Back to Contents

GIF-012-00

Generation IV Roadmap

Final System Screening Evaluation Methodology R&D Report

Issued by the Nuclear Energy Research Advisory Committee
and the Generation IV International Forum

December 2002

MEMBERS OF THE EVALUATION METHODOLOGY GROUP

William Rasin	Co-chair	Consultant, formerly Duke Engineering and Services, United States
Jordi Roglans	Co-chair	Argonne National Laboratory, United States
Deborah Bennett		Los Alamos National Laboratory, United States
Evelyne Bertel		OECD-Nuclear Energy Agency, Europe
Dennis Bley		Buttonwood Consulting, United States
Douglas Crawford		Argonne National Laboratory-West, United States
Brent Dixon		Idaho National Engineering and Environmental Laboratory, United States
Michael Golay		Massachusetts Institute of Technology, United States
William Halsey		Lawrence Livermore National Laboratory, United States
Kazuaki Matsui		Institute of Applied Energy, Japan
Keith Miller		British Nuclear Fuels Limited, United Kingdom
Per Peterson		University of California - Berkeley, United States
Geoffrey Rothwell		Stanford University, United States
Thomas Shea		International Atomic Energy Agency, Europe
Michel Vidard		Electricite de France, France
Jean-Claude Yazidjian		Framatome, France

Other Contributors

Burrus Carnahan		U.S. Department of State, United States
Gian-Luigi Fiorini	RIT Representative	Commissariat à l'Energie Atomique, France
Hussein Khalil	RIT Representative	Argonne National Laboratory, United States
Salomon Levy	GRNS Representative	Levy & Associates, United States
Stephen Mladineo		Pacific Northwest National Laboratory, United States
Jon Phillips		U.S. Department of Energy NA-241, United States
Rob Versluis	DOE Representative	U.S. Department of Energy Nuclear Energy, United States

CONTENTS

FIGURES

TABLES

Final System Screening Evaluation Methodology R&D Report

1. BACKGROUND

The purpose of the Generation IV Roadmap is to identify nuclear energy systems that offer the greatest potential for meeting the goals of the Generation IV initiative. A *system* is defined as a complete nuclear system, consisting of the power-producing plant and its associated fuel cycle. The Roadmap will set forth a long-term research, development and demonstration plan that will form a basis for international collaboration on those systems. The Roadmap process is expected to stimulate innovative and critical thinking on new nuclear energy systems that could, in the long term, offer substantial advances and breakthroughs.

Nuclear energy systems proposed for Generation IV (Gen IV) are evaluated at different stages for their potential to meet the Gen IV goals. A screening for potential began in July 2001. The screening was carried out with relatively limited information about the proposed systems. The purpose of the screening for potential was to identify for further consideration those nuclear energy systems that meet the purpose and principles of the Generation IV initiative and have the potential for significant progress toward the established goals. The basic philosophy for the screening for potential was to avoid discarding systems with potential because of limited information available.

After the screening for potential, the technical working groups (TWGs) acquired additional information about the remaining systems and further assessed their performance characteristics. Moreover, the TWGs defined sets comprising systems with engineering and performance similarities. [Note: the method described applies both to a system or a system set, even though only the term *system* will be used.] To complete the Roadmap, the most promising systems or sets need to be selected. For those systems, R&D needs to be identified and prioritized. To do this, the TWGs will begin the process by performing a second evaluation of the systems, which is called *final screening*. The selection of systems will be accomplished with input and guidance from the Generation IV International Forum (GIF) and the Generation IV Roadmap NERAC Subcommittee (GRNS).

This document presents the methodology developed by the Evaluation Methodology Group (EMG) to perform the final screening. The following sections describe the process, the method, and the criteria to be applied in evaluating the systems.

Note: Although this document uses the term *system*, it is acknowledged that the use of the term concept to refer to a nuclear energy system has been in wide use in the Roadmap.

2. PURPOSE

The purpose of the final screening system evaluation is twofold:

1. Identify the leading Generation IV systems. The screening for potential already identified the systems that met the general Gen IV principles and showed potential for meeting the Generation IV Goals. The final screening needs to identify the most promising systems for Generation IV in order to continue their development. The notion of a promising system has two aspects, namely its performance potential, and the development challenge for the system to meet that potential. The systems need to be differentiated by a combination of these two characteristics.

2. Develop information to support the R&D planning. The TWGs and crosscut groups (CGs) will have identified the R&D needs for the different systems. The evaluation methodology provides input to prioritize the R&D and to support the development of an R&D plan.

To meet the first purpose, the method in the final screening needs to go beyond the method applied in the screening for potential, as it will now be necessary to perform a comparison between the systems.

The basic philosophy of the final screening is that, unlike in the screening for potential, where a justification was needed to eliminate a system from further consideration, the TWGs need to provide a good justification for selecting a system in comparison with the others.

3. PROCESS

The United States Department of Energy (DOE) and the Roadmap Integration Team (RIT), in combination with the EMG, have defined the process that will be followed during the final screening:

The TWGs will, if necessary, collect additional information about the system proposals. The information will need to cover aspects of the system performance used in determining if the system has good potential to meet the Gen IV Goals. The criteria used in the screening for potential provides the basic indication to the TWGs for the information that will be needed to evaluate the systems in the final screening. The information available will depend on the state of development of the system. Promising systems are not to be penalized for earlier states of development. TWG judgment is to be used when complete information is not available.

The TWGs and CGs will also need to assess the R&D needs associated with the different systems. Some basic parameters for the required R&D will need to be determined by the TWGs, like magnitude of the R&D program (cost, length of time, complexity) and uncertainty about the outcome of the R&D. Crosscutting technology R&D will be determined by the CGs in a similar manner.

The TWGs may have grouped the individual proposed systems into sets. The sets consist of proposed nuclear energy systems that share the same basic engineering solutions and have comparable performance. For simplicity, the TWGs may choose to compile the performance and R&D information on a set basis rather than on an individual system basis.

> Note: The TWGs will determine their own criteria in establishing the sets. However, if an individual system within a set has a significantly different potential with respect to one or more criteria than the rest of the set, the TWGs may want to reconsider the definition of the set and rearrange it so that all systems within a set have a similar potential and evaluate similarly.

The EMG will provide the methodology and the set of criteria and metrics to be used in evaluating the systems and input to support the selection of the preferred systems (this document). The evaluation will be summarized in an electronic evaluation form (MS Excel spreadsheet) provided by the EMG. The evaluation method will apply to individual systems or to sets.

The TWGs will systematically apply the methodology to all the systems (or sets of systems, at the discretion of the TWGs) within their group and will fill the evaluation forms indicating the potential of the systems and the perceived uncertainty about that potential.

The EMG will monitor the evaluation to ensure a uniform application of the methodology and a consistent use of the criteria/metrics across the working groups.

The TWGs will thoroughly document their system evaluation.

After the TWGs have completed the evaluation of the systems in their technology group, the selection process will be implemented. The RIT, with input from the TWGs and methodology guidance from the EMG, will advance the selection of the systems on the bases of their potential and development challenge. The GIF and GRNS will advise and provide input to the process and endorse the results.

After the selection process has been completed, an R&D plan for the development of the selected systems will be generated. The RIT, with input from GRNS, GIF, the TWGs, and CGs will develop the R&D Plan. The EMG will only provide recommendations for the process.

The results of the evaluation, selection, and R&D prioritization will be documented in the Gen IV Roadmap report.

In this process, the role of the TWGs is to perform the system or set evaluation and to assess the R&D requirements for the systems. The RIT will have the main responsibility for advancing the system selection after the evaluation. The TWGs will naturally provide input to the selection process. The EMG does not take part in the evaluations or selection, except for providing the methodology and for ensuring that the method is used as intended. Because evaluations will be conducted in four different technology groups, it will be important that EMG members are available to provide guidance in the interpretation of the criteria and metrics. This will help to provide consistency across the technology groups. EMG members, however, will not provide expert opinions in the actual evaluation of system attributes or characteristics. The process and responsibilities of the different groups is illustrated in Figure 1.

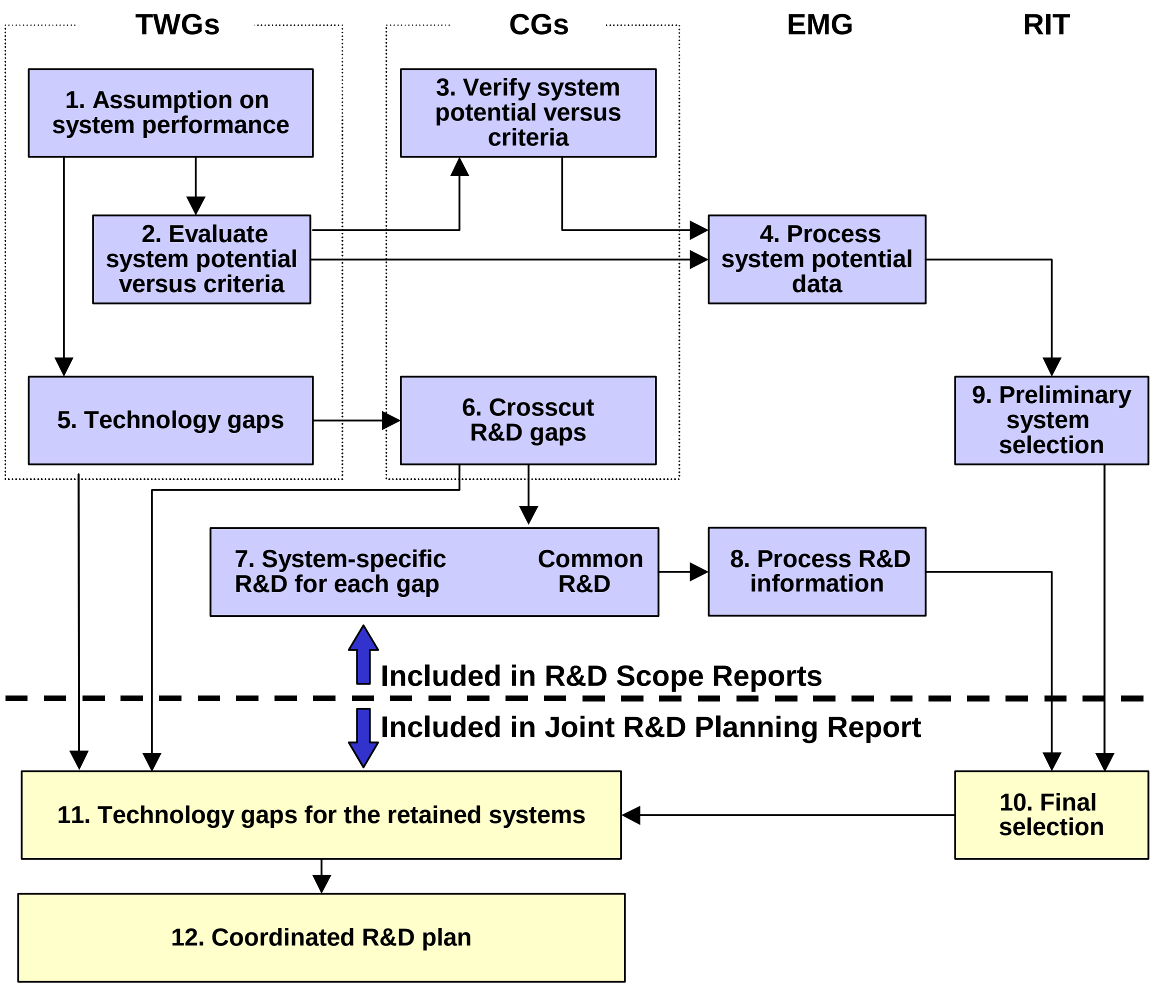


Figure 1. Final screening and R&D prioritization steps and responsibilities.

4. FINAL SCREENING: EVALUATION METHOD

This section describes the methodology for the final screening. As part of the methodology, a set of system evaluation criteria has been developed. Appendix A lists and describes the criteria and metrics, grouped by Gen IV Goal. Appendix B summarizes the reference values corresponding to the Generation III [assumed to be the Advanced Light Water Reactor (ALWR) with a once-through cycle] performance with respect to the Gen IV goals. The ALWR reference values correspond to the mid-scale of the metrics of the criteria presented in Appendix A.

4.1 Overall Approach

The purpose of the methodology is to establish a process to select the systems with the highest potential to meet the Gen IV goals with an acceptable development difficulty or challenge. To that effect, the systems need to be first evaluated for their potential to meet the Generation IV goals, and their R&D needs must be identified and assessed in order to determine the degree of difficulty in developing the systems to their estimated potential.

The basic principle for selecting the preferred Generation IV systems is based on the premise that the most desirable systems are those that offer a high potential to meet the Gen IV goals and have a reasonable risk associated with achieving that potential. In other words, the selection process is a trade off between the system potential and a measure of uncertainty about that potential.

The process of selecting a subset of nuclear systems for further research and development is not a simple mechanical task of choosing the systems that scored higher than a certain threshold in the evaluation of potential. On the contrary, the process needs to address additional considerations such as issues related to the likelihood that a system can realize its potential given all possible technology gaps associated with it, and possibly the desire for a portfolio of technologies to meet diverse scenarios or specific desirable missions. The introduction of these considerations needs to be done gradually, as more detailed information about the systems and their R&D needs is being developed by the TWGs. The process proposed is a multistep approach that starts with emphasis on the potential of the systems to meet the Gen IV goals, gradually shifts the stress to the system R&D issues, including R&D overlaps among systems, and leads to the selection of a small set of systems and an R&D plan for their development. The detailed process proposed for the final screening is shown in Figure 2.

The terminology used in this section is as follows:

System Performance. A measure of the ability of a system to meet the Gen IV goals in terms of the preestablished criteria and metrics. The true value of the metrics is highly uncertain, so the system performance is presented in terms of probability density functions.

System Potential. Figure of merit characterizing the system performance, a measure of its ability to meet the Generation IV goals. The performance potential is indicated as the value corresponding to a high (optimistic) confidence level, typically 75%.

Development Costs. An overall estimate of the costs of developing the system from its current status to commercialization.

R&D Costs. An overall estimate of the costs of research and development that sponsoring agencies would incur in developing the system from its current status to the point of preparation for the system design certification. Design certification expenditures and demonstration costs are excluded.

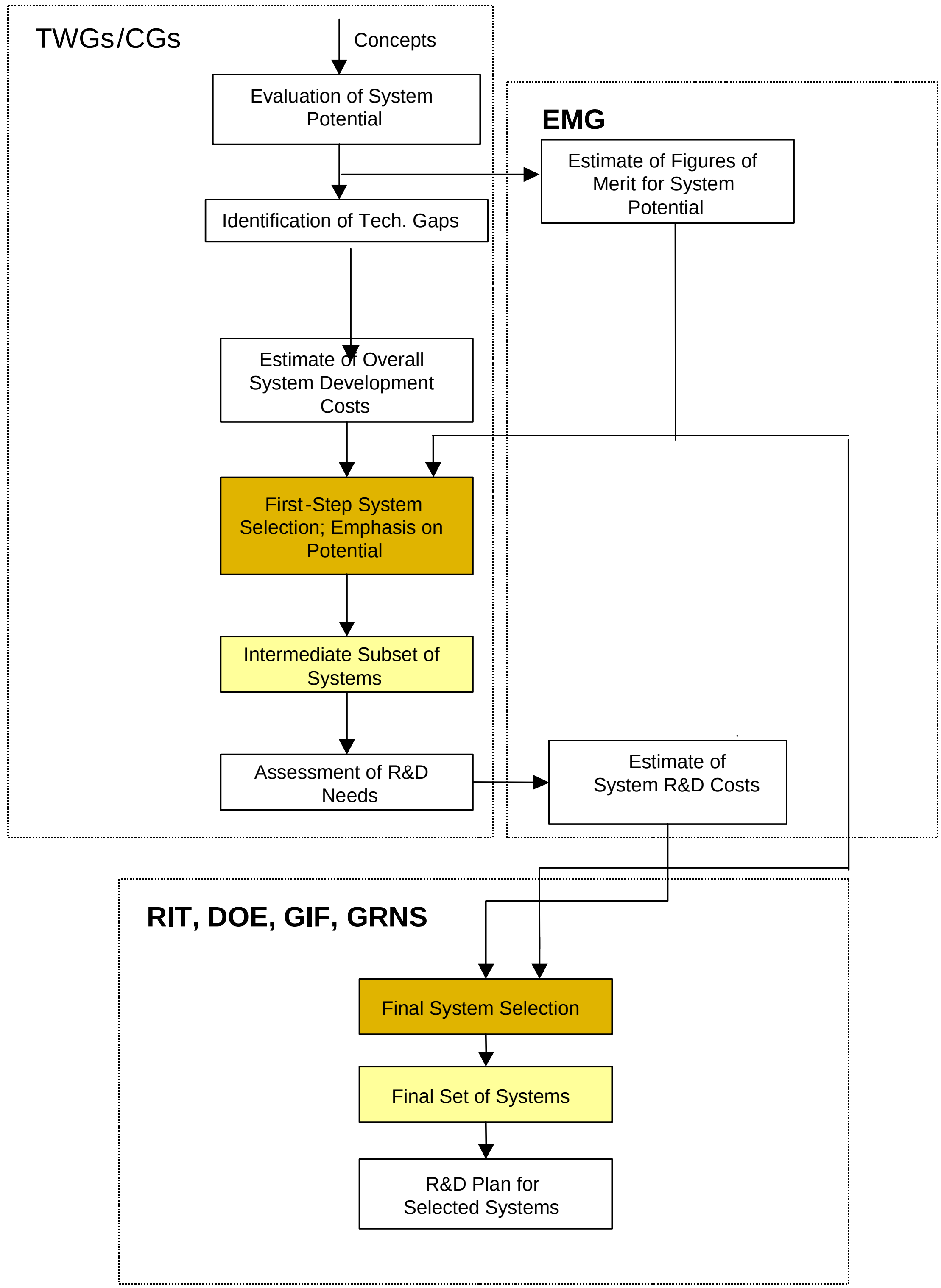


Figure 2. Detailed process for final screening.

The guidance developed in this section provides direction on:

1. Development and use of the figures of merit representing system potential
2. Initial system selection
3. Definition of the R&D costs associated with a system
4. EMG recommendation on final system selection and consideration for an R&D plan.

In the initial selection step, system performance is evaluated in terms of potential (i.e., at the 50% confidence level of performance), and in the final selection stage it is evaluated in terms of a range of upper and lower performance variable values. Similarly, R&D costs are evaluated in response to the values of the performance variable, taking the most and least favorable values.

4.2 System Evaluation with Respect to Generation IV Goals

A probability-based approach will be used for expressing the potential of the systems and the uncertainty associated with that potential. Probability distributions will be used to represent the potential and uncertainty of a system with respect to each criterion. The distributions can then be propagated to obtain

1. A figure of merit representing the performance potential of the system under each goal, to provide a quick global view of the system performance with respect to the Gen IV goals
2. An overall integrated figure of merit for each system, under each goal area, that supports a relative ranking of systems and subsequent comparison.

The mechanics of performing the evaluation are discussed below. Details of the theoretical basis supporting the approach and the mathematical treatment are presented in Appendix C.

The TWGs need to evaluate the systems with respect to a set of criteria representative of the Gen IV goals. The list of criteria, with their definition and interpretation, is presented in Appendix A. A metric is provided for each criterion. Whenever possible, the metric has been defined in a numerical manner, but in some instances a qualitative definition of the metric is given.

The evaluation for each criterion will be performed by comparing the performance of the system with respect to a reference. The reference for comparison is primarily based on an ALWR system with a once-through cycle, although economic data has been updated, as specified in Appendix B. The metric for each criterion provides several reference points (numerical for most criteria, descriptive for the rest). The TWGs are asked to evaluate the performance of the system with respect to the reference and place the score in the corresponding range. Along with a *neutral* value (performance comparable to the reference ALWR system), each criterion range also allows for scores representing values better/worse, and much better/much worse than the reference.

The scale provided with the metric allows for performance ratings above the much better value or below the much worse, if needed. When feasible, and in particular for numerical criteria, the full range is defined in the metric, and it provides a wider set of options of the performance of the system with respect to the reference.

Note that the scales for the criteria metrics are discretized, so that in evaluating the systems the TWGs do not need to choose specific values in the scale, but only discrete intervals.

	Much worse	Worse	Reference Value	Better	Much better	

For each criterion, the TWGs need to evaluate the potential of the system as well as the uncertainty about that potential. The score will be assigned in a single step in the form of a probability distribution that represents the range of potential of the system with respect to the criterion considering all sources of uncertainty.

The upper end of the probability distribution assigned to a system represents the performance potential of the system that would be realized if all the required R&D to validate the design performance assumptions proves to be successful, while the lower end represents performance that would be obtained should the assumptions be proven to be invalid by the R&D.

It is recognized that the TWGs have very limited information to estimate a probability distribution for the performance of a system with respect to each criterion. On the other hand, the EMG desires to avoid excessive constraints. The final screening method aims at minimizing the burden on the TWGs while providing reasonable flexibility.

The TWGs are asked to estimate the probability distribution for the performance of a concept with respect to each criterion, as follows:

1. Assess the level of information/understanding of the system relative to the criterion and select the appropriate distribution (i.e., probability density function) from the following list:

    - *Triangular*. Used when a "most likely" performance value can be determined relative to the criterion. When this distribution is selected, the TWGs will be asked to specify the upper and lower bounds and the most likely value.

      This should be regarded as the default distribution type. The TWGs should make an effort to define the system performance in terms of this distribution. Only when this is not feasible because the TWGs cannot identify the most likely value (or can clearly differentiate between two well-defined possibilities depending on the outcome of the R&D) will they use one of the other distributions suggested.

