INSTITUTO TECNOLÓGICO DE AERONÁUTICA

Brian Clark Zanfelice

METHODOLOGY FOR DEVELOPING A LEVEL 1 DIGITAL TWIN WITH AI INTEGRATION FOR GAS TURBINES

Final Paper

2025

Course of Mechanical Aeronautics Engineering

"No que diz respeito ao empenho, ao compromisso, ao esforço, à dedicação, não existe meio-termo. Ou você faz uma coisa bem-feita ou não faz."

AYRTON SENNA

Resumo

As turbinas a gás são cruciais na geração de energia e na aviação, tornando sua modelagem precisa essencial para a otimização do desempenho e manutenção preditiva. Abordagens de modelagem tradicionais podem ser computacionalmente intensivas ou, se puramente baseadas em dados, podem carecer de consistência física. Esta tese apresenta o desenvolvimento de um gêmeo digital de primeiro nível para uma turbina a gás em escala de laboratório utilizando Redes Neurais Informadas pela Física (PINNs). Este trabalho envolve o aproveitamento de dados experimentais coletados da microturbina de laboratório do Instituto Tecnológico de Aeronáutica (ITA), incluindo parâmetros operacionais chave como temperaturas e pressões em vários estágios, fluxo de combustível, velocidade do compressor e empuxo. Estes dados empíricos são integrados com princípios termodinâmicos fundamentais e as equações algébricas governantes que descrevem o comportamento dos principais componentes da turbina a gás: o compressor, a câmara de combustão, a turbina e o bocal. Ao incorporar estas leis físicas diretamente na função de perda da rede neural, a PINN é treinada não apenas para se ajustar aos dados experimentais observados, mas também para aderir à física inerente do sistema. O gêmeo digital desenvolvido visa prever com precisão as principais saídas de desempenho, como a temperatura de saída da câmara de combustão, a temperatura dos gases de escape e o empuxo, sob diversas condições operacionais definidas pelo fluxo de combustível e pela velocidade do compressor. Além disso, o modelo é projetado para prever temperaturas e pressões intermediárias ao longo da turbina, garantindo consistência termodinâmica em todo o processo simulado. Esta pesquisa contribui com uma metodologia para a criação de um gêmeo digital robusto, computacionalmente eficiente e fisicamente fundamentado. Espera-se que o modelo resultante sirva como uma ferramenta valiosa para análise de desempenho aprofundada, compreensão das sensibilidades do sistema e, potencialmente, para o desenvolvimento futuro de estratégias de controle e otimização para a turbina a gás.

Palavras-chave: Turbina a Gás. Gêmeo Digital. Inteligência Artificial. Aprendizado de Máquina Informado pela Física. First-Principle Modeling.

Abstract

Gas turbines are crucial in power generation and aviation, making their accurate modeling essential for performance optimization and predictive maintenance. Traditional modeling approaches can be computationally intensive or, if purely data-driven, may lack physical consistency. This thesis presents the development of a Level 1 (L1) digital twin for a laboratory-scale gas turbine using Physics-Informed Neural Networks (PINNs). This work involves leveraging experimental data collected from the Aeronautical Institute of Technology (ITA) laboratory microturbine, including key operational parameters such as temperatures and pressures at various stages, fuel flow, compressor speed, and thrust. These empirical data are integrated with fundamental thermodynamic principles and the governing algebraic equations describing the behavior of the main gas turbine components: the compressor, combustion chamber, turbine, and nozzle. By incorporating these physical laws directly into the neural network's loss function, the PINN is trained not only to fit the observed experimental data but also to adhere to the inherent physics of the system. The developed digital twin aims to accurately predict key performance outputs, such as combustor outlet temperature, exhaust gas temperature, and thrust, under various operating conditions defined by fuel flow and compressor speed. Furthermore, the model is designed to predict intermediate temperatures and pressures throughout the turbine, ensuring thermodynamic consistency across the simulated process. This research contributes a methodology for creating a robust, computationally efficient, and physically grounded digital twin. The resulting model is expected to serve as a valuable tool for in-depth performance analysis, understanding system sensitivities, and potentially for the future development of control and optimization strategies for the gas turbine.

Keywords: Gas Turbine. Digital Twin. Artificial Intelligence. Physics-Informed Machine Learning. First-Principle Modeling.

