Conference Paper Manuscript
APT#: 187390

Markov Decision Processes for Intelligent, Risk-Informed Asset-Management Decision-Making

prepared by

David Grabaskas[1], Yunfei Zhao[2], Carol Smidts[2], Vera Moiseytseva[1], Roberto Ponciroli[1], Pascal Brocheny[3]

[1]Nuclear Science and Engineering Division
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 208
Argonne, IL 60439-4854

[2]The Ohio State University

[3]Framatome

ANS Advanced Reactor Safety Conference
June 16-19, 2024
Las Vegas, Nevada

Markov Decision Processes for Intelligent, Risk-Informed Asset-Management Decision-Making

David Grabaskas[1], Yunfei Zhao[2], Carol Smidts[2], Vera Moiseytseva[1], Roberto Ponciroli[1], Pascal Brocheny[3]

[1]Argonne National Laboratory, Lemont, IL, USA, [2]The Ohio State University, Columbus, OH, USA, [3]Framatome, Lynchburg, VA, USA

[leave space for DOI, which will be inserted by ANS]

ABSTRACT

Advanced nuclear reactors are a promising option for aiding the world in achieving its net-zero carbon emission goals, however, there are significant challenges to attaining and maintaining economic competitiveness with other sources of electricity. To improve the economic competitiveness of advanced reactor designs, a project was initiated to explore the use of Markov Decision Processes (MDPs) to guide asset-management decision-making during advanced reactor operation. MDPs are a powerful tool for optimizing decision-making in complex environments and their application to advanced reactors can aid in planning maintenance and repair activities to minimize downtime and maximize generation. The described approach expands on previous work regarding the use of MDPs for operational decision-making through the direct incorporation of real-time plant information. The integral MDP analysis includes information from online component diagnostic tools and the plant's real-time generation risk assessment (GRA) and probabilistic risk assessment (PRA), which evaluate plant risk from both an economic and safety perspective. The result is an asset-management optimization framework that is based on real-time data regarding plant component status and the current best-estimate of plant risk. The paper presents an overview of the theoretical framework to incorporate the different information pathways into an integral MDP analysis, along with example analyses.

Keywords: Probabilistic Risk Assessment, Generation Risk Assessment, Asset-Management, Diagnostics

1. INTRODUCTION

There has been a recent resurgence of interest in advanced reactors designs as a potential avenue for addressing rapid climate change. However, advanced reactors face an increasingly challengingly economic environment in countries with competitive energy markets. Approaches are necessary to reduce the cost of advanced reactors, including both upfront investments and the costs associated with operation and maintenance. The objective of the current project is to improve the economic competitiveness of advanced reactors through the optimization of cost and plant performance, which can be achieved by coupling online monitoring with intelligent asset-management decision-making. The effort reviewed here is a multi-year collaboration between Argonne National Laboratory (Argonne), the Ohio State University (OSU), and Framatome, funded by the U.S. Department of Energy (USDOE) Nuclear Energy Enabling Technologies (NEET) program.

As advanced reactors are early in the development life-cycle, online monitoring systems and associated sensor networks can be incorporated directly into the design without the constraints related to retrofitting and system upgrades. Due to their innovative designs and lack of operating experience, advanced reactors have large uncertainties regarding component reliability, potential failure modes, and long-term maintenance needs. The project has focused on two objectives for the optimization of advanced reactor operation and asset management using online monitoring and diagnostics. First, during the reactor design phase, it is necessary to develop a sensor network that can properly detect and diagnose important faults and component degradation. This is a difficult task as there are many unknowns regarding long-term operational reliability and the associated costs of additional sensors and system penetrations can be prohibitive. Secondly, once reactor operation begins, the asset management approach must seamlessly integrate online monitoring information and the plant's risk profile to develop an optimized plant operation and maintenance plan. The challenges of this task include cost-benefit decision-making in multivariate space while ensuring the plant does not approach risk or safety limits.

