Institute of Fundamental Technological Research
Polish Academy of Sciences

Department of Intelligent Technologies

A method for predicting the size of damage to gas turbine components based on machine learning algorithms, applicable at the early stages of parts' lifetime

Ph.D. dissertation

Maciej Badora, M.Sc.Eng.

Supervisor: Prof. Tomasz Szolc, Ph.D., D.Sc., Mech.Eng.

Warsaw, 2024

This research was cofounded by the Ministry of Science and Education and conducted under the “Industrial Doctorate” program (decision number 0080/DW/2018/02) in cooperation between the Institute of Fundamental Technological Research Polish Academy of Sciences and BH Poland sp. z o.o. (Baker Hughes Company).

Assistant advisor designated by the entrepreneur, in accordance with the rules of the “Industrial Doctorate” program: Przemysław Bartosik, M.Sc.Eng.

Special thanks

I express my sincere gratitude to Professor Tadeusz Burczyński
for his invaluable remarks and suggestions, which significantly improved this dissertation.

Acknowledgments

I want to express my gratitude to the supervisor, Professor Tomasz Szolc,
for his guidance, continuous support, optimism and motivation to work.

I also want to thank the co-supervisor, Przemysław Bartosik, M.Sc.Eng.,
for the effective collaboration and help offered in the most difficult stages of this research.

I want to acknowledge Antonino Graziano, Marcin Bielecki and Marzia Sepe,
my colleagues from Baker Hughes Company, for mentoring and challenging me and for the many discussions and guidance that helped me overcome countless technical difficulties.

I thank my managers, Fausto Carlevaro, Katarzyna Przybyłowicz and Leszek Torbicz,
for creating a comfortable research environment and their care and patience.

I want to give special thanks to my parents, Małgorzata and Zbigniew,
for the effort put into my education, constant help and teaching me that hard work pays off.

Above all, I would like to thank my wife, Anna.
These have been exhausting years for both of us. Thank you for your constant support, patience,
strength and caring for our family during this period.

Contents

Abstract

This dissertation focuses on the problem of effectively predicting damage size to gas turbine components using machine learning algorithms when the availability of empirical data on damage size measurements is limited or the data are missing. This research addresses the problems of estimating the maximal length of fatigue cracks found on the trailing edges of high-pressure nozzles and predicting the wall thickness reduction due to the oxidation of transition pieces. The analyzed objects are the components of the turbine and combustion systems of two different heavy-duty gas turbines manufactured by Baker Hughes Company. The research objective is to formulate a method based on machine learning algorithms for predicting the damage growth that can be applied for various types of failures, regardless of the empirical data limitations, allowing to obtain estimates consistent with prior knowledge and physical laws about the analyzed phenomena.

The first stage of the research focuses on the application of popular statistical learning algorithms to the problem of fatigue crack propagation using a sample comprising 25 damage size measurements. A custom cost function and a controlled approach to preparing the training, test and validation sets are applied. The final predictions based on actual operational data can be accurate despite the limited sample size. In the following research stage, physics-informed neural networks are applied to the same technical problem to build predictive models with better generalization and extrapolation capabilities and obtain estimates consistent with the relevant equations describing crack growth due to low cycle fatigue. A recurrent neural network with an embedded multilayer perceptron is trained using a dynamically changing custom cost function, which coefficients depend on the consistency with the underlying physical laws. The obtained predictions are accurate regardless of training based on ten, two, or only one observation. A normalized error evaluated against unseen data and characteristics of the perceptron were used to evaluate the predictions' credibility. The novel method of single-source domain generalization and cross-domain knowledge transfer in regression analysis based on physics-informed neural networks is proposed. It enables the effective application of an unchanged cumulative damage model built using a physics-informed neural network, which was trained in a domain where measurement data on damage size are available in domains, where such data are unavailable. The procedure consists of eight steps, but the so-called shifting of the operational data and neural network's inputs is essential. It is applied to predict fatigue crack growth in a simulated scenario where damage size measurements are unavailable. The final predictions are similar to those obtained using a predictive model trained from the ground up in the target domain. In the last research stage, physics-informed neural networks and the proposed method are applied to estimate the metal loss due to oxidation. The obtained results are satisfactory and provide evidence that this method has a universal character.

The main novelty of this research is the method of single-source domain generalization and cross-domain knowledge transfer based on physics-informed neural networks. It allows for predicting the size of damage to gas turbine parts in domains without any measurements, applying appropriately a physics-informed neural network trained in a source domain where damage size measurements are available. It is intended to make predictions for fielded products, but it can also provide valuable outcomes to support the design phase, combining knowledge extracted from historical empirical data and numerical simulation results.

Streszczenie

Niniejsza rozprawa doktorska poświęcona jest problemom związanym ze skutecznym przewidywaniem rozmiaru uszkodzeń komponentów turbin gazowych za pomocą algorytmów uczenia maszynowego, gdy liczba danych dotyczących pomiarów uszkodzeń jest ograniczona lub są one niedostępne. Zrealizowane prace badawcze dotyczą estymacji maksymalnej długości pęknięć zmęczeniowych występujących na łopatkach kierowniczych wysokiego ciśnienia oraz przewidywania redukcji grubości ścianki elementu łączącego komorę spalania z sekcją turbiny w wyniku utleniania. Analizowane obiekty to części systemów turbiny oraz spalania dwóch różnych przemysłowych turbin gazowych wytwarzanych przez firmę Baker Hughes Company. Celem badań jest opracowanie metody predykcji rozmiaru uszkodzeń bazującej na algorytmach uczenia maszynowego, którą można zastosować do różnego rodzaju defektów, niezależnie od liczby dostępnych pomiarów uszkodzeń. Rezultaty wyznaczone za pomocą tej metody powinny być zgodne ze stanem wiedzy na temat badanego zjawiska i jego opisem teoretycznym.

Pierwszy etap prac badawczych dotyczy szacowania długości pęknięć zmęczeniowych za pomocą popularnych algorytmów uczenia maszynowego, bazując na próbce zawierającej 25 pomiarów uszkodzeń. Utworzono niestandardową funkcję kosztu, a elementy zbioru uczącego, testowego oraz walidacyjnego dobrano w sposób nadzorowany. Rzeczywiste dane operacyjne są użyte do wyznaczenia końcowych rezultatów, które mogą być dokładne pomimo nielicznej próbki. Kolejny etap badań dotyczy tego samego problemu technicznego, ale w celu uzyskania lepszej generalizacji i zdolności do ekstrapolacji oraz otrzymania przewidywań zgodnych z równaniami opisującymi wzrost długości pęknięć zmęczeniowych, zastosowano sztuczną sieć neuronową opartą na prawach fizyki (ang. physics-informed neural network). Źródło tego tłumaczenia stanowi opublikowana w 2021 roku praca zbiorowa[1], lecz może ono budzić pewne kontrowersje, gdyż sztuczne sieci neuronowe mają inspirację biologiczną, a nie fizyczną. Utworzony model bazuje na rekurencyjnej sieci neuronowej, w którą wbudowano perceptron wielowarstwowy. W trakcie trwania procesu uczenia współczynniki funkcji kosztu zmieniają się dynamicznie zależnie od zgodności otrzymanego rozwiązania z równaniem teoretycznym opisującym badane zjawisko. Dokładność końcowych rezultatów jest satysfakcjonująca mimo zbiorów uczących składających się z dziesięciu, dwóch czy jednego elementu. Wiarygodność tych wyników oceniano na podstawie wartości znormalizowanego błędu wyznaczonego na zbiorach walidacyjnych, które nie służyły do uczenia sieci neuronowych, oraz na podstawie zgodności charakterystyk perceptronu z prawami fizyki opisującymi analizowany problem. W pracy tej zaproponowano nową metodę generalizacji dziedziny oraz transferu wiedzy pomiędzy dziedzinami w analizie regresji, która wykorzystuje sieci neuronowe oparte na prawach fizyki. Umożliwia ona skuteczne zastosowanie modelu predykcyjnego bazującego na wspomnianych sieciach neuronowych i wyuczonego w dziedzinie, gdzie dane pomiarowe dotyczące rozmiaru uszkodzeń są dostępne, w dziedzinach, w których takich danych brakuje. Całość procedury stanowi osiem kroków, lecz najistotniejszym z nich jest translacja wartości danych operacyjnych i danych wejściowych do sieci neuronowej. Metodę tą zastosowano do predykcji długości pęknięć zmęczeniowych, symulując brak pomiarów uszkodzeń. Otrzymane rezultaty są zgodne z obliczeniami modelu predykcyjnego, który został wyuczony od podstaw w dziedzinie docelowej. W ostatnim etapie badań, sztuczne sieci neuronowe oparte na prawach fizyki wraz z proponowaną metodą użyto do szacowania ubytku materiału w wyniku utleniania. Uzyskane rezultaty są satysfakcjonujące i potwierdzają uniwersalny charakter tej metody.

[1] Moszyński, M., Borzyszkowski, B., Damaszke, K., Romankiewicz, J., and Świniarski, M., 2021, "Sieci neuronowe oparte na prawach fizyki," *Uczenie maszynowe i systemy rozproszone*, Wydawnictwo Politechniki Gdańskiej, Gdańsk, pp. 110–119.

Proponowana metoda generalizacji dziedziny i transferu wiedzy pomiędzy dziedzinami w analizie regresji, która wykorzystuje sieci neuronowe oparte na prawach fizyki, jest oryginalnym rozwiązaniem problemu naukowego i stanowi główny element nowości. Metoda ta umożliwia przewidywanie rozmiaru uszkodzeń komponentów turbin gazowych w przypadku braku pomiarów w wybranej dziedzinie, poprzez odpowiednie wykorzystanie sieci neuronowej opartej na prawach fizyki, którą wyuczono na podstawie danych empirycznych z innej dziedziny. Poza zastosowaniami do już pracujących turbin gazowych, metoda ta może posłużyć do wsparcia procesu projektowania, umożliwiając skuteczne łączenie danych empirycznych z rezultatami symulacji numerycznych.

List of abbreviations

ANN	Artificial neural network
CV	Cross-validation
DAM	Damage accumulation model
ESD	Emergency shutdown
FEA	Finite element analysis
FH	Fired hours
FNN	Feedforward neural network
FS	Fired starts
HGP	Hot gas path
LCF	Low cycle fatigue
MAE	Mean absolute error
ML	Machine learning
MLP	Multilayer perceptron
MSE	Mean squared error
OEM	Original equipment manufacturer
PINN	Physics-informed neural network
ReLU	Rectifier linear unit
RM&D	Remote Monitoring & Diagnostics
RMSE	Root mean squared error
RNN	Recurrent neural network
S1N	1st stage nozzle
SVR	Support vector regression
TBC	Thermal barrier coating
TP	Transition piece
XGBoost	Extreme gradient boosting

1. Introduction

Global energy demand grew by 1% in 2022, exceeding the 2019 pre-pandemic levels by around 3% [1]. In 2022, global electricity generation increased by 2.3%, oil consumption raised by 3% and global natural gas production was constant compared to 2021. Petrochemical and power plant operators are forced to maximize production due to constantly growing energy needs. Concerning turbomachinery, this target can be achieved by investing in state-of-the-art rotating equipment with the highest efficiency and long mean time between maintenance, improving the performance of currently available assets through uprates, or optimizing their operation without significant capital expenditures on modernization. The latter may include analyses aimed at maximizing uptime and availability without jeopardizing reliability. Such requests concerning gas turbines are often raised by operators of natural gas liquefaction facilities, where the production process has a continuous character, and each hour of downtime results in a notable revenue loss. Such optimization may be related to single or multiple extensions of time between maintenance, shortening the duration of overhauls and inspections, alignment of a maintenance schedule with production plans, or applying modern strategies like predictive or prescriptive maintenance instead of traditional ones like corrective and preventive maintenance. Other requests relate to the feasibility of gas turbine output increase remaining the mean time between maintenance unchanged or the ability to compensate service at peak load by the operation with lower loads in other periods. Such requirements exert pressure on turbomachinery manufacturers, which must constantly push the boundaries of their products.

Gas turbines are continuous-flow internal combustion engines comprising the following main sections: a compressor system, a combustion system and a turbine system. These devices are complicated in terms of design and maintenance mainly, but not exclusively, due to the high temperatures of gases in the combustor and turbine sections. Components that operate in such temperatures or are subjected to significant temperature changes will deteriorate due to thermo-mechanical fatigue, oxidation, overheating, or creep, which are the most common causes of failures of gas turbine parts. Therefore, the turbine's mean time between maintenance results from the endurance and maximum capabilities of combustion and turbine section components. These parts' capabilities are mainly determined during the design process, the selection of materials, protective coatings and manufacturing methods. Nevertheless, also a lot can be done to optimize the engine's operation after the design phase, when the units are deployed in the field and some empirical data are already available. Gas turbines are equipped with numerous sensors recording selected operating parameters. These large volumes of data are primarily used for monitoring, diagnostics and anomaly prediction purposes. Moreover, one of the most important data relates to the technical condition of hardware after service. This information is recorded during disassembly inspections, including visual, dimensional and non-destructive checks preceding repair activities. Considering a certain number of parts for which the damage size measurements and the historical operational data are available, these two datasets can be combined in order to approximate the damage size by means of a function of selected operating parameters. Such models allow departing from fixed intervals between planned outages enabling predictive and prescriptive maintenance. They can be applied to estimate the damage growth, the percentage of components expected to pass an inspection without a need to be repaired, the probability of an unplanned outage, or the expected percentage of items scrapped during the repair. The data-driven insights are used to support decisions on the operation and maintenance of a single unit and the entire fleet, helping to maximize revenues while assuring the highest availability and reliability levels. Therefore, original equipment manufacturers (OEMs) invest in remote monitoring systems, their predictive capabilities and platforms for

asset performance management and optimization to satisfy their customers. Advanced machine learning (ML) and artificial intelligence techniques are applied in each of these areas.

Machine learning algorithms are utilized to predict future data based on the knowledge extracted from historical data. These algorithms are applied for classification or regression problems, depending on the output variable type (i.e. whether categorical or continuous, respectively). Concerning gas turbines, machine learning methods are used to detect anomalies and faults based on recorded operating parameters, create synthetic sensors, predict degradation and damage to components due to various failure types, evaluate residual useful life of parts and the entire engine, or adjust gas turbine controls in real time. Despite many successful applications of these methods available in the literature, the most visible limitations are related to the lack of use of actual operational data, a limited number of papers focusing on finding root causes of damage and modeling the deterioration of engine parts due to them, a small number of studies about heavy-duty gas turbines and the lack of deployment of the proposed solutions in industrial applications. Moreover, many of these methods require much historical data to extract the knowledge effectively, may give results inconsistent with prior knowledge about the analyzed phenomenon and cannot be used to extrapolate. However, combining machine learning algorithms with theoretical knowledge into a hybrid model can be effective even if the availability of empirical data is significantly limited. Generally, a model represents numbers, structures, mathematical operations, rules and other algorithm-specific attributes required to make predictions given proper input data. These attributes are determined during training by applying the specific machine learning algorithm to the historical data. In this research, hybrid models are based on physics-informed neural networks (PINNs), which are neural networks with embedded equations describing the analyzed problem. The equations limit the space of allowable solutions and should improve the quality and accuracy of the final estimates. The effectiveness of PINNs and several data-driven machine learning algorithms is evaluated concerning the problem of fatigue crack propagation found on the trailing edges of high-pressure nozzles of a heavy-duty gas turbine. Furthermore, a PINN is applied to predict the wall thickness reduction due to oxidation in industrial engine transition pieces (TPs). The predictions are based on actual operational data and refer to gas turbines manufactured by Baker Hughes Company. Considering the limitations in the quantity and quality of the operational data and damage size measurements, the accuracy of the obtained damage size predictions is satisfactory for both problems.

In case of limited availability of training data, transfer learning methods can be applied to improve prediction capabilities in the so-called target domain by using knowledge about the source domain, empirical data available there and the predictive function (i.e. the model) created based on that data. Generally, the analyzed object, the damage type and its location and the entire data set available to create a new model or apply an existing one are referred to as a domain. The most popular form of transfer learning, widely used for deep neural networks, applies a valid model trained in the source domain and fine-tunes it in the target domain. This approach assumes the availability of data in the target domain necessary to adjust the neural network weights. However, in practical applications, this assumption may not be satisfied. In such cases, domain generalization methods are used to extract domain-invariant representations from multiple source domains and achieve good out-of-distribution generalization capabilities. Nevertheless, studies on domain generalization focus mainly on classification problems, with a limited number of papers related to rotating equipment. This research addresses this gap. The primary novelty resulting from this study is a method of single-source domain generalization and cross-domain knowledge transfer in regression analysis, which is based on physics-informed neural networks. The method aims to effectively apply a cumulative damage model built using a PINN, i.e. the so-called source hybrid model, which was trained in a source domain

where measurement data on damage size are available, in a target domain, where such data are unavailable. This method can be applied if the damage type in the target domain, the physical phenomena causing it and the applicable theoretical or experimentally determined equations for estimating the damage increment in a single time step are the same as in the source domain. The so-called shifting of the operational data and PINN's input variables in the target domain is the essential step that enables an effective application of the unchanged source hybrid model in the target domain. The predictions in the target domain for the original set of operational data are obtained by applying the source hybrid model to the set of shifted operational data. The proposed method is effectively applied for the analyzed problems of fatigue crack growth and wall thickness reduction due to oxidation, assuming no damage size measurements are available in the target domains. The results obtained using the proposed method are comparable with the predictions determined by means of hybrid models trained from the ground up in the target domains, i.e. with the same structure as the source PINNs, and are considered satisfactory.

The proposed method can be applied effectively at the early stages of the part's lifetime when the availability of damage size measurements is limited and the number of observations is small. It has a universal character and can be used for modeling damage due to various failure types found on the main components of gas turbines. Nevertheless, this method can also be applied to other turbomachinery assets, including compressors, pumps or steam turbines, if the minimal requirements for data availability are satisfied (i.e. they are listed in Chapter 4). Experience gained during this research will be used in order to optimize the operation of rotating equipment manufactured by Baker Hughes Company, assuring the highest levels of safety and reliability.

2. Objectives, theses and scope of the dissertation

An ability to estimate in real-time and predict damage to critical components of gas turbines is crucial to ensuring the reliability and availability targets are met. Empirical data describing the extent of damage and operating conditions under which parts worked may be utilized to build a purely empirical model, or to adjust a physics-based model of the analyzed phenomenon. However, the availability of data may be limited due to various reasons. Building a credible model based on a small sample is challenging, particularly if the applied algorithm performs well in an abundance of data, like in the case of many machine learning methods. Nevertheless, hybrid models that combine such algorithms with theoretical knowledge about the analyzed failure type may provide valuable results even in the small data regime. The effectiveness of selected machine learning algorithms in such scenarios has been evaluated and presented in this dissertation.

2.1 Objectives of the dissertation

The main objective of this research is to create a method based on machine learning algorithms to predict the size of damage to gas turbine parts. The specific aims are as follows:

- models created by means of this method shall respect prior knowledge and physical laws that govern the analyzed phenomenon;
- the proposed method shall be universal and applicable for modeling damage due to various failure types found on the main components of gas turbines;
- the method shall be applicable at the early stages of the components' lifetime, e.g. during the design phase, first tests of the part, or first uses in the market before the product reaches maturity;
- application of the proposed method shall be possible and provide valuable insights even if the availability of damage measurements is limited and the number of observations is small, i.e. less than 30 data points to train and test the model.

These objectives were achieved by creating hybrid models based on physics-informed neural networks combined with a novel method of domain generalization.

This research was conducted in cooperation between BH Poland sp. z o.o. (Baker Hughes Company) and the Institute of Fundamental Technological Research of the Polish Academy of Sciences. The objectives defined by the company reflect the voice of its customers. Increasing energy demands require higher availability of gas turbines and longer intervals between planned outages without jeopardizing reliability. The application of machine learning algorithms in maintenance optimization aligns with state-of-the-art solutions applied in the industry and the company's digital strategy.

2.2 Theses of the dissertation

The research presented in this dissertation was aimed at proving the following theses:

1) Models based on machine learning algorithms can accurately predict the size of damage to gas turbine parts in domains characterized by the limited availability of damage measurements or lack of them using the proposed method of domain generalization.

2) Damage accumulation models (DAMs) based on machine learning algorithms can be trained effectively, generalize well and provide accurate estimates even if the sample is composed of several observations only, e.g. less than or equal to 10 data points.
3) The extent of damage to gas turbine parts can be predicted accurately by hybrid models combining machine learning algorithms with theoretical knowledge about the analyzed failure type, even if the time series of turbine operating parameters are incomplete.

2.3 Scope and structure of the dissertation

The research was divided into the following subprojects:

- Prediction of the length of fatigue cracks found on the trailing edges of high-pressure nozzles of a heavy-duty gas turbine manufactured by Baker Hughes Company.
- Prediction of the wall thickness reduction due to oxidation in transition pieces of industrial gas turbines manufactured by Baker Hughes Company.

The analyzed objects, i.e. the high-pressure nozzles and transition pieces, are the parts of two different models of heavy-duty gas turbines, two separate systems (i.e. combustion system and high-pressure turbine system, respectively), perform different functions and are subjected to different loads. Moreover, the analyzed failure types, i.e. cracks due to low cycle fatigue (LCF) and base material depletion due to oxidation, are governed by different physical laws. Such objects and failure types were chosen intentionally to demonstrate that the proposed predictive method gives satisfactory results for various problems and setups. Nevertheless, a mandatory prerequisite to applying this method is the availability of damage measurements and the corresponding gas turbine operational data captured by a data acquisition system.

The first stage of this research focused on the fatigue cracks of the 1st stage nozzles (S1N). The scope was to apply several popular statistical learning algorithms (i.e. multiple linear regression, polynomial regression, kernel-based methods, random forest, AdaBoost and extreme gradient boosting and artificial neural networks) on a small dataset composed of 31 observations and understand if those methods are sufficient to achieve the objectives of this research. The most important contributions of this part of the research are as follows:

- Selection of operating parameters, i.e. time series captured by the Remote Monitoring and Diagnostics (RM&D) system, that are meaningful in the process of modeling the growth of fatigue cracks.
- A fully controlled manner of splitting the small dataset into training, validation and test subsets was applied, which reduced ambiguity during the cross-validation, testing and interpretation of results.
- A custom cost function that favors solutions that accurately predict the longest cracks was created and applied.
- The obtained results confirm that some of these algorithms can accurately predict the length of analyzed cracks, even if the sample is composed of 31 data points only.

A detailed description of this part of the research is given in Chapter 4.

The second stage of this research was related to the same object and failure type. However, the objectives evolved and were about building models based on machine learning algorithms with better generalization and extrapolation capabilities, which are more consistent

with theoretical formulations describing the propagation of fatigue cracks. Therefore, it was decided to proceed with physics-informed neural networks, considering the many advantages this class of algorithms may offer in the analyzed scenario. The most important contributions of this part of the research are as follows:

- An approach was proposed and applied to calculate thermal stresses at shutdown based on operating parameters at steady-state conditions and by using selected results of the numerical analysis.
- A custom cost function was created with a variable parameter regulating the balance between empirical data and physical laws during the training process.
- An approach to expand the source domain with synthetically created data was proposed and applied, thus, enlarging the subspace in which the hybrid model based on the physics-informed neural network obeys the theoretical formulations.
- The results prove that PINNs can be trained effectively and provide credible estimates when the training subset comprises ten, two, or only one observation.

This part of the research is described in the first three sections of Chapter 5.

The first application of the proposed method for predicting the size of damage to gas turbine components is documented in the remaining sections of Chapter 5. The analyzed object and the failure type remained the same. The main objective of this stage of the research was to train a model and extract knowledge in the source domain, in which damage measurements are available, and apply this knowledge in the target domain to obtain accurate and credible estimates of the damage size, despite the lack of observations in that domain. Both domains refer to the same nozzle segment but two different positions on the airfoils' trailing edges. The most important contributions of this part of the research are as follows:

- Formulation of the method for predicting the size of damage to gas turbine parts, which is based on building physics-based dependency between the source and target domains for domain generalization purposes coupled with a physics-informed neural network.
- Prediction of the nozzles fatigue crack lengths by means of the proposed method in case of the absence of crack length measurements in the target domain.
- The obtained results prove that the method is effective and enables predicting in the target domain, even if damage size measurements are unavailable.

The third stage of this research focused on modeling the reduction of transition pieces' wall thickness due to oxidation. It is described in Chapter 6. The main objective was to evaluate the flexibility and versatility of the proposed method when applied to a different object and failure type. The scope was to create a model based on a PINN, considering only a small dataset of 11 observations. Subsequently, the method of domain generalization and knowledge transfer was applied. In this case, the source and target domains refer to two different areas of the same component. The most important contributions of this part of the research are as follows:

- Creation of a hybrid model based on a physics-informed neural network that is capable of accurately estimating material depletion due to oxidation using the actual operational data captured by the Remote Monitoring & Diagnostics system.
- The results prove that such a PINN based on a recurrent neural network (RNN) with an embedded feedforward neural network (FNN) can be trained effectively and provide consistent outputs even when the training set has only seven or two observations.

- Prediction of the Transition Pieces' wall thickness reduction due to oxidation using the proposed method based on machine learning algorithms in case of the absence of wall thickness measurements in the target domain.
- The results prove that the proposed method of domain generalization and knowledge transfer is robust and versatile, with successful applications to different kinds of gas turbine parts and failure types, even though it is not free of disadvantages.

2.4 Data availability statement

Data supporting this research are confidential and are the proprietary property of Baker Hughes Company. Due to these reasons, the data cannot be shared openly and were anonymized in this dissertation. Nevertheless, the anonymized published data are sufficient to describe the analyzed technical problems and characteristics of input, intermediate and output variables and to understand the context of various decisions taken during the research.

3. Background

3.1 Maintenance of gas turbines

Gas turbines are internal combustion engines that work according to the Brayton thermodynamic cycle. Their efficiency depends on the maximum cycle temperature, the so-called firing temperature, which is the average temperature of exhaust gases at the outlet of high-pressure nozzles. However, the maximum temperature cannot exceed specific values due to limitations imposed by the resistance of available materials. The most severe operating conditions are in the combustion and high-pressure turbine systems. Thus, the endurance of components from these sections determines the frequency of planned maintenance activities.

A well-designed maintenance program reduces operating costs while maintaining the high availability of the equipment. Maintenance of heavy-duty gas turbines is typically based on fixed intervals between planned outages. Such a strategy based on regularly executed inspections is known as preventive maintenance. It is a proactive approach to asset life optimization, which results in fewer maintenance events, increased performance of the fleet and lower and more predictable operating expenses in the long term in comparison with the reactive approach known as corrective maintenance. The scheduled outages vary in terms of work scope, but the most common types of disassembly inspections are as follows:

- combustion inspection,
- liners and transition pieces inspection,
- hot gas path inspection,
- major inspection,
- modular replacement.

Time until the upcoming planned outage may be reduced based on many criteria influencing the equipment's life. The most common hours- and starts-based maintenance factors concern the following:

- fuel type,
- peak load operation,
- operation above 100% of the nominal speed,
- injection of steam or water for emissions reduction or power augmentation,
- operation in suboptimal combustion mode,
- startup type, i.e. standard or peaking-fast,
- startup load (i.e. part, base, or peak load achieved during the startup),
- shutdown type, i.e. normal or emergency.

These factors are applied to calculate so-called factored hours and factored starts based on the actual hours and starts and the unit’s operational profile. Detailed considerations regarding the operation and maintenance of heavy-duty gas turbines are available in [2]. In addition, examples of upgrade packages that may extend the baseline inspection intervals are provided in [3].

Aeroderivative gas turbines are composed of an aircraft engine serving as a gas generator coupled with a non-aviation power turbine. Such units are widely applied for marine propulsion, power generation or mechanical drive applications, e.g. as drivers of compressors or pumps. The maintenance philosophy of such units is different compared to heavy-duty gas

turbines. The disassembly inspections, i.e. a hot section inspection or major inspection, are performed when needed, based on the actual technical condition of internal parts determined during periodical borescope inspections. Some factors may influence the components' life, but maintenance factors are generally not defined and do not impact the frequency of teardown inspections. Additionally, the lightweight construction enables on-site replacement of the entire package within 48 hours. A general description of aeroderivative gas turbines manufactured by General Electric is available in [4].

Nevertheless, such an on-condition concept is also applied to heavy-duty gas turbines. A condition-based maintenance approach aims to make data-driven decisions based on collected and appropriately processed information about the analyzed object's health [5,6]. Components eligible for condition-based maintenance at outage (CBMO) are subjected to detailed visual, dimensional and non-destructive checks during disassembly inspections. The obtained data are compared with the service limits defined based on engineering models and knowledge of subject-matter experts. The parts that satisfy all the acceptance criteria do not require repair and can remain in the gas turbine until the next planned outage. For clarity, the eligibility is evaluated based on predefined policies, an analysis of past repairs and the components' residual life, and an analysis of operational data describing the parts and the unit. The same concept can be extended to repair activities aiming at scrapping only components that do not satisfy the acceptance criteria, i.e. condition-based replacement life (CBRL). Summaries of numerous papers about gas turbine condition-based maintenance and gas path diagnostics are available in [7,8]. Based on them, the following areas need continuous research and further improvements:

- Creation of effective hybrid methods for performance degradation and gas path analysis that can provide both qualitative and quantitative estimates of the fault level.
- Data-efficient, accurate prognostics methods in health monitoring, including multiple and combined failure types scenarios.
- Standardization in terms of the definition of problems, principles and terminologies related to gas path diagnostics.
- Development of monitoring and diagnostics platforms capable of integrating various techniques applied in this area, including machine learning algorithms and providing actionable insights to the end-users.

The research described in this dissertation addresses these needs.

A maintenance strategy designed for the prevention of failures based on continuous monitoring and prognosis of hardware degradation is known as predictive maintenance. A thorough literature review on predictive maintenance in Industry 4.0 is available in [9]. In the context of gas turbines maintenance, such a strategy brings the following advantages over the preventive maintenance approach:

- Continuous monitoring, evaluation of the assets' health status and rapid remediation assured by the Remote Monitoring & Diagnostic services.
- Real-time evaluation and prediction of damage to gas turbine parts and the associated failure probability based on actual operational data.
- Making knowledgeable, data-driven decisions or building autonomous expert systems based on the result given by various models fed by data gathered by the RM&D system.

- Resignation from the fixed intervals between planned inspections in favor of a flexible maintenance strategy with variable intervals between the outages based on the actual and foreseen status of the fleet.
- Optimization of the work scope executed during planned shutdowns focusing mainly on the components indicated by the diagnostics and prognostics models.
- Optimization of capital and operating expenses related to spare parts management.

Finally, a combination of predictive capabilities with descriptive analytics is the main concept behind prescriptive maintenance. It is an extension of the predictive maintenance approach and is known as the highest maturity level of knowledge-based maintenance [10]. This strategy does not focus on predictive performance only but attempts to prescribe the optimal decision based on historical and real-time data [11]. The recommendations drawn from the comprehensive analysis and optimization of maintenance measures may be especially relevant in the case of complex systems composed of many elements. Automated agents could be effective in situations where a variety of actions need to be taken quickly or repeatedly. The paper [12] shows an application of deep reinforcement learning to train a decision-making agent in a simulated industrial plant of 100 machines. The study [13] describes an extreme, entirely data-driven maintenance strategy that learns the maintenance effects based on machine characteristics without any upfront assumptions based on expertise.

Original equipment manufacturers use the available in-house knowledge and build software platforms to optimize the performance of their fleet. The paper [14] presents an example of a physics-informed analytics framework created by Baker Hughes Company for the predictive maintenance of gas turbines. Combining data-driven techniques with physics-based modeling improves the prognostic capabilities and allows for an accurate description of these complex systems. Such high-fidelity virtual representations of physical assets capable of replicating original physical processes with abilities to optimize and predict are known as digital twins. This term was first introduced in 2003, [15], and has gained great popularity in industry and science, especially since 2015 [16]. Creating an accurate replica in the digital space is enabled by continuous data transfer and synchronization between the two counterparts. The benefits of such a virtual representation span the entire life of a product, starting from the design and manufacturing phases through service and disposal [17]. Therefore, a digital twin can enable predictive or prescriptive maintenance strategies for fielded products. The study [18] provides a systematic literature review on digital twins, while [19] reviews their applications for maintenance. Digital twins can be grouped into three classes depending on their magnitude [20]. The smallest replicas represent individual materials, products or units that cannot be further divided. They can be combined into system replicas to represent complex products, production lines or factories. The most extensive replicas are the system of systems, which refers to an enterprise-wide integration of multiple systems or even cross-enterprise exchange of information and optimization of different plants or factories. Such a bottom-up approach provides a comprehensive and holistic overview and allows for optimizing actions considering each piece of equipment. Several examples of unit digital replicas aimed at monitoring and optimizing gas turbine maintenance (e.g. adjusting the axial compressor water wash frequency of aeroderivative engines or estimating the next inspection of fuel nozzles in heavy-duty gas turbines) are presented in [21]. The paper [22] presents a digital twin for rotating machinery designed to quantify and localize rotor unbalance. In [23] a multilayer digital representation of a two-spool high bypass turbofan engine was presented. The top layer is the cycle model adapted to test data of the engine, the intermediate layer models the modules (i.e. the fan and

booster section, high-pressure compressor, combustion chamber and turbine section) using one-dimensional (1D) mean line models to extend the component digital replicas from the third layer, which is based on three-dimensional computational fluid dynamic models. The study [24] presents a digital system of systems for diagnostics, monitoring and prognostics of the health of a synthetically created fleet of 1000 aircraft engines. Of course, such studies are not limited only to scientific research. Integrated solutions for asset performance management and process optimization are commercially offered by the leading manufacturers of gas turbines in oil & gas and energy sectors, including Baker Hughes Company [25], General Electric [26] or Siemens Energy [27].

3.2 Methods for predicting damage to gas turbine components

An analysis and prediction of the life of gas turbine parts starts at the design phase and continues for the entire lifetime of the hardware. Manufacturers of turbomachinery equipment apply standardized processes for predicting damage and evaluating the life of critical parts. For example, the life expectancy of the main combustion and hot gas path (HGP) components is defined by the repair (or inspection) and replacement intervals, the percentage of items expected to pass the inspection without a need to be repaired and the expected percentage of items scrapped during the repair. Typically, the damage observed on those parts is due to one of the following failure types or combinations of them:

- thermo-mechanical, low cycle fatigue;
- oxidation of coatings and base material;
- plastic deformation due to creep;
- wear;
- corrosion due to fuel or air contaminants;
- others, like indentations, damage due to inappropriate disassembly or assembly, or nonconformities related to the manufacturing or repair.

According to [28], more than 60% of damage observed in aircraft engines was due to cracks, while almost 7% was caused by material overheating. It shows the importance of the capability to accurately model and predict the life of gas turbine components subjected to damage due to low cycle fatigue or oxidation. This research is focused on both failure types. Additionally, it provides evidence that the proposed method of domain generalization and knowledge transfer can also be applied to predict the size of damage due to the remaining failure types.

Following the standard process, designing and predicting the life of gas turbine components is an interdisciplinary project which integrates numerous efforts, e.g.:

1) Definition of operating characteristics of loading, steady-state operation and unloading of the gas turbine and estimation of performance parameters during such cycles.
2) Execution of a thermal analysis aimed at the calculation of metal temperatures of the analyzed object.
3) Execution of a structural analysis aimed at the calculation of stresses, strains and fatigue life prediction of each critical location.
4) Selection of materials, coatings, and methods of manufacturing.
5) Definition of limits in terms of damage detectability, serviceability and repairability of the analyzed object.

6) Estimation of costs associated with the manufacturing and repair of the part.

The structural analysis combines the inputs related to the operational parameters, material properties and thermal maps to estimate the extent of damage due to thermo-mechanical fatigue, creep and oxidation. The use of finite element analysis (FEA) in the stress analysis process is a standard approach in industry and academia. In some cases, a two-dimensional (2D) model may be sufficiently accurate and simultaneously reduce the computational cost. For example, in [29] a 2D axisymmetric model for the prediction of metal temperatures of the whole engine during a transient state was proposed. The paper [30] presents a method to predict the burst speed of a turbine disk made of Inconel 718 nickel-based superalloy based on two-dimensional finite element simulations. Nevertheless, for complex geometries, three-dimensional loading, or both, a 2D analysis may be inaccurate. The study [31] describes an approach that combines three-dimensional FEA with multiple attribute decision-making methods to select materials for combustor liners. In [32] the authors integrated a thermal analysis with a 3D finite element model to perform a life assessment on the combustor liner. In [33] the finite element method was applied to calculate the stress concentration factor for a turbine blade of an aircraft engine with an indentation due to foreign object damage (FOD). A hybrid approach that merges FEA with the Monte Carlo method [34] to predict the life of a single-crystal gas turbine blade made of CMSX-4 nickel-based superalloy was presented in [35]. Based on the literature review, it is visible that the most common application of finite element analysis for designing gas turbines is related to the prediction of life due to low cycle fatigue. Such focus is consistent with the statistics of damage observed in commercially operating engines, as already described. LCF life prediction of rotor blades is the most common research area [36,37], including interactions with other failure types, like creep [38] or high cycle fatigue (HCF) and creep [39]. Other examples of using FEA for predicting the LCF life of gas turbine parts are related to a compressor disc [40], turbine section disc [41] or turbine casing [42]. Finally, the finite element method was applied to model thermal barrier coatings (TBCs), starting from the design of TBC thickness distribution on a rotor blade [43] through the prediction of deformation near cooling holes of a blade [44] and the estimation of TBC life on a combustor liner of an aircraft engine [45].

Detecting and identifying cracks in rotating machinery components has been and still is a significant area of basic and applied research. Transverse cracks in rotor shafts play a particularly important role. So many works have been devoted to this research area that it is almost impossible to cite in this dissertation, even those considered the most representative. In addition to various data-driven statistical approaches used to detect this type of cracks, it is common to use methods based on physical and computer models of the tested object, where the cause-and-effect relationships between the crack impact and the dynamic response of the rotor system result from the physical properties of the mechanical vibration phenomenon accompanying the rotational motion of the rotor shaft with at the same time possibly accurate knowledge of the structure properties and numerical parameters of the machine under study. In most cases of these methods, the symptom of a crack occurrence is a change in the natural and forced vibration parameters of such an object, which, due to the generally low sensitivity of this change to the depth of cracks, especially in their initial stages, makes this type of approaches not always completely satisfactory for practical purposes.

A quite original alternative to the above-mentioned commonly used methods of detection, positioning and depth identification of transverse cracks in rotor shafts is an application of the

phenomenon of propagation and reflection of elastic waves in these elements, described, among others, in the paper [46]. In turn, in works [47,48], an identification of transverse cracks in rotor shafts is carried out using a stochastic approach when solving the model-based inverse problem. Namely, by means of an advanced structural mathematical model of the tested rotating system using Monte-Carlo computer simulations, a cloud of the most probable impact scenarios caused by the crack with different depths and positions on the rotor shaft is generated. Then, during on-line monitoring of an operation of the real rotating machine, by the use of the probability density functions, the most similar theoretical scenario of the impact of a given crack in relation to the scenario recorded by measurements is successfully identified.

Continuous diagnostics of gas turbine performance is one of the enablers of predictive maintenance. Gas path analysis (GPA) is the most popular data-driven diagnostic method for monitoring the health of gas turbines. Introduced 50 years ago as described in [49], this model-based method relies on a thermodynamic model of a gas turbine. According to this approach, explicit mathematical and thermodynamic equations determine the relationship between the gas-path measurements (e.g. temperatures, pressures, flow rates, positions and rotational speeds) and the performance parameters, such as pressure ratio, firing temperature, power output and component efficiencies. The method assumes that physical faults, like erosion, foreign object damage of turbine blades, or axial compressor fouling, result in measured parameters deviation. Therefore, it is possible to identify the faulty component by knowing the relationship between the measured variables and the performance parameters. Initially, a linear dependency between the vector of health parameters $\vec{x}$ and the vector of measurements $\vec{z}$ was proposed:

$$\vec{z} = \mathbf{H} \cdot \vec{x} \,, \tag{1}$$

where **H** is known as the influence coefficient matrix (ICM). The deviations of components' performance parameters from their reference values corresponding to the new and clean engine $\Delta\vec{x}$ can be calculated as follows:

$$\Delta\vec{x} = \mathbf{H}^{-1} \cdot \Delta\vec{z} \,, \tag{2}$$

where $\mathbf{H}^{-1}$ is the inverted ICM, known as the fault coefficient matrix (FCM), while $\Delta\vec{z}$ represents the deviations of measurements from their reference values. This simple linear model can detect and isolate faults, handling multiple fault diagnostics. Nevertheless, its reliability depends on the accuracy of the influence coefficient matrix, the measurement system's setup and the sensors' accuracy [50]. Additionally, the inversion of the ICM matrix requires that the number of measurements is higher than or equal to the number of performance parameters to get a unique solution.

In reality, the assumption of linearity is correct only for small deviations from the reference state and becomes invalid as the engine deteriorates and the number of faults increases [51]. Therefore, in [52] the iterative Newton-Raphson algorithm was applied to eliminate the linear method's drawbacks and solve a nonlinear relationship between the health parameters and measurements. Furthermore, to overcome the limitations resulting from the quantity and quality of available measurements, weighted linear regression, linear quadratic estimation (known as Kalman filter), or their variants were applied by OEMs.

Kalman filter is a model-based iterative algorithm to estimate the state of the analyzed system based on uncertain measurements and the system's previous state. This predictor-corrector method was proposed in 1960 [53] and has numerous industrial applications, including real-time control of systems, signal processing, robotics and data fusion techniques or applications for distributed generation and energy storage systems [54]. The original concept applies to problems with linear state transition and linear measurements characterized by a Gaussian distribution. In the prediction step, the algorithm estimates the state along with the associated uncertainties as follows:

$$\hat{\vec{x}}_k = \mathbf{F}_k \hat{\vec{x}}_{k-1} + \mathbf{B}_k \vec{u}_k \,, \tag{3}$$

$$\mathbf{P}_k = \mathbf{F}_k \mathbf{P}_{k-1} \mathbf{F}_k^T + \mathbf{Q}_k \,, \tag{4}$$

where vector $\hat{\vec{x}}_k$ is the estimate of the system state at time k, $\vec{u}$ is the control vector, **F** is the state-transition matrix, **B** is the control matrix, **P** is the covariance matrix of state variables and **Q** is the covariance matrix of the system noise. Once the estimates are available, a measurement is taken:

$$\vec{z}_k = \mathbf{H}_k \vec{x}_k + \vec{v}_k \,, \tag{5}$$

where $\vec{z}$ is the measurement vector, **H** is the model matrix and $\vec{v}$ is the measurement noise vector. In the corrector step, the estimates are updated utilizing a weighted average, with less emphasis on the estimates with higher uncertainty:

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} , \tag{6}$$

$$\hat{\vec{x}}'_k = \hat{\vec{x}}_k + \mathbf{K}_k \left(\vec{z}_k - \mathbf{H}_k \hat{\vec{x}}_k \right), \tag{7}$$

$$\mathbf{P}'_k = \mathbf{P}_k - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_k \,, \tag{8}$$

where **K** is the Kalman gain that determines how much the filter follows the measurements or the model, **R** is the covariance matrix of measurement noise, vector $\hat{\vec{x}}'_k$ represents the corrected state of the system at time k and the matrix $\mathbf{P}'_k$ stores the corrected covariances. Since it is a recursive algorithm, the corrected estimates serve as the input data at the next time step $k + 1$.

An extension of this concept applicable to nonlinear systems, where nonlinear functions describe the state transition and measurement equations, is known as the extended Kalman filter. Following the modified algorithm, these functions are linearized by taking the first-order Taylor approximation around the current mean and covariance. Subsequently, the linear filter can be applied. Nevertheless, if the analyzed system is highly nonlinear or the local linearity assumptions are violated, even the extended Kalman filter may result in suboptimal estimates [55]. Moreover, when applied to model non-Gaussian processes, the algorithm's performance is unsatisfactory [56]. Finally, the Taylor expansion requires deriving the Jacobian matrices, which may be challenging and computationally expensive. Modifications of the Kalman filter algorithm to reduce the limitations and improve its performance are presented in [57].

Along with expanding the installed fleet and gathering experience, the technical condition of gas turbine components can be recorded during the disassembly inspections,

including visual, dimensional and non-destructive checks preceding the repair activities. With the availability of empirical data, statistical methods can be applied. Survival analysis is a part of statistics focusing on the analysis and prediction of time until a failure occurs. One of the most popular survival analysis approaches is the Weibull analysis. Since its introduction in 1951 [58], it has been successfully applied in various industries, like aerospace, automotive, materials and composites research, military, power generation, oil & gas and many others. The main advantages of Weibull analysis are as follows [59]:

- The capability to provide accurate estimates with small datasets composed of two or three failed data points or even without failures if the Weibayes method is applied.
- The output of the analysis is a Weibull plot which gives a concise representation of the inputs and outputs, allowing to assess the correctness and quality of the model and quickly interpret the results.
- The value of the shape parameter (or the slope) β provides information concerning the physics of failure (i.e. $\beta < 1$ describes early failures and so-called infant mortality, $\beta = 1$ is valid for randomly occurring failure types that are independent of time and $\beta > 1$ describes late failures due to wear of the analyzed object).
- The Weibull probability distribution can accurately describe many different types of data distributions.

The cumulative distribution function (CDF) of a three-parameter Weibull distribution has the following form:

$$F(t) = 1 - e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}}, \tag{9}$$

where $F(t)$ is the cumulative probability of failure up to time t, η is the scale parameter or the characteristic life when 63.2% of units are expected to fail and γ is the location parameter that determines the shift of the distribution, i.e. a positive value of γ represents the duration of a failure-free period. In the case of small samples, unreasonable values of the slope, or lack of failed data points, the value of β parameter can be assumed, and the characteristic life can be derived using the maximum likelihood method. The assumption can be based on historical data or models, design practices or literature, physics-based models or material properties or the experts' knowledge. The resultant one-parameter distribution is expected to have a narrower confidence interval than a two-parameter Weibull model based on just a few failures.

In order to take into consideration the impact of operational parameters on the probability of failure, Eq. (9) can be modified as follows:

$$F(t) = 1 - \exp\left[-\left(\frac{t-\gamma}{\exp(\alpha_0 + \alpha_1 z_1 + \cdots + \alpha_K z_K)}\right)^{\beta}\right], \tag{10}$$

where the characteristic life $\eta = \exp(\alpha_0 + \alpha_1 z_1 + \cdots + \alpha_K z_K)$ depends on the covariates z, $k = 1, 2, \ldots, K$ is the ordinal number assigned to each covariate and α denotes the coefficients. Usually, the maximum likelihood method is utilized to estimate the model's parameters. Such an extended model is known as the Weibull proportional hazards model. This technique always yields a reasonable measure of the importance of covariates but does not explain the failure mechanism. Thereby, it can be effectively applied for data exploration [60].

Finally, Eq. (10) can be reformulated to model the extent of damage as a function of time and other independent variables:

$$F(l) = 1 - \exp\left[-\left(\frac{l}{\exp\left(\alpha_0 + \alpha_1 z_1 + \cdots + \alpha_K z_K\right)} \right)^{\beta} \right], \tag{11}$$

where l is a continuous variable that describes the damage size, e.g. the length of cracks or the reduction of wall thickness due to oxidation. Such a model provides a distribution of damage size for each combination of the covariates. With a defined limit, which discriminates failed parts from acceptable ones, this model can be utilized to predict the probability of failure.

The most valuable and informative way to execute Weibull analysis is to split the failure data into distinct classes and prepare separate Weibull plots for each failure type. There are multiple examples of such analysis applied to turbomachinery assets. In [61] a Weibull distribution was fit to field data describing the fatigue life of a high-pressure turbine disc of an aeroderivative gas turbine. Another application of Weibull analysis for predicting the probability of failure due to low cycle fatigue is presented in [62]. The model was prepared based on data describing 520 gas turbine blades subjected to a time-truncated fatigue test. The study [63] describes a statistical analysis of empirical data related to the 1st stage nozzles of MS9001F gas turbines. Fatigue cracks are found in many locations of this stationary component because of high thermal stresses during engine shutdowns and tensile stresses during steady-state operation. The authors created a separate Weibull model describing crack dimensions for each critical location. In [64] Johnson-Weibull analysis of inspection data describing 1312 high-pressure turbine blades of aircraft engines was executed. Three separate models with the same slope were created, i.e. for failures due to thermo-mechanical fatigue, due to oxidation or erosion or both, and a model which pools the remaining failure types, e.g. creep, wear, rubbing and indentations. Another example of Weibull analysis applied to aircraft engines is presented in [65], where the authors used a three-parameter Weibull distribution to predict the reliability of a high-pressure compressor subjected to degradation. The item [66] describes applications of Weibull analysis for predicting failures due to a manufacturing process change and estimating thrust bearing wear rate on an industrial gas turbine. An unusual way to apply the Weibull distribution is presented in [67,68], where it was used to model the performance and predict the reliability of MS5002C and PGT10/2 gas turbines installed in a gas compression station without extracting the individual failure types. In the first paper, the authors proposed a model with a variable slope β, which gradually increases with time. In comparison, the reliability assessment presented in the second paper was based on the modified Weibull distribution [69]. It should be clarified that this dissertation did not cite many older papers about the application of Weibull analysis for predicting the life and damage to gas turbine parts. Nevertheless, it is still a widely applied standard in the reliability analysis of rotating equipment. However, in the era of big data, most state-of-the-art techniques for data-driven inference rely on machine learning algorithms.

3.3 Applications of machine learning for predicting damage to gas turbine components

Machine learning could be defined as a set of methods and algorithms that can learn and extract knowledge from historical data to predict future data. Such algorithms can be utilized to support decisions under uncertainty or even to make data-driven decisions. The theoretical background of many of these algorithms was formulated decades ago, but they gained enormous popularity with the exponential growth of the computational power of modern computers. This technological progress resulted in advanced low-cost sensors installed in each Internet of Things (IoT) object, continuously generating data and exchanging them with a remotely accessed server, commonly called a cloud. Such quantities of data hold many valuable, often unexpected relationships. Specific techniques should be applied to extract as many of them as possible, which is the objective of machine learning algorithms. A comprehensive literature review on machine learning applications within the industry field is provided in [70].

Depending on the learning approach, machine learning is typically classified into the following categories:

- supervised learning,
- unsupervised learning,
- reinforcement learning.

The objective of supervised learning is to learn and approximate the relationship between the input data **X** and the output **Y** based on some observations that represent the actual mapping f:**X**→**Y**:

$$\hat{f} : \mathbf{X} \rightarrow \mathbf{Y} , \tag{12}$$

where $\hat{f}$ is the learned function. If **Y** is a categorical variable, such a problem is a classification task (multiclass classification, if the number of classes is higher than two). If the output variable is continuous, it is known as regression analysis. In a simplified case, the analyzed dataset with observed input-output pairs is split into two subsets. The training set is used to extract the approximated relationship $\hat{f}$ during the learning process aimed at optimizing a predetermined cost function. In contrast, the test set is utilized to assess the accuracy and generalization capabilities of the extracted function $\hat{f}$, when applied to data not used during the training. The number of algorithms and their variations applied in supervised learning constantly increases. Some of the most popular (i.e. multiple linear regression, polynomial regression, kernel-based methods, random forest, AdaBoost and extreme gradient boosting and artificial neural networks) are described in detail in the next chapter concerning predicting the length of fatigue cracks found on the high-pressure nozzles.

The objective of unsupervised learning is to detect and extract patterns from the analyzed set of unlabeled data, i.e. there are no input-output pairs, and the **Y** matrix does not exist. According to [71], typical applications of unsupervised learning are related to clustering, dimensionality reduction, learning sparse graphs and imputation. Clustering aims to detect similar observations in the analyzed dataset and group them into clusters. A comprehensive review of clustering methods is given in [72]. The goal of dimensionality reduction is to project high-dimensional data to a lower-dimensional space to reduce the complexity and extract the most meaningful

information. In [73] the authors explained how feature extraction algorithms work and evaluated their performance on binary and multiclass datasets. They observed higher accuracy of the classifiers after applying dimensionality reduction techniques, especially in the case of nonlinear algorithms. The main applications for learning sparse graphs are related to discovering correlated variables or modeling correlations and making estimations of the joint probability. Finally, in the case of imputation, unsupervised learning is applied to extract the latent variables and complete missing entries of the analyzed matrix.

The objective of reinforcement learning is to create an automated agent or a set of agents capable of making good decisions, considering the current state of the environment and their experience from previous interactions with it. During interactions with the environment, the agent acts as specified by the predefined policy. Modifications of the system's state due to actions taken by the agent are quantified and provided as feedback to the agent. Each good decision of the agent which results in achieving the goal is rewarded. During the training, the agent aims to maximize the cumulative reward. Additionally, the process of exploration and exploitation of the environment can be optimized using supervised or unsupervised machine learning algorithms or both. It is essential to balance these characteristics properly since too much focus on the exploration may result in difficulties in converging to the optimal solution. At the same time, overexploitation may result in too much focus on a local, suboptimal solution. This autonomous, iterative trial-and-error learning concept is general and applied in many areas. A comprehensive literature review of applications of reinforcement learning for predictive maintenance is available in [74].

In addition, a hybrid approach, known as semi-supervised learning, should be mentioned. This branch of machine learning is dedicated to problems where only a small fraction of the analyzed dataset is labeled. During the training, large quantities of unlabeled data are combined with the labeled data under certain assumptions. Researchers have found that such an approach improves the results compared to the training based only on the small labeled dataset. The paper [75] provides an overview of semi-supervised learning algorithms, focusing mainly on classification problems.

The second part of this section presents the latest applications of machine learning methods for the predictive maintenance of turbomachinery assets. In [76] the authors applied the extreme gradient boosting algorithm (XGBoost) [77] to predict the damage of high-pressure turbine blades due to creep based on engine operational data. A hybrid approach was proposed that combines a physics-based model of the engine, a finite element analysis of the blade aimed at evaluating stress and strain fields and identifying the critical zones as well as a nonlinear regression model estimating the residual useful life. Such a surrogate model can be expanded to consider the effects of cyclic loads or vibrations, but they were not analyzed in that study. Additionally, the analysis was limited to steady-state conditions only. Another hybrid approach to predict the residual useful life of high-pressure gas turbine blades due to creep is described in [78]. The proposed framework comprises a two-shaft aircraft engine performance model, analytical models for calculating stresses and temperatures in critical locations and a prediction module based on the Larson-Miller parameter. Actual operational data describing 300 flight missions were processed through the framework and used to predict the creep rupture life based on gas temperatures at the exhaust section by means of a random forest algorithm. It was observed that the inclusion of categorical features describing environmental conditions increased the prediction accuracy significantly. Nevertheless, further improvements could be

achieved by utilizing computational fluid dynamics and finite element methods to precisely compute the stress and temperature fields. In [79] a convolutional neural network was trained based on blade tip timing (BTT) measurement data to predict the presence, position and depth of gas turbine blade cracks. BBT method is used for non-contact measuring and characterizing rotor blade vibration. A numerical model was used to generate the data utilized to train the neural network. Subsequently, the data-driven model was validated on measurement data recorded during a test of real cracked blades. The blades were machined accordingly on the root, mid or top span to simulate the presence of cracks. The proposed neural network provided more accurate predictions than the models based on other machine learning algorithms, i.e. support vector machine, extreme gradient boosting, K-nearest neighbor and random forest. However, the test and the validation would be more meaningful if real gas turbine blades were analyzed. In [80] a hybrid model to detect lean blowout in Dry Low NOx combustors was presented. It is composed of a physics-based model for detecting power output drop coupled with a binary, data-driven classifier to detect the loss of flame based on measurements of gas temperatures at the exhaust section. The authors applied dimensionality reduction techniques followed by two types of classifiers, employing linear regression and decision tree algorithms. The latter model performed better concerning all the analyzed metrics, i.e. accuracy, precision, recall and the F-score. The final model was deployed in Baker Hughes' monitoring and diagnostics platform. In [81] stacked denoising autoencoders [82] were applied for feature engineering, and the extreme machine learning classifier [83] was used to detect combustor anomalies in heavy-duty gas turbines manufactured by General Electric. The authors observed that the classifications based on extracted features are more accurate than those based on raw operational data. Therefore, an unsupervised machine learning method was applied to generate the features in an automated manner. As a result, the classifier's accuracy improved compared to the model based on knowledge-driven, handcrafted features. Another industrial application of machine learning methods to classify the state of a combustion system is described in [84]. Logistic regression and artificial neural networks were applied to detect anomalies and split them into three classes, i.e. sensor fault, cold spot, or hot spot. The latter two classes correspond to uneven distribution of gas temperatures at the exhaust section, with at least two readings far from the mean temperature. The data-driven algorithms achieved better performance compared to the physics-based model. In both papers, the authors emphasized that expanding the features' space should improve the models' capabilities. In [85] exhaust thermocouple readings were used to detect faults in combustor chambers by means of a model based on a convolutional neural network. It was trained based on abundant fault data related to Solar Turbines' Taurus 70 industrial gas turbine. Then, transfer learning was applied to retrain the last layers of the model in the target domain related to the Titan 130 units, where the availability of fault data was limited. The study successfully applies transfer learning and knowledge sharing between two similar domains. The performance of the retrained neural network was better compared to the models trained in the target domain only or trained based on a mixed dataset. However, the paper does not address the issue of fault isolation and does not explain what kind of faults were analyzed, which is a drawback of this study. In [86] a system based on a convolutional neural network for the real-time detection and prioritization of aircraft engine anomalies was proposed. Compared to recurrent neural networks, this class of ANNs requires less computational power during the forward pass. Thus, the proposed solution can run on low-power devices, and it was deployed on Rolls Royce Pearl 15 engines. An application of deep recurrent neural networks to evaluate and predict the residual useful life of a turbofan engine, expressed by a health indicator, is presented in [87]. The authors analyzed two datasets published by the National Aeronautics

and Space Administration (NASA) describing the degradation of aircraft engines. This regression analysis aimed at predicting the health status based on multivariate time series describing operational settings and sensor readings. The proposed model based on a deep RNN returned more accurate estimates compared to a nonlinear autoregressive neural network with exogenous inputs (NARX), multilayer perceptron (MLP) and cascade forward neural network. Such an approach may be applied for high-level fleet monitoring. However, it should be combined with a fault identification module to provide more specific suggestions about the required maintenance activities. In [88] the authors applied a convolutional neural network with long short-term memory (LSTM) layers to analyze the same dataset and predict the residual useful life of aircraft engines. In the proposed approach, relationships between the sensors were represented as flow charts and transformed into embedding vectors guiding the construction of the deep learning model and the arrangement of its input space. Similar embedding vectors were merged into a cluster representing the physical significance of the measuring instruments. Combining the knowledge about sensor relationships with a deep neural network was effective. The prediction accuracy was similar to or even better compared to other deep learning models trained on that dataset and reported in the literature. In [89] five different classification methods were applied for detecting and isolating faults due to the fouling and erosion of a three-shaft gas turbine. Single and multiple simultaneous faults were analyzed during the study. The faults' presence was simulated using an engine performance model, and the input data were generated, which is a shortcoming of this analysis. Models based on the K-nearest neighbor algorithm and an artificial neural network obtained the highest faults detection and isolation accuracy. Another study aimed at predicting aircraft engine degradation due to compressor fouling and turbine erosion is described in [90]. The goal was to predict the engine's health status expressed by the turbine output temperature up to 12 flights ahead using a nonlinear autoregressive neural network with exogenous inputs and an Elman neural network. The most significant limitation of this study, also mentioned by the authors, is related to the use of synthetic data for training and validation purposes. Nevertheless, multiple network structures were created to analyze various degradation scenarios. Regardless of the simulated setup, the Elman neural networks were more accurate and required less time to execute the training. In [91] a multilayer perceptron, gradient boosting regression, decision tree and random forest algorithms were applied to model the erosion of thermal barrier coatings, i.e. air plasma spray and electron beam physical vapor deposition TBCs, of gas turbine hot gas path components at elevated temperatures. The authors extracted the experimental data utilized to create the models from the literature, creating a dataset with 245 samples. The erosion rate, expressed in g/kg, was the output being modeled in the function of five input variables. The prediction accuracy was similar regardless of the applied algorithm and considered satisfactory. Additionally, it was observed that training the models on a merged dataset related to both coating types had a detrimental effect on the accuracy. In [92] the authors applied several popular algorithms (e.g. support vector regression (SVR), random forest, extreme gradient boosting, or feedforward neural networks) to predict the length of fatigue cracks found on high-pressure nozzles of heavy-duty gas turbines. They got accurate estimates even if the dataset was composed of only 31 observations. However, it was underlined that the responses from those data-driven models were inconsistent with the underlying physical laws and that the extrapolation capabilities were unsatisfactory. A comprehensive comparison of various machine learning methods (e.g. linear regression with its variants, SVR, random forest, gradient boosting decision trees or artificial neural networks) applied to predict the high-pressure compressor recoup pressure of aeroderivative gas turbines is presented in [93]. This parameter provides information about

axial loads acting on the engine's thrust bearing. It is not directly measured, but the capability to estimate it based on other operational parameters was proved in that study. The most accurate predictions were obtained using the tree-based algorithms, i.e. random forests and gradient boosting regression. However, despite a large number of operational data, the generalization capabilities of those data-driven models were not understood completely. Additionally, the authors admitted that a different way to process the data, treating them as sequences of discrete-time data, should raise the accuracy. A comparison of several machine learning methods with a particular emphasis on deep neural networks applied to predict the condition of a turbofan engine component is presented in [94]. Actual operational data, environmental data and outputs of thermodynamic models related to over 150 units under monitoring were used to train and test data-driven models. A 10-grade scale was used to evaluate the technical condition of the analyzed hot gas path part due to the analyzed failure type and reflect the life consumption. The models were built to predict these ranks based on the available input data. The best results on the validation subset were obtained using a multilayer perceptron and an ensemble of such neural networks. Depending on the algorithm, the authors applied some actions to prevent overfitting, but they were not always successful. The models based on extreme gradient boosting, extremely randomized trees and support vector regression had unsatisfactory generalization capabilities. In [95] three different classification methods (i.e. decision tree, random forest and the C5.0 decision tree algorithm) were utilized to determine and predict the state of gas control valves of MS7001FA gas turbines. Actual operational data related to more than 500 units were used to train these binary classifiers. The best results in terms of precision, recall and the F-score were obtained with the random forest classifier. Nevertheless, the C5.0 model was deployed in a production environment due to better scalability and interpretability. This model is capable of raising an alert even four months before an emergency shutdown caused by the valve. However, the classifier's performance and generalization capabilities were not evaluated against other than F-class gas turbines.

