Prioritizing Diverse Observations or Issues from Safety Reviews at Nuclear Power Plants According to Possible Safety Impacts

Ivan Vrbanić, Ivica Bašić

Abstract— Being a facility with potential for radioactive release, any nuclear power plant (NPP) is, over its operating life time, permanently subject to numerous safety reviews with different scopes and objectives. The reviews may be initiated and implemented by various stakeholders, including regulators, utilities or industry. Some of them are, by their nature, general and extensive in terms of different safety areas or safety attributes which are covered. An example of such a review is a Periodic Safety Review (PSR) which is promoted by the International Atomic Energy Agency (IAEA) and a number of national safety authorities in Europe and worldwide. The others may, depending on the objective, be targeted at particular safety area (e.g., ageing management or equipment qualification or safety analyses). Both of the mentioned cases (single general review or multiple targeted reviews over a time period) can generate an inventory of observations (“findings”) or “issues” which need to be addressed but may be very different in their nature and implications, as well as in benefits or resources associated with their resolutions. For some issues a resolution may be straightforward. For others, it may require a feasibility study and identification of options for possible resolution. Also, in some cases the resolution is simply a “must” (e.g., discrepancy from licensing basis) while in some other cases it may be a matter of balance (e.g., effectiveness of maintenance program). Furthermore, while some of the issues may be directly related to operational safety (e.g., non-compliance with single failure criterion or aging-related degradation of safety features), for some others the link to operational safety may not be explicit (e.g., comparison of safety bases against the newly emerging methodologies or issues observed with regard to so called “soft factors”). The paper discusses types of different observations or issues which may come from general or targeted safety reviews and outlines some basic principles for their comparison and prioritizing with regard to possible safety impacts, which is many times needed for the purpose of developing an action plan for safety improvements.

Keywords — operating NPP, safety review, safety issues, prioritization, ranking.

I. INTRODUCTION

ANY operating nuclear power plant (NPP), as a facility with potential for radioactive release, is subject to numerous safety reviews with different purposes and objectives. Some of the safety reviews are, by their nature, general and extensive in terms of different safety factors or safety attributes which are covered. An example of such a review is a Periodic Safety Review (PSR) which is promoted by the International Atomic Energy Agency (IAEA) and a number of national safety authorities, [1], [6], [14], [15] and [16]. PSR is many times used as a means for verifying whether a plant which has been operated in long term (e.g., over decades) is still as safe as originally intended and particularly in the context of new safety standards which have come into place after the time of plant’s initial operation. The other reviews may, depending on the objective, be targeted at particular safety factor (e.g., ageing management or equipment qualification or safety analyses). The reviews may be initiated and implemented by various stakeholders, including utilities, industry and regulators. Both of the mentioned cases (single general review or multiple targeted reviews over a time period) can generate an inventory of safety issues which need to be addressed but may be very different in their nature and implications, as well as in benefits or resources associated with their resolutions.

Safety issue generated by a review can usually be characterized in terms of the three general attributes:

• Directly connected to nuclear safety (for considered issue, a direct link to nuclear safety can be established; for example, an observed deviation from the requirements relevant for nuclear safety); the category will here be referred to as “DS”;
• Re-evaluation of nuclear safety basis (e.g., adequacy with regard to safety assessment standards or methods); the category will here be referred to as “RS”;
• Related to “soft factors” (e.g., human factors engineering, organization for safety, safety culture and similar) or non-nuclear safety issue (e.g., industrial hazard); the category will here be referred to as “SF”.

Each issue can usually be related to at least one of these three general attributes. Some issues may relate to more than one.

When assessing and prioritizing (ranking), a separate evaluation path would, in principle, need to be applied with regard to each of the three general attributes. The key is that ranking with respect to particular general attribute results with a certain (nume-
rical) score, and that the scores for the three general attributes are directly comparable to each other, on the same scale. The score can be assigned by passing the issue of concern through multiple layers of ranking evaluation, with respect to predefined criteria. In the case that more than one ranking evaluation path (general attribute) applies, a rule can be set to take the highest score as a final issue rank.

