OPERA HPC

OPEn HPC theRmomechanical tools for the development of eAtf fuels

Deliverable D6.1 - Numerical and mathematical approaches for computation time reduction

Version 2 – 04/02/2025

Funded by the European Union

Disclaimer

Document type	Deliverable
Document number	D6.1 version 2
Document title	Numerical and mathematical approaches for computation time reduction
Authors	R. Lo Frano (UniPi), S. A. Cancemi (UniPi), L. Giaccardi (NINE), P. Van Uffelen (JRC) S. Gianfelici (ENEA), T. Barani (CEA), I. Ramiere (CEA), D. Pizzocri (POLIMI), R. Largenton (EDF), M. Povilaitis (LEI), F. Feria (CIEMAT), L. Caveglia Curtil (VTT)
Release date	04/02/2025
Contributing partners	CEA, CIEMAT, EDF, ENEA, JRC, LEI, NINE, POLIMI, UNIPI, VTT
Dissemination level	Public

Version	Short description	Main author	PMO	WP leader	Coordinator
1	First release	R. Lo Frano (UniPi), S. A. Cancemi (UniPi) 28/06/2024	Silvia De Grandis (SINTEC) 27/08/2024	R. Lo Frano (UniPi) 30/07/2024	B. Michel (CEA) 19/09/2024
2	Second release: addition to I. Ramiere as co-author	R. Lo Frano (UniPi), S. A. Cancemi (UniPi) 28/06/2024	Silvia De Grandis (SINTEC) 04/02/2025	R. Lo Frano (UniPi) 04/02/2025	B. Michel (CEA) 04/02/2025

Abstract

This work, coordinated by UNIPI, was done with a direct contribution of all the Task 6.1 partners. This document provides first a review and a detailed analysis of existing models for Pellet-Cladding Mechanical Interaction (PCMI), the Fission Gas Behaviour (FGB) and Fuel Overfragmentation (FO) phenomena in industrial type fuel performance codes, such as TRANSURANUS, FINIX and CYRANO3, for which an improvement is expected by the substitution of an empirical formulation by machine-learning, surrogate modelling, or data-driven approaches. In a second part of the document a review is proposed for an identification and a presentation of the most powerful computation time reduction methods. The latter are decomposed in two categories with Machine Learning Methods (MLM) and surrogate models.

In the last part of document some existing preliminary application of computation time reduction method for PCMI and FGB are discussed with more details.

At the end of this review, we can conclude that the integration of MLM and surrogate models will bring significant progress in reducing computation time for the complex simulations needed for the fuel performance studies expected in the WP7 of the OperaHPC project. The implementation of the computation time reduction methods, specified in this document for PCMI, FGB and FO, can now start in the framework of the Tasks 6.2 and 6.3 of the WP6.

Table of contents

List of Acronyms

AI	Artificial Intelligence
ANNs	Artificial Neural Networks
BU	Burnup
CNN	Convolution Neural Networks
DOE	Design Of Experiment
DOF(s)	Degrees of Freedom
DNN	Deep Neural Networks
DR	Digital Replica
EIM	Empirical Interpolation Method
DEIM	Discrete Empirical Interpolation Method
ECSW	Energy-Conserving Sampling and Weighting method
FE	Finite Element
FGB	Fission Gas Behaviour
FFRD	Fuel Fragmentation Relocation and Dispersal
FOM	Full Order Model
FPs	Fission Products
FPCs	Fuel Performance Codes
GB	Grain Boundary
GPR	Gaussian Process Regression
HR	Hyper-Reduction
IO	Internet of Thing
HPC	High Performance Computing
MC	Monte Carlo
MAE	Mean Absolute Error
ML	Machine Learning
MOX	Mixed Oxide fuel
NTFA	Non-Uniform Transformation Field Analysis
OAT	One At Time
PB-ROM	Projection-Based Reduced Order Models
PCE	Polynomial Chaos Expansion
PCI	Pellet-Cladding Interaction
PCMI	Pellet-Cladding Mechanical Interaction

P_{EOL}	End-Of-Life Rod Internal Pressure
PINNs	Physics-Informed Neural Networks
POD	Proper Orthogonal Decomposition
RBF	Radial Basis Function
RID	Reduced Integration Domain
RMSE	Root Mean Squared Error
ROM	Reduced Order Model
RSM	Response Surface Method
RVE	Representative Volume Element
SVD	Singular Value Decomposition
SVM	Support Vector Machines
SVR	Support Vector Regression
TSO	Tangent Second Order
TFA	Transformation Field Analysis
UO2	Uranium Dioxide

List of Figures

List of Table

1. Introduction

The nuclear sector is characterized by complex systems and processes that require detailed and accurate simulations to ensure safety, efficiency, and reliability. These simulations are essential for various aspects of nuclear operations, including reactor core behaviour analysis, thermal-hydraulic assessments, structural integrity evaluations, and safety and risk analyses. Traditional high-fidelity simulations, while accurate, are often computationally intensive and time-consuming, presenting significant challenges, particularly in scenarios requiring real-time decision-making or extensive iterative analyses.

In nuclear power plants, high-fidelity simulations are critical for ensuring that reactors operate safely and efficiently under various conditions. To study the several phenomena, it needs the use of sophisticated mathematical models and computational methods, which, while providing detailed and precise results, demand substantial computational resources and time.

The demand for rapid and reliable simulations is further accentuated by the dynamic nature of nuclear operations. Real-time monitoring and control are essential to respond to transient events, system malfunctions, and emergency scenarios. In such situations, delays in obtaining simulation results can hinder timely decision-making and jeopardize plant safety. Additionally, extensive iterative analyses, such as those required for optimization and design studies, can become impractical with traditional high-fidelity simulations due to the prohibitive computational cost.

Consequently one of the strategic objectives of the OperaHPC project is to improve the numerical capabilities of 3D state-of-the-art fuel performance existing models, for which a significant improvement is expected by the substitution of an empirical formulation by rigorous data-driven approaches, by bringing the simulation of the thermo-mechanical behaviour of fuel rods several steps forward in terms of microstructure description, high performance computing (HPC) capabilities and quantification of uncertainties.

In link with these objectives, machine learning techniques and approaches for reducing computational time are essential to take advantage of the reference results produced by the 3D simulation to advance modelling used in industrial FPC., systematize, and contextualize the models.

The improvement process proposed in the OperaHPC project is based on the sequence described on the Figure 1.

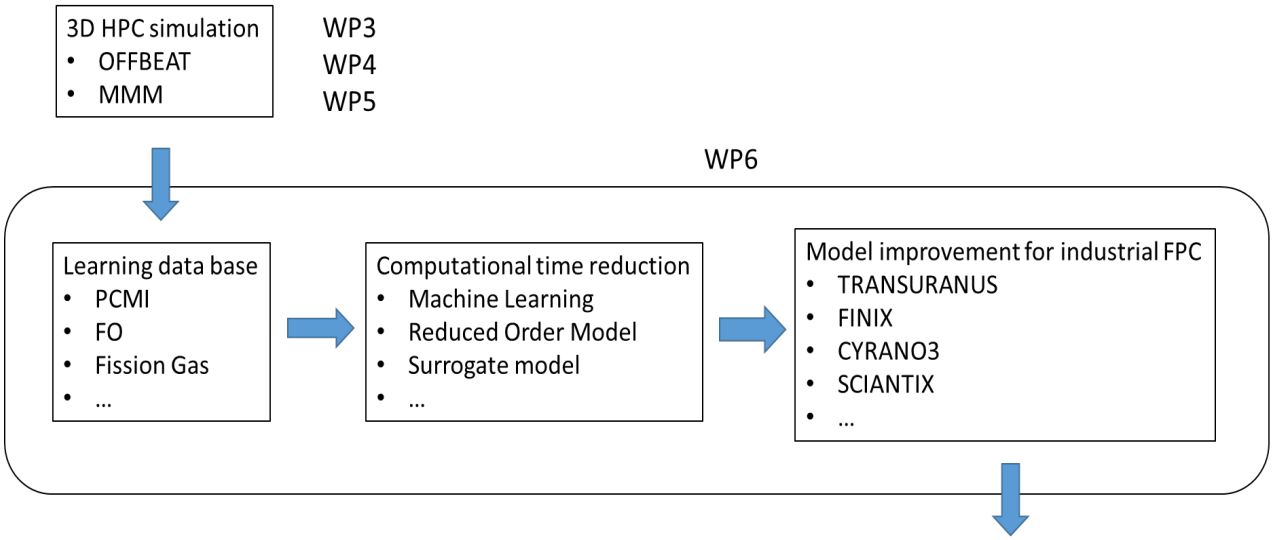

Figure 1: Modelling improvement process in the OperaHPC project

The industrial models targeted by the improvement process in Figure 1 focus on the cladding loading induced by the pellet-cladding mechanical interaction, termed Pellet-Cladding Mechanical Interaction (PCMI), the Fission Gas Behaviour (FGB) and the fuel pellet fragmentation in very small pieces of sub millimetric size, called Fuel Overfragmentation (FO) [1].

This document is decomposed as following :

- section 2 : description of the modelling currently used in industrial type fuel performance codes for PCMI, FGB and FO,
- section 3 : description of the computation time reduction methods based on MLM with the background, taking into account the Proper Orthogonal Decomposition (POD) and Hyper-Reduction (HR), the Non-uniform Transformation Field Analysis (NTFA), the Support Vector Machine approach (SVM) and the Physics-Informed Learning (PINN),
- section 4 : description of the computation time reduction methods based on surrogate models
- section 5 : existing preliminary application of computation time reduction method for PCMI and FGB

2. Modelling currently used for PCMI, FGB and FO

2.1 Pellet Cladding Mechanical Interaction

2.1.1 Phenomenological aspects and driving mechanisms

The PCMI can lead to the cladding failure and, consequently, to the release of gaseous fission products in the primary coolant. The mechanical loadings induced by PCMI are responsible of gap closure, pellet swelling (thermal and gaseous), pellet mechanical behaviour (accommodation of the volume expansion by the thermal creep of the fuel pellet), and stress and strain localization in the cladding (pellet fragmentation and hourglass shape).

PCMI is a relevant phenomenon during the operation of fuel rods in nuclear reactors. Mechanical stresses generated by the contact of fuel pellets, typically cylindrical uranium dioxide (UO2) or mixed oxide (MOX) fuel, and the surrounding cladding during normal and transient irradiation conditions may lead to early failures of fuel pins. Besides the safety implications, this phenomenon can affect the overall performance of the fuel pins in the reactor, limiting its technological burnup (BU) and operational capabilities.

The main factors driving the onset and evolution of the PCMI mechanism can be recognized in:

- thermal expansion of the fuel pellet; the low thermal conductivity of the oxide generates a significant temperature gradient from the centre of the pellet to its periphery. As a result, there will be a non-uniform expansion between the central and the outer regions of the fuel, leading to thermal stresses and variation in the pellet geometry.
- irradiation effects; the interaction of the neutrons with the fuel and cladding matrices alters the microstructure and the properties of both the materials. In particular, in the fuel, there is the additional effect of the fission reactions, with the consequent generation of fission products (FPs) and, generally, thermo-mechanical properties degradation. In their gaseous state, FPs in the grains and their migration towards the grain boundaries contributes to gaseous swelling, and stresses due to internal pressurization. In their solid state, FPs will modify the fuel matrix and cause solid swelling. In addition, densification phenomena will further affect the volume and geometrical variation of the pellet.
- fuel swelling; solid and gaseous swelling, modifying the volume of the pellet, increase the pressure on the cladding after gap closure.
- fuel cracking; due to the strong thermal stresses, the fuel pellet cracks and forms several fragments which, in contact with the claddings, may alter the local stress field.
- pellet deformation; pellet hourglassing may increase local stresses in proximity of the pellet ends.

• cladding deformation; in addition to the thermal and irradiation effects, after gap closure, the stress from the contact with the pellet fragments can lead to creep and additional plastic deformation.

A schematic representation of the pellet hourglassing and the effects of PCMI is shown in Figure 1, with the ridges formed in the clad after the gap closure and clad deformation. The accurate prediction of PCMI effects can enhance the confidence in assessing the fuel pin performances, giving the possibility to adapt the operational strategies to limit adverse effect and so possibly increasing the safety margins. Experimental and modelling approaches are being adopted to better understand the basic mechanisms of the onset and enhancement of the local stresses on the fuel and clad surfaces.

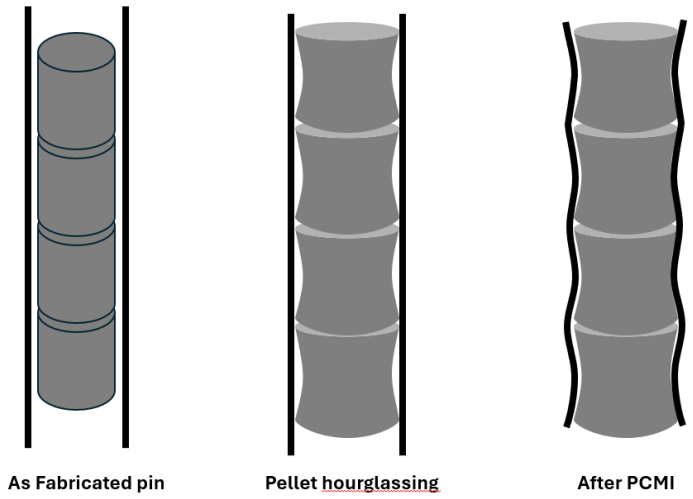

Figure 2: Schematic representation of PCMI

2.1.2 Full field simulation requirements

Considering the Multiphysics nature of the PCMI problem, which involves neutronic and thermos-mechanical effects, and its scale to the pellet-cladding level, simulations are being performed with fuel performance codes implementing specific models aiming to separate and mechanistically assess the concurring phenomena during PCMI. While computational power is made more and more available with time, suggesting wider adoption of full 3D simulations with fine meshes, larger scale computations (at core or reactor level) can still profit from fast and still accurate tools at lower dimensions. 1.5D tools (multi-axial meshes 1D solvers) are still widely adopted to evaluate fuel pin performances, evaluating the thermo-physical properties and the geometrical and mechanical evolution during irradiation for various pellet-cladding configurations.

TRANSURANUS [45] is a computational fuel performance code to simulate the behaviour of fuel rods undergoing operating conditions in nuclear reactors, both for normal operations and transient or accident scenarios. The code includes comprehensive material models for different types of fuel and

cladding materials. These models account for the thermal, mechanical, and chemical properties of materials and their variation over time due to irradiation.

TRANSURANUS performs detailed thermal analysis to calculate temperature distributions within the fuel and cladding and conducts mechanical analysis to predict stresses and strains, thereby assessing the structural integrity of fuel rods. The code also includes models for evaluating the effect of phenomena like swelling, creep, irradiation-induced growth, oxidation and corrosion. The release and transport of fission gases within the fuel, which impact internal rod pressure and thermal conductivity, can be assessed by and intra-, intergranular bubble formation and diffusion mechanistic model or by external codes such as SCIANTIX [46]. Potential failure modes of fuel rods are also estimated, such as cladding rupture or pellet-cladding interaction, under various conditions.

Even considering these features, specific 2D or 3D calculations at a smaller but a more detailed level can provide valuable information in simulating the PCMI evolution. Pellet fragments-cladding interactions may have multi-dimensional effects, with friction and bonding phenomena generating stresses in the cladding. 2D and 3D fine mesh analyses can better assess the impact of pellet hourglassing as well, better reproducing the interaction between the fuel pellets and the cladding.

Among the various available tools, the OFFBEAT code [47] is a fuel performance code capable to model both one-dimensional and multi-dimensional behaviour of fuel rods in normal and off-normal conditions, providing comprehensive analysis of axial and radial behaviour of both fuel and cladding materials.

Due to the intensive computational power and time necessary to perform these simulations, such detailed models cannot be used for entire rods, especially if the results are needed in real time in larger frameworks with other coupled codes.

For this reason, a strategy to build a training and validation database, as mentioned in section 2.1.1, is under development to take advantage of detailed 2D and 3D simulations of PCMI performed with the OFFBEAT code to setup a surrogate model to be implemented in the TRANSURANUS 1.5D code, enhancing its capabilities in simulating PCMI phenomena, while keeping the required computational time low.

One of the main limitations of the Fuel Performance Codes (FPCs) is their inability to perfectly describe the real system they refer to, being a simplified and therefore approximated representation of it. In fact, a fuel performance code is based on a system of equations characterized by several assumptions and hypothesis, which are necessary to perform the analysis in an efficient way considering mainly the computational cost.

1.5D FPCs owe their success to the fast execution time, which is possible thanks to the axisymmetric hypothesis that limits the analysis in the radial direction. This approximation is on one side

fundamental for the efficient analysis of the fuel rods, but on the other side is going to neglect lots of asymmetric phenomena which are typical of both normal and accidental irradiation conditions. One of these phenomena is the fuel cracking and relocation, which influences the thermo-mechanical behaviour of the fuel rod from the beginning of the irradiation until the closure of the gap, i.e. during Pellet Cladding Mechanical Interaction (PCMI).

Fuel cracking and relocation is characterised by strong heterogeneous and 3D evolution over time, which is challengingly represented by 1.5D codes. To reduce the error associated to the 1.5D code approximation, one can think about moving to 3D codes, which are in fact attracting more attention in the last years. The advantage of using a 3D code is immediately clear when considering those non symmetric and heterogenous effects and material properties. At the same time, moving to 3D mesh and PDEs implies a considerably higher computational cost of 3D codes limits their applicability in engineering scale analyses. Still, 3D codes are crucial for reducing the uncertainties associated to 1.5D codes, allowing to generate data which can be exploited in the improvement and refinement of 1.5D code models.

To better understand how this can be done, one should consider how 1.5D code models are typically built. In most of the cases, they are empirical or semi-empirical models fitted on experiments and measured data. Therefore, they are built under specific experimental conditions, data of which are used first for calibrating the model parameters and then for validating the model itself. As a direct consequence, the model is going to perform well in irradiations conditions for which it was calibrated and any extensions to other conditions should be carefully evaluated. The largest the database used to build and validate the model, the less the uncertainties associated to the results. Therefore, there are two possible solutions: to enlarge the experimental database, which is also affected by cost of experimental campaigns and post irradiation examination and data availability, or to find another source of data to exploit for the same purpose. The latter can be exactly a 3D code used to generate big amount of data more efficiently than an experimental campaign might do.

