ORNL/TM-2022/2453

Consolidated Hydropower Data Repository: Value and Opportunities

Brennan Smith
Scott Smyth
Alisa Peterson
Pradeep Ramuhalli
Isabella Hu
Andrew Bulmer

April 2022

OAK RIDGE National Laboratory

ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY

ORNL/TM-2022/2453

Nuclear Energy and Fuel Cycle Division

CONSOLIDATED HYDROPOWER DATA REPOSITORY: VALUE AND OPPORTUNITIES

Brennan Smith[†]
Scott Smyth*
Alisa Peterson*
Pradeep Ramuhalli
Isabella Hu*
Andrew Bulmer*

*Sapere Consulting, Hydropower Research Institute Operating Contractor
[†]HDR, formerly with ORNL

January 2022

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831-6283
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Hydropower is one of several types of generating assets that provides energy, capacity, and services to electric power systems. It does so under rubrics and objectives—market driven and regulated, internal and external to asset and fleet owners—that address reliability, cost, price, and, increasingly, flexibility of output. The aggregation of data from multiple hydropower units can provide insights into asset operations and maintenance practices and needs and assist in meeting hydropower objectives. This paper examines the concept and potential benefits of aggregating hydropower asset data—primarily supervisory control and data acquisition (SCADA) information—with examples of insights developed from data aggregated by the Hydropower Research Institute (HRI). Data aggregation as discussed herein, and as implemented by the HRI, extends beyond multiple units in a powerhouse and beyond multiple hydropower facilities in an electric utility fleet or river system. Examples of research and analytics from such aggregated datasets range from unit load dependency analyses to modeling sensor measurements to detect and diagnose anomalies in assets. These examples showed the benefits of utilizing the entire dataset for insights into how the sensor layout of a single unit or set of units compares to the hydropower industry overall. Such insights include whether additional sensors are needed to complete analyses or to make decisions. In addition, utilizing multiple sensors of the same kind within a unit can provide an indication of possible current or upcoming problems with equipment. Although other analyses are possible, their use requires the development of complex models and, potentially, access to types of data that are currently not included with the example dataset used in this study. However, the examples studied herein confirmed the value of data aggregation in the fleet and unit contexts, and the value extends beyond multiple units in a powerhouse and beyond multiple hydropower facilities in an electric utility fleet or river system. The assessments also provided insights into potential extensions to the data aggregation concept that could further add to their value to the hydropower community, and these are included in this document as a set of recommendations.

1. INTRODUCTION

Hydropower is one of several types of generating assets that provides energy, capacity, and services to electric power systems. It does so under rubrics and objectives—market driven and regulated, internal and external to asset and fleet owners—that address reliability, cost, price, and, increasingly, flexibility of output. Because they contribute to aggregate production, revenues, and expenses of an entire fleet, most hydropower assets are managed in a fleet-wide context. As constituents of the fleet, hydropower assets are subject to multiple policies, including some applicable to all facilities, such as site security and health and safety measures. However, individual assets are also subject to decisions that differentiate roles (power and water dispatch patterns) for assets and allocate scarce or finite resources in non-uniform ways for operations, maintenance, and rehabilitation to each asset. Even in the case of fleet-wide policies or uniform allocations, the responses and outcomes for individual assets are likely to be different because of differences in operations and maintenance histories. This requires individualized consideration when the issue is the potential for improvements to policies, practices, and decision-making.

This paper examines the concept and potential benefits of aggregating hydropower asset data—primarily supervisory control and data acquisition (SCADA) information—with examples of insights developed from data aggregated by the Hydropower Research Institute (HRI). Data aggregation as discussed herein, and as implemented by the HRI, extends beyond multiple units in a powerhouse and beyond multiple hydropower facilities in an electric utility fleet or river system. The result is the collation of data from multiple fleets, typically with different owners, into a rigorously structured and standardized database to enable research and analytics that can inform policies, practices, and decision-making across multiple fleets. This paper begins with a presentation of hydropower analytics concepts that are necessary for a definitive and rigorous discussion of how hydropower data can be aggregated, analyzed, and interpreted across multiple fleets. It then describes how these concepts are implemented by the HRI, provides multiple examples of analysis and modeling using HRI data, and provides recommendations for future efforts and research to enhance the value of the HRI dataset and activities.

1.1 HYDROPOWER ANALYTICS CONCEPTS

1.1.1 Hydropower Fleet Ontology

In the hydropower domain, the hydroelectric unit is the fundamental asset. It is defined in this document as an entity that can be dispatched or controlled separately from other assets to convert the energy of water flowing through it into electric power output. Multiple units comprise a fleet, but even within the asset management domain, *fleet* has several possible definitions[1]. The most general definitions articulate that "[a] fleet's units must share some characteristics that enable [one] to group them together according to a specific purpose"[2] and that "there needs to be a [common] interest to consider a group of assets as a fleet"[3]. The implication of these definitions is that when the term fleet is invoked, it should be accompanied by an indication of the characteristics, interests, and purposes that establish membership in the fleet. It will be helpful in the hydropower domain to replace the terms *purpose* and *interest* with *context* and refer to the environmental, power system, regulatory, or other context for a hydropower facility. With the definition of *ontology* as the "set of concepts and categories in a subject area or domain that shows their properties and the relations between them" (Oxford English Dictionary 2021), one may understand *hydropower fleet ontology* as the set of concepts and categories that show how characteristics and contexts are used for classification and study of the relations between hydroelectric assets. Hydropower fleet ontology is a special case of a taxonomy that addresses (i) the relationships between the parts or components and the whole of an asset (Section 1.1.2) and (ii) the relationships of identicality, homology, and similarity between two or more assets (Section 1.1.3).

The distinction between characteristic, context, and response (or outcome) is important in analyses and management of asset fleets and the data they produce for multiple reasons. In this work, the term *characteristic* implies a static physical property of an asset (usually a design parameter), whereas the term *context* implies a geographical, hydrologic, environmental, regulatory, economic, legal, or other policy-relevant feature that distinguishes an asset. Responses (or outcomes) are the dependent experiences produced by design, installation, operation, and maintenance of assets—typically evident or discerned as energy production, availability of power system services, normal and abnormal sensor readings, reliability and failure statistics, and operation and maintenance costs, among other indicators.

Characteristics are relatively static compared to contexts because significant effort, expense, and time are usually required to modify assets in ways that change characteristics. Such changes are typically under the control and decision-making authority of the asset owner, although regulatory approval is required in most cases. Only a subset of asset characteristics data is available publicly; many other important characteristics needed for fleet-wide analytics are available only within the asset owner organization. Conversely, with few exceptions, contexts (or contextual data) are typically dynamic (time series

[1] S.-K. Kinnunen, A. Happonen, S. Marttonen-Arola, and T. Kärri, "Traditional and extended fleets in literature and practice: definition and untapped potential," *International Journal of Strategic Engineering Asset Management,* vol. 3, no. 3, pp. 239-261, 2019/01/01 2019, doi: 10.1504/IJSEAM.2019.108467.
[2] G. Medina-Oliva, A. Voisin, M. Monnin, and J.-B. Leger, "Predictive diagnosis based on a fleet-wide ontology approach," *Knowledge-Based Systems,* vol. 68, pp. 40-57, 2014/09/01/ 2014, doi: https://doi.org/10.1016/j.knosys.2013.12.020.
[3] Kortelainen, H., Hanski, J., Kunttu, S., Kinnunen, S.-K. & Marttonen-Arola, S. (2017). Fleet service creation in business ecosystems – from data to decisions. VTT Technology 309. 70 p. + app. 6 p.

exhibiting long-term variation and short-term fluctuation in response to geophysical, economic, or political drivers) or subject to changes and decisions outside the control of the asset owner organization. Contextual data, however, tend to be publicly available with varying accuracy and time-specificity. For example, daily time series of hydropower facility aggregate flow are often publicly available, whereas sub-hourly time series data for individual turbine flows are rarely publicly available. Operational history (which includes aspects such as number of start-stop cycles, runtime, and ramping events) is itself a non-public contextual feature of an asset and is determined by characteristics and other types of contexts. Operational history may also be considered a response of the managers, dispatchers, and operators of an asset to the combined characteristics, capabilities, and demands placed upon an asset.

Multiple characteristics and contexts of hydroelectric assets may define fleet membership, depending on the objective in delineating a fleet. Table 1 lists several examples of high-level criteria for fleet membership, such as "Reclamation Francis units" specifying the fleet of hydropower assets of turbine type Francis that are owned by the Bureau of Reclamation; or "High-head (rated head > x feet), non-federal (ownership) assets within Region 06 (Ohio River Basin)." An important extension of this ontology is the selection of fleet members based on more detailed physical equipment and design characteristics, as in "Kaplan units with greaseless wicket gate bushings" or "large (capacity > X MW) units with upper and lower guide bearings."

Table 1. High-Level Examples of Fleet Membership Criteria.

Fleet Criteria Examples	Example Values
Characteristics	
Turbine type	Francis, Kaplan, Pelton, pump-turbine, other
Dam type	gravity, arch, buttress, coffer, other
Rated head	low, medium, high
Capacity	small, medium, large
Contexts	
Ownership	Corps of Engineers, Bureau of Reclamation, Tennessee Valley Authority, Lower Colorado River Authority, federal, non-federal, municipal, investor-owned electric utility, independent power producer
Manufacturer	Newport News Shipbuilding and Dry Dock Co., Allis-Chalmers, Baldwin-Lima-Hamilton Corp., S. Morgan Smith Co., I. P. Morris, James Leffel & Co.
Hydrologic	Region 06 (Ohio River Basin), Subregion 0108 (Connecticut River Basin), Subregion 1702 (Upper Columbia in the United States)
Geographic and political	State of California; tribal lands; public lands
Electric power context	MISO, PJM, ERCOT, WECC, a specific balancing authority
Environmental	Species presence, water quality concern

The preceding discussion focuses on the hydroelectric unit as the asset of interest. In the discussions and examples that follow, the terms *asset* and [hydroelectric] *unit* used without qualification have the same meaning. However, the hydropower fleet ontology concepts and benefits discussed herein are also applicable to higher levels (powerhouses), lower levels (governors, exciters, transformers, circuit breakers, and other subsystems), and non-powertrain components (auxiliary generators, dewatering pumps) in the hydropower equipment hierarchy, each of which can also be referred to as assets. For example, it may be useful to define a "fleet of transformers" or a "fleet of servomotors"—with the hydroelectric unit to which each member component (a transformer or servomotor in this example) belongs being one of many characteristics and contexts included in analyses used to understand how response varies throughout a fleet of components.

1.1.2 Physical Equipment Hierarchy

The previous section describes how assets may be categorized according to characteristics and contexts. This section complements the ontology by defining the physical arrangement of a hydropower asset. *Physical equipment hierarchy* (Figure 1) describes how a hydropower asset is arranged physically, in part and in whole. For example, the physical equipment hierarchy allows one to specify that a blade is a part of the turbine, which is a part of a hydroelectric unit. The hydroelectric unit is a part of the powerhouse facility, which is a part of the hydropower site.

The specifics of physical equipment hierarchy may vary from one facility or asset to another: for example, Kaplan turbines include blade tilt mechanisms, whereas Francis turbines do not; diversion facilities include a penstock or tunnel, whereas short intake run-of-river facilities do not. Physical equipment hierarchies also may differ by convention: for example, a guide bearing may be considered to be a part of the generator, shaft, or turbine depending on its vertical position and the convention established by asset managers. Physical equipment hierarchies provide consistent terminology for asset characterization, assessments, and reporting. They allow for efficient and clear exchange of human and data-based analytical insights, as well as provide the physical and functional hierarchy upon which multiple sources of asset data are organized (Section 1.2). They allow data scientists, analysts, and hydropower professionals to exchange insights efficiently and accurately for complex analyses. The key point is that the physical equipment hierarchy of each asset included in an analysis should be defined, documented, and made unambiguous for effective asset fleet management and fleet data analytics, and it should be easily interpreted by humans but amenable to processing by computer algorithms.

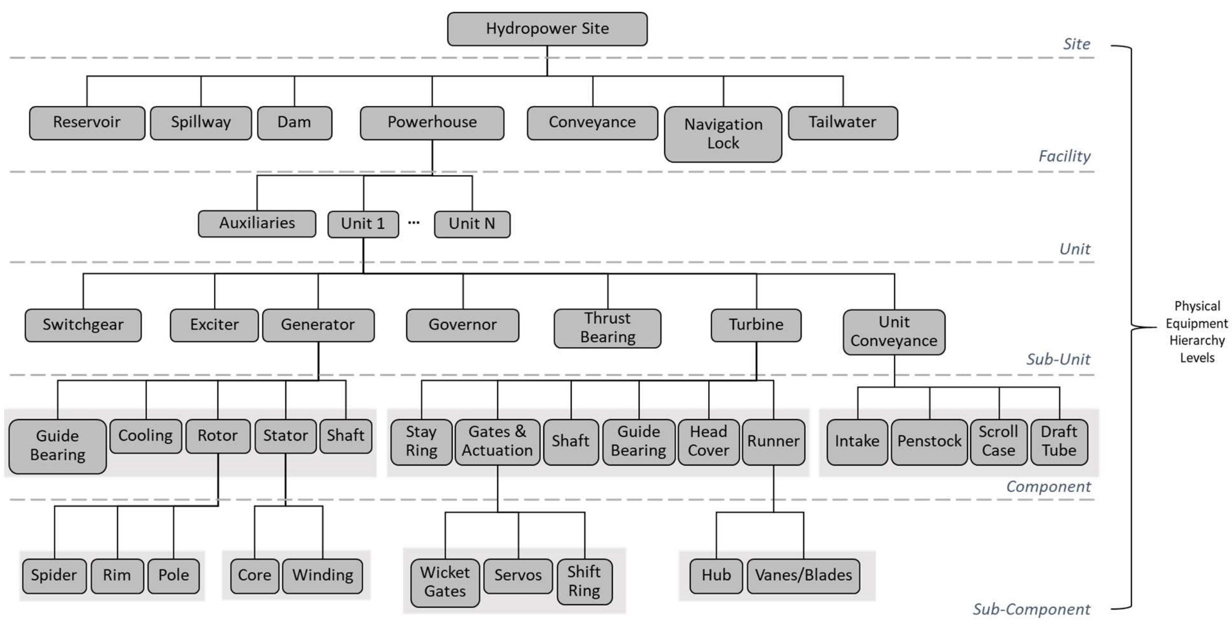

Figure 1. A partial example of physical equipment hierarchy for a hydropower facility.

More specifically, in hydropower data analytics, fleet-wide clarity of physical equipment hierarchy enables proper and consistent aggregation and correlation of characteristics, contexts, measurements, responses, and outcomes for hydropower assets. Discernment of patterns of response across fleets of hydroelectric units depends upon knowing that equipment occupies the same or a different position in the physical equipment hierarchies of the units being compared. Assessments of risk (for example, potential failure mode analyses), root-cause analyses, and reliability-centered maintenance all depend upon clarity of parent and child relationships of subcomponents, components, subunits, and units. This clarity enables top-to-bottom disaggregation and bottom-to-top aggregation of runtimes, start-stop cycles, risks, and other cumulative parameters throughout the physical equipment hierarchy.

1.1.3 Similarity of Hydropower Assets

As fleet-wide databases grow in their coverage of units and facilities and begin to incorporate multiple sources of information (Figure 2), systems specialists, analysts, and decision-makers will require a shared understanding of sameness between hydropower assets. Achieving the objective of the data aggregation discussed in this document—more insights of greater value through analytics that are possible only with multi-unit data—requires that the responses and outcomes of two or more assets (hydroelectric units) and equipment within the physical equipment hierarchy, as well as the characteristics and contexts assumed to influence those responses, be compared and contrasted. This section addresses the basic question of how to represent the extent to which two or more assets are the same. Specifically, this section proposes a mechanism for comparing two or more assets based on the degree to which their characteristics *and* contexts are the same—hereinafter referred to as *similarity*. Similarity as defined here does not rely on the characteristics of the asset alone (i.e., not just on the design and configuration) but includes the context of unit operations. Characteristics are mostly associated with the design of assets and do not reveal, by themselves, context. Responses and outcomes, however, are influenced by the combination of characteristics and contexts. Thus, it is possible that two hydroelectric units with exactly the same characteristics (design) have different contexts and may thus exhibit responses that are not the same. Thus, in seeking comprehensive insight and prognostics about how and why assets respond, it is insufficient to specify that two assets share the same or different characteristics; their contexts must be compared as well.

The concept of similarity, as described here, allows systems specialists, analysts, and decision-makers to determine the degree of similarity based on commonly used and accepted quantities (characteristics and contexts) and ranges between two extremes. One extreme is the scenario where two or more assets have exactly the same characteristics *and* contexts; this represents the case where the assets may be considered *identical*. The opposite extreme is where the assets are completely dissimilar, with different characteristics and context. In practice, one would expect most assets to fall somewhere between these two extremes, and the case of identical assets may be considered as an idealization for the purposes of analysis.

A related concept is that of homologous assets. These refer to the case where one asset may be a scaled version of another asset, but both are otherwise identical. Scaling may be in either the characteristics or context, or both. The homologous relationship between asset characteristics is most often encountered in turbine design and testing, wherein it enables comparison and prediction of prototype performance and response based on scaled physical model testing.

In a general sense, models and prognostics developed from unit-specific data can be expected to be valid and useful for a set of units that are exactly the same (identical, or complete similarity). Conversely, those models and prognostics would be expected to be inappropriate for units that are significantly different (very little sameness). For units that are similar (somewhat alike and sharing only some characteristics and contexts) but not identical, additional detail on which characteristics and contexts are the same or different would be required to understand the extent to which data-driven models and prognostics would be applicable amid the differences. The expectation is that the data-driven models and prognostics would be applicable through a set of scaling factors when dealing with homologous assets.

