



HAL
open science

Horizon Euratom ASSAS project: can machine-learning make fast and accurate severe accident simulators a reality?

Bastien Poubeau, Yann Richet, Lionel Chailan, Fulvio Mascari, Mattia Massone, Simone Gianfelici, Luis Enrique Herranz, Joan Fontanet, Terttaliisa Lind, Christophe D'alessandro, et al.

► To cite this version:

Bastien Poubeau, Yann Richet, Lionel Chailan, Fulvio Mascari, Mattia Massone, et al.. Horizon Euratom ASSAS project: can machine-learning make fast and accurate severe accident simulators a reality?. The 11 th European Review Meeting on Severe Accident Research (ERMSAR2024), Kungliga Tekniska Högskolan, May 2024, Stockholm, Sweden. 10.5445/IR/1000174165 . irsn-04588965

HAL Id: irsn-04588965

<https://irsn.hal.science/irsn-04588965v1>

Submitted on 27 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

HORIZON EURATOM ASSAS PROJECT: CAN MACHINE-LEARNING MAKE FAST AND ACCURATE SEVERE ACCIDENT SIMULATORS A REALITY?

Bastien POUBEAU, Yann RICHET

Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PSN-RES, PSE-ENV/SCAN
F-92262, Fontenay-aux-Roses, France
bastien.poubeau@irsn.fr, yann.richet@irsn.fr

Lionel CHAILAN

Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PSN-RES/SAM
F-13115, Saint-Paul-lez-Durance, France
lionel.chailan@irsn.fr

Fulvio MASCARI, Mattia MASSONE, Simone GIANFELICI

Agenzia Nazionale per le Nuove Tecnologie, l'Energia e lo Sviluppo
Economico Sostenibile (ENEA)
Via Martiri di Monte Sole 4 - 40129 Bologna Italy
fulvio.mascari@enea.it, mattia.massone@enea.it, simone.gianfelici@enea.it

Luis Enrique HERRANZ, Joan FONTANET

Centro de Investigaciones Energeticas, Medioambientales y
Tecnologicas (CIEMAT)
Avenida Complutense 40 - 28040 Madrid Spain
luisen.herranz@ciemat.es, joan.fontanet@ciemat.es

Terttaliisa LIND, Christophe D'ALESSANDRO

Paul Scherrer Institut (PSI)
Forschungstrasse 111 - 5232 Villingen PSI Switzerland
terttaliisa.lind@psi.ch, christophe.dalessandro@psi.ch

Jure BRENCÉ, Ivo KLJENAK, Saso DZEROSKI

Institut Jozef Stefan (JSI)
Jamova 39 - 1000 Ljubljana Slovenia
jure.brence@ijs.si, ivo.kljenak@ijs.si, saso.dzeroski@ijs.si

Fabrizio GABRIELLI, Anastasia STAKHANOVA, Jonas DRESSNER

Karlsruher Institut für Technologie (KIT)
Kaiserstrasse 12 - 76131 Karlsruhe Germany
fabrizio.gabrielli@kit.edu, anastasia.stakhanova@kit.edu, jonas.dressner@kit.edu

Isabel PARRADO RODRIGUEZ

Westinghouse Spain (formerly Tecnatom)
Avenida Montes de Oca 1 - 28703 San Sebastian de los Reyes - Madrid Spain
isabel.parradorodriguez@westinghouse.com

ABSTRACT

Interfacing reference severe accident (SA) codes with commercial nuclear simulators can help spread SA knowledge to a larger audience and build more accurate plant models. However, this requires improving the performance of calculation codes. The Horizon Euratom ASSAS project (Artificial intelligence for the Simulation of Severe AccidentS) addresses this issue and evaluates, among other strategies, the possible added-value of machine-learning (ML) to speed-up calculations.

Efficient programming will be applied to ASTEC (Accident Source Term Evaluation Code, developed by IRSN) [1], with negligible impact on accuracy. This might be sufficient to achieve a real time execution of ASTEC if combined with simplifications of some physical models and of the nodalization used in the input decks. Nevertheless, such modifications are hard to implement, and high acceleration factors are out of reach for complex legacy multiphysics codes like ASTEC.

ML could be a game-changer to reach significantly higher acceleration factors. The first step is the definition of the parts of the SA code that can be replaced by a data-driven surrogate model. Global models, replacing the code integrally, can be more efficient and do not require access to the source code. They will be tested on ASTEC for the ex-vessel phase of the accident. On the other hand, they are expected to be data-greedy since they need to capture a large diversity of physical phenomena. Therefore, ASSAS also explores hybrid approaches, for which surrogate models will replace only selected functions of ASTEC and will be interfaced with other modules.

The specifications of the basic-principles simulator to be developed during ASSAS must be selected at the beginning of the project, to ensure the consistency between the plant model and the training data for ML models. Exercise scenarios have been chosen to determine the systems to be modelled, their scope and the data to be displayed on the Human-Machine Interface.

The present paper aims at providing a description of the ASSAS project and its goals. The adopted strategy and the related challenges will be discussed.

