HPC Analytics of Fused Thermal Plants Data to Optimize Operating Envelope

Sangkeun Lee (ORNL)
Travis Johnston (ORNL)
Dongwon Shin (ORNL)
Salvatore Della Villa Jr (SPS)
Robert Steele (SPS)
Christopher Perullo (Turbine Logic)

November 2022

FINAL CRADA Report
NFE-19-07565
DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via OSTI.GOV.

Website www.osti.gov

Reports produced before January 1, 1996, may be purchased by members of the public from the following source:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone 703-605-6000 (1-800-553-6847)
TDD 703-487-4639
Fax 703-605-6900
E-mail info@ntis.gov
Website http://classic.ntis.gov/

Reports are available to US Department of Energy (DOE) employees, DOE contractors, Energy Technology Data Exchange representatives, and International Nuclear Information System representatives from the following source:

Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831
Telephone 865-576-8401
Fax 865-576-5728
E-mail reports@osti.gov
Website https://www.osti.gov/

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
HPC ANALYTICS OF FUSED THERMAL PLANTS DATA TO OPTIMIZE OPERATING ENVELOPE

Sangkeun Lee, Travis Johnston, and Dongwon Shin
Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

Salvatore Della Villa Jr., Robert Steele
Strategic Power Systems, Charlotte, NC 28277

Christopher Perullo
Turbine Logic, Atlanta, GA 30308

November 2022

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR2272
## CONTENTS

**ABSTRACT** 4

1. **INTRODUCTION** 5  
2. **IMPLEMENTATION** 6  
3. **RESULTS AND DISCUSSION** 7  
   3.1 LONG-TERM RELIABILITY PREDICTION 7  
   3.2 SHORT-TERM RELIABILITY PREDICTION 9  
4. **ISSUES OR CHALLENGES** 11  
5. **IMPACT** 11  
6. **FUTURE WORK** 12
ABSTRACT

Today, thermal power plant operators are facing a new operational reality. Renewable energy, including wind and solar energy, are typically given higher dispatch priority over gas turbines for electric power generation. Gas generators are more easily throttled than coal and nuclear plants, so they have become the spring that balances out power production and demand to keep the electric grid stable. In this cycling state caused by renewables on the grid, gas turbine power plants can become more susceptible to failures just like a spring can fatigue if stretched too often. Thus, power plant operators must understand the potential impact on Reliability, Availability, and Maintainability (RAM) of their power plant assets.

Strategic Power Systems (SPS) has been collecting detailed reliability performance data captured from over a thousand global operating power plant units (gas, steam, and wind) as well as reciprocating engines for over 33 years. The dataset is a commercial dataset named ORAP (Operational Reliability Analysis Program). The specific focus of ORAP is the tracking plant operational, performance and event data to benchmark RAM. Specifically, the dataset includes unit pedigree data, events (e.g., failures, scheduled and unscheduled maintenance, etc.) data, operation profiles data, and age data. Since the dataset covers a number of power plant units in various operational conditions, it is potentially very valuable for understanding power plant failures and correlations with changing operation profiles.

In this project, ORNL extensively reviewed the ORAP RAM data, and it guided us to develop machine learning models that can predict time to next failures and forecast failure trends, which will be useful for optimizing power plant operation strategies. More specifically, we trained multiple random forest models and evaluated the model accuracy to validate with 10+ years of historical data. In addition, we implemented a web-based graphical user interface system for the models to show how our models can be used in more intuitive ways. This proof of concept allowed exploration of model use with power plant operators in mind. Developed machine learning models will be helpful for managing risks, planning maintenance and operation, ultimately reducing the down time and increasing the service hours. For future work, there are several interesting research topics including but not limited to model enhancement, creating synergy with traditional failure modeling approaches, and data-driven actionable recommendation and suggestions.

