The 8th European Congress on Computational Methods in Applied Sciences and Engineering
ECCOMAS Congress 2022
5-–9 June 2022, Oslo, Norway

DATA-DRIVEN 3D SHRINKAGE POROSITY PREDICTION

Madyen Nouri[1], Julien Artozoul[1], Aude Caillaud[1], Amine Ammar[1], Francisco Chinesta[2] and Ole Köser[3]

[1] LAMPA Laboratory, Arts et Métiers Institute of Technology
Boulevard du Ronceray 2, CEDEX 01, BP 93525
F-49035 Angers, France
e-mail: madyen.nouri@ensam.eu
julien.artozoul@ensam.eu
aude.caillaud@ensam.eu
amine.ammar@ensam.eu

[2] PIMM Laboratory ESI Group Chair, Arts et Métiers Institute of Technology
CNRS, Cnam, HESAM Université
151 boulevard de l'Hôpital, 75013 Paris, France
email: francisco.chinesta@ensam.eu

[2] ESI Group
Gal. Benjamin Constant 1 c/o Fidurba SA, 1003 Lausanne, Switzerland
e-mail: ole.koeser@esi-group.com

Key words: Physics-based modelling, Model order reduction, Machine learning, Diagnosis and prognosis, Shrinkage porosity

Abstract. Casting is one of the most used processes to form metals like aluminium. A casting part can contain several defects that threaten its resistance. Shrinkage porosity is one of the major anomalies that designers try to avoid. For this purpose, rounds of numerical simulations should be performed with operating on a selection of parameters in order to minimize the presence of porosity in the casting part. In general, these approaches are time-cost with dependence on the complexity of the study case and the needed accuracy. In this paper, a methodology of data-driven porosity prediction for 3D parts is proposed in order to minimize the time-cost. A supervised learning algorithm is implemented to learn nodal porosity prediction using decision trees based method. A dataset is generated from a casting simulation software with operating on a selection of parameters. The training is realised on critical features vector extracted from nodal thermal history. Model order reduction method is used to interpolate thermal fields allover the parameter space. This interpolation is sufficiently accurate with minor errors. Promising results of shrinkage porosity prediction on a 3D study case are obtained. An evaluation of these results is performed with reference to the simulations results. This solution can contribute to open perspectives for more data-driven solutions that optimize the time-cost in the design stage.

1 INTRODUCTION

When designing a mechanical system, each component should be optimised to acquire a proper resistance depending on its role. There are various manufacturing processes that can be used to produce a metal part. Casting is one of the most used material forming methods to produce customized shapes with minimizing material consumption. However, this process can induce various types of defects that can threaten the resistance of the part and contribute on crack initialization like porosity. In casting process, alloys contract during their solidification and produce a lower volume in the final casting part compared to the liquid phase. This contraction occurs with a volume shrinkage varying from 1 to 10% depending on the alloy. It is characterized by surface depression called piping and by internal cavities called shrinkage porosity. In addition, microporosity can appear inside the part due to gas rejection. In general, shrinkage porosity is located in the last solidified regions where it cannot be compensated by residual liquid metal. The traditional solution to avoid their appearance is to add additional parts near to the last solidified zones called risers. The latter should be the last region that solidifies within the casting part. It serves as a liquid metal reserve that feeds the last solidifying region in the main part. There are some rules in the literature to apply in order to approximate the ideal position and size of these risers. Moreover, several rounds of numerical simulations might be run to optimize a selection of parameters (process parameters, chemical composition...) in order to minimize porosity inside the part or to get shrinkage porosity out of the main part or on a precise region like in the risers. Several research studies aimed at investigating the minimization of the total rate of porosity inside a cast part with reducing the time cost. For this purpose, Tsoukalas [1] and Hsu et al. [2] developped a solution based on predicting the overall rate of porosity using a multivariable linear regression from a selection of parameters (process parameters, chemical composition...). The idea that remains behind is to create an optimization loop that minimizes the total volume of porosity using Genetic algorithm that generates new optimised combination of parameters. Another solution proposed in the works of Anijdan et al. [3] and Gong et al. [4] adopts the same principle using Artificial Neural Network. Besides these methods that predict and minimize the total volume of porosity, an other efficient method proposed in [5] that allows to predict local value of shrinkage porosity depending on a selection of parameters in bi-dimensional cases. The present work is an adaptation of the same approach for tri-dimensional (3D) cases.