TABLE OF CONTENTS

1 Introduction

1.1 Motivation

Gas turbines are fundamental components in a wide range of critical applications, from power generation to aircraft propulsion. The ability to accurately model their behavior is essential for optimizing performance, ensuring operational safety, and implementing effective predictive maintenance strategies. However, traditional modeling techniques present a significant trade-off. High-fidelity simulations based on first principles (e.g., computational fluid dynamics) are computationally expensive and time-consuming, making them impractical for real-time analysis and control. On the other hand, purely data-driven models, while computationally efficient, often lack physical consistency and may produce unreliable predictions when extrapolating beyond the training data. This gap highlights the need for a new generation of models that can bridge the gap between physical fidelity and computational efficiency.

1.2 Objectives

The main objective of this work is to develop and validate a methodology for creating a Level 1 (L1) first-principle digital twin of a laboratory-scale gas turbine, integrating artificial intelligence techniques. This involves several specific aims. First, a mathematical model will be created based on the fundamental thermodynamic principles governing the main components: the compressor, combustion chamber, turbine, and nozzle. Concurrently, experimental data from a laboratory microturbine—including key parameters like temperatures, pressures, fuel flow, compressor speed, and thrust—will be utilized to inform and validate this model. A hybrid model will then be implemented using a Physics-Informed Neural Network (PINN), which integrates the first-principle equations directly into the network's loss function to ensure predictions are both accurate and physically consistent. The resulting trained PINN will serve as the digital twin, capable of predicting key performance outputs and intermediate thermodynamic states. Finally, the digital twin's accuracy and robustness will be rigorously assessed by comparing its predictions against experimental data not used during the training phase.

2 Literature Review

2.1 Gas Turbines

2.1.1 Gas Turbine: Overview

Gas turbines stand as fundamental components within a broad spectrum of critical applications, ranging from large-scale power generation to the propulsion of aircraft. Their widespread use underscores the imperative for precise modeling of their operational behavior (Boyce, 2012). Such accurate modeling is not merely a theoretical exercise but is essential for achieving optimal performance, ensuring operational safety protocols, and successfully implementing proactive predictive maintenance strategies (Boyce, 2012). These complex machines operate on fundamental thermodynamic principles, involving several main components that work in tandem to convert fuel energy into useful work (Cengel; Boles, 2019).

These core components typically include a compressor, which draws in and pressurizes air; a combustion chamber, where fuel is mixed with the compressed air and ignited; a turbine, which extracts energy from the hot, high-pressure combustion gases; and a nozzle, which accelerates the exhaust gases to produce thrust or direct them for other purposes (Saravanamuttoo et al., 2017). Understanding the intricate interplay between these components and their adherence to fundamental physical laws is crucial for their effective design, analysis, and operation (Saravanamuttoo et al., 2017). The fundamental principles governing the design and operation of gas turbine components, including detailed thermodynamic cycles and performance characteristics, are extensively documented in specialized literature (Saravanamuttoo et al., 2017).

2.1.2 Traditional Modeling Approaches and Their Limitations

Traditional modeling approaches for gas turbines present a significant trade-off in terms of computational resources and physical consistency. High-fidelity simulations, often based on first principles such as computational fluid dynamics (CFD), are known to be computationally intensive and time-consuming (Verstraete; Van den Braembussche, 2010). This characteristic makes them impractical for real-time analysis and control applications (Verstraete; Van den Braembussche, 2010). Conversely, purely data-driven models offer computational efficiency but frequently lack physical consistency (Gurney, 2010). Such models may also yield unreliable predictions when extrapolated beyond the dataset they were trained on (Gurney, 2010). This inherent limitation in traditional methods highlights a crucial gap, emphasizing the need for a new generation of models capable of bridging the divide between physical fidelity and computational efficiency (Kurz; Brun, 2009).

2.2 Foundations for Advanced System Representation

2.2.1 Digital Twins Technology

Digital Twin Technology represents a paradigm shift in system modeling and management, moving beyond traditional simulation to create a dynamic, virtual replica of a physical system or process (Grieves, 2011). This virtual counterpart is continuously updated with real-time data from its physical twin, allowing for high-fidelity mirroring of the physical entity's state, behavior, and performance throughout its lifecycle (Tao et al., 2019). The core concept involves the seamless integration of physical and virtual worlds, enabling predictive analytics, proactive maintenance, and optimization strategies that were previously unattainable. For complex systems like gas turbines, a digital twin provides an invaluable tool for monitoring operational parameters, diagnosing anomalies, and even predicting future performance degradation. This capability allows operators to make informed decisions, optimize efficiency, and extend the lifespan of costly assets, by running simulations and analyses on the virtual model that directly reflect the real-world conditions (Schluse et al., 2018). The utility of digital twins extends across various stages, from design and manufacturing to operation and decommissioning, offering a comprehensive and integrated approach to system representation and control.