ML techniques or model calibration techniques can be adopted together with experimental data or realistic, i.e. 3D data, to calibrate the 1.5D code model parameters and to reduce the uncertainty associated to code models, following data driven approaches. Chapter 2.1.3 identifies the partial reversibility of relocation during PCMI as one of the models for which this technique may be applied and describes some potentially valuable implementations into the TRANSURANUS code that will be further developed in the framework of Task 6.2.

2.1.3 Modelling Partial Reversibility of Relocation in TRANSURANUS

Under the coordination of JRC, NINE is responsible for the implementation of a data-driven model for partial recovery of fuel fragment relocation during PCMI into the TRANSURANUS code: to understand the reason why it is necessary to have such a model in FPC, a brief description of the physical phenomenon, and how it influences the results is provided in the following paragraphs.

During the first start-up in a reactor, the fuel pellet cracks because of the thermal stress gradient evolving in the oxide fuel characterised by poor thermal conductivity. The resulting pellet fragments deform and move towards the cladding, effectively contributing to the cladding-fuel gap reduction at beginning of life. This is taken into consideration in FPCs by means of a so-called relocation strain, which comes in addition to the other strain components (elastic and non-elastic such as thermal expansion, densification, etc.) as for example suggested by Oguma [48].

Subsequent fission product induced swelling in the cracked pellets and the simultaneous cladding creep down under the influence of the pressure difference between the coolant and the inner gas pressure, causes the pellets fragments to come into contact with the cladding. Beyond this point in time, a soft PCMI is established during which part of the relocation strain can be reversed or recovered, which implicitly accounts for the partial reclosing of the radial cracks. At the same time, other phenomena as the gaseous swelling continue to cause the pellet volume increase, until no more relocation strain can be recovered and a strong or true PCMI occurs.

Different codes have introduced partial reversal of the relocation, sometimes referred to as relocation accommodation. Some examples of different modelling approached are describe in the following list:

- The modified FRACAS mechanical model of the FRAPCON code [49] allows 50% of the original fuel relocation to be recovered due to fuel swelling before hard contact is established between the fuel and the cladding. As a result, the FRACAS model uses the relocated fuel-cladding gap size for the thermal calculations and makes partial use of the fuel surface relocation in the mechanics calculation (i.e., when 50% of the relocation is recovered, the code assumes the pellet to be a rigid structure, and, therefore, hard contact is assumed between the fuel and cladding).
- In FRAPTRAN [50] the relocation model is only dependent on burnup (Heaviside function) and there is no accommodation.
- Zahoor et al introduced fuel relocation recovery in the BISON code recently [51]. For this purpose, they introduced an additional negative strain component that is somewhat equivalent to the densification strain. Therefore, they have on one hand a cumulative relocation strain that stops as soon as pellet-clad contact is established. On the other hand, they introduce an additional (negative) cumulative strain that is initiated when pellet-clad contact is established

until 50% of the maximum absolute value of the relocation strain is reached. After initial contact between the fuel and cladding, if the gap reopens, they assume that relocation is not reinitiated, which is not consistent with the initial consideration of fragment relocation. However, the relocation recovery strain is cumulative and gets activated again after contact is re-established following gap reopening, with the limiting value of 50% of the maximum relocation strain as described above.

- Most recently, Deng et al. [52] introduced relocation recovery in annular fuel (i.e. both inward and outward relocation are considered) modelled by the FROBA code on the basis of finite element modelling. The estimated maximum allowable recovery fraction to be 55.1% in annular fuel.

The current TRANSURANUS code considers the increasing relocation strain as a function of power, burnup and initial gap size by means of different empirical correlations. The partial reversal of relocation is not considered, leading to over-prediction of cladding deformation and stress under certain circumstances, as pointed out by UJV for example [53] Accurate stress and strain predictions are crucial in FPC to evaluate cladding integrity and safety calculations. Therefore, there is the need of developing a new model in the TRANSURANUS code to overcome the results distortions and avoid excessive conservatism and over-estimations for safety analyses.

The modelling approach suggested by JRC and further developed by NINE is to calculate the accommodation strain as a separate non-elastic strain, similarly to the approach adopted in BISON [51].

The new accommodation strain is cumulatively increasing (but negative) until the maximum possible relocation is achieved, and it is defined as a fraction of the relocation strain, through a multiplication factor. This could lead to less oscillations, since both the relocation and accommodation would be cumulative increasing, without oscillating back and forth (because after gap opening, relocation is not supposed to be restored). Rather than introducing a Heaviside function that can lead to convergence problems as well, a smooth function that depends only on the contact pressure has been introduced to calculate the multiplication factor:

$$y = f_1\left(1 - \exp\left(-\left(\frac{x-f_2}{f_3}\right)^{f_4}\right)\right) \qquad (23)$$

Where x is the contact pressure [MPa] and y is the relocation strain accommodation factor. The equation contains 4 model parameters that must be adjusted, i.e.:

- f_1 is the maximum value of the accommodation factor;

- f_2 is the threshold value for the contact pressure above which the cladding starts to push the fragments back;
- f_3 is the value for the contact pressure around which approximately 50% of the maximum accommodation is achieved;
- f_4 is the power factor that determines how rapidly the transition of the accommodation factor is reached;

Like FRAPCON and BISON, as a first tentative, f_1 can be set at 0.5 (i.e. 50% maximum accommodation) and f_2 can be taken consistently with the threshold contact at which gaseous swelling is assumed to stop, i.e. 0.01, although intuitively one would expect a lower value.
What remains to be done is to fit f_3 and f_4 based on a comparison with some selected dataset. Following the approach described in the introduction of Chapter 1, 3D calculations should be considered to generate the data used to calibrate the model parameters.
A database of 3D test cases under different irradiation conditions should be defined and used for fine-tuning of the model parameters with ML or data adjustment methods (e.g. Levenberg Marquardt algorithm, least squares method). After that, another database generated by 3D calculations, or by experimental data should be used for comparison and final validation of the new model for accommodation of relocation.
Further development may consider adopting a different algorithm and calibrate all the four model parameters. The practical development of the improved model, i.e. the application of the methodology to calibrate the model parameters, will be the topic of Task 6.2.

2.2 Fuel Overfragmentation

The FO, associated to a cladding rupture during RIA (and RIA-postDNB) or LOCA, can lead to a loss of integrity of the fuel column and induce the relocation of fuel fragments in fuel rod and dispersal in the coolant (FFRD, Fuel Fragmentation Relocation and Dispersal). The FO is the results of a microstructural damage process that can be induced by over pressurized fission gas bubbles, stress and strain heterogeneities, e.g., at grain boundaries of the polycrystalline material, and hydro-dynamic loading induced in the pellet by the depressurization in case of cladding rupture [1].

The first FO process, related to an over-pressurization of the fission gas bubbles, focuses on a LOCA situation where the fuel experienced a quasi-homogeneous temperature below its fragile-ductile transition. The fragmentation process is linked to crack initiation and growth from over-pressurized bubbles. During the LOCA, the fuel temperature in the external part of the pellet, which can exhibit a High Burnup Structure with submicronic grain size and porosities, can increase beyond

the value experienced during the nominal irradiation. This will lead to an over-pressure loading in fission gas bubbles compare to the equilibrium stage reached during the steady state nominal irradiation. The fragmentation process has to be described at the volume element scale with a set of input data coming from post-irradiation examination or computation with Fuel Performance codes. The main inputs needed are:

- The characteristic of the microstructure and the fission gas inventory at the beginning of the accident (in link with fission gas behaviour as discussed for broader in the previous subsection 2.1).
- The temperature ramp during the accident.

The second FO process targeted for model improvement is the microstructural damage induced by viscoplastic stress and strain heterogeneities under a thermal expansion gradient. The loading conditions correspond to a power transient with a high temperature (> 1400° C) where the fuel experienced thermal creep behaviour. Consistently with fuel studies, expected in WP7, the targeted case is a UO2 fuel with large grain to compare the risk of FO with a standard UO2 microstructure. In a polycrystalline material, the anisotropy of the crystal behavior induced strain incompatibility at grain boundary. Consequently, it results stress heterogeneities potentially responsible for micro-cracks. In this second FO process the first order aspects to consider are: the fuel porosity at grain boundary , the grain size effect in LWRs UO2 enhanced Accident TolerantFuel pellets with large grain and viscoplastic anisotropy induced by dislocation gliding in UO2 fluorite crystalline structure. To describe all the related mechanisms an advanced physical description is expected at the volume element scale with a representative microstructure and a set of input data coming from experimental examinations and computation with Fuel Performance codes. The main inputs needed are:

- The characteristic of the microstructure (geometry and thermo-mechanical behaviour of each elementary material composing the microstructure).
- The total strain loading and corresponding free strain prescribed locally in the pellet during the power transient.

For these two types of FO the learning data base will be built with the MMM code developed in WP4, and the computational time reduction question will be addressed in via dedicated reduced order models (see the review in section 3) maintaining physical description of the system at the macroscopic scale but also at the local scale of heterogeneities.

3. Computation time reduction methods based on MLM: Background

Over the past ten years, there has been a surge in machine learning (ML) research, marked by the development and successful implementation of numerous algorithms across various fields of application. Nonetheless, optimizing the performance of many machine learning algorithms necessitates attention to several factors, including meticulous hyper-parameter tuning and data cleansing, which demand considerable human intervention to yield satisfactory outcomes. Consequently, the implementation of machine learning algorithms calls for in-depth domain knowledge and the expertise of data scientists.

In the nuclear field, the maintenance and safety of nuclear power plants is the key priority. A crucial aspect of ensuring operational efficiency and safety in nuclear facilities is the identification and management of the anomalies that may occur during the normal operation.

The development of digital replicas (DR), which are high-fidelity simulations that mirror physical systems, has provided an unprecedented opportunity for improving predictive safety strategies in various industries, particularly in the nuclear sector. These digital replicas, create a virtual representation of physical components, systems, or entire facilities, allowing for detailed analysis and experimentation without the inherent risks associated with real-world testing.

DR technology offers a chance to test and validate predictive methods without risking real-world assets. By creating a virtual environment that accurately replicates the physical conditions and behaviours of the actual systems, engineers and data scientists can conduct a wide range of tests and simulations to explore different scenarios, identify potential failure points, and develop robust safety protocols. This reduces the need for physical prototypes, which are often costly and time-consuming to produce. This innovative technique, based on advanced AI algorithms, can detect data in a multivariate timeseries that can deviate from nominal patterns, commonly designated as outliers, which can indicate early signs of potential issues, such as material fatigue, pellet-cladding-interaction or unexpected operational anomalies. In particular, the pellet-cladding interaction (PCI) and the pellet-cladding mechanical interaction (PCMI) are still important phenomena to be addressed.

Since the 1960s, fuel failure in water-cooled reactors has been seen as a recurring issue caused by PCI or PCMI.

Fission gas release significantly impacts reactor fuel designs. Gases like Xenon and Krypton, which are low in solubility and generated during fission, can either be retained in the fuel pellet or released to the fuel gas gap. When retained, they contribute to fuel swelling, potentially causing PCMI. Conversely, if these gases are released into the fuel gas gap, they can elevate fuel temperature and potentially cause cladding to crack, due to their relatively low thermal conductivities. These phenomena are mechanistically complicated and hard to model as it is coupled with fuel neutronic

and thermal responses. Many fuel codes focus on simulating PCMI and implement different physical-chemical phenomena. These codes often apply axial symmetry in the domain between the pellet gap and cladding, treating the rod at various heights to analyse different points of thermo-mechanical interaction in 1.5D, 2D and 3D layout. Additionally, it is common to employ coupled code systems, using Finite Element Method (FEM) packages to simulate thermo-mechanical interactions and integrating them with fuel codes [2].

The simulation of PCMI scenarios is extremely important in nuclear safety and risk assessment field. Fuel code and FEM software been proven to be accurate, but it is known to be computationally challenging especially if 3D model is implemented. The ability to identify such deviations in real-time is crucial for preventive maintenance and for avoiding catastrophic failures, thereby enhancing the overall safety and reliability of the system. The DR of a component or system generates synthetic datasets that can be utilized to train ML based models, thus aiding in the timely detection of potential faults in real systems. These synthetic datasets are important because they can be tailored to include a wide variety of operating conditions and potential fault scenarios that may not be easily replicated in real-world environments. This ensures ML based models are well-equipped to recognize and respond to a broad spectrum of issues.

Furthermore, the use of synthetic data generated by DRs allows for extensive training of ML models without the constraints of data availability, which is a significant challenge in the nuclear field. In the nuclear industry, acquiring real-world data for training machine learning models is challenging. Due to the critical nature of nuclear sector, data collection processes are highly regulated to ensure safety and security. This often limits the availability of detailed operational data, especially data related to rare fault conditions or catastrophic events, which are crucial for developing robust predictive models. DRs simulate a comprehensive range of operational scenarios, including normal operations, transient states, and potential fault conditions, without any risk to actual physical systems. This simulated data provides a rich and varied dataset that encompasses the wide array of conditions needed to train machine learning models comprehensively.

In the nuclear field, the rarity of certain failure events makes it nearly impossible to gather sufficient real-world examples for effective ML training. Digital replicas can artificially generate these rare event scenarios in a controlled virtual environment, ensuring that the ML models are exposed to and learn from these critical situations. This capability significantly enhances the predictive power of the models, enabling them to accurately detect and respond to anomalies that they might not encounter frequently in real-world operations. The synthetic data from digital replicas can be meticulously labelled and categorized, providing clear and accurate training data for ML models. In contrast, real-world data often requires extensive preprocessing and cleaning to remove noise, which is both time-

consuming and resource intensive. The use of synthetic data streamlines this process, allowing for more efficient and effective model training.

Moreover, the continuous feedback loop between the physical system and its digital replica enables ongoing refinement and improvement of predictive models. As new data is collected from the physical system, it can be fed back into the digital replica to update and enhance the simulations. This ensures that the reduced order model provides an accurate and up-to-date representation of the physical system, and a reliable platform for ongoing analysis and decision-making.

By leveraging high-fidelity simulations and AI-driven analytics, organizations can achieve a higher level of operational efficiency, safety, and reliability, ultimately contributing to the advancement of their respective industries.

However, one of the major challenges in deploying these advanced ML models, especially in critical fields like nuclear power, is the computational time required. High computational demands can lead to degrade the model's performance, potentially compromising safety and operational efficiency. Therefore, developing methods to reduce computation time without compromising the accuracy and reliability of ML models is of paramount importance.

Several strategies can be employed to address the computation time reduction in ML based models. They include, but are not limited to:

• Model Compression: Techniques such as pruning, quantization, and knowledge distillation can significantly reduce the size and complexity of ML models, leading to faster inference times. Pruning involves removing redundant parameters from the model, while quantization reduces the precision of the numbers used, both of which help in speeding up the computation. Knowledge distillation transfers knowledge from a large model to a smaller one, maintaining performance while reducing complexity.

• Efficient Algorithms: Implementing more efficient algorithms that can perform the same tasks with less computational overhead. For instance, using approximate algorithms or those that can exploit hardware accelerations. Approximate algorithms trade off some accuracy for speed, and hardware accelerations, such as those provided by GPUs, can massively parallelize computations, drastically reducing processing time.

• Parallel and Distributed Computing: Leveraging the power of parallel and distributed computing to break down complex computations into smaller, more manageable tasks that can be processed simultaneously. By distributing tasks across multiple processors or even different machines, the overall computation time can be significantly reduced.

• Hardware Optimization: Utilizing specialized hardware such as GPUs and TPUs that are optimized for ML computations to achieve faster processing times. These processing units are specifically

designed to handle the heavy mathematical computations involved in ML, providing significant speed-ups over general-purpose CPUs.

• Hybrid Approaches: Combining traditional statistical methods with ML to offload some of the computational burdens. For example, using statistical methods to filter out non-critical data before applying ML models. This pre-processing step can reduce the amount of data that needs to be analysed, speeding up the overall process.

• Algorithmic Optimizations: Refining the core algorithms to be more computationally efficient. This can include optimizing the data structures used, improving the convergence rates of iterative methods, and reducing the dimensionality of the input data through feature selection and extraction techniques. Better data structures can lead to more efficient data access and storage, faster convergence rates can reduce the number of iterations needed, and reducing dimensionality can lessen the computational load.

• Real-time Data Processing: Developing methods for real-time data processing that can handle data streams efficiently and update models on-the-fly without the need for extensive retraining. Techniques like online learning, where the model is continuously updated as new data comes in, can be particularly effective in maintaining up-to-date models with minimal computation.

By implementing these methods, it is possible to significantly derive a reduced order model (ROM), based on parametrized solutions and in turn reduce the computational time required for ML-based models' simulation. As described in the following sub-section 3.1, ROM, which are based on parametrized solutions, may allow integrating lower–scale information into an engineering scale without losing the simulation accuracy.

3.1 Proper Orthogonal Decomposition (POD) and Hyper-Reduction (HR)

To accurately represent the behaviour of UO_2 fuel in accidental conditions, and especially fuel over-fragmentation due to over-pressurized bubbles, a detailed thermo-mechanical simulation at the microstructure scale is needed. Indeed, the spatial effects acting at microstructural level cannot be assessed by industrial fuel performance codes. On the other hand, such detailed models are impractical for use in engineering studies, since they entail a computational burden beyond the limits of industrial codes. One way of integrating lower–scale information into an engineering scale is to derive a reduced order model (ROM), based on parametrized solutions obtained by the more comprehensive but computationally–expensive one, to be integrated into the industrial one. The ROM retains an accuracy and a burden in–line with the industrial needs, whilst bringing about a description of the lower–scale phenomena.