Table 2 expands upon this concept and formalizes it within hydropower fleet ontology by providing a preliminary set of rubrics for *identical*, *homologous*, and *similar* hydropower assets based on the hydroelectric unit characteristics and contexts. Several aspects of Table 2 warrant reiteration or additional explanation.

- The similarity of assets has two aspects: (1) sameness of characteristics and (2) sameness of contexts.
- Sameness of characteristics does not always correspond to sameness of contexts. This is the situation for powerhouses with multiple units of identical design and characteristics—the so-called "family" of units. Strict identicality of context requires the exact same operational history, maintenance record, and other contextual indicators. This is seldom the case, even for adjacent units in a powerhouse. In most cases, **the term *family of units* means a group of assets within the same powerhouse that have identical characteristics but only similar contexts.** As the examples that follow show, explaining the differing responses of assets within a family requires consideration of how their contexts diverge from identicality into similarity or beyond to dissimilarity.
- The homologous relationship is included in Table 2 in anticipation of analyses of assets with characteristics that differ only in scale and not in shape. Even when the corresponding characteristics (primarily geometry) of two assets are not strictly homologous, there may be utility in comparing their physical asset hierarchies (Section 1.1.2) to determine that their structure and hierarchy are the same. As with model and prototype, two or more homologous assets in the field are likely to have at least some responses, mostly kinematic and hydrodynamic in nature, that are very appropriately analyzed using theoretical relationships calibrated with data from one of the assets.
- Conversely, even non-homologous, dissimilar units and subunits may have similar or identical components and subcomponents that could be analyzed with a shared model of performance or reliability calibrated from aggregated data. This is a hypothesis that the HRI dataset can help to test in the future.
- The homology of contexts is notional at present and will require further development to be useful for hydropower assets.

Table 2 is preliminary because there are few available data, even within the HRI dataset, with which to validate the rubrics for a robust and quantitative assessment of how characteristics and contexts differ (or not) for two or more units and determine their sameness. As more data from more assets become available in the HRI and other datasets, the rubrics and classifications of Table 2 can be tested and refined. The authors anticipate a general methodology that includes (1) identifying a potential set of assets for analysis, (2) tabulating and comparing the characteristics of the assets in the set, (3) defining and computing (from asset-specific data) statistics that describe the contexts of the set, (4) using the rubrics of Table 2 to assess and classify the sameness of assets in the set, (5) defining and computing statistics (from asset-specific data) that describe the responses of each asset in the set, and (6) asserting and testing hypotheses about the consistency of responses from subsets of units that are identical, homologous, or similar, based on their individual characteristics and contexts.

Table 2. Preliminary rubrics for identicality, homology, and similarity among hydropower assets.

Relationship	Characteristic	Context

Similar Assets
 	
Assets with similar design or contextual features
	Similar characteristics:
	• Geometries are not strictly identical or homologous but have the same types of features
	• Turbine type is the same
	• Similarity must be considered throughout the physical equipment hierarchy and linked to the physics and mechanisms of concern:
	  ○ number of vanes, gates, and blades
	  ○ intake and draft tube bays
	  ○ stator and rotor designs
	  ○ shaft and bearing designs
	  ○ structure, hierarchy, and logic of controls, subsystems, and components
		Similar contexts:
		• Shared maintenance practices
		• Similar maintenance histories
		• Similar operating histories
		• Similar dispatch strategies
		• Similar water management contexts:
		  ○ reservoir residence times
		  ○ hourly, daily, or seasonal flow release patterns
		  ○ facility flow fractions to each unit
		  ○ spill, sluice, and bypass timing and fractions

Homologous Assets
 	
Scaled versions of assets that are otherwise identical
	Homologous characteristics:
	• Corresponding geometries (shapes, dimensions, tolerances, and roughness) in all three spatial coordinates have the same linear scale ratio
	  ○ Identical number and arrangement of vanes, gates, blades
	  ○ Homologous intake and discharge geometries
	• All angles are the same within the geometries of the assets
	• Homologous bearing and shaft designs
	• Homologous rotor and stator designs
	• Identical structure, hierarchy, and logic of controls, subsystems, and components
		Homologous contexts:
		• Orientation of assets with respect to the surroundings must be the same.
		• Homologous setting above or below tailwater
		• Assets have the same kinematic and dynamic ratios (e.g., Froude number)
		• Homologous water management contexts:
		  ○ hourly, daily, or seasonal flow release patterns
		  ○ facility flow fractions to each unit
		  ○ spill, sluice, and bypass timing and fractions

Identical Assets
 	
Assets that are exactly the same
	Identical design and configuration:
	• Identical ratings and specifications
	• Identical geometries throughout the assets (shapes, dimensions, tolerances, and roughness)
	  ○ Identical number and arrangement of vanes, gates, blades
	  ○ Identical intake and discharge geometries
	• Identical bearing and shaft designs
	• Identical rotor and stator designs
	• Identical controls, subsystems, and components
		Identical contexts:
		• Maintenance history and practice
		• Identical upstream/downstream hydraulics and water levels
		• Identical environmental conditions
		• Identical limits and constraints
		• Identical operational preferences and history
		• Identical setting above or below tailwater

1.2 SOURCES AND BOUNDARIES OF HYDROPOWER ASSET DATA

Optimizing the outcomes for a fleet of hydropower assets, or merely discerning and explaining the varied conditions and responses of assets, requires insights into how the characteristics and contexts of each asset influence the responses and outcomes for that asset. Such insights may be unrealized because the body of archived data for an asset remains unexamined or unavailable because of the absence of adequate sensors and data acquisition. There are also "human insights"—knowledge and experience of the operators, technicians, engineers, and managers who have operated, inspected, serviced, managed, and designed the asset. Such human insight is valuable input to operations and maintenance and asset management decisions, but it is limited in two respects: (1) it does not endure beyond the tenure of the staff that possess it unless there are effective mentoring and on-the-job training programs in place to capture and transfer knowledge; and (2) its use may require the staff that possess it to be present within decision-making forums to provide the information and logic that lead to the insights.

Quantitative and qualitative information recorded about an asset, including its characteristics and the context within which it operates, become "analytical insight" only if the data are amenable to automated retrieval and analyses. Manual intervention by a human to request, receive, align, and synthesize multiple data streams in an ad hoc way will be increasingly infeasible as the rate and volume of data available grow. Automated data synthesis and associated analytical insights extracted from it can be retained by an organization and shared easily within decision-making forums to complement human insight. Several distinct sources of recorded information exist for most hydropower assets: (1) design and configuration information, (2) operations or SCADA system data, (3) condition monitoring data, (4) inspection and testing reports, (5) maintenance records, (6) dispatch, forecast, and contextual histories for energy demand, ancillary services demand, meteorology, and hydrology, (7) production and performance reporting, (8) event logs, and (9) expense (cost) information (Figure 2). This categorization of recorded data by source is neither abstract nor academic; data from each of these sources are typically created, processed, and archived by distinct staff sections and IT systems within an owner organization. When the linkages and compatibility of these IT systems are not robust enough to support complex queries for insight within an owner organization, the sources are what Galar et al.[4] refer to as "information islands." With or without the challenge of information islands, analytical insights and their benefits do not follow automatically from recorded data. Effort and expense are required to retrieve, qualify, and collate data from IT systems and disparate sources, analyze the combined information, interpret the patterns that are discerned, validate the insights against similar assets and data, and communicate the insights to decision-makers.

[4] D. Galar, M. Kans, and B. Schmidt, "Big Data in Asset Management: Knowledge Discovery in Asset Data by the Means of Data Mining," in *Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015)*, Cham, K. T. Koskinen *et al.*, Eds., 2016// 2016: Springer International Publishing, pp. 161-171.

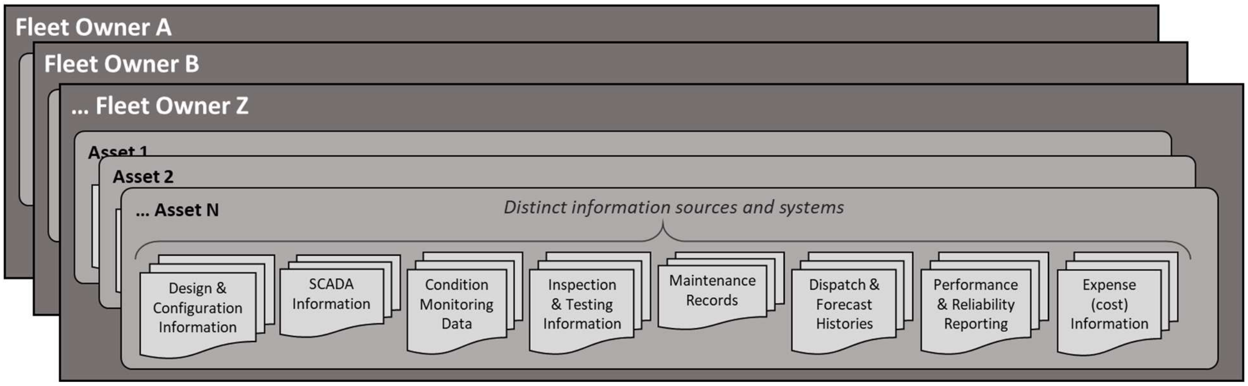

Figure 2. An illustration of the universe and typical boundaries of data for hydropower assets.

To this end, few would argue with a best practice that maintains and improves a robust body of data for each asset as well as uses that body of data to inform and justify decisions about how and when to operate, service, rehabilitate, and upgrade the asset. Furthermore, some owners of multiple assets have found value in centralizing individual asset data into internal databases, so that the data are readily available when needs for analyses and decisions arise. Efficient and coordinated operations and maintenance practices typically prioritize organizing information by source: for example, all maintenance records for all assets are maintained in the work order system for facility managers and maintenance supervisors; performance and reliability reports for all assets may be prepared, submitted, and archived by central engineering staff. Such single-source databases, in addition to the workflow they support, do have value in hydropower analytics (even if they are information islands), but they can be of greater analytical value when integrated (linked by asset identification number) across multiple sources. Linkages, integration, and the ability to automatically access the data through those links can enable analysts and algorithms to efficiently and systematically correlate potential causes and effects, evaluate and summarize tradeoffs between costs and benefits, and examine potential outcomes of alternative strategies for allocating resources and activities across hydropower fleets.

2. THE HYDROPOWER RESEARCH INSTITUTE DATASET

HRI aggregates measurement data from multiple hydroelectric units residing in multiple facilities owned by multiple utilities and agencies (Figure 3). The impetus for this structured aggregation of data is to enable and accelerate research, insights, and technology development contingent on the availability of empirical data from many similar (and dissimilar) subcomponents, components, subunits, and units.

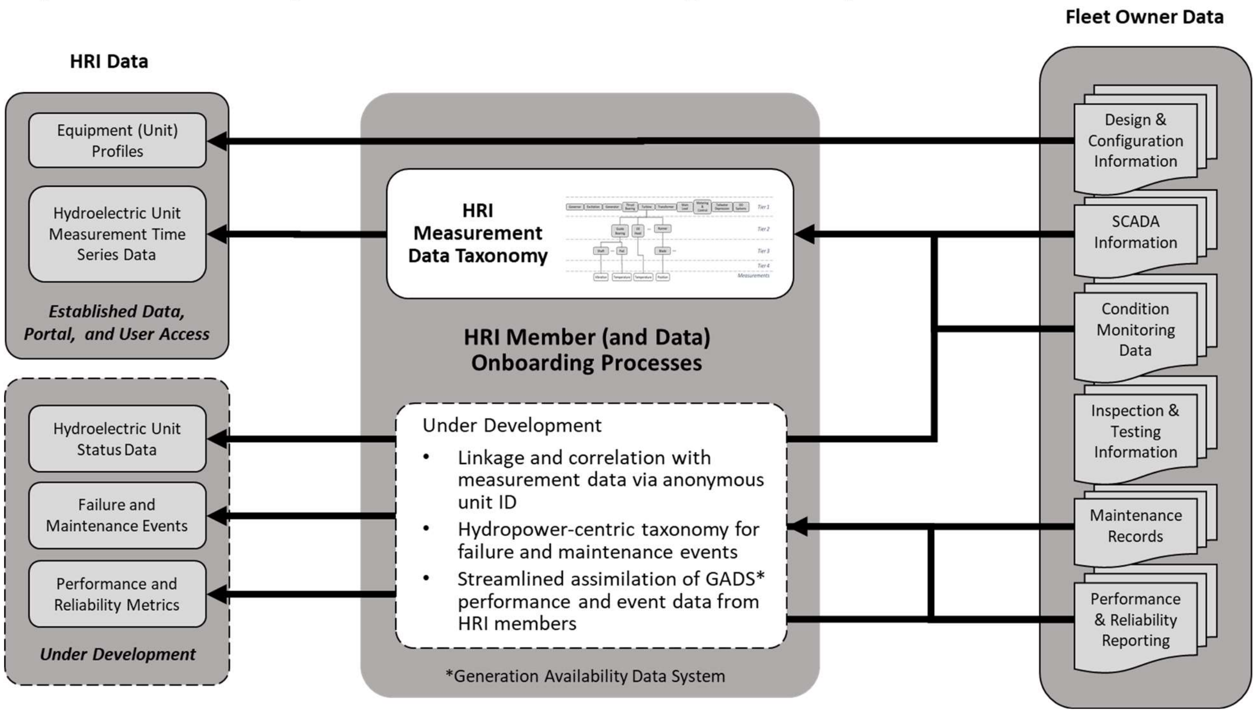

Figure 3. The relationship between typical hydropower asset data systems and the HRI dataset.

2.1 HRI DATA STRUCTURE

HRI anticipates that its program and products of data aggregation will enable research and development, with at least two practical outcomes:

1. Capabilities of hydroelectric units to respond or be adapted to changing power system demands and values ("flexible operations"), as well as the impacts of doing so, are discerned and generalized over time and fleet-wide through analyses of consolidated data, ultimately enabling forecasting of capabilities and optimized scheduling of services to power systems.
2. Factors that most influence subcomponent, component, subunit, and unit reliability are discerned and generalized over time and fleet-wide through analyses of consolidated data, ultimately enabling greater reliability and reduced costs through equipment selection and optimized maintenance activities.

In this context, the aggregation of data that HRI accomplishes is an ongoing process of receiving, standardizing, cataloging, and archiving data from multiple contributing members. In its initial phase of development, HRI is gathering what are most aptly termed *hydroelectric unit measurement time series data.* These data, which are a mixture of contextual and response data for hydroelectric units (primarily the powertrain portions), are the core of the HRI dataset and are typically extracted from SCADA systems and condition-monitoring systems, and they are augmented by equipment profiles (typically design characteristics, using the terminology of Section 1.1.1) pertaining to each unit in the dataset. HRI establishes rules and permissions for HRI members and users to search and retrieve these data systematically in support of hydropower asset analytics. HRI does not yet systematically aggregate event and statistical response information (for example, failure and maintenance events, performance and reliability metrics), although HRI can facilitate requests for such data from individual members to augment analyses of the core time series data. HRI does not collect broader contextual information (hydrologic time series, for example) that are not available within powerhouse SCADA and condition monitoring systems. Hydropower facilities and units are assigned unique and anonymous identifiers in the HRI dataset. Identity information, such as facility name, owner, and location, remains unknown to HRI data users, but the unique identifiers can be used to link measurements to units and units to facilities within the dataset so that contexts and responses from multiple units within an anonymous facility can be correlated.

Anonymization is necessary to preserve the proprietary status of HRI member data. Although the aggregate flow, headwater elevation, and tailwater elevation data and power production statistics of most hydropower facilities are available publicly (apart from HRI) in non-anonymized form (typically at a daily or monthly time step to support public uses of rivers and streams), the unit to subcomponent level data that HRI aggregates from members are proprietary and non-public. These data in non-anonymized form are valuable for predicting future operational status, energy production, capacity, ancillary services, failures, costs, and needs for maintenance and repairs within the competitive markets and commercial procurement environments in which HRI members operate. Thus, it is necessary and appropriate for HRI members to retain control of non-anonymized unit- to subcomponent-level data to preserve their business value and competitive position. This necessary anonymization of HRI member data does limit the capability of HRI users to correlate asset response with the broader hydrologic and power system contextual information that HRI does not collect for each asset. Anonymized equipment profiles are available for assets in the HRI dataset and can provide some additional context for interpreting responses. With greater use of HRI data, there may emerge needs and justifications for hydrologic and power system contextual statistics that would be of high value and could be gathered and reported by HRI members in anonymized form for assets within the HRI data.

2.1.1 The HRI Measurements Data Taxonomy

As a consortium of hydropower asset owners seeking to aggregate operating data for greater insight, HRI requires a set of defined concepts and an implementable process for characterizing and categorizing the data it receives from its members. The *HRI Measurements Data Taxonomy* (hereinafter, *HRI Data Taxonomy*, or HDT) follows from these concepts. It resembles, in part, the physical equipment hierarchies described in Section 1.1.2. Unlike those hierarchies, which support operations and maintenance of individual or limited numbers of assets, the HDT must support a broader set of purposes focused on analyses and comparison of hydropower asset responses across many assets and fleets. Although the ultimate focus of HRI-enabled analytics is on enhancing the value and reliability of assets and components, insights for doing so come through analyses of measurements (observations) made by sensors installed within assets and components. Thus, the HDT is essentially a characterization and categorization scheme for measurement data rather than a universal taxonomy for physical components.