2.2.2 Artificial Intelligence

Artificial Intelligence (AI) represents a broad field of computer science dedicated to creating intelligent agents capable of performing tasks that typically require human intelligence. These tasks include learning, problem-solving, perception, and decision-making (Russell; Norvig, 2010). In the context of advanced system representation, AI plays an important role, particularly through its subfields such as machine learning and neural networks. Machine learning algorithms enable systems to learn patterns and make predictions from data without being explicitly programmed, which is crucial for handling complex, non-linear relationships often found in engineering systems. The integration of AI capabilities allows for enhanced predictive power, adaptive behavior, and the ability to extract insights from vast amounts of data, significantly contributing to the development of sophisticated models like digital twins (Kreuzer; Papapetrou; Zdravkovic, 2024).

Among the various machine learning techniques, Neural Networks (NNs) are particularly prominent due to their ability to model complex, non-linear relationships and learn from large datasets. Inspired by the structure and function of the human brain, NNs consist of interconnected layers of nodes (neurons) that process information through weighted connections (Goodfellow; Bengio; Courville, 2016). Their capacity for pattern recognition and function approximation makes them highly effective for tasks such as prediction, classification, and control in engineering applications. The adaptability of NNs allows them to capture intricate dynamics within the system, making them a powerful tool for building data-driven components of hybrid models and digital twins.

2.3 Hybrid Modeling with Physics-Informed Machine Learning

2.3.1 The Role of First-Principle Models

First-principle models, grounded in fundamental physical laws such as thermodynamics, fluid dynamics, and mechanics, serve as the bedrock for understanding and predicting the behavior of complex engineering systems like gas turbines (Incropera et al., 2007). These models provide inherent physical consistency and interpretability, as their predictions are directly derived from established scientific principles rather than solely from observed data. They offer a strong foundation for analysis, enabling accurate predictions even outside the range of typical operating conditions, which is a significant advantage over purely empirical approaches. Furthermore, first-principle models can capture the underlying mechanisms driving system behavior, offering deep insights into component interactions and overall performance (Serway; Jewett, 2018). However, their development often involves intricate mathematical formulations and can be computationally expensive, particularly for high-fidelity simulations of complex geometries or transient phenomena. This computational burden can limit their utility for real-time applications or scenarios requiring rapid iteration.

2.3.2 Integrating Physics and Data with PINNs

Physics-Informed Neural Networks (PINNs) emerge as a powerful solution to overcome the limitations of both purely data-driven and purely physics-based models by integrating physical laws directly into the machine learning framework (Raissi; Perdikaris; Karniadakis, 2019). This hybrid approach leverages the universal approximation capabilities of neural networks to learn from experimental data, while simultaneously enforcing adherence to the governing physical equations of the system. In PINNs, the neural network is trained not only to minimize the error between its predictions and observed data points but also to satisfy the underlying partial differential equations (PDEs), ordinary differential equations (ODEs), or algebraic equations that describe the system's physics. This is achieved by incorporating the residuals of these physics equations into the network's loss function. The result is a model that is both data-driven and physically consistent, leading to enhanced predictive accuracy, improved generalization to unseen conditions, and the ability to handle sparse or noisy data more effectively (Karniadakis et al., 2021).

For gas turbines, PINNs offer a promising avenue for creating robust digital twins that can accurately predict performance, diagnose faults, and optimize operations while respecting fundamental thermodynamic and fluid dynamic principles. A study by Wang; Lu; Zhang (2023) demonstrates how PINNs can be effectively employed to model steam turbine performance, enabling more accurate predictions and condition monitoring essential for digital twin functionalities. This approach addresses the computational intensity of traditional physics-based models while ensuring the physical consistency often lacking in purely data-driven methods, leading to reliable insights for predictive maintenance and operational optimization.

3 Materials and Methods

This chapter details the methodology for developing a Level 1 digital twin for a laboratory-scale gas turbine. The approach integrates first-principle thermodynamic models with a data-driven framework using Physics-Informed Neural Networks (PINNs). The primary objective is to create a robust and physically consistent model capable of predicting the turbine's performance under various operating conditions. The methodology encompasses several key stages: establishing the mathematical model from fundamental thermodynamic laws, leveraging experimental data for training and validation, implementing the hybrid PINN architecture, and rigorously evaluating the final digital twin's predictive accuracy and physical fidelity. This structured approach ensures that the resulting digital twin is not only accurate but also grounded in the underlying physics of the gas turbine system.