In the framework of the OperaHPC project, detailed simulations on a Representative Volume Element (RVE) on FO, due to bubble over-pressurization, are tackled by performing a parametrized set of simulations with the MMM code, developed in Work Package 4, constituting the offline database on which the industrial model will be built. A ROM based on POD and HR is intended to be derived and made available (by CEA) for inclusion into fuel performance codes.
In the following, fundamentals of the reduction order techniques considered are briefly presented.

3.1.1 *A posteriori* Projection-Based Reduced Order Models (PB-ROM)

Two main families of model order reduction methods exist in the literature: *a priori* methods for which the approximation space is constructed on the fly and *a posteriori* methods which use computation results obtained with the full order model (FOM), commonly called snapshots according to Sirovich's terminology [3], to build the reduced approximation space.
A posteriori methods are generally easier to implement in an existing solver. A part of the snapshots (i.e., the solutions of the problem), called the training set, is used to construct the reduced basis (generating the reduced approximation space included in the approximation space of the FOM) denoted $\mathbf{V}$ in the sequel. The remaining part of the snapshots is used for validation (errors, stability, ...) of the ROM and is hence generally called validation set.
For example, for a generic Finite Element (FE) problem of size N given by Equation (1), a reduced basis $\mathbf{V}$ is built, composed of ℓ column vectors. The FOM is reduced by a Galerkin projection onto $\mathbf{V}$ and the ROM of Equation (2) is obtained.

FOM:		**ROM**	
Find $\mathbf{u} \in \mathbb{R}^N$ such that		Find $\boldsymbol{\gamma} \in \mathbb{R}^\ell$ such that	
$\mathbf{K}\mathbf{u} = \mathbf{f}$	(1)	$\mathbf{V}^T\mathbf{K}\mathbf{V}\boldsymbol{\gamma} = \mathbf{V}^T\mathbf{f}$	(2)
with $\mathbf{K} \in \mathbb{R}^{N\times N}$, $\mathbf{f} \in \mathbb{R}^N$.		with $\mathbf{V}^T\mathbf{K}\mathbf{V} \in \mathbb{R}^{\ell\times\ell}$, $\mathbf{V}^T\mathbf{f} \in \mathbb{R}^\ell$.	

The order of the ROM is then ℓ. The ROM is therefore efficient if $\ell \ll N$ and accurate if $\mathbf{u}_{ROM} = \mathbf{V}\boldsymbol{\gamma}$ approximates well the FOM solution $\mathbf{u}$.
The matrix $\widetilde{\mathbf{K}} = \mathbf{V}^T\mathbf{K}\mathbf{V} \in \mathbb{R}^{\ell\times\ell}$ is of smaller dimension than the matrix $\mathbf{K}$, but it is not sparse. Thus, even though ROM preserves the form of the FE problem (cf. $\widetilde{\mathbf{K}}\boldsymbol{\gamma} = \tilde{\mathbf{f}}$, with $\tilde{\mathbf{f}} = \mathbf{V}^T\mathbf{f}$), it requires a solver capable of solving a full matrix system.
A posteriori, ROM methods are generally divided into two parts.

- The first part, called offline part, consists in building the database (snapshots) and the reduced basis. Parameter-independent matrix projections and validation tests are also carried out in the offline part.
- The second part, called online part, is where the parameter dependent operations are performed: remaining matrix projections and the reduced problem solving.

The time required for the offline part can be considerable but is often disregarded because the offline part is only carried out once and for all, unlike the online part which can be executed many times in a parametric study.

There are many techniques for constructing reduced bases, the aim being to obtain good accuracy in approximating solutions while keeping the basis dimension small (for computational efficiency). To achieve this, data compression methods are often used. The most common is the POD-snapshot method [2] because it constructs an optimal reduced basis for the Euclidean norm, i.e. it provides the smallest error in Euclidean norm for a given base dimension.

3.1.1 Proper Orthogonal Decomposition

The POD is a technique to reduce the dimensionality of a problem by operating a change of the basic functions, adopting an orthogonal base that is optimal according to a certain metric (e.g., Frobenius norm). POD has been widely utilized in the simulations of turbulent flows, in the analysis of complex data, and to build ROMs for time-dependent data and parametrized PDEs [2].

The POD method relies on the Singular Value Decomposition (SVD), a particular factorization of a rectangular matrix that can be interpreted as a generalization of the diagonalization of square ones (Quarteroni, Manzoni, & Negri, 2016). The SVD decomposition of the $\mathbf{S_u} \in \mathbb{R}^{N \times Ns}$ matrix writes

$$\mathbf{S_u} = \mathbf{V_u} \mathbf{\Sigma_u} \mathbf{W_u}^T \tag{3}$$

with $\mathbf{V_u} \in \mathbb{R}^{N \times N}$, $\mathbf{W_u} \in \mathbb{R}^{Ns \times Ns}$, the orthogonal matrices of the left (resp. right) singular vectors. The singular value matrix $\mathbf{\Sigma_u} \in \mathbb{R}^{N \times Ns}$ is written as:

$$\mathbf{\Sigma_u} = \begin{pmatrix} \mathbf{D_u} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4}$$

with $\mathbf{D_u} = diag(\sigma_1, \dots, \sigma_d) \in \mathbb{R}^{d \times d}$, such that the singular values are ordered as $\sigma_1 \geq \cdots \geq \sigma_d > 0$. The SVD allows to find the eigenvectors of the matrix to which it is applied, associated with the

singular values ranked in decreasing order. These eigenvectors are representative for a part of the energy of the original matrix.

In the POD method, the $\mathbf{S_u}$ matrix is the so-called snapshot matrix, i.e., a matrix whose columns correspond to solutions of the problem for different values of the parameters $\boldsymbol{\mu}$, see for example the following Equation (5):

$$\mathbf{S_u} = \left[\mathbf{u}(\boldsymbol{\mu}_1) \dots \mathbf{u}(\boldsymbol{\mu}_{N_S})\right] \in \mathbb{R}^{N \times N_S} \tag{5}$$

with N_S the number of snapshots (i.e. problem realizations), and N the number of degrees of freedoms (DoFs) of the problem under consideration.

Applying the SVD on the snapshot matrix, the singular vectors associated to the largest singular values are the best candidates to approximate the matrix. These vectors compose the reduced basis $\mathbf{V}$:

$$\mathbf{V} = \mathbf{V_u}[:, 1:\ell] \tag{6}$$

with ℓ the number of POD modes.

The FOM is projected on the basis $\mathbf{V}$ to obtain the ROM, see Equation (*2*). The number of modes ℓ is either *a priori* given or automatically obtained as the first ℓ modes satisfying the energy criterion defined by the Equation (7):

$$\mathcal{E}(\ell) = \frac{\sum_{i=1}^{\ell} \sigma_i^2}{\sum_{i=1}^{d} \sigma_i^2} > 1 - \varepsilon \tag{7}$$

where ε is a given threshold.

One can demonstrate (e.g., [2] that ε is directly linked to the approximation error (in Euclidean/Frobenius norm) made on the snapshot matrix after reconstruction using to the ℓ first POD modes:

$$\| \mathbf{S_u} - \mathbf{S_u}(l) \|^2 \leq \varepsilon \tag{8}$$

with

$$\mathbf{S_u}(l) = \mathbf{V_u}[:, 1:\ell]\boldsymbol{\Sigma_u}[1:\ell, 1:\ell]\mathbf{W_u}[1:\ell, :]^T \tag{9}$$

A graphical illustration of SVD and POD is given in Figure 1.

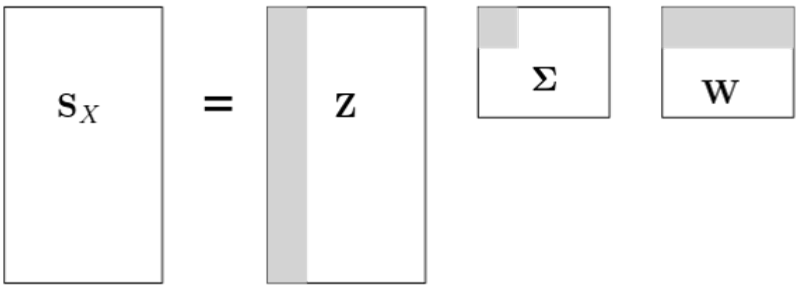

Figure 3: Sketch of the POD method based on a SVD decomposition.

An example to show how SVD and POD effectively work is given in the following.

If we consider e.g., the OperaHPC logo, it can be seen as superposition of three 196x405-pixel matrices (the R, G, and B channels) having integer values ranging between 0 and 255, representative for the intensity of the colour of each channel. We want to perform a POD on the logo and then store the result, which will be a compressed version of the original image, with a degree of approximation that could be chosen *a priori* or determined based on the criterion defined in Equation (7).

The most efficient way (in term of compression) to perform the POD is to consider a unique snapshot matrix defined as the vertical concatenation of the three RGB matrices. Hence the snapshot matrix $\mathbf{S_u}$ is of dimension 588 x 405.

By applying the POD technique explained above on $\mathbf{S_u}$, we obtain the results shown in Figure 3: alongside the original logo, we can see its reconstruction using different levels of accuracy ($\varepsilon = 10^{-2}, 10^{-3}, 5 \cdot 10^{-4}, 10^{-4}$ and 10^{-5}), thus an increasing number of POD modes. One could see how employing roughly only 17% of the initial information (70 modes), the POD procedure allows to represent the initial image with an approximation error inferior to $5 \cdot 10^{-4}$.

Figure 4: Example of the POD technique application to image compression using different error thresholds.

The decreasing of the reconstruction error (cf. Equation (8)) as a function of the POD modes is reported in Figure 4.

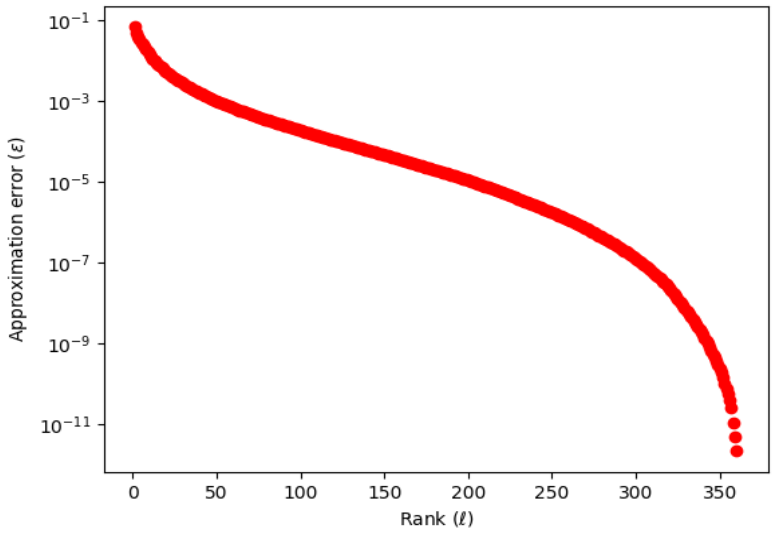

Figure 5: Example of the decrease of the approximation error with respect to the POD modes for the OperaHPC logo.

One can moreover remark that the snapshot matrix was initially not of full column ranked (around 360 over 405). That is due to the smart concatenation of the RBG vectors, which makes it easier to detect redundancies.

3.1.2 Hyper Reduction Methods

HR methods had been initially developed to design efficient ROM for non-linear problems. Indeed, for these problems, matrices must be projected into the online part. Particularly, in the case of mechanical problems with non-linear material behaviour solved through Newton-type methods, the Jacobian matrix must be projected at each iteration of the iterative algorithm.

HR methods make the reduced problem less dependent on the dimensions of the complete FE mesh by using sampling or condensation methods. More generally, these approaches speed up the online projection of non-linear or parameter-dependent terms, since this projection is limited to the terms in the sample.

Among hyper-reduction methods, we can cite the Empirical Interpolation Method (EIM) [4]and its discrete version (DEIM) [5], the Energy-Conserving Sampling and Weighting (ECSW) method [6] or the Hyper-Reduction method based on Reduced Integration Domain (HR-RID) [7].

An example of mesh sampling by the HR-RID method is shown in Figure 6. This method is of interest because it is both low intrusive and has already proved its effectiveness for mechanical problems of interest for the nuclear fuel simulation, implying non-linear materials and/or contact conditions [8][9][10][11].

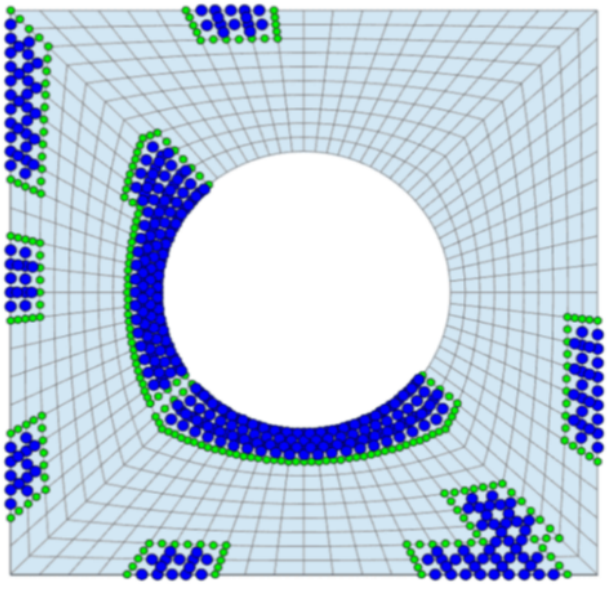

Figure 6: Example of mesh sampling as part of the HR-RID method. Figure from [12] Set of nodes associated to A in blue and to I in green.

In this method, the computational mesh is sampled into a Reduced Integration Domain (RID), i.e., a sub-set of nodes on the whole domain. The RID is constructed by a hybrid automatic/manually

approach, i.e., some nodes correspond to Modal Interpolation Points [5] of POD bases of fields of interest (displacement, contact forces, strain, stress...) while some others are added where it is deemed important (so-called region of interest).

The DoFs associated to the RID are denoted $\mathcal{A}$, while the DoFs linked to $\mathcal{A}$ outside the RID are denoted $\mathcal{I}$ (for interface). Finally, the HR problem associated to the example problem (1) is given by Equation (10):

$$\text{Find } \boldsymbol{\gamma} \in \mathbb{R}^{\ell} \text{ such that}$$

$$\mathbf{V}^T[\mathcal{A},:]\mathbf{K}[\mathcal{A},\mathcal{A}\cup\mathcal{I}]\mathbf{V}[\mathcal{A}\cup\mathcal{I},:]\boldsymbol{\gamma} = \mathbf{V}^T[\mathcal{A},:]\mathbf{f}[\mathcal{A}] \tag{10}$$

$$\text{with } \mathbf{V}^T\mathbf{K}\mathbf{V} \in \mathbb{R}^{\ell\times\ell},\ \mathbf{V}^T\mathbf{f} \in \mathbb{R}^{\ell}$$

Compared to the ROM problem given in Equation (2), the HR problem exhibits reduced terms in matrix vector products since the matrices and vectors implied in Equation (10) are restricted to the RID.

3.2 Non-uniform Transformation Field Analysis (NTFA)

The NFTA which is a reduction technique introduced in the field of Multiscale problems in non-linear solid mechanics to achieve scale transition for materials showing a non-linear behaviour. It is indeed well recognised that the non-linearity introduces a strong coupling between the problems at different scales which, in full rigour, remain coupled. To avoid the computational cost of the scale coupling, reduced models have been developed.

To improve on the predictions of Transformation Field Analysis (TFA) where the plastic strain field is assumed to be uniform in each domain, the reduced model, called the NFTA, assumes the plastic strain fields follow shape functions which are not piecewise uniform.

In the international literature, there are various approaches to develop reduced-order models. As part of the OPERA-HPC project, ROM based on NTFA will be developed and used in Tasks 6.2 and 6.3 of WP6 (see [1] and [13] for more details). Ideally, to simulate in the OPERA-HPC project the phenomena such as the viscoplasticity and grain boundary decohesion in the porous UO_2 polycrystal and the pellet over fragmentation during incidental and accidental transients (see [1] for more details), we would like to perform FEM^2 simulations (square Finite Element Method: [14][15]) in nuclear fuel rod performance code in order to have both the macroscopic behaviour (macroscopic stresses and strains at the first FEM simulation) and the local phenomena: in particular local stresses and strains to deal with the damage and cracking issues at the second FEM resolution. However, this kind of

resolution is time consuming. The second full-field resolution, which is at the polycrystalline scale as illustrated on Figure 3 for UO_2 fuel, can be replaced by a model reduction approach.

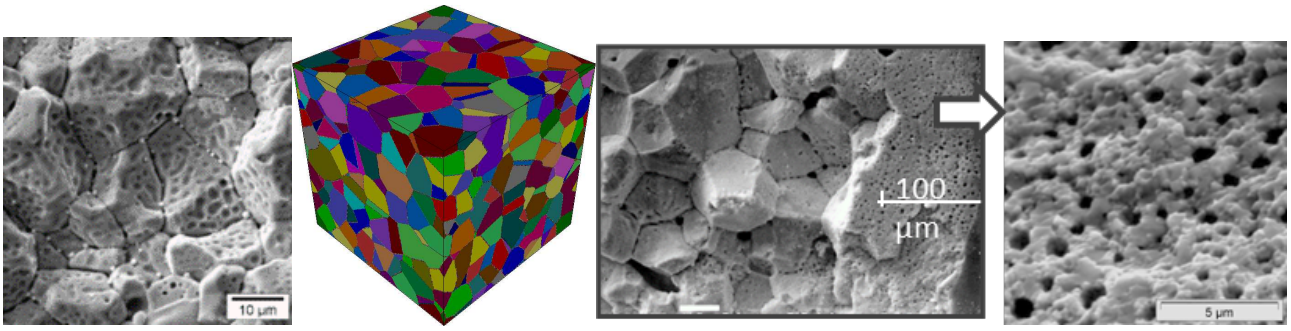

Figure 7: Microstructure of UO_2 fuel pellet center (on the left of the figure) and periphery (on the right of the figure) with standard grain after steady-state irradiation [16]and RVE of the UO_2 polycrystal (in the center of the figure) for simulations at the second FEM resolution [17]

The reduced order models drastically reducing the computational time while keeping a maximum of information at the local level of the RVE.

