

POLITECNICO DI TORINO

Master of Science in Energy and Nuclear Engineering
Joint M.Sc. Program in Nuclear Engineering with Politecnico di Milano

Master Thesis

An Efficient Metamodel-based Exploration Framework for Characterizing the Critical Failure Regions of a Nuclear Passive Safety System



Supervisors:

Prof. Pedroni Nicola
Prof. Bertani Cristina
Dr. Bersano Andrea
Prof. Di Maio Francesco (PoliMi)
Prof. Zio Enrico (PoliMi)

Candidate:

Lorenzo Puppo

Abstract

Passive Safety Systems (PSSs) are increasingly employed in advanced Nuclear Power Plants (NPPs) since they are considered, in general, more reliable than active systems. Their safety performance is evaluated through computationally expensive Thermal-Hydraulic (T-H) simulations models and the identification of the operational conditions which lead to unsafe conditions (the so-called Critical failure Regions, CRs) may be challenging.

In this view, the computationally expensive T-H models simulating the PSSs behavior can be replaced by fast-running surrogate models (also called metamodels), coupled with adaptive sampling techniques for speeding up the exploration and efficiently focusing the analysis on the most interesting regions of the domain, i.e., the CRs boundary (limit surface). However, when the PSS state-space also shows a non-smooth and/or multimodal nature, even the previously mentioned metamodel-based approaches may not suffice. In such cases, suitable techniques, like Finite Mixture Models (FMMs) or clustering methods, should be sought and effectively combined to tackle these issues.

In the present thesis, a passive Decay Heat Removal (DHR) system of a NPP is considered and its CRs are characterized with respect to two safety-critical variables of interest (used to evaluate the success of the PSS operation): the amount of energy exchanged by the PSS and the maximum pressure value reached inside the pressure vessel. A time-demanding Best-Estimate Thermal-Hydraulic (BE-TH) code is employed to simulate the PSS operation.

In the analysis of the energy exchanged, which shows a smooth trend, the well-known Adaptive Kriging Monte Carlo Sampling (AK-MCS) is employed. This methodology is based on a fast-running Kriging regression model, iteratively constructed and progressively refined in proximity of the PSS CR boundary by means of an adaptive sampling technique.

Abstract

To the best of the author's knowledge this is the first time that the AK-MCS technique is implemented for the identification of the CRs of a PSS of an NPP. The results show that a satisfactory level of accuracy (estimation error around 2%) can be reached with less than 200 BE-TH simulations. Thus, the Kriging metamodel can be exploited to accurately predict the outcomes of many new PSS configurations (several thousands) in few minutes, with the aim of finding the critical ones for the CRs characterization, instead of directly using the BE-TH which takes some hours for each simulation.

On the other side, in the analysis of the maximum vessel pressure, which shows a non-smooth and multimodal behavior, a novel methodological framework is proposed, combining Finite Mixture Models (FMMs) and AK-MCS. In particular, 1) FMMs are employed to reduce the dimensionality of the non-smooth and multimodal PSS state-space, while identifying "prototypical clusters" of PSS functional failure configurations; 2) a metamodel (namely, AK-MCS) is adaptively trained on the reduced space to mimic the time-demanding T-H model; and, eventually, 3) the trained metamodel is used to evaluate new PSS configurations and retrieve information about CRs. Finally, a comparison with an alternative approach of literature based on the use of a classifier to cluster the output domain is presented to support the framework as a valid approach in challenging CRs characterization. The results show that the FMM-based framework allows overcoming the issue of PSS state-space non-smoothness and multimodality, indeed, a satisfactory metamodel accuracy (estimation error $< 0.5\%$) is reached with only 300 BE-TH simulations. Moreover, the proposed framework always outperforms the alternative technique with the classifier if an equal number of BE-TH simulations is used. In the end, the Kriging metamodel predictions, which take few seconds, are adopted, again, to speed up the calculation for the CRs characterization.

Contents

List of Figures	VI
List of Tables	VIII
List of Acronyms	IX
List of Symbols	XII
1 Introduction.....	1
1.1 Nuclear Passive Safety Systems.....	1
1.2 The Problem: Nuclear Passive Safety System Critical Failure Regions Characterization	6
1.3 Advanced Computational Methods for Efficient Critical Regions Identification ..	7
1.3.1 Methods for the Reduction of the Computational Cost	7
1.3.1.1 Advanced Sampling Strategies.....	7
1.3.1.2 Metamodeling.....	10
1.3.2 Methods for Tackling the Output Non-smoothness and Multimodality	13
1.3.2.1 Clustering the Output Space	13
1.3.2.2 Dimensionality Reduction	15
1.4 Methodological and Applicative Contribution of the Present Thesis.....	17
2 Case Study	19
2.1 Description of the Passive Safety System and its Failure Criteria	22
2.2 Synthetic Description of the PSS T-H model.....	26
2.3 Description of the Two Outputs Behavior.....	27

Contents

2.3.1	Energy Output Behavior	28
2.3.2	Pressure Output Behavior.....	28
2.4	Other Possible Relevant Parameters	30
3	Metamodel-based AK-MCS Framework for the CRs Exploration of a PSS	31
3.1	AK-MCS Framework Presentation.....	32
3.1.1	Kriging Metamodel	32
3.1.2	Metamodel-based AK-MCS framework for CRs exploration	34
3.1.3	Metamodel Accuracy Evaluation and CRs Characterization	35
3.2	AK-MCS Framework Application for CRs characterization of the PSS considered – Energy Output.....	37
3.2.1	I/O Training Set and Kriging Metamodel Construction	37
3.2.2	Metamodel-based AK-MCS Application.....	39
3.2.3	Metamodel Accuracy Evaluation and CR Characterization	43
3.2.3.1	Kriging Accuracy Evaluation	44
3.2.3.2	CRs Characterization with respect to the FC “Low Heat Removal”	48
4	Novel Framework for System CRs Characterization in case of State-space Non-smoothness and Multimodality	54
4.1	Novel Framework Presentation	57
4.1.1	Dimensionality Reduction.....	57
4.1.1.1	Sensitivity Analysis.....	57
4.1.1.2	Finite Mixture Models Approximation	59
4.1.2	Iterative Metamodel Training (AK-MCS)	62
4.1.3	CRs Representation & Information Retrieval.....	63
4.2	Novel Framework Application for the CRs Characterization of the PSS considered – Pressure Output.....	64
4.2.1	Dimensionality Reduction.....	64
4.2.2	Iterative Metamodel Training (AK-MCS)	68
4.2.3	CRs Representation & Information Retrieval.....	72

Contents

4.2.3.1 Metamodel Accuracy Evaluation	72
4.2.3.2 CR Characterization	73
4.2.4 Comparison with the Results obtained with SVC + AK-MCS procedure.....	75
5 Conclusion.....	81
References.....	86

List of Figures

<i>Figure 1: examples of searching techniques for adaptive sampling [20]</i>	9
<i>Figure 2: a set of data points (DoE) evaluated by the simulator with the corresponding outputs; a metamodel will be fitted to this DoE [20]</i>	11
<i>Figure 3: Two-stage surrogate modelling technique [18]</i>	15
<i>Figure 4: PRHR heat exchanger system based on a natural circulation loop and designed for advanced PWRs [2]</i>	20
<i>Figure 5: IC heat exchanger system based on a condensation-driven loop designed for advanced BWRs [2]</i>	21
<i>Figure 6: Simplified sketch of the PSS under study</i>	23
<i>Figure 7: Eex, % regular distribution</i>	28
<i>Figure 8: pmax non-smooth and multimodal distribution</i>	29
<i>Figure 9: Identification of the best candidates (Step 5): random selection of input combinations with $U < 1.96$ divided in equally-spaced bins</i>	43
<i>Figure 10: RMSE evaluated with respect to two validation sets (energy output)</i>	47
<i>Figure 11: Q1 (a) and Q2 (b) predictivity indicators evaluated with respect to the two validation sets (energy output)</i>	47
<i>Figure 12: Scatter plots of the PSS CR (energy output)</i>	49
<i>Figure 13: PCP with 0.25 quantiles (energy output)</i>	52
<i>Figure 14: flow diagram of the exploration framework proposed for tackling the output non-smoothness and multimodality</i>	55
<i>Figure 15: pmax clustering according to $k = 3$ Gaussian distributions</i>	66

List of Figures

<i>Figure 16: Hellinger distance for each input parameter (x_m) evaluated with respect to high-pressure cluster</i>	68
<i>Figure 17: Bins selection technique (see Section 3.2.2) applied also for the AK-MCS procedure referred to pmax output</i>	71
<i>Figure 18: QIs evaluated with respect to a given validation set (pressure output)</i>	73
<i>Figure 19: CR for pmax output</i>	74
<i>Figure 20: QIs evolution in FMM+AK-MCS strategy compared with the QIs values at the end of SVC+AK-MCS (pressure output)</i>	80

List of Tables

<i>Table 1: Comparison between different SA methods to tackle the state-space non-smoothness or multimodality</i>	16
<i>Table 2: Range of variation of the inputs</i>	25
<i>Table 3: I/O reference conditions</i>	25
<i>Table 4: Average CV error in trend type estimation (energy output)</i>	38
<i>Table 5: Average CV error in correlation function family estimation (energy output)</i>	38
<i>Table 6: results of the AK-MCS procedure applied to the energy output</i>	44
<i>Table 7: QIs at the end of AK-MCS procedure - 14th iteration (energy output)</i>	47
<i>Table 8: FMM components parameters</i>	65
<i>Table 9: Average CV error in trend type estimation (with Matérn 5/2 corr. function family)</i>	69
<i>Table 10: Average CV error in trend type estimation (with Exponential corr. function family)</i>	69
<i>Table 11: Average CV error in trend type estimation (with Exponential corr. function family)</i>	70
<i>Table 12: Average CV error in correlation function family estimation (with Ordinary trend)</i>	70
<i>Table 13: QIs values at the end of AK-MCS procedure – 10th iteration (pressure output)</i>	73
<i>Table 14: QIs at the end of SVC+AK-MCS iterative procedure (pressure output)</i>	80

List of Acronyms

AIC	Akaike Information Criterion
AK-MCS	Adaptive Kriging Monte Carlo Sampling
ANN	Artificial Neural Network
AV	Activation Valve
BE-TH	Best-Estimate Thermal-Hydraulic
BIC	Bayes Information Criterion
BWR	Boiling Water Reactor
CEA	Commissariat à l'Energie Atomique
CR	Critical Region
CV	Cross Validation
DBSA	Distribution-Based Sensitivity Analysis
DHR	Decay Heat Removal
DoE	Design of Experiments
E-HX	Emergency Heat eXchanger
EFF	Expected Feasibility Function
EM	Expectation Minimization
ENEA	Energia Nucleare ed Energie Alternative
FC	Failure Criterion
FMM	Finite Mixture Models
GA	Genetic Algorithm

List of Acronyms

GP	Gaussian Process
I/O	Input/Output
IAEA	International Atomic Energy Agency
IC	Isolation Condenser
LHS	Latin Hypercube Sampling
LOO	Leave-One-Out
MARS	Multivariate Adaptive Regression Splines
MCS	Monte Carlo Sampling
ML	Maximum Likelihood
MML	Minimum Message Length
MSIV	Main Steam Isolation Valve
NPP	Nuclear Power Plant
NRMSE	Normalized Root-Mean-Square Error
PCC	Partial Correlation Coefficients
PDF	Probability Density Function
PCP	Parallel Coordinates Plot
PNN	Polynomial Neural Network
PRHR	Passive Residual Heat Removal
PSS	Passive Safety System
PV	Pressure Vessel
PWR	Pressurized Water Reactor
QI	Quality Indicator
RBF	Radial Basis Function
RBSA	Regression-Based Sensitivity Analysis
RELAP	Reactor Excursion and Leak Analysis Program
REPAS	Reliability Evaluation of Passive Safety Systems
RMSE	Root-Mean-Square Error
SA	Sensitivity Analysis

List of Acronyms

SBO	Station Black-Out
SBWR	Simplified Boiling Water Reactor
SPLOM	Scatter PLOt Matrix
SRC	Standardized Regression Coefficients
SRV	Safety Relief Valve
SVC	Support Vector Classifier
SVR	Support Vector Regression
T-H	Thermal-Hydraulic
TM	Target Mission
TPI	Transient Performance Indicator
VBSA	Variance-Based Sensitivity Analysis

List of Symbols

A_{AV}	Activation Valve flow area
A_{MSIV}	Main Steam Isolation Valve flow area
α	Lagrange multiplier
β	Trend coefficients of Kriging approximation
$\hat{\beta}$	Trend coefficients least square estimates
C	Constant
D_X	Input domain
D_Y	Output domain
DEL_{AV}	Delay of Activation Valve opening
DEL_{MSIV}	Delay of Main Steam Isolation Valve closure
d	Distance
E	Expected value
E_{ex}	Amount of energy exchanged by the Passive Safety System
$E_{ex,\%}$	Percentage of energy exchanged by the Passive Safety System
ϵLOO_{abs}	Absolute Leave-One-Out error
ϵLOO_{norm}	Normalized Leave-One-Out error
F	Functions space
f	Generic model function
f^{-1}	Inverse model function
\tilde{f}	Approximated model function
Φ	Mapping function for the Support Vector Classifier construction
H	Information matrix of Kriging metamodel

List of Symbols

\mathbf{h}	Kriging metamodel trend
h	Arbitrary function of the Kriging trend
H_{jm}	Hellinger distance measured for the m -th input with respect to the j -th cluster
I	Encoding of information for the for the Minimum Message Length calculation
i	Input combination index
j	Cluster index
K	Number of partitions of a set of samples
k	Number of components of the Finite Mixture Models approximation
ker	Kernel
ℓ	Classifier label
M	Problem original dimensionality
m	Input variable index
μ	Mean value of a probability distribution
$\mu_{\hat{y}}$	Mean value of a metamodel prediction
\mathcal{N}	Normal Gaussian distribution
N	Number of elements
N_{cand}	Number of candidates selected at each iteration of the AK-MCS procedure
N_{MCS}	Number of samples generated by Monte Carlo Sampling
N_{train}	Number of training samples
N_{val}	Number of validation samples
$NC\%$	Non-condensable gases percentage
n	Iteration number
n_{fin}	Final number of iterations
ξ	Slack term of the Support Vector Classifier construction process
O	Upper Asymptotic Limit
P	Number of arbitrary functions of the Kriging trend
p	Probability distribution

List of Symbols

p_{max}	Maximum value of pressure inside the Pressure Vessel of the Passive Safety System
Pr	Probability
π	Set of weights in the Finite Mixture Models approximation
π	Probability distribution weight in the Finite Mixture Models approximation
$\hat{\pi}$	Estimate of the weight of a probability distribution
Q	Q-function in the Finite Mixture Models approximation
Q	Predictivity indicator
\mathcal{R}	Correlation matrix of the Kriging metamodel
R	Problem reduced dimensionality
\mathbf{r}	Vector of cross correlations between input vectors in the Kriging construction process
r	Generic element of the cross correlations vector
$sign$	Sign function
σ	Standard deviation
$\bar{\sigma}$	Average standard deviation
$\sigma_{\hat{y}}$	Estimation error of a given metamodel prediction
t	Index of the mixture parameters estimate
Θ	Set of mixture parameters
$\hat{\Theta}$	Estimate of the mixture parameters
Θ	Probability distribution parameters
$\hat{\Theta}$	Estimate of the probability distribution parameters
θ	Kriging metamodel hyperparameters
U	U learning function
Var	Variance
W	Conditional expectation of the set of Finite Mixture Models labels
\mathbf{w}	Vector of the Support Vector Classifier hyperplane coefficients
w	Conditional expectation of the
\mathbf{X}	Generic input vector
\mathbf{X}^R	Generic input vector of reduced dimensionality

List of Symbols

\mathcal{X}	Set of model input vectors
\mathcal{X}^*	Input vectors of the set of best candidates selected during AK-MCS procedure
\mathcal{X}^{Krig}	Set of input vectors to be evaluated with Kriging metamodel after their classification
\mathcal{X}_{train}	Set of training input vectors
$\mathcal{X}_{train}^{SVC}$	Set of training input vectors for the Support Vector Classifier
\mathcal{X}_{val}	Set of validation input vectors
\mathbf{x}	Model input vector
x	Model input parameter
Y	Generic output
Y_{thres}	Threshold output value
\mathbf{y}^*	Outputs of of the set of best candidates selected during AK-MCS procedure
\mathcal{Y}_{train}	Set of training outputs
$\mathcal{Y}_{train}^{FMM}$	Set of training outputs for the Finite Mixture Models approximation
$\mathcal{Y}_{train}^{SVC}$	Set of training outputs for the Support Vector Classifier construction
\mathcal{Y}_{val}	Set of validation outputs
$\hat{\mathcal{Y}}$	Set of metamodel predictions
$\hat{\mathcal{Y}}_{val}$	Set of predictions of the validation outputs
y	Model output
\hat{y}	Metamodel prediction output
\hat{y}^{SVC}	Support Vector Classifier prediction output
\bar{y}_{val}	Average validation output value
Z	Set of Finite Mixture Models labels
Z	Unit variance stationary Gaussian process of Kriging metamodel construction
\mathbf{z}	Finite Mixture Models labels vector
z	Component of the Finite Mixture Models labels vector

1 Introduction

The objective of the present Master Thesis is to effectively address the problem of identifying and characterizing the Critical failure Regions (CRs) of Passive Safety Systems (PSSs) designed for advanced Nuclear Power Plants (NPPs). The work has been carried out within a cooperation between Polytechnic of Turin and Polytechnic of Milan (namely, Poly2Nuc), involving Professors Nicola Pedroni and Cristina Bertani (Turin), Dr. Andrea Bersano (Turin) and Professors Francesco di Maio and Enrico Zio (Milan).

The remainder of the Introduction is organized as follows: in [Section 1.1](#), an exhaustive description of nuclear PSSs is offered together with their strengths and weaknesses; in [Section 1.2](#), the issues related to PSS CRs characterization are discussed in detail; [Section 1.3](#) presents several advanced computational methods available in literature for CRs exploration; finally, [Section 1.4](#) explains how some of these innovative techniques are implemented, modified and originally combined in this work for the efficient CR characterization of a generic PSS of a NPP designed for Decay Heat Removal (DHR).

1.1 Nuclear Passive Safety Systems

In recent years, important efforts have been made by industries, research organizations and utilities for the design and development of PSSs to increase the safety level of NPPs.

The use of PSSs within the nuclear field was addressed in 1991 at the International Atomic Energy Agency (IAEA) Conference on “The Safety of Nuclear Power: Strategy for the Future” [1]. Innovative PSSs are implemented in new advanced reactors to cover the main safety functions, e.g., decay heat removal, reactivity control and fission product containment; moreover, they also eliminate the costs associated with the installation, maintenance and operation of active systems that typically require multiple pumps with independent and redundant electric power supplies. According to IAEA definition [2] the term “passive system” represents a system entirely composed of passive elements and structures or a system that uses active components in a very limited way, e.g., just to initiate subsequent passive actions [3]. Passive components do not rely on external inputs (e.g., signal, human action, forces or power) to work; indeed, they just exploit natural physical laws (e.g., gravity, heat conduction, natural convection etc.) and/or inherent characteristics of such systems (e.g., internally stored energy, properties of materials etc.) and/or energy inherently available in such systems (e.g., decay heat, chemical energy etc.) [4]. When deliberating over the differences between active and passive systems, it was realized that, within these two general categories, a spectrum of possibilities exists. Indeed, in [1] four categories are established to distinguish the different degrees of passivity, which are reported in what follows.

Category A:

- no signal inputs of “intelligence”,
- no external power sources or forces,
- no moving mechanical parts and
- no moving working fluid.

Physical barriers against the release of fission products (e.g., nuclear fuel cladding), hardened building structures for the protection against seismic or external events and static components of safety related PSSs (e.g., tubes, accumulators, surge tanks) all belong to this category.

Category B:

- no signal inputs of “intelligence”,
- no external power sources or forces,
- no moving mechanical parts, but
- moving working fluid.

Examples of safety features included in this category are reactor emergency cooling systems based on the natural circulation of air or water in heat exchangers immersed in water pools to which the decay heat is discharged or containment cooling systems based on natural circulation of air flowing around the containment walls.

Category C:

- no signal inputs of “intelligence”,
- no external power sources or forces, but
- moving mechanical parts, whether or not moving working fluids are also present.

Example of safety features belonging to this category are emergency injection systems consisting of accumulators or storage tanks and discharge lines equipped with check valves; overpressure protection and/or emergency cooling devices of pressure boundary systems based on release of fluid through safety relief valves; filtered venting systems of containment building triggered by rupture disks.

Category D:

- signal inputs of “intelligence” to initiate the passive process;
- energy to initiate the process must be provided by stored sources (e.g., batteries or elevated fluids);
- active components are limited to controls, instrumentation and valves to initiate the passive system;
- manual initiation is excluded.

Examples of safety features included in this category are emergency core cooling and injection systems based on gravity that are initiated by battery-powered electric or electro-pneumatic valves; emergency reactor shutdown systems based on gravity driven control rods. This last category addresses the intermediate zone between active and passive systems, indeed the safety function is executed through passive methods, but, since internal energy is not available to initiate the process, the triggering is provided by an external signal. To sum up, the spectrum of possibilities from passive to active systems may have even more categories, but the essential point is that the more self-contained and self-supported a system, the higher its degree of passivity.

PSSs are generally considered as more reliable than active systems due to the simplified architecture (fewer pieces of equipment, less dependence on external power sources, lower need for human intervention). Moreover, the determination of the failure conditions of PSSs (which are mainly composed of static components) may seem less complex than for active systems (which instead typically comprise a higher number of components) [5]. However, other relevant issues arise when adopting PSSs. For example, there is the need to demonstrate the understanding of the key Thermal-Hydraulic (T-H) phenomena at the basis of the PSSs operation: this implies the identification of physical parameter ranges, the availability of proper experimental programs and the demonstration of predictive capabilities for the computational tools. Within this framework, some research activities comprehensive of experiments and code development have been carried out in the last decades, mainly addressed to existing nuclear reactors; however, they appear to be limited and in general lacking of operational experience and structured procedures. A more systematic effort seems necessary for assessing the level of understanding of the T-H phenomena and the connected code capabilities for PSSs future applications [2]. Moreover, PSSs rely on natural forces (e.g., natural circulation) and some deviation from the design conditions may impair the performance of the system itself. This is especially remarkable for PSSs involving moving working fluids due to small engaged driving forces. The magnitude of these natural forces is, in general, relatively small and counterforces (e.g., friction) can be of comparable magnitude; thus, they may have a greater impact with respect to the case of active systems (using, e.g., electric pumps). This represents an issue

if we also consider that there are several uncertainties affecting the magnitude of these forces and counterforces (e.g., heat transfer coefficients values and pressure losses). Additionally, the dependence of the magnitude of natural driving forces on the specific plant conditions and configurations existing when the PSS is called to intervene is complex to be assessed [5], [6]. There is, thus, the need for a proper treatment of all these uncertainties, even under scarce operational data, because they strongly affect the T-H performance of PSSs. In particular, such uncertainties must be propagated through the PSS models, which typically requires a large number of calculations conducted with Best-Estimate Thermal-hydraulic (BE-TH) codes: this procedure may become prohibitive from the point of view of the computational cost. For all these reasons, the determination of the conditions leading a PSS to fail its function is a challenging task.

A pioneering activity aimed at the evaluation of the reliability of PSSs was proposed in the mid-1990s within the framework of bilateral contacts between CEA and ENEA (French and Italian research bodies). Later on, a cooperation between ENEA, University of Pisa and Polytechnic of Milan led to the proposal of a methodology called Reliability Evaluation of Passive Safety Systems (REPAS) [7], [8]. REPAS represents a structured methodology for conducting a T-H reliability assessment (i.e., the reliability concerning the occurrence of T-H phenomena, like natural circulation) of PSSs accounting for uncertainties. For example, one question that may arise dealing with a system based on natural circulation regards the possible disruption of natural circulation when the system, or the thermal-hydraulic circuit, and the corresponding geometric and material configurations do not vary. However, in its original formulation REPAS does not explicitly include the detailed characterization of the PSS CRs. This is necessary to define the configurations of critical operation for the system, i.e., those combination of values of PSS design and/or operational parameters which lead the PSS to fail providing its safety function.

1.2 The Problem: Nuclear Passive Safety System Critical Failure Regions Characterization

In mathematical terms, given the Input/Output (I/O) representation, $Y = f(\mathbf{X})$, of the PSS behaviour, a specific combination of input parameters values \mathbf{x} is critical, if the resulting output value (e.g., the decay heat removed by the PSS or the maximum of pressure inside the pressure vessel) is lower (higher) than a predefined threshold, $y = f(\mathbf{x}) \leq (\geq) Y_{thres}$, representing the limit value for the system operation. These combinations define the PSS CRs, whose identification and characterization can be addressed with computational methods: see, e.g., [9], [10], [11], [12]. In these methods, BE-TH models are not directly adopted to numerically compute the PSS response in the several accidental conditions that need to be explored due to: i) the high computational cost often related to BE-TH models (e.g., hours or even days for a single simulation), and ii) the high number of code runs typically needed for a deep exploration of the PSS state-space that could require excessive computational resources. For this reason, advanced intelligent methods should be sought and combined to address these computational issues. On one side, fast-running surrogate models may be adopted to mimic the behaviour of the computationally demanding original T-H simulator and replace it in the analysis [13]. On the other side, intelligent adaptive sampling strategies may be implemented to efficiently trace the CR boundary (i.e., the PSS limit surface), with minimum waste of computational time for samples far from the CR [14].

Beside the significant computational cost that may be associated to the CRs characterization, another relevant issue that could be faced is the possible non-smoothness and/or multimodality of the output function that must be approximated by the metamodel. In the general setting, when dealing with metamodel approximations, the computational model function that is approximated is frequently assumed to have “accommodating properties” of regularity, such as continuity and smoothness. However, several engineering problems showing non-smooth and/or multimodal functional behavior can be found, e.g.,

in structural and mechanical engineering phenomena like snap-through, buckling or others [15] [16] [17], challenging traditional smooth metamodels, like Kriging, that might lead to large errors [18].

In view of the relevant issues mentioned above, advanced computational strategies must be sought for an efficient characterization of the CRs of PSSs.

1.3 Advanced Computational Methods for Efficient Critical Regions Identification

1.3.1 Methods for the Reduction of the Computational Cost

In the following subsections we provide an exhaustive description of the two most widespread techniques for reducing the computational cost related to PSS CRs identification in the presence of time-demanding BE-TH models: advanced sampling strategies (1.3.1.1) (e.g., adaptive sampling) to efficiently select random input configurations close to the CR limit surface; and metamodel techniques (e.g., Kriging) to replace the original long-running models (1.3.1.2).

