DYNAMIC BEHAVIOUR OF A CONCRETE DAM: DEVELOPMENT OF STATISTICAL AND MACHINE LEARNING MODELS FOR INTERPRETATION OF MONITORING DATA

ANDRÉS MAURICIO GUZMÁN SEJAS

Dissertation submitted in partial fulfillment of the requirement or the degree of

MASTER OF CIVIL ENGINEERING — SPECIALIZATION IN STRUCTURES AND GEOTECHNICS

Supervisor: Professor Dr. Álvaro Alberto de Matos Ferreira da Cunha

---

Co-Supervisor: Dr. Sérgio Bouça Pereira

Co-Supervisor: Dr. Juan Tomé Caires da Mata

JUNE 2024

MESTRADO EM ENGENHARIA CIVIL 2023/2024

DEPARTAMENTO DE ENGENHARIA CIVIL

Tel. +351-22-508 1901

✉ m.ec@fe.up.pt

Editado por

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Rua Dr. Roberto Frias

4200-465 PORTO

Portugal

Tel. +351-22-508 1400

✉ feup@fe.up.pt

http://www.fe.up.pt

To my family and friends

Al final, todo estará bien y si no está bien, no hemos llegado al final.

Pelusa Caligari

ACKNOWLEDGMENTS

The completion of this dissertation marks the end of another chapter in my life, another goal achieved and completed on a path filled with difficulties, where many people stood by me to whom I would like to express my gratitude.

I would like to thank the Faculty of Engineering of the University of Porto, which was part of this entire process, shaping me and welcoming me with open arms from the beginning. I could not be more grateful for this university and the people within it.

To the National Laboratory for Civil Engineering (LNEC), which always offered me their full support.

To Movhera - Hidroelétricas do Norte, S.A and ENGIE Douro - Hidroeléctricas do Douro, for the support and authorization of the use of experimental data related with continuous dynamic monitoring of Baixo Sabor dam.

To my supervisor, Professor Dr. Alvaro Cunha, for providing me the opportunity, trust, support, and knowledge to carry out this thesis.

To my co-supervisor, Professor Dr. Sérgio Pereira, for all his time, dedication, support, wide knowledge, and patience throughout the last part of this amazing journey.

To my co-supervisor, Professor Dr. Juan Mata, for his knowledge, availability, dedication, patience, unconditional support, and motivation during my last cycle of studies.

To my parents for their unwavering faith in me, for their affection, love, and motivation day and night, always there for me with their unconditional support.

To my siblings who have always been there for me, watching me grow throughout this process, never ceasing to motivate and support me throughout my life.

To Julia and Sidney, whom I know I can always count on and who I know will be there when I need them to guide my path.

To my nephews Mateo, Fernando, and niece Clara, who despite the distance always bring a smile to my face, and to whom my support will always go.

To my friends in my beloved Bolivia Ariel, Daniela, Fernanda, Irvin, Joel, and Katy who despite the distance offered me their support, motivation, and encouragement in times when I needed true friendship. Thank you very much to each of them.

To the good friends I have had the pleasure of meeting during this long and arduous journey because, with their company, support, and hours together, they made this experience the best at FEUP, Amanda, André, Camille, Catalina, Cesar, Florián, João, José, Kenny, Lana, Lucas, Sahil, and Susana. Thank you very much to each of you, without a doubt, you will always have my sincerest support.

I would also like to thank every person who has ever shared time with me and enjoyed good moments together.

"Last but not least, I want to thank myself, I want to thank me for believing in me, for doing all this hard work, for not taking days off, I want to thank me for not giving up and stopping believing in myself, and most importantly I want to thank me for just being me at all times."

RESUMO

Esta dissertação centra-se na monitorização dinâmica de barragens de betão, com um foco específico na barragem do Baixo Sabor como caso de estudo. O objetivo principal desta monitorização dinâmica é observar continuamente o comportamento da barragem, garantindo que permanece dentro dos limites de variação esperados e possibilitar a emissão de alertas em caso de desvios, o que é crucial para manter a funcionalidade e integridade da estrutura. O processo de monitorização baseia-se em instrumentação in-situ e modelos de comportamento que utilizam reconhecimento de padrões.

O trabalho realizado teve como objetivo desenvolver, calibrar e comparar modelos estatísticos e de aprendizagem automática para apoiar a interpretação do comportamento dinâmico observado de uma barragem de betão em funcionamento. A metodologia inclui várias etapas-chave: análise modal operacional de séries temporais de aceleração, caracterização da evolução temporal de variáveis ambientais e operacionais influentes, construção e calibração de modelos preditivos usando métodos estatísticos e de aprendizagem de máquina, e a comparação das suas eficiências.

Foram desenvolvidos e testados modelos de regressão linear e de redes neuronais. Os resultados indicaram que, embora ambos os modelos sejam capazes de prever eficazmente o comportamento dinâmico da barragem, a rede neuronal superou ligeiramente o modelo de regressão em termos de precisão de previsão. No entanto, o modelo de regressão linear mostrou-se mais interpretável. Para o caso de estudo da barragem do Baixo Sabor, a rede neuronal demonstrou desempenho robusto na captura de relações complexas entre variaveis dependentes e independentes, que a torna uma opção sólida para compreender o comportamento dinâmico influenciado por efeitos ambientais e operacionais.

Em conclusão, tanto os modelos de regressão linear como os de redes neuronais são viáveis para projetos de monitorização dinâmica. A escolha entre eles depende da complexidade dos dados e da relevância dos resultados.

O trabalho realizado ao longo deste estudo inclui uma compreensão abrangente da monitorização e controlo de segurança de barragens, exploração de diferentes abordagens de modelação e um caso de estudo detalhado da barragem do Baixo Sabor. Os resultados contribuem para o conhecimento e metodologias no campo da monitorização da condição estrutural, enfatizando a importância de modelos preditivos avançados para garantir a segurança e funcionalidade de infraestruturas críticas.

PALABRAS-CHAVE: Barragem do Baixo Sabor, Barragem de Betão, *Machine Learning,* Regressão Linear Múltipla, Comportamento Estrutural.

ABSTRACT

This dissertation focuses on the dynamic monitoring of concrete dams, with a specific emphasis on the Baixo Sabor dam as a case study. The main objective of this dynamic monitoring is to continuously observe the dam's behaviour, ensuring it remains within expected variation limits and issuing alerts if deviations occur. This is essential for maintaining the structure's functionality and integrity. The monitoring process relies on on-site instruments and behaviour models that use pattern recognition, thereby avoiding explicit dependence on mechanical principles.

The undertaken work aimed to develop, calibrate, and compare statistical and machine learning models to aid in interpreting the observed dynamic behaviour of a functional concrete dam. The methodology included several key steps: operational modal analysis of acceleration time series, characterization of the temporal evolution of observed magnitudes and influential environmental and operational variables, construction and calibration of predictive models using both statistical and machine learning methods, and the comparison of their effectiveness.

Both linear regression and neural network models were developed and tested. The results showed that while both models effectively captured and predicted the dam's behaviour, the neural network slightly outperformed the regression model in prediction accuracy, particularly with complex variables and data. However, the linear regression model was more interpretable. For the Baixo Sabor dam, the neural network demonstrated strong performance in capturing complex relationships between input features and the target variable. This makes it a solid choice for understanding the dynamic behaviour influenced by natural frequencies and environmental effects.

In conclusion, both linear regression and neural network models are suitable for dynamic monitoring projects. The choice between them depends on the complexity of the data and the relevance of the results. For large-scale projects like the Baixo Sabor dam, neural networks offer significant advantages in handling intricate data relationships, thus providing better insights into the dam's dynamic behaviour.

The work conducted throughout this study includes a thorough understanding of dam monitoring and safety control, an exploration of various modelling approaches, and a detailed case study of the Baixo Sabor dam. The findings contribute to knowledge and methodologies in the field of structural health monitoring, highlighting the importance of advanced predictive models for ensuring the safety and functionality of critical infrastructure.

KEYWORDS: Baixo Sabor Dam, Concrete Dam, Machine Learning, Multiple Linear Regression, Structural Behaviour.

TABLE OF FIGURES

TABLE INDEX

SYMBOLS, ACRONYMS, AND ABBREVIATIONS

ANN	Artificial Neural Network
ANPSD	Average Normalized Power Spectral Density
APA	Agência Portuguesa do Ambiente
EMA	Exponential Modal Analysis
FEUP	Faculdade de Engenharia da Universidade do Porto
Fitrnet	Fit Regression Neural Network
FTP	File Transfer Protocol
GPS	Global Positioning System
HST	Hydrostatic, Seasonal, Time
HTT	Hydrostatic, Thermal, Time
ICOLD	International Commission on Large Dams
LBFGS	Limited-memory Broyden-Fletcher-Goldfarb-Shanno
LiDAR	Light Detection and Ranging
LNEC	Laboratório Nacional de Engenharia Civil
MAC	Modal Assurance Criterion
MATLAB	Matrix Laboratory
MLP	Multilayer Perceptron
MLR	Multiple Linear Regression
N	Number of nodes
NN	Neural Networks
NUC	Normalized Unbiased Cross-correlation
OMA	Operational Modal Analysis
PNBEPH	Plano Nacional de Barragens com Elevado Potencial Hidroeléctrico
p-LSCF	Poly-reference Least Squares Complex Frequency
RMD	Manual Data Collection
RPB	Regulamento de Pequenas Barragens
RSB	Regulamento de Segurança de Barragens
RMS	Root Mean Square
SNIRH	Sistema Nacional de Informação de Recursos Hídricos
SSI-Cov	Stochastic Subspace Identification - Covariance
Vibest	Laboratory of Vibrations and Structural Monitoring

PH.	Hydrostatic Pressure
S	Uplift Pressure
h	Height of the water in the reservoir
kN	Kilonewton
m	Meters
mm	Millimeters
rpm	Revolutions per minute
$°C$	Degrees Celsius
R^2	Determination Coefficient
σ	Standard deviation of residuals
K	Window size for spectral analysis
Hz	Hertz
MW	Megawatts
$\delta(h_i, S_i, t_i)$	Observed value of the quantity under analysis in observation
$U_i(h_i) + U_s(S_i) + U_t(t_i)$	Parcels of the quantities corresponding to the elastic effect of the water level in the reservoir, the elastic effect of seasonal temperature variation and the effect of time on the i^{th} observation.
k	Constant that corresponds to the difference between observed and calculated values at the beginning of the calibration period.
ε_i	Residual of the i^{th} observation given by the difference between the estimated value and the observed value.
a_n	Number of days between the first filling and the start date of the analysis
t_d	Number of days elapsed from the beginning of the year to the date of observation
β_j	Regression coefficient corresponding to the term X_j
X_j	J^{th} term associated with the actions
t	Number of days between the campaign and the start of the monitoring

1.

INTRODUCTION

1.1. CONTEXT

In the field of civil engineering, the safety and stability of concrete dams are critical concerns that demand continuous attention. These structures, designed to contain large volumes of water for human supply, irrigation, flood control and production of hydro-electricity, represent a vital asset to any country's infrastructure. However, over time, exposure to adverse environmental conditions, the ageing of the dams and sometimes some of them with swelling problems, and the dynamic loading of natural forces, such as earthquakes, can affect their structural integrity and functionality.

Monitoring static quantities in concrete dams are crucial for ensuring their stability and safety over time. Key aspects of this monitoring include observing displacements and deformations, measuring strain and stress, and assessing the hydraulic flow in the foundation.

Displacements and deformations involve tracking the horizontal and vertical movements of the dam structure, which can result from changes in water pressure or foundation (ICOLD, 2011). Monitoring crack widths is also essential to detect any structural weaknesses. Strain gauges are used to measure how the dam's concrete deforms under various loads, providing insight into internal stress levels (ASCE, 2020). This information is vital for identifying areas where the material might be approaching its stress limits.

Hydrostatic pressure, the force exerted by the water the dam holds back, is another critical factor. Consistent monitoring of these pressures helps in understanding the forces acting on the dam and identifying potential risks (Journal of Hydraulic Engineering, 2018).