The concepts that influence the HDT structure arise from the needs of HRI members and users and focus on the characterization and categorization of HRI data to enable useful, credible, and reproducible comparisons of measurements. Thus, the characterization and categorization scheme of the HDT must:

1. enable identification (selection) of measurements from two or more assets or components that are appropriate for inclusion in a comparison, statistical summary, or model of a "user-defined fleet" (for example, according to user-defined criteria for similarity of characteristics and contexts as described in Section 1.1.3);

2. allow for differentiation of measurements so as to avoid inappropriate comparisons and the wrongful inferences that may result;
3. enable analysts unfamiliar with the physical structure or details of an individual asset that produced a measurement to use that measurement appropriately within a multi-fleet analysis (the nature and location of the sensor and measurement are unambiguous while remaining anonymous in terms of geographic location or asset ownership).
4. enable HRI members (with support from HRI technical staff) to map (correlate) and update measurement data for their assets consistently across time and across their fleet; and
5. be adaptable and expandable to accommodate the measurements, assets, and analytics priorities of future members of HRI while maintaining compatibility with existing data, documentation, and analytics tools stewarded by the HRI.

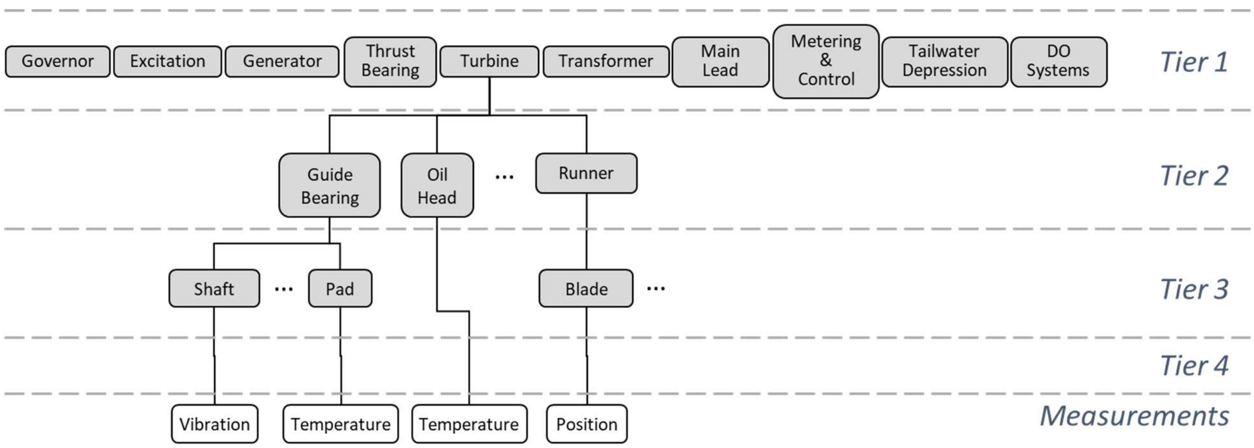

Figure 4. An excerpt from the HRI hydroelectric unit measurements data taxonomy.

The HDT (Figure 4) is reviewed and revised periodically to accommodate new data that become available from a variety of assets while preserving the capability to characterize measurement type and physical location within a hydroelectric unit. The HDT allows data to be mapped to specific asset locations in a way that is consistent across the industry, using tiered components to identify increasingly specific locations for, or categories of, data collection. The current Tier 1 categories of the HRI Data Taxonomy are (1) governor, (2) excitation system, (3) generator, (4) thrust bearing, (5) hydraulic turbine, (6) transformer, (7) main lead, (8) metering and control, (9) tailwater depression, and (10) dissolved oxygen (DO) systems. Tiers 2 through 4 allow measurement data to be mapped to subcomponents and more specific locations within the powertrain. The schema for each measurement within a dataset includes a measurement type (for example, vibration or temperature), sensor details (for example, accelerometer or thermistor), linkage to a component (and any equipment profile data available for that component), and location with respect to the geometry of the hydroelectric unit. The HDT scope includes the hydroelectric unit powertrain and measurements from balance of plant equipment or sensors, although balance of plant equipment and measurements are not currently available in the HRI database and are expected to be included in the future.

2.1.2 Equipment Profile Data

HRI data gathering from its members includes equipment profile information (configuration details) for hydroelectric assets. The focus of these data is the unit level of the physical equipment hierarchy, with greater detail for the hydraulic turbine subunit. Unit-level details include installation or in-service dates and out-of-service dates (when applicable). Hydraulic turbine details include manufacturer, turbine name, type (for example, Kaplan, Francis, or Pelton), number of blades, orientation (of the axis of rotation), pump parameters (for pump-turbine units), discharge diameter, rated speed, rated power, nameplate capacity, rated net head, and runner material. Ongoing efforts supplement these data with analogous details for generators (including rotor and stator details) and governors. Equipment profile data are linked to anonymous unique identifiers for individual units so that HRI users can correlate configuration information with unit measurement time series data.

2.1.3 Other Data

Consistent with member priorities and data availability, the HRI aggregates other types of data for preliminary analysis and potential future integration into the aggregated dataset available to HRI users. In this preliminary phase of data gathering, HRI must consider the extent to which such data can be standardized, characterized, and categorized to support fleet-wide analyses and insights. HRI is currently considering the feasibility of aggregating time series data of unit status (for example, on, off, starting, running, or unavailable), mode of operation (auto or manual), generation setpoints and limits under automatic generation control, and main breaker status (open or closed).

HRI does not yet systematically aggregate other types of hydropower asset data, such as inspection and testing reports, maintenance records, or reliability and performance data (generation statistics and failure, outage, and other significant events). These data are held by most HRI members in internal databases and are not available generally to HRI or HRI users. In cases where such data may contribute to explanation of and insight from results of HRI time series data analyses, users can request that HRI staff discuss with the appropriate members the need and benefit of sharing additional data to augment users' analyses. The additional data can then either be shared anonymously through HRI or, at the discretion of the member, shared directly between the member and the user.

2.2 PREVALENCE OF HRI DATA

2.2.1 Unit Size in the HRI Dataset Compared to the US Fleet

As of summer 2021, the HRI dataset included data from 160 hydroelectric units (including 3 units that had been decommissioned) at 34 geographic sites, with existing commitments from over 500 additional units. Compared to the entire fleet of over 5,300 hydroelectric units within the United States, a greater percentage of the 160 units within the HRI dataset are large hydroelectric units (Figure 5). There are multiple reasons why large hydropower units are more prevalent in the HRI dataset than in the whole of the US hydropower fleet. First, there are many very small (less than 1 MW) aging units in the US fleet that have almost no instrumentation and no data stewardship (see Section 4 for a recommendation to address this issue). Second, a large hydropower unit represents a greater capacity for risk from a single failure than does a small hydropower unit. Thus, a logical risk management strategy for hydropower asset fleets, which typically would employ monitoring to trigger inspections and maintenance to avert failures, should prioritize the limited resources of fleet owners on instrumenting large units before small units. Even without consideration of risk, the greater revenue attributable to large hydropower units is consistent with larger operations and maintenance budgets and staff resources dedicated to the unit (or a family of units), which is consistent with increased availability of sensors and the measurement data produced. Third, HRI membership, at present, includes several owners of large fleets with mostly large units and few small units. These large fleet owners typically have systematic fleet-wide programs and procedures for instrumentation, data acquisition, and analyses, although the focus of those programs and procedures is typically on installation and operational economies of scale and optimizing real-time operations rather than providing for fleet-wide trending and data analytics. It is likely that these large fleet owners can use their economies of scale to contribute data to HRI at a lower cost per unit and per measurement than can small fleet or single-site owners.

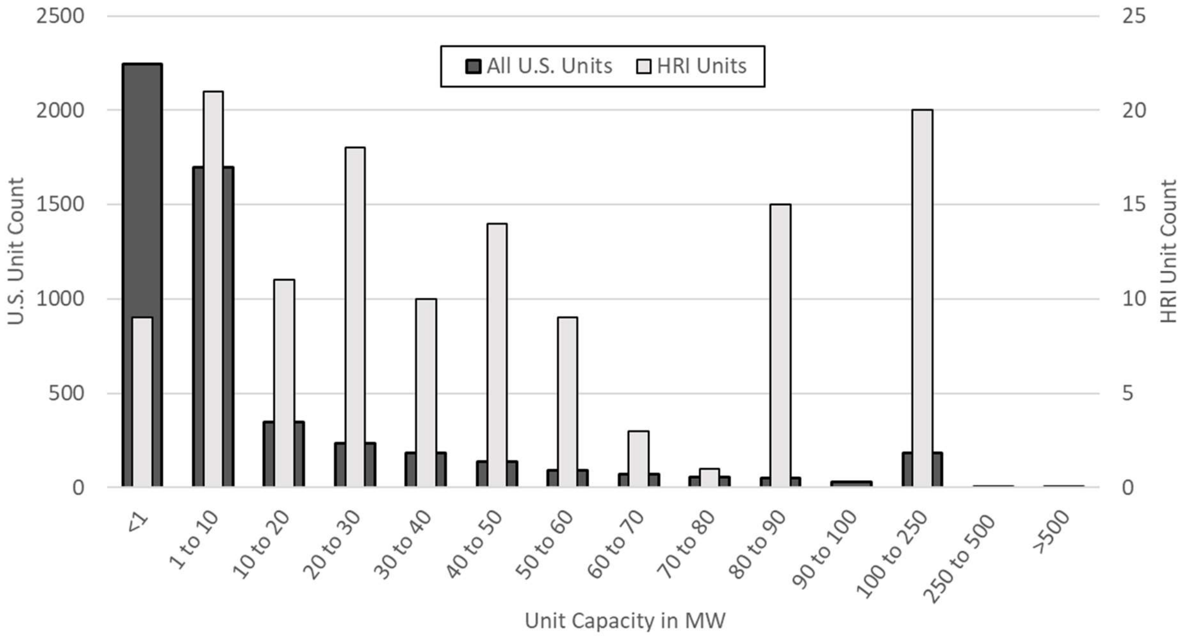

Figure 5. HRI unit MW capacity compared to US unit MW capacity.

2.2.2 Occurrence of Measurements in the HRI Dataset

The capabilities, resources, and opportunities to instrument hydroelectric units and curate the resulting data vary among fleet owners, including HRI members. There are 4,429 measurement points in the HRI dataset, but they are not uniformly distributed among the 160 units in the dataset. It appears that 15 to 20 measurements per unit is most common, but some units have almost no measurements, whereas a few have more than 100 measurements (**Error! Reference source not found.**).

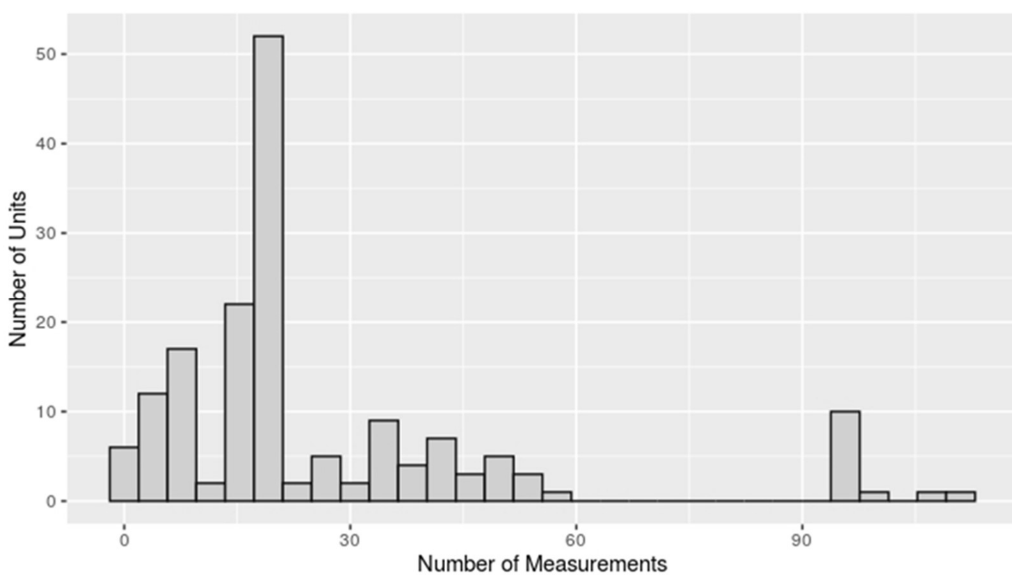

Figure 6. Histogram of measurement occurrence within the HRI data set.

Further detail on the distribution of measurements is possible by identifying and discussing three regions of the plot in Figure 6:

- **"Baseline" Units** (113 units reporting less than 25 measurements per unit): units in this category are the most common in the HRI database and represent an entry-level suite of measurements to be used in data analytics by HRI members and users. Table 3 provides a measurement schedule typical for units in the Baseline category. These measurements are available for most units in the HRI dataset.

Table 3. Typical measurement schedule for baseline units in the HRI dataset.

HDT Element	Measurement Details	Number of Measurements
Generator/Guide Bearing/Non-Drive End	Metal and Oil Temperature	2
Generator/Rotor/Field	Current and Voltage	2
Generator/Stator/Winding	Metal Temperature Ambient	1 to 20
Turbine/Headcover/Wicket Gate	Gate Position (Relative)	1
Metering & Control	Reactive and Real Power	2

- **"Intermediate" Units** (39 units reporting 25 to 60 measurements per unit): relative to baseline units, intermediate units exhibit an increased number of measurements in two regions of the HDT: (1) an increased number of Generator/Stator temperature measurements (typically 15 to 30 distinct measurements in the stator windings and other areas) and (2) an increased number of temperature measurements in the Thrust Bearing.

- **"Highly Instrumented" Units** (13 units reporting more than 90 measurements per unit)—relative to baseline and intermediate units, highly instrumented units exhibit (1) an increased number of measurements pertaining to the Governor and Turbine/Guide Bearing elements of the HDT and (2) additional measurements within the Metering & Control/Hydraulic Passage and Metering & Control/Runner Blade elements of the HDT.

2.2.3 Measurement Distribution Throughout the HRI Data Taxonomy

Figure 7 shows the relative prevalence of data in the four most populated Tier 1 categories and the Tier 2 elements within those categories. The statistics are grouped and compared by Tier 1 category (Generator, Metering & Control, Thrust Bearing, and Turbine). For example, the figure shows that approximately 60% of units in the HRI dataset report measurements for the Generator/Rotor element, whereas approximately 40% of units report measurements for the Thrust Bearing/Oil System element. The "N/A" category in each plot of Figure 7 shows the percentage of units reporting measurements that are associated only with a Tier 1 element and are not linked to any Tier 2 elements. For example, nearly all (almost 100%) of units report measurements associated with the Tier 1 element "Metering & Control" (but not linked to any Tier 2 element), whereas almost no units report measurements associated only with the Thrust Bearing (Tier 1) but not associated with either the Thrust Bearing/Bearing or Thrust Bearing/Oil System Tier 2 elements. The high number of "N/A" measurements in the Metering & Control category is due to the MW Output measurement that collected and stored for nearly all units, but it is mapped only to the Tier 1 Metering & Control category of the HDT and not to any lower tier elements.

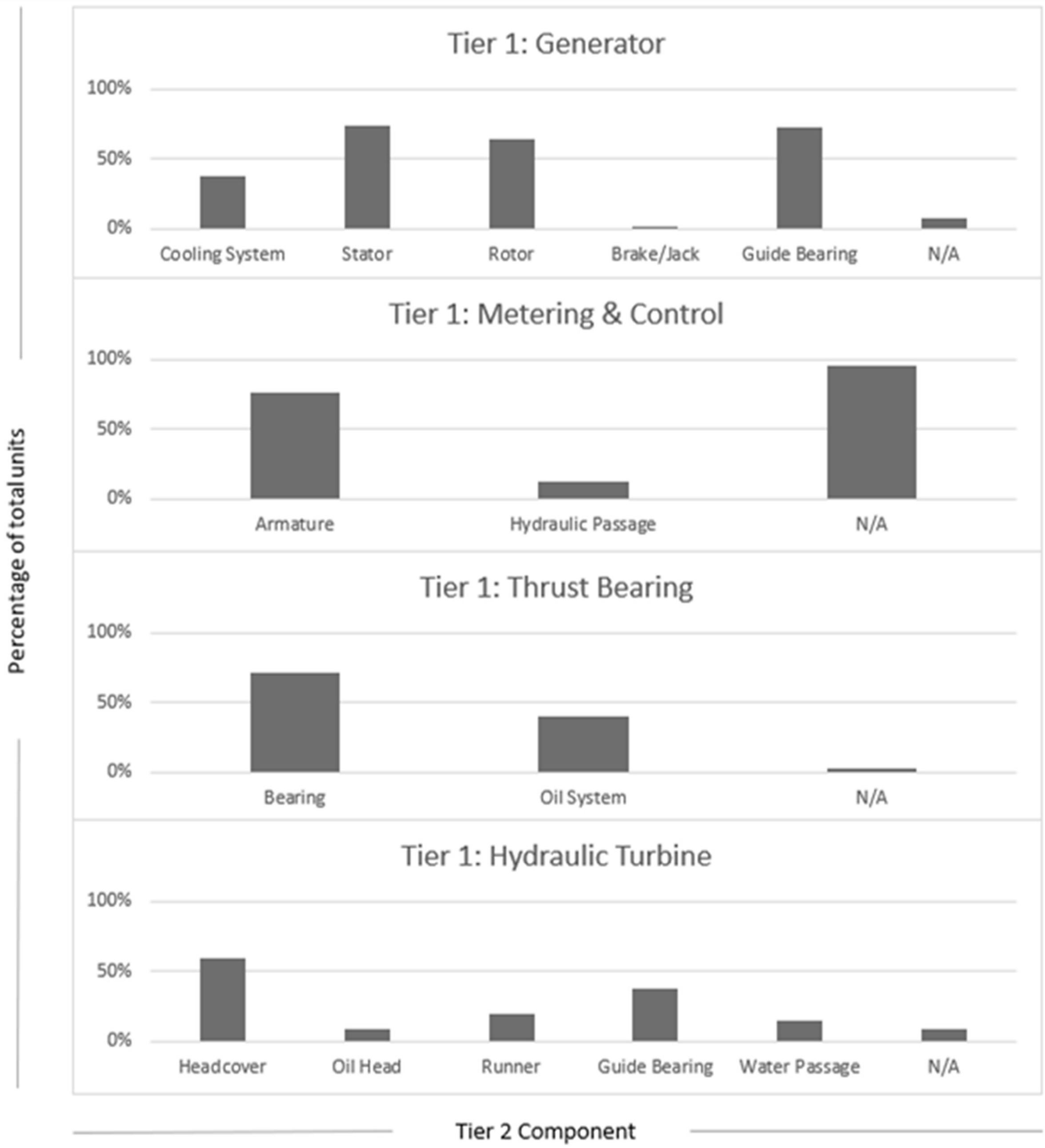

Figure 7. Data availability for selected elements of Tiers 1 and 2 of the HDT.