KEYWORDS

Simulator, hybrid machine-learning, ASTEC, nuclear energy, severe accident, training, efficient programming

1. INTRODUCTION

1.1. General objectives of the ASSAS project

Severe accident (SA) calculation codes usually require computer programming skills to launch calculations and visualise the results. The properties of the reactor, of its safety systems, and of its Instrumentation and Control (I&C) must be implemented manually in the input decks. Interfacing SA codes with commercial simulation environments can overcome those limitations. An ergonomic human-machine interface (HMI) makes it possible to interact with the simulation and visualise the results easily, even during the calculation. Simulators usually include computer aided design tools for safety systems and I&C, which generate data for the input deck from graphical representations. Those features can improve the learning curve and efficiency of SA engineers and researchers and limit the risk of errors. Simulators are also powerful tools for education and training, making SA phenomenology more accessible. The use of realistic simulators for accident management training has been recommended for a long time [2], especially since the Fukushima-

Daiichi accidents [3]. Finally, nuclear simulators tend to become a central asset for data centric approaches in the nuclear industry. It can be anticipated that utilities would not develop simulators dedicated to SA modelling, but they would rather extend the capacities of their existing simulators to SA, benefiting from the already existing detailed description of the power plant.

The main objective of the Horizon Euratom Artificial intelligence for the Simulation of Severe AccidentS (ASSAS) project, coordinated by IRSN (France), is to make substantial progress in the deployment of SA simulators. Only few such simulators exist in the world, and they often have limitations [4]. ASSAS plans to develop a proof-of-concept for a SA simulator based on ASTEC, with a level of accuracy close to best estimate models. The prototype will feature a generic basic-principles simulator for a 4-loop Western-type pressurized water reactor (PWR). The HMI will be developed by Westinghouse Spain benefiting from its simulation environment TEAM_SUITE®. It will prove the feasibility to interface ASTEC with industrial simulators. Actions are also planned in the project to prepare the design of VVER (Water-Water Energetic Reactor) simulators in the future. ASSAS focuses on the main scientific and technological challenges, while keeping a simplified description of systems. It is hoped that the prototype will give confidence to industrials to develop more complex simulators, as described in [5].

1.2. Specifications of the simulator

The simulator developed in ASSAS is meant to support educational activities about SAs. Specifications have been defined by project partners to show the different phenomena happening during a SA and the effect of the main mitigation actions. Two sequences have been selected to be implemented in the simulator:

- A 6-inch cold-leg Loss-of-Coolant Accident (LOCA) with Safety Injection (SI) failure and Containment Spray System (CSS) temporary failure,
- A Station Black-Out (SBO) with Steam Generator (SG) Auxiliary FeedWater (AFW) failure.

In both sequences, Portable Equipment Systems (PES) can be used after some delay to inject water into the SGs and the Reactor Coolant System (RCS). The instructor, who manages the failure and recovery of the different systems, will also have the possibility to trigger malfunctions like an induced Steam Generator tube rupture (SGTR).

The systems connected to the reactor will be simplified to focus on the most critical equipment to manage the SA. The Residual Heat Removal System (RHRS), the Chemical and Volume Control System (CVCS), and the balance-of-plant will not be considered for the accident management. Boundary conditions will be used to calculate the steady-state conditions of the reactor before the postulated initiating event. Containment isolation is also assumed to be successful. The most energetic phenomena of a SA, like steam explosion or direct containment heating, are out of the scope of the simulator. The simulation will end at the failure of the containment, caused by overpressure or basement erosion.

Different screens will be accessible from the HMI, for example the SA dashboard partially reproduced in Figure 1. The main sensors and alarms of the Main Control Room (MCR) will be displayed. Additional information will be provided to show the evolution of the SA, including the degradation of the core, the corium-concrete interaction, the pressurization of the containment (including combustible gases), and the release of fission products. It is also planned to develop a Virtual Reality (VR) display with an interactive three-dimensional representation of the plant. The trainee and the instructor will be able to manage the operator actions and interact with the simulation (freezing the simulation, loading new initial conditions, plotting variables, *etc.*) directly from the HMI. A plug-in has also been developed to interact with the simulator directly from a spreadsheet.

The coupling of ASTEC with TEAM_SUITE® is well advanced. The development of the simulator interface and the corresponding physical description in ASTEC and TEAM_SUITE® should be completed before the end of 2024.