1.3.1.1 Advanced Sampling Strategies

Sampling strategies have gained much popularity with the advent of modern Design of Experiments (DoE), which substituted monetarily expensive and time-demanding physical experiments with cheaper and faster computer simulations. They are historically divided into two categories, namely *static* and *adaptive*. Static sampling represents the earliest kind of designs. It ignores the system under study and focuses only on the spatial distribution of samples in the domain (space-filling). Such techniques are also known as one-shot sampling since the number of samples is decided one time above all according to the information available before the first simulation and, then, the simulations are carried out through one single session. In this way, many simulations can be run in parallel on different

computer to save computational time [14]. The most famous among static sampling strategies is the Latin Hypercube Sampling (LHS) [19] which owes its popularity due to its capability of dealing with high-dimensional problems.

Over the time, researchers began incorporating the system knowledge to tailor the DoE for the specific system under analysis, coming to the more flexible and dynamic approaches called adaptive sampling or active learning. The idea of adaptive sampling rises from the need of reducing the number of simulations with respect to the alternative sampling methods, i.e., static sampling. Because the simulator is usually considered a black box (in the present work the BE-TH model of the PSS), it is infeasible in practice to a priori predict the correct number of samples to achieve a desired accuracy; for this reason, in static sampling the number of samples is often far bigger than what is really necessary (oversampling) inducing useless computational efforts. Active learning solves this problem by selecting the samples iteratively, while, at the same time, updating the model they refer to and re-assessing the overall quality. If the target accuracy is reached, the procedure is stopped and no more samples are selected [20]. Thus, the great novelty is represented by the exploitation of the information gathered from previous samples to determine the location of new points [21]. The goal of adaptive sampling strategies becomes to find the best DoE with the smallest number of samples [22]. In general terms, samples are selected iteratively to fill the search domain (in this case, the PSS input parameters space) in such a way that any discontinuity or key features are not missed (namely, good exploration), and, at the same time, the search is focused on regions that have been identified as potentially interesting because close to the CR (namely, good exploitation). A good trade-off between exploration and exploitation is necessary, indeed, if exploitation is stressed too much the choice of the initial DoE becomes too influential on the quality of the results (e.g., the initial DoE needs to be large enough to catch all the regions of interest from the beginning, otherwise they are missed). On the other hand, if exploration searching strategy is prevalent, then the purpose of adaptive sampling is a bit lost because the responses of the already simulated samples are less relevant in the decision making process; hence, it could result oversampling in simple/linear regions and under-sampling in complex/nonlinear regions. One illustrative example is provided in [Fig. 1](#) which shows an unknown function

that needs to be approximated with some samples initially extracted (see Fig. 1(a)); only one of these initial samples suggests that the real trend is not linear as it could seem from the others. Exploration technique in Fig. 1(b) allows to discover the other non-linearities, but without properly focusing on them; on the other side, in Fig. 1(c), thanks to the exploitation technique, more attention is put where the non-linearity is individuated from the initial samples, but the other function key features are missed.

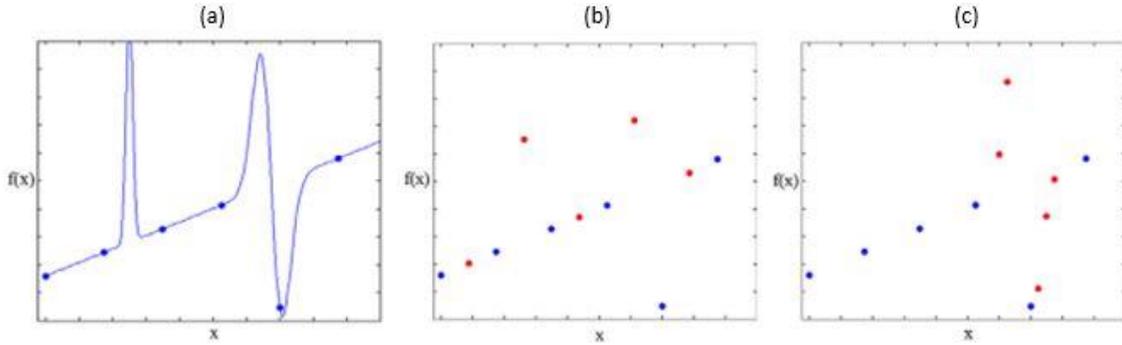


Figure 1: examples of searching techniques for adaptive sampling [20]

When speaking about adaptive sampling it is also meaningful to dwell on the role of the so-called “learning functions”, representing the criterion according to which the different searching techniques (exploration or exploitation) are put in practice. There are several examples of learning functions in literature, as described in [23]. One example is represented by the Expected Feasibility Function (EFF) [24], introduced within an efficient global reliability analysis with the aim of searching new points (i.e., system input combination) in proximity of the limit state over the entire variable space. Other learning functions have been introduced coupled with metamodel approximations, e.g., the U -function [25] or the H -function proposed by [26] (see details below).

One last element that plays an important role in determining the performance of adaptive sampling is the concept of *Granularity*, i.e., the number of new samples added at each iteration. A fine-grained method (few points added) is preferable if we want to avoid oversampling; indeed, in case of coarse granularity it might occur that too many samples are simulated at the last iteration, while instead only a few of them are really necessary to reach the desired accuracy level [20]. On the other hand, sometimes adding many samples

is favorable because of computational efficiency. For example, in case of iterative techniques based on adaptive sampling, if very few points are sampled at each iteration, small improvements typically result between one iteration and the successive one: hence, many iterations could be required, with a significant impact on the computational cost of the whole procedure [27].

1.3.1.2 Metamodeling

Beside the use of smart sampling strategies to face time-demanding computer simulations, another interesting and widespread approach is represented by the so-called surrogate models or metamodels approximations. The general idea consists in finding an approximation function that is constructed on multiple simulations (e.g., time-demanding T-H simulations) at key points of the design space (training set or DoE) and on the analysis of the outcomes of such simulations [20]. This function manages not only to mimic the results of the samples in the DoE, but also to approximately predict the (true) model output Y (e.g., the amount of heat removed by the PSS) in correspondence of other input values of the domain, depending on the quality of the training dataset.

From a mathematical perspective, the simulator function (in the present thesis, the T-H model of the PSS), which is unknown, can be defined as $f : \mathbb{R}^M \rightarrow \mathbb{R}$ (where M is the dimensionality of the problem), mapping a set (vector) of real inputs into a real output (that can be a vector likewise). f is evaluated at certain input combinations, numbered N_{train} , grouped as $\mathcal{X}_{train} = \{\mathbf{x}_1, \dots, \mathbf{x}_{N_{train}}\}$ and the corresponding output values are $\mathcal{Y}_{train} = \{f(\mathbf{x}_1), \dots, f(\mathbf{x}_{N_{train}})\} = \{y_1, \dots, y_{N_{train}}\}$. $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ constitutes the training set according to which a suitable approximation function \tilde{f} is chosen from a set of candidates \mathbf{F} (functions space), representing the different metamodeling techniques, with the scope of closely resembling the original f [20], [21], [28].

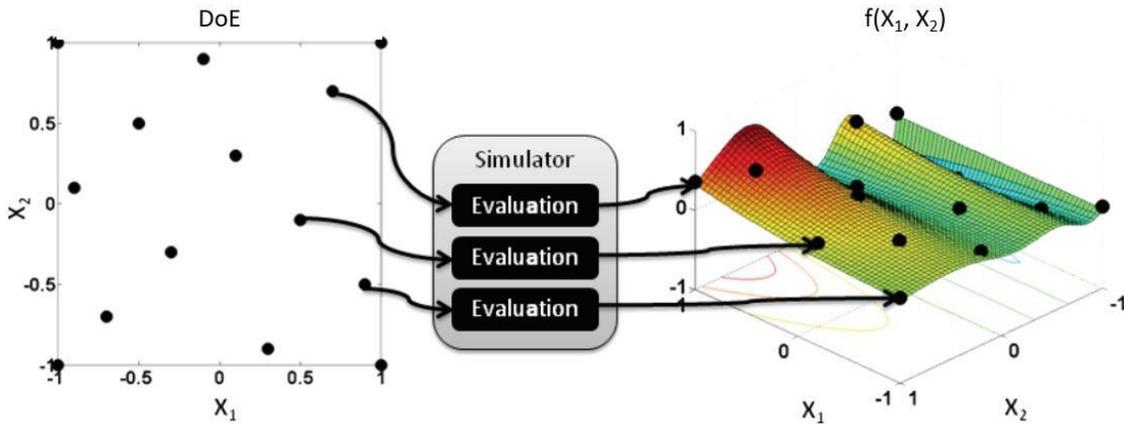


Figure 2: a set of data points (DoE) evaluated by the simulator with the corresponding outputs; a metamodel will be fitted to this DoE [20]

The quality of the approximation depends on F (ideally the simulator function would be contained in F and, hence, it could be theoretically achieved an exact approximation), but mostly on the choice of $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$. The training samples should be spread over the input space to provide the maximum amount of information about f , but, at the same time, have a density proportional to the local level of complexity/non-linearity of f . This is also the reason why metamodel approximations are often coupled with adaptive sampling to ensure such a smart distribution of training points in the domain. In this case, the choice of the initial training set, i.e., $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$, becomes fundamental; indeed, if a poor metamodel prediction is obtained at the beginning, then, this may lead to focus the adaptive sampling in inappropriate regions. Moreover, $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ size should not be too large, otherwise a significant amount of computational time is wasted simulating just space-filling samples rather than being spent more effectively on the interesting locations identified by the adaptive sampling strategy. To sum up, the choice of the most suitable training set size is challenging as it depends on many factors: the available computational budget (i.e., the maximum number of long-running model simulations initially foreseen), the complexity and dimensionality of the target function f , the features of the metamodel selected and, potentially, the coupling between metamodeling and adaptive sampling [27].

Several metamodeling techniques have been developed through the years, as described in [13]. The first historical examples regard classical DoE theory with polynomial functions;

then, from [29] onward, stochastic models gained a lot of popularity in the field of deterministic computer responses, especially one category of them, i.e., the Gaussian Processes (GPs); also Artificial Neural Networks (ANN) have become rapidly famous for many applications, as explained in [30] and [31]. Other promising statistical techniques developed during the time are: Radial Basis Functions (RBF) approximations [32], Multivariate Adaptive Regression Splines (MARS) [33], Support Vector Regression (SVR) [34] and Polynomial Neural Network (PNN) [35]. In general, it is not possible to claim that one metamodel always outperforms the others in all the applications; the choice of the best metamodel is still strongly problem-dependent and usually not a priori-known.

GPs, which belong to the huge field of stochastic processes, have been extensively used in the machine learning community for problems of regression, especially when dealing with time-demanding model regressions (like the BE-TH models typically used for PSS applications) due to their interesting properties: they have been shown capable of reproducing numerous system responses, even in case of non-linearity and without adding further complexity to the problem; furthermore, they are *exact interpolators*, i.e., the prediction at training input combinations is exactly the observed output with zero prediction error [36]. However, despite all these remarkable qualities, standard GPs show some drawbacks. First, inference on the GPs scales poorly with the number of training samples (N_{train}), typically requiring computing time in $O(N_{train}^3)$ for calculating the inverses of $N \times N$ covariance matrixes. Second, GPs are usually stationary, i.e., same estimation confidence in the whole domain. Indeed, it is not possible to be more accurate in some regions of the output space, which are potentially more interesting than others [37].

One specific category of GPs, called Kriging metamodels, allows to overcome the issue of stationarity typically associated to standard GPs. They have been originally developed in Geostatistics, indeed geology and soil science (like basically any other discipline working with data collected at different locations) need a model to indicate when there is interdependence among the measurements and to efficiently exploit such measurements for spatial prediction [38]. In practical terms, the metamodel prediction is based on a weighted combination of all the outputs already observed. These weights are defined

according to the distance (in the input space) between the new location to predict and the locations of the previous observations in a way that, the closer the input data are, the more correlated the corresponding prediction errors [39]. Later, beyond the applications in the geology field, Kriging metamodels became really widespread in many kinds of deterministic problems. Moreover, Kriging metamodels (like all the GPs in general) return the estimation confidence together with the response prediction at any site and this, coupled with the previously described non-stationarity, makes them particularly suitable to be implemented with adaptive sampling. On the other hand, Kriging metamodels still scale poorly with the number of training samples like all the GPs and they may also become very time-demanding in case their training has to be repeated many times.

1.3.2 Methods for Tackling the Output Non-smoothness and Multimodality

Besides the high cost of computation associated to the BE-TH models reproducing the PSSs responses, another relevant issue that may be faced when characterizing the CRs of a PSS is the possible non-smooth and/or multimodal behaviour of the system output considered. Indeed, this could lead to large approximation errors if not properly treated. In the following subsections two of the most innovative techniques to deal with non-smoothness and multimodality of a specific approximated output are presented: on one side, clustering of the output domain through a classifier with the aim of separating the different output behaviours in different “prototypical groups” (1.3.2.1). On the other side, the reduction of the problem dimensionality by Sensitivity Analysis (SA) methods to focus separately on the different output clusters and restrict the analysis only to those input parameters significantly affecting the clusters of interest (e.g., in this work, the clusters connected with the failure of the PSS function) (1.3.2.2).

1.3.2.1 Clustering the Output Space

One possible approach developed in the recent years to tackle non-smoothness and multimodality suggests a clustering of the output domain: this allows separating the

regions (also called partitions) of different behaviours and/or discontinuities in different output clusters and isolating them for a successive analysis [40]. Then, multiple metamodels, possibly coupled with adaptive sampling, can be fitted to these different domain regions to obtain a better approximation (rather than constructing a unique metamodel for the whole output space).

For this purpose, a classifier (namely, state-selecting model) represents a viable solution to identify the output states and cluster the output domain. One of the most common type of classifier is the Support Vector Classifier (SVC) [41]. It owes its popularity to the capability to identify disjoint regions and bring to an optimal decomposition through a maximization of the margins between them, thanks to an optimal separating function, even in case of multidimensional space. The SVC requires the construction of a training set (or DoE) constituted by samples properly spread over the input space to catch all the corresponding output behaviours on the PSS state-space (e.g., they can be generated with LHS). In particular, the training data allow identifying the hyperplanes separating the different regions: the training data directly belonging to these hyperplanes are called support vectors. The classification of training samples can be obtained by expert judgement or by some data clustering techniques, e.g., K-means or Hierarchical clustering. A label is assigned to each training sample according to the corresponding region. In general, the classification of an input combination can be binary or probabilistic: binary classification is the most widespread method and it implies a hard classifier according to which the sample may belong to one single class; on the other side, the use of a fuzzy classifier leads to probabilistic classification which implies that a sample contemporary belongs to all the classes with a certain degree of membership (weight).

One recent and interesting solution exploiting a SVC for clustering the output domain against the problem of non-smoothness and multimodality is represented by the so-called “two stage-surrogate modelling”, introduced in [18] (see Fig. 3). After identifying the different domain partitions and constructing an SVC, a metamodel is trained for each partition considered interesting to explore, i.e., in our case only for those partitions that are related to the failure of the PSS function. Then, a new system configuration (\mathbf{x}) (whose output needs to be predicted and identified if it is critical or not) is, first, classified by the

SVC (1st stage); then, it is evaluated by the metamodel (2nd stage) specifically constructed for the region x belongs to (prediction phase).

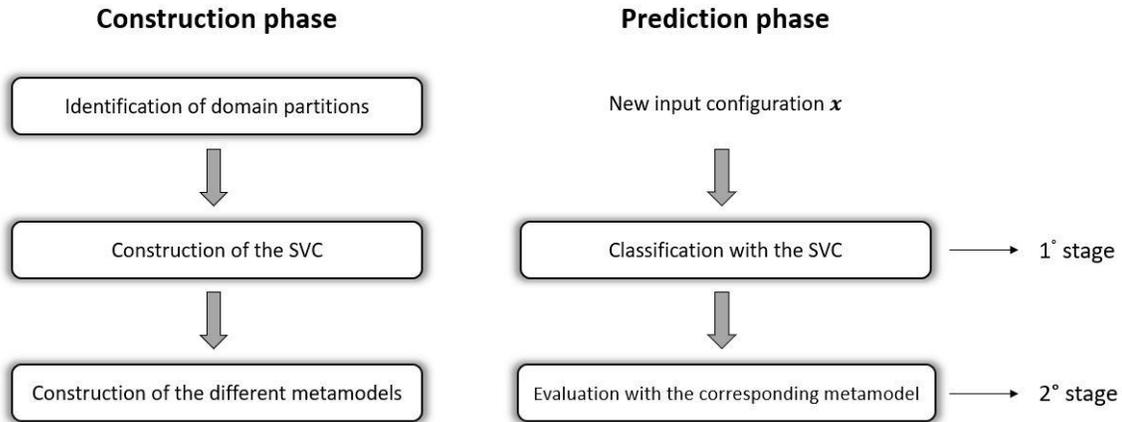


Figure 3: Two-stage surrogate modelling technique [18]

1.3.2.2 Dimensionality Reduction

A different approach to deal with a metamodel performing poorly due to the output non-smoothness and/or multimodality consists in circumventing the problem by means of feature selection [42] in order to reduce the dimensionality of the problem. Indeed, any metamodel, in general, greatly benefits from a dimensionality reduction; moreover, if the analysis is restricted only to those parameters significantly affecting the output clusters of interest (e.g., the clusters connected with failure behavior), also the specific issue of output multimodality can be overcome. Feature selection techniques for dimensionality reduction usually rely on many I/O relations, which might become an issue when the system model is time-demanding. Alternatively, SA methods can be employed to achieve the same goal of feature selection by ranking the inputs in terms of their contribution to the model output [43], [44]. Several SA techniques are available in literature and they can be subdivided into Local and Global methods [45], with the latter that are more suitable for feature selection since they quantify the contribution of each input to the variability of the output over the entire range of both the input and the output [46]. Global SA can be also divided into Regression-Based Sensitivity Analysis (RBSA) methods, also known as non-parametric techniques [47], such as Standardized Regression Coefficients (SRC) or Partial Correlation

Coefficients (PCC) [48], Variance-Based Sensitivity Analysis (VBSA) methods, such as Sobol' method [49], [50], [51], and Distribution-Based Sensitivity Analysis (DBSA) methods, also known as moment-independent methods [44], such as δ indicator [52], input saliency [53], Hellinger distance [54] and Kullback-Liebler divergence [54]. However, both the RBSA and VBSA methods, in general, suffer from the output non-smoothness and/or multimodality (as explained in detail in Section 4.1). On the other hand, DBSA methods become suitable to overcome this issue [55], e.g., when they are based on Finite Mixture Models (FMM), which provide a natural “clustering” of the output (e.g., subdividing the data in groups of large safety margin, low safety margin, failure) that can be used to calculate the sensitivity indexes [56], [57].

Table 1: Comparison between different SA methods to tackle the state-space non-smoothness and/or multimodality

Method	Low cost	Non-smoothness	Multimodality
RBSA	YES	NO	NO
VBSA	NO	YES	NO
DBSA	NO	YES	YES

In particular, FMMs are a flexible and powerful modeling tool for univariate and multivariate data, providing a formal approach to unsupervised learning for statistical pattern recognition. Indeed, FMM analyze a set of output variables (training set), each one assumed to be generated by a certain random model, i.e. a certain distribution of the mixture (also called component). Then, it infers the distributions parameters and identifies the distribution that has originated each training output, leading to the clustering of the training output variables. Moreover, FMM can be used in support of DBSA methods, aiming at identifying the most relevant input variables affecting the output clusters and hence to perform a feature selection [57]. The choice of the most appropriate model space (i.e., the space generated by a linear combination of known distributions of a specific kind) and the extraction of the right number of components to approximate the output multimodal distribution are challenging. Different metrics, based on Maximum Likelihood (ML) estimation, have been developed in the past to guide the model space selection: Minimum

Message Length (MML) [58], Akaike Information Criterion (AIC) [59] and Bayes Information Criterion (BIC) [60].

1.4 Methodological and Applicative Contribution of the Present Thesis

In the present thesis, a CRs characterization has been carried out for a generic PSS of a NPP, based on an Emergency Heat eXchanger (E-HX) and designed for DHR after the reactor shut down due to an accident initiation (e.g., Station Black-Out (SBO)). The PSS operation during the accidental transient is modelled through a time-demanding BE-TH code (more specifically, RELAP5-3D code). A single code run takes about 4.30h of computation. In particular, the PSS operation is analysed with respect to two outputs representative of the system response during the transient: 1) the amount of energy exchanged by the PSS (E_{ex}); 2) the maximum value reached by the pressure evolution inside the Pressure Vessel (PV) (p_{max}). The two system outputs are analysed separately to obtain a better accuracy in the metamodel-based approximation. Moreover, diversified approaches are required because of the different nature of the two outputs (see [Section 2.3](#) for details): in particular, E_{ex} shows a generally smooth trend, whereas p_{max} is characterized by a non-smooth and multimodal behaviour.

As far as the analysis of the exchanged energy output (E_{ex}) is concerned, an integration of adaptive sampling and Kriging metamodeling, known as Adaptive Kriging Monte Carlo Sampling (AK-MCS) [25], is here adopted and tailored to efficiently identify the CRs of the PSS. The objective of the analysis is, thus, to present the computational framework and show its feasibility and advantages for the PSS CRs exploration task. On the contrary, the aim is not to carry out the complete reliability analysis of a specific PSS. To the best of the author's knowledge, this is the first time that AK-MCS technique is implemented for the identification of the CRs of a PSS of an NPP.

In the case of p_{max} , the direct application of the metamodel-based AK-MCS procedure leads to poor results in terms of approximation, because a traditional smooth Kriging metamodel is not suitable to tackle its non-smoothness and multimodality. To address this problem, a novel method is proposed, inspired by the framework introduced in [11], which comprises three steps: i) “dimensionality reduction” is first carried out through a DBSA method supported by FMM, to tackle the output non-smoothness and multimodality; ii) an “iterative metamodel training” based on AK-MCS, is then implemented to substitute the computationally expensive I/O BE-TH model relationships on the reduced space by means of an accurate Kriging metamodel; iii) “CRs representation and information retrieval”: the Kriging obtained at the previous step is finally exploited to evaluate a large number of new input combinations used to visualize the CRs and retrieve useful information about them.

The remainder of the thesis is organized as follows. In [Chapter 2](#) the case study of the PSS for DHR is presented with a detailed description of the two output variables selected to monitor the system response. In [Chapter 3](#) the metamodel-based AK-MCS methodology, introduced for the first time in [25], is presented in the form tailored by [11] for the CRs exploration and, then, applied to the case study with respect to the analysis of the exchanged energy output (E_{ex}). [Chapter 4](#) offers an exhaustive description of the innovative framework based on a DBSA method and supported by FMM to carry out the CRs characterization in presence of non-smoothness and multimodality. The framework is then applied for the analysis of the PSS pressure output (p_{max}). In the last part, a comparison is carried out (in terms of accuracy and computational cost) with an alternative approach of literature [18], based on the combination of a SVC, to identify the different output clusters, and adaptive Kriging metamodeling (AK-MCS), applied to those clusters connected with the PSS functional failure. Finally, in [Chapter 5](#) the main achievements of the present work are summarized, and some conclusions are drawn.

2 Case Study

The PSS under study can be adopted for removing the decay heat from the core after the reactor shut down due to an accident initiation (e.g., SBO) through a natural circulation closed loop, including a shell and tube E-HX immersed in a large liquid pool elevated above the core. Several examples are available in literature about PSSs for DHR, with different degrees of passivity, implemented in advanced nuclear reactors: elevated tanks natural circulation loops (core make-up tanks), gravity drain tanks, passively cooled steam generator natural circulation, passive residual heat removal heat exchangers, passively cooled core isolation condensers and sump natural circulation. Among these PSSs, two of them are based on the idea of exploiting a heat exchanger immersed in an elevated pool to remove the decay heat, like in the case under study: the Passive Residual Heat Removal (PRHR) heat exchanger adopted in several advanced Pressurized Water Reactors (PWRs) and the passively cooled Isolation Condenser (IC) designed for advanced Boiling Water Reactors (BWRs) [6].

In particular, PRHR heat exchanger (also called emergency heat exchanger) is constituted by a single-phase liquid natural circulation loop (Fig. 4) designed to provide DHR to the reactor core for extended periods of time (at least few hours) [2]. The PRHR heat exchanger loop is normally pressurized at the same pressure level of the primary coolant and ready for service. The operation is triggered by the opening of an isolation valve typically located at the bottom of the heat exchanger. When the isolation valve opens the natural circulation starts thanks to the relative elevation between the vessel (heat source)

and the heat exchanger (heat sink) which generates a buoyancy force eliminating the need of a pump. The natural circulation loop consists in the following elements: at first, the decay heat is removed from the fuel in the core and transferred to the single-phase liquid flowing in the vessel. Then, the heat stored in the liquid is carried from the core to the elevated heat exchanger thanks to the buoyancy force and, successively, it is transferred to the heat exchanger tubes. Finally, the cooled liquid returns back to the vessel bottom, ready to remove further decay heat. The crucial point of this phenomenology is represented by the heat transfer inside the emergency heat exchanger that is constituted by three different mechanisms: single-phase convective heat transfer at the inner surface of the heat exchanger tubes, conduction through the tube walls, and nucleate boiling at the tubes outer surface [6]. This PSS belongs to the category D of passivity.

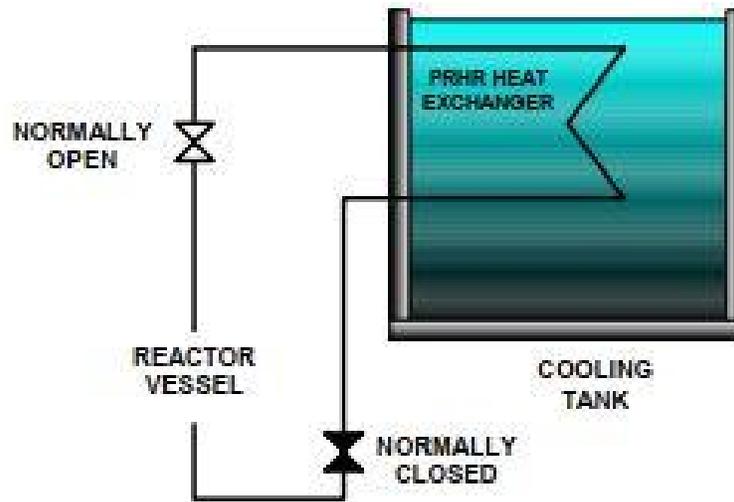


Figure 4: PRHR heat exchanger system based on a natural circulation loop and designed for advanced PWRs [2]

On the other side, the IC is used to provide cooling to some advanced BWRs when the reactor is shut down and isolated from the primary heat sink (i.e., the turbine/condenser set). The operation is triggered by two valves located on the IC lines that are normally closed. At accident initiation, the valves open and the steam is diverted to the IC heat exchanger where it is condensed in its vertical tubes [2]. The driving mechanism for the steam flow inside the IC lines is represented by the steam condensation process; indeed, the steam condensing inside the IC heat exchanger creates a low-pressure region which

draws additional steam in. The heat transfer through the IC heat exchanger tubes into the water pool is based on three different mechanisms: two-phase steam condensation (phase change) at the tubes inner surface, heat conduction through the tubes walls, and convective heat transfer or nucleate boiling at the tubes outer surface [6]. The efficiency of the condensation process can be negatively affected by the deterioration of the heat transfer coefficient due to the possible presence of non-condensable gases inside the IC lines. After condensation, the steam has turned into single-phase liquid and leaves the IC heat exchanger to return to the vessel by gravity draining. The IC belongs to the category D of passivity, as the PRHR system of the PWR.