The most visible limitations of the reviewed papers about the application of machine learning algorithms for predicting damage to gas turbine components could be summarized as follows:

- Numerous studies are based on data generated by means of thermodynamic models. However, even though noise is added to the synthetic time series, such datasets may not reflect the complexity of actual operational data and various relationships between measured and calculated parameters. Thus, the simulated scenarios may be an oversimplified representation of the reality.
- Most papers aim to detect and predict anomalies based on a time series analysis. Classifiers are created to estimate an entire gas turbine's health status, but faults are often not isolated, and the root cause is not highlighted. Such an output is too general to plan precisely the required maintenance activities.
- Lack of studies dedicated only to a specific component and failure type, focusing on the accumulation of damage during part lifetime due to that failure type and evaluation of residual useful life. With the exception of rotating blades, the remaining life-limited parts from combustion and turbine sections are rarely studied.
- The number of papers related to heavy-duty gas turbines is significantly smaller than the number of studies about aircraft engines. The likely reason for this is the availability of operational data on industrial gas turbines.

- Despite the good performance of predictive models based on machine learning methods, most of the reviewed solutions have not been implemented in industrial applications.

The research described in this dissertation addresses all these limitations.

3.4 Physics-informed neural networks

Numerous successful applications of machine learning algorithms for the predictive maintenance of gas turbines were referenced in the previous section. Such methods are applied to build data-driven models capable of approximating complex relationships between the input data **X** and the output **Y**. In several cases, more complex hybrid structures were proposed using physics-based models to prepare input data for a data-driven predictive model. Nevertheless, typically the latter does not take directly into account the underlying physical laws or prior knowledge, including empirical rules and experts' knowledge about the analyzed phenomenon. It is a significant shortcoming because such validated information could be used during the training to process data in a specific manner, limit the space of allowable solutions, guide the algorithm in the right direction during the cost function optimization and improve the generalization capabilities of the model thanks to the physics-based backbone. The advantages of including additional knowledge will be particularly visible when analyzing small datasets. Despite the digital transformation arising from the fourth industrial revolution, the small data regime is still a common problem, especially in the case of complex problems, where data gathering requires much time and cost or when the analyzed events rarely occur.

Artificial neural networks proved their effectiveness in various disciplines, but typically large quantities of data are required to achieve satisfactory accuracy, especially in the case of deep learning. However, ANNs can be successfully applied in the small data regime if they are appropriately combined with physical laws that describe the analyzed phenomenon. Neural networks with embedded equations governing the studied problem, like partial differential equations (PDEs), are known as physics-informed neural networks. This class of machine learning algorithms was introduced in 2017 [96] and applied to approximate solutions of nonlinear PDEs of the following form:

$$u_t + N[u] = 0\,, \tag{13}$$

where $u(t,x)$ is the state of the system, N is a nonlinear differential operator and the subscripts symbolize partial differentiation in time t or space x. Subsequently, $f(t,x)$ given as the left side of Eq. (13) is defined as follows:

$$f := u_t + N[u]\,. \tag{14}$$

Since artificial neural networks are known as universal function approximators [97], the latent solution $u(t,x)$ can be estimated by means of a neural network. The learning objective is to minimize the cost function that could be represented by the mean squared error *MSE*:

$$MSE = MSE_u + MSE_f\,, \tag{15}$$

where MSE_u is the error related to the initial and boundary training data, while MSE_f relates to the collocation points sampled from the solution domain. In [96] this approach was applied to

get data-driven solutions of the one-dimensional Burgers' equation, the one-dimensional Schrödinger equation and the Allen-Cahn equation. In the second part of the study [98], the authors focused on the data-driven discovery of PDEs given some scattered measurement data $u(t,x)$ across the analyzed domain. Solving these inverse problems was aimed at learning the parameters of partial differential equations and predicting the solution in the entire spatio-temporal domain. The one-dimensional Burgers' and Korteweg-de Vries equations and two-dimensional Navier-Stokes equations were presented as the examples. The obtained predictions were very accurate for both the forward and the inverse problems. The success of the authors of [96,98] can be represented by the constantly growing number of citations of these papers and the number of studies and successful applications of physics-informed neural networks in many disciplines. It can be thought as a mesh-free approach for finding PDE solutions by changing it to an optimization problem. However, from another perspective, it can be considered as an unsupervised learning algorithm that assures consistency with the underlying physical laws described by PDEs when labeled data are unavailable, or their availability is limited. Such models should have good generalization capabilities and return credible estimates in other domains governed by the same equations. Therefore, validated physics-informed kernels could serve as a medium to transfer knowledge between two domains. Nevertheless, this class of algorithms has certain limitations. A physics-informed neural network does not guarantee a unique solution since the analyzed task is converted to a non-convex optimization problem. The results depend on the neural network's structure and hyperparameters. Thus, choosing the optimal parameters and training the neural network may be more time-consuming compared to other methods. Additionally, the effects of weight initialization, loss estimation and the assignment of weights to the loss terms should also be considered. A comprehensive literature review on physics-informed neural networks is provided in [99].

The majority of PINNs described in the literature were based on a multilayer perceptron. However, applications of multiple feedforward neural networks [100], single hidden layer neural networks [101], convolutional neural networks [102] and recurrent neural networks [103] can be found in the literature. During the training, automatic differentiation is applied to adjust the neural network weights. This technique applies the chain rule to calculate derivatives iterating through layers of the neural network from the last layer to the input layer. The most popular deep learning libraries, like TensorFlow [104] or PyTorch [105], support the automatic calculation of gradients for any computational graph representing a neural network. Therefore, the default practice is to apply automatic differentiation during the learning process of artificial neural networks. Initial and boundary constraints can be implemented in a soft manner with an additional loss term defined on collocation points sampled from the boundaries. However, this approach does not guarantee that the boundary conditions are satisfied accurately, and it requires assigning a proper weight to the additional loss component. In the alternative, i.e. the so-called hard approach, a specific component is embedded into the custom structure of the neural network, which assures that the boundary condition is satisfied [106]. The approach to selecting collocation points and enforcing the initial and boundary constraints may affect the solution's accuracy and generalization error. Besides PDEs, physics-informed neural networks were applied to approximate solutions of ordinary differential equations [107], integro-differential equations [108] and stochastic differential equations [109]. The popularity of PINNs is also visible in the number of dedicated software packages and libraries created during recent years. They were summarized and concisely described in [99].

The second part of this section presents examples of industrial applications of physics-informed neural networks that could be adopted for the predictive maintenance of gas turbines. In [103] the Euler method was applied to approximate the solution of an ordinary differential equation describing the fatigue crack growth. The authors created a recurrent neural network with an embedded multilayer perceptron estimating the stress intensity range based on available values of cyclic stresses. In this setup, the MLP can compensate for some inaccuracies related to the values of stresses resulting from a simplified modeling approach or assumptions made. The crack length increments were evaluated after each cycle based on the stress intensity range estimates, applying the empirical equation. The effectiveness of this framework was shown in the paper, but the analyzed dataset was relatively large and synthetically generated. In [107] the same authors expanded the physics-informed neural network cell with an additional MLP to account for the impact of corrosion on the fatigue crack growth rate. The objective of this data-driven element was to capture the contribution of corrosion to the damage size and compensate for the fact that corrosion is not modeled using any empirical or theoretical equation. Such a structure allows for isolating failure types and quantifying their impact on the damage accumulation rate. The addition of the data-driven element significantly increased the prediction accuracy if the operation in a corrosive environment was simulated. In [101] the authors applied that hybrid framework based on a recurrent neural network to create a nonlinear damage accumulation model of a wind turbine main bearing. In this case, a reduced-order physics-based model in the form of an ordinal differential equation was embedded in the RNN's cell to estimate the damage increments due to fatigue. A multilayer perceptron was applied to model grease degradation, including the impact of the degraded properties on the bearing's fatigue life directly. The data-driven component was designed to model the contributions that are difficult to describe by means of simple mathematical equations. However, it was integrated with a physics-based component into the hybrid model. Additionally, purely data-driven models with a long short-term memory cell were prepared, but they performed poorly in that simulated scenario compared to the hybrid model. In [110] the authors applied a physics-informed recurrent neural network to predict the propagation of fatigue cracks observed on high-pressure nozzles made of FSX-414 cobalt-based superalloy of an industrial gas turbine manufactured by Baker Hughes Company. A feedforward neural network embedded into the RNN was utilized to estimate the stress intensity factor based on actual operational data and selected outcomes of a numerical simulation. An attempt was made to capture the contribution of creep on the crack propagation rate. The analyzed sample comprised only 13 crack length measurements, with the operational data available for just 54% of startup-shutdown cycles. Regardless of the limited dataset, three hybrid models sharing the same architecture were trained respectively based on ten, two and only one observation. A custom cost function was created, which changes dynamically during the training process in how the physical and empirical terms contribute to the final cost value. The fully-specified hybrid models had good generalization capabilities and satisfactory accuracy, also outside the original training domain. Moreover, the authors proposed a novel method of single-source domain generalization and cross-domain knowledge transfer in regression analysis leveraging physics-informed neural networks. The method's purpose is to effectively use a cumulative damage model based on a physics-informed neural network, i.e. the so-called source hybrid model, which was trained in a domain, where measurement data on damage size are available, in another domain, where such data are unavailable. This method was successfully applied, providing accurate damage size predictions in a simulated scenario, where damage size measurements were unavailable. In [111] a PINN based on a multilayer perceptron with two hidden layers was utilized to predict the fatigue life of AMS 5707 nickel-

based superalloy and 2024-T4 aluminum alloy in the small data regime. The maximum number of training observations was equal to 31, i.e. experimental data extracted from the literature were used. Furthermore, physical models describing the analyzed phenomenon were embedded into the neural network as the activation functions of some hidden neurons, referred to as physical neurons. These nodes yielded low-fidelity results, which were subsequently adjusted by the outputs of the remaining hidden neurons with the sigmoid activation function, using the training data that relate the fatigue life with the stress amplitude and ratio. As a result, the PINN's prediction accuracy was higher compared to several data-driven models trained on the same input data. A similar problem related to the fatigue life prediction of different metallic materials using a physics-informed neural network is described in [112]. The objective was to approximate a relationship between the stress amplitude and fatigue life expressed by the mean and standard deviation parameters, assuming that a normal distribution could describe the log-transformed fatigue life. The PINN was based on a feedforward neural network with one hidden layer and an output layer with two nodes. Thus, during the learning process, the parameters of the continuous probability distribution are extracted from available experimental data while satisfying several physics-based constraints reflected by additional loss terms in the cost function. However, the simultaneous minimization of the physics-related loss terms was observed to be difficult. Therefore, modifying the proposed architecture by directly embedding the constraints using the hard approach may help to overcome this problem. In [113] a PINN based on a feedforward neural network with five hidden layers was applied to predict the creep-fatigue life of 316 austenitic stainless steel at high temperatures. The authors extracted 145 datasets of strain-controlled creep-fatigue data from the literature, including loading characteristics and chemical composition. Feature engineering was performed to extract more significant variables based on the input data while using knowledge about the analyzed phenomenon. Physics-based constraints were enforced in a soft manner by expanding the cost function with two additional loss terms limiting the estimates to the range from 0 to 10^5. As a result, the proposed PINN's prediction accuracy was higher compared to the purely data-driven models based on support vector regression or random forest algorithms, or a multilayer perceptron with the same structure as the PINN. In [114] six different problems of the growth and propagation of cracks in brittle materials were solved using a physics-informed neural network based on a multilayer perceptron with three or four hidden layers. Instead of minimizing the residuals of the underlying differential equations, the authors proposed minimizing the system's total variational energy. This approach requires less computational power since lower-order derivatives are used. Furthermore, the boundary conditions were enforced in a hard manner by modifying the neural network's output. Consequently, no additional terms related to the boundary losses are present in the cost function. Additionally, transfer learning was applied during the retraining of the physics-informed neural network after each load step. The estimates of the hybrid models were satisfactory and comparable with those available in the literature. Finally, the authors concluded that such an approach could be used to create efficient low-fidelity surrogate models for high-fidelity numerical simulations. In [115] given some operational parameters' measurements, a PINN was built to monitor diesel engine dynamics. Empirical, engine-specific first-order differential equations describe the output variables representing the dynamics. Solutions to these equations were approximated using feedforward neural networks composed of two or three hidden layers. The cost function is a weighted sum of multiple loss terms using the self-adaptive weights [116]. The neural networks were pre-trained based on synthetic data and embedded into the physics-informed neural network to mimic the empirical equations. The results obtained by means of the model

were satisfactory, even when Gaussian noise was added to the measurement data. Using these estimates to monitor and predict the engines' health status could significantly improve this study. Applications of physics-informed neural networks to solve several heat transfer problems are presented in [117]. The objective was to predict temperature, velocity and pressure fields in the case of forced and mixed convection with thermal boundary conditions not fully known and for the two-phase Stefan problem with a moving phase-change boundary. Depending on the analyzed problem, the PINNs were based on fully-connected neural networks composed of ten or five hidden layers. It was assumed that some measurement data were available. Their values were obtained by means of numerical simulations. Using the conservation laws embedded in the neural network structure and the limited number of measurements, the proposed PINNs could accurately estimate the output variables on the boundaries and in the entire domain. In [118] the capabilities of the NVIDIA Modulus framework [119], previously known as SimNet, were shown for solving conjugate heat transfer problems related to integrated circuits' heat sinks. The PINNs were based on deep neural networks composed of six layers. No training data were used in the analyzed examples. The first problem aimed at predicting pressure, velocity and temperature fields with the Fourier feature network [120] and its modification. In the second problem, a fully-connected neural network was used to estimate the same outputs and optimize the heat sink design under a specific constraint, considering more than half a million heat sink configurations. Two separate neural networks estimated the temperature and flow fields in both cases. The obtained results were aligned with the results of numerical simulations. Furthermore, the application of PINN significantly reduced the time required for the design optimization. In [121] a physics-informed neural network was applied to predict the fatigue life of additively manufactured samples. The samples made of AlSi10Mg aluminum alloy were produced by selective laser melting. Experimental data related to 12 samples were used to train and validate the model. The PINN was based on a multilayer perceptron with two hidden layers. The prior knowledge embedded in the model was based on semi-empirical laws of linear elastic fracture mechanics. The hybrid model's prediction capabilities were significantly better compared to an equivalent MLP without the physics-based constraints, i.e. the coefficient of determination was around 83% higher. Concerning gas turbines, future research should focus on applying this framework for fatigue life modeling of nickel- and cobalt-based alloys, which are used to manufacture combustion and hot gas path components. In [122] a PINN based on a multilayer perceptron with five hidden layers was used to predict the melt pool dynamics and temperature field during selective laser melting of Inconel 625 nickel-based superalloy. Due to the limited availability of experimental data, a finite element analysis was used to generate additional data to train the neural network. The conservation laws of energy, mass and momentum in the form of partial differential equations were integrated into the model. A Heaviside function was used to enforce the Dirichlet boundary conditions in a hard manner, resulting in faster optimization compared to the soft approach. The proposed hybrid model accurately predicted the pressure, temperature and velocity fields regardless of the number of experimental data. A similar study about the application of PINN for predicting a three-dimensional temperature field during the laser metal deposition process, but in the absence of labeled data, is presented in [123]. Such studies are of great importance, considering the growing number of additively manufactured gas turbine components. In [124] physics-informed neural networks were applied to model one- and two-dimensional high-speed aerodynamic flows described by the Euler equations, which consist of the equations for the conservation of energy, mass and momentum for compressible flow. In the forward problem, the boundary constraints were enforced in a soft manner through additional loss terms in the

cost function. Two PINNs based on multilayer perceptrons with seven and four hidden layers were created to predict the density, pressure and velocity fields. The training points were limited to the initial and boundary conditions only. In the inverse problem, given the density gradient and some pressure measurements in the analyzed spatio-temporal domain, the hybrid model estimated the density, pressure and velocity fields with no initial or boundary conditions. It is explained in the paper that the position of points for which the measurements are available has an impact on the learning process. The advantages of such hybrid models were especially visible for the inverse problems, where PINNs could accurately approximate based on a few observations only. However, the hybrid models were less accurate for the forward problems compared to numerical simulations. In [125] a physics-informed neural network was applied to approximate the solution of the Reynolds partial differential equation and estimate the pressure field and gas film thickness of a gas bearing. The proposed PINN was based on a fully-connected neural network with ten linear and nonlinear layers. The boundary constraints were enforced in a soft manner by expanding the cost function with an additional loss term. Three different scenarios were considered in the study, dependent on the availability of measurement data, namely unsupervised, semi-supervised and supervised learning. Regardless of the learning approach, estimates from the neural networks were comparable with the results of numerical simulation. In case of the purely data-driven supervised learning approach, the conservation equations were not satisfied. Finally, certain limitations of PINNs are described in [126]. The authors applied a multilayer perceptron with four hidden layers to approximate solutions of linear and nonlinear partial differential equations of one-dimensional convection and reaction-diffusion problems. In this unsupervised learning setup, the collocation points were randomly sampled while the boundary constraints were enforced using the soft approach. It was observed that increasing or adding physics-based regularization constraints makes the optimization problem challenging and may result in learning failure. In order to overcome these difficulties, the authors proposed pre-training of the neural network for proper initialization of the weights and predicting in a step-by-step manner instead of predicting all timesteps at once.

3.5 Transferring knowledge in machine learning projects

The effectiveness of physics-informed neural networks in combining machine learning methods with prior knowledge about the analyzed phenomenon was described in the previous section. The embedded equations regularize the neural network. Thus, the resultant models have good generalization capabilities, even if the number of training observations is limited or they are unavailable. Thereby, PINNs can be considered as a method to extract knowledge in one domain, transfer it, and apply it in another domain for a problem governed by the same set of equations. However, the mathematical description can be imperfect due to the complexity of the problem or assumptions taken during the analysis. In such situations, the model may predict accurately in the source domain in which it was trained, but the generalization capabilities and accuracy in other domains will be worse. Transfer learning methods can be applied to address these concerns and effectively share knowledge between similar domains.

Transfer learning aims at improving the learning of the target predictive function $\hat{f}_T(\cdot)$ in the target domain D_T leveraging knowledge about the source domain D_S and the task in the source domain T_S, given that $D_S \neq D_T$ or $T_S \neq T_T$. Here, the notion of "leveraging" can be defined as the action of using something already available in order to achieve something new

or better and maximize advantages. A domain $D = \{\mathcal{X}, P(\mathbf{X})\}$ consists of the *n*-dimensional feature space $\mathcal{X}$ and the probability distribution $P(\mathbf{X})$, where $\mathbf{X} \in \mathcal{X}$ [127]. A task $T = \{\mathcal{Y}, \hat{f}(\cdot)\}$ consists of the *n*-dimensional output space $\mathcal{Y}$ and the predictive function $\hat{f}(\cdot)$, which relates the inputs **X** with the outputs **Y**, where $\mathbf{Y} \in \mathcal{Y}$. $\hat{f}(\cdot)$ can be seen as the conditional probability distribution $P(\mathbf{Y} \mid \mathbf{X})$. Following these definitions, the difference between the two domains relates either to the selected features, i.e. $\mathcal{X}_S \neq \mathcal{X}_T$, or to their marginal distributions, i.e. $P(\mathbf{X}_S) \neq P(\mathbf{X}_T)$. While the difference between the tasks relates either to the output spaces (i.e. $\mathcal{Y}_S \neq \mathcal{Y}_T$, for example, due to different dimensions) or to the conditional probability distributions (i.e. $P(\mathbf{Y}_S \mid \mathbf{X}_S) \neq P(\mathbf{Y}_T \mid \mathbf{X}_T)$, for example, due to different proportions of classes). Despite differences, some similarities between the domains or tasks must be present and measured to apply transfer learning.

Depending on the availability of labeled data in the source and target domains, the strategies of transfer learning can be categorized as follows:

- inductive transfer learning,
- transductive transfer learning,
- unsupervised transfer learning.

In the first scenario, the learning tasks are different between the domains, but labeled data in the target domain are available and utilized to induce the target predictive function $\hat{f}_T(\cdot)$. If labeled data are available in the source domain, this setup is similar to multi-task learning when the source and target tasks are learned simultaneously. If labeled data are unavailable in the source domain, this setup is similar to self-taught learning [128]. In the second scenario, the learning tasks are the same, but the domains are different, with labeled data available only in the source domain. However, some unlabeled data are required in the target domain at training time to find the marginal probability distribution $P(\mathbf{X}_T)$. If the feature spaces are the same, i.e. $\mathcal{X}_S = \mathcal{X}_T$, then the difference between the domains is due to the different marginal probability distributions, i.e. $P(\mathbf{X}_S) \neq P(\mathbf{X}_T)$. For such problems, domain adaptation methods are applied, which are described in more detail later in this section. The last scenario is similar to the first one, except that no labeled data are available for training in both domains. Applications of unsupervised transfer learning are related to clustering or dimensionality reduction.

According to [129], studies about the application of transfer learning for machinery diagnostics and prognostics can be grouped as follows:

1) Transfer of knowledge between domains related to different operating conditions of the same unit, which can be referred to as cross-domain transfer.
2) Transfer of knowledge between different parts with unique characteristics, which can be seen as a more challenging type of cross-domain transfer.
3) Transfer of knowledge from a simulated environment to real objects and installations. In this scenario, the difficulty results from an incomplete description of reality in the simulated environment.

The most popular approach of cross-domain transfer leverages models pre-trained in the source domain, which are subsequently fine-tuned in the target domain. This method is most widely used for deep neural networks since their layered structure allows for simple implementation. For example, weights of a pre-trained network can be utilized to set the starting point for the optimization in the target domain [130]. Alternatively, the first layers with the corresponding weights can be fixed, while the training process in the target domain could be aimed at fine-tuning the last layers only [131,132]. Thus, knowledge from the source domain can be leveraged for the weights initialization, reduction of the number of trainable parameters and time required for the training and creation of effective models even if the number of labeled data in the target domain is limited. Nevertheless, retraining the last layers may be insufficient if the difference between the marginal probability distributions is significant. In such situations, the source or the target features can be transformed to match each other better and reduce the difference between the distributions. Distribution adaptation described in [133] and feature subspace learning are the most popular methods for shallow-structured architectures. The first method aims at minimizing the distance between the marginal probability distributions, conditional probability distributions, or both by some transformations, e.g. applying a kernel function. Maximum mean discrepancy (MMD) defined in [134] is the most widely used statistical test to determine the distance between two probability distributions in a universal reproducing kernel Hilbert space. The second method assumes that the source and target data are similarly distributed in some subspace. Therefore, data from both domains can be projected into a shared subspace to align their statistical distributions, as considered in [135–137]. Alternatively, manifold learning [138] can be used, which is a nonlinear dimensionality reduction technique. This method generates a set of intermediate data representations and identifies the shortest path that connects both subspaces. These feature-matching methods applied to shallow-structured architectures can be leveraged in the case of deep neural networks. Such neural networks comprise additional adaptation layers to reduce the difference between the source and target data distributions and transfer knowledge between the domains. Thus, the subdomain in which the model can be effectively utilized is enlarged.

As previously written, deep neural networks require large quantities of data to achieve satisfactory accuracy, which may be unavailable in practical applications. Additionally, the assumption of independent and identically distributed training and testing data is considered for most machine learning algorithms. In reality, this assumption can be false, for example, due to different operating conditions or characteristics of the target object. Domain adaptation tackles the problem of the data distribution shift between the source and target domains and reduces limitations resulting from the assumption of independent and identically distributed datasets. It aims at maximizing the performance in the target domain by leveraging knowledge from the source domain and reducing the distribution differences between the domains. However, most domain adaptation methods assume access to either sparsely labeled or unlabeled target data. Thus, the knowledge of the marginal probability distribution of target data $P(\mathbf{X}_T)$ is used for the model adaptation. According to [139], the major advantages resulting from the use of domain adaptation methods are related to the high effectiveness of training and accuracy obtained with small datasets, increased generalization capabilities of models and low computational costs. Nevertheless, a negative transfer may occur if the domains differ significantly or the selected domain adaptation method is insufficient. Thus, leveraging knowledge from the source domain may undesirably reduce the model's performance in the target domain. A survey on negative transfer is available in [140]. Domain adaptation methods

can be divided into adversarial-based methods, reconstruction-based methods and discrepancy-based methods. The latter one, aimed at minimizing the distance between the distributions of source and target data, was mentioned in the case of shallow-structured architectures. Nevertheless, considering deep neural networks, the first two methods gained popularity and are described in more detail below.

Adversarial learning is utilized to extract domain-invariant features. It is inspired by generative adversarial networks (GANs) proposed in [141], which attracted many researchers in recent years, as follows from [142]. GAN comprises a generator and a discriminator, i.e. a generative and discriminative network. The former aims to learn the input data distribution and generates fake samples that are as close as possible to the real data. The latter is a classifier that tries to distinguish whether a sample is real or created by the generator. The training objectives are adversarial since the two neural networks have competing goals and try to confuse each other. Such architectures are mainly utilized to create synthetic data that plausibly come from the distribution of real input data, with numerous applications in image generation, processing and reconstruction. However, the adversarial structure can be used for transfer learning. In such a setup, the generator serves as a feature extractor, while the discriminator classifies whether the extracted features come from the source or target domain. The objective of the generative network is to learn such features which the discriminative network cannot distinguish if they come from one domain or another. Therefore, the loss function l may have the following form:

$$l\left(D_S, D_T\right) = l_S\left(D_S\right) + \alpha l_D\left(D_S, D_T\right), \tag{16}$$

where l_S is the loss related to the source model's performance on data from the source domain, l_D is the discriminative loss reflecting the discriminator's capabilities to distinguish the domains and α coefficient sets the importance of the second loss term. The two neural networks with competing goals are an example of a zero-sum game where an increase in the generator's performance harms the discriminator's performance. The training objective is to reach the Nash equilibrium. Nevertheless, applying commonly used gradient descent methods might make the training process unstable compared to the traditional approach of training artificial neural networks.

Reconstruction-based methods use encoder-decoder neural network architectures to reconstruct the input data based on latent feature representations. Such a neural network can be embedded into a generative adversarial network and serve as the feature extractor. An encoder, which is shared between the source and target domains, assures that the domain-invariant features are extracted while keeping the domain-specific feature representations needed to reconstruct the available source and target data. An example of this approach is the domain separation networks proposed in [143]. The authors showed that separating information unique to each domain may improve the network's capabilities to extract domain-invariant features.

As already described, most domain adaptation methods assume the availability of data from the target domain and knowledge of the marginal probability distribution $P\left(\mathbf{X}_T\right)$. However, in practical applications, this assumption may not be satisfied. For example, this may happen in the case of newly introduced products for which no operational data is available until the first tests and commissioning of the unit. It is a significantly more challenging situation than typical domain adaptation problems. Domain generalization methods are applied to achieve out-of-distribution generalization and accurately predict in unseen target domains. In a typical

scenario, known as multi-source domain generalization, the training is based on several distinct but related domains. The objective is to leverage knowledge from multiple domains, extract domain-invariant representations and train a model with good generalization capabilities. The most difficult situation assumes a single source domain with homogenous data. Such a setup, known as single-source domain generalization, is less commonly studied in literature but is present in practical applications. The proposed solutions focus on applying data generation methods to make the training data more heterogeneous and informative, as follows from [144]. A comprehensive literature review on domain generalization methods is provided in [145].

In accordance with [144], domain generalization methods can be classified into the following three groups:

- methods focusing on data augmentation [146,147] and data generation [148];
- methods applying specific learning strategies to obtain better generalization capabilities, e.g. ensemble learning [149,150], meta-learning [151] or self-supervised learning [152];
- methods focusing on the extraction of domain-invariant representations [153,154] or learning disentangled representations, i.e. identifying domain-specific and domain-invariant features [155].

Most of the papers referenced here are related to rotating equipment. Nevertheless, the majority of studies about domain generalization refer to computer vision, natural language processing or speech recognition. Additionally, most of them focus on classification problems with limited applications to regression analysis [156], which is a gap that should be addressed appropriately. Naturally, data limitations and a desire to utilize knowledge from other sources do not concern classification problems only. Domain shifts are frequent in actual applications of regression analysis, and they can be related to dependent variables, independent variables, or both of them. The research presented in this dissertation contributes to the advancement of domain generalization methods in regression problems. In the proposed method, a physics-based relationship between the source and target domains is built and coupled with a physics-informed neural network to get accurate estimates on unseen data. The research works leading to the creation of this method with the applications to predict the size of damage to selected gas turbine components and the analysis of results are described in the following chapters.

4. Predicting length of fatigue cracks using machine learning methods

The main objective of this research is to successfully apply machine learning algorithms for predicting the size of damage to gas turbine components, considering the limited availability of damage measurements. The so-called small data regime is a common problem in practical applications, including the maintenance of aeroderivative and heavy-duty gas turbines. In that field, the limited availability of empirical data is due to rare inspection opportunities resulting from the extended intervals between planned maintenance activities, a rare occurrence of failures and high costs associated with the data gathering. Regardless of the data limitations, there is a need to leverage the gathered experience and create data-driven predictive models to continuously optimize the operation and maintenance of the fleet, assuring the highest levels of reliability and availability. Additionally, accurate predictive models are expected to provide unit-specific predictions using actual operational data of each engine recorded by a data acquisition system. Therefore, taking into account the need to predict the size of damage to gas turbine parts accurately, the minimal requirements for data availability are as follows:

- A certain number of damage measurements must be available.
- The operational history of the components with the damage observed and measured must be restored, considering the scope of past repair activities and the number of fired hours, fired starts and emergency shutdowns accumulated.
- Operational data captured by the data acquisition system should be available for the historical service periods of the analyzed components.
- Basic characteristics of the analyzed parts must be restored, e.g. component codes or drawing numbers.
- Configuration details describing the gas turbines, where the analyzed hardware operated, must be restored.

It is necessary to satisfy these requirements to begin creating a predictive model. Nevertheless, it does not guarantee that the model will have satisfactory accuracy and provide valuable and credible insights that may impact the maintenance planning. However, considering these constraints, it was decided to select a high-pressure nozzle of a heavy-duty gas turbine as the analyzed object, focusing on predicting the length of fatigue cracks. The additional arguments in favor of this choice are as follows:

- Problems related to thermo-mechanical, low cycle fatigue are common for all types of gas turbines (as already described, over 60% of damage observed in aircraft engines was due to cracks [28]).
- Considering hot gas path components, first stage nozzles of a high-pressure turbine section are subjected to the highest temperatures of exhaust gases and limit the time between planned maintenance activities, especially in the case of the analyzed family of heavy-duty gas turbines.
- The investigated family of industrial gas turbines represents the second most numerous group of units covered by long-term service agreements (LTSAs) offered by Baker Hughes Company.

- Low cycle fatigue is the leading cause of the analyzed cracks, but the contributions of oxidation and creep should not be neglected. Previous attempts to model this scenario with multiple combined failure types were unsuccessful.
- The input data needed to create the model are available, and no additional efforts and costs are required to gather them.

Details of this part of the research are provided in the following sections of this chapter. At the outset, it should be highlighted that the study and the results were described in [92].

4.1 The analyzed object and problem setup

Gas turbines generally are composed of the following three systems: a compressor, a combustor and a turbine section. Depending on the number of shafts, the compressors and turbine sections can be further divided into high- and low-pressure sections. Single-shaft heavy-duty gas turbines installed in stationary applications typically serve as generator drivers to produce electrical energy. In comparison, two-shaft units in stationary applications are mainly utilized as drivers of compressors or pumps. Gas turbines operate according to the Brayton thermodynamic cycle, which is based on the following four processes:

1) The compressor draws air from the surrounding environment and raises its pressure.
2) The energy level of the compressed air is increased in the combustor, where fuel is injected and burned.
3) The hot gases are directed to the turbine section, where thermal energy is converted to mechanical energy by means of stationary vanes, also known as nozzles, and rotating blades, known as buckets.
4) The expansion in the turbine section reduces the pressure and temperature of the exhaust gases, which are expelled into the atmosphere.

The compressor absorbs more than 50% of the mechanical energy created during the expansion. The efficiency of modern heavy-duty gas turbines may be higher than 37% when operating in the simple cycle, higher than 62% in the combined cycle coupled with a steam turbine [157], and up to 84% of thermal efficiency in combined heat and power [158].

The analyzed high-pressure nozzles come from two-shaft heavy-duty gas turbines offered by Baker Hughes Company, specifically designed for mechanical drive applications. The installed fleet comprises more than 400 units worldwide (status of 06-Jun-2023). Over the years, multiple upgrades were offered for this class of engines. The analyzed sample comprises three different models, which are referenced as Type A, Type B and Type C. It should be noted that the latter has a higher firing temperature, i.e. the average temperature of exhaust gases at the outlet of high-pressure nozzles, than the first two. The entire 1st stage nozzle assembly comprises 18 segments, each with two airfoils. The segments are single-piece castings from FSX-414 cobalt-based superalloy. A single segment is presented in Fig. 1 and Fig. 2 [159]. The component is cooled using pressurized air discharged from the compressor. Inside each airfoil, a cooling insert enables internal impingement cooling, while the cooling holes on the leading and trailing edges enable film cooling of the external surfaces. The number of cooling holes on the trailing edge may vary depending on the component code. Additionally, a thermal barrier coating may be applied on the airfoils' pressure side, but such a configuration is not included in the analyzed sample. Based on the applicable maintenance policy, the nozzles are expected

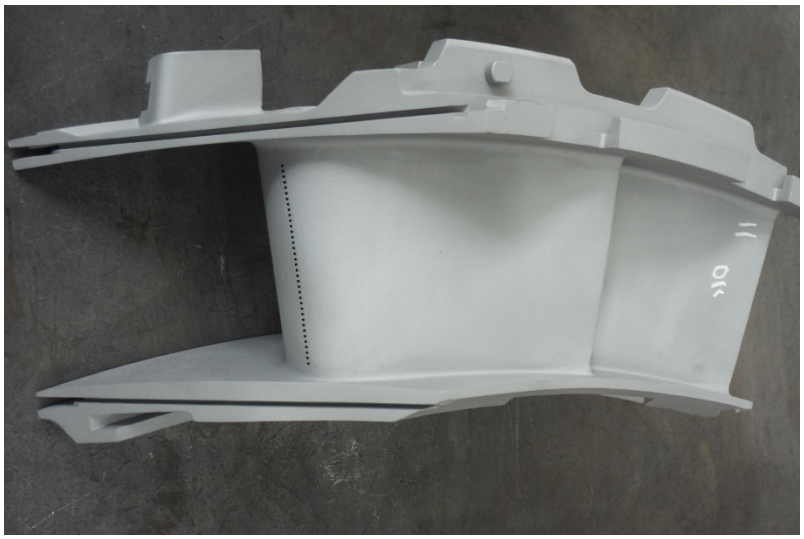


Fig. 1 Suction (convex) side view of the analyzed high-pressure nozzle.

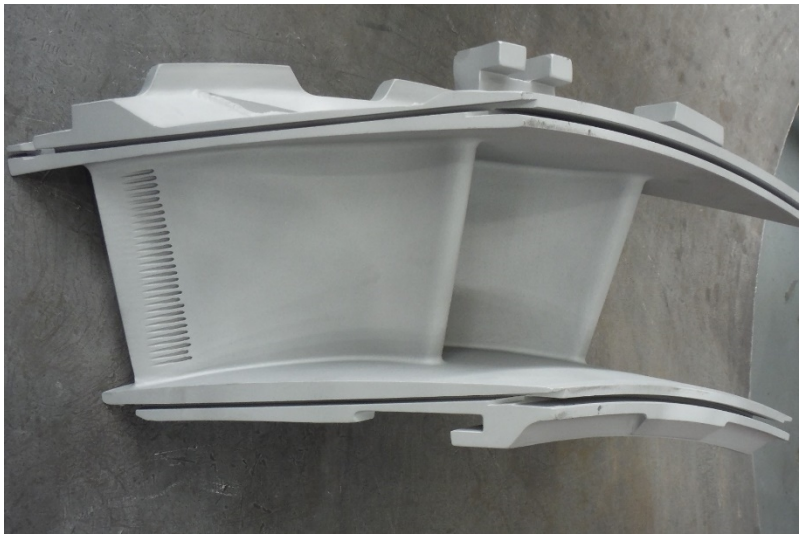


Fig. 2 Pressure (concave) side view of the analyzed high-pressure nozzle.

to be disassembled and repaired during each major inspection. The part should be repaired just once, and it should be scrapped after the second interval. Nevertheless, the OEM may propose less stringent maintenance rules for units covered by the long-term service agreement.

Based on data gathered during visual and dimensional inspections executed during disassembly inspections and subsequent repair activities, the following failure types can be found on the airfoils and platforms of the analyzed nozzles:

- cracks and craze cracks,
- oxidation or erosion,
- corrosion due to fuel or air contaminants,
- deformation of the airfoils caused by creep,
- indentation due to domestic object damage.

An overview of failure types observed on the first stage nozzle of a heavy-duty gas turbine is available in [160]. Due to the high temperatures of exhaust gasses surrounding the nozzles, this damage is not monitored nor directly measured continuously during the service. Nevertheless, as the number of service hours increases, the gradual degradation of the component (e.g. due to oxidation, erosion, or corrosion of the airfoils, inner/ outer platforms or both) should be reflected in gas turbine performances. It is possible to detect and isolate the fault based on operational data gathered by the data acquisition system and applying, for example, the gas path analysis. Based on this knowledge extracted from the data, a cross-functional team may decide

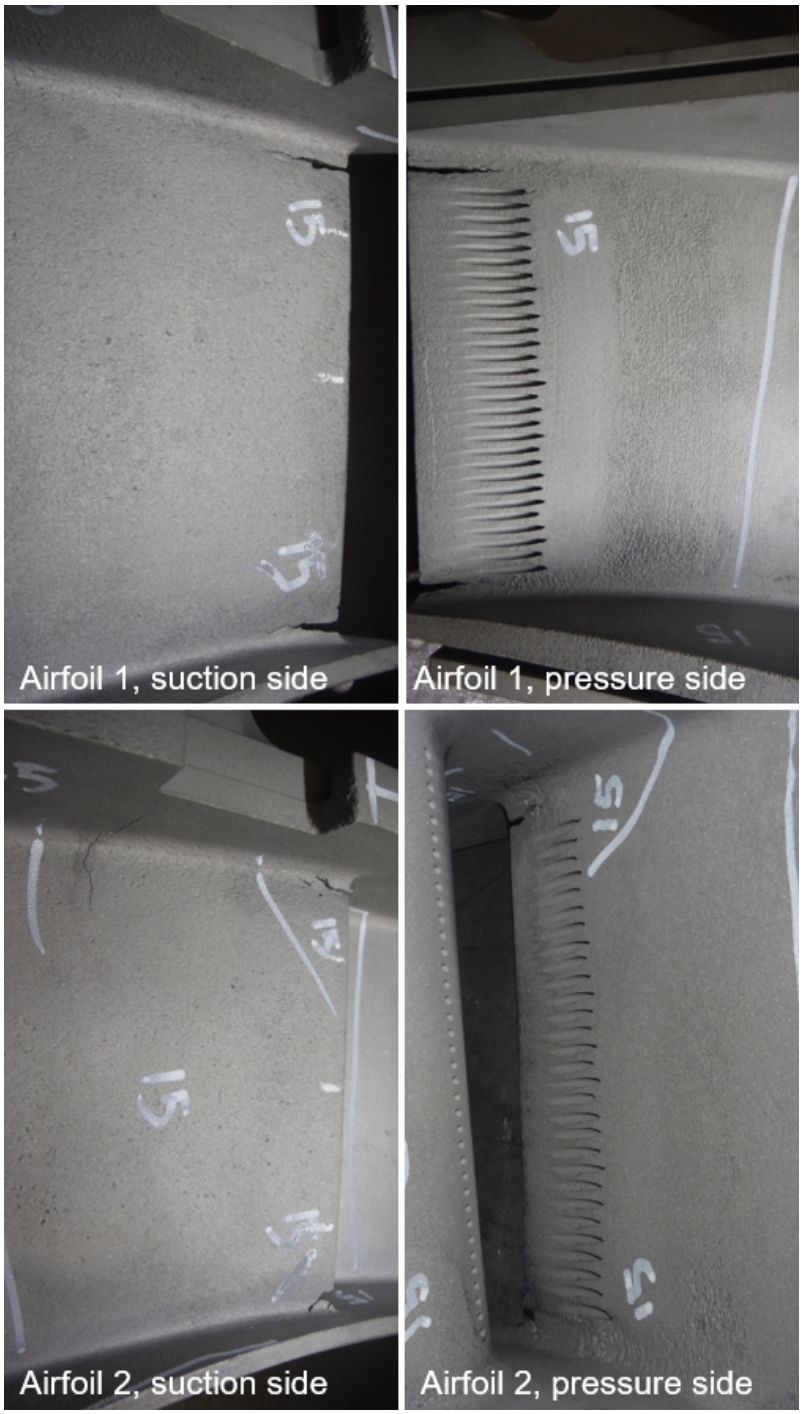


Fig. 3 Examples of cracks found on the analyzed high-pressure nozzles after service.

on the required corrective actions and further operation of the unit. However, the standard set of instruments cannot provide any direct or indirect indications about the presence and size of cracks. Examples of trailing edge cracks recorded on the analyzed high-pressure nozzles are shown in Fig. 3. It should be clearly stated that if a gas turbine is not maintained following the OEM standards, cracks of the nozzle may lead even to unplanned maintenance, for instance, if a piece of the airfoil is liberated due to converging cracks. Therefore, the ability to estimate the growth of nozzle cracks is of great importance for the availability and reliability of the analyzed unit. The potential advantages of such a capability are the same as the benefits of the predictive maintenance approach. It can be used to optimize the maintenance plan by applying data-driven insights from the model, e.g. continuous evaluation of the nozzles' health status, damage size and the associated failure probability based on actual operational data. Thus, it can support decisions related to the extension of repair or replacement interval, reduction of the work scope of a disassembly inspection or repair activity, or elimination of the need to repair the component.

Based on the above considerations, this part of the research aims to estimate the length of cracks found on the airfoils' trailing edges of the analyzed high-pressure nozzles, depicted in Fig. 3. Considering the airfoils' trailing edges, in accordance with the dedicated thermal analysis the highest metal temperatures are expected in locations where the cracks are found. The cracks are caused mainly by thermal stresses, which take the highest values during engine shutdowns. Additionally, based on the structural analysis, some of these locations are subjected to tensile stresses during steady-state operation, with the stresses acting in the direction perpendicular to the crack propagation plane. Furthermore, the crack growth rate may be affected by the base material thickness reduction due to oxidation. Therefore, considering the analyzed problem's complexity and the minimum requirements for data availability are met, it was decided to build data-driven models using several popular statistical learning algorithms and assess their performance in this scenario.

4.2 An overview of available empirical data

Data supporting this research are confidential and are the proprietary property of Baker Hughes Company. Due to these reasons, the data cannot be shared openly and were anonymized in this dissertation. The analyzed positions on the trailing edges of the nozzles are referenced as Positions 1, 2, 3 and 4 without disclosing further details.

Input data used to train and test the predictive models comes from three different databases of Baker Hughes Company. The first set is extracted from the Parts Life Database (PLDB) and contains the following information separately for each nozzle segment:

- The operational history of the analyzed components, including the scope of historical repair activities and the numbers of fired hours (FH), fired starts (FS) and emergency shutdowns (ESD) accumulated since the last repair (commonly referred as to interval counters) and since the part was manufactured (commonly referred as to total counters).
- Measurements of the cracks observed on the used components, which were subjected to a visual inspection preceding repair activities.
- The component codes and drawing numbers required to distinguish different nozzle configurations.

The repository contains 865 records with complete (i.e. data about the crack presence and size available for all four positions) and credible measurements of the analyzed cracks, where each record corresponds to a single nozzle segment. Only 13% of the parts, i.e. 114 out of 865 segments, did not have cracks in the analyzed locations. Considering the damaged components, the longest cracks, or less than 5 mm shorter than the longest, were observed at Position 2 in 85% of the cases (see Fig. 4). Therefore, the regression analysis focuses on this position and the maximal crack length at Position 2 is considered as the dependent variable, estimated by the models. Additionally, since the analyzed subsets are not normally distributed, a nonparametric Mood's median test was used to test the equality of the samples' medians. The resultant p-value equal to zero indicated that the null hypothesis should be rejected and that the medians are not equal. The median of measurements at Position 2 is several times higher than the remaining positions, as shown in Fig. 4.

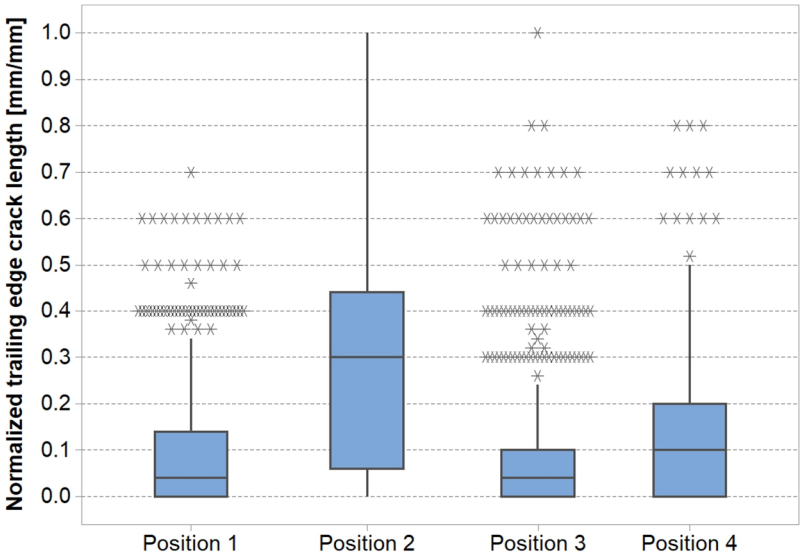


Fig. 4 Variation of the analyzed crack length measurements. The data normalized by dividing each value by the maximal measured crack length to protect proprietary information.

The second set of inputs contains operational data recorded by the Remote Monitoring & Diagnostics system. The system collects measurement data from gas turbine instrumentation about temperatures, translational displacements (vibrations), linear or angular positions, pressure and pressure differences, speeds, flows and others. Besides raw measurements, outputs of synthetic sensors and calculated parameters are also available. The following time series were utilized to create the models:

- position of the inlet guide vanes and nozzle guide vanes,
- pressure, temperature and relative humidity of ambient air,
- pressure and temperature of the air at the axial compressor's discharge,
- pressure losses in the inlet and exhaust ducts,
- rotational speeds of the high-pressure and low-pressure rotors,
- the axial compressor pressure ratio,
- the average temperature of exhaust gases at the inlet and outlet of the nozzles and the gas turbine's output (these three parameters are calculated based on other signals),
- temperatures of exhaust gases and their spread.

Components with low availability of the operational data, covering less than 25% of service hours since the last repair, i.e. interval fired hours (IFH), are excluded from further processing. Thus, the sample is reduced to 555 segments out of 31 nozzle sets. The sampling interval of the data extracted from the RM&D system can be as short as one second, meaning that the interval between two consecutive samples equals one second. Because some of the analyzed parts accrued more than 50 000 fired hours since the last repair, a very short sampling interval would increase the dataset significantly. Additionally, the analyzed problem is not characterized by such dynamic changes. Therefore, the sampling interval is equal to one hour, allowing to model the failure type accurately and simultaneously keeping the dataset size reasonable. Raw data downloaded from the system must be preprocessed to handle measurement errors, inconsistencies and incomplete data. The objectives of the data cleansing are as follows:

- Complete the data on ambient pressure and relative humidity based on the data related to other units from the same locations or using the data from the public repository [161].
- Complete the data on inlet and exhaust pressure losses using a regression model based on a random forest algorithm or information available in gas turbines' data sheets.
- Verify the data on high- and low-pressure shaft speeds.
- Remove the non-numerical data from the dataset, including strings, NaN (i.e. Not a Number) and positive or negative INF (i.e. infinity).
- Calculate the missing values of the compressor pressure ratio and corrected speed of the high-pressure shaft based on analytical formulas.
- Considering the previously accomplished tasks listed above, remove from the dataset records containing erroneous values outside the acceptable ranges and records related to the transient states, i.e. startup, acceleration, load step, load rejection, deceleration or shutdown. A range of acceptable values was defined for each analyzed operational parameter. Records with values outside these ranges are excluded from the dataset.
- Estimate values of the calculated parameters, i.e. the average temperature of exhaust gases at the inlet and outlet of the nozzles and the gas turbine's output. An in-house estimator based on an artificial neural network that rebuilds the thermodynamic cycle of the analyzed gas turbines is applied for this task.

Finally, the preprocessed dataset of the operational data comprises 1 029 215 records in total. Nevertheless, the data is incomplete for most of the 31 analyzed nozzle sets. The coverage, defined as the percentage of service hours with the operational data available, ranges from 26% to 100% depending on a particular set, with the mean coverage equal to only 73%. Therefore, the fraction of missing data is not insignificant. Fortunately, the analyzed gas turbines have a very stable operating profile. They are the drivers of centrifugal compressors in several natural gas liquefaction facilities worldwide. The units typically operate with their nominal power in continuous duty, with a high number of fired hours per each startup-shutdown cycle. Consequently, it is assumed that the available operational data could also be utilized to describe the missing periods.

The third set of inputs contains mainly categorical data describing the following configuration details of the analyzed gas turbines:

- the model, i.e. Type A, Type B or Type C;
- the type of combustion system, i.e. either Standard or Lean Head End combustor;
- primary and secondary fuel (if any) types, including the lower heating values;

- the geographical position of each unit.

The database also contains basic identification details, including serial numbers of the assets and customer names. Parameters that do not differentiate the investigated units or should not have any impact on the analyzed phenomenon were identified but not listed above.

The final set of input data comprises 32 variables. 21 out of 32 are the time series recorded by the RM&D system or calculated based on data from the system. 8 out of 32 were extracted from the Parts Life Database (i.e. two discrete variables describing the nozzle configuration and the number of past repairs, and six continuous variables with the interval and total counters). The remaining three discrete variables describe the engines' configuration details.

Concerning the dependent variable, it should be made clear that considering a particular set of nozzles, the crack lengths at Position 2 vary and are not the same for all 18 segments. Some differences in the base material properties may cause it. However, it cannot be ruled out that segments at a specific angular position may suffer more due to temperature pattern factors at the high-pressure turbine section's inlet. Thus, the dependent variable is distributed for a particular nozzle set, but the explanatory variables are the same for all segments within that set. Nevertheless, viewing the reliability and availability of the analyzed gas turbines, it should be enough to focus on the worst segment and accurately estimate the size of the longest crack. Therefore, to simplify the problem, the maximal crack size at Position 2 is chosen as the response variable. Thus, each set of the nozzles is described by a single value, composing the set of observations with 31 elements.

4.3 Feature selection

The set of features comprises 32 variables, but only 31 observations are available. In this case, it is necessary to reduce the dimension of the feature space to make the models more interpretable and avoid overfitting [162,163]. This process is known as feature selection. Many recommendations regarding the number of observations per feature to do a regression analysis are summarized in [164,165]. However, these suggestions vary significantly, from two hundred subjects (observations) per feature to just two. Considering these studies, the complexity of the analyzed phenomenon and the number of exploratory and response variables, it is decided to create the models using five features at the most, i.e. it gives around six subjects per feature. The models will share the same predictors regardless of the applied machine learning algorithm. Such an approach should allow for a direct comparison of the effectiveness of the algorithms and make it easier to interpret the obtained results. Filter and wrapper methods will be utilized during the feature selection process.

In order to combine the inputs from the PLDB, the gas turbines' configuration details and the time series recorded by the RM&D system, the latter is simplified to medians for each observation separately (i.e. considering a selected operating parameter, if 10 000 measurements are available for a particular observation, they are reduced to a single value, which is the median of that sample composed of 10 000 values; it is executed for each parameter separately). This transformation unifies the dimension of inputs and significantly reduces the quantity of data that will be processed during the training of the models. Such an approach is justified in the analyzed case since the investigated gas turbines operate continuously at their nominal power. Nevertheless, this assumption would not be valid for units characterized by a high load variance.

A filter method is applied in the first phase of the feature selection process. The goal is to remove the least important variables and intercorrelated features based on a simultaneous analysis of Pearson's and Spearman's correlation coefficients and the mutual information. The first is a measure of a linear relationship between two variables, while the second measures if the variables are monotonically related. The mutual information $I(X;Y)$ is a measure of the dependence between two variables X and Y, defined as the reduction in uncertainty of a single random variable due to another random variable [166]:

$$I(X;Y) = H(X) - H(X \mid Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)\,p(y)}, \tag{17}$$

where $H(X)$ is the measure of the average uncertainty in the random variable X (i.e. it is the entropy of that random variable), $H(X \mid Y)$ is the conditional entropy of X given the knowledge of Y and p represents a (joint) probability mass function. This measure is symmetric in X and Y, takes nonnegative values only, and equals zero if and only if the two random variables are independent. The scikit-learn Python library [167] is utilized to calculate the three measures. Additionally, assuming that the response variable is categorical, a chi-square test is applied to examine whether there is a dependency between the target and the categorical input variables. The least important variables are identified and removed iteratively, eliminating around 20% of the remaining covariates during each iteration until ten variables with the best measures are identified. As a result of the filtering, the following independent variables are identified:

- number of interval fired starts;
- numbers of total fired starts and emergency shutdowns;
- temperature and relative humidity of ambient air;
- the fuel stroke reference, which represents the amount of fuel supplied to the combustion chambers;
- the axial compressor pressure ratio;
- the average temperature of gases at the inlet of the nozzles and the gas turbine's output;
- the average temperature of gases measured in the exhaust duct.

The first two positions from the above list refer to the variables from the Parts Life Database, while the remainder represents the time series reduced to median values.

The next phase of the process aimed to identify the final set of predictors is known as feature engineering. The objective is to create new, calculated parameters based on the existing ones. The dependence between these new variables and the response may be higher compared to the measures obtained during the filtering. However, considering that the filtered independent variables correspond to physical quantities that are either measured or calculated, the new variables should also have a physics-based meaning and be easy to interpret. Only the previously filtered operating data are considered during the feature engineering, excluding the fuel stroke reference and relative humidity of ambient air. The new variables shall reflect the distributions of those operating parameters in a simplified way. Based on the entire operational data set with 1 029 215 records, certain statistical measures (i.e. 50^{th}, 70^{th} and 90^{th} percentile) are calculated for each parameter separately. Based on these measures, four new variables are introduced per each operating parameter and observation, i.e.:

- the number of service hours with the reading lower than the 50th percentile;
- the number of service hours with the reading between the 50th and 70th percentiles;
- the number of service hours with the reading between the 70th and 90th percentiles;
- the number of service hours with the reading higher than or equal to the 90th percentile.

This procedure temporarily increases the number of exploratory variables to 30, but by applying the same approach of filtering as previously, the list is limited as follows:

- number of total fired starts;
- temperature and relative humidity of ambient air;
- the axial compressor pressure ratio;
- the average temperature of gases at the inlet of the nozzles and the gas turbine's output;
- the average temperature of gases measured in the exhaust duct.

Finally, a wrapper method is utilized in the last phase of the feature selection process. The objective is to find the optimal combination of independent variables, given the previously defined limit of five predictors at the most, using a specific machine learning algorithm to select them. A random forest is chosen, considering the high prediction capabilities of an ensemble of decorrelated decision trees [168] and applying the law of large numbers. Features are removed recursively based on the importance described in [169] and defined as the total decrease in node impurity, weighted by the proportion of samples reaching that node in each decision tree and averaged over all 100 trees of the random forest. Additionally, a 5-fold cross-validation (CV) with 50 repetitions is applied during each iteration. Therefore, the feature importance values are also averaged over 250 random forests trained on different subgroups of the dataset. Such an approach could be applied and is not computationally expensive due the effective filtration of features at the previous phases of the feature selection process. As a results, the following independent variables are selected for the modeling purposes:

- number of total fired starts *TFS*,
- median ambient air temperature $\widetilde{T_{AMB}}$,
- number of service hours with the average temperature of gases at the inlet of the nozzles between the 50th and 70th percentiles $T_{INLET\ P50\text{-}P70}$,
- median gas turbine's output $\widetilde{P}$,
- median average temperature of gases measured in the exhaust duct $\widetilde{T_{EXH}}$.

The set of predictors is very diverse in terms of the source. It pools the information from the Parts Life Database, i.e. *TFS*, the variables based on the operational parameters recorded by the RM&D system, i.e. $\widetilde{T_{AMB}}$ and $\widetilde{T_{EXH}}$, the variable based on the parameter calculated using an artificial neural network, i.e. $\widetilde{P}$, and the variable created during the feature engineering, i.e. $T_{INLET\ P50\text{-}P70}$. Even though the predictors were determined based on the analysis of statistical measures, they reflect well knowledge about the causes of the analyzed phenomenon. The temperatures of the nozzles' base material on the hot and cold sides (i.e. on the external surface of the airfoils and from the inside of the cooling cavity) can be approximated as follows:

$$T_{HOT} = T_{INLET} - \phi_{HOT}\left(T_{INLET} - T_{COOL}\right), \tag{18}$$

$$T_{COLD} = T_{INLET} - \phi_{COLD}\left(T_{INLET} - T_{COOL}\right), \tag{19}$$

where T_{HOT} and T_{COLD} are the metal temperatures on the hot and cold sides, ϕ_{HOT} and ϕ_{COLD} are the cooling effectiveness coefficients on both sides and T_{COOL} is the air temperature at the axial compressor discharge. The thermal stresses σ resulting from the difference in material temperatures between both sides ΔT_{MATL} can be estimated as follows:

$$\sigma = E\alpha\Delta T_{MATL} = E\alpha\left(T_{HOT} - T_{COLD}\right) = E\alpha\left(\phi_{COLD} - \phi_{HOT}\right)\left(T_{INLET} - T_{COOL}\right), \tag{20}$$

where E is the Young's modulus and α is the thermal expansion coefficient. It should be noted that one of the selected predictors is calculated based on the values of T_{INLET}, while the values of gas turbines output are correlated with the values of T_{COOL}. Furthermore, the number of fired starts accumulated since the manufacturing can represent the number of load cycles, an essential term of any crack propagation rate equation. Such consistency between these basic equations describing the analyzed problem and the feature selection process results can be thought of as a confirmation that the process was executed correctly and that the set of predictors is composed optimally, considering all data limitations.

4.4 Training, tuning and testing of the models

Creating models using machine learning algorithms is strictly connected with the bias-variance tradeoff. The bias error is introduced in the case of an oversimplified approximation of the analyzed real-life problem, failing to capture the relationship between the input and output variables resulting in inaccurate estimates, so-called underfitting. The variance error is introduced in the case of models sensitive to peculiarities of the training subset, which perform well on the training data but have poor generalization capabilities and low prediction accuracy when applied to previously unseen data. This phenomenon is known as overfitting. According to [170], given a test value x, the expected mean squared error can be expressed as the sum of variance of $\hat{f}(x)$, the squared bias of $\hat{f}(x)$ and variance of the irreducible error ε:

$$E\left[y - \hat{f}(x)\right]^2 = Var\left[\hat{f}(x)\right] + \left[bias\left(\hat{f}(x)\right)\right]^2 + Var(\varepsilon), \tag{21}$$

where $\hat{f}$ is the learned function and y is the observation corresponding to the input variable x. Based on Eq. (21), it is visible that during the error minimization, values of the variance and bias terms are reduced simultaneously to find a trade-off between them. Therefore, to create models that generalize well and have optimized values of hyperparameters, a robust approach is to split the sample into the following three subsets:

- the training set utilized to fit the trainable model's parameters;
- the validation set utilized to tune hyperparameters of the model, which values are set by the modeler and used to control the training process;
- the testing set utilized to assess the model's performance on the unseen data.

In the analyzed case, the sample comprises just 31 data points. It should be split cautiously to balance the number of elements in each subset appropriately and to prevent errors that may negatively impact the entire analysis. As part of preparations to divide the sample, the available crack measurements are split into three classes based on the threshold values, L_{LOW} and L_{HIGH}:

- "short" class comprising the measurements lower than L_{LOW} (9 observations out of 31, which is 29% of the sample);
- "medium" class comprising the measurements that are higher than or equal to L_{LOW} and lower than L_{HIGH} (10 observations, 32%);
- "long" class comprising the measurements that are higher than or equal to L_{HIGH} (12 observations, 39%).