    - *Uniform* (flat). Used when the available information is insufficient to specify a "most likely" performance value for the criterion. When this distribution is selected, the TWGs will be asked to specify only the upper and lower bounds on the potential performance.

    - *Bi-modal*. Similar to the triangular distribution, but used when the outcome of a critical assumption could significantly alter the performance relative to the criterion. When this distribution is selected, the TWGs will be asked to specify the most likely value if the key assumption proves true, the most likely value if the assumption proves false, and a weighting of the relative probabilities of the favorable and unfavorable outcomes.

2. Assess the potential performance of the system with respect to the reference, upper and lower values provided for the criterion. Based on the distribution selected, indicate a range and/or most likely value(s) for the system performance by choosing locations in the scale provided by the metric associated with the criterion:

   Select a discrete interval (out of 7) in the scale corresponding to each point required by the selected shape (e.g., for the triangular distribution the high, low and most likely performance values).

   The TWGs will be able to enter this information directly into an electronic evaluation form. As the above values are provided, the resulting distribution will be displayed as a bar chart.

   Information for use by the TWGs regarding bias in this type of process is provided in Appendix D.

3. Using the text field provided, the TWGs will document the key assumptions in assessing the performance potential and uncertainty range of the system. These assumptions can be used later to ensure consistency with the estimated R&D Costs. The “justification” text field is to be used to explain the rationale behind the scoring for each criterion, while the “comment” text field may be used to record additional information. In particular, when a bi-modal distribution is selected the TWG should document the associated critical assumption in the justification field.

The outcome of the three steps above is a probability distribution and explanatory text for the potential of the system with respect to the specific criterion. The upper range of the distribution represents the maximum potential; the width of the distribution represents variability and R&D-related uncertainty.

The TWGs need to perform these steps for each of the criteria to complete the evaluation form for a system. The results for all systems will be combined and provided to the RIT for the selection process. Appendix C presents details on the methods used to generate the distributions and combine them to produce integrated figures of merit for each system. This will complete the evaluation.

Figures of merit for the system potential are obtained for the four goal areas: Sustainability, Proliferation Resistance and Physical Security, Safety and Reliability, and Economics. A figure of merit representing the potential of the systems to meet each goal is estimated automatically after the TWGs provide the information for the system to meet each of the criteria. The system score under each goal is rolled up using the criteria weights provided by the EMG. The weights for the criteria under each goal (presented in Appendix E) are uniform unless a specific criterion is believed to be more relevant to the goal than the rest. When criteria weights are nonuniform, the justification is presented in Appendix E.

A default estimate for the figure of merit for each of the goal areas is calculated using equal weights for the goals under each area. Relative weights among the goals can express specific policies, missions, or deployment scenarios. Therefore, the EMG has only provided a default equal-weight value and has provided the means for performing sensitivity studies. At their discretion, the RIT can perform these sensitivity analyses by changing the relative weights of the goals. The 50th percentile is recommended for use as the figure of merit for the potential.

Four measures of potential (sustainability, proliferation resistance and physical protection, safety and reliability, and economics) are obtained for each system using this process.

4.3 System Selection

The system selection in the final screening is performed in a two-step process.

4.3.1 Initial System Selection

The selection of the most desirable systems is based on a trade off between the system potential and the expense and difficulties involved in developing the system to the expected high performance (i.e., the potential and the associated R&D costs). A significant amount of information is needed about the R&D required to develop the system before a final set of systems can be selected on the basis of their potential and R&D challenges. To prevent overloading, the TWGs will define all aspects of the R&D requirements for each system under consideration. After the screening for potential, an initial selection of systems will be conducted, primarily on the basis of their potential. The detailed R&D information will then be developed for the selected systems, to be used in the final step of this Final Screening.

The purpose of this initial selection is therefore to identify a reduced number of systems, among those that survived the screening for potential, for which to develop a detailed R&D assessment that will be a key factor in the selection of the final set of systems and an R&D plan for their development.

The initial selection is performed primarily on the basis of the system potential (i.e., system performance evaluated at a high confidence level), as represented by the figures of merit described above. It is recommended to display the systems' potential against an estimate of the total cost of development, to provide information during the selection about the maturity of the systems selected. The specific reduction in the number of systems will be a function of the target burden on the TWGs (development of detailed R&D information for selected systems), and the distribution of systems performance, i.e., identification of groupings or clusters in potential.

The method for this selection consists in displaying the system potential in the four goal areas (sustainability, proliferation resistance and physical protection, safety and reliability, and economics) against the estimate of the total development costs. The selection will be based on those systems that offer the highest potential. The development costs will indicate the maturity of these systems and will not play any further role in the selection if there is a distribution of maturity in the systems selected on potential. However, if all systems selected on potential alone are associated with development costs in the highest range, the RIT may want to consider also choosing some systems with lower development costs.

This initial selection process, with emphasis on the system potential, avoids early biases against highly innovative, less mature systems before their associated R&D has been properly assessed. In addition, it reduces the burden of the TWGs and allows timely progress toward identification of the best Gen IV candidates.

4.3.1.1 Development Costs

The estimate of the overall development costs will be obtained by the TWGs. It is intended that these development costs be an estimate on a coarse range for development costs, since it is also to be used before a detailed R&D assessment is performed.

Although development costs are usually the smallest expenditure during the life cycle of an energy system, because they are spent early, their present value is a higher portion of total cost when discounted to the present. Also, given that research and development expenditures are made on competing technologies, only a few of which will be commercialized, the allocation of expenditures on technologies that are not marketed is problematic. Further, even for technologies that are commercialized, is it often difficult to allocate development expenditures over the entire fleet of commercial units. This leads to high *first-of-a-kind* costs that discourage early adoption of the technology. Therefore, because of these barriers to market entry, we treat development costs separately from other economic criteria. The focus here is on the reduction of technological uncertainties at each stage of screening and evaluation. The suggestion is to identify information on the following characteristics and determine whether the new nuclear technology will be more or less costly to develop than other similar Generation III technologies. The EMG suggests following the definition of research, development, demonstration, and first-of-a-kind costs in Oak Ridge National Laboratory, *Cost Estimate Guidelines for Advanced Nuclear Power Technologies* (May 1993, ORNL/TM-10071/R3). An electronic version of this report can be found at the following Web site: http://www.ornl.gov/~webworks/cppr/y2001/rpt/64453.pdf

The Development costs will normally include (sunk costs are excluded; first-of-a-kind (FOAK) equipment costs are not included):

- R&D, including cost and demonstration facilities
- Demonstration facilities if needed
- Generic verification tests for new design components
- Engineering sufficient to permit prelicensing and commercial application
- Prelicensing up to and including the Design Certification
- Detailed engineering costs up to FOAK plant
- Additional fuel development costs.

Development costs for the AP600 and ABWR can be used as guidelines. It is assumed that with industry-sponsored research and development and DOE matching funds that developing each of these technologies cost about $500M. This does not include the research and development costs of similar Generation I & II LWRs. These costs could be recovered as FOAK costs on the first set of power plants constructed or amortized over all units. For example, assuming 4,000 MW in the first set, this would add $125/kW (if $500M). The following scale for development costs is provided. The TWGs need to estimate the development cost range for the system being evaluated.

>$2000M	$2000–1000 M	$1000–550 M	$450–550 M	$450–350 M	$350–250 M	<$250 M

If dollar values are unknowable, use qualitative evaluation, e.g., research, development, and demonstration costs are much lower than the guideline. To aid in the estimation of RD&D costs at each stage of screening and evaluation, consider the following plant characteristics:

- Identify time, cost, and (demonstration) facilities needed to resolve major technological uncertainties during viability R&D phase. Identify major technical uncertainties.
- To facilitate the prioritization of R&D, estimate time, cost, and (demonstration) facilities needed to resolve major technological uncertainties. Those technologies with uncertainties that can be resolved more quickly (given an R&D budget constraint) and have higher probabilities of success should be cheaper to develop.

4.3.1.2 Display of System Potential against Development Costs

The summary information about the different systems can be displayed in plots of system potential versus development costs. Four such plots (see example below) can be generated: one each for Sustainability, Proliferation Resistance and Physical Protection, Safety and Reliability, and Economics.

The plots, if used, would display the following information:

- Ordinate. Figure of merit of the system potential for the four goal categories (a different figure of merit for each of the goal categories)
- Abscissa. Overall estimate of the development costs—a single value for all four plots.

4.3.1.3 Selection of Systems with Higher Potential

The selection will be based on those systems that offer the highest potential. It is recognized that systems may not score consistently across the four goal areas. The RIT and TWGs may want to estimate, in addition, a single figure of merit for each system by combining the four figures of merit under the goal areas. This will require using a set of weights. In this case, it is recommended that several weight combinations be considered, corresponding to different scenarios, in order to understand the sensitivity of the combined score to the weighting scheme. If all systems selected based upon potential alone are associated with development costs in the highest end of the range, the RIT/TWGs may want to consider adding some systems with lower development costs, pending the more detailed R&D assessment of the next step.

4.3.2 Final Selection

Once the initial set of systems has been selected, on the basis of their potential to meet the Gen IV goals, the TWGs will assess the R&D Costs in order to provide additional information to support the final selection.

The outcome of this final selection step will be a small group of systems for research and development that have high potential for meeting the Gen IV goals. Because the selected systems will be proposed for further R&D, the expected R&D costs to the sponsoring agencies are an important consideration in the selection.

The final selection process will use, among other considerations, information about the system potential to meet the Generation IV goals (figures of merit for system potential) and a measure of the degree of difficulty in attempting to demonstrate that the system can satisfy that potential (represented by an estimate of the R&D costs.)

R&D Programs are very dynamic and evolve as the different stages of R&D are performed. The costs associated with R&D will vary as R&D evolves. Recognizing that the true value of the R&D costs is unknown, R&D costs will be estimated with expert judgment with the current state of knowledge.

In the initial selection step the total expenses required for the development of the system up to the point of commercialization (including demonstration plants) were considered. In this final selection step, however, since the decision results in an R&D plan for a set of systems, the R&D costs for which sponsoring agencies would be responsible are more relevant.

4.3.2.1 Technology Readiness Level

The first two stages of R&D as a function of the characteristics and the nature of the R&D as the system design development matures through the successive stages of innovation are defined as follows:

1. Viability R&D. That R&D necessary for proof of the basic concepts, technologies, and processes at relevant conditions. Potential show-stoppers are identified and resolved in this phase.

   The information generated at this stage is sufficient for the conceptual design stage of a prototype.

2. Performance R&D. That R&D necessary for engineering-scale verification of processes, phenomena, and material capabilities in representative conditions.

   The information generated during this phase is sufficient to allow a detailed design and performance specifications for a prototype or demonstration plant and allow beginning the development of a certification application.

The demonstration phase will follow the Performance R&D. The demonstration phase may require the construction of a prototype or a demonstration plant and involved partnerships between R&D and commercial (vendors, utilities) organizations. The system design will be optimized in that phase at that time.

The degree of development of the system can be expressed in terms of technology readiness level. The use of a TRL provides a single parameter characterizing the state of development of a particular technology. The TRL levels defined in this context are shown in Table 1:

Table 1. The TRL levels defined in the Generation IV Roadmap context.

Level	Title	Description
TRL1	Basic research in new technologies	Scientific research begins to be translated into applied R&D; no experimental proof exists yet
TRL2	Proof of phenomena	Analytical and experimental demonstration of critical function and/or characteristic proof of concept
TRL3	Technology development	Small-scale (laboratory) demonstration in a relevant environment
TRL4	Proof of practicality	Subsystem or separate effects test completed in representative conditions; concept is proved to be practical in representative conditions
TRL5	Proof of concept	Large-scale (integral facilities) tests in representative conditions

The Viability R&D is expected to take the system to a TRL3 level. By the end of the Performance R&D phase, a system should have been developed to a TRL5. Specifically, the endpoints of the two R&D Phases are summarized in Table 2.

Table 2. Endpoints for the viability and performance R&D phases.

Viability Phase	Basic concepts, technologies and processes proven at relevant conditions; potential technical "show-stoppers" identified and resolved.
	System has advanced to the point that the following information has been developed: • Conceptual design of nuclear island • Simplified PRA • Definition of testing and analytical tools needed • Nominal interface requirements for power and support systems • Basic fuel cycle process flow sheets established by testing at reasonable scale • Preconceptual design of process facilities, with established pathways for disposal of all process waste streams • Simplified environmental impact statement for system • Preliminary Business Plan based on conceptual design • Preliminary safeguards and security strategy discussing intrinsic proliferation resistant features and identifying necessary extrinsic controls and physical protection needs
Performance Phase	Engineering-scale verification of process, phenomena, and material capabilities in prototypical conditions.
	System has advanced to the point that the following information has been developed: • PRA • Demonstration of safety features through testing, analysis, experience • Validation of analytical tools • Conceptual design sufficient to show interface requirements for power and support systems • Fuel cycle process flow sheets validated at scale sufficient for commercial demonstration • Conceptual design of process facilities, with validated acceptability for disposal of all process waste streams • Environmental impact statement for system • Detailed business plan for system • Safeguards and security strategy for system, including cost estimate for extrinsic features and physical protection • Performance requirements and design information for nuclear island, sufficient for procurement specifications for construction of a prototype or demonstration plant

The level of development of interest in the system selection process includes the R&D involved in both the viability and the performance R&D phases. Therefore, the R&D costs of interests are those required to take a system to a state of development represented by a TRL5.

The R&D costs will include:

- Fuel development costs
- Fuel cycle development costs
- R&D for materials development
- Applicable R&D for reactor development
  - Physics
  - Thermohydraulics
  - Instrumentation and control
  - Transient analysis
  - Containment systems
- Applicable Nonreactor facilities development
  - Power plant development
  - Production facilities for alternative energy products
  - Fuel cycle facilities
- Engineering sufficient to permit design of a prototype or demonstration plant
- Engineering sufficient to allow beginning the development of certification application
- Analysis in support of the system engineering (as listed in the endpoints above) during the two R&D phases.

4.3.2.2 System R&D

R&D costs will be proportional to the number of major issues that need to be resolved. The detailed assessment of technology gaps and R&D needs (including detailed costs and schedules for each R&D item) will only be performed for the selected concepts in order to develop the R&D plan. However, it will be useful for assessing the R&D costs in this section to review some of the main areas of R&D needed for a system. This will verify that the R&D estimated costs are consistent with the identified major development requirements for a system.

Of particular importance in this estimate is the consideration of the entire system, even for those concepts that have primarily provided information about the power plant. Some of the major development needs may occur in the area of the fuel cycle, both in the fuel development, and in the back end of the cycle for both recycle systems that require the process development and for once-through systems that may need to develop disposal forms.

The following is a sample verification list (not meant to be exhaustive) of R&D items:

- Fuel development
  - Does a new fuel form need to be developed for the system?
  - Are new cladding materials to be developed?

- Discharged fuel
  - Does a process need to be developed for the recycling of spent fuel?
  - Has the recycling process been demonstrated? At what scale and TRL level?
  - Do disposal forms for the spent fuel/waste streams need to be developed?
- For symbiotic systems
  - Have the all the fuel/waste material forms that allow the interface of the different elements of the symbiotic system been developed?
- Reactor Materials
  - Do advanced materials (e.g., high-temperature) need to be developed?
- Reactor systems
- Do safety systems and transient response for the system require new developments in design or in analysis tools?
  - Do primary systems require new development in design or in analysis tools?
  - Do performance and safety responses require separate effects testing?
  - Do performance and safety responses require integral effects testing?
- Balance of system
  - Does concept include advanced systems that require new development (e.g., hydrogen generation systems)?
  - Have advanced systems been demonstrated? At what scale and TRL level?
- Overall
  - Does the system (power plant, fuel cycle, facilities related to alternative energy products, if any) include any specific element or characteristic that requires new development?

4.3.2.3 R&D Costs

On the basis of the estimated development needed for the system to reach a status of proof of performance, the TWG will estimate the R&D costs. It is expected that information produced in the estimate of the overall development costs for the first step of the selection can be used here. Indeed, the R&D costs are a subset of the total development costs already estimated.

The TWGs will need to estimate the R&D costs on the basis of the major development needs for each system. As acknowledged above, it is difficult to estimate R&D costs for a large development program and R&D costs change as the program evolves. It is understood that the TWGs will provide an estimate relative to the current system information and based on the knowledge of the experts in the working groups.

In estimating the magnitude of the R&D expenses given the identified areas of major development, the TWGs need to rely on the expertise of their members with previous R&D programs. It is possible that a single TWG will not have the necessary expertise to estimate some of the R&D cost components. Crosscutting activities will therefore become a more important factor in estimating R&D costs. R&D costs for the AP600 and ABWR or other programs can be used as guidelines. It is estimated that with industry-sponsored research and development and DOE matching funds, developing each of these technologies cost about $500M. This does not include the research and development costs of similar Generation I & II LWRs.

As with the parameters estimated for the system potential, the TWGs will provide a range of R&D costs, rather than a single value. Since the R&D costs will not be combined with other parameters, there is no need to use utility functions, and the scale for the costs in $ will not be converted to a value function.

The scale for use in the R&D costs estimate is as follows:

>$750 M	$650–750 M	$550–650 M	$450–550 M	$350–450M	$250–350 M	<$250 M

After system selection is complete, the TWGs will evaluate the R&D needs in detail. The R&D costs estimated for a system after all the R&D information has been developed should be verified with the R&D costs estimated here, for consistency.

4.3.2.4 Considerations for Selecting the Preferred Concepts

Selection of the preferred systems should be based on the trade off between the potential of the systems to meet the four goal categories and the R&D Costs to develop the systems to meet that potential. A potential visual aid in the decision-making could be the four plots mentioned in the previous subsection, but the DOE/RIT and GIF will select the preferred systems as a balance between potential and the R&D costs, pus additional considerations, such as international priorities and missions.

The most desirable systems are those that offer very high potential with low R&D costs. It is likely, however, that those concepts offering the highest potential are also those that present the biggest costs in R&D. Moreover, the best performers (high potential, small R&D costs) in one goal category will not necessarily coincide with the best performers in the other two goal categories.

Choosing only the systems with smallest costs, even if their potential is not very good, is not the goal of the Generation IV Roadmap, which aims at developing truly innovative concepts that represent a step improvement, rather than a modest evolution, with respect to the current generation.

A relevant solution for selecting the systems for further development may be to choose a mixture that includes systems with very high potential, even if their uncertainty is high, and systems with low uncertainty, even if the potential is lower.

DOE/RIT and GIF will likely take into account the following considerations in selecting the final set of systems:

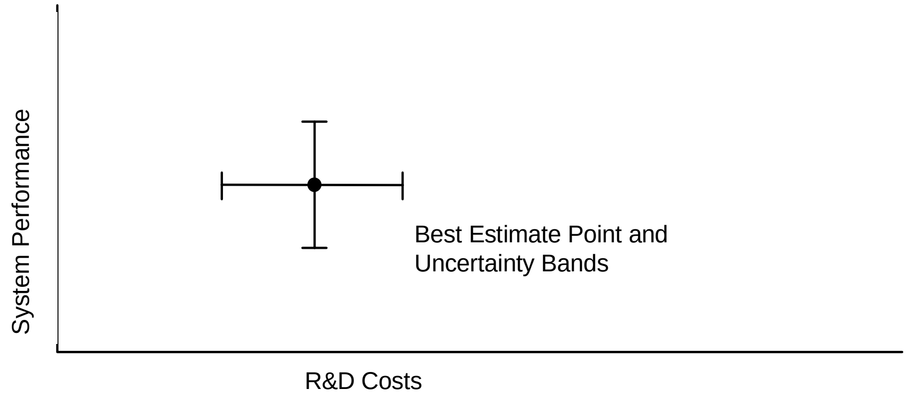
- The number of systems desired for further development
- How the concepts have scored in potential
- Level of R&D costs that the sponsoring agencies are willing to undertake
- Degree of innovation that GIF members agree on sponsoring
- Common interests in systems and technologies pursued by the different GIF participants.

4.3.2.5 Graphic Display of the Figures of Merit for Potential and R&D Costs

The summary information about the different concepts can be displayed in plots of concept performance versus R&D costs. Four such plots can be generated: for Sustainability, Proliferation resistance and Physical Protection, Safety and Reliability, and Economics.

The plots will display the following information:

- Ordinate. Figure of merit of the system performance for the four goal categories (a different figure of merit for each of the goal categories). A range will be displayed (25 to 75 percentile).
- Abscissa. Figure of merit for the R&D challenge (a single figure of merit for all four plots). A range will be displayed (25 to 75 percentile).



4.4 R&D Planning

The TWGs and the CGs will define the R&D needed for the selected nuclear systems to reach the R&D endpoints and technology readiness level, as discussed in Section 4.3.2. R&D should be identified as either viability R&D or performance R&D.

It is important to recognize that selected systems will begin from different starting points with respect to technology readiness level. However, it is important to have all systems at Technical Readiness Level 3 as a minimum by the time the viability evaluations are performed. It is anticipated that such viability evaluations will be performed in FY 2007. The information available for the selected systems will allow for a more rigorous evaluation than was possible for final screening.

An R&D pathway from the current TRL to TRL 3 and from TRL 3 to TRL 5 should be developed. Attention should be given to identifying any *show-stoppers*, those crucial viability issues that would affect the successful development of a system. However, at TRL 3, all information specified in the R&D Endpoints (see Table 1) should be available to facilitate a comprehensive system evaluation.