One possible approach is generally discussed below.

II. OUTLINES OF A GENERAL APPROACH

So-called ‘defense in depth’ (DID) is a basic concept in nuclear safety. It was defined as a philosophy to ensure that successive measures are incorporated into the design and operating practices for nuclear plants to compensate for potential failures in protection and safety measures, [2],[3], [4] and [5]. All safety activities, whether organizational, behavioral or equipment related, are subject to layers of overlapping provisions, so that if a failure should occur it would be compensated for or corrected without causing harm to individuals or the public at large. This idea of multiple levels of protection is the central feature of DID.

It then comes as natural to establish the approach for ranking of diverse safety issues in a way that it is based on assessing potential impact on DID. In its essence, the approach would consist of “measuring” or assessing the depth of remaining defense or remaining mitigation capability, provided that considered safety issue remains unaddressed. The approach would map the issue of concern into the DID structure (e.g., by failing or reducing the affected barrier capability) or into the accident sequences (e.g., by failing or reducing the affected mitigation function capability). In principle, both deterministic and probabilistic methods can be used for the purpose. Particular attention is to be paid to robustness of individual levels of defense and to mutual independence of levels of defense as important properties of DID.

The principles and elements of such approaches were, to various extents and levels of detail, used worldwide. For example, the approaches where remaining defense depth is estimated by counting of levels of defense were used to assess the safety of existing nuclear power plants and their elements are described in the IAEA publications such as [7], [8] and [9]. Similarly, the approaches where remaining mitigation capability is estimated (qualitatively, in terms of orders of magnitude) by mapping of the issue of concern into the relevant accident sequences are used in the US NRC Significance Determination Process (SDP), as originally described in references [10], [11] and [12]. Also, a similar approach is used in the industry risk informed applications (recognized also by the regulators). An example is risk informed in-service inspection where such approach is used to determine the remaining mitigation capability (or conditional risk) following an assumed failure or degradation of a pipe segment, [13].

In the case of a larger number of issues to be addressed (ranked) it can be expected that the process would be, for practical purposes, divided into two major steps, which can here be referred to as:

1. Broad ranking evaluation; and
2. Detailed ranking evaluation.

Broad ranking evaluation can be used for grouping of similar issues as well as for reviewing any particular issue in the light of other issues (from different areas of concern, such as PSR safety factors) which may be co-related. Sometimes, this may provide a different perspective on issue importance. By broad ranking, all issues from the inventory would be typically classified into several general categories of importance such as, for example: high (H), medium (M) and low (L).

The results of the broad ranking can also be used for initial pre-screening done in order to identify the issues which can be directly sent to a corrective action program (CAP) normally existing in any NPP. Pre-screened issues which can be a direct input to a CAP (without further ranking evaluation) usually are of the two types:

- Issues requiring immediate attention and short-term resolution such as those representing Technical Specification violation or violation of current licensing basis; in principle, such issues, if any, would usually be broadly ranked with high (H) importance; for such issues no ranking is needed as their implementation is a “must”; this group of issues will here be referred to as “IRR” (immediate resolution requirement);

- Issues desirable to be resolved, which can be resolved at minimum effort and in a short time frame. Although these can come from any importance category, it is expected that most of them would come from low significance (L) category, relating to matters such as changes to procedures in non-safety domain, corrections to plant drawings or documents and similar. This group of issues will here be referred to as “LSE” (low significance and effort).

All the remaining issues (i.e., those surviving pre-screening) would be subject to a detailed ranking. Some of them would be later input into a CAP, based on the rank and used criteria. General flow chart is illustrated by Fig. 1.