Considering these clusters, the sample's composition can be described as follows:

- 24 observations out of 31 (77%) correspond to a natural gas liquefaction facility located in a marine, salty environment. 95% of records from the "short" and "medium" classes are related to that plant, i.e. 18 observations out of 19.
- 6 observations out of 12 (50%) classified as "long" correspond to units installed in two other gas liquefaction facilities located in a tropical, humid environment.
- 19 records out of 31 (61%) are related to Type C gas turbines making up the majority of the "medium" and "long" classes (i.e. 90% and 58%, respectively).
- 7 records out of 31 (23%) are related to Type A units making up 33% of the "short" and "long" classes.
- 5 records out of 31 (16%) are related to Type B units making up 33% of the "short" class and minorities of the remaining two classes.

In summary, the available dataset is imbalanced, with an overrepresentation of records related to just one facility and a high number of data points corresponding to Type C units. However, regardless of the limited and imbalanced sample, the models should be capable of providing accurate estimates across the entire range of observations. In general, the composition of the training, validation and testing subsets should reflect the structure of the available sample. Therefore, the test set is composed considering the following constraints:

- The proportions between the "short", "medium" and "long" classes valid for the sample are applied to select elements of the test subset. Thus, each of the classes constitutes 1/3 of the test set.
- In order to avoid a significant reduction of the training and validation subsets, the test set comprises six elements. 81% of the data will be used to train the models and optimize their hyperparameters, while the remainder will be utilized to assess the generalization capabilities.
- 5 observations out of 6 (83%) correspond to the gas liquefaction facility located in a marine, salty environment.
- 4 observations out of 6 (67%) are related to Type C gas turbines. The remaining test data points correspond to Type A and Type B units.
- In order to enlarge the training subspace, records for which the predictors take moderate values are preferred to be assigned to the test set. The remainder, characterized by higher variance, is used to train and validate the models.

As described in the previous chapter, the relationship between the inputs and outputs is approximated during the learning process aimed at optimizing a predetermined cost function. This function represents the prediction error aggregated for all elements of the training dataset, while for a single element, a loss function is used to calculate the error. The squared error is frequently selected as the loss function in regression analysis. Consequently, the mean squared error (MSE) or root mean squared error (RMSE) is commonly chosen as the cost function.

Nevertheless, their popularity does not mean nor assure that their use will lead to optimal results always. The way the cost function is structured defines what the model will attempt to learn. Consistently, it significantly impacts the results determined by a fully-specified model, can even determine whether the model is valuable, and provides credible data-driven insights to support, e.g. business decisions. Therefore, a deep understanding of the analyzed problem and expectations of future users of the predictive model is a mandatory requirement to define the cost function appropriately. Due to the various peculiarities of many practical problems, creating a customized cost function may be necessary to obtain satisfactory results.

As already described, in the analyzed case of fatigue cracks found on the nozzles, the capability to accurately estimate the size of the longest cracks is the most important in ensuring that the reliability and availability targets are not compromised. In practice, a certain value of absolute error is allowed for the "short" class since such cracks do not jeopardize the part's integrity or require immediate corrective maintenance. Simultaneously, the same absolute error cannot be accepted for the "long" class, which pools the cracks that should be carefully examined. All these considerations should be better captured and reflected in a custom cost function compared, for example, to the (root) mean squared error. The proposed loss function is based on a variable-width scoring interval defined by two scoring bounds, presented in Fig. 5. Estimates falling within the scoring interval are considered successful predictions with sufficient accuracy. The lower bound $b(x)$ and the upper bound $t(x)$ get closer to each other as the crack size increases. The minimum distance between the bounds, corresponding to the "long" class, equals around 40% of the maximal one, related to the "short" class. The loss function takes discrete values and is defined as follows:

$$l\left(x,\hat{f}\left(x\right)\right)=\begin{cases}0, & x\notin\left[b\left(x\right),t\left(x\right)\right]\\ 1, & x\in\left[b\left(x\right),t\left(x\right)\right] \text{ and } x<L_{HIGH}\\ 2, & x\in\left[b\left(x\right),t\left(x\right)\right] \text{ and } x\geq L_{HIGH}\end{cases}. \tag{22}$$

If the crack length predicted by the model is within the scoring interval, the loss function has a positive value. It should be emphasized that accurate predictions of the cracks classified as "long" with the length higher than or equal to L_{HIGH} get a two-times higher score than accurate predictions of cracks classified as "short" or "medium". This bonus aims to promote solutions capable of accurately predicting the length of the longest cracks. Nevertheless, the magnitude of the bonus should not be raised artificially, as it is a quantitative representation of the higher importance of some attributes over others. During the creation of the predictive models, it was observed that raising the bonus too high results in models providing heavily biased predictions for the "short" and "medium" classes, regardless of their excellent capabilities to predict the longest cracks accurately. Finally, the cost function is defined as the average loss:

$$c\left[l\left(x,\hat{f}\left(x\right)\right)\right]=\frac{1}{N}\sum_{i=1}^{i=N}l\left(x_i,\hat{f}\left(x_i\right)\right), \tag{23}$$

where $i=1,2,\ldots,N$ is the observation's ordinal number, N is the total number of observations utilized in the training set. The objective of the optimization problem is to maximize the cost function's value by adjusting the model's trainable parameters.

The validation set is composed considering the following constraints:

- The proportions between the "short", "medium" and "long" classes valid for the sample are applied to select elements of the validation set. Thus, each of the classes constitutes 1/3 of the validation subset.
- In order to avoid a significant reduction of the training set, the validation set comprises only three elements, one from each class, i.e. it represents 10% of the sample. Thus, in the case of validation, the codomain of the cost function is a set composed of five elements, i.e. $\left\{0, \frac{1}{3}, \frac{2}{3}, \frac{3}{3}, \frac{4}{3}\right\}$.

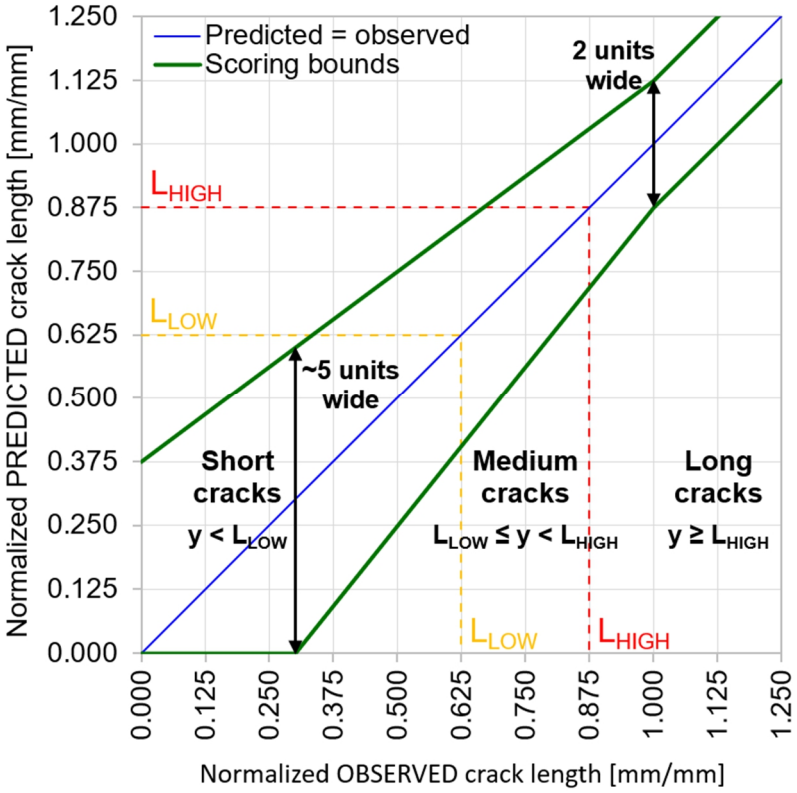


Fig. 5 The two scoring bounds composing the variable-width scoring interval of the custom cost function presented on the predicted vs. observed plot.

Due to the low cardinality of the validation set, a modified leave-one-out cross-validation procedure will be applied, considering all the possible train-validation splits. Cross-validation is necessary in this situation to avoid drawing conclusions based on the results obtained from a few randomly chosen data points. In the proposed approach, 560 distinct train-validation splits are generated based on the available data since the dataset comprises seven elements from the "short" class, eight elements from the "medium" class and ten elements from the "long" class. Values of hyperparameters controlled by the modeler are fixed for all 560 repetitions. Unique datasets are used for the training and validation (i.e. they comprise 22 and 3 observations, respectively) during each iteration. The final cross-validation score is the arithmetic average of the cost values from all the repetitions. The set of hyperparameters with the highest average CV score is considered optimal. Finally, the parameters are fixed, the model is trained on the train-validation set with 25 observations and is evaluated against the unseen data from the test subset.

The training, validation and test sets were not drawn randomly. They were prepared in a controlled manner considering multiple constraints, which were defined based on the analysis of the composition and clusters of the available sample. This approach is expected to improve the training effectiveness and the models' accuracy and simplify the interpretation of obtained results. Additionally, the proposed customized loss and cost functions reflect the peculiarities of the objectives better than the commonly utilized metrics, like MSE or RMSE. Therefore, the use of these functions should increase the models' usefulness and the chances of deploying them in real applications.

4.5 An overview of utilized machine learning algorithms

The following methods will be applied to create the data-driven models predicting the maximal length of fatigue cracks at Position 2 of the analyzed high-pressure nozzles:

1. multiple linear regression,
2. polynomial regression,
3. support vector regression,
4. kernel ridge regression,
5. random forest algorithm,
6. AdaBoost.R2 algorithm,
7. Extreme Gradient Boosting algorithm,
8. artificial neural network.

A concise overview of the theoretical foundations of these methods is given in this section.

4.5.1 Multiple linear regression

The key assumption of models based on multiple linear regression is that the target variable is linearly dependent on the predictors. Given the multiple inputs $x_1, x_2, \ldots, x_K$ and the output *y*, the linear relationship between them has the following scalar form:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + \varepsilon \,, \tag{24}$$

where $i = 1, 2, \ldots, N$ is the observation's ordinal number, $k = 1, 2, \ldots, K$ is the ordinal number assigned to each predictor, β_0 is the intercept, β_k are the coefficients of the regression equation and ε is the error term. It is assumed that $\beta_0 = 0$ for the analyzed problem since the nozzles do not have any cracks at the beginning of the service period. Then, using matrix notation, Eq. (24) can be simplified as follows:

$$\vec{y} = \vec{\beta}\mathbf{X} + \vec{\varepsilon} \,, \tag{25}$$

where vector $\vec{\beta}$ comprises the equation coefficients and **X** is a $K \times N$ matrix of the input data.

4.5.2 Polynomial regression

The key assumption of models based on polynomial regression is that the target variable is nonlinearly dependent on the predictors. The relationship between the inputs and the output has the following scalar form:

$$
\begin{aligned}
y_i = \beta_0 &+ \beta_{11}x_{i1} + \beta_{12}x_{i1}^2 + \beta_{13}x_{i1}^3 \cdots + \beta_{1p}x_{i1}^p + \\
&+ \beta_{21}x_{i2} + \beta_{22}x_{i2}^2 + \beta_{23}x_{i2}^3 \cdots + \beta_{2p}x_{i2}^p + \\
&+ \beta_{31}x_{i3} + \beta_{32}x_{i3}^2 + \beta_{33}x_{i3}^3 \cdots + \beta_{3p}x_{i3}^p + \\
&+ \cdots + \\
&+ \beta_{K1}x_{iK} + \beta_{K2}x_{iK}^2 + \beta_{K3}x_{iK}^3 \cdots + \beta_{Kp}x_{iK}^p + \varepsilon \,,
\end{aligned}
\tag{26}
$$

where p is the degree of the polynomial equation. Assuming $\beta_0 = 0$, Eq. (26) in matrix notation is the same as Eq. (25). However, the size of matrix **X** depends on the degree of the polynomial and on the number of interaction features of the degree $d = 1, 2, \ldots, p$. These additional features are calculated as products of several independent variables, which are distinct from each other (i.e. terms used to compute the interaction features have different values of the subscript k). The degree of the polynomial $p \in \{2, 3, 4\}$ is the only parameter assigned by the modeler and tuned during this analysis.

4.5.3 Support vector regression

Support vector regression generalizes the concept of support vector machines [171], which is frequently applied to solve classification problems. The SVR algorithm formulates Eq. (25) as an optimization task to find the narrowest margin around the approximated surface while minimizing the distance between the predictions and observations [172]. The maximum error ε determines the width of the interval of acceptable error and is a hyperparameter tuned during the cross-validation. The optimization is aimed at minimizing the Euclidean norm of the coefficients vector $\vec{\beta}$, i.e. the magnitude of the normal vector to the approximated surface, subjected to $\left| y_i - \vec{\beta}\vec{x}_i \right| \le \varepsilon$:

$$
\min \frac{1}{2}\left\| \vec{\beta} \right\|^2 . \tag{27}
$$

The algorithm penalizes only those predictions that are outside the ε-insensitive region, which are referred to as support vectors. Consequently, smaller values of ε result in more support vectors, while increasing the error has the opposite effect. This constrained optimization problem has the following solution:

$$
\hat{f}\left(\vec{x}\right) = \vec{\beta}\vec{x} = \sum_{i=1}^{i=N} \left(\alpha_i - \alpha_i^* \right) \vec{x}_i^T \cdot \vec{x} \,, \tag{28}
$$

where α_i and α_i^* are Lagrange multipliers and $\cdot$ denotes the scalar product in the space of input data $\mathcal{X}$. The coefficients vector $\vec{\beta}$ is expressed as a linear combination of the training vectors $\vec{x}_i$ and it is not required to compute $\vec{\beta}$ explicitly. Therefore, the so-called support vector expansion does not depend on the dimension of the input space $\mathcal{X}$, but it is driven by the number of support vectors only [173]. In the case of nonlinear relationships between the input and output variables, the data can be mapped into a higher dimensional feature space $\mathcal{F}$, to make them linearly separable in that space and to apply the SVR algorithm. In space $\mathcal{F}$, the scalar product of the vectors visible in Eq. (28) can be expressed using a similarity function

called a kernel function $k(\vec{x}_i, \vec{x})$. Thus, the explicit mapping $\Phi: \mathcal{X} \to \mathcal{F}$, requiring high computational power in the case of multidimensional problems, is not needed if the kernel satisfies:

$$k(\vec{x}_i, \vec{x}) := \Phi(\vec{x}_i)^T \cdot \Phi(\vec{x}). \tag{29}$$

The solution to this constrained optimization problem in the feature space $\mathcal{F}$ is as follows:

$$\hat{f}(\vec{x}) = \vec{\beta}\Phi(\vec{x}) = \sum_{i=1}^{i=N} (\alpha_i - \alpha_i^*) \Phi(\vec{x}_i)^T \cdot \Phi(\vec{x}) = \sum_{i=1}^{i=N} (\alpha_i - \alpha_i^*) k(\vec{x}_i, \vec{x}). \tag{30}$$

The kernel type may significantly impact the model's performance and is considered one of the hyperparameters being tuned during the training. The SVR class of the scikit-learn Python library is used to create the model, which has the linear, polynomial, radial basis function and sigmoid kernels embedded.

4.5.4 Kernel ridge regression

The kernel ridge regression combines the ridge regression algorithm, which is a linear model with one regularization parameter, with a kernel function. The inclusion of the regularization parameter in the cost function $\lambda_2 \geq 0$ reduces the chances of overfitting and the impact of insignificant features or eliminates such input variables from the regression equation (i.e. the latter could be achieved with the lasso regression method). The objective function minimized during the training process has the following form:

$$\min\left(\left\| \vec{y} - \mathbf{X}\vec{\beta} \right\|^2 + \lambda_2 \left\| \vec{\beta} \right\|^2 \right), \tag{31}$$

where the first term is the squared error and the second is the ridge penalty. In this case, the vector of coefficients can also be expressed as a linear combination of the training vectors $\vec{x}_i$:

$$\begin{aligned} \vec{\beta} &= \left(\mathbf{X}^T\mathbf{X} + \lambda_2 \mathbf{I}_N \right)^{-1} \mathbf{X}^T \vec{y} = \mathbf{X}^T \left(\mathbf{X}\mathbf{X}^T + \lambda_2 \mathbf{I}_N \right)^{-1} \vec{y} = \\ &= \mathbf{X}^T \left(\mathbf{G} + \lambda_2 \mathbf{I}_N \right)^{-1} \vec{y} = \mathbf{X}^T \boldsymbol{\alpha} = \sum_{i=1}^{i=N} \alpha_i \vec{x}_i \,, \end{aligned} \tag{32}$$

where $\mathbf{G} = \mathbf{X}\mathbf{X}^T$ is the Gram matrix, $\mathbf{I_N}$ is the $N \times N$ identity matrix and $\boldsymbol{\alpha} = \left(\mathbf{G} + \lambda_2 \mathbf{I}_N \right)^{-1} \vec{y}$ is the dual variable. For a test vector of input data $\vec{x}$, the prediction is calculated as follows:

$$\hat{f}(\vec{x}) = \vec{\beta}^T \vec{x} = \sum_{i=1}^{i=N} \alpha_i \vec{x}_i^T \cdot \vec{x} = \vec{y}^T \left(\mathbf{G} + \lambda_2 \mathbf{I}_N \right)^{-1} \vec{k} \,, \tag{33}$$

where $k_i = \vec{x}_i^T \cdot \vec{x}$. For nonlinear relations, the predictive function involves the Gram matrix with elements $\mathbf{G}_{ij} = \Phi(\vec{x}_i)^T \cdot \Phi(\vec{x}_j)$ and the vector $\vec{k}$ containing values $k_i = \Phi(\vec{x}_i)^T \cdot \Phi(\vec{x})$. The explicit mapping into the higher dimensional feature space $\Phi: \mathcal{X} \to \mathcal{F}$ is unnecessary if the kernel satisfies Eq. (29). Similar to the support vector regression, the kernel type is one of the tuned hyperparameters. The KernelRidge class has the same kernels as the SVR class, plus the (additive) chi-squared, cosine and Laplacian kernel functions.

4.5.5 Random forest algorithm

A random forest is an ensemble of decision trees. A single decision tree partitions the features space $\mathcal{X}$ into J high-dimensional rectangles R_j trying to minimize the sum of squared residuals (RSS) at each split of the tree:

$$\min \sum_{j=1}^{j=J} \sum_{i \in R_j} \left(y_i - \hat{y}_{R_j} \right)^2, \tag{34}$$

where $j = 1, 2, \ldots, J$ is the rectangles' ordinal number and $\hat{y}_{R_j}$ is the mean of the predictions in the R_j rectangle. Decision trees are created in the top-down approach, meaning that the features at the top of the tree minimize the RRS better than the remaining variables and determine the high-level divisions of the dataset. Nevertheless, a single decision tree may have a too simple structure to model more complex problems accurately. An ensemble of trees is created to raise the accuracy and generalization capabilities of the model. Each split of each decision tree is determined based on a random subsample of available data. Therefore, trees composing the random forest are not correlated with each other. The final output of the model is obtained by averaging the responses of all the decision trees:

$$\hat{f}(\vec{x}) = \frac{1}{M} \sum_{m=1}^{m=M} \hat{f}_m(\vec{x}), \tag{35}$$

where $m = 1, 2, \ldots, M$ is the decision tree's ordinal number, M is the number of trees in the random forest and $\hat{f}_m(\vec{x})$ is the prediction of the mth decision tree given the input data $\vec{x}$.

4.5.6 AdaBoost.R2 algorithm

The utilized AdaBoost.R2 algorithm described in [174] is a modification of the original AdaBoost.R presented in [175]. In this method, the predictive model is an ensemble of decision stumps, i.e. one-level decision trees. The responses of these weak learners are combined into a weighted median to get the final output of the model. Each new decision stump added to the ensemble focuses on training observations with the most inaccurate estimates obtained at the previous iteration. Based on the algorithm, higher weights are assigned to such observations, thus increasing the probability of selecting them into a new training dataset used to train a weak learner during the next iteration. The weights depend on the confidence in the decision stump Θ, which is a function of the average loss arising from the use of that one-level decision tree. For a test vector of input data $\vec{x}$, the prediction is calculated as the weighted median:

$$\hat{f}(\vec{x}) = \inf \left\{ y \in \mathcal{Y} : \sum_{m: \hat{f}_m(\vec{x}) \le y} \log\left(\frac{1}{\Theta_m}\right) \ge \frac{1}{2} \sum_m \log\left(\frac{1}{\Theta_m}\right) \right\}, \tag{36}$$

where $m = 1, 2, \ldots, M$ is the weak learner's ordinal number, M represents their total number in the ensemble and $\hat{f}_m(\vec{x})$ is the prediction of the mth decision stump given the input data $\vec{x}$.

4.5.7 Extreme gradient boosting algorithm

The extreme gradient boosting algorithm [77] is another approach based on an ensemble of decision trees. In this method, the responses from gradient-boosted decision trees are added up to obtain the final output of the model:

$$\hat{f}\left(\vec{x}\right) = \sum_{m=1}^{m=M} \hat{f}_m\left(\vec{x}\right), \tag{37}$$

where the meanings of m, M and $\hat{f}_m\left(\vec{x}\right)$ are the same as in Eq. (35). A characteristic feature of these decision trees is a similarity score assigned to each tree leaf. During the learning process, the following cost function is minimized at each iteration t:

$$c^{(t)} = \sum_{l=1}^{l=T}\left[G_l w_l + \tfrac{1}{2}\left(H_l + \lambda_2\right) w_l^2 \right] + \gamma T\,, \tag{38}$$

where $l = 1,2,\ldots,T$ is the leaf's ordinal number, T is the number of leaves in the tree, G_l and H_l represent the first and second-order derivatives of the loss function (i.e. the gradient and the Hessian), w_l is the similarity score assigned to each leaf, λ_2 is the regularization parameter and γ is the minimum loss reduction required to make a further partition of a node. The structure and size of decision trees are reduced in order to lower the likelihood of overfitting and improve the generalization capabilities of the final model. For the XGBoost algorithm, the so-called decision tree pruning is based on the value of gain, which is defined as follows:

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} + \frac{\left(G_L + G_R\right)^2}{H_L + H_R + \lambda} \right] - \gamma = \frac{1}{2}\left[L + R + N\right] - \gamma\,, \tag{39}$$

where L is the score on the new left leaf, R is the score on the new right leaf and N is the score on the new node, which was the original leaf. The new branch is pruned if the gain is negative.

4.5.8 Artificial neural network

Artificial neural networks are nonlinear statistical models whose structure is inspired by how human brains analyze and process information. A typical graphical representation of an artificial neural network is depicted in Fig. 6, where the units $Z_1, Z_2, \ldots, Z_M$ of the single hidden layer mimic neurons, while connections between neurons of the input, hidden and output layers represent synapses. The output from a single hidden neuron Z_m is a linear combination of the input variables:

$$Z_m = \sigma\left(\alpha_{0m} + \vec{a}_m^T \vec{x}\right), \tag{40}$$

where $m = 1,2,\ldots,M$ is the neuron's ordinal number, σ is the neuron's activation function, α_{0m} is the bias term and $\vec{\alpha}_m$ is the vector of K weights used to multiply each input variable. Consistently, results obtained by means of the model are a linear combination of the outputs determined by the hidden neurons:

$$\hat{f}_h\left(\vec{x}\right) = g\left(\beta_{0h} + \vec{\beta}_h^T \vec{Z}\right), \tag{41}$$

where $h = 1, 2, \ldots, H$ is the output's ordinal number, H is the dimension of the output space, β_{0h} is the bias term and $\vec{\beta}_h$ is the vector of weights comprising M elements. The activation function g is typically an identity function for regression problems.

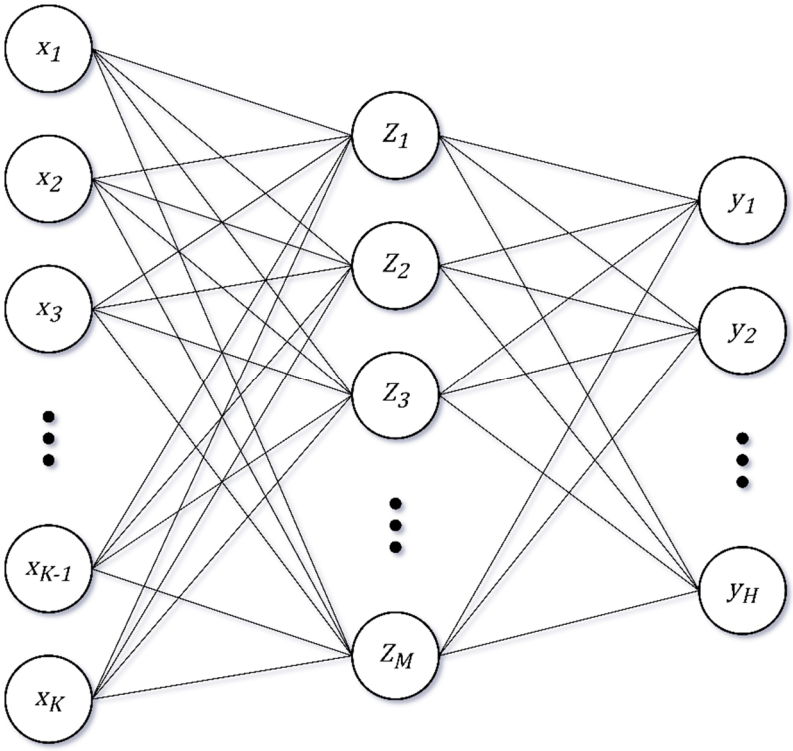


Fig. 6 A feedforward neural network with K input variables, a single hidden layer with M neurons and H output variables.

If the cost function is based on the squared error, the objective function minimized during the training process has the following form:

$$R(\theta) = \sum_{i=1}^{i=N} R_i = \sum_{i=1}^{i=N} \sum_{h=1}^{h=H} \left(y_{ih} - \hat{f}_h(x_i) \right)^2, \tag{42}$$

where θ is the complete set of weights composed of two subsets $\{\alpha_{0m}, \alpha_m; \quad m = 1, 2, \ldots, M\}$ and $\{\beta_{0h}, \beta_h; \quad h = 1, 2, \ldots, H\}$ for the structure presented in Fig. 6. The objective function is nonconvex with many local minima. The predictions $\hat{f}_h(x_i)$ are calculated during a forward pass through the network, while a backward pass is used to compute the errors and propagate them to the hidden layer. The backpropagation algorithm, being a special case of reverse accumulation technique of automatic differentiation, is applied to minimize $R(\theta)$. As already described in the previous chapter, this algorithm applies the chain rule to calculate gradients of $R(\theta)$ with respect to the networks' weights, iterating backward through the layers from the output to the input layer. The adjusted weights at the $t + 1$ iteration are calculated as follows:

$$\beta_{hm}^{(t+1)} = \beta_{hm}^{(t)} - \eta \sum_{i=1}^{i=N} \frac{\partial R_i}{\partial \beta_{hm}^{(t)}}, \tag{43}$$

$$\alpha_{mk}^{(t+1)} = \alpha_{mk}^{(t)} - \eta \sum_{i=1}^{i=N} \frac{\partial R_i}{\partial \alpha_{mk}^{(t)}}, \tag{44}$$

where η is the learning rate and the subscripts denote the ordinal numbers of the nodes connected by a synapse. The simplicity and local nature are the main advantages of this two-pass procedure. Nevertheless, the algorithm can be very slow in the case of multidimensional input data with numerous observations or extensive neural network structure. The state-of-the-art optimization algorithms are based on (stochastic) gradient descent.

The predictive models will be prepared using Python programming language. The most important open-source libraries that are needed for the study are as follows:

- Numpy [176] and Pandas [177] for data manipulation;
- Keras [178] application programming interface for the TensorFlow library will be used to create the structure and train the model based on an artificial neural network;
- Scikit-learn for the creation and training of the remaining models;
- SciPy [179] for solving the optimization problem utilizing the differential evolution method [180] in the case of models based on multiple linear and polynomial regression.

Finally, before starting the training process, the features and the target variable are standardized using the standard score:

$$z_{k,i} = \frac{x_{k,i} - \mu_k}{\sigma_k}, \tag{45}$$

where $k = 1, 2, \ldots, K$ is the features' ordinal number, $i = 1, 2, \ldots, N$ is the observation's ordinal number, $x_{k,i}$ is the raw value, μ_k is the mean value of a particular variable (i.e. either independent or dependent one) and σ_k is the corresponding value of standard deviation.

4.6 Results of the regression analysis

The performance of the data-driven models created to predict the maximal length of fatigue cracks found on the analyzed high-pressure nozzles is described in this section. The following set of information is provided for each data-driven predictive model:

- A list of optimized hyperparameters, including an interval or set of analyzed values.
- A plot presenting the model's performance on the entire train-validation set comprising 25 observations and the unseen data from the test set composed of six elements.
- A concise description of the obtained results.

In order to protect proprietary information, the data presented in the plots were normalized by dividing each value by the maximal measured crack length.

The results determined by the models based on multiple linear and polynomial regression are presented in Fig. 7 and Fig. 8, respectively. The most important considerations regarding the obtained results are as follows:

- It was decided to proceed with a quadratic polynomial, i.e. $p = 2$, since the final cross-validation score was equal to 1.200 regardless of the polynomial's degree, where

$p \in \{2,3,4\}$. Therefore, the simplest expression was selected due to the lack of significant differences. However, since all the possible interaction features are included in the final regression equation, the polynomial can be simplified by removing the least important predictors, e.g. based on an analysis of variance or focusing on the variables

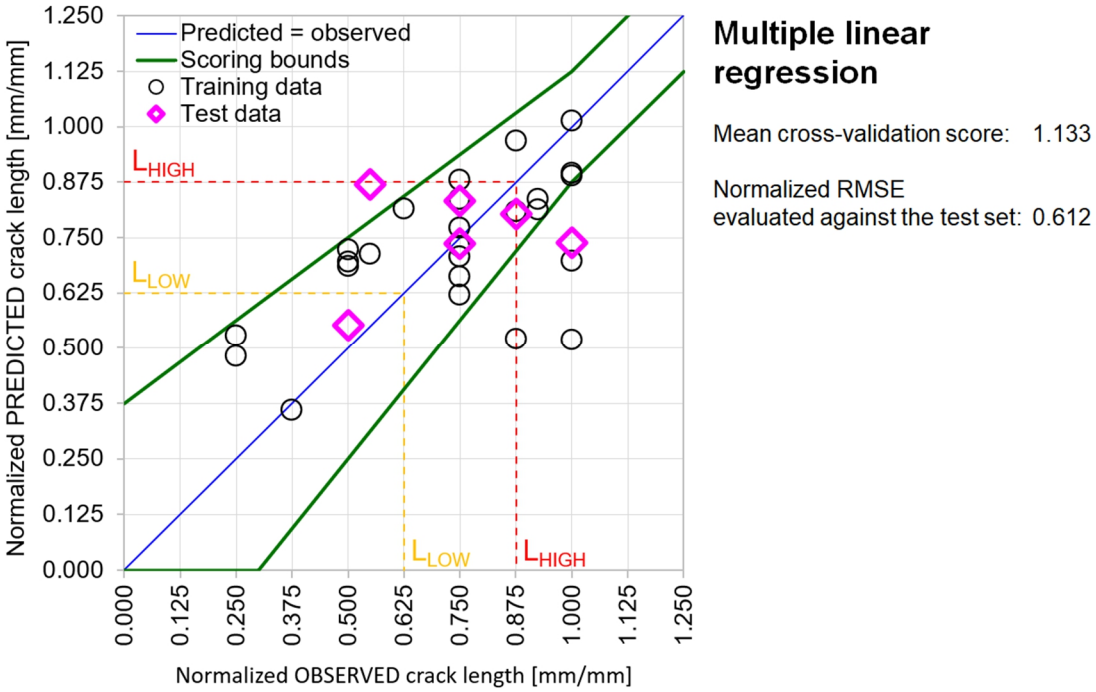


Fig. 7 Evaluation of the multiple linear regression model against the training and test data presented on the predicted vs. observed plot.

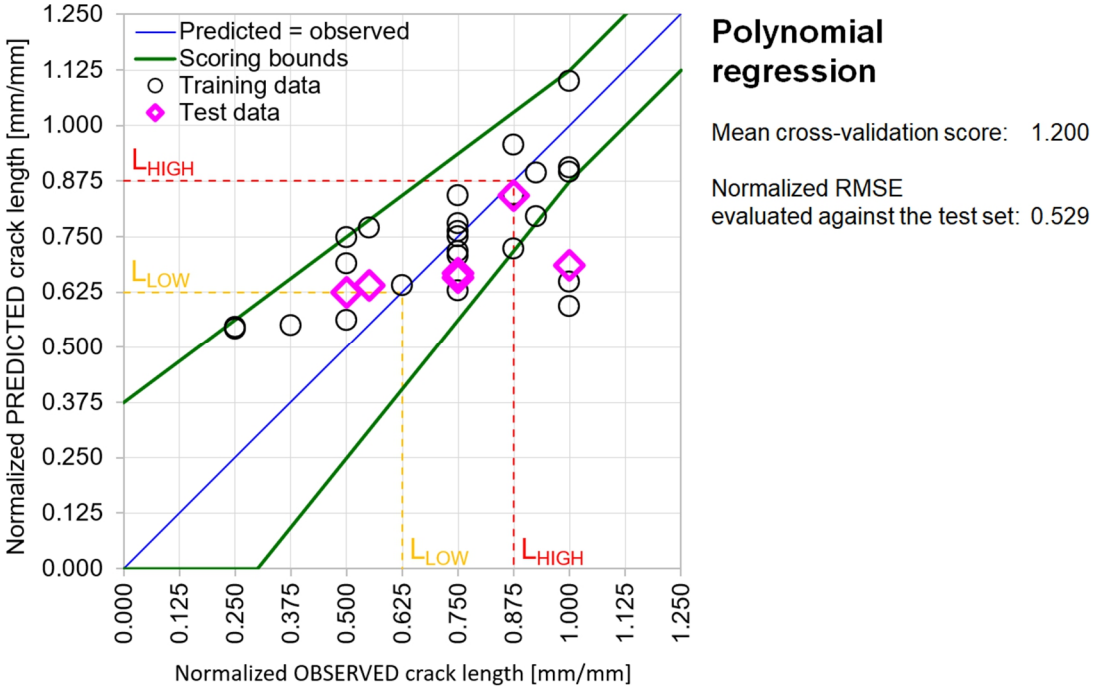


Fig. 8 Evaluation of the polynomial regression model against the training and test data presented on the predicted vs. observed plot.

with the regression equation coefficients close to zero. In the case of multiple linear regression, the final cross-validation score is equal to 1.133.

- Based on an analysis of the regression equation coefficients stored in the $\vec{\beta}$ vector, the median average temperature of gases measured in the exhaust duct $\widetilde{T_{EXH}}$ and the number of service hours with the average temperature of gases at the inlet of the nozzles between the 50th and 70th percentiles $T_{INLET\ P50\text{-}P70}$ are the most important variables in the multiple linear regression model. The product of these features and the squared number of total fired starts are the most significant variables in the polynomial model.
- The normalized root mean squared error (*NRMSE*) evaluated against the test data equals 0.529 for the polynomial model and 0.612 for the multiple linear regression model. Nevertheless, with all the interaction features, the original version of the former has a more complex structure and is less interpretable than the latter. *RMSE* values are normalized by dividing them by the maximal value of *RMSE*, which was obtained for the model based on extreme gradient boosting (i.e. *NRMSE* = 0.529 means that the error evaluated against the test data represents 52.9% of the error obtained in the case of XGBoost-based regression).
- The accuracy of these models is satisfactory only for observations from the "medium" class, where $y \in \left[L_{LOW}, L_{HIGH}\right)$. The models tend to overestimate the shortest cracks, but the estimates are within the scoring interval for most data points. The lowest prediction accuracy concerns the longest observations, with several predictions underestimated and below the lower scoring bound.

The results determined by the models based on support vector and kernel ridge regression are presented in Fig. 9 and Fig. 10, respectively. The list of hyperparameters optimized during the training process is as follows:

- the maximum error $\varepsilon \in [0,1)$;
- the regularization parameters $C, \lambda_2 \in [0.001,100)$;
- the kernel function, the degree of the polynomial kernel $p \in \{2,3,4,5\}$ and the kernel function coefficients $\gamma, r \in [0.001,1)$;
- the tolerance for stopping criterion from the interval $[0,1)$.

Firstly, applying the RandomizedSearchCV class from the scikit-learn library, a predetermined number of combinations of hyperparameter values is drawn and evaluated. The objective is to limit the hyperparameters' space to subspaces, where the cost function takes the maximal values. Thus, it should reduce computational costs associated with the next cross-validation step. After the preliminary filtering, all the remaining combinations are evaluated using the GridSearchCV class. The highest mean cross-validation score is obtained using the polynomial kernel defined as $k\left(\vec{x}_i, \vec{x}\right) = \left(\gamma \vec{x}_i^T \cdot \vec{x} + r\right)^p$. The most important considerations regarding the obtained results are as follows:

- The final cross-validation score of these models is comparable to the model based on multiple linear regression. It equals 1.126 for support vector regression and 1.097 for

kernel ridge regression. The *NRMSE* values are also comparable (i.e. 0.635 and 0.611, respectively).

- An extensive hyperparameters optimization can be executed since training a single model is very short.

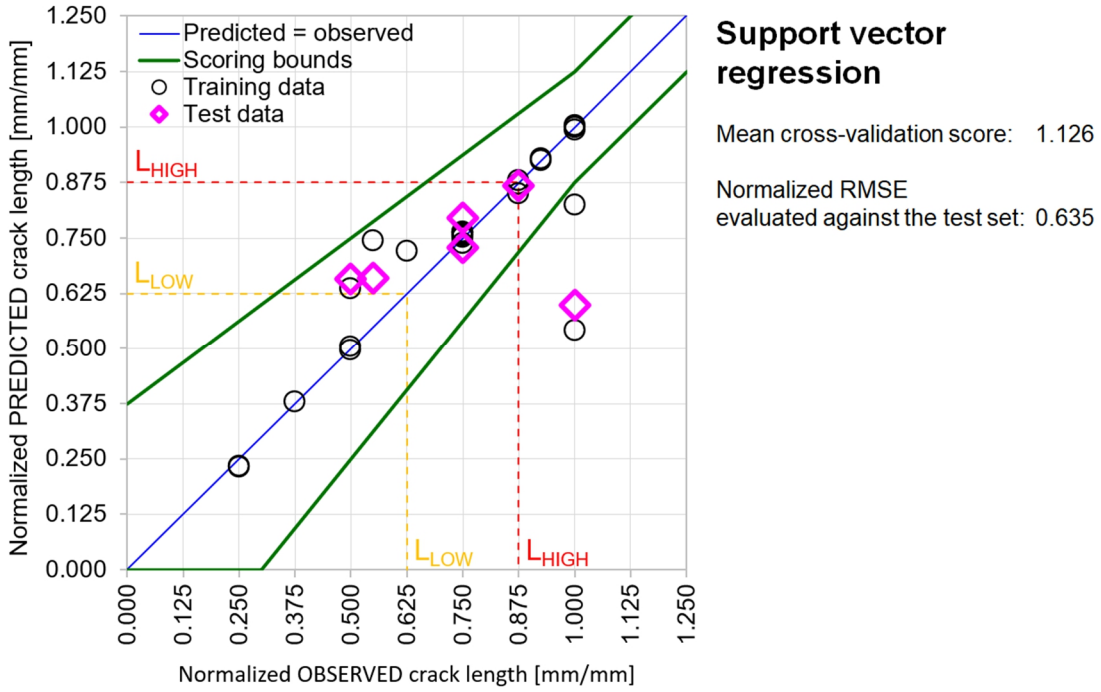


Fig. 9 Evaluation of the support vector regression model against the training and test data presented on the predicted vs. observed plot.

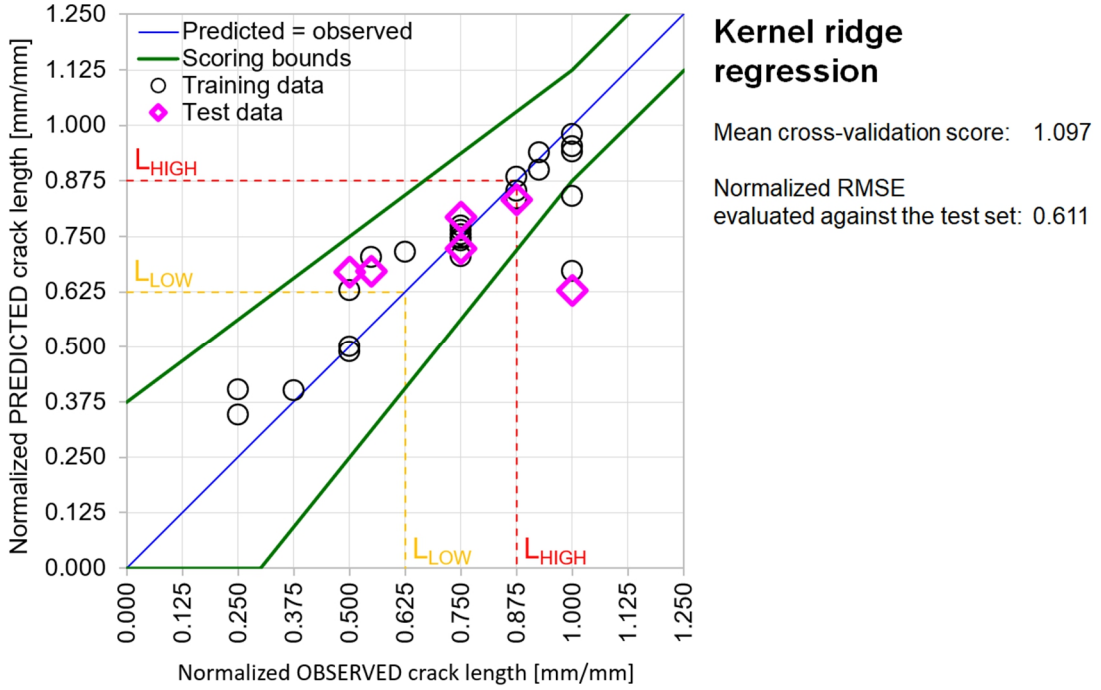


Fig. 10 Evaluation of the kernel ridge regression model against the training and test data presented on the predicted vs. observed plot.

- Estimates determined by these models are in the same range as the observations. The models are capable of capturing the impact of input data variation on the target variable. Nevertheless, despite high accuracy for the observations from the "low" and "medium" classes, these models gave significantly underestimated predictions for several data points from the "long" class.

In the case of decision tree-based methods (i.e. random forest, AdaBoost.R2 and extreme gradient boosting algorithms), there is a set of common hyperparameters optimized during the training process, which comprises the following parameters:

- the maximum number of decision trees M and their maximum depth;
- whether to apply random sampling with replacement or not;
- the number of independent variables utilized during each split;
- the minimum decrease in impurity, or the minimum loss reduction required to make a further partition of a node γ;
- the minimum number of samples or the instance weight required to split a node, or the minimum number of observations remaining on the leaf after the split;
- the learning rate η, in the case of AdaBoost.R2 and XGBoost algorithms.

Concerning these parameters, the analyzed intervals were not unified for these three models. The additional hyperparameters considered in the case of XGBoost algorithm are as follows:

- the booster type, which is either gbtree or dart;
- the $L1$ and $L2$ regularization parameters $\lambda_1 \in [0,5]$ and $\lambda_2 \in [0,500]$;
- the percentage of decision trees ignored, i.e. dropped out, during each boosting step $d_{rate} \in \{0.2, 0.4, 0.6, 0.8\}$, which applies to the dart booster only to reduce overfitting;
- the probability of skipping the dropout procedure $d_{skip} \in \{0, 0.2, 0.4, 0.6\}$.

The results determined by these models are presented in Fig. 11, Fig. 12 and Fig. 13. The most important considerations regarding the obtained results are as follows:

- The median average temperature of gases measured in the exhaust duct $\widetilde{T_{EXH}}$ is the most important independent variable in these models. The Gini importance (i.e. the mean decrease in impurity) for random forest and AdaBoost.R2 regression and the values of gain for each decision tree in the case of XGBoost are used to calculate the importance.

- The normalized root mean squared error evaluated against the test data equals 0.544 in the case of random forest regression and 0.500 for the model based on the AdaBoost.R2. The latter is the lowest value obtained during the analysis and can be considered satisfactory. However, these models may have worse generalization capabilities than the remaining ones since the obtained final cross-validation scores are one of the lowest, i.e. 0.962 and 0.980, respectively.

- The model built using the XGBoost algorithm achieved a better score during the cross-validation, i.e. equal to 1.096, but it has the highest error when evaluated against the test set, *NRMSE* = 1.000, due to the significantly underestimated prediction for the longest observation in the test subset (Fig. 13). Efforts aimed at reducing the error related to that data point by tuning of hyperparameters of the model result in substantial lowering of the average cross-validation score. Additionally, even though several regularization

parameters have nonnegative values, it is visible that the model's accuracy evaluated against the training set is exceptionally high, indicating possible overfitting. However, attempts to apply stronger regularization or reduce the number of decision trees have a negative impact on the final CV score and do not reduce the normalized error.

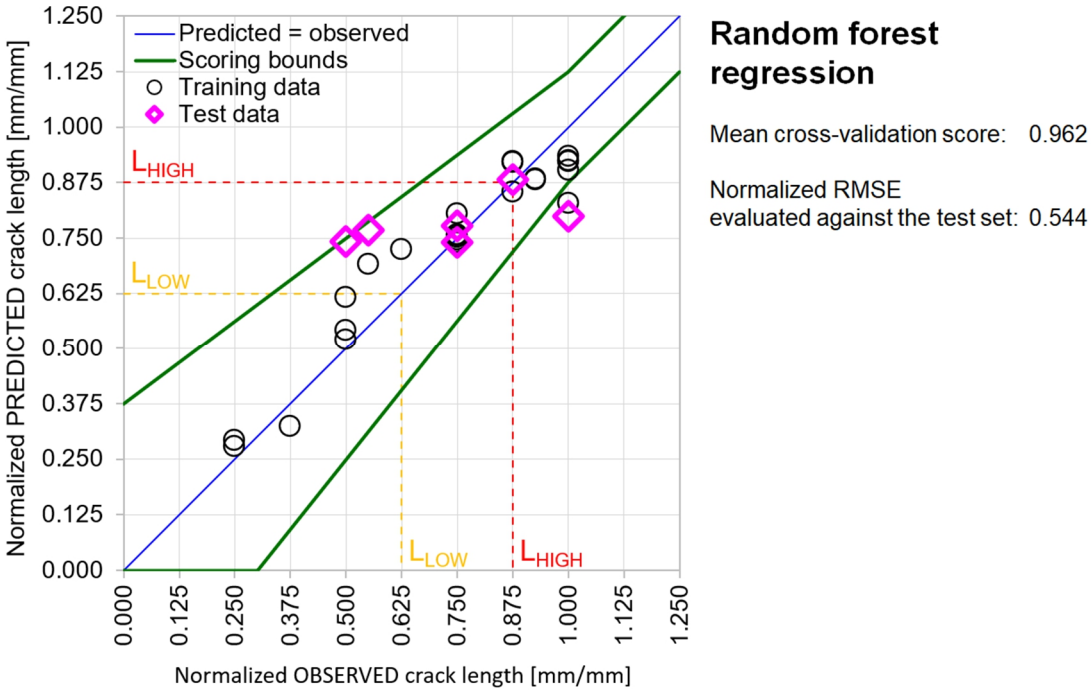


Fig. 11 Evaluation of the random forest regression model against the training and test data presented on the predicted vs. observed plot.

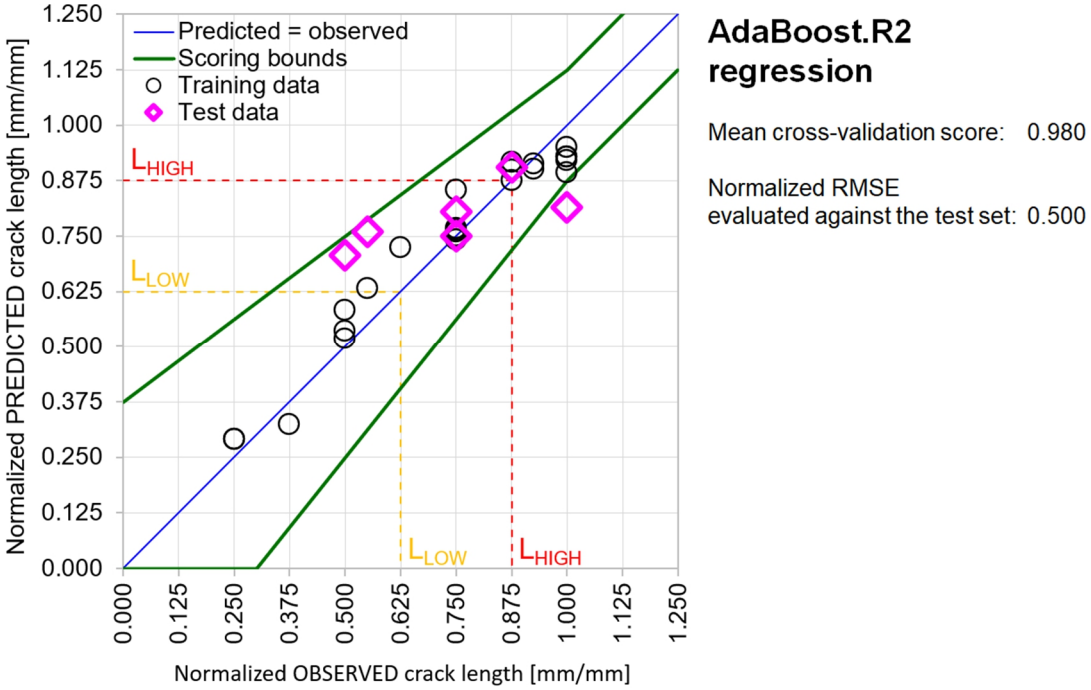


Fig. 12 Evaluation of the AdaBoost.R2 regression model against the training and test data presented on the predicted vs. observed plot.

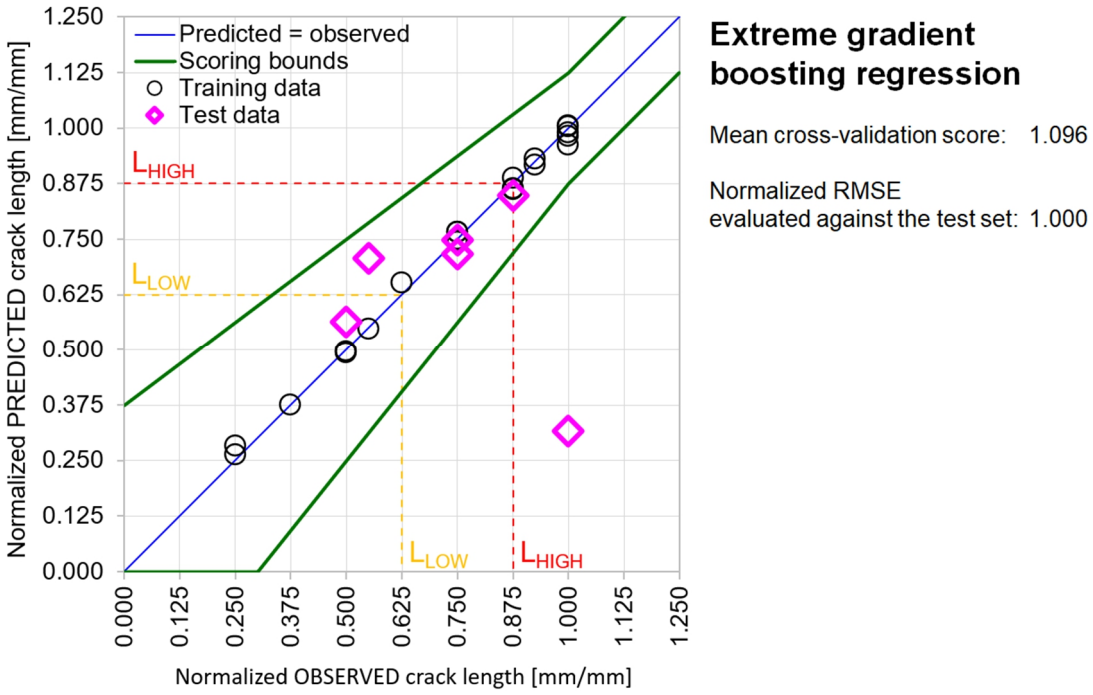


Fig. 13 Evaluation of the extreme gradient boosting regression model against the training and test data presented on the predicted vs. observed plot.

- Compared to the previously described algorithms, tuning of hyperparameters takes more time since these tree-based models have more parameters controlled by the modeler.
- Estimates determined by these models are in the same range as the observations. The models are capable of capturing the impact of input data variation on the target variable. Nevertheless, despite high accuracy for the observations from the "low" and "medium" classes, these models tend to underestimate the longest cracks and have one of the worst cross-validation scores, indicating potentially lower generalization capabilities.

For the model based on a feedforward neural network, the list of hyperparameters tuned during the training process is as follows:

- the number of hidden layers of the network from the set $\{2,3,4,5,6,8,10\}$ and the number of neurons in the input and hidden layers $M_{input}, M \in \{5,10,15,20\}$;
- the optimization algorithm from the set {Adadelta, Adam, Adamax, Nadam}, the learning rate $\eta \in [0.0001, 0.1]$ and the number of training epochs (i.e. an epoch refers to one complete pass of the algorithm through the data) from the interval $[100,1300]$;
- the activation function for the hidden layers $\sigma \in$ {relu, exponential, hard sigmoid, sigmoid, sofplus, tanh};
- the initializer used to determine the initial weights of the model from the set {he_normal, he_uniform, GlorotNormal, GlorotUniform, lecun_normal, lecun_uniform};
- the number of training samples presented to the neural network before each update of the weights from the interval $[4,10]$;

- the percentage of hidden neurons dropped out during each update of the weights $d_{rate} \in \{0, 0.2, 0.25, 0.3, 0.4\}$;
- the maximum norm of the weights vector $\|\vec{a}_m\|, \|\vec{\beta}_h\| \in \{3, 4, 5\}$;
- in the case of Adamax optimizer, the exponential decay rates for the 1st and 2nd moment estimates from the interval $[0, 0.999]$.

The highest averaged cross-validation score equal to 0.910 is obtained with an architecture comprising three hidden layers with 15 neurons each, the weights initialized using the Lecun initializer that draws them from the uniform distribution $U\left(-\sqrt{3/M_{input}}, \sqrt{3/M_{input}}\right)$, the rectifier linear unit (ReLU) activation function defined as $\sigma(x) = \max(0, x)$ and the Adamax optimization algorithm [181]. The results obtained by means of this model are presented in Fig. 14. The most important considerations regarding the obtained results are as follows:

- The normalized root mean squared error evaluated against the test data equals 0.519, the lowest of all the prepared models. However, simultaneously it has the worst final cross-validation score, indicating potentially lower generalization capabilities.
- The importance of the independent variables is unknown since artificial neural networks are so-called black boxes whose internal operations are difficult to be interpreted.
- Due to the number of parameters configured by the modeler and the time needed to train a single model, an extensive optimization of the hyperparameters requires much time.
- Estimates determined by this model are in the same range as the observations, and the model can capture the impact of input data variation on the target variable. Nevertheless,

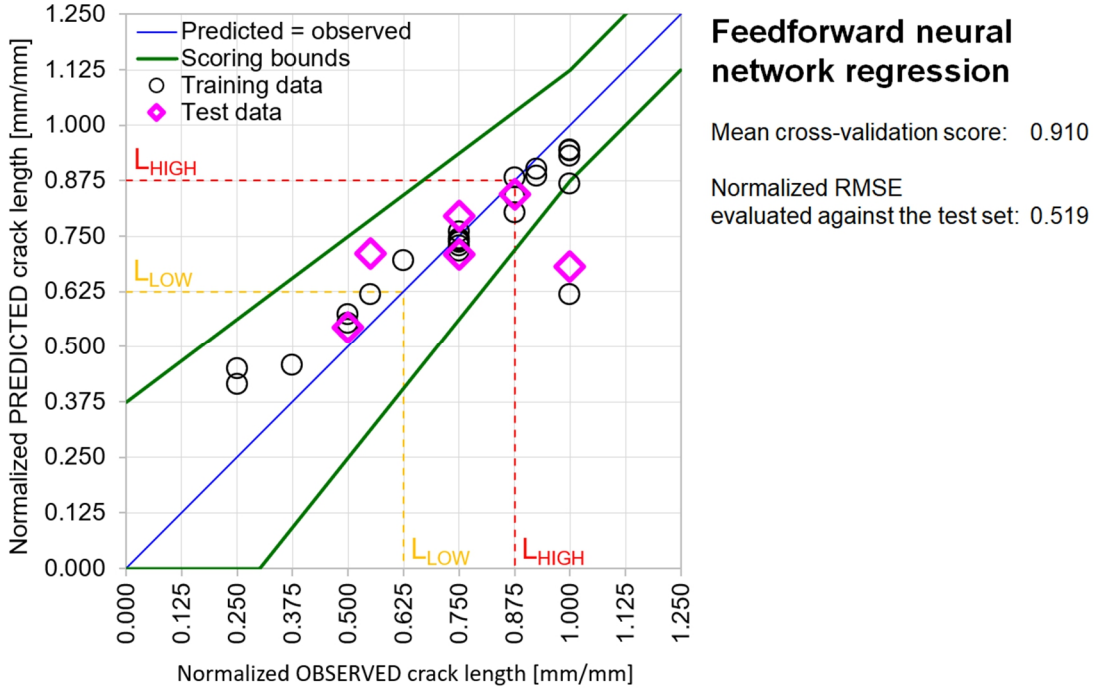


Fig. 14 Evaluation of the feedforward neural network regression model against the training and test data presented on the predicted vs. observed plot.

as for the previously described models, the outputs corresponding to the longest cracks are underestimated.

The results of the regression analysis are summarized in Table 1.

Table 1 A summary of metrics describing the models based on machine learning algorithms created to predict the length of fatigue cracks found on the high-pressure nozzles.

Algorithm type	Averaged cross-validation score	Normalized RMSE evaluated against the test set
Polynomial regression	1.200	0.529
Multiple linear regression	1.133	0.612
Support vector regression	1.126	0.635
Kernel ridge regression	1.097	0.611
XGBoost regression	1.096	1.000
AdaBoost.R2 regression	0.980	0.500
Random forest regression	0.962	0.544
Feedforward neural network regression	0.910	0.519

4.7 Conclusions concerning this stage of the research

The purpose of the first stage of this research was to apply several machine learning algorithms on a small sample comprising 31 observations and assess whether the objectives of this research, described in Section 2.1, can be achieved through these methods. The primary outcomes of this study are as follows:

- Five independent variables, essential in modeling the maximal length of fatigue cracks recorded on the analyzed 1st stage nozzles, were identified during the selection and extraction of features. The operating parameters used to create the models are related to gas temperatures at the gas turbine's inlet, exhaust and inlet of the nozzles, the engine's output and the number of total fired starts. The median average temperature of gases measured in the exhaust duct $\widetilde{T_{EXH}}$ was the most important predictor in the case of polynomial and multiple linear regression and for the tree-based models.
- Based on the metrics used to evaluate the models' performance (i.e. the averaged cross-validation score and the normalized RMSE evaluated against the test observations), the polynomial model is characterized by the lowest error and should generalize the best. However, this model does not capture sufficiently precisely the impact of input data variation on the dependent variable, tending to overestimate the shortest cracks and underestimate some of the longest observations. The latter is a common drawback of all the models created, but it applies only to single data points from the "long" class. The model based on the AdaBoost.R2 algorithm can be considered an alternative. However, before considering it a valid source of data-driven insights to support decisions, it should

be subjected to additional validation to evaluate mainly its generalization capabilities. For example, the predictions can be compared with the results obtained using a different approach, e.g. a finite element analysis. Additionally, a staggered deployment linked with intermediate visual inspections of the hardware, e.g. by means of a borescope, acquisition of new measurement data and cyclic updates of the model are suggested.

- It was proved that these machine learning algorithms could be applied at the early stages of the component's lifetime when the availability of damage measurements is limited, and the number of observations is small. It was confirmed that the length of analyzed fatigue cracks could be predicted accurately using some of these methods. Additionally, the applied algorithms are universal and can be utilized for modeling damage due to various failure types found on the main parts of gas turbines. However, even though the predictors reflect well knowledge about the cause of the analyzed phenomenon, the predictive models are not aligned with any fatigue crack growth equation. In the worst-case scenario, it may happen that a model characterized by high prediction accuracy is not entirely consistent with the underlying physical laws. In practice, it eliminates the possibility of using such a model to extrapolate and generates doubts regarding the interpolation capabilities in the entire training domain.
- Because of the limited number of observations, the training, validation and test subsets were not drawn randomly. These sets were prepared in a controlled manner considering multiple constraints resulting from the analysis of the composition and clusters of the sample. As a result, each set represents the entire dataset in a quantitative and qualitative way. This approach was applied to improve the training effectiveness, make the cross-validation more meaningful and simplify the evaluation and interpretation of results.
- Customized loss and cost functions were created and used during the training process. Compared to the (root) mean squared error, the utilized cost function more directly reflects the particular objective to accurately predict the longest cracks, which may jeopardize the part's integrity if not appropriately maintained. Therefore, according to the definition of the loss function, the prediction accuracy should rise as the crack lengths increase, which is represented by the variable-width scoring interval defined by two scoring bounds. Additionally, accurate predictions of the cracks classified as "long" are rewarded with a bonus, promoting models with specific demanded characteristics.