Methods of monitoring these static quantities typically involve regular geodetic surveys, which use precise measurements to track any movements or changes in the dam's structure over time. Remote sensing technologies, such as satellite imagery and LiDAR, offer detailed views and can detect surface deformations that might not be visible from the ground (USBR, 2014).

The importance of monitoring lies in its ability to provide early warning signs of potential issues, allowing for timely maintenance and repair, and ensuring the overall safety and functionality of the dam. Regular data collection helps engineers take preventive actions, thereby safeguarding communities and environments downstream.

Dynamic monitoring of concrete dams is becoming an essential tool to ensure their optimal performance and safety over time. This approach, involving continuous observation of parameters such as vibrations, natural frequencies, and deformations, allows for the detection and response to potential anomalies or changes in the structure's behaviour. The evolution of technology, especially in the field of machine learning, has opened up new possibilities for improving the interpretation and prediction of observed behaviour in these structures.

This study focuses on the development of statistical and machine learning models to interpret the dynamic behaviour of concrete dams. By integrating data obtained from continuous monitoring with machine learning algorithms, the aim is not only to understand the current behaviour of the structure but also to identify potential changes or deviations that could compromise its safety. This innovative approach not only strengthens the monitoring and control capabilities of concrete dams but also provides a solid foundation for informed decision-making regarding maintenance, repair, and risk management.

In this context, the development of statistical and machine learning models usually adopted in the analysis of “quasi-static” quantities emerge as a powerful tool to enhance the understanding and management of concrete dams in a dynamic and changing environment. By combining traditional knowledge with innovative approaches, this study aims to contribute to the development of more efficient engineering practices for these essential structures in civil infrastructure.

1.2. OBJECTIVES AND CONTRIBUTIONS

The present dissertation focuses on the topic of dynamic monitoring of structures, specifically concrete dams. The main objective of the dynamic monitoring of a concrete dam is to continuously track the evolution of its dynamic behaviour, allowing the characterization of the observer pattern behaviour, ensuring that it remains within expected variation limits, and issuing alerts when deviations occur. This approach is designed to guarantee that the structure satisfactorily performs the functions for which it was built, maintaining both functionality and structural integrity.

This activity relies on on-site observations through instrumentation installed in situ, as well as on behaviour models. Some of these models are based on pattern recognition that allows for characterizing the normal behaviour of structures without explicitly resorting to the principles of mechanics. In this dissertation, the aim is to construct, calibrate, and compare statistical and machine learning models to support the interpretation of observed dynamic behaviour in an operational concrete dam.

The use of these models involve not only collecting and processing data from the installed instruments but also using advanced statistical and machine learning techniques to derive insights. By doing so, the work seeks to establish reliable methods to detect anomalies or unusual behaviour in the structure, thereby contributing to the safety and longevity of the dam.

To achieve these goals, the following objectives are addressed in this thesis:

- Operational modal analysis of time series of accelerations. The identification of modal parameters in dynamic systems based on their response to ambient excitation can be achieved through several methods, each with different working principles. One common approach is the 'peak picking' method, which was the technique adopted in this study.
- Characterization of the temporal evolution of observed quantities and the environmental and operational variables that affect them. This includes factors like the hydrostatic effect, and seasonal effects (such as annual temperature variations and time-dependent). These variables will be applied throughout the observations, allowing for a combined analysis of the observed behaviour and environmental effects. This comprehensive approach will provide a more accurate picture of the dynamic system's performance under varying conditions.
- Construction and calibration of prediction models using statistical methods and machine learning techniques. This involves building HST (hydrostatic-seasonal-time) models using multiple linear regression and neural networks. The data used for these models will be gathered through automatic collection processes. The data from the monitoring system of a concrete dam under exploitation will serve as a basis for developing models that can capture the natural frequencies of the system.
- Comparison of the efficiency of various methodologies studied under the HST approach. The focus here is on assessing the performance of a multiple linear regression model and neural network, with a detailed analysis to provide conclusions based on the effectiveness and manageability of each approach. By comparing these methods, the study aims to identify if both techniques are most suitable for predicting the behaviour of dynamic systems under unseen data, ultimately contributing to improved safety and reliability in dams.

The project was carried out using the MATLAB software (Version R2022a), which provides the capability to implement and evaluate models with precision and efficiency. By adopting MATLAB software as a reference and potential software, every team gains access to advanced tools for data processing and result validation. If properly implemented, this ensures that the models are developed with high accuracy and that the outcomes are thoroughly tested to meet the project's requirements.

1.3. THESIS OUTLINE

This section describes how the document is organised. The present thesis is divided into five chapters, providing a summary of each chapter covering different aspects of the study on the monitoring and safety control activities of concrete dams to give the readers a clear idea of the document's flow and how it is structured throughout its development.

Chapter 1 lays the groundwork for the project. In the framework section, the topic is contextualised by highlighting the importance of structural safety in concrete dams and the necessity of having effective monitoring and observation systems. The objectives section outlines the primary goal of the project, which is the major challenge, and also details the specific objectives—these are the incremental steps that must be completed to meet or achieve the thesis's main goal.

Chapter 2 addresses the topic of safety control and health monitoring of concrete dams, introducing the concept of safety monitoring, and emphasizing its importance in preventing structural failures and potential disasters. This section also discusses some of the main challenges faced by concrete dams and why continuous monitoring is essential. "Concrete dams" section gives an overview of concrete dams, describing their key characteristics and how they differ from other types of dams. The "Safety control activities" section outlines the specific activities related to the safety control of dams. This can include regular inspections, analysis of historical data, and preventive measures to avoid structural problems. The "Structural behaviour monitoring" section explores the techniques and technologies used to monitor the structural behaviour of dams. This discusses observation systems for measuring the main loads, such as the water level and the air temperature variations, as well as the static and dynamic variables measured in the concrete dam and its foundation, such as stress, strain, and temperature in the concrete dam body, relative movement between joints, seepage and leakage, uplift pressure in the rock mass foundation and accelerations, among other quantities.

Chapter 3 focuses on interpreting the observed behaviour in concrete dams. The "Introduction" section sets the framework for the analysis, emphasizing the importance of understanding how structures respond to different conditions. The "Structural Behavior Interpretation" section outlines two main approaches to interpret structural behaviour: the HST approach and the HTT approach. However, the HTT models generally provide better performance due to their inclusion of additional data, such as temperatures in the dam body. Nevertheless, this project was focused on the HST approach, since it allows the use of existing data without the need for additional temperature measurements. Moreover, a multiple linear regression model will be provided, demonstrating how these models can be useful for identifying patterns in monitoring data and predicting future trends. The "Neural Network Models" section introduces the concept of neural networks, a type of artificial intelligence used for analyzing and predicting complex data, highlighting their capacity to learn from large datasets and perform sophisticated analyses to help interpret structural behaviour. Finally, the "Operational modal analysis" section introduces the concept of operational modal analysis, a technique used to understand how structures respond to vibrations and other dynamic behaviour. It describes the Peak-Picking method, a technique used to identify a structure's natural frequencies from real-world data. This method will also be applied to experimental data.

Chapter 4 focuses on a specific case study: the Baixo Sabor dam. In "Baixo Sabor hydroelectric scheme," section the dam and its monitoring system are described, and then the different chapters already studied will help to put into practice what has been learned and obtain interpretations of the dam's behaviour based on real-time monitored data. In this specific case study, both methods will be applied based on the HST approach: multiple linear regression, where the model is primarily learned and then tested to predict unseen events, and neural networks, where the same procedure is followed using a different technique. Finally, in "Comparison between regression and neural network models," a comparison is made between the two models to determine which is more effective in this case study. The advantages and disadvantages of each approach are discussed, and recommendations are provided for their use.

Chapter 5 the final chapter, offers a summary of the work conducted and presents general and specific conclusions about each proposed objective. It provides a summary of the work done throughout the project, highlighting the main achievements and findings. In the following section of the same chapter, the results are discussed, reflecting on what has been learned and suggesting possible interpretations of the data. Finally, some directions for future research are suggested, indicating where further studies or additional improvements are needed to increase the safety and efficiency of concrete dam monitoring.

2.

SAFETY CONTROL AND HEALTH MONITORING OF CONCRETE DAMS

2.1. INTRODUCTION

Concrete dams are critical structures for water supply, hydroelectric power generation, and flood control worldwide. The safety and operational efficiency of these structures depend on continuous monitoring and observation of their state and condition. These structures have been fundamental in water and energy management throughout history, from ancient times to the present day.

Dams have a long history that dates back to ancient times. The earliest water storage systems were built in Mesopotamia over 3,000 years ago, and since then, dams have evolved to meet the needs of societies and technology. In the 19^{th} century, the construction of dams became more widespread with the invention of concrete technology, allowing for the building of larger and safer structures. Dams have become an integral part of several hydroelectric and irrigation infrastructures worldwide, providing efficient and safe water supply and hydroelectric power generation (Tata & Howard, 2021).

Throughout the 20^{th} century, concrete dams became fundamental in global water and energy management. The construction of dams enabled flood control, hydroelectric power generation, and water supply for agriculture, industry, and urban populations. However, the construction of dams has also faced criticism and controversies due to the environmental and social impacts they can have on communities along the river (Tata & Howard, 2021).

The history of concrete dams in Portugal is set within the broader context of hydraulic development and hydroelectric power generation in Europe. Throughout the 20th century, Portugal, like many other European countries, faced the challenge of managing its water resources for multiple purposes, including potable water supply, agricultural irrigation, and energy production. While the use of dams in Portugal dates back to ancient times, the use of concrete as the primary material began to gain traction in the mid 20^{th} century. This period coincided with the rise of hydroelectric power in Europe, where many countries began constructing large dams to generate electricity and regulate water flows (Batista, 2022).

Portugal's diverse topography, with major rivers like the Tagus and Douro, provides significant potential for hydroelectric power. During the first half of the 20th century, Portugal began constructing several significant dams to harness this energy source. Concrete dams, because of the topography of the valleys, the knowledge about the construction process, as well as their durability and capacity to withstand high water pressure, became a popular choice. One of Portugal's most iconic concrete dams is the Castelo do Bode dam, located on the Zêzere river, a tributary of the Tagus. Built in the 1950s, it is one of the country's largest dams and plays a key role in Portugal's hydroelectric power system and the water supply for the region of Lisbon. Another significant dam is the Alqueva dam on the Guadiana river, completed in 2002[1] (gypsywarrior.com). More recently there is the example of the Alto Tâmega dam, the Daivões dam, and the Gouvães dam, which are part of the Alto Tâmega hydropower scheme.

This project, led by Iberdrola, represents a total investment of 1,500 million euros and aims to produce approximately 1,760 gigawatt-hours per year, which is equivalent to around 6% of the country's electrical consumption[2] (Industriaeambiente.pt).

Catastrophes such as those in Banqiao (Zipper, 2020) and Malpasset (Carmo, 2013) dams highlight the critical importance of safety control and monitoring activities for dams.

The Malpasset dam, located on the Reyran river near Fréjus in southern France, collapsed in December 1959. The cause of the failure was primarily water infiltration through the foundation of the dam, weakening its structure. The collapse occurred overnight, releasing approximately 50 million cubic meters of water. The resulting flood caused the deaths of 423 people and destroyed local infrastructure, including bridges and homes. This tragedy underscored the importance of geology in dam construction and the need for constant and careful monitoring[3] (Lynch, ASDSO).

In all operating structures, regulatory safety requirements must be met, making the use of rigorous observation techniques and the development of credible reference models for interpreting and predicting dam behaviour increasingly essential. These models help guide decisions regarding operational adjustments, maintenance, rehabilitation, or even demolition, if necessary (Mata, 2007).

Safety control activities of a dam are a multidisciplinary and multivariate process, given the variety of factors monitored and the complexity of the analyses involved. This process requires expertise in multiple disciplines such as structural engineering, materials engineering, geology, hydraulics, chemistry, and more. Automation of measurements, although introduced several years ago in some

---

[1] https://gypsywarrior.com/es/geograf%C3%ADa-y-recursos-naturales-en-portugal/, website visited 22/05/2024.
[2] https://tamega.iberdrola.pt/projeto/, website visited 22/05/2024.
[3] https://damfailures.org/case-study/malpasset-dam-france-1959/, website visited 14/05/2022.

projects in Portugal, is not yet fully established. Measurement errors still occur, and they must be addressed to ensure data reliability, compromising the effectiveness of timely safety control responses.