This paper is organized as follow: Section 2 revisits the adopted methodology to predict shrinkage porosity in 3D cases. Section 3 presents the results of temperature history interpolation and nodal porosity prediction with evaluating the efficiency of the proposed method. Finally, the achievements of this work are summarized in section 4.

2 Methodology

The implementation of our porosity prediction approach for 3D parts begins with identifying the selection of parameters to be optimized. A Design Of Experiment (DOE) is built with the variation of these parameters in a defined range. The considered study case is configured in the casting simulation software (ProCAST by ESI Group). Then, the defined number of simulations are ran following the DOE. The obtained results are composed mainly of the nodal thermal history and the corresponding nodal porosity. A modal representation of thermal his-

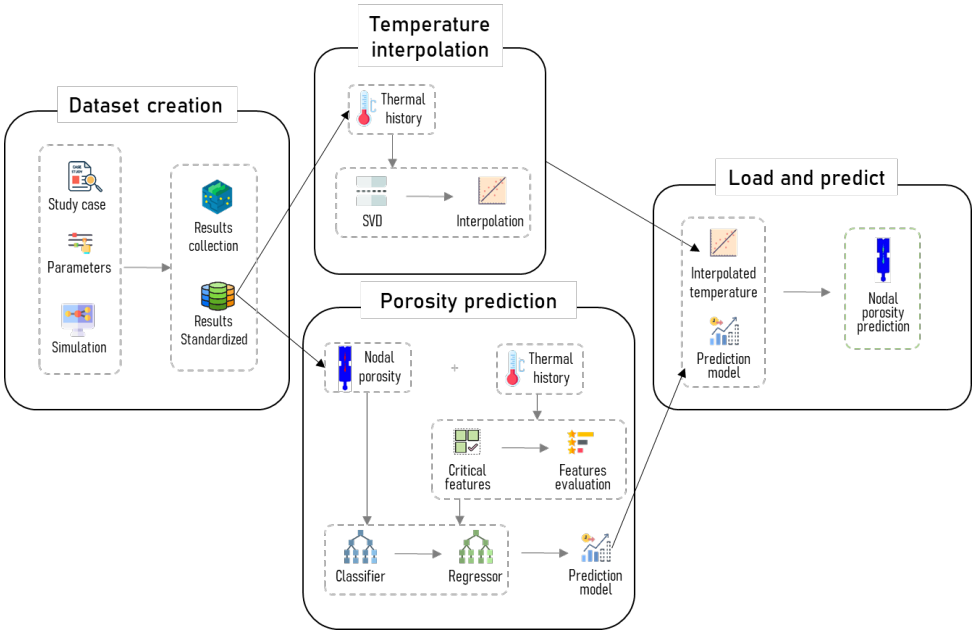

Figure 1: The proposed porosity prediction methodology on 3D cases

tory is implemented using model reduction technique to interpolate the thermal fields in other combinations of parameters. The local shrinkage porosity values are predicted using supervised learning applied on a selection of features that are extracted from the thermal history fields.

The mentioned steps of our approach are developed in the following subsections.

2.1 Dataset creation

The proposed methodology in the present paper is tested on an aluminium die casting study case. The latter is the industrial part referred in figure 2 and composed of the main body shape represented in green, the attacking part of the liquid metal colored in red, the mold is represented in blue and the added little risers above the main part in purple. The molten alloy in this cast part is an AlSi7Mg0.3 and the mold material is an AISI P20. The configuration of the study case including the process parameters and the boundary conditions are illustrated in table 1.