Highest priority for R&D should be given to crucial viability issues. Formal review points should be scheduled to assess the R&D progress for such crucial viability issues to ensure that continued development is justified. Such reviews should also consider any new issues that may have been discovered in the course of the work to date, so that the R&D plan can be revised, if necessary, in a timely manner. Other R&D associated with a system, particularly issues requiring long lead-time research, should be scheduled in parallel with the viability-related research. However, the pace and funding of such R&D should recognize the crucial nature of the viability questions under study.

For systems satisfying crucial viability issues early, performance R&D should proceed even ahead of the formal viability evaluations. This approach is important to prevent a loss of momentum in the development process. Again, however, the pace and funding should be established with consideration that other systems might prove superior at the end of the viability phase.

In addition to that R&D necessary to reach the potential performance of a system, as established by the final screening assessments, R&D to improve performance in goal areas where a system may have shown some weaknesses should be considered. Such R&D may be scheduled in either the viability or the performance phase.

Finally, R&D plans for the systems should be reviewed to identify R&D that might contribute to more than one system, such as fuel cycle or materials development. This R&D should be given priority to ensure that progress is compatible with individual system R&D schedules.

APPENDIX A

Criteria and Metrics for Final Screening

Appendix A

Criteria and Metrics for Sustainability Goal 1

Goal statement:

Sustainability–1 (SU1). Generation IV nuclear energy systems including fuel cycles will provide sustainable energy generation that meets clean air objectives and promotes long-term availability of systems and effective fuel utilization for worldwide energy production.

Evaluation for goal SU1 focuses on recognizing nuclear energy systems that make sustainable demands upon the existing mineral resource base and ecosystems. The basic principle is that such systems will have longer natural time scales of use and their disruptions of natural systems will be smaller for a given amount of energy production. Proposed metrics measure the satisfaction of these criteria, by comparing to a reference system, that of the LWR, the rate of resource consumption compared to the known resource base.

Summary table of criteria and metrics.

Criterion		Final Screening and R&D Prioritization	*Viability and Performance Evaluations*
SU1-1	Fuel utilization	*F = specific fuel resource consumption (MtU/GWyr)*	
SU1-2	Fuel cycle impact on environment	NA	$M2 = [(A/R)_I/(A/R)_0]$
SU1-3	Utilization of other resources	NA	$M3 = [(m_k/R_k)_I/(m_k/R_k)_0]$

Criteria Definitions

The justification for the forms of metrics suggested below is that each compares that of the system of interest to that of the once-through ALWR.

SU1-1—Fuel utilization

Definition: Generation IV systems will reduce the depletion of nuclear fuel resources.

Discussion: Assessment of the Sustainability Criterion 1 for a nuclear energy system is concerned with its depletion of fuel. The basic principle is that Gen IV systems will have longer natural time scales of use for a given amount of energy production. The attributes or factors to be considered in determining the degree to which a system satisfies this criterion are its specific demands (consumption per unit of energy (either electrical or thermal produced from a reactor) for fuel compared to the economically accessible resource inventory of such fuel.

Proposed Metrics

Use of Fuel Resources

Final Screening Metric: Quantitatively assess the use of fuel resources compared to the ALWR once-through cycle.

Systems that make better use of fuel than the reference once-through ALWR cycle, giving longer natural time scales t, will be rated positively. This will include either achieving higher burnups, increased conversion ratios, or recycling of fuel material.

Metric: Specific fuel resource consumption: F = R/Qt

where

F	=	specific fuel resource consumption (Mt U/GWeYr electric or equivalent product) for reactor/fuel cycle
R	=	total economically recoverable fuel resource inventory (Mt U) used by system
Q	=	the total installed nuclear capacity (GW electric or equivalent)
t	=	natural time scale (yrs).

Uncertainty exists for what the specific values of F, R and Q will be in the time period that Gen IV systems will operate. For final screening only the specific fuel consumption F is considered, because it is a specific attribute of a nuclear energy system. The scales, however, are based on analysis performed by the Fuel Cycle Cross Cut Group:

- The "better than reference" scale covers the range of the specific fuel consumption required to achieve a natural time scale of 100 years (to the end of life of Gen IV systems deployed starting in 2030), for an intermediate value of projected deployment Q (average of 1,250 to 2,500 GWe from 2,030 to 2,100), for an intermediate projected value of economically recoverable uranium resources (8,500,000 to 15,000,000 Mt U).
- The "much better than reference" scale point is set at 10 times the "better than reference" scale point, as representing a reasonable value for a fully sustainable fuel cycle.

This does not result in a score that is directly proportional to the percent utilization of fuel resource. As utilization percentage grows, there is an economic 'diminishing return' on further efficiency (i.e., improving from 1 to 10% resource utilization is more valuable that improving from 90% to 100%).

In addition, for final screening TWGs should provide written discussion of the potential for symbiosis with other systems that could affect the average fuel utilization of the entire Gen IV reactor fleet.

Use of fuel resources: final screening metric scale.

Much worse than reference	Worse than reference	Slightly worse than reference	Similar to reference	Slightly better than reference	Better than reference	Much better than reference
>300 Mt U feed/GWyr	250–300 Mt U feed/GWyr	200–250 Mt U feed/GWyr	150–200 Mt U feed/GWyr	100–150 Mt U feed/GWyr	10–100 Mt U feed/GWyr	<10 Mt U feed/GWyr

SU1-2—Fuel Cycle Impact on Environment

Definition: Generation IV systems will reduce their impact on the environment.

Discussion: Assessment of the Sustainability Criterion 2 for a nuclear energy system is concerned with the amount of environmental disruption associated with the fuel cycle. The basic principle is that Gen IV systems will produce smaller disruption of natural systems for a given amount of energy production. The attributes to be considered in determining the degree to which a fuel cycle satisfies this criterion are its associated specific amount of environmental disruption (e.g., areas of habitat for affected species of biota) compared to the total inventory of such habitat, and the specific disruption of areas of scenic land enjoyed by humans for recreation or aesthetic enjoyment compared to the total inventory of such categories of land.

Proposed Metrics

Fuel Cycle Compatibility With Environment

Final Screening Metric: Qualitatively assess the use of specific habitat or scenic area compared to the ALWR once-through cycle.

Insufficient information is available at final screening to differentiate systems.

SU1-3—Utilization of other resources

Definition: Generation IV systems will reduce the depletion of other specific resources.

Discussion: Assessment of the Sustainability Criterion 3 for a nuclear energy system is concerned with its depletion of identified specific material resources. Specific materials that need to be considered need to be identified among those used in a nuclear energy system that are particularly scarce. The basic principle is that Gen IV systems will have longer natural time scales of use for a given amount of energy production. The attributes or factors to be considered in determining the degree to which a nuclear energy system satisfies this criterion are its specific demands (consumption per unit of energy (either electrical or thermal) produced from a reactor) for unique materials compared to the economically accessible resource inventory of the identified specific materials. Utilization of specific scarce resources applies to the whole energy system.

Proposed Metrics

Use of Other Specific Material Resources

Final Screening Metric: Qualitatively assess the use of other specific resources compared to the ALWR once-through cycle.

Insufficient information is available at final screening to differentiate systems.

Criteria and Metrics for Sustainability Goal 2

Goal statement:

Sustainability–2 (SU2). Generation IV nuclear energy systems will minimize and manage their nuclear waste and notably reduce the long term stewardship burden in the future, thereby improving protection for the public health and the environment.

Evaluation of systems with respect to SU-2 will focus on waste minimization, environmental impacts and stewardship burden for waste disposal. Waste minimization compares unit waste generation, decay heat production and long-lived hazard, as well as opportunities for optimization in waste management. While spent fuel and/or high-level wastes are a clear focus, all wastes should be considered. For environmental impacts, the broad range of emissions considered in a typical EIS is reviewed for unique system features that provide advantages or disadvantages. For stewardship burden, the length and intensity of societal responsibility is considered. At the time of Final Screening and R&D Prioritization, quantitative details are expected to be available for some but certainly not all of these metrics. Qualitative evaluation of the unique features of a system relative to existing LWRs and fuel cycles will be used to evaluate the potential of systems to provide benefit against the criteria. Where more design information is available, a more quantitative approach may be used.

Summary table of criteria and metrics.

Criterion		Final Screening and R&D Prioritization	*Viability and Performance Evaluations*
SU2-1	Waste Minimization	Quantify: • t/GWyr • m^3/GWyr • long-term decay heat kW/GWyr • long-term radiotoxicity/GWyr	*Same, but add* • *waste form performance* • *dose (repository specific)* • *short term decay heat*
SU2-2	Environmental impact of waste management and disposal	Compare to reference: --/-/=/+/++	*Semi-quantitative environmental evaluation (EIS waste issues)*
SU2-3	Stewardship burden	NA	*Evaluate length and level of societal responsibility*

This results in five SU2 metrics for Final Screening and R&D Prioritization. The relative importance of the metrics varies with national perspective and system priorities. A default of equal weighting for each of the five metrics is adopted at this stage of the evaluation.

Criteria Definitions

SU2-1—Waste minimization

Definition: Generation IV systems will offer the opportunity for minimization and improved management of all wastes compared to the ALWR once-through reference system.

Discussion: Considering that management of high-level radioactive waste (HLW) and/or spent nuclear fuel (SNF) is a major issue for current nuclear energy systems, it is expected that Generation IV systems will address these topics. Production of less SNF/HLW per unit electric power than typical current systems is desirable. The reference for comparison varies from nation to nation. In the U.S. for example, comparison would be to a once-through fuel cycle with geologic disposal of intact SNF of intermediate burnup. Because there is no single “quantity” measure to encompass all aspects of waste management that would be uniformly appropriate for the range of potential Gen-IV systems and the range of possible geologic repository settings, several metrics are considered.

The Final Screening and R&D Prioritization is a quantitative assessment of positive and negative features of a system's waste streams, based on metrics representing the amount and properties of HLW/SNF sent to final disposal. It is possible that some systems may not have sufficient basis for quantitative values until further R&D is conducted, and may rely on qualitative or semi-quantitative evaluation.

Proposed Metrics

Mass and Volume of HLW/SNF Sent to Final Disposal

Final Screening Metric: Compare HLW/SNF quantity per GWyr to ALWR O-T; Quantify mass and/or volume of HLW/SNF per GWyr): (MT/GWyr) and/or (m^3/GWyr)

Mass and volume are gross measures of waste quantity and capture some, but not all, of the difficulty in management/disposal of the waste. The absolute and relative importance of mass and volume depend on the waste form and disposal concept details (i.e., repository design). Mass is often used in reference to SNF. Mass and volume of HLW may be dependent on the waste forms selected and the matrix forming materials used to create the waste form and the concentration of radionuclides achieved. Many advanced fuel cycles may vary substantially in mass, volume or both, per unit of generation. For example, higher burnup may result in modest reduction in spent fuel per unit generation, while fuel recycle may result in greater reductions. Advanced fuel and waste forms offer a range of specific density so mass and volume may vary independently.

Mass of waste: metric scale.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>80 MT/GWeYr	40-80 MT/GWeYr	20-40 MT/GWeYr	15-20 MT/GWeYr	10-15 MT/GWeYr	5-10 MT/GWeYr	<5 MT/GWeYr

Volume of waste: metric scale.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>100 m^3/GWeYr	50-100 m^3/GWeYr	20-50 m^3/GWeYr	15-20 m^3/GWeYr	10-15 m^3/GWeYr	5-10 m^3/GWeYr	<5 m^3/GWeYr

Decay Heat Thermal Output

Final Screening Metric: Specific heat output in KW/GWyr in HLW/SNF sent to final disposal compared to reference once-through fuel cycle. For later evaluations, the time variation of decay heat differs significantly from the reference, the system may be scored separately at 50 and 500 years out of core to reflect operational and geologic post-closure times: (kW/GWyr)

In geologic disposal, management of decay heat can be a critical driver for repository design, capacity and performance. The long-term thermal output tends to dominate repository performance, while short-term thermal output is an operational issue, and less discriminating for sustainability of the system. For Final Screening, only the long-term thermal output is evaluated. However, for some fuel cycles, the short-term versus long-term decay heat may be a discriminating feature, so decay heat as a function of time may be the desired metric for later evaluations. For example actinide recycle would result in equal short-term fission product heat but far less longer-term actinide heat. For typical LWR spent fuel the fission products contribute 55% of decay heat at 50 years, but fission products drop to only 10% of total decay heat at 120 years and become negligible by 250 years. Thermal output during the operational period of repository disposal is a design and operations issue while thermal output during post-closure times may be a performance issue. Fuel cycles with partitioning of waste streams (reprocessing) may also offer options for optimization of disposal of wastes with differing thermal properties. If insufficient system detail is available for quantification of this metric, a qualitative evaluation may be used.

Long-time (500 years out-of-core) waste decay heat: scale.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>10 kW/GWeYr	5-10 kW/GWeYr	3-5 kW/GWeYr	1-3 kW/GWeYr	0.5-1 kW/GWeYr	0.1-0.5 kW/GWeYr	<0.1 kW/GWeYr

Radiotoxicity Measures

Final Screening Metric: Long-lived radiotoxicity per GWyr compared to reference once-through fuel cycle.

The radioactivity of waste produced is another gross measure of waste production. But total activity produced per unit of fission energy does not vary greatly between nuclear systems, and activity is often not a major discriminator. However, some fuel cycles vary in the production and consumption of long-lived radionuclides sufficiently to affect the potential health hazard represented by the waste as measured by radiotoxicity. Such fuel cycle characteristics as high actinide consumption or actinide recycle resulting in less long-lived radiotoxicity. Radiotoxicity is a general measure of the potential hazard represented by the material, and can be measured in several ways. One simple representation of long-term toxicity is the sum of the specific activity of each radionuclide remaining 500 at years out-of-core multiplied by a biological dose factor such as the Sv/Bq factors from ICRP72, normalized per GWeYr. Major variations in this measure will be dominated by production and destruction of actinides because of the high dose conversion factors for alpha emitting isotopes.

How activity ultimately relates to repository dose is specific to the combination of waste form performance, repository design and the specific repository site. Such analysis may be applied for later selection steps, but is not relevant at the Final Screening and R&D Prioritization.

Long-lived (500 years out-of-core) radiotoxicity MSv/GWeYr: metric scale.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>3,500 MSv/GWeYr	2,500-3,500 MSv/GWeYr	1,500-2,500 MSv/GWeYr	500-1,500 MSv/GWeYr	100-500 MSv/GWeYr	20-100 MSv/GWeYr	<20 MSv/GWeYr

SU2-2—Environmental impact

Definition: Environmental and health impacts will be assessed relative to current nuclear systems. It is likely that many of these will not be discriminators for most Generation IV systems. However, systems may include unique features or processes that affect environmental issues.

Discussion: The environmental impacts of the complete reactor and associated fuel cycle must be considered. SU1 considers resource utilization, including land, minerals, etc, and SU2-1 considers high-level waste and spent fuel. This criterion considers all other wastes, emissions and operational environmental impacts. Systems should identify unique features and processes with either beneficial or detrimental environmental implications. The typical list of potential impacts considered in an EIS provides a useful guide to the range of issues to consider, while acknowledging that it is likely that many of these impacts will not be clear discriminators for most systems at this stage. The Final Screening and R&D Prioritization will rely on qualitative assessment, as in the Screening for Potential, but with refined input where available. For the later evaluations it may be possible to define and apply quantitative or semi-quantitative measures for certain topics.

Proposed Metrics

Environmental Impacts

Final Screening Metric: Same or similar to Screening for Potential: Qualitative ranking of the major positive and negative features of the environmental impact issues and unique characteristics of a system compared to a reference such as the once-through LWR system: Score ++ / + / = / - / -- for much less / less / same / more / much more environmental impacts expected than in ALWR O–T

This metric will measure the impact on the environment of a specific system as compared to the impact of the reference once-through ALWR system. This includes the generating plant and the fuel cycle (including transportation, etc.) and any other facilities or operations needed to implement the system. The following characteristics will be considered in the comparison with the reference:

- **Construction of facilities.** Construction wastes, emissions and environmental disruption.
- **System operation.** Environmental impacts from both normal and off-normal operations including all waste categories (except HLW), worker and non-worker exposure, emissions, traffic, noise, visual impact, etc.
- **Decommissioning.** Facility decommissioning, decontamination, removal and remediation processes, including exposures, emissions, wastes, etc.
- **Disposal.** Low-level wastes, toxic and mixed wastes, non-toxic waste, etc.

Environmental Impact: metric scale.

Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference
--	-	Equivalent =	+	++

SU2-3—Stewardship burden

Definition: Extent that a system leaves a stewardship burden on future generations will be assessed. This includes facilities, wastes and repository monitoring and/or safeguards.

Discussion: The long-term responsibility incurred by society from implementing a power generation system must be considered. Systems should identify unique features and processes with either beneficial or detrimental stewardship implications. The Final Screening and R&D Prioritization will not likely have additional information that will discriminate among systems. For the later evaluations it may be possible to define and apply quantitative or semi-quantitative measures for certain topics.

Proposed Metrics

Long-term Stewardship Burden

Final Screening Metric:

Insufficient information is available at final screening to differentiate systems.

Criteria and Metrics for Proliferation Resistance and Physical Protection

Goal Statement:

Proliferation Resistance and Physical Protection (PR&PP). Generation IV nuclear energy systems will increase the assurance that they are a very unattractive and least desirable route for diversion or theft of weapons-usable materials and that they provide increased physical protection.

Evaluation of goal PR&PP addresses the relative lifecycle proliferation resistance and physical-protection robustness of proposed Generation IV nuclear energy systems, including both intrinsic and extrinsic features. The sustainability of advanced nuclear fuel cycles will be dependent on cycles that are highly resistant to nation-state proliferation involving the diversion or undeclared production of nuclear materials at all points and all times in the fuel cycle, including the potential for diversion of weapons-usable material following geologic disposition as waste. Likewise, sustainability will also depend on robust physical protection to resist nuclear terrorism, involving theft or sabotage by sub-national groups. The final screening focuses on intrinsic characteristics, which minimize the vulnerability of materials used in reactors. The viability and performance evaluations will consider fully integrated nuclear energy systems including transportation, and will credit integrated systems with strong intrinsic proliferation resistance and physical protection, by measuring and minimizing the lifecycle resources required to provide materials control and accounting (MC&A), physical protection, and international safeguards.

Summary table of criteria and metrics.

Criterion		Final Screening and R&D Prioritization	*Viability and Performance Evaluations*
PR&PP-1	Minimize life-cycle susceptibility to the diversion or undeclared production of weapon-usable material; facilitate implementation of effective IAEA safeguards	Compare with once-through LWR (++/+/=/-/--)	*Life-cycle accessibility of weapon-usable material; safety implications and detectability of undeclared irradiation; life-cycle costs of IAEA inspections, including provision of essential safeguards equipment, per GWyr*
PR&PP-2	Minimize vulnerability to theft of weapon-usable material or hazardous radioactive material; minimize vulnerability of installations and transport systems to acts of terror or sabotage	Compare with once-through LWR (++/+/=/-/--)	*Life-cycle accessibility of weapon-usable material; life-cycle accessibility of hazardous radioactive material; robustness of facilities and transport systems against acts of sabotage instigated by insiders and/or external attacks by force or stealth; minimization of MC&A and physical protection costs per GWyr*

Criteria Definitions

Nation-state proliferation (PR&PP-1) involves the acquisition of weapon-usable nuclear material[a] in amounts sufficient for the production of one or more nuclear weapons through:

a. The *diversion* of weapon-usable nuclear material from declared inventories or flows associated with normal use in a nuclear energy system, or produced as a normal and declared byproduct of such a system; and/or

b. Misuse by the *undeclared production* of plutonium, ^{233}U or other weapon-usable material through clandestine irradiation of undeclared fertile material in a nuclear power reactor, or production of undeclared highly enriched uranium in an enrichment plant.

Diversion and undeclared production are deterred by a high probability of *timely detection* by international safeguards.

Nuclear terrorism (PR&PP-2) involves either the:

a. *Theft* of weapon-usable nuclear material from nuclear installations or transport systems for the production of one or more nuclear explosive devices;

b. *Theft* of hazardous radioactive material from nuclear installations or transport systems for the production of one or more radiation dispersal weapons; or

c. *Damage or sabotage* of a nuclear energy installation or transport system with the intention of causing the release of radioactive material.

Theft is deterred by the intrinsic characteristics of the nuclear material and the barriers that confine it, as well as by extrinsic MC&A and physical protection measures applied by the state. Characteristics that deter sabotage include the following: intrinsic characteristics of the fuel matrix; additional mechanical barriers that impede radiological dispersal; intrinsic characteristics of storage and processing facilities that impede access to or dispersal of stored materials, including wastes; and intrinsic characteristics of reactor safety systems that impede access to vital equipment, impede compounding failures of such equipment, and mitigate any radiological dispersal event. Sabotage is further deterred by extrinsic physical protection measures applied by the state that prevent unauthorized access to vital equipment and to stored materials.

PR&PP-1 gives credit to Generation IV nuclear energy systems that present formidable barriers to nation-state proliferation by limiting the availability of attractive forms of weapon-usable material and by enhancing the detectability of diversion or undeclared production of such material, sufficiently that a nation-state would be highly unlikely to choose a Generation IV nuclear energy system as a route to acquire weapons usable material.

PR&PP-2 gives credit to Generation IV nuclear energy systems that present formidable barriers to attempts at theft or sabotage, sufficiently that a terrorist group would be highly unlikely to choose a Generation IV nuclear energy system as a route to nuclear terrorism.