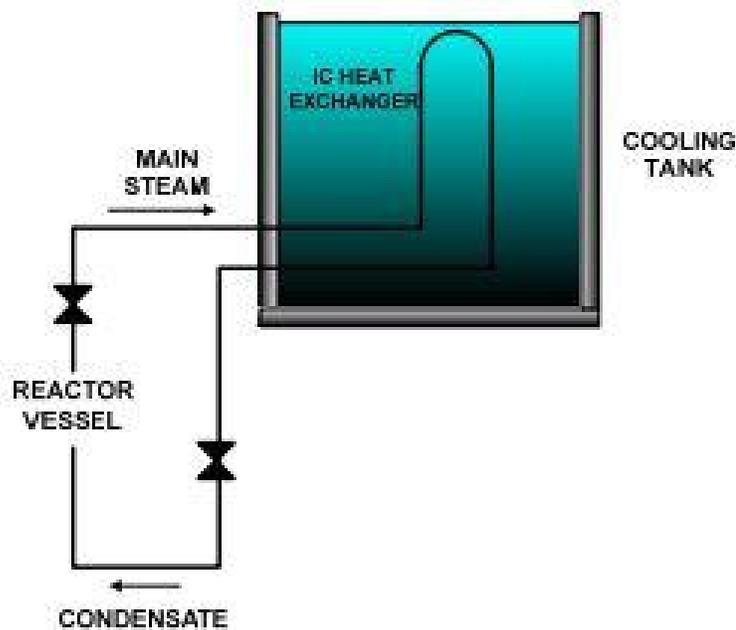


Figure 5: IC heat exchanger system based on a condensation-driven loop designed for advanced BWRs [2]

The work done in the present thesis aims at showing the effectiveness of the approaches proposed for the CRs exploration of a generic PSS for advanced nuclear reactors, without addressing any specific case of modern PSS. Some PSSs relying on natural circulation can be taken as a reference, but the methodologies developed here can be theoretically implemented for many examples of PSSs.

2.1 Description of the Passive Safety System and its Failure Criteria

The PSS under analysis is a generic DHR system which draws inspiration from the IC heat exchanger system used in advanced BWRs. It relies on a natural circulation closed loop including a heat exchanger immersed in a large water pool. In particular, this specific PSS function is to remove the decay heat after an SBO accident when the steam is prevented from going to the steam turbine, but diverted into the PSS lines. The PSS consists of (see Fig. 1):

- an E-HX condensing the steam, composed by two cylindrical horizontal headers and many vertical straight pipes, immersed in a large pool elevated with respect to the PV;
- a steam line bringing the steam from the PV to the E-HX;
- a condensate line bringing the liquid back to the PV after condensation in the E-HX tubes, with an Activation Valve (AV) that opens to trigger the PSS operation;
- a Main Steam Isolation Valve (MSIV) located on the primary circuit, closing at the accident initiation to isolate the PV from the turbine.

A PSS based on water natural circulation in a heat exchanger immersed in a pool is classified according to [1] as belonging to the category B of passivity, i.e., no signal input of “intelligence” and no external power source, no moving mechanical parts, but moving working fluid. However, due to the opening of the AV on the condensate line necessary to start the PSS operation, the system under analysis should be classified in the category D of passivity, which addresses the intermediate zone between active and passive behaviour (category D is also called “passive execution – active initiation”) [3].

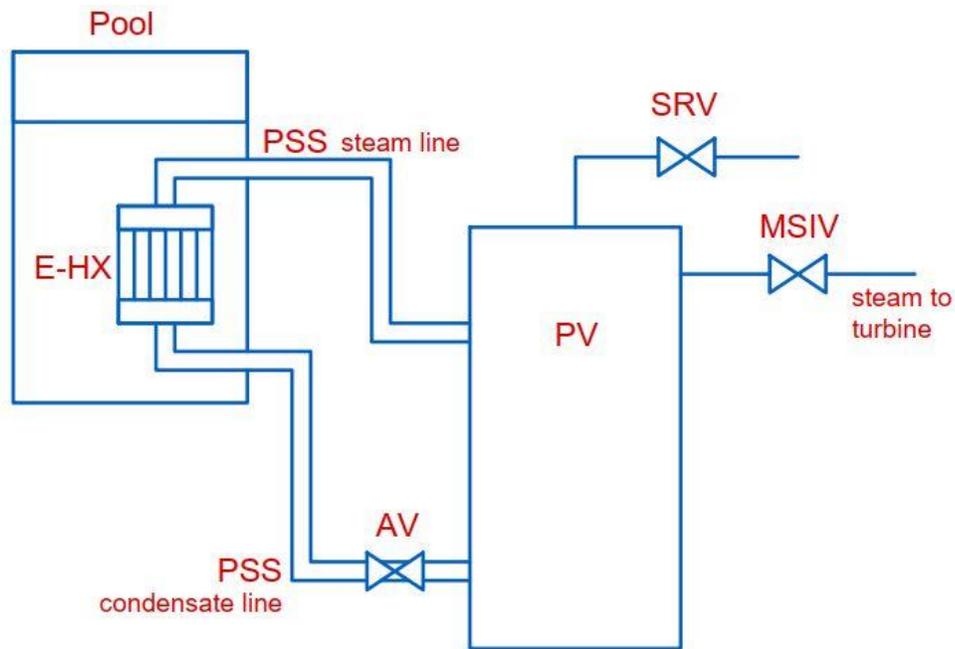


Figure 6: Simplified sketch of the PSS under study

The E-HX function is to remove the decay heat during an SBO accident after reactor shut down. This should prevent the energy increase within the PV, which may lead to over-pressurization and over-heating of the various components. The reactor is supposed to normally operate in steady state conditions at a pressure of about 70 bar and the E-HX is completely submerged in the pool. The PSS steam line is initially filled with saturated steam at 70 bar, with possible presence of a certain amount of non-condensable gases. The condensate line initially contains subcooled liquid assumed at 40°C and 70 bar. The AV is initially closed, preventing the connection of the PSS condensate line with the PV. The pool is filled with water initially assumed at 40°C.

We consider that an SBO accident occurs. The reactor is, then, shut down and the MSIV closes simultaneously with the opening of the AV, so that the vapor from the PV is not sent to the steam turbine, but it is directed into the PSS. Vapor condenses inside the E-HX due to the heat transferred to the pool and flows back, as liquid, to the PV through the condensate line.

Five input parameters $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$ have been identified as most relevant for the response of the PSS during the SBO accident. They are mainly associated to the

unreliability of components (e.g., mechanical malfunctions), but also to the uncertainty of some parameters (e.g., the amount of non-condensable gases that may be present in the PSS lines).

1. AV flow area (A_{AV}): the opening of the AV triggers the PSS operation when the accident occurs.
2. AV opening delay (DEL_{AV}): the AV may open with a certain delay with respect to the beginning of the accidental sequence.
3. MSIV residual flow area (A_{MSIV}): the MSIV should close completely when the accident occurs, but some leakage may be present (i.e., normalized flow area $> 0\%$).
4. MSIV closure delay (DEL_{MSIV}): the MSIV may close with a certain delay with respect to the beginning of the accidental sequence.
5. Percentage of non-condensable gases ($NC\%$): in the PSS lines a certain amount of non-condensable gases may build up during the system operation. These gases tend to accumulate in the coldest regions of the system where the vapor partial pressure is the lowest. Therefore, they could be present in the steam line of the PSS before the activation of the PSS. The presence of non-condensable gases leads to a deterioration of the heat transfer mechanism, indeed their build up near the condensate film in the E-HX tubes inhibits the diffusion of vapor from the bulk mixture to the liquid film. The net effect is a reduction of the driving force for heat and mass transfer [6]. Their quantity is expressed in terms of percentage of volume occupied by the non-condensable gases with respect to the total volume of the steam line.

To generate different combination of values of the five parameters, x_m ($m = 1, \dots, 5$), uniform probability distributions have been considered to span their ranges of variation, with equal sampling probability of any value and, thus, explore their possible combination of values in the search for CRs. The range of each input has been properly adjusted through a preliminary analysis not to sample too far from the limit surface. For example, if it has been discovered that $NC\% > 30\%$ leads to failure, also considering the possible interactions with the other input parameters, and knowing that the scope is to focus the

analysis nearby the Y_{thres} , it is not worth it to simulate also $NC\%$ values much higher than 30%. Table 1 lists the range of variations of the input parameters.

Table 2: Range of variation of the inputs

Input		Symbol	Range of variation
AV flow area	(%)	A_{AV}	0 ÷ 100
AV opening delay	(sec)	DEL_{AV}	0 ÷ 720
MSIV residual flow area	(%)	A_{MSIV}	0 ÷ 0.15
MSIV closure delay	(sec)	DEL_{MSIV}	0 ÷ 7200
Non-condensable gases percentage	(%)	$NC\%$	0 ÷ 40

The DHR system successful response to the accident is measured in terms of its heat removal function and, specifically, in relation to the amount of heat removed during the accidental transient (which has been simulated for a duration of about 8h). If the heat is not removed adequately, the temperature and pressure may dangerously rise inside the PV and if pressure increases beyond the Safety Relief Valve (SRV) set-point assumed at 75.5 bar, the valve opens to discharge the vapor inside the containment building (not simulated in the model). An output vector with two output parameters, i.e., $Y = (Y_1, Y_2)$, namely Transient Performance Indicators (TPIs) [8], is considered to evaluate the PSS functional response:

1. Energy exchanged (E_{ex}): the total amount of energy removed by the DHR during the transient;
2. Maximum of PV pressure (p_{max}): maximum value reached by the pressure evolution inside the PV during the transient.

Table 3 lists the values of the input and output parameters for the reference transient, i.e., the “reference conditions” of nominal functioning of the DHR system. The energy exchanged output is measured calculating the percentage ($E_{ex,\%}$) with respect to the value obtained in the reference conditions.

Table 3: I/O reference conditions

Variable symbol	A_{AV}	DEL_{AV}	A_{MSIV}	DEL_{MSIV}	$NC\%$	$E_{ex,\%}$	p_{max}
Reference Value	100%	0 sec	0.00 %	0 sec	0 %	100%	70.0 bar

Reference conditions allow to define the Target Missions (TMs) the system needs to accomplish during its operation: 1) to ensure $E_{ex,\%} > 90\%$, and 2) to keep p_{max} below 75.5 bar. In the present thesis, safe operation of the system is assumed when the TM under analysis is accomplished; on the contrary, if the TM is not met the system is said to fail its function. Therefore, two Failure Criteria (FC) are identified:

1. Low heat removal: if $E_{ex,\%} < 90\%$ [8].
2. Steam release in the containment: if $p_{max} > 75.5 \text{ bar}$ (i.e., pressure rise in the PV causes the SRV to open, which leads to vapor release in the containment of the NPP)

In particular, in [Chapter 3](#) the metamodel-based AK-MCS procedure is applied for the characterization of the CRs related with $E_{ex,\%}$ output with respect to the FC “Low heat removal”. Whereas, in [Chapter 4](#), a novel exploration framework is proposed for the analysis of the CRs related with p_{max} with respect to the FC “Steam release in the containment”, due to the non-smooth and multimodal nature of such output.

2.2 Synthetic Description of the PSS T-H model

A RELAP5-3D model of the generic PSS connected to a simplified reactor PV has been adopted. The model has been developed in cooperation by University of Pisa and Polytechnic of Turin [61]. The RELAP5-3D model that simulates the behaviour of the generic PSS described in [Section 2.1](#) is composed by two hydrodynamic regions: the primary side (with the PV, the E-HX and the pipe connections) and the pool side.

The PV is modelled using pipe and branch components, whereas its connections to the feedwater line and steam supply line are represented by two time-dependent volumes. On the steam supply line, the MSIV is located and modelled as a servo-valve, while the SRV at on the top of the PV is modelled as a trip valve. The E-HX is constituted by two headers represented by branch components and a pipe component for the heat exchanger tubes. PSS steam and condensate lines are represented by a series of pipe components. On the condensate line, the AV is located and modelled as a motor valve. For what concerns the

pool side, branch and pipe components laterally connected through crossflow junctions have been adopted.

A more detailed description of the RELAP5 model can be found in [61]. Some closure equations, relevant for the operation of the PSS (e.g., condensation heat transfer within the HX tubes), have been revised and correction factors have been applied to properly simulate the occurring phenomena [62].

RELAP5-3D constitutes the tool adopted to evaluate I/O representations of system behaviour, also named I/O observations, in terms of $Y = f(\mathbf{X})$, where $\mathbf{X} \in D_{\mathbf{X}} \subset \mathbb{R}^M$ represents the input vector of dimension M related to a given system configuration and whose output $Y \in D_Y \subset \mathbb{R}$ expresses the system state/condition. Each transient simulation with the RELAP5-3D code takes about 4.30h on a PC with *CPU Intel Core i7-7500U CPU @ 2.70GHz dual core* (Simulations carried out at the Energy Department of Polytechnic of Turin by faculty members.). The RELAP5-3D simulations are used as the training simulations (i.e., relations known with absolute certainty) for the here developed modelling techniques like Kriging, FMM and SVC.

2.3 Description of the Two Outputs Behavior

To analyse the behaviour of the two outputs identified to describe the PSS response in case of SBO is fundamental for the choice of the most appropriate approximating method which is exploited for the CRs exploration. Indeed, as explained in [Section 1.2](#), common metamodels used to replace time-demanding model simulations usually requires the underlying model function to show accommodating properties of regularity (e.g., smoothness and continuity). For this reason, in case the approximated output exhibits an irregular trend, as p_{max} that is non-smooth and multimodal, an alternative approach is necessary to avoid large errors.

2.3.1 Energy Output Behavior

The trend of $E_{ex,\%}$ is illustrated in Fig. 7 which collects the outcomes of 100 RELAP5-3D simulations generated with LHS. In general, all the five input parameters negatively impact on the amount of energy exchanged by the PSS when they deviate from their nominal conditions causing $E_{ex,\%}$ to decrease. Nevertheless, $E_{ex,\%}$ shows a regular trend (i.e., no discontinuities, no sharp changes, etc.) with a higher concentration of output values around $E_{ex,\%} = 70 \div 80 \%$. This is the reason why a smooth Kriging metamodel, within a properly developed AK-MCS methodology, is suitable for approximating such behaviour.

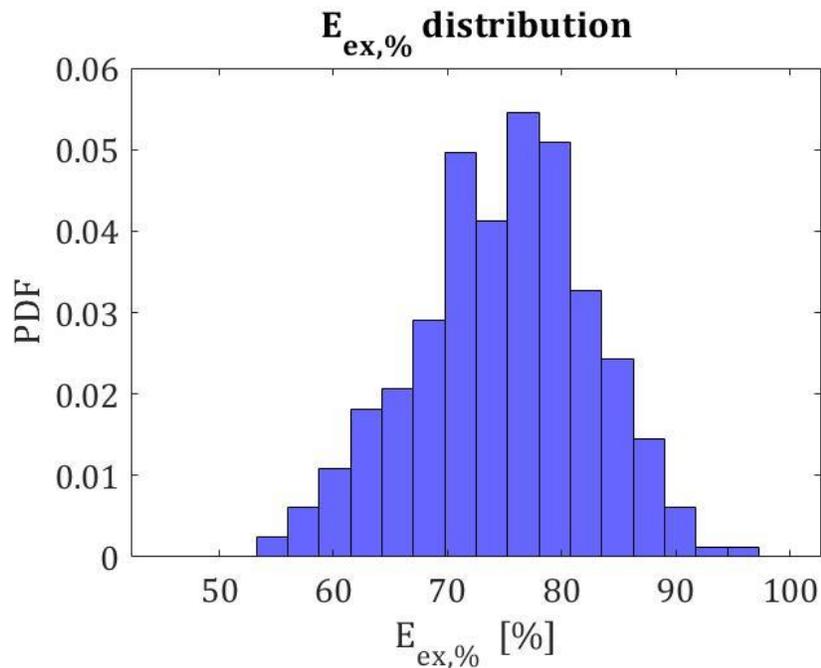


Figure 7: $E_{ex,\%}$ regular distribution

2.3.2 Pressure Output Behavior

The irregular trend of p_{max} is illustrated in Fig. 8 which collects the outcomes of 200 RELAP5-3D simulations.

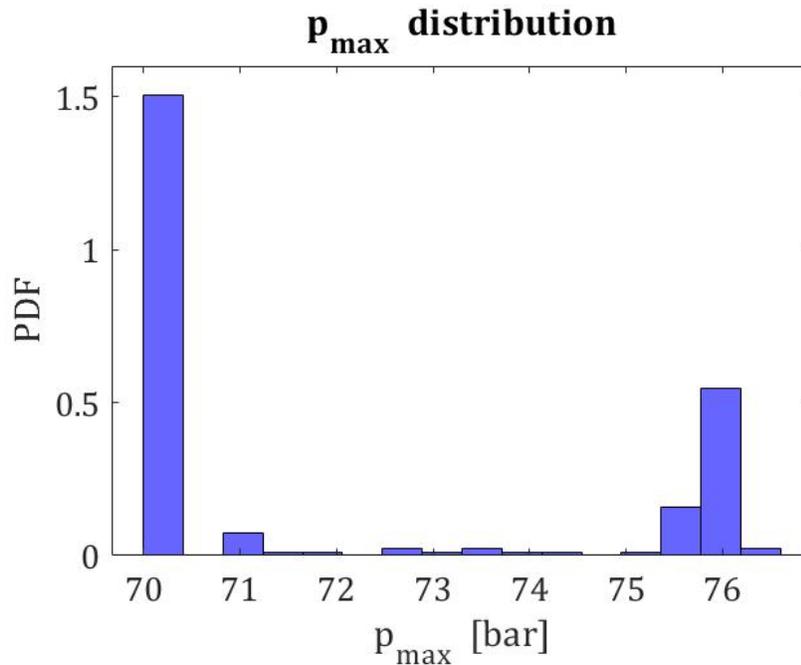


Figure 8: p_{max} non-smooth and multimodal distribution

At least two modes can be identified and hence two corresponding clusters are defined: a first cluster with low pressure values (70.0 bar), which is associated to the majority of the simulations collected; if the decay heat is correctly removed the pressure should never increase during the accidental transient and hence p_{max} coincides with the pressure value at the beginning of the transient, i.e., $p_{max} = 70.0 \text{ bar}$. A second cluster is concentrated around $Y_{thres} = 75.5 \text{ bar}$; if the MSIV closes before the AV opening, the decay heat cannot be removed through the E-HX and the vapor builds up in the PV, causing the pressure to rise. In this case, a quite short time interval is sufficient, in which the PV remains without outlets for vapor discharge, to cause a sharp pressure increase towards Y_{thres} with consequent SRV opening. Finally, very few points fall in the middle region showing intermediate values of pressure.

To conclude, to use a traditional smooth metamodel without an appropriate approach to tackle p_{max} non-smoothness and multimodality leads to large errors, like in the case of the direct application of AK-MCS technique.

2.4 Other Possible Relevant Parameters

There are some input and output parameters associated to the PSS that, in the first place, it has been decided not to include in the analysis of the present thesis because they are thought not to give a significant contribution to the accidental scenario or not have a primary role in terms of their capability to describe the PSS response. Anyway, some of them could be taken into account for future works and ranked through a SA, especially for a deeper analysis on the facility.

The input parameters that have not been considered in the present thesis are mainly associated to uncertain quantities, e.g., forward and reverse energy loss coefficients accounting for pressure losses in the components, but also to some malfunctions that have been rejected because they are considered too unlikely (e.g. unexpected closure of the AV during the transient). The most suitable for future analysis are:

- PSS steam line pipes inclination: the design foresees that pipes are perfectly horizontal or vertical. A pipe inclination could provoke water accumulation in the angles where the direction changes with “syphon effect”, hindering the vapour flux.
- Component rupture: a rupture may occur in many locations of the plant, like in the pool container or along the PSS lines. Such event is, in general, more likely in a pipe rather than a pool and, particularly, if the pressure value rapidly changes (e.g. in case of a strong pressure wave). In the case study, the PSS lines before accident initiation are pressurized at the same pressure level of the PV, thus, no pressure wave is generated when the AV opens.

3 Metamodel-based AK-MCS Framework for the CRs Exploration of a PSS

The amount of energy exchanged (E_{ex}) by the E-HX has been identified as relevant for the analysis of the PSS response to accidental conditions. In particular, the energy exchanged is measured in terms of percentage ($E_{ex,\%}$) of respect to the energy exchanged in reference conditions (see [Section 2.1](#)) and its failure threshold has been fixed to $E_{ex,\%} = 90\%$ [8]. In [Section 2.3](#), it has been demonstrated how the energy output shows a general smooth trend in the state-space. For this reason, metamodel-based AK-MCS technique [25] combining a fast-running Kriging metamodel and adaptive sampling has seemed suitable for the characterization of the CRs with respect to the FC “Low heat removal”, according to the exploration framework proposed in[11].

In particular, [Section 3.1](#) offers an exhaustive description of the framework applied: at first, it is presented the Kriging formalism together with the advantages of its use ([3.1.1](#)); then, the steps concerning AK-MCS procedure are presented in detail according to ([3.1.2](#)); finally, it is explained how the Kriging metamodel obtained at the end of AK-MCS iterations can be exploited to retrieve useful information about the CRs and, thus, how to graphically represent them ([3.1.3](#)).

On the other side, in Section 3.2, the framework has been applied to the CRs exploration of the PSS under study with respect to the specific FC “Low heat removal”: a Kriging metamodel has been constructed according to the available RELAP5-3D simulations (3.2.1) and, then, the Kriging has been refined through the AK-MCS procedure properly tailored for the analysis of $E_{ex,\%}$ (3.2.2). Finally, the results of the application are presented both in terms of the level of accuracy reached by the metamodel and graphical representation of the CRs (3.2.3).

3.1 AK-MCS Framework Presentation

3.1.1 Kriging Metamodel

A fast-running surrogate metamodel is constructed to replace the time-demanding model simulations. Indeed, given that many input combinations need to be simulated to retrieve enough information about the system CRs, it is fundamental to speed up the calculation. Among the several options available in literature [63] (already discussed in 1.3.1.2), we resort to GPs and, particularly, to one specific category of GPs, i.e., Kriging metamodels [64], since they present numerous interesting features, as described in 1.3.1.2. The most interesting property for this case study is that they can fit numerous response functions with a diversified level of accuracy throughout the domain (non-stationarity). Indeed, this makes them suitable for the specific aim of CRs characterization, since the metamodel can be refined in proximity of CRs limit surface to correctly discriminating between critical and safe system conditions. This can be achieved by training the Kriging with original model simulations whose outputs are concentrated nearby the limit surface, indeed, adaptive training strategies have been recently developed to this aim. In the present thesis, it is followed the metamodel-based AK-MCS framework developed in [25]. Before the application of the AK-MCS procedure, an initial Kriging is built using a small number of RELAP5-3D simulations whose input combinations can be generated with LHS, and the corresponding I/O values constitute the initial I/O training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$. Then, the

following training simulations for the metamodel refinement are iteratively added according to the AK-MCS procedure.

From a mathematical perspective, Kriging is defined as a stochastic interpolation algorithm, which assumes that the model output $y = f(\mathbf{X})$ is the realization of a Gaussian process indexed by $\mathbf{X} \in D_X \subset \mathbb{R}^M$, where D_X is the metamodel domain of validity and M is the dimensionality of the problem [11], [65]:

$$y = f(\mathbf{X}) = \mathcal{N}(\mathbf{h}(\mathbf{X})^T \boldsymbol{\beta}, \sigma^2 Z(\mathbf{X})), \quad (1)$$

where the first term, $\mathbf{h}(\mathbf{X})^T \boldsymbol{\beta}$, is the mean value of the Gaussian process, also known as trend, consisting of P arbitrary functions $\{h_j; j = 1, \dots, P\}$ and the corresponding coefficients $\{\beta_j; j = 1, \dots, P\}$; the second term consists of the variance of the Gaussian process, i.e., σ^2 , and a zero mean, unit variance stationary Gaussian process, i.e., $Z(\mathbf{X})$. The correlation function underlying $Z(\mathbf{X})$ is represented by $\mathcal{R}(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta})$, where \mathcal{R} is the correlation matrix (given a certain correlation function family) and $\boldsymbol{\theta}$ represents its hyperparameters. In particular, $\mathcal{R}(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta})$, describes the correlation between two vectors \mathbf{x}, \mathbf{x}' : the closer they are the higher their correlation. The Gaussian process assumption states that every set of realizations of the model output can be described by a Gaussian vector, whose relation between a single realization $y(\mathbf{x})$ and the rest of the set $\mathcal{Y}_{train} \in \mathbb{R}^{N_{train}}$ follows a Gaussian distribution defined by:

$$\begin{bmatrix} y(\mathbf{x}) \\ \mathcal{Y}_{train} \end{bmatrix} \sim \mathcal{N}_{N_{train}+1} \left(\begin{bmatrix} \mathbf{h}(\mathbf{x})^T \boldsymbol{\beta} \\ H \boldsymbol{\beta} \end{bmatrix}; \sigma^2 \begin{bmatrix} 1 & \mathbf{r}^T(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) & \mathcal{R} \end{bmatrix} \right). \quad (2)$$

In detail, H is the information matrix of the Kriging metamodel trend and in each row there are the regressors related to the corresponding observation \mathbf{x}_i (i.e., $H_i = \mathbf{h}(\mathbf{x}_i)$, $i = 1, \dots, N_{Krig}$); \mathcal{R} is the correlation matrix with elements $\mathcal{R}_{ij} = \mathcal{R}(\mathbf{x}_i, \mathbf{x}_j; \boldsymbol{\theta})$, $i, j = 1, \dots, N_{train}$, and $\mathbf{r}(\mathbf{x})$ is the vector of cross correlations between \mathbf{x} and each of the other vectors, whose elements reads as $r_i = \mathcal{R}(\mathbf{x}, \mathbf{x}_i; \boldsymbol{\theta})$, $i = 1, \dots, N_{train}$.

