Bridging Equipment Reliability Data and Risk-Informed Decisions in a Plant Operation Context

D. Mandelli[a], C. Wang[a], S. Lawrence[a], D. Morton[b], I. Popova[c], S. Hess[d]

[a] Idaho National Laboratory, Idaho Falls, USA, {diego.mandelli, congjian.wang}@inl.gov
[b] Northwestern University, Evanston, USA, david.morton@northwestern.edu
[c] Texas State University, Austin, USA, ip12@txstate.edu
[d] Jensen Hughes, West Chester, USA, SHess@jensenhughes.com

Abstract: Industry equipment reliability and asset management programs are essential elements that help ensure the safe and economical operation of nuclear power plants. The effectiveness of these programs is addressed in several industry-developed and regulatory programs. The Risk-Informed Asset Management project is tasked to develop tools in support of the equipment reliability and asset management programs at nuclear power plants. These tools are designed to create a direct bridge between component health and lifecycle data and decision-making (e.g., maintenance scheduling and project prioritization). This article provides a guide for specific use cases that the Risk-Informed Asset Management project is targeting. We have grouped uses cases into three main areas. The first area focuses on the analysis of equipment reliability data with a particular emphasis on condition-based data, such as test and surveillance reports and component monitoring data. The second area focuses on the integration of equipment reliability into system-plant reliability models to determine system-plant health and identify components critical to maintaining an operational system. Lastly, the third area manages plant resources, such as maintenance activities and replacement scheduling using optimization methods. Here, the primary focus is on supporting typical system engineer decisions regarding maintenance activity scheduling and component aging management. This is performed in a risk-informed context where the term "risk" is broadly constructed to include both plant reliability and economics. This framework combines data analytics tools to analyze equipment reliability data with risk-informed methods designed to support system engineer decisions (e.g., maintenance and replacement schedules, optimal maintenance posture) in a customizable workflow.

---

1. INTRODUCTION

The Risk-Informed Systems Analysis (RISA) Pathway1 [1] of the United States Department of Energy Light Water Reactor Sustainability2 [2] program is conducting collaborative research that applies risk-informed technology to assist operating nuclear power plants (NPPs) to reduce costs and support their adaptation to the changing economic and power generation environment. The research is being performed within the framework of specific use cases, which are intended to enable rapid technology development, deployment, and dissemination throughout the operating U.S. NPP fleet to address pressing economic, operational, or safety significance issues.

One area of research in the Risk-Informed Systems Analysis Pathway focuses on developing methods and tools to optimize plant operations (e.g., maintenance and replacement schedules, optimal maintenance postures for plant structures, systems, and components [SSCs]) in a more cost effective manner than current approaches and makes better use of available SSC health and cost data. This is accomplished under the Risk-Informed Asset Management (RIAM) project by creating a direct bridge between component equipment reliability (ER) data and decision-making (e.g., maintenance scheduling

[1] Risk-Informed Systems Analysis website: https://lwrs.inl.gov/SitePages/Risk-Informed%20Systems%20Analysis.aspx
[2] Light Water Reactor Sustainability website: https://lwrs.inl.gov/

and project prioritization). Here, we are supporting typical system engineer decisions regarding maintenance activity scheduling and component aging management.

2. USE CASE OVERVIEW

RIAM project capabilities can be grouped into three main areas:

1. *ER data analytics (see Section 4)*. Targeting the analysis of ER data to adequately measure component health. This area includes all methods designed to analyze numeric and text ER data (i.e., monitoring data, maintenance activities, work orders, and maintenance and issue reports [IRs]), employ historic data to detect abnormal behavior (anomaly detection) and the cause of such abnormal behavior (diagnostic), and predict SSC future performances (prognostic).
2. *Plant health digital modeling (see Section 5)*. Designed to model from a reliability perspective the considered system and plant and, more importantly, fully integrate ER data into such models. The main feature of these reliability models is that they encompass not only the reliability and availability of the system and plant but also economic aspects. The real challenge is to directly inform these models with the available ER data (historic and current).
3. *Resource optimization (see Section 6)*. Targeting the optimization of plant resources. Here, plant resources include multiple entities, such as plant operation, maintenance, and capital budgets; workforce tasks; and SSC lifecycles. This area is directly linked to the plant decision-making process and considers both short- and long-term decisions.