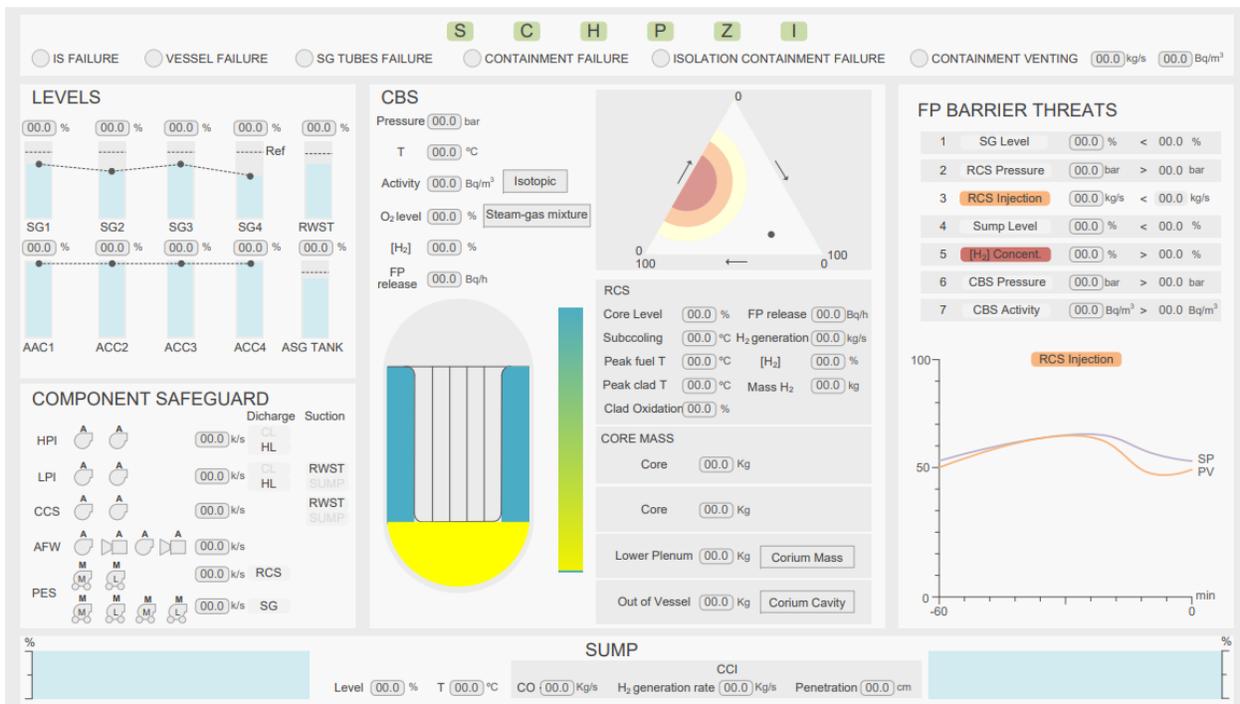


Figure 1: Draft SA screen of the basic-principles simulator (partial view of the SA display)

2. CLASSICAL COMPUTATIONAL OPTIMISATION OF ASTEC

In its current implementation, ASTEC is composed of several modules that are called sequentially and exchange data at a meeting point in time called the macro timestep. The main modules are:

- CESAR, which computes thermal-hydraulics in the vessel, primary and secondary circuits;
- ICARE, which calculates the core degradation in the vessel and the release of fission products;
- SOPHAEROS, which models the transport and speciation of fission products in the primary circuit and the containment;
- MEDICIS, which models the corium-concrete interaction in the reactor cavity;
- CPA, which computes thermal-hydraulics in the containment.

The different modules read and write data in a central database managed by the ODESSA tool. Most physical models need several internal timesteps to calculate the updated values of variables at the end of the macro timestep. In addition, algorithmic considerations impose CESAR's timestep to be a fraction of ICARE's timestep, which is in turn a fraction of the macro timestep. Therefore, three levels of timesteps exist in ASTEC. They are managed dynamically, which means that they will automatically be decreased or increased depending on the convergence rate of the solver.

2.1. Identification of the most time-consuming operations in ASTEC

Identifying the most computationally demanding functions in ASTEC is a prerequisite to launch any action to reduce the execution time of the code, either with classical optimisation, or with ML models. Profiling tools have been used to measure the CPU (central processor unit) time consumed by the different modules, routines, and subroutines of ASTEC, for the different phases of the accident. New data are generated periodically to measure the effects of source code or input deck modifications.

The most time-consuming module of ASTEC is CESAR, followed by ICARE and SOPHAEROS. The relative contribution of the different modules depends on the sequence and the nodalization of the plant. For example, SOPHAEROS becomes the first contributor to the CPU time for a SBO sequence calculation with a simplified input deck (see sub-section 2.3 for more details). CESAR solves a large system of non-linear partial differential equations, which accounts for most of its computation time. On the contrary, ICARE is composed of several different models that have a more balanced time consumption. Hence, optimising or replacing only one of those models cannot speed-up the code drastically.

2.2. Efficient programming

The objective of efficient programming is to improve the performance of ASTEC with a negligible impact on the results accuracy.

Different optimisations have been implemented in ASTEC, enabling a significant reduction in CPU time. Low-level optimisations have been carried out to accelerate memory access and variable search functions. Different post-treatment operations and database savings have been removed because there are not relevant for the simulator. The MEDICIS module has been fully reengineered to reduce its execution time significantly. Besides, several functions of the code have been parallelised thanks to OpenMP [6]. It has been checked that calculation results are not affected by parallelisation.

Additional actions are planned during the ASSAS project. New optimised linear algebra solvers will be considered. Moreover, the adaptive timestep management in CESAR may be improved. New strategies will also be explored for the initialisation of the Newton-Raphson solver in CESAR, to limit the number of necessary iterations to converge to the solution. Code parallelisation will also be tested more extensively, especially for SOPHAEROS and ICARE.