Detailed ranking would, in principle, be based on assessing each issue against the three general attributes discussed above, “DS”, “RS” and “SF” and corresponding predefined ranking criteria. In the process, particular issue would be initially related to one of these three general attributes (usually, most of the issues) or to more than one (usually, limited number of issues). It can be assumed that for each of those issues which are related to multiple general attributes the dominant general attribute can be identified, i.e., the one which would result with the highest score with regard to nuclear safety. Certain regrouping / subsuming of the issues may be done with regard to this. The effect would be that each issue in its finally defined form would be, before entering the detailed ranking evaluation, related to a single general attribute, “DS”, “RS” or “SF”. Thus, the issues would be effectively divided into three groups (as related to the three general attributes) and for each attribute a separate ranking evaluation path would apply. Passing through the respective evaluation path would result with a certain score assigned to the issue considered. A range of significance (a range of numerical values for the score achieved) would be defined for each of the three paths / general attributes in a way that the same scale would apply for all three paths. (For example, numerical value “3” for a score would represent the same safety significance regardless of whether the path was “DS”, “RS” or “SF”.)

![Fig. 1. General Flow Chart for Issue Ranking](https://example.com/fig1.png)
For resolution of any issue which would, as a result of the detailed ranking process, be input into a CAP, an appropriate corrective measure would need to be foreseen. Thus, in principle, each particular issue which will be, in the context of Fig. 1, forwarded to CAP, can be associated with particular corrective measure. Thus, the issue ranking and corrective measures prioritization become single process. In this process the issue achieves its numerical score through the evaluation path defined by the general attribute which was assigned to the issue, “DS”, “RS” or “SF”. Whichever the general attribute, this numerical score (final rank) would be achieved by passing the issue through several layers of ranking evaluation.

Having in mind the discussion at the beginning of this section, the first-layer evaluation in each path would be related to DID impact. The first two general attributes, “DS” and “RS”, relate to DID either directly (“DS”) or through the adequacy of DID assessment (“RS”).

Therefore, the “DS” and “RS” evaluation paths can be associated with DID impact in a rather straightforward manner: by deterministically considering the status of “lines of defense” (LOD) (e.g., evaluation of safety margins) or by risk assessment (as assessed risk reflects the status of the LODs), whichever is considered more appropriate. Either evaluation would assess the significance of an issue on “depth of defense” (DOD) and it can be evaluated qualitatively or quantitatively, depending on the issue addressed.

For the “SF” evaluation path, the relation to DID may not always be straightforward. However, it can usually be established through an assessment of issue’s impact on potentially affected plant’s “operational safety features” (OSF) which would in turn reflect on DID.

All these terms are further characterized in the sections below. But, in order to summarize the first-layer evaluation which relates to DID, it can be said:

- In the “DS” or “RS” path the significance of an issue is evaluated with respect to its DOD / risk impact;
- In the “SF” path the significance of an issue is evaluated with respect to its OSF impact.

The above mentioned first-layer evaluation provides means for primary ranking of safety issues. Finer sub-ranking can be achieved through the second-layer evaluation. As an example, the basis for the second-layer evaluation may be:

- Evidence from operating experience.

This kind of basis can be applied to all three paths.

Yet finer sub-ranking may be done by the evaluation of resources needed to resolve the issue of concern (i.e., budget for the respective corrective measure). This can be done assuming that all the “must-be-resolved” issues were input directly into CAP as the “IRR-type” issues discussed above. Thus, a third-layer evaluation of an issue can be foreseen on the basis of:

- Cost category.

For the purpose of issue ranking / corrective measures prioritization process this kind of evaluation would apply to all three paths and may be done as a simple qualitative cost category estimate.

Altogether, the process would result with a final score or priority for each issue / corrective measure. It is illustrated with generalized flow chart shown in Fig. 2 and Fig. 3.

### III. Broad Ranking Evaluation

During the broad ranking process, particular safety issue from any review area (e.g., safety factor in [1]) needs to be viewed in the light of other issues (from the same or from different review areas / safety factors) which may be co-related. This can be done through the process of grouping of issues from different review areas / safety factors, with correlated subjects (considering, also, associated possible corrective measures). The process would, usually, involve expert engineering judgment to certain extent, which would be facilitated if a ranking analysts’ team includes members with experi-

---

Ivan Vrbanić, Ivica Bašić, Prioritizing Diverse Observations or Issues from Safety Reviews at Nuclear Power Plants According to Possible Safety Impacts, Journal of Energy, vol. 71 Number 1 (2022), 27–33

https://doi.org/10.37798/2022711407
se and experience related to all review areas / safety factors. Fig. 4 shows an example of a general flow chart for the interface impact assessment for this purpose.