3. RISK ANALYTICS PLATFORM WORKFLOW

This section provides a more detailed guide on how the methods and tools developed under the RIAM project can be used to bridge ER data with decisions. We refer here to Figure 1, which graphically describes a complete risk analytics platform data workflow. The starting point is the set of ER data generated by the SSC monitoring system (bottom left of Figure 1), available for example from the plant monitoring and diagnostic center. The steps of this workflows are as follows:

1. Analyze ER data using methods presented in Section 4 to track SSC health (i.e., performance and degradation) and identify possible anomalies in the SSC behavior (see Section 4 and lower left portion of Figure 1).
2. Measure SSC health by determining the margin of specific SSC failure modes given current SSC conditions and historic data (see Section 5 and top-left portion of Figure 1).
3. Determine margin values for the failure modes of the chosen SSC and propagate them through classical reliability models (e.g., fault trees) to determine the margin at the system-plant level (i.e., plant system and health) and the risk importance measure for each SSC failure mode (see Section 5 and top-left portion of Figure 1)
4. Choose the optimal set of projects (e.g., maintenances activities) given plant system and health information determined in Step 3 and follow one of two possible paths (see mid-left portion of Figure 1):
    a. Ranking-based path: Select the failure modes with highest consequences and importance measures and use economic constraints to filter the chosen project list
    b. Multi-objective optimization path: Choose projects based on both reliability and economic factors simultaneously (e.g., through a Pareto frontier analysis as indicated in Section 6.1)
5. Set the optimal schedule for the projects chosen in Step 4 (see mid-right portion of Figure 1) and use the methods presented in Section 6.2 and [3] to set the optimal project actuation schedule based on reliability and economic constraints (i.e., medium- and long-term decisions)
6. Partition each project into tasks and determine the optimal schedule of each task (i.e., short-term decisions) (see right portion of Figure 1) using the methods presented in Section 6, where tasks that might be provided by plant system engineers are added to the list of tasks for the projects chosen in Step 5.

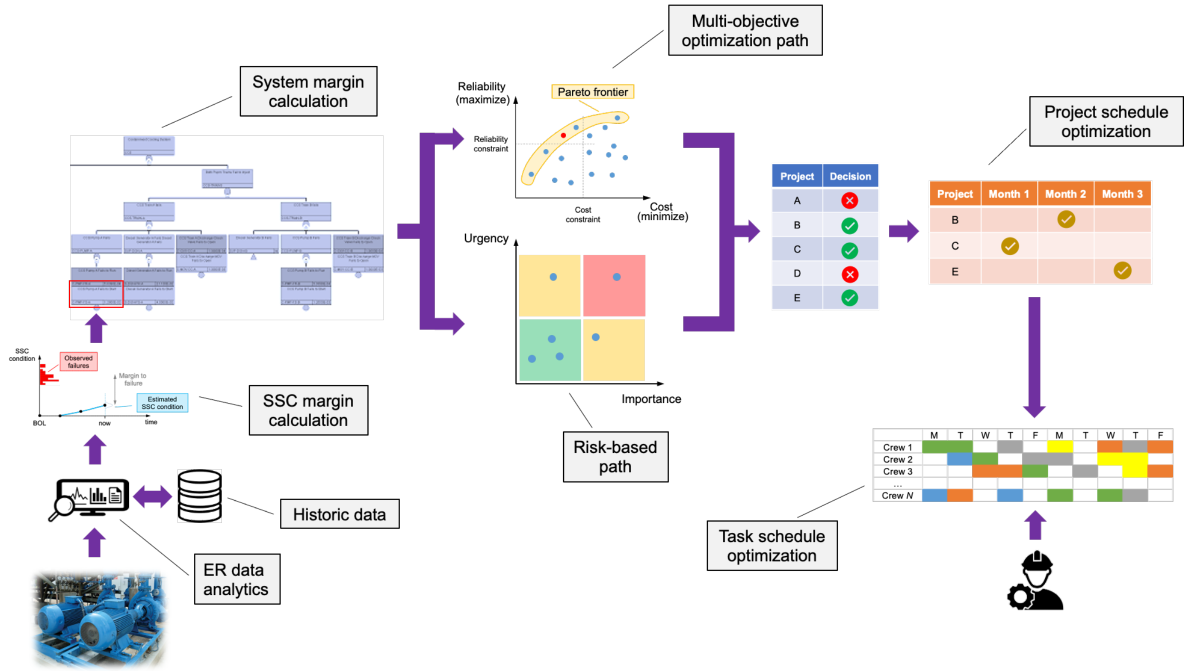

Figure 1. RIAM toolkit as bridge between ER data to decisions: graphical representation of the workflow starting with SSC monitoring data (bottom left) to reliability modeling (top left) to project prioritization (center) and task scheduling (right).