2.2.4 Measurement Distribution for Specific Use Cases and Equipment

The large number of measurements in the HRI dataset allows for comparison of the number of measurements to see which measurements are widely available and which are not. Figure 8 below shows a breakdown of how select measurements are mapped using the HRI taxonomy. This figure is not a comprehensive picture of the full HRI taxonomy; rather, it highlights common measurement types and includes specific examples used in later sections of this document. The numbers included on the diagram indicate how many sensors (which provide a single stream of data) are in the HRI for that measurement. For example, Figure 8 shows that thrust bearing pad temperature can be found in the HRI under Thrust Bearing/Bearing/Pad, and there are 276 sensors across all units providing that type of data. It should be noted that the HRI taxonomy does not require the use of all available tiers to map the data; Figure 8 also shows that there are temperature data mapped only to the Stator. This means the location of data collection was best represented at the Stator level rather than the available subsequent tiers.

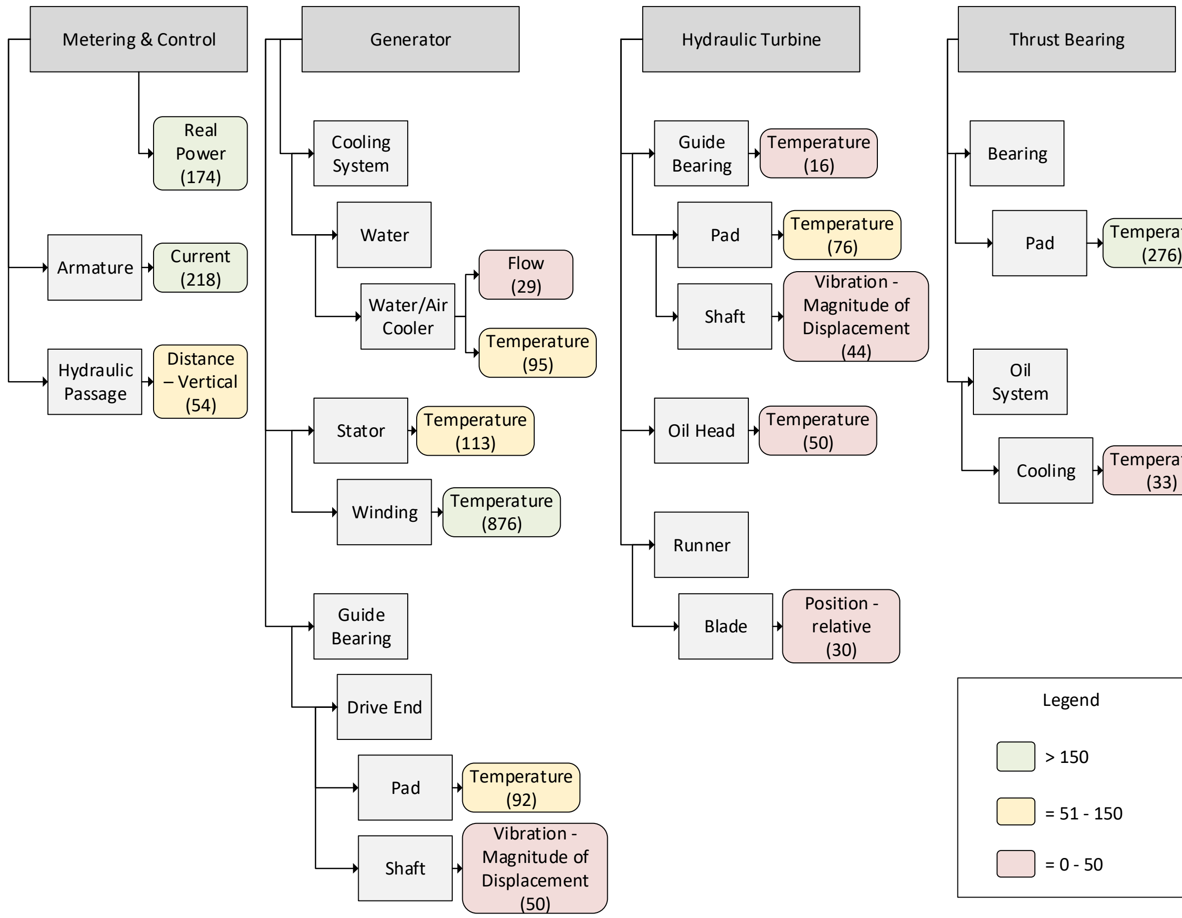

Figure 8. A diagram of HRI measurement data prevalence throughout portions of the HRI Data Taxonomy.

The HRI dataset allows a comparison of the maximum capacity of all the units. All units record megawatt output with 157 total measurements of the following HRI representation.

- Mapping: Metering & Control
- Measurement Details: Real Power, Gross

One of the most common (and important) temperature measurements is stator winding temperature, which is mapped into the HDT as follows.

- HDT Element: Generator/Stator/Winding
- Measurement Details: Temperature, Ambient, Metal

The HRI dataset includes 876 measurements for this element of the HDT, representing a significant portion (19.9%) of all measurements in the dataset and consistent with Figure 7, which shows that approximately 75% of all units in the dataset report stator winding measurements.

Figure 9 shows the availability of stator winding temperature data across the HRI dataset, sorted from left to right by the average of the top 5% of MW output for each unit (proxy for MW capacity). This demonstrates that most units have stator winding temperature data, and most have more than one sensor for this piece of data. The addition of MW capacity into the chart shows that in HRI data, there is **not** a clear trend in which higher capacity leads to more sensors. However, stator winding temperature and other such important measurements are common enough to be generally good candidates for standardized machine learning models of equipment and component operation and degradation, as demonstrated in Section 3.

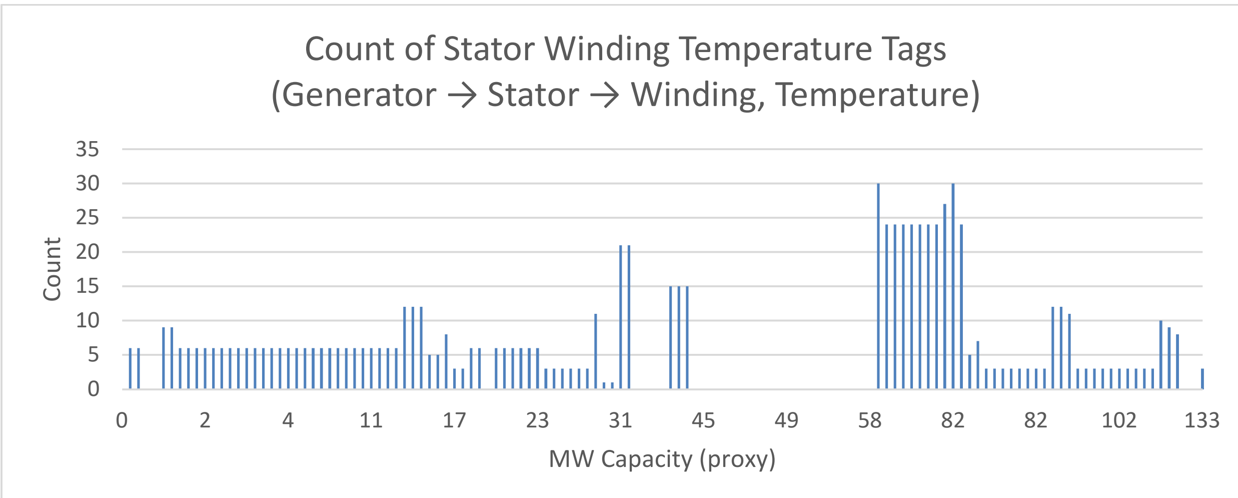

Figure 9. Stator winding temperature availability by unit.

3. ANALYSIS OF AGGREGATED MEASUREMENTS

An important use case for HRI aggregated measurements data from many units across the hydropower industry is provision of a robust context and norms for analyses of measurements from, and potentially anomalous responses of, specific units. Included here are examples of such context and analysis based on temperature and vibration measurements in the HRI database. The range of temperatures and vibrations throughout the HRI data provides the robust context for exploring the application of machine learning and pattern recognition algorithms to specific units. To the extent that the whole or subsets of HRI data are representative of hydropower fleets, families of units, or individual peer units of interest, they provide a basis or standard for comparison and, ultimately, prediction of responses. The sections that follow describe benchmarking analyses that can be performed using HRI data across families of units—those with similar characteristics and varying degrees of similarity among their contexts.

3.1 UNIT LOAD DEPENDENCY

One of the most immediate and valuable insights of HRI data aggregation is comparison of responses as a function of unit load among units within the same family and, ultimately, among similar units not in the same family or site. The example provided here uses stator winding temperature, which is an important factor in stator winding insulation lifetime (increased temperature duration is known to decrease stator winding insulation lifetime). This example is important and relevant for most hydroelectric units, as stator winding temperature and unit load measurements are available to some extent on most units (Figure 9). This example analyzes temperature response as a function of load, and it demonstrates how the shape of the temperature versus load fraction curve for a unit can be used as a more powerful indicator of similar or dissimilar response among units (one explanation for such [dis]similarity of response would be the relative [in]effectiveness of the generator cooling systems for the units being compared).

Figure 10 shows the average stator winding temperature of unit 36922-LO-44 (HRI reference identification) versus the load percentage relative to the overall dataset in blue for 2020 (HRI, 2020). Figure 10 also shows the range of the average stator winding temperature for a family of units at site 36922 plus or minus one standard deviation in the yellow shaded region. For the sake of comparison, Figure 10 also includes the average stator winding temperature for unit 36922-LO-44, which appears to be within the one standard deviation of temperatures experienced by this family of units. Figure 11 shows similar data but includes a different unit from the same site with a higher average stator winding temperature as a function of load fraction.

The types of comparisons shown in Figures 10 and 11 could be valuable, indicating that some units are exceeding the average stator winding temperature relative to load fraction. In this case, operational changes may be needed, or stator winding replacement planning may be expected to be sooner relative to the industry standard. This kind of comparison indicates that interested parties may need to complete further investigation into the causes of data falling out of normal range, allowing them to make operational decisions accordingly. These examples also demonstrate that the HRI dataset can be used to compare a family of units at a facility against similar units outside the facility. In this case, it can be observed that this particular family of units has a relatively high average stator winding temperature relative to load fraction compared to the overall HRI dataset (blue shaded region). These kinds of observations allow users to further investigate the causes of results that fall outside the typical range across the HRI set. The average stator winding curves could be separated by maximum capability in megawatts, which may explain the deviation in temperature curves versus load fraction.

Given the percentage of stator winding temperature measurements in the HRI data relative to other measurements, the HRI dataset also suggests that knowing stator winding temperatures are important to operations and maintenance planning. In this case, comparing single and multiple location temperature sensors (e.g., average temperature by location) could be advantageous for future analysis for making operational changes or identifying degradation in related components.

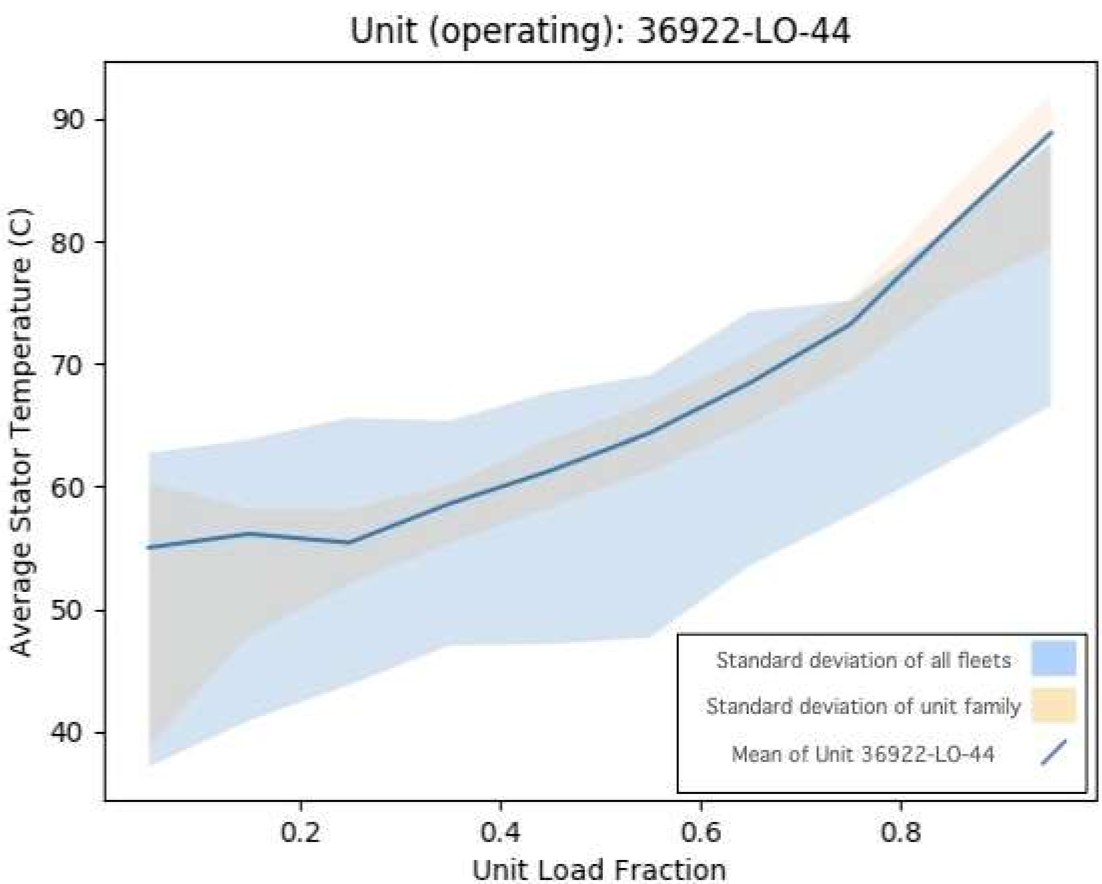

Figure 10. Average stator winding temperature of unit 36922-LO-44 relative to entire dataset.

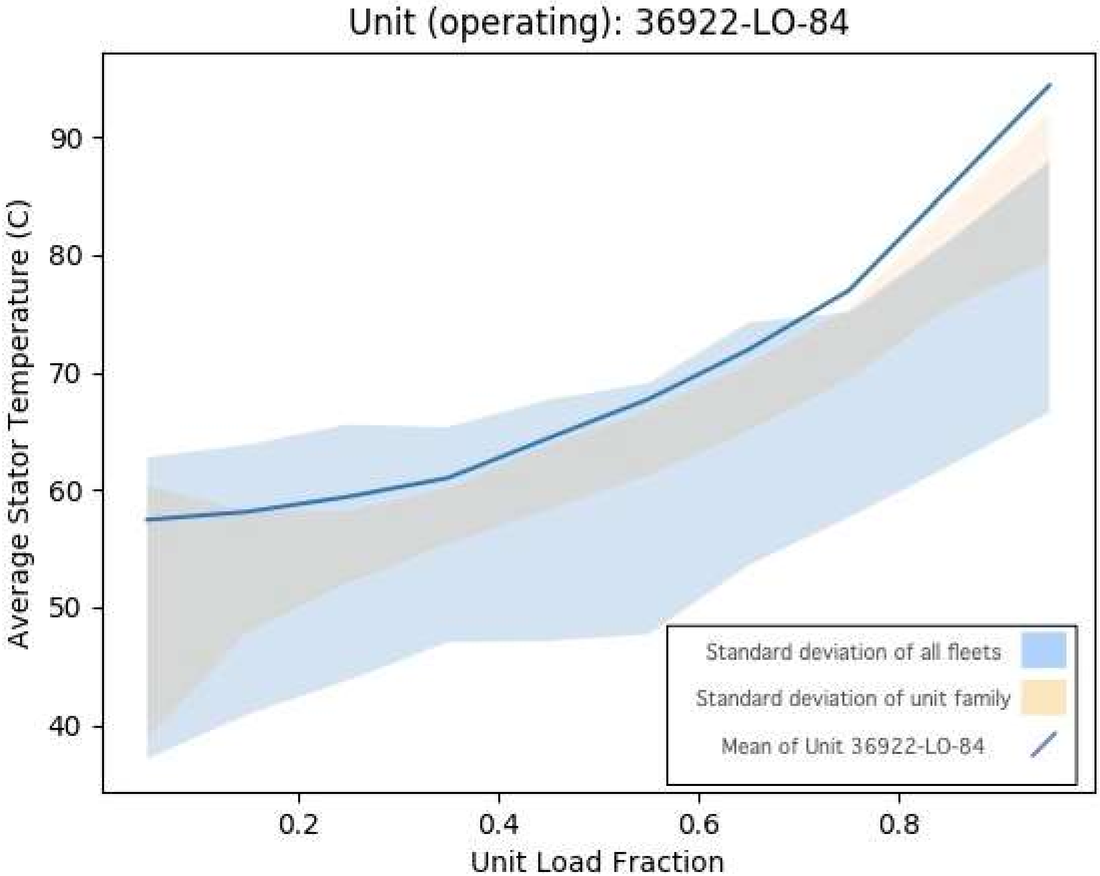

Figure 11. Average stator winding temperature of unit 36922-LO-84 relative to entire dataset.