The current paper focuses on the latter objective, which is optimization of asset-management decision-making during operation using an integrated Markov decision process (MDP) analysis. This approach directly incorporates information from the following:

- Markov component models utilizing degradation and failure rates from past experience
- Online monitoring and diagnostic information from the Argonne tool PRO-AID
- Analyses from a real-time plant Generation Risk Assessment (GRA)

In addition, the approach also includes an analysis of the potential impact on plant safety and licensing acceptability through an assessment utilizing a real-time Probabilistic Risk Assessment (PRA). Section 2 provides background information on MDPs and their solution methods, while Section 3 details the incorporation of each of the above factors into the MDP framework. The approach discussed here is currently #

2. MARKOV DECISION PROCESSES

A central aspect of the project is the utilization of an intelligent decision-making approach to optimize asset-management strategies. The system of interest (an advanced nuclear power plant) can be in any of a finite number of states, and the transitions between system states follow a Markov process. At discrete time steps, the decision-maker can take actions to influence system state transition. So, the transitions between system states depend not only on "nature," i.e., the inherent randomness in system state transition, but also on decision-maker actions. At each time step, different decision-maker actions and different system state transitions lead to varying rewards for the decision-maker. The decision-maker's objective is to maximize the sum of the rewards that will be received from the current time step into the future. For the current project, this problem is formally formulated as and solved by an MDP.

A Markov decision process can be formally defined by the following five elements, i.e., $\mathbb{S}$, $\mathbb{A}$, T, R, γ.

- $\mathbb{S}$: a discrete and finite space for the states of a system under study. A specific state in $\mathbb{S}$ at time step t is denoted by s_t.
- $\mathbb{A}$: a discrete and finite space for decision-maker actions. A specific action in $\mathbb{A}$ at time step t is denoted by a_t.
- T: $\mathbb{S} \times \mathbb{A} \times \mathbb{S} \rightarrow [0,1]$, the system state transition probability function, where $[0,1]$ is the interval between 0 and 1. For example, $T(s_{t+1}|s_t, a_t)$ denotes the probability of system state $s_{t+1} \in \mathbb{S}$ at time step $t+1$ given the system is in state $s_t \in \mathbb{S}$ at time step t and the decision-maker takes action $a_t \in \mathbb{A}$ at time step t.
- R: $\mathbb{S} \times \mathbb{A} \rightarrow (-\infty, +\infty)$, the reward function. For example, $R(s_t, a_t)$ denotes the reward that the decision-maker receives at time step t if the system is in state $s_t \in \mathbb{S}$ and the decision-maker takes action $a_t \in \mathbb{A}$.
- $\gamma \in (0,1)$: the discount factor used in the calculation of the cumulative rewards.

Take the application to the maintenance of a valve used in nuclear power plants as an example to illustrate the above notations, with decisions regarding the maintenance of the component made on a monthly basis. The state space describes the performance level of the valve, i.e., *perfect*, *degraded*, and *failed*. At each time step, the possible actions that maintenance staff can take may include *do nothing* and *repair*. The transition between system states depends on both the inherent randomness of the degradation of the valve, and the action taken by the maintenance staff. For example, if at time step t the valve is in the *degraded* state and the maintenance staff decides to *do nothing*, then the valve will be in the *failed* state at time step $t+1$ with probability 10^{-2}. However, for this same situation, if the maintenance staff takes action *repair*, the transition probability may be reduced to 10^{-3}. Depending on the system state and the decision-maker action at time step t, the decision-maker will receive a certain reward. For example, for the same decision-maker action, the reward in the case of a *perfect* state will typically be higher than the one in the case of a *failed* state. The discount factor γ may be determined as the real number that discounts future rewards back to the present value by referring to financial models used in maintenance management.

The decision-maker actions not only influence the reward for the current time step, but there are also the long-term repercussions of the action. In an MDP, given the system state s_0 at time step 0, the decision-maker aims to develop a policy σ that maximizes the expected discounted cumulative reward, i.e., the objective function.