Before developing ROMs based on the NTFA, in the next section several models developed, referenced, based on the NTFA approach and the history of these developments and progressive improvements are presented.

3.2.1 TFA Approach

Model reduction approaches are used to reduce computational time while allowing access to both macroscopic and local quantities. These approaches generally focus on the second resolution of FEM^2 resolutions, as above mentioned: e.g. the UO_2 fuel polycrystalline scale.

From these approaches we can retain the TFA of [18][19] enables to treat behaviours deriving from two potentials, is to consider the viscoplastic strain tensor piecewise uniform: i.e. to break the RVE into several phases and to consider viscoplastic strain tensor uniform per phase. This idea is taken up more recently in the work of [20].

Although TFA is originally designed to solve micro-mechanical problems (second scale of the FEM^2 resolution or polycrystalline scale), Zhang and Oskay, (2015) retain the general idea of TFA, i.e. a uniform constant field per phase, to deal with problems at higher scale (first resolution of the FEM^2 resolution, i.e. the coarsest scale: e.g. the UO_2 fuel pellet).

Unfortunately for the TFA of [18] and [19], in some problems, the viscoplastic strain tensor is strongly dependent on the spatial variable, which tends to drastically increase the number of internal variables. It has also been seen that the model can be relatively stiff, see [21] and [22].

3.2.2 NTFA Approach

In the WP6 of OPERA-HPC project, the NTFA is retained for several reasons. Compared to phenomenological approaches, such as those described by EDF in milestone 4 (MS4) of WP6 task 6.1 (see [13] for more details), the NTFA approach allows to rely on full-field work, where authors try to produce more physics-based models.

Compared to full field approaches, developed for example in WP2 of OPERA-HPC project (see [23] for more details), it allows to drastically reduce the computational time. Semi-analytical methods (as mean-field approaches), although reducing the computational time, do not allow access to local fields, unlike the NTFA approach (see [24] for more details).

Finally, the NTFA was preferred to the TFA because in the latter approach the internal variables are taken piecewise uniform.

When the heterogeneity is important it is then necessary to subdivide the RVE, which tends to consider a significant number of internal variables.

In WP6 of OPERA-HPC project, we seek to reduce the number of internal variables. Eventually the approach retained can deal with field which varying strongly in the RVE. This model reduction method (NTFA) allows access to local fields as well as to macroscopic quantities. It reduces the computational time by three orders of magnitude. It is also possible, if desired, not to localize and to be satisfied only with the macroscopic response. This approach is presented in the following section 3.2.2.1.

3.2.2.1 The history of NTFA

Within the NTFA approach, it is proposed to decompose the viscoplastic strain tensor on shape functions, called modes, which support some heterogeneity of the RVE. The analysis relies on the principle of superposition to find the strain tensor.

The reduced mechanical problem to be solved on the RVE must be specified. To do so, it is necessary to specify the expression of the evolution of the new reduced variables introduced in the decomposition or, by duality, the expression of the evolution of the reduced forces thermodynamically associated with them. Michel and Suquet, (2003) rely on the effective constitutive relations resulting from the use of the Generalized Standard Materials framework [25] and more particularly of the effective free energy, to explain the expression of the reduced thermodynamic forces.

Then, the expression of the reduced evolution laws is obtained through several approximations: their expression can be retrieved using modified secant linearization of the local dissipation potential (see Appendix B of this reference [25] for more details).

The first step, proposed by the authors of [26], consists in considering the effective force potential equal to the average of the local force potentials over the whole RVE. Here it is a first approximation due to the space-time decomposition specific to the NTFA. This decomposition approximates the viscoplastic strain tensor because it is based on a finite number of modes. This approximation is carried over to the strain and stress tensor. The reduced force evolution relations can be deduced in a systematic way.

To reduce the computational time and to get rid of the average of the potentials over the RVE at each iteration for non-linear behaviour, another approximation is made. The authors of [27] working with two potentials (free energy and force potential) proposed a linearization of the effective force potential around a certain operating point. The operating point is the average of the thermodynamic force variables per phase. This second-order expansion, called here Tangent Second Order (TSO), is inspired by the work on one potential (force potential) of [28] and in its second-order linearization. This development allows to radically reduce the computation time compared to the solely approximation of [26].

Roussette et al. (2009) detail the construction of the modes. In a first step, a certain number of snapshots are taken from full-field simulations previously performed. The snapshots are confounded with the internal variable decomposed at a certain time of the full-field simulation. From these snapshots, a global correlation matrix is constructed.

Finally, the POD is used to deduce from this correlation matrix the modal basis to consider. To do this, the authors of [29] define a criterion, which is like the rates of information restored on the information provided. They use it to consider only a reduced number of eigenvalues resulting from the decomposition, and by extension of modes. In the same vein, the authors of [30] propose to define the modes no longer by phase (local modes) but over the whole RVE (global modes).

Comparisons of results from NTFA models considering local and global modes differ only slightly, both at the macroscopic and local scales. This allows to reduce the number of internal variables and by extension the computational time. To treat any loading, the authors of [24] propose to build a modal basis by applying the POD on elementary loadings and then to perform a second POD to deduce the final modal basis. The results are rather satisfactory in the case of linear viscoelastic behavior. Michel and Suquet [31] generalize the use of global modes to non-linear case.

Three works studied among other, [31] ,[32] and [17] extend the NTFA to polycrystalline cases. In the first two works, the authors of [31], and [17] use all the advances made. The latest work [30] differs from [31] and [17] in the considered linearization. The TSO approach is replaced by a quadrature method to calculate the effective force potential. A parallel can be made with the FOSO method of [28]. It allows to go up in order, unlike the TSO, which is limited to an order 2, at the

expense of the computational time. It is also possible to consider field fluctuations, for example the second moment.

3.3 Support Vector Machine approach (SVM)

Support Vector Machines can be adapted for regression tasks (Support Vector Regression, SVR). SVR finds a function that deviates from the actual observed values by a value no greater than a predefined margin and is as flat as possible. The model uses a subset of the training data (support vectors) to define the decision boundary [57], [58].

- Advantages:
  - Effective in high-dimensional spaces.
  - Robust to overfitting, especially in high-dimensional space.
- Disadvantages:
  - Computationally intensive for large datasets.
  - Requires careful tuning of hyperparameters and kernel selection.

Regression is used in machine learning to learn a mapping between predictor variables and continuous response variables of a given system, based on observed example data. A typical regression model is a function $\hat{f}(\vec{x})$ of a vector of predictors (input features) $\vec{x} = \{x_1, \dots, x_d\}$ of dimension d, that yields an approximation $\hat{y}$ (target or output feature) of the scalar response of the original system $y = f(x)$ [77].

Regression involves a training phase and a prediction phase. During training, $\hat{f}$ is obtained by minimizing a cost function J that depends on $\hat{y}$ and y for a set of input-output pairs (x_i, y_i) of size n, called training dataset. In the prediction phase, the model is used to predict unknown outputs from new inputs [88].

Training is the computationally intensive step that limits scaling with the number of training samples, while prediction is comparatively much cheaper. Since typical training dataset sizes vary between 10^2 and 10^5, training cost is a key factor in choosing the appropriate regression method for a given problem. Prediction cost is relevant to a lesser extent, mainly when implementing the model in engineering codes where runtime performance is important. Model performance is measured on its ability to generalize over novel inputs and uses metrics of distance between model predictions and known outputs, relative to a test set excluded from training, such as mean absolute error (MAE) and root mean squared error (RMSE).

Regression methods have seen widespread adoption in research and industry in the past years, partly due to the availability of open-source implementations for Python and R, such as scikit-learn. In the

nuclear field, for instance, regression models trained on experimental and synthetic data were used to predict decay heat in spent fuel from burnup, cooling time, enrichment and uranium mass [78]. In the context of PCMI modelling, we are going to use regression models to transfer features between 3D and 1.5D simulations.

Two state-of-the-art regression methods that are especially suited to learn non-linear mappings are Kernel Ridge Regression (KRR) and Support Vector Machines (SVMs). Both methods use non-linear kernel functions that can be seen as measures of distance between input features of training points. The two methods have distinct formulations but can considered special cases of a common model based on linear least-squares (LLS) regression [89]. In LLS, the output is obtained as linear combination of inputs through a vector of weights $\vec{w}$ of dimension d:

$$\hat{y} = \vec{w}^T \vec{x} \tag{11}$$

The weights are obtained by minimizing the cost function J, which is the l^2-norm of the error vector between true response value and prediction, for all points in the training set [77]:

$$\vec{w} = \min_w J \tag{12}$$

$$J(\vec{w}) = \left\| \vec{y} - \vec{\hat{y}} \right\|^2 \tag{13}$$

A basic improvement on LLS that prevents overfitting and increases performance in ill-posed problems is ridge (regularized) regression. In this method, large weights are penalized via an additional term that depends on the l^2-norm of $\vec{w}$.

$$J(\vec{w}) = \left\| \vec{y} - \vec{\hat{y}} \right\|^2 + \lambda \|\vec{w}\|^2 \tag{14}$$

The parameter λ controls the regularization strength, i.e. the tradeoff between small weights and small residuals. LLS and ridge LLS models can be written in an equivalent formulation with respect to a vector of dual variables $\vec{\alpha}$, with dimension n, that is related to $\vec{\alpha}$ (also called primal variables) and the matrix of training features $\vec{X}$ as follows:

$$\vec{w} = \vec{X}^T \vec{\alpha} = \sum_{i=1}^{n} \alpha_i \vec{x_i} \tag{15}$$

$$\hat{y} = \sum_{i=1}^{n} \alpha_i \vec{x_i}^T \vec{x} \tag{16}$$

This dual formulation allows the modeler to choose a tradeoff between computational cost in training and prediction [88].

KRR and SVMs perform regularized LLS on a new set of features, obtained by applying the kernel function $k(\vec{x_i}, \vec{x})$ between training points $\vec{x_i}$ and input features $\vec{x}$:

$$\hat{y} = \sum_{i=1}^{n} \alpha_i k(\vec{x_i}, \vec{x}) \tag{17}$$

The choice of kernel function is usually based on prior knowledge of the data or direct performance comparison on the available test dataset [88]. Commonly used kernels are radial basis function (RBF), sigmoid and polynomial.

In KRR, the vector of dual variables $\vec{\alpha}$ is obtained by solving the linear system:

$$\left(\vec{K} + \lambda \vec{I_n}\right) \vec{\alpha} = \vec{y} \tag{18}$$

where $\vec{K} = \{K_{ij}\}, K_{ij} = k(x_i, x_j)$ is the matrix obtained by applying k pairwise between elements of $\vec{x}$ in the training set and $\vec{I_n}$ is the $n \times n$ identity matrix.

The solution algorithm requires inversion of a $n \times n$ dense matrix, which has computational cost of $O(n^3)$. This imposes a first limit on training dataset size. With the two alternative formulations for $\hat{y}$ defined above, prediction scales between $O(d)$ or $O(nd)$ [88]. Hence, a second limit on n is imposed by prediction performance.

Support Vector Machines employ a different cost function that ensures a sparse solution, resulting in faster training and prediction compared to KRR in most regression problems. In SVMs, the cost function J depends on the ϵ-insensitive loss function $L_\epsilon(\hat{y}, y)$:

$$J(\lambda, \vec{w}) = C \sum_{i=1}^{n} L_\epsilon(\hat{y_i}, y_i) + \lambda \|\vec{w}\|^2 \tag{19}$$

$$L_\epsilon(\hat{y_i}, y_i) = \begin{cases} 0, & if\ |y - \hat{y}| < \epsilon \\ |y - \hat{y}| - \epsilon, & otherwise \end{cases} \tag{20}$$

which penalizes only the subset of training points lying outside of a region of width ϵ (ϵ-insensitive tube) around the prediction, called support vectors.

A representation of support vectors is shown in Figure 7. The tunable regularization factor here is $C = 1/\lambda$ [88].

Unlike in KRR, training is performed by solving a constrained optimization problem that does not have closed-form solution and scales generally as $O(n^3)$. Specialized algorithms, however, can improve training performance up to $O(n^2)$. One such algorithms is sequential minimal optimization (SMO), which is implemented in the LIBSVM library [77]. For a given training dataset, SVMs are computationally cheaper than KRR also during prediction. This is due to the sparsity of the $\vec{\alpha}$ vector, that is non-null only for the fraction of training points constituting support vectors [88].

In addition to training and prediction cost, the two kernel methods show differences also in robustness to outliers. This is determined by the error norm used in the cost function J. KRR uses l^2-norm, which penalizes error quadratically and is thus sensitive to outliers, unlike the l^1-norm of $|y - \hat{y}|$ used in SVMs [88].

Given the above considerations, Support Vector Machines are selected as regression method to employ in tasks 6.2 and 6.3 for the improvement of fuel performance codes with machine learning.

3.4 Physics-Informed Learning (PINN)

If enough data is available, one of the ways to reduce computation times is replacement of PDE system solvers with ML methods together with physics-based constraints (physics-informed neural networks, PINNs) [33].

PINNs are a type of enhanced Artificial Neural Networks (ANNs) tailored for a specific physics. Many types of ANNs are widely used to accurately estimate complex functions [34], including physical behaviours, for example, Convolution Neural Networks (CNN), Deep Neural Networks (DNN) and others. Yet, for complicated physical behaviours PINNs should provide more accurate predictions then.

PINNs were introduced in 2017 by Raissi et al. [33], as a new type of method, defined as “Neural networks” that are trained to solve supervised learning tasks while respecting any given laws of physics described by general non-linear partial differential equations. In general, simplest case PINN is an ANN of any type with loss function modified to consider physical constrains. Then a lower error will most of the times lead to a more physical solution, which is not necessary the case with common ANN error criteria.

A common example of simple PINN would be ANN with a heat loss restriction, which can be obtained by following steps:

1. Training a neural network on the heat loss data. NN can be of DNN, CNN or similar types.

2. Describing a standard loss function used for training ANN such as mean squared error. Any loss function evaluating the difference between output and target values can be used.
3. Describing PDE for heat loss like in equation 1, where *F* is a differential operator:

$$F\left[u, \frac{\partial u}{\partial r_i}, \frac{\partial u}{\partial t}, \ldots\right] = 0 \quad (21)$$

4. Describing a physics-based loss function (any loss-function based on applicable physical constrain can be used) as:

$$Loss_{PDE} = \frac{1}{N}\sum_{i=1}^{N}\left\|F\left[\hat{u}\left(t^i, r^i\right)\right]\right\|^2 \quad (22)$$

5. Combining two loss functions from step 2 and 4 into one by summing them.

Advanced design of PINNs provides several advantages for physical process modelling, in addition to general advantages of neural networks. Like ANNs, PINNs are capable of learning almost any pattern, assuming that number of neurons and layers are sufficient.

PINNs inherit and can use all improvements and method variety developed for the ANNs, including but not limited to custom activation, error and limiter functions, variety of architectures and training methods, making PINNs as much versatile and flexible as any other ANN family algorithm.

Training process of PINNs can also be accelerated compared to other NNs, as physics-informed loss function could improve convergence and significantly reduce number of required iterations [36].

However, as all ANN models, PINNs also have a few limitations:

1. ANNs require a large amount of data for training, and their performance is highly influenced by the size and complexity of the dataset. Research shows that on limited data other machine learning models can outperform them [37].
2. Some physics-based loss functions could be computationally costly and slow down training process, which could be especially significant when hyper-parameter optimization is required to find more optimal network architecture.
3. As most other NN-based algorithms, PINNs are also considered to be black-boxes [38][39]. Recent research indicates that with additional transformations ANNs can be considered gray-boxes [40].
4. Like other NNs, PINNs can stuck in local minimum [37].

In addition, when working with PINNs the physically crucial evaluation areas should be considered. Care should be taken in cases when small absolute errors might signify different physics, transient or regimes, e.g., reaction rate 0 and a small number is qualitatively different. If reaction is self-supporting, even small initial error might result in a different transient. In such cases relative errors could result in a more correct model than absolute ones [41][42][44].

A common example of that can be combustion. Too early ignition or a slightly larger combustion rate at the start of the transient might have a larger effect on simulations compared to a larger absolute combustion rate error at the later stage of the transient. For a simulation of physical processes, PINNs should be a suitable algorithm for creating surrogate ML model, when enough training data is available, and physics can be modelled with PINN of a reasonable size.

4. Computation time reduction methods based on surrogate models

Surrogate models have emerged as a powerful tool to speed up simulations without significantly compromising accuracy. Surrogate models, also known as metamodels, are simplified models that approximate the behaviour of more complex, high-fidelity simulations. These models are designed to replicate the essential characteristics of detailed simulations but at a fraction of the computational cost. Surrogate models achieve this by learning from a limited set of high-fidelity simulation data and then making predictions for new, unseen scenarios.

The concept of surrogate modelling is not new, but its application in the nuclear sector has gained significant traction in recent years due to advances in computational techniques and the increasing availability of high-fidelity simulation data. The development and application of surrogate models in the nuclear industry offer several advantages, including reduced computational time, cost efficiency, real-time applicability, enhanced understanding of system behaviour, and uncertainty quantification. The adoption of surrogate models is particularly advantageous in scenarios where high-fidelity simulations are impractical or impossible to perform regularly.

In safety analysis and risk assessment, surrogate models can accelerate the evaluation of various fault scenarios and their potential impacts on reactor operations. This allows for the timely identification of critical risks and the implementation of preventive measures to mitigate these risks. The ability to perform rapid safety assessments is crucial for maintaining high safety standards. Furthermore, surrogate models play a vital role in enabling uncertainty quantification, which is essential for robust decision-making. High-fidelity simulations often involve numerous input parameters with inherent uncertainties.

Quantifying the impact of these uncertainties on simulation outcomes requires extensive sampling and repeated evaluations of the high-fidelity model. Surrogate models, by providing quick

approximations, enable efficient exploration of the uncertainty space, allowing for comprehensive uncertainty quantification and sensitivity analysis. The integration of surrogate models with traditional high-fidelity simulations represents a transformative approach to addressing the computational challenges in the nuclear sector. By leveraging the strengths of both methods, this hybrid approach ensures that simulations are both accurate and computationally efficient.