3.2 REPLICATE AND ARRAYED SENSOR RESPONSES

Benchmarking across a family of units (those with identical characteristics and similar contexts) can be performed based on the relationship between multiple measurements of the same type and location. One example of those relationships is looking at turbine oil bushing temperatures. Figure 12 shows three oil bushing temperatures on a single unit from February to June in 2016, where there is virtually no difference between the temperatures. The traces cover each other in Figure 12 because there is little variability across the three locations (1A, 2A, and 3A) for unit 94140-JC-99.

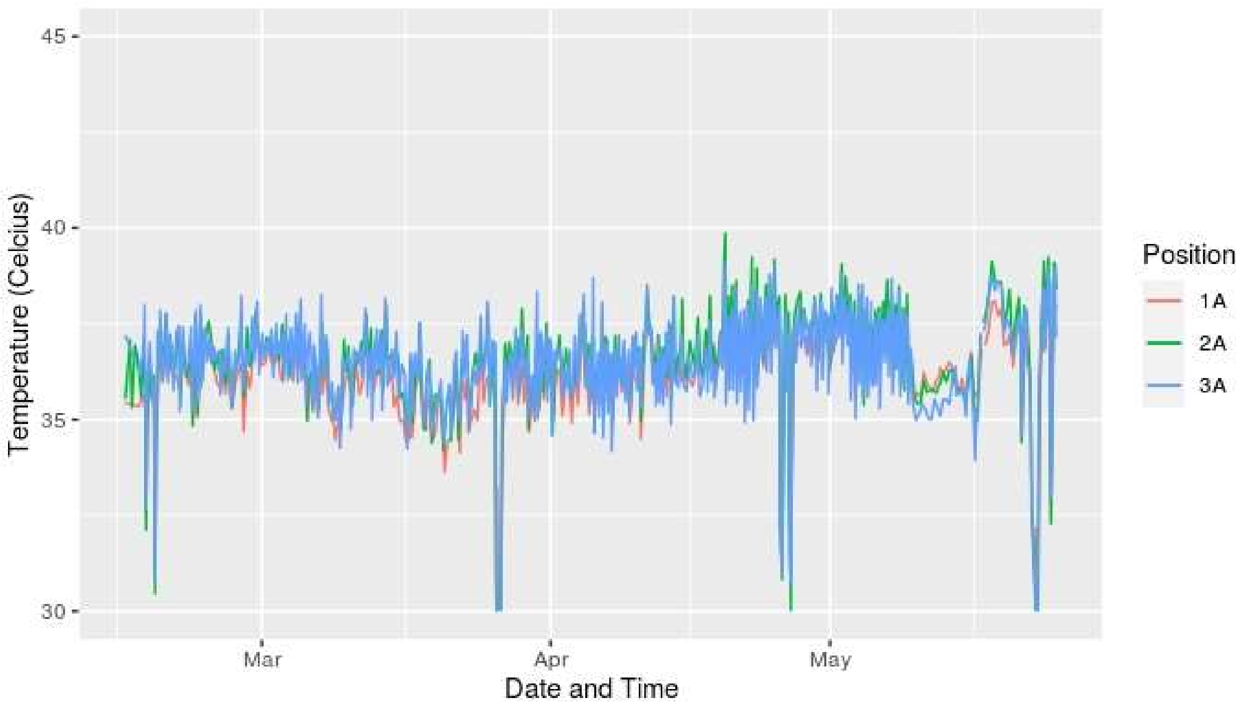

Figure 12. HRI unit 94140-JC-99 turbine oil bushing temperatures in 2016.

In 2017, unit 94140-JC-99 failed with no pre-failure alarms. Figure 13 shows the traces for 2017, with the unit failing at the end of the traces. Comparison of the data from the same period (February through June) in 2016 and 2017, location 2A (where the failure occurred), shows an elevated temperature in 2017 compared to the other two locations. This elevated temperature did not set off any alarms, but the data from these multiple variables show a measurable difference between these three locations.

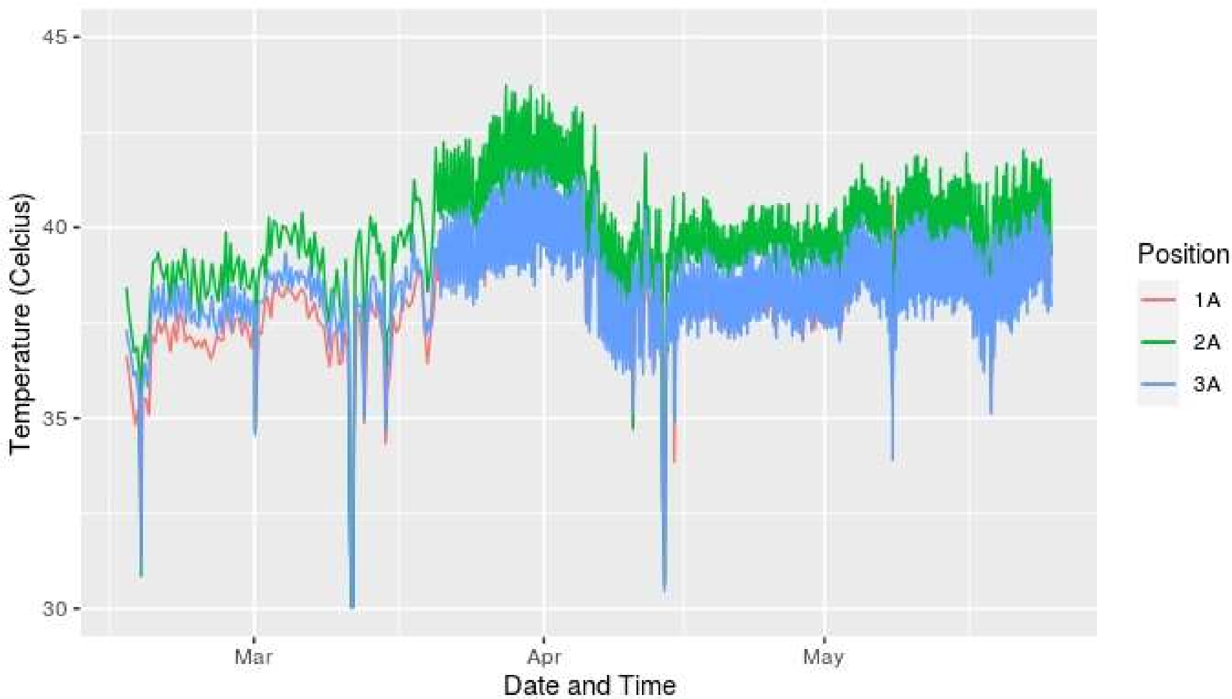

Figure 13. HRI unit 94140-JC-99 oil bushing temperatures in 2017 nearing the failure at the trace end.

Another unit in the family, 94140-JC-52 (Site-PH-Unit) shows a different set of signatures of the three oil bushing temperatures (Figure 14). Although location 2A in this unit shows a much wider range of temperatures with overall higher temperatures than those observed in Figure 12, the location 2A temperatures also move down into the range of measurements at bushing locations 1A and 3A. These measurements from unit 94140-JC-52 do not appear to be indicative of any upcoming failures, though the wider range of temperatures and higher overall temperatures compared to other units in this family point to a need for further investigation of their causes.

Such multivariate comparisons in the HRI dataset can establish numerous signatures to provide different views into what may be abnormal (and potentially detrimental) within those signatures across a family of units. Such comparison can be used later in anomaly detection by machine learning techniques, which is discussed in subsequent sections.

Extending multivariate comparisons across units with only similar (not identical) characteristics is more challenging and tenuous. There are fewer multivariate measurements that can be compared over the range of measurements per unit as shown in Figure 7. However, since stator winding temperatures are prevalent across most units, complex models of stator winding temperature predictions can be compared across similar units with cooling support measurements (water/air) and current (see Figure 5 for which units have these variables in the HRI). These comparisons could provide useful information and are discussed next. Future research with HRI data will be needed to establish metrics and thresholds for similarity among units that are necessary for valid use of these complex models.

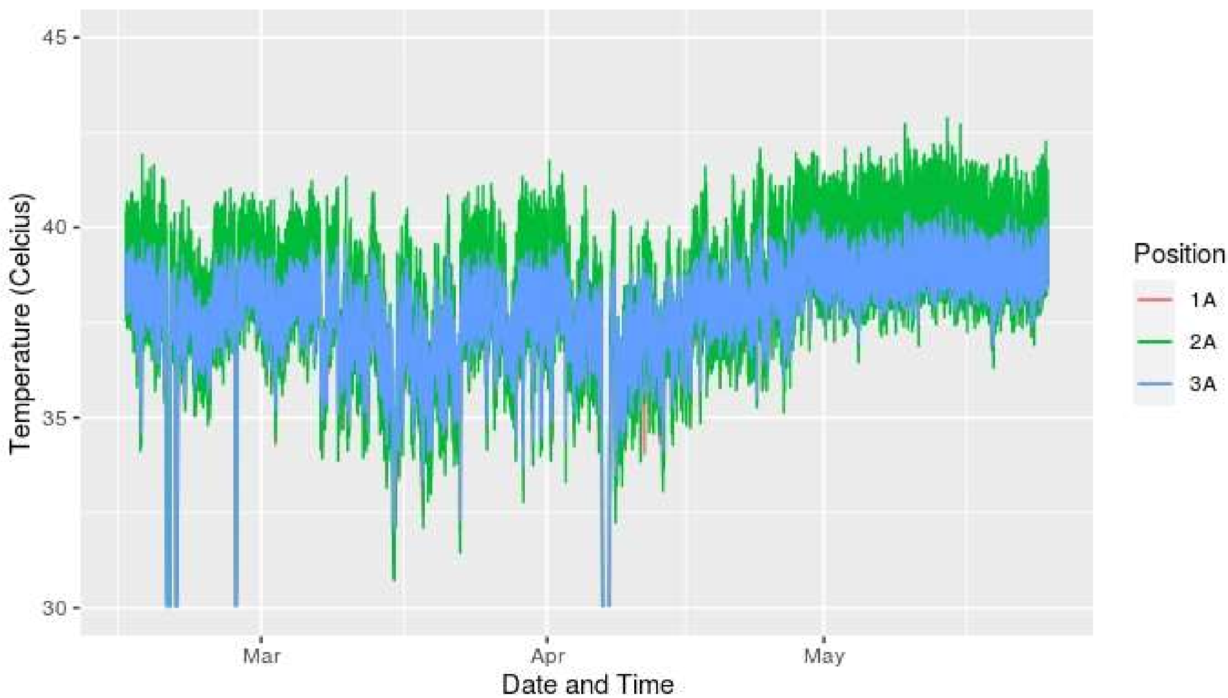

Figure 14. HRI unit 94140-JC-52 oil bushing temperatures for 2020.

3.3 MULTI-SENSOR RELATIONSHIP MODELING

Single sets of variables such as stator winding temperatures are generally limited in fully capturing the operational relationships that exist in a component (for instance, relationships among stator winding temperature and cooling air or water temperature and flow). The HRI dataset can allow comparison of more complex models that depend on several measurements to explain a dependent variable. The HRI dataset, with its aggregation of data from multiple sensors across the unit, allows benchmarking with multiple variables and analysis of the relationship between those multiple variables with single traces of multiple variables or complex models of multiple variables.

One such example is a stator winding temperature predictive algorithm (Yale, John [personal communication], 2020) that utilizes several inputs found in the HRI dataset to predict stator winding temperature across similar units. The dataset was used to develop two machine learning models, and the HRI dataset distribution of the input measurements point to whether such models could be applied across sites based on data availability alone. The results of using such a model can be used to benchmark one similar system to another to find unusual signatures or changes in operation. This model may be applicable across a broad range of units that are only similar and not identical.[5]

Developing such a model requires historical data covering a broad range of operational conditions. The HRI dataset includes data covering up to 14 years for some units. Using 12 years of data ending in 2018, a stator winding temperature machine learning model was developed, with the goal of predicting more recent (post-2018) stator winding temperatures. The model predicted average stator winding temperature from the inputs in Figure 8. Table 4 below shows the prediction accuracy (measured using mean absolute error [MAE] and root mean squared error [RMSE]) of this model for data from units within a family, as

[5] Machine Modeling System and Method, US Patent 20190279108-A1.

well as dissimilar units and the results of the overall model (using all the units' data) per unit. These results represent how accurately the model predicted the actual average stator winding temperature data from each unit after training the model on a combined dataset of historical data from all units. Some units were somewhat dissimilar from others, and maintenance or other anomalous events were not filtered out of the model training, leading to some variability in results across units. In this case, the operators made the decision that the error values fell within an acceptable range to utilize the model moving forward. The assumption is that the value of having a model that could predict temperature for 13 different units was high enough to justify the predicted value being off by a couple of degrees in the worst-case scenario.

Table 4. All Unit Stator Winding Temperature Model Metrics

Unit Identifier	Mean Absolute Error (MAE)	Root Mean Squared Error (RMSE)
01	2.36	5.04
02	2.05	4.77
03	2.91	6.45
04	1.84	2.75
05	1.70	2.61
06	2.01	3.96
07	1.54	2.43
08	0.77	1.34
09	0.24	1.01
10	0.36	1.17
11	0.59	1.74
12	0.23	0.66
13	0.27	0.79

Benchmarking these units together based only on the metrics in Table 4 provided information that Unit 03 contained a temperature sensor that was not functioning as expected, leading to a higher MAE and RMSE compared to the other units. The model metrics indicated a possible problem, demonstrating the value in using these results holistically to indicate the need to further investigate results that stray from the norm.

Identifying specific variables of interest allows for additional insight into the availability of data across the HRI. Figure 15 shows the availability of the inputs to the stator winding temperature predictive model. Most units have measurements for current, though very few units have all inputs needed for this model. The units with all available inputs come from only two sites. Further analysis could be completed to determine whether a subset of the needed inputs could lead to sufficiently accurate results. If it is assumed that all inputs used in this example are necessary to predict stator winding temperature (i.e., accuracy is lost by eliminating some of the inputs) then other sites interested in predicting stator winding temperature may need to add sensors to their equipment to capture the missing variables to compare the same model technique. However, additional multivariate models of the stator winding temperature may be possible that use a subset of these variables or other variables not considered in this initial study.

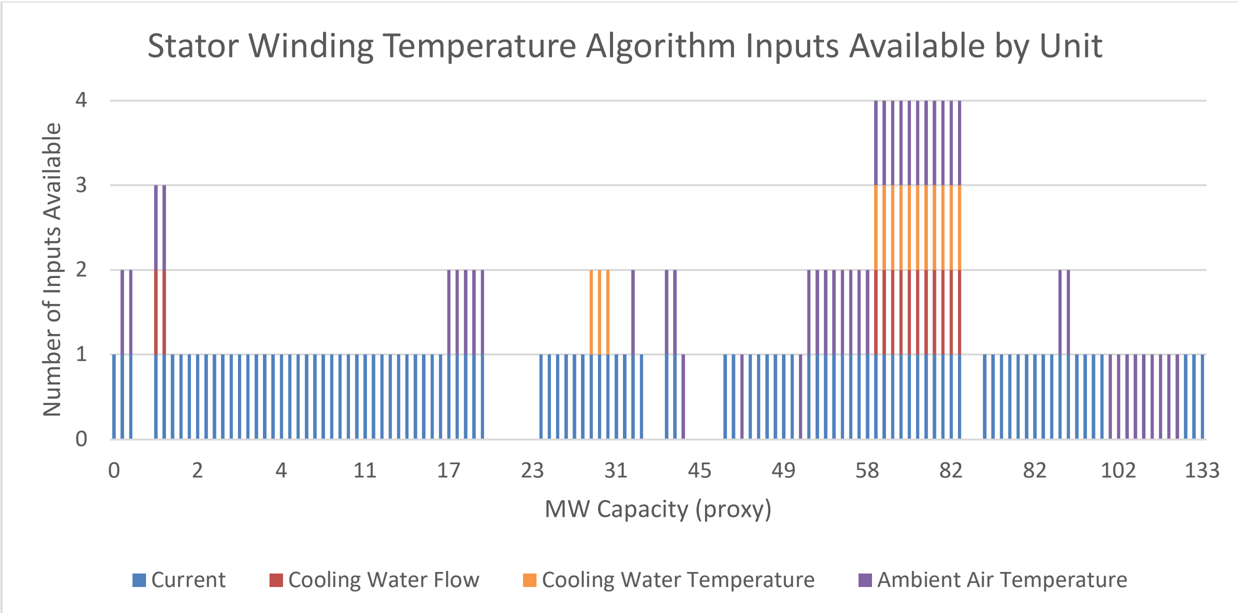

Figure 15. Number of available inputs for the stator winding temperature algorithm.

3.4 ANOMALY DETECTION

The HRI dataset provides several similar unit anomaly detection capabilities by providing a broader dataset for the hydropower industry. *Anomaly detection*, as defined in this section, refers to detecting where sensor measurements deviate from expected norms based on a historical record. Anomaly detection can be useful in identifying off-normal behavior of a unit and, if detected in advance of a failure, allow for mitigating actions. Anomaly detection may be performed using measurements from a single sensor or multiple sensors, and it can be achieved using statistical methods (simple anomalies) or complex models derived using machine learning. In all cases, the presence of a sufficiently long historical record of data is important. The HRI data goes back to 2010, so there are over 10 years for continuous conditions to establish historical expectations for most sensor measurements. In this section, anomaly detection from single measurements is discussed, followed by examples of anomaly detection using multiple measurements, and then those using a machine learning model.