Today, the risks of nation-state proliferation and nuclear terrorism are addressed through a spectrum of measures, most of which are common to any nuclear energy system. The nuclear threat

---

a. "Weapon-usable nuclear material" includes plutonium with any isotopic composition with the exception of "heat-source" plutonium containing 80% or more of the isotope ^{238}Pu, uranium enriched in the isotopes ^{233}U and/or ^{235}U to 20% or more, or other fissionable material having physical properties suitable for such purposes, including ^{237}Np.

response spectra include: measures to detect acts of proliferation or terrorism; efforts to defeat such acts, once detected; efforts to mitigate the impact of such acts, if they cannot be prevented; the pursuit and apprehension of perpetrators; and the legal processes of bringing the perpetrators to justice. All of these current activities are also likely to be applicable at the time that Generation IV nuclear energy systems enter the market place, and the scope and intensity of each is likely to continue to keep pace with perceived threats.

The features available to counter threats of proliferation and terrorism can be grouped into the categories of *intrinsic* and *extrinsic barriers*:

- *Intrinsic barriers* are defined by material qualities (isotopic composition, chemical separability, mass and bulk, fuel matrix radiation level, dilution and detectability characteristics), and by technical impediments that are inherent to a nuclear system, such as facility unattractiveness and accessibility, mechanical impediments to material and vital equipment access, skill requirements, and

- *Extrinsic barriers* involving institutional controls, such as materials control and accounting (MC&A) and physical protection performed by the nation-state to prevent theft and sabotage; and the detection of diversion and misuse performed by international safeguards; which may be facilitated by inherent characteristics of the system, and by the specific agreements that a nation is signatory to.

Complementary intrinsic features which facilitate the application of extrinsic measures, such as the installation of verification systems to be used by the IAEA when applying safeguards to detect diversion or undeclared production, and State MC&A and physical protection, are classified here as *intrinsic barrier elements* because they are provided as physical, integral components in the system.

The selection of promising Generation IV candidate systems for future R&D will give credit to intrinsic barriers. Later, for nuclear energy systems approaching commercialization, the evaluation framework may expand to include a broader scope of considerations. For example, evaluations encompassing a broader range of considerations should be anticipated in the context of specific export approvals and for the implementation of IAEA safeguards.

Credit can also be gained by using a *compensating intrinsic feature* to offset an apparent deficiency. For example, for a nuclear power reactor that might allow the introduction of bogus fuel elements facilitating undeclared plutonium production, the installation of instruments for use by the IAEA to verify that all fuel elements are bona fide to offset a deficiency that might otherwise jeopardize the prospects for a promising candidate.

In addition to the intrinsic barriers identified above, the susceptibility of Generation IV nuclear energy systems to nation-state proliferation and nuclear terrorism will be further reduced through the implementation of extrinsic measures. Such additional complementary measures include the institutional framework under which a State might acquire a Generation IV nuclear energy system, and the verification measures applied to ensure that once made available, the Generation IV nuclear energy system remains dedicated to peaceful use, and State measures of MC&A and physical protection. The obligations on a State developing or importing a Generation IV nuclear energy system are likely to be established in part by international non-proliferation undertakings, including those arising as a Party to the Treaty for the Non-Proliferation of Nuclear Weapons (the NPT) or regional nuclear weapon-free zone treaties, safeguards agreements with the International Atomic Energy Agency[b] (the IAEA), including the

---

b. Preferably comprehensive IAEA safeguards agreements based upon INFCIRC/153.

Additional Protocol to such safeguards agreements[c], by the Physical Protection Convention, and such other instruments as may come into force prior to the availability of Generation IV nuclear energy systems. Similarly, any exports of Generation IV nuclear energy systems are expected to be in compliance with Zangger Committee and Nuclear Suppliers Group (NSG) Guidelines in effect at the time of such exports.

The risks of proliferation and terrorism may also be reduced through appropriate implementation arrangements undertaken within specified terms. Examples might include all-in / all-out fuel supply arrangements and multi-national energy parks.

PR&PP-1—Minimize life-cycle susceptibility to the diversion or undeclared production of weapon-usable material; facilitate implementation of effective IAEA safeguards

Definition: Generation IV systems will employ features that facilitate the application of effective and inexpensive international safeguards over the entire lifecycle of materials, to effectively deter diversion and undeclared production.

Discussion: The ability of the IAEA to apply effective safeguards is measured by its ability to detect, in a timely manner, the diversion or undeclared production of one or more significant quantities of nuclear material from peaceful activities, and to deter such diversion or undeclared production through the risk of early detection.[2] The effectiveness of safeguards is measured by the probability of detection of an attempt at diversion or undeclared production, if perpetrated, and the estimated time available between the detection of material diversion and the potential fabrication of a nuclear explosive device. Longer time periods (between detection of diversion or undeclared production and the completion of one or more weapons) increase the available range of international responses to diversion, and thus increase how effectively diversion might be deterred. The safeguards effort must be evaluated for the full lifecycle of nuclear materials, from mining until (and if) the materials are no longer useful for any nuclear application, or are determined to be consumed or diluted in such a way that they are practicably irrecoverable.[e]

---

c. Ref: INFCIRC/540.

d. The IAEA has established parametric values for IAEA safeguards for the different types and forms of nuclear materials:

Nuclear Material	Significant Quantity	Timeliness	Detection Probability
Pu, ^{233}U, HEU	8 kgs Pu or ^{233}U; 25 kg ^{235}U in U containing ≥20% ^{235}U	Separated: 30 days; In spent fuel: 3 months	for monthly conclusions: 50% for gross and partial defects; for annual determinations: 90% for gross, partial and bias defects
LEU	75 kg ^{235}U in U containing <20% ^{235}U	1 year	50% for annual determinations for gross, partial and bias defects
Natural uranium	10 t	1 year	50% for annual determinations for gross and partial defects
Depleted uranium, thorium	20 t	1 year	50% for annual determinations for gross and partial defects

e. These are the conditions for termination of IAEA safeguards under comprehensive safeguards agreements (ref: INFCIRC/153, Article 11). In such cases, the determinations are to be made by the IAEA.

A host of factors determine safeguards effectiveness and the costs for implementation. These include the physical and chemical form of the material and its degree of irradiation; the type of operations carried out on the material; the number of facilities, their capacities and design features; the effectiveness of the State system of accounting for and control of the material, the features provided in a facility for inspection access to the material, the availability of operator measurement systems that can be shared by the IAEA, and other capabilities within the State existing at the time the Generation IV nuclear energy system(s) are introduced, or added later on.

Viability and performance evaluations will determine the lifecycle resources required to provide safeguards meeting Zangger Committee and NSG guidelines for candidate fuel cycles. Current guidelines for international safeguards will be used as a baseline for evaluation. Because these guidelines can be expected to evolve prior to Generation IV deployment, evaluation will also perform sensitivity studies for the effects of potential future enhancements of the guidelines, such as potential adoption of a requirement for signing of an Additional Protocol. The evaluations will also quantify the resulting time period between detection of diversion or misuse and possible fabrication of nuclear explosives, and translate that timeliness into a measure of effectiveness for all stages in the lifecycles of the materials.

Proposed Metrics

Avoid Separated Weapons-Usable Materials

Final Screening: Separated materials (i.e. separated cleanly from fission products) require larger resource investments for international safeguards.

Separated weapons-usable material requires more intensive international safeguards and provides less time for international response if diversion occurs. While international safeguards resource requirements for separated weapons-usable materials can be minimized by systems which collocate facilities and minimize inventories, final screening credits those systems which maintain or exceed intrinsic barriers comparable to those provided by LWR fresh fuel.

Scoring: The reference is the case where the fresh fuel is refractory low enrichment uranium oxide fuel clad in zircalloy, typically containing uranium of < 5% enrichment. The metric goes from “much lower than reference” to “better than reference”, taking into account alternative fuel materials and differences in physical, radiological and chemical characteristics.

Relative risk.

Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference
HEU, Pu, Np in forms than can be readily separated	HEU, Pu, Np fuels in matrices designed to resist reprocessing	LEU fuels, natural uranium, fuels containing Th; HEU, Pu, Np fuels with intense radiation barriers	DNLEU,[f] Th fuels providing additional barriers to material access/recovery

f. DNLEU = depleted, natural or low enrichment uranium

Credit Spent-Fuel Characteristics that Impede Handling and Recovery of Fissionable Material that is Well Suited for Use in Manufacture of Nuclear Weapons

Final Screening: High burn-up fuels increase the difficulty of handling and chemical separations, and reduce the attractiveness of the recovered material. [g]

High fuel burn-up, to levels equal or greater than current high-performance LWR fuels, decreases the total quantity of fuel requiring handling and management, increases the concentration of the Cs-137 which provides the primary radiation barrier in interim spent fuel storage, and increases heat generation rates in the residual plutonium in spent fuel, significantly delaying chemical recovery of weapons-usable material.

Scoring*:* The reference is the case where spent low-enrichment uranium-oxide fuel has been irradiated to 50,000 MWd/MTHM or more.

Relative risk.

Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference
Low burn-up fuels in forms that can be readily separated; any very low burn-up fuel or blanket assemblies	Low burn-up fuel or blanket assemblies	>50,000 MWd/MTHM		Long-lived core integrated with reactor vessel, no onsite spent fuel storage

PR&PP-2–Minimize Vulnerability to Theft of Weapon-Usable Material or Hazardous Radioactive Material; Minimize Vulnerability of Installations and Transport Systems to Sabotage

Definition: Generation IV fuels and facilities will have intrinsic characteristics that minimize the life-cycle vulnerability to theft and sabotage.

Discussion: Protection against theft and sabotage is the responsibility of the State. The resources required for materials control and accounting (MC&A) and physical protection that meet current U. S. Nuclear Regulatory Commission requirements and INFCIRC/225 guidelines, as well as relevant Zangger Committee and NSG guidelines for nuclear materials and facilities will be assessed through all phases of a proposed fuel cycle, including for materials placed in permanent geologic storage. Because these guidelines can be expected to evolve prior to Gen IV deployment, evaluation will also perform sensitivity studies for the effects of potential future enhancements of the guidelines for MC&A and physical protection, such as potential adoption of periodic international inspections of the adequacy of the physical protection measures. Credit will be given to nuclear energy systems that minimize the lifecycle resources required for MC&A and physical protection, per GWyr of electricity (or equivalent) production. The major characteristics of fresh and spent fuel that provide intrinsic resistance to theft are already covered

g. It is recognized that higher burn-up fuels require higher enrichment in the fresh LEU fuels, and that increasing the LEU enrichment would decrease the separative work required to bring the LEU to high enrichment levels suitable for its use in nuclear weapons. However, the overriding interest is to make the plutonium as unsuitable as possible.

by PR&PP-1. Thus, for final screening, PR&PP-2 focuses on the characteristics of reactor facilities which create higher passive resistance to sabotage, that reduces the size and response time requirements for physical protection forces required to meet the design-basis threats specified by regulatory requirements (i.e. 10 CFR 73 in the United States).

Proposed Metric

Reactors Have Passive Features That Resist Acts of Terror or Sabotage.

Final Screening: Passive systems for reactivity shutdown, heat removal, and radionuclide confinement require fewer resources for extrinsic physical protection measures than do active safety systems.

Passive reactivity shutdown and decay-heat removal systems can eliminate the requirement for safety-grade AC power and cooling water supply. Passive systems require only infrequent surveillance and can therefore be placed in enclosures that greatly impede rapid personnel access, which reduces the size and response time requirements for physical protection forces.

Physically robust aboveground structures, or below-grade construction, are also important to protect vital equipment from external missiles such as aircraft. However, such robust structures or below-grade construction can be provided for all systems, and thus does not differentiate between systems. Therefore final screening focuses on the intrinsic safety-system features that minimize access to vital equipment.

Scoring: The reference is an advanced light water reactor with an active emergency core cooling system.

Relative risk.

Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference
Substantially easier physical access to vital equipment than an advanced LWR	Somewhat easier physical access to vital equipment than an advanced LWR	Emergency cooling system using safety-grade AC power and an external cooling water source	Passive safety systems that require external control signals for activation	Fully passive safety systems isolated from rapid personnel access, no control activation signals required

Criteria and Metrics for Safety and Reliability Goals

General Comments for the Sections on the Safety and Reliability Goals

The Generation IV Roadmap "Technology Goals for Generation IV Nuclear Energy Systems" document introduces the safety and reliability goals in a way that provides an organizing principle to the full set of safety and reliability goals and criteria:

> *Safety and reliability are essential priorities in the development and operation of nuclear energy systems*. During normal operation or anticipated transients, nuclear energy systems must preserve their safety margins, prevent accidents, and keep accidents from deteriorating into more severe accidents. At the time, competitiveness requires a very high level of reliability and performance.
>
> There has been a definite trend over the years to improve the safety and reliability of nuclear power plants, reduce the frequency and degree of off-site radioactive releases, and reduce the possibility of significant damage. Generation IV systems have goals to achieve the highest levels of safety and reliability and to better protect workers, public health, and the environment through further improvements. The three safety and reliability goals continue the past trend and are in accord with the regulatory policy to have designs that are safe and minimize the potential for severe accidents and their consequences.

It is important to recognize that the safety and reliability goals are in accord with the regulatory policy of all GIF partners. In particular, the following discussion draws upon the defense in depth policy of IAEA and a generalized view of risk that together provide the unifying logic for all the safety and reliability goals and their respective criteria.

The Generation IV goals related to safety and reliability seek a global and comprehensive improvement of the safety related architecture (i.e. engineered and passive safety systems, inherent characteristics, etc.). This underlying goal translates into a recommendation for the improvement of the entire defense in depth system.

In the framework of safety the final objective is the reduction of the risk (frequency and consequences) linked to the installation under examination. The improvement in safety and reliability for Generation IV systems will be most transparent and convincing, when there is a reduction to all the accident categories/families, starting from the frequency of operational occurrences, including anticipated transients, and the probability of design extension conditions (former "beyond design basis") that include the "severe plant conditions" (i.e., core melting.)

The INSAG 10 - Defense in Depth in Nuclear Safety–document provides indications on how to achieve an improvement on defense in depth:

"*5.1 IMPROVEMENTS IN DEFENSE IN DEPTH*

122. The approach for further improvement of defense in depth is similar for existing and for future plants. However, for future plants such improvements can be achieved in a more systematic and complete way. This includes:

- *improving accident prevention, in particular by optimizing the balance between the measures taken at different levels of defense in depth and by increasing their independence;*
- *improving the confinement function.*

124. Possible means for strengthening accident prevention are:

- *increased thermal inertia;*
- *optimized human-machine interfaces;*
- *extended use of information technology;*
- *reduced complexity;*
- *improved maintainability;*
- *expanded use of passive features;*
- *more systematic consideration of the possibilities of multiple failures in the original plant design.*

125. The confinement function for advanced reactors will be strengthened by approaches and initiatives consistent with the following systems:

- *For advanced designs, it would be demonstrated, by deterministic and probabilistic means, that hypothetical severe accident sequences that could lead to large radioactive releases due to early containment failure are essentially eliminated with a high degree of confidence.*
- *Severe accidents that could lead to late containment failure would be considered explicitly in the design process for advanced reactors. This applies to both the prevention of such accidents and mitigation of their consequences, and includes a careful, realistic (best estimate) review of the confinement function and opportunities for improvement in such scenarios.*
- *For accident situations without core melt, it will need to be demonstrated for advanced designs that there is no necessity for protective measures (evacuation or sheltering) for people living in the vicinity of a plant. For those severe accidents that are considered explicitly in the design, it would be demonstrated by best estimate analysis that only protective measures that are very limited in scope in terms of both area and time would be needed (including restrictions in food consumption).*

5.2. LEVELS OF DEFENSE IN DEPTH FOR THE NEXT GENERATION OF PLANTS

126. Meeting the safety objectives set for the next generation of nuclear power plants will necessitate improving the strength and independence of the different levels of defense. The aim is to strengthen the preventive aspect and to consider explicitly the mitigation of the consequences of severe accidents consistent with the initiatives stated in Section 5.1. This development would include the following trends:

- *Level 1, for the prevention of abnormal operation and failures is to be extended by considering in the basic design a larger set of operating conditions based on general operating experience and the results of safety studies. The aims would be to reduce the expected frequencies of initiating failures and to deal with all operating conditions, including full power, low power and all relevant shutdown conditions.*
- *Level 2, for the control of abnormal operation and the detection of failures, is to be reinforced (for example by more systematic use of limitation systems, independent from control systems), with feedback of operating experience, an improved human-machine interface and extended diagnostic systems. This covers instrumentation and control capabilities over the necessary ranges and the use of digital technology of proven reliability.*
- *Level 3, for the control of accidents within the design basis, is to consider a larger set of incident and accident conditions including, as appropriate, some conditions initiated by multiple failures, for which best estimate assumptions and data are used. Probabilistic studies and other analytical means will contribute to the definition of the incidents and accidents to be dealt with; special care needs to be given to reducing the likelihood of containment bypass sequences.*
- *Level 4, for the prevention of accident progression, is to consider systematically the wide range of preventive strategies for accident management and to include means to control accidents resulting in severe core damage. This will include suitable devices to protect the containment function such as the capability of the containment building to withstand hydrogen deflagration, or improved protection of the basemat for the prevention of melt-through.*

Level 5, for the mitigation of the radiological consequences of significant releases, could be reduced, owing to improvements at previous levels, and especially owing to reductions in source terms."

The following relationship between the Generation IV Goals and defense in depth can be suggested:

<table>
<tr><th>Levels of Defense in Depth</th><th>Objective</th><th>Essential Means</th><th>Gen IV Goals</th></tr>
<tr><td>Level 1</td><td>Prevention of abnormal operation and failures</td><td>Conservative design and high quality in construction and operation</td><td rowspan="2">Safety and Reliability –1. Generation IV nuclear energy systems operations will excel in safety and reliability.</td></tr>
<tr><td>Level 2</td><td>Control of abnormal operation and detection of failures</td><td>Control, limiting and protection systems and other surveillance features</td></tr>
<tr><td>Level 3</td><td>Control of accidents within the design basis</td><td>Engineered safety features and accident procedures</td><td>Safety and Reliability –2. Generation IV nuclear energy systems will have a very low likelihood and degree of reactor core damage.</td></tr>
<tr><td>Level 4</td><td>Control of severe plant conditions, including prevention of accident progression and mitigation of the consequences of severe accidents</td><td>Complementary measures and accident management</td><td rowspan="2">Safety and Reliability–3. Generation IV nuclear energy systems will eliminate the need for offsite emergency response.</td></tr>
<tr><td>Level 5</td><td>Mitigation of radiological consequences of significant releases of radioactive materials</td><td>Off-site emergency response</td></tr>
</table>

CRITERIA AND METRICS FOR SAFETY AND RELIABILITY GOAL 1

Goal Statement:

Safety and Reliability – 1 (SR1). Generation IV nuclear energy systems operations will excel in safety and reliability.

Evaluation for goal SR1 focuses on safety and reliability during normal operation of all facilities in the nuclear energy system, from mining to the final disposal of waste. Thus the focus is on those high to medium probability events that set the forced outage rate, control routine worker safety, and result in routine emissions that could affect workers or the public. Assessment during screening focuses on unique system characteristics that can impact reliability and unusual design aspects that could significantly affect worker safety and routine emissions. Quantitative measures of parameters affecting plant reliability are added in the viability and performance evaluations, when a system design is sufficiently developed. Goal SR1 considers facility attributes operable at the first two levels of defense in depth, as described above in the introduction to "Criteria and Metrics for Safety and Reliability Goals", i.e., those features that can reduce the frequencies of initiating failures for all potential operating conditions and that can control abnormal operation and detect failures.

Summary table of criteria and metrics.

Criterion		Final Screening & R&D Prioritization	*Viability and Performance Evaluations*
SR1-1	Reliability	Screen for unique characteristics affecting forced outage rate and lines of defense (++/+/=/-/--)	*Forced outage rate probability distribution*
SR1-2	Public and worker safety–routine exposures	Screen for the possibility of unique routine exposure to radiation, chemical, and toxic hazards (+/+/=/-/-)	None
SR1-3	Public and Worker safety – accidents	Screen for unique radiation, chemical, toxic, handling hazards (++/+/=/-/--)	None

Criteria Definitions

SR1-1–Reliability

Definition: Generation IV nuclear energy systems will excel in reliability.

Discussion: Plant reliability affects both safety and economics. The impact on economics occurs primarily through capacity factor and maintenance costs; it is discussed at the end of this section. The role of plant reliability in safety aligns with defense in depth levels 1 and 2, prevention and control. Factors that lessen the chance of forced outage avoid the opportunity for accident sequences to develop. Forced outages can occur due to failures that directly preclude operation of the plant, and by failures that cause the plant to operate outside the limits set by its technical specifications. Low forced outage rates imply excellence in system design, maintenance, and operation, to prevent failure events of relatively high frequency from occurring, and from propagating to create conditions requiring plant shutdown. Under appropriate regulatory oversight, low forced outage rates also imply excellent performance in maintaining plant parameters and safety system availability and reliability inside the design safety limits specified by the plant technical specifications. High availability and reliability of safety systems reduces the probability that initiating events can lead to core damage.

The problem with emphasizing prevention in safety is that initiating events are not all created equal. The interactions among systems, given an initiating event and the impact of those interactions on plant thermal-hydraulic performance play a crucial role in identifying risk significant scenarios. Furthermore, attempting to identify major plant impact from simple metrics (e.g., the number of safety systems) or from actual or hypothesized event descriptions, without plant-specific design details, as-built configuration, O&M practice, or plant-specific PRA is flawed. It ignores all the lessons we have learned in 25 years of doing plant-specific PRA. Risk impact is plant-specific and requires careful analysis. In well-designed facilities, risk comes primarily from unexpected interactions among systems and from internal[h], external[i], or combinations of events that defeat designed redundancy or expected systems responses. Another way of saying this is that risk does not come from combinations of best estimate or most likely conditions, but from less likely and more challenging situations. Therefore, a systematic, integrated examination of facility response against the safety criteria of interest, such as PRA, is needed, if the subtleties that affect risk are to be evaluated.