A second thrust of this project focused on leveraging real-time data (collected at 1 Hz frequency) from numerous sensors across the power plants. The primary goal of this thrust was to determine the feasibility of forecasting fine-grained detail of power plant operations several minutes in advance enabling an early warning system for plant operators which could allow them to mitigate issues which might otherwise lead to unexpected failures or other forced downtime. ORNL leveraged convolutional neural networks to learn from historical sensor data and make predictions about the future state of power systems. ORNL also built a system for visualizing the predictions (second by second) from a vast array of sensors to better understand the efficacy of the neural network making predictions. We leveraged an in-house software system called MENNDL (Multinode Evolutionary Neural Networks for Deep Learning) to design the neural network architecture, tailoring it specifically for this novel task. We were able to demonstrate that by using the most recent 30 seconds of history we could make predictions 5 minutes into the future with an average error of less than .2 standard deviations from the signal mean. For future work it would be interesting to explore how the choice of training data affects the accuracy of the model, how much historical data to provide when making predictions, how far into the future a prediction can be made, and most importantly, how to deal with sensors which periodically malfunction or report inaccurate readings.
1. INTRODUCTION

Renewable energy is the fastest increasing energy source in the United States, and the use of renewable energy increased doubled from 2000 to 2018. Globally, renewables made up 26.2 percent of electricity generation in 2018, and it is expected to rise to 45 percent by 2040 according to C2ES (the Center for Climate and Energy Solutions). Unlike traditional energy sources, renewable energy sources heavily depend on climate. For instance, excessive amount of rain or low, or too fast wind can reduce the production of wind or solar energy; however, the electricity demand will be still the same. As gas generators are more easily throttled than coal and nuclear power plants, they become more and more responsible for keeping the grid stable as depicted in Figure 1.

A new operational reality that gas generator operators are facing is that gas turbine power plants are now required to cycle and load follow, meaning that the power generation units need to be operated at varying load levels. More cycles means more starts and stops – as well as varying load. Gas turbine generators now have become the spring that balances out the power production and demand. Just like we know springs can fatigue when they stretched often or overly stretched, power plant operators empirically expect that this environment will have a potentially high impact on the Reliability, Availability, and Maintainability of their power plant assets. It is very important to note that more cycles may influence events (e.g., failures) as well as repair and restore times. However, there has been little data-driven, quantifiable, and measurable analysis that can guide power plant operators in this changing new environment.

Machine learning has shown its value in various domains. For instance, many industries have shown that how learning from data, identifying patterns and making decisions with minimal human intervention can be extremely useful to their business (e.g., image classification, recommending a product to a customer, finding friends in a social network, predicting customers actions, etc.). It is no doubt that good quality of data is the key element of the successful machine learning application. Strategic Power Systems (SPS) has been collecting detailed reliability performance data captured from over a thousand global operating power plant units (gas, steam, and wind) as well as reciprocating engines for over 33 years. Availability of good quality big data, SPS’s expertise in the domain and ORNL’s expertise in machine learning and high-performance computing were big motivation for this project.
This project demonstrated how real-world power plant data and machine learning can be synergistically leveraged to help optimizing thermal power plant operating envelopes by forecasting failures and other forced downtime events. Figure 2 shows an overview of our approach in this research. We reviewed the ORAP RAM data and performed an exploratory analysis to understand the data set. The result guided us to preprocess the dataset and formally define our thermal power plant failure forecasting problem. Based on our data analysis and problem definition, we came up with machine learning models that use the ORAP RAM data for training. We trained multiple machine learning models, more specifically, random forest models that can be used to predict time to next failures and forecast failure trends. We evaluated our model accuracy to show their usefulness with historical data. In addition, we implemented a web-based graphical user interface system for the models to show how our models can be used in more intuitive ways.

### 2. IMPLEMENTATION

**Preparing and sharing data sets:** SPS exported data sets from its Operational Reliability Analysis Program (ORAP®) database and transferred them to ORNL. The transferred data is composed of pedigree, event, aging, and operation data collected from various global operating gas, steam, and wind power plants. Data dictionary which explains the data set structures and contents has been authored by SPS and shared with ORNL.

**Exploratory and visual data analysis:** ORNL has performed initial exploratory data analysis to understand the structural/statistical characteristics and limitation of data sets. Based on the result of exploratory data analysis and various criteria such as prevalence, quality, granularity, and representativity, SPS and ORNL prioritized data sets and objectives.