Each anomaly detection strategy requires determining appropriate parameters that lead to a clear definition of what is considered “normal” data and identifying the conditions under which normal is defined. The three case studies below include operational and repair information that is not currently in the HRI dataset to provide further insight into what data may be considered normal. Longer term, the HRI strives to continue growing the dataset to integrate new types of data that provide deeper insights through analysis. Although some pieces of information (maintenance data, repair information, etc.) included in the sections below are not yet available in the HRI, members can request such information from participants in the HRI.

3.4.1 Single-Sensor Anomaly Detection

Anomaly detection is used regularly in monitoring and diagnostic (M&D) centers in the industry to identify data indicators that show a change from previous or “normal” observations. The HRI dataset enables advanced analysis of long-term single-sensor time series to detect anomalies. Figure 16 shows a single sensor time series (turbine guide bearing vibration) plotted in tandem with the result of an anomaly detection algorithm (the "anomalize" algorithm within the R software language).

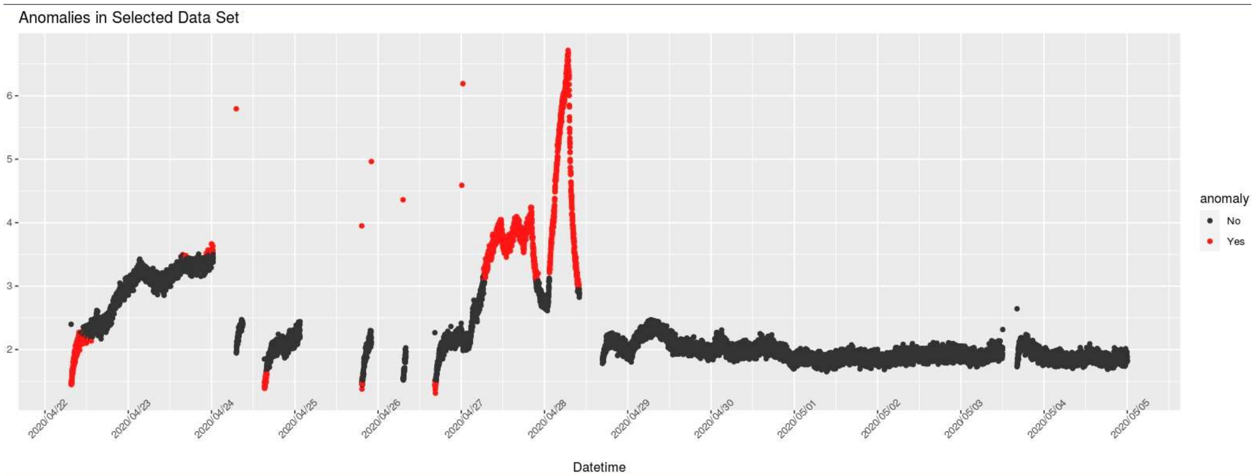

Figure 16. HRI turbine guide bearing vibration in x-axis mils for Site 36922, Unit LO-72.

Foremost in the plot is the "peak" anomaly occurring on April 28, 2020, and appearing in red because it was identified as anomalous by the algorithm. The event is obvious as an anomaly with only two weeks of record shown in the figure as baseline behavior, but the HRI dataset allows for inspection of such vibration data extending back several years for additional insight into historical baselines and trends. In this simple case, only two were chosen to show the immediate change that altered the trace because the event is linked in time with historical repair records describing head cover maintenance for Unit LO-72, for which the head cover studs were re-torqued on April 28 after the anomaly occurred. The post-maintenance time series does show that TGB x-axis sensor readings of 2 to 3 mils are typical, and that excursions of vibration sensor readings outside of this range that were occurring prior to the event and maintenance activity were also atypical. Following further experimentation, this particular anomaly detection result did not change with a longer historical record. Although the classification of the April 27–28 data as anomalous by R-anomalize algorithm did not directly trigger maintenance (other circumstances initiated the maintenance), it could have been used to do so. The head cover maintenance seems to have stabilized the TGB and reduced vibration, but it is alternatively possible that the vibration sensor alone was stabilized rather than the TGB component. The two possible conclusions are indicative of the limitations of these types of analyses: though data can be used to identify a component of interest, it is limited in the information that can be known using the component's available sensors. Further investigation into completed repairs would be needed to determine the precise effects of the head cover maintenance on the component vs. the sensor.

Other conclusions may be drawn from this type of analysis by overlaying other information that can also be found in the HRI dataset. Prior to the "peak" anomalous event, there are brief periods in which data are unavailable. MW Output (Figure 17) could be utilized to gain a greater understanding of when a unit is running, shut down, or in the process of starting up or shutting down, allowing for greater insight into the reasoning behind unavailable and/or anomalous data as well as confirming periods of data that may be considered "normal." In this specific example, it is clear that the periods with missing TGB vibration data correspond to times when the unit was shut down (no MW Output) with the unit generating within nominal parameters before and after the maintenance action in April 2020.

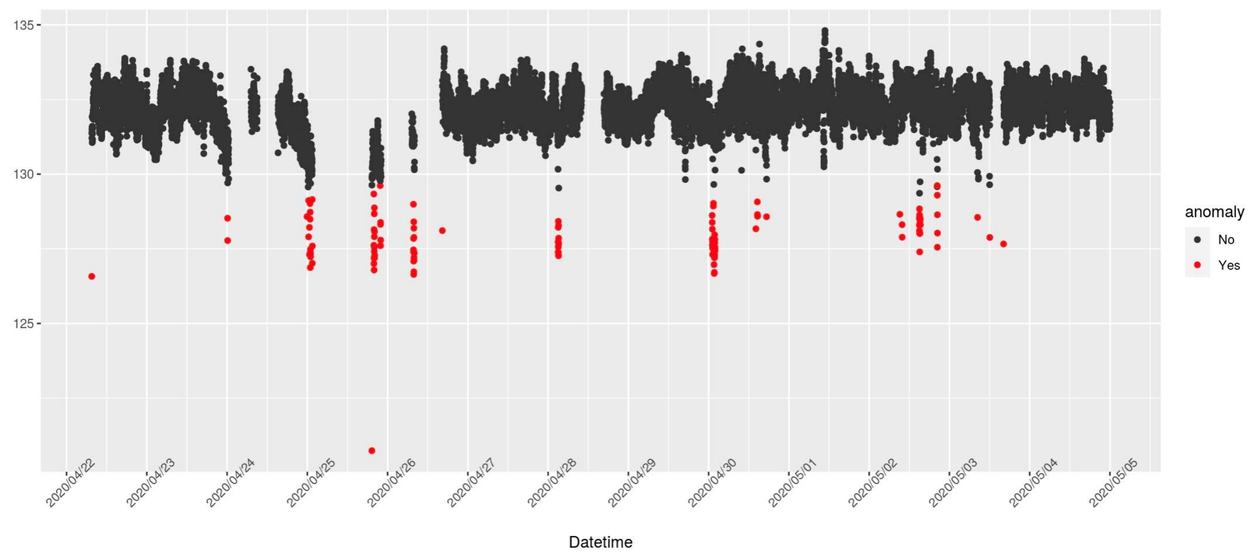

Figure 17. HRI megawatt output for Site 36922, Unit LO-72.

This case shows that the historical HRI data can be used to link to (yet unavailable in the HRI dataset) maintenance events and anomaly detection using off-the-shelf, open-source software tools. As HRI grows and maintenance data become available with outage data, linkages between events and signals identified as anomalies may be automated easily.

3.4.2 Multiple-Sensor Anomaly Detection Modeling

Analysis of time-synchronized multiple sensor time series from the HRI dataset enables a more sophisticated anomaly detection methodology. The HRI data often have the same type of sensor in multiple locations and across units. Finding anomalies in a combined dataset provides a more robust detection capability across identical, family, and similar units and possibly across sensor systems, and it allows for comparison across those measurements.

Figure 18 shows the use of anomaly detection using similar sensors across the turbine shaft that measures the metal temperature of an oil bushing. These bushing temperatures typically have a ratio at or very close to 1, as the deviations from each other are rather small. However, Figure 18 shows that the ratio of oil bushing temperatures deviated from the historical data in 2017. This happened before a catastrophic failure event (point of failure was at the end of the trace shown in Figure 18). The MW output for this unit shown in Figure 19 indicates that the unit appeared to be functional until the point of failure (data identified as anomalies in 2015 correspond to the unit being shut down, likely resulting in missing data). There was no alarm based on the single traces, but the leadup to the failure does show an anomaly for most of 2017 (Figure 18) when comparing the measurements from multiple similar sensors. The location of the point of failure is highlighted in Figure 20. This figure illustrates the key components surrounding the pipe at the point of failure, demonstrating why the bushing temperatures were indicative of failure.

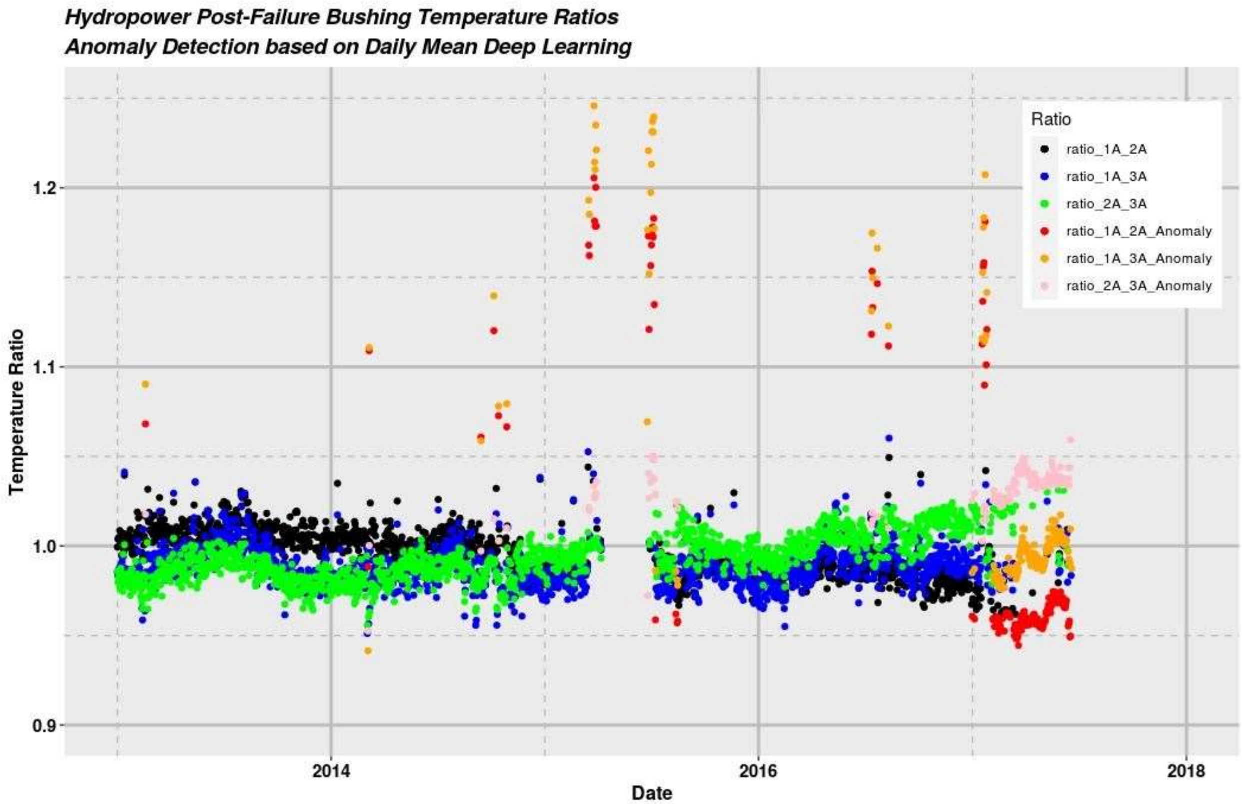

Figure 18. HRI oil bushing temperature ratio anomaly prior to a failure event at Site 94140, Unit JC-99.

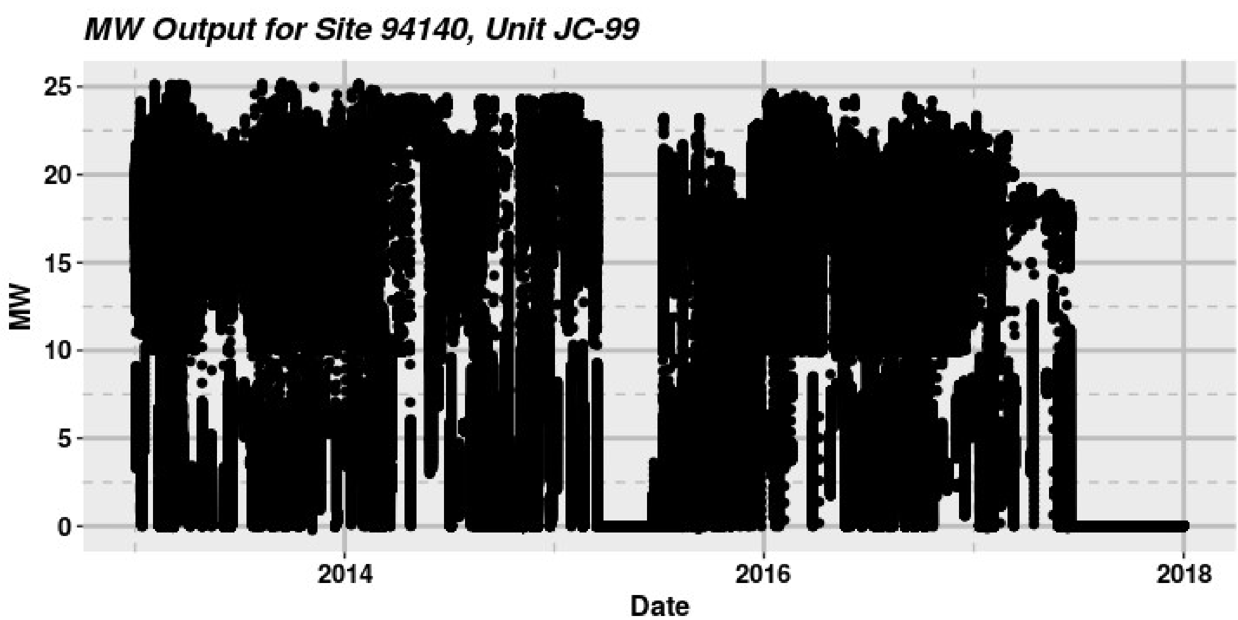

Figure 19. MW output for Site 94149, Unit JC-99.

In this case, anomaly detection could have played a unique role due to the long lead time, but the problem with anomaly detection is often that detected anomalies are not mapped to an actual cause or a pending event. However, the HRI dataset has multiple sets of families of units that have the same architecture and capacity at the same site. This may not seem useful for more complex models, but it is possible to use a similar unit with similar characteristics to train the model and then deploy it for anomaly detection. To test this hypothesis, a deep learning model was trained using data from a similar unit and applied to analyze the data from the failed unit.

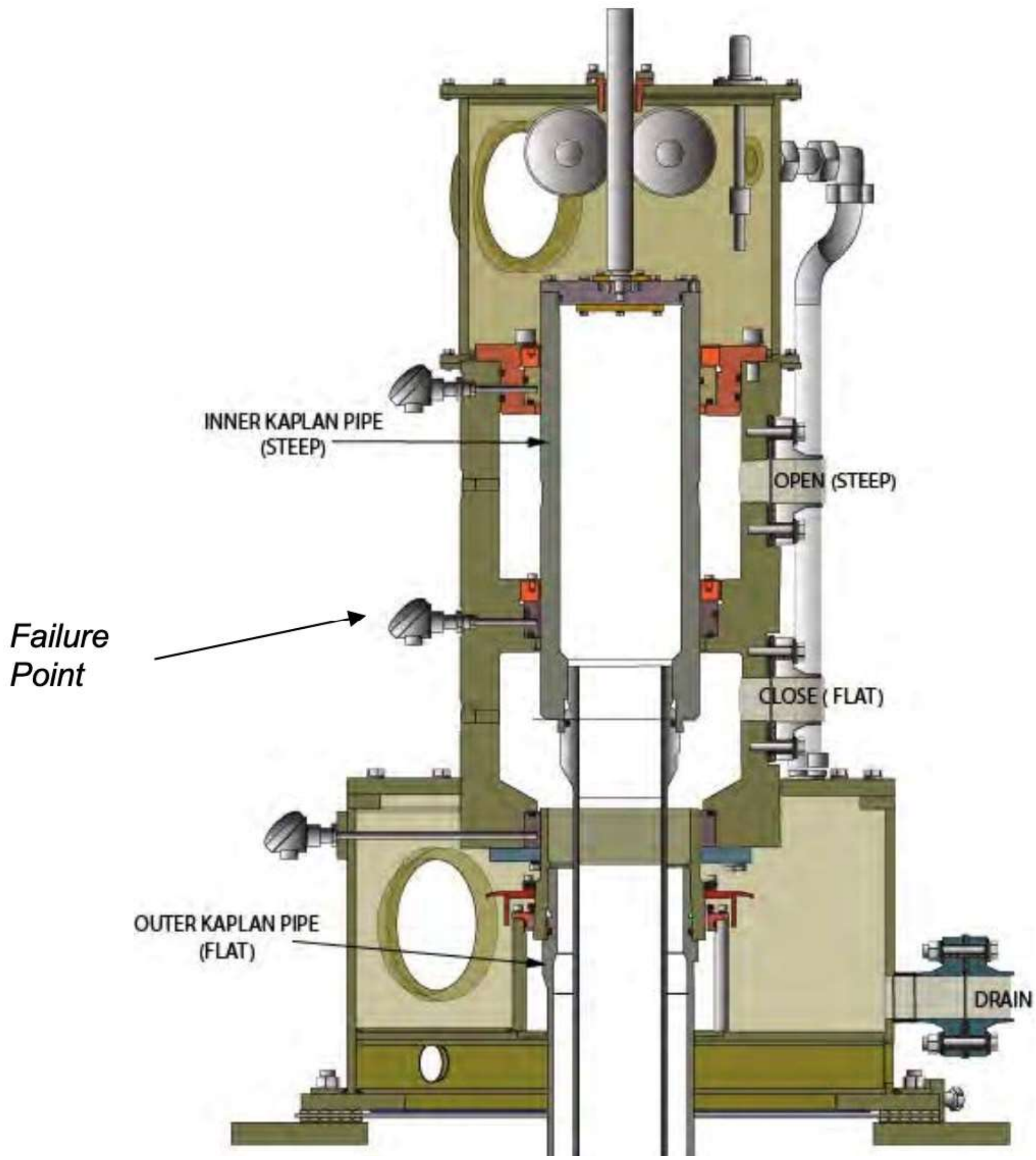

Figure 20. Illustration of failure point in Site 94140, Unit JC-99.

