.11.004

Mu, B., Peng, C., Yuan, S., & Chen, L. (2019, July). ENSO forecasting over multiple time horizons using ConvLSTM network and rolling mechanism. In *2019 international joint conference on neural networks (IJCNN)*. IEEE. doi: 10.1109/ijcnn.2019.8851967

Petersen, N. C., Rodrigues, F., & Pereira, F. C. (2019). Multi-output bus travel time prediction with convolutional lstm neural network. doi: 10.48550/ARXIV.1903.02791

Ramasamy, V., Zuboy, J., O'Shaughnessy, E., Feldman, D., Desai, J., Woodhouse, M., … Margolis, R. (2022, September). *U.s. solar photovoltaic system and energy storage cost benchmarks, with minimum sustainable price analysis: Q1 2022* (Tech. Rep.). Retrieved from `https://doi.org/10.2172/1891204` doi: 10.2172/1891204

Reid, N., & Cox, D. (2018). *Analysis of survival data*. Chapman and Hall/CRC.

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., … Muller, K.-R. (2021, May). A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, *109*(5), 756–795. doi: 10.1109/jproc.2021.3052449

Sapitang, M., Ridwan, W. M., Kushiar, K. F., Ahmed, A. N., & El-Shafie, A. (2020, July). Machine learning application in reservoir water level forecasting for sustainable hydropower generation strategy. *Sustainability*, *12*(15), 6121. Retrieved from `https://doi.org/10.3390/su12156121` doi: 10.3390/su12156121

Smith, B., Smyth, S., Peterson, A., Ramuhalli, P., Hu, I., & Bulmer, A. (2022, April). *Consolidated hydropower data repository: Value and opportunities* (Tech. Rep.). doi: 10.2172/1870208

Smith, B. T., Samu, N., Curd, S. L., Wei, Y., & Wei, Z. (2019, 3). Hydrosource data dictionary. doi: 10.2172/1619046

Stehly, T., & Duffy, P. (2022, December). *2021 cost of wind energy review* (Tech. Rep.). Retrieved from `https://doi.org/10.2172/1907623` doi:

10.2172/1907623

Sun, J., Wang, X., Xiong, N., & Shao, J. (2018). Learning sparse representation with variational auto-encoder for anomaly detection. *IEEE Access*, *6*, 33353–33361. doi: 10.1109/access.2018.2848210

Tayeh, T., Aburakhia, S., Myers, R., & Shami, A. (2022, April). An attention-based ConvLSTM autoencoder with dynamic thresholding for unsupervised anomaly detection in multivariate time series. *Machine Learning and Knowledge Extraction*, *4*(2), 350–370. Retrieved from https://doi.org/10.3390/make4020015 doi: 10.3390/make4020015

Tubeuf, C., Birkelbach, F., Maly, A., & Hofmann, R. (2023, February). Increasing the flexibility of hydropower with reinforcement learning on a digital twin platform. *Energies*, *16*(4), 1796. Retrieved from https://doi.org/10.3390/en16041796 doi: 10.3390/en16041796

Water power technologies office: Multi-year program plan. (2022, 3).
doi: 10.2172/1863492

Xiao, Y., Yin, H., Zhang, Y., Qi, H., Zhang, Y., & Liu, Z. (2021, January). A dual-stage attention-based conv-LSTM network for spatio-temporal correlation and multivariate time series prediction. *International Journal of Intelligent Systems*, *36*(5), 2036–2057. Retrieved from https://doi.org/10.1002/int.22370 doi: 10.1002/int.22370