The impact of plant reliability on public safety is quantified under SR2, where PRA can calculate the frequency of plant states that can challenge the core and the plant. Identification of those states is an iterative part of the PRA, between mechanistic (thermal-hydraulic, neutronic, electrical, and mechanical) analysts and systems analysts. Mechanistic analyses are used to set success criteria on system functions, such that meeting those success criteria will ensure no serious challenge to the core or the plant. Note that the challenging plant states are generally a small subset of those that contribute to plant forced outages; they are the states that also partially disable mitigating systems (either unannounced failures of safety systems or conditions outside technical specifications). Thus improving plant reliability for production may have little or no impact on safety, unless the frequency of the challenging plant states is reduced (and reducing that frequency may have no discernible effect on availability, because the challenging states are often relatively minor contributors to shutdowns). However, if the frequencies of all contributors to forced outage rate are reduced (including those associated with challenging plant states), improvement in safety and reliability will result.

In the final screening process, which occurs well before the design is fully specified and before operations and maintenance practices are established, it is not possible to make a meaningful calculation of forced outage rate. The uncertainties are so great that the results could not discriminate among systems. Nevertheless, unique features in design may offer early insight into the factors that could have major impact on forced outage rate. For example, redundancy in major secondary plant equipment (e.g., main turbine generator, condenser) is minimal in Generation II and III plants to keep capital costs low and consequently they have been significant contributors to downtime (forced and planned). So unique features that offer improvements over Generation III plants can provide early indication of potential. Examples of such features can affect forced outage rate (and capacity factor, i.e., economics) could include:

- Enhanced redundancy and diversity (functional redundancy) that can improve both reliability and capacity factor

h. Relevant internal events could include operator errors, operation outside technical specifications, equipment failure, fire, flooding, missile generation, pipe whip, jet impact, or release of fluid from failed systems or from other installations on the site. Note that plants with fewer and simpler technical specifications would have less chance of such events.

i. Relevant external events could include earthquakes, floods, high winds, tornadoes, tsunami (tidal waves) and extreme meteorological conditions.

- Advanced control and monitoring systems that can reduce the cognitive challenges to operators, can improve forced outage rate and can minimize routine maintenance (improving capacity factor by eliminating the need for shutting down and opening equipment for inspection)
- Advanced control and monitoring systems that can flag oncoming failure thereby minimizing the chance of catastrophic failure (improving forced outage rate) and reducing repair time by replacing forced outages with planned outages (improving capacity factor)
- Design features that can improve average thermal efficiency over Generation II and III plants (fluctuations in efficiency introduce de-ratings, i.e., departures from 100% power; reducing them can improve capacity factor)
- Design features that facilitate and simplify maintenance while optimizing the use of building space thereby reducing the chance of errors (improving forced outage rate) and making maintenance more efficient (improving capacity factor)
- Plant simplification offers the opportunity to improve reliability by reducing chances for failure and error; however, it can also reduce the opportunities for recovery
- Safety system simplifications that reduce the number and complexity of technical specifications

Proposed Metrics

Plant Forced Outage Rate

Final Screening Metric: Screen for unique characteristics affecting forced outage rate and lines of defense.

Only unique features need be considered; features that are common to all systems cannot discriminate among them. Focus should be on design features and lines of defense that distinguish a system from Generation III facilities, during all phases of operation. The following table lists a number of features that should be considered when evaluating this criterion. A vulnerability leading to a higher forced outage rate calls for improvements in the number or quality of the lines of defense or an R&D program to reduce the vulnerability.

Design Features Affecting Forced Outage Rate that Could Distinguish the System from Generation III	
• Frequency of initiating events • Experience with key components, materials, thermal cycling, corrosion, and aging • Scaling in size of components • Vulnerability to common cause failure • Ease of maintenance; low vulnerability to error • Load following capability • Accommodate loss of offsite power without reactor trip	• Sensitivity of operating plant to external events such as earthquakes, floods, and fires • Degree of use of advanced control systems; clarity of these systems to operators and cues to operators under various modes of system failure • Time for operators to take actions or intervene before the plant trips • Fewer or simpler technical specifications
Characteristics of Lines of Defense that Could Distinguish the System from Generation III	
For all systems (e.g., reactivity control, reactor	• Redundancy and diversity

Design Features Affecting Forced Outage Rate that Could Distinguish the System from Generation III	
heat removal, power conversion) that can affect forced outages directly or by technical specifications	• Flexibility to cover a wide range of conditions • Simplicity of configuration • Degree of reliance on external power sources versus passive systems • Response of safety systems to external events such as earthquakes, floods, and fires • Time for operators to take actions or intervene before damage results • Optimized human-machine interfaces

Each TWG must use its judgment to assess the likelihood that the unique factors affecting the system under evaluation improves or degrades the forced outage rate.

Plant forced outage rate: metric scale.

Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference
Forced outage rate increased by more than a factor of 5 and number or quality of lines of defense degraded	Forced outage rate increased by a factor of 5 or number or quality of lines of defense degraded	Forced outage rate unchanged and lines of defense unchanged	Forced outage rate unchanged and number or quality of lines of defense improved	Forced outage rate reduced by factor of 5 and number or quality of lines of defense improved

SR1-2: Worker and Public Safety and Routine Exposures

Definition: Generation IV nuclear energy systems will excel in safety and will not expose workers and the public to significant risk via routine exposure to radiation or hazardous material.

Discussion: Given the premise of high quality design, monitoring, and operation, routine exposure should be minimal. This is true in most well designed and managed industrial facilities, including nuclear facilities, today. It is important to identify unique hazards. However, risk is a matter of hazards and safeguards. Generally, routine exposure is a consequence of poor management and practices, rather than inherent in design concept. Nevertheless, evaluators must be alert to special aspects of each system. The role of routine exposure in safety aligns with defense in depth levels 1 and 2, prevention and control.

Even if worker safety is protected, a unique hazard could cause additional maintenance cost associated with time delays and staff hours associated with controlling unusual hazards. While that is not a reliability issue, it will be most efficient to identify such potential costs during this evaluation.

Proposed Metrics

Routine Exposure to Radiation or Hazardous Materials

Final Screening Metric: Screen for unique routine radiation, chemical, and toxic hazards, during handling, transport and all other phases of operations (+/=/-).

Evaluators must be alert to unusual potential for routine exposure from each system. Possible hazards would include coolant compatibility with humans and environment. Evaluators must also separate design issues from management issues. Designs that avoid or minimize management control can be advantageous.

Routine exposure: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Significantly greater risk of routine personnel or public exposure compared to Generation III	Risk of routine personnel or public exposure about the same as Generation III	Significant reduction of risk of routine personnel or public exposure compared to Generation III

NOTE: *Given the low level of exposure in current designs, a very strong case must be made to support the evaluation of "better than reference"; note that improvements that affect O&M costs must be credited in the economic criteria. For this reason, only three ranges are considered*.

SR1-3: Worker/Public Safety–Accidents

Definition: Generation IV nuclear energy systems will excel in safety and will not expose workers to significant accident hazard, involving radiation, hazardous materials, or severe physical conditions. Radiological releases from major plant accidents is the subject of Safety and Reliability 2 and 3.

Discussion: As in routine exposure, personnel accidents are more often a function of company management and culture than inherent to the design. Still, it is clear that the hazard presented by one facility may be greater than that presented by another. When that is true, evaluators must examine the protection against that hazard, the safeguards, to see if risk is balanced by such measures. So the first step is to identify any unique hazards, those not present in other facilities. The hazards may be radioactive, chemically active, toxic, or physical (e.g., high temperature or pressure). Note that the role of plant reliability in safety aligns with defense in depth level 3, control of accidents within the design basis. While it could, therefore, align better with SR2, it is retained under SR1 because it follows similar evaluation steps with other SR1 criteria.

When hazard-screening analysis identifies unique hazards, evaluators must follow up with safeguards screening to ensure workers and the public are protected at a level commensurate with the potential for harm. In difficult cases, calculation of the risks, probabilities and consequences, may be necessary to discriminate among systems, but this is not expected. As in other criteria, both intrinsic and extrinsic protection is possible. Intrinsic, “designed in” protection can be more convincing to observers and may be more reliable.

Proposed Metrics

Accidental Exposure to Radiation, Hazardous Materials or Physical Conditions

Final Screening Metric: Screen for unique radiation, chemical, toxic, and physical hazards, during handling, transport and all other phases of operations (+/=/-).

Evaluators must be alert to unusual potential for accidental exposure to radiation.

Accidental exposure: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Significantly greater risk of accidental personnel or public exposure compared to Generation III	Risk of accidental personnel or public exposure about the same as Generation III	Significant reduction of risk of accidental personnel or public exposure compared to Generation III

NOTE: *Given the low level of exposure in current designs, a very strong case must be made to support the evaluation of "better than reference"; note that improvements that affect O&M costs must be credited in the economic criteria. For this reason, only three ranges are considered.*

Economics Notes:

1. Capacity factor should also be calculated to support economics evaluation.
2. Improved reliability could also lead to reduced manning requirements, reducing the O&M costs.
3. Maintenance costs associated with special hazards, should be identified during this evaluation and considered under the economics criteria.

Criteria and Metrics for Safety And Reliability Goal 2

Goal statement:

Safety and Reliability–2 (SR2). Generation IV nuclear energy systems will have a very low likelihood and degree of reactor core damage.

Evaluation for goal SR2 identifies facility attributes that, using models and experiments, create high confidence that all design basis accidents (DBA) are correctly managed and that reactor core damage will have a very low likelihood or can be excluded or practically excluded by design (and in other facilities, that the release of radioactive material from its most immediate confinement or nuclear criticality can not occur.) For performance evaluations much of the information required for a Preliminary Safety Analysis Report (PSAR) will be available, so the likelihood of core (or other facility) damage can be evaluated quantitatively for specific Design Extension Conditions. Results can be presented as a frequency probability distribution which reflects all sources of uncertainty in models and experiments. For preliminary design information available at final screening, an approximate analysis of the safety related architecture using the level of defense (LOD) analysis, as described above in the introduction to “Criteria and Metrics for Safety and Reliability Goals,” can identify conflicts with safety fundamentals. The final screening identifies major design characteristics that are likely to robustly bound potential transient power, temperature, chemical reaction, and mechanical stresses well inside damage thresholds. Equally important, the screening credits design approaches that facilitate the modeling and experiments required to predict quantitatively the uncertainty of safety margins.

Summary table of criteria and metrics.

Criterion		Final Screening and R&D Prioritization	*Viability and Performance Evaluations*
SR2-1	Robust engineered safety features	(++/+/=/-/--)	*Probability distribution for core damage frequency (or release from normal configuration for non-reactor facilities), combined with number/quality of levels of defense*
SR2-2	System models have small and well-characterized uncertainty (physical models / well-scaled experiments)	(++/+/=/-/--)	

Criteria Definitions

SR2-1: Robust Engineered Safety Features

Definition: Generation IV facilities will have engineered safety features and/or inherent features (for reactors: power control, heat removal, and radionuclide confinement) that will transparently bound the accessible range of operating and accident conditions and will allow the facility state to be predicted with very low uncertainty, inside this range of conditions.

Discussion: To provide high confidence that damage is precluded or has very small probability for all the plausible plant conditions [i.e.: design-basis accidents and design extension conditions, DEC, (see the introduction to the S&R Goals above)], the accessible range of facility operating and accident conditions must be bounded by inherent design characteristics and by simplicity of the technical specifications that guide facility operation. Inside these boundaries the performance and reliability of safety related design features depends on the application of excellent design practice:

- Redundancy
- Prevention of common mode failure due to internal or external hazards, by physical or spatial separation and structural protection
- Prevention of common mode failure due to design, manufacturing, construction, commissioning, maintenance or other human intervention, by diversity or functional redundancy
- Automation to reduce vulnerability to human failure, at least in the initial phase of an incident or an accident
- Testability to provide clear evidence of system availability and performance
- Qualification of systems, components and structures for specific environmental conditions that may result from an accident or an external hazard.

Some of these design features can be introduced during detailed design; others are inherent to fundamental characteristics of systems. Final screening focuses on fundamental features that are likely to support high confidence and transparency in predictions of low core damage probability. For reactors, simple (often passive) reactivity control, heat removal and radionuclide confinement methods reduce the complexity of system interactions and require less intensive surveillance to confirm operability. Robust fuel and core designs with long thermal time constants maintain more predictable geometry and thermophysical properties over the full range of accessible plant states.

Proposed Metrics

Reliable Reactivity Control in Reactors

Final Screening Metric: Identify simplicity and robustness of systems for reactivity control in reactors.

Systems receive higher ranking when core damage from reactivity insertion is precluded, even for inadvertent removal of multiple reactivity control elements. For all systems at least two additional independent and diverse mechanisms for reactivity shutdown must also be provided.

Reliable reactivity control: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Positive temperature or void reactivity coefficient can exist under some operating modes, and/or less than two independent and diverse mechanisms are provided for reactivity shutdown	Power, temperature and void reactivity coefficients are inherently negative, or inherently have no safety-related effects, during normal and anticipated transients and all modes of operation	Inherent design features preclude core damage from reactivity insertion due to inadvertent movement of multiple reactivity control elements, and temperature and void reactivity coefficients are inherently negative

Reliable Heat Removal for Reactors

Final Screening Metric: Identify simplicity and robustness of systems for decay heat removal in reactors.

This metric credits decay heat removal approaches that maximize simplicity and minimize the number and complexity of decay-heat-removal subsystems, while still being capable of achieving the redundancy, diversity, and other design goals listed above.

Reliable heat removal: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference
Decay heat removal system is more complex than current evolutionary LWRs	Decay heat removal system is similar to current evolutionary LWRs	Decay heat removal system uses no AC power	Decay heat removal system always operates and has no moving parts

SR2-2: System Models Have Small and Well-Characterized Uncertainties

Definition: Generation IV systems will be governed by dominant phenomena and phenomena interactions that can be predicted with very high and well-bounded certainty using models and experiments.

Discussion: DBA analysis and calculation of damage-frequency probability distributions requires physically based models with uncertainties that have been accurately characterized by comparison with separate effects and integral experimental data. This screening criterion identifies system attributes that are likely to reduce modeling uncertainty. Some phenomena, such as conduction and single-phase convective heat transfer in channels, can be predicted with low uncertainty using appropriate data from well-designed and instrumented separate effects experiments. Other phenomena, such as critical heat flux and strongly multi-dimensional flows, are more complex and introduce greater uncertainty in modeling. Well-scaled integral experiments are required to confirm the completeness and accuracy of integral models.

Proposed Metrics

Dominant Phenomena Can Be Modeled with Small and Well Characterized Uncertainty

Final Screening Metric: Identify and penalize systems with dominant phenomena that are difficult to characterize or bound with high accuracy.

Systems where all dominant phenomena can be accurately characterized and studied experimentally will receive highest ranking. Systems where dominant phenomena are stochastic or difficult to characterize and bound will be ranked lower. Systems will be preferred where separate effects experiments for dominant phenomena can be performed at full scale, and where the initial conditions or time-dependent boundary conditions can be replicated accurately in separate effects experiments.

Low uncertainty in modeling dominant phenomena: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Some dominant phenomena are difficult to accurately characterize and predict, and must be treated with bounding analysis	Some dominant phenomena can be studied only in scaled separate-effects experiments requiring extrapolation of experimental results	All dominant accident transport phenomena can be studied in well-instrumented separate effects experiments with negligible scale distortions and characterized with well understood probability distributions, for the full range of environmental conditions that may result from an accident or an external hazard

Interactions Between Phenomena in Different Spatial Regions Are Simple to Characterize in Integral Modeling

Final Screening Metric: Long fuel thermal response time.

A key source of uncertainty in predicting potential fuel damage comes from the fuel response to the evolution of phenomena in other regions of the reactor system. Systems with long fuel thermal response time constants are preferred, because the fuel response becomes insensitive to the detailed transient evolution of phenomena in other regions. For gas coolants this requires a large fuel and internals thermal inertia. For liquid coolants this requires designs that preclude core uncovery.

Long fuel thermal response time: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Fuel and coolant thermal inertia lower than current evolutionary PWRs	Fuel and coolant thermal inertia has characteristics similar to current evolutionary PWRs	Inherent fuel and coolant thermal inertia provides much longer core thermal response times than evolutionary PWR fuel under design-basis accident transients

NOTE: *An evolutionary ABWR does not experience core uncovery during a large-break loss of coolant and would be evaluated as better than reference*

Validation With Integral Experiments

Final Screening Metric: Minimal scaling distortion in integral experiments.

Systems are preferred that permit integral testing at prototypical scale, under prototypical operating conditions, of all safety systems designed prevent core damage. Such testing reduces uncertainties about magnitudes of scaling distortions and increases confidence that systems codes include all relevant and important phenomena.

Minimal scaling distortion in integral experiments: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference
Scaling or other constraints generate significant distortions in integral testing	Integral experiments are performed in well-scaled facilities at reduced geometric scale	Integral experiments to study power plant DBAs can be performed at prototypical scale	All safety systems function continuously during normal operation of the power plant, and all dominant safety-related parameters can be monitored continuously during plant operation

SR2 Reference

1. B.E. Boyack et al., "An overview of the code scaling, applicability, and uncertainty evaluation methodology," *Nuclear Engineering and Design*, Vol. 119, pp. 1-16, 1990.

Criteria and Metrics for Safety and Reliability Goal 3

Goal Statement:

Safety and Reliability-3 (SR3). Generation IV nuclear energy systems will eliminate the need for offsite emergency response.

Evaluation for goal SR3 considers system attributes that allow demonstration, with high confidence, that the radioactive release from any scenario results in doses that are insignificant for public health consequences. Such confidence must come from the knowledge that reactor core damage (Design Extension Conditions, DEC, as described in the above introduction to “Criteria and Metrics for Safety and Reliability Goals”) has very low probability (SR1 and 2) and that mitigation features provide additional lines of defense to account for any significant residual risk. This confidence comes from three sources: accurate bounding prediction of the timing and magnitude of radioactive source terms and energy releases; accurate assessment of the effectiveness of the confinement system in accommodating the all bounding energy releases and providing holdup of radioactive material; and assessment of the resulting off-site dose probability distribution and comparison against appropriate standards for individual and societal risk. For screening, a system is qualitatively ranked according to accident and release potential relative to present nuclear energy systems. This screening includes assessment of how well severe-accident phenomena can be characterized and modeled for the system. For later viability and performance evaluations, quantitative evaluation of damage, release and transport, and comparison of resulting dose relative to public health criteria, are used.

Summary table of criteria and metrics.

Criterion		Final Screening and R&D Prioritization	*Viability and Performance Evaluations*
SR3-1	Radioactive source/energy release magnitude and timing understood and bounded by inherent features	+/=/-	*Offsite dose probability distribution*
SR3-2	Confinement or containment provides robust mitigation of bounding source and energy releases	+/=/-	
SR3-3	No additional individual risk	n/a	*Quantitative*
SR3-4	Societal risk comparable to competing technology	n/a	*Quantitative*

Criteria Definitions

SR3-1: Radioactive Source/Energy Release Magnitude and Timing Understood and Bounded By Inherent Features

Definition: In Generation IV systems, the potential timing and magnitude of the release of radioactive material from the core, and energy from all potential internal and external sources, will be understood, minimized, and bounded by inherent features of the fuel and confinement structures including thermal inertia and chemical stability.

Discussion: The principal of defense-in-depth, as discussed in the introduction to S&R goals, requires that Generation IV reactor systems include independent confinement or containment systems that can provide sufficient hold-up of radioactive materials to meet off-site dose goals, for the physically possible timing and magnitude of releases of radioactive material. This independent confinement or containment system must be designed to function appropriately when subjected to the full range of physically plausible energy source magnitudes and timing that could be derived from internal or external sources.

- **Excellent fuel damage resistance.** Those fission products which are gases, or become vapors at high temperatures, can be mobilized by fuel damage at elevated temperatures. Highly robust fuel forms can delay and reduce fission product releases after fuel is damaged by high temperatures and/or oxidation, and can prevent the propagation of failure to neighboring fuel material. These fuel features can delay the release and reduce the fraction of the fission products that can be mobilized, compared to the source term from fuel damage in current LWRs. Thus, these features mitigate the effects of reaching the conditions for fuel damage in metric 1.1.2 below.

- **Bounded and controllable energy releases.** The confinement/containment structures can be subjected to a range of energy releases with the potential to damage the structures' capability to hold up radioactive materials. Systems are preferred where the timing and magnitude for potential release of all internal stored energy sources, and external energy sources, can be predicted and bounded with high confidence as indicated in metric 1.2.2 below.

- **Predictability of source term/energy release.** Systems are preferred where detailed, mechanistic models can accurately predict time-dependent probability distributions for the fractional release, and physical and chemical form, of radionuclides released from fuel subjected to overheating and/or oxidation damage (or for damage to the most immediate confinement or criticality for non-reactor facilities), and predict the timing and magnitude of all energy releases that could damage confinement and containment structures These factors also affect metric1.2.2 below. For reactors these models will be based on phenomenological models for structural and fuel damage, and for fission product transport to the reactor coolant system, for a spectrum of accident sequences. These models will be validated against well scaled and instrumented experiments.

Proposed Metrics

Excellent Fuel Damage Resistance

Final Screening Metric: Lower source term than predicted for LWR severe accidents.

