Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira, Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria

doi: 10.3850/ESREL2026061419_esrel26-p26289-cd

Tackling the complexity of PRA Models: measure a Fault Tree

François Sauvaget

EDF R&D, 7 Bd Gaspard Monge, 91120, Palaiseau, France. E-mail: francois.sauvaget@edf.fr

Claudia Picoco

EDF R&D, 7 Bd Gaspard Monge, 91120, Palaiseau, France. E-mail: claudia.picoco@edf.fr

Mohamed Hibti

EDF R&D, 7 Bd Gaspard Monge, 91120, Palaiseau, France. E-mail: mohamed.hibti@edf.fr

Probabilistic Risk Assessment (PRA) models are becoming increasingly complex. This growing complexity is due to the expanding coverage perimeter (level 1, 2 and 3 PRA, hazards, combination of hazards, etc.) as well as the need for enhanced safety and finer levels of detail.
This increased complexity of PRA models makes them challenging to manage, verify, maintain, and quantify. Many stakeholders in PRA domain are working over this issue, exploring potential model simplifications without compromising modelling details. Several publications on this topic can already be found in the literature.
In 2023, we initiated an R&D project to tackle this complexity. This work consists of several steps: defining « model complexity » and its different types, identifying the sources of such complexities and, eventually, addressing them.
The initial task of defining and identifying the different types of complexity have already been presented in previous works. In this study, we focus on computational time.
One potential cause of long computational times might be the complexity of Fault Trees (FTs). To assess this complexity, the computational time should be put in relation with the size and the structure of FTs. But which metrics can be used to characterize and measure it? This paper aims at answering this question. The final goal of the activity being to determine the impact of the FT metrics on the computational time.

Keywords: structure, fault trees, complexity, PRA/PSA.

1. Introduction

Due to the enhanced needs for safety and the resulting increase in regulation, in recent years the nuclear industry has faced growing complexity of systems and procedures. This real increased complexity, ranging from operators' procedures to Probabilistic Risk Assessment (PRA), inevitably impacts models, as models are a representation of reality (though with some important conservatism).

Within PRA, this complexity manifests in larger safety systems (represented by larger event trees and fault trees), increased in number, more extended scope (that since multiple years encompasses the internal and external hazards). This inevitably leads to more and more complex models, making them harder to read, quantify and derive insights from, but also to develop and maintain.

Many works, such as (Pragya et al., 2024), (Ikegawa et al., 2024), (Farag et al., 2023), (Jung, 2024), (Jung, 2009) and (Zhong et al., 2006), can be found in literature concerning the topic of PRA model complexity, indicating the need for addressing this aspect. However, most of these studies focus on a single aspect at a time (quantification algorithms and engines, computational complexity, etc.).

Since 2023, at EDF R&D, research focused on exploring possible ways to reduce PRA model complexities. The approach is twofold: a theoretical approach, the PRA model complexity is determined by using measures coming from complexity theory (Hibti, 2026) and a practical approach, using feedback and insights from PRA analysts.

In this work, the goal is to look at the problem of PRA model at large by not focusing on a single particular aspect but rather aiming at decreasing the global complexity. In other words, this work has aimed to address the problem of PRA model complexity comprehensively.

This implies identifying the sources of complexity (such as modeling choices and computational demands) and then tackling them individually (or together), while ensuring that the overall complexity does not increase.

Focusing on the practical approach to the problem, in (Picoco et al., 2025), the initial steps performed to address the complexity of PRA models and the associated EDF R&D program were described. In this paper, these first steps are summarized, then the next steps executed in 2025 and focused on computational time and fault trees' (FTs) structure are described.

In the follow-up, in fact, the fault tree is considered as the main element of the PRA model, and different features are defined allowing to measure its size and structure. If the computational time is considered as an indicator of complexity[a] (the longer the computational time, the more complex is the FT structure), then this will allow to:

- Determine whether the FTs size are directly responsible for computational time (and/or if one or few particular features are directly responsible for it).
- Provide a mean to compare FTs structures in terms of modeling choices (for example belonging to different PRA models) and ideally determine a structure that is more optimal than another.
- Test different strategies to reduce the FT complexity and measure their efficiency.

The paper is organized as follows: Section 2 briefly presents the R&D program associated to this activity and summarizes the first steps, namely the PRA Model ID card; Section 3, present the work performed on fault trees structure; Section 4 presents the example of an application for the analysis of one fault tree; finally, in Section 5 some conclusions are drawn and next steps are presented.