4. ER DATA ANALYTICS

Typically, a single component SSC is part of a system of components (see Figure 2 [left]) where such a system is designed to provide a designed function, that is, *emergence* (such as electric power generation for a power plant). Each component contributes to the system emergence by providing a specified functionality used by other components through a set of connections where *operands* (e.g., mass, energy, or data) are exchanged. The goal of a system health program is to monitor not only the correct operation of each component but also health parameters, such as aging and degradation (indicated as $\underline{F}(t)$ in Figure 2 [right]). In addition, a system health program is designed to perform appropriate actions to assure component functionality (indicated as $\underline{T}(t)$ in Figure 2 [right]). In this paper, $\underline{T}(t)$ includes all the external stressors that contribute to altering component aging and degradation.

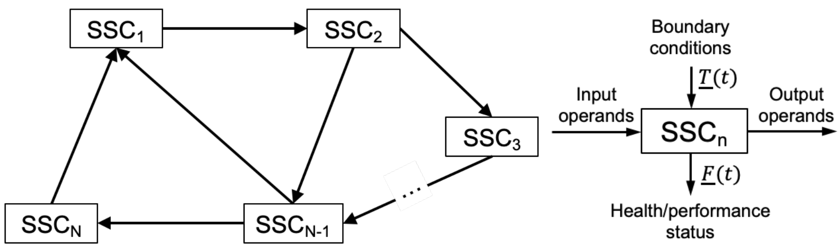

Figure 2. System (left) and component (right) representation.

When moving in more detail to the component level, it is vital to understand the relationships between monitoring and testing data, maintenance activities (MAs), and failure modes (FMs). Figure 3 provides a detailed functional-form description of a generic SSC by employing an object process methodology (OPM) diagram [4]. An SSC OPM diagram provides an essential description of the SSC from both a form and functional perspective. This diagram explicitly indicates how SSC internal functions ($Func_f, f = 1, \dots, F$) process and act upon operands and how the elemental components ($ssc_r, r = 1, \dots, R$) support these functions.

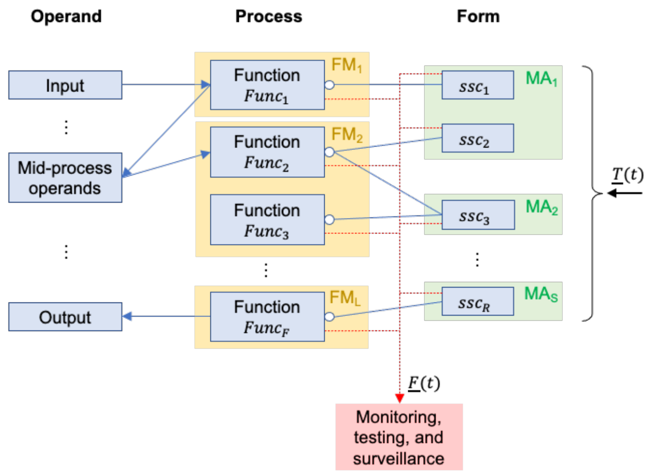

Figure 3. SSC representation through an OPM diagram.

From an ER perspective, monitoring and testing activities (i.e., $\underline{F}(t)$) act on both SSC functional (e.g., rpm recorded for an induction motor) and form (e.g., blade corrosion of centrifugal pump) elements. Degradation processes (i.e., $\underline{T}(t)$) directly alter the form-related elements of the component (i.e., ssc_r) that consequently affect SSC functional elements (i.e., $Func_f$,). Typically, from a reliability perspective, component FMs are described in term of loss of function, and, hence, in the OPM diagram, FMs are only directly linked to the functional elements of the component (i.e., $Func_f$,). Lastly, note that MAs (such as component replacement, refurbishment, or reconditioning), indicated as MA in Figure 3, act on the form elements of components (i.e., ssc_r).