Table 1: Study case configuration on ProCAST

Gravity direction	-Z
Alloy initial temperature	$780°C$
Mold initial temperature	$60°C$
HTC of the interface alloy/mold	$1000\ W/m^2.K$
Ambient temperature	$50°C$

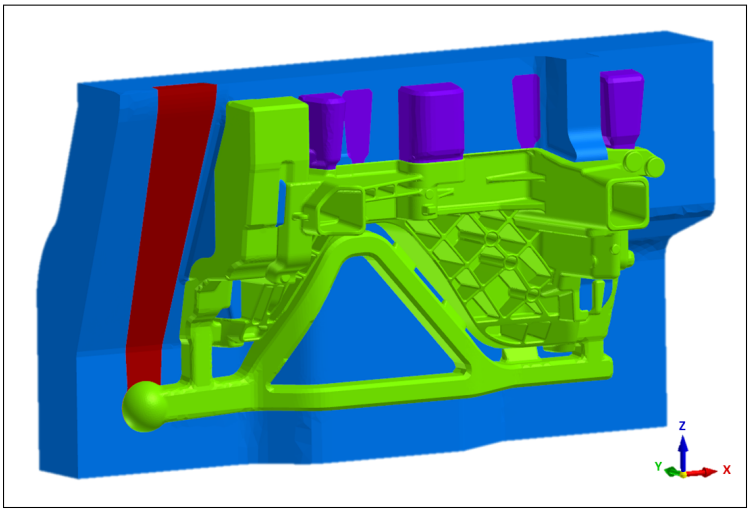

Figure 2: 3D study case

Die casting is an industrial process that allows to produce complex shapes usually in a reusable mold. This process is composed mainly of the following steps:

- Clamping: Assembling and closing the two parts of the mold. In the initial step, the mold should be cleaned and sprayed with a die coating.
- Injection: generally, in manufacturing processes the liquid metal is maintained in a set temperature. The latter is transferred to a chamber and injected to the mold with a defined pressure.
- Ejection: After waiting for the casting part to cool-down (reaching the solidus temperature at least), the mold is opened and the produced part is ejected.
- Spraying: a layer of die coating is sprayed to avoid problems if the mold is scratched or a part of the coating is accidentally damaged during the process.
- Blowing: this step serves to dry quickly the applied coating and cools-down the mold.

This manufacturing process can be implemented in an automated or semi-automated chain that performs a sequence of cycles with iterating the above tasks in order to produce a required number of casting parts with a minimum time cost. This process can be simulated using the casting simulation software ProCAST with defining the start and the end times of each task in the cycle. The configuration of our study case is illustrated in table 2.

This configuration is used in the simulation mainly to perform a thermal simulation of the mold all over the cycle that should influence the thermal field inside the cast part. The resulting temperature map of the mold is considered in the next cycle simulation. Thus, the variation of the number of the performed number affect directly the temperature field of the part and consequently the porosity distribution. Thence, in this study the number of cycles is the considered parameter to variate in the DOE. Three values of the number of cycles are defined in the

Table 2: Example of the construction of one table

Mold opening time	500 s
Part ejection time	510 s
Start time of the die spraying	515 s
End time of the die spraying	530 s
Start time of blow	550 s
End time of blow	560 s
Duration of the cycle	600 s

DOE: 1, 5 and 9. The resulting data from these three simulations are used in the next section to interpolate temperature history for other values of the identified parameter.

2.2 Modal representation of thermal history

In this section, the method of interpolation of nodal thermal history is explained.

The continuous form of the temperature field $T(P, x, t)$ is represented as a function that depends on the parameter P, the space x and the time t. This function is decomposed using Sigular Value Decomposition (SVD) [6] into a piece-wise linear functions $f(x, t)$ that depends on time-space and $g(P)$ that is function of the selected parameter:

$$T(P, x, t) = \sum_{j=1}^{j=m} f_j(x, t) \times g_j(P) \tag{1}$$

where m represents the number of the performed modes.