In the context of metamodeling, the interest is to predict a new point response, hence, given an experimental design or training set, i.e., $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$, with $\mathcal{Y}_{train} = (y_1, \dots, y_{N_{train}})$ and with an associated information matrix H and correlation matrix R , the prediction of the output, i.e., \hat{y} for a given input configuration \mathbf{x} is given by:

$$\hat{y}(\mathbf{x})|\mathcal{Y}_{train}, \sigma^2, \boldsymbol{\theta} \sim \mathcal{N}(\mu_{\hat{y}}; \sigma_y^2), \quad (3)$$

where $\mu_{\hat{y}}(\mathbf{x})$ and $\sigma_y^2(\mathbf{x})$ are respectively the mean value and the variance of the Gaussian random variate \hat{y} , defined by:

$$\mu_{\hat{y}}(\mathbf{x}) = \mathbf{h}(\mathbf{x})^T \boldsymbol{\beta} + \mathbf{r}(\mathbf{x})^T \mathcal{R}^{-1}(\mathcal{Y}_{train} - H\boldsymbol{\beta}), \quad (4)$$

$$\sigma_y^2(\mathbf{x}) = \sigma^2(1 - \mathbf{r}(\mathbf{x})^T \mathcal{R}^{-1} \mathbf{r}(\mathbf{x})) + (\mathbf{h}(\mathbf{x})^T - \mathbf{r}(\mathbf{x})^T \mathcal{R}^{-1} H) (H^T \mathcal{R}^{-1} H)^{-1} (\mathbf{h}(\mathbf{x})^T - \mathbf{r}(\mathbf{x})^T \mathcal{R}^{-1} H)^T \quad (5)$$

And the least square estimates of $\boldsymbol{\beta}$:

$$\hat{\boldsymbol{\beta}} = (H^T \mathcal{R}^{-1} H)^{-1} H^T \mathcal{R}^{-1} \mathcal{Y}_{train}. \quad (6)$$

3.1.2 Metamodel-based AK-MCS framework for CRs exploration

The Kriging metamodel initially constructed according to $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ is, then, refined through the AK-MCS procedure, introduced in [25] and further developed by [11]. In particular, The AK-MCS framework here proposed it has been properly tailored to the specific aim of characterizing the CRs of a PSS for nuclear application and it consists in the following steps, for each n -th iteration:

1. **Construction:** a Kriging metamodel is built with the available I/O training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ (see Section 3.1.1). The first I/O training set is $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$; then, the set is progressively updated and enriched in the successive iterations.
2. **Generation of random input configuration:** a large number N_{MCS} of new input configurations $\mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{N_{MCS}})$, is generated by means of LHS so as to efficiently span the input parameters space.
3. **Metamodel Evaluation:** the Kriging metamodel is used to evaluate the outputs corresponding to the \mathcal{X} combinations: $\hat{\mathcal{Y}} = (\hat{y}_1, \dots, \hat{y}_{N_{MCS}})$.
4. **Convergence check:** Convergence of the metamodel construction is verified through an a priori defined convergence (e.g., a certain error metric) or stopping criterion

(e.g., a limited computational budget expressed in the form of a maximum number of BE-TH simulations).

5. **Selection:** if convergence criterion at step 4 is not satisfied, the best candidate subset $\mathcal{X}^* \subset \mathcal{X}$ of input combinations is added to the current training set by evaluating the corresponding outputs \mathcal{Y}^* through the long-running BE-TH model. The N_{cand} best candidates are selected on the basis of their learning function values. Among the several examples of learning functions provided in literature [23], the U -function [11], [25] is adopted:

$$U(\mathbf{x}) = \frac{|Y_{thres} - \mu_{\hat{y}}(\mathbf{x})|}{\sigma_{\hat{y}}(\mathbf{x})}. \quad (7)$$

The $U(\mathbf{x})$ value represents the distance, expressed relative to the standard deviation, of the metamodel prediction (whose mean value is $\mu_{\hat{y}}(\mathbf{x})$ and estimation error is $\sigma_{\hat{y}}(\mathbf{x})$) from the contour of the CR, defined by Y_{thres} . The smaller is $U(\mathbf{x})$, the closer is the metamodel prediction to the failure threshold and the higher the interest in adding the observation corresponding to \mathbf{x} to the current training set, since the main scope is to focus on the limit state and to increase the metamodel accuracy in that area. However, notice that the choice of \mathcal{X}^* should *not* be made *only* among the N_{cand} combinations with the lowest U -function values. In fact, in this way the corresponding inputs could result too close to each other in their domain due to a high correlation function, bringing a small amount of information to the Kriging training process; some techniques (e.g., clustering) are proposed in literature to face this issue by evenly “spreading” the candidates along the limit state surface [11].

Once the new I/O observations $\{\mathcal{X}^*, \mathcal{Y}^*\}$ have been simulated with the original model and added to the training set, steps 1 to 5 are repeated until step 4 is verified.

3.1.3 Metamodel Accuracy Evaluation and CRs Characterization

The Kriging metamodel obtained at the end of the AK-MCS iterations is expected to provide predictions of the output with a satisfactory level of accuracy, especially in proximity of the CRs limit surface. This can be verified in different ways, e.g., by exploiting

an external validation set: the metamodel is tested on validation input combinations which are selected separately with respect to the set of training input combinations and, then, the corresponding metamodel predictions are compared to the real outputs which have been obtained from the BE-TH simulations. Thus, the final Kriging metamodel is exploited to replace the time-demanding model for the characterization of the CRs.

A large number of new input combinations \mathbf{x} (e.g., several thousands) can, then, be generated, again with LHS, and sent in input to the Kriging metamodel and the critical ones, i.e., $\hat{y} = f(\mathbf{x}) \leq Y_{thres}$, are retained for characterizing the shape and cardinality of the CRs. In mathematical terms, this corresponds to solving the inverse problem $\mathbf{x} = f^{-1}(\hat{y})$, with $\hat{y} \leq Y_{thres}$. Once this is done, a graphical representation of the CRs can be provided, with specific attention to their boundaries, by high dimensional data visualization techniques, e.g., scatter plots or Parallel Coordinates Plots (PCPs).

In brief, scatter plots show the two-dimensional projections of the CRs over all the possible pairs of inputs (this is useful to visualize the shape of the CRs). Moreover, in case of many input parameters involved, multiple scatter plots can be collected together in the so-called Scatter PLOt Matrix (SPLOM) providing a more complete view [66].

On the other hand, PCP [67] allows representing all the input combinations belonging to the CRs in a unique plot: all the M input variables, normalized on their respective ranges, are reported on vertical axes and lined up horizontally; then, each input combination is represented by a line connecting in the horizontal direction the corresponding input variables values on the vertical axes. In this way, the analyst is provided with exemplary patterns of typical critical conditions for the system operation.

3.2 AK-MCS Framework Application for CRs characterization of the PSS considered – Energy Output

3.2.1 I/O Training Set and Kriging Metamodel Construction

A first rough Kriging metamodel has been constructed with a possibly small I/O training set in order to be sent in input to the AK-MCS procedure, which automatically refines it for CRs exploration. The training I/O combinations used for its construction have been initially generated by simulations with varying values of each input x_m within its range (see [Table 2](#)). Unfortunately, no definite recommendations exist about the choice of the most suitable size of the training set [27], [68]. The criterion proposed for Kriging metamodels in [69] suggests a number of training combinations equal to about $10M$, where M is the dimensionality of the problem; hence, about 50 RELAP5-3D runs were necessary in this case ($M = 5$). Following this criterion, we built an initial I/O training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ made by 64 RELAP5-3D simulations in correspondence of the input values combinations generated by LHS to evenly cover the input domain.

The UQLab Software Framework for Uncertainty Quantification [70] has been used to fit the Kriging metamodel to the training set. UQLab provides straightforward parametrization of the Kriging (see [Section 3.1.1](#)): constant, linear, polynomial, or arbitrary trends, associated to elliptic and separable correlation kernels, based on many possible one-dimensional distribution families (e.g., Exponential, Gaussian, Matérn, or user-defined). The metamodel hyperparameters can be estimated through Cross-Validation (CV) or Maximum-Likelihood (ML) methods, using different optimization techniques (local or global) [65].

The best Kriging setting for the specific case study has been established by testing different options with the CV procedure. The initial I/O training set (64 RELAP5-3D simulations) has been split into K partitions (with $K = 4$) of the same size (16 simulations

each): a Kriging metamodel has been trained on $K - 1$ partitions and the CV error (which is actually a Root-Mean-Square Error, RMSE) has been evaluated by comparing the metamodel predictions on the input combinations of the left out partition k with the RELAP5-3D model outputs. The process has been repeated for $k = 1, 2, \dots, K$ and the CV error has been averaged among the trials. Then, the whole procedure has been repeated to calculate the average CV error with other Kriging options and, finally, the best option of a certain kind (e.g., the best trend option) has been selected according to the lowest average CV error. In particular, two Kriging features have been tested: the trend type and correlation function family (see Tables 4 and 5), while other features have been set to their default options defined in UQLab.

Table 4: Average CV error in trend type estimation (energy output)

Trend type	Ordinary	Linear	Quadratic
Average CV error [%]	8.82	5.13	9.53

Table 5: Average CV error in correlation function family estimation (energy output)

Corr. Function family	Exponential	Gaussian	Matérn 3/2	Matérn 5/2
Average CV error [%]	5.58	5.31	5.55	4.99

Note that the average CV error used to rank the different options is expressed in percentage because it has the same unit of measure of the predicted output, i.e., the percentage of energy exchanged ($E_{ex,\%}$). The best trend type has been evaluated with all the other Kriging features set to their default options defined in UQLab (the correlation function family set by default is *Matérn 5/2*). The same default options have been used for the estimation of the best correlation function family, with the trend type set to *Linear* (optimal setting found at the previous step). In the end, The Kriging best setting has resulted to be:

- Trend type: *Linear*
- Family of correlation functions: *Matern-5_2*
- Type of correlation functions: *Ellipsoidal* (default)
- Estimation method: *CV* (default)

- Optimization method: *Genetic Algorithm (GA)* (default)

3.2.2 Metamodel-based AK-MCS Application

The first rough Kriging metamodel constructed with $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ is here adaptively refined within the metamodel-based AK-MCS framework introduced in [Section 3.1.2](#). The steps concerning the framework application are illustrated in the way they have been applied to the PSS considered, in relation to the energy exchanged (E_{ex}) by the E-HX during an SBO accidental transient.

1. **Construction:** a new Kriging metamodel is constructed at each n -th iteration using an I/O training set of increasing dimension, starting from the initial one $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ made by 64 RELAP5-3D simulations. The Kriging features tailored on the initial training set (see [Section 3.2.1](#)), can be adjusted at each iteration of the adaptive procedure to improve the fit with the new training sets. The metamodel accuracy is improved specifically in proximity of the failure threshold ($Y_{thres} = 90\%$ of the energy exchanged during the reference transient).
2. **Generation of random input combinations:** $N_{MCS} = 10.000$ new combinations, $\mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{N_{MCS}})$, are generated by LHS (see [Table 2](#)). The number of combinations N_{MCS} ($= 10.000$) gives a satisfactory trade-off between thoroughness of PSS state space exploration and computational cost.
3. **Metamodel evaluation:** the sampled input combinations \mathcal{X} are run through the metamodel to predict the corresponding outputs of energy exchanged: $\hat{\mathcal{Y}} = (\hat{y}_1, \dots, \hat{y}_{N_{MCS}})$.
4. **Convergence check:** a double convergence criterion is defined. On the one hand, the level of accuracy of the metamodel should be increased as much as possible; on the other hand, the computational cost of the successive iterations (and corresponding RELAP5-3D simulations) should be kept to a feasible level.

Accuracy is evaluated with respect to the N_{val} combinations of the validation set (\mathcal{X}_{val}): $\hat{\mathcal{Y}}_{val} = (\hat{y}_1, \dots, \hat{y}_{N_{val}})$ different from the training set. The predicted $\hat{\mathcal{Y}}_{val}$

values are compared to the corresponding RELAP5-3D outputs \mathcal{Y}_{val} through the construction of Quality Indicators (QIs). No definitive guidelines are found in literature about the correct size N_{val} of the validation set. [71] suggests $N_{val} \gg N_{train}$, since a small N_{val} can be misleading in case validation samples are taken, by chance, too close to training samples where the metamodel is clearly more refined [72]. However, this approach becomes extremely expensive in case of time-demanding simulators. [73] proposes a “sequential validation design” to get to a meaningful validation while keeping N_{val} small: validation is carried out gradually by adding validation samples in the unfilled regions of the input space to optimize the distance between the validation set and the training set. Here two validation sets are considered for the analysis of the energy output. The first one derived from 55 simulated transients, with output $E_{ex,\%} = 85 \div 95 \%$, and it is used to verify the metamodel accuracy around the limit surface. The second one includes 138 I/O simulated transients, with $E_{ex,\%}$ values spreading all over the domain and it is employed to obtain an indication of the metamodel accuracy over the entire domain. The QIs used to quantify metamodel accuracy with respect to both validation sets are the well-known RMSE and two different predictivity indicators, Q_1 , defined in [73], and Q_2 presented by [65]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_{val}} (\hat{y}_i - y_i)^2}{N_{val}}} \quad (8)$$

$$Q_1 = 1 - \frac{\sum_{i=1}^{N_{val}} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{N_{val}} (\bar{y}_{val} - y_i)^2} \quad (9)$$

$$Q_2 = \frac{N_{val} - 1}{N_{val}} \left(\frac{\sum_{i=1}^{N_{val}} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{N_{val}} (\bar{y}_{val} - y_i)^2} \right), \quad (10)$$

where y_i is the i -th output of $\{\mathcal{X}_{val}, \mathcal{Y}_{val}\}$, \hat{y}_i is the corresponding metamodel prediction and \bar{y}_{val} is the mean value of all the real model outputs in the validation set. RMSE and Q_2 should be as low as possible, whereas Q_1 tends to 1 as the

prediction accuracy increases. Notice that the RMSE has the same unit of measure of the physical quantity of interest ($E_{ex,\%}$) and, thus, it can be progressively compared to the exchanged energy output to understand whether the predictions are satisfactory. It can be also normalized (Normalized Root-Mean-Square Error, NRMSE) dividing it by \bar{y}_{val} . Q_1 and Q_2 have similar expressions and, differently from RMSE, they take into account also the variability of the output in the validation set. The values of these QIs should be improved as much as possible through the successive iterations of the algorithm. In this work the convergence (stopping) criterion related to the metamodel accuracy is considered satisfied when the NRMSE evaluated on the “local” validation set constructed around Y_{thres} becomes about 2%. On the other hand, if the metamodel quality is still unsatisfactory, the second convergence (stopping) criterion needs to be checked: further computational time is required to add a new algorithm iteration and to simulate new configurations by the BE-TH code in order to enrich the training set. The computational budget, i.e., the maximum number of simulations initially foreseen, has been here fixed to 100 RELAP5-3D simulations in addition to the initial ones (i.e., the 64 simulations belonging to $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$). When the computational budget is completely run out, even if Kriging accuracy is still not satisfactory, the procedure stops.

5. Selection: if algorithm convergence has not been reached at step 4, new I/O simulations related to the so-called best candidates \mathcal{X}^* are run and the inputs and outputs added to the training set to refine the metamodel. The N_{cand} best candidates are selected among the \mathcal{X} generated at step 2 according to their U -function values. Combinations with $U < 1.96$ are sorted in ascending order according to their predicted output value \hat{y} and, then, organized in N_{cand} *equally-spaced bins*. Then, one candidate is randomly picked from *each bin*. This procedure is implemented to avoid selecting candidates “clustered” in the same area of the input space (i.e., too similar to each other). Actually, combinations that are close in the input space share similar U values; hence, selecting the candidates only according to the N_{cand} lowest U values would cause them to be restricted in the same area of the domain, instead of

spanning the whole input space. The selection procedure is illustrated in Fig. 2. $N_{cand} = 7$ or 8 samples have been chosen as a satisfactory trade-off between computational cost, number of iterations of the algorithm and metamodel accuracy (see 1.3.1.1). Indeed, lower values of N_{cand} would require a larger number of algorithm iterations and training repetitions (i.e., higher computational cost) to obtain the same Kriging accuracy; also, an excessively small number of candidates implies a rougher exploration and “mapping” of the area close to the limit surface. On the other hand, limiting the number N_{cand} is useful in particular in the first iterations, when the metamodel is still inaccurate. Selecting many candidates according to its predictions may lead to a waste of computational time: actually, some candidates, simulated with the time-demanding BE-TH model, may later reveal to be not so useful for the scope of the analysis (e.g., they may lie far from the limit surface). In Fig. 2, the value $E_{ex,\%}$ is reported on the x -axis, whereas y -axis displays the corresponding U -function values. It is clear from the dashed vertical lines how the bins are constructed by subdividing the x -axis in segments of the same length. \mathcal{Y}^* values are represented by diamonds, whereas all the samples with $U < 1.96$ are shown as crosses. The shape of the graph shows that the closer is a point to $Y_{thres} = 90\%$, the lower its U value is; this was easily foreseeable looking at equation (7). Once the best candidates \mathcal{X}^* have been selected and the corresponding I/O transients simulated with the BE-TH model to obtain the output \mathcal{Y}^* , the training set is enriched and steps 1 to 5 are repeated until convergence at step 4 is reached.

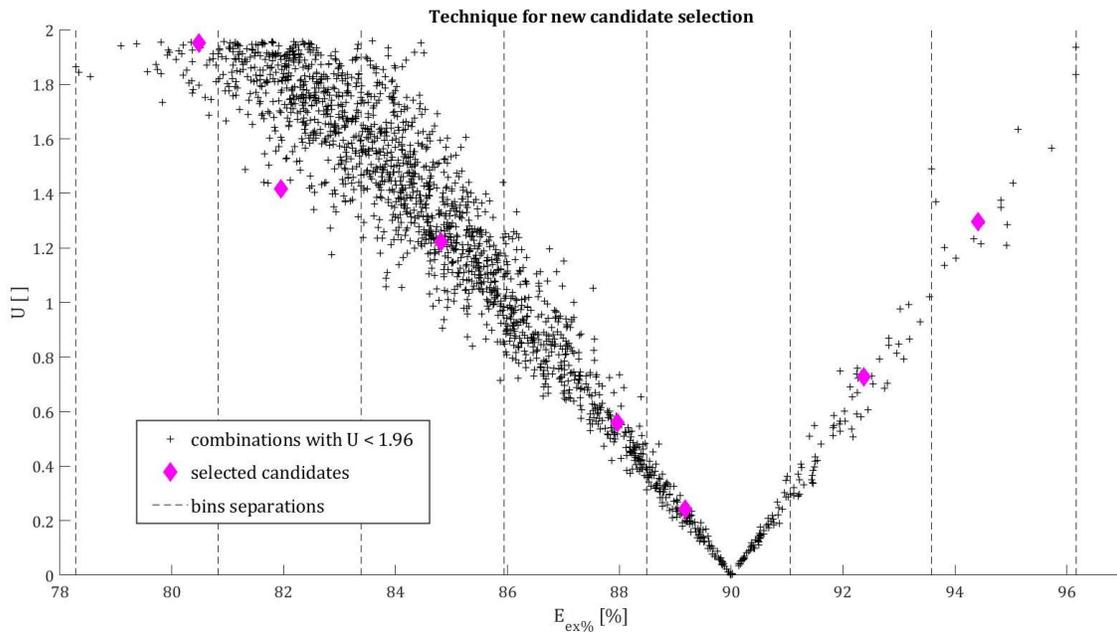


Figure 9: Identification of the best candidates (Step 5): random selection of input combinations with $U < 1.96$ divided in equally-spaced bins

3.2.3 Metamodel Accuracy Evaluation and CR Characterization

In the present section, the results of the application of the AK-MCS procedure are measured and, then, they are exploited to identify the CRs with respect to the FC “Low heat removal” concerning the energy exchanged output (E_{ex}). In particular, at first, the accuracy level reached by the Kriging metamodel at the end of the AK-MCS iterations is quantified by exploiting some indicators (e.g., standard deviations of metamodel predictions, RMSE, etc.) (3.2.3.1). Second, the Kriging metamodel is used to predict the outputs of a large number of new input combinations to find the critical ones and retrieve information about the shape of CRs. In the end, a graphical representation of the CR identified is provided (3.2.3.2).

3.2.3.1 Kriging Accuracy Evaluation

The AK-MCS procedure has been stopped at iteration $n_{fin} = 14$, after enriching the initial training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ with the inputs and outputs of 100 RELAP5-3D (i.e., maximum computational budget available). In Table 6 the salient aspects of each n -th iteration are reported.

Table 6: results of the AK-MCS procedure applied to the energy output

n	N_{train}	$\bar{\sigma}_1$	$\bar{\sigma}_2$	ϵLOO_{norm}	ϵLOO_{abs}
0	64	11.76	11.67	0.128	20.62
1	71	11.09	10.99	0.130	19.39
2	78	9.41	9.38	0.138	19.05
3	85	8.83	8.80	0.148	18.76
4	92	8.39	8.35	0.152	18.36
5	99	7.10	7.06	0.159	18.24
6	106	7.60	7.55	0.169	18.18
7	113	7.50	7.46	0.169	17.36
8	121	7.35	7.30	0.167	16.37
9	129	7.41	7.36	0.166	15.61
10	136	7.23	7.18	0.176	16.61
11	143	7.01	6.96	0.175	15.98
12	150	4.56	4.21	0.177	15.66
13	157	4.61	4.10	0.126	10.81
14	164	4.47	3.92	0.119	9.69

In the 3rd and 4th columns, two average standard deviations ($\bar{\sigma}$) are reported; they are calculated with respect to two different sets of metamodel outputs $\hat{\mathcal{Y}}$: $\bar{\sigma}_1$ in column 3 is evaluated with respect to the outputs of all the combinations generated at step 2 of the AK-MCS procedure, which are spread throughout the domain; instead, $\bar{\sigma}_2$ in column 4 is calculated using only the outputs of those combinations for which $U < 1.96$. Both the measures have been used in the successive iterations of the algorithm in order to check the Kriging gradual refinement, without resorting to the computationally expensive validation set. In particular, $\bar{\sigma}_1$ allows following the improvement in the metamodel accuracy over the entire domain, whereas $\bar{\sigma}_2$ is used to focus on the accuracy increase nearby the limit surface. As expected, the metamodel general improvement during the iterations makes both average standard deviations decrease. However, $\bar{\sigma}_2$ diminishes more rapidly due to

the nature of the AK-MCS algorithm, which adds new I/O data with outputs close to Y_{thres} (where U is lower), thus making the predictions more accurate in proximity of the limit surface than elsewhere in the domain (e.g., at iteration $n_{fin} = 14$, $\bar{\sigma}_1$ is equal to 4.47, while $\bar{\sigma}_2$ is equal to 3.92). The Kriging settings (see Section 3.2.1) have been adjusted from iteration 12 onwards by changing the correlation function family from *Matérn 5/2* to *Exponential*, in order to improve the fit with the new, expanded training set; indeed, looking at the evolution of the two average standard deviations up to that point, it can be noticed that the corresponding values were not decreasing anymore and the metamodel improvement seemed stuck.

The last two columns of Table 6 report the Leave-One-Out (LOO) error evolution with iterations: column 5 shows the LOO error directly returned by the UQLab tool, also called normalized LOO error (ϵLOO_{norm}), whereas column 6 reports its absolute version (ϵLOO_{abs}):

$$\epsilon LOO_{norm} = \frac{1}{N_{train}} \left[\frac{\sum_{i=1}^{N_{train}} (y(x_i) - \hat{y}_{(-i)}(x_i))^2}{Var[\mathcal{Y}_{train}]} \right] \quad \epsilon LOO_{abs} = \epsilon LOO_{norm} \cdot Var[\mathcal{Y}_{train}], \quad (11)$$

where $\hat{y}_{(-i)}(x_i)$ is the prediction made by the metamodel in correspondence of the i -th combination $x_i \in \mathcal{X}_{train}$ and obtained using all the $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ pairs of values available, except $\{x_i, y_i\}$, and $Var[\mathcal{Y}_{train}]$ is the variance of the training outputs. The only difference is in the term $Var[\mathcal{Y}_{train}]$ representing the output variability in the training set. The LOO error is generally used (as the $\bar{\sigma}$ values previously introduced) to assess model accuracy when there is no availability of an external validation set due to its high computational cost; thus, the evolution of these two quantities has been followed since it gives an idea about the Kriging progressive refinement. A gradual decrease was expected, but what occurs in reality is that ϵLOO_{norm} initially rises and, then, sharply drops reaching its lowest value at $n_{fin} = 14$. This behaviour is justified by equation (11): the metamodel becomes progressively more refined, causing the numerator of ϵLOO_{norm} to decrease; however, at the same time, also $Var[\mathcal{Y}_{train}]$ at the denominator diminishes, because the I/O data are all selected with outputs close to Y_{thres} and hence the variability of \mathcal{Y}_{train} reduces. On the other hand, ϵLOO_{abs} , not containing the $Var[\mathcal{Y}_{train}]$ term, shows a more regular (decreasing) trend.

The best way to follow the Kriging accuracy improvement with iterations, if enough computational power is available, is to construct an external validation set and to evaluate the corresponding QIs by computing the Kriging predictions with respect to the validation data (as explained at step 4 of the AK-MCS procedure). For the analysis of the energy output, two validation sets have been constructed (see Section 3.2.2), one with output values close to the limit surface, named “local”, and the other with output values more spread over their domain, named “global”. Three QIs have been computed for each validation set: RMSE, Q_1 and Q_2 (see equations (8), (9) and (10)). The QIs evolution is illustrated in Figs. 10 and 11. Note that the RMSE in Fig. 10 is expressed in percentage because it has the same unit of measure of the predicted output, i.e., the percentage energy exchanged ($E_{ex,\%}$); but, it should not be confused with the NRMSE.

The two curves in each plot (Fig. 10, and Fig. 11(a) and 11(b)) are referred to different validation sets (local and global), but they all show the same general trend: a decrease in RMSE and Q_2 , and an increase towards 1 for Q_1 , representing the improvement of the metamodel accuracy. The three curves with circles associated to the validation set whose outputs are near Y_{thres} (local) display a faster improvement in accuracy because of the nature of the metamodel-based AK-MCS procedure, which gradually makes the metamodel more refined around the failure threshold. The diamond of different color in correspondence of the 12th iteration symbolizes the change of Kriging setting (as already explained the correlation function family has been changed from *Matérn 5/2* to *Exponential*). The two curves in Figs. 11(a) and 11(b) cross because of Q_1 and Q_2 mathematical expressions. For example, for what concerns Q_2 , being the metamodel at the beginning still inaccurate, the numerator in equation (10) is small for both the local and global validation sets; on the contrary, the denominator is obviously larger for the global validation set, with data spread all over the domain than the local validation set. This is the reason why at the beginning the Q_2 value is lower (and hence better) when evaluated with respect to the most various validation set (curve with crosses in Fig. 11(b)) than with the local validation set (curve with circles in Fig. 11(b)), unlike what is observed at the end of the AK-MCS procedure. The same behavior with respect to the two validation sets is observed also for Q_1 , but in the opposite way (see equation (9)).