Systems are preferred where design features of the fuel reduce the released source term, compared to the source term for LWRs (NUREG-1465 "Accident Source Terms for Light-Water Nuclear Power Plants: Final Report" US NRC, Feb 1995,) for a spectrum of potential damage mechanisms. The source term is defined as the ratio of the activity-weighted fractional release of fission products from the fuel into the confinement, compared to the values for LWR fuel, under conditions of substantial core damage. The source term provides the key input for analyzing the effectiveness of the containment/confinement barrier.

Source term: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Bounding fractional release from degraded core is a significantly greater than LWR fuel	Bounding fractional release from degraded core is similar to LWR fuel	Bounding fractional release from degraded core is a factor of 10 less than LWR fuel

Well Bounded, Understood, and Limited Number of Mechanisms for Significant Energy Releases

Final Screening Metric: Assess the potential for R&D to minimize the number of energy release mechanisms and provide high certainty in energy release predictions used to design the containment/confinement system.

Systems are preferred where the number of significant mechanisms for energy release are limited, and where approaches exist, with reasonable resource investment, to provide high certainty in the timing and magnitude of the energy-releases for all potential internal and external energy sources. For comparison, for Gen. III LWRs, the potential significant severe-accident energy release mechanisms of relatively high complexity include blow-down, hydrogen generation and combustion, core-coolant interaction, direct-containment heating, and core-concrete interaction. The number of significant energy release mechanisms is used as a surrogate for the probability of containment/confinement damage and bypass given a core damage event.

Mechanisms for energy release: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Number of significant, complex energy release mechanisms from internal sources is greater by a factor of 2 than a Gen. III LWR	Number of significant, complex energy release mechanisms from internal sources is similar to a Gen. III LWR	Number of significant, complex energy release mechanisms from internal sources is reduced by a factor of 2 from Gen. III LWR

SR3-2: Confinement or Containment Provides Robust Mitigation of Bounding Source and Energy Releases

Definition: Generation IV systems will provide confinement or containment systems that provide sufficient hold-up to reduce off-site doses to levels that preclude harm to the public, for the bounding range of radioactivity source terms and energy releases.

Discussion: Generation IV systems will provide robust, independent mitigation features that will preclude harm to the public even in the event of any significant damage to reactor cores that might be generated by a spectrum of very-low probability event sequences. These systems will be designed to accommodate the release of stored energy in the system, as well as external energy sources. These features will include inherent mechanisms which create long delays for any release, and will make the magnitude of any residual release sufficiently small to meet off-site risk goals.

- **Long time constants.** For reactors inherent reactivity shutdown mechanisms and very long fuel thermal time constants can delay the onset of any fuel damage for long periods of time following the occurrence of conditions that would lead to fuel damage, i.e., delays until well after initiating events such as loss-of-coolant. Because highly improbable combinations of system failures would be required to cause fuel damage, substantial periods of time are available to diagnose and correct failures.

- **Long and effective holdup.** For reactors, fission product barriers or additional mitigation features, independent of the fuel robustness, will provide effective retention of any aerosols formed from volatile fission products and will greatly delay and control any residual release of gaseous fission products. For all facilities, the structural integrity of all mitigation systems will be robust against damage by all stored energy sources present in the system, and the system design will effectively dissipate or eliminate stored energy sources to reduce the probability of damage to mitigation systems. Fission product barriers and mitigation systems will withstand the effects of external events such as earthquakes, fires and floods. Particular attention will be paid to eliminating the potential for bypass of mitigation systems or confinement.

- **Transport.** During performance evaluation, assessment of the transport of radionuclides following fuel damage, will be performed using the same general methods, to the same level of quality, as residual heat removal modeling described in SR2. Phenomenological models to assess the effectiveness of mitigation features will include time-dependent models for transport and deposition of radionuclides, as well as models for mechanisms where stored energy sources could damage mitigation systems. These models will be validated using data from well-scaled separate effects experiments. While integral experiments can not be performed in the prototypical plant, all transport related integral experiments will be designed to have small and well characterized scaling distortions.

- **Dose.** Offsite dose calculations should be performed using the most technically realistic source terms. Licensing level diffusion calculations should be used to determine the dose at the site boundary to the average individual with no sheltering or evacuation. The dose must be predicted to be well below the EPA Emergency Protection Guidelines. The acceptability of the EPA EPG criterion as an ultimate licensing criteria is questionable and is, therefore, only used in the screening steps. Evaluation steps must employ a rigorous analysis showing attainment of the U. S. NRC Safety Goal Policy (modified for a complete nuclear energy system) as discussed separately below. Systems that rely on a large exclusion zone to accomplish this low dose result will imply requirements for remote siting and are, therefore, less preferable than systems employing more robust fuel or inherent mitigation features.

Proposed Metrics

Long System Time Constants

Final Screening Metric: Long time constants for all potential core damage mechanisms.

Systems are preferred where long time constants exist for the occurrence of all potential severe fuel damage events, including thermal, chemical, and mechanical damage.

Long time constants for potential core damage mechanisms: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Intrinsic features could permit severe core damage directly following initiating event	Intrinsic features could permit severe core damage one hour after initiating event	Intrinsic features delay severe core damage by over 24 hours following initiating event

Long and Effective Holdup

Final Screening Metric: Containment/confinement systems include robust independent mitigation features to greatly delay and reduce any release following core damage.

Systems are preferred that employ an effective, independent containment or confinement system that is robust in performance, provides small release fraction, and is resistance to damage or bypass for a spectrum of core damage event sequences. The release fraction is defined as the fraction of radioactive material released into the containment/confinement volume that is subsequently transported to and released to the environment.

Long and effective holdup: metric scale.

Worse Than Reference	Similar to Reference	Better Than Reference
Release fraction for containment/confinement system is greater than Gen. III LWR by factor of 10	Release fraction for containment/confinement system is comparable to Gen. III LWR	Release fraction for containment/confinement system is less than Gen. III LWR by a factor of 10

Criteria and Metrics for Economic Goals 1&2

Goal Statement:

Economics 1 (EC1): Generation IV nuclear energy systems will have a life-cycle cost advantage over other energy sources.

Economics 2 (EC2): Generation IV nuclear energy systems will have a level of financial risk comparable to other energy projects.

We define a cost metric: (EC1) Average Cost as a function of (EC1-1) Overnight Construction Costs and (EC1-2) Production Costs (including fuel, labor, materials, and waste management and disposal costs). We define a risk metric: (EC2) Capital-at-Risk (during construction) as a function of unit and plant size, (EC1-1), and (EC2-1) Construction Duration. These are discussed below in the order of data collection and calculation, i.e., EC1-1, EC1-2, EC2-1, EC2, and EC1. Research and Development costs are considered elsewhere.

Although much of the quantitative information on these costs and measures will not be available until R&D has been completed, we suggest comparing projected costs at each stage of screening and evaluation with the costs and construction duration of Advanced Light Water Reactors. The goal of the economic roadmap is to develop a business plan for each potentially successful Generation IV technology. Each step of the screening and evaluation process is one step closer to creating this business plan.

Summary table of criteria and metrics.

Criterion	Final screening and R&D prioritization	*Viability evaluation*	*Performance evaluation*
EC1-1: Low Overnight Construction Costs	Compare with ALWR costs of ~$1500/kW; identify cost uncertainties	Project ranges of overnight construction cost and operating lifetime	Estimate overnight construction cost distribution
EC1-2: Low Production Costs	Compare with ALWR production costs of ~$15/MWh	Project cost ranges for O&M and Fuel (including waste management expenses)	Estimate production and fuel cycle cost distributions per unit of output
EC2-1: Short Construction Duration	Compare with ABWR construction duration of 48 months	Project range of construction duration for first unit and plant	Estimate distribution of construction durations for single and multiple units
EC2: Low Capital-at-Risk	Compare with ALWR Capital at Risk of $1,800M	Project range for Capital-at-Risk, optimal plant size, and common costs	Estimate distribution of Capital-at-Risk
EC1: Low Average Cost	Compare average cost (including capital additions and decommissioning costs) with a reference market price of ~$32/MWh	Project ranges of average cost and market prices of electricity	Estimate distributions of average cost and market clearing prices

Criteria Definitions

EC1-1: Low Overnight Construction Costs

Definition: Generation IV systems will minimize the cost of constructing generating units.

Discussion: Construction costs have been the most costly aspect of generating electricity from the current generation of nuclear power plants. These costs have been driven by many characteristics. Four of these cost drivers have been:

- The lack of simplicity
- Large structural volumes
- The lack of scalability
- The lack of standardization and modularity.

Here, we suggest using a step-by-step approach to assess overnight capital costs. These include the costs of land, structures, and equipment. Other capital costs, including the costs of financing will be addressed below. (We discuss construction duration in EC2-1.) We suggest identifying information on construction cost drivers to determine whether overnight capital costs for Generation IV technologies will be more or less than for Generation III technologies.

Metric:

We rely on NEA, *Projected Costs of Generating Electricity: Update 1998* (Paris: OECD). NEA's Table 7 describes "Nuclear power plant investment costs discounted to the date of commissioning (US$ of 1st July 1996/kWe)." These plants are either commercially available or are expected to be commercially available between 2005–2010. We assume these are "Nth-of-a-kind" (NOAK) costs, defined in ORNL, *Cost Estimate Guidelines for Advanced Nuclear Power Technologies* (Oak Ridge, 1993). (Because of different interpretations of what should be included in "contingency," we do not consider contingency in the Final Screening; instead, we focus on developing an understanding of the probability distribution of Generation IV costs.)

In NEA (1998), the median "base construction cost" is $1557. The value for the US observation is $1441. Therefore, we define the range of ALWR overnight construction cost to be between $1400 and $1600/kW. (While inflation has decreased the value of the dollar since 1996, and normally we would inflate the OECD values to 2001 dollars, we believe that the OECD values are still relevant in today's dollars and provide reasonable values with which to benchmark Generation IV technologies.) We propose the following linear scoring for anticipated overnight construction costs:

EC1-1.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>$2000/kW	$1800–2000/kW	$1600–1800/kW	$1400–1600/kW	$1200–1400/kW	$1000–1200/kW	<$1000/kW

If dollar values are unknowable, use qualitative evaluation, e.g., "construction costs are much better (lower) than ALWR." To aid in the estimation of overnight construction costs at each stage of screening and evaluation, consider the following plant characteristics. (Note: To convert capital costs per kW into costs per MWh see discussion in EC1.)

EC1-1. Final Screening

1. *Primary equipment and components of major generating unit subsystems*

   Because technologies with fewer primary components are usually cheaper to build, generating units with fewer major components and structures should be cheaper to build than an ALWR.

2. *Containment volume and structural footprint of anticipated generating unit as a function of size*

   Because the containment volume and structural footprint of the generating unit drives the cost of structures, generating units that involve smaller containment volumes and structural footprints per unit of output should be cheaper to build than an ALWR.

3. *Identify minimum efficient size of the generating unit*

   Minimum efficient size describes the size at which most (90%) of the scale economies are exhausted. Power systems that exhibit lower minimum efficient sizes should be able to compete in more markets. (Also, smaller plants usually involve lower Capital-at-Risk during construction, see EC2.) Identify both unit size and anticipated plant size.

4. *Generating unit subsystems that can be manufactured in modules*

   Because of the declining costs of mass (serial) production, and improved quality control, technologies with more subsystems that can be produced in central locations and transported to generation sites should receive a higher score. (Modules need not be identical to achieve economies.)

5. *Describe how generating unit construction costs change with the number of same-technology generating units constructed (or total capacity).*

   Technologies that exhibit continually declining capital costs (i.e., few production capacity constraints) in the number of units constructed should receive a higher score.

EC1-2: Low Production Costs

Definition: Generation IV systems will minimize production costs.

Discussion: This section addresses all non-depreciable production inputs, including the cost of waste management. (Depreciable costs, i.e., for equipment with a service life of more than one year, are considered capital additions and are included in EC1.) While fuel inputs vary with electricity production for most generating technologies, nuclear fuel accounting is more complex. Therefore, although it has many characteristics of a depreciable capital asset, we treat it as a production cost. Also, while budgets and labor contracts are set annually (or over longer periods), we treat all non-depreciable operations and maintenance (O&M) expenditures as production costs. In the Final Screening evaluators should compare O&M plus Fuel (including waste management costs) to Generation III experience and projections. The Viability and Performance Evaluation should improve on these estimates.

Metric:

We rely on NEA (1998) Table 16 “Projected generation costs calculated with generic assumptions at 10% p.a. discount rate.” The median “O&M + Fuel” (including waste management) projection is $13.74/MWh, the mean is $14.61/MWh, the standard deviation is $5.10, and the US estimate is $14.78/MWh. We define the range of ALWR production costs to be between $14.00 and $16.00/MWh.

These costs represent approximately $9/MWh for O&M and $6/MWh for Fuel (including $1/MWh for spent nuclear fuel management). Therefore, if the power plant is likely to have Fuel and O&M costs that diverge from these values, explain how each of these costs is similar to or different from these reference values. Also, we assume that all plants have a lifetime capacity factor of 90%. We propose the following linear scoring for anticipated production costs:

EC1-2

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>$20/MWh	$18–20/MWh	$16–18/MWh	$14–16/MWh	$12–14/MWh	$10–12/MWh	<$10/MWh

If dollar values are unknowable, use qualitative evaluation, e.g., “production costs are much lower than ALWR.” To aid in the estimation of production costs at each stage of screening and evaluation consider the following plant characteristics.

EC1-2. Final Screening:

1. *Estimate expected fuel requirements per unit of energy output.*

   For the primary fuel form, estimate expected fuel requirements per unit of output (e.g., tons per GW-year). Technologies that are more fuel efficient, holding other costs equal, should have lower fuel costs. (Information comes from evaluation of SU1-1.)

2. *Describe refueling process.*

   Descri A. W. be refueling activities in terms of labor, materials, and downtime. Technologies that are easier to refuel should be cheaper to operate.

3. *Identify anticipated major chemical, radioactive, or mixed wastes.*

   Discuss any special maintenance or personnel requirements. Technologies with smaller, less toxic, and easier to manage waste volumes should have lower waste management costs. (Information comes from evaluation of SU2-1.)

EC2-1: Short Construction Durations

Definition: Generation IV systems will minimize construction duration.

Discussion: Non-construction capital costs are dominated by Interest During Construction (IDC), which depends primarily on construction duration (and expenditure profile) and the cost of capital charged by financial markets. Construction expenditures are addressed in EC1-1. The cost of capital is addressed in EC2. Here, we address construction duration: the time between "Construction Start" (defined by IAEA as "Date when first major placing of concrete, usually for the base mat of the reactor building, is done.") to Commercial Operation (defined by the IAEA as "Date when the plant is handed over by the contractors to the owner and declared officially to be in commercial operation.") See, for example, IAEA, *Nuclear Power Reactors in the World* (April 2001). The Viability and Performance Evaluations should improve on these estimates.

Metric:

We rely on the construction duration of the ABWR in Japan of 48 months within a 10-month range. We assume linear scoring in a range between about 2 and 6 years. Again, if construction duration is unknowable, use qualitative evaluation, e.g., "construction duration is much lower than ABWR."

EC2-1.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>75 months	65–75 months	55–65 months	45–55 months	35–45 months	25–35 months	<25 months

EC2-1. Final Screening:

1. *Estimate construction duration*.

   Estimate construction duration for the generating unit. Note: There are three phases in nuclear plant construction. Phase 1 includes site preparation and excavation, design engineering, and equipment procurement. Phase 2 begins with "construction start" and ends with start of fuel loading. Phase 3 includes fuel loading and safety testing. Discuss the length of each phase for the first unit and subsequent units. Estimate the durations of Phase 2 and 3 (during which most of the construction expenditures are spent).

2. *Estimate construction duration; consider possibilities for modular production*

   Because of shorter construction durations with mass production, technologies with more subsystems that can be produced in central locations should be quicker to build. Describe how modular production will shorten construction duration.

EC2: Low Capital-at-Risk

Definition: Generation IV systems will minimize Capital-at-Risk.

Discussion: During the period of construction, typically the owner is responsible for financing plant construction costs. Capital-at-Risk measures the total accumulated investment in the project during the construction period. It includes in constant dollars both the overnight cost (the physical cost of building the plant and indirect costs) and Interest During Construction (IDC), which depends primarily on *construction duration* (and expenditure profile) and the *cost of capital* charged by financial markets. Capital-at-Risk measures the total amount of capital that must be obtained to finance the complete construction of the first unit, i.e., until the project is capable of generating power and earning a return. Both bankers (providing loans) and private investors (providing equity) are interested in this measure because it indicates the total funding that needs to be dedicated to a specific project before revenues are generated. In general, the lower the total investment required, the lower the risk.

For single-unit plants, the Capital-at-Risk is equal to all overnight construction costs plus IDC. For multiple-unit plants Capital-at-Risk includes common costs, overnight construction costs, and IDC on the first unit. Therefore, to calculate Capital-at-Risk, evaluators must identify the anticipated size of the generating unit and size of the plant. The software will calculate Capital-at-Risk as a function of Overnight Construction Costs and Construction Duration for the first unit. (It assumes that common costs are $150 million for a 1,000 MW plant. Adjustment to EC1-1 will be made to account for common costs.)

IDC and the present value of other capital costs is a function of the cost of capital. Throughout this analysis we use real costs, and thus we use real costs of capital, abstracting from the general level of inflation. As in ORNL (1993), we assume the escalation rate is equal to the inflation rate. This is equivalent to assuming that real "escalation during construction" is zero.

Finally, risk premiums charged to owners of nuclear power plants after construction reflects the probability of losing asset value due to an accident (see discussion associated with SR1-1) or regulatory action, and the risk of default on capital obligations.

Metric:

First, we assume a 10% *real* discount rate (following suggested practice by the US Office of Management and Budget; this discount rate should decrease as financial markets become more familiar with Generation IV technologies).

Second, we rely on NEA (1998) Table 7. In addition to the "base construction cost," the table presents estimates of "Contingency, Interest during construction, Major refurbishments, and Decommissioning" at discount rates of 5% and 10%. (We discuss refurbishments and decommissioning in EC1.) To aid in the estimation of IDC, we use the following "rule of thumb."

We assume a uniform spending rate. So, IDC is approximately equal to the discount rate times one-half the construction duration. For example, with a 10% discount rate, if the construction duration were 4 years, IDC would be approximately 20% of the construction cost. If construction costs were $1500, then IDC would be $300.

Capital-at-Risk = [(First Unit Size in kW · Overnight Cost/kW) + Common Costs] ·

(1 + [10% · (Construction Duration/2)])

Considering a 1,000 MW (single-unit) ALWR as the standard, with a $1,500/kW overnight construction cost and a construction duration of 4 years, the Capital-at-Risk would be

1,000 MW · $1,500/kW · (1 + 20%) = $1,800M

The highest and lowest estimates for overnight construction costs and construction duration will be combined to obtain the range for Capital-at-Risk.

EC2 Capital-at-risk.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>$3000M	$2500–3000M	$2000–2500M	$1500–2000M	$1000–1500M	$1000–500M	<$500M

NOTE: *These values have been changed in Version 3 to provide a wider range with a linear value function.) Again, if dollar values are unknowable, use qualitative evaluation, e.g., "Capital-at-Risk is much lower than ALWR." The Viability and Performance Evaluation should improve on these estimates.*

EC2. Final Screening:

1. *Estimate Capital-at-Risk*

 This will be calculated by the software with information on unit and plant size and on other economic metrics.

EC1: Low Average Cost:

Definition: Generation IV systems will be have average costs lower than the market clearing price of electricity.

Discussion: This section addresses the bottom line. Cost estimates per unit of electricity for each cost category are summed by the software to determine the life-cycle cost per unit of energy and scored with respect to the reference. Some nuclear technologies can integrate the production of electricity with the production of other commercial products. The Viability and Performance Evaluations will assess the affect of these other commercial products on competitiveness.

Metric:

Determine Average Cost:

Average Cost is equal to the sum of

a. Overnight Construction Costs per MWh (from EC1-1),

b. Interest During Construction per MWh (see below),

c. Production Costs per MWh (from EC1-2),

d. Capital Additions per MWh (see below), and

e. Contributions per MWh to a Nuclear Decommissioning Trust Fund (or equivalent, see below).

1. **Overnight Construction Costs plus IDC per MWh**: Here, EC1-1 is converted into a Capital Cost per MWh. This is done by multiplying Overnight Construction Costs and IDC by the Capital Recovery Factor (CRF) and dividing by the number of MWh generated annually. (We assume that all plants have an 90% Capacity Factor.)

$$\text{Capital} = [(\text{ Construction Cost} + \text{IDC}) \cdot \text{CRF}] / (\text{CF} \cdot 8760)$$

$$\text{CRF} = [r \cdot (1 + r)^T] / [(1 + r)^T - 1] = (0.1 \cdot 1.1^{40}) / (1.1^{40} - 1) = 10.2\%$$

$$\text{Capital} = \frac{[(\text{Construction Cost} \cdot [1 + (10\% \cdot \text{Construction Duration} \cdot 0.5)]) \; 10.2\%]}{(90\% \cdot 8760)}$$

$$\text{Capital} = (\$1{,}800{,}000/\text{MW} \cdot 10.2\%) / (90\% \cdot 8760) = \$23.29/\text{MWh}$$

NOTE: *The capital recovery level per MWh for the Overnight Construction Cost (without IDC) would be $19.40 in this example.*

2. **Production Costs**: Production Costs were calculated on a per MWh basis in EC1-2. They were compared with \$15/MWh.

3. **Capital Additions:** Capital additions include all production costs that have a productive life of more than one year. We assume Capital Additions of the Generation IV plant are \$2/MWh.