The temperature field on a new value of the parameter denoted P' can be interpolated using the same formula (1) with a new vector $G_j^{P'}$ computed as follows:

$$G_j^{P'} = \sum_{i=1}^{i=K} \alpha_i G_j^{P^i} \tag{2}$$

where K is the number of the considered values of P in the DOE (the number of the performed simulations) and α_i is a coefficient that depends on the interpolation method on the parametric space (piece-wise linear, Chebyshev, ...).

2.3 Machine learning

The aim of this part is to develop a methodology that allows to predict nodal porosity from the nodal thermal history for 3D study cases using supervised learning. The main idea is to prepare a set of reduced learning samples that contain an aggregation of the most relevant features extracted from the thermal history. The latter are used to implement supervised learning methods instead of training the algorithm directly using the raw time series that represent the thermal fields. In order to build a vector of physics-informed critical features, let us begin with an intuition to the process of porosity formation.

During solidification, the liquid and solid phases are separated by an interface called mushy zone where 4 feeding mechanisms may be active to compensate shrinkage porosity. They are distinguished with the solid fraction (or temperature) at the instant when it occurs. These feeding mechanisms are the liquid feeding which may be active until the Liquidus, the mass feeding that occurs until reaching the Dendritic Coherency (DC), the interdendritic feeding that appears until the Dendritic Rigidity (DR) and the solid feeding which is related to the distortion of casting. The temperatures at the isosurfaces that separate these feeding mechanisms $\{T_L, T_{DC}, T_{DR}, T_S\}$ are called in this work critical temperature and the corresponding instants are the critical times $\{t_L, t_{DC}, t_{DR}, t_S\}$. When a liquid metal is solidifying, the mushy zone grows up towards the coolest non-solidified regions driven by the gradient of temperature. This underlines the importance of considering the gradient and the Laplacian of temperature in the mushy zone especially in the isosurfaces the separate the feeding mechanisms. Thus the ith vector of critical features writes:

$$a_i = [t_1^i, .., t_4^i,\ \mathbf{U}^{t_1^i}, .., \mathbf{U}^{t_4^i},\ \mathbf{V}^{t_1^i}, .., \mathbf{V}^{t_4^i},\ \mathbf{W}^{t_1^i}, .., \mathbf{W}^{t_4^i}, \mathcal{L}^{t_1^i}, .., \mathcal{L}^{t_4^i},\ T_1 - T_{max}^{t_1^i}, ..,\ T_4 - T_{max}^{t_4^i}]$$

where $\{t_1^i, .., t_4^i\}$ correspond to the critical instants $\{t_L, t_{DC}, t_{DR}, t_S\}$ of the considered node i, $\{T_1, .., T_4\}$ are the critical temperatures $\{T_L, T_{DC}, T_{DR}, T_S\}$ related to the used alloy, $T_{max}^{t_j^i}$ is the maximum nodal temperature at the instant t_j^i, $\mathbf{U}^{t_j^i}$, $\mathbf{V}^{t_j^i}$, $\mathbf{W}^{t_j^i}$ and $\mathcal{L}^{t_j^i}$ are respectively the gradients along X, Y and Z and the Laplacian of temperature in the considered node at the instant t_j^i.
Each learning sample a_i of a considered node will be correlated with the corresponding nodal porosity using supervised learning.

The built vectors of critical features are used to implement supervised learning and mimic the porosity simulation method on ProCAST software. The later is a conditional algorithm that evaluates the solid fraction of the existing nodes at each critical time. A decision tree can be seen as a piece-wise constant approximation that can fit the best this type of algorithms.

The prediction of nodal porosity is performed incrementally with training a classifier to localize porosity nodes and then a regressor to predict the value of porosity in the localized nodes.