As important part of the broad ranking evaluation, the issues would be grouped according to the three general attributes, i.e., for each issue it would be defined which detailed ranking evaluation path (“DS”, “RS” or “SF”) will apply. As shown in Table I below, different assessment and criteria at broad ranking would be applied for general attributes “DS” / “RS” as compared to “SF”. The same table outlines high-level criteria for assigning low (L), medium (M) or high (H) safety significance to issues considered.

<table>
<thead>
<tr>
<th>General Attribute</th>
<th>Description of assessment criteria:</th>
<th>Score:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low (L)</td>
<td>Medium (M)</td>
</tr>
<tr>
<td>DS or RS</td>
<td>Impact on plant safety systems (a)</td>
<td>Small (S)</td>
</tr>
<tr>
<td></td>
<td>Impact of issue on the physical barrier to the release of radioactivity (b)</td>
<td>Affected (A)</td>
</tr>
<tr>
<td></td>
<td>Impact of issue on one or more levels of defense (c)</td>
<td>Affected (A)</td>
</tr>
<tr>
<td>SF</td>
<td>Level of safety culture or organizational safety or associated subject (d)</td>
<td></td>
</tr>
</tbody>
</table>

It is expected that broad ranking would involve subject matter experts for particular review area (e.g., safety factor in [1]).

IV. DETAILED RANKING EVALUATION – FIRST LAYER

IV.1 GENERAL ATTRIBUTES (RANKING EVALUATION PATHS) “DS” AND “RS”

As mentioned above, for “DS” and “RS” path the first-layer evaluation represents an assessment of significance of an issue with regard to its DOD / risk impact. As an example, for DOD impact assessment a limited set of qualitatively defined end-states (pre-defined outcomes of an assessment) can be defined such as:

- “D1”, Non-Relevant or Awareness Needed;
- “D2”, Tolerable Long Term;
- “D3”, Tolerable Short Term;
- “D4”, Not Tolerable.

The assessment (and respective criteria) for DOD impact evaluation can be based on the approaches such as counting LODs, e.g. [7] or [8]. In this context an LOD is defined as a system, barrier or human action (or combination of those) needed for providing protection against an initiating event. In principle, LODs exist at any or several of the levels of defense in depth. Considered issue is imposed upon the existing LODs and the remaining LODs (accounting for the issue impact) are then identified and counted. The result of an evaluation is then expressed as one of the above end-states.

If risk impact assessment is used instead, it can be based on the approaches such as Significance Determination Process, [10], [11] and [12], or on industry approaches such as the lookup tables in [13]. Same or similar set of predefined end-states can be used as above.

Fig. 5 below is an illustrative example of a set of criteria for assigning the end-states in the DOD impact assessment, inspired by [7]. (“S” refers to a “strong” LOD while “W” refers to a “weak” LOD.) It is noted that the figure is illustrative and without pretension to establish the actual criteria or definitions of “strong” and “weak”.

In the case that risk impact assessment is used instead, a set of criteria can be derived from the mentioned US NRC SDP or directly based on Regulatory Guide 1.174, [17].

![Fig. 5. Illustration of Criteria for DOD Impact Assessment](image-url)
Fig. 6. Concept for Counting LODs

Once the DOD end-state (e.g., D2) is determined for particular issue, it would be translated to numerical score, as indicated by Fig. 7. (The analogous would be done in the case that risk impact states were used, instead.) Here, it is noted that in the case of D1 (green color in Fig. 5) additional sub-ranking criteria are provided since, based on the past experience, it is expected that considerable portion of the issues would be assigned this end-state. It is important to point out here that for the same end-state a numerical score on the “RS” path would, in principle, be lower than on the “DS” path. For example, for the state D2 a numerical score on “RS” path would be lower than on “DS” path. This is because in the case of “DS” an issue relates to the actual status of DOD (e.g., aging issue or environmental qualification issue) while in the case of “RS” an issue relates to potential or indicated status of DOD (which may or may not be confirmed when actual re-evaluation of safety basis is done).