During this part of the research, it was proved that the models based on machine learning methods can provide accurate predictions, even if the available sample is small. Despite several advantages, those data-driven models have restricted extrapolation capabilities, mainly because of the lack of direct representation of the underlying physical laws in the model. These drawbacks are addressed in the second stage of this research.

5. Predicting fatigue cracks growth by physics-informed neural networks

As described in Section 3.4, physics-informed neural networks can be considered a method of unsupervised learning that assures consistency with physical laws describing the analyzed problem, which are embedded into the cost function, even when the available number of observations is limited. Many successful applications of such networks, combining data with underlying theoretical equations, were referenced. The main benefits resulting from the use of physics-informed neural networks are as follows:

- A PINN can be trained effectively if the training subset comprises only a few elements or even when no observations exist.
- Prior knowledge about the analyzed phenomenon is integrated into the neural network's structure and utilized to regularize the training process and limit the output space only to solutions consistent with the embedded physical laws. Thus, compared to purely data-driven models, PINNs should be more robust in the small data regime and have better generalization and extrapolation capabilities.
- A validated physics-informed neural network can be used as a source model adjusted in the target domain. Otherwise, if the governing equation is the same in both domains, applying the source model directly in the target domain is possible. Therefore, from this perspective, PINNs can be considered a means to transfer knowledge between domains.
- PINNs are universal function approximators applied to problems described mainly by partial or ordinary differential equations. Regardless of the mathematical description's complexity, a physics-informed neural network can be built and trained using open-source software and popular, easy-to-use libraries.

Considering the limitations of the predictive models created during the first stage of the research and the advantages of PINNs, it is decided to apply physics-informed neural networks to the analyzed problem of fatigue crack length prediction. The dependent variable remains the same: it is the maximal crack size at Position 2 on the nozzle's trailing edge. Nevertheless, the sample is reduced again and limited only to Type C gas turbines, for which numerical analysis results are available. Thus, the sample comprises 19 observations, for which 584 434 records with the operating parameters are available. The set of operational data is again incomplete. Fig. 15 presents the values of the coverage, defined as the percentage of startup-shutdown cycles for which at least one complete record is available. Its values range from 10% to 100%, with the mean coverage equal to only 54%. The highest absolute value of cycles with the operational data available equals 48 (~81% of 59 interval fired starts) for the nozzle set with id. S1N_5. The approach applied for the data imputation is described in the next section. The operating parameters utilized to prepare the input of the physics-informed neural network are as follows:

- the ambient air temperature T_{AMB},
- the average air temperature at the axial compressor's discharge T_{COOL},
- the average temperature of gases at the outlet of the nozzles T_{FIRE}, which is calculated based on other signals the data acquisition system recorded.

Moreover, the fuel stroke reference, the mean air pressure at the axial compressor's discharge, the high-pressure shaft's rotational speed and the gas temperatures measured in the exhaust duct are used to identify startups and shutdowns of the analyzed gas turbines.

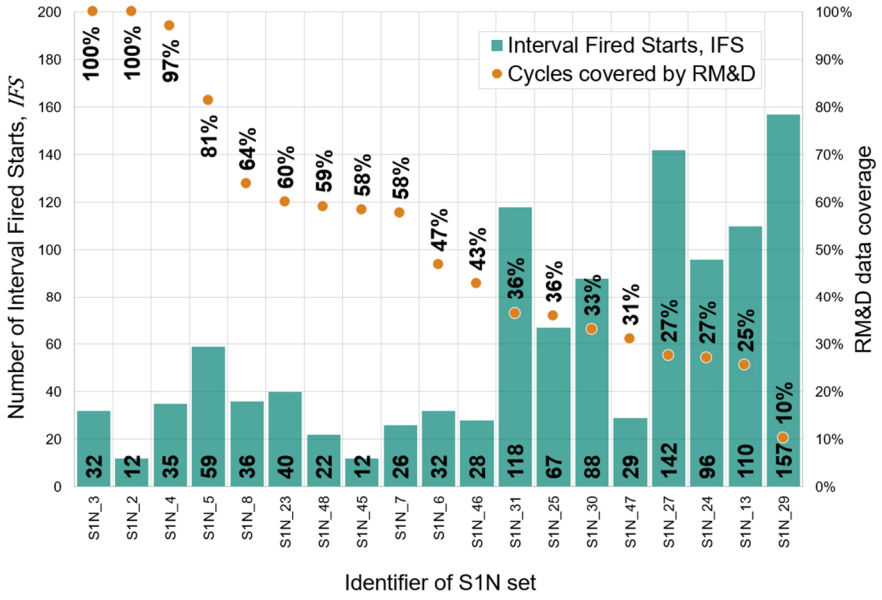


Fig. 15 The number of fired starts accumulated since the last repair *IFS* for each nozzle set and the percentage of startup-shutdown cycles, for which at least one complete record with the operational data is available.

Details of this research stage are provided in the following sections of this chapter. At the outset, it should be highlighted that a significant part of the conducted research and obtained results were described in [110].

5.1 The architecture of the hybrid model

In [182] the authors applied a PINN based on a recurrent neural network for modeling the propagation of fatigue cracks in a corrosive environment. The model's structure was simple, but the obtained estimates were accurate despite this. Consequently, based on these promising results, it is decided to apply a similar architecture based on a recurrent neural network. RNNs are suitable for processing variable-length time series, allowing to use as the input at the current time step t, the output of the previous time step t - 1. Such a layer can be presented as a for loop that iterates through the series of data points while keeping information from the previous time step in the internal state. In the case of progressive, irreversible damage accumulation, for example, caused by fatigue, oxidation, creep, wear, corrosion, or erosion, a recurrent neural network is a logical choice for building cumulative damage models. Equations that describe the damage increment during a single cycle or time step can be applied multiple times for all records in the sequence. Consequently, the total damage to a component at time t is represented as the sum of damage increments until that moment.

The equations describing the damage increment, which are utilized to process the data during each time step, are embedded inside the recurrent neural network's cell. Fig. 16 presents the custom RNN cell designed to approximate the solution to the analyzed problem. The cell comprises the following major elements:

- The data-driven layer is based on a feedforward neural network estimating values of the stress intensity factor range at engine shutdown ΔK_{shdn} and the stress intensity factor during steady-state operation K_{ss}. The multilayer perceptron takes σ_{shdn} as the input,

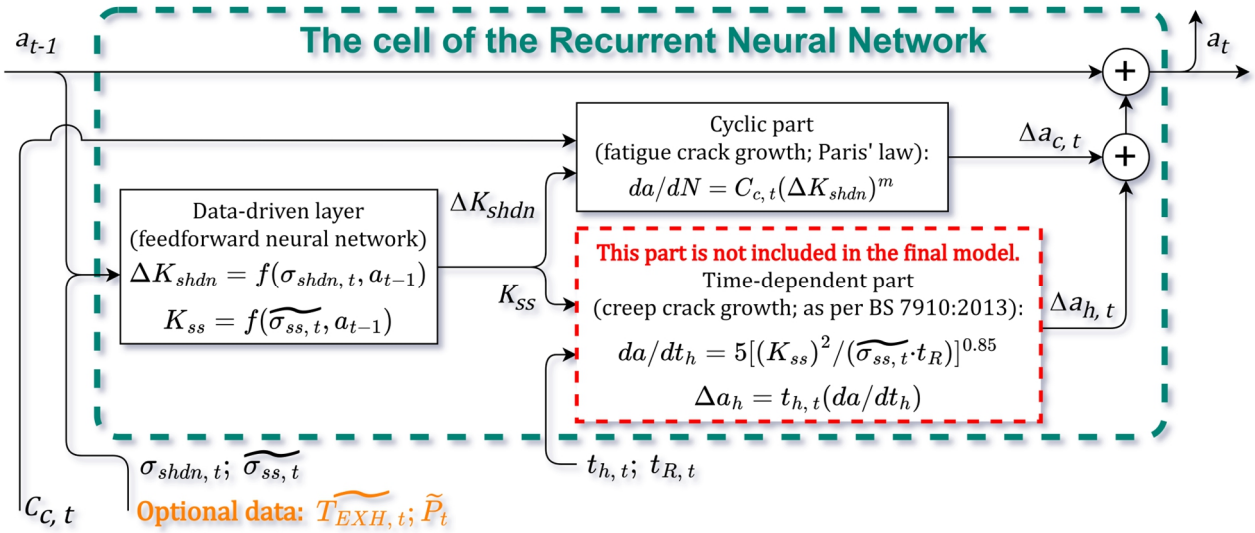


Fig. 16 The custom cell of the recurrent neural network applied to estimate crack length increments.

which is the value of maximum thermal stresses at each shutdown, then the median of thermal stresses at steady-state operation conditions for each startup-shutdown cycle $\widetilde{\sigma}_{ss}$, and the estimated crack length at the previous time step a_{t-1}. Simultaneously, it is assumed that for each cycle, the minimum value of thermal stresses σ_{MIN} equals zero. Consequently, the corresponding value of the stress intensity factor $K_{MIN} = 0$ and the stress intensity factor range at shutdown can be evaluated as follows:

$$\Delta K_{shdn} = K\left(\sigma_{shdn}, a\right) - K_{MIN} = K\left(\sigma_{shdn}, a\right) = K_{\sigma_{shdn}} . \tag{46}$$

- The cyclic part is based on the Paris' law [183], which is used to calculate the crack length increment Δa_c during each startup-shutdown cycle using the following formula:

$$\frac{da}{dN} = \Delta a_c = C_c \left(\Delta K_{shdn}\right)^m , \tag{47}$$

where C_c is the linear function of $\overline{T}_{S,ss}$, which is the mean temperature of the trailing edge tip at steady-state operating conditions and m has a constant value. The results of fatigue tests conducted at Baker Hughes Company were used to determine the formulas utilized to estimate these material parameters.

- The time-dependent part, which is used to calculate the crack length increment due to creep Δa_h following the British Standard BS 7910:2013+A1:2015 [184]:

$$\frac{da}{dt_h} = 5\left[\frac{\left(K_{ss}\right)^2}{\widetilde{\sigma}_{ss} t_R}\right]^{0.85} , \tag{48}$$

$$\Delta a_h = \frac{da}{dt_h} t_h , \tag{49}$$

where t_R is the time to rupture (i.e. at $\widetilde{\sigma_{ss}}$ and $\overline{T}_{S,ss}$) estimated by means of the Larson-Miller parameter and t_h is the duration of a startup-shutdown cycle expressed in fired hours. The standard does not provide a specific formula for FSX-414 or other cobalt-based superalloys. Therefore, Eq. (48) is a universal formula applicable to various materials and temperature ranges.

Predicting accurately and effectively the initiation and propagation of cracks in areas of stress concentrations is one of the main objectives of fracture mechanics. This problem is of primary importance to ensure the reliable operation of structural elements and machines under multiaxial loading, including various components of gas turbines. In 1957, Irwin [185] introduced the stress intensity factor parameter, which characterizes the stress field near the tip of a tensile fracture of a linear elastic material. Since then, the stress intensity factor has become one of the most essential parameters for crack propagation analysis. Many different methods were proposed to determine the parameter's value. The propagation of cracks with singular concentrations at their tip can be estimated by applying, e.g. the strain energy density criterion [186], the crack tip opening displacement approach [187], or the criterion of critical value of the elastic energy release. At the regular areas of stress concentrations, like holes, grooves, or notches, the material strength is typically estimated based on the local stress conditions. In [188,189] the authors proposed a non-local stress criterion for determining the initiation and propagation of cracks, which can be applied to regular and singular stress concentrations. This brittle failure criterion is represented by a non-local measure of stress intensity over the finite damage zone, which the material grain size can specify. Since this condition can be employed to determine the crack initiation, the growth rate, or propagation direction, it can significantly simplify the theoretical description of the problem under analysis. Moreover, in the case of small stress gradients, this non-local condition can be converted to local stress criteria or, in the case of singular stress distributions, to energy criteria. The authors successfully applied the proposed condition, considering that the analyzed object was a plate subjected to tension with a wedge-shaped notch, an elliptical hole, or hyperbolic notches. In the case of multiaxial fatigue loading, the non-local stress criterion is applied together with a local failure function, which is a homogeneous function of non-dimensional values of the normal and shear stresses on the physical plane. In this setup, the same type of non-local condition is utilized to generate the initiation and propagation rules for cyclic loading. The crack growth condition due to low cycle tensile stresses was generated based on the crack initiation condition. Moreover, as emphasized in the paper [189], the former condition has a form similar to the Paris' law, i.e. Eq. (47), which is used to determine the crack length increment Δa_c during each startup-shutdown cycle in the recurrent neural network presented in Fig. 16. However, an important reason that contributed to the decision to choose the Paris' law relates to the availability of fatigue crack growth rates of FSX-414 cobalt-based superalloy. The raw data were recorded during fatigue tests performed at two distinct material temperatures and subsequently have been approximated and described by means of a power function. Only the final, post-processed values of the stress intensity factor range and the corresponding crack growth rates are available. Therefore, these data can be directly used to determine the material-specific coefficients m and C_c of the Paris' equation, where the latter depends on the nozzle's trailing edge tip temperature at steady-state operating conditions.

Conditions of crack initiation and propagation, which were established 50 years ago or more, are a point of reference often used by today's researchers. In [190] the authors applied the strain

energy density criterion and the theory of critical distances, which is concisely described in [191], combining them with the equivalent material concept. In accordance with this concept introduced in [192], a virtual brittle linear elastic material and brittle fracture conditions can be applied to analyze fractures of ductile materials showing elasto-plastic behavior. Therefore, by employing the equivalent material concept, the strain energy density criterion and the theory of critical distances can be applied to objects with nonlinear elastic behavior. According to [190], such a combined approach provides accurate predictions of fracture loads for large radii of the notch, but it underestimates them when the radii are small. Nevertheless, a significant drawback of purely theoretical methods for predicting the stress intensity factor is their limited applicability related to several examples of crack types and loadings, which do not reflect the wide range of cases observed in engineering practice. Similar limitations characterize traditional experimental methods, which are insufficient for components with complex shapes and multiaxial loadings. In such situations, a numerical simulation can be run and solved by employing the boundary element method or finite element method. Application of the latter is the most popular approach. However, special techniques and a properly created mesh in the crack tip vicinity are required to estimate the stress intensity factor accurately. Increasing the mesh density makes the simulation more computationally expensive and time-consuming. Thus, multiple variations of the finite element method for estimating the stress intensity factor have been proposed. A concise overview of these methods is given in [193]. As with many other scientific and engineering challenges, applying machine learning algorithms to determine the stress intensity factor is an alternative to computationally expensive numerical simulations. These data-driven approaches may require substantial time to create, tune, train and test the predictive model. However, a fully-specified model does not require much computational power to be applied and provides results significantly faster than a numerical simulation. Numerous studies on the prediction of stress intensity factors by means of machine learning methods are available, including the use of multiple linear regression [194], support vector regression [195], extreme gradient boosting [196], artificial neural networks [197], or deep convolutional neural networks [198]. Models based on such algorithms allow for continuous estimation of the stress intensity factor using the gas turbine operational data or digital images of the analyzed object and location. As already described, for the analyzed problem of fatigue cracks observed on the high-pressure nozzles, the stress intensity factor is estimated using the feedforward neural network, whose input data are calculated based on the selected operating parameters of gas turbines. The structure of this multilayer perceptron, being the data-driven layer, is as follows:

- A data normalization layer, where the min-max normalization is applied. The predictions' accuracy was lower when the input data was standardized using the standard score instead of normalized.
- The input layer comprises 16 neurons with the Scaled Exponential Linear Unit (SELU) activation function and the lecun_normal initialization function. Reducing the number of neurons to eight, four, or even two results in longer cost function optimization. The activation function returns λx if $x \geq 0$ or $\lambda\alpha\left(e^{x}-1\right)$ if $x < 0$, where λ = 1.05070098 and α = 1.67326324. Attempts to use the hyperbolic tangent activation function reduce the effectiveness of the training process and extend its duration. The initial weights are drawn from a truncated normal distribution centered on zero with a standard deviation $\sigma_{ln} = \sqrt{1 - fan\_in}$, where fan_in is the number of input units in the weight tensor.

- Two hidden layers comprising nine neurons, with the SELU activation function and the lecun_normal initialization function. During the optimization of the neural network's structure, the numbers of hidden layers and the corresponding neurons were drawn from sets {1, 2, 4, 8} and {2, 3, 5, 9}, respectively. Considering the accuracy of the predictive model and the time required to execute the learning process, the selected configuration is the optimal one.
- The output layer comprises one neuron with the rectifier linear unit activation function to ensure that the stress intensity factor estimates are nonnegative.

The feedforward neural network requires information about the stress and actual crack length to estimate the stress intensity factor. However, based on the feature selection process described in Chapter 4, also the following variables were previously chosen as the predictors:

- median gas turbine's output $\widetilde{P}$,
- median average temperature of gases measured in the exhaust duct $\widetilde{T_{EXH}}$.

The latter was the most important independent variable in the case of polynomial and multiple linear regression and for the tree-based data-driven models. Although these parameters should not affect the stress intensity factor, an attempt was made to consider them as inputs of the feedforward neural network. The objective was to check if their inclusion could increase the accuracy of the final estimates determined by the physics-informed neural network. Unfortunately, the prediction error did not decrease while the model's complexity increased. Therefore, these two parameters are not included in the input data set for the final, fully-specified model. An alternative approach is to include in the RNN's cell an additional block to correct the crack length estimates based on values of $\widetilde{P}$ and $\widetilde{T_{EXH}}$, i.e.:

$$a_t = a_{t-1} + \Delta a_c + \Delta a_{corrector} = a_{t-1} + \Delta a_c + f\left(\Delta a_c, \tilde{P}, \widetilde{T_{EXH}}\right), \tag{50}$$

where $\Delta a_{corrector}$ is the correction applied to the crack length increment estimation Δa_c. The set of values the correction function may take should be constrained to keep the model interpretable. Physics-based equations embedded in the cost function can be used to restrict the space of allowable solutions. Additionally, the time-dependent part is excluded from the RNN's cell. With the time-dependent part included, it was impossible to simultaneously reduce the error related to the theoretical equation embedded in the cost function and the error associated with the empirical data. The resultant models estimated high crack length increments due to creep, but the contribution of fatigue was negligible due to low estimates of the stress intensity factor (i.e. compared to models with the cyclic part only). Consequently, those models did not reflect prior knowledge about the analyzed problem and provided very inaccurate estimates of the final crack lengths. Thus, the RNN's cell of the final model comprises only the data-driven layer and the cyclic part. It should be highlighted that no element is intended to model the crack initiation process in the structure. However, on the basis of the available empirical data, the analyzed cracks should be visible after the first few startup-shutdown cycles. Therefore, such a simplification should not affect the model's accuracy significantly. The initial crack length a_0 is assumed in accordance with the company's standard. It has to be explained that the analyzed component is subjected to a non-destructive penetrant inspection after manufacture and each repair. Therefore, the initial length of the analyzed edge crack is approximately equal to the length of the smallest indication that can be detected using a fluorescent penetrant inspection.

Finally, with the architecture of the hybrid model determined, it has to be explained how the thermal stresses at shutdown σ_{shdn} are estimated based on the gas turbine operational data, i.e. T_{AMB}, T_{COOL} and T_{FIRE}. During engine shutdowns the nozzle's trailing edge cools down faster than the inner and outer platforms. These changes in the base material temperature create tensile thermal stresses and open the analyzed cracks. Concerning the analyzed Position 2, the tensile stresses in the direction perpendicular to the crack propagation plane are nearly five times higher at the emergency shutdown from full load in comparison with the stresses at the steady-state operating conditions, as shown in Fig. 17. It confirms that the analyzed cracks are mainly caused by low cycle fatigue. In addition, it should be clarified that the utilized numerical simulation dedicated to the high-pressure nozzles of Type C gas turbines was not prepared specifically for this research, as it was prepared a few years earlier for a different New Product Development program. A particular limitation resulting from this decision is that the transient analysis was executed assuming only an emergency shutdown from the full load, and no results

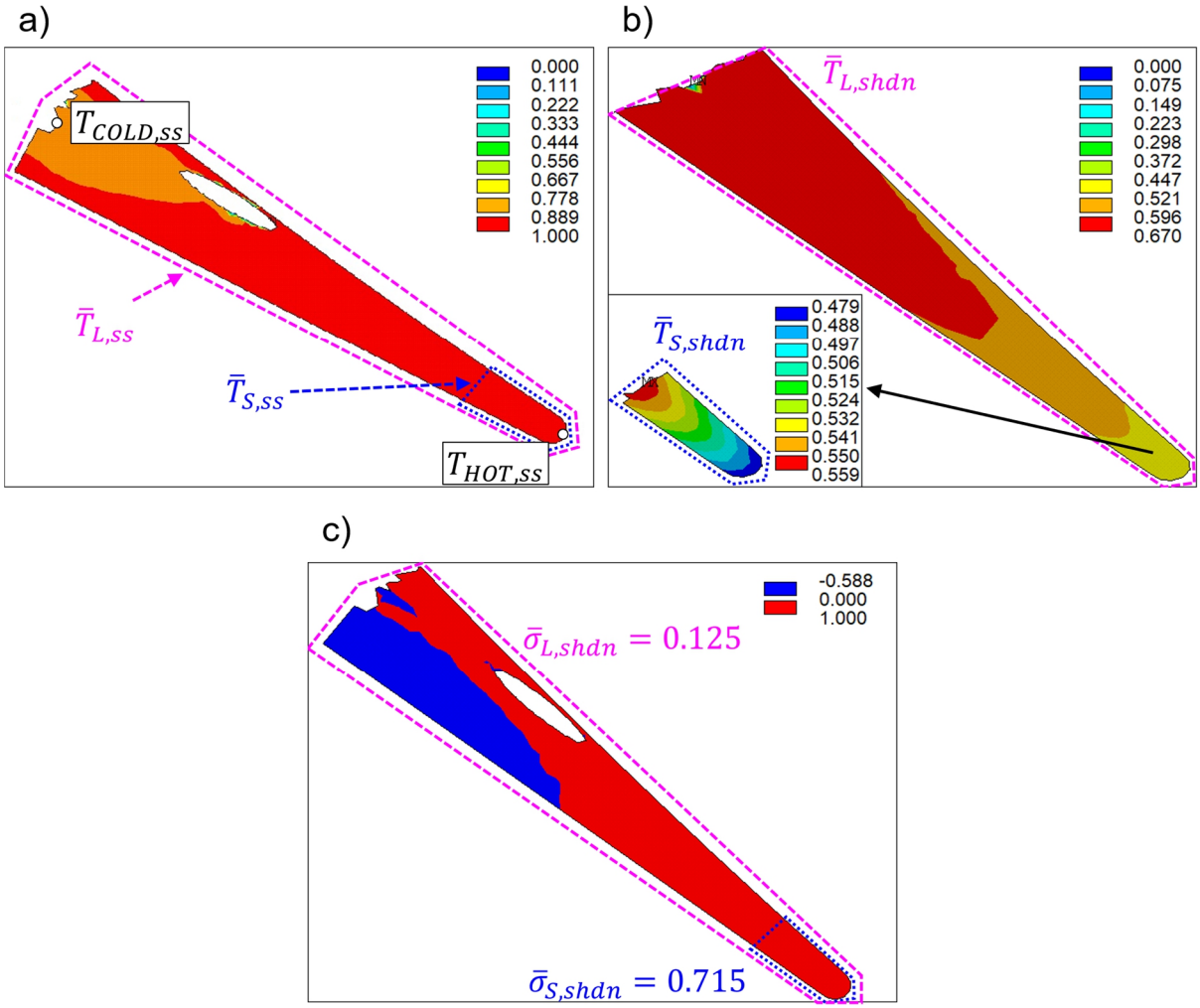


Fig. 17 The numerical simulation results for Position 2. Distribution of the base material temperature at steady-state operating conditions (a) and emergency shutdown from the full load (b) for a time step, when the thermal stresses take the maximum value. The temperatures were normalized by dividing each value by $T_{HOT,ss}$. (c) Thermal stress field in the direction perpendicular to the crack propagation plane at emergency shutdown from the full load (tensile stresses are marked in red and compressive stresses are in blue). The stresses were normalized by dividing each value by the maximum value of the thermal stresses.

are available for a normal shutdown mission. Consequently, the temperature and stress fields corresponding to the emergency shutdown from the full load are assumed for each cycle that the analyzed parts accumulated, even though a substantial number of them finished with a normal shutdown.

The thermal stresses are estimated based on the difference in thermal growths of the two cross-sections of the airfoil close to Position 2. The smaller area is around the trailing edge tip (marked by the blue dotted line in Fig. 17), while the larger area spans from the tip to the cooling cavity (marked by the pink dotted line in the same figure). The thermal stresses are calculated based on the following formulas:

$$\sigma_{shdn} = \left(\varepsilon_L - \varepsilon_S\right) E_{\overline{T}_{S,shdn}} = \left[\alpha_{\overline{T}_{L,shdn}} \left(\overline{T}_{L,shdn} - T_{AMB}\right) - \alpha_{\overline{T}_{S,shdn}} \left(\overline{T}_{S,shdn} - T_{AMB}\right)\right] E_{\overline{T}_{S,shdn}} \quad (51)$$

and

$$\overline{T}_{L,shdn} = \lambda_1 \overline{T}_{L,ss} = \lambda_1 \left[\lambda_3 T_{HOT,ss} + \left(1-\lambda_3\right) T_{COLD,ss}\right], \quad (52)$$

$$\overline{T}_{S,shdn} = \lambda_2 \overline{T}_{S,ss} = \lambda_2 \left(\lambda_4 T_{HOT,ss}\right), \quad (53)$$

$$T_{HOT,ss} = T_{FIRE} - \phi_{HOT} \left(T_{FIRE} - T_{COOL}\right), \quad (54)$$

$$T_{COLD,ss} = T_{FIRE} - \phi_{COLD} \left(T_{FIRE} - T_{COOL}\right), \quad (55)$$

where the large and small areas are denoted by subscripts *L* and *S*, respectively. Subscripts *ss* and *shdn* refer to the operation in steady-state and transient (i.e. at shutdown) operating conditions, respectively. The strain due to thermal expansion is denoted as ε, E is the Young's modulus, α is the thermal expansion coefficient, $\overline{T}$ is the mean material temperature calculated over the area indicated in the subscript, $T_{HOT,ss}$ and $T_{COLD,ss}$ are computed in the nodes shown in Fig. 17 (a). The coefficients λ_1, λ_2, λ_3 and λ_4, which are necessary to calculate material temperatures at shutdown based on the temperatures during steady-state operation, and the cooling effectiveness coefficients ϕ_{HOT} and ϕ_{COLD} were determined based on the numerical simulation. For clarity, for each cycle, the temperatures at shutdown are computed considering only the last timestamp, which was recorded at steady-state operating conditions. For cycles without at least one complete record with the required operational data, values of σ_{shdn} and C_c are calculated as an arithmetic average of the values describing the adjacent cycles (i.e. if the operational data are available for the preceding and subsequent cycles) or all cycles with the required operating parameters available. Values of the modulus of elasticity and the thermal expansion coefficient are determined on the basis of the FSX-414 characteristics, created based on the results of material properties tests at various temperatures.

In accordance with Equation (51), the thermal stresses at shutdown are calculated assuming a purely elastic deformation. The obtained stress values fluctuate around the 0.2% yield strength values estimated for the same material temperatures. Nevertheless, for some startup-shutdown cycles, the stresses exceed the 0.2% proof stress, which contradicts the assumption applied. Additionally, it should be considered that the yield strength of FSX-414 cobalt-based superalloy decreases under cyclic loading. Therefore, an attempt was made to account for plasticity and determine the true stress by applying the Neuber's correction method [199], which is often used

for monitoring thermal stresses and fatigue life prediction of steam turbine rotors [200–202]. This method relates the nominal elastic stress and the corresponding strain with the elasto-plastic stress and strain at the notch root, applying a function referred to as Neuber's hyperbola. It can be easily employed in the analyzed case but is also known to overestimate the notch tip stress distribution [203]. The corrected thermal stress values were insignificantly lower, on average, by 5%, compared to the elastic stresses. Simultaneously, the variation of the stress values for a single set of nozzles has considerably reduced after applying the Neuber's rule, which resulted in assigning a nearly constant value to each operating cycle. This more complex way of preparing input data for the hybrid model did not improve the accuracy of the final crack length estimates. Because of the lack of tangible benefits, it was decided to follow the initial approach, assuming a linear elastic behavior of the analyzed object.

5.2 Configuration and execution of the training process

Leveraging Eq. (15), the custom cost function, which is minimized during the learning process, is defined as follows:

$$c = \gamma RMSE_{PHY} + \left(1-\gamma\right) RMSE_{EMP}\,, \tag{56}$$

where $RMSE_{PHY}$ is referred to as the physical term, which represents how well the hybrid model respects a theoretical relationship embedded in the cost function, $RMSE_{EMP}$ is referred to as the empirical term representing the model's capability to predict the final length of the analyzed fatigue cracks accurately and $\gamma \in \left[0,1\right]$ determines how these terms contribute to the cost value. Considering the function's structure and the values that γ coefficient can take, the physical and empirical terms must have the same magnitude. In this regard, the root mean squared error is better than the mean squared error or the normalized metrics, i.e. NRMSE or NMSE.

The physical term is based on the load ratio R, which for a specific geometry and size of a crack a = const is defined as follows:

$$R = \frac{\sigma_{shdn,min}}{\sigma_{shdn,max}} = \frac{K\left(\sigma_{shdn,min}, a\right)}{K\left(\sigma_{shdn,max}, a\right)}\,, \tag{57}$$

where K is the stress intensity factor and the minimal and maximal values of thermal stresses are denoted by *min* and *max* subscripts, respectively. In the created structure of the hybrid model, the stress intensity factor is estimated using the data-driven layer based on a multilayer perceptron. Equation (57) can be rearranged as follows:

$$\frac{\sigma_{shdn,min}}{\sigma_{shdn,max}} - \frac{K\left(\sigma_{shdn,min}, a\right)}{K\left(\sigma_{shdn,max}, a\right)} = 0 \tag{58}$$

and modified, assuming that estimates of the feedforward neural network $FNN\left(\sigma_{shdn}, a\right)$ are perfectly accurate:

$$\frac{\sigma_{shdn,min}}{\sigma_{shdn,max}} - \frac{FNN\left(\sigma_{shdn,min}, a\right)}{FNN\left(\sigma_{shdn,max}, a\right)} = 0\,. \tag{59}$$

Based on Eq. (59), the physical term is defined as follows:

$$RMSE_{PHY} = \sqrt{\frac{1}{J}\sum_{j=1}^{j=J}\left(\frac{\sigma_{shdn,min,j}}{\sigma_{shdn,max,j}} - \frac{FNN\left(\sigma_{shdn,min,j}, a_j\right)}{FNN\left(\sigma_{shdn,max,j}, a_j\right)}\right)^2} = \sqrt{\frac{1}{J}\sum_{j=1}^{j=J}\left(R_j - \hat{R}_j\right)^2}, \qquad (60)$$

where $j = 1,2,\ldots,J$ is the ordinal number of triples created synthetically to train the multilayer perceptron, $J = 15\ 125$ is the number of triples and $\hat{R}_j$ is the load ratio calculated based on the estimates determined by the feedforward neural network. The triples are created based on the available empirical data, with each triple comprising a, $\sigma_{shdn,min}$ and $\sigma_{shdn,max}$. The entire set is created as follows:

- The unique values of σ_{shdn} are the basis for creating the first 375 records. $\sigma_{shdn,min}$ takes these values, which subsequently are randomly permutated to determine the values of $\sigma_{shdn,max}$. The crack lengths associated with them are drawn from the uniform distribution $U\left(\min(a), \max(a)\right)$.
- The next 1 000 triples are outside the space composed of the available empirical data. The values of $\sigma_{shdn,min}$ are randomly drawn from $U\left(\frac{9}{10}\min\left(\sigma_{shdn}\right), \frac{11}{10}\max\left(\sigma_{shdn}\right)\right)$, the corresponding values of $\sigma_{shdn,max}$ are drawn from $U\left(\sigma_{shdn,min}, \frac{11}{10}\max\left(\sigma_{shdn}\right)\right)$ and the crack lengths are randomly sampled from $U\left(a_0, \frac{5}{4}\max(a)\right)$. Training on these triples expands the original space of features and the output space. Thus, the space where the hybrid model can be applied effectively and safely, satisfying Eq. (59), enlarges.
- For each of 1 375 values of $\sigma_{shdn,min}$ determined during the two previous steps, ten values of $\sigma_{shdn,max}$ are randomly drawn from $U\left(\frac{9}{10}\min\left(\sigma_{shdn}\right), \frac{11}{10}\max\left(\sigma_{shdn}\right)\right)$, while the associated crack lengths are drawn from $U\left(a_0, \frac{5}{4}\max(a)\right)$. In this way, additional 13 750 triples are created.

The value of the physical term is calculated based on all the triples. Therefore, the constraints resulting from Eq. (57) are implemented in a soft manner through the additional term in the cost function. Attempts to reduce the number of triples and execute the training process using only the first subset, comprising 375 triplets, were unsuccessful. The stress intensity factor estimates were inconsistent with the results calculated using the theoretical formulation (65), and the predictions for particular combinations of input data were equal to zero, i.e. a minor change of the input variables resulted in a significant drop in the FNN's output. Thus, the remaining two subsets of triples are necessary to train the feedforward neural network to respect the theoretical relationship embedded in the cost function, i.e. Eq. (57). The additional triplets are not created randomly but in a way to improve the effectiveness of the training process and the multilayer perceptron's extrapolation capabilities. The process applied to generate these additional training data can be adjusted depending on particular objectives. Finally, it should be remembered that the physical term's impact on the cost value depends on the γ coefficient, which should be set based on the evaluation of the quantity and quality of empirical data and the accuracy of the theoretical description of the analyzed problem, which is embedded into the physics-informed neural network. If γ equals one, the cost reduces to the physical term only, while with γ equal

to zero the learning objective is to minimize only the empirical term, limiting the extrapolation capabilities of the model, ignoring consistency with the theoretical relationship and resulting in unreliable and erratic predictions.

The empirical term of the custom cost function is calculated in the following way:

$$RMSE_{EMP} = \sqrt{\frac{1}{N}\sum_{i=1}^{i=N}\left(a_i - \hat{a}_i\right)^2}\,, \tag{61}$$

where $i = 1,2,\ldots,N$ is the observation's ordinal number, N is the total number of observations in the training set and $\hat{a}_i$ is the final crack length estimate obtained using the hybrid model.

An attempt was made to embed into the cost function an additional physical term, which is based on the following inequality:

$$\frac{K\left(\sigma_{shdn}, a_{min}\right)}{K\left(\sigma_{shdn}, a_{max}\right)} - 1 < 0\,. \tag{62}$$

This mathematical relationship is valid for a specific stress value σ_{shdn} = const, and when a_{min} < a_{max}. Based on Inequality (62), the additional physical term is defined as follows:

$$\begin{aligned} RMSE_{PHY,ADD} &= \sqrt{\frac{1}{H}\sum_{h=1}^{h=H}\left[ReLU\left(\frac{FNN\left(\sigma_{shdn,h}, a_{min,h}\right)}{FNN\left(\sigma_{shdn,h}, a_{max,h}\right)} - 1\right) - 0\right]^2} = \\ &= \sqrt{\frac{1}{H}\sum_{h=1}^{h=H}\left[ReLU\left(\hat{R}_h - 1\right)\right]^2}\,, \end{aligned} \tag{63}$$

where $h = 1,2,\ldots,H$ is the ordinal number of triples created synthetically to train the multilayer perceptron, H = 13 750 is the number of triples and $\hat{R}_h$ is the load ratio calculated based on the estimates determined by the feedforward neural network. With $RMSE_{PHY,ADD}$ defined, the cost function is modified in the following way:

$$c = \gamma\left(\delta RMSE_{PHY,ADD} + \left(1-\delta\right) RMSE_{PHY}\right) + \left(1-\gamma\right) RMSE_{EMP}\,, \tag{64}$$

where $\delta \in [0,1]$ determines how $RMSE_{PHY,ADD}$ contributes to the final cost value with reference to $RMSE_{PHY}$. The expansion of the cost function was aimed at changing the inclination of $K_{\sigma_{shdn}} = FNN\left(\sigma_{shdn} = const, a\right)$ characteristics, which should be monotonically increasing functions according to the theoretical Equation (65). Including the additional physical term in the cost function helped to modify the characteristics. Nevertheless, the slopes were significantly smaller compared to the theoretical curves, and the empirical term increased considerably. In this setup, the final crack length estimates obtained by means of the hybrid model were inaccurate, and the FNN's characteristics were not completely consistent with the theoretical ones. Therefore, following Eq. (56), $RMSE_{PHY,ADD}$ is not included in the custom cost function used to train the physics-informed neural network.

As explained in [182], starting the learning process with poorly initialized neural network weights may increase the time required to train the model or even cause a divergence

of the optimization. Thus, the feedforward neural network composing the data-driven layer is pre-trained using the results determined by the following empirical formula as the reference that applies to a single-notch test specimen subjected to tension:

$$K_{THEOR}\left(\sigma_{shdn},a\right)=\sigma_{shdn}\sqrt{a}Y= \\ =\sigma_{shdn}\sqrt{a}\left[1.99-0.41\frac{a}{W}+18.7\left(\frac{a}{W}\right)^{2}+38.48\left(\frac{a}{W}\right)^{3}+53.85\left(\frac{a}{W}\right)^{4}\right], \tag{65}$$

where Y is the geometrical factor and W is the specimen width assumed to be equal to the nozzle airfoil's chord length. This equation connecting the stress intensity factor with the specific type of loading and the specimen dimensions is based on the report [204], the results of which are also published in [205]. The authors of the report used the boundary value collocation procedure applied to the Williams stress function, which was proposed by Gross et al. [206]. However, this analytical method assumes that the specimen is uniformly loaded across the width and that the distance between the loaded cross-section and the notch plane is not shorter than the width. One thousand records were generated for the pretraining. The crack lengths and stresses are evenly spaced numbers over $\left[\min\left(a_{train}\right),\max\left(a_{train}\right)\right]$ and $\left[\min\left(\sigma_{shdn,train}\right),\max\left(\sigma_{shdn,train}\right)\right]$ intervals, respectively. Additionally, the former are randomly permutated. The mean squared error is considered as the cost function. Typically, less than 300 epochs are required to achieve satisfactory accuracy, and the training process duration is shorter than 20 seconds.

During the main part of the training process the value of γ, which determines how the physical and empirical terms contribute to the final cost value, is not constant and changes depending on the mean absolute error value defined by means of Eq. (67). Attempts to train the hybrid model using a fixed value of the coefficient resulted in biased estimates. Due to a lack of balance between the $RMSE_{PHY}$ and $RMSE_{EMP}$, the predictions were either inaccurate but with very low values of the physical term or very accurate but with high values of the physical term, indicating poor generalization capabilities of the model. Therefore, to find a proper balance between both terms and reduce their values during the training simultaneously, the value of γ coefficient changes dynamically during the learning process using the following formula:

$$\gamma=\begin{cases}0, & MAE_{PHY}\le 0.0001\\ \left(777+\frac{7}{9}\right)MAE_{PHY}+\frac{7}{90}, & MAE_{PHY}\in\left(0.0001,0.001\right]\\ \left(33+\frac{1}{3}\right)MAE_{PHY}+\frac{2}{3}, & MAE_{PHY}\in\left(0.001,0.01\right]\\ 1, & MAE_{PHY}>0.01\end{cases}, \tag{66}$$

where the mean absolute error MAE_{PHY} is defined as follows:

$$MAE_{PHY}=\frac{1}{J}\sum_{j=1}^{j=J}\left|\frac{\sigma_{shdn,min,j}}{\sigma_{shdn,max,j}}-\frac{FNN\left(\sigma_{shdn,min,j},a_{j}\right)}{FNN\left(\sigma_{shdn,max,j},a_{j}\right)}\right|=\frac{1}{J}\sum_{j=1}^{j=J}\left|R_{j}-\hat{R}_{j}\right|, \tag{67}$$

where the meanings of j, J and $\hat{R}_j$ are the same as in Eq. (60). Fig. 18 presents the piecewise continuous function $\gamma = f\left(MAE_{PHY}\right)$. The first stage of the training is aimed at assuring that the final solution will respect the theoretical relationship embedded in the cost function. Thus, the training process is initiated with $\gamma = 1$, and as the error decreases, the coefficient γ reduces slowly to 0.7. Further decrease of the mean absolute error results in a significant drop in the value of the coefficient, reducing the contribution of $RMSE_{PHY}$ on the cost value and putting more emphasis on $RMSE_{EMP}$. The objective of this stage of the training is to reduce as much as possible the empirical term but within the subspace of solutions, which are consistent with the considered theoretical relationship. During the execution of the analysis, it was observed that at this stage of the learning process, values of the empirical and physical terms decrease simultaneously. For clarity, the value of γ coefficient does not change after each iteration but after a certain number of training epochs. Initially, that number is equal to 100, but if the product of MAE_{PHY} and $MSE_{EMP,test}$ (i.e. it is the mean squared error evaluated against the empirical data from the test set) does not reduce consecutively, the number of epochs doubles. The Adam optimization algorithm [181] with a constant learning rate $\eta = 0.001$ is applied to minimize the cost function. This algorithm performed better when the data-driven layer comprises one or two hidden layers and is comparable to the Nadam algorithm when the multilayer perceptron has eight hidden layers. This kind of iterative training of the physics-informed neural network with dynamic cost function changes is more time-consuming than the standard continuous approach to execute the learning process. It may take up to one week on an Intel Core i7-9850H Central Processing Unit with six cores and a 2.60 GHz base frequency. Nevertheless, it enables finding the right balance between the physical and empirical terms of the cost function, simultaneously reducing their values and creating accurate predictive models that respect prior knowledge. Open-source Python packages are used during the study, i.e. Keras application programming interface for the TensorFlow library is used to create the structure and train the hybrid model, Numpy and Pandas are used for data manipulation and Matplotlib [207] for visualizations.

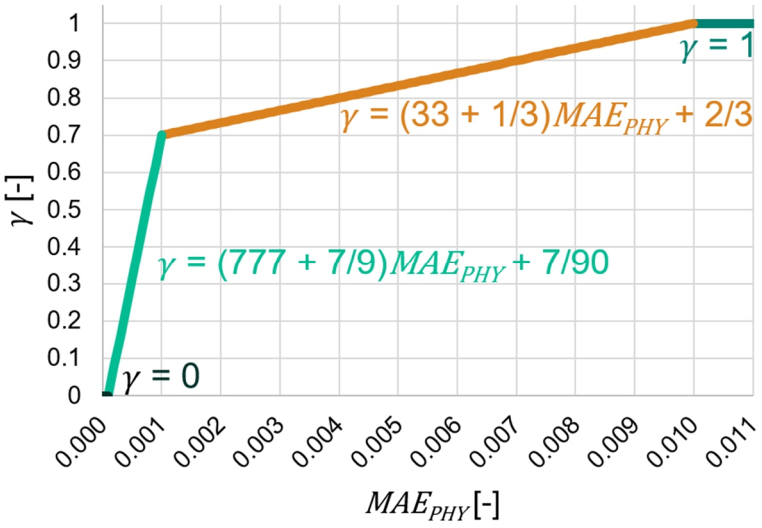


Fig. 18 The piecewise continuous function used to calculate the values of coefficient γ, which determines how the physical and empirical terms contribute to the final cost value.

5.3 Results of the regression analysis

In the introduction to this chapter, it was written that the available sample comprises 19 observations since it is limited only to Type C gas turbines. However, during the analysis of available data and attempts to effectively train the hybrid model, differences in crack growth rates were observed between the new and repaired nozzles. The consequence is a further reduction of the sample and removal of data related to components that were repaired in the past. The removed nozzle sets have the following identifiers (see Fig. 15 for details): S1N_2, S1N_3, S1N_5, S1N_6, S1N_24 and S1N_45. Finally, the sample comprises only 13 observations and it can be described as follows:

- Considering the thresholds defined in the previous chapter, i.e. L_{LOW} and L_{HIGH}, 4 out of 13 observations (31%) are classified as "short", the "medium" class is represented by 5 observations (38%) and the remaining 4 observations are classified as "long".
- 8 observations out of 13 (62%) correspond to the natural gas liquefaction facility located in a marine, salty environment. 89% of records from the "short" and "medium" classes are related to that plant, i.e. 8 observations out of 9.
- All the observations classified as "long" correspond to units installed in another gas liquefaction facility located in a tropical, humid environment.
- For the available observations, the interval fired start values are constrained within the interval $[22,157]$, with a median of 40 (the lower quartile equals $Q_1 = 29$ and the upper $Q_3 = 110$). The percentage of startup-shutdown cycles for which at least one complete record with the operational data is available ranges from 10% to 97%, with a median of 36% ($Q_1 = 31\%$ and $Q_3 = 59\%$). Thus, the fraction of missing data is significant.

The target variable is the maximum crack size at Position 2 on the nozzle's trailing edge.

In order to not reduce the training set considerably, the first, reference model is trained based on ten observations. As described in the previous chapter, the training and testing subsets' composition should reflect the available sample's structure. Therefore, the test set is prepared considering the following constraints:

- The proportions between the "short", "medium" and "long" classes valid for the sample are applied to select elements of the test subset. Thus, each of the classes constitutes 1/3 of the test set.
- In order to avoid a significant reduction of the training and validation subsets, the test set comprises three elements. 77% of the data will be used to train the model, while the remainder will be utilized to assess the generalization capabilities.
- 2 observations out of 3 (67%) correspond to the gas liquefaction facility located in a marine, salty environment.
- The test data points should be diversified in terms of the number of interval fired starts and the availability of the operational data. Actually, for the selected observations, the *IFS* and coverage values are between the 40^{th} and 90^{th} percentiles. An emphasis is placed on parts that accumulated more interval fired starts. Moreover, operational data availability in future applications is expected to be better than in this research, based on the oldest available data in many cases.

Such a controlled manner of preparing the training and test sets considering multiple constraints was successfully applied in the case of purely data-driven models based on popular machine learning algorithms described in the previous chapter. It should improve the effectiveness of the learning process, facilitate the interpretation of obtained results, help to identify conditions in which the model returns both accurate and inaccurate estimates and simplify the overall evaluation of the hybrid model. The estimates obtained by means of the reference model and selected characteristics of the multilayer perceptron, which constitutes the data-driven layer, are shown in Fig. 19, Fig. 20, Fig. 21 and Fig. 22. In order to protect proprietary information, the data presented in the plots were normalized by dividing each value of crack length and thermal stress by the maximal measured crack length and the maximal calculated thermal stress, respectively, which were determined based only on the data related to the analyzed 13 observations. The most important considerations regarding the obtained results are as follows:

- The pre-trained model, which was trained on the basis of the results calculated using the theoretical formula (65) as the reference, provides significantly underestimated predictions of the final crack lengths (Fig. 19).
- The normalized root mean squared error evaluated against the testing data is equal to $NRMSE_{EMP,TEST}$ = 9%, while for the training data, it equals $NRMSE_{EMP,\ TRAIN}$ = 7%. Considering the limitations in terms of the sample size and the availability of the operational data and the simplified approach applied to compute the thermal stresses, the accuracy of the final crack length estimates is satisfactory. These two values are comparable with the accuracy of recording those cracks during a visual inspection preceding repair activities, where the maximum normalized error may reach 7%. The most accurate predictions were obtained for the longest cracks, while the least accurate ones were for the two shortest observations in the sample, which the model overestimated. Nevertheless, such a characteristic of the model is valued since, as already explained, long cracks may jeopardize the component's integrity or require immediate corrective maintenance.
- Regardless of the very stable operating profile of the analyzed gas turbines, which is mainly characterized by a continuous duty at nominal power, the model is capable of differentiating the crack growth rates and the final estimates based on subtle differences between the units, which were extracted from actual operational data during the training.
- The feedforward neural network responsible for the stress intensity factor calculations obeys the theoretical equation embedded in the cost function, i.e. Eq. (57). The estimates determined by the data-driven layer for a few selected crack lengths and the full range of thermal stresses are presented in Fig. 21.
- The characteristics of the multilayer perceptron at the constant thermal stresses after the training, i.e. $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn} = const, a'\right)$, which are presented in Fig. 22, are completely different compared to the pre-trained model. At the beginning of the crack propagation process, the stress intensity factor estimates are the highest, and then they decrease as the crack length increases. Regardless of the normalized thermal stress value, the characteristics before and after the training cross each other at $a' \cong 0.82$ for a specific value of σ'_{shdn}. It may indicate that the actual values of thermal stresses at the beginning of the crack growth process (i.e. until the normalized crack size is smaller

than 0.82) are higher than the values utilized to train the model, which were estimated using the previously described approach. In fact, it is expected that the applied stresses should decrease as the crack length increases since the crack propagates toward areas subjected to lower stresses. Simultaneously, the yield strength of FSX-414 cobalt-based superalloy decreases under cyclic loading, which introduces additional complexity. It

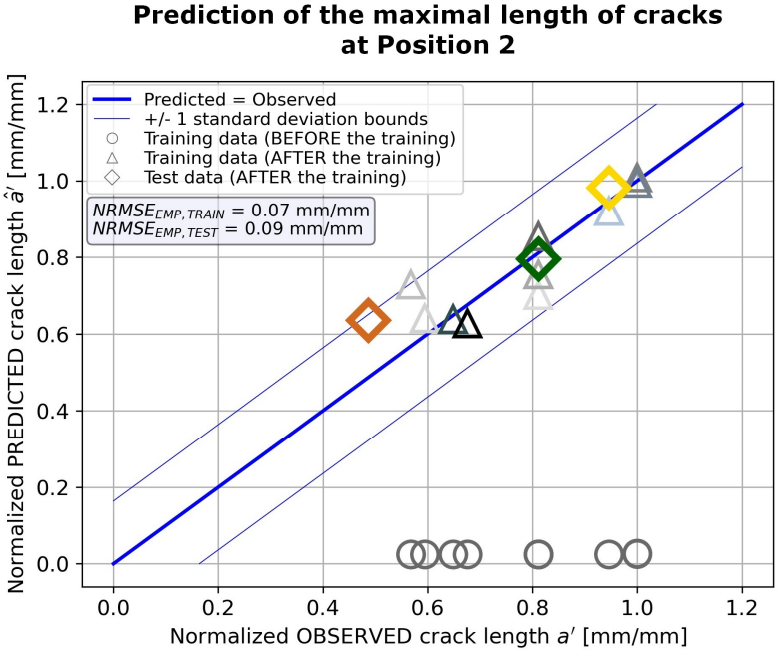


Fig. 19 Evaluation of the hybrid model based on the physics-informed neural network, trained using ten observations, against the training and test data presented on the predicted vs. observed plot.

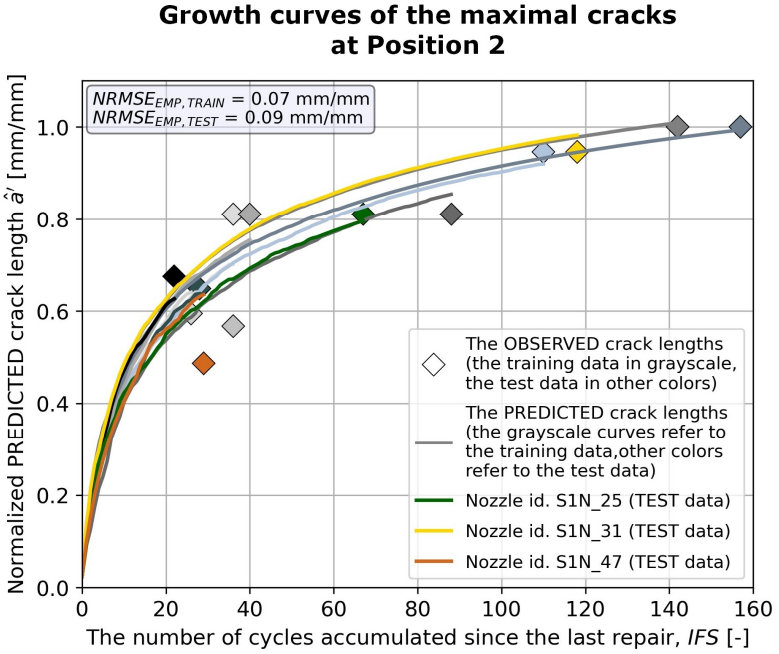


Fig. 20 Evaluation of the hybrid model based on the physics-informed neural network, trained using ten observations, against the training and test data.

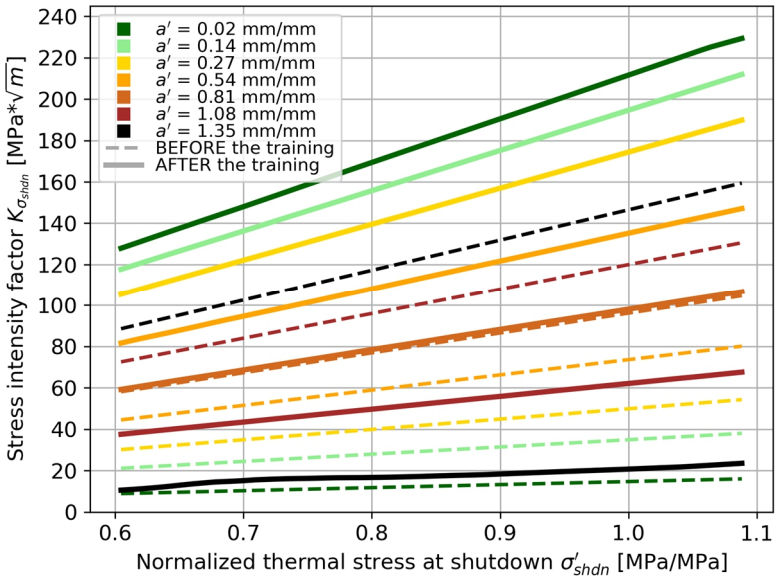


Fig. 21 Stress intensity factor estimates at the constant crack lengths $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn}, a' = const\right)$ determined by the data-driven layer before and after the training using ten observations.

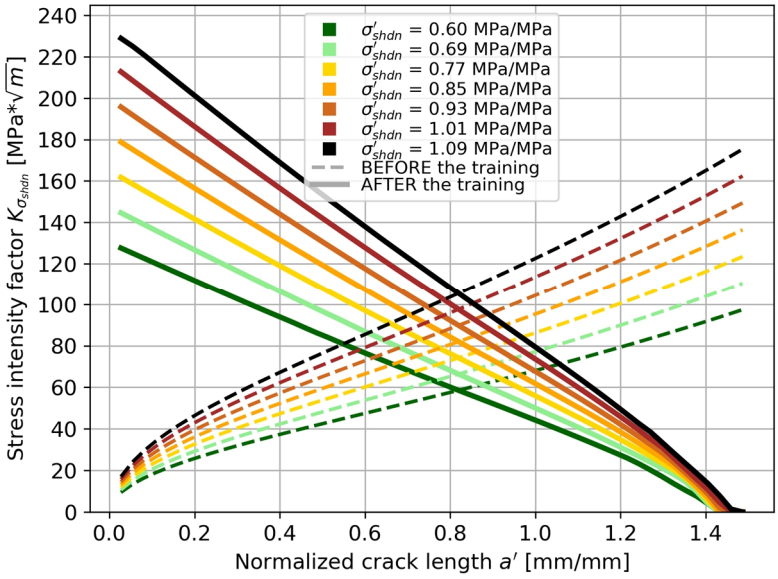


Fig. 22 Stress intensity factor values at the constant thermal stresses $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn} = const, a'\right)$ determined by the data-driven layer before and after the training using ten observations.

can be presumed that an improved approach to estimating the thermal stresses can change the characteristics and align them to ones based on the theoretical relationship. Nevertheless, as already described, the attempts to include the additional physical term $RMSE_{PHY,ADD}$ in the cost function or the time-depended part estimating crack length increments due to creep did not improve the model's accuracy or consistency with the applicable physical laws.

- Based on the final, fully-specified model, the highest crack growth rates are expected at the beginning of the crack propagation process during the first 20 startup-shutdown

cycles. It is consistent with the available empirical data since the cracks with lengths $a' = 0.68$ and $a' = 0.81$ were recorded after only 22 and 35 cycles, respectively.

- Increasing the number of hidden layers of the feedforward neural network from two to eight has an insignificant impact on the final crack length estimates.
- Application of the reference model to the nozzles subjected to repair activities in the past leads to significantly underestimated predictions of the final crack lengths. Compared with brand-new components, these parts are characterized by different crack propagation rates, and it is recommended to train a separate model using a dedicated set of empirical data.
- Fig. 21 and Fig. 22 show estimates of the data-driven layer outside the original space of input data when $a' > 1$ or $\sigma'_{shdn} > 1$. The presented characteristics are continuous, with no peaks or drops, regardless of whether the input variables are from inside or outside the original input space. The curves show directly how the feedforward neural network extrapolates and interpolates. Thereby, knowing the range of values of the input data, the predictions can be made in a fully conscious way, using those characteristics. Additionally, the space in which the model should respect the theoretical relationship can be enlarged by synthetically created data, as described in the previous section.
- The same architecture was successfully used to train models predicting other statistical measures, i.e. 25th, 50th and 75th percentiles and the minimal crack length. Considering a single set of nozzles, such a collection of hybrid models may be utilized to predict the distribution of trailing edge cracks in all sectors. It can also be applied to evaluate if the component is eligible for a repair interval extension through an on-condition check, to determine a serviceable limit for such an inspection and to predict the percentage of parts that will pass the inspection.

In order to better understand the limitations of the proposed architecture in terms of the minimal number of available damage measurements, the next attempt is to train the hybrid model with the training set comprising two observations only and use the remaining data points for testing purposes. Two different train-test splits are prepared and subsequently processed. The first training set comprises the components, which accumulated many interval fired starts, but the availability of the operational data for those cycles is very limited. The identifiers of these nozzle sets are as follows: S1N_27 with coverage equal to 27% (i.e. at least one complete record with the required operating parameters is available for 39 out of 142 cycles) and S1N_30 with 33% coverage, i.e. 29 out of 88 cycles. The second training set comprises the parts that accumulated a significantly lower number of startup-shutdown cycles, but the availability of the operational data is doubled compared to the previous set. The identifiers of these nozzle sets are S1N_7 with 58% coverage, i.e. 15 out of 26 cycles, and S1N_8 with 64% coverage, i.e. 23 out of 36 cycles. The learning process was executed without any difficulties for both sets. The training duration was significantly shorter compared to the reference model. Regardless of the analyzed training set, the obtained results are very similar. Therefore, to avoid excessive reporting, only the estimates and the characteristics of the model trained using the second training set are presented in the dissertation. These results are shown in Fig. 23, Fig. 24, Fig. 25 and Fig. 26. The most important considerations regarding the obtained results are as follows:

- Despite the small number of training observations, the normalized root mean squared error evaluated against the test set comprising 11 data points equals to $NRMSE_{EMP,TEST}$ = 11% and is comparable to the performance of the reference model on the unseen data.

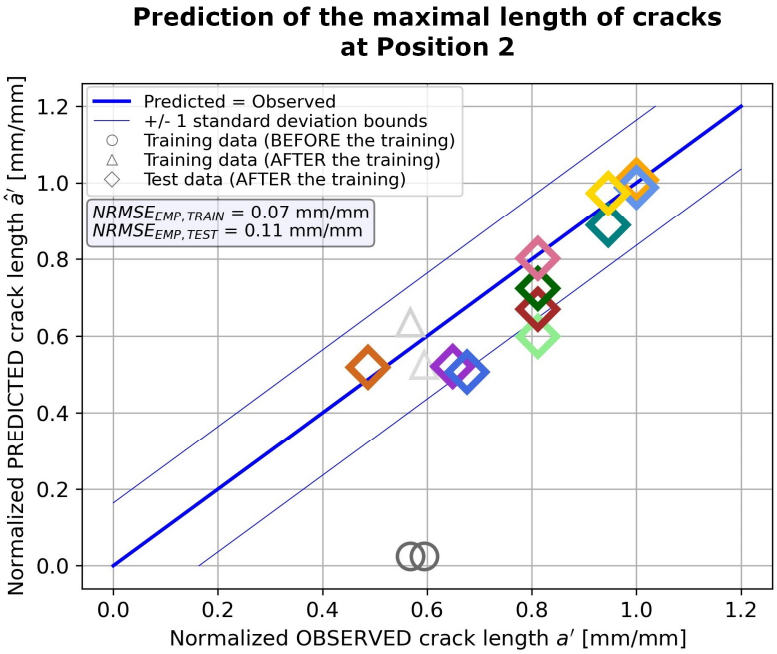


Fig. 23 Evaluation of the hybrid model based on the physics-informed neural network, trained using two observations, i.e. S1N_7 and S1N_8, against the training and test data presented on the predicted vs. observed plot.