2.2. CONCRETE DAMS

Dams can be classified according to their function, to the materials that compose them and to their forms and structural schemes. The types of concrete dams have evolved mainly to suit the topography of the site and the different uses of water. These two features help to determine the best type of dam, always bearing in mind that they must meet two main functions: to withstand the pressure of the water and to release it when necessary, as can be seen in Fig. 1 describing the types of concrete dams.

Concrete dams can be classified based on the construction material and according to their shape, taking into account their structural schemes can be classified as gravity dams, buttress dams, arch dams, multiple arch dams, and taking into account the building processes such as the traditional and the roller-compacted concrete dams.

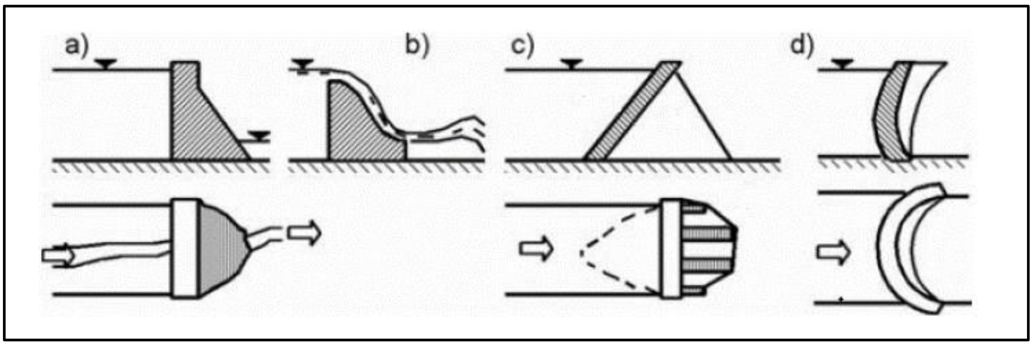


Fig. 1 - Types of concrete dams: a) b) gravity dam, c) buttress dam, d) arch dam, adapted from (Tanchev, 2014).

Mass concrete for use in dams should present satisfactory density, strength, durability, and resistance to avoid cracking and low thermal volume change and permeability. Concrete dams, the main study object of this thesis, are generally large and massive structures whose construction requires enormous volumes of concrete, poured at high placing rates, and subjected to lower compressive stresses when compared to other structures.

2.3. SAFETY CONTROL ACTIVITIES

In Portugal, the regulations governing the safety control of dams throughout their entire lifespan are defined by the Dam Safety Regulation (RSB) for large dams and the Small Dams Regulation (RPB) for smaller dams. According to the Dam Safety Regulation, dams are classified into two groups (APA, 2018):

- Large dams: These are dams with a height of 15 meters or more (measured from the lowest point of the foundation surface to the top of the dam) or dams with a height of 10 meters or more where the reservoir has a capacity exceeding 1,000,000 cubic meters. Furthermore, according to the latest data, there are more than 260 large dams in Portugal[4] (apambiente.pt).
- Small dams: Dams shorter than 15 meters that are not included in the first group but have a reservoir capacity greater than 100,000 cubic meters.

Any dams not covered by the Dam Safety Regulation are automatically subject to the provisions of the Small Dams Regulation. The framework for safety control of Portuguese dams currently in force was established by Decree-Law N. 21/2018, dated March 28. This decree consolidates the following provisions into a single document[5] (apambiente.pt):

- Dam Classification
- Dam Safety Regulation (RSB)
- Small Dams Regulation (RPB)
- Technical Support Documents (to support the application of dam safety regulations).

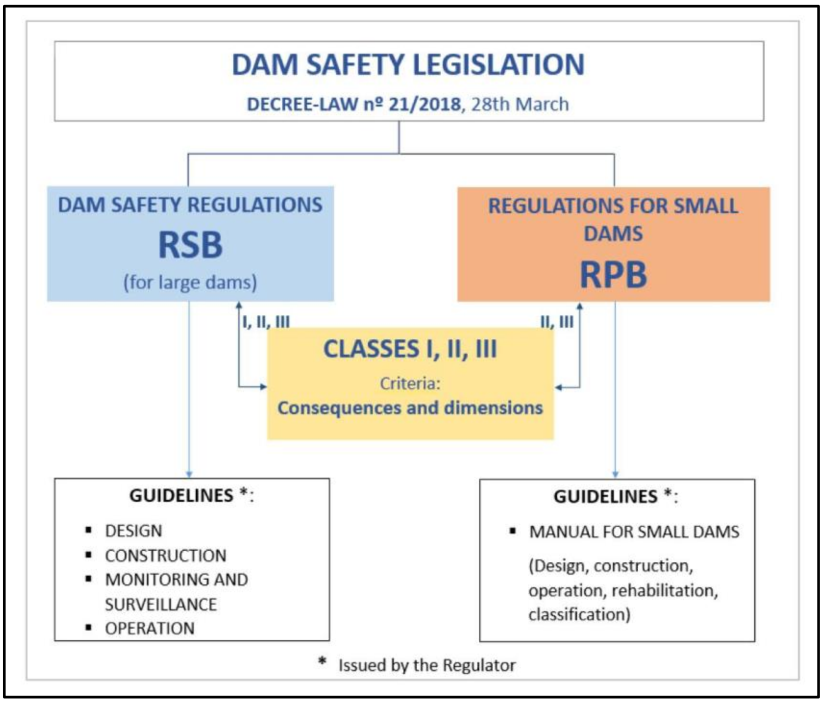


Fig. 2 - Legislation and national regulatory framework for dam safety (ICOLD, 2020).

This regulatory framework ensures that all dams in Portugal, regardless of size, are subject to safety standards designed to protect public safety and environmental integrity. The classification system

[4] https://apambiente.pt/prevencao-e-gestao-de-riscos/barragens-de-portugal, website visited 22/05/2024.
[5] https://apambiente.pt/prevencao-e-gestao-de-riscos/legislacao-e-guias, website visited 22/05/2024.

helps to define the level of oversight and requirements needed for each dam, providing a structured approach to dam safety.

Safety control must be conducted at three levels, encompassing both individual and overall control across all the components that make up the dam-foundation-reservoir system. Safety control must be checked in the following aspects:

Environmental Safety

This aspect is crucial for protecting the natural environment and nearby communities. Key considerations include (APA, 2018):

- Ecological impact: Assessing how the dam affects aquatic ecosystems, flora, and fauna. This involves considering habitat alterations, water quality, and biodiversity.
- Waste management: Properly handling construction materials and waste generated during dam operation.
- Regulatory compliance: Complying with environmental regulations and obtaining necessary permits before construction.

Structural Safety

The structural safety control of dams is carried out in a planned manner, supported by an observation plan. In Portugal, the observation plan is a binding document mandated by the Regulation of Dam Safety (RSB). This regulation is obligatory for both new and existing structures and primarily aims at ensuring safety control throughout the different phases of a dam's lifecycle. According to this, the safety control process spans the following phases (Mata, 2007):

- **Project Phase:** This initial phase involves verifying safety for both normal and failure scenarios. It includes the definition of preliminary observation plans and the initial reservoir filling plan.
- **Construction Phase:** During this period, the observation system is installed, material properties are characterized, and safety control is conducted for scenarios specific to the construction period.
- **Reservoir First Filling Phase:** This phase represents the first load test of the dam-foundation system, requiring a specific observation plan due to its critical importance.
- **Operational Phase:** This phase is divided into an initial period (typically the first 5 years), during which the structure adapts to operating conditions, followed by a normal operation period where most of the analyses are concentrated.
- **Demolition Phase:** This final phase involves the procedures for safely decommissioning and demolishing the dam.

One of the main elements that the Observation Plan must include is the definition of the observation system. This system allows for the measurement of variables representing actions, structural properties, and structural responses throughout the dam's lifecycle (Pedro, 2001).

A correct definition of the observation system and its use during the operational phase allows for several important functions (Portela, 1999):

- **Characterize Structural Performance:** The observation system enables the characterization of the structure's performance in response to permanent, variable, and cyclic loads.
- **Verify Structural Integrity:** It allows for the verification that the dam maintains satisfactory performance over time.
- **Facilitate Maintenance and Repairs:** The system provides the means to monitor and accompany any repair and rehabilitation measures necessitated by structural deterioration. This includes controlling the structure’s behaviour during the implementation of such measures to prevent further damage or significant modifications to the structural properties that could adversely affect the dam’s performance.
- **Suggest Operational Changes:** It can also be used to recommend changes to the operational regime based on observed data and performance metrics.

By implementing a robust observation system, it is possible to ensure ongoing monitoring and maintenance of the dam, thereby enhancing its safety and operational efficiency throughout its lifecycle.

Structural dam safety control activities consist of decision-making throughout the various phases of the project's life, based on observation activities. Thus, the main objective of dam observation activities is the multiple evaluation of behaviour simulation models, measurements carried out, and parameters that characterize and condition this behaviour, to draw up and justify a judgment on safety. What is at stake is the real behaviour of the structure, in the real conditions in which it is operated, with a view to the possibility of detecting possible anomalies.

Hydraulic/Operational Safety

This aspect relates to managing water flow and operating the concrete dam. Key considerations include:

- Gates and valves: Ensuring that gates and valves function correctly to regulate water flow.
- Flood control: Evaluating the dam's capacity to control floods and release water safely.
- Water levels: Monitoring water levels and adjusting operations based on weather conditions and water demand.
- Emergency plan: Having a clear action plan in case of failure or emergency.

By addressing these safety aspects comprehensively, dam operators can ensure that their structures remain safe, reliable, and compliant with regulations, thereby protecting both the environment and nearby communities.

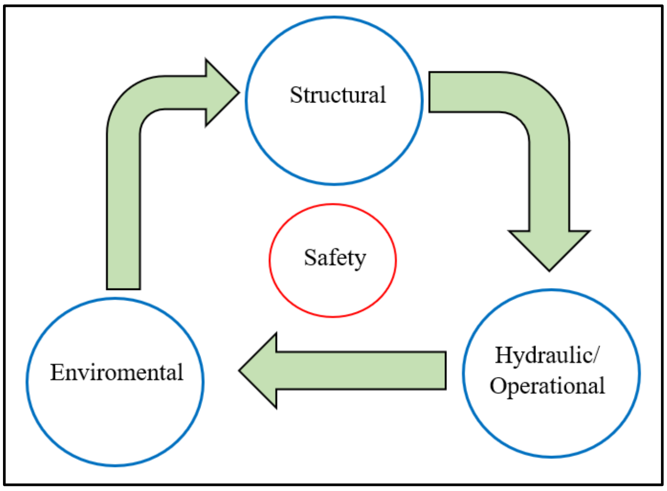


Fig. 3 - Dam safety components adapted from (Silva, 2022).

2.4. STRUCTURAL DAM BEHAVIOUR MONITORING

The structural behaviour of dams is unique for each project because the characteristics of the reservoir dam-foundation system are distinct. The materials used in a dam are specified according to the environmental context, the available materials in the region, and the demands that the structure will face. The foundation has its specificities, whether it's a more or less heterogeneous rock mass or the geological conditions that compose it. Consequently, characterizing the mechanical and hydraulic properties of the dam's materials and its foundation is crucial to estimate the key parameters that define the dam-foundation system. The loads on the structure also vary depending on the surrounding conditions, with the most critical ones related to changes in water levels upstream and downstream and temperature variations in air and water (Silva, 2022).

The monitoring systems must be tailored to each project, adhering to established legislation and documentation, to ensure the measurement of physical quantities representing the loads, material properties, and structural responses that characterize and influence the dam's real behaviour during various phases, from construction to first filling and operation, until its demolition (Martins et al., 2012).

The implementation of automated measurement systems aims to improve the accuracy and timeliness of safety monitoring. Automation of measurements, although introduced several years ago in some projects in Portugal, is not yet fully established. The current level of automation is often hindered by technical challenges, including sensor malfunctions and data transmission errors. Measurement errors from automated monitoring systems still occur, and they must be addressed to ensure data reliability. Addressing these issues requires continual technological advancements and quality control processes to ensure that the safety monitoring system functions as intended.