In this case, a neural network model was trained for the well-behaved (no failures) unit from the same family, JC-52, at the same site, 94140. The nearly identical unit was a rebuilt unit with well-behaved bushing temperature ratios and a relatively long history from 2019 to 2021. Figure 21 shows the daily average bushing temperature ratios for JC-52, and Figure 22 shows the corresponding MW output for that unit. The trained model was then used to diagnose the historical data from the failed unit to uncover the anomalies. This anomaly detection identification of the bushing temperature ratios appears to be a possible signature for similar types of units that will result in failure, though their use as a real-time signature of a precursor to failure must be confirmed with additional data sets.

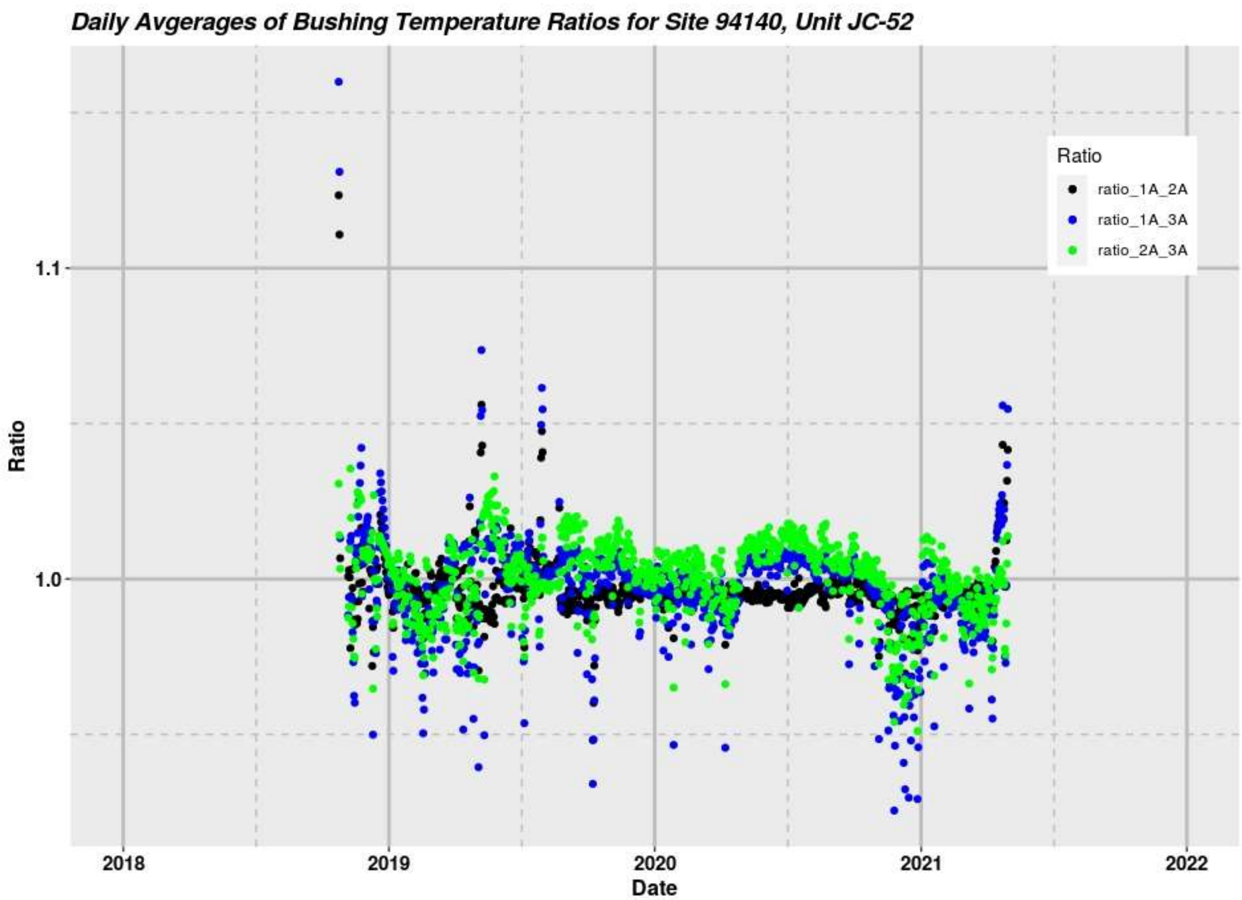

Figure 21. Daily average bushing temperature ratios for Site 94140, Unit JC-52.

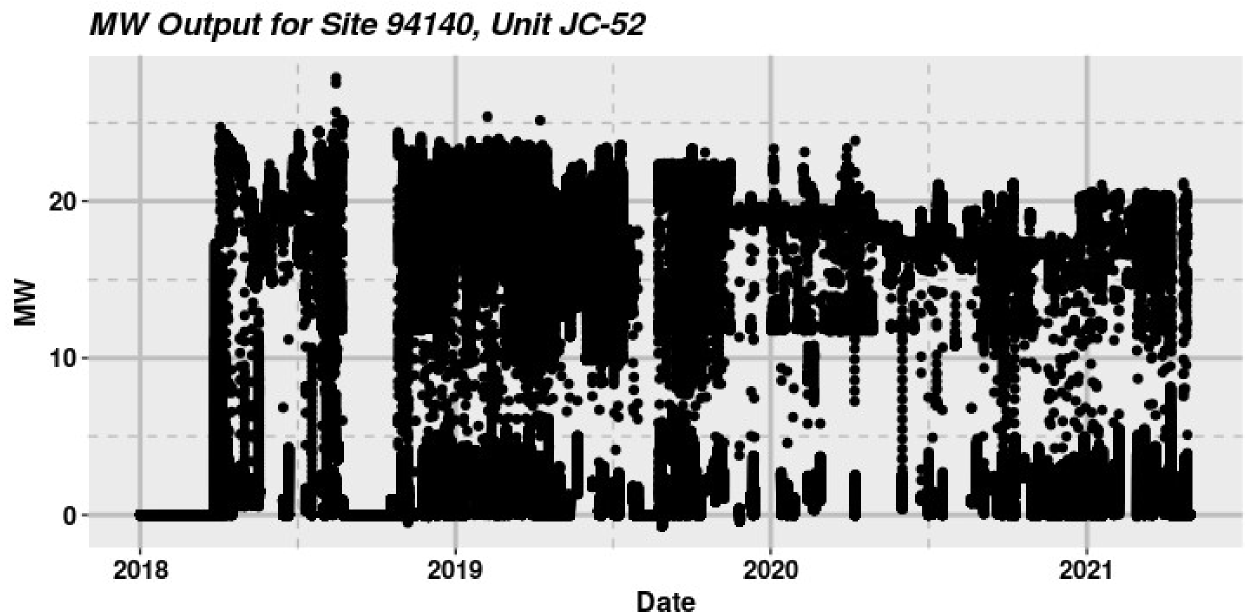

Figure 22. MW output for Site 94140, Unit JC-52.

3.4.3 Anomaly Detection with Machine Learning Models

One of the other methods to find anomalies enabled by the use of HRI's multi-source dataset is to train a model with multiple units that may or may not be from the same site. Earlier in the paper, the machine learning model was introduced to calculate the average stator winding temperature based on a short list of input variables and several stator winding temperatures. The HRI dataset has several units with extensive stator winding temperature measurements that can be used in such a model. The refenced model has been patented due to its accuracy.[5]

The figure below notes the HRI mapping criteria needed to obtain each variable used in the stator winding temperature analysis; this information can be used to extract each variable from the HRI dataset by selecting the relevant components for each tier. The figure shows the HRI structure and calls out where each variable is located. These measurements were selected because of their known correlation to stator winding temperature, per recommendation from engineers at the site used for developing the model.

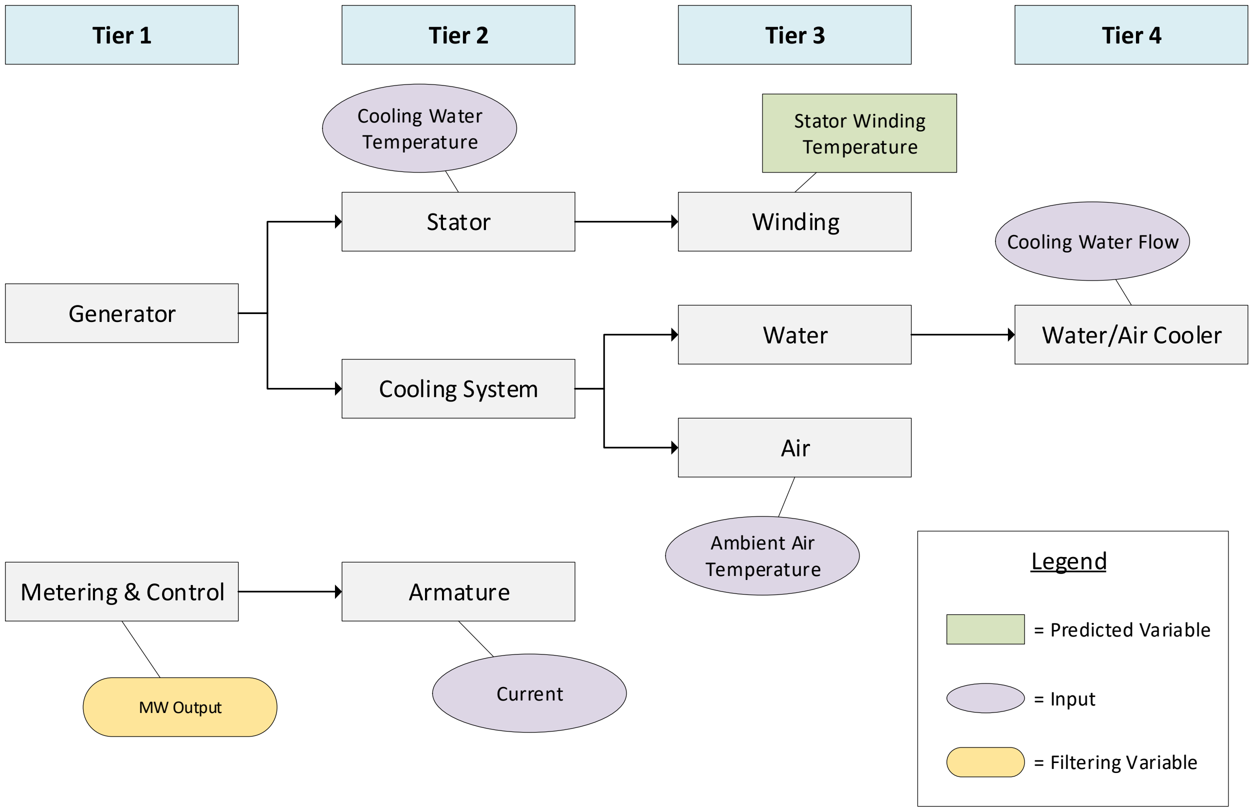

Figure 23. HRI schema for stator winding temperature model.

Seven units in the same family were used together to train a machine learning model to predict the average stator winding temperature based on eight years of historical data from 2010 to 2017. The cumulative model has average MAEs of less than 4°C. The trained model was then applied to 2019–2021 data to see whether the model-predicted values were still similar to the measurements.

Most the units still had a reasonable set of evaluation metrics (MAE: ~4–5°C), but a few units did not, and this led to an anomaly detection on these units manually to see where the biggest differences occurred. Figure 24 shows the measured and modeled average stator winding temperatures for site 36922 and unit LO-93, which had a MAE of 9.6°C. What is obvious in the figure is the discrepancy during the summer months in 2018 and 2019, driving the higher than expected MAE. In this case, the MAE value is used to register an anomaly between the measured and modeled values.

The data from Unit 36922-LO-93 was then compared to the data from the remaining family of units to see whether the data from Unit LO-93 had a significant deviation from data from low MAE units. Figure 25 shows that the main contribution to the high MAE was a rather high measured average stator winding temperature, as shown for the summer 2019. The family of units did not have as high a measured average stator winding temperature (e.g., LO-84 and LO-44).

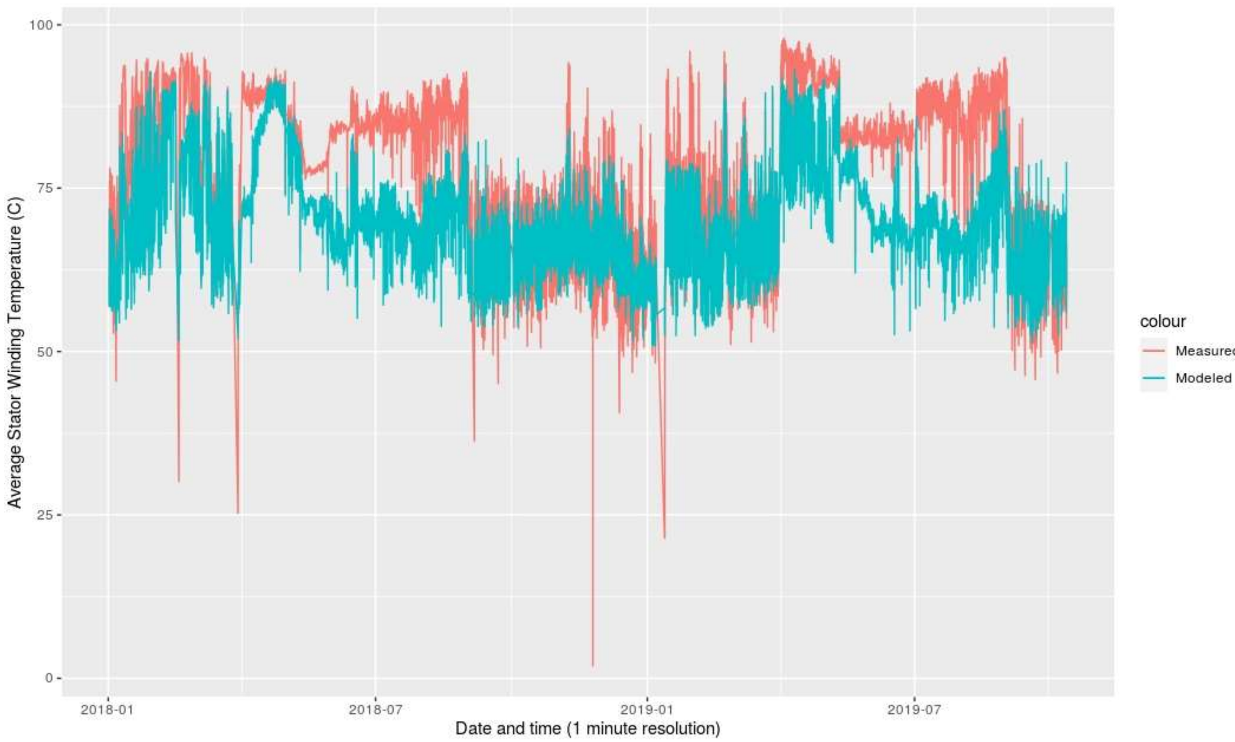

Figure 24. Measured versus modeled average stator winding temperature for Site 36922 Unit LO-93.

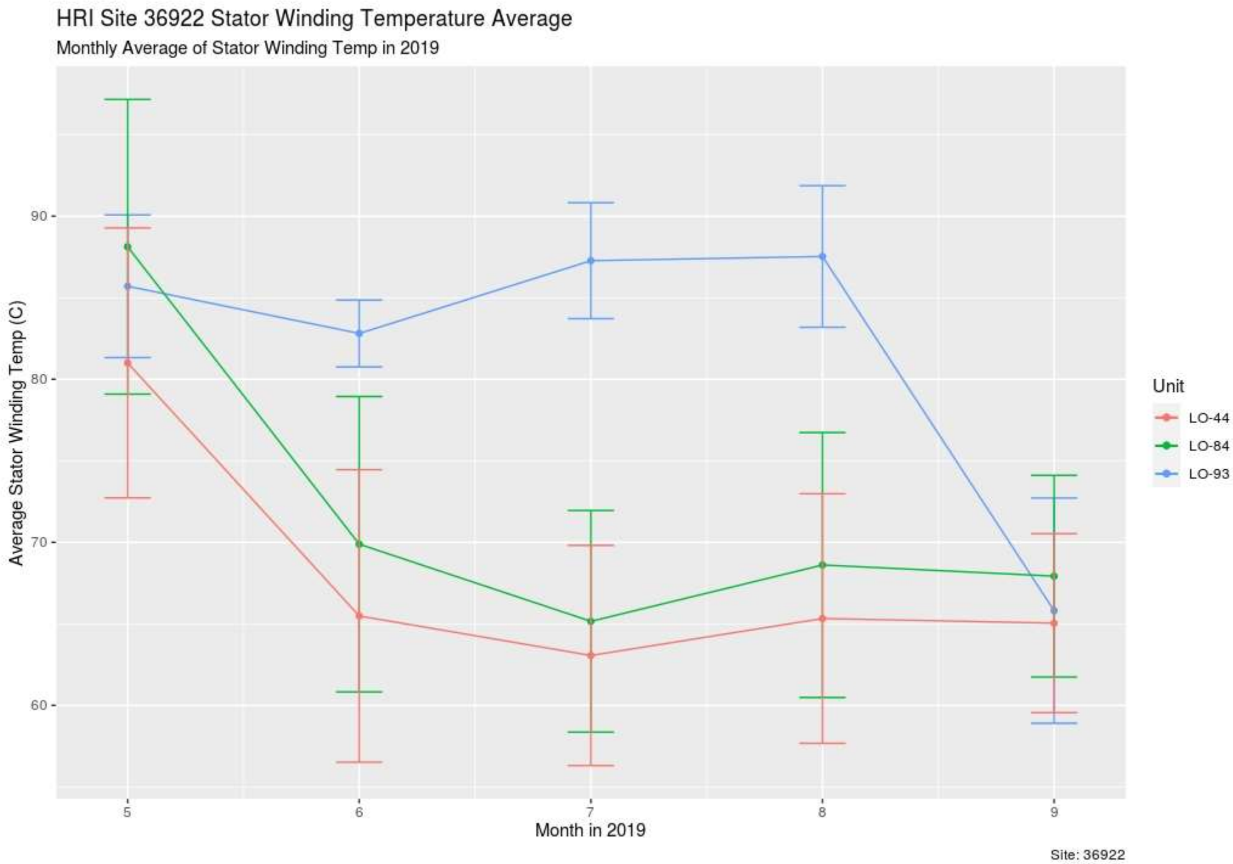

Figure 25. Measured average stator winding temperature across similar units.