4.1 Different types of surrogate models: background

4.1.1 Polynomial Regression

Polynomial regression extends linear regression by fitting a polynomial equation to the data. The relationship between the dependent variable and the independent variables is modelled as an nth-degree polynomial. This method can model curved relationships, but as the degree of the polynomial increases, the model can become prone to overfitting, especially with limited data[54].

- Advantages:
  - Simple to implement and interpret.
  - Computationally efficient for low-degree polynomials.
- Disadvantages:
  - Can become unstable and overfit with high-degree polynomials.
  - Limited ability to capture complex, high-dimensional relationships.

4.1.2 Kriging (Gaussian Process Regression)

Kriging, or Gaussian Process Regression (GPR), is a non-parametric, probabilistic model. It assumes that the underlying function is a realization of a Gaussian process and provides predictions with associated uncertainties. The predictions are derived from a weighted sum of known data points, where weights are determined by a covariance function (kernel) that defines the similarity between points [55].

- Advantages:
  - Provides uncertainty estimates for predictions.
  - Can capture complex patterns and handle small datasets well.
- Disadvantages:
  - Computationally expensive, especially for large datasets.
  - Requires careful selection and tuning of the covariance function.

4.1.3 Radial Basis Function (RBF) Networks

RBF networks are a type of artificial neural network that uses radial basis functions as activation functions. The network consists of an input layer, a hidden layer with RBF neurons, and an output layer. The RBF neurons compute a distance metric (usually Euclidean) between the input and a center, and apply a radial basis function (e.g., Gaussian) to this distance [56].

- Advantages:
  - Effective for interpolation in high-dimensional space.
  - Can model non-linear relationships well.
- Disadvantages:
  - Requires selection of the number and positions of RBF centers.
  - Can be sensitive to the scale of input data.

4.1.4 Neural Networks

Neural networks, especially deep neural networks, consist of multiple layers of interconnected neurons. Each neuron applies a non-linear transformation to its inputs and passes the result to the next layer. The network learns by adjusting the weights of the connections through backpropagation, minimizing the difference between predicted and actual outputs [59].

- Advantages:
  - Can model highly complex and non-linear relationships.
  - Scalable to large datasets and high-dimensional inputs.
- Disadvantages:
  - Requires a large amount of training data and computational resources.
  - Can be prone to overfitting and requires careful tuning and regularization.

4.1.5 Polynomial Chaos Expansion (PCE)

PCE models represent the system response as a series of orthogonal polynomials of the input random variables. It is particularly useful for uncertainty quantification as it can efficiently propagate input uncertainties through a model to obtain the probability distribution of the output [60].

- Advantages:
  - Efficient for uncertainty quantification and sensitivity analysis.
  - Can handle correlated input variables.
- Disadvantages:
  - Requires assumptions about the distribution of input variables.

- Complexity increases with the number of input variables and the degree of the expansion.

4.1.6 Bayesian Networks

Bayesian networks are probabilistic graphical models that represent a set of variables and their conditional dependencies using a directed acyclic graph. Each node represents a variable, and each edge represents a conditional dependency. The network provides a compact representation of the joint probability distribution of the variables [61].

- Advantages:
  - Handles uncertainty and probabilistic relationships well.
  - Provides a clear graphical representation of dependencies.
- Disadvantages:
  - Requires expert knowledge to define the network structure.
  - Computationally intensive for inference in large networks.

- Ensemble Methods: Ensemble methods combine multiple models to improve overall performance. Common ensemble techniques include Bagging (Bootstrap Aggregating), Boosting, and Random Forests. These methods enhance accuracy by reducing variance (bagging), bias (boosting), or both (Random Forests) [62].
- Advantages:
  - Often achieve higher accuracy than individual models.
  - Robust to overfitting, especially when using diverse base models.
- Disadvantages:
  - Increased computational cost and complexity.
  - Harder to interpret than single models.

4.1.7 Co-Kriging

Co-Kriging is an extension of Kriging that handles multiple correlated outputs simultaneously. It models the correlation between different outputs to improve the prediction accuracy of each output by leveraging information from related outputs [63].

- Advantages:
  - Efficiently handles multi-output problems.
  - Improves accuracy by exploiting correlations between outputs.
- Disadvantages:
  - More complex and computationally intensive than standard Kriging.

- Requires a good understanding of the correlation structure between outputs.

4.1.8 Sparse Grid Techniques

Sparse grid techniques are designed to handle high-dimensional problems efficiently by selectively sampling the input space. They reduce the number of grid points required for accurate approximation, focusing computational resources on the most critical areas of the input space [64].

- Advantages:
  - Efficient for high-dimensional problems.
  - Reduces the computational burden compared to full grid methods.
- Disadvantages:
  - Implementation complexity increases with dimensionality.
  - May require tuning to balance accuracy and computational efficiency.

These algorithms provide a range of options for creating surrogate models, each with its own strengths and suitable applications. The choice of algorithm depends on the specific characteristics of the problem, such as the complexity of the data, the need for uncertainty quantification, and the available computational resources. By selecting the appropriate surrogate modelling technique, it is possible to achieve significant speedups in simulations while maintaining a high level of accuracy.

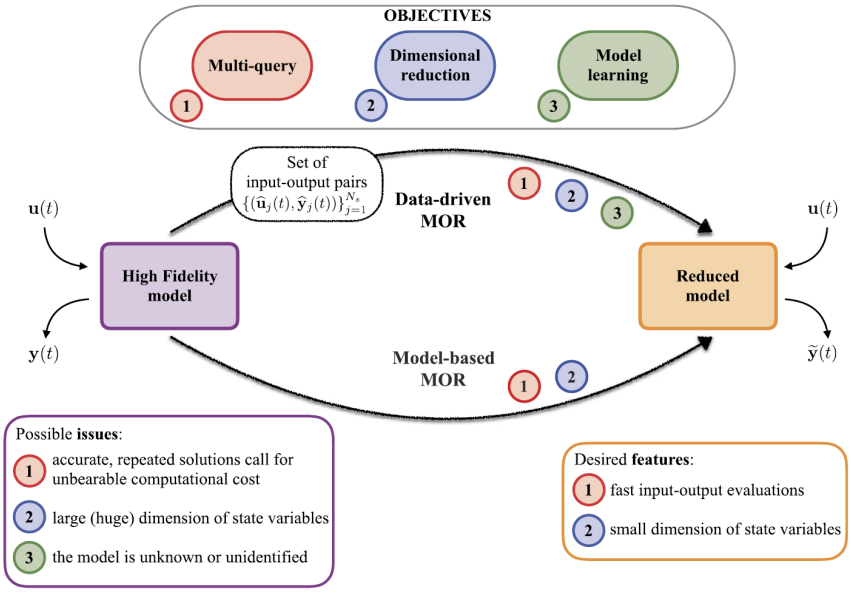

Figure 8: Surrogate models development

Among the descriptions of computation time reduction methods based on ML models, CIEMAT contributes with the analysis of the surrogate modelling as a special case of supervised machine learning.

4.2 Surrogate Modelling Requirements

The use of a surrogate model to replace the code simulations is a technique frequently used, particularly in uncertainty studies [65][66][67]. This method allows saving time of computing while providing enough information about the system. In addition to the uncertainty range, it is possible to obtain other magnitudes, such as accuracy of the predictions. The surrogate model method is especially appropriate if few variables dominate the whole system behaviour.

The methodology described here is based on a generic one proposed by the IAEA for addressing uncertainties in computer modelling regardless the application field [68].

A priori, it is valid for any code modelling, however it is specifically adapted to deal with scenarios modelled using fuel behaviour codes. This methodology, that can be defined as a statistical uncertainty analysis based on reduction methods, is sketched in Figure 9.

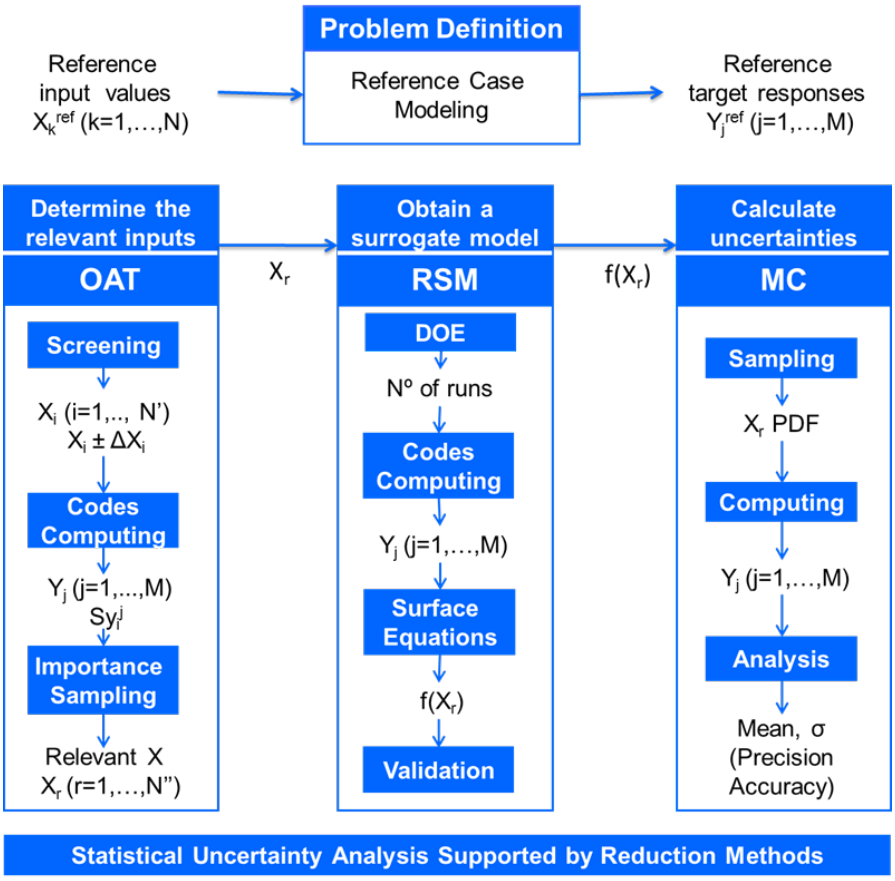

Figure 9: Scheme of uncertainty calculations by using a surrogate model.

It is divided in three steps: reduce the number of parameters to deal with, construct a model capable of replacing the code simulations and perform a Monte Carlo (MC) analysis by using the model. To find out which of the input variables play a key role in the outputs of interest the OAT (One At Time) technique has been selected. Then, the RSM (Response Surface Method) has been applied to build up a simple equation relating the target responses and the selected uncertain inputs.

4.3 Mathematical Model

The mathematical model (RSM in this case) is built following the steps shown in Figure 9 and described below through an application carried out in the previous work [69]. Particularly, a surrogate model of the end-of-life rod internal pressure (P_{EOL}) predicted by the fuel performance code FRAPCON [74] was derived and used for carrying out uncertainty analysis. The scenario was the irradiation of a 17x17 PWR fuel rod [69] up to a burnup level of 65 GWd/tU (based on four different enveloping power histories).

4.3.1 OAT

The OAT analysis is a procedure used for quantifying the influence of an input variable uncertainty (X) on a selected output uncertainty (Y).

It consists in perturbing only one input variable per computer run: each input is evaluated at its upper or lower value, whereas the rest of the input deck is kept at their nominal values. This approach is not powerful enough to give quantitative insights, nonetheless it is suitable to identify the variables and models causing larger effects on the target responses [70], [71], [72], [73].

Note that, by using the OAT technique is not possible to detect the potential synergies between the uncertain inputs, that is, the cross effects have been neglected when selecting the relevant variables.

Going to the application abovementioned, the influence of the uncertain inputs on P_{EOL} prediction was assessed. As a reference case, the scenario previously explained was considered.

The inputs selected to perform the analysis can be divided into three groups:

- Manufacturing variables. The uncertain variables and their variation range are shown in Table 1.
- Power history. The analysis was carried out with the curves shown in Figure 10. As upper and lower bounds, a q'max variation of 10% is considered in each case.
- Models. The FRAPCON code provides the option to bias eight models related to: fuel thermal conductivity (ftc), fuel thermal expansion (ftex), fission gas release (fgr), fuel swelling (swell), cladding creep (creep), cladding axial growth (gro), cladding corrosion (cor) and hydrogen pick up (h2). Except the latter, the rest of the models may affect P_{EOL}, so the corresponding bias was used in this study (a confidence level greater than 90% was considered).

Once defined the uncertain inputs and their range of variation, the OAT approach was performed for each power history from Figure 10.

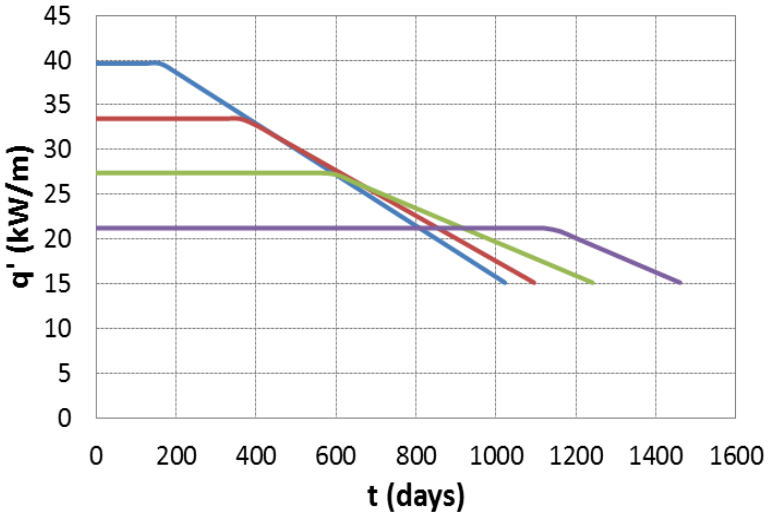

Figure 10: Average linear power, q', vs time [71].

Variable	Reference Value	Tolerance
Cladding inner diameter, d_{ci} (mm)	8.18	0.04
Pellet diameter, d_p (mm)	8	0.013
Plenum length, l_p (mm)	254	11.4
Dish radius, r_d (mm)	2.005	0.25
Dish depth, d_d (mm)	0.287	0.05
Pellet density, %TD (%)	95	0.91
Pellet roughness, r_p (μm)	1	-0.9/+2
Density increase, ΔD (kg/m^3)	98.64	43.84
Cladding roughness, r_c (μm)	0.5	-0.4/+0.5
Fill gas pressure, P_{fill} (MPa)	2.41	0.069

Table 1. Manufacturing uncertain inputs.

The results are collected in Figure 11 in terms of P_{EOL} relative deviation, RD, with respect to the reference value (only RD greater than 5% are displayed). The highest sensitivities were obtained with ftc and fgr (maximum RD of around 18% and 16%, respectively).

Furthermore, these models gave rise to non-negligible RD (greater than 5%) with all the power histories. Therefore, ftc and fgr uncertain inputs were selected to build up the surrogate model (RSM).

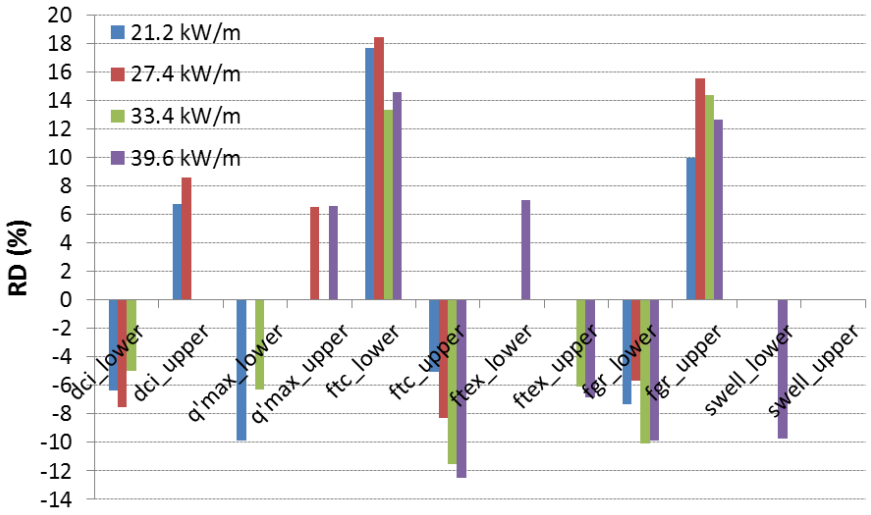

Figure 11: Results of OAT with RD greater than 5% (q'_{max} represented by colours)

4.3.2 RSM

As explained above, the RSM technique is used in order to build up a model able to replace the code simulations [66], [70], [72]. The RSM finds out the relationships between one or more responses and a group of variables. Each run is regarded as an experiment; a particular combination of runs defines a DOE (Design Of Experiment).

The objective of a DOE is determining the points where the responses of interest will be evaluated to generate the database used to build up the RSM approximation. Then, common regression techniques are applied to calculate the relationship between the responses and the input variables.

The DOE can have an influence on the accuracy of the approximation and the cost of constructing the RSM equations. The choice of DOE (i.e. full or fractional factorial design, central composite design, Audze-Eglais, Box–Behnken, etc.) depends on the selected RSM model, the number of variables to be considered and the desired accuracy (Montgomery, 2013)[75].

The last step to be able to use the model instead of the code simulations is to check its performance. The ability of the RSM equations to reproduce the outputs given by the code must be ensured. To do so, the tests for the significance of the regression are applied. Besides, the scatter plots of residuals and the normal probability plots of residuals are examined.

It is important to underline that the RSM equations are only valid in the selected ranges of the input variables.

In the application shown, the DOE chosen was a full factorial design, since only two variables were retained after the OAT application (ftc and fgr). As shown below, a second-degree model was selected; therefore, three levels were required. Thus, this DOE consisted in considering all the

combinations of the three levels of the selected variables, which covered the whole studied region and ensured a high degree of resolution.