The main loads and quantities representing the structural dam behaviour that is observed through monitoring systems are described below.

2.4.1. OBSERVATION OF MAIN LOADS

Concrete dams are subject to many simultaneous loads throughout their lifespan, with special emphasis on the following:

Self-weight load

Given the low variability in the specific weight of concrete, an estimate of self-weight is made based on the properties of the materials and the dam's geometry (Mata, 2007). Self-weight is due to gravitational force, which the magnitude is determined by the density of the materials comprising the structure. In the case of concrete dams, the primary material is concrete, with a specific weight ranging from γc=23 to 25 [kN/m³], depending on the type of aggregate (Chen and Chen, 2015).

Furthermore, it should be noted that observation results are generally reset upon completion of the construction phase, following the installation of the structure's self-weight. Therefore, these actions are not considered in data-based models, unless there are specific situations where creep effects may significantly contribute over time and in that case should be addressed properly. Creep refers to the tendency of concrete to deform permanently under sustained load, a phenomenon that can alter stress distribution within the structure and consequently affect its response to external loads. Hence, accounting for these time-dependent effects is crucial for accurate long-term structural modelling.

Hydrostatic loads

These refer to mechanical loads resulting from the presence of water in the reservoir. Depending on the area of application, they can be referred to as hydrostatic pressure when applied to upstream and downstream faces, and uplift pressure when resulting from hydraulic seepage through the dam's foundation (Mendes, 2023). Hydrostatic pressure varies over the dam's lifespan according to reservoir water levels, usually measured by limnimeters scales, radar or submersible ultrasonic level sensors, among others. The variation between upstream and downstream water levels depends on reservoir operation regimes and weather conditions. The impact of water on structural safety is significant, potentially leading to mechanical or chemical effects, such as:

- Hydrostatic pressures on dam faces.
- Internal pressures in concrete dam cracks and foundation partition surfaces.
- Forces due to water seepage in the dam's foundation.
- Chemical alterations of foundation materials.
- Erosion of materials and devices within the dam.
- Deterioration of rock masses due to water seepage through discontinuities.

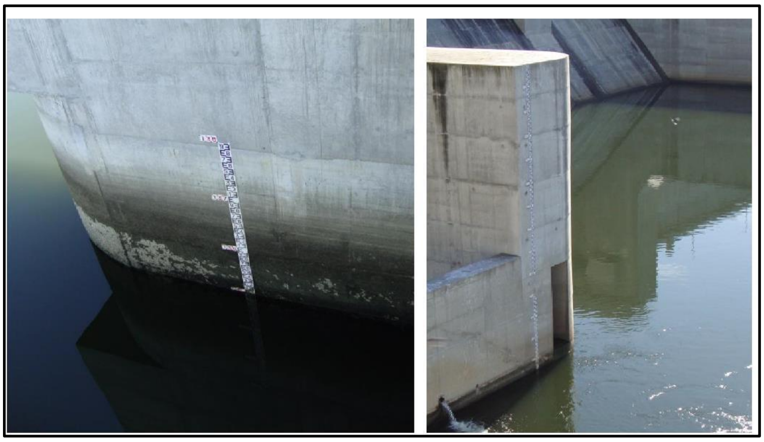


Fig. 4 - Limnimetric scales upstream and downstream of a dam (LNEC, 2020).

Thermal loads

In Portugal, internal temperatures within the dam's body are usually obtained through electric resistance instruments embedded in the concrete, such as electric resistance thermometers, and other electrical devices, such as strain gauges and joint movement meters (Castro, 1998). During dam construction, given the large volumes of concrete, high internal temperatures are reached due to heat released during cement hydration. This heat causes a significant temperature increase in the first few days after pouring, followed by a slow decrease until equilibrium with the external temperature is reached after several months (Batista, 1998).

In the operation phase, air and water temperature variations are the main thermal loads. In this phase, characterizing air temperature variations can be approximated by overlaying two sinusoidal waves, one with an annual period and another with a daily period, being the annual wave the most important for the thermal effect. While the annual thermal wave affects the entire dam, the daily wave affects only a superficial area about 0.5 to 0.8 meters deep along the air-exposed faces (Mata, 2007).

In general, the temperature of reservoir water is hardly influenced by the daily thermal wave of the air, and at greater depths, it is minimally affected by the water level in the reservoir (Silveira, 1961). However, in regulation reservoirs, the temperature of the water varies significantly from the surface, where it is close to the air temperature, down to a depth of about 20 to 30 meters, where it remains constant at greater depths. The phase shift of the annual thermal wave of water temperature, relative to the air temperature, also increases with depth (Mata, 2007).

Dynamic loads

Concrete dams and safety and operating components are subject to dynamic loads of various types, such as earthquakes, explosions, vibrations from plant operation, and discharge systems. Seismic actions can be of tectonic origin or induced by reservoir creation. Generally, concrete dams withstand earthquakes well. However, in cases of high acceleration values, significant damage could occur, potentially affecting the structure's functionality. To characterize seismic actions, especially in zones with seismic risk or when dealing with higher and more structurally significant dams, a network of seismic stations should be installed across the entire dam-foundation-rock mass system, considering the seismic-tectonic characterization of the region (Mata, 2017). Moreover, important reservoirs are also common to install accelerometers across their borders to evaluate the seismic activity related to the formation of the reservoir.

Other type of loads

There are also other types of actions, less common due to their rarity and low probability of occurring in Portugal, such as explosions, ice action (including ice sheet impact on the dam body), and landslides (Silva, 2022a). In addition, for old and deteriorated concrete dams it might be necessary to consider internal volumetric changes due to chemical reactions, for instance, alkali-silica reaction between the concrete components.

A general set of the referred loads is schematically represented shown in Fig. *5*, applied to a gravity dam

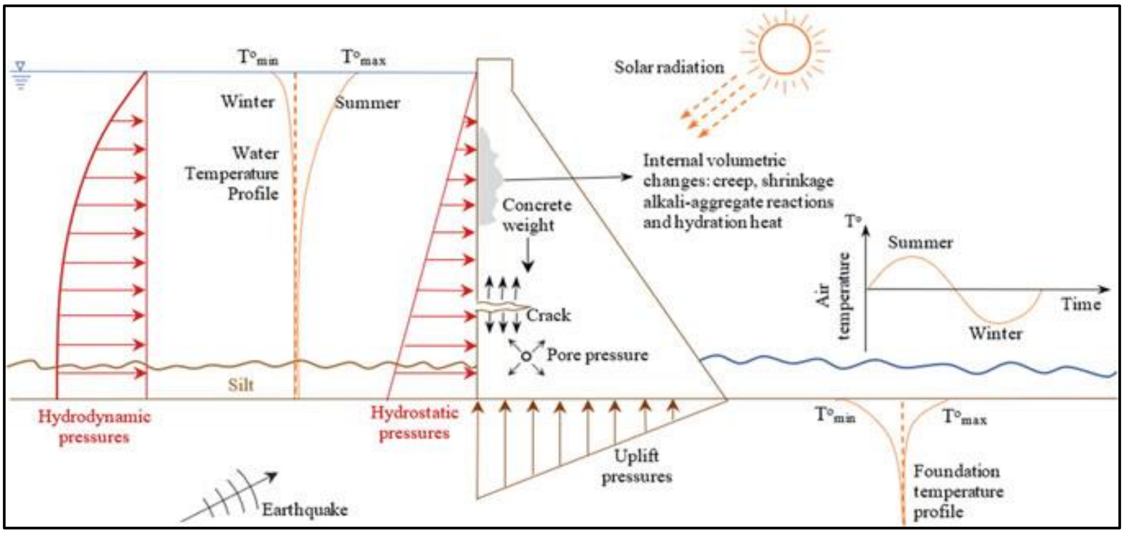


Fig. 5 - Schematic representation of the main loads on a gravity dam. (Sundar, 2023).

2.4.2 OBSERVATION OF THE STRUCTURAL RESPONSE

The observation of large structures is a planned activity where reference values should be established. During the first five years of operation, these values are adjusted to characterize the normal behaviour of the structure, becoming the long-term reference values (Marcelino et al., 2008).

The structural responses commonly observed in concrete dams are addressed in the following paragraphs.

Displacements

Displacements, whether absolute or relative, are usually determined at key points representing the behaviour of the dam, its foundation, and abutments. The measurements are recorded through geodetic methods, plumb lines, and foundation extensometers, among other measuring methods (Mata, 2007). Vertical displacements in the structure can be measured through precise geometric levelling at points located on the dam crest or within galleries. This involves determining differences in height between points connected to the structure and fixed points (or points with known displacements) using appropriate equipment, like precision levels and invar staffs (Gomes, 1986).

Additionally, in the last years, there have been important developments considering the use of GNSS (Global Navigation Satellite System) devices to measure horizontal and vertical displacements, usually in the dam crest. GNSS devices play a crucial role in precisely measuring both horizontal and vertical displacements along the crest of dams. These measurements are critical for detecting any movements in the dam's position, which is essential for maintaining its stability and ensuring safety. By continuously collecting data from GNSS devices, engineers can closely monitor the structural integrity of dams and promptly implement preventive measures at the earliest indication of potential risks[6] (CHCNAV, 2024).

Horizontal displacements are measured using plumb lines, a common method in most large dams. Depending on the importance of the dam, more or fewer bases for optical coordinometers meters and plumb lines, whether straight or inverted, are installed. Straight plumb lines work like a pendulum, fixed at the highest point and suspended by a weight at the other end, keeping the line taut. Inverted plumb lines work the opposite way, with the lower end fixed by injecting cement grout into a deep borehole in the foundation, considered a fixed point, and the upper end floating in a tank with antifreeze water.

Manual measurements are taken with optical coordinometers placed on bases on the structure (Bartholomew and Haverland, 1987; Martins et al., 2012), while automated measurements are usually done with tele coordinometers as shown in Fig. 6.

In the straight plumb lines the displacements obtained at various access points to the plumb line (e.g., in galleries or platforms) are defined relative to the point of attachment at a high elevation as presented in Fig. 6. In the inverted plumb lines relative displacements are measured at accessible points in relation to the deep point in the foundation. If this deep point is considered fixed, the measured displacements are absolute. The installation of plumb lines in the dam body is carried out in areas where the information is most relevant for behaviour assessment. (Martins et al., 2011; Mata, 2007).

[6] https://www.linkedin.com/pulse/enhanced-dam-safety-use-gnss-technology-imsqc, website visited 17/07/2024.

Fig. 6 - Plumb line and optical coordinometer (left) and tele-coordinometer (right), taken from (LNEC, 2020).

Joint movements

Cracks are random discontinuities, whereas joints are predefined discontinuities. Almost all concrete structures develop cracks over time, while there is no standard scheme for selecting the locations of instruments to measure cracks, they can be installed based on the anticipation and assumptions of crack formation made during the design phase. To monitor crack movement, measurements typically focus on the crack width and/or the relative slip. This movement can be measured by installing crack meters or joint meters across the crack. These instruments provide essential data for understanding the behaviour of the structure and assessing its integrity (Mathur, 2017).

In addition, relative movement is observed at contraction joints with the help of instrumentation embedded in the concrete or at the surface. Inside the concrete, relative displacements between blocks are obtained from embedded electric resistance joint movement meters installed during construction (Mata, 2007).

As for surface instrumentation over the last years, the measurement of relative movements between blocks in three-dimensional deformeter bases has increased. The traditional (2 dimensions) deformeter base has been less used in new dams because they only allow for the measurement of two physical quantities (opening-closing movement and only one of vertical and horizontal relative displacement between blocks) and the reading devices are sensitive and expensive (Mata, 2013).

Fig. 7 - Strain gauge bases - Flat rosette with bases (left), and use of the strain gauge (right) (Mata,2007).

Stresses and strains in concrete

Stresses in a dam are caused by static loads such as the weight of the dam, earth, and water loadings. These locations might include the dam body, the interface between the dam and its foundation, or between specific components of the dam. It is crucial to understand the behaviour of concrete dams and their foundations. Direct measurement of stress can be achieved by installing various devices to measure the static total pressure in a dam. In addition, the measurement of stresses can be indirectly performed using electrical resistance strain gauges. Electrical resistance strain gauges work by measuring changes in electrical resistance that occur when the material deforms under stress.