4. **Decommissioning Costs**: The implicit assumption in NEA (2000) for decommissioning US plants is one-third of the construction cost, discounted 40 years to the start of operations. We adopt this assumption. If construction costs were \$1500/kW, decommissioning costs would be \$500/kW. These costs must be accumulated over the life of the plant, T (e.g., 40 years), so that decommissioning could begin at the end of the operating life. We assume that the return on Nuclear Decommissioning Trust Funds is 5% (real).

   $$\text{Decomm} = (1/3 \cdot \text{Construction Cost}) \cdot (r / [(1 + r)^{T} - 1]) / (CF \cdot 8760)$$

   $$\text{Decomm} = (1/3 \cdot \$1500) \cdot (0.05 / [(1.05)^{40} - 1])] / (90\% \cdot 8760) \qquad = \$0.52/\text{MWh}$$

   In the example values used here, Average Cost is equal to \$23.29/MWh plus \$15/MWh plus \$2/MWh plus \$0.52/MWh, or \$40.81/MWh. These are much higher than the reference market-clearing price. Therefore, for Generation IV costs to be competitive, they must be significantly lower than ALWR costs in one or more cost categories.

   We use a \$20/MWh range of electricity. Scoring is linear within this \$20 range.

EC1.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>\$42/MWh	\$42–38/MWh	\$38–34/MWh	\$34–30/MWh	\$30–26/MWh	\$26–22/MWh	<\$22/MWh

EC1. Final Screening:

1. Estimate range of Average Cost

   Average Cost for highest Overnight Construction Cost, longest Construction Duration, highest Production Cost, Capital Additions, and Decommissioning Cost; will scored at the left end of the range. Average cost for lowest Overnight Construction Cost, shortest Construction Duration, lowest Production Cost, Capital Additions, and Decommissioning cost will be scored at the right end of the range.

2. Compare approximate Average Cost/MWh with current electricity prices.

   The software will estimate Average Cost/MWh and determine the appropriate scores.

REFERENCES

1. International Atomic Energy Agency, *Nuclear Power Reactors in the World* (Vienna: IAEA, April 2001).

2. International Energy Agency/Nuclear Energy Agency, *Projected Costs of Generating Electricity: Update 1998* (Paris: OECD, 1998).

3. ORNL, *Cost Estimate Guidelines for Advanced Nuclear Power Technologies* (Oak Ridge, TN: ORNL/TM-10071/R3, May 1993).

Appendix B

Attributes for Generation III Nuclear Energy Systems – Final Screening Criteria

Appendix B

Attributes for Generation Iii Nuclear Energy Systems – Final Screening Criteria

Criterion	Reference
	ALWR - 4.2–5.0% Initial Enrichment 50,000MWd/MtU, 32%thermal efficiency, 1GWe system
SU1-1: Fuel Utilization	~150–200 tonnes natural U (feed)/GWeYr
SU2-1: Waste minimization: mass	15–20 Mt/GweYr
SU2-1: Waste minimization: volume	15–20 m^3/GweYr
SU2-1: Waste minimization: decay heat, long-term	1–3 kW/GweYr after 500 yrs
SU2-2: Waste minimization: radiotoxicity (long-term)	500–1,500 MSv/GweYr after 500 yrs
SU2-3: Stewardship burden	N/a
PR&PP-1: Minimize material and facility vulnerability: Separated materials	LEU fuels, natural U, fuels containing Th; HEU, Pu, Np fuels with intense radiation barrier
PR&PP-1: Minimize material and facility vulnerability: Spent fuel characteristics	~ 50,000 MWd/MtHM
PR&PP-2: Passive features to resist sabotage	Emergency cooling system using safety-grade AC power and an external cooling water source
SR1-3: Reliability	1 forced outage/year
SR1-1:Routine worker exposure	82 man-rem/unit-Yr
SR1-2: Worker-accident exposure	Design features presenting worker accident risks comparable to Gen III reactor systems
SR2: Safety and Reliability	Design features comparable to Gen III reactor systems
SR3: Safety and Reliability	Design features comparable to Gen III reactor systems
EC1: Overnight Construction Cost	$1,400–1,600/kWe
EC1: Production Cost	$14–16/kWe
EC2: Construction duration	45–55 months
EC2: Low capital at Risk	$1500–2000M
EC1: Average Cost	$30–34/MWh

Appendix C

Theoretical Basis Supporting the Probability Based Methodology

Appendix C

Theoretical Basis Supporting the Probability-Based Methodology

The concept potential will be evaluated with respect to the Gen IV Goals by using the criteria and metrics developed for this screening.

The method described here leads to the formulation of an integrated performance score for a concept, based upon the contribution of each performance criterion to the overall Gen IV Goals. The method supports uncertainty of concept scoring against each criterion through the use of probability distributions.

The individual performance attributes of Table C.1 can be viewed as elements in our overall performance vector, $\overline{\mathrm{X}}$, where

$$\overline{\mathrm{X}} = \left[\mathrm{X}_1,\ \mathrm{X}_2,\ \ldots,\ \mathrm{X}_n\right] \tag{1}$$

For example, one could select

$$X_1 = \text{SU1} - 1\ \text{Fuel utilzation}$$

$$\vdots$$

$$\mathrm{X}_n = \text{EC - 5 High profitability.}$$

For a particular concept the values of its performance metrics will only be known within some uncertainty bounds. However, for each metric, X_i, a subjective probability density function, f(x), can be formulated, stating the beliefs of an evaluator of the relative likelihood of the various possible values of X_i (see Figure C.1).

The value of a particular performance metric can be categorized as summarized in Table C.2, where the label of a category reflects the degree of contribution to successful performance in the area of interest.

Because the value of any performance metric can only be known subject to a level of uncertainty, as reflected in the functional dependence of f(x), one can state the probability that a concept will have a performance score falling in a particular performance quality range, k, as

$$\text{Prob.}_{\cdot i}(k) \equiv p_k \equiv \int_{X_{K_{\min}}}^{X_{K_{\max}}} f(x)dx \tag{2}$$

Table C.1. Categorization of Performance Metric Values to Success in the Area of Interest.

Range Number (k)	Performance Quality	Range of Values of X in That Quality Range
1.	Much better	$X_{max} = X_{E_{max}} \geq X > X_{A_{max}}$
2.	Better	$X_{A_{max}} \geq X > X_{M_{max}}$
3.	Similar	$X_{M_{max}} \geq X > X_{F_{max}}$
4.	Worse or much worse	$X_{F_{max}} \geq X > X_{min}$

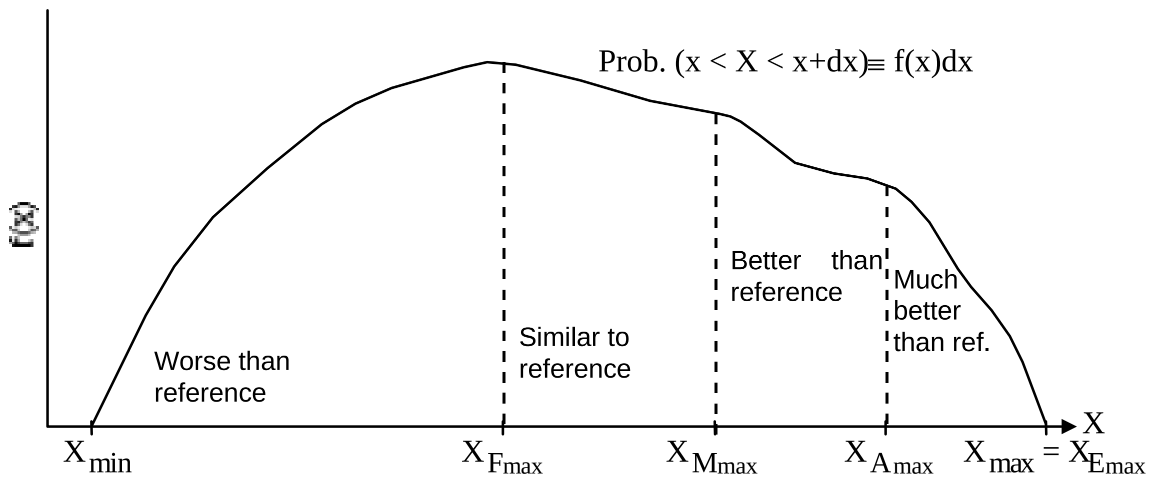


Figure C.1. Probability Density Function of Random Variable, X

To determine the overall measure of the potential of the concept, an overall success likelihood is determined. Each performance area (criterion) will have a weighted importance for overall success, W_i, which will need to be pre-established. The EMG has identified a set of weights for the relative importance of the criteria under a goal.

In order to combine the probability density functions of the estimated performance of a system for various criteria, a value function system is used. The EMG has provided a uniform linear value function for all the criteria. To roll up several criteria distributions into a single figure of merit for the performance potential under a goal, the probability density functions are discretized. The discrete probabilities under each segment and the value function for the segment are used in the roll up operation. Table C.2 shows an example of discretized distributions and value functions for sample criteria.

Table C.2. Example of development of scores for the different criteria.

	S_k - value function										
Criterion	*-1*	*-0.8*	*-0.6*	*-0.4*	*-0.2*	*0*	*0.2*	*0.4*	*0.6*	*0.8*	*1.*
1					0.1	.03	.04	.02			
1								0.35	.055	0.1	

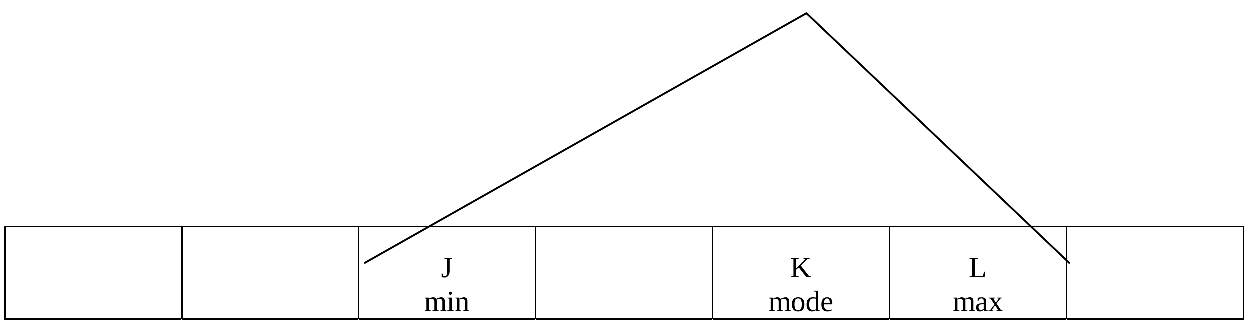
N						0.2	0.2	0.2	0.2	0.2	

Combining Potential Under Different Criteria into Overall Concept Potential

Once the TWGs have determined the distributions of performance potential of the concept with respect to each criterion (expressed only as a set min, mode, and max value locations), a success probability of the concept with respect to each goal will be estimated.

The values provided by the TWGs will be used to build either a flat, triangular, or bi-modal distribution. Assume that the TWG has selected a triangular distribution and the discrete intervals j, k, and l for the minimum, mode, and maximum values:

:



A distribution is developed with the lower end set at the left edge of the “min” interval, the peak at the center of the “mode” interval, and the upper end at the right edge of the “max” interval. Note that the same interval may be selected for two or even all three of the values.

For the uniform distribution, the same process is used, but the distribution is flat. For the bi-polar distribution, two equal triangles are developed, each one interval wide, located at the upper and lower “modes” selected.

The spreadsheet sets the height of the distribution such that the area under the distribution is normalized to 1. Next, it converts to a discrete distribution by subdividing each interval into three equal parts and calculating the associated areas. The result is the “bar chart” distribution shown on the evaluation spreadsheet. (Note: Three sub-intervals were determined to be the minimum number required to differentiate between distributions that are only one interval wide.)

Propagation of the Distributions of Potential

A value function is associated with each of the criteria metric scales. The value function varies from -1 to +1. In general, the function is linear (-1 at the low end of the range and +1 at the high end, with 0 corresponding to the center of the discrete interval of the reference value).

-1.	-.9	-.8	-.7	-.6	-.5	-.4	-.3	-.2	-.1	0.0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
			Much worse than reference			Worse than reference			Similar to reference			Better than reference			Much better than reference					

The probability distribution is then expressed internally in the spreadsheet as a discrete probability distribution (DPD), x:

$x = \{< p_i, x_i >\}$,

where p_i are the probabilities associated with the discrete x_i values of the value function enveloped by the distribution. $\sum p_i = 1$. These probabilities correspond to the area under the distribution for each discrete sub-interval.

To roll up the score for all the criteria under one goal into a single "goal score" a new DPD is calculated. A criterion weight is assigned by the EMG to each of the n criteria under the goal, w_k, k=1 to n; $\sum w_k = 1$.

The DPD at the goal level, g, is estimated as follows:

$$g = x\, w_1 + y\, w_2 + \ldots + z\, w_n$$

where x, y, …, and z are the DPDs for each of the criteria under the goal

Two DPDs are combined as follows:

$$g = x + y; \quad x=\{<p_i, x_i>\},\ y=\{<q_j, y_j>\}$$

$$g = \{<r_{ij}, g_{ij}> \}$$

$$r_{ij} = p_i q_j$$

$$g_{ij} = x_i\, w_1 + y_j\, w_2$$

Note that the number of "doublets" expressing the resulting DPD is i times j. Normally, this would result in an exponential growth in the number of doublets as successive criteria are combined. To avoid this, a condensation procedure to reduce the number of doublets is carried out. The value of each doublet is used to prorate the probability into one or two of the original discrete values.

For example, assume two criteria:

Criteria 1:

weight = 60%

distribution = [{0.2, 0.2}, {0.6, 0.3}, {0.2, 0.4}]

where the distribution is defined by sets of {prob, value}

Criteria 2

weight = 40%

distribution = [{0.7, 0.4}, {0.3, 0.5}]

To calculate a weighted pair, we take { (PiCi * PjCj), ((ViCi * WCi) + (VjCj * WCj)) }

where P = probability, C = criteria, V = value, and W = weight

The resulting weighted pairs (before condensation) for all combinations of i & j would be:

[{0.14, 0.28}, {0.42, 0.34}, {0.14, 0.40}, {0.06, 0.32}, {0.18, 0.38}, {0.06, 0.44}]

Sorting gives:

[{0.14, 0.28}, {0.06, 0.32}, {0.42, 0.34}, {0.18, 0.38}, {0.14, 0.40}, {0.06, 0.44}]

Next, we condense back into the original value increments by prorating the probabilities:

For example, the set {0.14, 0.28} is prorated between the values of 0.2 (20%) and 0.3 (80%) -> {0.028, 0.2}, {0.112, 0.3}

Prorating all the sets and adding the probability contributions to each discrete value increment gives the condensed result:

combined distribution = [{0.028, 0.2}, {0.448, 0.3}, {0.5, 0.4}, {0.024, 0.5}]

This process is repeated for each criterion within a goal to develop a figure of merit for the goal, expressed as a distribution. With default (equal) weights for the goals, or alternatively with sets or relative goal weights, figures of merit can be obtained for each goal area (Sustainability, Safety and Reliability and Economics), or a single figure of merit for each system.

The scores obtained in this manner (either the global score or the goal-level scores) characterize both the systems potential (top 25% of the distribution) and provide an indication of the degree of uncertainty (spread of the distribution).

APPENDIX D

Control of the Expert Elicitation Process

Appendix D

Control of the Expert Elicitation Process

FRAMEWORK OF THE EVALUATION QUANTIFICATION PROCESS

The Methodology Document, "Generation IV Roadmap Evaluation of Generation IV Concept Proposals, Final Screen and R&D Prioritization: Methods," guides the TWGs to collect information about the concept proposals and to evaluate each concept set (nuclear energy system) against the Generation IV goal criteria using the metrics provided in Appendix A. This evaluation requires a probability-based approach in which the TWG must assess a probability distribution to represent the potential and uncertainty of a concept with respect to each criterion.

Each probability distribution is intended to be a subjective statement of the TWG's understanding of the consensus of the technical community on the likelihood of how the concept will perform compared to a reference Generation III system. In evaluating their uncertainty, the TWG should consider both randomness (aleatory uncertainty) and gaps in the state-of-knowledge (epistemic uncertainty). Closing the gaps in our state-of-knowledge will be the goal of future research.

The use of probability as the language of uncertainty, as a statement of our state-of-knowledge, has a sometimes controversial history. A brief summary of the idea is taken from (Kaplan 1981):

> *People have been arguing about the meaning of probability for more than 200 years. The major polarization is between the "objectivist" or "frequentist" school that views probability as something external, the result of repetitive experiments, and the "subjectivists" who view probability as an expression of an internal state of knowledge. There is no need to take sides here, because both sides are right -- they are just talking about different ideas. And both ideas are useful. Unfortunately both use the same word, probability, for these two ideas. Over the years, there have been many suggestions for clarifying the language, using different words for each. In this note, what the objectivists have been talking about will be called "frequency." What the subjectivists have been talking about will be called "probability." Thus, "probability," as used here, is a numerical measure of a state of knowledge, a degree of belief. "Frequency," on the other hand, refers to the outcome of an experiment of some kind involving repeated trials. Thus "frequency" is a "hard" measurable number. This is so even if the experiment is only a thought experiment or an experiment to be done in the future. At least in concept then, a frequency is a well-defined, objective, measurable number.*
>
> *Probability is a notion of a different kind. Defined as a number used to communicate a state of mind, it is thus inherently subjective and changeable as new information arrives. To make this notion useful, we must clearly define the correlation between the numbers and the state of mind.*
>
> *This can be done in several ways. The most direct is to use frequency in the following way. Suppose we have a lottery basket containing coupons numbered from 1 to 1,000. Suppose the basket is thoroughly mixed and that you are about to draw a coupon blindfolded.*
>
> *We ask: Will you draw a coupon numbered 632 or less? With respect to this question you experience a certain state of confidence. Similarly, I experience a state of confidence with respect to this same question. Let us agree to call this state of confidence, "probability 0.632," equal to the frequency of such draws in an infinitely repeated experiment. Now we both know exactly what we mean by p = 0.632.*
>
> *So if you now say that the probability of your latest business venture succeeding is 0.632, I know exactly what your experiential state of confidence is. We have communicated.*

In the same way, we define or "calibrate" the entire probability scale, from zero to one, using frequency as a standard of reference. Note that the process used here is entirely parallel to the way by which we define "red," "chair," "seventeen," and all the other words or symbols.

This method of definition shows the intimate connection between probability and frequency. This connection needs to be recognized always and at the same time not allowed to obscure the fundamental difference. Frequency is used to calibrate the probability scale. Once the calibration is established, we then use probability to discuss our state of confidence in areas where we are dealing with one-time events and have no frequency information at all.

[This approach allows us to analyze situations beyond those permitted by] *the restrictions of the relative frequency school of thought (i.e., that only mass repetitive phenomena can be analyzed probabilistically) and…create…a systematic, disciplined theory and language for dealing with rare events, for quantifying risks, for making decisions in the face of the uncertainties that are inevitably present.* [And we can take action] *with the knowledge that these are the best decisions and actions possible in the light of all the information available to us.*

This then is the definition of probability adopted in this [note and for the Gen IV evaluation process]*. For additional insight, we quote the following paragraph by A. DeMorgan:*

> We have lower grades of knowledge, which we usually call degrees of belief, but they are really degrees of knowledge…It may seem a strange thing to treat knowledge as a magnitude, in the same manner as length, or weight or surface. This is what all writers do who treat of probability, and what all their readers have done, long before they ever saw a book on the subject…By degree of probability we really mean, or ought to mean, degree of belief…belief is but another name for imperfect knowledge, or it may be, expresses the mind in a state of imperfect knowledge.

and as further elaboration [consider] *the following paragraph from unpublished notes by E.T. Jaynes:*

> Probability theory is an extension of logic, which describes the inductive reasoning of an idealized being who represents degrees of plausibility by real numbers. The numerical value of any probability (A|B) will in general depend not only on A and B, but also on the entire background of other propositions that this being is taking into account. A probability assignment is "subjective" in the sense that it describes a state of knowledge rather than any property of the "real" world; but it is completely "objective" in the sense that it is independent of the personality of the user; two beings faced with the same total background of knowledge must assign the same probabilities.

[This statement, especially the final sentence, is seen by some as controversial. Presumably all would agree that assigning probabilities without full knowledge of key facts would be flawed. Our goal in the elicitation of evidence, rather than opinion, is to bring all participants to a similar level of recognition of "the…total background of knowledge" that could be relevant to the question at hand. Experience in this work indicates that consensus generally comes easily, once all evaluators have access to all the information.]

Corresponding with the…definitions of frequency and probability as numbers, we may say that [classical or frequentist] *statistics, as a subject, is the study of frequency type information. That is, it is the science of handling experimental data. On the other hand, probability as a subject we might say is the science of handling the lack of experimental data.*

Thus, one often hears it said that we cannot use probability distributions because we have insufficient data. In light of our current definitions, we see that this is a misunderstanding. When one has insufficient data, there is nothing else one can do but use probability. [When there are sufficient data, there may or may not be uncertainty. If there is variability in the data, then we have aleatory uncertainty and a frequency or variability distribution. When we do not have sufficient data, we must use a probability distribution.]

If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts, he shall end in certainties.

– Francis Bacon

So, for decision-making purposes, the subjective notion of probability is appropriate and provides the language for dealing with uncertainty. The literature includes many discussions of techniques for the elicitation of uncertainty, including formalisms for facilitating information exchange, ensuring complete sharing of relevant information, reaching agreement on the meaning of a community consensus distribution, and protecting against biases. A recent and thorough exposition relevant to the nuclear industry can be found in (Budnitz 1998) or its source document (Budnitz 1995). Although these documents were developed for evaluating seismic risk, their discussion of elicitation and methods to ensure its reliability are quite general.