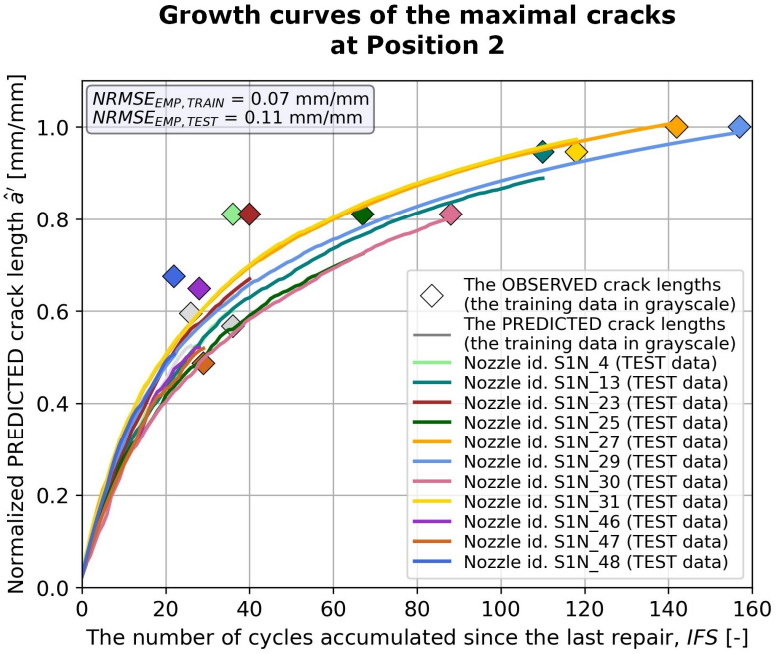


Fig. 24 Evaluation of the hybrid model based on the physics-informed neural network, trained using two observations, i.e. S1N_7 and S1N_8, against the training and test data.

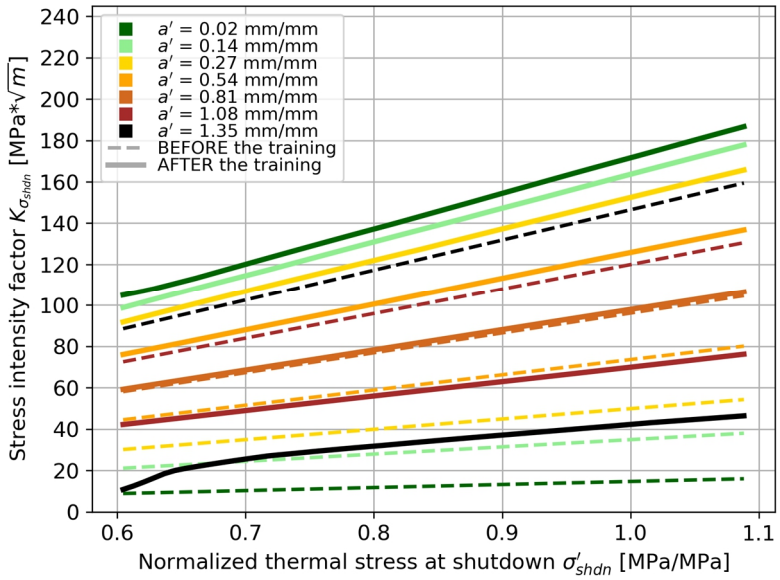


Fig. 25 Stress intensity factor estimates at the constant crack lengths $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn}, a' = const\right)$ determined by the data-driven layer before and after the training using two observations, i.e. S1N_7 and S1N_8.

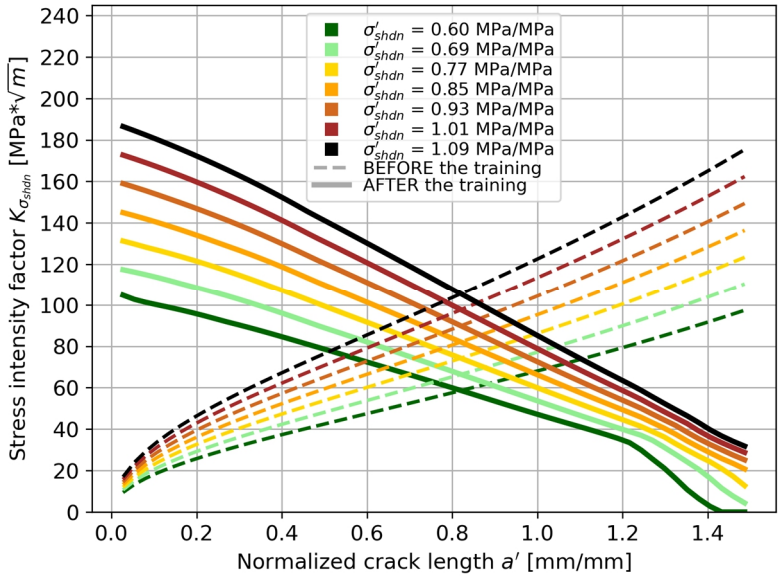


Fig. 26 Stress intensity factor values at the constant thermal stresses $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn} = const, a'\right)$ determined by the data-driven layer before and after the training using two observations, i.e. S1N_7 and S1N_8.

- The components that constitute the training set accumulated less than 40 interval fired starts and had relatively short cracks, i.e. $a' \cong 0.58$, which were classified as “short”. Despite the limited training subspace, the model has good generalization capabilities and extrapolates accurately if $IFS > 36$ and $a' > 0.58$. It is capable of providing accurate predictions of the final crack lengths for parts that accrued over four times more startup-shutdown cycles and had around 40% longer cracks.

- The characteristics of the multilayer perceptron presented in Fig. 25 and Fig. 26 are comparable to the characteristics of the reference model. Nevertheless, the crack length increments estimated by means of this model during the first ten cycles are smaller in comparison with the reference model, which is caused by lower stress intensity factor estimates at the beginning of the crack propagation process. In addition, in accordance with the model, the process does not stop at $a' \cong 1.48$ as it was observed for the reference model.

Considering the successful attempt to train the hybrid model based on two observations, the last trial aims to execute the learning process with the training set comprising just one data point. The results presented in Fig. 27, Fig. 28, Fig. 29 and Fig. 30 were obtained using nozzle set S1N_8 as the training observation. It should be clarified that it is the only successful attempt to train the hybrid model using a single observation. During the remaining trials, the learning process was executed using S1N_4, S1N_27, S1N_30, S1N_31 and S1N_48 nozzle sets as the single training data point, but the model's accuracy was unsatisfactory. Additionally, due to difficulties with the optimization, the number of hidden layers of the multilayer perceptron was increased from two to eight. The cost function optimization proceeded in a different way in comparison with the previously described models. In this case, the root mean squared error evaluated against the test data quickly reduced to a satisfactory value. In the latter stage of the training process, the objective was to reduce the physical term's value. For the models trained using ten and two data points, $RMSE_{PHY}$ reduced to a certain value in the first stage of the process, and subsequently, both terms decreased as the training progressed. The most important considerations regarding the obtained results are as follows:

- The normalized root mean squared error evaluated against the test set comprising 12 data points equals $NRMSE_{EMP,TEST}$ = 17% and is worse compared to the previously described model trained using two data points. The lowest prediction accuracy is found for the components that accrued 40 or fewer interval fired starts, for which the model underestimates the final crack lengths.

- As already described, the part selected for the training set accumulated 36 cycles and had a crack with the length $a' \cong 0.58$. Despite the training subspace's limitations, the model provides accurate predictions of the final crack lengths for parts that accrued over four times more startup-shutdown cycles and had around 40% longer cracks. It is significantly more accurate than the model trained using the theoretical formula results as the reference. Therefore, the capability to extract and use knowledge even from one empirical data point is very valuable and may bring tangible benefits.

- In the interval $a' \in \left[a'_0, 0.3\right]$ the estimates determined by the feedforward neural network are lower compared to the model trained using two observations. The characteristics of the multilayer perceptron at the constant thermal stresses after the training, i.e. $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn} = const, a'\right)$, which are presented in Fig. 30, do not decrease monotonically as in the case of the previously described models, and have a visible peak at $a' \cong 0.3$. However, the prediction interval is comparable with the characteristics of the pre-trained model.

- This attempt proves that the physics-informed neural network based on the architecture described in Section 5.1 can be trained effectively using a single observation, even if

the required measurements of operating parameters are not available for all startup-shutdown cycles. Nevertheless, with such a limited training set, a successful execution of the learning process cannot be guaranteed, and the results obtained by means of the fully-specified model may vary significantly depending on the particular observation used during the training process. Additionally, it should be highlighted that evaluating

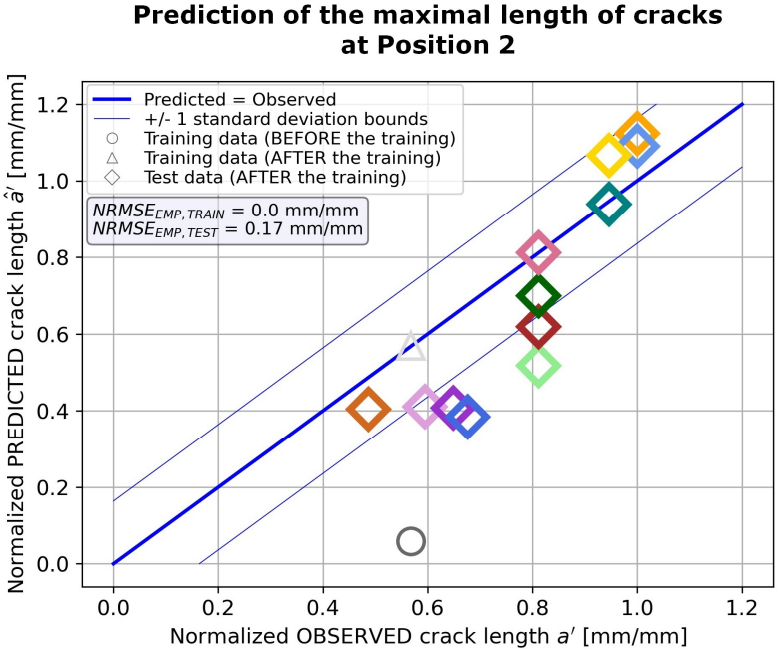


Fig. 27 Evaluation of the hybrid model based on the physics-informed neural network, trained using a single observation , i.e. S1N_8, against the training and test data presented on the predicted vs. observed plot.

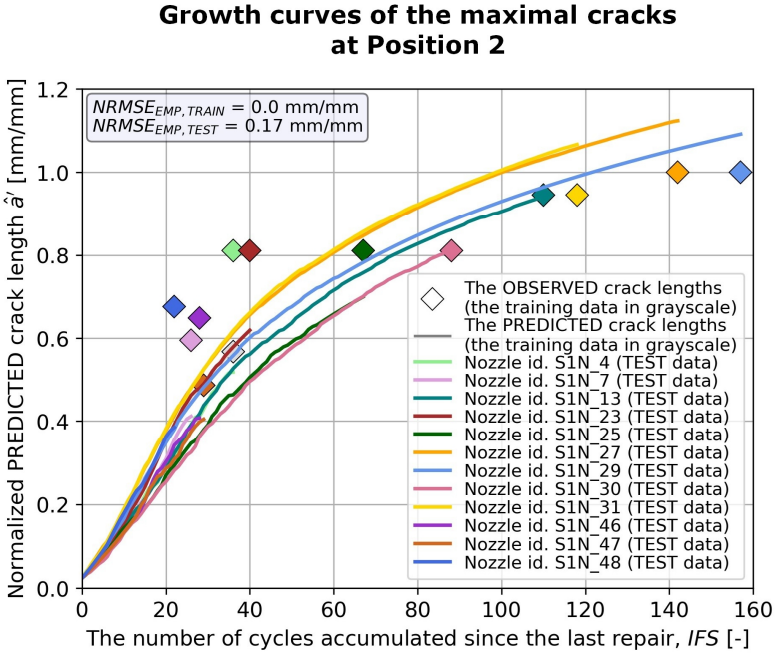


Fig. 28 Evaluation of the hybrid model based on the physics-informed neural network, trained using a single observation, i.e. S1N_8, against the training and test data.

the model's performance against unseen data in the analyzed case was possible since the test set comprises 12 data points. Without data to build the test set, a standardized set of metrics and quality indicators can be defined and used for the assessment.

Considering the normalized root mean squared error evaluated against the test subset $NRMSE_{EMP,TEST}$ as the reference metric, the results of this research aimed at applying machine learning methods for predicting the size of damage to gas turbine parts are summarized in Table

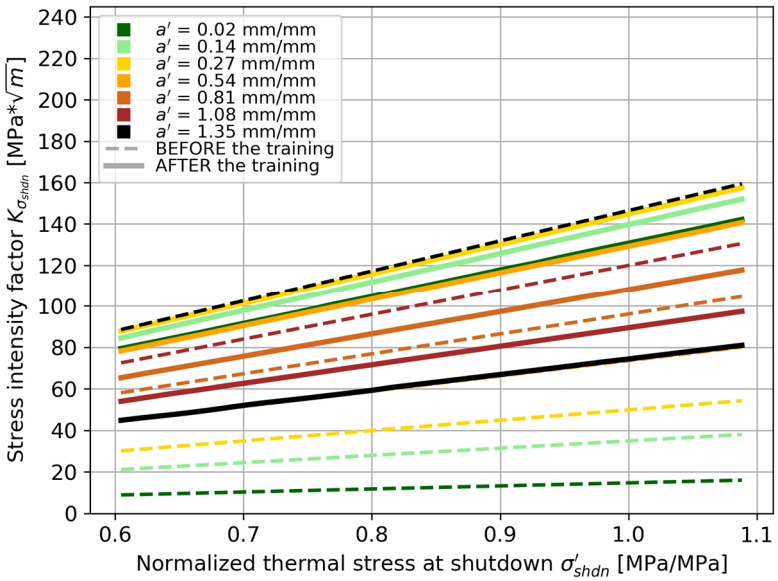


Fig. 29 Stress intensity factor estimates at the constant crack lengths $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn}, a' = const\right)$ determined by the data-driven layer before and after the training using a single observation, i.e. S1N_8.

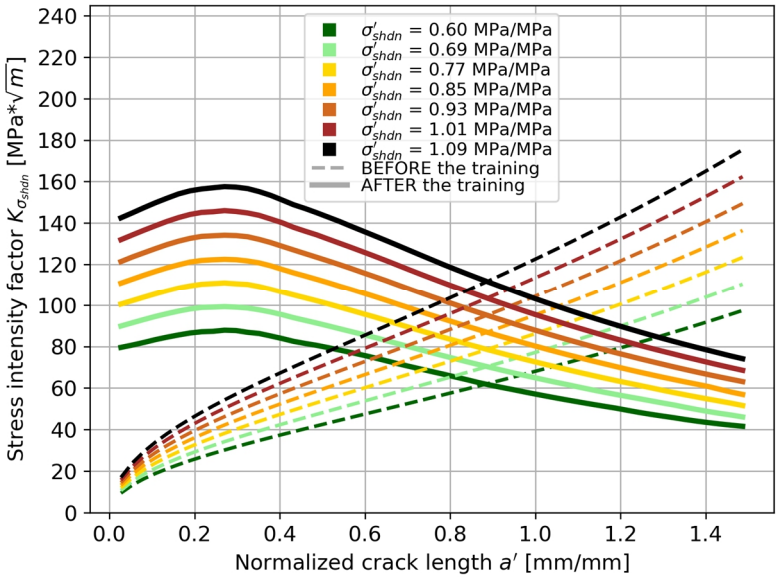


Fig. 30 Stress intensity factor values at the constant thermal stresses $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn} = const, a'\right)$ determined by the data-driven layer before and after the training using a single observation, i.e. S1N_8.

2. The error values were normalized by dividing them by the maximal measured crack length but considering only the subset used to create the models based on a PINN (i.e. 13 sets of the nozzles that were not subjected to repair activities and operated in Type C gas turbines). The table presents the results obtained by applying the physics-informed neural network and the popular statistical learning algorithms described in the previous chapter. Regardless of the small number of training observations, the highest prediction accuracy was achieved with the hybrid models based on the PINN. However, it should be reminded that the sample utilized to create these models was the most homogeneous, i.e. it consisted of components that operated in Type C gas turbines only and were not subjected to repair. The accuracy of the hybrid models based on the physics-informed neural network is satisfactory, especially if the limitations in empirical data availability are considered. Moreover, the estimates determined using these models respect the theoretical relationships embedded in the cost function, and the models can be consciously used to extrapolate since their behavior outside the space of available empirical data is known.

Table 2 The normalized root mean squared error evaluated against the test data $NRMSE_{EMP,TEST}$ describing the models based on machine learning algorithms created to predict the length of fatigue cracks found on the analyzed high-pressure nozzles.

Algorithm type	Description of the sample	Train/test split	Normalized RMSE evaluated against the test set
Physics-informed neural network	The parts that have not been repaired and operated in Type C gas turbines.	10/3	0.089
		2/11	0.110
		1/12	0.173
AdaBoost.R2 regression	The brand-new and repaired parts that operated in Type A, Type B and Type C gas turbines.	25/6	0.156
Feedforward neural network regression			0.161
Polynomial regression			0.165
Random forest regression			0.169
Kernel ridge regression			0.190
Multiple linear regression			0.190
Support vector regression			0.198
XGBoost regression			0.311

5.4 The novel method of domain generalization and knowledge transfer

As presented in this and the previous chapter, creating predictive models based on a limited number of empirical data is a challenge. However, in many practical applications the so-called small data regime is normal since failures are rare, inspection intervals are extended, or data acquisition costs are too high. The Weibull analysis is commonly used in such situations, including applications to the survival analysis of turbomachinery. The main advantages of

Weibull analysis are listed in Section 3.2. Interestingly, all of them can be used to characterize the previously described hybrid models based on the physics-informed neural network, i.e.:

- The capability to provide accurate estimates with small datasets comprising two or three failed data points.
- The method is data-driven, but the obtained results provide information concerning the physics of failure, which can be the basis for assessing whether the model is consistent with prior knowledge about the analyzed phenomenon.
- The core of a fully-specified model can be presented on a plot, enabling assessment of the model's correctness and quality and quickly interpreting the results. In the case of a Weibull distribution, the shape parameter can be considered as the core. This parameter determines the slope of the straight line representing the Weibull distribution's cumulative density function on a log-log plot, informing how the failure rate will change over time. In the case of created hybrid models, the multilayer perceptron responsible for the stress intensity factor estimates can be seen the model's core, which determines the crack propagation rate.
- This method has a universal character and is applicable for modeling damage caused by various physical phenomena.

As already mentioned, a Weibull analysis can be done even if no failed data points are available by applying the Weibayes method, which is a one-parameter Weibull distribution with an assumed value of the shape parameter β. According to [208], Weibayes is the best choice if failures have not been recorded or their number is limited, but only if a reasonable slope estimate is available. If there are no failed data points, it is assumed that the first failure is imminent, and the scale parameter estimate is the conservative 63% confidence bound on the true η value. If failures have been recorded, the scale parameter is a maximum likelihood estimate. The sources used to assume the slope value can be as follows:

- a Weibull analysis related to another object but the same type of damage, including models created in the past;
- recommendations available in standards, design practices, handbooks and literature;
- physics-based models or models based on material property data;
- knowledge of experts about the physics of failure.

Therefore, if the sample comprises only a few failed data points and the shape parameter value is assumed with the best available knowledge, only the scale parameter must be estimated, and the resultant Weibayes will have smaller uncertainties. It can be seen as a transfer of knowledge from a source domain, in which the shape parameter value was determined, to the target domain, in which the knowledge is leveraged to obtain more accurate predictions. It should be reminded that the notion of "leveraging" was defined as the action of using something already available in order to obtain something new or better and maximize advantages. The knowledge extracted from empirical data in the source domain is captured in the value of β parameter, and it can be easily transferred and applied in other domains.

The Dauser shift [209] is another method in which the slope value is fixed, and the scale parameter is estimated. It can be applied to adjust η parameter estimate when the time to failure is known for each unit that failed, but it is unknown how much the remaining units that did not fail operated. In this metod the shape parameter value is estimated based on the failed data points only, while the characteristic life is estimated based on the mean time to failure (MTTF)

and the percentage of failed units. The original model, which is the source of the slope value, provides a very conservative estimate of the characteristic life. The adjusted model is represented on a log-log graph by a straight line parallel to the original model's line. The shifted line is at the intersection of the MTTF value and the percentage of failed units. In accordance with [209], the method was successfully applied in many cases, but it is not a best practice. Nevertheless, again, the value of β parameter is determined based on a subset of empirical data and may be leveraged for the entire fleet of units to obtain more accurate predictions.

The last method of adjusting survival models based on a Weibull distribution will be described using the example of high-pressure turbine blades of a heavy-duty gas turbine. The model's objective is to predict the number of components scrapped during repair activities due to radial cracks found on the blade's squealer tip. In the analyzed scenario, a new blade configuration is introduced, made from a different material compared to the baseline configuration but with the same geometry. Without empirical data on the new type of blades, the analysis aims to predict the number of upgraded blades scrapped due to the radial cracks, leveraging empirical data on the old configuration. The survival model for the baseline configuration is based on a two-parameter Weibull distribution and is presented in Fig. 31 (author's own plot based on the analysis results described in [210]). The sample comprising 11 failures (i.e. these data points refer to components that were scrapped due to the radial cracks) and 2818 suspensions (i.e. data points related to the parts that were not scrapped due to the analyzed failure type) was used to create the baseline model. The first assumption taken to create a model for the new blade is that the change in failure rate over time of the new configuration is similar to the old one. The Weibayes method is applied to build the model for the new configuration, for which empirical

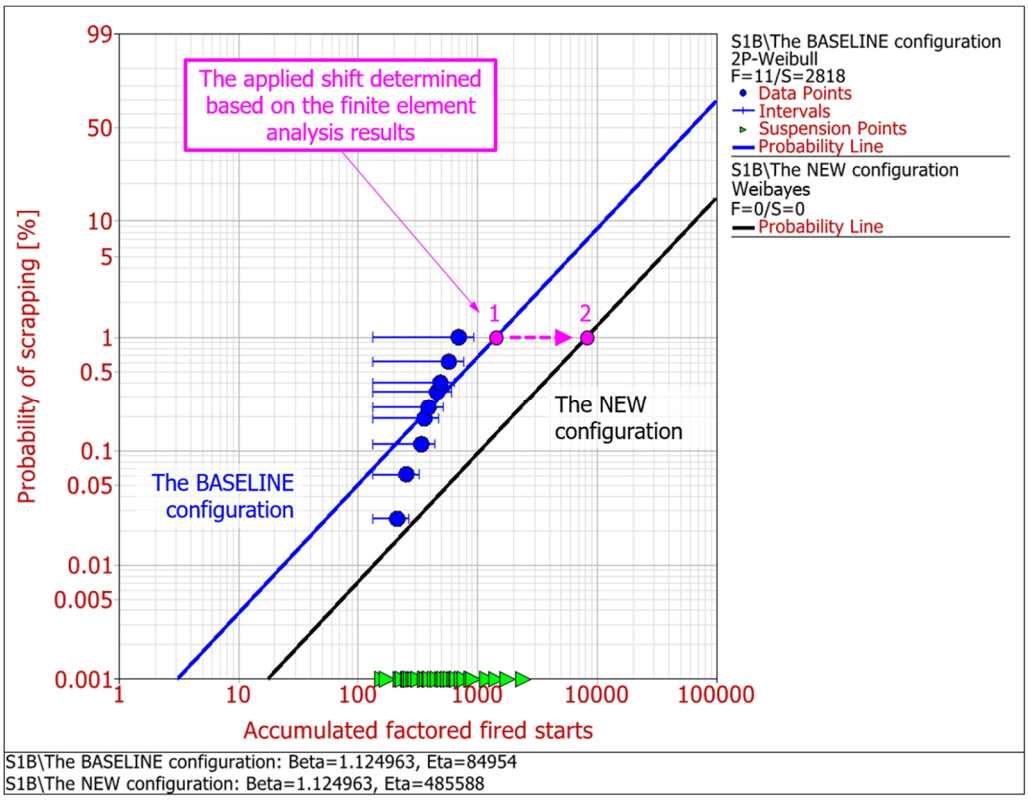


Fig. 31 The probability of scrapping a high-pressure turbine blade due to the radial cracks on the blade's squealer tip.

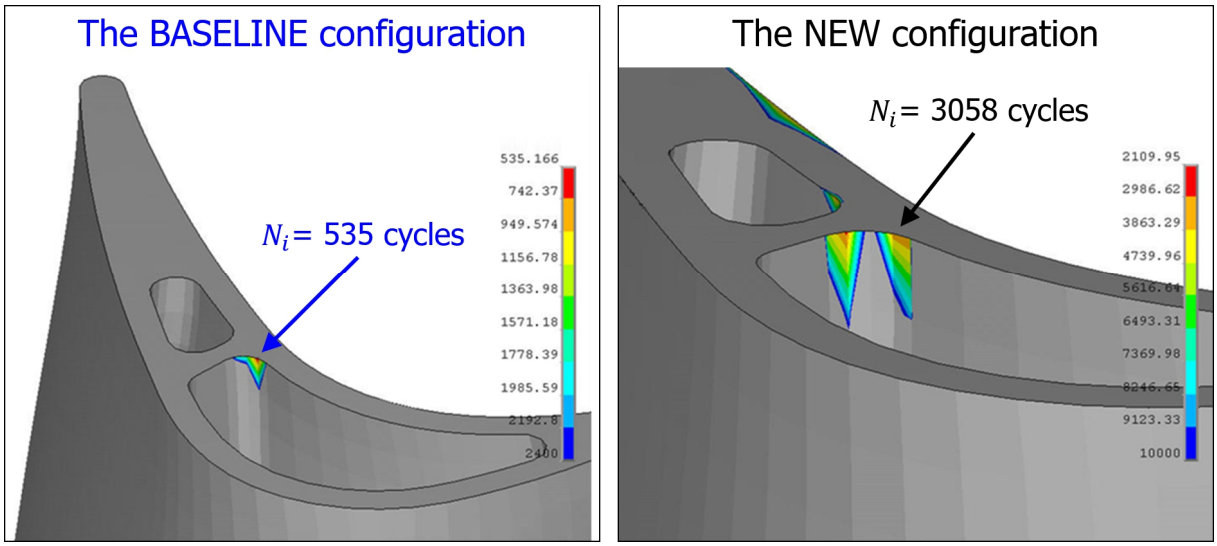


Fig. 32 Results of the numerical simulation at the critical location of the squealer tip: the number of cycles to initiate the fatigue crack N_i for the baseline and new configurations.

data are unavailable. Thus, the shape parameter of the new model is the one from the model dedicated to the baseline configuration, i.e. $\beta \cong 1.125$, meaning that failures will randomly occur. The second assumption is taken to determine the scale parameter value of the new model. It is based on the numerical simulation which results are presented in Fig. 32 [210]. At the critical location of the blade's tip the number of cycles to initiate the fatigue crack equals $N_{i,old}$ = 535 for the baseline configuration and $N_{i,new}$ = 3058 for the part made of the new material. The ratio between these two values $N_{i,new}/N_{i,old} \cong 5.715888$ is used to multiply the value of η parameter of the baseline model and determine the scale parameter value of the new model, i.e. $\eta_{new} = \eta_{old}\left(N_{i,new}/N_{i,old}\right) \cong 84954 \cdot 5.715888 \cong 485588$ accumulated factored fired starts. The characteristic life is the estimated time when 63.2% of components are expected to fail, but the same ratio applies to all other percentiles. Thereby, on a log-log plot the new model is represented by a straight line parallel to the baseline model's line (see Fig. 31), and the value of $N_{i,new}/N_{i,old}$ determines the shift between them. It is another example in which knowledge is extracted in one domain, where empirical data are available, captured in β parameter value, and subsequently transferred and applied in another domain, where empirical data is not yet available. Moreover, in comparison with the two data-driven methods described previously, this approach is more comprehensive since it leverages the results of finite element analysis.

Despite a simple mathematical description of those methods, their effectiveness was proved in practical applications, allowing the creation of survival models in case of shortage or lack of empirical data. It is an inspiration for further research related to physics-informed neural networks. Considering many similarities between models based on a Weibull distribution and PINN-based models, the question can be posed, whether the latter can be leveraged in other domains in such a simple yet effective way. This question can be answered affirmatively based on the results of further research described in this and the next chapter of this dissertation.

Further research activities intend to assess whether a hybrid model based on a physics-informed neural network aimed at predicting the size of damage of a specific component based on actual operational data, which is trained in a domain where damage size measurements are

available, can be appropriately shifted or scaled, similarly to the previously described methods applicable to a Weibull distribution in order to obtain accurate damage size predictions in another domain, in which the number of damage size measurements is limited, or even the data are unavailable. The idea is to assume that the hybrid model's core remains unchanged and is the same in both domains (i.e. analogous to the methods related to a Weibull distribution, where the value of β parameter does not change during the shifting process) and properly scale the damage increments determined by the model using a factor that generally captures the differences between the two domains, to get accurate estimates in the domain characterized by a limited number of the damage size measurements. As already explained, in the case of the models described in this chapter aimed at predicting the maximal length of fatigue cracks at Position 2 of the analyzed 1st stage nozzles, the feedforward neural network responsible for the stress intensity factor estimates is considered as the core of the hybrid model. Training a neural network can be seen as extracting latent information and knowledge from the available data. An effectively trained neural network can be applied to make predictions leveraging that knowledge. For the created hybrid models, the multilayer perceptron was trained based on actual empirical data but also considering the constraints resulting from Eq. (57), which were implemented in a soft manner in the cost function. Therefore, the intention is to apply this hybrid kernel, which unites knowledge extracted from the empirical data with the utilized theoretical relationship for predicting fatigue crack lengths in different locations of the analyzed high-pressure nozzles. In further consideration, the analyzed object, the damage type and its location and the entire set of data used to train and test the fully-specified hybrid model will be referred to as the source domain. The hybrid model successfully created in the source domain will be called the source model. Since the reference model, which was trained using ten observations, gives the most accurate predictions, it is selected as the source model. While the object to which the source model will be applied, the damage location and the set of operating parameters required to prepare input of the physics-informed neural network will be referred to as the target domain. Simultaneously, it is assumed that the damage type and physical phenomena causing it are the same in both domains and the theoretical relationships embedded into the source model also apply to the target domain. In order to reduce the time and costs needed to gather additional empirical data, the target domain refers to fatigue cracks found at Position 4 of the previously analyzed high-pressure nozzles. The target location is not on the same airfoil as Position 2. The maximum crack size at Position 4 is the dependent variable in the target domain. It is consistent with the approach applied to train the source model. The intent is to avoid applying the damage size measurements available in the target domain, i.e. 13 observations, for training purposes but only to test the model performance on the unseen data.

The conceptual diagram describing hypothetical steps needed to be taken to apply the source model in the target domain effectively is shown in Fig. 33. Terms related to the source and the target domains are denoted by subscripts *SRC* and *TGT*, respectively. These steps can be described as follows:

1) Determine how to estimate the damage increment in a single time step t in the source domain $\widehat{\Delta y}_{THEOR,SRC,t}$ based on the available operational data $\vec{x}_{SRC,t}$ recorded by the data acquisition system using prior knowledge about the analyzed phenomenon. The input data should be processed to get the final estimate based on theoretical or experimentally determined equations describing the damage growth. In the fourth step, these equations will be leveraged either to prepare input data for the physics-

informed neural network or directly embed them into the neural network's structure in a soft or hard manner. The complete set of equations used to estimate the damage increment is referred to as the theoretical model. Finally, the damage increments should be calculated for each available time step.

2) The theoretical model created in the previous step should be applied in the target domain to estimate the damage increments $\widehat{\Delta y}_{THEOR,TGT,t}$ for each time step based on the available operating parameters $\vec{x}_{TGT,t}$ recorded in this domain.
3) The objective of this step is to determine a map M, which purpose is to capture the proportions between $\widehat{\Delta y}_{THEOR,TGT,t}$ and $\widehat{\Delta y}_{THEOR,SRC,t}$ (or the shift, in analogy to the methods applicable to a Weibull distribution) based on the damage increment estimates obtained for each time step in both domains using the theoretical model. Considering the analyzed case of high-pressure nozzles fatigue cracks, it is assumed that the two domains refer to the same gas turbine, that the available operational data were recorded in the same period and are available for each timestamp in both domains. With these assumptions set, the proportion between the increments can be calculated individually for each time step. A generalized form of the map can also be determined based on the individual values using a function approximation if it does not introduce a significant error.
4) Design a physics-informed neural network architecture leveraging the same set of equations describing the damage growth, which was used to create the theoretical model at the first step. Applying the same framework of equations should align the hybrid model with the theoretical one and impose the same way of data processing. Identify variables for which the theoretical model provides estimates which are not accurate or uncertain and they can be directly approximated using the physics-informed neural network. Such inaccuracies may result from general problems

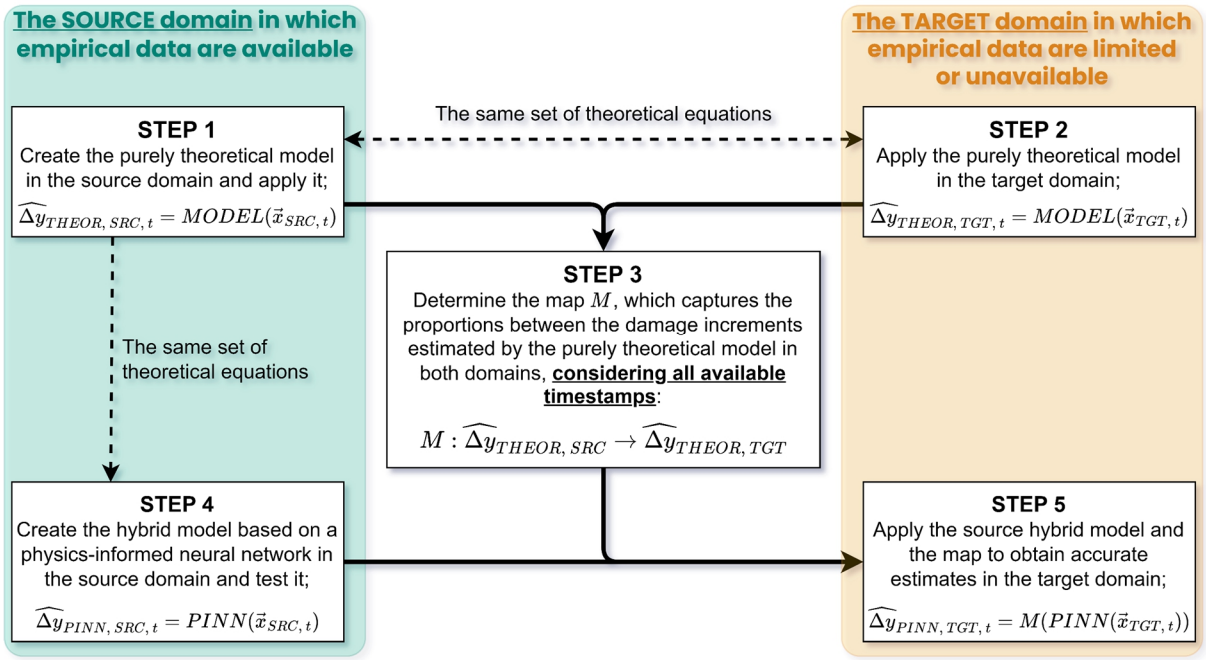


Fig. 33 The conceptual diagram describing hypothetical steps needed to be taken to effectively apply the source model based on a physics-informed neural network in the target domain.

related to the availability and quality of the operational data, various assumptions taken during the analysis and simplification applied during the data processing and preparing the set of equations composing the theoretical model. In the analyzed case of high-pressure nozzles the stress intensity factor at shutdown $K_{\sigma_{shdn}}$ was selected as the variable to be approximated by means of the PINN. Actual measurement data on the damage size available in the source domain must be used during the training process. The learning process is aimed at extracting latent information and knowledge from the available empirical data while respecting the constraints imposed by the applied framework of theoretical equations. The inclusion of actual measurement data in the training process should increase the prediction's accuracy compared to the purely theoretical model, as observed in the analyzed case of 1st stage nozzles. In this step, assessing the hybrid model's performance on the unseen data and the generalization capabilities is necessary. If the test results are positive, the last step can be executed.

5) The hybrid model created in the previous step should be applied in the target domain to estimate the damage increments $\widehat{\Delta y}_{PINN,TGT,t}$ for each time step based on the available operating parameters $\vec{x}_{TGT,t}$ recorded in this domain. The model created in the source domain is employed in the target domain in an unaltered form. Afterward, the map M is applied to each damage increment $PINN\left(\vec{x}_{TGT,t}\right)$ determined by means of the physics-informed neural network to appropriately adjust the values estimated by the model and obtain the final estimates. Since the hybrid and theoretical models are based on the same set of equations describing the damage growth, and both of them were applied in an unchanged form in the source and target domain, it is assumed that the proportions between the damage increments calculated by means the theoretical model in both domains described with the map $M : \widehat{\Delta y}_{THEOR,SRC} \rightarrow \widehat{\Delta y}_{THEOR,TGT}$ are the same in the case of damage increments determined by the hybrid model in the two domains. It is the key assumption of the entire process described herein. Consequently, for each time step, the final estimate in the target domain is calculated as $\widehat{\Delta y}_{PINN,TGT,t} = M\left(PINN\left(\vec{x}_{TGT,t}\right)\right)$.

This five-step process was successfully applied for the analyzed 1st stage nozzles to predict the maximal crack length at Position 4 using the reference hybrid model dedicated to Position 2. This part of the research is described in the next section. Nevertheless, a specific mathematical operation exists, which is an enabler of this process and determines its effectiveness. Fig. 34 presents how the thermal stresses at shutdown σ_{shdn}, which are the main input of the physics-informed neural network, depend on the values of the average air temperature at the axial compressor's discharge T_{COOL} and the average temperature of gases at the outlet of the nozzles T_{FIRE}. The actual values of the variables were divided by the maximum value of T_{COOL}, T_{FIRE}, or σ_{shdn} in order to obtain the normalized variables, i.e. T'_{COOL}, T'_{FIRE} and σ'_{shdn}. The presented results (i.e. the same data are shown in each view) refer to the source and the target domain. The scatter plots present the normalized values of the operating parameters recorded by the data acquisition system, which are the same for a single time step in both domains, and the thermal stress values in the two domains. In the source domain, the thermal stresses are estimated using

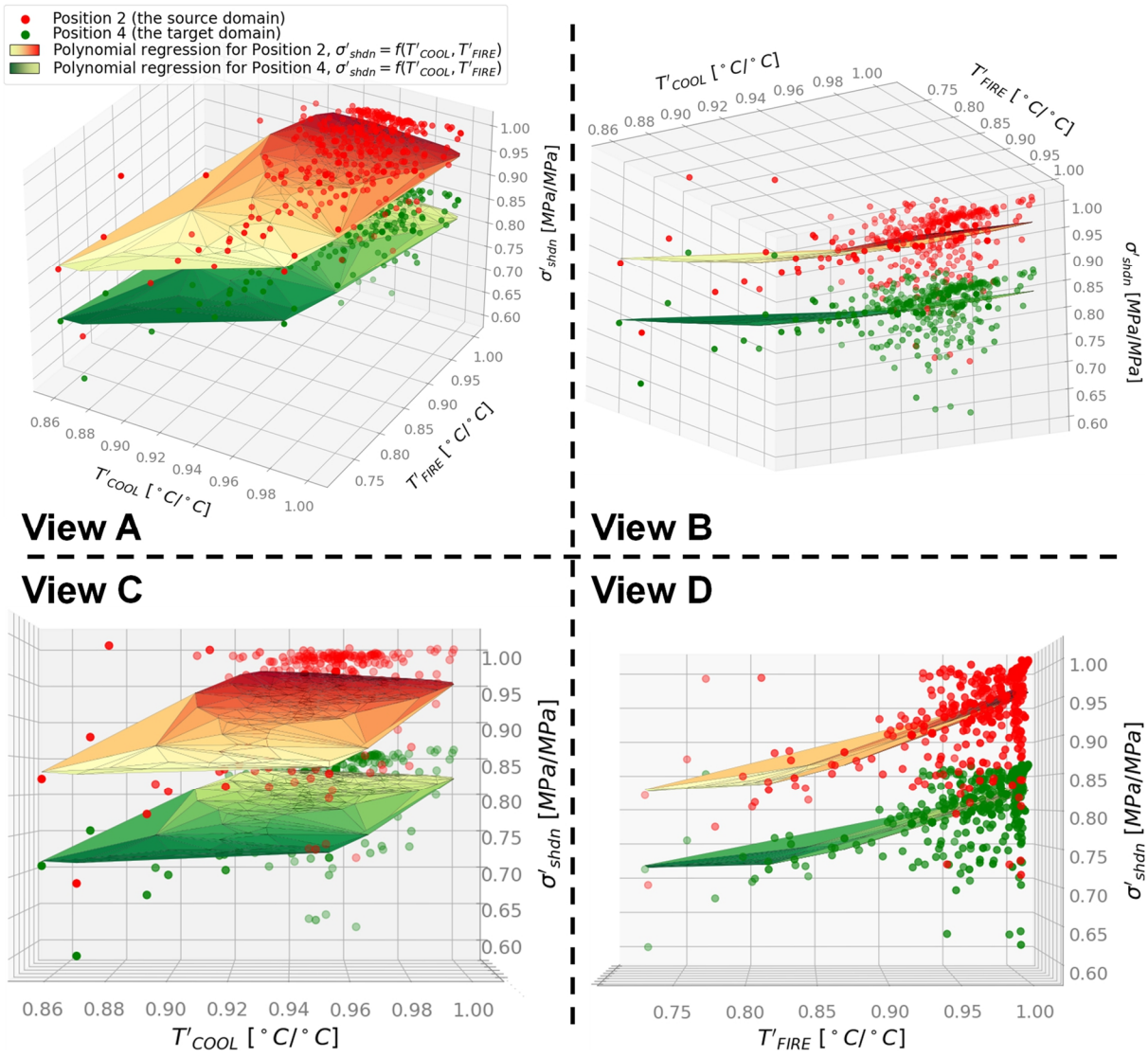


Fig. 34 The relationship between the average air temperature at the axial compressor's discharge T_{COOL}, the average temperature of gases at the outlet of the nozzles T_{FIRE} and the thermal stresses at shutdown σ_{shdn} in the source and target domains, i.e. Position 2 and Position 4, respectively.

Eq. (51). The same approach is applied to calculate the stresses in the target domain, which is based on the following formulas:

$$\begin{aligned}\sigma_{shdn,TGT} &= \left(\varepsilon_{L,TGT} - \varepsilon_{S,TGT}\right) E_{\overline{T}_{S,shdn,TGT}} = \\ &= \left[\alpha_{\overline{T}_{L,shdn,TGT}} \left(\overline{T}_{L,shdn,TGT} - T_{AMB}\right) - \alpha_{\overline{T}_{S,shdn,TGT}} \left(\overline{T}_{S,shdn,TGT} - T_{AMB}\right)\right] E_{\overline{T}_{S,shdn,TGT}}\end{aligned} \tag{68}$$

and

$$\overline{T}_{L,shdn,TGT} = \lambda_5 \overline{T}_{L,ss,TGT} = \lambda_5 \left[\lambda_7 T_{HOT,ss,TGT} + \left(1-\lambda_7\right) T_{COLD,ss,TGT}\right], \tag{69}$$

$$\overline{T}_{S,shdn,TGT} = \lambda_6 \overline{T}_{S,ss,TGT} = \lambda_6 \left(\lambda_8 T_{HOT,ss,TGT}\right), \tag{70}$$

$$T_{HOT,ss,TGT} = T_{FIRE} - \phi_{HOT,TGT}\left(T_{FIRE} - T_{COOL}\right), \tag{71}$$

$$T_{COLD,ss,TGT} = T_{FIRE} - \phi_{COLD,TGT}\left(T_{FIRE} - T_{COOL}\right), \tag{72}$$

where the subscript *TGT* indicates variables dedicated to the target domain, while the general meaning of all the symbols is the same as in Equations (51)−(55). It should be clearly stated that the values of the coefficients λ_5, λ_6, λ_7, λ_8, $\phi_{HOT,TGT}$ and $\phi_{COLD,TGT}$ were determined by means of the numerical simulation. They are specific to Position 4, and none of them apply to the source domain, i.e. Position 2. In order to better visualize, how the thermal stresses change depending on the variation of T_{COOL} and T_{FIRE}, polynomial regression models are prepared to approximate the relationship between these two operating parameters and the input variable of the physics-informed neural network:

$$\hat{\sigma}_{shdn} = f\left(T_{COOL}, T_{FIRE}\right), \tag{73}$$

where $\hat{\sigma}_{shdn}$ is the estimated value of the thermal stresses at shutdown. Due to the low variation of values of the ambient air temperature T_{AMB} (i.e. the mean equals $\overline{T_{AMB}}$ = 28.2 °C with a standard deviation $\sigma_{T_{AMB}}$ = 2.7 °C), a relatively small impact of T_{AMB} changes on the thermal stresses, and because the polynomial regression models were created for visualization purposes only, the ambient air temperature is not considered as the models' input variable. One model is dedicated to the source domain and another to the target domain. The surface plots shown in Fig. 34 present the normalized results obtained using the two polynomial regression models. Based on these plots, it is visible that the surfaces are shifted relative to each other with the length of the displacement vector approximately constant independently of the selected pair of T_{COOL} and T_{FIRE}. Using the actual operational data and the thermal stress values calculated by means of Eq. (51) and Eq. (68) (i.e. the data presented with the scatter plots in Fig. 34), a scaling factor between the values of σ_{shdn} and $\sigma_{shdn,TGT}$ can be determined for each startup-shutdown cycle. The mean value of the factor equals $\bar{\varphi}$ = 0.8632 with a standard deviation σ_{φ} = 0.0016. Thus, considering a specific nozzle segment and operating cycle, the calculated value of maximum thermal stresses at shutdown related to Position 4 is equal, on average, to 86.32% of the stress value calculated for Position 2. It can be assumed that the factor is constant in the analyzed training space for any pair of the input operating parameters. Therefore, as a result of this assumption, the applicability of the factor is no longer limited only to a single time step, but it can be applied independently of time. Thus, for any triple comprising T_{AMB}, T_{COOL} and T_{FIRE}, the value of thermal stress at shutdown in the target domain $\sigma_{shdn,TGT}$ can be calculated by means of Eq. (68). Moreover, with additional constraints taken into account, it can be determined, which triplet of the operating parameters results in the same value of thermal stress $\sigma_{shdn} = \sigma_{shdn,TGT}$ in the source domain. Applying again Eq. (68) to this newly determined triple results in another value, which will be referred to as the shifted value of thermal stress in the target domain and denoted by $\sigma^{*}_{shdn,TGT}$. In summary, in the analyzed case, it is possible to determine a relationship between the thermal stress values in the source and the target domain using the Equations (51)−(55) and (68)−(72), relying on the common set of operating parameters in both domains. It should be reminded that the thermal stress at shutdown is the main input variable of the hybrid model based on the physics-informed neural network.

Although the source model respects the constraints resulting from Eq. (57) and has good generalization capabilities, it is dedicated to Position 2. During the optimization process, the empirical term $RMSE_{EMP}$ was minimized, which is based on the crack length measurements recorded at Position 2. Thus, the crack length increments and final estimates are adjusted to the source domain. Consequently, applying the source model directly in the target domain on the unchanged input data from the target domain will not yield accurate estimates for that domain. As explained, these estimates will be adjusted to the source domain, despite the input data from the target domain. Nevertheless, it is possible to increase the prediction's accuracy by leveraging the relationship between σ_{shdn} and $\sigma_{shdn,TGT}$ described in the previous paragraph. For further consideration, the set of Equations (51)−(55) is denoted by f_{SRC}, while the set of Equations (68)−(72) is denoted by f_{TGT}. These two functions are defined as follows:

$$\sigma_{shdn} = f_{SRC}\left(T_{AMB}, T_{COOL}, T_{FIRE}\right), \tag{74}$$

$$\sigma_{shdn,TGT} = f_{TGT}\left(T_{AMB}, T_{COOL}, T_{FIRE}\right). \tag{75}$$

Considering any startup-shutdown cycle with the required operational data available (i.e. a triplet comprising T_{AMB}, T_{COOL} and T_{FIRE}), it is proposed to execute the following steps in order to increase the prediction's accuracy in the target domain by applying the unchanged physics-informed neural network created and trained in the source domain:

1) Calculate the value of thermal stress at shutdown in the target domain $\sigma_{shdn,TGT}$ using function f_{TGT} in accordance with Eq. (75).
2) Determine the values of operating parameters resulting in the same value of the thermal stresses but in the source domain, i.e. $\sigma_{shdn} = \sigma_{shdn,TGT}$. The resultant values are denoted as T^{*}_{AMB}, T^{*}_{COOL} and T^{*}_{FIRE}, and referred to as the shifted values of the operating parameters. In a generalized case, in this step, it is necessary to determine the inverse of f_{SRC}, denoted by f^{-1}_{SRC}, if it exists. In the analyzed case, function f_{SRC} is not bijective since many different triples may result in the same σ_{shdn} value. Therefore, this function is not invertible. However, the inability to determine the inverse function f^{-1}_{SRC} is not a problem and does not prevent the execution of the next steps since, in the analyzed case, the ratio between the values of $\sigma_{shdn,TGT}$ and σ_{shdn} is constant and equal to $\overline{\varphi}$ = 0.8632. In a generalized case, to determine the values of T^{*}_{AMB}, T^{*}_{COOL} and T^{*}_{FIRE} unambiguously and reduce the degrees of freedom, it will be required to define certain assumptions regarding the values of some of the operating parameters or prepare additional equations describing relationships between the selected operating parameters and other dependent variables, different than σ_{shdn}. Such cases are described in the next chapter.
3) Considering T^{*}_{AMB}, T^{*}_{COOL} and T^{*}_{FIRE} as the input data, calculate the shifted value of the thermal stress in the target domain $\sigma^{*}_{shdn,TGT}$ again using function f_{TGT}.
4) Use $\sigma^{*}_{shdn,TGT}$ as the input variable of the unchanged source model to estimate the crack length increment during a single startup-shutdown cycle. Therefore, the final estimate in the target domain for the original triple (i.e. T_{AMB}, T_{COOL} and T_{FIRE}) is obtained by means of the source model applied to the shifted triple comprising T^{*}_{AMB}, T^{*}_{COOL} and

T^*_{FIRE}. In addition, it should be clarified, why the second input variable of the physics-informed neural network is neglected in these considerations. For a specific time step, the ratio between the values of C_c coefficient in both domains is close to one in the analyzed case. Thus, the prediction's accuracy should be maintained if the original, unshifted values of $C_{c,TGT}$ calculated in the target domain are used for estimation purposes. Nevertheless, it should be reminded that C_c is a monotonically increasing function of $\overline{T}_{S,ss}$, which inverse can be easily determined.

Functions f_{SRC} and f_{TGT}, representing the set of equations, respectively (51)−(55) and (68)−(72), were introduced to prepare more specific input data for the physics-informed neural network, simplify the processing of data within the PINN's structure and make the training process and the interpretation of the obtained results easier. These functions constitute a simple physics-based model used to estimate the thermal stresses on the basis of the actual operational data, which is also adjusted based on the numerical simulation results. Simultaneously, these functions determine the relationship between the inputs of the hybrid model in the source and the target domain, leaning on the same set of operating parameters shared across the domains. Functions f_{SRC} and f_{TGT} are the enablers of the four-step process described above, which is expected to provide accurate damage size estimates in case of limited availability or even lack of actual damage measurements. This process can be seen as a method of transferring knowledge between domains related to different operating conditions of the same unit, and also between different parts or units with unique characteristics, assuming that the damage type and physical phenomena causing it are the same in the two domains, and the theoretical or experimentally determined relationships embedded in the source model based on a physics-informed neural network apply to the target domain as well. Moreover, in the described process, the independent variables from the target domain, i.e. the operating parameters recorded by the data acquisition system, are not utilized in the training of the source hybrid model, similarly to the damage size measurements, which may not even be available. Thus, this process can also be seen as a method of single-source domain generalization in regression analysis, in which a specific strategy is applied to create a model with good generalization capabilities. Fig. 35 presents the first three steps of the process described above, considering a situation in which the single input variable of a physics-informed neural network depends on one operating parameter only and functions f_{SRC} and f_{TGT} are bijective and invertible. This simple case illustrates, why shifting the neural network's input data is necessary and how to accomplish it. These three steps, which are denoted in the figure as (1), (2) and (3), are as follows:

1) For a single time step, for which the operating parameter *x* considered as the predictor has a value equal to x_1, i.e. this reading was recorded in the target domain, calculate the corresponding value of *v*, which is the input variable of the physics-informed neural network trained in the source domain, applying the equation below:

$$v_{TGT} = f_{TGT}\left(x_1\right). \tag{76}$$

2) Using the inverse function of f_{SRC}, determine the operating parameter value resulting in the same value of the PINN's input variable, i.e. v_{TGT}, but in the source domain, following the formula below:

$$x_1^* = f_{SRC}^{-1}\left(v_{TGT}\right). \tag{77}$$

The resultant value x_1^* is referred to as the shifted value of the operating parameter.

3) Calculate the adjusted value of the neural network's input in the target domain using the equation below:

$$v_{TGT}^* = f_{TGT}\left(x_1^*\right). \tag{78}$$

The resultant value v_{TGT}^* is referred to as the shifted value of the hybrid model's input. For the analyzed time step, for which the operating parameter's value equals x_1, v_{TGT}^* instead of v_{TGT} should be considered as the input of the unchanged source model when applied directly in the target domain. Such an adjustment of the PINN's input variable should increase the prediction's accuracy in the target domain.

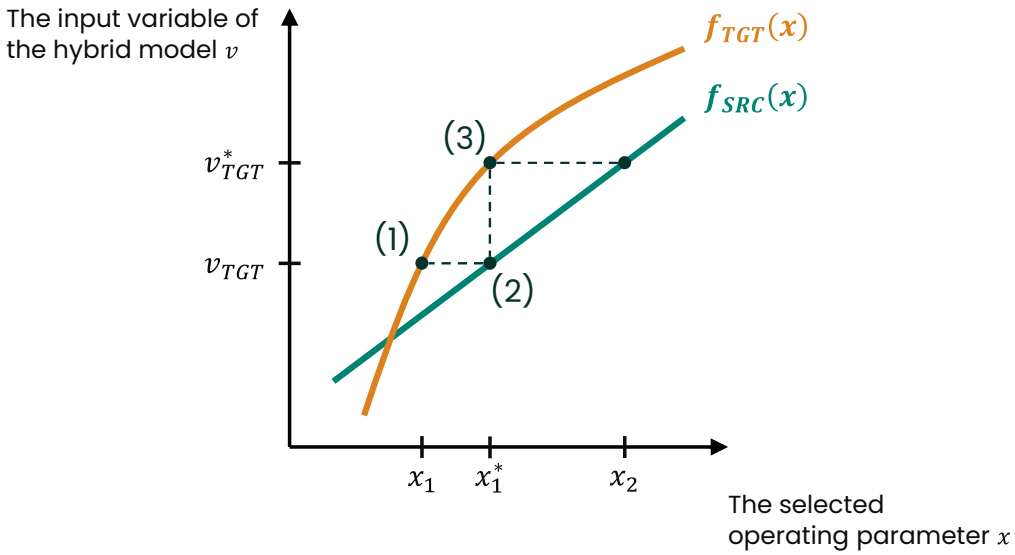


Fig. 35 An illustration of the three-step process applied to adjust the hybrid model's input in the target domain from the original value $v_{TGT} = f_{TGT}\left(x_1\right)$ to the shifted value $v_{TGT}^* = f_{TGT}\left(x_1^*\right) = f_{TGT}\left(f_{SRC}^{-1}\left(v_{TGT}\right)\right) = f_{SRC}\left(x_2\right)$, where the inverse of f_{SRC} function is denoted as f_{SRC}^{-1}.

As already mentioned, predictions in the target domain are obtained using the hybrid model based on a physics-informed neural network created and trained in the source domain. Thus, if the operating parameter in the target domains equals x_1 and the PINN's input variable is equal to v_{TGT}, then the estimated damage increment has the same value as the prediction obtained in the source domain when the operating parameter equals x_1^* since $f_{TGT}\left(x_1\right) = v_{TGT} = f_{SRC}\left(x_1^*\right)$. Simultaneously, when the shifted value of the operating parameter is determined as $x = x_1^*$, function f_{TGT} can be applied to calculate the corresponding value of the PINN's input variable, denoted as v_{TGT}^*, where $v_{TGT}^* = f_{TGT}\left(x_1^*\right) = f_{SRC}\left(x_2\right)$. Finally, if the operating parameter's value recorded at time step t in the target domain is equal to x_1, then the resultant damage increment can be estimated by the source hybrid model as follows:

$$\begin{aligned}\widehat{\Delta y}_{PINN,TGT,t} &= PINN\left(v_{TGT}^{*}\right) = PINN\left(f_{TGT}\left(x_1^{*}\right)\right) = \\ &= PINN\left(f_{TGT}\left(f_{SRC}^{-1}\left(v_{TGT}\right)\right)\right) = PINN\left(f_{TGT}\left(f_{SRC}^{-1}\left(f_{TGT}\left(x_1\right)\right)\right)\right),\end{aligned} \tag{79}$$

where function *PINN* symbolically denotes the physics-informed neural network trained in the source domain.

The above considerations regarding the hypothetical steps necessary to apply the source hybrid model in the target domain effectively, which are illustrated in the conceptual diagram in Fig. 33, and the steps needed to appropriately adjust the input data values of that model in the target domain can be combined to form a systematic procedure, which can be applied in cases similar to the analyzed one, aimed at modeling the process of progressive, irreversible damage accumulation based on real operational data recorded by a data acquisition system. This procedure establishes a novel method of single-source domain generalization and cross-domain knowledge transfer in regression analysis leveraging physics-informed neural networks. In general, the main steps of this method are as follows:

1) Determine how to estimate the damage increment in a single time step t based on the operational data $\vec{x}_{SRC,t}$ recorded in the source domain, i.e. it is assumed that the feature selection process was executed following best practices, using prior knowledge about the analyzed phenomenon. Theoretical or experimentally determined equations should be applied to process the input data accordingly and obtain the final estimate. As a reminder, the source domain is the analyzed object, the damage type and its location and the entire set of data available for training and testing a regression model.

2) Design a physics-informed neural network architecture leveraging the equations formulated in the prior step. Identify variables whose values are insufficiently accurately estimated by means of the theoretical equations or whose prediction uncertainty is high and can be directly approximated using the physics-informed neural network. It is essential to derive a complete sequence of functions that are applied to calculate the PINN's input variable $v_{SRC,t}$ based on the selected operating parameters $\vec{x}_{SRC,t}$. The composition of all these functions applicable to the source domain is denoted by f_{SRC}. Any constraints resulting from prior knowledge about the analyzed phenomenon restricting the space of allowable solutions can be implemented either in the soft or the hard manner, i.e. using an additional loss term in the cost function or customizing the neural network's structure to ensure a constraint is satisfied. Synthetically created data can enlarge the training space in which the model should respect the constraints.

3) Execute the training process of the physics-informed neural network using the actual operational data and measurement data on the damage size, which are available in the source domain. The optimization aims to extract latent information and knowledge from the available empirical data, maximizing the prediction's accuracy while respecting the constraints embedded in the neural network's structure. Additionally, leveraging the availability of the empirical data in this domain, it is necessary to assess the hybrid's model performance on the test data and the generalization capabilities. A particular emphasis should be put on those parts of the neural network used to estimate the values of the variables identified in the prior step, i.e. the variables inaccurately estimated by

means of the theoretical formulas. Characteristics showing the relationship between the input data and the output must be prepared and evaluated, focusing on consistency with the prior knowledge about the analyzed phenomenon. The fully-specified PINN is called the source model or the source hybrid model.

4) For explanation purposes, an assumption is taken that applying the source model in another domain is requested, where data on the same operating parameters are available, but the damage size measurements are unavailable. Therefore, this step aims to evaluate if the source hybrid model can be safely employed in that domain and provide credible estimates. The evaluation's result determines whether that domain is a valid target domain. As a reminder, the target domain is the specific object to which the source model will be applied, the damage location and the set of operating parameters required to prepare the input for the model. Nevertheless, the damage type in the target domain, the physical phenomena causing it and the applicable theoretical or experimentally determined equations must be the same as in the source domain.

5) Considering all available operational data recorded during the service period in the target domain $\mathbf{X}_{TGT}$ as the input data, calculate the values of the PINN's input variable $\vec{v}_{TGT}$. The selection of the operating parameters taken as the independent variables should be the same as in the source domain. It is expected that the composition of functions applied to calculate the values of $\vec{v}_{TGT}$ components should be similar to the one applied in the source domain, i.e. f_{SRC}, and simultaneously these functions should capture all specific properties of the target domain, which significantly affect the value of $\vec{v}_{TGT}$ elements. The composition of all these functions applicable to the target domain is denoted by f_{TGT}. Numerical simulation results can be leveraged in order to adjust functions f_{SRC} and f_{TGT} to their corresponding domains. A particular emphasis should be put on the verification of the range of values of the $\vec{v}_{TGT}$ vector components. These values should be inside the space of input data used to train the physics-informed neural network in the source domain. This approach aims to consciously and prudently predict in the target domain, using the hybrid model for interpolation only. Otherwise, it is suggested to retrain the model enlarging the training space by synthetically created data, i.e. to enlarge the space where the PINN respects the theoretical constraints, rather than using the original version of the hybrid model for extrapolation.