The first adopted method is Extremely Randomized Trees [7] (know as Extra-Trees). An ensemble trees method which builds a multitude of independent tree estimators using the entire dataset to train each tree. The final result is an aggregation of the outputs of all the built trees with voting for classification and averaging for regression. In general, a decision tree is a flowchart-like architecture that is built with evaluating the features of the learning dataset using a criterion like Geni index. The latter evaluates the features and puts the one with the best score in the root. Then, it serves to choose the best node value to split the data. This process is repeated iteratively until defining paths for all the learning samples to get from the root to the leafs of the tree. The particularity of Extra-Trees is that split nodes are randomly generated, then the most efficient one is picked to be implemented in the tree. The second algorithm tested in this paper is Extreme Gradient Boosting [8] (known as XGBoost). The latter is distinguished with building dependent trees using Gradient Boosting. After building one decision tree estimator using the entire learning dataset, a second tree is implemented to predict the residue of the previous one. A differentiable loss function is defined. Then, an

optimisation of this function is performed using gradient descendant. The process is repeated iteratively until reaching a predefined loss or the maximum number of estimators.

The results of the performed tests using these two methods are illustrated in the next section.

3 Results

After performing a modal representation of thermal history on our 3D study case, the interpolated temperature fields of the 3^{rd} and the 7^{th} cycles using the prepared dataset are illustrated in figure 3 with the corresponding ProCAST simulations.