The expectation in the objective function is over the possible transitions between system states. At each time step t, action a_t is taken according to the policy σ. The policy σ is defined as a mapping from the state space to the action space, i.e., $\sigma: \mathbb{S} \rightarrow \mathbb{A}$. The objective function takes into account both the reward to be gained for the current time step $t = 0$, but also the rewards to be received in future time steps. The discount factor γ is used to account for future versus current costs/rewards. The policy that maximizes the objective function is denoted by σ^*. The maximum expected discounted cumulative reward following the optimal policy σ^* is denoted by $J^*(s_0)$ for the starting state s_0. This maximum number is also called the value for state s_0, which is a key concept in algorithms for obtaining σ^*.

Dynamic programming serves as one of the most promising methods for solving an MDP problem, i.e., to obtain the optimal policy σ^*. Dynamic programming methods can typically be classified into two categories, value iteration based methods, such as the Bellman equation, and policy iteration based methods.

In the above discussion, the system state is fully observable or known. In practical applications however, the actual system state may only be partially observable, and can only be inferred through observations, such is often the case with equipment monitoring at nuclear power plants. Decision optimization under such situations can be formulated as and solved by partially observable Markov decision processes (POMDPs). A POMDP can be formally defined by the following seven elements, i.e., $\mathbb{S}$, $\mathbb{A}$, $\mathbb{O}$, P, O, R, γ. The only differences between a basic MDP and a POMDP is the addition of $\mathbb{O}$ and O, as below, with the other elements defined in the same way.

- $\mathbb{O}$: the observation space. The observations can be either discrete and finite, or continuous. A specific observation in $\mathbb{O}$ at time step t is denoted by o_t.
- O: the observation probability function. It can be defined as either $\mathbb{S} \rightarrow [0,1]$ for discrete observations where $[0,1]$ denotes the space of all possible probability mass values for discrete observations, or $\mathbb{S} \rightarrow [0, +\infty)$ for continuous observations where $[0, +\infty]$ denotes the space of all possible probability density values for continuous observations.

In a POMDP, the actual system state is not directly observable. So, the objective function in a POMDP is defined based on the belief in the system states. Denote the observations collected from time step 0 to time step t by $o_{0:t}$. The belief $b_t(s_t)$ in system state s_t at time step t is defined as the posterior probability of system state s_t at time step t, i.e., $p(s_t | o_{0:t})$.

The objective function for a POMDP problem is also defined as the expected discounted cumulative reward with initial belief b_0 at the beginning time step 0, i.e., the probability distribution over system states at time step 0. The expectation in the objective function is over the possible observations at each time step which in turn affect belief $b_t(s_t)$ at each time step. In an MDP, a policy is defined as a mapping from the system state space to the action space, i.e., $\sigma: \mathbb{S} \rightarrow \mathbb{A}$. But in a POMDP, a policy is defined as a mapping from the belief state space $\mathbb{B}$ to the action space $\mathbb{A}$, i.e., $\sigma: \mathbb{B} \rightarrow \mathbb{A}$. In the above objective function, $\sigma(b_t)$ denotes the action a_t provided by the policy σ for belief state b_t. Similar to the notations in an MDP, the optimal policy is denoted by σ^* and the maximum expected cumulative reward $J(b_0)$ following policy σ^* is denoted by $J^*(b_0)$, and is called the value for belief state b_0.

The optimal policy in a POMDP can also be obtained using value iteration or policy iteration, similar to the methods for an MDP. However, the major difference between solving an MDP and solving a POMDP is that in a POMDP the policy is a function of the belief state, i.e., the posterior probability distribution over system states, instead of the actual system state. In a value iteration algorithm for an MDP, the key to obtaining the optimal policy is to obtain the values for all system states. In the case of an MDP, the system state space is discrete and finite, so the values for system states can be enumerated. However, in the case of a POMDP, even for discrete and finite system state space, the belief state space is continuous. So, it is not feasible to enumerate the values for an effectively infinite number of belief states. A typical solution to this problem is to sample the belief state space and use these samples in the value iteration algorithm. The details of such algorithms are not provided in this paper, but can be found in the literature [1].

3. INTEGRATED MARKOV DECISION PROCESS APPROACH

As mentioned in Section 1, the approach for asset-management decision-making during plant operation requires the incorporation of multiple factors into the MDP/POMDP optimization framework. The steps before the MDP are necessary to supply the MPD calculation with the information required to form a real-time assessment of plant conditions. Fig. 1 provides an overview of the inputs and outputs of the different analysis segments.