By using tools and techniques, engineers can obtain accurate data on the stress states within the dam, helping to ensure the structure's safety and integrity while also informing future design enhancements. Electrical resistance sensors shown in Fig. 8, measure deformations in concrete by detecting changes in electrical resistance. These sensors are useful for monitoring small deformations in the structure, providing continuous data on the behavior of the concrete under loading conditions.

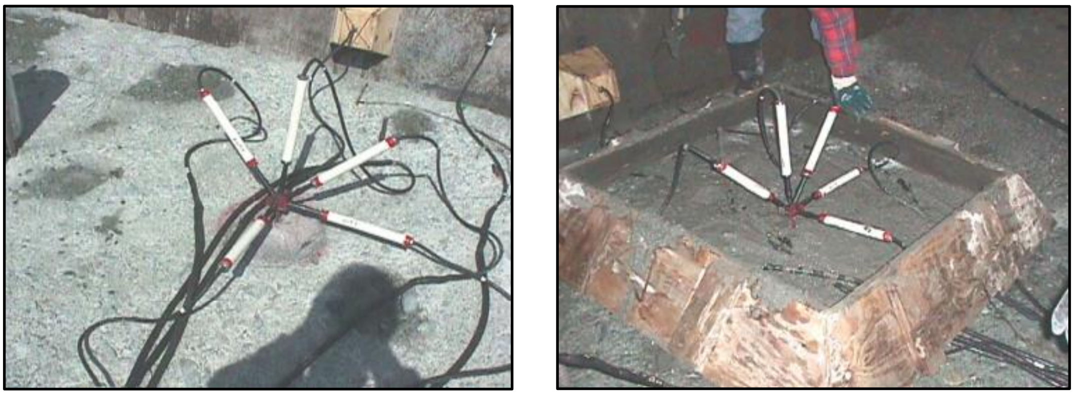


Fig. 8 - Placing a flat group of strain gauges (Moura, 2005).

Vibrating wire piezometers are another technique to measure water pressure through the deformation of an internal diaphragm. The deflection of this diaphragm is measured by a vibrating wire sensor installed perpendicularly to the diaphragm's plane. The height of the water column is added to the installation elevation, providing the piezometric level at the point. Currently, vibrating wire piezometers are widely used in dam monitoring due to their precision, sensitivity, ability to be read remotely, and integration with automatic data acquisition systems. Furthermore, wire sensors are especially valued for their durability and ability to operate in extreme environmental conditions, making them ideal for applications in concrete dams[7] (GEOTHRA, 2024).

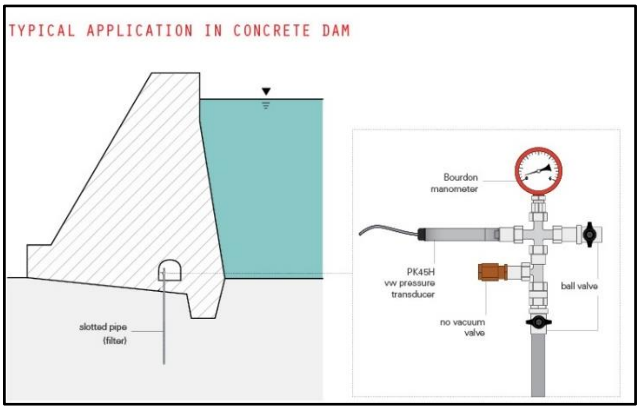


Fig. 9 - Application of vibrating wire sensor in a concrete dam (Mikkelsen, 2003).

Seepage and leakage

The water-retention ability of any dam is crucial not only for conserving valuable stored water but also for monitoring any threats to the dam's stability caused by seepage or leakage. Seepage refers to the interstitial movement of water through the dam, its foundation, or its abutments. In contrast, leakage is the flow of water through caverns and/or cracks. Both seepage and leakage must be controlled to prevent internal erosion, which can lead to undermining or piping in embankment dams and loss of material strength or density in concrete and masonry dams (Adamo et al., 2020).

To measure seepage and leakage, engineers commonly use weirs, Parshall flumes, and calibrated containers. This is done through a drainage network where pipes are installed in boreholes drilled into the foundation from drainage galleries, collecting water that seeps through the rock mass. The flow rate is calculated by measuring the time it takes to fill a reservoir of known volume. The flow is directed to totalizing weirs that allow the quantification of drained and infiltrated flows in predetermined areas (Mendes, 2023).

[7] https://geothra.com.br/instrumentacao-geotecnica/piezometro-de-corda-vibrante/, web visited 17/07/2024.

Fig. 10 - Terminal of a drain in a gallery (Moura, 2005).

Control of drained and infiltrated flows is crucial, as it is one of the key elements of safety control. Combined with piezometers, it allows for evaluating the efficiency of the dam's waterproofing and drainage systems (Mata, 2007).

Accelerations

Vibrations in dams are caused by surrounding environmental conditions, such as earthquakes and small tremors induced by the reservoir's weight, wind action, and nearby or dam crest road traffic, as well as by the normal operation of hydroelectric plants associated with the reservoir (Pereira, 2019). Although no accidents have been recorded due to this type of action, vibration-based monitoring systems are increasingly used in large civil engineering structures to characterize their behaviour over time and identify changes resulting from structural degradation (Mendes, 2023).

Continuous monitoring systems are typically composed of uniaxial and triaxial accelerometers where the most common and used are triaxial accelerometers arranged horizontally along the dam's galleries, allowing for the recording of time series of accelerations and subsequent characterization of vibration levels and modal parameters such as natural frequencies, damping ratios, and vibration modes (Pereira et al., 2018).

Fig. 11 - Accelerometer (Mata,2013).

2.4.3. ANALYSIS AND INTERPRETATION OF THE OBSERVED STRUCTURAL BEHAVIOUR

As previously discussed, dams are considered high-potential risk structures. The critical period in the life of a dam occurs during construction and the first complete filling of the reservoir, extending through a span of 5 to 7 years at the beginning of the dam's normal operational period. During this critical phase, there isn't much data on the dam's behaviour, since it coincides with the construction phase and the early stages of operation. This is why deterministic and statistical models are used, based on the limited measurement data available, to try to characterize the observed behaviour.

Validation of the results is done by comparing measurements with the values predicted by these behavioural models. If there are significant deviations from what the models predict, a thorough investigation is required to determine the causes of these discrepancies. This could lead to the reworking of the model and a consequent re-evaluation of the dam's safety conditions (Mata, 2007).

During this early phase, it's crucial to gather as much information as possible to refine the models and ensure that the dam's operation is within safe limits. Engineers and dam safety professionals must pay close attention to any signs of abnormal behaviour, such as unexpected displacements, stress points, or changes in the flow and pressure of the reservoir. By closely monitoring these variables and comparing them with the model's predictions, it is possible to quickly identify any issues that could compromise the dam's structural integrity or operational safety.

In summary, the initial years after construction are a critical time for a dam's safety and require specific monitoring and analysis. The use of both deterministic and statistical models helps engineers to assess the structural dam behaviour and anticipate potential problems and take corrective actions when necessary. This proactive approach to dam safety is essential for preventing accidents and ensuring the long-term stability and functionality of these crucial structures.

3. CHARACTERIZATION OF OBSERVED STRUCTURAL BEHAVIOUR AIDED BY DATA-BASED MODELS

3.1. INTRODUCTION

As previously mentioned, dams are considered structures with high potential risk. The critical period in the life of a dam occurs during the construction phase and the first complete filling of the reservoir, extending through a span of 5 to 7 years during the initial phase of the dam's normal operation. During this critical phase, there's limited data on the structure's behaviour, since it coincides with the construction phase and the beginning of the dam's operation. Consequently, the data is limited during the filling of the reservoir.

One of the main issues with data-based models during the initial phases of the dam's life is not only the consideration of effects related to the construction, such as the release of hydration heat. More importantly, there is not enough historical data, meaning insufficient training data.

When enough data is available, one of the methods used is the Effect Separation Method, which is employed to analyze results obtained through monitoring of parameters that impact the dam. As the name suggests, this method allows you to analyze the data for a given response over time, separating the effects on that parameter due to various loads, given that the variation of these loads over time is known. This model establishes semi-empirical functional relationships between the responses and the variables causing them (Morais, 2017). Statistical models used to predict the structural response are based on the relationships between loads and the structure's response. These statistical models, also known as quantitative interpretation models, are based on building multiple linear regression models from monitoring data obtained during dam operations (Mendes, 2023).

3.2. INTERPRETATION OF STRUCTURAL BEHAVIOUR BASED ON HST AND HTT APPROACHES

3.2.1. THE HST APPROACH

The most common approach for developing quick models in the field of dam engineering is called HST (Hydrostatic, Seasonal, Time). This approach is based on the hypothesis that the variable under study, such as horizontal displacements, depends on a combination of hydrostatic load effects, seasonal temperature variations represented in a simplified manner through sinusoidal functions, and the effect of time (Silva et al, 2022b).

This statistical method is only applicable if there are enough number of observations and the generated functions can predict the behaviour of the variable in question. The methods aim to approximate effects associated with a limited period at a specific point using the equation (1):

$$\delta(h_i, S_i, t_i) = U_i(h_i) + U_s(S_i) + U_t(t_i) + k + \varepsilon_i \qquad \text{Eq. 1}$$

where:

- $\delta(h_i, S_i, t_i)$ - The observed value of the variable under analysis in observation i, which depends on hydrostatic pressure, temperature, and the point in time when the observation is made.
- $U_i(h_i), U_s(S_i), U_t(t_i)$ - Components of the variable that correspond to the elastic effect of the water level in the reservoir, the elastic effect of seasonal temperature variations, and the effect of time in the i^{th} observation.
- k - A constant that corresponds to the difference between observed and calculated values at the beginning of the calibration period.
- ε_i - The residual of the i^{th} observation, given by the difference between the estimated value and the observed value.

The effect of the water level in the reservoir

The effects of the water level in the reservoir can be represented by a polynomial function. The approximation can be made using the following equation (2):

$$U_h(h) = a_1 * h^4 + a_2 * h^3 + a_3 * h^2 + a_4 * h \qquad \text{Eq. 2}$$

where:

h - Height of the water in the reservoir.

a_1, a_2, a_3, a_4 – Coefficients to adjust.

This approach underscores the importance of considering historical data and the cumulative impact of water level fluctuations over time.

The temperature effect

In the normal operation phase, the temperature effects are related to variations in air and reservoir water temperatures and their impact on the thermal field of the structure. In the HST (hydrostatic, seasonal, time) approach, the effect of temperature change can be considered as a proportional attenuation of air temperature changes, with a phase shift that depends on the depth along the analyzed section. Straightforward data-based models typically do not use temperature measurements, as it is assumed that the thermal effect U_θ can be represented by a sum of sinusoidal functions. Thus, the effect of temperature variations is defined by a linear combination of sinusoidal functions, depending only on the day of the year (Mata et al., 2023).

The annual heat wave is represented by:

$$U_\theta(\theta_1)^{annual} = b_1 * cos(\theta_1) + b_2 * sin(\theta_1) + b_3 * sin^2(\theta_1) + b_4 * cos(\theta_1) * sin(\theta_1) \qquad \text{Eq. 3}$$

with:

$$\theta_1 = \frac{2*\pi*t_d}{365} 1 \leq t_d \leq 365 \qquad \text{Eq. 4}$$

where:

t_d - Number of days elapsed from the beginning of the year to the date of observation.

b_1, b_2 – Coefficients to adjust.

The time effect

The effect of time on the structure is an irreversible component associated with the effects of inelastic actions, such as creep and/or concrete stress relaxation, as well as phenomena related to deterioration. Under normal operating conditions, where variations in primary stresses are well known, a significant portion of the 'effects of time' can be attributed to the delayed behaviour of concrete (Mata, 2007).

The combination of polynomial functions is often used in the context of this component, with various ways to write this equation (5), depending on the author and the phenomena under study.

$$U_t(t) = c_1 * t^3 + c_2 * t^2 c_3 * t + c_4 * ln * (1 + \frac{t}{a}) \qquad \text{Eq. 5}$$

where:

t - Number of days between the observation campaign and the start of the monitoring.

a - Number of days between the first filling and the date of the start of the analysis.

c_1, c_2, c_3, c_4 – Coefficients to be adjusted.