It is appropriate to define dimensionless variables (X_1, X_2) for describing the uncertain inputs in the DOE. In this case, ftc and fgr were described in the [-1, 1] interval; 0 was assigned to the midpoints (reference case) and ±1 was assigned to the lower and upper bounds. The results of the simulations for each power history considered are gathered in Table 2.

Based on the DOE shown, common regression techniques were applied to calculate the relationship between P_{EOL} and the uncertain inputs (as coded variables X_1 and X_2):

$$P_{EOL} = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_{11} X_1^2 + \alpha_{22} X_2^2 + \alpha_{12} X_1 X_2 \quad (24)$$

The fitting parameters are listed in Table 3. The relative influence of ftc and fgr on P_{EOL} can be evaluated by examining these parameters. Although the weight of both is comparable, ftc seems to weigh more than fgr in terms of uncertainty, which is in accordance with OAT results.

Concerning the quadratic and the interaction terms involved, most of them have a significant contribution.

The accuracy of the model was checked by verifying the ability of the RSM equation to reproduce P_{EOL} given by the code. To do so, the test for the significance of the regression was applied.

The results obtained for each q'_{max} ($R^2 > 0.99$) confirm that the model derived is a suitable tool.

run	X_1 (ftc)	X_2 (fgr)	P_{EOL} (MPa)			
			21.2 kW/m	27.4 kW/m	33.4 kW/m	39.6 kW/m
1	0	0	4.208	4.383	4.984	5.401
2	0	-1	3.900	4.133	4.483	4.868
3	0	1	4.625	5.063	5.701	6.083
4	-1	0	4.950	5.191	5.649	6.187
5	-1	-1	4.445	4.695	5.177	5.735
6	-1	1	5.570	5.831	6.391	6.973
7	1	0	3.996	4.019	4.408	4.728
8	1	-1	3.996	4.019	4.018	4.500
9	1	1	3.996	4.346	4.725	5.477

Table 2. DOE results for each q'_{max}.

Parameter	Value			
	21.2 kW/m	27.4 kW/m	33.4 kW/m	39.6 kW/m
α_0	4.219	4.426	5.010	5.339
α_1	-0.496	-0.555	-0.678	-0.698
α_2	0.308	0.399	0.523	0.572
α_{11}	0.247	0.157	0.005	0.149
α_{22}	0.038	0.150	0.069	0.167
α_{21}	-0.281	-0.202	-0.127	-0.065

Table 3. RSM fitting parameters for each $q'_{max.}$

Based on the surrogate model derived in this application, a MC analysis was conducted. A 2000 sample was generated by varying ftc and fgr simultaneously, using simple random sampling with uniform distributions defined between the upper and lower bounds.

The frequency plots obtained are shown in Figure 12. In order to set the uncertainty bands of P_{EOL} predictions, the lower and upper bounds were assumed to be the 0.025 and 0.975 percentiles of the output distributions (95% of the distribution). The results are displayed in Figure 13.

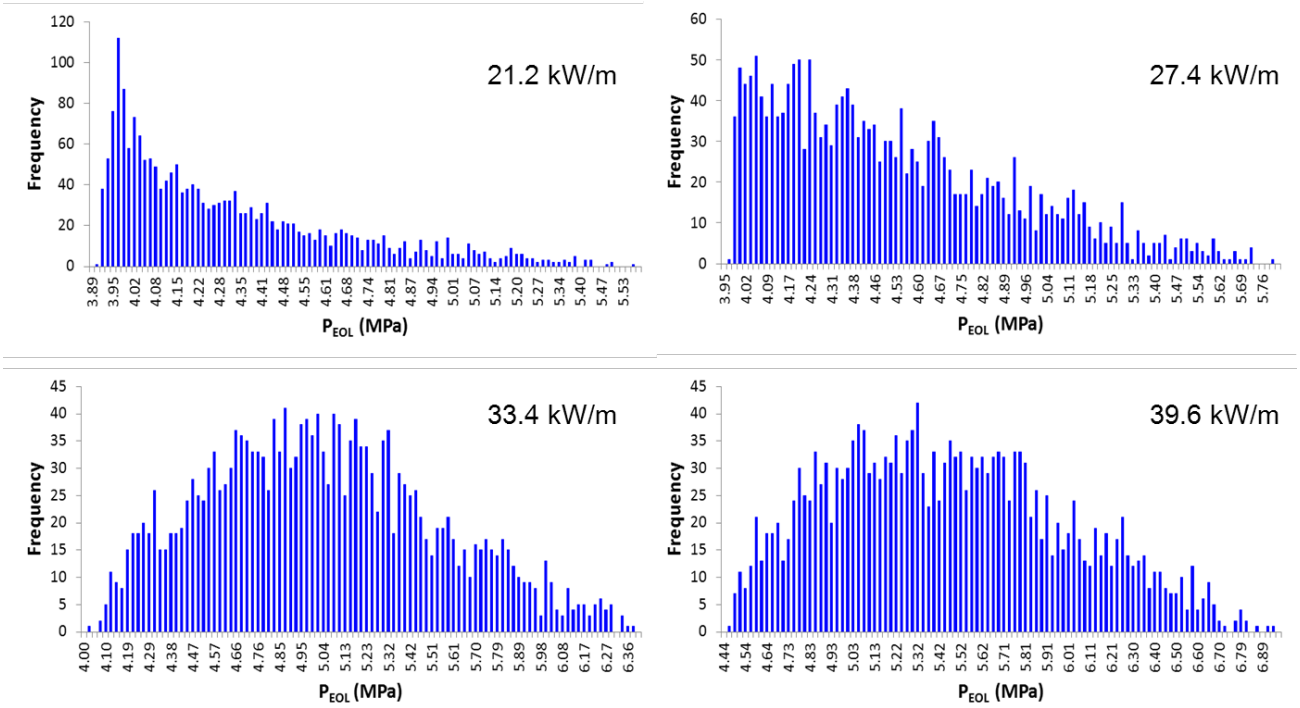

Figure 12: Output distributions of P_{EOL} for each q'_{max}

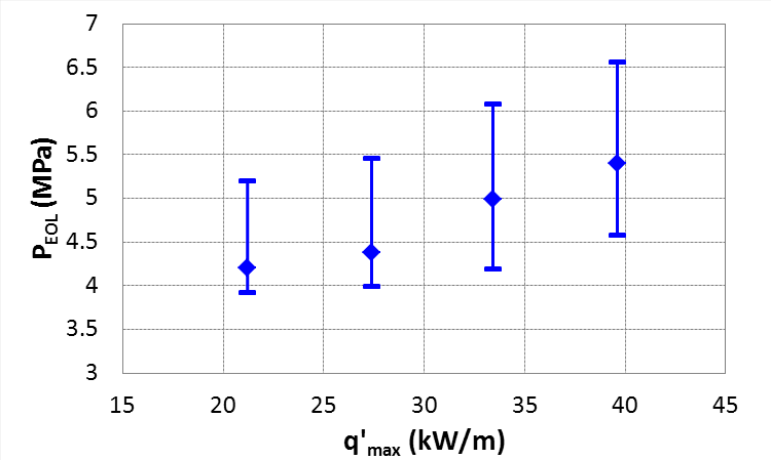

Figure 13: P_{EOL} predictions with uncertainty bands.

5. Preliminary application

5.1 Pellet-cladding mechanical interaction

5.1.1 The FINIX fuel performance code

FINIX is the transient fuel behaviour simulation module developed in VTT since 2012. It was originally designed to be integrated into larger simulation codes, so it prioritizes lightweight operation over a comprehensive modelling of fuel behaviour phenomenology [77]. It has since evolved to be a standalone fuel simulation code, while maintaining the same design goals as the original release.

The module is designed to simulate conventional LWR fuel rods in operational and accidental transients. It consists of coupled thermal and mechanical models for pellet, cladding and gas gap, solved with fixed point iterations at subsequent time steps. Material properties are computed from publicly available correlations [77][78].

The solution scheme for both models is generally referred to as 1.5D in cylindrical, axisymmetric geometry. In this scheme, pellet and cladding are discretized radially in cylindrical elements, and axially in independent nodes coupled via pressure and thermal conductivity of the pellet-cladding gas gap. The azimuthal dependence of temperature and material properties is neglected, while the axial dependence is included by solving uniform axial slices. The radial discretization of each slice employs the finite element method; thus each cylindrical element is homogeneous [84].

The mechanical model in FINIX is based on a set of simplifying assumptions compatible with the 1.5D solution scheme just presented. Several of these assumptions are shared with similar 1.5D fuel performance codes, such as FRAPCON/FRAPTRAN and TRANSURANUS, while others are unique to the simplified model chosen for FINIX. In particular, it is assumed that [83], [84]

- strain and displacement are linearly related; this assumption provides adequate result and numerical stability only for small strains (5% to 10%).
- a generalized plain strain condition holds, i.e. the axial deformation is constant along the radius.
- the pellet is completely rigid under external stress, but deformable due to thermal expansion, radial relocation, irradiation swelling and densification.
- half of the radial pellet displacement due to fragment relocation is allowed to recover, before the pellet is considered rigid to cladding interaction.
- the pellet is perfectly cylindrical and centered within the cladding, neglecting manufacturing features, with mechanical contact at the centerline.
- pellet, cladding and cladding oxide are isotropic and their thermomechanical properties (i.e., specific heat, thermal conductivity, Meyer's hardness, Young's modulus, Poisson's ratio) can be expressed as function of temperature and burnup by semi-empirical correlations.
- physical phenomena leading to dimensional changes in pellet and cladding (i.e., thermal strain, radial relocation, irradiation swelling, densification and creep) can be expressed by semi-empirical correlations.
- the cladding can be modelled as thin cylindrical shell, neglecting radial differences in temperature and stress and assuming no significant bending in the axial direction.
- pellet cracking is not modelled explicitly.
- the plastic behavior of cladding can be described based on quasi-static incremental plasticity theory and strain-hardening hypothesis; yielding occurs according to the von Mises yield function, and the components of plastic strain increments are related through the Prandtl-Reuss flow rule.

The mechanical model is responsible for calculating stresses σ, strains ϵ and corresponding displacements u in the cladding under different loading regimes throughout a simulated transient. At each timestep, cladding mechanical equilibrium is determined from a consistent set of equations and suitable boundary conditions, in terms of main components of the mechanical variables. In cylindrical geometry, they are r, θ and z.The thin shell approximation further allows to translate radial displacements into changes of inner radius R_i and outer radius R_o. These are primarily used in FINIX in place of u. The complete set of equations is detailed in [84]

As known from literature and industrial practice, the pellet-cladding gap is introduced during manufacturing and closes progressively during operation, disappearing completely when pellet and cladding are fully in contact. Two calculation regimes are considered in the mechanical model

depending on the status of the pellet-cladding gap. The appropriate regime is selected by comparing cladding inner radius and fuel outer radius R_f, corrected with the respective values of roughness ρ. If the condition $R_i > R_f + \rho_c + \rho_f$ is satisfied, the open gap regime is selected. Pellet and cladding are separated by a volume of internal gas at pressure P_i, determined by temperature and concentration of gas species.

The cladding is loaded by internal gas pressure on its inner surface, and primary coolant pressure on its outer surface. The unknowns R_i and P_i are related through the pellet-cladding gap volume and are found iteratively at each time step. Fixed point iteration is used to determine P_i from an initial guess, which is then used to get the components of σ and ϵ that guarantee mechanical equilibrium in the cladding.

The effect of creep and plastic cladding behavior is introduced in the equations for ϵ according to incremental plasticity theory. The mechanical variables are computed in the following order until convergence is reached [84].:

$$P_i \rightarrow \sigma_r, \sigma_\theta, \sigma_z \rightarrow \epsilon_r, \epsilon_\theta, \epsilon_z \rightarrow R_i, R_o \tag{25}$$

In closed gap regime, pellet and cladding are in mechanical contact and develop an interfacial pressure (contact pressure) P_c, acting on the inner cladding surface. Depending on the value of P_c relative to P_i, the closed gap regime is thus further distinct into strong contact and weak contact. The distinction considers that the pressure exerted by the gas allow pellet and cladding to slide against each other, when P_c is sufficiently small. The value of P_c is computed first assuming strong contact, then compared to P_i and recomputed if the assumption is incorrect. In both cases, R_i is fixed and equal to $R_f + \rho_c + \rho_f$ [84].

Strong contact occurs when $P_c > P_i$ and requires the additional constraint of axial locking between pellet and cladding, which translates to imposed strain in the axial direction ϵ_z. R_o can be obtained explicitly and used to get stresses and strains:

$$R_i, \epsilon_z \rightarrow R_o \rightarrow \epsilon_r, \epsilon_\theta \rightarrow \sigma_r, \sigma_\theta, \sigma_z \rightarrow P_c \tag{26}$$

In the weak contact condition, $P_c = P_i$ and all mechanical variables are determined explicitly:

$$R_i, P_c \rightarrow R_o \rightarrow \sigma_r, \sigma_\theta, \sigma_z \rightarrow \epsilon_r, \epsilon_\theta, \epsilon_z \tag{27}$$

Resulting stresses and strains are averaged over the cladding thickness, due to the thin shell approximation [84].

Cladding failure can be predicted from the mechanical state of the cladding by applying suitable failure criteria. FINIX implements two simple failure criteria based on the effective strain ϵ_e and its rate of change $\dot{\epsilon_e}$:

- the plastic instability criterion predicts rod failure when $\epsilon_e > 0.02$ and $\dot{\epsilon_e} > 0.0278\ \mathrm{s}^{-1}$
- the overstrain criterion predicts failure when $\epsilon_e > 0.5$.

The criteria are suitable for the analysis of common accidental scenarios, including RIA and LOCA [81].

5.1.2 Application of Regression Methods PCMI modelling

The combined effect of pellet-cladding mechanical interaction and chemical attack from corrosive fission products was identified as cause of cladding failure in fuel elements subjected to power ramps already in the 1970s. The process is known as pellet-cladding interaction (PCI) assisted, iodine-induce stress corrosion cracking (I-SCC), or PCI/SCC, and extensively documented in literature. PCI/SCC is primarily of interest in fuel operated at reduced power level with respect to nominal power [79].

During reduced power operation, iodine accumulates naturally due to fission, while cladding undergoes creep-down past the limit reached at nominal power. A subsequent return to nominal power causes additional mechanical load on the cladding, that combined with a simultaneous anticipated operational occurrence (AOO) can result in cracking.

The tangential (hoop) component of cladding stress σ_θ is known to be a key parameter in PCI/SCC failure because it drives the opening of longitudinal cracks initiated at the inner cladding surface [79].

In view of the transition of operating reactors to flexible power operation, fuel performance codes are required to predict fuel failure to PCI/SCC by implementing suitable failure criteria.

For example, the reference failure criterion formulated by Mattas et al. [85]and available in TRANSURANUS [82], depends explicitly on σ_θ for estimating the chemical crack growth rate of a SCC induced fracture. Accurate estimation of σ_θ is thus crucial in modelling fuel under flexible power operation.

Full 3D fuel modelling codes such as BISON or ALCYONE can compute the non-uniform cladding hoop stress distribution due to the hourglass shaping of the pellet fragments, depicted in Figure 14 [86]. Traditional 1.5D codes, on the other hand, produced values of σ_θ averaged over the cladding thickness, under the set of simplifying assumptions described previously.

In the work planned for Task 6.2, we propose to bridge the gap between 3D and 1.5D modelling of PCMI by employing regression models to estimate stress concentration in the cladding. In particular,

we plan to define a dimensionless quantity representative of the stress concentration caused by PCMI, such as the ratio between maximum and average hoop stress in the cladding:

$$R_\theta = \frac{\sigma_\theta^{max}}{\sigma_\theta} \tag{28}$$

This quantity would be used as target feature of a regression model implemented in FINIX.
Given a set of loading conditions during a power transient, the model would provide a value of R_θ that the code can use to compute a mechanical variable for a PCI/SCC failure criterion, e.g., σ_θ^{max}. The training dataset would be based on full 3D simulations with the OFFBEAT code, with input features characteristic of fuel performance calculations (e.g., linear heat rate, burnup). The different assumptions of the mechanical models of OFFBEAT and FINIX require the output feature to be relative, in order to ensure a consistent mechanical state.
A natural evolution of the work could extend to a simultaneous prediction of shear stress at the cladding surface, that was suggested by Michel [87] to be equally relevant in the definition of a more comprehensive failure criterion for PCI/SCC.