For the scope of this article, the OPM diagram of a component represents the key point to automatically understand and analyze health data $\underline{F}(t)$. In particular, it clearly links monitored and recorded data with FMs that might affect component performance and MAs that would restore component functionality. We are employing model-based data analysis methods with the goal of linking component models with data rather than using machine-learning methods, which solely rely on the available data in order to perform diagnostic and prognostic operations. Note that an OPM diagram extends FMs and effect analysis tables by providing a form and functional description of the considered system in a graphical form.

The next step is to characterize a generic component SSC from a data scientist point of view. This is shown in Figure 4 where three levels are identified: the component level (which would correspond to what is shown in Figure 3), a sensor and monitoring level (which retrieves and records portions of $\underline{T}(t)$ and $\underline{F}(t)$ in digital form), and data level. Data retrieved from $\underline{T}(t)$ (i.e., $\underline{\theta}(t)$) can be either textual (e.g., work orders) or numeric (e.g., environment temperature). We indicate here with "num" the portion of $\underline{\theta}(t)$ that is numeric while we indicate with "NL" the portion of $\underline{\theta}(t)$ that is textual (NL here stands for natural language). Data retrieved from $\underline{F}(t)$ has been portioned in two portions, component health and performance monitoring ($\underline{\varrho}(t)$ and $\underline{\gamma}(t)$), which can be numeric or textual in nature as well.

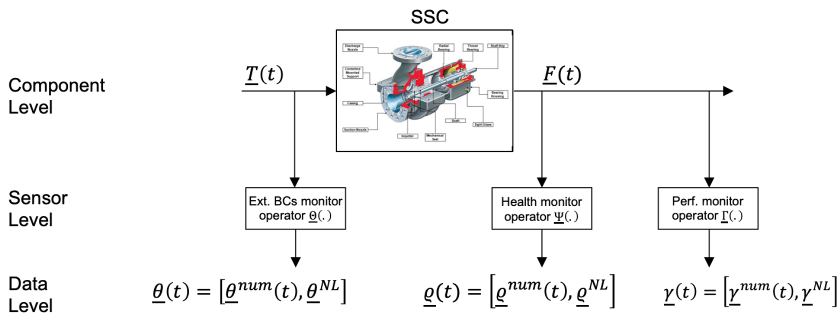

Figure 4. System health program: component representation from a data point of view.

ER data can be of different formats (i.e., numeric and textual). In addition, the events and logs recorded in $\underline{\theta}^{NL}(t)$ or $\underline{\gamma}^{NL}(t)$ can be defined over an interval or a single time instant. These two observations led to a challenge when we analyzed ER data: identifying a common data structure that can be employed to represent numeric and textual data and events defined over time instants and time intervals. The advantage of having a common data structure is that it considerably simplifies the causal representation of events and monitoring data for condition-based monitoring applications.

This challenge has been resolved here by representing all elements of $\underline{\theta}(t)$, $\underline{\varrho}(t)$, and $\underline{\gamma}(t)$ (numeric and textual) in symbolic form (i.e., a series of symbols). This approach has the advantage that it simplifies the integration of numeric data with recorded events to identify patterns and outliers. In more detail, the method is structured in the following four steps:

1. Create a symbolic conversion of numerical time series using the SAX method [5]. Data preprocessing (e.g., identification of anomalous behavior) may be required depending on the situation.
2. Create a symbolic representation of textual data by characterizing events and logs into a graph form using natural language processing (NLP) methods [6]. A graph form has the advantage that it easily captures the structural relationship among text objects (i.e., nouns, verbs).
3. Combine data from Steps 1 and 2 into a common symbolic data structure. In our case, this is performed by creating a multivariate symbolic time series.
4. Apply model-based causal inference methods on the structure generated in Step 3 by coupling data analysis methods with component OPM diagrams to infer component health, its FMs, and related maintenance activity that should be performed.

4.1 Analysis of Textual Data

Most methods found in the literature [6] process text reports using supervised learning to predict the report nature (e.g., failure, operating). Here we are following a different path by analyzing the sentence structure of logs and reports, organizing information in a structured form, and creating a structural relationship among text objects (i.e., understand who and what did what, when, why, and where). This is being accomplished by employing NLP methods to perform two main tasks, syntactic and semantic analysis.

As a starting point, we are characterizing the content of a generic IR or a maintenance report. Note that maintenance report content is fairly straightforward since it basically reports component replacement or restoration activity. Thus, we can define two classes of IRs:

- Class 1 IR: When the IR reports either an event (e.g., SSC malfunction) or data regarding the component health (e.g., excessive corrosion on pump impeller).
- Class 2 IR: When the IR reports a causal relation between two nodes, the content of these nodes can be any combination between events and SSC health information linked by a causal relationship.