3.1 Experimental Setup and Data Acquisition

The physical asset central to this investigation is the Mini-Lab Gas Turbine Power System, which is based on the SR-30 turbojet engine. This system is purpose-built for educational and experimental applications, allowing for the simulation and detailed analysis of core gas turbine operations and thermodynamic cycles on a laboratory scale.

3.1.1 System Overview and Instrumentation

The SR-30 engine is configured as a single-shaft turbojet, comprising a centrifugal compressor, an annular combustion chamber, an axial turbine, and a nozzle. To facilitate comprehensive operational monitoring and data acquisition, the engine is extensively instrumented with a variety of sensors. These sensors provide high-resolution, real-time data on critical thermodynamic and performance parameters.

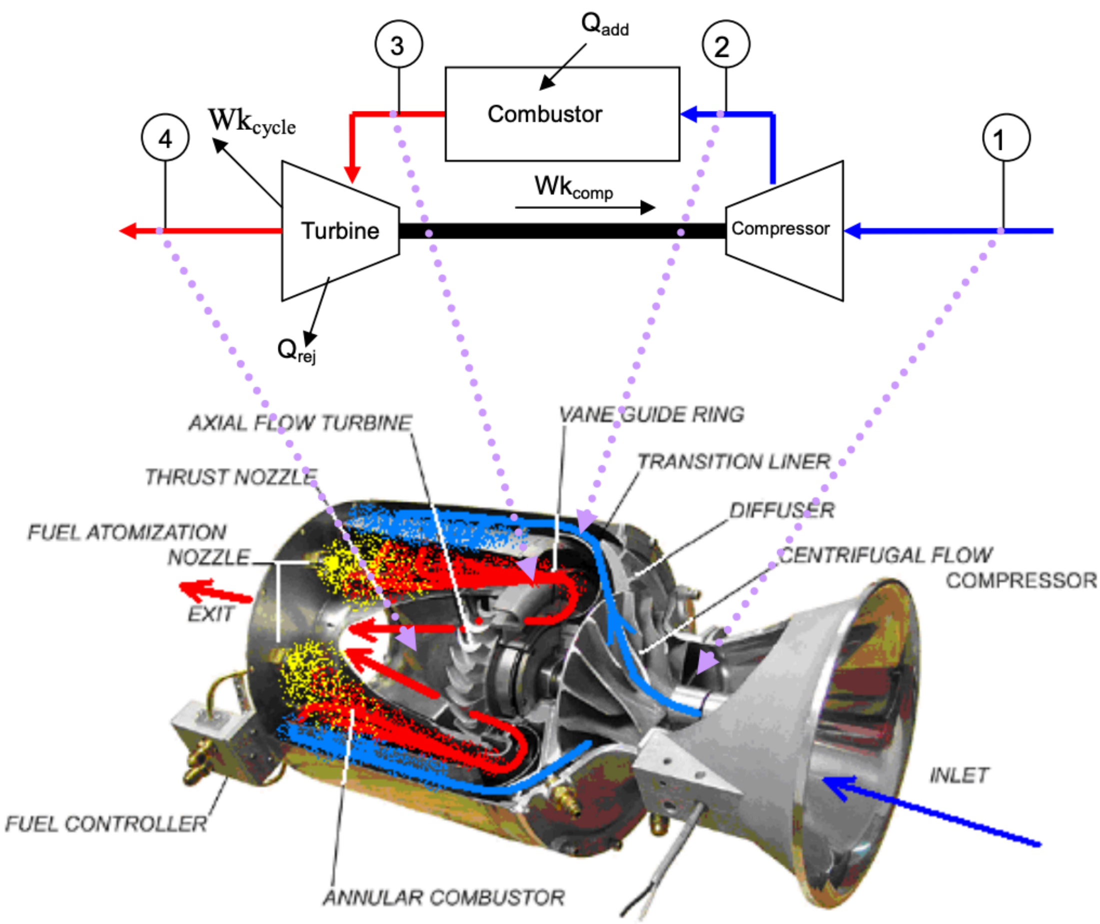
The instrumentation setup includes:

- Compressor Inlet Pressure (P_1) and Temperature (T_1).
- Compressor Exit Pressure (P_2) and Temperature (T_2).
- Turbine Inlet Pressure (P_3) and Temperature (T_3).
- Turbine Exit Pressure (P_4) and Temperature (T_4).
- Exhaust Gas Pressure (P_5) and Temperature (T_5).
- Compressor Rotational Speed (RPM), measured by a tachometer generator.