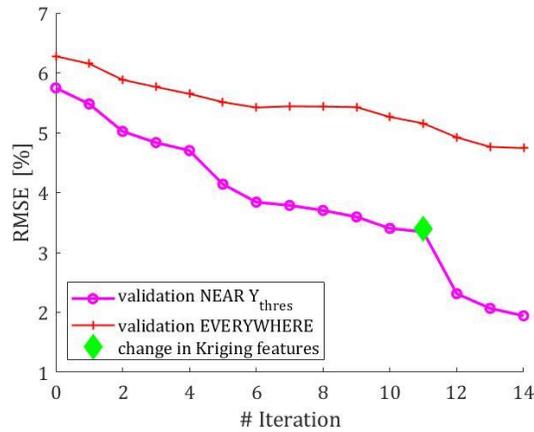


Figure 10: RMSE evaluated with respect to two validation sets (energy output)

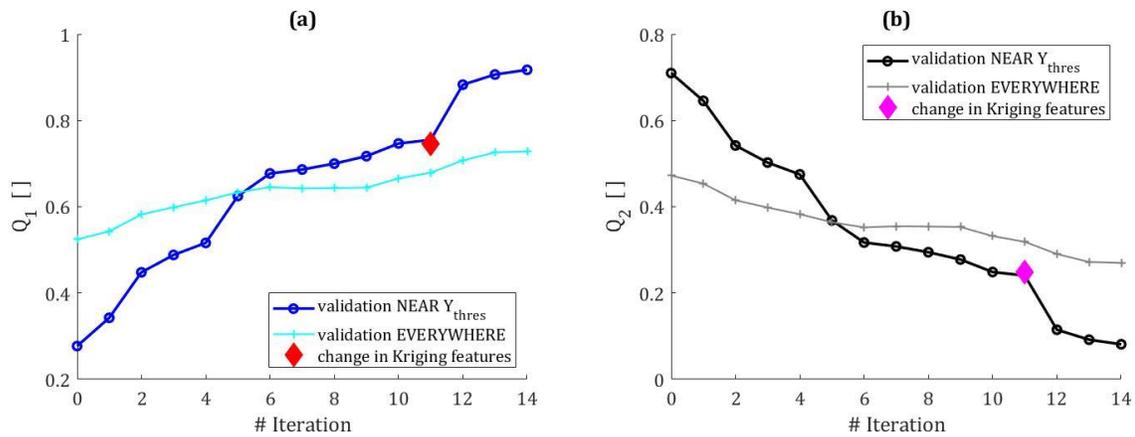


Figure 11: Q_1 (a) and Q_2 (b) predictivity indicators evaluated with respect to the two validation sets (energy output)

Table 7 reports all the QI values at the last iteration, with the RMSEs computed in both the absolute and normalized forms:

Table 7: QIs at the end of AK-MCS procedure - 14th iteration (energy output)

Quality indicator (QI)	RMSE [%]	NRMSE [%]	Q_1	Q_2
Validation EVERYWHERE	4.74	5.95%	0.728	0.270
Validation NEAR Y_{thres}	1.94	2.24%	0.917	0.081

The final results are generally satisfying: a *RMSE* of 1.94 is acceptable if compared to $E_{ex,\%}$, which usually varies from 70% to 100% in the simulated transients. A value of 2.24% for the *NRMSE* is remarkable since it could be taken, at first instance, as a measure of the percentage error in prediction: an error around 2% near the limit surface is considered satisfactory.

3.2.3.2 CRs Characterization with respect to the FC “Low Heat Removal”

The Kriging metamodel obtained at the end of the AK-MCS iterations has been demonstrated capable of predicting the outputs in correspondence of new input combinations with a good accuracy level. Thus, the Kriging has been exploited to predict the outcomes of a large number (10.000) of new input combinations x in order to: (i) find the critical ones, with reference to the PSS function, i.e. $\hat{y} = f(x) \leq Y_{thres}$; and (ii) retrieve information about the shape of the CRs of the PSS operation. Given that the input space dimensionality is $M = 5$, a graphical representation of the unique CR identified is provided through a series of scatter plots with paired inputs representing the CR two-dimensional projections (Fig. 12): green diamonds are used to indicate combinations leading to a safe operation (i.e., $E_{ex,\%} \geq 90\%$), whereas red crosses represent critical inputs when the PSS fails its function.

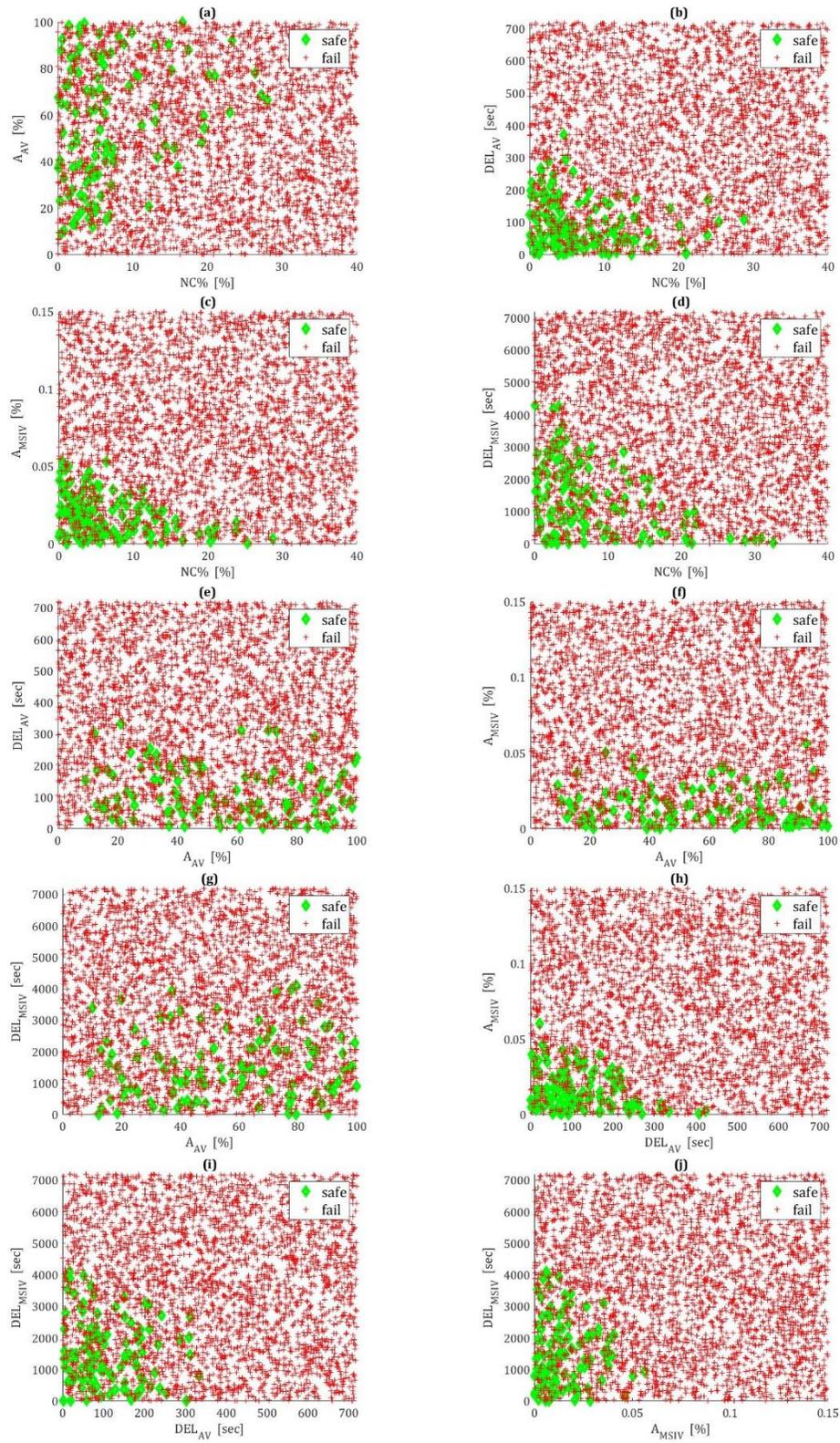


Figure 12: Scatter plots of the PSS CR (energy output)

The subplots in Fig. 12 allow catching the influence of the five input parameters on the amount of energy exchanged (E_{ex}) by the PSS; in particular, each scatter plot provides information about the effect of the interaction between the two input parameters therein represented (whatever the values of the other parameters). The results show that only four of the five input parameters have significant influence on E_{ex} . Indeed, A_{AV} is not very relevant for driving the PSS response in terms of the energy exchanged: in fact, whatever its value, the DHR function may or may not be successfully accomplished (green diamonds or red crosses, indifferently). The interaction of $NC\%$ with the other input parameters (except A_{AV}) is shown in Fig. 12(b), (c) and (d). In particular, in all the subplots the maximum $NC\%$ value corresponding to a combination of functional success is around 30%, suggesting that PSS fails to provide its function whenever $NC\% > 30\%$, independently of the values of the other parameters. This is coherent with the underlying physics: the presence of non-condensable gases leads to a reduction in the heat transfer coefficient during condensation in the E-HX (see Section 2.1) and, in fact, the higher $NC\%$, the worse the impact on E_{ex} . According to Fig. 12(b), (c) and (d), the upper limit for $NC\%$ is generally reduced in case of interactions with other parameters that reduce the E_{ex} value, i.e., variations of DEL_{AV} , DEL_{MSIV} and A_{MSIV} from the values in the reference conditions (see Table 3). The results are represented by triangle-shaped safe region (green diamonds). For example, a value $A_{MSIV} = 0.035\%$ represents a leakage in the MSIV that reduces the amount of steam directed into the PSS, thus lowering E_{ex} : in this situation, the maximum value of $NC\%$, for which the PSS function can still be successfully accomplished, is about $NC\% = 15\%$ (whatever the values taken by the other three parameters). Also, DEL_{AV} plays a central role in E_{ex} determination. Indeed, if AV opens with a certain delay, i.e., $DEL_{AV} > 0$, the whole heat transfer process is delayed and this impacts severely on E_{ex} , especially because the largest amount of energy is exchanged in the first part of the accidental transient. Moreover, if AV is not open, the PV pressure may rise, which causes the opening of the SRV and, hence, vapor discharging into the containment instead of condensing inside the PSS. Looking at subplots 12(b), (e), (h) and (i), the maximum DEL_{AV} for which the PSS function can still be accomplished is about 400 s: however, this upper limit is, in general, lowered in case of interactions with other parameters (as for input $NC\%$). For example, when varied together with A_{MSIV} (see subplot 12(h)), again a triangle-shaped region of safe

function is observed: e.g., if $A_{MSIV} = 0.025\%$, the maximum DEL_{AV} value for successful function is about 200 sec. For what concerns DEL_{MSIV} , the observed upper limit is about 4000 s. The MSIV closure is necessary to isolate the turbine side and start sending the vapor into the PSS for condensation; hence, if the closure is delayed, less vapor enters the PSS in the first part of the transient and E_{ex} is reduced. Whereas a priori the expected interaction of DEL_{MSIV} with other input parameters negatively affecting E_{ex} could have been a decrease in the value of DEL_{MSIV} leading to the PSS functional failure. Instead, what is observed from the simulations in case of interaction with, e.g., DEL_{AV} or A_{MSIV} is different (see subplots 12(i) and (j)). In such cases, the safe region is square-shaped. For example, focusing on the interaction between DEL_{MSIV} and DEL_{AV} , the upper limit of DEL_{MSIV} should, in general, decrease if $DEL_{AV} > 0$ s, independently of the values of the other three parameters; instead, the upper limit remains about 4000 s (except in the extreme case where DEL_{AV} reaches its own upper limit causing PSS functional failure by itself). This behaviour (and the consequent square-shaped regions) is probably due to the influence of DEL_{MSIV} on E_{ex} , which is more significant than that of other input parameters. For example, considering the interaction between DEL_{MSIV} and A_{MSIV} , the possible presence of a certain leakage in the MSIV after its closure (i.e., $A_{MSIV} > 0\%$) is less relevant in terms of contribution to the amount of energy exchanged, if it occurs in case of a significant delay in the MISV closure, which is much more influential.

Other interesting conclusions about the CR can be inferred from the PCP (Fig. 13). PCP allows displaying in a unique plot all the five input values corresponding to each combination; each of the five vertical axes reports the values of one input parameter (normalized on its range) and, hence, one input combination is represented by a line connecting in the horizontal direction the corresponding input values on the different axes.

In particular, Fig. 13 shows the quantile representation of all the predicted combinations together: the solid blue lines are representative of the PSS safe operation, whereas the dashed orange lines represent the input combinations leading the PSS to fail its function. In particular, in both cases (solid or dashed lines) the line in the middle stands for an “average” combination (average safe combination or average failure combination), whereas the other two external lines are its 0.25 upper and lower quantiles (also called quartiles).

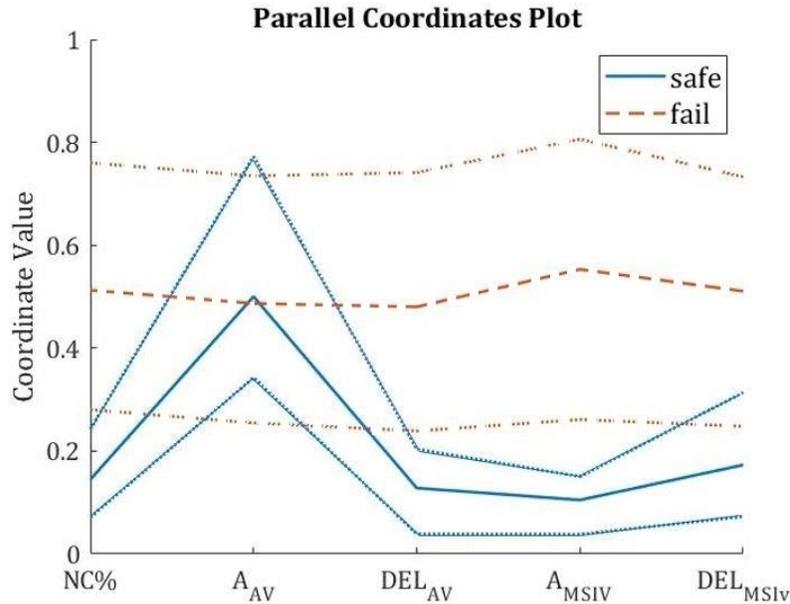


Figure 13: PCP with 0.25 quantiles (energy output)

The contribution of A_{AV} to safe function immediately stands out: the normalized value associated to the blue curve in the middle is about the same as the one observed for functional failure (orange curve in the middle). This confirms the scarce importance of A_{AV} in the providing of the PSS function, as already seen from Fig. 12, but here in a more quantitative way. Moreover, the solid blue lines are mainly located in the lower part of the graph, close to the 10% of the range of variation of each input parameter (except for A_{AV} , as previously mentioned); on the contrary, the dashed orange lines occupy quite a large portion (in particular, the middle part) of these intervals, meaning that the majority of the input combinations leads to PSS functional failure. The same result can be deduced from Fig. 12, where the red crosses are far more abundant than the green diamonds. A word of caution is in order in this respect. This result does not mean that the PSS analysed is more prone to fail than to succeed, since this type of conclusion should be supported by a *probabilistic analysis* of the occurrence of the input combinations, which is not carried out in this work. Actually, the probability estimated for the event that the PSS fails its function strongly depends on: (i) the characteristics of the system itself, and (ii) the (data and/or expert-based) probability distributions of the PSS input parameters. In this application, as mentioned in the Introduction and Section 2.1, no realistic probability distributions are

assigned to the PSS parameters, since the objective is not to carry out a reliability analysis of the PSS, but to describe how the metamodel-based AK-MCS procedure can be exploited for critical regions characterization.

To sum up, the metamodel-based AK-MCS has been demonstrated capable of approximating E_{ex} output with good accuracy, within the aim of exploring the PSS CRs relative to the FC “Low heat removal”, by resorting to an I/O training set of 164 RELAP5-3D simulations. A unique CR has been identified and correctly visualized with multiple scatter plots which show how the PSS is more prone to failure in case of *combined action* of the input parameters varying together, rather than in case of one-at-a-time variations. Moreover, most of the input combinations predicted by the Kriging metamodel have been found to lead the PSS to failure, suggesting that, in a hypothetical design phase, the variations of such relevant inputs should be limited only on a small portion of the explored ranges. To conclude, these results could be even exploited within a more general reliability analysis for PSSs.

4 Novel Framework for System CRs Characterization in case of State-space Non-smoothness and Multimodality

The exploration of a system state-space and the characterization of the conditions leading to failure (CRs) is a challenging problem, not only due to the need of replacing the computationally demanding system codes typically available with fast-running metamodels to speed up the calculation, but also due to the possible irregularity of the state-space. Indeed, in case of non-smoothness and/or multimodality of the state-space, traditional metamodeling techniques, like the Kriging metamodel developed within an AK-MCS procedure (see [Chapter 3](#)), lead to large errors because they require the approximated function to show accommodating properties of regularity (e.g., continuity and smoothness) [18]. This is the case of the pressure output (p_{max}) identified as relevant for the analysis of the PSS described in [Chapter 2](#). The maximum value reached by the pressure inside the PV is calculated during the accidental transient (lasting about 8h) and it has shown a strong non-smooth and multimodal distribution (see [Fig. 8](#)), with at least two peaks of concentration: one in correspondence of $p_{max} = 70.0 \text{ bar}$ and the other around the failure threshold (75.5 bar).

Chapter 4 / Novel framework for System CRs characterization in case of state-space non-smoothness and multimodality

In the first part of this chapter, a novel framework for exploring the state-space of a generic system, for which a time-demanding model is available and under the specific condition of non-smooth and multimodal output is proposed. The main goal is to iteratively run a (possibly low) number of time-demanding simulations to construct an accurate metamodel not suffering for the output non-smoothness and multimodality. Then, the metamodel is exploited to predict the outputs corresponding to a large number of input combinations, which are then used to retrieve information about the CRs characteristics.

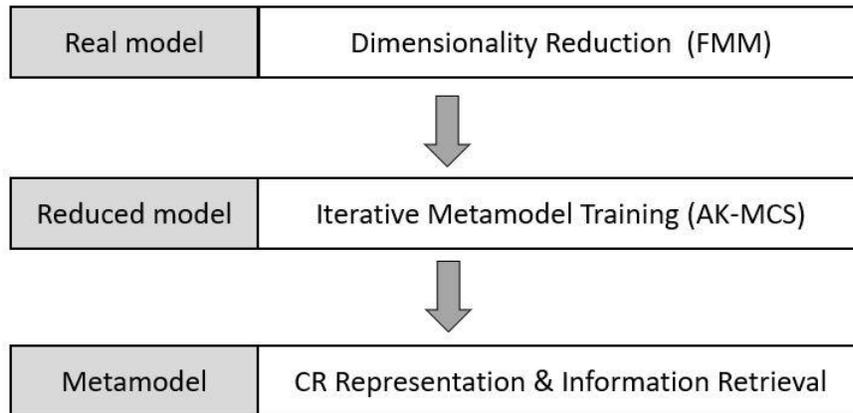


Figure 14: flow diagram of the exploration framework proposed for tackling the output non-smoothness and multimodality

The framework consists in three main steps, as reported in [Fig. 14](#). In short, the first step, i.e., “dimensionality reduction”, aims at identifying the input parameters most affecting the output distribution, specifically those related to the output clusters in correspondence of the failure threshold, being the ones related to the CRs. We resort to a DBSA method supported by FMM technique. The second step, i.e., “iterative metamodel training”, aims at iteratively constructing an accurate and fast-running metamodel to replace the real model simulations on the reduced input space, with specific attention to the boundary of the CR (limit surface), to identify the critical input combinations providing a limit for the system safe operation. The metamodel accuracy is verified (e.g., exploiting a validation set) and, then, in the third step, i.e., “CR representation & information retrieval”, the metamodel is employed to evaluate the outcomes of a large number (several thousands) of new input combinations to retrieve information about the CRs, like their number and shape. Finally,

the CRs are graphically represented by exploiting high-dimensional data visualization techniques, like scatter plots or PCP.

The details about this framework are provided in [Section 4.1](#) which is organized in three subsections, one for each step of the framework: at first, dimensionality reduction step is presented with the details of the FMM construction ([4.1.1](#)); then, a brief recap of the AK-MCS technique (described in detail in [Section 3.1.2](#)) is provided ([4.1.2](#)) and, finally, it is explained how the Kriging metamodel obtained at the end of AK-MCS iterations can be exploited for CRs characterization ([4.1.3](#)).

In [Section 4.2](#), this novel framework has been applied for the characterization of the PSS CRs with respect to the FC “Steam release in the containment”, defined according to p_{max} . In this case, the time-demanding system model to be replaced is a BE-TH model developed in RELAP5-3D (see [Section 2.2](#)); the DBSA method chosen for dimensionality reduction is the Hellinger distance method [46], [54] and the metamodel is iteratively trained, again, according to an AK-MCS procedure (see [Chapter 3](#)). Finally, scatter plot technique is exploited for the CR visualization. In detail, at first, the FMM approximation has been constructed and used in support of the Hellinger distance method to rank the five input parameters initially selected for the analysis of the PSS (see [Section 2.1](#)) ([4.2.1](#)). Then, after reducing the dimensionality of the problem, a Kriging metamodel has been constructed and refined through an AK-MCS procedure properly tailored for the analysis of p_{max} ([4.2.2](#)). The results of the application are presented both in terms of the accuracy level reached by the metamodel and graphical representation of the CR ([4.2.3](#)). Finally, an alternative approach found in literature [18], based on the combination of a SVC and AK-MCS technique, is briefly presented and the results of its application compared to those of the framework previously introduced (FMM+AK-MCS), in terms of accuracy and computational cost ([4.2.4](#)).

4.1 Novel Framework Presentation

4.1.1 Dimensionality Reduction

The purpose of dimensionality reduction is to find a lower-dimensional subspace of variables, i.e., $\mathbf{X}^R \in D_{\mathbf{X}^R} \subset \mathbb{R}^R$ (where $R < M$ is the reduced dimensionality of the problem), to build a reduced model still capable of correctly representing the system behavior [42]. From the point of view of exploring the state-space for CRs characterization, reducing the dimensionality means finding a more effective I/O training set to construct a more accurate metamodel. In case of non-smooth and/or multimodal output this is still valid, but with specific attention given to those variables contributing most to the output clusters corresponding to system failure conditions. The dimensionality reduction step of the proposed framework combines the use of a SA method to rank the input parameters according to some sensitivity indices, and the application of FMM technique to approximate the output non-smooth and multimodal distribution. For the sake of clarity, these two fundamental ingredients are treated separately in the following subsections (4.1.1.1 and 4.1.1.2).

4.1.1.1 Sensitivity Analysis

Several examples are available in literature about how to carry out a dimensionality reduction and many of them rely on SA techniques. Among the SA techniques, it is possible to identify two families: Local and Global. The Local approach to SA considers small variations of each input parameter around its nominal value, whereas Global SA [45] allows to quantify the contribution of an input to the variability of the output, computed over the entire range of both the input and the output. Global SA offers higher capabilities, especially when model responses are not regular (e.g., non-linear and non-monotone), but at a higher computational cost [46].

Global SA methods can be divided into three categories: 1) RBSA methods, 2) VBSA methods and 3) DBSA methods (See Table 1). RBSA or non-parametric methods [47] exploit

regression techniques to fit a regression model on a set of I/O relations and to use the regression coefficients as indices of sensitivity. RBSA methods are typically the simplest ones, also associated to the lowest computational cost, but their success strongly depends on the output form which is often required to be linear. Indeed, if the regression model does not fit the underlying I/O relationships (e.g., in case of non-smoothness), the SA performs really poorly. Examples of widespread RBSA methods are the SRC and PCC methods already introduced in 1.3.2.2. VBSA methods [74] quantify the contribution of each input parameter (first-order effect) and each possible two- or high-order interactions among multiple parameters to the total output variance; hence, the ratio of each contribution to the total variance is interpreted as a sensitivity coefficient [75]. VBSA methods are the most widespread, because they do not introduce any hypothesis on the model since they do not carry out any approximation of it. Anyway, VBSA methods are unable to distinguish between output structures (i.e., how the output values are organized in the state-space) having identical global variance, but different distributions and spatial organizations [75]. Thus, they may suffer for output multimodality since, by definition, the calculation of variance in case of a multimodal variable is not trivial. The most famous VBSA method is the Sobol' method [49], [50], [51]. Finally, DBSA or moment-independent methods rank the input variables most affecting the *entire* output distribution and they may overcome the issue of non-smoothness and multimodality if the output distribution is properly approximated, despite its irregular form, by means of FMM technique [57]. FMM are classically implemented for pattern recognition to approximate the output distribution, even in case of multimodality, by identifying the output clusters (corresponding to the different output modes) and, hence, representing the output as a linear combination of known distributions, also called components (e.g., Gaussian, Exponential, etc.). Anyway, FMM can be also adopted as a support for SA: indeed, the output clustering is mapped to the input space and, in the end, the contribution of each input to the clustering of the output is ranked according to the different DBSA methods. Examples of DBSA methods are the already mentioned δ indicator [52], input saliency [53], Hellinger distance [54] and Kullback-Liebler divergence [54].

4.1.1.2 Finite Mixture Models Approximation

FMM application for SA entails following at the beginning the same procedure adopted in case of the more general pattern recognition: the primary goal is to find the appropriate type and number of components (k) to approximate the output distribution, given a set of I/O relations. The best k is historically determined through the application of the Expectation Minimization (EM) algorithm [76].