6) Determine the inverse function f_{SRC}^{-1}, which exists if and only if f_{SRC} is bijective. Else, f_{SRC} is not invertible, for example, because it is a non-injective function or a function of several variables, i.e. $v_{SRC,t} = f_{SRC}\left(\vec{x}_{SRC,t}\right)$. In such a case, additional operations will be needed to determine the components of $\vec{x}_{SRC,t}$ unambiguously, given any value of $v_{SRC,t}$. These operations may relate to defining certain assumptions regarding the values of the PINN's input variables or the values of some of the operating parameters in the source and target domains or preparing additional equations, which describe relationships between the selected operating parameters and other dependent variables, different than v. The ultimate objective of this step is to well-define the components of $\vec{x}_{SRC,t}$ for each time step given $v_{TGT,t}$ as the input, using the following equation:

$$\vec{x}_{SRC,t} = \vec{x}^{*}_{TGT,t} = f_{SRC}^{-1}\left(v_{TGT,t}\right) = f_{SRC}^{-1}\left(f_{TGT,t}\left(\vec{x}_{TGT,t}\right)\right), \tag{80}$$

where the elements of vector $\vec{x}^{*}_{TGT,t}$ are referred to as the shifted values of the operating parameters in the target domain. Considering the entire service period, the matrix with the shifted values of the operating parameters is denoted as $\mathbf{X}^{*}_{TGT}$. Based on Eq. (80), the knowledge of $\vec{x}_{TGT,t}$ can be leveraged in defining the assumptions mentioned above regarding the values of the vector $\vec{x}_{SRC,t}$ components. As already mentioned, such cases are described in the next chapter.

7) Considering $\mathbf{X}^{*}_{TGT}$ as the input, apply function f_{TGT} to calculate $\vec{v}^{*}_{TGT} = f_{TGT}\left(\mathbf{X}^{*}_{TGT}\right)$, which is referred to as the vector of shifted values of the PINN's input variable in the target domain.

8) Use $\vec{v}^{*}_{TGT}$ as the input data of the unchanged source hybrid model to estimate the damage increment in the target domain for all the time steps in the analyzed service period. Thus, the final estimate in the target domain for the original set of operational data $\mathbf{X}_{TGT}$ is obtained by applying the source model to the set of shifted operational data $\mathbf{X}^{*}_{TGT}$.

Additionally, for clarity, if more than one variable calculated based on operational data is used as the input of the source hybrid model, then each variable is expected to be determined by means of a dedicated set of equations. Therefore, functions f_{SRC}, f_{TGT} and f^{-1}_{SRC} are specific to each physics-informed neural network input variable, and it is necessary to determine the shifted values for each variable separately unless some of the input variables have the same values in both domains for any set of operational data.

5.5 Application of the proposed method for predicting fatigue cracks growth

As described in the previous section, the selected target domain refers to fatigue cracks observed at Position 4 of the airfoil's trailing edge of the analyzed high-pressure nozzles. The maximum crack size is considered the dependent variable. The reference model used to predict the maximal crack length at Position 2, which was trained using ten observations, is selected as the source hybrid model. The target location is not on the same airfoil as Position 2. The target domain is considered valid since the physical phenomena causing the cracks and the applicable framework of equations are the same as in the source domain. The thermal stresses in the target domain, which are the input of the source model based on the physics-informed neural network, are calculated by means of Equations (68)−(72). The shifted values of the PINN's input variable in the target domain are estimated using the factor $\overline{\varphi} = 0.8632$, which is the mean value of the quotient of $\sigma_{shdn,TGT}$ and σ_{shdn} determined considering all startup-shutdown cycles. Therefore, the calculated value of maximum thermal stresses at shutdown related to Position 4 is 86.32% of the stress value calculated for Position 2. Consequently, given a vector with the operating parameter's readings $\vec{x}_{TGT,t}$ recorded at time step t in the target domain, the shifted value of the thermal stresses is estimated in the following way:

$$\sigma^{*}_{shdn,TGT,t} = \overline{\varphi} \cdot \sigma_{shdn,TGT,t} = 0.8632 \cdot f_{TGT}\left(\vec{x}_{TGT,t}\right). \tag{81}$$

Subsequently, applying Eq. (47), the crack length increment for that startup-shutdown cycle is calculated based on the Paris' law as follows:

$$\left(\frac{da}{dN}\right)_{TGT,t} = \Delta a_{c,TGT,t} = C_{c,TGT,t}\left(\Delta K^{*}_{shdn,TGT,t}\right)^{m} = \\ = C_{c,TGT,t}\left[FNN\left(\sigma^{*}_{shdn,TGT,t}\right)\right]^{m} = C_{c,TGT,t}\left[FNN\left(0.8632 \cdot f_{TGT}\left(\vec{x}_{TGT,t}\right)\right)\right]^{m}, \tag{82}$$

where $\Delta K^{*}_{shdn,TGT,t}$ is the stress intensity factor value calculated based on the shifted value of the thermal stresses in the target domain using the feedforward neural network, symbolically denoted as function *FNN*. Considering the entire service period, the matrix with the shifted values of the thermal stresses is denoted as $\mathbf{S}^{*}_{shdn,TGT}$, and it can be used as the input data of the source model to estimate values of the dependent variable in the target domain. Thus, the final estimate in the target domain for the original values of the thermal stresses $\mathbf{S}_{shdn,TGT}$ is obtained by applying the source hybrid model to the set of shifted stresses $\mathbf{S}^{*}_{shdn,TGT}$. Finally, the damage size measurements, which are available in the target domain (i.e. 13 observations related to the same nozzle sets as in the source domain), will be used to evaluate the accuracy of the estimates obtained through the proposed method of single-source domain generalization and cross-domain knowledge transfer in regression analysis, which is based on physics-informed neural networks.

Moreover, it can be shown how to determine a map *M* and apply the process presented in Fig. 33. The map captures the proportions between the damage increments determined by means of the so-called theoretical model in the source and target domains, i.e. $\widehat{\Delta a}_{c,THEOR,TGT}$ and $\widehat{\Delta a}_{c,THEOR,SRC}$. For a single time step *t*, the proportion between the increments is calculated as follows:

$$M_t = \frac{\left(\frac{da}{dN}\right)_{THEOR,TGT,t}}{\left(\frac{da}{dN}\right)_{THEOR,SRC,t}} = \frac{\widehat{\Delta a}_{c,THEOR,TGT}}{\widehat{\Delta a}_{c,THEOR,SRC}} = \frac{C_{c,TGT,t}\left(\Delta K_{shdn,TGT,t}\right)^{m}}{C_{c,SRC,t}\left(\Delta K_{shdn,SRC,t}\right)^{m}} = \\ = \frac{C_{c,TGT,t}\left(\sigma_{shdn,TGT,t}\sqrt{a_{TGT,t-1}}Y_{TGT,t}\right)^{m}}{C_{c,SRC,t}\left(\sigma_{shdn,SRC,t}\sqrt{a_{SRC,t-1}}Y_{SRC,t}\right)^{m}}, \tag{83}$$

where the geometrical factor *Y* is calculated in accordance with Eq. (65). As described in the previous section, the ratio between the C_c coefficient values in both domains is close to one in the analyzed case. Additionally, the estimates obtained by means of the theoretical model are very similar in the two domains. Thus, assuming $C_{c,TGT,t} = C_{c,SRC,t}$, $Y_{TGT,t} = Y_{SRC,t}$ and $\sqrt{a_{TGT,t-1}} = \sqrt{a_{SRC,t-1}}$, Eq. (83) simplifies to:

$$M_t = \left(\frac{\sigma_{shdn,TGT,t}}{\sigma_{shdn,SRC,t}}\right)^{m}. \tag{84}$$

Leveraging the mean value of the quotient of $\sigma_{shdn,TGT}$ and σ_{shdn} equal to $\bar{\varphi} = 0.8632$, the map can be approximated by the following constant:

$$\bar{M} = 0.8632^{m}. \tag{85}$$

Subsequently, for a single time step t, the crack length increment is estimated as follows:

$$\left(\frac{da}{dN}\right)_{TGT,t} = \bar{M}\Delta a_{c,TGT,t} = \bar{M}\left[C_{c,TGT,t}\left(\Delta K_{shdn,TGT,t}\right)^{m}\right] = \\ = \bar{M}\left[C_{c,TGT,t}\left(FNN\left(\sigma_{shdn,TGT,t}\right)\right)^{m}\right] = \bar{M}\left[C_{c,TGT,t}\left(FNN\left(f_{TGT}\left(\vec{x}_{TGT,t}\right)\right)\right)^{m}\right]. \tag{86}$$

Differences between the final predictions of the maximal crack lengths at Position 4 determined by applying Eq. (86) resulting from the process illustrated schematically in Fig. 33 and the predictions obtained by means of Equation (82), which results directly from the use of the proposed method of domain generalization, are insignificant. Thus, to avoid excessive reporting, only the latter set of results is presented in the dissertation.

The results obtained by means of the created method of domain generalization and cross-domain knowledge transfer in regression analysis, which leverages physics-informed neural networks, are presented in Fig. 36 and Fig. 38. They can be compared with the results obtained by means of a hybrid model with the same architecture as the source PINN, which was trained from the ground up in the target domain using ten observations (i.e. these are related to the same nozzle sets that were selected to train the source model), which are presented in Fig. 37 and Fig. 39. The learning process of that model was executed using the actual operational data, the domain-specific values of $\sigma_{shdn,TGT}$ and $C_{c,TGT}$ and the measurement data on the damage size recorded in the target domain. In order to protect proprietary information, the data presented in the plots were normalized by dividing each value by the maximal measured crack length recorded in the target domain, which was determined based only on the data related to the analyzed 13 observations. The most important considerations regarding the obtained results are as follows:

- The pre-trained model provides significantly underestimated predictions of the final crack lengths in the target domain (the details are presented in Fig. 37), similarly to the attempt to apply it in the source domain.
- The accuracy of the predictions on the unseen data obtained using the proposed method is similar to the prediction accuracy of the hybrid model trained in the target domain. When the proposed method is applied, the normalized root mean squared error evaluated against all 13 observations is equal to $NRMSE_{EMP,TEST}$ = 15.5%, while considering nozzle sets with identifiers S1N_25, S1N_31 and S1N_47 as the reference, the error equals $NRMSE_{EMP,TEST}$ = 18.7% (i.e. these nozzle sets constitute the test subset for the model trained in the target domain). When the model trained from the ground up in the target domain is used, the error equals $NRMSE_{EMP}$ = 15.9% (with $NRMSE_{EMP,TRAIN}$ = 14.3% and $NRMSE_{EMP,TEST}$ = 20.2%). Thus, the obtained results provide evidence that the created method of domain generalization and cross-domain knowledge transfer in regression analysis, which leverages physics-informed neural networks, can be applied effectively even if no damage size measurements are available in the target domain, providing accurate damage size predictions. Such limitations may occur at the early stages of the parts' lifetime, e.g. during the design phase, first tests of the part or until the product reaches maturity. It can be presumed that leveraging a good-quality source hybrid model applying the created method of domain generalization may result in more

accurate predictions than training a model based on a limited number of observations, low-quality empirical data or both.

- The stress intensity factor estimates determined by the feedforward neural network trained in the target domain are lower than estimates obtained using the source hybrid

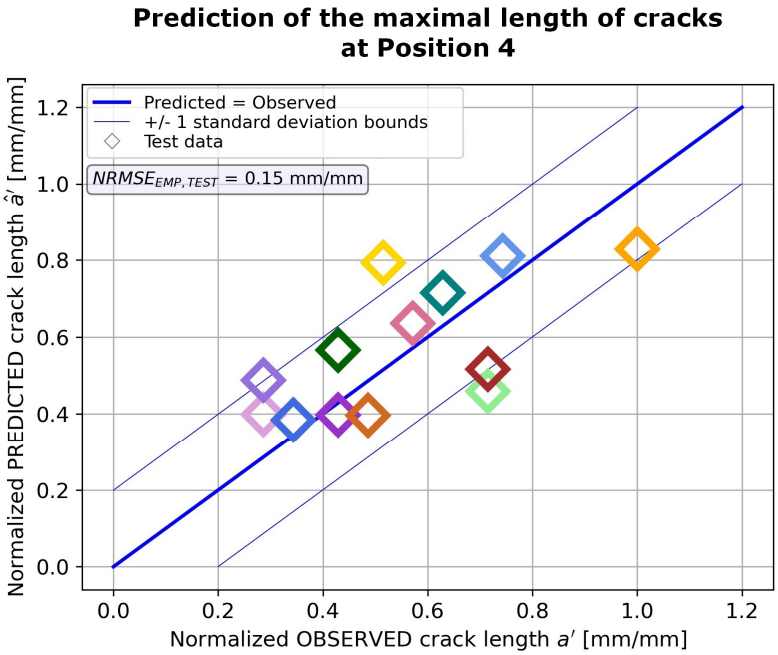


Fig. 36 Evaluation of the source hybrid model based on the physics-informed neural network, trained in the source domain using ten observations, against the test data from the target domain presented on the predicted vs. observed plot.

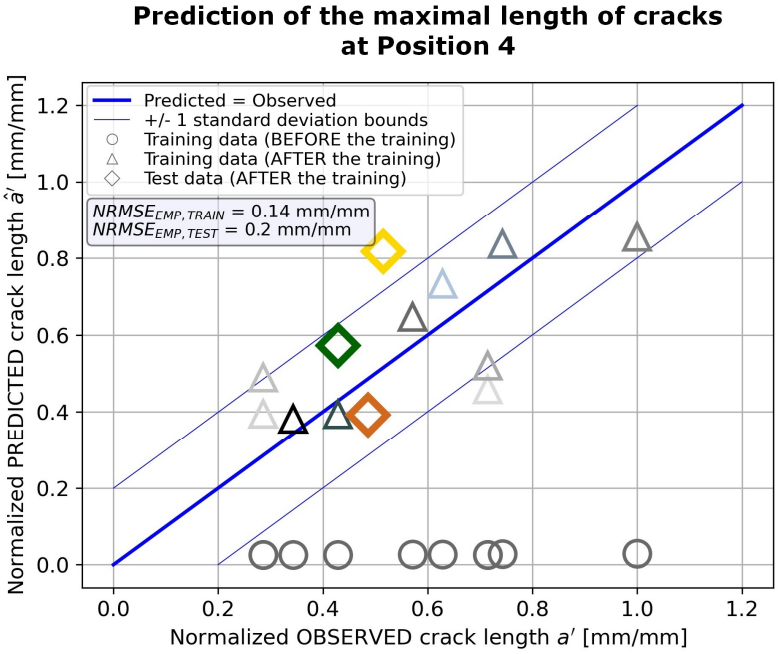


Fig. 37 Evaluation of the hybrid model based on the physics-informed neural network, trained from the ground up in the target domain using ten observations, against the training and test data presented on the predicted vs. observed plot.

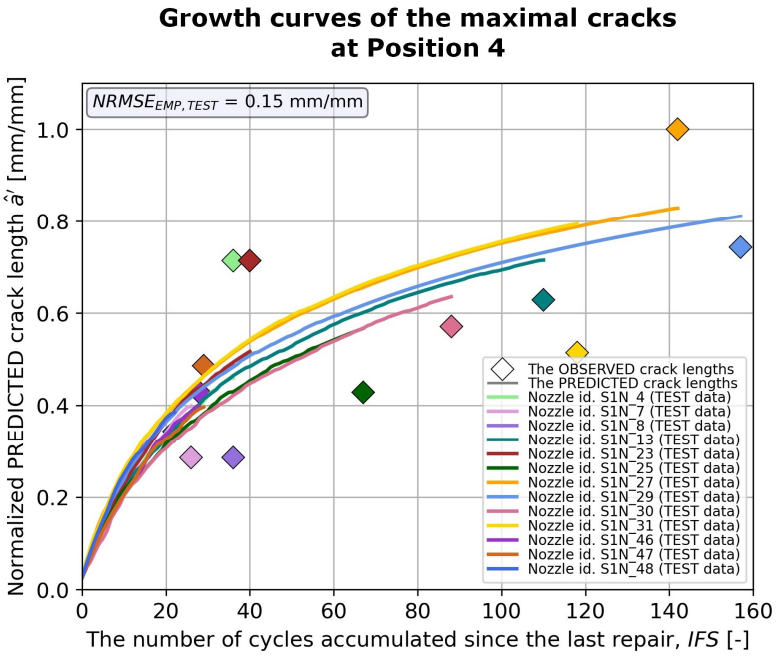


Fig. 38 Evaluation of the source hybrid model based on the physics-informed neural network, trained in the source domain using ten observations, against the test data from the target domain.

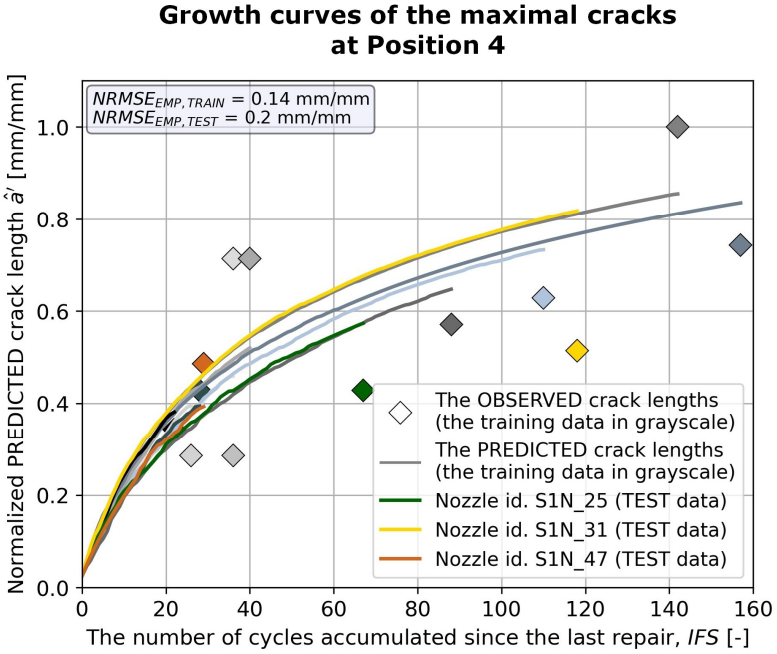


Fig. 39 Evaluation of the hybrid model based on the physics-informed neural network, trained from the ground up in the target domain using ten observations, against the training and test data.

model. It is because the same set of operational data describes both domains, but the cracks detected at Position 4 are shorter than the cracks recorded at Position 2 as shown in Fig. 4. Thereby, the stress intensity factor values obtained using the data-driven layer must be smaller to avoid overestimation of the final crack lengths. Fig. 40 presents the

characteristics of the multilayer perceptron at the constant thermal stresses, i.e. $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn} = const, a'\right)$, of the hybrid model trained in the target domain. The values presented in the plot were normalized by dividing them by the maximal measured crack length or the maximal calculated thermal stress related not to the target domain but to the source one, i.e. from Position 2. It simplifies a comparison with the results presented in Fig. 22 related to the source hybrid model. Comparing the results shown in these plots, it is visible that differences between the estimates decrease as the crack length increases and the thermal stresses reduce. Consequently, at the beginning of the crack propagation process, the crack length increments obtained using the source model are larger in coparison with the values obtained by means of the model trained in the target domain. Nevertheless, it does not significantly impact the final crack length estimates.

- Regardless of the very stable operating profile of the analyzed gas turbines, which is mainly characterized by a continuous duty at nominal power, the source hybrid model can differentiate the crack growth rates and the final estimates also in the target domain.
- Direct application of the source model in the target domain using the original values of the operating parameters (i.e. T_{AMB}, T_{COOL} and T_{FIRE}) instead of the shifted values (i.e. T^*_{AMB}, T^*_{COOL} and T^*_{FIRE}) results in overestimating the final crack lengths. As described in the previous section, regardless of the generalization capabilities of the source hybrid model, the learning process executed in the source domain was aimed at reducing the empirical error value related specifically to the data recorded in that domain. Thus, unless the same sequence of functions is applied in the source and target domains to calculate a PINN's input based on selected operating parameters, i.e. $f_{SRC} = f_{TGT}$, it is needed to determine shifted values of the neural network's input and make predictions in the target domain based on them, applying the source hybrid model effectively.

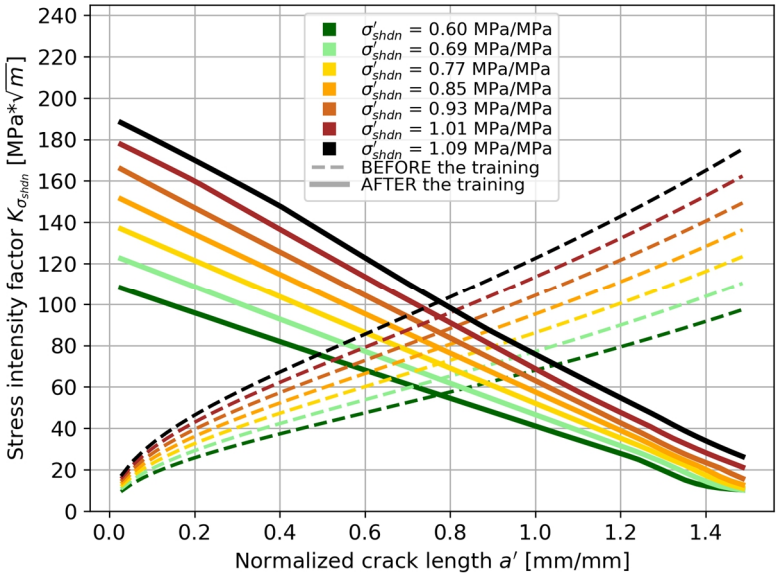


Fig. 40 Stress intensity factor values at the constant thermal stresses $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn} = const, a'\right)$ determined by the data-driven layer before and after the training from the ground up in the target domain using ten observations.

5.6 Conclusions concerning this stage of the research

The purpose of the second stage of this research was to apply physics-informed neural networks on a small sample comprising 13 observations, and assess whether the objectives of this research, described in Section 2.1, can be achieved using this class of machine learning algorithms. The specific aims were to create models based on machine learning algorithms, which are more consistent with theoretical or experimentally determined equations describing the analyzed phenomenon, and have better generalization and extrapolation capabilities than the data-driven models created during the first stage of the research and described in Chapter 4. The primary outcomes of this part of the research are as follows:

- An approach was proposed and applied to estimate thermal stresses at shutdown based on readings of the selected operating parameters recorded at steady-state conditions. These tensile stresses act in the direction perpendicular to the crack propagation plane and open the analyzed fatigue cracks. Actual readings of the ambient air temperature T_{AMB}, the average air temperature at the axial compressor's discharge T_{COOL} and the average temperature of gases at the outlet of the nozzles T_{FIRE}, which are recorded by the data acquisition system, are utilized to estimate the thermal stress values. The finite element analysis results were used to determine the values of several coefficients of the equations applied to calculate the stresses. Additionally, selected characteristics of the FSX-414 cobalt-based superalloy are used, which were created based on the results of material properties tests conducted at Baker Hughes Company. The thermal stresses are utilized as the input data to estimate the stress intensity factor values.

- A physics-informed neural network was designed and successfully applied to model the propagation of the fatigue cracks observed on the analyzed high-pressure nozzles. The PINN is based on a recurrent neural network, which is a rational choice for modeling progressive, irreversible damage accumulation. The custom cell of the RNN comprises a data-driven layer and a physics-based cyclic part. The former is a feedforward neural network that estimates the stress intensity factor based on the thermal stress values. The latter calculates the crack length increment during each startup-shutdown cycle based on the Paris' law. An attempt to embed in the RNN cell a time-dependent part to capture the contribution of creep on the crack propagation rate was unsuccessful, resulting in a significant reduction of the prediction's accuracy. The constraints resulting from the theoretical Equation (57), which the fully-specified PINN respects, were implemented in a soft manner through the additional term in the customized cost function, i.e. the physical term $RMSE_{PHY}$.

- A custom cost function was designed and applied, with a variable coefficient $\gamma \in [0,1]$ determining how the physical and empirical, i.e. $RMSE_{EMP}$, terms contribute to the final cost value during the training process in accordance with Eq. (56). If γ is equal to zero the cost function reduces to the empirical term only. In contrast, if it equals one, the cost function simplifies to the physical term only, and when it has intermediate values, both terms are included accordingly. The coefficient value changes dynamically during the training process following the predefined function (described by Eq. (66)) to find a proper balance between the two terms. As the consistency of the data-driven layer estimates with the theoretical equation embedded in the cost function increases, the value of the coefficient decreases. Thus, the first stage of the learning process aims to

restrict the space of allowable solutions only to a subspace where the theoretical equation is respected, and the second stage aims to reduce the empirical term's value as much as possible while staying in that subspace. The set of input data utilized as the basis to calculate the value of the physical term was created synthetically. It was generated in a systematic way aimed at improving the effectiveness of the training process and the extrapolation capabilities of the multilayer perceptron. The original input space, which is based on the empirical data available in the source domain, was enlarged with synthetic data, which simultaneously enlarged the training space where the physics-informed neural network should respect the theoretical constraints. The approach applied to generate these additional training data can be adjusted depending on particular objectives.

- Three hybrid models sharing the same architecture but trained respectively based on ten, two and only one observation were prepared, with the maximal crack size at Position 2 on the nozzle's trailing edge as the dependent variable. Considering the limitations regarding the availability of the operational data and damage size measurements and the simplified theoretical description of the fatigue crack propagation process applied to prepare the hybrid models, the accuracy of the final estimates is satisfactory. The models trained using ten and two data points are the most accurate, with the normalized root mean squared error evaluated against the test sets equal 9% and 11%, respectively. These values are comparable with the accuracy of recording those cracks during a visual inspection preceding repair activities, where the maximum normalized error may reach 7%. The most accurate estimates were obtained for nozzles with more than 40 interval fired starts accumulated and with the longest cracks. The models trained using two and one observation can provide accurate crack length estimates for parts with 40% longer cracks and over four times more startup-shutdown cycles in comparison with the utilized training data points. The prediction accuracy of the hybrid models is significantly better than that of the model pre-trained on the results obtained using the theoretical formula (65). When the normalized RMSE evaluated against the test set is considered the reference metric, those two hybrid models are also better than the data-driven models created during the first stage of the research (the details are provided in Table 2). However, even though the sample used to train the PINNs is significantly smaller, it is also more homogeneous, comprising only parts that operated in Type C gas turbines and have not been subjected to repair. Moreover, an attempt to apply the reference hybrid model, i.e. the one trained using ten observations, to the nozzles subjected to repair activities in the past resulted in a significant underestimation of the final crack lengths. Finally, the models are capable of differentiating the crack growth rates and the final estimates based on actual operational data, despite the very stable operating profile of the analyzed units. The obtained results are consistent with each other, regardless of the cardinality and the composition of the training set. Thus, cross-validation is not necessary.

- Effective execution of the optimization process based on a single training observation was accomplished in one case only and failed for five other data points. Extracting and leveraging knowledge from a single damage size measurement is valuable and may give tangible benefits. However, with such a small sample, difficulties in converging to the optimal solution may be encountered, and the final predictions may vary significantly depending on the characteristics of the selected training observation.

- The core of the hybrid model is the feedforward neural network responsible for the stress intensity factor estimates. It determines, to a great extent, the final predictions obtained by means of the fully-specified hybrid model. This neural network cannot be considered a black box in its entirety since the characteristics of the data-driven layer are available, i.e. $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn}, a' = const\right)$ and $K_{\sigma_{shdn}} = FNN\left(\sigma'_{shdn} = const, a'\right)$, presenting directly the relationship between the inputs and the output. For example, the characteristics of the reference model are shown in Fig. 21 and Fig. 22. They present estimates of the multilayer perceptron outside the original space of input data when $a' > 1$ or $\sigma'_{shdn} > 1$. These characteristics are continuous, without peaks or drops inside and outside the original input space. They illustrate that the feedforward neural network obeys the theoretical equations embedded in the cost function and how it interpolates and extrapolates. Thus, leveraging these characteristics, the predictions can be made in a fully conscious way. Otherwise, they can be utilized to limit the input space, where the fully-specified model can be safely applied. Nevertheless, if necessary, the training space can be enlarged with synthetically created data.

- A novel method of single-source domain generalization and cross-domain knowledge transfer in regression analysis leveraging physics-informed neural networks has been proposed. The method's purpose is to effectively apply a cumulative damage model based on a physics-informed neural network, i.e. the so-called source hybrid model, which was trained in a domain, where measurement data on damage size are available, in another domain, where such data are unavailable. Following the definitions given in Section 5.4, the former domain is called the source domain, and the latter is the target domain. The method can be applied if the damage type in the target domain, the physical phenomena causing it and the applicable theoretical or experimentally determined equations for estimating the damage increment in a single time step, are the same as in the source domain. The eight-step procedure describing the main steps of the proposed method is provided at the end of Section 5.4. The so-called shifting of the operational data and PINN's input variables in the target domain is the essential step that enables an effective application of the unchanged source hybrid model in the target domain. The composition of functions used to calculate the input data of the physics-informed neural network based on the selected operating parameters shared across the domains must be known in both domains to accomplish that step. Eventually, the final estimate in the target domain for the original set of operational data $\mathbf{X}_{TGT}$ is obtained by applying the source hybrid model to the set of shifted operational data $\mathbf{X}^{*}_{TGT}$.

- The created method of domain generalization was applied to predict the maximal crack length at Position 4, leveraging the reference hybrid model dedicated to Position 2. The target location is not on the same airfoil as the source location. The value of maximum thermal stresses at shutdown corresponding to Position 4 is 86.32% of the stress value calculated for Position 2 for any vector of the chosen operating parameters. The damage size measurements available in the target domain were used to evaluate the prediction's accuracy. Applying the proposed method, the normalized error evaluated against all 13 data points equals $NRMSE_{EMP,TEST}$ = 15.5%, and is similar to the error resulting from the application of a model trained from the ground up in the target domain that is equal to $NRMSE_{EMP}$ = 15.9%. Thus, this method was successfully applied, providing accurate damage size predictions in a simulated scenario, where damage size measurements are

unavailable. The latent information and knowledge extracted from the empirical data available in the source domain were leveraged and effectively used in the target domain.

During this part of the research, it was proved that the models based on physics-informed neural networks respect prior knowledge and physical laws that govern the analyzed phenomenon and can be trained effectively when only two or even one training observation is available. These features are inherited by the created method of single-source domain generalization and cross-domain knowledge transfer in regression analysis since it leverages physics-informed neural networks. It was presented that accurate damage size predictions can be obtained using the proposed method, even if no damage size measurements are available in the target domain. As already described, such limitations may occur at the early stages of components' lifetime. The next part of this research will focus on evaluating the flexibility and versatility of the method when applied to a different object and failure type. The main objective will be to prove that this method is universal and can be applied for modeling damage due to various failure types found on gas turbine components.

6. Predicting metal loss due to oxidation by physics-informed neural networks

One of the specific objectives of this dissertation is to create a method based on machine learning algorithms for predicting the size of damage to gas turbine components, which will be universal, with the capability of applying it effectively to the most common failure types found on the main combustion and hot gas path parts. Thus, after the successful application of physics-informed neural networks and the proposed method of single-source domain generalization for predicting the maximal length of fatigue cracks, another object and failure type is selected to evaluate the effectiveness of this modeling approach. Considering the minimal requirements for data availability, which are listed at the beginning of Chapter 4, it is decided to select a transition piece of a heavy-duty gas turbine as the analyzed object, focusing on predicting the component wall thickness reduction due to oxidation. The additional arguments in favor of this selection are as follows:

- Problems related to oxidation and overheating are common for all types of gas turbines (according to [28], almost 7% of damage observed in aircraft engines was caused by material overheating).
- This type of failure is among the most common causes of scrapping the transition pieces. Thus, the capability to predict the damage size may be leveraged to foresee the number of scrapped components, optimize parts management and predict operating expenses.
- The analyzed part is made of Nimonic 263 nickel-based alloy, which is widely used to manufacture transition pieces and combustion liners of other types of gas turbines. Therefore, many opportunities exist to use the source hybrid model in another domain.
- The presence of thermal barrier coating on the inner side of the part, possible effects of erosion on the wall thickness reduction and the limited number of empirical data make this a challenging problem in the context of cumulative damage model creation.
- The input data needed to create the model are available, and no additional efforts and costs are required to gather them.

Details of this part of the research are provided in the following sections of this chapter.

6.1 The analyzed object, problem setup and an overview of available empirical data

The analyzed component comes from a family of industrial gas turbines offered by Baker Hughes Company, which are available in different configurations to support mechanical drive and power generation applications. The installed fleet is composed of almost 200 units worldwide (status of 06-Jun-2023). The available sample comprises three configurations, which are referenced as Type I, Type II and Type III. The latter has a lower firing temperature and compressor pressure ratio than the first two, i.e. consequently, the air temperature at the axial compressor's discharge is also lower. In gas turbines equipped with can-type combustors, transition pieces serve as a kind of collector, which transforms the cylindrical flow of exhaust gases from a combustion liner into a continuous annular flow at the inlet of the high-pressure nozzles. The number of transition pieces is equal to the number of combustion cans. However, the analyzed gas turbines are equipped with a single combustion chamber. The analyzed object

is presented in Fig. 41, Fig. 42 and Fig. 43 [211]. The transition piece is divided at the centerline for assembling purposes. Over the years, a few versions of the part were designed and deployed, but only the latest configuration is being considered in this analysis. Thus, from this perspective, the analyzed sample is homogeneous. This part is manufactured from Nimonic 263 nickel-based superalloy. From the outside it is cooled by the air discharged from the axial compressor that flows in an inverse direction to that of the hot gases inside the transition piece. This component has numerous holes to inject high-pressure air into the hot gas path, creating a thin

Fig. 41 The outlet of the analyzed transition piece.

Fig. 42 A view from the compressor side on the analyzed transition piece.

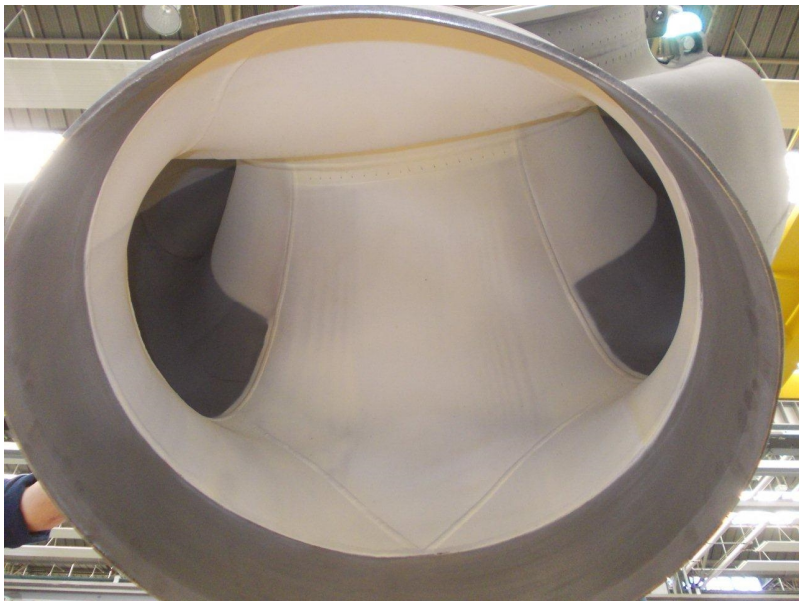


Fig. 43 The inlet mouth of the analyzed transition piece.

coolant layer along the internal surfaces. This technique is known as film cooling. Additionally, as shown in Fig. 43, the internal surfaces are partially covered with a thermal barrier coating to provide thermal insulation from the hot gases and improve oxidation resistance. In accordance with the applicable maintenance policy, the transition pieces should be disassembled and repaired during each combustion inspection. This part should be repaired just once, and it should be scrapped after the second interval. However, it should be reminded that less stringent maintenance rules may be applied for units covered by the long-term service agreement offered by the original equipment manufacturer.

Based on empirical data recorded during visual and dimensional checks executed during disassembly inspections and subsequent repair activities, the following failure types can be found on the analyzed transition pieces:

- deformation and damage occurred during disassembly caused by the deformation,
- oxidation or erosion,
- cracks,
- fretting of the outlet flange,
- surface degradation and spallation of the thermal barrier coating.

As for the previously analyzed high-pressure nozzles, this damage is not monitored nor directly measured during the operation because of the high temperature of exhaust gases. Additionally, the impact on gas turbine performances of surface degradation due to oxidation or erosion is negligible and not reflected in operational data gathered by the data acquisition system. Thus, a more interdisciplinary approach should be applied to predict the component wall thickness reduction due to oxidation. Examples of surface degradation due to oxidation and spallation of the thermal barrier coating are shown in Fig. 44 and Fig. 45 [212]. During a dimensional check preceding repair activities, the thickness of the transition piece wall is measured in three planes, denoted as 1, 2 and 3 in Fig. 46 [213], with eight measurements taken per plane, denoted with letters A to H in the same figure. Nevertheless, the completed inspection form provides only the minimal thickness measurement for each plane. Thus, for a specific transition piece, three

Fig. 44 Green nickel oxide found on the external surface of the analyzed object after service.

Fig. 45 Thermal barrier coating spallation found on the internal surface of the analyzed object after service.

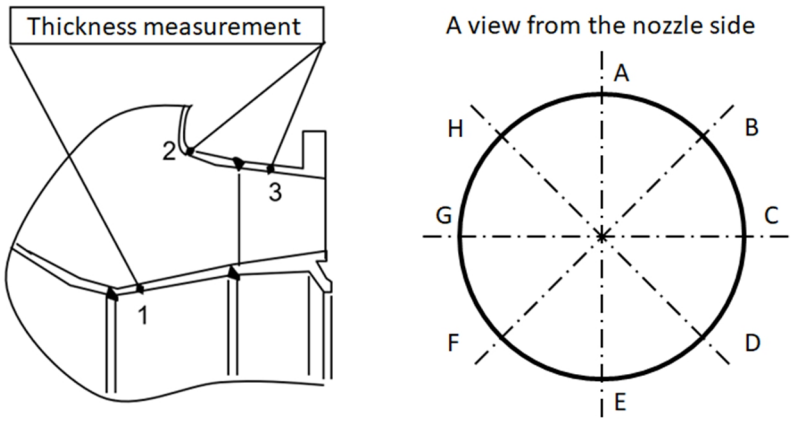


Fig. 46 Areas on the transition piece outlet where dimensional inspection is executed.

measurements of the minimal wall thickness are available, separately for planes 1, 2 and 3. Consequently, if a set of measurements related to a particular plane is analyzed, comprising data on various components, it should be remembered that it refers to the same plane but not necessarily to the same position. This additional source of uncertainty regarding the available measurements further complicates the task of modeling effectively the wall thickness reduction due to oxidation for the analyzed transition pieces.

Before an overview of available empirical data, it should be reminded that data used for this research are confidential and are the proprietary property of Baker Hughes Company. Due to these reasons, the data cannot be shared openly and were anonymized in this dissertation.

Input data used to train and test the predictive models comes from previously utilized sources for the analyzed high-pressure nozzles. The first set of data is extracted from the Parts Life Database and contains the following information for each transition piece:

- The operational history of the analyzed components, including the scope of historical repair activities and the numbers of fired hours, fired starts and emergency shutdowns accumulated since the last repair and since the part was manufactured.
- Wall thickness measurements of components after service, which were subjected to a visual inspection preceding repair activities.
- The component codes and drawing numbers required to select records related only to the latest configuration of the transition piece.

The analysis is limited only to parts that were not repaired in the past to avoid introducing noise resulting from the inclusion in the dataset of repaired components, which could be subjected to the stripping of thermal barrier coating and recoating after completion of repair activities.

The second set of inputs contains actual operational data recorded by the Remote Monitoring & Diagnostics system. Concerning the analyzed phenomenon, the oxidation rate depends on the material temperature and increases as the temperature rises. Therefore, the operational data are used to estimate the material temperature. The following operating parameters are utilized:

- the average air temperature at the axial compressor's discharge T_{COOL},
- the average temperature of gases at the inlet of the nozzles T_{INLET}, which is calculated based on other signals recorded by means of the data acquisition system.

Parts with less than 50% availability of T_{INLET} values are excluded from further processing. The sampling interval is set to one hour. Therefore, the damage increment will be estimated every hour. Raw data downloaded from the system are subjected to data cleansing that has the following objectives:

- Remove the non-numerical data from the dataset, including strings, NaN (i.e. Not a Number) and positive or negative INF (i.e. infinity).
- Remove from the dataset records containing erroneous values outside the acceptable ranges and records related to the transient states, i.e. startup, acceleration, load step, load rejection, deceleration or shutdown. A range of acceptable values was defined for each analyzed operational parameter. Records with T_{COOL} values outside the acceptable range are excluded from the dataset, while improper values of T_{INLET} are completed in the next stage of data processing.

Missing values of the selected operating parameters are calculated as an arithmetic average of the values describing the adjacent cycles if the operational data are available for the preceding and subsequent cycles. Otherwise, they are calculated as medians of measurements recorded during the same day, or week (if the measurements are unavailable for that day), month, quarter, or finally, the entire service period. Generally, the interval used to calculate the median expands if no readings are available in the shorter interval. Such a gradual approach is necessary since the analyzed gas turbines operate in industrial applications, typically at partial load, which may change dynamically.

The third set of inputs comprises the following configuration details of the analyzed engines:

- the model, i.e. Type I, Type II or Type III;
- the type of combustion system (either diffusion or premixed combustor);
- the geographical position of each unit.

Parameters that do not differentiate the investigated gas turbines or should not have any impact on the analyzed phenomenon were identified but not listed above.

Considering all the limitations resulting from the availability of empirical data and the assumptions made, the final sample comprises only 11 observations. Those transition pieces have accumulated in total 148 442 service hours, but the required operating parameters are available for 127 532 records, which is 86% of the total number. Fig. 47 shows the values of the coverage, defined as the number of fired hours for which the operational data are available. Its values range from 60% to 99%. The data imputation approach was described in the previous paragraph. Fig. 48 presents the variation of the wall thickness measurements related to the analyzed 11 components. The median values are very similar for each position. However, based on the box plot, it is visible that the lowest values were recorded at Position 1, while the highest correspond to Position 3. It is an unexpected finding since, in accordance with the numerical simulation results presented in the next section, material temperatures in the areas considered as Position 1 should be comparable with those at Position 3. Additionally, these two areas

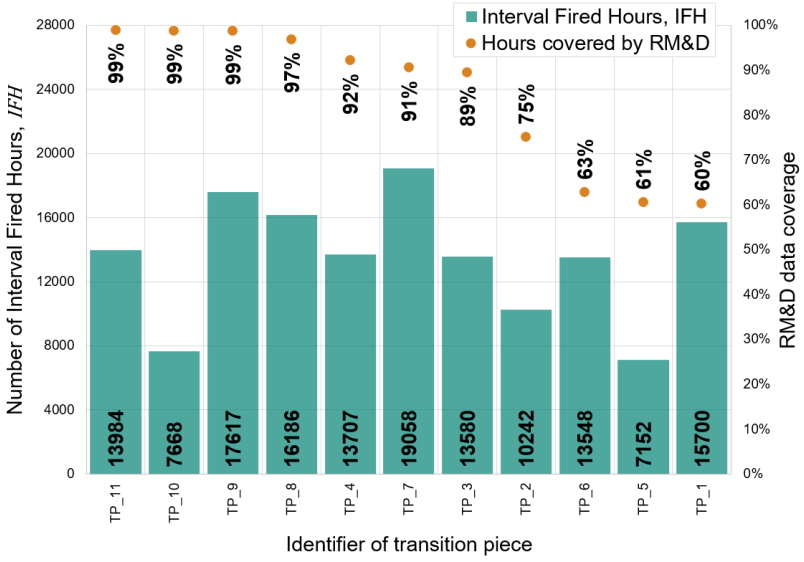


Fig. 47 The number of fired hours accumulated since the last repair *IFH* for each transition piece, and the percentage of service hours for which values of the required operating parameters are available.

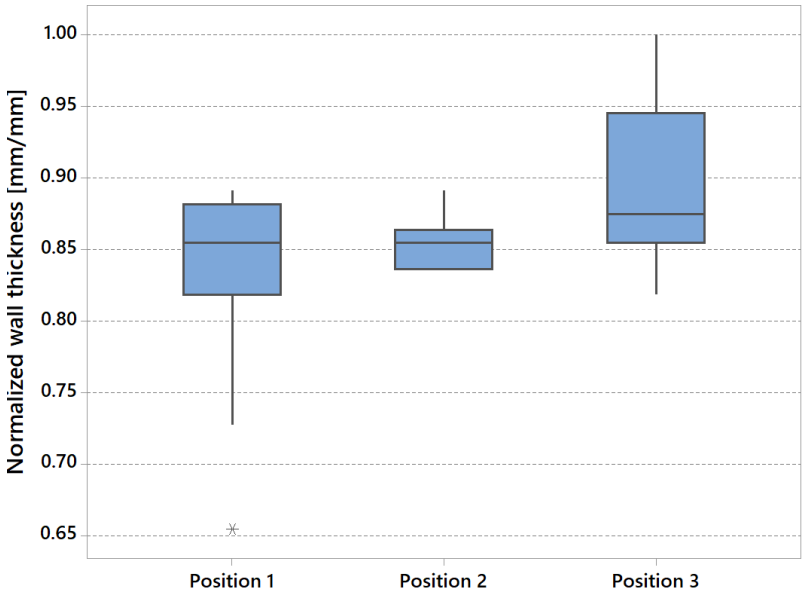


Fig. 48 Variation of the wall thickness measurements. The data normalized by dividing each value by the maximal measured thickness to protect proprietary information.

should be significantly colder compared to the area considered as Position 2. Thus, based on the results of analysis performed by means of the finite element method, the lowest wall thickness measurements are expected at Position 2, while at the remaining two positions the material should be thicker with the measurements comparable to each other. It can be presumed that higher metal loss values recorded at Position 1 may be caused by erosion. Therefore, considering that the measurement data related to Position 1 are not consistent with the numerical simulation results and that associated physical phenomena may not be completely known, it is decided to focus on the remaining two positions. Consequently, the source domain refers to Position 2, and the target domain refers to Position 3. The reduction in base material thickness due to oxidation is the dependent variable, which will be predicted using a hybrid model based on a physics-informed neural network. Finally, it should be clarified that the initial thickness measurements done after the manufacturing process are not available. In accordance with the applicable drawing, the nominal metal thickness equals $w_{nom} \pm \frac{5}{100} w_{nom}$, while for the thermal barrier coating it is equal to $\frac{8}{100} w_{nom} \pm \frac{14}{1000} w_{nom}$. Thereby, focusing only on base material thickness, the difference between the upper and lower limits equals 10%. More than 80% of the analyzed measurements, presented in Fig. 48, can be contained in an interval with such width. The absence of the initial thickness measurements introduces additional uncertainties and significantly increases the complexity level of the analyzed problem. In order to train the model and make predictions in the source and target domains, it is assumed that the initial material thickness is equal to $w_{init} = w_{nom} + \frac{5}{100} w_{nom}$. Attempts to use different values of w_{init} did not improve the prediction's accuracy.

6.2 The architecture of the hybrid model

The physics-informed neural network is based on a recurrent neural network, similarly to the previously analyzed problem related to the high-pressure nozzles. The equations describing the damage increment are embedded into the recurrent neural network's cell, as presented in Fig. 49. The cell comprises the following major elements:

- The physics-based part leverages the parabolic law that was first derived by Wagner [214] and that has the following form:

$$p = C1\left(t\right)^{0.5}, \tag{87}$$

where p is the oxide scale thickness, $C1$ is the positively defined parabolic rate constant and t denotes the operation time. Based on this law, the reaction rate $dp/dt = C1/p$ is inversely proportional to the oxide scale thickness and it decreases with time. A more generalized form of Eq. (87) is based on the following power function:

$$p = C1\left(t\right)^{C2}, \tag{88}$$

where the positively defined exponent $C2$ depends on the metal temperature T, i.e. $C2 = f\left(T\right)$. Derivating Eq. (88) with respect to time results in:

$$\frac{\partial p}{\partial t} = C2 \cdot C1\left(t\right)^{C2-1}, \tag{89}$$

which substituting $t = \left(p/C1\right)^{1/C2}$ according to Eq. (88) can be applied to approximate the wall thickness reduction due to oxidation in each time interval $\Delta t = 1$ hour as follows:

$$\begin{aligned} \Delta p &= \left(C2 \cdot C1 \left({}^{p}\!/_{C1} \right)^{\frac{C2-1}{C2}} \right) \Delta t = \left(C2 \cdot C1 \left({}^{p}\!/_{C1} \right)^{1-\frac{1}{C2}} \right) \Delta t = \\ &= \left(C2 \cdot C1 \frac{{}^{p}\!/_{C1}}{\left({}^{p}\!/_{C1} \right)^{\frac{1}{C2}}} \right) \Delta t = \left(\frac{C2 \cdot p}{\left({}^{p}\!/_{C1} \right)^{\frac{1}{C2}}} \right) \Delta t . \end{aligned} \tag{90}$$

The physics-based part is used to calculate the damage increment on the external, colder side of the transition piece Δp_{COLD} using Eq. (90) based on the estimated oxide scale thickness at the previous time step t - 1 denoted as $p_{COLD,t-1}$ and values of $C1$ and $C2_{COLD,t}$, which are approximated by means of feedforward neural networks.

- The data-driven layer is based on a multilayer perceptron estimating values of the exponent $C2_{COLD}$ considering T_{COLD} as the input, which is the wall's temperature at the analyzed position on the external, colder side of the transition piece. The relationship between T_{COLD} and $C2_{COLD}$ should be a monotonically increasing function, i.e. the exponent value increases as the temperature rises. Thus, a monotonic feedforward neural network is applied to satisfy this constraint, inspired by the architecture presented

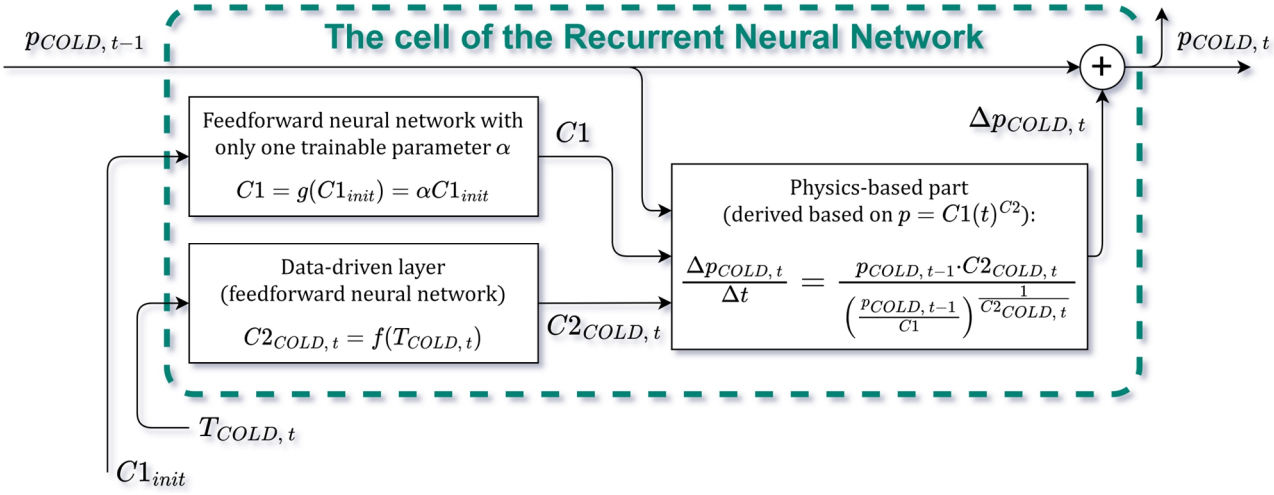


Fig. 49 The custom cell of the recurrent neural network applied to estimate the material loss due to oxidation during each time step.

in [215]. It implements a continuous piecewise linear function by taking maximum and minimum operations on groups of two-dimensional hyperplanes. The structure of the data-driven layer is shown in Fig. 50. The first hidden layer is composed of 35 neurons with linear activation functions and a constraint applied to the weights, which limits their values only to nonnegative numbers. The neurons are divided into seven groups, each containing five neurons. The second hidden layer calculates the maximum from the results obtained in the first layer, separately for each group. Therefore, the second layer comprises seven units, each transforming five values into one. The third hidden layer computes the minimum from the outputs of the second layer, transforming seven values into one. Finally, the output layer has no trainable parameters and applies the rectifier linear unit function to ensure that the $C2_{COLD}$ estimates are nonnegative. Attempts to simplify the structure by reducing the number of neurons in the first layer and dividing them into three groups, as in the reference paper [215], resulted in lower prediction accuracy of the hybrid model. The min-max scaling is applied to normalize the input of the data-driven layer.

- The second perceptron comprises a single neuron and is applied to estimate the constant $C1$. The initial value of the constant $C1_{init}$ is determined using the results of a burner-rig test for oxidation resistance of Nimonic 263 nickel-based superalloy, which was conducted at Baker Hughes Company. The final value of the constant is computed as $C1 = \alpha C1_{init}$, where the approximated value of α minimizes the cost function. No bias is used in this case, while the ReLU activation function is used to ensure that the factor will have nonnegative values.

By applying this architecture of the physics-informed neural network, it is assumed that the oxides form only on the external, colder side of the transition piece and that the oxide scale thickness on the internal, hotter side is equal to zero. The thermal barrier coating is assumed to effectively protect against metal loss due to oxidation throughout the entire service period until the component is disassembled and subject to inspection and repair activities. This assumption substantially simplifies the model's architecture. Nevertheless, it may be too strong and false, especially since the thermal barrier coating spallation was observed on the transition pieces

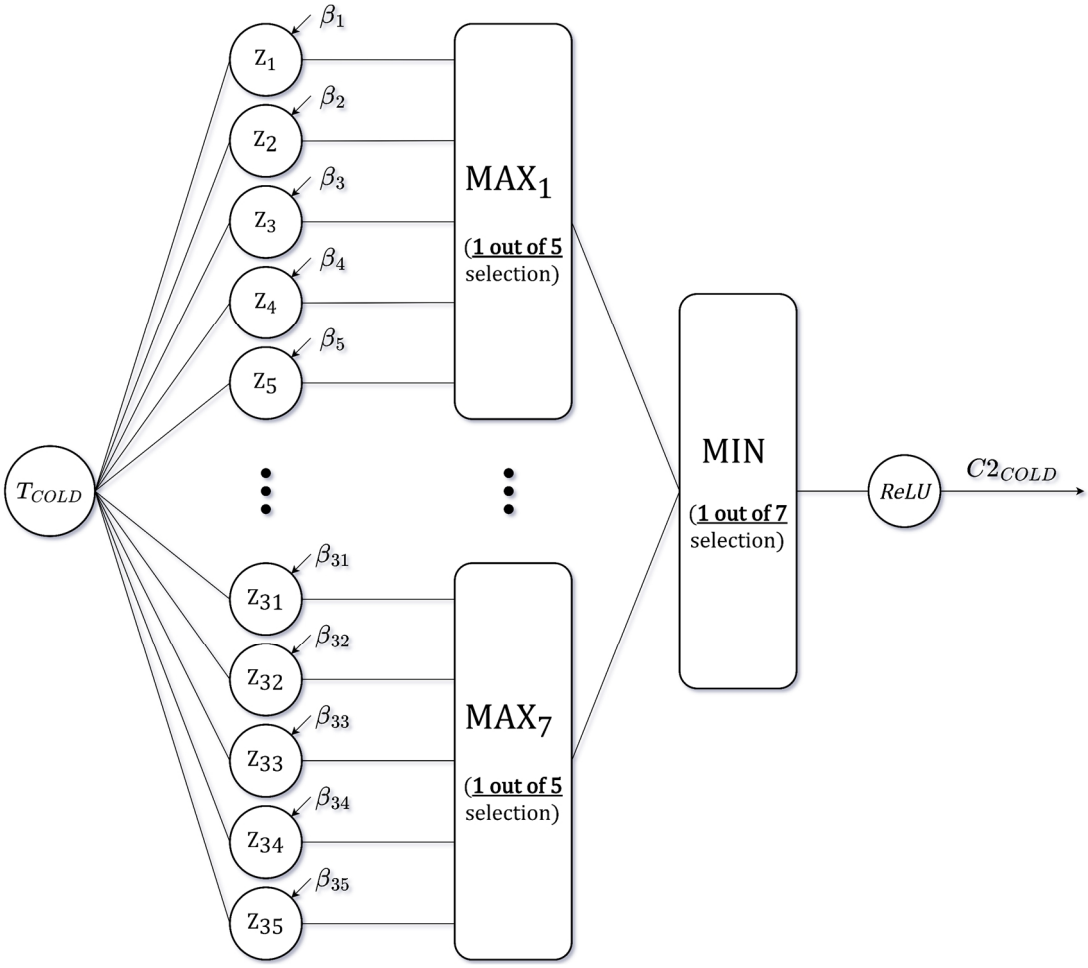


Fig. 50 The structure of the monotonic feedforward neural network, being the data-driven layer, which estimates the exponent $C2_{COLD}$ considering T_{COLD} as the input, which is the wall's temperature on the colder side of the transition piece.

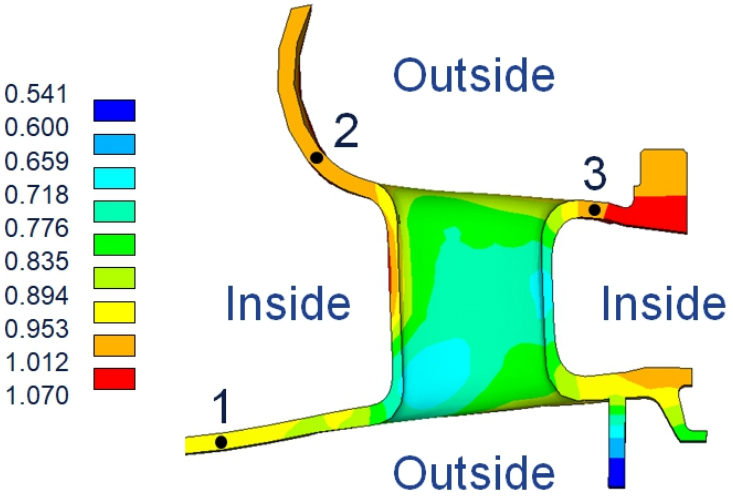
included in the analyzed sample. The company possesses a model that is applied to estimate the number of startup-shutdown cycles until the coating spalls off completely, exposing the base material. Several attempts were made to consider the results of that model as an additional input of the physics-informed neural network, but all of them were unsuccessful. The applied approach assumed that the oxide scale on the internal side of the part does not grow until the coating detaches and the base material is exposed. The wall temperature values on the hotter side of the transition piece T_{HOT} were determined based on the operational data. The corresponding values of the exponent $C2_{HOT}$ were estimated using the same data-driven layer, or an additional multilayer perceptron, which estimated the values dedicated only to the internal side of the part. However, the increased complexity of the hybrid model's architecture did not increase the prediction's accuracy and caused some optimization process disruptions. Despite these problems, future research should be related to the inclusion of thermal barrier coating presence in the hybrid model and appropriate modeling of the oxide scale growth on the internal side of the part.

For each time step, the material temperature values on the internal and external sides of the transition piece are calculated based on the actual operating data as follows:

$$T_{HOT} = T_{INLET} - \phi_{HOT}\left(T_{INLET} - T_{COOL}\right), \tag{91}$$

$$T_{COLD} = T_{INLET} - \phi_{COLD}\left(T_{INLET} - T_{COOL}\right), \tag{92}$$

where ϕ_{HOT} and ϕ_{COLD} are the cooling effectiveness coefficients on the internal and external sides, respectively. Their values were determined using the numerical simulation results related to steady-state operating conditions, shown in Fig. 51. It was not prepared specifically for this research but for a different New Product Development program finalized a few years earlier. However, it should be clarified that the utilized numerical simulation is dedicated to Type I gas turbines equipped with a premixed combustor. Two observations in the available sample are related to Type III units equipped with a diffusion combustor and characterized by lower nominal values of T_{INLET} and T_{COOL}. A dedicated numerical simulation is not available for this configuration. Therefore, to avoid further reduction of the sample, the values of ϕ_{HOT} and ϕ_{COLD} coefficients applicable to Type I and Type II units are also assumed valid for Type III gas turbines.



	Normalized temperature of the wall	
	Inside	Outside
Position 1	0.939	0.914
Position 2	1.019	1.000
Position 3	0.939	0.927

Fig. 51 The material temperature field at steady-state operating conditions obtained by means of the finite element analysis with the values normalized by dividing them by the temperature related to the colder side of Position 2.

6.3 Configuration and execution of the training process

The general form of the custom cost function, which is minimized during the training process, is the same as for the previously analyzed high-pressure nozzles and defined by Eq. (56). Considering that $C2_{COLD} = f\left(T_{COLD}\right)$, the physical term of the cost function $RMSE_{PHY}$ is based on the following inequality:

$$f\left(T_{COLD,min}\right) - f\left(T_{COLD,max}\right) < 0,$$

$$C2_{COLD,min} - C2_{COLD,max} < 0, \tag{93}$$

which is valid when $T_{COLD,min} < T_{COLD,max}$. Based on Inequality (93) and considering that in accordance with the created architecture, the values of $C2_{COLD}$ are estimated using the feedforward neural network, the physical term is defined as follows:

$$RMSE_{PHY} = \sqrt{\frac{1}{J}\sum_{j=1}^{j=J}\left[ReLU\left(FNN\left(T_{COLD,min,j}\right) - FNN\left(T_{COLD,max,j}\right)\right) - 0\right]^2} = \\ = \sqrt{\frac{1}{J}\sum_{j=1}^{j=J}\left[ReLU\left(\widehat{C2}_{COLD,min,j} - \widehat{C2}_{COLD,max,j}\right) - 0\right]^2}, \tag{94}$$

where $j = 1,2,\ldots,J$ is the ordinal number of couples created synthetically to train the multilayer perceptron, J = 6 830 is the number of couples, $\widehat{C2}_{COLD,min}$ and $\widehat{C2}_{COLD,max}$ are the estimates determined by the feedforward neural network (i.e. respectively for $T_{COLD,min}$ and $T_{COLD,max}$), which is symbolically denoted as function *FNN*. The ordered pairs are created basing on the available empirical data with each couple comprising $T_{COLD,min}$ and $T_{COLD,max}$. The 683 unique values of $T_{COLD,min}$ are evenly spaced numbers over the interval $\left[\min\left(T_{COLD}\right), \max\left(T_{COLD}\right)\right]$. For each unique value of $T_{COLD,min}$, ten values of $T_{COLD,max}$ are randomly drawn from the uniform distribution $U\left(T_{COLD,min}, \frac{5}{4}\max\left(T_{COLD}\right)\right)$. The original input space is enlarged with synthetic data, simultaneously enlarging the training space, where the physics-informed neural network should respect the theoretical constraints. In this way, all 6 830 ordered pairs are created, which are used to calculate the value of the physical term. Thus, the constraints resulting from Inequality (93) are implemented in a hard manner, applying the monotonic multilayer perceptron, but also in a soft manner, through the additional term in the cost function. Such a redundant approach is used since the constraints resulting from Inequality (93) are less stringent than, for example, the restrictions resulting from Equation (57) considered during the second stage of the research related to the high-pressure nozzles.