The driver for the higher average stator winding temperature was the sum of the total current of LO-93 shown in Figure 26. Information provided by the owner indicated that the unit was running longer and at a higher capacity than similar units. However, this is expected for LO-93, as the unit was on for fish passage during the summer months and is left running with very little start and stop time, which is reflected in the smaller standard deviation bars. Data about the capacity of each unit are available in the HRI, allowing others this same insight.

However, this does not explain why it deviates from the stator winding temperature model results, and why unit LO-93 appears as an anomaly when using MAE as an indicator. Based on additional information provided by the system owner, the reason for the anomaly between modeled and measured average stator winding temperature was identified as the inability of the cooling water flow to keep LO-93 at its setpoint temperature during summer (some set point data are available in the HRI; however, it is not consistently available for all variables). In this instance, the cooling water flow reached its maximum possible value. This was not observed during the winter even near maximum capacity in these similar units because the cooler water and air temperatures keep the units at their temperature set point. Thus, the anomaly detected when using the stator winding temperature model and the separation observed in the measured values are limitations of a unit running near capacity without the ability to cool where the model would expect it.

In this case, a machine learning model trained on years of similar unit data can determine an unusual operating condition of one of the units on more recent data. It is worth noting that machine learning models such as those used in this specific anomaly detection model do well when the training and test data are similar. Thus, using data from multiple units and multiple operating conditions allows owners of hydropower units to better identify the range of measurements for most operating conditions and determine unusual conditions quickly. However, such identification of potential anomalies must be followed up with additional analysis to determine whether the identified condition is a true anomaly (potentially one that can lead to failure of the unit) or a previously unrecognized but normal operating condition.

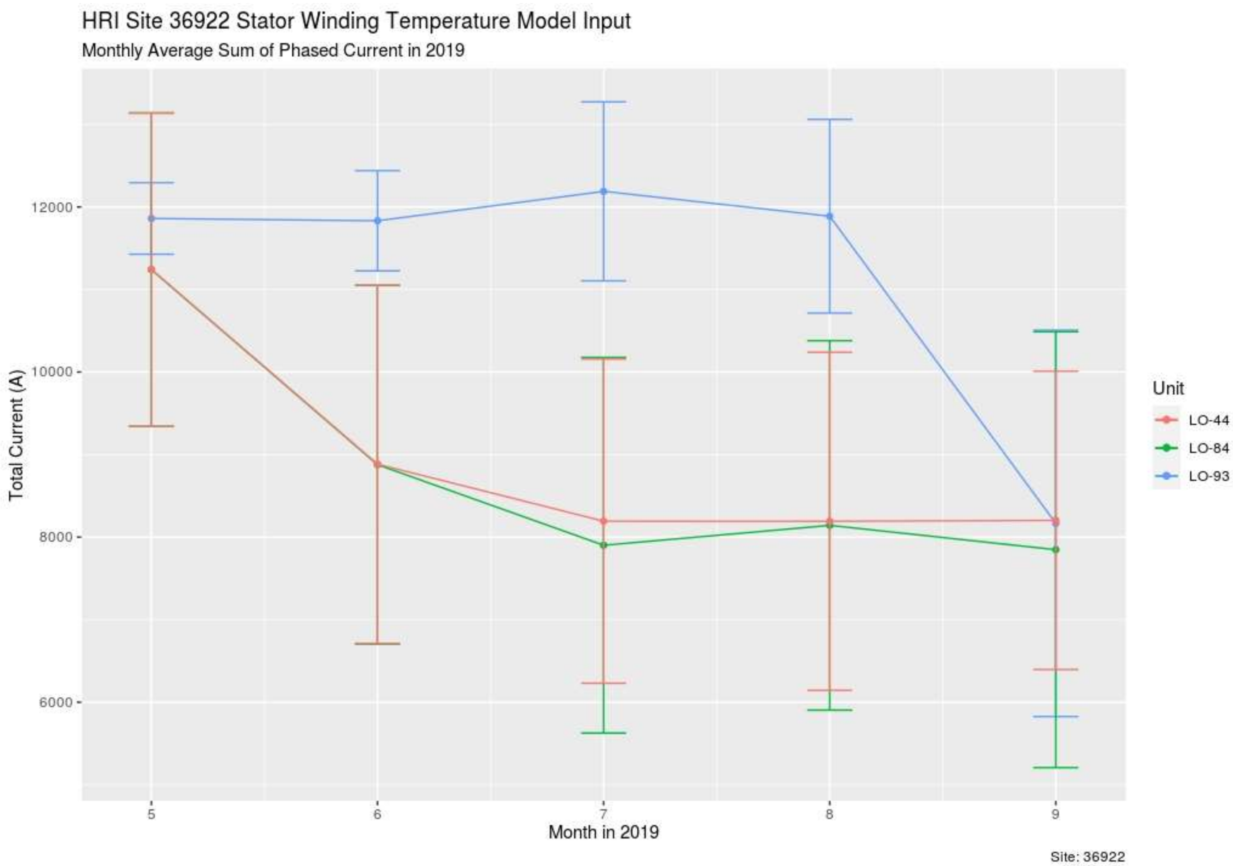

Figure 26. Measured total current model input in summer 2019.

3.5 OTHER OPPORTUNITIES FOR ANALYSIS

The examples discussed in this section focus on anomaly detection as a use case for data aggregation. Although anomaly detection is likely to be the most important near-term use of data aggregation, there are a number of other use cases for the type of data aggregation represented by the HRI dataset. Anomaly detection is related to the broader issue of asset health monitoring and diagnosis if relevant measurements are available. For instance, guide bearing wear may manifest itself as an increase in bearing vibration, and measurements of bearing vibration may provide not only information for anomaly detection (bearing failure) but also sufficient information to diagnose the cause (insufficient lubrication, cracking in bearing components, etc.). A related use case is the prognosis of asset health issues, where signatures computed from the data (such as stator winding temperature ratios, discussed in Section 3.4.2) may be leading indicators of asset failure. Such leading indicators can enable the prediction of upcoming asset failure, along with allowing asset managers to estimate the remaining life of the asset, enabling predictive maintenance strategies. Such analyses, if combined with insights on failure mechanisms of the assets, may also be able to provide information to asset managers on specific sensors and measurements that may be useful to include in the future as part of the monitoring system for the unit.

Apart from such asset health analysis, data aggregation also enables assessment of data quality—through comparative analyses of data from multiple similar units at the same facility or across multiple facilities. These assessments may help with evaluating data consistency across similar units and addressing any data quality issues that may exist.

4. CONCLUSIONS AND RECOMMENDATIONS

Hydropower is one of several types of generating assets that provides energy, capacity, and services to electric power systems. Although most hydropower assets are managed in a fleet-wide context, individual assets are also subject to decisions that differentiate roles (power and water dispatch patterns) for assets and allocate scarce or finite resources in non-uniform ways for operations, maintenance, and rehabilitation to each asset. Fleet- and unit-level data, with information on responses and outcomes for individual assets, can enable potential improvements to policies, practices, and decision-making. Data aggregation was explored in this context, and examples of analysis, using the HRI dataset, include the following:

- Utilizing the dataset as a whole for insights about how the sensor layout of a single unit or set of units compares to the hydropower industry overall. Such insights include whether additional sensors are needed to complete analyses or to make decisions.
- Utilizing multiple sensors of the same kind within a unit to provide indication of possible current or upcoming problems with equipment. When one sensor has a different trace from the others, it may be worthwhile to investigate the cause and address any issues that are found.

Data aggregation is valuable in the fleet and unit contexts, and as demonstrated using HRI data, extending beyond multiple units in a powerhouse and beyond multiple hydropower facilities in an electric utility fleet or river system. This paper discusses hydropower analytics concepts for hydropower data successfully aggregated, analyzed, and interpreted across multiple fleets. It describes these concepts and provides multiple examples of analysis and modeling using HRI data confirming the value in data aggregation.

There are many other potential analytics applications of aggregated data, including the use of complex models that can be analyzed for outliers and anomalies. Predictive models of asset performance or changes in condition can highlight a unit functioning differently from others, simply by having different metrics from other units. Again, this signals that further investigation may be needed to determine whether it is time for repairs or upgrades. These applications are the focus of ongoing efforts.

4.1 RECOMMENDATIONS

In the long-term context of deriving value from large, aggregated datasets like HRI, there are three broad constituencies: data contributors, data users, and data aggregators. Data contributors may serve two purposes: providing data and using aggregated data. Both data contributor and data user constituencies bear costs beyond membership and licensing fees when realizing the benefits of fleet-wide data aggregation. The categorized recommendations that follow are aimed at increasing the value of using aggregated datasets as a data contributor or a data user, or providing aggregated data as a data aggregator.

Some recommendations are based specifically on the characteristics observed using the HRI dataset; therefore, they are specific to HRI and may or may not apply to other data aggregators. These recommendations are intended to promote the establishment of fleet-wide aggregation of hydropower asset data as a normative best practice in the hydropower community.

4.1.1 Make Data Aggregation More Predictable, Intuitive, and Measurable

- Consider classifying and tagging hydroelectric units within aggregated datasets as *minimally instrumented, normal*, or *highly instrumented*. Interview asset managers or, through analyses, determine and document the drivers for units being minimally instrumented or highly instrumented and provide such insights to members or potential members. Summaries of the prevalence and sensor inventory of minimal, normal, and highly instrumented units can provide asset managers with insights about how the sensor layout of a single unit or set of units compares to the hydropower industry overall, and whether additional sensors are needed to complete analyses or to make decisions.
- Data aggregators should publicly provide checklists or readiness evaluations to support a data contributor's ability to estimate the level of effort to make their data available by a data aggregator as well as highlighting common challenges to success.
- Data contributors should define and track expenses related to data archival, stewardship, sharing, and integration with others to support aggregation. Tracking these expenses provides a basis for understanding the cost benefit of leveraging aggregated data and allows comparison of costs to identify the most cost-efficient way to manage and share data with others willing to share their costs.

4.1.2 Expand Data Content, Further Assess Data Quality, and Enhance Data Interoperability

- HRI data users will want information about the quality of the data they are using. HRI could implement minimum standards for data quality or, as a less intrusive measure, simply establish rubrics for the quality of data that are contributed by members. Definitions and rubrics for data quality can take several forms: documentation of features of the instrumentation and control systems that produced the data, quality control activities undertaken by a member prior to submitting data to HRI, and statistical indicators computed from data already within the HRI dataset.
- The representativeness of an aggregated dataset (i.e., "all hydroelectric units" versus "hydroelectric units in the HRI data set") should be addressed through process and comparison so that data users and decision-makers can make appropriate inferences about how closely an aggregated data-based inference is related to their individual unit or fleet decision-making context. Data aggregators have access only to data from those who participate in the data aggregation. Data aggregators could report on a regular basis a comparison of units within the dataset to a broader fleet of units (e.g., comparison to all units within the United States; see Figure 5 of this report).
- HRI could consider developing and publishing a strategic plan that identifies the "universe of hydropower asset data" (see Figure 2) and establishes cost/benefit estimates, priorities, necessary alliances, and timelines for enabling correlated analysis of data across the universe of hydropower asset data. HRI began its efforts with time series operating data and is likely moving toward aggregation and unit-by-unit correlation with reliability and event data (e.g., GADS data). Members and users alike will benefit from knowing what types of data may be slated for future correlation as a part of the HRI dataset.

- HRI should consider coordination with peer consortia (EPRI, CEATI-AMIG/HPLIG, NERC-GADSWG, EUCG-HPC) that develop or maintain hydropower asset configuration hierarchies (also known as *hydropower taxonomies*) for multiple but related purposes. Standardization of equipment terms, hierarchical structure, and system definitions and boundaries for hydropower assets and equipment will avoid duplication of effort and allow HRI members and users to perform data analytics and communicate insights more easily within technical forums.
- HRI could consider a joint proposal with research institutions for a federally funded effort to develop standardized "anonymous contextual information" (ACI) for hydroelectric units that HRI could correlate with its operating data. Such ACI could include meteorological, hydrologic, and power system statistics for every hydroelectric unit in the United States and Canada and provide valuable auxiliary data for explaining trends and patterns within the operating data that HRI is aggregating. Given the substantial ongoing and increasing efforts funded by DOE and other agencies to address potential future climate impacts on hydropower production and capabilities, such an effort may be possible without major additional expense.
- HRI should consider partnerships or alliances with organizations that could promote, research, develop, or otherwise enhance small hydropower instrumentation, monitoring, data collection, and aggregation to help address the relative lack of small unit data within the HRI dataset. The relatively small revenue streams of small hydropower units seldom allow for extensive instrumentation, monitoring, and data analytics. However, recent advances in low-cost sensors, telemetry, remote operation, automation, and data analytics capabilities may be able to unlock the potential for data aggregation and data-driven decisions within small hydropower fleets.

4.1.3 Expand Value Reporting, Standards, and Partnerships to Grow Membership

- Track "post-onboarding" experiences of members to establish a history of return on investment for membership in a data aggregation service and for the incorporating fleet-wide data aggregation and analyses into decision-making.
- Convene data users to refine concepts, statistics, rubrics, and visualizations of *similarity* (see Section 1.1.3) for hydroelectric unit characteristics, contexts, and outcomes (computed from operational data). Develop useful, repeatable analyses of how [dis]similarity of machine response compares to [dis]similarity of characteristics, contexts, and operational history.
- Develop standards for "computed data" provided by HRI members to augment raw data already in the HRI dataset. In many cases, members (and/or their engineering consultants) are the only entities with access to internal information necessary to perform computations of more complex indicators that could be useful for fleet-wide comparison.
- Consider high-level discussions and technical exchanges/workshops with peer consortia (EPRI, CEATI-AMIG/HPLIG, NERC-GADSWG, EUCG-HPC) that are also engaged in data aggregation to identify opportunities for complementary activities (e.g., correlating EUCG cost data with HRI operating and reliability data, or establishing an HRI template to feed data to the CEATI-HPLIG Start-Stop Cost Working Group Model).

4.1.4 Expand HRI's Service Offering to Include Defining and Implementing Use Cases, Standards, Statistics, and Analytics Tools

This section of the recommendations is specifically based on the aggregated datasets used (HRI dataset). As such, not all of the recommendations below may apply to other data aggregation systems.

- Create a set of canonical analyses of unit load dependencies (see Section 3.1) for the most common parameters in the dataset (i.e., how temperature, pressure, vibration, displacement, and other indicators vary with unit load) for major classes of unit design (e.g., turbine types). Enabling data users to base analyses on divergence of individual unit response from these canonical relationships is likely to make individual unit analyses more efficient and accelerate the development of insights into normal and abnormal operations.
- Develop templates for recurring and common use cases (see Sections 3 and 4) that specify necessary instrumentation, sufficient duration and granularity of operating and auxiliary data, availability of peer data within the dataset, equipment profile data required, spreadsheets, and codes for statistical analyses and AI/ML, and examples of the findings, visualizations, and decision-making uses of findings and visualizations that are possible for each use case.
    - A simple template to develop and test with multiple data users would be the replicate or arrayed sensor analysis of Section 3.2 that could help users detect and investigate anomalies among such sensors.
    - Templates to determine the number and type of sensors needed to address other problems, failures, or other business questions, such as assessing asset health.
    - Templates for more complex models that can be analyzed for outliers and anomalies. The results of predictive modeling can highlight a unit functioning differently from others, simply by having different metrics from other units. Again, this signals that further investigation may be needed to determine whether it is time for repairs or upgrades.
    - Templates for machine learning methods (supervised methods such as neural networks) that require larger datasets for which the training and test data are similar. The HRI dataset contains data from multiple units and multiple operating conditions, which may allow for better identifying the range of measurements for most operating conditions. Further exploration may allow for determining whether the dataset could be used to verify that training data contains data covering an expected range (indicating that it falls within a certain set of operating conditions).

5. ACKNOWLEDGEMENTS

The authors acknowledge the support, review, and feedback from the Hydropower Research Institute Management Committee and Hydropower Research Institute Technical Steering Committee. The authors also gratefully acknowledge the technical editing support from John Batson III, and the technical feedback from the peer reviewers of this document.