Here is provided a description of the FMM construction process through the EM algorithm according to [77]. Let assume a set of N_{train} output variables $\mathcal{Y}_{train}^{FMM} = \{y_1, \dots, y_{N_{train}}\}$, the generic y_i is said to follow a k -component finite mixture distributions if its PDF can be written as:

$$p(y_i|\Theta) = \sum_{j=1}^k \pi_j p_j(y_i|\Theta_j), \quad (12)$$

where $\{\pi_j, j = 1, \dots, k\}$ are the mixing parameters or weights, Θ_j are the parameters of each j -th component and $\Theta = \{\pi_1, \dots, \pi_k, \Theta_1, \dots, \Theta_k\}$ is the complete set of mixture parameters; being probabilities, π_j must satisfy:

$$\sum_{j=1}^k \pi_j = 1. \quad (13)$$

Considering the set of samples \mathcal{Y}_{train} , the log-likelihood corresponding to a k -component mixture is:

$$\log p(\mathcal{Y}_{train}^{FMM}|\Theta) = \log \prod_{i=1}^{N_{train}} p(y_i|\Theta) = \sum_{i=1}^{N_{train}} \log \sum_{j=1}^k \pi_j p(y_i|\Theta_j), \quad (14)$$

and the related ML estimate reads:

$$\hat{\Theta} = \arg \max_{\Theta} \{ \log p(\mathcal{Y}_{train}^{FMM}|\Theta) \}. \quad (15)$$

$\hat{\Theta}$ cannot be found analytically since it implies to solve a non-linear equations system. Hence, the solution is provided through the application of EM Algorithm which interprets $\mathcal{Y}_{train}^{FMM}$ as a set of incomplete data. The “missing part” is represented by a set of labels, i.e.,

Chapter 4 / Novel framework for System CRs characterization in case of state-space non-smoothness and multimodality

$\mathcal{Z} = \{\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(N_{train})}\}$, associated to the y_i values numbered N_{train} , where each i -th label is a binary vector, i.e., $\mathbf{z}^{(i)} = \{z_1^{(i)}, \dots, z_k^{(i)}\}$, whose components are all zeros except for $z_j^{(i)} = 1$, i.e., the $\mathbf{z}^{(i)}$ component associated to the j -th distribution of the mixture that has generated y_i . Now, the complete log-likelihood for the estimation of $\hat{\Theta}$ can be written as:

$$\log p(\mathcal{Y}_{train}^{FMM}, \mathcal{Z} | \Theta) = \sum_{i=1}^{N_{train}} \sum_{j=1}^k z_j^{(i)} \log[\pi_j p(y_i | \Theta_j)]. \quad (16)$$

The EM Algorithm provides a sequence of estimates $\{\hat{\Theta}(t) \text{ with } t = 0, 1, 2 \dots\}$ through the alternate realization of two steps, until some convergence criterion is satisfied:

- **E-step:** given the $\mathcal{Y}_{train}^{FMM}$ estimate through the current $\hat{\Theta}(t)$, and considering that $\log p(\mathcal{Y}_{train}^{FMM}, \mathcal{Z} | \Theta)$ is linear with respect to \mathcal{Z} , the conditional expectation of the log-likelihood is computed through the construction of the so-called Q -function by simply evaluating the conditional expectation, i.e., $W \equiv E[\mathcal{Z} | \mathcal{Y}_{train}^{FMM}, \hat{\Theta}(t)]$, and plugging it into $\log p(\mathcal{Y}_{train}^{FMM}, \mathcal{Z} | \Theta)$:

$$Q(\Theta, \hat{\Theta}(t)) \equiv E[\log p(\mathcal{Y}_{train}^{FMM}, \mathcal{Z} | \Theta) | \mathcal{Y}_{train}^{FMM}, \hat{\Theta}(t)] = \log p(\mathcal{Y}_{train}^{FMM}, W | \Theta). \quad (17)$$

Knowing that $z_j^{(i)}$ coefficients are of binary kind, Bayes law can be exploited to calculate their conditional expectation:

$$w_j^{(i)} \equiv E[z_j^{(i)} | \mathcal{Y}_{train}^{FMM}, \hat{\Theta}(t)] = \Pr[z_j^{(i)} = 1 | y_i, \hat{\Theta}(t)] = \frac{\hat{\pi}_j(t) p(y_i | \hat{\Theta}_j(t))}{\sum_{m=1}^k \hat{\pi}_m(t) p(y_i | \hat{\Theta}_m(t))}, \quad (18)$$

where π_j and $w_j^{(i)}$ are the respectively the a priori probability and a posteriori probability, after observing y_i , that $z_j^{(i)} = 1$.

- **M-step:** the mixture parameters are updated, under the constraints introduced by [equation \(13\)](#), according to:

$$\hat{\Theta}(t+1) = \arg \max_{\Theta} \{Q(\Theta, \hat{\Theta}(t))\} \quad (19)$$

Anyway, classical EM algorithm presents several drawbacks: it is a local method, thus, it is sensitive to initialization and, for certain kinds of mixtures, it may converge toward an

estimate at the boundary of parameter space where the likelihood is unbounded. One possible solution found in literature is represented by the so-called “merge and split” technique [78], which eliminates the necessity of a careful initialization of EM algorithm and avoids the problem of getting stuck in local minima. Searching for the best mixture of models, local minima may arise if there are too many components in one area and too few in another; indeed, points with lower likelihood are found when trying to move one component from the overpopulated region to the underpopulated region. Split and merge technique manages to tunnel through log-likelihood barriers by merging two components in the overpopulated region and simultaneously splitting a component in the underpopulated region. For the exploration framework presented in this thesis, it is proposed the SNOB algorithm, introduced for the first time in [58] and then updated through the years and implemented in MATLAB by [79]. This SNOB version is based on the EM algorithm supported by the merge and split technique and exploits the MML inference criterion to select the best number of components k :

$$I(\Theta | \mathcal{Y}_{train}^{FMM}) = I(k) + I(\boldsymbol{\pi}) + \sum_{j=1}^k I(\Theta_j) + I(\mathcal{Y}_{train}^{FMM} | \Theta), \quad (20)$$

where $\mathcal{Y}_{train}^{FMM} = \{y_1, \dots, y_n\}$ are the output values of the transients simulated and $\Theta = \{\pi_1, \dots, \pi_k, \Theta_1, \dots, \Theta_k\}$ are the mixture parameters (π_j and $p_j(y|\Theta_j)$ are respectively the weight and PDF of the j -th component). The output approximation is encoded in a message which comprises all its terms. The lower is the encoding of information, i.e., $I(\Theta | \mathcal{Y}_{train}^{FMM})$, the lower is the message length and, hence, the more accurate is the output distribution approximation with that mixture of components [80]. In particular, $I(k)$ represents the encoding of the number of components (k), $I(\boldsymbol{\pi})$ is the encoding of the weights, $\sum_{j=1}^k I(\Theta_j)$ is the encoding of the component parameters (Θ_j) and $I(\mathcal{Y}_{train}^{FMM} | \Theta)$ is the encoding of the data. All these terms are logarithmic terms and in the most favorable situations they could be even negative. To sum up, the MML criterion is a trade-off between the complexity of the model and the goodness of fit [81]; indeed, when a new component is added the encoding of the new component parameters increases the message length, while $I(\mathcal{Y}_{train}^{FMM} | \Theta)$ term reduces it due to the improved fit quality.

The SNOB algorithm allows the user to choose among several types of distributions (i.e., model space), e.g., Gaussian, Weibull, Exponential etc. The algorithm automatically finds the best k according to that type and provides in output the MML metric that can be used to justify the model space selection. The solution associated to the lowest MML value is the most accurate for the case study.

Once the parameters of the mixture of models are known, the output distribution is completely characterized and the clusters obtained may be representative of safe conditions, whereas others of failure conditions. For Global SA, the focus is shifted to the input space and the output clustering is exploited to cluster also the inputs. The PDFs of each input variable (x_m) with the conditioning on the different j -th clusters are constructed, i.e., $p(x_m | \Theta_{jm})$, and, then, the difference between $p(x_m | \Theta_{jm})$ and the input common distribution, i.e., $p(x_m)$ is measured according to one of the DBSA methods introduced before (e.g., Hellinger distance, Kullback-Liebler divergence etc.). These measures allow to rank the input variables contribution to the different output clusters, with special attention to the clusters of interest, e.g., those related to the failure of the system, and, finally, the most important inputs are selected.

4.1.2 Iterative Metamodel Training (AK-MCS)

After reducing the number of input parameters through dimensionality reduction, a surrogate metamodel is constructed to approximate the real model I/O relationships on the reduced input space, i.e., $Y = f(\mathbf{X}^R)$, where $\mathbf{X}^R \in D_{\mathbf{X}^R} \subset \mathbb{R}^R$ ($R < M$ is the dimensionality of the reduced space). Among the several options available in literature [63], we resort, again, to Kriging metamodels (as in [Chapter 3](#)). Indeed, Kriging metamodel has already proved suitable for the problem of CRs characterization, being capable of approximating many response functions and non-stationary, which is really useful for the specific aim of making the metamodel more refined in proximity of the CR limit surface. The idea is that, even if Kriging metamodel usually requires the approximated functions to be smooth, in this case the metamodel should not suffer for the output non-smoothness and

multimodality thanks to the more effective I/O training set that is constructed after dimensionality reduction (see [Section 4.1.1](#)).

A Kriging metamodel is initially built according to a small I/O training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$, whose simulated inputs have been generated by LHS, but now varying only the input parameters selected at the previous step ([Section 4.1.1](#)), whereas the others are set to their nominal values. Then, the metamodel refinement is carried out through the AK-MCS iterative procedure, following the same steps described in [Section 3.1.2](#), which now are just briefly summarized. For each n -th iteration:

1. **Construction:** a Kriging metamodel is built with the available I/O training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$.
2. **Generation of random input configuration:** a large number N_{MCS} of new input configurations $\mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{N_{MCS}})$, is generated by means of LHS.
3. **Metamodel Evaluation:** the Kriging metamodel is used to evaluate the outputs corresponding to the \mathcal{X} combinations: $\hat{\mathcal{Y}} = (\hat{y}_1, \dots, \hat{y}_{N_{MCS}})$.
4. **Convergence check:** convergence of the metamodel construction is verified through an a priori defined convergence or stopping criterion.
5. **Selection:** if convergence at the previous step is not satisfied, new input combinations are selected according to the U learning function and, then, simulated with the time-demanding model. After that, they are added to the I/O training set for the metamodel refinement and the procedure goes back to step 1, until step 4 is verified.

4.1.3 CRs Representation & Information Retrieval

The Kriging metamodel obtained at the end of the AK-MCS iterations is expected to provide accurate predictions of new input combinations and, hence, it can be now exploited for the final objective of CRs exploration, instead of the far more computationally demanding original model. In particular, the level of accuracy reached by the metamodel can be quantified by exploiting an external validation set. Then, the Kriging metamodel is

used to predict the outcomes of many new input combinations to find the critical ones and characterize the CRs. Finally, the system CRs can be graphically represented by exploiting the high-dimensional data visualization techniques, like scatter plots and PCP, already introduced in [Section 3.1.3](#).

4.2 Novel Framework Application for the CRs Characterization of the PSS considered – Pressure Output

The novel framework for the CRs characterization in presence of state-space non-smoothness and multimodality, illustrated in [Section 4.1](#), has been here implemented to identify and explore the CRs of the PSS introduced in [Chapter 4](#), with respect to the FC “Steam release in the containment”. Indeed, this FC is defined according to the pressure output (p_{max}) which has shown a strong non-smooth and multimodal distribution (see [Fig. 8](#)). In the following subsections the relevant steps of this application are illustrated in detail, without loss of generality.

4.2.1 Dimensionality Reduction

With the aim of finding an effective I/O training set to construct an accurate metamodel for the approximation of the PSS response with respect to p_{max} output, the input vector dimensionality has been reduced from M to R ($R < M = 5$); hence, a reduced model dealing with the reduced input vector $\mathbf{X}^R \in D_{\mathbf{X}^R} \subset \mathbb{R}^R$ can be obtained. A DBSA method supported by FMM technique has been implemented to tackle p_{max} non-smoothness and multimodality by catching the output different clusters and, finally, selecting the most relevant inputs contributing to the output distribution. In particular, the analysis has been restricted only to those input parameters significantly affecting the output cluster connected with the critical conditions, i.e., those with p_{max} around 75.5 bar.

200 RELAP5-3D simulated transients have been used for the FMM implementation with the SNOB algorithm, introduced in 4.1.1.2. The SNOB algorithm is based on the EM algorithm to select the best number of components (k), and this selection is guided by the MML criterion (see equation (20)).

The final goal of this specific FMM application was not to approximate p_{max} distribution in the best possible way, whatever the number of components, but to obtain a good fit while ensuring that the k components still reproduced the underlying physics of the problem. The SNOB algorithm has reached the optimal fitting of the p_{max} multimodal distribution with $k = 3$ Gaussian distributions (whose characteristic parameters, i.e., mean value μ and standard deviation σ , are reported in Table 8). Within the aim of ranking the most relevant inputs by means of one of the DBSA methods introduced in Section 4.1.1, firstly, it was necessary to assign each output variable in the set of 200 RELAP5-3D simulations to the cluster that was assumed to have generated it. In particular, it has been considered that the sample y_i belongs to the j -th cluster if it returns the highest probability value when substituted into the PDF expression of that cluster.

Table 8: FMM components parameters

Cluster name		μ [bar]	σ [bar]
Low-pressure	(green)	70.0	1E-3
Medium-pressure	(orange)	72.6	2.48
High-pressure	(red)	75.9	0.04

The three Gaussian distributions together with the related output clusters are reported in Fig. 15. A “low-pressure” cluster on the left associated with the system safe conditions (in green in Fig. 15) and is approximated by a Dirac’s delta distribution: it represents the p_{max} concentration around 70.0 bar, corresponding to all the transients in which the decay heat is correctly removed by the PSS and the pressure never increases (122 simulations out of 200). Thus, in these simulations, p_{max} is always equal to the pressure value at the beginning of the transient, i.e., 70.0 bar. The remaining 78 outputs are almost equally split among the “medium-pressure” cluster in the middle (safe conditions, but with lower safety margin) and the “high-pressure cluster” on the right (critical conditions): they are

Chapter 4 / Novel framework for System CRs characterization in case of state-space non-smoothness and multimodality

associated to two Gaussian distributions (respectively orange and red in Fig. 15), with the second that is more peaked. Both clusters include those transients in which the pressure initially increases beyond 70.0 bar, due to the AV delayed opening with respect to the MSIV closure, which causes the PV to remain without vapor outlets. The only difference is that p_{max} values in the “high-pressure” cluster reach $Y_{thres} = 75.5 \text{ bar}$, causing the SRV opening, while in most of the transients assigned to the “medium-pressure” cluster the pressure increases, but without reaching Y_{thres} .

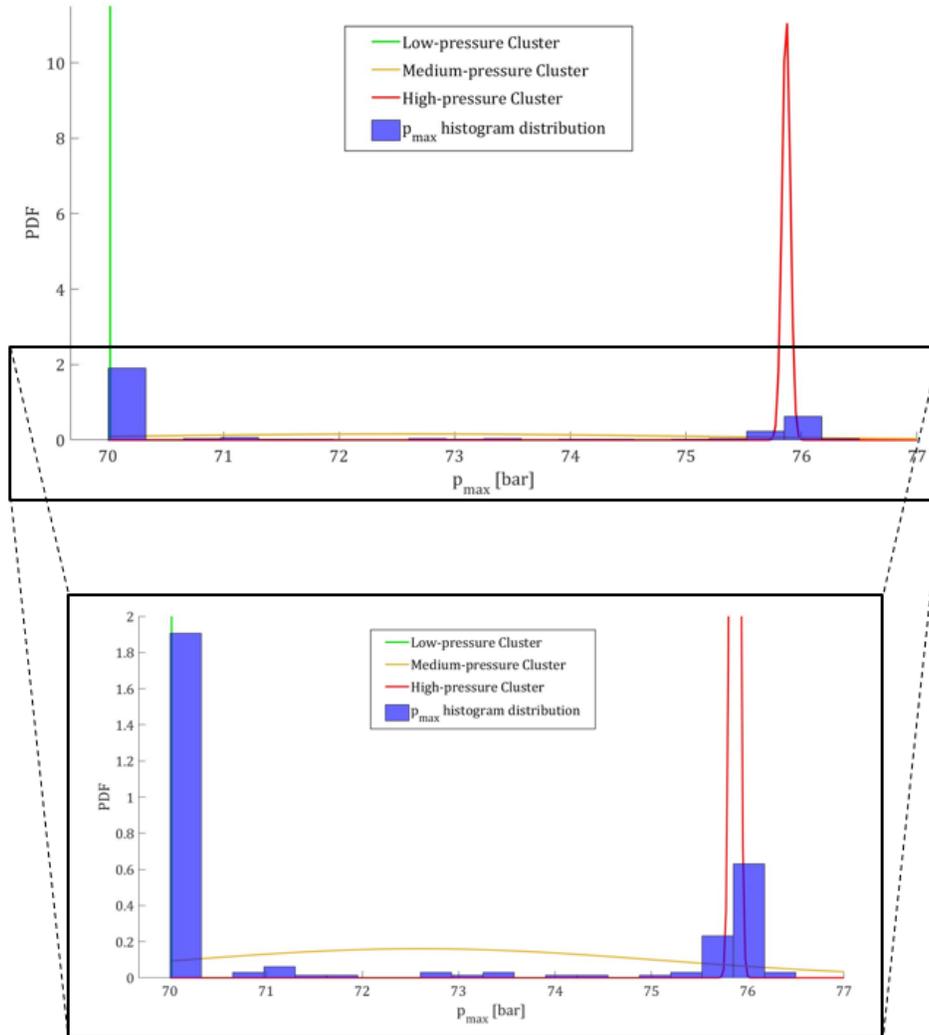


Figure 15: p_{max} clustering according to $k = 3$ Gaussian distributions

The output clustering has been exploited to rank the input variables most affecting the output clusters (DBSA) by constructing the PDF of each input x_m conditioned on each j -th cluster, i.e., $p(x_m | \Theta_{jm})$. In particular, these PDFs have been constructed by simply assigning the input variables belonging to the set of 200 RELAP5-3D simulations to the same cluster of the associated outputs and, then, the PDF $p(x_m | \Theta_{jm})$ has been created using only the x_m inputs assigned to the j -th cluster. In this way, it is possible to measure the difference between $p(x_m | \Theta_{jm})$ and the original input distribution of x_m , i.e., $p(x_m)$, and to use this difference to rank x_m . For the case study, the Hellinger distance method for SA [46], [54] has been adopted:

$$H_{jm} = \left[\frac{1}{2} \int \left(\sqrt{p(x_m)} - \sqrt{p(x_m | \Theta_{jm})} \right)^2 dx_m \right]^{1/2}, \quad (21)$$

with H_{jm} that needs to satisfy the inequality $0 \leq H_{mj} \leq 1$. The quantity H_{jm} represents the importance of the m -th input in affecting the j -th cluster of the output distribution. The higher the H_{jm} value with respect to the one of the other input parameters, the greater the relative importance of x_m .

For the analysis of p_{max} , special attention has been paid to the “high-pressure” cluster, since it is the one connected with the failure of the PSS function (critical conditions). Hence, for each input parameter, the corresponding H_{mj} value referred to this cluster (i.e. with $j = 3$) has been exploited as a sensitivity index. Fig. 16 reports a comparison between the H_{3m} values calculated for each of the five input parameters.

As it can be deduced from Fig. 16, the two valves delays, i.e., DEL_{AV} and DEL_{MISV} , mostly affect p_{max} “high-pressure” cluster and hence they are more likely to generate scenarios in which the pressure increases towards Y_{thres} with the consequent SRV opening. Therefore, the problem dimensionality has been reduced from $M = 5$ to $R = 2$ and a reduced model dealing with a reduced input vector has been obtained, i.e., $f(\mathbf{X}^R) = Y$, with $\mathbf{X}^R \in D_{\mathbf{X}^R} \subset \mathbb{R}^R$ and Y still equal to p_{max} .

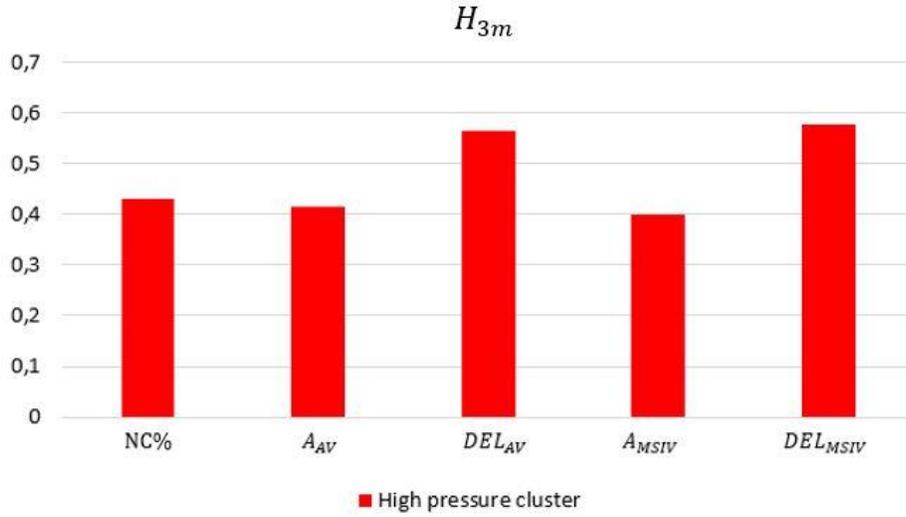


Figure 16: Hellinger distance for each input parameter (x_m) evaluated with respect to high-pressure cluster

4.2.2 Iterative Metamodel Training (AK-MCS)

After the dimensionality reduction previously presented, the input parameters used to model the generic PSS behaviour with respect to p_{max} are only DEL_{AV} and DEL_{MSIV} ; Thus, an accurate Kriging metamodel has been built to mimic the RELAP5-3D model I/O relationships on the reduced and more effective space of dimensionality $R = 2$.

For the purpose of CRs exploration, knowing that p_{max} can approach $Y_{thres} = 75.5 \text{ bar}$ only if $DEL_{AV} > DEL_{MSIV}$, with a quite significant interval of time between the two valves actions, has led to adjust the range of variation of DEL_{MSIV} from $DEL_{MSIV} = 0 \div 7200 \text{ s}$ to $DEL_{MSIV} = 0 \div 480 \text{ s}$: this has allowed to be coherent with $DEL_{AV} = 0 \div 720 \text{ s}$ (see Table 2) and to avoid sampling far from the limit surface.

Following the criterion proposed in [69], which suggests a number of training simulations $N_{train} \geq 10R$, a Kriging metamodel has been initially constructed with an I/O training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ of 25 RELAP5-3D runs (obtained in correspondence of the input values generated by LHS). In particular, the construction has been performed by means of the UQLab Software Framework for Uncertainty Quantification [70], already introduced in Section 3.2.1. The best Kriging setting for the specific study of p_{max} output

has been established, again, by testing different Kriging features with the CV procedure. As in the case of the PSS energy output (E_{ex}), only two features, i.e., the trend type and correlation function family, have been tested (See [Tables 9, 10, 11](#) and [12](#)), whereas the other features have been set to their default options defined in UQLab. At first, the best trend type with all the other Kriging features set to their default options (the correlation function family set by default is *Matérn 5/2*) has revealed to be *Linear* type ([Table 9](#)). Then, *Linear* trend has been used for the calculation of the best correlation function family which has resulted to be *Exponential* family ([Table 10](#)). For this reason, the calculation has come back looking for the best trend type, but now using the *Exponential* correlation function family instead of *Matérn 5/2* one. The new best trend type has resulted to be *Ordinary* (see [Table 11](#)), for which the best correlation function family is still *Exponential* ([Table 12](#)) so the calculation has been stopped. In the end, the Kriging best setting for p_{max} approximation is:

- Trend type: *Ordinary*
- Family of correlation functions: *Exponential*
- Type of correlation functions: *Ellipsoidal* (default)
- Estimation method: *CV* (default)
- Optimization method: *Genetic Algorithm (GA)* (default)

Table 9: Average CV error in trend type estimation (with Matérn 5/2 corr. function family)

Trend type	Ordinary	Linear	Quadratic
Average CV error [bar]	1.57	1.53	2.23

Table 10: Average CV error in trend type estimation (with Exponential corr. function family)

Corr. Function family	Exponential	Gaussian	Matérn 3/2	Matérn 5/2
Average CV error [bar]	1.50	2.27	1.56	1.56

Chapter 4 / Novel framework for System CRs characterization in case of state-space non-smoothness and multimodality

Table 11: Average CV error in trend type estimation (with Exponential corr. function family)

Trend type	Ordinary	Linear	Quadratic
Average CV error [bar]	1.17	1.48	1.98

Table 12: Average CV error in correlation function family estimation (with Ordinary trend)

Corr. Function family	Exponential	Gaussian	Matérn 3/2	Matérn 5/2
Average CV error [bar]	1.19	3.61	1.35	1.49

Then, the Kriging metamodel has been adaptively refined, with a focus on the CR limit surface, by enriching $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ within the AK-MCS framework introduced in Section 3.1.2, but here properly tailored to the specific case study of the PSS presented in Section 2.1, in relation to p_{max} output. Here are reported the details of the steps concerning the AK-MCS application, for each n -th iteration:

1. **Construction:** a Kriging metamodel is constructed with the available I/O training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ which increases its size with the iterations to better refine the metamodel. The metamodel accuracy is improved specifically nearby $Y_{thres} = 75.5 \text{ bar}$.
2. **Generation of random input configurations:** $N_{MCS} = 10.000$ new inputs combinations $\mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{N_{MCS}})$ (of reduced dimensionality $R = 2$) are sampled with LHS (see the input ranges defined at the beginning of this Section).
3. **Metamodel evaluation:** the sampled input combinations \mathcal{X} are run through the Kriging metamodel to predict the corresponding output values (i.e., maximum vessel pressure): $\hat{\mathcal{Y}} = (\hat{y}_1, \dots, \hat{y}_{N_{MCS}})$.
4. **Convergence check:** a convergence or stopping criterion regarding the computational cost has been defined. The maximum number of simulations foreseen for the metamodel training has been set to 100, due to the significant computational cost associated to the dimensionality reduction procedure carried out before (200 RELAP5-3D simulations required). Thus, considering that $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ is constituted by 25 simulations, only 75 simulations can be iteratively added during

the AK-MCS procedure. When the size of $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ reaches its maximum value, the procedure stops.

5. **Selection:** if the convergence criterion at step 4 is not verified, new I/O simulations related to the so-called best candidate subset, i.e., $\mathcal{X}^* \subset \mathcal{X}$, are conducted and the corresponding inputs and outputs $\{\mathcal{X}^*, \mathcal{Y}^*\}$ are added to $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ to refine the metamodel. The N_{cand} best candidates \mathcal{X}^* are selected among the \mathcal{X} combinations according to their U -function values (see [equation \(7\)](#)), in order to choose them close to Y_{thres} . The “bins selection technique”, introduced in [Section 3.2.2](#) for the analysis of E_{ex} output, is here followed to avoid selecting \mathcal{X}^* from the same restricted portion of the input domain (i.e., candidates too similar to each other) due to the correlation function. $N_{cand} = 7$ or 8 candidates are selected at each n -th iteration, according to the same rationale presented in [Section 3.2.2](#) (see [Fig. 17](#)). Once \mathcal{X}^* combinations have been selected and the corresponding transients simulated with the RELAP5-3D code to obtain the output \mathcal{Y}^* , $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ is enriched and steps 1 to 5 are repeated until convergence at step 4 is verified.