Moreover, a study has been carried out to identify variables in CESAR that require more iterations of the Newton-Raphson solver to reach convergence criteria. It appears that a few variables are responsible for most of the solver iterations. In some cases, these variables face an undesired behaviour of the governing equations: strong non-linearity, non-derivability, non-monotonousness, coexistence of different solutions, absence of solutions, *etc.* If the maximum number of iterations is reached, ASTEC will restart the calculation of the timestep with a shorter time horizon. Different strategies must be explored to limit the occurrence of such situations, for example thanks to smoothing functions.

If the above-mentioned actions are not sufficient, other strategies will be explored. For example, different modules could be executed in parallel, or the thermal-hydraulics equations may be solved separately in the vessel and the circuits. Optimisations offered by the compiler and currently deactivated may improve performances as well. These approaches may have drawbacks on the accuracy or the reproducibility of results and on its numerical stability, so they must be evaluated carefully.

2.3. Reactor model optimisation

Another option to improve the performances of ASTEC is to simplify the reactor model, while checking that the effects on the calculation results are acceptable.

The first simplifications that have been carried out concern the physical models used by the simulator. A simplified criterion will be used for the vessel failure, which is supposed to occur at the bottom of the lower plenum. It is assumed that the corium will be integrally poured into the reactor pit, leading to a clear transition between the in-vessel and the ex-vessel phases of the accident. Therefore, all calculations in the vessel and the circuits will be stopped after the vessel failure, switching off CESAR and ICARE completely. CO₂ and BHO₂ have also been removed from the list of non-condensable gases in CESAR, because of their

low influence on SA phenomenology. Hence, the boron oxidation model of ICARE, BCOX, has been disabled, with limited impact on accuracy. Other possible improvements include the identification and removal from the model of less significant elements and chemical species released during the degradation.

The other actions concern the simplification of the nodalization used for the reactor design featured by the simulator. The input decks have been evaluated on large-break LOCA (hot leg and cold leg breaches) and SBO sequences, which are representative of the sequences used in the simulator.

The description of the reactor core in ASTEC is two-dimensional and axisymmetric. It is divided into several concentric cylinders, piled up in the vertical axis. The best-estimate input deck is composed of 6 concentric rings and 24 axial nodes vertically. Different nodalizations have been investigated, reducing the number of rings down to 3, and the number of axial nodes down to 12. The study has been combined with two descriptions of the primary and secondary circuits: a best-estimate one with 276 control volumes, and a simplified one with 158 control volumes. Introducing both simplifications increases ASTEC's performances substantially while preserving an acceptable accuracy for a basic principles simulator.

The simplification of the containment has been evaluated in parallel. The number of control volumes has been decreased from 28 to 18, the number of walls from 107 to 62, and the number of junctions between control volumes from 79 to 54, without significant impact on the results.

The final step was to evaluate the performances and accuracy of the input deck combining the simplified circuit and containment descriptions, compared to the best estimate nodalization. The increase of performance is larger than a factor of two for all sequences. However, dispersion in results led to investigate the numerical sensitivity of ASTEC. It has been chosen to evaluate the input decks on 50 runs with non-significant modifications of the volume of the sumps and to compare the dispersion of results. The relative variation of the input parameters was small enough to ensure that variations had no physical meaning and could be attributed to numerical sensitivity. Several conclusions can be drawn from this study. First, results are more dispersed for long-lasting sequences, especially the SBO. Second, the simplified input deck usually yields to a higher dispersion of results: the standard deviation of results can be up to a few times higher. Third, for most figures of merit (e.g. time of vessel rupture or activity released to the environment), the mean value obtained with the simplified description differs from the mean value calculated with the best estimate description by less than a few standard deviations. One noticeable exception concerns the time of vessel rupture: it occurs in average after 10 hours with the reference input deck but is delayed by 40 minutes with the simplified one. The long-term efforts spent to reduce the numerical sensitivity of ASTEC need to be pursued. It must also be considered for the development of ML models.

Thanks to the above-mentioned improvements and simplifications, ASTEC is an order of magnitude faster than real time for the ex-vessel phase of the accident. It is in average close to a real-time execution for the in-vessel phase, despite a few strong local slowdowns. Figure 2 depicts the calculation time versus simulated time for the LOCA scenario included in the simulator and compares it to the real time ($X=Y$ curve).

The discrepancies caused by the simplified nodalization have been considered acceptable as regards to the objectives of the project. The simplifications will not only make the simulator faster, but also make the generation of the training database for ML applications easier. However, a general use of simplified input decks requires further investigations.

3. MACHINE-LEARNING APPROACHES

The employment of ML surrogate models aims at reproducing the behaviour of a calculation tool based on a database of pre-computed cases [7]. Instead of solving first principles equations, a direct relationship is established between the inputs and outputs of the code, which can be seen as a high-dimensional regression

task. This can lead to a reduction in calculation time by orders of magnitude, enabling for instance uncertainty propagation [8]. However, the complexity of the problem increases exponentially with the number of degrees of freedom of the calculation tool, a phenomenon called the “curse of dimensionality”. Complex regression functions and large training databases (sets of inputs and outputs) are then required. The use of state-of-the-art deep neural network architectures with large data infrastructures has made it possible to develop advanced surrogate models in physics. One of the most remarkable achievements is related to the development of data-driven weather forecasting models [9], that can outperform classical approaches both in terms of accuracy and calculation time, thanks to the combination of real and simulated data. Deep learning methods proved they could tackle the high dimensionality of the physical description of the atmosphere and the chaotic behaviour of the underlying equations.