3.2.2. THE HTT APPROACH

The HTT (Hydrostatic, Temperature, Time) approach differs from the HST approach in the component related to the thermal effect. In this case, the HTT approach represents the thermal effect through information recorded in thermometers embedded in the concrete dam body. These instruments are connected to control panels where the data can be monitored or may even have technology for automatic data collection. These thermometers should be placed along the dam in groups distributed across its profile to measure temperature variations within the same cross-sectional plane. Measurements from these devices start as soon as they are installed, aiding in the characterization and evaluation of the concrete temperature during construction. The most commonly used instruments are electrical resistance thermometers (Silva, 2022).

Both HST and HTT approaches are based on the principle of effect separation, although some correlation between the variables is acknowledged (Silva et al, 2022b). When compared to HST models, HTT models may show a higher determination coefficient R^2 and lower standard deviations of the residuals[8] σ_ε (Mata et al., 2014).

In this sense, a better performance based on the HTT models approach than HST models is usually expected. However, additional data (temperatures in the dam body) is required to perform HTT models. In this work, the main focus was on the analysis of the performance of the HST models as it was referred in the thesis outline.

3.3. MULTIPLE LINEAR REGRESSION MODELS

The primary goal of Multiple Linear Regression (MLR) models is to predict response values (dependent variables) based on a set of predictor data (independent variables) (Johnson & Wichern, 2007). Initially, the model can be used to understand the relationship between predictor variables and the dependent variable. Once this relationship is established, the model can be used to predict dependent data that is not readily available by manipulating the predictor values.

Multiple Linear Regression allows for the consideration of several predictor variables simultaneously, giving a more comprehensive picture of the factors influencing the dependent variable.

In the context of dam monitoring, certain loads trigger responses from the reservoir dam-foundation system, as mentioned throughout this document. The set of actions that stand out results from variations in environmental and operational conditions, such as the water level in the reservoir, temperature, and others.

If it can be confirmed that: i) there were no significant structural changes during the period under analysis; and ii) the structure operates with low stress levels, displaying elastic, linear, and reversible behaviour, then the principle of superposition can be considered valid. This allows for simplification in analysis by using models built with the Multiple Linear Regression method.

[8] Residuals are the difference between the observed values and the predicted values

In summary, these models aim to characterize the observed behaviour pattern based on the history of the structure's observations, which translates into a statistical relationship between key environmental variables (such as the water level in the reservoir and temperature variations), the effect of time, and the observed quantities (Mata, 2007).

There are various options for selecting the independent variables used in constructing quantitative interpretation models (Castro, 1998), each with its characteristics, tailored to the needs of each case study.

Quantitative interpretation models are statistical models in which the user leverages knowledge of the structure and its behaviour to define the independent terms used to establish the relationship between each solicitation's variation and the corresponding response. This relationship is created through parametric analytical expressions containing functions associated with coefficients determined by a statistical relationship. The regression coefficients and the constant are traditionally obtained using the least squares method (Mata, 2007).

3.4. MACHINE LEARNING MODELS

The role of artificial intelligence in society has grown increasingly significant, helping to address current and future needs in the development and production of goods and services. When a particular field has a structure capable of collecting data, it can be the catalyst for applying artificial intelligence methodologies and technologies to address its challenges. These data are managed, processed, and interpreted with the help of Data Science, where data engineers apply various algorithms to solve specific problems. Machine Learning emerges as a tool that uses these data to create algorithms based on predefined rules. These algorithms lead to models like pattern recognition and others. Scenarios are developed that require responses. Machine Learning techniques apply the concept of artificial intelligence, whose fundamental idea is to give machines the capabilities to assist specialists in making more informed decisions. Fig. 12 illustrates the various levels at which artificial intelligence can be organized.

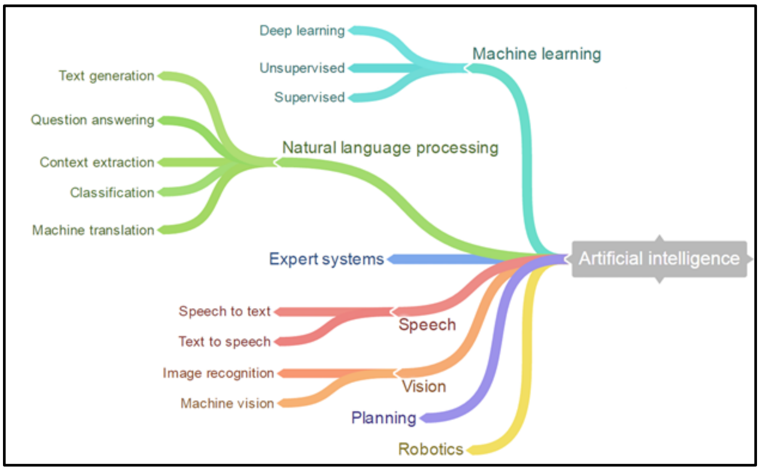


Fig. 12 - Branches of artificial intelligence (Sysoi, 2017).

These Machine Learning models, developed for each type of problem, are built on algorithms that use data from a previously collected and processed sample to make predictions for new data.

Machine Learning methods are used in programs and applications to enable them to learn from collected data and identify patterns, allowing them to make decisions with minimal human intervention while having the capacity to adapt and improve behaviour and responses.

This leads the algorithm to iteratively adjust its outputs for better performance. Although Machine Learning algorithms are not a new concept, the ability to automatically perform complex mathematical calculations with reliable and repeatable results, at increasingly faster speeds, is a recent advancement in the field. Machine Learning can be broadly divided into two major types: supervised and unsupervised learning.

3.4.1. SUPERVISED LEARNING

Supervised learning is a Machine Learning task that involves learning from examples. In this type of learning, a learner (typically a computer software) receives a training dataset and a testing dataset. The goal is for the learner to 'learn' from the labeled examples in the training set to correctly identify the unlabeled examples in the test set. The training set consists of n ordered pairs $(x_1, y_{1)}, (x_2, y_2), \dots, (x_n, y_n)$, where each x represents a measurement or a set of measurements at a single data point, and y is the corresponding label for that data point.

In general, supervised learning uses algorithms that, during the training phase, track all inputs that lead to desired outputs. In this type of learning, both the inputs and the outputs used in the learning process are known. Regarding its application, supervised learning can be divided into regression and classification models (Ayodele, 2010).

Classification is the process of identifying the category or class of an observation through a supervised learning method, where the categories of the objects are already known for training and testing purposes. After completing the training process for a given classification task, the algorithm can assign a category to a newly observed object (Lee, Shin, 2020).

When machine learning is applied to the linear regression model, this involves training algorithms to find the best solution (or hyperplane in higher dimensions) that fits the observed data, thereby enabling the prediction of future values based on linear relationships between variables (Hastie, 2009).

Functionality:

Initially, data containing the variable to be predicted (the dependent variable) and the variables believed to influence it (the independent variables) are gathered. Subsequently, these data are prepared for analysis, cleaning them of errors and ensuring the data consist of a high quality to be based on. Once the data is ready, they are divided into two sets: one for training the model and another for testing it. During training, the model adjusts the best-fit line using the training set. This line is adjusted to minimize the difference between the model's predictions and the actual values in the training set. After training, the model's performance is evaluated using the test set. This indicates how well the model generalizes to unseen data. Metrics such as root mean squared error (RMSE) and determination coefficient (R^2) help to measure the model's accuracy. Once the model is trained and evaluated, it can be used to make predictions on new data (Hastie, 2009).

Neural Network:

Neural networks are important example of supervised learning techniques. Neural networks consist of interconnected layers of nodes (neurons) that process inputs through multiple layers to learn complex patterns in data. They are particularly useful for tasks that involve non-linear relationships and can handle large datasets effectively. Neural networks are often used in applications such as image, speech recognition, and learn data where they excel at identifying intricate patterns and features that traditional methods might miss. The training of neural networks involves adjusting weights and biases through techniques such as backpropagation to minimize the difference between predicted and actual outcomes[9].

Important Considerations:

It is essential to carefully select the independent variables that will be used to predict the dependent variable. This may require prior knowledge of the problem domain. Additionally, it is important to consider the possibility of overfitting, where the model fits too closely to the training data and does not generalize well to new data. Techniques such as validations, adjustment of weights and regularization can help mitigate this issue.

[9] https://scikit-learn.org/stable/modules/neural_networks_supervised.html, web visited 17/07/2024.

3.4.2. UNSUPERVISED LEARNING

In unsupervised learning, the model learns patterns from the input data while the output data are unknown. The algorithm clusters the input data into groups or creates clusters to classify/identify them, and when a new set of input data is provided, the model finds the best group to which this new set belongs to produce an output (Haykin, 2009).

Unsupervised learning aims to identify similarities in the data to group or organize them accordingly. Clustering tools are used to group sets of objects based on similarities in a multidimensional space. Objects in the same cluster are more similar to each other than those in different clusters.

Clustering is considered unsupervised Machine Learning because class labels for the objects are not known beforehand. In clustering, objects are grouped in an n-dimensional space, and the similarity between two objects is measured using a similarity distance function (Lee, Shin, 2020).

3.4.3 MULTILAYER PERCEPTRON NEURAL NETWORK

A Multilayer Perceptron Neural Network model (MLP-NN) is a type of feedforward neural network in which all nodes are interconnected across different layers. The concept of the perceptron, a fundamental unit of an artificial neural network, was first introduced by Frank Rosenblatt. A perceptron is a supervised learning algorithm that includes node values, activation functions, inputs, and weights to compute outputs (Banoula, 2023).

The MLP-NN model operates only in the forward direction, meaning each node passes its value to the next node in the forward direction only. The backpropagation algorithm is used to enhance the training model's accuracy by adjusting the weights based on the error between the predicted and actual outputs[10].

The general architecture of a neural network consists of three main layers: input layer, hidden layer, and the output layer. The input layer is where the network first receives data. The number of nodes in this layer corresponds to the number of features or variables in the dataset. Each node in the input layer represents a distinct feature of the input data.

The hidden layers are where the network performs its computations and learns to transform the input data. The complexity of the relationships in the input data determines the number of hidden layers required. Some models, a single hidden layer may be enough, whereas more complex models may require multiple hidden layers to effectively capture patterns in the data

Finally, the output layer provides the final prediction or classification. After processing through the hidden layers, the network's output is presented in this layer.

[10] https://www.simplilearn.com/tutorials/deep-learning-tutorial/multilayer-perceptron website visited 23/05/2024.

Fig. 13 provides a detailed graphical representation of a neural network architecture, illustrating how data flows through each layer and how each component contributes to the network's learning and prediction processes (Banoula, 2023):

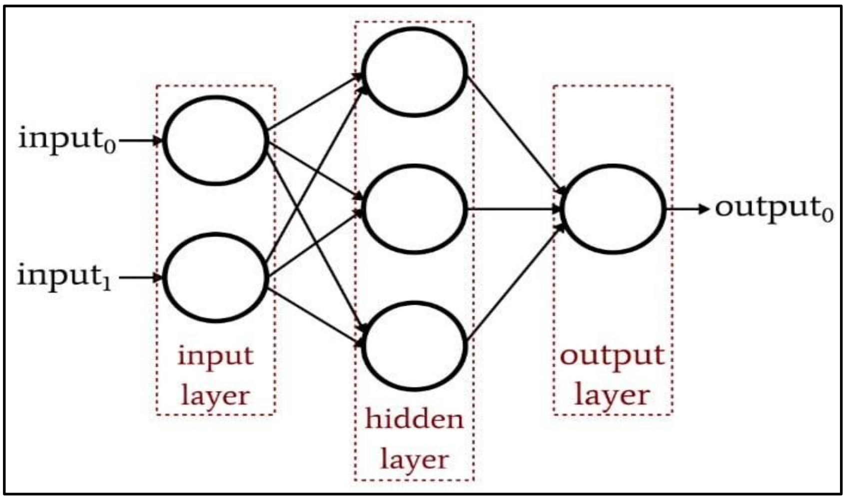


Fig. 13 - Multilayer perceptron architecture[11] (Keim, 2019).