First, sensor information from the operating plant is provided to the Argonne online diagnostic tool PRO-AID [2], which assesses component status based on the sensor data and physical system models. To inform this calculation, Markov component models provide additional insights regarding component behavior (component state probabilities based on estimated degradation/failure rates). Both PRO-AID and the Markov component models work in tandem to assess the condition of components within the system. The output of PRO-AID are real-time probabilities regarding current component status (healthy, degraded, failed, etc.). The output from PRO-AID and the Markov component models are utilized to develop a real-time plant risk profile, which consists of a PRA and a GRA. The PRA assesses plant risk from a safety perspective, while the GRA assesses economic risk. The output from PRO-AID, the Markov component models, and the real-time GRA are fed to the MDP analysis. The MDP analyzes different operational strategies to determine the optimal asset-management strategy to maximize revenue (in green). The selected asset-management strategy(ies) is assessed by the real-time PRA to ensure acceptability from a safety and licensing perspective. How these differing aspects are integrated into the integrated MDP analysis is detailed in the following subsections.

3.1. Integration of Markov Component Models

In probability theory, a Markov model is a stochastic model used to model randomly changing systems. It is assumed that future states depend only on the current state, not on the events that occurred before it (that is, it assumes the Markov property). For any given component, a Markov model consists of a list of the possible states the component could be at any specific time, the transition paths between those states, the rate parameters of those transitions, and the initial conditions describing the chance of the component to be in the states at some initial time point.

The Markov model of a real component typically includes a fully operational state (*Healthy*) and a set of intermediate states representing partially failed condition (*Degraded*), leading to the fully failed state, i.e., the state in which the component is unable to perform its design function (*Failed*). The model may include repair transition paths as well as failure transition paths. An example of how Markov analyses can be used to model nuclear reactor components can be found in ref [3].