Note that this classification scheme defined by these two mutually exclusive classes needs to be validated with NPP actual data to measure its validity (i.e., the degree to which the two classes are in actuality mutually exclusive). In the validation process, we can measure the percentage of actual NPP IRs that falls in each class and, more importantly, the percentage of IRs that do not fall in either class. Note that the classifications provided above are relevant to IRs related to plant equipment performance. Since IRs can be written on a wide variety of topics (e.g., issues related to programmatic performance, human performance, etc.), we expect that a substantial fraction of IRs would not be classified as one of the two classes related to the plant system and equipment health defined above.

The first step in the analysis of text data is to perform a syntactic analysis [6] of the raw text by employing the rules of formal grammar. Here we have assumed that the text is in a digital form (typically in a string form) and performed the syntactic analysis through the following main steps:

1. Sentence segmentation and word tokenization: translate each sentence into a list of string elements
2. Part of speech tagging: identify grammatic elements in each string (e.g., nouns, verbs)

3. Named entity recognition: classify text entities (e.g., names, dates, events) and identify them (e.g., component ID, event occurrence time)
4. Relation extraction: create a knowledge graph where entities identified in Step 3 are linked together in a graph that reflects the structure of the original sentence.

While Steps 1–3 listed above are common in any NLP analysis, our approach deviates from the standard NLP method in Step 4 where we identify the elements of the SSC OPM model (i.e., operands, forms or functions) in the text. From each SSC OPM model, we can generate a set of textual elements that lists not only all OPM elements but also their relationship.

In Step 4, we infer the causal relationship between elements of the IR. These relationships are in the form of cause and consequence. Here, we exploit the observations reported in the IR by plant system engineers and trace back causal relationship with other IRs using the SSC OPM models.

The methods designed to extract information from IRs that belong to Class 1 has been structured in a similar way to the one presented in [7]. We, in fact, based our methods on a new set of rule templates based on specific trigger words and relations. At its initial stage, our work focused on developing status nouns and verbs that would indicate a degradation of SSC functions or SSC internal elements. The chosen set of status words includes verbs, adjectives, and nouns obtained again from the `WordNet3` database. For Class 1 IRs, we have identified an initial list of status relations encoded using `STANZA`, which are listed in Table 1.

Table 1. Set of status relations.

Relation
A (noun) “status verb” “status adjective”
A (noun) “status verb” “status verb-ing”
“Status adjective” B (noun) “status verb”
“Status noun” “status verb” prep. B (noun)

These status relations were coded in a Python-based code that relies on the Stanford NLP library `STANZA`. Once the IR has been processed using all the NLP steps listed above, a set of tuples is created from each sentence in the form (SSC, form, function, health status). These tuples are designed to represent the sentence in digital form as follows:

SSC; subject = ‘OPM function/form’; health status = ‘ok, ‘degraded’ or ‘anomalous’

As an example of Class 1 IR is provided as follows:

Oil puddle was found in proximity of CCW Pump 1B.

By using the NLP analysis in Steps 1–4 listed above using `STANZA` and `NLTK` Python libraries, the resulting grammatical structure of the IR is shown in Figure 5. This figure shows the part of speech tags represented on top of each word and the grammatical dependencies[4] between words (represented with arrows).

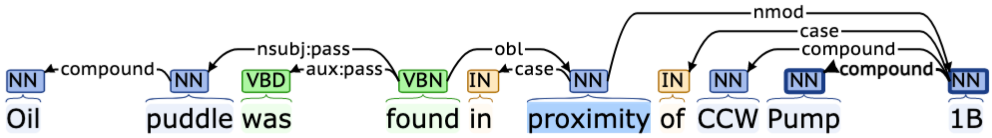

Figure 5. Grammatical decomposition and analysis of the example Class 1 IR.