Input layer: It is the initial or starting layer of the Multilayer perceptron. It takes input from the training data set and forwards it to the hidden layer. There are n input nodes in the input layer. The number of input nodes depends on the number of dataset features. Each input vector variable is distributed to each of the nodes of the hidden layer.

Hidden layer: It is the heart of all MLP-NN. This layer comprises most of all computations of the neural network. The edges of the hidden layer have weights multiplied by the node values. This layer uses the activation function. Several hidden layer nodes should be accurate as few nodes in the hidden layer make the model unable to work efficiently with complex data. More nodes may result in an overfitting problem.

Output layer: This layer gives the estimated output of the MLP-NN. The number of nodes in the output layer depends on the type of problem. For a single targeted variable, only one node is used. For N classification problems, N nodes are used in the output layer.

Advantages of MLP-NN models:

- Effectively handles non-linear problems.
- Suitable for complex problems with large datasets.
- Useful for solving fitness problems in neural networks.
- High accuracy and reduced prediction error via backpropagation.
- Quick prediction after model training.

[11] https://www.allaboutcircuits.com/technical-articles/how-to-train-a-multilayer-perceptron-neural-network/ website visited 23/05/2024.

Disadvantages of MLP-NN models:

- Performance depends heavily on the quality of training.
- Tight connections can lead to increased parameters and node redundancy.

In the present work, MLP-NN models have been employed to predict the dynamic response of a dam based on the loads it experiences.

The neural network model and consequently the architecture of the model has to be taken into account, as this is quite important to understand how the software performs the learning. certain aspects were taken into account which are mentioned below.

In the neural network, the weights were used to calculate the output of each neuron in a given layer. For an input layer, the weights represent the influence of each input feature in the neural network. In the hidden and output layers, the weights determine how the outputs of the neurons in the previous layer were combined to produce the outputs of the current layer (Leung, 2008).

During the training set, weights and biases of the connections between neurons in each layer are updated in order to minimize the loss function, which represents the difference between the model predictions and the actual values in the training set. In addition, the weights are important parameters that determine the strength and direction of the influence that one neuron has on another in the neural network model.

Bias helps the neural network to avoid linearity problems by allowing a neuron to become active even if the weighted combination of inputs is zero. With biasing, the model is better able to represent nonlinear functions, which is essential for solving complex problems.

The “s” form of the sigmoid function is a crucial component in the architecture of neural networks, transforming any input into a value between 0 and 1. This transformation is essential as it introduces nonlinearity into the neural network, enabling the model to learn complex, nonlinear relationships within the data. The sigmoid function is characterized by its smooth curve and well-defined slope, which significantly aids in the backpropagation process during the training phase of the neural network.

One of the primary applications of the sigmoid function is within the hidden layers of neural networks. By introducing nonlinearity, it allows these layers to capture and model intricate patterns and dependencies that would be unattainable with linear functions alone. Furthermore, the sigmoid function is commonly employed in the output layer of neural networks designed for binary classification problems. Since the output of the sigmoid function ranges from 0 to 1, it can be conveniently interpreted as a probability, making it ideal for scenarios where the model needs to distinguish between two classes.

In addition to activation functions like the sigmoid, biases play a pivotal role in the functionality of neural networks. Biases are additional values added to the input of each neuron, providing the network with greater flexibility and enhancing its ability to learn nonlinear relationships. When used in conjunction with activation functions such as the sigmoid, biases enable the neural network to better adapt to the underlying complexities of the data, ultimately leading to more accurate and reliable predictions (Jurafsky, 2024).

By incorporating both biases and activation functions like the sigmoid, neural networks are equipped to handle a wide range of tasks, from simple classifications to more sophisticated pattern recognition and predictive modelling. These elements work together to ensure that the neural network can effectively learn and generalize from the training data, making them important tools in the field of machine learning and artificial intelligence.

3.5. OPERATIONAL MODAL ANALYSIS

3.5.1. INTRODUCTION

The ageing and deterioration of structures have driven the development of accurate vibration-based damage detection techniques, supported by structural health monitoring systems (Cunha & Caetano, 2006), which can play a critical role in managing maintenance programs for civil engineering structures, improving their durability, safety, and operational standards. This progress is largely due to the emergence of new transducers and analogue-to-digital converters, allowing for the identification of a structure's modal properties without relying on any additional excitation besides the natural vibrations caused by ambient and operational conditions.

Operational Modal Analysis (OMA) is an advanced technique in structural engineering used to identify the dynamic properties of a structure without the need for controlled external excitations. Instead of relying on forced vibration tests, OMA harnesses the natural vibrations of the structure occurring under normal operational conditions or due to environmental stimuli (Brincker, 2015).

The operational modal analysis process involves installing vibration sensors, such as accelerometers, on the structure to record dynamic responses in the form of accelerations, velocities, or displacements. These data can be acquired over extended periods and processed using advanced signal analysis algorithms to extract modal properties of the structure, such as natural frequencies, mode shapes, and damping ratios. The predicted values are then compared to those obtained directly from modal identification algorithms (Brincker, 2015). This method is applied to a real-world scenario in chapter 4, with the objective of isolating the effects of reservoir water level, temperature, and concrete hardening over time on a dam's natural frequencies. Furthermore, it is used again to predict the optimal reference value for tracking a specific vibration mode in unseen test data

One of the key advantages of OMA is its ability to perform non-invasive and real-time assessments of the dynamic behaviour of structures under actual operating conditions. This enables engineers to continuously monitor the structural integrity of facilities and detect any significant changes in their modal properties that may indicate potential safety issues or structural degradation (Magalhães, 2010).

In summary, the use of operational modal analysis (OMA) is a valuable tool in structural engineering, particularly for large-scale structures, as it allows for the identification of dynamic properties using natural vibrations induced by normal operational conditions. This non-invasive approach provides critical information about the structural health of structures and facilitates the detection of any potential deterioration or damage, thereby improving maintenance planning and structural safety.

3.5.2. PEAK PICKING METHOD

The Peak Picking is a simple and straightforward identification method, and the first one developed to perform the dynamic identification of structures based on their responses to ambient vibration. It is still commonly used for an initial assessment of the data since it can quickly provide good results. Though some previous works had already been developed, (Bendat & Piersol, 1980) are considered responsible for the current methodology.

In addition, considering time series measured in situ commonly based on accelerations this method basically consists of locating peaks within a power spectral density function (PSD), associated with the resonant frequencies of the structure.

An introduction to modal identification using the peak picking method is presented as follows, using experimental data recorded at Baixo Sabor dam. It is important to note that a set of time series of accelerations with a frequency sampling rate of 50 [Hz] was considered for this application, which means capturing 50 points per second. This was done for a period of 30 minutes, with accelerometers placed at different points along the dam, specifically 12 locations (the dynamic monitoring structure system is presented in Fig. 19 showing the locations of the accelerometers throughout the dam. This resulted in a time series of 90,000 data points per channel (90,000 x 12 channels).

During this period, the accelerometers captured the vibrations of the structure due to ambient and operational excitation. It's important to note that measurements will never be exactly the same from one day to the next. This is because, even under similar conditions, the dynamic and changing nature of the environment and operational influences on the dam cause variability in the results. Even if conditions seem identical, the measurements will always differ due to these constant changes.

The Peak-Picking method is used to analyse the accelerations and to identify the most relevant natural frequencies. Once the time series data were already obtained, the methodology applied involved breaking the time series into several windows here denoted as K, whose sizes were tested, and the final results were reached through the averaging of the results across all windows. Defining a larger window size results in greater frequency resolution but increases the effect of noise on the spectra, making it more difficult to select the correct natural frequency accurately.

Fig. 14 presents the analyzed model, where the Y-axis represents the intensity of the Averaged Normalized Power Spectral Density (ANPSD) function and the X-axis represents the natural frequency [Hz]. The frequencies shown are the most representative natural frequencies. Considering higher order vibration modes implies a greater level of uncertainty. In each figure, a different window size was used.

Moreover, Fig. 14 shows two figures illustrating the effects of different time windows on data sensitivity. The top figure uses a time window of 8,192 data points, equivalent to 163 seconds of recorded samples, resulting in not very smooth lines. In contrast, the bottom figure employs a time window of 16,384 data points, equivalent to approximately 328 seconds or 6 minutes, leading to a more detailed and sensitive graph with more recorded data points but having a figure where the precision increases but the noise also increases. The discrepancy between the two figures highlights how the variation in window length affects the sensitivity of the data.

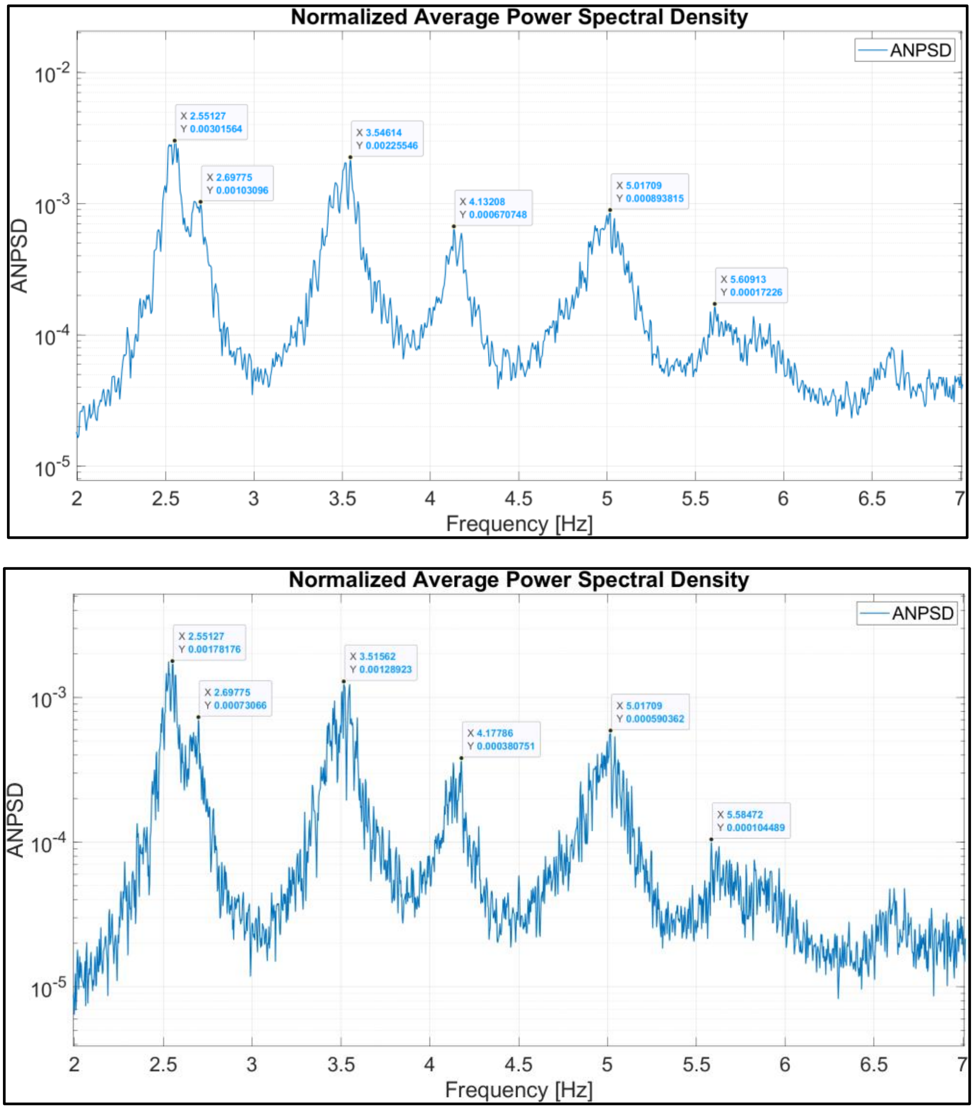


Fig. 14 - Application of Peak Picking method to time series of accelerations measured in Baixo Sabor dam: Frequency vs ANPSD.

In the displayed Fig. 14, the precision of peak values changes as the size of the window is adjusted as it was mentioned before. This results in some peaks having very small differences in value, while others show significant variations.