2. R&D Program on PRA Model Complexity

The problem of PRA models' complexity is broad and not straightforward particularly when one wants to tackle it in the most comprehensive and exhaustive way. Therefore, it cannot be addressed rapidly but it rather requires a multiple years R&D program. The EDF R&D program to tackle the complexity of PRA models includes, as mentioned earlier, a two ways approach from a theoretical and from a practical perspective. The "practical" ways involved the PRA analysts and included:

a. Interviews with PRA analysts from engineering divisions to catalogue the difficulties, the practices and the perceived complexities.
b. Analysis of their feedback and categorization in classes.
c. Search for the source of the complexities and, if possible, propositions of actions to reduce it in the perspective of the development of new models in the future.
d. Applications of the identified actions to multiple models to verify their general character.
e. Perform of the analysis and development of tools (if possible and suitable).

This research program started in EDF R&D in 2023, and it is supposed to hit its first milestone in 2026.

Beyond this activity, the program of tackling complexity in PRA is much broader and impacted training program and tools analysis and use.

[a] Eventually a calculation is always reconducted to the solution of a Master Fault Tree.

2.1.PRA Model Identity Card

(Picoco et al., 2025) presents the results of the first steps, namely steps (a) and (b) of the program.

In 2024, we went interviewing the PRA analysts from the EDF engineering divisions in order to review and track all the complexities and the perceived difficulties of the analysts in their everyday life.

In preparing for these interviews, we came up with the idea of a sort of structure, kind of template, which allowed us (i) to standardize the questions to be asked to the different analysts in a coherent manner and (ii) to facilitate the comparison of models after each had been analysed. This template was namedthe Identity Card (ID) of a PRA Model.

The ID Card of PRA Model covers in one place the different aspects of the model including those that defines their structure, quantification parameters, contributors, etc.

It consists of six tabs, each with several questions as described in detail in (Picoco et al., 2025):

(i) Identity
(ii) Results
(iii) Physiognomy and Structure
(iv) Tools
(v) Others
(vi) Comments

The interview would consist of filling the ID card with the analysts and then let the analysts express himself concerning the difficulties that he/she faces in the everyday life.

A development is ongoing to allow the analyst to automatically generate the PRA Model ID Card, helping the maintenance of the documentation and the process of "Getting started with a model" (which was one of the identified classes of complexity) for newcomers.

2.2. Source of complexities

Upon analyzing the various collected elements, it became evident that they could be categorized into distinct classes, and these classes represented "source of complexities". Once all the "sources of complexities" were identified, each verbatim element from the analysis was categorized into one of these "sources". The identified "sources of complexities" of the PRA model, also called "Class of PRA Model Complexity", are presented in Fig.1.

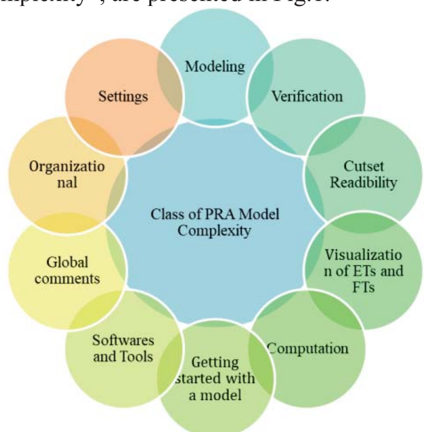

Fig.1. Class of PRA Model Complexity

3. Next step: who's responsible for the model computational time

Once identified the different class of complexity and attributed each element coming from the interviews to the class, some specific actions per class were identified aiming at tackling the complexity at large. To determine the effectiveness of these actions in reducing complexity, a key performance indicator was needed. Computational time was chosen as this measure.

Our initial goal was to pinpoint model features contributing to computational time, and then to identify actions that could reduce it without increasing other sources of complexity. For example, it is known that reducing the number of sequences of an event trees (by smartly aggregating the function events) can reduce the computational time. This action, on one hand, improves the "Computation" complexity but it may worsen the "Verification" and "visualization of ETs and FTs" sources if not correctly traced and documented.

In EDF, the tool used for quantifying the PRA models is RiskSpectrum®. An in-house tool, Andromeda, is also available and used for editing PRA Models.