The empirical term of the custom cost function is calculated in the following way:

$$RMSE_{EMP} = \sqrt{\frac{1}{N}\sum_{i=1}^{i=N}\left(p_i - \hat{p}_i\right)^2}, \tag{95}$$

where $i = 1,2,\ldots,N$ denotes the observation's ordinal number, N is the total number of observations in the training set, p_i denotes the observed reduction of the wall thickness and $\hat{p}_i$ is the final reduction of the base material thickness due to oxidation estimated by means of the hybrid model.

The multilayer perceptron composing the data-driven is pre-trained using the results obtained by means of the following formula:

$$p_{EXP} = C1_{init}\left(t\right)^{C2EXP} = C1_{init}\left(t\right)^{\left(AT_{COLD}+B\right)}, \tag{96}$$

where the values of constants $C1_{init}$, A and B are based on the results of the burner-rig test for oxidation resistance of Nimonic 263 nickel-based superalloy, which was conducted at Baker Hughes Company. One thousand records were generated for the pretraining. The values of

T_{COLD} are evenly spaced numbers over the interval $\left[\min\left(T_{COLD}\right), \max\left(T_{HOT}\right)\right]$. The training process duration is shorter than 20 seconds.

During the main part of the training process, the value of γ, which determines how the physical and empirical terms contribute to the final cost value, changes dynamically using the following formula:

$$\gamma = \begin{cases} 0.2, & RMSE_{PHY} \leq 0.0086 \\ \left(10 + {}^{219}\!/_{563}\right) RMSE_{PHY} + {}^{1}\!/_{9}, & RMSE_{PHY} \in \left(0.0086, 0.0856\right] \\ 1, & RMSE_{PHY} > 0.0856 \end{cases}. \tag{97}$$

Fig. 52 presents the piecewise continuous function $\gamma = f\left(RMSE_{PHY}\right)$. The pretraining ensures that the main part of the learning process is initiated with $\gamma = 0.2$. The value of γ coefficient is recalculated after each epoch, i.e. after one complete pass of the algorithm through the entire dataset, to react quickly and increase the coefficient value if the value of $RMSE_{PHY}$ rises. During the execution of this analysis, it was observed that the implemented constraints work effectively, maintaining the low value of the physical term while reducing the empirical term's value as the training process progresses. The Adam optimization algorithm is applied to minimize the cost function. Initially, the learning rate's value was set to 0.001, but it was increased to η = 0.01 in order to reduce the training process duration without negatively impacting the prediction's accuracy. Finally, the optimization process should not take more than a few hours using an Intel Core i7-9850H Central Processing Unit with six cores and a 2.60 GHz base frequency. As for the previously analyzed high-pressure nozzles, open-source Python packages are used to execute the analysis, i.e. Keras application programming interface for the TensorFlow library, Numpy, Pandas and Matplolib.

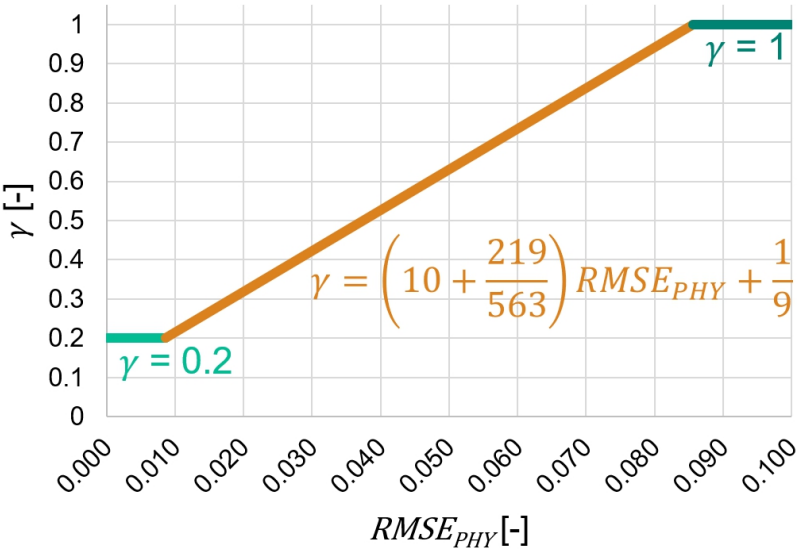


Fig. 52 The piecewise continuous function used to calculate the values of coefficient γ, which determines how the physical and empirical terms contribute to the final cost value.

6.4 Results of the regression analysis

Several hybrid models were created during the execution of analysis considering different training and test subsets compositions. However, the most accurate predictions are obtained using the model trained basing on seven observations characterized by the highest availability of the operational data as shown in Fig. 47. The test set comprises four data points with the following identifiers: TP_1, TP_2, TP_5 and TP_6. Regardless of the main criterion applied to create the two sets, the composition of the test subset is heterogeneous and can it be described as follows:

- 2 observations out of 4 (50%) are related to Type I gas turbines, while the remaining test data points correspond to Type II and Type III units. Considering the entire sample, 7 data points out of 11 (64%) are related to Type I gas turbines and 2 observations each are available for the remaining units.
- 3 observations out of 4 (75%) are related to the gas turbines equipped with premixed combustors. Considering the entire sample, 8 data points (73%) correspond to such units.
- 2 observations each on gas turbines driving compressors and generators. Considering the entire sample, 7 data points (64%) are related to power generation applications.
- 3 observations out of 4 (75%) are related to units installed onshore. Considering the entire sample, 6 data points (55%) correspond to such location.
- For the available observations, the interval fired hours values are constrained within the interval [7152, 19058], with a median of 13707 (the lower quartile is equal to Q_1 = 11895 and the upper Q_3 = 15943). The selected test data points are between the minimum and 70th percentile values. Moreover, as for the high-pressure nozzles, operational data availability in future applications is expected to be better than in this research, based on the oldest available data in many cases.

Despite the low cardinality of the test set, it has a very diverse composition, which should help to evaluate the model's performance on unseen data and identify its advantages and disadvantages. However, the characteristics of the multilayer perceptron, which is responsible for the estimates of the exponent $C2_{COLD}$, must be included during such an evaluation. Finally, this hybrid model trained basing on seven observations, used to predict the reduction in base material thickness due to oxidation, is called the baseline model. The estimates obtained using the baseline model and the characteristic of the data-driven layer $\widehat{C2}_{COLD} = FNN\left(T_{COLD}\right)$ are presented in Fig. 53, Fig. 54 and Fig. 55. The data presented in these plots were normalized by dividing each value of material loss due to oxidation and component wall temperature by the maximal measured wall thickness reduction and the maximal calculated temperature of the material on the external side at Position 2, respectively. The most important considerations regarding the obtained results are as follows:

- The physics-informed neural network, pre-trained basing on the results calculated by means of the experimentally determined Eq. (96), provides underestimated predictions of the final wall thickness reduction due to oxidation as shown in Fig. 53.

- The normalized root mean squared error evaluated against the unseen test data is equal to $NRMSE_{EMP,TEST}$ = 25%, and for the training data, it equals $NRMSE_{EMP,TRAIN}$ = 15%.

Considering the test dataset, the least accurate predictions are related to the parts with identifiers TP_5 and TP_6, for which the operational data coverage, i.e. the number of fired hours for which the required operating parameters are available, is equal to 63% and 61%, respectively. Nevertheless, attempts to train the hybrid model using a training set with a modified composition comprising TP_5 and TP_6 do not reduce the empirical

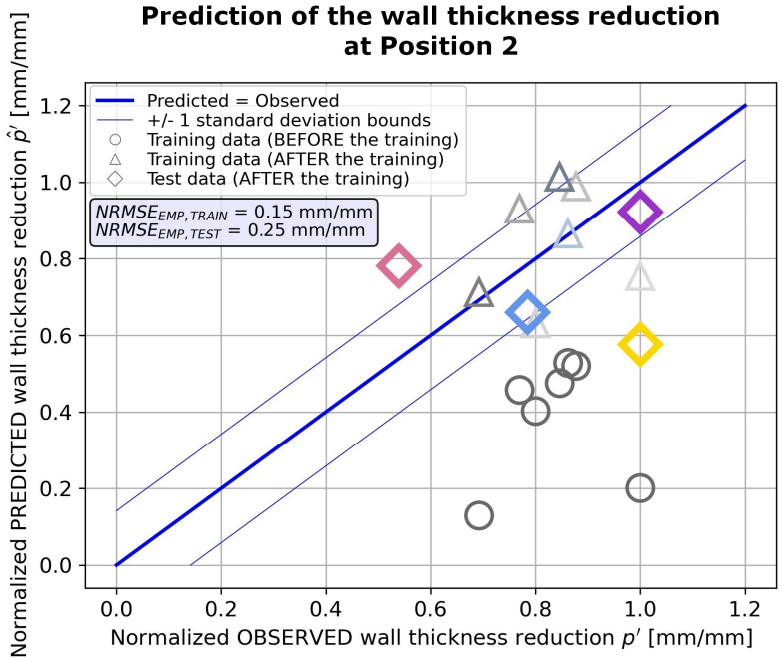


Fig. 53 Evaluation of the hybrid model based on the physics-informed neural network, trained using seven observations, against the training and test data presented on the predicted vs. observed plot.

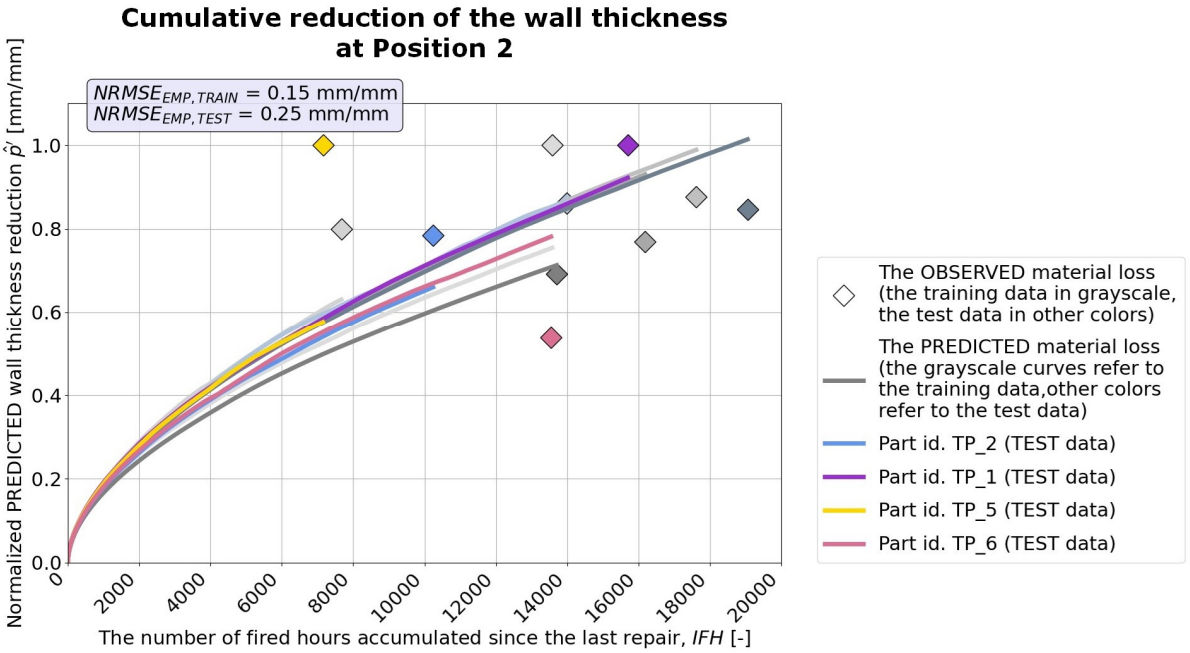


Fig. 54 Evaluation of the hybrid model based on the physics-informed neural network, trained using seven observations, against the training and test data.

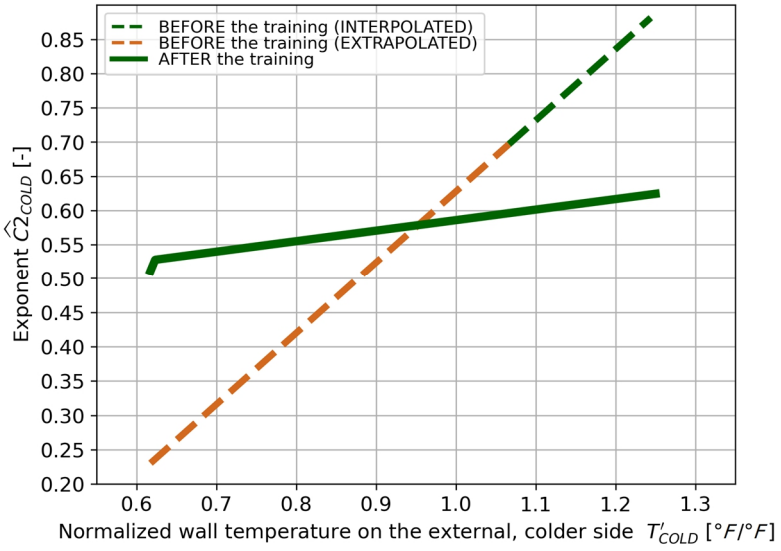


Fig. 55 Estimates of the $C2_{COLD}$ exponent determined by the data-driven layer before and after the training using seven observations $\widehat{C2}_{COLD} = FNN\left(T_{COLD}\right)$.

error value for these two observations. Thus, the lower prediction accuracy may be due to the low availability of the operational data, but also due to the assumption regarding the initial thickness of the transition piece. The difference between the upper and lower limits of the base material thickness is equal to $w_{nom}/10$. This difference can be converted to an interval with a width equal to 0.77, applying the same normalization approach as for the data shown in Fig. 53 and Fig. 54. The available measurements are within the interval with a width equal to 0.46. Thus, the measurement's range is smaller than the tolerance of the nominal base material thickness. Taking this into consideration and the limitations in terms of the sample size and the availability of the operational data, the accuracy of the final wall thickness reduction estimates is satisfactory and significantly better compared to the pre-trained model. However, it should be noted that for observations characterized by the highest material loss, the predictions obtained by means of the baseline model are underestimated.

- The final value of the constant $C1$ approximated using the single neuron perceptron is higher than the initial value and equal to $C1 = 1.377 \cdot C_{init}$.

- The feedforward neural network estimating values of the exponent $C2_{COLD}$ satisfies the implemented constraints resulting from Inequality (93), as presented in Fig. 55. The values of $\widehat{C2}_{COLD}$ obtained using the data-driven layer are linearly dependent on the values of the base material temperature T_{COLD}. However, the inclination of $\widehat{C2}_{COLD} = FNN\left(T_{COLD}\right)$ characteristic decreased after the training in comparison with the model pre-trained basing on the results calculated by means of the experimentally determined Equation (96). Thus, the impact of the material temperature changes on the value of $\widehat{C2}_{COLD}$ is lower. Simultaneously, it should be highlighted that the experiment was conducted at higher temperatures, i.e. $T'_{COLD} > 1.066$, than those analyzed. Executing the burner-rig test for oxidation resistance at lower temperatures, comparable

with the analyzed ones, i.e. $0.617 \leq T'_{COLD} \leq 1$, where the oxidation rate is expected to be lower, could result in a characteristic with a different inclination. As described previously, the monotonic multilayer perceptron implements a continuous piecewise linear function. It should be noted that the function $\widehat{C2}_{COLD} = FNN\left(T_{COLD}\right)$ valid for the original space of input data, i.e. when $T'_{COLD} \leq 1$, is applicable also when $T'_{COLD} > 1$ in the additional training subspace comprising synthetic data only.

- The hybrid model is capable of differentiating the estimates of damage increments and the final wall thickness reduction basing on the actual operational data. Nevertheless, due to the lower impact of the material temperature changes on the value of $\widehat{C2}_{COLD}$, the final material loss due to oxidation estimated by the physics-informed neural network depends more on the number of fired hours accumulated by each part than on the actual operating conditions.
- Fig. 55 presents estimates of the data-driven layer outside the original space of input data when $T'_{COLD} > 1$. This continuous piecewise linear function shows directly how the multilayer perceptron interpolates and extrapolates outside that space. Thereby, knowing the range of values of the input data, the predictions can be made in a fully conscious way, using this characteristic. If necessary, synthetically created data can enlarge the space in which the model should respect the prior knowledge about the analyzed phenomenon.

In order to evaluate the effectiveness of this modeling approach in an even more challenging scenario, the next attempt is to train the physics-informed neural network with the training set comprising two observations only and use the remaining data points for testing. Three different train-test splits are prepared and subsequently processed. The first training set comprises the transition pieces with identifiers TP_3 and TP_6, for which the baseline model gave one of the least accurate estimates. Additionally, these parts have a similar number of interval fired hours and are characterized by moderate availability of the operational data. The second training set comprises the transition pieces with identifiers TP_3 and TP_9, which are characterized by relatively long service periods, good availability of the operational data (i.e. for TP_9, the coverage equals 99%) and a significant difference in T_{COLD} values. The last training subset comprises the parts with identifiers TP_2 and TP_10, which are characterized by short service periods, a difference in terms of operational data availability and a significant difference in T_{COLD} values. The learning process was executed without any difficulties for each training set, and regardless of their composition, the obtained results are very similar. Therefore, to avoid excessive reporting, only the estimates and the characteristic of the model trained using the first subset are presented in the dissertation. These results are shown in Fig. 56, Fig. 57 and Fig. 58. The most important considerations regarding the obtained results are as follows:

- Despite the small number of training points, the final estimates of the wall thickness reduction due to oxidation are very similar to the predictions obtained by means of the baseline model trained using seven data points. The normalized root mean squared error evaluated against the entire sample is equal to $NRMSE_{EMP}$ = 19.6%, while concerning the baseline model, it equals 19.4%. The obtained results are consistent with each other, regardless of the cardinality and the composition of the training set. Therefore, cross-validation is not necessary.

- The final value of constant $C1$ approximated using the single neuron perceptron is higher than the initial value and equal to $C1 = 1.298 \cdot C1_{init}$.

- For $T'_{COLD} >\sim 0.714$ the characteristic of the multilayer perceptron presented in Fig. 58 is very similar to the characteristic of the baseline model. The difference between them

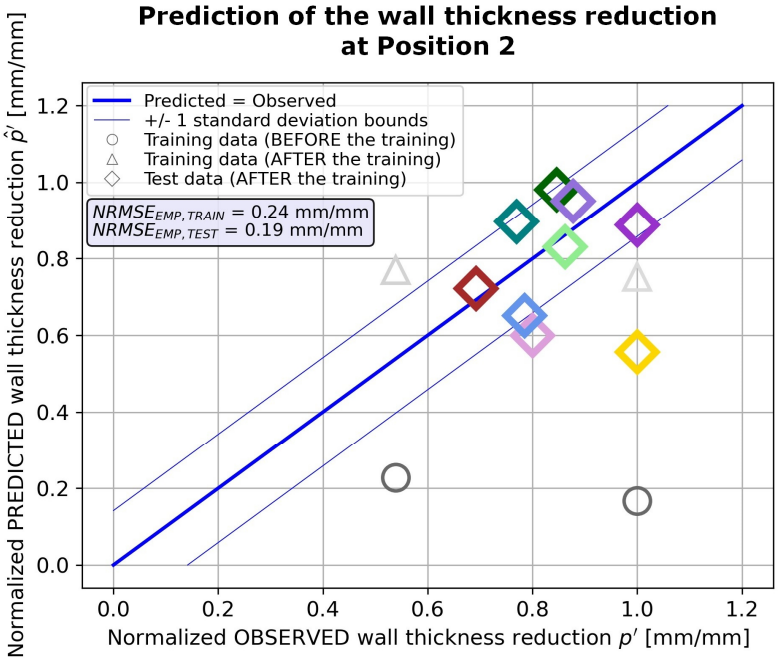


Fig. 56 Evaluation of the hybrid model based on the physics-informed neural network, trained using two observations, against the training and test data presented on the predicted vs. observed plot.

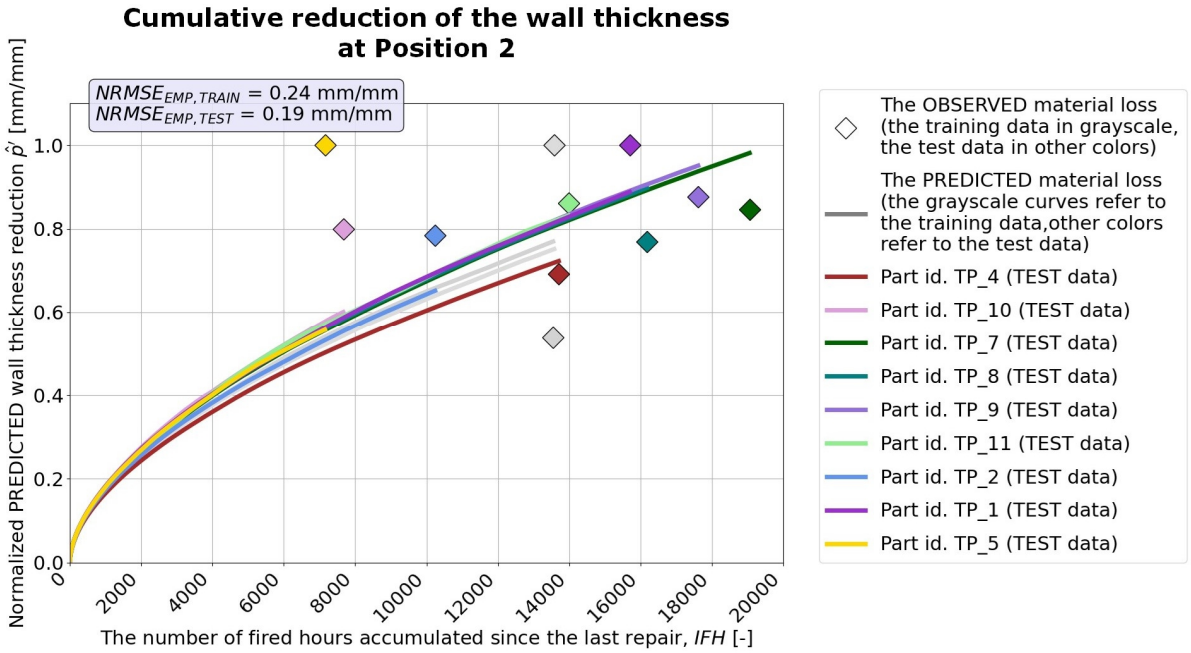


Fig. 57 Evaluation of the hybrid model based on the physics-informed neural network, trained using two observations, against the training and test data.

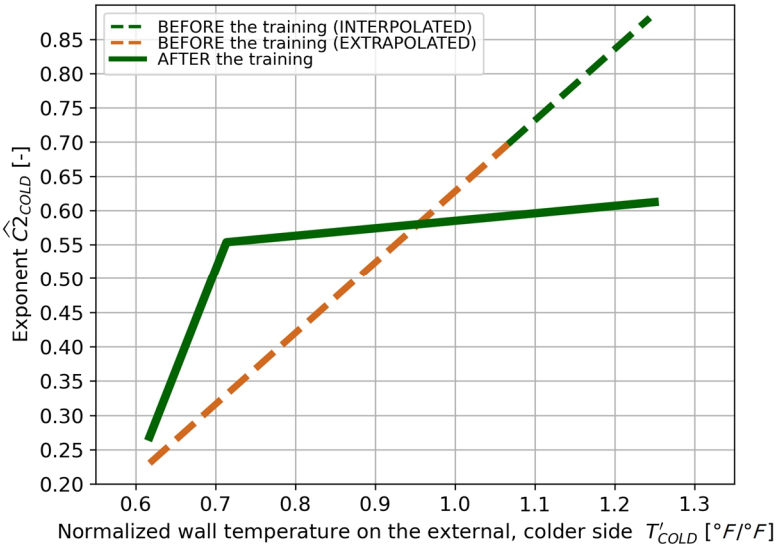


Fig. 58 Estimates of the $C2_{COLD}$ exponent determined by the data-driven layer before and after the training using two observations $\widehat{C2}_{COLD} = FNN\left(T_{COLD}\right)$.

for values of T'_{COLD} lower than ~0.714 can be disregarded since less than 0.08% of the data are within this interval.

The last objective is to train the physics-informed neural network with the training set comprising just one data point. Several attempts were made to modify the composition of the training subset (i.e. the components with identifiers TP_2, TP_3, TP_5, TP_6, TP_10 and TP_11 were considered) and execute the learning process. However, the obtained results were not completely satisfactory, with the model's accuracy higher than the pre-trained model's accuracy, i.e. the root mean squared error evaluated against the entire sample reduced up to 38%, but also significantly lower compared to the previously described models trained using seven and two observations, i.e. when the value of $NRMSE_{EMP}$ was at least 60% higher in comparison with the baseline model. Therefore, these results are not presented in this dissertation.

Finally, it should be noted that the value of the empirical term $RMSE_{EMP}$ can be further reduced if the constraints resulting from the prior knowledge about the analyzed phenomenon are implemented partially (i.e. the monotonic feedforward neural network is not applied, and the physical term $RMSE_{PHY}$ is either included or excluded from the cost function by setting $\gamma = 0$). Nevertheless, the reduction of the empirical term's value has a significant impact on the $\widehat{C2}_{COLD} = FNN\left(T_{COLD}\right)$ characterisitic of the multilayer perceptron, making it constant or even decreasing in some intervals of the domain. Thus, the characteristic may be inconsistent in some subdomains with the prior knowledge about the analyzed phenomenon and, consequently, are difficult to interpret. Despite the higher prediction accuracy, such models are considered invalid and cannot be applied to support business decisions or apply them in other domains.

6.5 Application of the proposed method for predicting metal loss due to oxidation

As described in Section 6.1, the selected target domain refers to the damage due to oxidation recorded at Position 3 of the analyzed transition pieces. The reduction in base material thickness due to oxidation is the dependent variable. The baseline model used to predict the wall thickness reduction at Position 2, which was trained using seven data points, is selected as the source hybrid model. The target domain is considered valid since the physical phenomenon causing the damage and the applicable framework of equations are the same as in the source domain. In the source domain the material temperature values on the internal and external sides of the transition piece are estimated using Eq. (91) and Eq. (92). The same approach is applied to calculate the temperature in the target domain, basing on the following formulas:

$$T_{HOT,TGT} = T_{INLET} - \phi_{HOT,TGT}\left(T_{INLET} - T_{COOL}\right), \tag{98}$$

$$T_{COLD,TGT} = T_{INLET} - \phi_{COLD,TGT}\left(T_{INLET} - T_{COOL}\right), \tag{99}$$

where $\phi_{HOT,TGT}$ and $\phi_{COLD,TGT}$ are the cooling effectiveness coefficients on the internal and external sides of the component wall at Position 3, respectively. Their values were determined specifically for the target domain basing on the numerical simulation results. In order to better visualize how the material temperatures change depending on the variation of T_{COOL} and T_{INLET}, polynomial regression models are prepared to approximate the relationship between these two operating parameters and the input variable of the physics-informed neural network:

$$\hat{T}_{COLD} = f\left(T_{COOL}, T_{INLET}\right), \tag{100}$$

where $\hat{T}_{COLD}$ is the estimated value of the material temperature on the external side of the part. One model is dedicated to the source domain and another to the target domain. The surface plots shown in Fig. 59 present the normalized results obtained using the two polynomial regression models (i.e. the same data are shown in each view). The actual values of the variables were divided by the maximum value of T_{COOL}, T_{INLET}, or T_{COLD} in order to obtain the normalized variables, i.e. T'_{COOL}, T'_{INLET} and T'_{COLD}. Based on these plots, it is visible that the surfaces are shifted relative to each other with the length of the displacement vector approximately constant independently of the selected pair of T_{COOL} and T_{INLET}. Using the actual operational data recorded by the data acquisition system and the material temperature values calculated by means of Eq. (92) and Eq. (99), i.e. the data presented with the scatter plots in Fig. 59, a scaling factor between the values of T_{COLD} and $T_{COLD,TGT}$ can be determined for each time step. The mean value of the factor equals $\overline{\varphi} = 0.9274$ with a standard deviation $\sigma_{\varphi} = 0.0047$. Thus, independently of time, it can be assumed that the calculated value of the material temperature on the external side of the transition piece wall related to Position 3 is equal, on average, to 92.74% of the material temperature value calculated for Position 2. Consequently, given a vector with the operating parameter's readings $\vec{x}_{TGT,t}$ recorded at time step t in the target domain, the shifted value of the material temperature is estimated in the following way:

$$T^{*}_{COLD,TGT,t} = \overline{\varphi} \cdot T_{COLD,TGT,t} = 0.9274 \cdot f_{TGT}\left(\vec{x}_{TGT,t}\right). \tag{101}$$

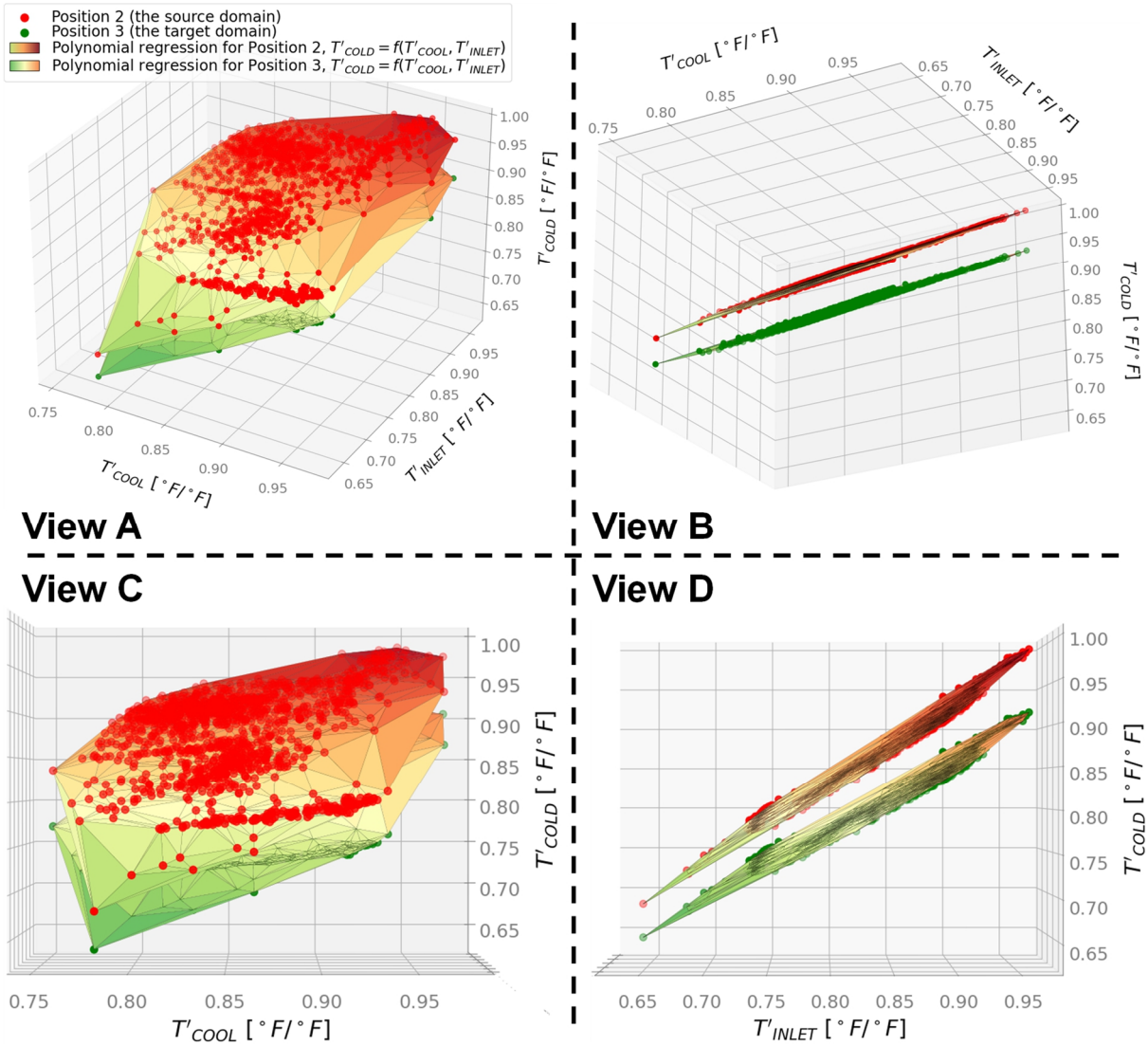


Fig. 59 The relationship between the average air temperature at the axial compressor's discharge T_{COOL}, the average temperature of gases at the inlet of the nozzles T_{INLET} and the material temperature on the external, colder side of the transition piece T_{COLD} in the source and target domains, i.e. Position 2 and Position 3, respectively.

Subsequently, in accordance with Eq. (90), the wall thickness reduction due to oxidation in each time interval $\Delta t = 1$ hour is calculated as follows:

$$\Delta p_{COLD,TGT,t} = \left(\frac{C2^*_{COLD,TGT,t} \cdot p_{COLD,TGT,t-1}}{\left(p_{COLD,TGT,t-1} / C1 \right)^{\frac{1}{C2^*_{COLD,TGT,t}}}} \right) \Delta t = \left(\frac{FNN\left(T^*_{COLD,TGT,t}\right) \cdot p_{COLD,TGT,t-1}}{\left(p_{COLD,TGT,t-1} / C1 \right)^{\frac{1}{FNN\left(T^*_{COLD,TGT,t}\right)}}} \right) \Delta t = \tag{102}$$

$$= \left(\frac{FNN\left(0.9274 \cdot f_{TGT}\left(\vec{x}_{TGT,t}\right)\right) \cdot p_{COLD,TGT,t-1}}{\left(p_{COLD,TGT,t-1} / C1 \right)^{\frac{1}{FNN\left(0.9274 \cdot f_{TGT}\left(\vec{x}_{TGT,t}\right)\right)}}} \right) \Delta t ,$$

where $C2^{*}_{COLD,TGT,t}$ is the exponent value calculated basing on the shifted value of the material temperature using the feedforward neural network, symbolically denoted as function *FNN*.

However, considering the entire service period, the matrix with shifted values of the operating parameters $\mathbf{X}^{*}_{TGT} = \left[\vec{T}^{*}_{COOL} \quad \vec{T}^{*}_{INLET} \right]$ can be determined differently. For example, it can be assumed that the proportion between the values of $T_{HOT,t}$ and $T_{COLD,t}$ calculated in the source domain using Equations (91) and (92) at specific time step *t* and denoted as ψ_t, is valid also for the shifted values of these temperatures in the target domain as specified by the following formula:

$$\frac{T_{HOT,t}}{T_{COLD,t}} = \frac{T^{*}_{HOT,TGT,t}}{T^{*}_{COLD,TGT,t}} = \psi_t \,. \tag{103}$$

Subsequently, the shifted values of the operating parameters, i.e. $T^{*}_{COOL,t}$ and $T^{*}_{INLET,t}$, can be determined by solving the following system of equations based on Eq. (91) and Eq. (92):

$$\begin{cases} T^{*}_{COOL,t} = \dfrac{T_{COLD,TGT,t} - \left(1-\phi_{COLD}\right) T^{*}_{INLET,t}}{\phi_{COLD}} \\ \psi_t T_{COLD,TGT,t} = \left(1-\phi_{HOT}\right) T^{*}_{INLET,t} + \phi_{HOT} T^{*}_{COOL,t} \end{cases}, \tag{104}$$

where $\psi_t T_{COLD,TGT,t} = T_{HOT,TGT,t}$ according to Eq. (103). The resultant shifted value of the average temperature of gases at the inlet of the nozzles is as follows:

$$T^{*}_{INLET,t} = \frac{\left(\psi_t - \dfrac{\phi_{HOT}}{\phi_{COLD}}\right) T_{COLD,TGT,t}}{\left(1-\phi_{HOT}\right) - \dfrac{\left(1-\phi_{COLD}\right)\phi_{HOT}}{\phi_{COLD}}}, \tag{105}$$

while the corresponding value of the average air temperature at the compressor's discharge is calculated by applying the first equation of relationships (104). Additionally, if more operating parameters are used for prediction purposes, the system of equations can be enlarged with further equations basing on the numerical simulation results. Such equations may have a form similar to Eq. (98). However, they will describe the transition piece body temperature in other areas by employing a dedicated value of the cooling effectiveness coefficient ϕ for each position.

The shifted values of the operational data can also be calculated by assuming that the values of a specific operating parameter in the source and target domains are proportional to each other. For example, if the source domain refers to gas turbines characterized by a nominal value of the average temperature of gases at the inlet of the nozzles equal to $T_{INLET,nom}$, and the target domain refers to units, where this temperature is 10% higher, i.e. $\frac{11}{10} T_{INLET,nom}$, then for any time step *t*, it can be assumed that the shifted value of this operating parameter is equal to:

$$T^{*}_{INLET,t} = \tfrac{11}{10} T_{INLET,t} \,, \tag{106}$$

when the material temperature of the target object is higher than that of the source object for any ordered pair of T_{COOL} and T_{INLET} (i.e. if the material temperature is estimated using Eq. (98), it is true when $\phi_{TGT} < \phi_{SRC}$). Otherwise, when the temperature of the target object is 10% lower in comparison with the source object, the value of $T^{*}_{INLET,t}$ is determined as follows:

$$T^{*}_{INLET,t} = \tfrac{9}{10} T_{INLET,t} \,. \tag{107}$$

Subsequently, the corresponding value of the average air temperature at the compressor's discharge is calculated by applying the first equation of relationships (104). Nevertheless, compared to the two approaches described before, this approach applied to determine the shifted values of the operating parameters is less appropriate in the analyzed case since both domains refer to the same gas turbines. Finally, the results presented in this section are based on Eq. (102), but practically the same predictions are obtained when the shifted values of the operating parameters are calculated in accordance with relationships (104).

The results obtained using the proposed method of single-source domain generalization and cross-domain knowledge transfer in regression analysis, which leverages physics-informed neural networks, are presented in Fig. 60 and Fig. 62. They can be compared with the results obtained by means of a hybrid model with the same architecture as the source PINN, which was trained from the ground up in the target domain using seven observations. These results are presented in Fig. 61 and Fig. 63. The learning process of that model was executed using the actual operational data, the domain-specific values of $T_{COLD,TGT}$ and the measurement data on the damage size recorded in the target domain. In order to protect proprietary information, the data presented in the plots were normalized by dividing each value by the maximal measured wall thickness reduction recorded in the target domain. The most important considerations regarding the obtained results are as follows:

- The pre-trained model provides underestimated predictions of the final wall thickness reduction due to oxidation in the target domain (the details are presented in Fig. 61), similarly to the attempt to apply it in the source domain.
- The accuracy of the predictions on the unseen data obtained using the proposed method is similar to the prediction accuracy of the hybrid model trained in the target domain. Nevertheless, additional clarifications are needed since the error changes significantly depending on the composition of the test subset. As shown in Fig. 62, it should be noted that the parts with identifiers TP_3, TP_4 and TP_6 accumulated more than 13 500 fired hours with negligible, i.e. $p' = 0.067$ for TP_4, or even without material loss observed. Moreover, for the part with identifier TP_7 and the highest value of interval fired hours IFH = 19 058, the normalized wall thickness reduction is equal to $p' = 0.200$. These measurements are unexpected, differ significantly from the remaining ones and are difficult to be interpreted. Presumably, they are erroneous and caused by incorrectly taken measurements or improper recordkeeping. Including these observations in the training subset significantly affects the estimates of the fully-specified model. Thus, these four data points are in the test set, but they should eventually be disregarded. The model created from the ground up in the target domain is trained basing on the remaining seven data points. Because the model was built only to evaluate the accuracy of the proposed method of domain generalization, its performance was not verified on a valid

test set, and it may be too sensitive to the peculiarities of the training set. The normalized root mean squared error evaluated against the training data points equals $NRMSE_{EMP,TRAIN} = 16.6\%$, where $NRMSE_{EMP,TEST} = 58.6\%$. When the proposed method is applied, the error evaluated against the same seven data points is equal to

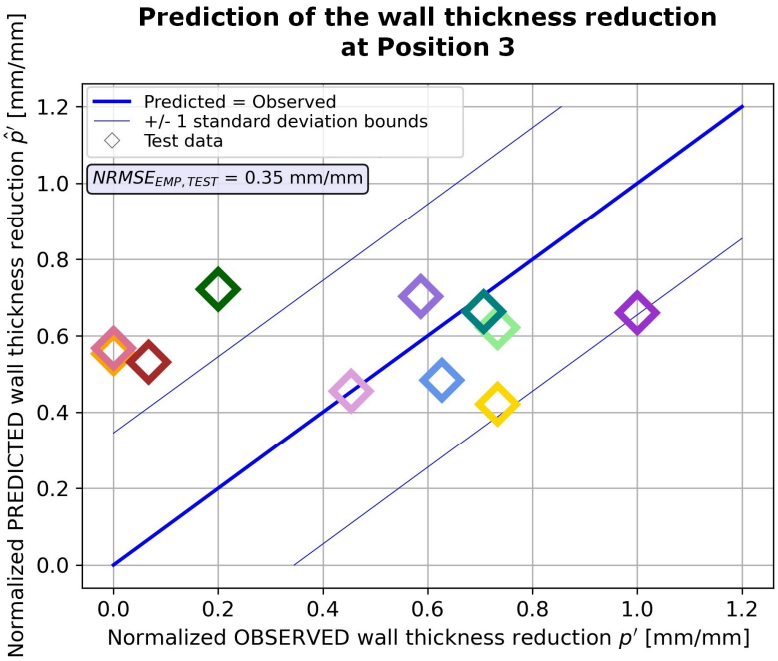


Fig. 60 Evaluation of the source hybrid model based on the physics-informed neural network, trained in the source domain using seven observations, against the test data from the target domain presented on the predicted vs. observed plot.

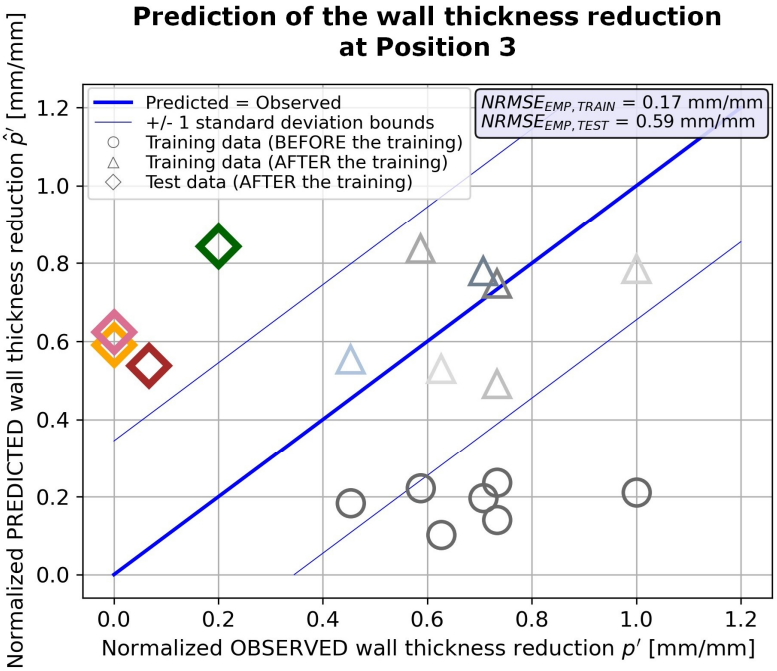


Fig. 61 Evaluation of the hybrid model based on the physics-informed neural network, trained from the ground up in the target domain using seven observations, against the training and test data presented on the predicted vs. observed plot.

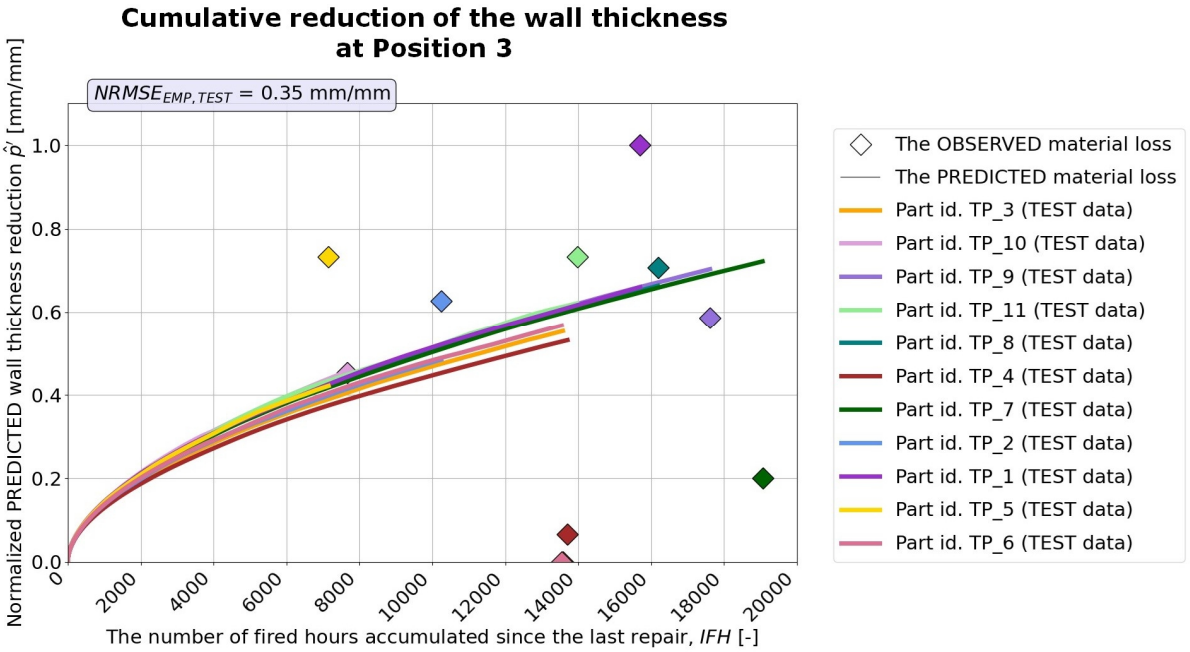


Fig. 62 Evaluation of the source hybrid model based on the physics-informed neural network, trained in the source domain using seven observations, against the test data from the target domain.

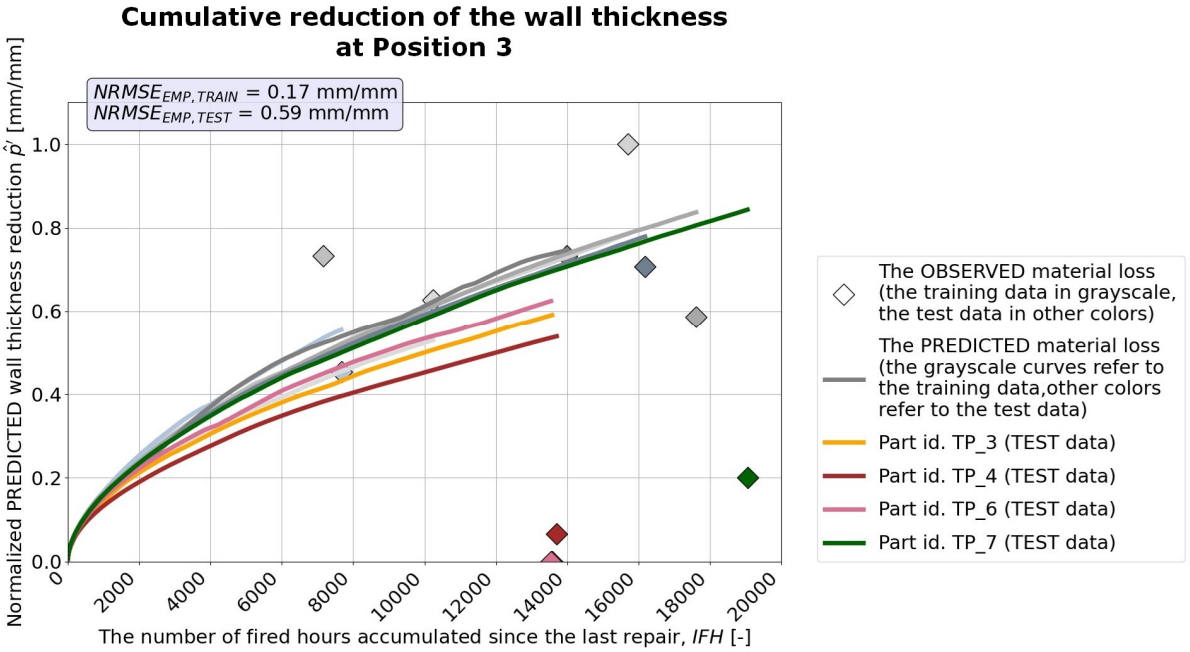


Fig. 63 Evaluation of the hybrid model based on the physics-informed neural network trained, from the ground up in the target domain using seven observations, against the training and test data.

$NRMSE_{EMP,TEST}$ = 19.3% and 52.8%, when evaluated against the remaining four observations. Thus, the obtained results provide further evidence that this method of domain generalization and cross-domain knowledge transfer in regression analysis, leveraging physics-informed neural networks, can be applied effectively even if no

damage size measurements are available in the target domain, providing accurate damage size predictions.

- The final value of constant $C1$ of the model trained in the target domain is greater than the initial value and equal to $C1 = 1.517 \cdot C1_{init}$. The obtained value is similar to that of the source hybrid model.
- For $T'_{COLD} \geq\sim 0.675$ the characteristic of the multilayer perceptron trained in the target domain presented in Fig. 64 is comparable to the characteristic of the source hybrid model. Considering the small distance between the analyzed positions, such a similarity between the characteristics was expected. Because of the lower impact of the material temperature changes on the estimated values of the exponent $C2_{COLD}$ (i.e. compared to the model trained on the results calculated using the experimentally determined Equation (96) as the reference), the final wall thickness reduction due to oxidation estimated by the physics-informed neural network depends more on the number of fired hours accumulated by each part than on the actual operating conditions. The difference between the characteristics of the feedforward neural network for values of T'_{COLD} smaller than ~0.675 can be disregarded since less than 0.09% of the data in the target domain are within this interval. The values presented in this figure were normalized by dividing them by the maximal calculated temperature of the material on the external side related not to the target domain but to the source one, i.e. of Position 2. It simplifies a comparison with the results shown in Fig. 55 related to the source hybrid model.

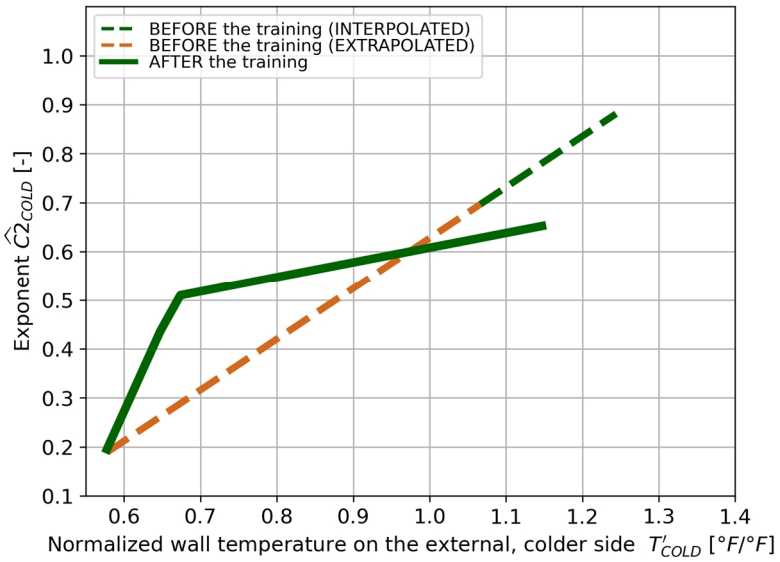


Fig. 64 Estimates of the $C2_{COLD}$ exponent determined by the data-driven layer before and after the training from the ground up in the target domain using seven observations $\widehat{C2}_{COLD} = FNN\left(T_{COLD}\right)$.

6.6 Conclusions concerning this stage of the research

The last stage of the research aimed to evaluate the flexibility and versatility of the proposed method of single-source domain generalization and cross-domain knowledge transfer in regression analysis, which leverages physics-informed neural networks. In accordance with the specific objectives of the dissertation, this method shall be universal and applicable for modeling damage due to various failure types found on the main components of gas turbines. Thus, after the successful application of the method for predicting the maximal length of fatigue cracks found on the high-pressure nozzles, it was used to predict the reduction in the transition piece wall thickness due to oxidation. The primary outcomes of this part of the research are as follows:

- A physics-informed neural network was designed and successfully applied to model the wall thickness reduction due to oxidation based on actual operational data recorded by the data acquisition system. The PINN is basing on a recurrent neural network. The cell of the RNN comprises two feedforward neural networks and a physics-based part. The first multilayer perceptron, referred to as the data-driven layer, estimates the values of $C2_{COLD}$ considering as the input material temperature T_{COLD} on the external, colder side of this component, which is calculated basing on the selected operating parameters. The second perceptron comprises a single neuron only and is applied to estimate the value of $C1$ constant. The estimates of $C2_{COLD}$ and $C1$ are used in the physics-based part, which computes the material loss due to oxidation in each time interval Δt = 1 hour, assuming that the rate of oxide scale formation decreases with time and can be described using a power function. It was assumed that the thermal barrier coating on the transition piece's internal, hotter side effectively protects against metal loss due to oxidation throughout the entire service period until the component is disassembled and subject to inspection and repair activities. Attempts to embed estimates of the number of startup-shutdown cycles until the coating spalls off into the hybrid model, exposing the base material, and then to model the oxide scale formation on the internal side of the component were not successful. This concern should be addressed during future research activities. The constraints resulting from Inequality (93), which the fully-specified PINN respects, are implemented in a hard manner, applying the monotonic neural network responsible for estimating exponent $C2_{COLD}$, but also in a soft manner, through the additional term in the customized cost function, i.e. the physical term $RMSE_{PHY}$.
- The general form of the applied custom cost function is the same as for the previously analyzed high-pressure nozzles, with a variable coefficient $\gamma \in [0,1]$ determining how the physical and empirical, i.e. $RMSE_{EMP}$, terms contribute to the cost value during the training process in accordance with Eq. (56). The value of this coefficient is recalculated after each epoch using the predefined function described by Eq. (97) to find a proper balance between the two terms and quickly increase the coefficient value if the value of $RMSE_{PHY}$ rises. The applied constraints are effective, maintaining the low value of the physical term while reducing the empirical term's value as the training process progresses. The original input space, which is based on the empirical data available in the source domain, was enlarged with synthetic data, which simultaneously enlarged the

training space, where the physics-informed neural network should respect the theoretical constraints.

- The analyzed sample comprises only 11 data points. The same initial thickness of the wall is assumed for all the observations since the material thickness measurements, made after manufacturing, are unavailable. The lack of this information introduces significant uncertainties since the measurements' range, i.e. the range of the dependent variable values, is smaller than the tolerance of the base material thickness. In this complicated case two hybrid models sharing the same architecture but trained respectively basing on seven and two observations were prepared. Considering the limitations in the quantity and quality of available empirical data, the accuracy of the final wall thickness reduction estimates is satisfactory and better compared to the model pre-trained based on the results calculated using the experimentally determined Equation (96) as the reference. The least accurate predictions are related to the parts with the lowest availability of the actual operational data. The obtained results are consistent with each other, regardless of the cardinality and the composition of the training set. Therefore, cross-validation is not necessary. Finally, attempts to train the model using a single training data point were not completely satisfactory, regardless of the selected observation. The accuracy of such models was greater than that of the pre-trained model but significantly smaller compared to the hybrid models trained using seven and two data points.

- The final value of constant $C1$ approximated using the single neuron perceptron is higher than the initial value $C1_{init}$ by up to 38%, considering the models trained in the source domain and by 52% for the hybrid model created in the target domain.

- The characteristics $\widehat{C2}_{COLD} = FNN\left(T_{COLD}\right)$ of the monotonic multilayer perceptrons trained in the source and target domains are similar to each other. They illustrate that the data-driven layer obeys the constraints resulting from Inequality (93) and show how the feedforward neural network estimates inside and outside the original input space, i.e. when $T'_{COLD} \le 1$ and $T'_{COLD} > 1$, respectively. Generally, the values of $\widehat{C2}_{COLD}$ determined by the data-driven layer depend linearly on the values of the base material temperature T_{COLD}. The inclinations of the characteristics obtained after the training process are smaller compared to the pre-trained model. Nevertheless, the pretraining was based on the results of a burner-rig test for oxidation resistance, which was conducted at significantly higher temperatures, i.e. $1.066 < T'_{COLD} < 1.244$, than those analyzed. The created hybrid models can differentiate the estimates of damage increments and the final wall thickness reduction basing on the actual operational data. However, due to the smaller impact of the metal temperature changes on the value of $\widehat{C2}_{COLD}$, the final material loss due to oxidation estimated by the physics-informed neural network depends more on the number of interval fired hours accumulated by each transition piece than on the actual operating conditions.

- The created method of domain generalization was applied to predict the wall thickness reduction at Position 3, leveraging the baseline hybrid model dedicated to Position 2. Three different approaches were described to estimate the shifted values of the material temperature on the external side of the wall, which is the input variable of the physics-

informed neural network. The material temperature value corresponding to Position 3 is 92.74% of the temperature value calculated for Position 2 for any vector of the chosen operating parameters. Seven wall thickness measurements recorded in the target domain were used to evaluate the prediction's accuracy, while the remainder was considered erroneous and disregarded. When the proposed method is applied, the normalized error evaluated against the seven valid data points equals $NRMSE_{EMP,TEST}$ = 19.3%, and is similar to the error resulting from the application of a model trained from the ground up in the target domain that is equal to $NRMSE_{EMP,TRAIN}$ = 16.6%. Additionally, it should be noted that the latter's performance was not evaluated against any valid test set, and it may be too sensitive to the peculiarities of the training data. Thus, the outcomes obtained by means of the proposed method are considered satisfactory. Compared to the research described in the previous chapter, the method was successfully applied to another type of gas turbine component and a different type of failure in a simulated scenario, where damage size measurements are unavailable. Thus, the final results provide evidence that this method of single-source domain generalization and cross-domain knowledge transfer in regression analysis, which leverages physics-informed neural networks, has a universal character and it can be effectively applied for modeling damage due to various failure types found on the main parts of gas turbines.

7. Summary and conclusions

The main objective of this research was to create a method based on machine learning algorithms to predict the size of damage to gas turbine components. In accordance with the specific objectives of the dissertation, this method shall be applicable at the early stages of the parts' lifetime for modeling damage due to various failure types found on the main components of turbines. Furthermore, it should be effective even if the availability of damage size measurements is limited. Finally, models created by means of this method will respect prior knowledge and physical laws governing the analyzed phenomenon. These objectives were achieved by creating a novel method of single-source domain generalization and cross-domain knowledge transfer, which leverages physics-informed neural networks. In this context, the notion of "leveraging" is considered as the action of using something already available in order to achieve something new or better and maximize advantages. The conducted research addressed two different technical problems. The first one was related to predicting the maximal length of fatigue cracks found on the trailing edges of 1st stage nozzles of a heavy-duty gas turbine. The second one was related to predicting the wall thickness reduction due to oxidation in transition pieces of industrial turbines. The analyzed objects perform different functions, are subjected to various loads and are the parts of two different models of gas turbines manufactured by Baker Hughes Company. Additionally, the analyzed failure types are governed by different physical laws.

7.1 Conclusions

The application of several popular statistical learning algorithms to the first technical problem is described in Chapter 4, with specific, detailed conclusions in Section 4.7. The application of physics-informed neural networks to this problem is described in Chapter 5, with the procedure establishing the novel method of single-source domain generalization given in Section 5.4 and the dedicated conclusions in Section 5.6. The application of PINNs and the proposed method to the second technical problem is described in Chapter 6, with the dedicated conclusions provided in Section 6.6. The general conclusions of this research are as follows:

- It was shown that data-driven machine learning algorithms can be applied effectively at the early stages of the part's lifetime when the availability of damage size measurements is limited and the number of observations is small. Despite these limitations, they can provide accurate estimates of the damage size based on actual operational data gathered by the data acquisition system. Furthermore, the applied algorithms are universal and can be utilized for modeling damage due to various failure types found on the main parts of gas turbines. However, even though the selected features reflected well knowledge about the cause of the analyzed phenomenon, the created data-driven predictive models did not obey theoretical or experimentally determined equations describing the damage growth. This shortcoming creates doubts regarding the interpolation capabilities in the entire training domain and eliminates the possibility of using such models to extrapolate.
- A controlled approach to preparing the training, validation and test sets improved the training effectiveness, made the cross-validation more meaningful and simplified the evaluation and interpretation of results. Due to the limited number of available damage size measurements, elements of these sets were picked considering multiple constraints

resulting from the analysis of the composition and clusters of the entire sample instead of drawing them randomly. As a consequence, each set represents the whole dataset in a quantitative and qualitative way.