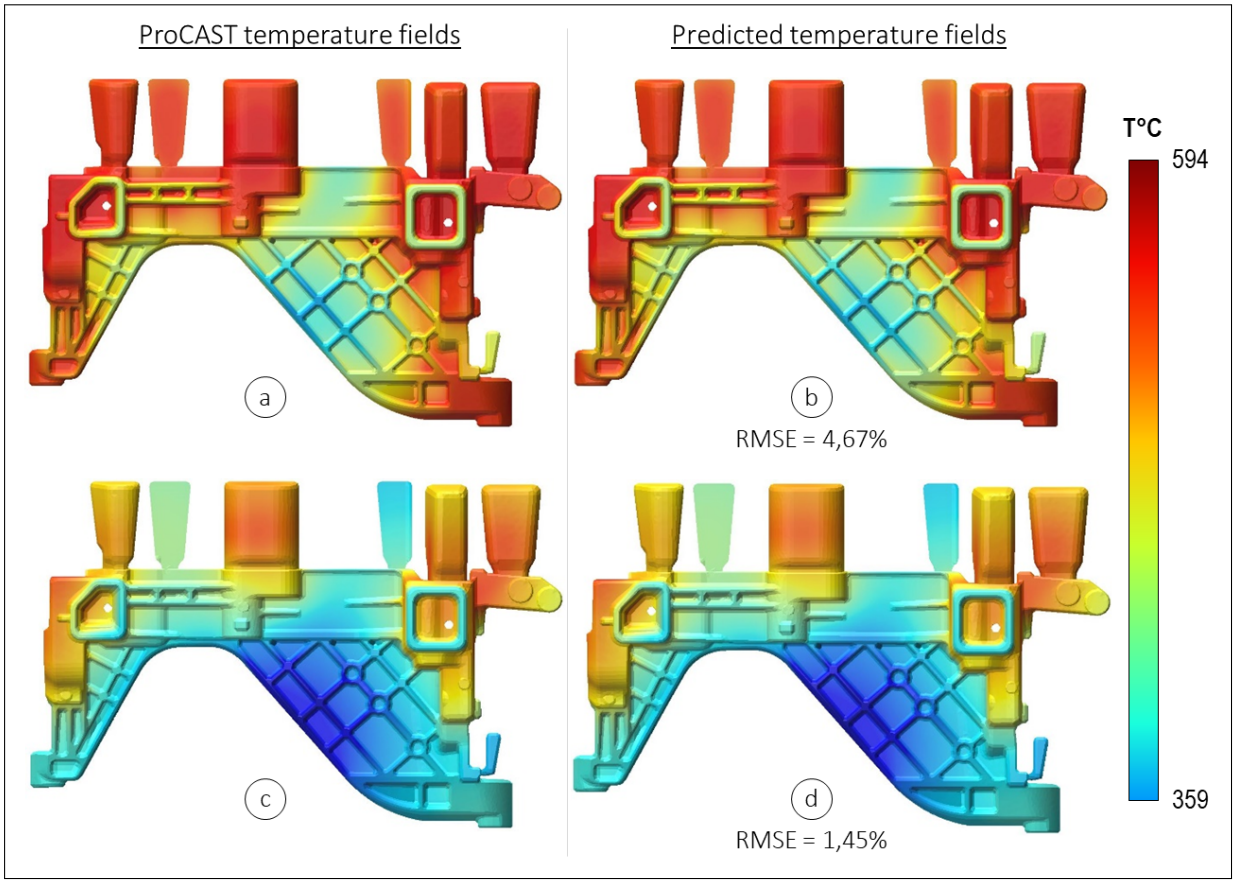

Figure 3: Temperature history interpolation
a and b are the results on the 3^{rd} cycle. c and d are the results on the 7^{th} cycle

The cast part starts to down from the thin parts as can be seen on the ribs towards the thicker regions until getting to the risers. Minor difference can be seen visually between the simulated and the interpolated temperatures fields. Root mean squared error (RMSE) of the interpolations is computed allover each part and represented in the same figure 3. The value of the obtained errors can be considered as low and sufficient to exploit the interpolated fields in nodal porosity prediction.

For supervised learning stage, 6 simulations (instead of 3 for the temperature history inter-

polation) are performed and used to train the identified supervised learning algorithms. The simulated number of cycles values are: 1, 2, 4, 5, 6, 8 and 9. The obtained results from learning nodal porosity using Extra-Trees and XGBoost are represented in figure 4.

The predicted porosity distribution using XGBoost method is visually more accurate than Extra-Trees in localizing shrinkage porosity compared to the simulated porosity distribution generated by ProCAST. For better evaluation of the predicted classification, confusion matrices and Precision-recall curves are represented in figure 5. A confusion matrix shows the reference classification against the predictions. Besides, Precision represents the true detected porosity among all the detected porosity and the recall (or sensitivity) is the fraction of relevant detected porosity and the reference porosity. A high area under the curve represents both high recall and high precision. In case of unbalanced data as our study case, this representation helps to evaluate the efficiency of a prediction method.

The confusion matrices show that the classification predictions are more accurate in detecting the class 0 (nodes that not containing porosity) that class 1 (nodes containing porosity). However, this metric is less accurate in evaluating the porosity detection because it does not consider an efficient porosity prediction with an offset position. The illustrated Precision-recall curves show that XGBoost approach was more accurate in predicting porosity distribution.

Unlike the classification prediction where XGBoost were more efficient especially in predicting porosity distribution on the 3^{rd} cycle, the performance of the two tested methods in regression was roughly equal. The RMSE of Extra-Trees and XGBoost in regression evaluated on the 3^{rd} and the 7^{th} cycles are illustrated in table 3.

Table 3: RMSE of the regression applied to test data

	RMSE of 3^{rd} cycle prediction	RMSE of 7^{th} cycle prediction
Extra-Trees	9,59%	8,9%
XGBoost	9,30%	8,54%

4 Conclusion and perspectives

Machine learning can help to reduce the time cost of the design stage with mimicking some numerical simulation methods and generating real-time results with a defined accuracy.

In this paper, a porosity prediction methodology for 3D parts is proposed. The major achievements in this work are enumerated below:

- Nodal porosity prediction methodology for 3D study case is proposed with testing two supervised learning approaches: Extra-Trees and XGBoost.
- An efficient temperature history interpolation method is implemented for 3D parts.
- XGBoost is more efficient in porosity localisation and produces promising results that can be improved in future works.