[3] WordNet official website: https://wordnet.princeton.edu/
[4] Refer to https://downloads.cs.stanford.edu/nlp/software/dependencies_manual.pdf for a complete description of each Stanford dependency

Step 4 is accomplished by looking in the IR for specific SSC tags (i.e., CCW pump 1B). Here we have also assumed that SSC tags are unique and given. Once the SSC tag has been identified, its OPM model can identify OPM elements in the sentence that refer to such a model. Next, NLP analysis identifies the verb "find" (i.e., verb being part of anomalous status). The following tuple is constructed:

(SSC = CCW Pump 1B; subject = ISO VG100 oil; health status = anomalous)

5. PLANT HEALTH DIGITAL MODELING

Current reliability models are based on Boolean logic structures [8] (e.g., fault trees), which describe the deterministic functional relationship between SSCs and human interventions. Each basic event in a reliability model represents a specific elemental occurrence (e.g., failure of a component, failure to perform an action by the plant operators, recovery of a safety system, etc.), and a probability value is associated with each basic event, which represents the probability that the basic event can occur. However, maintenance and surveillance operations are typically not completely integrated into a probabilistic risk analysis (PRA) structure. In addition, a probability value associated with an event is thus an integral representation of the past operational experience for such an event, and it does not incorporate information on the present health status of SSCs (e.g., from diagnostic and condition-based data) and health projections (when available from prognostic data) on anticipated changes in SSC condition and performance in the near future.

A possible alternate path can start by redefining the word "reliability" to encompass a broader meaning that better reflects the needs of a system health and asset management decision-making process. Rather than focusing on how likely an event is to occur (in probabilistic terms), we think in terms of how far this event is from occurring [9]. This new interpretation of risk transforms the concept from one that focuses on the probability of occurrence to one that focuses on assessing how far away (or close) an SSC is to an unacceptable level of performance or failure. This transformation has the advantage that it provides a direct link between the SSC health evaluation process and standard plant processes used to manage plant performance (e.g., the plant maintenance and budgeting processes). The transformation also places the question into a form that is more familiar and readily understandable to plant system engineers and decision makers. When dealing with condition-based data (actual and archived data), margin $\widetilde{M}$ is defined here as the distance between actual SSC observed past conditions (e.g., oil temperature, vibration spectrum) that lead to failure (see Figure 6).

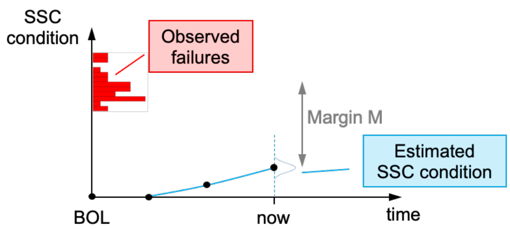

Figure 6. Margin in a condition-based maintenance context: evolution of an SSC condition as a function of time and margin definition.

Consider now two components (A and B). The $\widetilde{M}$ for both components can be visualized in a 2D space, as shown in Figure 7. Starting with brand-new components (i.e., $\widetilde{M}_A, \widetilde{M}_B = 1$), aging degradation that affects both can be represented by the blue line of Figure 7, which parametrically represents the combination of the normalized margins $(\widetilde{M}_A(t), \widetilde{M}_B(t))$ at a point in time t. Note that, if no maintenance (whether preventive or corrective) was ever performed on either component, this path would move from the coordinates $(1,1)$, components A and B at the beginning of life to the coordinates $(0,0)$ where both components had failed. We can identify these regions in Figure 7: the occurrence of both events where $\widetilde{M}_A = 0$ and $\widetilde{M}_B = 0$ and the occurrence of either event when $\widetilde{M}_A = 0$ or $\widetilde{M}_B = 0$. Now we can calculate the $\widetilde{M}$ for the events listed above. This is accomplished by following the definition of margin:

by measuring the distance between the actual condition of components A and B and $\widetilde{M}$ conditions identified by the event under consideration (e.g., the occurrence of both or either events):

$$\begin{gathered}\widetilde{M}(A\ AND\ B) = dist\left[\left(\widetilde{M}_A, \widetilde{M}_B\right), (0{,}0)\right] \\ \widetilde{M}(A\ OR\ B) = min\left(\widetilde{M}_A, \widetilde{M}_B\right)\end{gathered} \tag{1}$$

The function $dist[X, Y]$ is designed to calculate the Euclidean distance between points X and Y.

Hence, exact solutions can be obtained extremely fast. More precisely, reliability calculations using $\widetilde{M}$-based data can be performed by completing these four steps:

1. Construct the fault tree (FT): at this point, an FT only contains deterministic information about the architecture of the system under consideration (i.e., it simply models how the basic events are related to each other from a functional perspective).
2. Generate the minimal cut-sets (MCSs) from the FT: as also indicated in Step 1, an MCS still represents the minimal combinations of basic events that lead to the TE.
3. Assign $\widetilde{M}$ to each basic event.
4. Calculate the $\widetilde{M}$ of the union of the MCSs.