**Initial machine learning model training and test:** We selected a specific category of units whose Duty Cycle is Baseload and Plant Arrangement is Combined Cycle - Multi Shaft with Bypass and trained a Random Forest model to predict failure trends. We start building GUI-interface for better communication among teams and intuitive testing and demonstration.

Predictive models for powerplant operation strategy optimization have been improved: ORNL updated pre-developed predictive machine learning models based on the discussion. We used all category of unit data with properly defined data filter to optimize the training of machine learning models. A data bootstrapping technique is developed and applied to increase the size of training data for the model. Also, the predictive model is updated to use not only numerical values but also categorical values (e.g., the duty cycle of a powerplant unit) as inputs for models.

**Extensive validation performed for understanding the accuracy of developed models:** Many numbers of training, predicting, and measuring the prediction accuracy have been performed to understand the
accuracy of developed models. Every unit in the database with at least 1-year worth of historical data has been used for model validation. Accuracy of predictive models was measured in error metrics in terms of days, starts, and fired hours. A total of 112,257 predictions were validated.

**Convolutional neural network training on real-time data:** We built a training and validation workflow with a 1D-convolutional neural network at its core. The neural network trains on real-time sensor data and is designed to predict near term future sensor readings (5 minutes into the future). We also created visualization tools to analyze and understand the error in the models—which sensor predictions are most accurate, and which have highest variability.

### 3. RESULTS AND DISCUSSION

#### 3.1 LONG-TERM RELIABILITY PREDICTION

Figures 3. Examples of forecasting failures for the next 365 days

With preprocessed ORAP RAM data, we trained machine learning models that predict expected failure count increase for given time period. We chose to use the random forest model, which is an ensemble model that trains many numbers of decision trees and combines the prediction of the trees to minimize the overfitting and maximize the model accuracy. Figure 3 shows examples of forecasting results generated by the trained model for a selected unit.
Figure 4. (a) Distribution of failure errors in terms of fired hours
(b) Our model’s prediction error changes over more data (regarding failure) collection

Figure 4 (left) shows the distribution of our model’s prediction error (error in predicting the next failure) in terms of cumulative fired starts. (e.g., if our model predicts a failure will be at fired hours $h$ and if the failure actually happens at fired hour $h'$, the error will be $h' - h$). If the error is a positive value, then it means that the actual failure was later than our prediction. On the other hand, if the error is a negative value, it means that the actual failure happened earlier than our model’s prediction. The averaged error was 1402.560 hours (58.44 days) with standard deviation 2850.577 hours (118.774 days), the distribution is skewed, so the median should be measured, which was 507.5 hours (21.145 days). It is important to notice that machine learning model needs to be trained again to keep the model up to date when new data is available. Table 1 shows how prediction error changes over plant age. We observed that the error of our prediction of a unit tend to be lower if the unit has more failure records in the past, in other words, our model’s accuracy will improve over time as we collect more data about the operated power plant units (see Table 1).

<table>
<thead>
<tr>
<th>curr_fail_cnt</th>
<th>Binned</th>
<th>Mean(age_fired_hours)</th>
<th>Mean(age_fired_starts)</th>
<th>Std Dev(age_fired_hours)</th>
<th>Std Dev(age_fired_starts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 — 5</td>
<td>5</td>
<td>8585.612903</td>
<td>977.3384615</td>
<td>9621.486291</td>
<td>2724.007889</td>
</tr>
<tr>
<td>5 — 10</td>
<td>10</td>
<td>5197.02924</td>
<td>795.7297297</td>
<td>7282.893197</td>
<td>2619.190043</td>
</tr>
<tr>
<td>10 — 15</td>
<td>15</td>
<td>4238.869919</td>
<td>1174.48062</td>
<td>6028.37575</td>
<td>4170.025452</td>
</tr>
<tr>
<td>15 — 20</td>
<td>20</td>
<td>2349.088608</td>
<td>799.0568182</td>
<td>3491.526931</td>
<td>3426.53149</td>
</tr>
<tr>
<td>20 — 25</td>
<td>25</td>
<td>1173.774194</td>
<td>838.5942029</td>
<td>2363.531363</td>
<td>2360.981079</td>
</tr>
<tr>
<td>25 — 30</td>
<td>30</td>
<td>1218.666667</td>
<td>173.733333</td>
<td>3111.462087</td>
<td>450.137272</td>
</tr>
<tr>
<td>30 — 35</td>
<td>35</td>
<td>1582.107143</td>
<td>11.9642857</td>
<td>2432.749798</td>
<td>49.41695776</td>
</tr>
<tr>
<td>35 — 40</td>
<td>40</td>
<td>1725.555555</td>
<td>-3.16666667</td>
<td>1604.40897</td>
<td>216.2597159</td>
</tr>
<tr>
<td>40 — 45</td>
<td>45</td>
<td>1470.333333</td>
<td>76.93333333</td>
<td>1901.076525</td>
<td>89.83832039</td>
</tr>
<tr>
<td>45 — 50</td>
<td>50</td>
<td>1245.625</td>
<td>97.625</td>
<td>1968.607335</td>
<td>146.3566461</td>
</tr>
</tbody>
</table>
We also developed a prototype web GUI (graphical user interface)-based thermal power plant failure forecasting system for subject matter experts (see Figure 5).