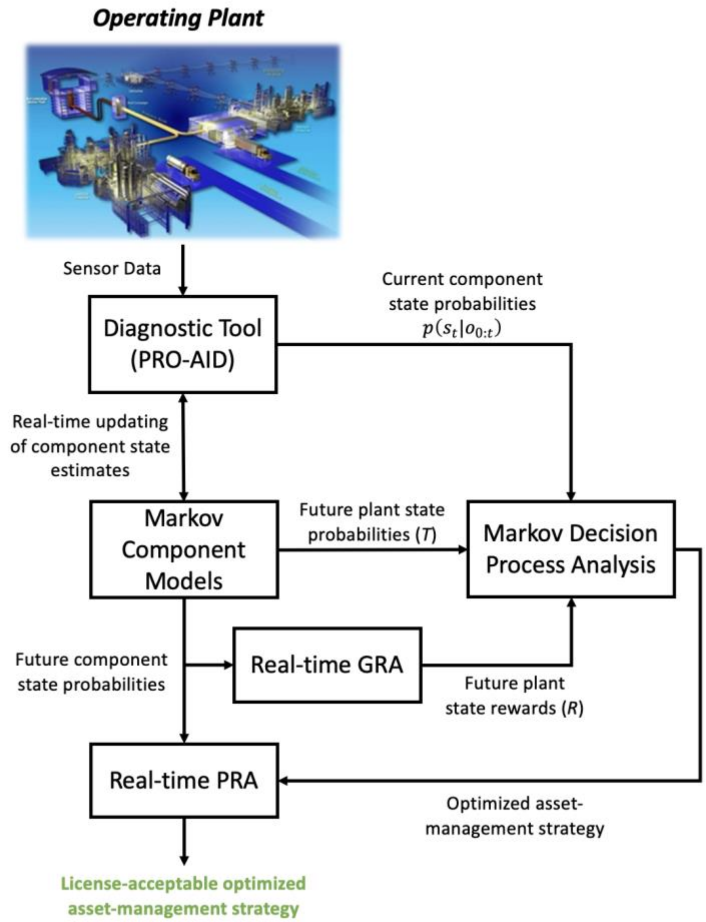

Figure 1: Overview of Integrated MDP Analysis Approach

The ability of Markov models to estimate the future state(s) of components is an important factor in the MDP asset-management framework, as the change in component status will impact the likelihood of being in different plant states in the future. As outlined in Section 2, MDPs include a set of transition probabilities (*T*) that depict the likelihood of the system transitioning from one state to another during the next time interval. In the developed approach, the Markov component models are utilized to develop the transition probabilities. While the Markov component models only assess a single component in the plant, the results of the individual component models can be combined to provide a comprehensive assessment of the likelihood of the plant transitioning states. To complete this analysis, each individual Markov component model is solved for the next time interval under consideration, in a manner consistent with the assumptions of that action. For example, if the action is to repair that component, then that is reflected in the state of the Markov component model. It is important to note that the individual Markov component models can be developed using a variety of techniques, such as the use of historical operating data or mechanistic component models. This allows users to leverage pre-existing component models within the framework in a plug-and-play manner.

Within the overall framework, the derived state probabilities for each component at each time interval are integrated into a total plant model to become the transition probabilities in the MDP framework (transition function T in Section 2). In this way, the likelihood of the plant being in a future state is directly integrated with the associated Markov component models.

3.2. Integration of Generation Risk Assessment

Nuclear power plant operators require tools to help make decisions involving the operation and maintenance of equipment whose failure can cause reactor trips or down-power events. A GRA is the process of predicting the risk of generation loss during future operation by estimating the probability and duration of plant trip or derate due to equipment degradation or failure [4]. A GRA model is an important element of nuclear asset management risk-informed tools for analyzing effects of equipment reliability and availability on plant value and resource allocation decision-making.

Central to the assessment of generation risk is the development of a trip model. A trip model is similar in function and construction to that used for PRA with the exception that the end-state of the trip model is the frequency of plant trip as opposed to the frequency of core damage or offsite dose consequence. The trip model is generally used to estimate the frequency of instantaneous trip based on actual plant configuration and condition. Another model important for a GRA analysis is a derate power model, i.e., a model where the end-state is the frequency of plant to operate at decreased (derated) power level. When built, the two models help to identify different plant states and the awards (generation) associated with them which are key input parameters for any asset management decision at the plant level.

For the MDP asset-management decision-making framework, the integration of the GRA provides insight into the potential reward associated with different plant states (variable R in Section 2), in terms of generation and revenue. Consider a K-state plant that consists of N subsystems. At any particular time t the plant can be in any one of K plant states with corresponding probability $P_k(t)$ and the reward associated with k^{th} state r_k. The overall plant reward at time t can in general be estimated as

$$R(t) = \sum_{k=1}^{K} r_k \cdot P_k(t)$$

where,

r_k – plant reward for being in state k, $k = 1, 2, \dots, K$

$P_k(t)$ – probability for the plant to be in state k.

Strictly speaking, P_k is a function of not only time t, but also of $p_i(t)$ – probability for subsystem i ($i = 1, 2, .., N$) to be in a state that would allow the plant to be in state k. The exact formula to calculate $P_k(t)$ is strongly dependent on configuration and arrangement of the plant subsystems and is derived from that knowledge for each particular case. The aforementioned trip and derate models can help with estimating plant state probabilities. In the approach proposed in this project, the fault tree method to build the derate/trip model is applied (i.e., one fault tree for each plant state so, that probability of the plant to be in a particular derate/trip state would be the top event probability of fault tree corresponding to that state). One of the benefits of such approach is that such models can be constructed rather easily, and once built, are straightforward to use.

A real-time GRA is developed utilizing the component status information derived from the combination of the Markov component models and real-time diagnostic tool (discussed in the Section 3.3). The function P_k can be found from the Markov component models discussed in Section 3.1 and the associated reward with each state, r_k, is known from the real-time GRA. Therefore, $R(t)$ can be found for each time interval and integrated into the MDP framework.