- Fuel Pressure and Fuel Flow.
- Engine Thrust, measured via a dedicated load cell.

These parameters are displayed on the system's control panel and, more extensively, on a connected Data Acquisition Screen. Specifically, P_3 and T_3 (Turbine Inlet Temperature, displayed as TIT) are available on both the panel and screen, while T_5 (Exhaust Gas Temperature, displayed as EGT) is also accessible from both. Fuel Pressure is displayed on the panel. The precise location of these engine sensors is depicted in the SR-30 Gas Turbine Cutaway diagram within the Mini-Lab's operational manual.

Figure 1 – Schematic of Brayton Cycle for Gas Turbine and Cut Away of SR-30 Engine.



Source: Mini-Lab Gas Turbine Power System Manual Turbine Technologies, Ltd. (2011)

3.1.2 Data Acquisition System and Control Interface

The data acquisition process is managed through a dedicated Data Acquisition Computer running the MiniLab 1.1 software. This computer connects to the Mini-Lab system via a USB port, leveraging a National Instruments DAQ (Data Acquisition) system (specifically the NI

DAQ 6218 module) for real-time data capture and display. The MiniLab 1.1 software facilitates logging data to file, displaying plot features, and offers controls for unit toggling (e.g., Celsius to Fahrenheit for temperatures, Psig to kPa for pressures, Liters/hour or Gallons/hour for fuel flow, and Newtons to Pounds for thrust). The data sampling rate can be selected between 0.1 and 5 samples per second.

3.1.3 Experimental Procedure and Data Collection Protocol

The experimental procedure involved operating the Mini-Lab gas turbine through its pre-start, start-up, and operational phases, followed by a controlled shutdown. Prior to each run, essential checks were conducted, including verification of fuel properties and ambient barometric pressure $P_{amb} = 946.7mbar$. During operation, the MiniLab 1.1 software collected real-time data from various sensors, including temperatures, pressures, fuel flow, RPM, and thrust. This data was logged continuously to a file on the connected computer's hard drive via a USB connection to a National Instruments DAQ system. The software allowed for adjustable sampling rates, and the recorded data was stored in an ASCII format, enabling direct import into spreadsheet programs for subsequent detailed analysis.

3.2 Thermodynamic Modeling of the Gas Turbine System

3.2.1 System Overview

The system of interest is a small-scale gas turbine engine operating in steady-state. It consists of the following major components:

- Compressor: Increases the pressure and temperature of incoming ambient air.
- Combustor: Mixes compressed air with fuel, where combustion raises the temperature significantly.
- Turbine: Extracts energy from the hot combustion gases to drive the compressor and generate useful work.
- Nozzle (Exhaust Section): Accelerates the exhaust gases to produce thrust.

Each component operates under the assumption of quasi-one-dimensional, steady, adiabatic flow, with negligible heat loss to the surroundings unless explicitly modeled.

3.2.2 Assumptions and Idealizations

To derive tractable analytical expressions and guide neural network constraints, the following assumptions are adopted:

- The working fluid (air and combustion gases) behaves as a calorically perfect ideal gas:

$$P = RT, \quad c_p, c_v \text{ constant}, \quad = \frac{c_p}{c_v} \;\; 1.4$$

- Isentropic relations apply to ideal compression and expansion processes with specified efficiencies $_c$ (compressor) and $_t$ (turbine).
- Constant specific heats c_p and c_v are used, consistent with the assumption of ideal gases.
- No heat transfer or pressure losses in ducts, except where captured by model residuals.
- Steady-state and quasi-1D flow are assumed throughout.

3.2.3 Key Thermodynamic Equations

3.2.3.1 Compressor

- Isentropic temperature relation:

$$\frac{T_{2s}}{T_1} = \left(\frac{P_2}{P_1}\right)^{\frac{1}{\;}}$$

- Actual outlet temperature considering efficiency:

$$T_2 = T_1 + \frac{T_{2s} \;\; T_1}{_c}$$

- Power required by the compressor:

$$W_{\mathsf{comp}} = m_a c_p (T_2 \;\; T_1)$$

3.2.3.2 Combustor

- Energy balance (idealized with complete combustion and no heat losses):

$$m_f Q_{\mathsf{HV}} = (m_a + m_f) c_p (T_3 \;\; T_2)$$

where Q_{HV} is the lower heating value of the fuel and the gas mass flow is $m_g = m_a + m_f$.