As part of system reliability modeling, it is always important to determine the importance of each basic event. In a PRA setting, this is performed by relying on risk importance measures [8], such as Birnbaum or Fussell-Vesely. Given the different nature of $\widetilde{M}$, it is possible to perform a risk importance ranking by relying on a classical sensitivity measure (derivative based) for each basic event BE defined as: $S_{BE} = \frac{\partial\, \widetilde{M}(TE)}{\partial\, \widetilde{M}(BE)}$. In other words, S_{BE} indicates how a small variation of $\widetilde{M}(BE)$ directly affects $\widetilde{M}(TE)$.

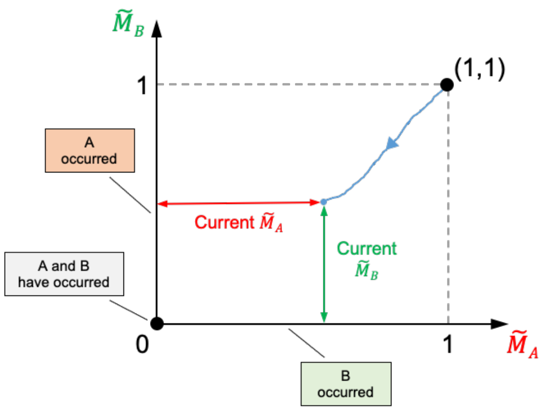

Figure 7. Graphical representation of event occurrences based on a margin framework.

6. PLANT RESOURCE OPTIMIZATION

6.1 Project Prioritization

The FMs with a higher S_{BE} (see Section 5) are the ones selected as candidates to be subject to MAs (see Figure 3). A list of possible options to address each failure mode is available where costs (e.g., procurement costs for a new or refurbished component) and benefits (e.g., increased margin for loss of production) are readily available or can be numerically determined. Given the candidate MAs and their options, we can now identify the best set of activities and options that give "the most bang for the buck."

This is accomplished by identifying the Pareto frontier [10] out of the all the possible MAs and options. The Pareto frontier is defined as the set of non-dominated solutions that are characterized by the highest value in at least one dimension. Let's assume a decision can be taken from a set of options by considering the utility and cost of each option. Using a graphical representation (see Figure 8), it is possible to plot each option as a point in a 2D space of cost vs. utility[5]:

- *Cost*: this axis represents the cost associated with each option ranging from 0 (i.e., cheapest option) to a maximum value C_{max} (i.e., the most expensive option).
- *Utility*: this axis represents the added value (or the performance) associated with each option ranging from 0 (i.e., lowest performance option) to a maximum value U_{max} (i.e., option with highest performance).

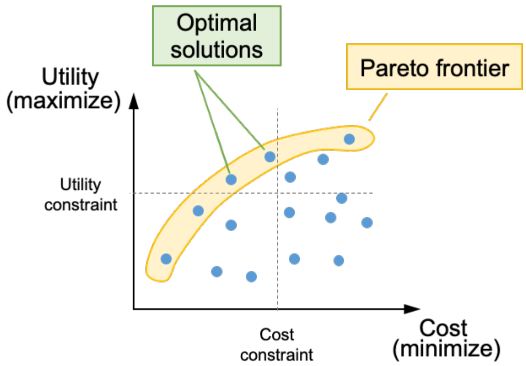

Figure 8. Pareto frontier obtained from a set of options plotted in a cost vs. utility space and imposition of cost and utility constraints (right).

Once the complete set of options have been generated and the utility and cost values have been determined for each option, the next step is the determination of the Pareto optimal frontier, which is fundamentally an envelope of options that dominates (in terms of both utility and cost) the set of remaining options (see Figure 8).

6.2 Long-Term Decisions: Project Scheduling Given Budget Constraints

The method described in Section 6.1 does not explicitly consider project actuation scheduling but instead focuses on the optimal subset of projects that provide higher value through a multi-objective optimization lens. In practical settings, project scheduling is done in phases (e.g., monthly, quarterly) wherein each phase budget is allocated, and the goal now is to choose the optimal project actuation schedule that minimize**s** costs and satisf**ies** budget constraints [11].