3.2 SHORT-TERM RELIABILITY PREDICTION

Leveraging the real-time sensor data (collected at frequency of 1 Hz, from over 100 sensors on a single power plant) we designed and trained a convolutional neural network to predict future sensor readings. We used a window of 30 seconds of history (the most recently observed 30 seconds) to predict sensor readings 5 minutes into the future. The goal of the model is to provide plant operators near real time advanced notice of potential problems so they can take appropriate action to either mitigate the problem (without causing an outage) or to prepare for a graceful shutdown if the outage is unavoidable. We demonstrated the feasibility of the approach by training on one week of data and validating the model on the subsequent day. The Figure 6 shows the evolution of training as the model learns to recognize and predict signals. Since the figure shows error, smaller values are better; the x-axis can be thought of as time spent training—so training has progressed as one reads from left to right. Focus should be placed on the solid black line (the validation loss) and specifically comparing it to the dashed line (a validation baseline). We demonstrate that after about 27,000 training instances the validation loss (using the neural network) begins consistently outperforming a baseline model. The chosen baseline model was to assume that the sensor values 5 minutes into the future would be approximately the average of the values for the most recent 30 seconds—essentially, that major changes in individual signals require larger timescales than 5 minutes to be apparent.
Figure 6: Example training results on real-time data using one week of training data. For Loss (error) lower numbers are better. After approximately 27,000 training instances, the Validation Loss is consistently better than the baseline comparison.

While Figure 6 shows the average error, Figure 7 illustrates the robustness of the prediction on a specific sensor—the overall load of the plant. Figure 7 shows the individual signals (in green) that were observed over the training days (1 week of data). The blue line shows how the signal tracked (and differed) on the validation day. In orange, we see that the predictions, being made 5 minutes in advance, track the observed data very accurately. Note that the y-axis depicts the distance from the mean plan load in standard deviations.

Figure 7: This figure shows the plant’s overall load on several days. The neural network was trained using days whose signals are shown in green lines. The validation day’s signal is shown in blue with the predicted signal closely matching the observed signal. Note that the units on the y-axis are in standard deviations from the average (MW).
4. ISSUES OR CHALLENGES

The effort to train a neural network on real-time data encountered one major challenge: data consistency. Neural networks are very brittle, so if we design our network to accept inputs from $N$ sensors, then every time we want to predict we must have consistent values for all $N$ sensors. Since sensors periodically malfunction—or there are periodic issues communicating the sensor readings with the server—we are stuck either training only on time frames where every sensor is working or attempting some mitigation strategy. Both cases present their own problems. If we only train (or predict) when all sensors are working, then the tool would be ineffective for substantial portions of time. If we try to mitigate the problem by either training multiple models (some using fewer signals) then we have an exponential growth in the number of models we need, and we build in an assumption that a failing sensor does NOT correlate with other failures in the system—i.e. we would have to assume that sensors failing are random and independent of everything else going on in the system, which is not likely always the case. If we try to mitigate the problem by inserting “fake” signals for missing sensors, then similar assumptions are being made (e.g. that the loss of a signal is not indicative of broader issues) and we may give the neural network the false impression that everything is “normal” if we insert a standard value for missing sensors. We ultimately decided to train only on a short interval of time where all sensors were working in order to show a proof-of-concept. Research around mitigating these issues and/or making neural networks less brittle is left as an intriguing future research area.