3.2.3.3 Turbine

- Isentropic temperature relation:

$$\frac{T_{4s}}{T_3} = \left(\frac{P_4}{P_3}\right)^{\frac{1}{\;}}$$

- Actual outlet temperature considering turbine efficiency:

$$T_4 = T_3 \;\; _t(T_3 \;\; T_{4s})$$

- Power generated by the turbine:

$$W_{\mathsf{turb}} = (m_a + m_f) c_p (T_3 \;\; T_4)$$

3.2.3.4 Nozzle (Exhaust)

Assuming isentropic expansion and ambient back pressure P_0, the exhaust velocity is:

$$V_{\text{exit}} = 2c_p T_5 \left(1 \ \left(\frac{P_0}{P_5}\right)^{\frac{1}{}}\ \right)$$

The thrust is then calculated via momentum balance:

$$F = (m_a + m_f)(V_{\text{exit}} \ \ V_{\text{inlet}})$$

If inlet velocity is negligible (static tests), this simplifies to:

$$F \ \ (m_a + m_f)V_{\text{exit}}$$

3.2.4 Station Numbering Convention

To standardize variables across the PINN and thermodynamic sections:

1	Ambient conditions (inlet to compressor)
2	Compressor exit / combustor inlet
3	Combustor exit / turbine inlet
4	Turbine exit / nozzle inlet
5	Nozzle exit (exhaust gas temperature station)

This convention allows the PINN to learn both the observable outputs (T_3, T_5, F) and intermediate station states (T_2, P_2, T_4, P_4, etc.), constrained by physics-based residuals.

3.3 Physics-Informed Neural Network (PINN) Framework

To bridge the gap between purely data-driven models and first-principle simulations, this work employs a Physics-Informed Neural Network (PINN) to model the gas turbine system.

3.3.1 Core Concept

PINNs are neural networks that embed physical laws directly into their training process. Rather than minimizing only the error between model predictions and experimental data, the PINN loss function also includes terms that penalize violations of governing physical equations. This approach ensures that predictions are both data-accurate and physically consistent, even in regions where data is sparse.

3.3.2 PINN Architecture

The core model is a fully connected Multi-Layer Perceptron (MLP) trained to learn mappings from inputs to thermodynamic states and performance metrics. The architecture is guided by simplifying assumptions from thermodynamics:

- Air is modeled as a calorically perfect ideal gas with constant specific heats (c_p, c_v) and a constant heat capacity ratio $= c_p/c_v \;\; 1.4$.
- Isentropic efficiencies for the compressor ($_c$) and turbine ($_t$) are assumed constant.
- Heat losses, frictional effects, and pressure drops outside of defined station points are neglected.

Inputs: Key operational parameters that determine the thermodynamic state of the system:

- Fuel flow rate m_f (kg/s)
- Compressor rotational speed N_1 (RPM)
- Ambient temperature T_1 (K)
- Ambient pressure P_1 (Pa)

Outputs: Predicted thermodynamic variables and performance indicators:

- Station states: $T_{2,\text{pred}}, P_{2,\text{pred}}, T_{3,\text{pred}}, P_{3,\text{pred}}, T_{4,\text{pred}}, P_{4,\text{pred}}, T_{5,\text{pred}}, P_{5,\text{pred}}$
- Net thrust F_{pred} (N)
- Air mass flow rate $m_{a,\text{pred}}$ (kg/s)

3.3.3 Hybrid Loss Function

The total loss used to train the PINN is a weighted sum of data-driven and physics-based components:

$$L_{\text{total}} = w_{\text{data}} L_{\text{data}} + w_{\text{physics}} L_{\text{physics}}$$

- Data Loss (L_{data}): A supervised learning loss (Mean Squared Error) that penalizes deviation from available experimental measurements:

$$L_{\text{data}} = \frac{1}{N} \, {}_{i=1}^{N} (y_i \;\; y_{i,\text{pred}})^2$$

- Physics Loss (L_{physics}): A sum of residuals derived from physical principles, each enforcing thermodynamic consistency.