CONTROL OF BIAS IN THE EXPERT ELICITATION PROCESS

When a subjective process is adopted for estimating key parameters in a decision process, such as the probability distributions the TWGs must develop, strong controls are needed to prevent bias from distorting the results. The TWGs and the crosscut groups are the first defense against intentional bias. Unintentional bias is more troublesome and must also be addressed. Perhaps the best approach is to thoroughly understand how unintended bias can occur. With that knowledge, the TWGs can guard against its influence in their deliberations.

A number of studies present substantial evidence that people [both naïve judges and subject matter (domain) experts] are not naturally good at estimating probability (including uncertainty in the form of probability distributions or variance). [(Hogarth 1975), (Tversky 1974)] For example, Hogarth notes that psychologists conclude that man has only limited information processing capacity. This in turn infers that his perception of information is selective, that he must apply heuristics and cognitive simplification mechanisms, and that he processes information in sequential fashion. These characteristics in turn lead to a number of problems in assessing subjective probability, such as:

- Often ignore uncertainty (a simplification mechanism; uncertainty is uncomfortable and complicating, and beyond most people's training)

- Lack concept of the impact of sample size on uncertainty

- Lack of understanding of independence (prefer to balance a sequence of events; alternating sequences appear to be "more random)

- A need to structure the situation leads to imagining patterns, even when there are none

- Fairly accurate at judging central tendency, especially the mode

- Estimates of variability in data are influenced by the mean; they more often estimate the coefficient of variation (standard deviation divided by the mean) than the variance.

- People significantly underestimate the range of uncertainty; e.g., in half the cases people's estimates of the 98% intervals fail to include the true values

Lest we agree prematurely that people are irretrievably poor at this evaluation task, it is significant to realize that there are many successful applications. Hogarth himself points out that studies of experienced meteorologists have shown excellent agreement with actual facts. So we need to understand what techniques can help make good assessments.

(Winkler 1968) make a useful distinction between two kinds of expertise or "goodness." "Substantive" expertise refers to knowledge of the about subject matter of concern. "Normative" expertise is the ability to express opinions in probabilistic form. Hogarth points out that the subjects in most of the studies were neither substantive nor normative experts. A number of studies have shown that normative experts can generate appropriate probability distributions, but that substantive experts require significant training and experience to do well.

Our purpose here is to understand how these biases occur and to use that information to combat its influence.

Biases and Heuristics. Another view of these problems can be seen in the following discussion of psychological difficulties in elicitation (Tversky 1974), (Bodily 1976):

- *Inadequacies of individuals*. Societal influence, sampling distributions, sequential information processing, anticipations and emotions

- *Inadequacies of the process*. Response mode matters (order of information and consequences of outcomes), type of feedback may affect response

- *Meaningfullness*. To the assessor, better than other statistical methods (models, simulations, experiment), better than seat of the pants?

- *Representiveness*. Insensitive to prior information and sample size, misconceptions of chance, insensitive to predictability, illusions of validity, misconceptions of regression

- *Availability*. Retrievability, effectiveness of a search set, bias of imaginability

- *Anchoring and Adjustment*. Hard to change existing (first) estimates, biases (conjunctive and disjunctive)

While some of these are self-evident, others require a bit of explanation. People tend to rely on a number of heuristics to simplify the process of assessing probability distributions. Some of these introduce bias into the assessment process in ways that can be difficult to overcome. In particular, the three heuristics of representativeness, availability, and anchoring and adjustment will be examined.

In using the *representativeness* heuristic, people assess probabilities by the degree to which A is representative of B. A simple example presented by Tversky and Kahneman can illustrate how this approach can cause serious errors. If we describe traits of an individual as a "meek and tidy" man and then ask if we think he is a "farmer, salesman, airline pilot, librarian, or physician" naïve assessors often give the highest probability to librarian, because he fits the stereotype.

Representativeness also ignores the prior probability. Clearly the prior should have an impact on the posterior probability, but basing our judgment on similarity alone ignores that point. Representativeness is also insensitive to sample size and many of the experimental subjects give the same answer, regardless of sample size. Other failings of representativeness include:

- Misconception of chance – many people expect that the general characteristics of a sequence will be represented in each of its parts, no matter how small

- Insensitivity to predictability – descriptions not relevant to the question at hand can bias results if a match is seen with those characteristics

- Illusion of validity – confidence in the degree of match is confused with probability

- Misconceptions of regression to the mean.

Sometimes people assess the probability of an event by the ease with which instances can be recalled. This *availability* of the information is confused with its occurrence rate. Availability is a useful cue, but is affected by other factors than probability. Several associated biases have been observed:

- Biases due to the retrievability of instances **-** recency, familiarity and salience

- Biases due to the effectiveness of a search set – the mode of search may affect the ability to recall

- Biases of imaginability – the ease of constructing inferences is not always connected with the probability (of course, building models to help with such searches can help

- One of the most common approaches is *anchoring and adjustment*. People start with an initial value and adjust it to account for other factors affecting the analysis. The problem is that it appears to be difficult to make appropriate adjustment. It is easy to imagine being locked to ones initial estimate, but anchoring is much more sinister than that alone. A number of experiments have shown that even when the initial estimates are totally arbitrary, and represented as such to the participants the effect is strong. Two groups are each told we pick a starting point randomly just to have a place to work from. The one given the higher arbitrary starting point generate higher probability.

Given all these problems, is there a way to develop "good" estimates? The following paragraphs give some evidence and guidance.

Who is successful at this process? Many of the above studies ask judges to sketch their probability distributions for general questions such as: How long is the Mississippi River? What is the ratio between the suicide rates in the U.S. and in Japan? When first attempting tests of this kind, (1) we all do very poorly – *most* of the correct answers fall well outside our 10^{th} and 90^{th} percentiles (of course, only 20% should fall outside those limits) and (2) we usually believe that the reason we do so poorly is that they did not limit the questions to our area of expertise. After sufficiently additional experience, it becomes clear that the main problem is not the domain of the questions, but our lack of normative experience. This is supported by studies of substantive experts and normative experts, where only the normative experts – statisticians, domain experts with experience at normative tasks, and domain experts supported by normative expert facilitators – were successful in developing distributions such that they covered the span of the real world answers and covered that span consistent with their proffered distributions.

How can we be successful? If we can understand the heuristics people use to develop subjective probability distributions and the biases that attend those techniques, that awareness can help us avoid the same traps. If we can learn which framings for eliciting distributions cause problems, we can use those that work better. In his comments published with the Hogarth paper, Ward Edwards objects to his fellow psychologists' focus on unaided, untrained judges. He observes that humans use tools in all tasks and there are tools that can help us do a very good job in the elicitation process. (Hogarth 1975).

So we need to develop a tool kit to help in the elicitation process. The first tool is simply an awareness of the problems discussed above. (Tversky 1975) gives many detailed examples useful for developing such familiarity.

Some new methods have proved helpful. They offer a structured, facilitated process. Because the facilitator is familiar with the potential biases, she can test the group's ideas and push them in the right direction. Some of the simplest and best aids include:

- Modeling – construct simple models of the maximum and minimum points of the distribution, avoiding focus on the central tendency until the end points are studied to avoid anchoring; test these models to examine the evidence supporting them rather than relying on opinion alone.

- Seek consensus on the evidence considered by the technical community. (Bley 1992)

- Test distributions against lotteries, e.g., see if the assessor agrees it is equally likely for the real answer to lie between the 25–75%-tiles or outside them. Or between the 40–60%-tiles and outside the 10% and 90%-tiles. Would the assessor be indifferent between the two lotteries: i.e., he sees no difference between a 50-50 bet on $100 and getting $100 if the true answer is above the median and nothing if below.

- Establish a strong facilitator who ensures each participant must individually put his evidence on the table and justify it. (Budnitz 1995)

- Beware of generating overly narrow prior distributions. Overly narrow prior distributions can also arise in other ways than cognitive bias discussed above. A typical example concerns situations where generic industry data are used to generate a distribution for a specific concept. If a great deal of generic data are available, and if concept-to-concept variability is present but not recognized, an extremely and unreasonably peaked prior distribution can result. (Siu 1998)

- Ensure that the evidence used to generate the prior distribution is relevant to the estimation problem. Obvious as this may sound, there are practical situations where this can be a concern. The assessment team should be alert to the validity of their sources of evidence. If there is a concern with the applicability of a particular body of evidence, the TWG should consider generating a distribution that excludes this evidence and then updating the distribution. (Siu 1998)

- Be careful when assessing parameters that are not directly observable. The distribution is supposed to reflect the analyst's evidence concerning a particular parameter. If the analyst has little direct experience with the parameter, it can be difficult to justify an informative prior distribution. (Siu 1998)

- Beware of conservatism. The natural tendency of engineers, when faced with uncertainty, is to employ conservative assumptions. The problem is that the degree of conservatism can vary from evaluator to evaluator, thereby upsetting the ranking of alternatives. More generally, a conservative evaluator injects his/her own values into the analysis, and to some extent usurps the decision-maker's role. (Siu 1998) Require defense of judgments.

References

1. (Bley 1992) Bley, D. C., S. Kaplan, and D. H. Johnson, "The Strengths and Limitations of PSA: Where We Stand," Reliability Engineering and Systems Safety, vol. 38, 1992.

2. (Bodily 1976) Bodily, S., notes from Massachusetts Institute of Technology subject 15.065 Decision Analysis, spring 1976.

3. (Budnitz 1995) Budnitz, R.J., D.M. Boore, G. Apostolakis, L.S. Cluff, K.J. Coppersmith, C.A. Cornell and P.A. Morris, “Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts,” NUREG/CR-6372, U.S. Nuclear Regulatory Commission, Washington, D.C., 1995.

4. (Budnitz 1998) Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Cluff, K.J. Coppersmith, C.A. Cornell and P.A. Morris, “Use of Technical Expert Panels: Applications to Probabilistic Seismic Hazard Analysis,” *Risk Analysis*, Vol. 18, No. 4, pp. 463–469, 1998.

5. (Hogarth 1975) Hogarth, R.M, “Cognitive Process and the Assessment of Subjective Probability Distributions,” *Journal of the American Statistical Association*, Vol. 70, No. 350, pp 271–294, June 1975.

6. (Kaplan 1981) Kaplan, S., G. Apostolakis, B.J. Garrick, D.C. Bley, and K. Woodard, “Methodology for Probabilistic Risk Assessment of Nuclear Power Plants,” Pickard, Lowe and Garrick, Inc., Newport Beach, June 1981.

7. (Siu 1998) Siu, N.O., and D.L. Kelly, “Bayesian Parameter Estimation in Probabilistic Risk Assessment,” *Reliability Engineering and System Safety*, vol. 62, pp 89–116, 1998.

8. (Tversky 1974) Tversky, A., and D. Kahneman, “Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty.” *Science*, Vol. 185, pp 1124–1131, 27 September 1974.

9. (Kahneman 1982) Kahneman, D., P. Slovic, and A. Tversky, ed., *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, 1982.

10. Winkler, R.L., and A.H. Murphy, “’Good’ Probability Assessors,” *Journal of Applied Meteorology*, 7 pp 751–758, October 1968.

APPENDIX E

Summary of Criteria Weighting Factors Within Each Gen IV Goal and Criteria Metrics - Final Screening

APPENDIX E

Summary of Criteria Weighting Factors Within Each Gen IV Goal And Criteria Metrics - Final Screening

CRITERIA WEIGHTING FACTORS

Sustainability-1: Resource utilization

SU1-1	Fuel utilization	1

Sustainability-2: Waste minimization and management

SU2-1	Waste minimization	
	Mass of waste	0.1
	Volume of waste	0.1
	Long-term heat output	0.3
	Long-lived radiotoxicity	0.3
SU2-2	Environmental impact	0.2

The relative importance of the criteria varies with national perspective and system concept priorities. A default of equal weighting for each of the five metrics is adopted at this stage of the evaluation.

Proliferation Resistance and Physical Protection

PR&PP-1	Separated Materials	0.35
	Spent fuel characteristics	0.35
PR&PP-2	Passive resistance to sabotage	0.3

The characteristics of fresh fuel (separated materials metric) and spent fuel are considered equally important; both increase time and difficulty of nation-state diversion and both contribute to increasing the difficulty of subnational theft

Safety and Reliability-1:

SR1-1	Reliability	0.6
SR1-2	Worker safety - routine exposures	0.2
SR1-3	Worker safety - accidents	0.2

Safety and Reliability-2:

SR2-1	Robust safety features	
	Reliable reactivity control	0.2
	Reliable heat removal	0.2
SR2-2	Models with well characterized uncertainty	
	Dominant phenomena models have low uncer.	0.2
	Long fuel thermal response time	0.2
	Integral experiments scalability	0.2

Safety and Reliability-3:

SR3-1	Source term	0.25
	Mechanisms for energy release	0.25
SR3-2	Robust mitigation	
	Long system time constants	0.25
	Long and effective holdup	0.25

Economics-1:

EC1-1 Overnight construction cost

EC1-2 Low production costs

Value function for Economics-1 (EC1) is calculated as a function of EC1-1 and EC1-2 and no criteria weights are necessary.

Economics-2:

EC2-1 Short construction duration

Value function for Economics-2 (EC2) is calculated as a function of EC1-1 and EC2-1 and no criteria weights are necessary.

E1. Value Functions

Value functions are linear and range from -1 to +1. The center of the criteria metric scales corresponds to 0.

E2. Criteria Metric Scales

SU1-1 Use of fuel resources.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>300 Mt U feed/GWyr	250–300 Mt U feed/GWyr	200–250 Mt U feed/GWyr	150–200 Mt U feed/GWyr	100–150 Mt U feed/GWyr	10–100 Mt U feed/GWyr	<10 Mt U feed/GWyr

SU2-1 Waste minimization.
Mass.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>80 MT/GWeYr	40–80 MT/GWeYr	20–40 MT/GWeYr	15–20 MT/GWeYr	10–15 MT/GWeYr	5–10 MT/GWeYr	<5 MT/GWeYr

Volume.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>100 m3/GWeYr	50–100 m3/GWeYr	20–50 m3/GWeYr	15–20 m3/GWeYr	10–15 m3/GWeYr	5–10 m3/GWeYr	<5 m3/GWeYr

Long-term (500 years out of core) decay heat.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better than Reference	Much Better Than Reference
>10 kW/GWeYr	5–10 kW/GWeYr	3–5 kW/GWeYr	1–3 kW/GWeYr	0.5–1 kW/GWeYr	0.1–0.5 kW/GWeYr	<0.1 kW/GWeYr

Long-term (500 years out of core) radiotoxicity.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>3,500 MSv/GWeYr	2,500–3,500 MSv/GWeYr	1,500–2,500 MSv/GWeYr	500–1,500 MSv/GWeYr	100–500 MSv/GWeYr	20–100 MSv/GWeYr	<20 MSv/GWeYr

SU2-2 Environmental impact.

	Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference	
	--	-	Equivalent =	+	++	

PR&PP-1 Minimize diversion or undeclared production.
Avoid weapons grade separated materials.

	Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference		
	HEU, Pu, Np in forms than can be readily separated	HEU, Pu, Np fuels in matrices designed to resist reprocessing	LEU fuels, natural uranium, fuels containing Th; HEU, Pu, Np fuels with intense radiation barriers	DNLEU, Th fuels providing additional barriers to material access/recovery		

Spent fuel characteristics.

	Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference	
	Low burn-up fuels in forms that can be readily separated; any very low burn-up fuel or blanket assemblies	Low burn-up fuel or blanket assemblies	>50,000 MWd/MTHM		Long-lived core integrated with reactor vessel, no onsite spent fuel storage	

PR&PP-2 Reactors have passive features that resist sabotage.

	Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better Than Reference	Much Better Than Reference	
	Substantially easier physical access to vital equipment than an advanced LWR	Somewhat easier physical access to vital equipment than an advanced LWR	Emergency cooling system using safety-grade AC power and an external cooling water source	Passive safety systems that require external control signals for activation	Fully passive safety systems isolated from rapid personnel access, no control activation signals required	

SR1-1 reliability.

	Much Worse Than Reference	Worse Than Reference	Similar to Reference	Better than Reference	Much Better Than Reference	
	Forced outage rate increased by more than a factor of 5 and number or quality of lines of defense degraded	Forced outage rate increased by a factor of 5 or number or quality of lines of defense degraded	Forced outage rate unchanged and lines of defense unchanged	Forced outage rate unchanged and number or quality of lines of defense improved	Forced outage rate reduced by factor of 5 and number or quality of lines of defense improved	

SR1-2 routine exposure.

	Worse Than Reference		Similar to Reference	Better Than Reference		
	Significantly greater risk of routine personnel or public exposure compared to Generation III		Risk of routine personnel or public exposure about the same as Generation III	Significant reduction of risk of routine personnel or public exposure compared to Generation III		

SR1-3 accidental exposure.

	Worse Than Reference		Similar to Reference	Better Than Reference		
	Significantly greater risk of accidental personnel or public exposure compared to Generation III		Risk of accidental personnel or public exposure about the same as Generation III	Significant reduction of risk of accidental personnel or public exposure compared to Generation III		

SR2-1 robust engineered safety features.
Reactivity control.

	Worse Than Reference		Similar to Reference		Better Than Reference	
	Positive temperature or void reactivity coefficient can exist under some operating modes, and/or less than two independent and diverse mechanisms are provided for reactivity shutdown		Power, temperature and void reactivity coefficients are inherently negative, or inherently have no safety-related effects, during normal and anticipated transients and all modes of operation		Inherent design features preclude core damage from reactivity insertion due to inadvertent movement of multiple reactivity control elements, and temperature and void reactivity coefficients are inherently negative	

Heat removal.

	Worse Than Reference		Similar to Reference	Better Than Reference	Much Better Than Reference	
	Decay heat removal system is more complex than current evolutionary LWRs		Decay heat removal system is similar to current evolutionary LWRs	Decay heat removal system uses no AC power	Decay heat removal system always operates and has no moving parts	

SR2-2 System models have small and well-characterized uncertainties.
Phenomena models.

	Worse Than Reference		Similar to Reference		Better Than Reference	
	Some dominant phenomena are difficult to accurately characterize and predict, and must be treated with bounding analysis		Some dominant phenomena can be studied only in scaled separate-effects experiments requiring extrapolation of experimental results		All dominant accident transport phenomena can be studied in well-instrumented separate effects experiments with negligible scale distortions and characterized with well understood probability distributions, for the full range of environmental conditions that may result from an accident or an external hazard	

Long fuel thermal response time.

	Worse Than Reference		Similar to Reference		Better Than Reference	
	Fuel and coolant thermal inertia lower than current evolutionary PWRs		Fuel and coolant thermal inertia has characteristics similar to current evolutionary PWRs		Inherent fuel and coolant thermal inertia provides much longer core thermal response times than evolutionary PWR fuel under design-basis accident transients	

Scalability.

	Worse Than Reference		Similar to Reference	Better Than Reference	Much Better Than Reference	
	Scaling or other constraints generate significant distortions in integral testing		Integral experiments are performed in well-scaled facilities at reduced geometric scale	Integral experiments to study power plant DBAs can be performed at prototypical scale	All safety systems function continuously during normal operation of the power plant, and all dominant safety-related parameters can be monitored continuously during plant operation	

SR3-1 radioactive source term.
Source term.

	Worse Than Reference		Similar to Reference		Better Than Reference	
	Bounding fractional release from degraded core is a significantly greater than LWR fuel		Bounding fractional release from degraded core is similar to LWR fuel		Bounding fractional release from degraded core is a factor of 10 less than LWR fuel	

Energy release.

	Worse Than Reference		Similar to Reference		Better Than Reference	
	Number of significant, complex energy release mechanisms from internal sources is greater by a factor of 2 than a Gen. III LWR		Number of significant, complex energy release mechanisms from internal sources is similar to a Gen. III LWR		Number of significant, complex energy release mechanisms from internal sources is reduced by a factor of 2 from Gen. III LWR	

SR3-2 robust confinement or containment.
Time constants.

	Worse Than Reference		Similar to Reference		Better Than Reference	
	Intrinsic features could permit severe core damage directly following initiating event		Intrinsic features could permit severe core damage one hour after initiating event		Intrinsic features delay severe core damage by over 24 hours following initiating event	

Holdup.

	Worse Than Reference		Similar to Reference		Better Than Reference	
	Release fraction for containment/confinement system is greater than Gen. III LWR by factor of 10		Release fraction for containment/confinement system is comparable to Gen. III LWR		Release fraction for containment/confinement system is less than Gen. III LWR by a factor of 10	

EC1-1 low overnight construction costs.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>$2000/kW	$1800–2000/kW	$1600–1800/kW	$1400–1600/kW	$1200–1400/kW	$1000–1200/kW	<$1000/kW

EC1-2 low production costs.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>$20/MWh	$18–20/MWh	$16–18/MWh	$14–16/MWh	$12–14/MWh	$10–12/MWh	<$10/MWh

EC2-1 short construction duration.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>75 months	65–75 months	55–65 months	45–55 months	35–45 months	25–35 months	<25 months

EC2 low capital at risk.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>$3000M	$2500–3000M	$2000–2500M	$1500–2000M	$1000–1500M	$1000–500M	<$500M

EC1 average cost.

Much Worse Than Reference	Worse Than Reference	Slightly Worse Than Reference	Similar to Reference	Slightly Better Than Reference	Better Than Reference	Much Better Than Reference
>$42/MWh	$42–38/MWh	$38–34/MWh	$34–30/MWh	$30–26/MWh	$26–22/MWh	<$22/MWh