6.3 Short-Term Decisions: Maintenance Activity Scheduling Given Personnel Constraints

Once the project schedule has been finalized (see Section 6.2), each project is **divided** into tasks where each task is characterized by a set of parameters (e.g., duration, number of personnel required, list of required tasks that need to be performed prior to start**ing** this task, skillset required, completion deadline). Plant resources are now constituted by a set of crews where each crew is characterized by the number of people, schedule availability, and available skill set. The goal is to minimize the time to complete all jobs [3] and tasks provided the constraints that the order of tasks for each job must be preserved and one task can be assigned to each crew [11].

7. CONCLUSION

In this paper, we have presented a series of methods and models designed to create a direct bridge between ER data and ER related decisions. Even though this type of bridging is not new, we are here presenting a different structure for such a bridge. First, we have introduced a novel approach to analyze ER data that integrates logs and event data with numeric data available from plant monitoring and diagnostic centers in a common data structure (symbolic in nature). Rather than focusing on machine-

[5] As indicated earlier, the number of attributes considered in complex settings can be $N > 2$. Thus, in such cases, the space would be N-dimensional.

learning heuristics, the system view of the component (through a OPM diagram) provides the required knowledge to our data analysis methods to extract knowledge from text data retrieved by logs or workorders. A challenge of this class of methods is their reliability to effectively analyze large amounts of reports and correctly extract the information contained in them. Processed data can then be integrated into classical reliability models (e.g., fault trees) that are solved not by using a probability-based but a margin-based language. The main advantages of this method are that it allows a much better use of ER data and provides a more adequate risk importance ranking of the FMs for the considered set of SSCs. Lastly, the decision-making step is carried through by determining the set of projects and operations that provide "the most bang for the buck" (i.e., the Pareto frontier), prioritizing the actuation schedule for the selected projects (medium-term decisions), and identifying the optimal schedule that minimizes the completion time of the required maintenance tasks (short-term decisions). This is performed in a risk-informed context where the term "risk" is broadly constructed to include both plant reliability and economics.

Acknowledgements

This manuscript has been authored by Battelle Energy Alliance, LLC under Contract No. DE-AC07-05ID14517 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes.

References

[1] R. Szilard, H. Zhang, S. Hess, J. Gaertner, D. Mandelli, S. Prescott, and M. Farmer, "RISA Industry Use Case Analysis," Idaho National Laboratory Technical Report, INL/EXT-18-51012 (2018).

[2] U.S. Department of Energy, "Light Water Reactor Sustainability Program Integrated Program Plan," Idaho National Laboratory Technical Report, INL/EXT-11-23452 (2020).

[3] D. Mandelli, C. Wang, M. Abdo, K. Vedros, J. Cogliati, J. Farber, A. Al Rashdan, S. Lawrence, D. Morton, I. Popova, S. Hess, C. Pope, J. Miller, S. Ercanbrack, "Industry Use Cases for Risk-Informed System Health and Asset Management," Idaho National Laboratory Technical Report, INL/EXT-21-64377 (2021).

[4] D. Dori, E. Crawley, *Object-Process Methodology: A Holistic Systems Paradigm*, Springer ed. (2002).

[5] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, "A Symbolic Representation of Time Series, with Implications for Streaming Algorithms," *8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, San Diego, June, pp.2-11 (2003).

[6] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*, O'Reilly Media (2009).

[7] S. Doan, E. W. Yang, S. S. Tilak, P. W. Li, D. S. Zisook, M.Torii, "Extracting health-related causality from twitter messages using natural language processing," *BMC Medical Informatics and Decision Making*, **19**, pp. 71–84 (2019).

[8] J. Lee and N. J. McCormick, *Risk and Safety Analysis of Nuclear Systems*, Wiley edition (2011).

[9] D. Mandelli, S. M. Hess, and C. Wang, "On the Language of Reliability: A System Engineer Perspective," in *Proceedings of PSA conference* (2021).

[10] C. A. Mattson, A. Messac, "Pareto Frontier Based Concept Selection Under Uncertainty, with Visualization," *Optimization and Engineering*, **6**, pp. 85–115 (2005).

[11] D. Mandelli and C. Wang, S. M. Hess, D. Morton, I. Popova, "Development and Application of Risk Analysis Toolkit for Plant Resource Optimization," in *Proceedings of PSA conference* (2021).