3.3.3.1 Compressor Temperature Rise Loss (L_{T2})

Derived from the isentropic temperature relation:

$$L_{T2} = \left({}_{c}(T_{2,\text{pred}} \ T_1) \ T_1 \left(\left(\frac{P_{2,\text{pred}}}{P_1} \right)^{\frac{1}{\ }} \ 1 \right) \right)^2$$

3.3.3.2 Turbine Temperature Drop Loss (L_{T4})

Enforcing the relationship between pressure drop and temperature drop across the turbine:

$$L_{T4} = \left((T_{3,\text{pred}} \ T_{4,\text{pred}}) \ {}_{t}T_{3,\text{pred}} \left(1 \ \left(\frac{P_{4,\text{pred}}}{P_{3,\text{pred}}} \right)^{\frac{1}{\ }} \right) \right)^2$$

3.3.3.3 Shaft Power Balance Loss (L_{power})

For a single-shaft turbojet, the power generated by the turbine (W_{turb}) is consumed entirely by the compressor (W_{comp}). This constraint, $W_{\text{turb}} = W_{\text{comp}}$, provides a fundamental physical link between the component states. The loss residual enforces this balance.

$$L_{\text{power}} = ((m_a + m_f)c_p(T_{3,\text{pred}} \ T_{4,\text{pred}}) \ m_a c_p(T_{2,\text{pred}} \ T_1))^2$$

3.3.3.4 Combustor Energy Balance Loss (L_{comb})

Assuming ideal heat release with no losses:

$$L_{\text{comb}} = (m_f Q_{\text{HV}} \ (m_a + m_f)c_p(T_{3,\text{pred}} \ T_{2,\text{pred}}))^2$$

3.3.3.5 Thrust Loss

A simplified thrust estimation (using momentum conservation) could also be encoded if exit velocity and intake velocity are known or modeled:

$$L_{\text{thrust}} = (F_{\text{pred}} \ (m_a + m_f)(V_{\text{exit}}))^2$$

While air mass flow rate m_a could be derived from a physics-based model based on N_1 and ambient conditions, in this PINN-based approach, it is treated as a learnable parameter. This allows the model to infer the air mass flow rate directly from the operational data, providing a more flexible and potentially more accurate estimation by capturing complex relationships that a simplified physics model might overlook.

3.4 **Model Implementation and Evaluation**

This section outlines the strategy for the development, training, and rigorous evaluation of the gas turbine digital twin.

3.4.1 Problem Definition

The primary objective is to develop a digital twin capable of predicting key performance indicators—steady-state thrust (F_{pred}), combustor outlet temperature ($T_{3,pred}$), and exhaust gas temperature ($T_{5,pred}$)—based on operational inputs (fuel flow m_f, compressor speed N_1) and ambient conditions (T_1, P_1).

To enforce physical realism, the model simultaneously predicts a set of intermediate thermodynamic variables: - Air mass flow: $m_{a,pred}$ - Station temperatures: $T_{2,pred}, T_{4,pred}$ - Station pressures: $P_{2,pred}, P_{3,pred}, P_{4,pred}, P_{5,pred}$

These intermediate predictions are constrained during training by the physics-informed loss functions derived in the previous chapter.

3.4.2 Data Preparation and Preprocessing

The experimental dataset undergoes a standard preprocessing pipeline before being used for model training:

- Unit Conversion: All variables are converted to base SI units (e.g., temperatures to Kelvin, pressures to Pascals, fuel flow to kg/s) to ensure consistency in the physics-based calculations.
- Normalization: Input and output features are scaled to a common range (e.g., [0, 1]) using min-max scaling. This improves numerical stability during training and helps the optimization algorithm converge more efficiently.
- Data Splitting: The complete dataset is randomly partitioned into three subsets: 80% for training, 10% for validation (hyperparameter tuning), and 10% for final, unbiased testing of the trained model.

3.4.3 Model Architecture and Hyperparameters

The core of the digital twin is a Physics-Informed Neural Network (PINN) based on a Multi-Layer Perceptron (MLP).

- Baseline Architecture: The initial architecture for exploration is a fully connected MLP with 4 input neurons, 2 hidden layers containing 32 neurons each, and 10 output neurons.
- Inputs (4): Fuel flow (m_f), compressor speed (N_1), ambient temperature (T_1), and ambient pressure (P_1).
- Outputs (10): Thrust (F_{pred}), air mass flow ($m_{a,pred}$), and the eight station states ($T_{2,pred}, P_{2,pred}, T_{3,pred}, P_{3,pred}, T_{4,pred}, P_{4,pred}, T_{5,pred}, P_{5,pred}$).

- Activation Functions: The ReLU (Rectified Linear Unit) activation function is used for hidden layers, while a Linear activation is used for the output layer to allow for unbounded physical values.
- Hyperparameter Tuning: A hyperparameter search will be conducted using the validation set to optimize the network architecture (e.g., number of layers and neurons), learning rate, and loss function weights to achieve the best performance. The search will be using a random search with 10 iterations.