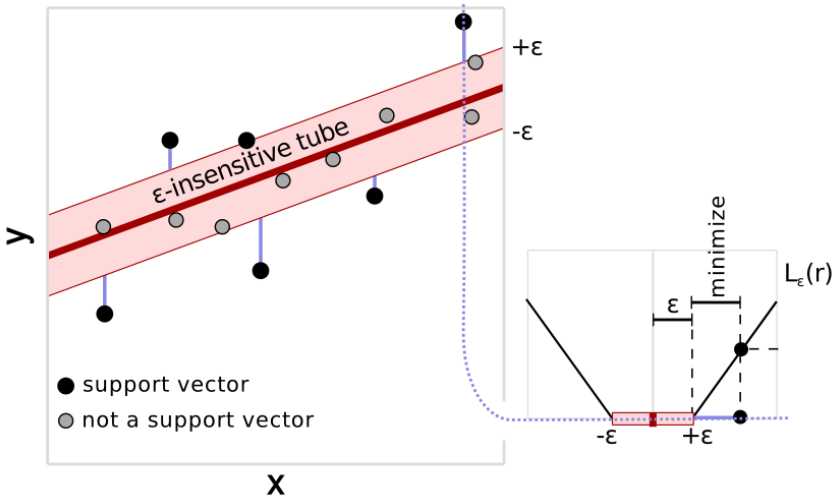

Figure 14: Representation of support vectors for a linear kernel SVM. Points in grey are located within the ϵ-insensitive tube and have zero penalty. Points outside the marked area are support vectors [89]

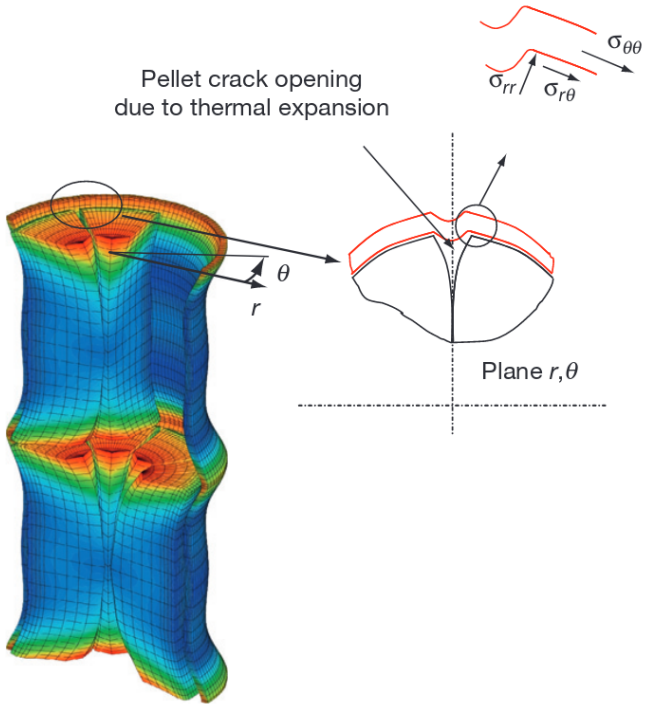

Figure 15: Qualitative stress distribution in pellet and cladding, due to pellet hourglassing and pellet-cladding interaction [86]

5.2 Fission gas behaviour

In the frame of WP4, POLIMI is developing models for athermal fission gas release, evolution of the dislocation density in uranium dioxide under irradiation, and gas diffusion in non-spherical domains. Each of these model developments involves the use of surrogate models to reduce their computational requirements and thus being applicable in fuel performance code simulations.

The content of this subsection has been disseminated within OperaHPC through three contributions to NuFuel2023 and two Master of Science thesis projects at Polimi [90][91][92][93].

5.2.1 Athermal fission gas release

Athermal fission gas release is relevant in light water reactors during normal operation, where low fission gas release (roughly < 1.0%) is usually observed. Various mechanisms of athermal release exist, e.g., induced by recoil, knockout, or venting through the open porosity [93].

In the frame of WP4, POLIMI and JRC-Ka are targeting the improvement of an available model describing the transport of gas atoms from within the fuel grains to the open porosity tunnels, and from there to the rod free volume [94]. In the standard model for athermal release, the amount of gas reaching the rod free volume is set to be proportional to the open porosity through a geometrical

athermal venting factor. A more physical description of the process requires to describe the diffusion of fission gas towards grain edges (i.e., the open porosity tunnels), which is a complex phenomenon and depends on several local geometrical and physical properties at the sub-grain scale.

This rather complex behaviour at the sub-grain scale has been modelled with an high-fidelity tool accounting for the local geometry in the proximity of a grain edge (in terms of angle between consecutive grain faces, length of the grain edge, fabrication porosity and its open fraction), the local amounts of fission gas available for diffusion (in terms of initial intra-granular gas concentration, local burnup, and local fission rate), and the local transport properties of fission gas (in terms of local temperature and fission rate determining the fission gas diffusivity). This local high-fidelity model provides an accurate description of the phenomenon but has a computational cost which is not compatible with its direct inclusion in fuel performance codes. For this reason, in WP6 we developed a surrogated model based on an artificial neural network trained on virtual data generated from the high-fidelity model.

The variation range of the input parameters is reported in Table 4, with the structure of the network being depicted in Figure 16. As can be seen, the output of the artificial neural network is *athermal venting corrective factor* which is in turn included in the rate theory model implemented in SCIANTIX and condenses in a single scalar value all the aspects explicitly accounted for by the high-fidelity model.

Future work will target the verification and the validation of such model, including the use of the artificial neural network within fuel performance simulations, in the frame of WP5 and WP6.

Quantity	Range
Grain-edge angle	90°-135°
Fabrication porosity	2 %-9 %
Grain-edge length	3.67-11 μm
Burnup	0-50 MWd/kgUO_2
Temperature	300-1600 K
Gas production rate (irradiation conditions)	10^{18}-10^{20} at/m^3-s
Initial gas concentration (annealing conditions)	10^{19}-10^{24} at/m^3

Table 4: Validity range for the quantities analysed with the high-fidelity 2D diffusion model

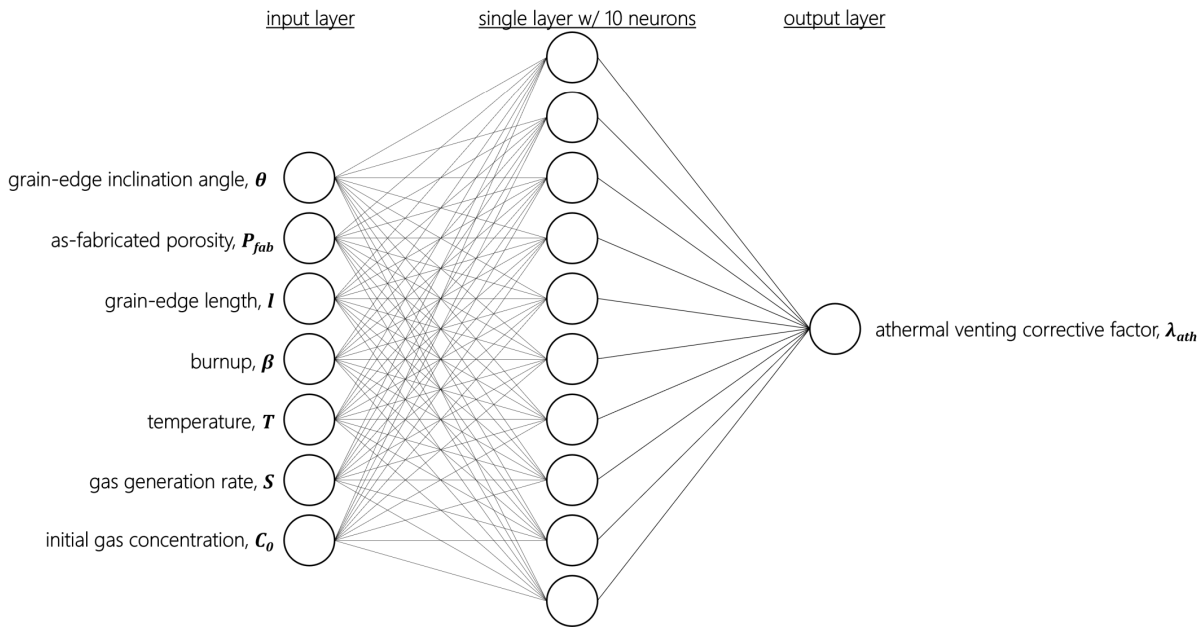

Figure 16: Structure of the artificial neural network used to surrogate the high-fidelity model describing the transport of fission gas towards the grain edges and thus contributing to athermal fission gas release.

5.2.2 Dislocation density in the fuel

The dislocation evolution is crucial in determining the amount of high burn-up restructuring at the periphery of oxide nuclear fuel irradiated in light water reactors. A mechanistic model for the evolution of point and extended defects in uranium dioxide under irradiation has been developed at this purpose in a collaborative activity by CEA and Polimi.

The model provides a physics-based description of the evolution of point and extended defects in the fuel and is constituted by a set of seven non-linear ordinary differential equations to be solved coupled to each other. Nevertheless, not all the seven state variables considered by the physics-based model are needed in a fuel performance calculation, and for a subset of them no separate effect validation is possible. For these reasons, to include the full coupled solution of the system of equations in a fuel performance code simulation represents an additional computational burden with limited added value to the predictive capabilities of the code.

Given these premises, we developed a surrogate model through an artificial neural network trained on a virtual dataset generated by the high-fidelity solution of the physics-based model.

The virtual dataset allows correlating inputs (local fission rate, local temperature, and local burnup) and outputs (local concentration of vacancies, local concentration of interstitials, local concentration of interstitial and vacancy loops, local concentration of voids, average size of interstitial loops, dislocation density) of the physics-based model.

For the targeted application, we select only the dislocation density as a relevant output, thus greatly reducing the dimensionality of the problem and its computational time. As inputs, we consider the local fission rate density varied between 5 10^{18} and 5 10^{19} fiss/m^3-s, the local temperature varied between 100 and 1200 °C, and the local burnup varied between 0 and 200 GWd/t.
As output, we extract from the virtual dataset the local increase rate of the dislocation density (see Figure 18). This value is then provided to SCIANTIX numerical solver which integrates it along the time step. Separate effect validation of the physics-based model is ongoing in collaboration between CEA and Polimi, together with the verification of the surrogate model in terms of approximation error (WP5\Task 5.1).
Further development will involve the construction of surrogate models with additional parameters among the outputs of the physics-based model which are relevant to other model options available both in SCIANTIX and in fuel performance codes besides the high burnup structure models, e.g., tracking the evolution of the local concentration of vacancies can provide valuable information on the fission product mobility in Chromium-doped fuels being analysed in WP7.

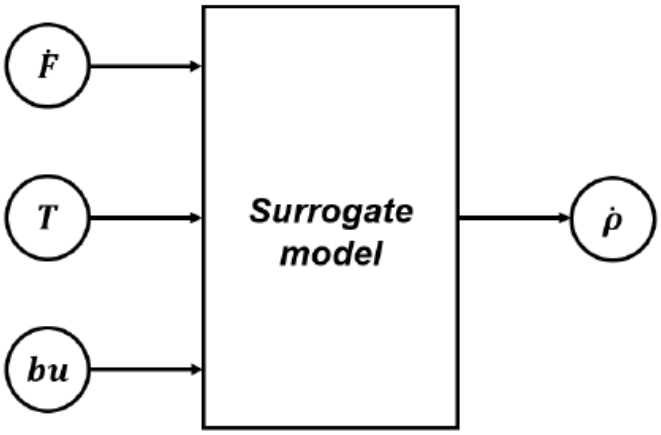

Figure 17: Schematic representation of the surrogate model. Input quantities are the local fission rate density ($\dot{F}$), fuel temperature (T) and burnup (bu), output quantity is the instantaneous local increase rate of the dislocation density.

5.2.3 Gas diffusion in non-spherical domains

As an improvement of the predictive capabilities of SCIANTIX, Polimi is progressively generalizing the numerical solvers available in the code for the treatment of fission gas diffusion.
Remarkably, fission gas diffusion equation is the only partial differential equation solved by the code, the rest being ordinary differential equations, and thus it accounts for a considerable fraction of the

computational time. Moreover, the typical values of the diffusivity of fission gas are low, which translates into steep gradients of gas concentrations and require a fine discretization of the domain.
In SCIANTIX, to solve this mathematical problem in spherical grains we adopt spectral algorithms, expanding the solution of the diffusion equation on the eigenfunction of the Laplacian operator. This numerical approach is suitable for standard geometries, e.g., spheres, for which the eigenfunctions are known analytically.
To extend this approach to generic geometries and to domains larger than the calculation domains used in the coupled fuel performance codes (e.g., the cylindrical grains observed after restructuring at high temperature, which can have a length of up to a millimetre and thus way larger than the mesh element adopted in a fuel performance simulation) we are developing reduced order models allowing for computational times in line with application in fuel performance codes, applicability in different geometries and non-uniform temperature fields, and seamless implementation in meso-scale codes.
The POD-FV-ROM algorithm adopted in this work is illustrated in [95].
The OFFLINE phase involves the solution of the full-order problem in OpenFOAM (finite volume discretization, or in general in a high-fidelity tool able to describe the diffusion problem in the domain of interest) and the proper orthogonal decomposition via Galerkin projection and is performed once and for all. The ONLINE phase is performed in SCIANTIX, allowing the reconstruction of the intra-granular gas concentration.
Current development target simplified geometries such as plates, spherical shells, and cylindrical shells, which are relevant for applications cases within OperaHPC WP5 and WP7. Detailed verification of the solvers to be obtained is targeted in T5.1.

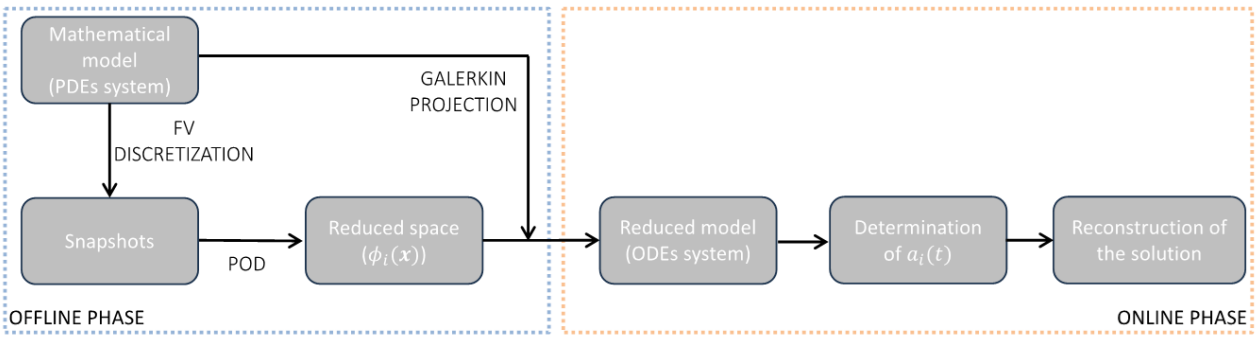

Figure 18. Schematic of the POD-FV-ROM algorithm used to generalize the solution of diffusion equations to non-spherical domains.

6. Summary and Outlook

This document provided first a review and a detailed analysis of existing models for Pellet-Cladding Mechanical Interaction (PCMI), the Fission Gas Behaviour (FGB) and Fuel Overfragmentation (FO) phenomena in industrial type fuel performance codes, such as TRANSURANUS, FINIX and CYRANO3, for which an improvement is expected by the substitution of an empirical formulation by machine-learning, surrogate modelling, or data-driven approaches. In a second part of the document a review was proposed for an identification and a presentation of the most powerful computation time reduction methods. The latter were decomposed in two categories with Machine Learning Methods (MLM) and surrogate models.

The adoption of MLM or surrogate models, such as Proper Orthogonal Decomposition (POD), Support Vector Machines (SVM) and Physics-Informed Neural Networks (PINNs), demonstrates significant potential in reducing computational demands while maintaining high levels of accuracy. These methods enable computation times consistent with the requirement of industrial safety studies. By replacing some of the existing industrial models with these new approaches, one can expect to take advantage of the advanced simulation tools and their physically based formulation more rapidly in the context of the industrial studies.

The document underscores the importance of digital replicas (DR), using 3D advanced simulation tools, in order to build the learning data bases needed for the identification or the training of MLM and surrogate models.

Beyond the methods review, a key achievement highlighted in this document is the existing successful implementation of surrogate modelling techniques, particularly the Response Surface Method (RSM) and Monte Carlo (MC) analysis, to predict the end-of-life rod internal pressure (PEOL) in fuel performance codes like FRAPCON. This approach has proven effective in managing the uncertainties associated with various input parameters, providing a robust framework for uncertainty quantification and sensitivity analysis.

The findings and methodologies outlined in this document cover the way for future advancements in computational efficiency, which are essential to achieve the Specific Technical Objective STO6 devoted to the development of improved models for industrial fuel performance codes using computational time reduction methods in the OperaHPC project.

7. References

[1] B. Michel et al. OperaHPC_D9.2 - Detailed Work Plan for technical work Packages. 2023.

[2] Quarteroni, A.; Manzoni, A.; Negri, F., Reduced basis methods for partial differential equations: an introduction, Springer International Publishing, 2016, https://doi.org/10.1007/978-3-319-15431-2

[3] Sirovich, L. (1987). Turbulence and the dynamics of coherent structures. Parts I--III. Quarterly of Applied Mathematics, 45(3), 561-590

[4] Barrault, M.; Maday, Y.; Nguyen, N. C.; Patera, A. T., An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations, Comptes Rendus Mathematique, 667-672, 2004

[5] Chaturantabut, S.; Sorensen, D. C. Nonlinear Model Reduction via Discrete Empirical Interpolation.SIAM Journal on Scientific Computing, 2737-2764

[6] Farhat, Charbel; Avery, Philip; Chapman, Todd; Cortial, Julien. Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency.Int J. for Numerical Methods in Engineering, 625-662, 2014

[7] Ryckelynck, David. Hyper-reduction of mechanical models involving internal variables.International Journal for Numerical Methods in Engineering,75-89, 2009

[8] J. Fauque de Maistre, Hybrid hyper-reduced modeling for contact mechanics problems, Paris: Ph.D. Thesis, Université Paris Science et Lettres, 2018.

[9] S. Le Berre, Condition Number and Clustering-Based Efficiency Improvement of Reduced-Order Solvers for Contact Problems Using Lagrange Multipliers, Paris: Ph.D. Thesis, Université Paris Sciences et Lettres , 2022.

[10] J. Fauque, I. Ramière et D. Ryckelynck, «Hybrid hyper-reduced modeling for contact mechanics problems,» International Journal for Numerical Methods in Engineering, vol. 115, n° %11, pp. 117-139}, 2018.

[11] S. Le Berre, I. Ramière, J. Fauque et D. Ryckelynck, «Condition Number and Clustering-Based Efficiency Improvement of Reduced-Order Solvers for Contact Problems Using Lagrange Multipliers,» Mathematics, vol. 10, n° 19, p. 1495, 2022.

[12] F. Fritzen, B. Haasdonk, D. Ryckelynck et S. Schöps, «An Algorithmic Comparison of the Hyper-Reduction and the Discrete Empirical Interpolation Method for a Nonlinear Thermal Problem,» Mathematical and Computational Applications, vol. 23, n° 11, p. 8, 2018.

[13] R. Lo Frano et al. OperaHPC: MS4 - Detailed analysis of industrial models to improve. 2024.