- The data-driven models were trained using customized loss and cost functions, which directly reflect the specific objective of obtaining the highest prediction accuracy for the most damaged parts, which may jeopardize the turbine's availability if not appropriately maintained. The applied loss function was based on a variable-width scoring interval defined by two scoring bounds, which get closer to each other as the observed damage size increases. Estimates falling within the scoring interval are considered successful predictions with sufficient accuracy. Moreover, accurate estimates for the most damaged components are rewarded with a bonus, promoting models with the specific demanded characteristics. Applying customized loss and cost functions results in models providing more valuable data-driven insights to support, e.g. business decisions, compared to the commonly applied cost function like the (root) mean squared error.

- In order to prepare more specific input for a physics-informed neural network or any predictive model based on machine learning algorithms, simplify the processing of data within the PINN's structure and make the training process and the interpretation of the obtained results easier, the actual operational data can be appropriately preprocessed by employing a simple physics-based model. Besides the theoretical knowledge about the analyzed phenomenon, such a model may be adjusted based on numerical simulation results, or it can use the outcomes of material properties tests. Nevertheless, it should be clarified that leveraging prior knowledge to prepare PINN's input data does not imply that the fully-specified hybrid model will respect any law describing the damage growth due to the analyzed phenomenon.

- Two different physics-informed neural networks were designed and effectively applied to address the analyzed technical problems, i.e. a separate model for predicting fatigue crack growth and another one for estimating the wall thickness reduction due to oxidation. They are based on a recurrent neural network, which is a rational choice for modeling progressive, irreversible damage accumulation. The custom cells of the RNNs comprise data-driven components and a physics-based component. The former ones are based on a feedforward neural network. They are used to approximate variables whose values are insufficiently accurately estimated by means of applied theoretical equations or whose prediction uncertainty is too high. The FNNs' outputs are the inputs of the physics-based component calculating the damage increment in each time step based on the applicable theoretical or experimentally determined law. The Paris' law was applied for the fatigue crack growth, while the rate of oxide scale formation is described with a power function. The constraints resulting from prior knowledge about the analyzed phenomena, which the fully-specified physics-informed neural networks respect, were implemented either in a hard manner, by applying a monotonic neural network as the data-driven component, or in a soft manner through the additional term in the customized cost function, i.e. the physical term $RMSE_{PHY}$, or both.

- A custom cost function was designed and effectively applied to train the models based on physics-informed neural networks. It has a variable coefficient $\gamma \in [0,1]$ determining

how the physical and empirical, i.e. $RMSE_{EMP}$, terms contribute to the final cost value during the training process. If γ equals zero, the function reduces to the empirical term only. In contrast, if it equals one, the function simplifies to the physical term only, and when it has intermediate values, both terms are included accordingly. The coefficient value depends on the physical term's value and changes dynamically during the learning process in conformity with the predefined function to find a proper balance between the two terms. In general, the value of γ reduces as the value of $RMSE_{PHY}$ decreases. Thus, the first stage of the training process aims to limit the space of allowable solutions only to a subspace where the implemented constraints resulting from prior knowledge about the analyzed phenomenon are satisfied, and the second stage is aimed at reducing the empirical term's value as much as possible while staying in that subspace. The coefficient value can be recalculated even after each epoch, i.e. after one complete pass of the algorithm through the entire dataset, to react quickly and increase it if the value of $RMSE_{PHY}$ rises. Such an approach, with dynamic changes of the weights assigned to the cost function terms, was applied to avoid biased estimates, which are either inaccurate but with very low values of the physical term or very accurate but with high values of the physical term, indicating poor generalization capabilities of the model. The PINNs were pre-trained to initiate the main part of the training process with the lowest possible value of γ. The pretraining is based on the data obtained using applicable theoretical or experimentally determined equations as a reference. Considering the main part of the training, the set of input data utilized as the basis to calculate the value of the physical term was created synthetically. It was generated in a systematic manner to improve the training process's effectiveness and the extrapolation capabilities of the data-driven components, which are embedded in the RNN cells. The original input space, which is based on the empirical data available in the source domain, was enlarged with synthetic data, which simultaneously enlarged the training space, where the physics-informed neural network should respect the theoretical constraints. This approach applied to generate these additional training data can be adjusted depending on particular objectives.

- In the analyzed cases concerning the so-called small data regime, the designed physics-informed neural networks were trained effectively based on ten, seven, two, or even one training observation, i.e. a damage size measurement. Considering the limitations in the quantity and quality of the operational data and damage size measurements and the simplified theoretical description of the analyzed phenomena applied to build the hybrid models, the accuracy of the obtained damage size estimates is satisfactory. The results are consistent with each other, regardless of the cardinality and the composition of the training set. Thus, cross-validation is not necessary. The resultant physics-informed neural networks have good generalization capabilities and may extrapolate accurately, despite the limited training subspace. For example, concerning the fatigue crack growth prediction, the models trained using two and one observation can provide accurate crack length estimates for parts with around 40% longer cracks and over four times more startup-shutdown cycles in comparison with the utilized training data points.

- Effective execution of the optimization process based on a single training observation was accomplished in only one case, which was related to the prediction of fatigue crack length. Considering this technical problem, five other attempts were unsuccessful, each

with a different training data point. Six different observations were analyzed for the second technical problem, but the results were unsatisfactory. The prediction accuracy of the resultant models was higher than that of the pre-trained model but significantly lower compared to the hybrid models trained using more data points. Extracting and leveraging knowledge from a single damage size measurement is valuable and may give tangible benefits. However, with such a small sample, difficulties in converging to the optimal solution may be encountered, and the final predictions may vary significantly depending on the characteristics of the selected training observation.

- The data-driven components approximating the key input variables of the physics-based parts are the core of the created physics-informed neural networks. These components determine, to a great extent, the final predictions obtained by means of the fully-specified hybrid models. Those feedforward neural networks cannot be considered as a black box since a characteristic of each data-driven component is available, presenting the relationship between the inputs and the output directly. Such a characteristic shows how the multilayer perceptron estimates inside the original space of input data but also outside it, i.e. in the training subspace created synthetically. Considering the two analyzed technical problems, the obtained characteristics illustrate that the implemented constraints are satisfied and that the data-driven components obey the underlying laws governing the investigated phenomena. Consequently, leveraging these characteristics, the predictions can be made in a fully conscious way. Otherwise, they can be utilized to limit the input space where the fully-specified model can be safely applied. However, the training space can be further enlarged with synthetically created data if necessary.

- A novel method of single-source domain generalization and cross-domain knowledge transfer in regression analysis leveraging physics-informed neural networks has been proposed. The method's purpose is to effectively apply a cumulative damage model based on a PINN, i.e. the so-called source hybrid model, which was trained in a domain, where measurement data on damage size are available, in another domain, where such data are unavailable. Following the definitions given in Section 5.4, the former domain is called the source domain, and the latter is the target domain. This method can be applied if the damage type in the target domain, the physical phenomena causing it and the applicable theoretical or experimentally determined equations for estimating the damage increment in a single time step are the same as in the source domain. The eight-step procedure describing the main steps of the proposed method is provided at the end of Section 5.4. The so-called shifting of the operational data and PINN's input variables in the target domain is the essential step that enables an effective application of the unchanged source hybrid model in the target domain. The composition of functions used to calculate the input data of the physics-informed neural network based on the selected operating parameters shared across the domains must be known in both domains to perform that step. In Section 6.5, three different approaches were described to estimate the shifted values of the PINN's input variable concerning the second of the investigated technical problems. Nevertheless, it does not exhaust all possibilities, which may vary depending on the peculiarities of the analyzed case. In the end, the final estimate in the target domain for the original set of operational data is obtained by applying the source hybrid model to the set of shifted operational data.

- The proposed method of domain generalization and cross-domain knowledge transfer in regression analysis, which leverages physics-informed neural networks, was applied to the analyzed technical problems. Concerning the fatigue crack growth, the source model related to a specific position on one airfoil was used to predict the maximal length of cracks on another airfoil of the high-pressure nozzle. Concerning the wall thickness reduction due to oxidation, the source model related to a specific area of the transition piece was used to predict the material loss in another area of the same component. In both cases, a constant shift between the input variables of the physics-informed neural networks in the source and target domains was assumed, independently of values of the operating parameters selected as the features. The damage size measurements available in the target domains were used to evaluate the prediction's accuracy. By applying the proposed method, the normalized root mean squared errors evaluated against the unseen data had similar values as the errors resulting from applying hybrid models trained from the ground up in the target domains, i.e. with the same structure as the source PINNs. Moreover, it should be noted that some elements of the dataset used for evaluating the performance of the applied method were used to train the hybrid models created in the target domains. Thereby, the results obtained by means of the proposed method are considered satisfactory, providing accurate damage size predictions in a simulated scenario where the training observations, i.e. damage size measurements, are missing. The latent information and knowledge extracted from the empirical data available in the source domain were leveraged and effectively used in the target domain.

Based on the obtained results, it can be concluded that the created method of single-source domain generalization and cross-domain knowledge transfer, which leverages physics-informed neural networks, can accurately predict the size of damage to gas turbine components in domains characterized by the limited availability of damage size measurements or lack of them. The damage accumulation models based on physics-informed neural networks were trained effectively using ten, seven, two, or even one training observation. Despite the limited training subspace, these models have good generalization capabilities and may extrapolate accurately outside the original input space. The hybrid models can accurately predict the extent of damage to gas turbine parts, even if the available operational data gathered by a data acquisition system are incomplete. For the analyzed technical problems, the required operational data were available on average for 54% of startup-shutdown cycles and 86% of service hours. Thus, the theses posed in this research have been proven.

7.2 Novelties resulting from this research

In summary, the primary novelties resulting from this research are as follows:

1) Creation of a physics-based method of single-source domain generalization and cross-domain knowledge transfer in regression analysis that enables accurate predictions with no observations available (i.e. damage size data), leveraging a physics-informed neural network trained in another domain with some observations existing.
2) A model based on a physics-informed neural network of fatigue crack propagation was created for the high-pressure nozzles made of FSX-414 cobalt-based superalloy. The analyzed dataset was not created synthetically and the predictions are based on actual operational data gathered by the Remote Monitoring & Diagnostics system.

3) A model based on a PINN of material depletion due to oxidation was created for the transition pieces made of Nimonic 263 nickel-based superalloy. The analyzed dataset was not created synthetically and the predictions are based on actual operational data gathered by the RM&D system.
4) A proof that the utilized structure of physics-informed neural network, i.e. composed of a recurrent neural network with a feedforward neural network embedded in the RNN cell, can be trained effectively and has good generalization capabilities, even if the number of training observations equals ten, two, or one.
5) Custom cost functions were created to train the models based on machine learning algorithms. The first function favors solutions that accurately predict the longest cracks and awards them with a bonus. The second function uses a variable parameter to bind the empirical error with the error term describing the consistency of the model with the applicable theoretical equations. The parameter's value can be changed dynamically to control and guide the learning process, depending on the quality and quantity of data and knowledge of the underlying laws governing the analyzed phenomena.
6) Computer programs were written in Python programming language. They were used to extract features based on time series captured by the RM&D system, create the synthetic data applied to expand the source domain, build the structures of the physics-informed neural networks, process the data through them and train the PINNs.

7.3 Future research

The obtained results are very promising and provide a motivation to explore further applications of physics-informed neural networks and the proposed method of single-source domain generalization for predicting damage to gas turbine components. Future research works should address the following points:

- Evaluating the effectiveness of the proposed method when the source and target objects refer to the same type of component, e.g. a rotating blade, but two different types of gas turbines characterized by significant differences in the values of critical operating parameters, e.g. the nominal firing temperature.
- Evaluating the effectiveness of the proposed method when the source and target objects are two different types of gas turbine components, e.g. a high-pressure nozzle and a rotating blade, respectively.
- Evaluating the effectiveness of the proposed method when applied to other failure types.
- Exploring and defining the method's limitations regarding its applicability, flexibility and capabilities. Especially in terms of evaluating if the target domain is valid and if the source hybrid model can be safely employed there, providing credible predictions. In this regard, literature on transfer learning can be leveraged.
- Defining a standardized set of metrics and quality indicators to evaluate the performance of the source hybrid model when applied in the target domain, where no damage size measurements are available for testing purposes.
- Exploring applications of PINNs and the proposed method for classification problems.
- Defining best practices related to creating, deploying and maintaining hybrid models based on physics-informed neural networks in an engineering practice.

Bibliography

[1] 2023, *Statistical Review of World Energy 2023*, 72nd edition, Energy Institute, London.

[2] Beagle, D., Moran, B., McDufford, M., and Merine, M., 2017, "GER-3620P Heavy-Duty Gas Turbine Operating and Maintenance Considerations."

[3] Johnston, J. R., 2000, "GER-3571H Performance and Reliability Improvements for Heavy-Duty Gas Turbines."

[4] Badeer, G. H., 2000, "GER-3695E GE Aeroderivative Gas Turbines - Design and Operating Features."

[5] Jardine, A. K. S., Lin, D., and Banjevic, D., 2006, "A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance," Mechanical Systems and Signal Processing, **20**(7), pp. 1483–1510.

[6] Ali, A., and Abdelhadi, A., 2022, "Condition-Based Monitoring and Maintenance: State of the Art Review," Applied Sciences, **12**(2), p. 688.

[7] Tahan, M., Tsoutsanis, E., Muhammad, M., and Abdul Karim, Z. A., 2017, "Performance-Based Health Monitoring, Diagnostics and Prognostics for Condition-Based Maintenance of Gas Turbines: A Review," Applied Energy, **198**, pp. 122–144.

[8] Fentaye, A. D., Baheta, A. T., Gilani, S. I., and Kyprianidis, K. G., 2019, "A Review on Gas Turbine Gas-Path Diagnostics: State-of-the-Art Methods, Challenges and Opportunities," Aerospace, **6**(7), p. 83.

[9] Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., and Li, G. P., 2020, "Predictive Maintenance in the Industry 4.0: A Systematic Literature Review," Computers & Industrial Engineering, **150**, p. 106889.

[10] Ansari, F., Glawar, R., and Sihn, W., 2020, "Prescriptive Maintenance of CPPS by Integrating Multimodal Data with Dynamic Bayesian Networks," *Machine Learning for Cyber Physical Systems*, J. Beyerer, A. Maier, and O. Niggemann, eds., Springer, Berlin, Heidelberg, pp. 1–8.

[11] Liu, B., Lin, J., Zhang, L., and Kumar, U., 2019, "A Dynamic Prescriptive Maintenance Model Considering System Aging and Degradation," IEEE Access, **7**, pp. 94931–94943.

[12] Goby, N., Brandt, T., and Neumann, D., 2023, "Deep Reinforcement Learning with Combinatorial Actions Spaces: An Application to Prescriptive Maintenance," Computers & Industrial Engineering, **179**, p. 109165.

[13] Vanderschueren, T., Boute, R., Verdonck, T., Baesens, B., and Verbeke, W., 2022, "Prescriptive Maintenance with Causal Machine Learning."

[14] Sepe, M., Graziano, A., Badora, M., Stazio, A. D., Bellani, L., Compare, M., and Zio, E., 2021, "A Physics-Informed Machine Learning Framework for Predictive Maintenance Applied to Turbomachinery Assets," Journal of the Global Power and Propulsion Society, **2021**(May), pp. 1–15.

[15] Grieves, M., 2014, "Digital Twin: Manufacturing Excellence through Virtual Factory Replication," White paper, **1**, pp. 1–7.

[16] Liu, M., Fang, S., Dong, H., and Xu, C., 2021, "Review of Digital Twin about Concepts, Technologies, and Industrial Applications," Journal of Manufacturing Systems, **58**, pp. 346–361.

[17] Thelen, A., Zhang, X., Fink, O., Lu, Y., Ghosh, S., Youn, B. D., Todd, M. D., Mahadevan, S., Hu, C., and Hu, Z., 2022, "A Comprehensive Review of Digital Twin — Part 1: Modeling and Twinning Enabling Technologies," Struct Multidisc Optim, **65**(12), p. 354.

[18] Semeraro, C., Lezoche, M., Panetto, H., and Dassisti, M., 2021, "Digital Twin Paradigm: A Systematic Literature Review," Computers in Industry, **130**, p. 103469.

[19] Errandonea, I., Beltrán, S., and Arrizabalaga, S., 2020, "Digital Twin for Maintenance: A Literature Review," Computers in Industry, **123**, p. 103316.
[20] Tao, F., Qi, Q., Wang, L., and Nee, A. Y. C., 2019, "Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison," Engineering, **5**(4), pp. 653–661.
[21] Carlevaro, F., Cioncolini, S., Sepe, M., Parrella, I., Allegorico, C., De Stefanis, L., Mastroianni, M., and Escobedo, E., 2018, "Use of Operating Parameters, Digital Replicas and Models for Condition Monitoring and Improved Equipment Health," *Proceedings of the ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition*, American Society of Mechanical Engineers Digital Collection, Oslo, Norway.
[22] Wang, J., Ye, L., Gao, R. X., Li, C., and Zhang, L., 2019, "Digital Twin for Rotating Machinery Fault Diagnosis in Smart Manufacturing," International Journal of Production Research, **57**(12), pp. 3920–3934.
[23] Kraft, J., and Kuntzagk, S., 2017, "Engine Fleet-Management: The Use of Digital Twins From a MRO Perspective," *Proceedings of the ASME Turbo Expo 2017: Turbomachinery Technical Conference and Exposition. Volume 1: Aircraft Engine; Fans and Blowers; Marine; Honors and Awards.*, American Society of Mechanical Engineers Digital Collection, Charlotte, North Carolina, USA.
[24] Zaccaria, V., Stenfelt, M., Aslanidou, I., and Kyprianidis, K. G., 2018, "Fleet Monitoring and Diagnostics Framework Based on Digital Twin of Aero-Engines," *Proceedings of the ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition. Volume 6: Ceramics; Controls, Diagnostics, and Instrumentation; Education; Manufacturing Materials and Metallurgy.*, American Society of Mechanical Engineers Digital Collection, Oslo, Norway.
[25] 2023, "Cordant | Baker Hughes," Baker Hughes Company [Online]. Available: https://www.bakerhughes.com/featured-capabilities/cordant. [Accessed: 04-May-2023].
[26] 2023, "Asset Performance Management Software | GE Digital," GE Digital [Online]. Available: https://www.ge.com/digital/applications/asset-performance-management. [Accessed: 04-May-2023].
[27] 2023, "Omnivise Digital Services Portfolio," Siemens Energy [Online]. Available: https://www.siemens-energy.com/global/en/offerings/services/digital-services.html. [Accessed: 04-May-2023].
[28] Kułaszka, A., Błachnio, J., and Borowczyk, H., 2023, "The Impact of Temperature on the Surface Colour of Gas Turbine Blades Heated in the Presence of Kerosene," Aerospace, **10**(4), p. 375.
[29] Giuntini, S., Andreini, A., Cappuccini, G., and Facchini, B., 2017, "Finite Element Transient Modelling for Whole Engine-Secondary Air System Thermomechanical Analysis," Energy Procedia, **126**, pp. 746–753.
[30] Squarcella, N., Firrone, C. M., Allara, M., and Gola, M., 2014, "The Importance of the Material Properties on the Burst Speed of Turbine Disks for Aeronautical Applications," International Journal of Mechanical Sciences, **84**, pp. 73–83.
[31] Shanian, A., Milani, A. S., Vermaak, N., Bertoldi, K., Scarinci, T., and Gerendas, M., 2012, "A Combined Finite Element-Multiple Criteria Optimization Approach for Materials Selection of Gas Turbine Components," Journal of Applied Mechanics, **79**(6).
[32] Tinga, T., van Kampen, J. F., de Jager, B., and Kok, J. B. W., 2006, "Gas Turbine Combustor Liner Life Assessment Using a Combined Fluid/Structural Approach," Journal of Engineering for Gas Turbines and Power, **129**(1), pp. 69–79.
[33] Marandi, S. M., Rahmani, Kh., and Tajdari, M., 2014, "Foreign Object Damage on the Leading Edge of Gas Turbine Blades," Aerospace Science and Technology, **33**(1), pp. 65–75.

[34] Metropolis, N., and Ulam, S., 1949, "The Monte Carlo Method," Journal of the American Statistical Association, **44**(247), pp. 335–341.
[35] Weiss, T., Voigt, M., Schlums, H., Mücke, R., Becker, K.-H., and Vogeler, K., 2009, "Probabilistic Finite-Element Analyses on Turbine Blades," *Proceedings of the ASME Turbo Expo 2009: Power for Land, Sea, and Air. Volume 6: Structures and Dynamics, Parts A and B*, Orlando, Florida, USA, pp. 1093–1102.
[36] Vo, D.-T., Mai, T.-D., Kim, B., Jung, J.-S., and Ryu, J., 2023, "Numerical Investigation of Crack Initiation in High-Pressure Gas Turbine Blade Subjected to Thermal-Fluid-Mechanical Low-Cycle Fatigue," International Journal of Heat and Mass Transfer, **202**, p. 123748.
[37] Almroth, P., and Gustafsson, D., 2017, *On Thermomechanical Fatigue Crack Growth Analysis in Gas Turbine Blading in a3D Finite Element Context*.
[38] Skamniotis, C., Grilli, N., and Cocks, A. C. F., 2023, "Crystal Plasticity Analysis of Fatigue-Creep Behavior at Cooling Holes in Single Crystal Nickel Based Gas Turbine Blade Components," International Journal of Plasticity, p. 103589.
[39] Riccius, J. R., and Zametaev, E. B., 2023, "HCF and LCF Analysis of a Generic Full Admission Turbine Blade," Aerospace, **10**(2), p. 154.
[40] Shlyannikov, V. N., and Ishtyryakov, I. S., 2019, "Crack Growth Rate and Lifetime Prediction for Aviation Gas Turbine Engine Compressor Disk Based on Nonlinear Fracture Mechanics Parameters," Theoretical and Applied Fracture Mechanics, **103**, p. 102313.
[41] Cláudio, R. A., Branco, C. M., Gomes, E. C., Harrison, G. F., and Winstone, M. R., 2004, "Fatigue Life Prediction and Failure Analysis of a Gas Turbine Disc Using the Finite-Element Method," Fatigue & Fracture of Engineering Materials & Structures, **27**(9), pp. 849–860.
[42] Poursaeidi, E., and Bazvandi, H., 2016, "Effects of Emergency and Fired Shut down on Transient Thermal Fatigue Life of a Gas Turbine Casing," Applied Thermal Engineering, **100**, pp. 453–461.
[43] Li, B., Fan, X., Li, D., and Jiang, P., 2017, "Design of Thermal Barrier Coatings Thickness for Gas Turbine Blade Based on Finite Element Analysis," Mathematical Problems in Engineering, **2017**, p. e2147830.
[44] Wang, J.-X., Sun, H.-T., Gong, Q.-T., Li, F.-X., and Li, Z.-Z., 2023, "Deformation Prediction Theory of Thermal Barrier Coatings near Cooling Holes under Thermal Cycling," ACS Omega, **8**(14), pp. 13048–13058.
[45] Ragupathy, R., Mishra, R. K., and Misal, R. D., 2012, "Life Estimation of TBC on an Aero Gas Turbine Combustor: A Finite Element Approach," *Proceedings of the ASME 2011 Turbo Expo: Turbine Technical Conference and Exposition. Volume 5: Heat Transfer, Parts A and B*, American Society of Mechanical Engineers Digital Collection, Vancouver, British Columbia, Canada, pp. 2037–2043.
[46] Szolc, T., 2007, "Continuous One-Dimensional Elastic Macro-Elements as a Natural Alternative Crack Detection Tool to the Spectral FInite Element Method," Archives of Mechanics, **59**(6), pp. 365–391.
[47] Szolc, T., Tauzowski, P., Knabel, J., and Stocki, R., 2009, "Nonlinear and Parametric Coupled Vibrations of the Rotor-Shaft System as Fault Identification Symptom Using Stochastic Methods," Nonlinear Dyn, **57**(4), pp. 533–557.
[48] Szolc, T., Tauzowski, P., Stocki, R., and Knabel, J., 2009, "Damage Identification in Vibrating Rotor-Shaft Systems by Efficient Sampling Approach," Mechanical Systems and Signal Processing, **23**(5), pp. 1615–1633.
[49] Urban, L. A., 1973, "Gas Path Analysis Applied to Turbine Engine Condition Monitoring," Journal of Aircraft, **10**(7), pp. 400–406.

[50] Li, Y. G., 2002, "Performance-Analysis-Based Gas Turbine Diagnostics: A Review," Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, **216**(5), pp. 363–377.
[51] Marinai, L., Probert, D., and Singh, R., 2004, "Prospects for Aero Gas-Turbine Diagnostics: A Review," Applied Energy, **79**(1), pp. 109–126.
[52] Escher, P. C., 1995, "Pythia: An Object-Orientated Gas Path Analysis Computer Program for General Applications," Ph.D. thesis, Cranfield University.
[53] Kalman, R. E., 1960, "A New Approach to Linear Filtering and Prediction Problems," Journal of Basic Engineering, **82**(1), pp. 35–45.
[54] Auger, F., Hilairet, M., Guerrero, J. M., Monmasson, E., Orlowska-Kowalska, T., and Katsura, S., 2013, "Industrial Applications of the Kalman Filter: A Review," IEEE Transactions on Industrial Electronics, **60**(12), pp. 5458–5471.
[55] Julier, S. J., and Uhlmann, J. K., 1997, "New Extension of the Kalman Filter to Nonlinear Systems," *Proc. SPIE 3068, Signal Processing, Sensor Fusion, and Target Recognition VI*, SPIE, Orlando, FL, United States, pp. 182–193.
[56] Kan, M. S., Tan, A. C. C., and Mathew, J., 2015, "A Review on Prognostic Techniques for Non-Stationary and Non-Linear Rotating Systems," Mechanical Systems and Signal Processing, **62–63**, pp. 1–20.
[57] Urrea, C., and Agramonte, R., 2021, "Kalman Filter: Historical Overview and Review of Its Use in Robotics 60 Years after Its Creation," Journal of Sensors, **2021**, p. Article ID 9674015.
[58] Weibull, W., 1951, "A Statistical Distribution Function of Wide Applicability," Journal of Applied Mechanics, **18**(3), pp. 293–297.
[59] Abernethy, R. B., 2004, "An Overview Of Weibull Analysis," *The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life, Safety, Risk, Support Costs, Failures, and Forecasting Warranty Claims, Substantiation and Accelerated Testing, Usin Weibull, Log Normal, Crow-AMSAA, Probit, and Kaplan-Meier Models*, Robert B. Abernethy, North Palm Beach, Florida, pp. 13–24.
[60] Newby, M., 1994, "Perspective on Weibull Proportional-Hazards Models," IEEE Transactions on Reliability, **43**(2), pp. 217–223.
[61] Mishra, R. K., and Dileep, S., 2017, "A Novel Methodology to Estimate Life of Gas Turbine Components Under Multiaxial Variable Amplitude Loading," J Fail. Anal. and Preven., **17**(4), pp. 731–739.
[62] Li, Y.-F., Zhu, S.-P., Li, J., Peng, W., and Huang, H.-Z., 2015, "Uncertainty Analysis in Fatigue Life Prediction of Gas Turbine Blades Using Bayesian Inference," International Journal of Turbo & Jet-Engines, **32**(4), pp. 319–324.
[63] Yau, J., and Kuang, H., 2009, "An Empirical Model for Predicting Crack Growth Behavior of Gas Turbine Stage 1 Nozzles," *Proceedings of the ASME Turbo Expo 2002: Power for Land, Sea, and Air. Volume 4: Turbo Expo 2002, Parts A and B*, American Society of Mechanical Engineers Digital Collection, Amsterdam, The Netherlands, pp. 691–697.
[64] Zaretsky, E. V., Litt, J. S., Hendricks, R. C., and Soditus, S. M., 2012, "Determination of Turbine Blade Life from Engine Field Data," Journal of Propulsion and Power, **28**(6), pp. 1156–1167.
[65] Ahsan, S., Lemma, T. A., and Gebremariam, M. A., 2020, "Reliability Analysis of Gas Turbine Engine by Means of Bathtub-Shaped Failure Rate Distribution," Process Safety Progress, **39**(S1), p. e12115.
[66] Salzman, R. H., 2005, "Applying Weibull Methods in Gas Turbine Component Data Analysis," *Proceedings of the ASME 2005 Power Conference. ASME 2005 Power*

Conference, American Society of Mechanical Engineers Digital Collection, Chicago, Illinois, USA, pp. 373–381.
[67] Zohair, D. A., Hafaifa, A., Abdelhamid, I., and Abdellah, K., 2022, "Gas Turbine Reliability Estimation to Reduce the Risk of Failure Occurrence with a Comparative Study between the Two-Parameter Weibull Distribution and a New Modified Weibull Distribution," Diagnostyka, **23**(1), pp. 1–18.
[68] Djeddi, A. Z., Hafaifa, A., Kouzou, A., and Abudura, S., 2017, "Exploration of Reliability Algorithms Using Modified Weibull Distribution: Application on Gas Turbine," Int J Syst Assur Eng Manag, **8**(2), pp. 1885–1894.
[69] Lai, C. D., Xie, M., and Murthy, D. N. P., 2003, "A Modified Weibull Distribution," IEEE Transactions on Reliability, **52**(1), pp. 33–37.
[70] Bertolini, M., Mezzogori, D., Neroni, M., and Zammori, F., 2021, "Machine Learning for Industrial Applications: A Comprehensive Literature Review," Expert Systems with Applications, **175**, p. 114820.
[71] Murphy, K. P., 2012, *Machine Learning: A Probabilistic Perspective*, MIT Press.
[72] Zhang, M., and Parnell, A., 2023, "Review of Clustering Methods for Functional Data," ACM Trans. Knowl. Discov. Data, **17**(7), p. 91:1-91:34.
[73] Anowar, F., Sadaoui, S., and Selim, B., 2021, "Conceptual and Empirical Comparison of Dimensionality Reduction Algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE)," Computer Science Review, **40**, p. 100378.
[74] Siraskar, R., Kumar, S., Patil, S., Bongale, A., and Kotecha, K., 2023, "Reinforcement Learning for Predictive Maintenance: A Systematic Technical Review," Artif Intell Rev.
[75] van Engelen, J. E., and Hoos, H. H., 2020, "A Survey on Semi-Supervised Learning," Mach Learn, **109**(2), pp. 373–440.
[76] Vasilyev, B., Nikolaev, S., Raevskiy, M., Belov, S., and Uzhinsky, I., 2020, "Residual Life Prediction of Gas-Engine Turbine Blades Based on Damage Surrogate-Assisted Modeling," Applied Sciences, **10**(23), p. 8541.
[77] Chen, T., and Guestrin, C., 2016, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, NY, USA, pp. 785–794.
[78] Liu, H., Sun, J., Lei, S., and Ning, S., 2021, "In-Service Aircraft Engines Turbine Blades Life Prediction Based on Multi-Modal Operation and Maintenance Data," Propulsion and Power Research, **10**(4), pp. 360–373.
[79] Zhu, G., Wang, C., Zhao, W., Xie, Y., Guo, D., and Zhang, D., 2023, "Blade Crack Diagnosis Based on Blade Tip Timing and Convolution Neural Networks," Applied Sciences, **13**(2), p. 1102.
[80] Iannitelli, M., Allegorico, C., Garau, F., and Capanni, M., 2018, "A Hybrid Model for On-Line Detection of Gas Turbine Lean Blowout Events," PHM Society European Conference, **4**(1).
[81] Yan, W., and Yu, L., 2015, "On Accurate and Reliable Anomaly Detection for Gas Turbine Combustors: A Deep Learning Approach," *Proceedings of the Annual Conference of the PHM Society 2015*, Coronado, California, USA.
[82] Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., and Manzagol, P.-A., 2010, "Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion," J. Mach. Learn. Res., **11**, pp. 3371–3408.
[83] Huang, G.-B., Zhu, Q.-Y., and Siew, C.-K., 2006, "Extreme Learning Machine: Theory and Applications," Neurocomputing, **70**(1), pp. 489–501.
[84] Allegorico, C., and Mantini, V., 2014, "A Data-Driven Approach for on-Line Gas Turbine Combustion Monitoring Using Classification Models," PHM Society European Conference, **2**(1).

[85] Bai, M., Yang, X., Liu, J., Liu, J., and Yu, D., 2021, "Convolutional Neural Network-Based Deep Transfer Learning for Fault Detection of Gas Turbine Combustion Chambers," Applied Energy, **302**, p. 117509.
[86] Hartwell, A., Montana, F., Jacobs, W., Kadirkamanathan, V., Mills, A. R., and Clark, T., 2021, "In-Flight Novelty Detection with Convolutional Neural Networks."
[87] Thakkar, U., and Chaoui, H., 2022, "Remaining Useful Life Prediction of an Aircraft Turbofan Engine Using Deep Layer Recurrent Neural Networks," Actuators, **11**(3), p. 67.
[88] Li, Y., Chen, Y., Hu, Z., and Zhang, H., 2023, "Remaining Useful Life Prediction of Aero-Engine Enabled by Fusing Knowledge and Deep Learning Models," Reliability Engineering & System Safety, **229**, p. 108869.
[89] Molla Salilew, W., Ambri Abdul Karim, Z., and Alemu Lemma, T., 2022, "Investigation of Fault Detection and Isolation Accuracy of Different Machine Learning Techniques with Different Data Processing Methods for Gas Turbine," Alexandria Engineering Journal, **61**(12), pp. 12635–12651.
[90] Kiakojoori, S., and Khorasani, K., 2016, "Dynamic Neural Networks for Gas Turbine Engine Degradation Prediction, Health Monitoring and Prognosis," Neural Comput & Applic, **27**(8), pp. 2157–2192.
[91] Liu, Y., Ravichandran, R., Chen, K., and Patnaik, P., 2021, "Application of Machine Learning to Solid Particle Erosion of APS-TBC and EB-PVD TBC at Elevated Temperatures," Coatings, **11**(7), p. 845.
[92] Badora, M., Sepe, M., Bielecki, M., Graziano, A., and Szolc, T., 2021, "Predicting Length of Fatigue Cracks by Means of Machine Learning Algorithms in the Small-Data Regime," Eksploatacja i Niezawodnosc - Maintenance and Reliability, **23**, pp. 575–585.
[93] Pawełczyk, M., Fulara, S., Sepe, M., Luca, A., and Badora, M., 2020, "Industrial Gas Turbine Operating Parameters Monitoring and Data-Driven Prediction," Eksploatacja i Niezawodnosc - Maintenance and Reliability, **22**, pp. 391–399.
[94] Matuszczak, M., Żbikowski, M., and Teodorczyk, A., 2021, "Predictive Modelling of Turbofan Engine Components Condition Using Machine and Deep Learning Methods," Eksploatacja i Niezawodność – Maintenance and Reliability, **23**(2), pp. 359–370.
[95] D'Amato, J., and Patanian, J., 2016, "Method and System for Predicting Hydraulic Valve Degradation on a Gas Turbine," Annual Conference of the PHM Society, **8**(1).
[96] Raissi, M., Perdikaris, P., and Karniadakis, G. E., 2017, "Physics Informed Deep Learning (Part I): Data-Driven Solutions of Nonlinear Partial Differential Equations," arXiv:1711.10561 [cs, math, stat].
[97] Hornik, K., Stinchcombe, M., and White, H., 1989, "Multilayer Feedforward Networks Are Universal Approximators," Neural Networks, **2**(5), pp. 359–366.
[98] Raissi, M., Perdikaris, P., and Karniadakis, G. E., 2017, "Physics Informed Deep Learning (Part II): Data-Driven Discovery of Nonlinear Partial Differential Equations," arXiv:1711.10566 [cs, math, stat].
[99] Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F., 2022, "Scientific Machine Learning Through Physics–Informed Neural Networks: Where We Are and What's Next," J Sci Comput, **92**(3), p. 88.
[100] Haghighat, E., Raissi, M., Moure, A., Gomez, H., and Juanes, R., 2021, "A Physics-Informed Deep Learning Framework for Inversion and Surrogate Modeling in Solid Mechanics," Computer Methods in Applied Mechanics and Engineering, **379**, p. 113741.
[101] Yucesan, Y. A., and Viana, F. A. C., 2022, "A Hybrid Physics-Informed Neural Network for Main Bearing Fatigue Prognosis under Grease Quality Variation," Mechanical Systems and Signal Processing, **171**, p. 108875.

[102] Zhao, X., Gong, Z., Zhang, Y., Yao, W., and Chen, X., 2023, "Physics-Informed Convolutional Neural Networks for Temperature Field Prediction of Heat Source Layout without Labeled Data," Engineering Applications of Artificial Intelligence, **117**, p. 105516.
[103] Nascimento, R. G., Fricke, K., and Viana, F. A. C., 2020, "A Tutorial on Solving Ordinary Differential Equations Using Python and Hybrid Physics-Informed Neural Network," Engineering Applications of Artificial Intelligence, **96**, p. 103996.
[104] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Jia, Y., Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, 2015, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems."
[105] Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A., 2017, "Automatic Differentiation in PyTorch."
[106] Yu, J., Lu, L., Meng, X., and Karniadakis, G. E., 2022, "Gradient-Enhanced Physics-Informed Neural Networks for Forward and Inverse PDE Problems," Computer Methods in Applied Mechanics and Engineering, **393**, p. 114823.
[107] Viana, F. A. C., Nascimento, R. G., Dourado, A., and Yucesan, Y. A., 2021, "Estimating Model Inadequacy in Ordinary Differential Equations with Physics-Informed Neural Networks," Computers & Structures, **245**, p. 106458.
[108] Yuan, L., Ni, Y.-Q., Deng, X.-Y., and Hao, S., 2022, "A-PINN: Auxiliary Physics Informed Neural Networks for Forward and Inverse Problems of Nonlinear Integro-Differential Equations," Journal of Computational Physics, **462**, p. 111260.
[109] Yang, L., Zhang, D., and Karniadakis, G. E., 2020, "Physics-Informed Generative Adversarial Networks for Stochastic Differential Equations," SIAM J. Sci. Comput., **42**(1), pp. A292–A317.
[110] Badora, M., Bartosik, P., Graziano, A., and Szolc, T., 2023, "Using Physics-Informed Neural Networks with Small Datasets to Predict the Length of Gas Turbine Nozzle Cracks," Advanced Engineering Informatics, **58**, p. 102232.
[111] Chen, D., Li, Y., Liu, K., and Li, Y., 2023, "A Physics-Informed Neural Network Approach to Fatigue Life Prediction Using Small Quantity of Samples," International Journal of Fatigue, **166**, p. 107270.
[112] Zhou, T., Jiang, S., Han, T., Zhu, S.-P., and Cai, Y., 2023, "A Physically Consistent Framework for Fatigue Life Prediction Using Probabilistic Physics-Informed Neural Network," International Journal of Fatigue, **166**, p. 107234.
[113] Zhang, X.-C., Gong, J.-G., and Xuan, F.-Z., 2021, "A Physics-Informed Neural Network for Creep-Fatigue Life Prediction of Components at Elevated Temperatures," Engineering Fracture Mechanics, **258**, p. 108130.
[114] Goswami, S., Anitescu, C., Chakraborty, S., and Rabczuk, T., 2020, "Transfer Learning Enhanced Physics Informed Neural Network for Phase-Field Modeling of Fracture," Theoretical and Applied Fracture Mechanics, **106**, p. 102447.
[115] Nath, K., Meng, X., Smith, D. J., and Karniadakis, G. E., 2023, "Physics-Informed Neural Networks for Predicting Gas Flow Dynamics and Unknown Parameters in Diesel Engines."
[116] McClenny, L., and Braga-Neto, U., 2022, "Self-Adaptive Physics-Informed Neural Networks Using a Soft Attention Mechanism."

[117] Cai, S., Wang, Z., Wang, S., Perdikaris, P., and Karniadakis, G. E., 2021, "Physics-Informed Neural Networks for Heat Transfer Problems," Journal of Heat Transfer, **143**(060801).
[118] Hennigh, O., Narasimhan, S., Nabian, M. A., Subramaniam, A., Tangsali, K., Fang, Z., Rietmann, M., Byeon, W., and Choudhry, S., 2021, "NVIDIA SimNet^{TM}: An AI-Accelerated Multi-Physics Simulation Framework," *Computational Science – ICCS 2021*, M. Paszynski, D. Kranzlmüller, V.V. Krzhizhanovskaya, J.J. Dongarra, and P.M.A. Sloot, eds., Springer International Publishing, Cham, pp. 447–461.
[119] 2019, "Modulus - A Neural Network Framework," NVIDIA Developer [Online]. Available: https://developer.nvidia.com/modulus. [Accessed: 22-May-2023].
[120] Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., and Ng, R., 2020, "Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains."
[121] Salvati, E., Tognan, A., Laurenti, L., Pelegatti, M., and De Bona, F., 2022, "A Defect-Based Physics-Informed Machine Learning Framework for Fatigue Finite Life Prediction in Additive Manufacturing," Materials & Design, **222**, p. 111089.
[122] Zhu, Q., Liu, Z., and Yan, J., 2021, "Machine Learning for Metal Additive Manufacturing: Predicting Temperature and Melt Pool Fluid Dynamics Using Physics-Informed Neural Networks," Comput Mech, **67**(2), pp. 619–635.
[123] Li, S., Wang, G., Di, Y., Wang, L., Wang, H., and Zhou, Q., 2023, "A Physics-Informed Neural Network Framework to Predict 3D Temperature Field without Labeled Data in Process of Laser Metal Deposition," Engineering Applications of Artificial Intelligence, **120**, p. 105908.
[124] Mao, Z., Jagtap, A., and Karniadakis, G., 2020, "Physics-Informed Neural Networks for High-Speed Flows," Computer Methods in Applied Mechanics and Engineering, **360**, p. 112789.
[125] Li, L., Li, Y., Du, Q., Liu, T., and Xie, Y., 2022, "ReF-Nets: Physics-Informed Neural Network for Reynolds Equation of Gas Bearing," Computer Methods in Applied Mechanics and Engineering, **391**, p. 114524.
[126] Krishnapriyan, A. S., Gholami, A., Zhe, S., Kirby, R. M., and Mahoney, M. W., 2021, "Characterizing Possible Failure Modes in Physics-Informed Neural Networks."
[127] Pan, S. J., and Yang, Q., 2010, "A Survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, **22**(10), pp. 1345–1359.
[128] Raina, R., Battle, A., Lee, H., Packer, B., and Ng, A. Y., 2007, "Self-Taught Learning: Transfer Learning from Unlabeled Data," *Proceedings of the 24th International Conference on Machine Learning*, Association for Computing Machinery, New York, NY, USA, pp. 759–766.
[129] Yao, S., Kang, Q., Zhou, M., Rawa, M. J., and Abusorrah, A., 2023, "A Survey of Transfer Learning for Machinery Diagnostics and Prognostics," Artif Intell Rev, **56**(4), pp. 2871–2922.
[130] Liu, S., Wang, H., Tang, J., and Zhang, X., 2022, "Research on Fault Diagnosis of Gas Turbine Rotor Based on Adversarial Discriminative Domain Adaption Transfer Learning," Measurement, **196**, p. 111174.
[131] Yang, X., Bai, M., Liu, J., Liu, J., and Yu, D., 2021, "Gas Path Fault Diagnosis for Gas Turbine Group Based on Deep Transfer Learning," Measurement, **181**, p. 109631.
[132] Mondal, S., Chattopadhyay, A., Mukhopadhyay, A., and Ray, A., 2021, "Transfer Learning of Deep Neural Networks for Predicting Thermoacoustic Instabilities in Combustion Systems," Energy and AI, **5**, p. 100085.

[133] Wang, J., Chen, Y., Hao, S., Feng, W., and Shen, Z., 2017, "Balanced Distribution Adaptation for Transfer Learning," *2017 IEEE International Conference on Data Mining (ICDM)*, New Orleans, LA, USA, pp. 1129–1134.
[134] Borgwardt, K. M., Gretton, A., Rasch, M. J., Kriegel, H.-P., Schölkopf, B., and Smola, A. J., 2006, "Integrating Structured Biological Data by Kernel Maximum Mean Discrepancy," Bioinformatics, **22**(14), pp. e49–e57.
[135] Fernando, B., Habrard, A., Sebban, M., and Tuytelaars, T., 2014, "Subspace Alignment For Domain Adaptation."
[136] Sun, B., and Saenko, K., 2015, "Subspace Distribution Alignment for Unsupervised Domain Adaptation," *Proceedings of the British Machine Vision Conference 2015*, British Machine Vision Association, Swansea, p. 24.1-24.10.
[137] Sun, B., Feng, J., and Saenko, K., 2017, "Correlation Alignment for Unsupervised Domain Adaptation," *Domain Adaptation in Computer Vision Applications*, G. Csurka, ed., Springer International Publishing, Cham, pp. 153–171.
[138] Olson, C. C., Judd, K. P., and Nichols, J. M., 2018, "Manifold Learning Techniques for Unsupervised Anomaly Detection," Expert Systems with Applications, **91**, pp. 374–385.
[139] Zhang, S., Su, L., Gu, J., Li, K., Zhou, L., and Pecht, M., 2023, "Rotating Machinery Fault Detection and Diagnosis Based on Deep Domain Adaptation: A Survey," Chinese Journal of Aeronautics, **36**(1), pp. 45–74.
[140] Zhang, W., Deng, L., Zhang, L., and Wu, D., 2023, "A Survey on Negative Transfer," IEEE/CAA Journal of Automatica Sinica, **10**(2), pp. 305–329.
[141] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., 2014, "Generative Adversarial Networks."
[142] Sabuhi, M., Zhou, M., Bezemer, C.-P., and Musilek, P., 2021, "Applications of Generative Adversarial Networks in Anomaly Detection: A Systematic Literature Review," IEEE Access, **9**, pp. 161003–161029.
[143] Bousmalis, K., Trigeorgis, G., Silberman, N., Krishnan, D., and Erhan, D., 2016, "Domain Separation Networks," *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Curran Associates Inc., Red Hook, NY, USA, pp. 343–351.
[144] Wang, J., Lan, C., Liu, C., Ouyang, Y., Qin, T., Lu, W., Chen, Y., Zeng, W., and Yu, P. S., 2022, "Generalizing to Unseen Domains: A Survey on Domain Generalization."
[145] Zhou, K., Liu, Z., Qiao, Y., Xiang, T., and Loy, C. C., 2022, "Domain Generalization: A Survey," IEEE Trans. Pattern Anal. Mach. Intell., pp. 1–20.
[146] Li, X., Zhang, W., Ma, H., Luo, Z., and Li, X., 2020, "Domain Generalization in Rotating Machinery Fault Diagnostics Using Deep Neural Networks," Neurocomputing, **403**, pp. 409–420.
[147] Ding, Y., Jia, M., Cao, Y., Ding, P., Zhao, X., and Lee, C.-G., 2023, "Domain Generalization via Adversarial Out-Domain Augmentation for Remaining Useful Life Prediction of Bearings under Unseen Conditions," Knowledge-Based Systems, **261**, p. 110199.
[148] Shi, Y., Deng, A., Deng, M., Xu, M., Liu, Y., Ding, X., and Bian, W., 2023, "Domain Augmentation Generalization Network for Real-Time Fault Diagnosis under Unseen Working Conditions," Reliability Engineering & System Safety, **235**, p. 109188.
[149] Amozegar, M., and Khorasani, K., 2016, "An Ensemble of Dynamic Neural Network Identifiers for Fault Detection and Isolation of Gas Turbine Engines," Neural Networks, **76**, pp. 106–121.
[150] Xiao, Y., Wang, Q., Wang, S., and Chen, W., 2022, "Research on a Multisource Domain Improved Fault Diagnosis Method of the Rotor System," IEEE Access, **10**, pp. 85399–85415.

[151] Feng, Y., Chen, J., Xie, J., Zhang, T., Lv, H., and Pan, T., 2022, "Meta-Learning as a Promising Approach for Few-Shot Cross-Domain Fault Diagnosis: Algorithms, Applications, and Prospects," Knowledge-Based Systems, **235**, p. 107646.
[152] Zhuang, J., Jia, M., Ding, Y., and Zhao, X., 2022, "Health Assessment of Rotating Equipment With Unseen Conditions Using Adversarial Domain Generalization Toward Self-Supervised Regularization Learning," IEEE/ASME Transactions on Mechatronics, **27**(6), pp. 4675–4685.
[153] Wang, Q., Yang, L., and Rao, Y., 2021, "Establishment of a Generalizable Model on a Small-Scale Dataset to Predict the Surface Pressure Distribution of Gas Turbine Blades," Energy, **214**, p. 118878.
[154] Farahani, H. S., Fatehi, A., Nadali, A., and Shoorehdeli, M. A., 2021, "Domain Adversarial Neural Network Regression to Design Transferable Soft Sensor in a Power Plant," Computers in Industry, **132**, p. 103489.
[155] Zhou, H., Lei, Z., Zio, E., Wen, G., Liu, Z., Su, Y., and Chen, X., 2023, "Conditional Feature Disentanglement Learning for Anomaly Detection in Machines Operating under Time-Varying Conditions," Mechanical Systems and Signal Processing, **191**, p. 110139.
[156] Ma, N., Liu, F., Wang, H., Zhou, S., Bu, J., and Han, B., 2023, "Domain Generalization in Regression."
[157] 2016, "Breaking the Power Plant Efficiency Record | GE Gas Power," Breaking the Power Plant Efficiency Record | GE Gas Power [Online]. Available: https://www.ge.com/gas-power/resources/articles/2016/power-plant-efficiency-record. [Accessed: 08-Jun-2023].
[158] 2023, "NovaLT16 | Baker Hughes," NovaLT16 | Baker Hughes [Online]. Available: https://www.bakerhughes.com/gas-turbines/novalt-technology/novalt16. [Accessed: 08-Jun-2023].
[159] Shanmugam, M., 2015, *Final Report - Repair Job Number 2901323*, FOR2901323, Baker Hughes Company, Qatar.
[160] Mirhosseini, A. M., Adib Nazari, S., Maghsoud Pour, A., Etemadi Haghighi, S., and Zareh, M., 2020, "Failure Analysis of First Stage Nozzle in a Heavy-Duty Gas Turbine," Engineering Failure Analysis, **109**, p. 104303.
[161] 2023, "Travel Weather Averages (Weatherbase)," Travel Weather Averages (Weatherbase) [Online]. Available: https://www.weatherbase.com/. [Accessed: 11-Mar-2021].
[162] James, G., Witten, D., Hastie, T., and Tibshirani, R., 2013, "Linear Model Selection and Regularization," *An Introduction to Statistical Learning: With Applications in R*, G. James, D. Witten, T. Hastie, and R. Tibshirani, eds., Springer, New York, NY, pp. 203–264.
[163] Karagiannopoulos, M., Anyfantis, D., Kotsiantis, S. B., and Pintelas, P. E., 2007, "Feature Selection for Regression Problems," *Proceedings of HERCMA 2007*, Athens.
[164] Green, S. B., 1991, "How Many Subjects Does It Take To Do A Regression Analysis," Multivariate Behavioral Research, **26**(3), pp. 499–510.
[165] Austin, P. C., and Steyerberg, E. W., 2015, "The Number of Subjects per Variable Required in Linear Regression Analyses," J Clin Epidemiol, **68**(6), pp. 627–636.
[166] Cover, T. M., and Thomas, J. A., 2006, "Introduction and Preview," *Elements of Information Theory*, John Wiley & Sons, Inc., Hoboken, New Jersey, pp. 1–12.
[167] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., and Cournapeau, D., 2011, "Scikit-Learn: Machine Learning in Python," Journal of Machine Learning Research, **12**, pp. 2825–2860.
[168] Breiman, L., 2001, "Random Forests," Machine Learning, **45**(1), pp. 5–32.

[169] Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A., 1984, *Classification and Regression Trees*, Taylor & Francis, New York.
[170] James, G., Witten, D., Hastie, T., and Tibshirani, R., 2013, "Statistical Learning," *An Introduction to Statistical Learning: With Applications in R*, G. James, D. Witten, T. Hastie, and R. Tibshirani, eds., Springer, New York, NY, pp. 15–57.
[171] Cortes, C., and Vapnik, V., 1995, "Support-Vector Networks," Mach Learn, **20**(3), pp. 273–297.
[172] Awad, M., and Khanna, R., 2015, "Support Vector Regression," *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, M. Awad, and R. Khanna, eds., Apress, Berkeley, CA, pp. 67–80.
[173] Smola, A. J., and Schölkopf, B., 2004, "A Tutorial on Support Vector Regression," Statistics and Computing, **14**(3), pp. 199–222.
[174] Drucker, H., 1997, "Improving Regressors Using Boosting Techniques," Proceedings of the 14th International Conference on Machine Learning, pp. 107–115.
[175] Freund, Y., and Schapire, R. E., 1997, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," Journal of Computer and System Sciences, **55**(1), pp. 119–139.
[176] Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E., 2020, "Array Programming with NumPy," Nature, **585**(7825), pp. 357–362.
[177] McKinney, W., 2010, "Data Structures for Statistical Computing in Python," *Proceedings of the 9th Python in Science Conference*, Austin, Texas, pp. 56–61.
[178] Chollet, F., 2015, "Keras" [Online]. Available: https://keras.io/. [Accessed: 18-Jun-2023].
[179] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. 1 0, 2020, "SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python," Nat Methods, **17**(3), pp. 261–272.
[180] Storn, R., and Price, K., 1997, "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," Journal of Global Optimization, **11**, pp. 341–359.
[181] Kingma, D. P., and Ba, J., 2015, "Adam: A Method for Stochastic Optimization.," *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, San Diego, CA, USA.
[182] Dourado, A., and Viana, F. A. C., 2020, "Physics-Informed Neural Networks for Missing Physics Estimation in Cumulative Damage Models: A Case Study in Corrosion Fatigue," J. Comput. Inf. Sci. Eng, **20**(6).
[183] Paris, P., and Erdogan, F., 1963, "A Critical Analysis of Crack Propagation Laws," Journal of Basic Engineering, **85**(4), pp. 528–533.
[184] The British Standards Institution, 2015, "S.1.5 Creep Crack Propagation Rate," *BS 7910:2013+A1:2015: Guide to Methods for Assessing the Acceptability of Flaws in Metallic Structures*, London, pp. 414–416.
[185] Irwin, G. R., 1957, "Analysis of Stresses and Strains Near the End of a Crack Traversing a Plate," J. Appl. Mech., (24(3)), pp. 361–364.

[186] Sih, G. C., 1974, "Strain-Energy-Density Factor Applied to Mixed Mode Crack Problems," Int J Fract, **10**(3), pp. 305–321.
[187] Burdekin, F. M., and Stone, D. E. W., 1966, "The Crack Opening Displacement Approach to Fracture Mechanics in Yielding Materials," Journal of Strain Analysis, **1**(2), pp. 145–153.
[188] Seweryn, A., and Mróz, Z., 1995, "A Non-Local Stress Failure Condition for Structural Elements under Multiaxial Loading," Engineering Fracture Mechanics, **51**(6), pp. 955–973.
[189] Seweryn, A., and Mróz, Z., 1996, "A Non-Local Stress Failure and Fatigue Damage Accumulation Condition," *Multiaxial Fatigue and Design*, Mechanical Engineering Publications, London, pp. 261–282.
[190] Fuentes, J. D., Cicero, S., Berto, F., Torabi, A. R., Madrazo, V., and Azizi, P., 2018, "Estimation of Fracture Loads in AL7075-T651 Notched Specimens Using the Equivalent Material Concept Combined with the Strain Energy Density Criterion and with the Theory of Critical Distances," Metals, **8**(2), p. 87.
[191] Taylor, D., 2008, "The Theory of Critical Distances," Engineering Fracture Mechanics, **75**(7), pp. 1696–1705.
[192] Torabi, A. R., 2012, "Estimation of Tensile Load-Bearing Capacity of Ductile Metallic Materials Weakened by a V-Notch: The Equivalent Material Concept," Materials Science and Engineering: A, **536**, pp. 249–255.
[193] Leung, A. Y. T., Zhou, Z., and Xu, X., 2014, "Determination of Stress Intensity Factors by the Finite Element Discretized Symplectic Method," International Journal of Solids and Structures, **51**(5), pp. 1115–1122.
[194] Keprate, A., Chandima Ratnayake, R. M., and Sankararaman, S., 2017, "Comparison of Various Surrogate Models to Predict Stress Intensity Factor of a Crack Propagating in Offshore Piping," Journal of Offshore Mechanics and Arctic Engineering, **139**(061401).
[195] Yuvaraj, P., Ramachandra Murthy, A., Iyer, N. R., Sekar, S. K., and Samui, P., 2013, "Support Vector Regression Based Models to Predict Fracture Characteristics of High Strength and Ultra High Strength Concrete Beams," Engineering Fracture Mechanics, **98**, pp. 29–43.
[196] Fakhri, D., Khodayari, A., Mahmoodzadeh, A., Hosseini, M., Hashim Ibrahim, H., and Hussein Mohammed, A., 2022, "Prediction of Mixed-Mode I and II Effective Fracture Toughness of Several Types of Concrete Using the Extreme Gradient Boosting Method and Metaheuristic Optimization Algorithms," Engineering Fracture Mechanics, **276**, p. 108916.
[197] Rubio, P., Muñoz-Abella, B., and Rubio, L., 2018, "Neural Approach to Estimate the Stress Intensity Factor of Semi-Elliptical Cracks in Rotating Cracked Shafts in Bending," Fatigue & Fracture of Engineering Materials & Structures, **41**(3), pp. 539–550.
[198] Long, X. Y., Zhao, S. K., Jiang, C., Li, W. P., and Liu, C. H., 2021, "Deep Learning-Based Planar Crack Damage Evaluation Using Convolutional Neural Networks," Engineering Fracture Mechanics, **246**, p. 107604.
[199] Neuber, H., 1961, "Theory of Stress Concentration for Shear-Strained Prismatical Bodies With Arbitrary Nonlinear Stress-Strain Law," Journal of Applied Mechanics, **28**(4), pp. 544–550.
[200] Banaszkiewicz, M., 2016, "On-Line Monitoring and Control of Thermal Stresses in Steam Turbine Rotors," Applied Thermal Engineering, **94**, pp. 763–776.
[201] Banaszkiewicz, M., 2018, "The Low-Cycle Fatigue Life Assessment Method for Online Monitoring of Steam Turbine Rotors," International Journal of Fatigue, **113**, pp. 311–323.
[202] Zhao, X., Ru, D., Wang, P., Gan, L., Wu, H., and Zhong, Z., 2021, "Fatigue Life Prediction of a Supercritical Steam Turbine Rotor Based on Neural Networks," Engineering Failure Analysis, **127**, p. 105435.

[203] Shin, C. S., Man, K. C., and Wang, C. M., 1994, "A Practical Method to Estimate the Stress Concentration of Notches," International Journal of Fatigue, **16**(4), pp. 242–256.
[204] Brown, W. F., and Srawley, J. E., 1966, *Plane Strain Crack Toughness Testing of High Strength Metallic Materials*, 410, American Society for Testing Materials, Philadelphia.
[205] Tada, H., Paris, P. C., and Irwin, G. R., 2000, *The Stress Analysis of Cracks Handbook, Third Edition*, ASME Press.
[206] Gross, B., Srawley, J., and Brown, W., 1964, *Stress-Intensity Factors for a Single-Edge-Notch Tension Specimen by Boundary Collocation of a Stress Function*, NASA TN D-2395, National Aeronautics and Space Administration, Cleveland, Ohio.
[207] Hunter, J. D., 2007, "Matplotlib: A 2D Graphics Environment," Computing in Science & Engineering, **9**(3), pp. 90–95.
[208] Abernethy, R. B., 2004, "Weibayes And Weibayes Substantiation Testing," *The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life, Safety, Risk, Support Costs, Failures, and Forecasting Warranty Claims, Substantiation and Accelerated Testing, Usin Weibull, Log Normal, Crow-AMSAA, Probit, and Kaplan-Meier Models*, Robert B. Abernethy, North Palm Beach, Florida, pp. 115–146.
[209] Abernethy, R. B., 2004, "Maximum Likelihood Estimates & Other Alternatives," *The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life, Safety, Risk, Support Costs, Failures, and Forecasting Warranty Claims, Substantiation and Accelerated Testing, Usin Weibull, Log Normal, Crow-AMSAA, Probit, and Kaplan-Meier Models*, Robert B. Abernethy, North Palm Beach, Florida, pp. 97–114.
[210] Pisegna, S., Annunziata, F., Graziano, A., and Frediani, S., 2013, *Design Record Book Article Number OG-5110.A*, OG-5110.A, Baker Hughes Company, Florence.
[211] Mugnai, D., 2014, *Final Report - Repair Job Number 1733439*, FOR4992, Baker Hughes Company, Florence.
[212] Sutter, T., 2018, *Condition Report - Repair Job Number 12227*, 12227, Baker Hughes Company, Houston.
[213] Giunti, D., Giorgetti, A., and Pisegna, S., 2022, "Repair Engineering Instruction IC-M-TP-1003: Transition Piece (Disassembly and Inspection)."
[214] Wagner, C., 1933, "Beitrag zur Theorie des Anlaufvorgangs," Zeitschrift für Physikalische Chemie, **21B**(1), pp. 25–41.
[215] Sill, J., 1997, "Monotonic Networks," *Advances in Neural Information Processing Systems*, MIT Press.