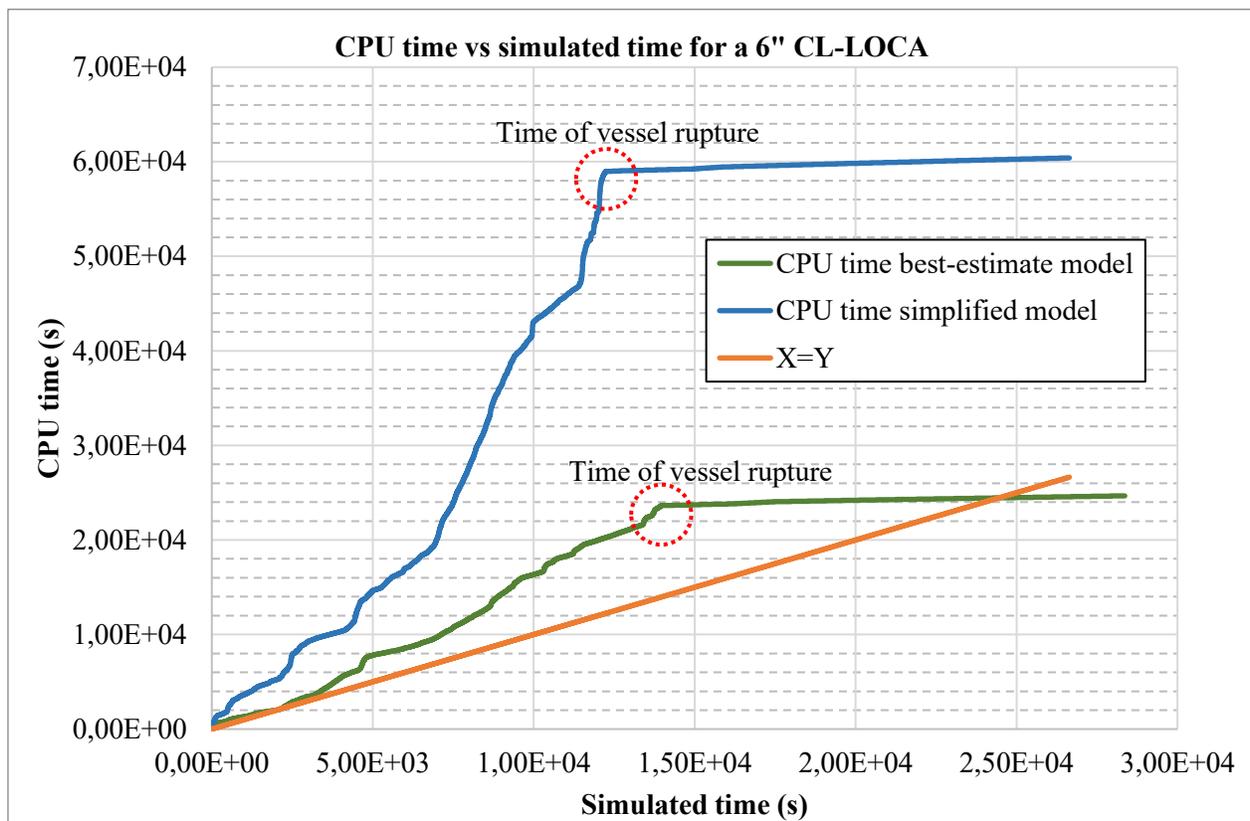


Figure 2: ASTEC CPU time versus simulated time for a 6-inch cold leg LOCA sequence with the best-estimate and simplified input decks

Anticipating the complexity of a ML task is almost impossible without a preliminary analysis on small data samples. This is necessary to evaluate the regularity of the functions to be modelled and their so-called intrinsic dimensionality, which corresponds to the minimal number of dimensions that must be considered to achieve a given accuracy [10]. It has been chosen in ASSAS to explore different surrogate modelling strategies to assess their respective advantages and drawbacks. Physics-informed models will be investigated to ensure that the most important conservation laws (mass and energy balance) are respected, improving the trustworthiness of the models.

To limit the complexity of the surrogate models developed in ASSAS, it has been decided to reduce the number of degrees of freedom of the simulator as much as possible, while preserving its representativeness

and pedagogical function. Hence, only full power initial conditions have been considered. Only a few operator actions will be possible and only best-estimate physical parameters will be considered. Such assumptions are in line with the objectives of a basic-principles simulator, that features a simplified reactor and gives a deterministic result to the trainee.

In addition to the already mentioned numerical noise propagation, numerical and physical oscillations of results can be encountered with ASTEC. This is a challenge for the training of the models. On the other hand, ML models may be able to solve those numerical issues and smooth calculation results.

3.1. Data generation

Quality of the generated data is decisive for the development of surrogate models. Using ML models in a region of the parameter space not sufficiently covered by training data is highly risky and not recommended. Data sampling must be all the denser as the phenomena to model are complex. It must also be representative of the diversity of situations in which the model will be called. Surrogate models will only be trained for a specific SA code and a specific reactor design.