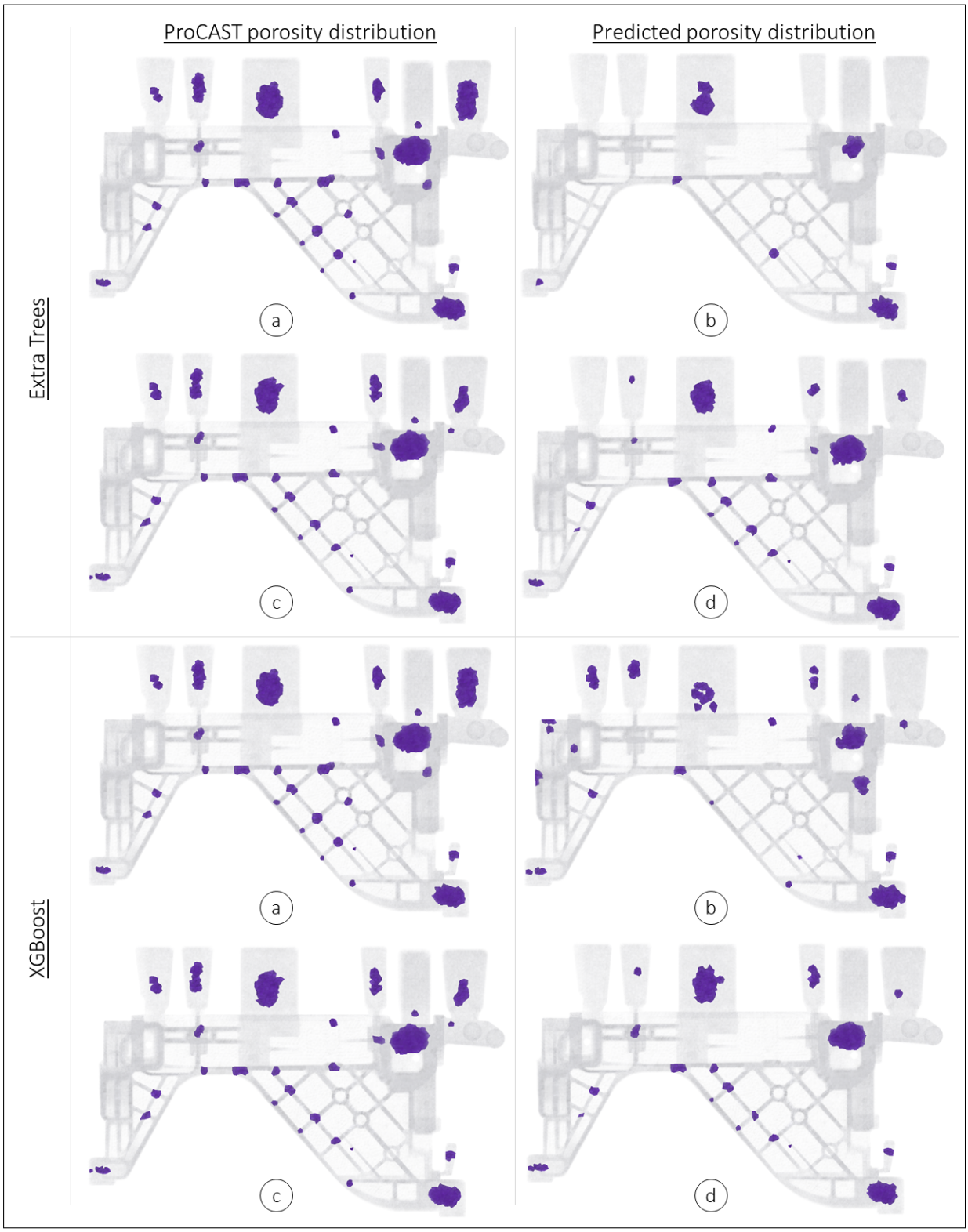

Figure 4: Porosity prediction using Extra-Trees and XGBoost
a and b are the results on the 3^{rd} cycle. c and d are the results on the 7^{th} cycle