5. IMPACT

Three use cases of our developed predictive models were discussed, but there can be more use cases of our developed models.

- **Next failure prediction:** While operating a power plant unit, understanding the risk of next failure is very helpful for operators as they can better optimize their parts, maintenance schedule, and operation. Normally we are provided by the unit manufacturer what is the ordinary lifetime and maintenance cycle of a unit. However, a unit’s failure can be earlier or later depending on how the unit has been operated. The developed model can be used to numerically quantify the risk of next failure by estimating the required increased age value to the next failure.

- **What-if scenario-based failure trend prediction:** Today, power plant operators often face situations that they need to change the operation profile (e.g., how many hours or starts will be operated) due to integration of renewables to the grid. There can be multiple future operating scenarios for an operator, for instance, he or she can keep the same number of cycles, double, or triple the cycles. The developed model can generate various failure trend prediction based on synthetically modified operation profile and allow us to compare with historical average operation profile.

- **Early warning system, potential for automated mitigation:** Currently, there are manual alarms that can be set for specific ranges of specific sensors which can alert operators when a problem is already occurring and, in some cases, can trip shutdowns. Using the real time data, a predictive model could alert an operator of a potential issue which has not yet become a serious problem, but which might become serious (in several minutes) if no action is taken. This early warning system could assist to reduce unscheduled outages. If taken to the next logical step, it may even be possible to train a “virtual operator” capable of recognizing these issues and taking appropriate action in real time to prevent unscheduled outages either without the need of constant human supervision or with reaction times much faster than the human operator.

**Company Impact:** Collaborating with the labs has enabled SPS to perform analysis and investigations on the data available in the ORAP System that would not have been feasible as a small business. Working with ORNL has provided access to the resources and talents, adding subject matter expertise in AI and ML. The ability to assess the tools and techniques developed as part of the project provide an opportunity
to incorporate the methodology into the ORAP system for commercial use. This will increase the value of the system to the power plant operators who participate in ORAP, across the various technologies (as shown in Figure 1) - allowing the operators to make better informed operational decisions, reduce disruptions or forced outages, and meet the needs of changing service demands based on increased needs for operational flexibility.

**National Impact:** If successfully commercialized the tools and techniques developed will enable both large and smaller, independent power producers to continue to use existing capital investments in fossil-based energy generation. This maintains highly skilled labor and may help to reduce operational and maintenance costs leading to upgrades of capital equipment and other activities to maintain competitiveness in the marketplace.

6. **FUTURE WORK**

There are several interesting research topics as follows, but not limited to:

- **Model enhancement:** There is a huge opportunity to improve the model we presented in this work by using more data and tuning. One data we particularly can use is narrative data which is a natural language text data that describes events recorded in ORAP data. Also, combinational usage of ORAP RAM data and real-time sensor data for training models will be **key for commercialization of the technology**. Also, testing various model approaches including neural network and finely tuning hyperparameters should be investigated.

- **Creating synergy with traditional statistical failure modeling approaches:** Machine learning models has advantages in that they can be trained using large volume of high dimensional data; however, they do not have useful interpretable parameters like Weibull [4] does. After training machine learning model, we think that fitting the traditional model curve with the predicted failure trends generated by the model can provide additional information, which was not available either solely using machine learning or traditional models.

- **Data-driven actionable recommendations and suggestions:** Although forecasting the number of failures and its trend are significantly useful, it is often challenging for power plant operators to figure out what to do to mitigate the impact. Ultimately, providing suggested actionable items to subject matter experts will be one of the crucial future objectives.

- **Robust models and missing data:** As mentioned in section on issues and challenges, neural networks are brittle—requiring consistent inputs. Since the availability of signals from individual sensors fluctuates often it is important to develop a model (or modeling technique) which is resilient to fluctuations in the input data (i.e. the shape or size of the input data). Success would have obvious practical implications for this project but would also have broad impact on the field of AI/ML.