Efforts to build-up the training database will be pooled up: several ASSAS partners will generate training data with ASTEC and upload it to a central database hosted by the Large-Scale Data Facility (LSDF) of KIT. Each partner will be able to retrieve the data it needs for the type of surrogate model it plans to develop. The database will consist of time-dependent raw data for different scenarios, with all ASTEC modules activated. The strategy for time sampling is still an open question. Keeping data for all timesteps is safer to avoid discarding data they may be valuable later in the project. However, data storage costs may be prohibitive if a large variety of scenarios is necessary to ensure training data representativity. In this case, only short series of timesteps will be saved at given time intervals.

In the context of a simulator, the inputs that influence the simulation are the malfunctions and operator actions defined in the training scenarios. The timing of these events will be sampled to generate the training data. A statistical analysis will be performed to evaluate if this approach leads to a balanced sampling of the physical conditions encountered by the reactor. If it is not the case, adaptative sampling methods will be used to generate more data in the under-sampled regions of the parameter space. Additional refinement of the database may be necessary to increase the precision of models for hard-to-predict conditions [11].

3.2. Global models

A first approach consists in replacing ASTEC by a surrogate model completely, at least for some phases of the accident. This black-box approach is the most common in the ML community, since it does not require access to the source code of the calculation tool to be replaced. In the context of ASSAS, it would mean that the simulator would interact directly with the surrogate model, while ASTEC would be switched off. This would make integration easier. In addition, a global model would be faster to execute. On the other hand, this approach is expected to be more data-greedy, because of the possibly higher dimensionality of the problem.

JSI will test the possibility to replace ASTEC by a global surrogate model for the ex-vessel phase of the accident. Two models will be developed respectively for the reactor design featured by the simulator and for a generic VVER-1000 design. The state of the plant at the time of the vessel rupture will constitute the initial conditions of the model. The outputs of the surrogate models must include at least the variables displayed to the user by the simulator. In the context of a simulator, operator and instructor actions may change the boundary conditions dynamically during the progress of the accident. So, two main approaches can be employed in this context, as described for example in [12].

The first method is time-stepping. The objective is to learn the mathematical operator that computes the next timestep(s) from the previous one(s). The timestep is imposed by the refreshment rate of the simulator. Such approaches usually require less simulations, because each timestep of the training database can be considered as a learning point. However, they are subject to error accumulation and, consequently, instability. Using several timesteps in the past and predicting several timesteps in the future, as well as advanced training strategies, may limit such undesired behaviour.

The second methodology is based on time-series prediction. In this case, the complete transient is calculated in advance based on the initial conditions. Such approaches are usually more stable and accurate, but they require more data. Indeed, each sequence is only one training point. The possibility for the operator and the instructor to make an action at any time during the simulation is another challenge. Different approaches can be investigated to address the bifurcations of the simulation at such events. The limited number of possible actions during the ex-vessel phase of the accident is an advantage for this strategy.

3.3. Hybrid approaches

ASSAS will take advantage of the modular structure of ASTEC to explore hybrid ML approaches, for which surrogate models will replace only a part of the code. Different criteria must be considered to choose their appropriate boundaries:

- The contribution of the corresponding function to the CPU time, which gives an upper limit of the achievable speed-up factor;
- The complexity of the surrogate model, depending primarily on the intrinsic dimensionality of the problem;
- The complexity of data generation;
- The complexity of the integration of the surrogate model into ASTEC.

One additional complexity to consider is the coexistence of three levels of timesteps in ASTEC. Until the version 3.1 of ASTEC, data could only be extracted at each macro timestep, making it impossible to develop surrogate models interacting with a module at each of its internal timestep. An additional feature has been added in version 3.1.1, which gives access to the variables generated by CESAR at every internal timesteps, broadening the possibilities for ML approaches.

Three strategies have been selected for the development of surrogate models. Contrary to the global approach mentioned in subsection 3.2, they address the in-vessel phase of the accident only. They are described in the next sub-subsections. In each case, the identification of the inputs and outputs of the models has been a complex task, requiring a deep understanding of ASTEC. It is strongly connected to the functional specifications of the simulator as well because the surrogate models must at least predict the data displayed by the HMI.

3.3.1. CESAR solver initialisation

The most time-consuming operation of the thermal-hydraulics module CESAR is to solve a non-linear system of partial differential equations, thanks to a Newton-Raphson iterative method. Such methods are efficient if they are initialised close to the solution. In the current version of ASTEC, the solver is initialised with the converged solution of the previous timestep. The idea is to use a ML algorithm to initialise the solver with a first guess of the solution, as it has been tested in [13]. This strategy could decrease the number of necessary iterations to find the converged solution, even allowing the use of longer timesteps. One major advantage is that the level of accuracy of calculations is not impacted since the convergence criteria of the solver are not modified. It is also relatively easy to integrate this approach into ASTEC's source code. On the other hand, the execution of the surrogate model must be efficient, as it will be called at each internal timestep of CESAR, so very frequently.