<table>
<thead>
<tr>
<th>DOD State</th>
<th>Functionality Impact</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>D4</td>
<td></td>
<td>xx</td>
</tr>
<tr>
<td>D3</td>
<td></td>
<td>xx</td>
</tr>
<tr>
<td>D2</td>
<td></td>
<td>xx</td>
</tr>
<tr>
<td>D1</td>
<td>D1.3 Affects safety system or feature involved in D6 accident sequence</td>
<td>xx</td>
</tr>
<tr>
<td></td>
<td>D1.2 Affects system or feature involved in BDB accident sequence</td>
<td>xx</td>
</tr>
<tr>
<td></td>
<td>D1.1 Impact limited to non-safety system not involved in BDB accident sequence</td>
<td>xx</td>
</tr>
</tbody>
</table>

Fig. 7. DOD States versus Numerical Score (Rank) for Path “DS” or “RS”

IV. II General Attribute (Ranking Evaluation Path) “SF”

As mentioned above, for the “SF” path the first-layer evaluation represents assessment of issue’s significance with regard to operational safety features (OSF) impact. For this purpose, many times the OSFs can be divided into a limited set of categories such as:

- Safety Management Systems;
- Radiological Protection and Other Occupational Hazards.

For assessment of an issue against each of the categories there are a number of techniques which can be used, such as those based on the gap analysis, check lists, and others. Once the assessment of particular issue is done, the issue of concern would be assigned a numerical score. An example of high-level criteria for OSF is given in Fig. 8. The range of the numerical score values needs to be calibrated against the numerical score values for “DS” and “RS” paths (shown in Fig. 7).

<table>
<thead>
<tr>
<th>OSF Impact</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant shortfall identified or major element missing</td>
<td>xx</td>
</tr>
<tr>
<td>Gaps or areas not fully addressed</td>
<td>xx</td>
</tr>
<tr>
<td>Minor areas to be improved in order to bring in line with best international practice</td>
<td>xx</td>
</tr>
<tr>
<td>Minor issue which does not influence safety-related SSC, plant security or workers safety</td>
<td>xx</td>
</tr>
<tr>
<td>Impact limited to operational excellence, external confidence or public image of the plant</td>
<td>xx</td>
</tr>
</tbody>
</table>

Thus, upon the completion of the first-layer evaluation the issue is given a first-layer score taken from the numerical values “xx” from Fig. 7 or Fig. 8 depending on the general attribute (“DS”, “RS” or “SF”) which was the basis for the evaluation. Since the scoring range is calibrated, scores of all different issues can be directly compared and, hence, prioritized.

V. Detailed Ranking Evaluation – Second Layer

For all issues input into the detailed ranking evaluation and being given a first-layer score finer sub-ranking can be achieved through the second-layer evaluation. As discussed above, a basis for the second-layer evaluation may be:

- Evidence from operating experience.

The evaluation of evidence would consider problems encountered with respect to a particular safety issue identified in plant-specific experience, plants of similar design (e.g., vendor owner’s group), or generic industry-wide sources.

Table II shows an example of the criteria and significance scale for the evidence from operating experience. The significance criteria / scale can be applied to issues from all ranking paths (i.e., “DS”, “RS” and “SF”).

TABLE II

IIllustration for Significance Scale for Evidence from Operating Experience

<table>
<thead>
<tr>
<th>Evidence from Operating Experience</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue identified as plant specific problem</td>
<td>3</td>
</tr>
<tr>
<td>Issue identified as vendor’s group generic problem (no plant specific identification)</td>
<td>2</td>
</tr>
<tr>
<td>Issued identified as generic industry-wide problem exclusively</td>
<td>1</td>
</tr>
</tbody>
</table>

Expert judgment and consultation with plant subject matter experts (SME) can be utilized to determine the plant specific, vendor-specific, and generic industry-wide information sources for a given safety issue. This evaluation would provide a second-layer
numerical score for the issue, which may be used for sub ranking of the issues with the same first-level score.