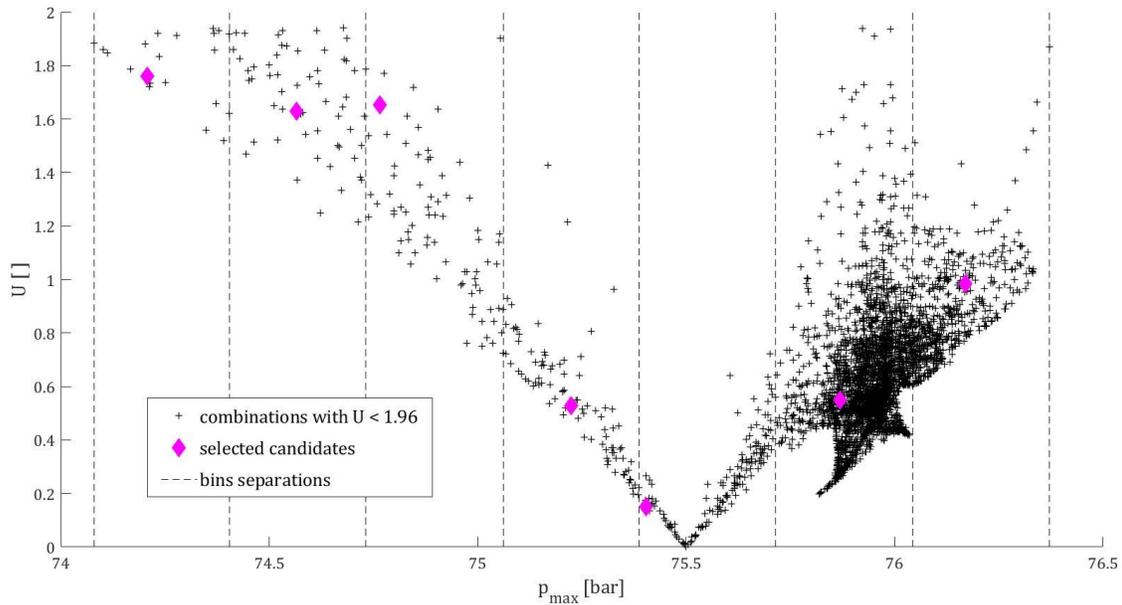


Figure 17: Bins selection technique (see [Section 3.2.2](#)) applied also for the AK-MCS procedure referred to p_{max} output

4.2.3 CRs Representation & Information Retrieval

The present Section is organized as follows: at first, the accuracy of the metamodel obtained at the end of the AK-MCS procedure (previous section) has been quantified thanks to the introduction of a validation set and some QIs (4.2.3.1). Second, the Kriging metamodel has been exploited to predict the pressure output of a large number of new input combinations on the reduced space and to find the critical ones. In this way, the CR can be identified and, finally, graphically represented (4.2.3.2).

4.2.3.1 Metamodel Accuracy Evaluation

The AK-MCS procedure presented in Section 4.2.2 has been stopped at iteration $n_{fin} = 10$, when the maximum number (100) of RELAP5-3D simulations allowed for the construction of the training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ has been reached. The evolution of the metamodel accuracy with the iterations has been followed through the introduction of a validation set $\{\mathcal{X}_{val}, \mathcal{Y}_{val}\}$, made by N_{val} I/O relations. As explained in Chapter 3 in case of the introduction of a validation set for the quantification of the accuracy of the metamodel used to approximate E_{ex} , no definitive guidelines are available in literature about the most suitable size N_{val} . In case of p_{max} analysis, also considering the available computational budget, a validation set of 50 RELAP5-3D simulations with the outputs mainly distributed around Y_{thres} has been constructed to measure the accuracy increase, especially in proximity of the limit surface. The metamodel has been used to predict the outcomes $\hat{\mathcal{Y}}_{val} = (\hat{y}_1, \dots, \hat{y}_{N_{val}})$ corresponding to the input combinations \mathcal{X}_{val} ; then, the accuracy has been quantified through QIs comparing $\hat{\mathcal{Y}}_{val}$ to the real outputs evaluated with the RELAP5-3D model. This comparison has been carried out with respect to the same three QIs introduced in Chapter 3, i.e., RMSE, Q_1 and Q_2 (see equations (8), (9) and (10)).

The progressive increase of accuracy is shown by the trends of the three QIs illustrated in Fig. 18. All the QIs considered show a significant improvement at the beginning, then, in the successive iterations, the relative improvement becomes negligible. Also for this reason, stopping the AK-MCS procedure at iteration $n_{fin} = 10$ represents a reasonable choice.

Chapter 4 / Novel framework for System CRs characterization in case of state-space non-smoothness and multimodality

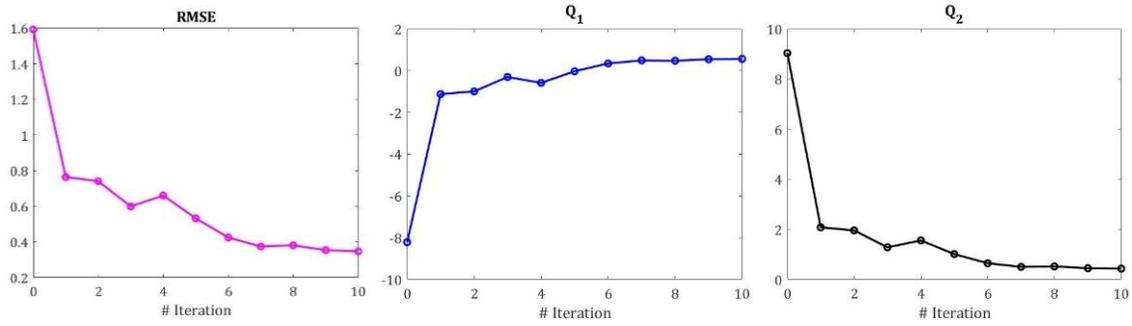


Figure 18: QIs evaluated with respect to a given validation set (pressure output)

The three QI values at the end of the AK-MCS procedure for the approximation of p_{max} output are reported in Table 13. The RMSE at the last iteration is satisfactory, indeed $RMSE = 0.35 \text{ bar}$ is really low if compared to the p_{max} values in the simulated transients ($p_{max} = 70.0 \div 76.5 \text{ bar}$). Moreover, a final $NRMSE = 0.46\%$ is remarkable, since it can be taken, in the first instance, as a measure of the percentage error of the Kriging predictions. For what concerns Q_1 and Q_2 , they show a significant improvement during the successive iterations, but their final values are not so satisfactory, especially for the final Q_1 which lies far from 1. This is probably due to the very low variability of the validation set chosen for the analysis: indeed, most of the p_{max} values of \mathcal{Y}_{val} are spread on a range of only 2 bar around $Y_{thres} = 75.5 \text{ bar}$.

Table 13: QIs values at the end of AK-MCS procedure – 10th iteration (pressure output)

Quality indicator	RMSE [bar]	NRMSE [%]	Q ₁	Q ₂
Final value	0.35	0.46%	0.56	0.43

4.2.3.2 CR Characterization

After demonstrating that the Kriging metamodel obtained at the end of the AK-MCS procedure presents a high accuracy, especially in proximity of Y_{thres} , it can be used for CR exploration instead of the more time-demanding RELAP5-3D model. For this purpose, 10.000 new input combinations have been generated by LHS and, then, predicted with the metamodel to: (i) find the critical ones, i.e., $\hat{y} = f(\mathbf{x}) \geq Y_{thres}$; and (ii) retrieve useful

information about the CRs, like their number and shape (same procedure followed for the characterization of the CRs related with the energy output).

One single CR has been identified (see Fig. 19); moreover, given that the analysis has been restricted only to two parameters after dimensionality reduction, no high-dimensional data visualizations techniques, like SPLOM or PCP (see Section 3.1.3), were needed. The CR has been represented in the two-dimensional input space through a scatter plot, in which green diamonds indicate combinations leading to safe operation (p_{max} is kept $< 75.5 \text{ bar}$), whereas red crosses represent the critical input combinations of PSS functional failure.

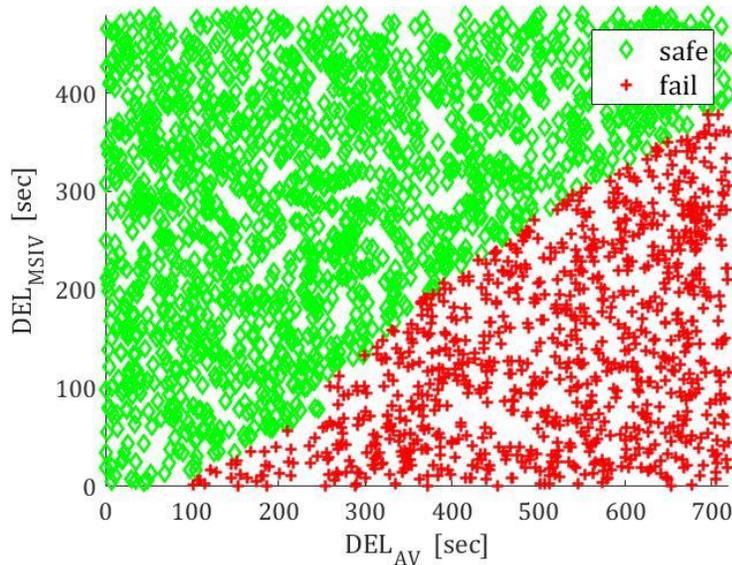


Figure 19: CR for p_{max} output

A triangle-shaped CR has been identified, showing the direct influence of both DEL_{AV} and DEL_{MSIV} on the FC “Steam release in the containment”; indeed, it is evident how p_{max} may exceed 75.5 bar *only when* the MSIV closes before the opening of the AV, i.e., when $DEL_{MSIV} < DEL_{AV}$ (as introduced in Section 2.3). This occurs because the PV remains without vapor discharge outlets and, hence, the vapor builds up causing the PV overpressurization. Also, Fig. 19 shows that not always $DEL_{MSIV} < DEL_{AV}$ leads the PSS to fail its function: e.g., even if the MSIV is supposed in its reference conditions (i.e., $DEL_{MSIV} = 0 \text{ sec}$), if $DEL_{AV} < 50 \text{ sec}$, p_{max} remains below Y_{tresh} . In general, the higher DEL_{MSIV} , the

lower the chances to lead to functional failure: eventually, if $DEL_{MSIV} > 380 \text{ sec}$ failure is never reached, whatever the value assumed by DEL_{AV} .

To sum up, the novel framework based on FMM and AK-MCS has been demonstrated capable of approximating p_{max} output, despite its strong non-smooth and multimodal nature, within the aim of exploring the PSS CRs relative to the FC “Steam release in the containment”. 200 RELAP5-3D simulations have been exploited for the FMM approximation, according to which the output multimodal distribution has been divided into three clusters exploited for dimensionality reduction through the Hellinger distance SA method. Then, a Kriging metamodel has been adaptively trained on the reduced (two-dimensional) input space, within an AK-MCS framework, by exploiting 100 additional RELAP5-3D simulations. A unique CR has been identified and characterized in terms of *combined action* of the two input parameters previously selected, i.e., DEL_{AV} and DEL_{MSIV} . The PSS has been demonstrated more prone to failure in case of large DEL_{AV} , especially if $DEL_{AV} > DEL_{MSIV}$. To conclude, as for the analysis of the energy output, these results could be used within a more general reliability assessment of the PSS considered; moreover, the novel framework proposed could be exploited for the CRs exploration of other systems for which a time-demanding model is available and characterized by a non-smooth and multimodal state-space.

4.2.4 Comparison with the Results obtained with SVC + AK-MCS procedure

An alternative approach to tackle the output non-smoothness and multimodality is represented by the use of a classifier, as explained in 1.3.2.1. After separating the regions with different output behaviours, a classifier can be trained and used to predict the partition a new input combination belongs to. In this way, only the outcomes of the input combinations belonging to the region of interest (i.e., the region connected with the system failure) are, then, predicted with a metamodel properly fitted to that specific region. In particular, in case of p_{max} approximation, we have applied the “two-stage surrogate modelling” technique [18], already introduced in 1.3.2.1. A hard SVC (i.e., where one input

combination cannot contemporary belong to different domain partitions) has been initially trained according to the domain partitions previously identified. After that, a new input combination \mathbf{x} whose output needed to be predicted (and identified if critical or not) has been, first, classified with the hard SVC (1st stage); then, the corresponding output predicted with the metamodel specifically built for the partition (cluster) of \mathbf{x} (2nd stage) (see Fig. 3). Differently from the framework proposed in Section 4.1, the metamodel has been fitted to the original input space of dimensionality M : indeed, the idea was that the metamodel would have not suffered for the state-space non-smoothness and multimodality anymore, due to the separate analysis of the different output behavioural regions.

At first, two output domain partitions have been identified according to expert judgement (see Fig. 8): a “low-pressure” region corresponding to $p_{max} = 70.0 \text{ bar}$ (which occurs in most of the transients simulated, see Section 2.3.2), and a “high-pressure” region with $p_{max} > 70.0 \text{ bar}$, representing those transients in which the pressure rises. Thus, a binary classification results: i.e., given a certain input combination \mathbf{x}_i , the corresponding label assigned by the classifier is $\ell_i = \{-1, +1\}$, with the $\ell_i = -1$ that is associated to the “low-pressure” region and $\ell_i = 1$ that represents the “high-pressure” region.

The SVC has been constructed according to the two output domain regions identified, thanks to N_{train} I/O training simulations $\{\mathcal{X}_{train}^{SVC}, \mathcal{Y}_{train}^{SVC}\}$, where $\mathcal{X}_{train}^{SVC} = \{\mathbf{x}_i \in \mathbb{R}^M, i = 1, \dots, N_{train}\}$ are the training input combinations and $\mathcal{Y}_{train}^{SVC} = \{y_i^{SVC} = \ell_i = \{-1, +1\}, i = 1, \dots, N_{train}\}$ are the corresponding labels. Here is provided a description of SVC construction process in case of binary classification, i.e., when only two classes have been identified [82]. The SVC classification is carried out according to the separating hyperplane that maximizes its distance (also known as margin) from the closest training combinations. The separating hyperplane can be defined as:

$$\{\mathbf{x} \in \mathbb{R}^M : \mathbf{w}^T \mathbf{x} + b\}, \quad (22)$$

where \mathbf{w} is the vector of hyperplane coefficients and b is the bias. The perpendicular distance of any input combination from this hyperplane is:

$$d(\mathbf{x}_i) = \frac{|\mathbf{w}^T \mathbf{x}_i + b|}{\|\mathbf{w}\|}. \quad (23)$$

Chapter 4 / Novel framework for System CRs characterization in case of state-space non-smoothness and multimodality

It turns out that maximizing the margin corresponds to the minimization of the norm of \mathbf{w} under some constraints. Therefore, determining the separating hyperplane reduces to the following optimization problem:

$$\min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|^2, \quad \text{subject to} \quad y_i^{SVC}(\mathbf{w}^T \mathbf{x}_i + b) - 1 \geq 0, \quad i = \{1, \dots, N_{train}\}, \quad (24)$$

where the constraints ensure that no samples can lie inside the area covered by the margin. The optimization problem is convex and it can be solved by introducing the Lagrange multipliers. After some algebra, the final optimization problem becomes:

$$\min_{\alpha} -\frac{1}{2} \sum_{i=1}^{N_{train}} \sum_{j=1}^{N_{train}} \alpha_i \alpha_j y_i^{SVC} y_j^{SVC} \mathbf{x}_i^T \mathbf{x}_j + \sum_{i=1}^{N_{train}} \alpha_i, \quad (25)$$

$$\text{subject to} \quad \sum_{i=1}^{N_{train}} \alpha_i y_i^{SVC} = 0, \quad \alpha_i \geq 0, \quad i = \{1, \dots, N_{train}\}.$$

After finding the Lagrange multipliers $\{\alpha_i, i = 1, \dots, N_{train}\}$ and the bias b , the SVC classification of a new configuration can be expressed in terms of training input combinations:

$$\hat{y}^{SVC}(\mathbf{x}_i) = \ell(\mathbf{x}_i) = \text{sign} \left(\sum_{i=1}^{N_{train}} \alpha_i y_i^{SVC} \mathbf{x}_i^T \mathbf{x} + b \right). \quad (26)$$

In some situations, the optimization problem becomes unfeasible. A new solution is provided by allowing misclassifications, i.e. by relaxing the inequality constraints through the introduction of the so-called slack terms ξ_i , which measures the distance of the misclassified sample from its actual class. A penalized objective function is obtained in which the slack terms are minimized. Two final expressions are obtained according to the type of penalization:

➤ Linear penalization

$$\min_{\alpha} -\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{N_{train}} \xi_i \quad \text{subject to} \quad y_i^{SVC}(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i \quad (27)$$

➤ Quadratic penalization

Chapter 4 / Novel framework for System CRs characterization in case of state-space non-smoothness and multimodality

$$\min_{\alpha} -\frac{1}{2} \|\mathbf{w}\|^2 + \frac{C}{2} \sum_{i=1}^{N_{train}} \xi_i^2 \quad \text{subject to} \quad y_i^{SVC} (\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i. \quad (28)$$

In the case where the data are not linearly separable, the training combinations are mapped into a higher dimensional space referred to as feature space and, therefore, the construction of the optimal separating hyperplane is shifted to this new space. A new classification formula is given by the sign of the following expression:

$$\mathbf{w}^T \Phi(\mathbf{x}) + b = \sum_{i=1}^{N_{train}} \alpha_i y_i^{SVC} \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}) + b, \quad (29)$$

where $\Phi(\bullet)$ is the mapping function and hence the components of \mathbf{x} in the feature space are $(\Phi_1(\mathbf{x}), \dots, \Phi_M(\mathbf{x}))$. The expression in [equation \(29\)](#) shows how, if one is able to calculate the inner product of the two vector images in the feature space, i.e., $\Phi(\mathbf{x}_i)^T \Phi(\mathbf{x})$, no further cumbersome operations need to be carried out in that space. This operation is named “kernel trick” since it is conducted thanks to kernel functions. Several examples of kernel functions are available in literature (e.g., Polynomial, Gaussian, Exponential etc.). Once the kernel function *ker* has been chosen, the final classification reads:

$$\hat{y}^{SVC}(\mathbf{x}_i) = \ell(\mathbf{x}_i) = \text{sign} \left(\sum_{i=1}^{N_{train}} \alpha_i y_i^{SVC} \text{ker}(\mathbf{x}_i, \mathbf{x}) + b \right). \quad (30)$$

In the case study, the training set $\{\mathcal{X}_{train}^{SVC}, \mathcal{Y}_{train}^{SVC}\}$ is constituted by the same 200 RELAP5-3D simulations exploited for the FMM-based approach (see [Section 4.2.1](#)). The only difference is that $\mathcal{Y}_{train}^{SVC}$ is a set of labels associated to the training input combinations $\mathcal{X}_{train}^{SVC}$, instead of the corresponding pressure output values used for the FMM approximation because the aim here is to classify the combinations (122 combinations are labelled $\ell_i = -1$ and 78 are labelled $\ell_i = +1$). This choice is justified by the criterion introduced in [40], which proposes convergence points instead of a far more expensive validation set to quantify the accuracy of the SVC. This criterion has proven that at least 180 simulations were necessary to construct an initial, but sufficiently accurate SVC for the case study.

Chapter 4 / Novel framework for System CRs characterization in case of state-space non-smoothness and multimodality

After the SVC training, a Kriging metamodel has been built to predict p_{max} in the “high-pressure” region, because it was not worth exploring also the “low-pressure” region where p_{max} is constant (70.0 bar). The metamodel training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ has been constructed the same I/O relations collected for the SVC training, but taking only those classified as belonging to the high-pressure region (78 simulations out of 200). In this case, no dimensionality reduction has been carried out and, hence, the metamodel has been used to mimic the RELAP5-3D model on the original input space of dimensionality $M = 5$, i.e., $f(\mathbf{X}) = Y$, with $\mathbf{X} \in D_{\mathbf{X}^M} \subset \mathbb{R}^M$. Thus, \mathcal{X}_{train} is a set of five-dimensional input combinations.

The Kriging metamodel has been adaptively refined in proximity of $Y_{thres} = 75.5 \text{ bar}$ with a sort of AK-MCS procedure (see [Section 3.1.2](#)), conveniently adjusted to be coupled with the SVC. At each n -th iteration, $N_{MCS} = 100.000$ new input combinations $\mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_{N_{MCS}})$ have been generated by LHS and classified by the SVC according to the two regions identified (1st stage). Only the combinations classified as belonging to the high-pressure region, i.e., $\mathcal{X}^{Krig} \subset \mathcal{X}$ have been, then, evaluated with the Kriging metamodel to find the corresponding outputs (2nd stage). The most interesting input combinations among \mathcal{X}^{Krig} , in terms of learning function U value (7/8 candidates at each iteration), have been selected to be simulated with the RELAP5-3D model and added to $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ for the metamodel refinement (the same “bins technique” introduced in [Section 3.2.2](#) can be exploited for this selection). This procedure could be repeated until the level of accuracy of Kriging predictions became satisfactory. The I/O relations simulated at each iteration to enrich the metamodel training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}$ could be labelled and exploited to enrich also the classifier training set $\{\mathcal{X}_{train}^{SVC}, \mathcal{Y}_{train}^{SVC}\}$. This procedure has been called SVC+AK-MCS.

The idea was to exploit the same number of RELAP5-3D simulations, i.e., the same computational budget, as the one used for the novel framework (FMM+AK-MCS) implemented in [Sections 4.2.1](#), [4.2.2](#) and [4.2.3](#), to refine both the Kriging metamodel and the SVC within the SVC+AK-MCS framework, with the aim of fairly comparing the final Kriging accuracy reached. The initial metamodel training set $\{\mathcal{X}_{train}, \mathcal{Y}_{train}\}_{in}$ has been adaptively enriched together with $\{\mathcal{X}_{train}^{SVC}, \mathcal{Y}_{train}^{SVC}\}$, up to the limit of the 300 simulations (the same limit as FMM+AK-MCS). Thus, starting from $\{\mathcal{X}_{train}^{SVC}, \mathcal{Y}_{train}^{SVC}\}$ made by 200 I/O samples (the same used for the FMM application), 100 simulations have been added (only

82 of them could be used for the Kriging training) and the Kriging training set size has been simultaneously increased from 78 to 160. The Kriging accuracy has been quantified with respect to a validation set of the same size of the one used in Section 4.2.3 (i.e., 50 I/O relations). Again, the validation set is constituted by samples mainly distributed around Y_{thres} , to verify the metamodel accuracy improvement with specific attention to the area close to the limit surface. Table 14 reports the values of three QIs (RMSE, Q_1 and Q_2) computed on this validation set according to the Kriging metamodel obtained at the end of SVC+AK-MCS procedure.

Table 14: QIs at the end of SVC+AK-MCS iterative procedure (pressure output)

Quality indicator	RMSE [bar]	NRMSE [%]	Q_1	Q_2
Final value	0.85	1.12%	0.16	0.82

All the QIs values are worse than the ones calculated at the end of the FMM+AK-MCS framework application (see Table 13). For example, the RMSE and NRMSE are more than twice larger, and Q_1 is even 3.5 times lower, meaning that the accuracy of the Kriging metamodel at the end of the SVC+AK-MCS procedure is lower. Indeed, Fig. 20 shows how the novel framework previously proposed, based on FMM+AK-MCS, outperforms the SVC+AK-MCS procedure in terms of Q_1 and Q_2 , after six iterations (i.e., with 270 simulations rather than 300) and, for what concerns the RMSE, just after one iteration (i.e., with 233 simulations rather than 300).

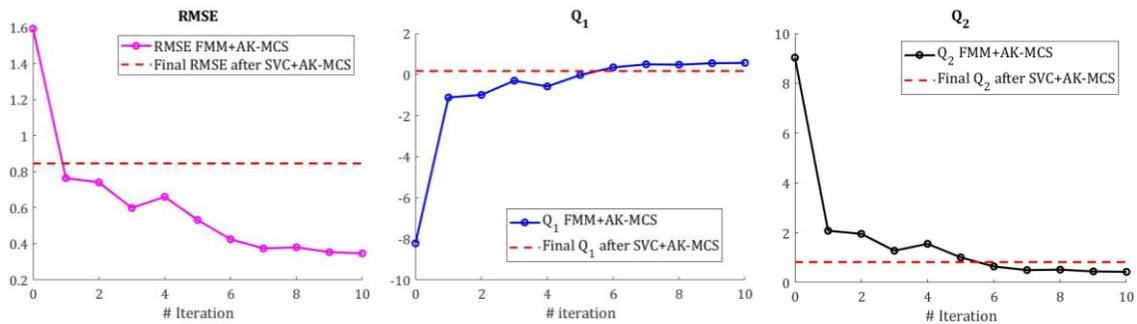


Figure 20: QIs evolution in FMM+AK-MCS strategy compared with the QIs values at the end of SVC+AK-MCS (pressure output)

5 Conclusion

The use of PSSs within the nuclear field was first addressed in 1991 in the IAEA Conference on “The Safety of Nuclear Power: Strategy for the Future” [1]. From that moment on, the adoption of PSSs has represented a promising way to increase the safety of advanced NPPs: indeed, they are generally considered as more reliable than active systems, since they present a simplified architecture and they do not rely on external energy inputs during the operation. However, i) the scarce operational experience of PSSs with respect to active systems and ii) the need to treat properly the uncertainties in their functioning and modeling have required a more thorough analysis of the possibly wide range of PSSs operating conditions and a detailed exploration of the critical combinations that may lead to PSSs functional failures.

A structured procedure for the reliability assessment of PSSs (namely, REPAS) has been developed in the past [7], [8]. However, it does not explicitly include an important step, i.e., the detailed characterization of the PSS CRs, which is necessary to identify those system configurations leading the PSS to fail its function. The identification and exploration of the CRs is a challenging problem, which can be addressed by computational methods [9], [10], [11], [12]. All these methods require the repeated simulation of the PSSs behavior by means of (typically time-demanding) BE-TH models, which makes the computational cost often prohibitive. For this reason, *new advanced* computational methods are being sought and combined to reduce the cost of computation. For example, fast-running metamodels are employed to mimic the behaviour of the time-consuming original codes and replace it in

the analysis; also, adaptive sampling strategies are used to intelligently trace the CR limit surface with the minimum waste of computational time for samples far from the CR.

In this framework, the main objective of this work has been to study and develop advanced computational methods for the efficient identification and exploration of the CRs of a nuclear PSS. Firstly, the metamodel-based AK-MCS, combining a fast-running Kriging metamodel and adaptive sampling [25], [11], has been adopted. The AK-MCS procedure automatically refines the Kriging in proximity of the PSS limit surface to predict – at an acceptable computational cost – those physical input combinations leading the PSS to functional failure (and hence to trace the boundary between safe and failed behavior). However, in case of the PSS state-space non-smoothness and/or multimodality, metamodel-based approaches (like Kriging-based) may not suffice. Indeed, they typically require accommodating properties of regularity of the underlying function to approximate, otherwise they might lead to large errors [18]. For this reason, we have also developed a novel adaptive exploration framework, based on FMM and AK-MCS, capable of tackling the state-space non-smoothness and multimodality, while searching for the system CRs.