As mentioned in subsection 2.2, specific variables need more iterations of the solver to converge, depending on the phase of the accident. Efforts should focus on them, while keeping simpler models (linear extrapolation or initialisation with the solution of the previous timestep) for other variables. Since the problematic variables change during the accident progression, it could be convenient to develop an ensemble of surrogate models, governed by a classifier that choose which one to apply, depending on the state of the reactor.

3.3.2. Modelling thermal-hydraulics in a portion of the circuits

An alternative strategy to accelerate thermal-hydraulic calculations consists in replacing CESAR completely in a specific zone of the circuits, to be selected cautiously.

Developing a surrogate model for thermal-hydraulics within the vessel has been excluded for computational and physical reasons. First, thermal-hydraulics and core degradation are tightly coupled phenomena. The runaway core oxidation at high temperature is an instable phenomenon because the chemical reaction is accelerated by the heat it releases. It has been anticipated that small prediction errors of the surrogate model would lead to unacceptable discrepancies in results. Second, practical reasons linked to the management of different levels of timesteps made data retrieval and integration of the surrogate model more complex, requiring several adaptations of the source code. It has been estimated that a global surrogate model for the vessel coupling thermal-hydraulics and core degradation was more promising, as described in subsection 3.3.3.

Therefore, it has been decided to concentrate efforts on thermal-hydraulics in the primary and secondary circuits. Contrary to the approach described in sub-subsection 3.3.1, the surrogate model will predict the thermal-hydraulic conditions in the concerned systems only at each macro timestep. A first attempt has been carried out to replace the SGs by a surrogate model, taking advantage of their four-fold symmetry [14]. A data pipeline has been built to integrate the new model in ASTEC:

- First, CESAR calculates thermal-hydraulic conditions in all circuits except the SG,
- Then, the surrogate model calculates variables in the SG,
- Finally, data is combined to reconstruct the flow in all circuits.

The last step is necessary for SOPHAEROS to access the flow conditions in the circuits, which determine the transport of fission products.

The scope of the surrogate models will be progressively complexified, starting with a limited region of the RCS, possibly stretching up to the complete primary and secondary circuits. Autoencoder neural networks have so far proved to be the most appropriate dimension reduction methods, that scale up with the large number of variables in the system and their non-linear behaviour.

3.3.3. Primary vessel modelling

The last surrogate modelling option explored by ASSAS focuses on the global phenomenology in the primary vessel, computing thermal-hydraulics and core degradation simultaneously, to account for their strong coupling. It is hoped that this approach will reduce the dimensionality of the problem. In fact, the vessel (as defined numerically in the ASTEC input decks) is connected to other control volumes only by the 4 cold legs of the primary circuits and the upper plenum. The variables exchanged at these junctions include only thermal-hydraulic conditions and the mass of the different chemical elements released by the core degradation. In ASTEC, the speciation of the fission products is re-calculated by SOPHAEROS at the outlet of the core. Finally, the data displayed by the simulator to represent the core degradation is quite synthetic.

3.4. Validation of machine-learning models

Since they do not compute calculation results from first-principles equations, the validation of ML models requires a different methodology than for physical models. The training data must be split into a training set (to train the models), a validation set (to tune the hyperparameters of the models, like the number of layers for a neural network) and a test set (to evaluate the generalisation capabilities of the models). For hybrid approaches, the interactions between the native physical models of ASTEC and the machine-learning models must be assessed.

The ML community generally relies on simple metrics to train and validate models, like the root-mean-squared (RMS) error. Such error functions may not be suited for SA simulations, which are subject to large uncertainties due to the complexity of the phenomenology and the lack of experimental validation data. It has been decided to complement the RMS metric with expert judgement and the Fast Fourier Transform Based Method (FFT-BM) metric, which has already been used to assess the agreement between experimental results and severe accident code results [15]. Other relevant metrics may be explored as a benchmark. Particular attention will be paid to the good timing of the transitions between the different phases of the SA: core uncovering, start of fission product release, start of molten material relocation, vessel rupture, containment failure, *etc.* The discrepancies between the results provided by the best-estimate calculations and the calculations performed with the simulator (which will include simplified physical models and possibly ML models) will be compared to the usual dispersion of results between different users and calculation codes, as well as the effect of numerical sensitivity.

4. CONCLUSIONS

ASSAS is a 4-year Horizon Euratom project focused on the development of a proof-of-concept SA simulator, which require to increase the performances of ASTEC. Several approaches will be explored during the project to reach a real-time execution of the simulator at any phase of the SA simulation. Improving the algorithmics of the code is appealing because it almost does not affect the physical models. However, the corresponding modifications of the source code are time-consuming. Simplifying the physical models and the plant nodalization is simple and effective but it degrades the accuracy of results and increases numerical sensitivity. Machine-learning may be a game-changer to develop fast surrogate models, replacing SA codes partially or totally. Preliminary tests must be performed to evaluate the complexity of the various options envisaged in ASSAS. The project will help to determine if such approaches can scale-up to industrial simulators, which are used for much more diverse situations than the prototype developed in ASSAS. The efforts associated with data generation and ML models training should not be prohibitive.