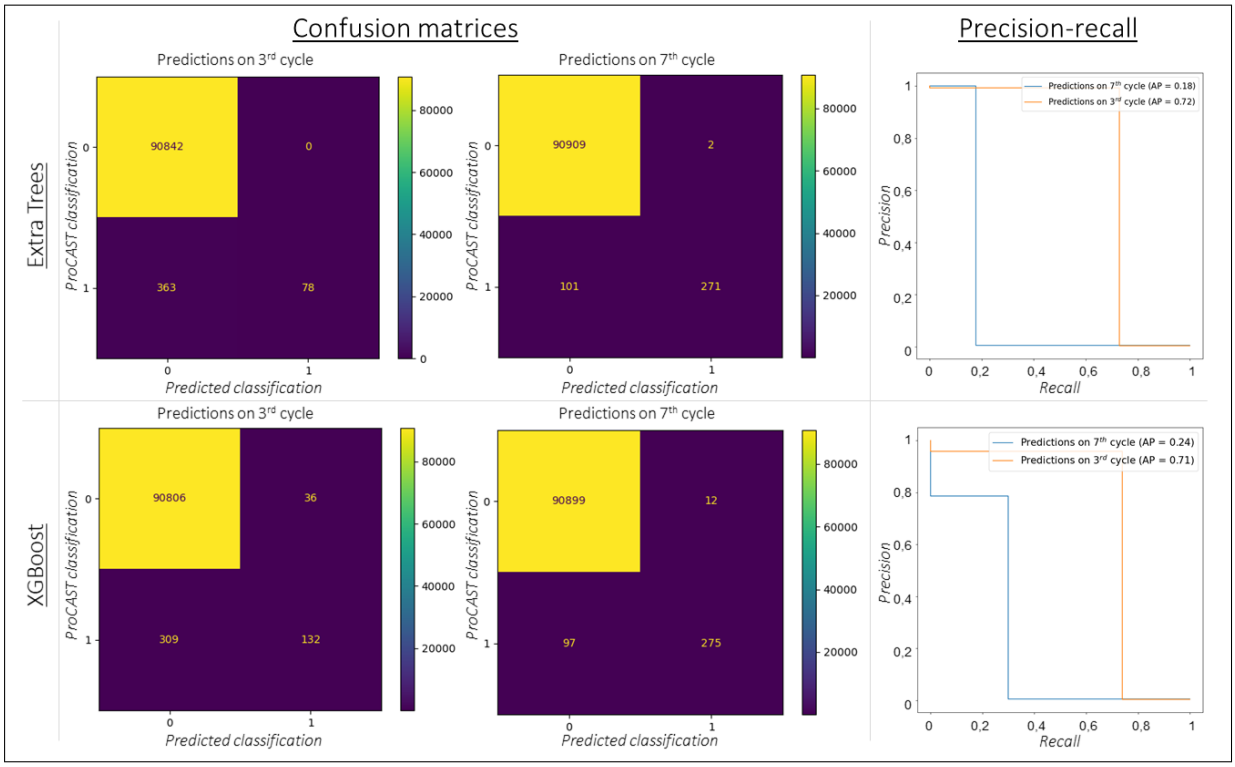

Figure 5: Temperature history prediction

- The supervised learning is applied on a critical features vector extracted from temperature history.

REFERENCES

[1] V. D. Tsoukalas. Optimization of porosity formation in AlSi9cu3 pressure die castings using genetic algorithm analysis. 29(10):2027–2033.

[2] Quang-Cherng Hsu and Anh Tuan Do. Minimum porosity formation in pressure die casting by taguchi method. 2013:e920865. Publisher: Hindawi.

[3] S. H. Mousavi Anijdan, A. Bahrami, H. R. Madaah Hosseini, and A. Shafyei. Using genetic algorithm and artificial neural network analyses to design an al–si casting alloy of minimum porosity. 27(7):605–609.

[4] Xue-dan Gong, Dun-ming Liao, Tao Chen, Jian-xin Zhou, and Ya-jun Yin. Optimization of steel casting feeding system based on BP neural network and genetic algorithm. 13(3):182–190.

[5] M. Nouri, J. Artozoul, A. Caillaud, A. Ammar, F. Chinesta and O. Köser. Shrinkage porosity prediction empowered by physics-based and data-driven hybrid models. 15 (3), 1-16.

[6] G. H. Golub and C. Reinsch. Singular value decomposition and least squares solutions. 14(5):403–420.

[7] Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine learning*, 63(1):3–42, 2006.

[8] Tianqi Chen, Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. 2016. p. 785-794.