When a Consequence Analysis Case (CAC) is defined, RiskSpectrum provides with the results some information about the CAC such as: number of gates, number of basic events, etc.

All these information where benchmarked with the computational time for all the CACs defined for the PRA model analyzed. This was done in order to verify if one of these "features" could be identified as directly responsible of the computational time.

An initial finding showed that the way a PRA analyst defines the CACs significantly impacts calculation times. Optimizations in CAC construction have led to substantial reductions for certain CACs that previously required relatively lon computations. This opens possibilities for future work aimed at automating this task for the analyst or, at very least, offering pre-optimized CACs.

The list of elements considered as possible responsible of the computation time is the following:

- Probability
- No. of Minimal Cut Sets
- No. of gates
- No. of basic events
- No. of CCF events
- No. of modules
- No. of basic event MCS saved
- Cutoff error
- MCS gen.time (sec)

This first analysis revealed that none of the considered features in the list hereabove could directly be correlated to the computational time.

It was therefore necessary to identify some other possible measures to investigate to find the responsible of the computational time. It was decided to focus the analysis on FTs and their structures as eventually all the CACs are reduced to the calculation of an MFT.

4. The analysis of the FTs structure

As the first analysis of the elements provided by RiskSpectrum® in the output file was unsuccessful for identifying one direct responsible, the element worth investigating as contributor to the computational time was identified to be the Fault Trees (FTs). The question arises of how to define the structure of a FT and to measure its size. Once defined and computed the size, it would have been possible to benchmark the size of different FTs with their computational time for a Gate Analysis.
The problem of measuring the size of a FTs is not straightforward. In this work, the following definitions to be referred to a given FT were formulated or considered.

Number of Basic Events (BE): all BEs attached to the top gate of the FT recursively down to the bottom level.

Number of House Events (HE): all HEs attached to the FT recursively down to the bottom level.

Base Gates: refers to gates whose inputs are composed exclusively of basic events or HE (no "child" gates).

Number of Gates: this is the number of gates within a given FT starting from the top gate and including the transfer gates. Note that the top gate is not counted.

Vertical Level: vertical position of the gates in the FT with respect to the top gate. Gates immediately below the top gate have the same level (level = 1).

Depth: the highest vertical level of the FT.

Horizontal Level: for each vertical level, this is the number of gates occurrences.

Width: the highest horizontal level of the FT. The width considered is the absolute width, each element is counted every time it appears.

It was considered these definitions to be a good starting point for "measuring" a FT.

4.1. Application to a small FT: example

Let's see the application of the definitions indicated in the previous paragraph to a simple case study, as an example.

Let's considered the FT of Fig.2 (where gates are represented in capital letters and BE and HE are in lowercase).

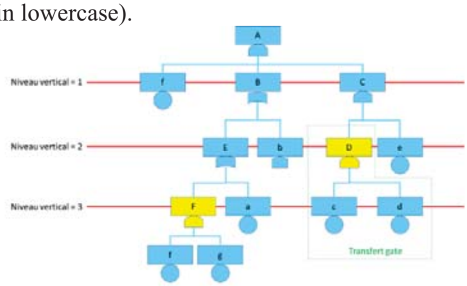

Fig.2. Example of FT

Then, the different features defining the FT are the following.

- Nb of BE = 6 - a, c, d, e, f, g, h
- Nb of HE = 1 – b
- Nb of gates = 5 – B, C, D, E, F
- Vertical levels =
    - B:1
    - C:1
    - E:2
    - D:2
    - F:3
- Depth = 3
- Horizontal level =
    - Level 1: 2
    - Level 2: 2
    - Level 3: 1
- Width = 2

In Fig.2, it is possible to notice the indication of the vertical levels (i.e., *Niveau vertical*) of the different gates and, in yellow, the "base gates".

4.2. Application to a real FT

To evaluate whether the elements defined were useful and easy to apply to a real FT analysis, we decided to isolate a FT representing a system of one of a real industrial model and analyze it accordingly.

Fig.3. The structure of a real FT

Fig.3 shows the structure of the selected FT representing the Auxiliary Feedwater system. In this representation only the gates are shown (each one as a different rectangle with a given ID).

The complete FTs (each transfer gate is itself a FT) has a depth of 16 and a width of 20. Recursively, from the top to very bottom, it presents 49 basic events, 15 house events and 134 "child" gates.