The novel framework consists of three steps: 1) “dimensionality reduction”, relying on a DBSA method (e.g., Hellinger distance in the present work) and supported by the FMM technique for output clustering, to approximate the non-smooth and multimodal output distribution [46], [57] and to restrict the analysis only to the input parameters affecting the output clusters connected with the PSS failure; 2) “Iterative metamodel training”, based on the AK-MCS technique for the construction of an accurate Kriging metamodel to replace the typically long-running system model codes and predict the system response on a space of reduced dimensionality. The metamodel is trained with a possibly small number of time-demanding code runs; 3) “CR representation and information retrieval”, using the Kriging metamodel obtained at the previous step to predict a large number of new input combinations and retrieve useful information about the system CRs. The CRs can be, then, visualized by exploiting high-dimensional data visualization techniques (e.g. scatter plots in the present work).

These innovative methodologies have been applied to the exploration of the CRs of a generic PSS of an NPP, designed for DHR in case of reactor shut down due to an SBO

accident (see Chapter 2). The PSS has been modelled through a time-demanding BE-TH code (RELAP5-3D model, see Section 2.2) and the success of its operation has been analyzed with respect to two outputs of interest, i.e., the amount of energy exchanged during the accidental transient (E_{ex}) and the maximum value of pressure reached inside the reactor PV (p_{max}). Two FC have been identified: i) “Low heat removal”, when $E_{ex,\%} < 90\%$ and ii) “Steam release in the containment”, when $p_{max} > 75.5 \text{ bar}$. The analysis of the CRs relative to these two FC has been conducted separately, because the two outputs have shown different behaviors: a regular smooth trend for $E_{ex,\%}$ (see Fig. 7) against the strong non-smooth and multimodal nature of p_{max} (see Fig. 8); thus, different computational methods have been applied in the two cases.

The metamodel-based AK-MCS has been applied in Chapter 3 for the characterization of the CRs relative to the FC “Low heat removal” and, hence, for $E_{ex,\%}$ approximation. The Kriging metamodel obtained at the end of iterations, by resorting to a limited number (less than 200) of computationally expensive BE TH code runs, has shown a satisfactory accuracy level (estimation error around 2%). Thus, the AK-MCS procedure has been demonstrated capable of accurately identifying the combinations leading to functional failure with good accuracy and at feasible computational cost. The final Kriging obtained takes *few minutes* to predict the outcomes of *several thousands* of new input combinations to find the critical ones and identify the CRs: on the contrary, the BE-TH would take about four hours for each simulation. The CR and safe regions have been visualized by scatter plots and PCPs (see Figs. 12 and 13). Moreover, it has been shown that one of the five input variables initially chosen to describe the PSS response, A_{AV} (i.e., the flow area of the Activation Valve that opens to trigger the PSS operation), does not play a significant role in the determination of the PSS energy output. The other four parameters, as expected, affect the amount of energy exchanged by the PSS: for example, a certain percentage of non-condensable gases ($NC\%$) inside the PSS steam line causes a deterioration of the heat transfer coefficient inside the E-HX (e.g., $NC\% > 30\%$ always leads to failure). Moreover, a possible Main Steam Isolation Valve closure delay (i.e. $DEL_{MSIV} > 0 \text{ s}$) or leakage (i.e., $A_{MSIV} > 0\%$), as Activation Valve opening delay (i.e., $DEL_{AV} > 0 \text{ s}$), negatively affect $E_{ex,\%}$ (e.g., $DEL_{MSIV} = 4000 \text{ s}$, $A_{MSIV} = 0.05\%$ and $DEL_{AV} = 400 \text{ s}$ have been found to be critical values of such parameters).

On the other side, the novel framework for CRs characterization in presence of non-smooth and multimodal system state-space has been applied in [Chapter 4](#) for the analysis of the FC “Steam release in the containment”. The FMM technique has been shown capable of approximating the p_{max} irregular distribution by identifying three different clusters (associated to three different kinds of responses with respect to the failure limit of 75.5 bar). Also, the Hellinger distance method for SA has been exploited to select the input parameters mostly affecting the output clusters associated to critical conditions. By so doing, the analysis has been restricted to two relevant input parameters out of the five ones initially identified: DEL_{AV} and DEL_{MSIV} . Then, the AK-MCS technique has allowed the adaptive construction of an accurate Kriging metamodel (with increased accuracy nearby the threshold $Y_{thres} = 75.5 \text{ bar}$) to replace the time-demanding RELAP5-3D model on the reduced (two-dimensional) input space, by resorting to a limited number of simulations. Indeed, the procedure (FMM+AK-MCS) has required 300 RELAP5-3D simulations in total. Thanks to dimensionality reduction, the Kriging metamodel has managed to accurately predict the output p_{max} , despite its non-smoothness and multimodality (the estimation error has turned out to be less than 0.5% when evaluated with respect to a validation set constructed around Y_{thres}). Finally, the (output) pressure values corresponding to many new input combinations (several thousands) have been predicted by the metamodel and a unique CR has been identified and visualized with a scatter plot (see [Fig. 19](#)).

A comparison with an alternative approach of literature, namely “two-stage surrogate modeling” [18], to tackle the non-smoothness and multimodality of a system response has been carried out in [Section 4.2.4](#). In particular, different regions characterized by different behaviors of p_{max} have been identified and, then, both the output and input space have been partitioned. A SVC has been trained and coupled with the AK-MCS technique: first, new input combinations have been classified and assigned to the correct domain partition and, then, the corresponding output has been predicted to identify if it was critical or not. The results, in terms of metamodel accuracy, have been compared with those obtained by the FMM-based exploration framework proposed in this thesis, considering the same computational budget (i.e., same number of RELAP5-3D simulations). The strategy

adopting an initial dimensionality reduction based on a DBSA method supported by FMM outperforms the one relying on a SVC (see [Fig. 20](#)).

In conclusion, the two computational methods here adopted can be considered adequate for tackling the problem of CRs characterization of PSSs, but in different conditions (regular or not regular response). However, both the methods inherit the intrinsic limits of the techniques employed: on one side, the AK-MCS procedure exploiting traditional Kriging metamodels require accommodating properties like continuity and smoothness of the approximated function (otherwise large prediction error might result). This is the reason why, for the analysis of the non-smooth and multimodal output p_{max} , the support of FMM has been necessary. On the other side, the novel framework proposed, even if capable of overcoming the problem of non-smoothness and multimodality, has some limitations: if the number of parameters identified after the dimensionality reduction is not sufficiently low to be managed by a Kriging metamodel, which suffers in the presence of high-dimensionality and irregular output behavior, the success of the entire framework may be compromised.

Many of the input combinations explored by Kriging metamodels have been found to lead to the PSS failure with respect to both the FC considered, especially in case of the analysis of E_{ex} (as well illustrated in [Fig. 13](#)). This suggests that, in a hypothetical design phase, the variations of such input parameters should be limited only to a small portion of the explored ranges. Moreover, there are some input and output parameters that have not been included in the analysis of the present thesis (e.g., pipes inclination, component ruptures etc., see [Section 2.4](#)). They can be included in future works if a more complete analysis of the PSS is of concern. In that case, a SA to rank their contribution would be mandatory. Finally, these results could be even exploited in the future within a more general reliability analysis for PSSs, by adding the information about probability distributions of the input parameters considered.

References

- [1] International Atomic Energy Agency, *Safety related terms for Advanced Nuclear Plants*. Vienna (AT): IAEA-TECDOC-626, 1991.
- [2] International Atomic Energy Agency, *Passive Safety Systems and Natural Circulation in Water Cooled Nuclear Power Plants*. Vienna (AT): IAEA-TECDOC-1624, 2009.
- [3] L. Burgazzi, "Passive system reliability analysis: A study on the isolation condenser," *Nucl. Technol.*, vol. 139, no. 1, pp. 3–9, 2002, doi: 10.13182/NT139-3-9.
- [4] L. Burgazzi, "Reliability evaluation of passive systems through functional reliability assessment," *Nucl. Technol.*, vol. 144, no. 2, pp. 145–151, 2003, doi: 10.13182/NT144-145.
- [5] C. Herer, B. Dimitrov, J. M. Evrard, A. Lejosne, and E. Wattelle, "IRSN Activities related to Passive Safety Systems Assessment," *Proceedings of the International Congress on Advances in Nuclear Power Plants (ICAPP-2019)*, no. chapter 2. Juan-les-pins, France, 2019.
- [6] International Atomic Energy Agency, *Natural Circulation Phenomena and Modelling for Advanced Water Cooled Reactors*. Vienna (AT): IAEA-TECDOC-1677, 2012.
- [7] J. Jafari, F. D'Auria, H. Kazeminejad, and H. Davilu, "Reliability evaluation of a natural circulation system," *Nucl. Eng. Des.*, vol. 224, no. 1, pp. 79–104, 2003, doi:

References

- 10.1016/S0029-5493(03)00105-5.
- [8] F. Pierro, D. Araneo, G. Galassi, and F. D’Auria, “Application of REPAS methodology to assess the reliability of passive safety systems,” *Sci. Technol. Nucl. Install.*, vol. 2009, 2009, doi: 10.1155/2009/768947.
- [9] V. Picheny, D. Ginsbourger, O. Roustant, R. T. Haftka, and N. H. Kim, “Adaptive designs of experiments for accurate approximation of a target region,” *J. Mech. Des. Trans. ASME*, vol. 132, no. 7, pp. 0710081–0710089, 2010, doi: 10.1115/1.4001873.
- [10] F. Cadini, F. Santos, and E. Zio, “Passive systems failure probability estimation by the meta-AK-IS2 algorithm,” *Nucl. Eng. Des.*, vol. 277, pp. 203–211, 2014, doi: 10.1016/j.nucengdes.2014.06.025.
- [11] P. Turati, N. Pedroni, and E. Zio, “Simulation-based exploration of high-dimensional system models for identifying unexpected events,” *Reliab. Eng. Syst. Saf.*, vol. 165, no. July 2016, pp. 317–330, 2017, doi: 10.1016/j.res.2017.04.004.
- [12] P. Turati, A. Cammi, S. Lorenzi, N. Pedroni, and E. Zio, “Adaptive simulation for failure identification in the Advanced Lead Fast Reactor European Demonstrator,” *Prog. Nucl. Energy*, vol. 103, no. April 2017, pp. 176–190, 2018, doi: 10.1016/j.pnucene.2017.11.013.
- [13] G. G. Wang and S. Shan, “Review of metamodeling techniques in support of engineering design optimization,” *J. Mech. Des. Trans. ASME*, vol. 129, no. 4, pp. 370–380, 2007, doi: 10.1115/1.2429697.
- [14] K. Crombecq, L. De Tommasi, D. Gorissen, and T. Dhaene, “A novel sequential design strategy for global surrogate modeling,” *Proceedings - Winter Simulation Conference*. IEEE, pp. 731–742, 2009, doi: 10.1109/WSC.2009.5429687.
- [15] S. Missoum, Z. Gürdal, and W. Gu, “Optimization of nonlinear trusses using a displacement-based approach,” *Struct. Multidiscip. Optim.*, vol. 23, no. 3, pp. 214–221, 2002, doi: 10.1007/s00158-002-0179-1.

References

- [16] G. A. Hrinda, "Snap-through instability patterns in truss structures," *Collect. Tech. Pap. - AIAA/ASME/ASCE/AHS/ASC Struct. Struct. Dyn. Mater. Conf.*, pp. 1–12, 2010, doi: 10.2514/6.2010-2611.
- [17] E. Boroson and S. Missoum, "Stochastic optimization of nonlinear energy sinks," *Struct. Multidiscip. Optim.*, vol. 55, no. 2, pp. 633–646, 2017, doi: 10.1007/s00158-016-1526-y.
- [18] M. Moustapha and B. Sudret, "A two-stage surrogate modelling approach for the approximation of models with non-smooth outputs," *Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering, UNCECOMP 2019*, no. January. Crete, Greece, pp. 357–366, 2019, doi: 10.7712/120219.6346.18665.
- [19] M. D. McKay, R. J. Beckman, and W. J. Conover, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979, doi: 10.1080/00401706.1979.10489755.
- [20] K. Crombecq, E. Laermans, and T. Dhaene, "Efficient space-filling and non-collapsing sequential design strategies for simulation-based modeling," *Eur. J. Oper. Res.*, vol. 214, no. 3, pp. 683–696, 2011, doi: 10.1016/j.ejor.2011.05.032.
- [21] K. Crombecq, D. Gorissen, D. Deschrijver, and T. Dhaene, "A Novel Hybrid Sequential Design Strategy for Global Surrogate Modeling of Computer Experiments," vol. 33, no. 4, pp. 1948–1974, 2011.
- [22] S. S. Garud, I. A. Karimi, and M. Kraft, "Design of computer experiments: A review," *Comput. Chem. Eng.*, vol. 106, pp. 71–95, 2017, doi: 10.1016/j.compchemeng.2017.05.010.
- [23] N. C. Xiao, M. J. Zuo, and C. Zhou, "A new adaptive sequential sampling method to construct surrogate models for efficient reliability analysis," *Reliab. Eng. Syst. Saf.*, vol. 169, no. 2006, pp. 330–338, 2018, doi: 10.1016/j.ress.2017.09.008.

References

- [24] B. J. Bichon, M. S. Eldred, L. P. Swiler, S. Mahadevan, and J. M. McFarland, "Efficient global reliability analysis for nonlinear implicit performance functions," *AIAA J.*, vol. 46, no. 10, pp. 2459–2468, 2008, doi: 10.2514/1.34321.
- [25] B. Echard, N. Gayton, and M. Lemaire, "AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation," *Struct. Saf.*, vol. 33, no. 2, pp. 145–154, 2011, doi: 10.1016/j.strusafe.2011.01.002.
- [26] Z. Lv, Z. Lu, and P. Wang, "A new learning function for Kriging and its applications to solve reliability problems in engineering," *Comput. Math. with Appl.*, vol. 70, no. 5, pp. 1182–1197, 2015, doi: 10.1016/j.camwa.2015.07.004.
- [27] H. Liu, Y. S. Ong, and J. Cai, "A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design," *Struct. Multidiscip. Optim.*, vol. 57, no. 1, pp. 393–416, 2018, doi: 10.1007/s00158-017-1739-8.
- [28] D. Busby, C. L. Farmer, and A. Iske, "Hierarchical nonlinear approximation for experimental design and statistical data fitting," *SIAM J. Sci. Comput.*, vol. 29, no. 1, pp. 49–69, 2007, doi: 10.1137/050639983.
- [29] H. P. . Sacks, Jerome; Welch, William J.; Mitchell, Toby J.; Wynn, "Design and Analysis of Computer Experiments," *Stat. Sci.*, vol. 4, no. 4, pp. 409–435, 1989.
- [30] J. V. Tu, "Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes," *J. Clin. Epidemiol.*, vol. 49, no. 11, pp. 1225–1231, 1996, doi: 10.1016/S0895-4356(96)00002-9.
- [31] M. Papadrakakis, N. D. Lagaros, and Y. Tsompanakis, "Structural optimization using evolution strategies and neural networks," *Comput. Methods Appl. Mech. Eng.*, vol. 156, no. 1–4, pp. 309–333, 1998, doi: 10.1016/S0045-7825(97)00215-6.
- [32] H. Fang and M. F. Horstemeyer, "Global response approximation with radial basis functions," *Eng. Optim.*, vol. 38, no. 4, pp. 407–424, 2006, doi:

References

- 10.1080/03052150500422294.
- [33] J. H. . Friedman, "Multivariate Adaptive Regression Splines," *Ann. Stat.*, vol. 19, no. 1, pp. 1–141, 1991, [Online]. Available: <http://projecteuclid.org/euclid.aop/1176996548>.
- [34] H. Dai, B. Zhang, and W. Wang, "A multiwavelet support vector regression method for efficient reliability assessment," *Reliab. Eng. Syst. Saf.*, vol. 136, pp. 132–139, 2015, doi: 10.1016/j.ress.2014.12.002.
- [35] S. Dey, T. Mukhopadhyay, and S. Adhikari, "Metamodel based high-fidelity stochastic analysis of composite laminates: A concise review with critical comparative assessment," *Compos. Struct.*, vol. 171, pp. 227–250, 2017, doi: 10.1016/j.compstruct.2017.01.061.
- [36] P. Ranjan, D. Bingham, and G. Michailidis, "Sequential experiment design for contour estimation from complex computer codes," *Technometrics*, vol. 50, no. 4, pp. 527–541, 2008, doi: 10.1198/004017008000000541.
- [37] R. B. Gramacy and H. K. H. Lee, "Bayesian treed Gaussian process models with an application to computer modeling," *J. Am. Stat. Assoc.*, vol. 103, no. 483, pp. 1119–1130, 2008, doi: 10.1198/016214508000000689.
- [38] N. Cressie, "Spatial prediction and ordinary kriging," *Math. Geol.*, vol. 21, no. 4, pp. 493–494, 1988, doi: 10.1007/BF00897332.
- [39] J. P. C. Kleijnen and W. C. M. Van Beers, "Application-driven sequential designs for simulation experiments: Kriging metamodeling," *J. Oper. Res. Soc.*, vol. 55, no. 8, pp. 876–883, 2004, doi: 10.1057/palgrave.jors.2601747.
- [40] A. Basudhar, S. Missoum, and A. Harrison Sanchez, "Limit state function identification using Support Vector Machines for discontinuous responses and disjoint failure domains," *Probabilistic Eng. Mech.*, vol. 23, no. 1, pp. 1–11, 2008, doi: 10.1016/j.probengmech.2007.08.004.
- [41] C. Cortes and V. Vapnik, "Support-Vector Networks," vol. 297, pp. 273–297, 1995.

References

- [42] A. Guyonne, Isabelle; Elisseeff, “An Introduction to Variable and Feature Selection,” *J. Mach. Learn. Res.*, 2003.
- [43] B. Sudret, “Global sensitivity analysis using polynomial chaos expansions,” *Reliab. Eng. Syst. Saf.*, vol. 93, no. 7, pp. 964–979, 2008, doi: 10.1016/j.ress.2007.04.002.
- [44] E. Borgonovo and E. Plischke, “Sensitivity analysis: A review of recent advances,” *Eur. J. Oper. Res.*, vol. 248, no. 3, pp. 869–887, 2016, doi: 10.1016/j.ejor.2015.06.032.
- [45] S. Saltelli, Andrea; Ratto, Marco; Andres, Terry; Campolongo, Francesca; Cariboni, Jessica; Gatelli, Debora; Saisana, Michaela; Tarantola, *Global Sensitivity Analysis. The Primer*. John Wiley & Sons, 2008.
- [46] F. Di Maio, G. Nicola, E. Zio, and Y. Yu, “Ensemble-based sensitivity analysis of a Best Estimate Thermal Hydraulics model: Application to a Passive Containment Cooling System of an AP1000 Nuclear Power Plant,” *Ann. Nucl. Energy*, vol. 73, pp. 200–210, 2014, doi: 10.1016/j.anucene.2014.06.043.
- [47] J. . Saltelli, A.; Marivoet, “Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques,” *Reliab. Eng. Syst. Saf.*, 1990.
- [48] A. Saltelli, T. H. Andres, and T. Homma, “Sensitivity analysis of model output. An investigation of new techniques,” *Comput. Stat. Data Anal.*, vol. 15, no. 2, pp. 211–238, 1993, doi: 10.1016/0167-9473(93)90193-W.
- [49] A. Saltelli and I. M. Sobol’, “About the use of rank transformation in sensitivity analysis of model output,” *Reliab. Eng. Syst. Saf.*, vol. 50, no. 3, pp. 225–239, 1995, doi: 10.1016/0951-8320(95)00099-2.
- [50] G. E. B. Archer, A. Saltelli, and I. M. Sobol, “Sensitivity measures, anova-like techniques and the use of bootstrap,” *J. Stat. Comput. Simul.*, vol. 58, no. 2, pp. 99–120, 1997, doi: 10.1080/00949659708811825.
- [51] J. Nossent, P. Elsen, and W. Bauwens, “Sobol’ sensitivity analysis of a complex environmental model,” *Environ. Model. Softw.*, vol. 26, no. 12, pp. 1515–1525,

References

- 2011, doi: 10.1016/j.envsoft.2011.08.010.
- [52] E. Borgonovo, "A new uncertainty importance measure," *Reliab. Eng. Syst. Saf.*, vol. 92, no. 6, pp. 771–784, 2007, doi: 10.1016/j.ress.2006.04.015.
- [53] M. H. C. Law, M. A. T. Figueiredo, and A. K. Jain, "Simultaneous feature selection and clustering using mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1154–1166, 2004, doi: 10.1109/TPAMI.2004.71.
- [54] A. L. Gibbs and F. E. Su, "On choosing and bounding probability metrics," *Int. Stat. Rev.*, vol. 70, no. 3, pp. 419–435, 2002, doi: 10.1111/j.1751-5823.2002.tb00178.x.
- [55] E. Borgonovo, W. Castaings, and S. Tarantola, "Model emulation and moment-independent sensitivity analysis: An application to environmental modelling," *Environ. Model. Softw.*, vol. 34, pp. 105–115, 2012, doi: 10.1016/j.envsoft.2011.06.006.
- [56] S. Carlos, A. Sánchez, D. Ginestar, and S. Martorell, "Using finite mixture models in thermal-hydraulics system code uncertainty analysis," *Nucl. Eng. Des.*, vol. 262, pp. 306–318, 2013, doi: 10.1016/j.nucengdes.2013.04.030.
- [57] F. Di Maio, G. Nicola, E. Zio, and Y. Yu, "Finite mixture models for sensitivity analysis of thermal hydraulic codes for passive safety systems analysis," *Nucl. Eng. Des.*, vol. 289, pp. 144–154, 2015, doi: 10.1016/j.nucengdes.2015.04.035.
- [58] C. S. Wallace and D. M. Boulton, "An information measure for classification," *Comput. J.*, vol. 11, no. 2, pp. 185–194, 1968, doi: 10.1093/comjnl/11.2.185.
- [59] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Trans. Automat. Contr.*, vol. 19, no. 6, pp. 716–723, 1974, doi: 10.1109/TAC.1974.1100705.
- [60] G. Schwarz, "Estimating the Dimension of a Model," *Ann. Stat.*, vol. 5, no. 2, pp. 461–464, 1978.
- [61] F. . Lanfredini, M; Bersano, A.; D’Auria, "A Demonstrative Application of a

References

- Methodology for Thermal-Hydraulics Passive Systems Reliability Assessment - Extreme Cases Analysis," *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*. 2020.
- [62] P. . Bersano, A; Bertani, C.; Falcone, N.; De Salve, M.; Mascari, F.; Meloni, "Qualification of RELAP5-3D code against the in-pool passive energy removal system PERSEO data," *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*, 2020.
- [63] R. Jin, W. Chen, and T. W. Simpson, "Comparative studies of metamodeling techniques under multiple modeling criteria," *8th Symp. Multidiscip. Anal. Optim.*, no. 1998, pp. 1–13, 2000, doi: 10.2514/6.2000-4801.
- [64] J. P. C. Kleijnen, "Kriging metamodeling in simulation: A review," *Eur. J. Oper. Res.*, vol. 192, no. 3, pp. 707–716, 2009, doi: 10.1016/j.ejor.2007.10.013.
- [65] B. . Lataniotis, C.; Wicaksono, D.; Marelli, S.; Sudret, "UQLab user manual - Kriging (Gaussian process modeling)." Zurigo (SW), 2019.
- [66] M. Sedlmair, T. Munzner, and M. Tory, "Empirical guidance on scatterplot and dimension reduction technique choices," *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 12, pp. 2634–2643, 2013, doi: 10.1109/TVCG.2013.153.
- [67] A. Inselberg, *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*. 2009.
- [68] L. Liu, "Could enough samples be more important than better designs for computer experiments?," *Proc. - Simul. Symp.*, vol. 1, no. Mi, pp. 107–115, 2005, doi: 10.1109/ANSS.2005.17.
- [69] J. L. Loeppky, J. Sacks, and W. J. Welch, "Choosing the sample size of a computer experiment: A practical guide," *Technometrics*, vol. 51, no. 4, pp. 366–376, 2009, doi: 10.1198/TECH.2009.08040.

References

- [70] B. . Marelli, S.; Sudret, “UQLab: a framework for uncertainty quantification in Matlab,” *Proceedings 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014)*. Liverpool, UK, pp. 2554–2563, 2014.
- [71] J. D. Martin and T. W. Simpson, “Use of kriging models to approximate deterministic computer models,” *AIAA J.*, vol. 43, no. 4, pp. 853–863, 2005, doi: 10.2514/1.8650.
- [72] X. Wu, T. Kozlowski, H. Meidani, and K. Shirvan, “Inverse uncertainty quantification using the modular Bayesian approach based on Gaussian process, Part 1: Theory,” *Nucl. Eng. Des.*, vol. 335, no. May, pp. 339–355, 2018, doi: 10.1016/j.nucengdes.2018.06.004.
- [73] B. Iooss, “Numerical Study of the Metamodel validation process,” *First Int. Conf. Adv. Syst. Simul.*, no. 1, pp. 100–105, 2009, doi: 10.1109/SIMUL.2009.8.
- [74] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, “Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index,” *Comput. Phys. Commun.*, vol. 181, no. 2, pp. 259–270, 2010, doi: 10.1016/j.cpc.2009.09.018.
- [75] V. . Razavi, Saman; Gupta, Hoshin, “What do we mean by sensitivity analysis? The need for comprehensive characterization of “global” sensitivity in Earth and Environmental systems models,” *J. Am. Water Resour. Assoc.*, 2015, doi: 10.1111/j.1752-1688.1969.tb04897.x.
- [76] A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum Likelihood from Incomplete Data Via the EM Algorithm* , vol. 39, no. 1. 1977.
- [77] M. A. T. Figueiredo and A. K. Jain, “Unsupervised learning of finite mixture models,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, 2002, doi: 10.1109/34.990138.
- [78] N. Ueda, R. Nakano, Z. Ghahramani, and G. E. Hinton, “SMEM algorithm for mixture models,” *Adv. Neural Inf. Process. Syst.*, pp. 599–605, 1999.

References

- [79] Statovic, "Flexible mixture models for automatic clustering," 2020. <https://it.mathworks.com/matlabcentral/fileexchange/72310-flexible-mixture-models-for-automatic-clustering>.
- [80] P. Kasarapu and L. Allison, "Minimum message length estimation of mixtures of multivariate Gaussian and von Mises-Fisher distributions," *Mach. Learn.*, vol. 100, no. 2–3, pp. 333–378, 2015, doi: 10.1007/s10994-015-5493-0.
- [81] J. J. Oliver, R. a Baxter, and C. S. Wallace, "Unsupervised Learning Using MML," *Proc. 13th Int'l Conf. Mach. Learn.*, no. January 1999, pp. 364–372, 1996.
- [82] B. . Moustapha, M.; Lataniotis, C.; Marelli, S.; Sudret, "UQLab user manual - Support vector machines for classification." Zurigo (SW), 2019.