Beside the prototype simulator, ASSAS will produce various results for the SA community, like a faster version of ASTEC, more detailed recommendations for the nodalization of reactors, or a higher modularity of the code. The training database generated during the project will be a central asset to explore new data-driven methodologies, like the design of expert systems supporting emergency response, pursuing the work of [16]. It may also interest data scientists to test new neural network architectures on real data.

ACKNOWLEDGMENTS

The authors acknowledge the support provided by the partners and stakeholders of the ASSAS project, in particular Alexei Miassoedov (International Atomic Energy Agency, Vienna, Austria), Alban Martinez-Delcayrou (Framatome, Courbevoie, France), and Francisco Chinesta (Ecole Nationale Supérieure d'Arts et Métiers, Paris, France).



Funded by the
European Union

Funded by the European Union under the grant agreement no. 101059682. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission-Euratom. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

1. Lionel Chailan, Loic Bosland, Laure Carénini, Julien Chambarel, Frederic Cousin, *et al.*, “Overview of ASTEC integral code status and perspectives,” 9th European Review Meeting on Severe Accident Research (ERMSAR2019), Prague, Czech Republic, March 2019, doi: {irsn-04106726}
2. *Application of simulation techniques for accident management training in nuclear power plants*, IAEA TECDOC SERIES, IAEA-TECDOC-1352 (May 2003).
3. European Nuclear Safety Regulators Group, *Compilation of recommendations and suggestions – Peer review of stress tests performed on European nuclear power plants* (26th July 2012).
4. Nuclear Energy Agency Committee on the Safety of Nuclear Installations, *Informing Severe Accident Management Guidance and Actions for Nuclear Power Plants through Analytical Simulation*, NEA/CSNI/R (2017)16 (2nd July 2018).
5. *Classification, Selection and Use of Nuclear Power Plant Simulators for Education and Training*, IAEA TECDOC SERIES, IAEA-TECDOC-1887 (2019).
6. OpenMP Application Programming Interface, Version 5.1 November 2020, www.openmp.org
7. S. Girard, V. Mallet, I. Korsakissok, and A. Mathieu, “Emulation and Sobol’ sensitivity analysis of an atmospheric dispersion model applied to the Fukushima nuclear accident”, *Journal of Geophysical Research: Atmospheres*, **121** (7), pp. 3484– 3496 (2016), doi:10.1002/2015JD023993.
8. Sergey Galushin, Dmitry Grishchenko, Pavel Kudinov, “Implementation of framework for assessment of severe accident management effectiveness in Nordic BWR”, *Reliability Engineering & System Safety* **203** (107049) (2020), doi: 10.1016/j.ress.2020.107049
9. Espeholt, L., Agrawal, S., Sønderby, C. *et al.* “Deep learning for twelve-hour precipitation forecasts”, *Nat Commun* **13** (2022) <https://doi.org/10.1038/s41467-022-32483-x>.
10. Bac J, Mirkes EM, Gorban AN, Tyukin I, Zinovyev, A., “Scikit-Dimension: A Python Package for Intrinsic Dimension Estimation,” *Entropy*, 19th October 2021, **23** (10), pp. 1368, doi: 10.3390/e23101368
11. Fahad Alsayyari, Zoltán Perkó, Danny Lathouwers, Jan Leen Kloosterman, “A nonintrusive reduced order modelling approach using Proper Orthogonal Decomposition and locally adaptive sparse grids”, *Journal of Computational Physics* **399** (2019), <https://doi.org/10.1016/j.jcp.2019.108912>.
12. Lucas Meyer, Louen Pottier, Alejandro Ribes, Bruno Raffin, “Deep Surrogate for Direct Time Fluid Dynamics”, *Proceedings of meeting NeurIPS 2021 - Thirty-fifth Workshop on Machine Learning and the Physical Sciences*, Vancouver, Canada, Dec 2021, pp.1-7 {hal-03451432v1}.
13. Antoine Lechevallier, Sylvain Desroziers, Thibault Faney, Eric Flauraud, Frédéric Nataf. “Hybrid Newton method for the acceleration of well event handling in the simulation of CO2 storage using supervised learning”, preprint (2023), fhal-04085358v3f.
14. Louis Proffit, “Internship report – ASSAS project” (25th July 2023), [https://assas-horizon-
euratom.eu/documents-and-data/](https://assas-horizon-euratom.eu/documents-and-data/)
15. Pietro Maccari, Andrea Bersano, Stefano Ederli, Fabrizio Gabrielli, Fulvio Mascari, “Validation and uncertainty analysis of ASTEC in early degradation phase against QUENCH-06 experiment”, *Nuclear Engineering and Design* **414** (2023), <https://doi.org/10.1016/j.nucengdes.2023.112600>
16. F. Mascari, F. Rocchi, P. Carny, L. Liptak, M. Adorni, J. Fontanet, L.E. Herranz, M. Shawkat, W. Raskob, F. Cousin, J. C. de la Rosa Blul, E. Urbonavicius, F. Di Dedda, T. Augustsson, M. Constantin, G. Arbaev, P. Isaksson, J. Kubicek, “FASTNET Scenarios Database Development and Status”. *Proceedings of the 9th European Review Meeting on Severe Accident Research (ERMSAR2019)*, Praga (Czech Republic), 18-20 March 2019.