3.3. Integration of Real-time Diagnostic Information

Optimized asset-management decision-making relies on an accurate, real-time picture of plant conditions. In the developed methodology, this process begins with the sensor network within the operating plant. The sensor network provides information to PRO-AID, which assesses the data and diagnoses component operating states. By combining sensor data with physics-based models, PRO-AID has the capability to discriminate sensor-level from component-level faults. A brief overview of the PRO-AID method and linking to the Markov component models is provided here, with further detail to be provided in future publications.

The PRO-AID framework consists of quantitative model-based diagnosis, statistical change detection and probabilistic reasoning that allows detecting both component faults and sensor faults. The use of physics-based diagnostic models provides high detection sensitivity and allows noise and measurement uncertainty to be incorporated robustly. For each component model, there are two major sources of uncertainty that need to be taken into account, i.e., (1) the measurement uncertainty (uncertainty in the reading value of each sensor), and (2) the modeling uncertainty (uncertainty in the output of each component model). As a result, the computed residuals, i.e., difference between the model predictions and the measurements, will be affected by uncertainty. To this aim, Bayesian inference is used to detect and localize possible faults starting from the observed fault symptoms, i.e., the relation between the posterior probability and prior probability of a fault can be written as the following equation. A generalized version of this formula is implemented to account for multiple-fault event scenarios.

$$P(S|O) = \frac{P(O|S) \cdot P(S)}{P(O)}$$

where,

O: The set of observed zero/non-zero residuals

P(S): The prior state probabilities

$P(O|S)$: The likelihood of the observed data, i.e. the probability to have observations described by O for state S to occur

$P(O)$: The probability to have the observations in O regardless of whether any faults occurred

$P(S|O)$: The posterior probability of the fault, i.e., the probability that state S has occurred given the observed residuals O

As a result of the online monitoring process, at every time-step, PRO-AID evaluates the conditions of sensors and components, i.e., operating, faulted, degraded. In case inconsistencies between the set of observations and the fault-free system model are detected, the compatible fault scenarios are identified. For each one of them, the posterior probabilities are calculated by using the Bayesian network method. Possible diagnoses are then ranked in a meaningful order so that unlikely events can be eliminated.

As was outlined in the Section 3.1, the probability of the component to be in a given state at a future time can be calculated from the Markov component model. The obtained component state probabilities are passed to PRO-AID to be utilized as the prior probabilities for its internal Bayesian models. PRO-AID then utilizes sensor readings to perform a Bayesian update and return posterior state probabilities to the Markov component models. The component state probabilities are then used within the Markov models to predict the state probabilities at the next time-step. The component status information is then passed to the real-time GRA and PRA, which are updated to reflect the real-time risk profile of the plant, and can then be used to estimate the risk profile of the plant at a future time.

The linking of PRO-AID and the Markov component models provides a real-time estimate of the status of plant components. This information can be utilized in the MDP to inform the current plant state (s_t in

Section 2). The result is a POMDP, since the current state is not known but only estimated. As described in Section 2, in a POMDP, observations *O* of the system state are made but are uncertain. This directly aligns with the estimated state probabilities that are obtained through the utilization of PRO-AID. The POMDP can still be solved to identify an optimal policy. However, now the policy is also dependent on the uncertainty regarding the current system state.

3.4. Integration of Probabilistic Risk Assessment Information

Incorporation of the PRA into the intelligent asset-management decision-making approach is a necessary step to ensure that plant operations remain within acceptable safety bounds. The insights from the PRA provide critical insights into the acceptability of proposed asset-management strategies. To accomplish this task, the real-time PRA is utilized in conjunction with the risk-informed performance-based licensing approach of the Licensing Modernization Project (LMP) [5].

The use of PRA in reactor licensing has evolved since its introduction in the 1970s. As the use of risk-informed decision-making in regulatory matters has gradually expanded, a new risk-informed performance-based licensing approach (the LMP approach) has been developed and recently endorsed by the NRC for use by advanced reactor vendors [6]. The LMP approach utilizes risk information to inform key licensing decisions, such as the identification and categorization of licensing basis events (LBEs), the safety classification of structures, systems, and components (SSCs), and the evaluation of the adequacy of defense-in-depth.