VI. Detailed Ranking Evaluation – Third Layer

The issues with the same first-layer score and second layer score may further be sub-ranked. As already mentioned above, this can be done by considering the resources needed for resolution of particular issues (i.e., budget for the respective corrective measure). Thus, a third-layer evaluation can be performed which would be based on:

- Cost category.

Like in the case of evidence from operating experience under the second layer, this kind of evaluation can be applied under all three paths (“DS”, “RS” and “SF”). It may be done as a simplified qualitative cost category estimate. An illustration is provided by Table III. (It should be noted that, from the cost perspective, favorable are those measures which are cheaper.)

Consideration of cost only at third layer reflects the fact that safety concerns (as well as problems from the past operating experience) are given higher priority over the budgetary concerns. Issues are always ranked by safety importance first.

<table>
<thead>
<tr>
<th>Qualitative Cost Evaluation</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than X1 (EUR)</td>
<td>4</td>
</tr>
<tr>
<td>Between X1 (EUR) and X2 (EUR)</td>
<td>3</td>
</tr>
<tr>
<td>Between X2 (EUR) and X3 (EUR)</td>
<td>2</td>
</tr>
<tr>
<td>Greater than X3 (EUR)</td>
<td>1</td>
</tr>
</tbody>
</table>

VII. Concluding Remarks

The above outlined process would result with each issue / corrective measure being assigned significance in the form “f.s.t” where:

- “f”: numerical score with regard to a first-layer attribute, DOD / OSF;
- “s”: numerical score with regard to a second-layer attribute, past experience;
- “t”: numerical score with regard to a third-layer attribute, cost of resolution.

Each of the three numerical scores is positive integer (e.g., 1, 2, 3…).

All the scores can then be sorted in descending order: first by “t”, then for a given “t” by “s”, and then for given “t,s” by “f”. The process would result with a list of safety issues which would be sorted according to their significance. The significance sub-rank at second and third layer can be very important in the case of larger total number of safety issues: in those cases the numbers of issues with same primary rank can also be considerable and second and third rank would then provide for their sorting in decreasing order.

Very important part of the process is defining the ranges of numerical scores for the three paths, “DS”, “RS” and “SF” in a calibrated manner, because the final order of the list may considerably depend on this. This may involve an iteration or two in the process. Consistency checks may also help (e.g.: issue A is listed as more significant than issue B, and issue B as more significant than issue C; is issue A demonstrably much more significant than issue C?). This and some other parts would inevitably involve some expert judgement and discussions among the reviewers.

It is very important to recognize that the ranking process as discussed in this paper is relative, i.e., the final result is a list of safety issues which are sorted by predefined significance. The results can be used to obtain an answer to the question: is issue X more than significant than issue Y or Z? However, no attempt was made in this paper to discuss the absolute importance, such as for example, at which place (item) can the sorted list be “cut off”. The answer to this particular question is not simple and it would ask for some kind of “global assessment” (e.g. [1]) which would consider joint impact (synergy) of issues on lower side of “cut off” and implementation of the corrective measures for the issues on the upper side, as well as their time schedule (i.e., corrective actions implementation plan). This can be based on principles of risk assessment and / or deterministic principles such as those related to adequacy of safety margins and fault tolerance. Additionally, PSR [1] requires global assessment to provide safety justification for proposed long term operation [2] by evaluating the cumulative effects of both ageing and obsolescence on the safety and reflecting the combined effects of all safety factors (findings and proposed improvements).

In practice, a process like this was used, rather successfully, for the initial inventories with several hundreds of diverse safety issues resulting from periodic safety reviews of nuclear power plants.

REFERENCES