[14] F. Feyel. A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua Comput. Methods Appl. Mech. Eng. (ISSN 0045-7825) 471 (28) (2003) 3233–3244

[15] I. Ramière, R. Masson, B. Michel, S. Bernaud. Un schéma de calcul multi-échelles de type Éléments Finis au carré pour la simulation de combustibles nucléaires hétérogènes, 13e colloque national en calcul des structures. Université Paris-Saclay. Giens, Var, France.2017

[16] D. Baron, L. Hallstadius, K. Kulacsy, R. Largenton, J. Noirot.Fuel Performance of Light Water Reactors (Uranium Oxide and MOX - Comprehensive Nuclear Materials 2nd Edition, 2020, Elsevier Ltd - Editors: Rudy Konings, Roger E Stoller

[17] J. Labat, R. Largenton, J.-C. Michel. Multiscale identification and NTFA reduction of the elasto-viscoplastic polycrystalline behaviour of uranium dioxide (UO2) for wide range of loading conditions, Journal of Nuclear Materials 582, 154471, 2023, 2023

[18] G.J. Dvorak. 1992. Transformation field analysis of inelastic composite materials. s.l. : Proc. R. Soc. Lond. Ser. A, Math. Phys. Sci. 437 (1992) 311–327, 1992. https://doi .org /10 .1098/rspa .1992 .0063..

[19] G.J. Dvorak, Y. Benveniste. 1992. On transformation strains and uniform fields in multiphase elastic media. s.l. : Proc. R. Soc. Lond. Ser. A, Math. Phys. Sci. 437 (1992) 291–310, 1992. https://doi .org /10 .1098 /rspa .1992 .0062.

[20] X. Zhang, C. Oskay. Eigenstrain based reduced order homogenization for polycrystalline materialsomput. Methods Appl. Mech. Eng. 297 (2015) 408–436, 2015

[21] J.-C. Michel, U. Galvanetto, P. Suquet. 2000. Constitutive relations involving internal variables based on a micromechanical analysis. s.l. : in: Continuum Thermomechanics: The Art and Science of Modelling Material Behaviour, Springer, Netherlands, 2000, pp. 301–312, 2000. https://doi .org /10 .1007 /0 -306 -46946 -4 _23.

[22] J.L. Chaboche, S. Kruch, J.F. Maire, T. Pottier. 2001. Towards a micromechanics based inelastic and damage modeling of composites. s.l. : Int. J. Plast. (ISSN 0749-6419) 17 (4) (2001) 411–439, 2001. https://doi .org /10 .1016 /S0749 -6419(00)00056 -5.

[23] R. Largenton et al. OPERA-HPC: MS3 - Common methodology for the rheological non-linear mechanical modelling and homogenization techniques for fuels and claddings. 2024.

[24] R. Largenton, J.-C. Michel, P. Suquet. 2014. Extension of the nonuniform transformation field analysis to linear viscoelastic composites in the presence of aging and swelling. s.l.: Mech. Mater. 73 (2014) 76–100, 2014. https://doi .org /10 .1016 /j .mechmat .2014 .02 .004.

[25] B. Halphen, Q. Son Nguyen. Sur les matériaux standard généralisés.J. Méc. 14 (1) (1975) 39–63,1975

[26] F. Fritzen, M. Leuschner.Reduced basis hybrid computational homogenization based on a mixed incremental formulation.Comput. Methods Appl. Mech. Eng. 260 (2013) 143–154

[27] J.-C. Michel, P. Suquet. A model-reduction approach in micromechanics of materials preserving the variational structure of constitutive relations.J. Mech. Phys. Solids 90 (2016) 254–285

[28] P. Ponte-Castaneda. Exact second-order estimates for the effective mechanical properties of nonlinear composite materials. J. Mech. Phys. Solids 44 (6) (1996) 827–862

[29] S. Roussette, J.-C. Michel, P. Suquet. Nonuniform transformation field analysis of elastic–viscoplastic composites. Compos. Sci. Technol.69 (1) (2009) 22–27

[30] R. Largenton, J.-C. Michel, P. Suquet. Extension of the nonuniform transformation field analysis to linear viscoelastic composites in the presence of aging and swelling Mech. Mater. 73 (2014) 76–100

[31] J.-C. Michel, P. Suquet. A model-reduction approach in micromechanics of materials preserving the variational structure of constitutive relations. J. Mech. Phys. Solids 90 (2016) 254–285, 2016

[32] J.-C. Michel, P. Suquet, Effective potentials in nonlinear polycrystals and quadrature formulae, Proc. R. Soc. A, Math. Phys. Eng. Sci. 473 (2017) 2204–2213

[33] M. Raissi, P. Perdikaris, and G. E. Karniadakis, 'Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations', *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb. 2019, doi: 10.1016/j.jcp.2018.10.045;

[34] X. Meng, Z. Li, D. Zhang, and G. E. Karniadakis, 'PPINN: Parareal physics-informed neural network for time-dependent PDEs', *Computer Methods in Applied Mechanics and Engineering*, vol. 370, p. 113250, Oct. 2020, doi: 10.1016/j.cma.2020.113250

[35] S. Lin and Y. Chen, 'Physics-informed neural network methods based on Miura transformations and discovery of new localized wave solutions', *Physica D: Nonlinear Phenomena*, vol. 445, p. 133629, Mar. 2023, doi: 10.1016/j.physd.2022.133629

[36] S. Gonzalez and R. Miikkulainen, 'Improved Training Speed, Accuracy, and Data Utilization Through Loss Function Optimization', in 2020 IEEE Congress on Evolutionary Computation (CEC), Jul. 2020, pp. 1–8. doi: 10.1109/CEC48606.2020.9185777.].

[37] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, 'Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What's Next', J Sci Comput, vol. 92, no. 3, p. 88, Jul. 2022, doi: 10.1007/s10915-022-01939-z

[38] F. S. Mjalli, S. Al-Asheh, and H. E. Alfadala, 'Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance', Journal of Environmental Management, vol. 83, no. 3, pp. 329–338, May 2007, doi: 10.1016/j.jenvman.2006.03.004.;

[39] Z. Z. Latt, 'Application of Feedforward Artificial Neural Network in Muskingum Flood Routing: a Black-Box Forecasting Approach for a Natural River System', Water Resour Manage, vol. 29, no. 14, pp. 4995–5014, Nov. 2015, doi: 10.1007/s11269-015-1100-1.

[40] K. Ahmed and P. Yan, 'Modeling and Identification of Rate Dependent Hysteresis in Piezoelectric Actuated Nano-Stage: A Gray Box Neural Network Based Approach', IEEE Access, vol. 9, pp. 65440–65448, 2021, doi: 10.1109/ACCESS.2021.3076403.;

[41] Z. Cen, J. Wei, and R. Jiang, 'A gray-box neural network-based model identification and fault estimation scheme for nonlinear dynamic systems', Int. J. Neur. Syst., vol. 23, no. 06, p. 1350025, Dec. 2013, doi: 10.1142/S0129065713500251.

[42] A. de Myttenaere, B. Golden, B. Le Grand, and F. Rossi, 'Mean Absolute Percentage Error for regression models', Neurocomputing, vol. 192, pp. 38–48, Jun. 2016, doi: 10.1016/j.neucom.2015.12.114;

[43] A. Rao, H. Gao, and F. Ma, 'Study of laminar burning speed and calibration coefficients of quasi-dimensional combustion model for hydrogen enriched compressed natural gas fueled internal combustion engine along with exhaust gas recirculation', Fuel, vol. 283, p. 119284, Jan. 2021, doi: 10.1016/j.fuel.2020.119284;

[44] A. G. Abdul Jameel, A. Al-Muslem, N. Ahmad, A. B. S. Alquaity, U. Zahid, and U. Ahmed, 'Predicting Enthalpy of Combustion Using Machine Learning', Processes, vol. 10, no. 11, Art. no. 11, Nov. 2022, doi: 10.3390/pr10112384

[45] Lassmann, K. TRANSURANUS: a fuel rod analysis code ready for use. Nuclear Materials for Fission Reactors, 295-302, 1992

[46] D. Pizzocri, T. Barani, L. Luzzi. SCIANTIX: A new open source multi-scale code for fission gas behaviour modelling designed for nuclear fuel performance codesJournal of Nuclear Materials 532,2020

[47] Scolaro, A.; Clifford, I.; Fiorina, C.; Pautz, A.The OFFBEAT multi-dimensional fuel behavior solver.Nuclear Engineering and Design 358 2020

[48] M. Oguma, "Cracking and Relocation Behavior of Nuclear Fuel Pellets During Rise to Power", Nuclear Engineering and Design, 76: p. 35-45, 1983

[49] K. J. Geelhood, W.G. Luscher, P. Raynaud, "FRAPCON-3.5: A Computer Code for the Calculation of Steady-State", Thermal-Mechanical Behavior of Oxide Fuel Rods for High Burnup 2014, NUREG/CR-7022, Volume 1, Revision 1

[50] D. Pizzocri et al., "Report describing the complete inert gas behaviour model with grain re-crystallization and the corresponding improved data", in INSPYRE project deliverables, M. Bertolus, Editor. 2020: Brussels, Belgium

[51] M. Zahoor, A. Casagranda, "Fuel relocation recovery implementation in Bison", Journal of Nuclear Materials, 511: p. 473-479, 2018

[52] Y. Deng et al., "Investigation on fuel cracking induced relocation behavior of dual-cooled annular fuel based on finite element simulation", Journal of Nuclear Materials, 567: p. 153779, 2022.

[53] J. Klouzal, "Preparation of the TRANSURANUS code for TEMELÍN NPP", in 9th International Conference on WWER Fuel Performance, Modelling and Experimental Support, Bulgaria 2011

[54] Rasmussen, C.E., Williams, C.K.I. (2006). Gaussian Processes for Machine Learning. MIT Press

[55] Duvenaud, D. (2014). The Kernel Cookbook: Advice on Covariance functions

[56] Wu, L., Weng, J. (2004). Probability estimates for multi-class classification by pairwise coupling. JMLR, 5, 975-1005

[57] Chang, C.C., Lin, C.J. LIBSVM: A Library for Support Vector Machines

[58] Neural Networks and Deep Learning. https://link.springer.com/book/10.1007/978-3-031-29642-0. Accessed 12 July 2024

[59] Xiu, D., Karniadakis, G.E. (2002). The Wiener--Askey Polynomial Chaos for Stochastic Differential Equations. SIAM Journal on Scientific Computing, 24(2), 619–644. https://doi.org/10.1137/S1064827501387826

[60] Bayesian Networks [Book]. https://www.oreilly.com/library/view/bayesian-networks/9781482225587/. Accessed 12 July 2024

[61] Ensemble Methods - Python: Real World Machine Learning [Book]. https://www.oreilly.com/library/view/python-real-world/9781787123212/ch20.html. Accessed 12 July 2024

[62] Myers, D.E. (1984). Co-Kriging — New Developments. Geostatistics for Natural Resources Characterization: Part 1, edited by Georges Verly et al., Springer Netherlands, pp. 295–305. https://doi.org/10.1007/978-94-009-3699-7_18

[63] Sparse Grids and Applications - Munich 2018. https://link.springer.com/book/10.1007/978-3-030-81362-8. Accessed 12 July 2024

[64] Danowsky, B.P., Chrstos, J.R., Klyde, D.H., Farhat, C., Brenner, M. (2010). Evaluation of Aeroelastic Uncertainty Analysis Methods. J. Aircr, 47, 1266–1273. https://doi.org/10.2514/1.47118

[65] Fouet, F., Probst, P. (2012). Sensitivity Analysis by the Use of a Neural Network Model during Large Break. LOCA on LOFT L2-5 Experiment with CATHARE-2 V2.5 Code, in: Proceedings of ICAPP '12. Presented at the ICAPP, Chicago, USA, p. 11045.

[66] Iooss, B., Van Dorpe, F., Devictor, N. (2006). Response surfaces and sensitivity analyses for an environmental model of dose calculations. Reliab. Eng. Syst. Saf, 91, 1241–1251. https://doi.org/10.1016/j.ress.2005.11.021

[67] IAEA (2008). Best estimate safety analysis for nuclear power plants: uncertainty evaluation. International Atomic Energy Agency, Vienna

[68] Feria, F., Herranz, L.E. (2015). Internal pressure of spent PWR fuel rod at high burnup: prediction enhancement through FRAPCON-3.5 uncertainty analysis. Top Fuel 2015

[69] Geelhood, K., Luscher, W., Raynaud, P., Porter, I. (2015). FRAPCON-4.0: A computer code for the calculation of steady-state, thermal-mechanical behavior of oxide fuel rods for high burnup. PNNL-19418, Vol.1 Rev.2.

[70] Kerrigan, J.D., Coleman, D.R. (1979). Application of the response surface method of uncertainty analysis to establish distributions of FRAP-S3 calculated stored energy for PWR-type fuels. Nucl. Eng. Des, 54, 211–224. https://doi.org/10.1016/0029-5493(79)90168-7

[71] O'Donnel et al.. Comparative Analysis of Fuel Designs. NUREG-1754 (2001)

[72] Ibrahim, H.M., Elkhidir, E.E. (2011). Response Surface Method as an Efficient Tool for Medium Optimisation. Trends Appl. Sci. Res, 6, 121–129. https://doi.org/10.3923/tasr.2011.121.129

[73] Kazeminejad, H. (2007). Uncertainty and sensitivity analyses for steady-state thermal–hydraulics of research reactors. Prog. Nucl. Energy, 49, 313–322. https://doi.org/10.1016/j.pnucene.2007.01.005

[74] Geelhood, K., Luscher, W., Raynaud, P., Porter, I. (2015). FRAPCON-4.0: A computer code for the calculation of steady-state, thermal-mechanical behavior of oxide fuel rods for high burnup. PNNL-19418, Vol.1 Rev.2.

[75] Montgomery, D.C. Design and analysis of experiments, Eighth edition. ed. John Wiley & Sons, Inc, Hoboken, NJ. (2013).

[76]

[77] C.-C. Chang and C.-J. Lin, 'LIBSVM: A library for support vector machines', ACM Trans. Intell. Syst. Technol., vol. 2, no. 3, pp. 1–27, Apr. 2011, doi: 10.1145/1961189.1961199.

[78] B. Ebiwonjumi, A. Cherezov, S. Dzianisau, and D. Lee, 'Machine learning of LWR spent nuclear fuel assembly decay heat measurements', Nuclear Engineering and Technology, vol. 53, no. 11, pp. 3563–3579, Nov. 2021, doi: 10.1016/j.net.2021.05.037.

[79] IAEA, 'Progress on Pellet-Cladding Interaction and Stress Corrosion Cracking', International Atomic Energy Agency, Vienna, IAEA-TECDOC-1960, 2021.

[80] T. Ikonen, H. Loukusa, E. Syrjälahti, V. Valtavirta, J. Leppänen, and V. Tulkki, 'Module for thermomechanical modeling of LWR fuel in multiphysics simulations', Annals of Nuclear Energy, vol. 84, pp. 111–121, Oct. 2015, doi: 10.1016/j.anucene.2014.11.004.

[81] T. Ikonen, E. Syrjälahti, V. Valtavirta, H. Loukusa, J. Leppänen, and V. Tulkki, 'Multiphysics simulation of fast transients with the FINIX fuel behaviour module', EPJ Nuclear Sci. Technol., vol. 2, p. 37, 2016, doi: 10.1051/epjn/2016032.

[82] JRC, 'TRANSURANUS Handbook', V1M1J17', 2017.

[83] X. Li, C. Allison, and J. Hohorst, Eds., Nuclear power plant design and analysis codes: development, validation, and application. in Woodhead Publishing series in energy. Duxford, UK: Woodhead Publishing an imprint of Elsevier, 2020.

[84] H. Loukusa, J. Peltonen, and V. Valtavirta, 'FINIX - Fuel behavior model and interface for multiphysics applications - Code documentation for version 1.19.12', VTT-R-01103-19, 2019.

[85] R. F. Mattas, F. L. Yaggee, and L. A. Neimark, 'Iodine stress-corrosion cracking in irradiated Zircaloy cladding', presented at the ANS topical meeting on light water reactor fuel performance, Portland, OR, 1979.

[86] B. Michel, J. Sercombe, C. Nonon, and O. Fandeur, 'Modeling of Pellet Cladding Interaction', in Comprehensive Nuclear Materials, Elsevier, 2012, pp. 677–712. doi: 10.1016/B978-0-08-056033-5.00074-4.

[87] B. Michel, J. Sercombe, and G. Thouvenin, 'A new phenomenological criterion for pellet–cladding interaction rupture', Nuclear Engineering and Design, vol. 238, no. 7, pp. 1612–1628, Jul. 2008, doi: 10.1016/j.nucengdes.2008.01.012.

[88] K. P. Murphy, Machine learning: a probabilistic perspective. in Adaptive computation and machine learning series. Cambridge, MA: MIT Press, 2012.

[89] F. Stulp and O. Sigaud, 'Many regression algorithms, one unified model: A review', Neural Networks, vol. 69, pp. 60–79, Sep. 2015, doi: 10.1016/j.neunet.2015.05.005.

[90] M. Di Gennaro, Reduced order modelling of fission gas diffusion in fuel performance codes; A. Pagani, A model for athermal fission gas release in SCIANTIX;

[91] A. Djonovic, Modelling the evolution of point and extended defects in UO2 under irradiation] and two Master of Science thesis projects at Polimi, 2023

[92] A. Djonovic, 2023, Modelling the evolution of dislocation density in UO2 under irradiation; A. Pagani, 2023, An AI-Enhanced Model of Athermal Fission Gas Release in SCIANTIX.

[93] L. Bernard, J. Jacoud, P. Vesco. An efficient model for the analysis of fission gas release. Journal of Nuclear Materials, 125-134, 2002

[94] Claisse, A.; Van Uffelen, P. Towards the inclusion of open fabrication porosity in a fission gas release model. Journal of Nuclear Materials 351-356, 2015

[95] D. Pizzocri, M. Di Gennaro, T. Barani, F.A.B. Silva, G. Zullo, S. Lorenzi, A. Cammi. A reduced order model for fission gas diffusion in columnar grains.Nuclear Engineering and Technology,3983-3995, 2023