From this representation, it was possible to notice how same transfer gates were treated as different, even though they seemed to be the same.

5. Next steps

The work on FTs presented in this paper was the starting point for tackling the question related to structure of the PRA model, the modeling choices, and the possible "responsible" for the computational time. The next step of this work will consist in:

- Analyze the size of the different FTs into a given model and compare different models,

- Benchmark the FT size vs the computational time of a Gate Analysis in order to determine whether a simplification (to be clarified in which sense: reduction of no. of BE, reduction of no. of gate, etc.) of the FT could lead to a reduction of the Gate Analysis computational time and to the global complexity of the PRA Model.

A development in Andromeda is currently ongoing allowing to operate this measuring analysis of FTs automatically in the entire model. This analysis is supposed to fasten the research on complexity and, over the long term, could be used by analysts as a diagnostic tool for their model. By doing this development, for example, it appeared clear that, the analysis of vertical and horizontal levels is impossible to be performed for the entire PRA model that contains several thousands of FTs. What is foreseeable however is that the PRA analysts can analyze all the FTs of the model in terms of Number of gates, Number of BEs, etc. and then focuses on the bigger ones (and/or the most complex) for a deeper analysis that includes the levels.

6. Conclusions

This paper represents the follow-up of the 2025 one (Picoco et al., 2025) on the EDF R&D activities on PRA Model Complexity.

Tackling the model complexity has become a necessity since the PRA models have become increasingly complex to quantify and maintain.

The goal of the R&D activity is to comprehensively address the problem by identifying the main responsible, with the aim of reducing them. Rather than simply referring to “complexity”, we prefer to talk about different types, or “sources of complexities” as these activities aims at reducing the complexities globally (meaning that addressing one source should not increase another, whether possible).

The first step of the activity was to interview the PRA analysts. This led to the developpment of PRA Model ID Card and the identification of the sources of complexities hereabove mentioned.

In this paper, we presented the how the problem of computational time is being addressed. This involved, firstly analyzing the impact of the distinctive features involved in calculations to determine whether they are responsible for computation time, and secondly, focusing on the structure of FTs and trying to simplify them.

The metrics for characterizing FT structures, for whose definition were given in this paper, will subsequently be applied to a real industrial model, with the aim of reducing its size and analyzing the impact on computational time.

References

Asmaa S. Farag, Egemen M. Aras, Arjun Earthperson, S. Ted Wood, Jordan Boyce, Mihai A. Diaconeasa, (2023) Preliminary Benchmarking of SAPHSOLVE, XFTA, and SCRAM Using Synthetically Generated Fault Trees with Common Cause Failures, Proceedings of the 18th International Probabilistic Safety Assessment and Analysis (PSA 2023), Knoxville, TN, July 15-20, 2023, Pages 40-49

Hibti M., Caleca E., Rifi M., Picoco C., Sauvaget F., (2026) Propositions for Measuring Complexity Probabilistic Safety Assessment Models, 6th World conference on Complex Systems (submitted).

Jung, W. S. (2024). Zero-suppressed ternary decision diagram algorithm for solving noncoherent fault trees in probabilistic safety assessment of nuclear power plants. Nuclear Engineering and Technology, 56(6), 2092–2098. http://dx.doi.org/https://doi.org/10.1016/j.net.2024.01.017

Jung, W. S. (2009). ZBDD algorithm features for an efficient Probabilistic Safety Assessment. Nuclear Engineering and Design, http://dx.doi.org/10.1016/j.nucengdes.2009.05.005

Picoco C., Sauvaget F., Hibti M., Rifi M., (2025), Tackle the complexity of PRA models: what does define it?, ANS PSA 2025, Chicago, June 2025

Pragya Vaishanav, Saran Srikanth Bodda, Abhinav Gupta, (2024) Computationally efficient approach for risk-informed decision making, Progress in Nuclear Energy, Volume 167, 2024, 104983, ISSN 0149-1970

Tomohiko Ikegawa, Naoki Hirokawa, Hiroki Nakamura, Yuki Hirai, Kazunobu Noriyasu, (2024) Improvement of Risk Quantification Procedure for High-Speed Computing, PSAM17 & ASRAM 2024, Sendai, 7-11 october 2024

Zhong, J., Tong, J., & He, X. (2006). An approach to use BDD during the fault tree editing and analyzing. In, Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management, PSAM 2006