As the current project focuses on the economics of advanced reactors, the implementation approach assumes the utilization of the LMP process for reactor licensing. Of particular importance for the current work is the LMP's use of the frequency versus consequence (F-C) curve, shown in Fig 2. When utilizing the LMP approach, event sequences or event sequence families from the PRA are plotted on the F-C curve. The frequency of the event sequences is utilized for LBE categorization (anticipated operational occurrence – AOO, design basis event – DBE, beyond design basis event – BDBE). In addition, the event sequences and their associated offsite dose are compared to a consequence target, which is developed based on applicable regulation and guidance. The distance from the event sequences to the consequence target helps demonstrate the available safety margin of the design. The F-C curve also has a designated "risk significant" region, which identifies LBEs that may be located close to the consequence target. The F-C curve is also utilized to aid in the safety classification of SSCs. The importance of different SSCs is determined through PRA sensitivity studies, such as not crediting an SSC (*i.e.,* assuming it is in the failed state) and assessing the impact of event sequence placement on the F-C curve. For example, if not crediting an SSC results in an event sequence exceeding the consequence target, then that SSC may be designated as "Safety Related," the highest safety classification for SSCs.

For the project approach, the results of the real-time PRA provide important insights into plant safety and licensing. The real-time PRA incorporates online monitoring and diagnostic information to provide updates of failure probabilities and component status. At a high-level, the real-time PRA provides an avenue for assessing the safety impact of different operational strategies. Since the LMP approach to licensing is fundamentally risk-informed, changes to the PRA and the associated plant risk profile can provide direct insights into the status of the plant within its licensing basis.

With the real-time PRA established, different asset-management strategies are assessed throughout reactor operation utilizing the MDP approach described in the previous sections. The result of the MDP analysis is one or more asset-management strategies that are optimized to maximize revenue for the plant. However, the MDP calculation itself does not provide a bound on available actions from a safety perspective. Therefore, after the optimization process has taken place, the strategy must be analyzed by the real-time PRA to assess whether safety/licensing constraints would be violated over the next operating window. The frequency and consequence of potential event sequences is updated based on the planned status of components and systems in the plant. This includes actions taken immediately but also those planned in the future. The full list of impacted parameters includes:

- Mission time of analysis (*i.e.*, operating period of the reactor, includes planned downtime associated with strategy under examination)
- Status of components (in service, out of service, degraded state, repaired, etc.)
- Failure rate of components (depending on state, repair/maintenance, etc.)

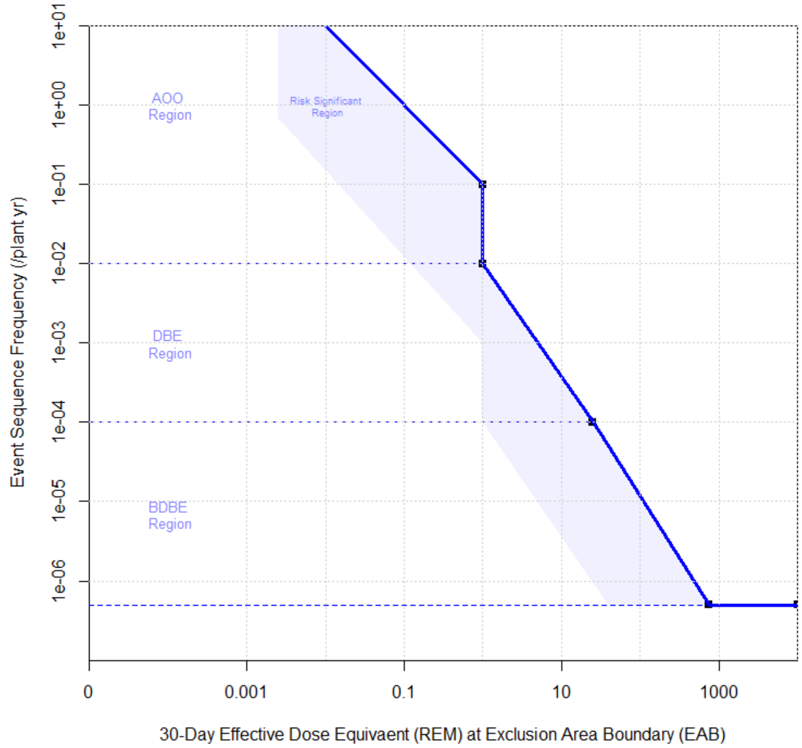

Figure 2: LMP Frequency versus Consequence Curve[1]