3.4.4 Loss Function and Weighting Strategy

The custom loss function guides the PINN to be both data-accurate and physically-consistent. The total loss is a weighted sum of two components:

$$L_{total} = w_{data} L_{data} + w_{physics} L_{physics}$$

- Data Loss (L_{data}): This is the Mean Squared Error (MSE) between the model's predictions and the measured experimental data for the observable outputs (T_3, T_5, F, etc.).
- Physics Loss ($L_{physics}$): This is the sum of the mean squared residuals from the governing thermodynamic equations (e.g., $L_{power}, L_{comb}, L_{T2}, L_{T4}$). This term penalizes solutions that violate physical laws.
- Weighting Strategy: The weights ($w_{data}, w_{physics}$) are critical hyperparameters that balance the influence of the experimental data and the physical laws. A sensitivity analysis will be performed to find a set of weights that minimizes prediction error on the validation set without compromising physical consistency.

3.4.5 Training and Evaluation Metrics

The model will be implemented and trained using the PyTorch framework with the AdamW optimizer. To ensure a thorough and unbiased assessment of the final digital twin, its performance will be evaluated on the unseen test set using three key criteria:

1. **Predictive Accuracy:** The model's ability to predict the primary performance indicators will be quantified using Mean Squared Error (MSE) for F_{pred}, $T_{3,pred}$, and $T_{5,pred}$.
2. **Physical Consistency:** The magnitude of the physics loss residuals ($L_{physics}$) will be calculated on the test set. A low residual value indicates that the model's predictions adhere well to the governing laws of thermodynamics, even for data it has never seen.
3. **Overall Fidelity:** The accuracy of the intermediate variable predictions (e.g., P_2, T_2, m_a) will also be assessed to ensure the overall health and internal consistency of the model.

Bibliography

BOYCE, Meherwan P. **Gas Turbine Engineering Handbook**. *[S.l.]*: Elsevier Science, 2012.

CENGEL, Yunus A.; BOLES, Michael A. **Thermodynamics: An Engineering Approach**. *[S.l.]*: McGraw-Hill Education, 2019.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. *[S.l.]*: MIT Press, 2016.

GRIEVES, Michael. Digital Twin: Manufacturing Excellence through Virtual Replica of Physical Product for Future Factory. **White Paper, Florida Institute of Technology**, 2011.

GURNEY, L. S. **Gas Turbine Performance**. *[S.l.]*: Butterworth-Heinemann, 2010.

INCROPERA, Frank P. *et al.* **Fundamentals of Heat and Mass Transfer**. *[S.l.]*: John Wiley & Sons, 2007.

KARNIADAKIS, George Em *et al.* Physics-informed machine learning. **Nature Reviews Physics**, v. 3, n. 6, p. 422–441, 2021.

KREUZER, Tim; PAPAPETROU, Panagiotis; ZDRAVKOVIC, Jelena. Artificial intelligence in digital twins—A systematic literature review. **Data & Knowledge Engineering**, 2024.

KURZ, R.; BRUN, K. **Gas Turbine System Design, Analysis, and Optimization**. *[S.l.]*: ASME Press, 2009.

RAISSI, Maziar; PERDIKARIS, Paris; KARNIADAKIS, George E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. **Journal of Computational Physics**, v. 378, p. 686–707, 2019.

RUSSELL, Stuart; NORVIG, Peter. **Artificial Intelligence: A Modern Approach**. *[S.l.]*: Prentice Hall, 2010.

SARAVANAMUTTOO, H. I. H. *et al.* **Gas Turbine Theory**. *[S.l.]*: Pearson, 2017.

SCHLUSE, Martin *et al.* Digital Twin: An Overview of the Current State. *In*: **Virtual and Augmented Reality in Industry**. *[S.l.]*: Springer, Cham, 2018. p. 253–272.

SERWAY, Raymond A.; JEWETT, John W. **Physics for Scientists and Engineers with Modern Physics**. *[S.l.]*: Cengage Learning, 2018.

TAO, Fei *et al.* **Digital Twin Driven Smart Manufacturing**. *[S.l.]*: Elsevier, 2019.

TURBINE TECHNOLOGIES, LTD. **Mini-Lab Gas Turbine Power System™ Sample Lab Experiment Manual**. *[S.l.]*: Turbine Technologies, Ltd., 2011.

VERSTRAETE, D.; VAN DEN BRAEMBUSSCHE, R. A. **CFD for Turbomachinery**. *[S.l.]*: VKI Lecture Series, 2010.

WANG, Yibin; LU, Beijia; ZHANG, Jian. A physics-informed neural network for modeling and prognosis of steam turbine performance. **Applied Thermal Engineering**, v. 223, p. 119934, 2023.