Each asset-management strategy is modelled in the real-time PRA for the future operational period, such as the time until next plant shutdown. The locations of LBEs on the F-C curve is tracked to ensure that no LBEs exceed the consequence target of the curve and therefore may violate the operational bounds of the reactor. In addition, the change in the frequency and consequence of LBEs is also assessed to gauge any potential impact on SSC classification, as this could have ramifications on plant technical specifications.

4. CONCLUSION AND NEXT STEPS

Advanced reactors offer a promising pathway to address rapid climate change but large-scale deployment faces a challenging economic environment in competitive energy markets. To improve the economic competitiveness of advanced reactors, an effort was undertaken to develop an approach to optimize asset-management strategies during reactor operation. The result is an integrated MDP analysis, which directly incorporates key information from Markov component models, online monitoring and diagnostics, and a real-time GRA. In addition, a real-time PRA provides an avenue for assessing the safety and licensing acceptability of the proposed strategies.

---

[1] Derived from ref [5].

The developed approach benefits from its integrated nature, which incorporates multiple streams of available information into the decision-making process to conduct a comprehensive assessment. However, the approach is also flexible and scalable, as it can be tailored to focus on individual systems or plant-wide behavior. Similarly, the approach can directly incorporate any available model for system/component behavior, such as reliability analysis based on historical data or mechanistic system analysis, or combinations of models.

A comprehensive demonstration problem utilizing an advanced reactor design has been completed, with a high-level summary provided in ref [7]. The demonstration problem conducted each step of the developed approach for a steam-cycle HTGR design, with a focus on operational strategies of the feedwater system. The operational flexibility of the steam-cycle HTGR provides a promising but also challenging setting for asset-management decision-making, given the numerous potential options and complexity of the analysis. Complete detail on the demonstration problem will be provided in future publications.

ACKNOWLEDGEMENTS

Argonne National Laboratory’s work is supported by the US Department of Energy, Office of Nuclear Energy under contract DE-AC02-06CH11357. This work was prepared as an account of work sponsored by the U.S. Department of Energy, Office of Nuclear Energy Advanced Sensors and Instrumentation program under DOE Contract DE-NE-19-17045. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof. The approach described in this paper is patent pending.

REFERENCES

[1] M. Spaan and N. Vlassis, "Perseus: Randomized point-based value iteration for POMPDs," *Journal of Artificial Intelligence Research,* vol. 24, pp. 195-220, 2005.

[2] T. N. Nguyen, T. Downar, and R. Vilim, "A Probabilistic Model-Based Diagnostic Framework for Nuclear Engineering Systems," *Annals of Nuclear Energy,* vol. 149, no. 15, 2020.

[3] K. Fleming, "Markov Models for Evaluating Risk-Informed In-service Inspection Strategies for Nuclear Power Plants Piping Systems," *Reliability Engineering and System Safety,* vol. 83, no. 1, pp. 27-45, 2004.

[4] Electric Power Research Institute, "Generation Risk Assessment (GRA) Plant Implementation Guide," 1008121, 2004.

[5] Nuclear Energy Institute, "Risk-Informed Performance-Based Guidance for Non-Light Water Reactor Licensing Basis Development," NEI 18-04 Rev 1, 2019.

[6] U.S. Nuclear Regulatory Commission, "Guidance for a Technology-Inclusive, Risk-Informed, and Performance-Based Methodology to Inform the Licensing Basis and Content of Applications for Licenses, Certifications, and Approvals for Non-Light-Water Reactors," RG 1.233, 2020.

[7] V. Moiseytseva *et al.*, "Cost-Benefit Analyses through Integrated Online Monitoring and Diagnostics: Final Report," ANL/NSE-23/22, 2023.