This leads to a critical observation; with a larger time window, there is an increase in resolution but also greater complexity in analyzing and selecting frequency values. Conversely, with a smaller time window, there is less noise, reduced resolution, and less complexity in identifying and selecting peak frequencies using the "Peak Peaking" method.

These factors should be considered when deciding on the final values and how much precision is needed. When the window size is adjusted, greater precision in the data is also achieved, but it also becomes more challenging to identify natural frequencies. Therefore, finding the optimal window size is crucial to ensuring good precision without compromising the quality of the natural frequencies.

4.

CASE STUDY: BAIXO SABOR DAM

4.1. BAIXO SABOR HYDROELECTRIC POWER PLANT

The Baixo Sabor dam is an important hydroelectric infrastructure located on the Sabor River in northeastern Portugal. This dam is part of a larger project aimed at harnessing the hydroelectric resources in the region. It is situated in the district of Bragança, within the Trás-os-Montes e Alto Douro region, known for its mountainous landscapes and rivers. The Sabor river is one of the main tributaries of the Douro river.

The Baixo Sabor dam construction began in 2008 and subsequently entered into operation in 2016, being the beginning of the first filling in 2015. This dam contributes to Portugal's energy supply, helping to reduce its reliance on non-renewable energy sources. The water retained by the dam is used to spin turbines that generate electricity[12].

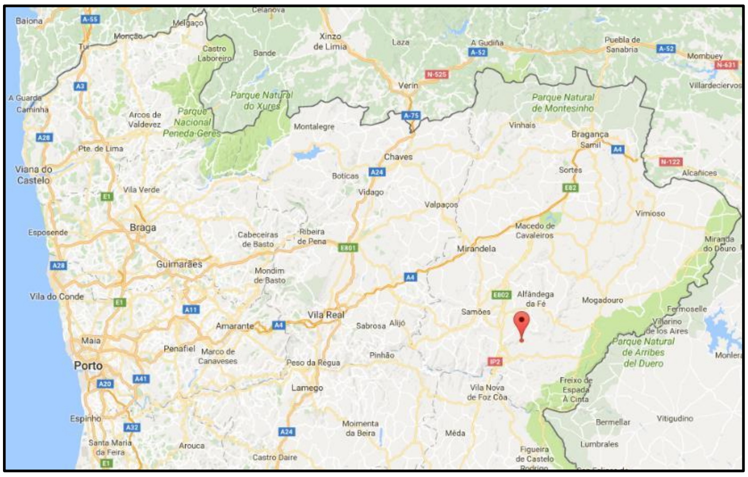


Fig. 15 - Location of Baixo Sabor hydroelectric power plant in the north-east of Portugal (Google Maps).

[12] https://www.power-technology.com/marketdata/baixo-sabor-portugal/ website visited 23/5/2024.

This project included the construction of two dams located about 12.6 km from each other. However, in this work, only the upstream step will be approached. This is constituted by a concrete double-curvature arch dam, embedded in a narrow valley zone. The dam is 123 m height and its crest, with a width of 6 m and 505 m long. Thirty-two concrete vertical blocks, separated by vertical contraction joints, compose the arch, which is crossed by six horizontal visit galleries and one general gallery for drainage that contacts with the foundation all along the upstream side of the structure (Pereira, 2019). Fig. 16 shows the construction of the dam.

Fig. 16 - Aerial view of the Baixo Sabor dam construction phase (Faria, 2022).

The dam is endowed with a spillway composed by four floodgates, 16 m long each. The full storage level is 234 meters above sea level and the turbines in the power plant adjacent to the dam rotate at 214 rpm which is equal to 3.57 Hz (Pereira, 2019). Fig. 17 shows an aerial view of the dam and the reservoir (on the left) and a closer picture of the spillway (on the right). In the aerial photography, it is possible to observe that the powerhouse containing the turbines for electricity production is quite close to the dam, which may influence the results.

Fig. 17 - Aerial view of Baixo Sabor dam and reservoir (left) and closer picture of the spillway (right) (Engie).

4.2. DYNAMIC MONITORING SYSTEM

When monitoring large civil engineering structures, it is advantageous to have digitizers distributed over the monitored structure in order to reduce cable length and thus electrical interferences that can corrupt the signals (Magalhães, Cunha, et al., 2008). Therefore, the dynamic monitoring system installed in the Baixo Sabor arch dam is divided into three subsystems connected by optical fiber. Subsystem 1 and subsystem 2 are composed by 6 uniaxial force balance accelerometers and a digitizer each, and both subsystems are connected to a field processor which is, in turn, linked to a computer (NUC). The two groups of accelerometers are radially installed over the upper gallery and disposed on each side of the spillway. Subsystem 3 is composed of 8 uniaxial force balance accelerometers, radially installed over the second and third upper galleries, and two digitizers that are connected to a computer, which is responsible for running the acquisition. To perform modal analysis a good synchronization of the data recorded by all digitizers must be assured, which required the installation of GPS antennas. In addition, the field computer is connected to the optical fiber network between the dam and the plant allowing remote access (Pereira, 2019). A scheme representing the monitoring system's main elements is presented in Fig. 18, while the position of the total 20 accelerometers is characterized in Fig. 19 by red and blue marks in both a scheme and a picture of the structure.

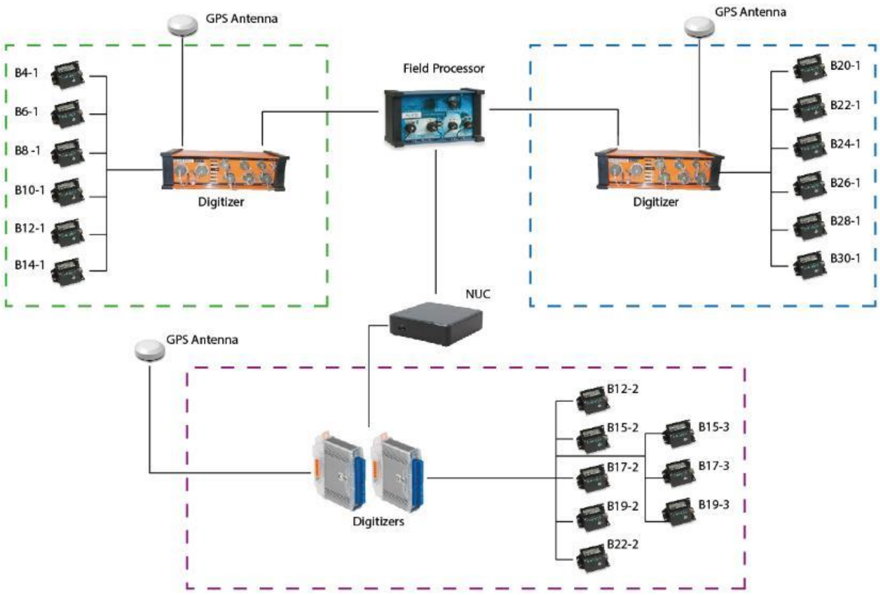


Fig. 18 - Scheme of the dynamic monitoring system installed in Baixo Sabor dam (Pereira, 2018).

The dynamic monitoring system is configured to continuously record acceleration time series with a sampling rate of 50 [Hz] and a duration of 30 minutes at all instrumented points, thus producing 244 groups of time series per day. This data is automatically stored in the field computer and then transferred remotely both to FEUP and LNEC, where they are stored and processed (Pereira, 2019).

The duration at Baixo Sabor dam continuous dynamic monitoring system was defined according to the techniques used in its data processing, and it can be adjusted if it is needed.

Additionally, in order to better characterize eventual seismic actions acting on the dam, a seismic monitoring system has also been implemented. This seismic system, which is working independently from the dynamic monitoring system previously described, is presented in (Pereira Gomes et al., 2018).

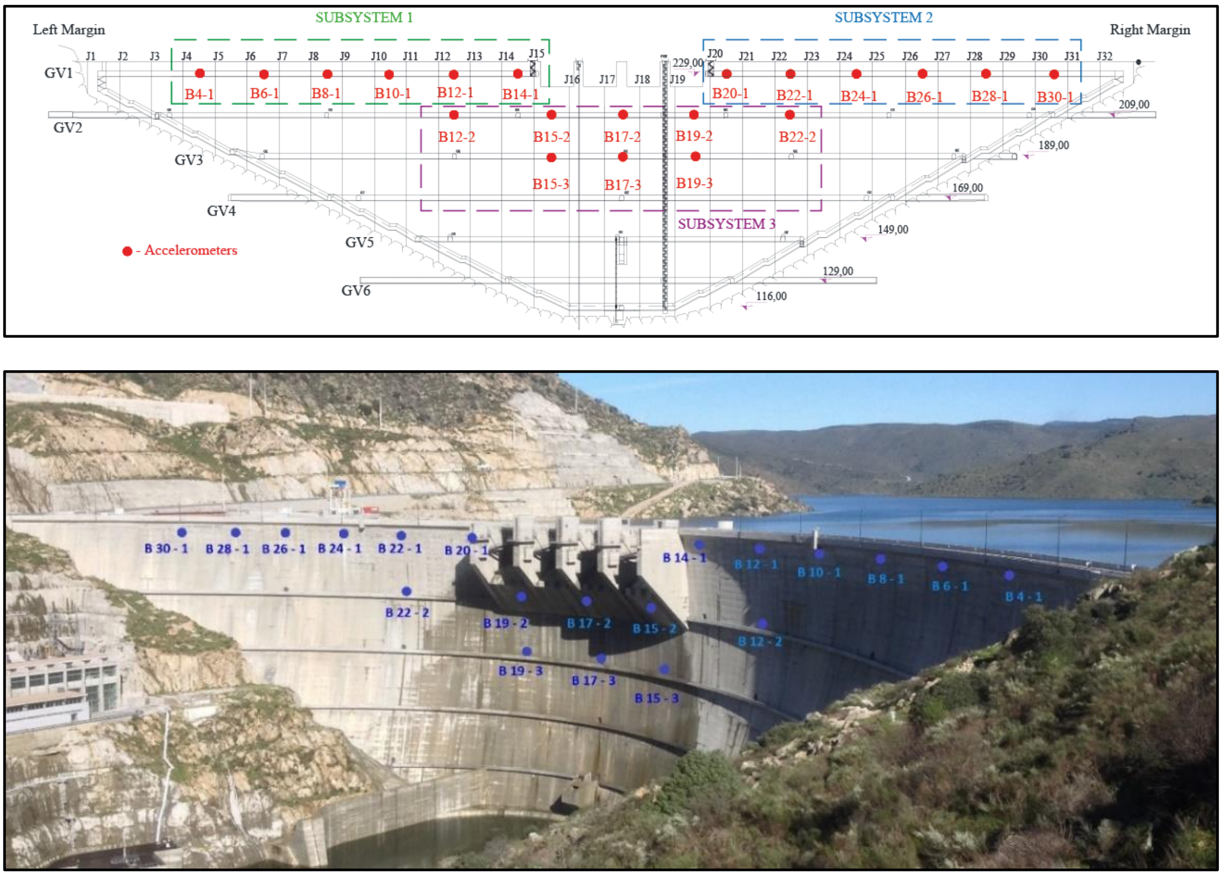


Fig. 19 - Position of measuring points and subsystem components of the dynamic monitoring system of the Baixo Sabor dam (Pereira, 2019).

As can be seen in Fig. 19, the positions of the accelerometers were strategically placed to identify the modal parameters such as natural frequencies, modal damping ratios, and modal configurations. This setup was employed under normal operating conditions for the dam.

4.3. CONTINUOUS DYNAMIC MONITORING

The continuous dynamic monitoring of the Baixo Sabor dam is conducted by the Vibration and Structural Monitoring Laboratory (ViBest) at the University of Porto and the Concrete Dams Department of the National Laboratory for Civil Engineering (LNEC). This monitoring aims to identify and track the evolution of the dam's dynamic characteristics under various operational and environmental conditions.

The modal identification method used was SSI-Cov, which is based on a stochastic state-space model using covariance matrices of the measured structural responses. This method involves singular value decomposition and solving a least squares equation using the Moore-Penrose pseudo-inverse (Peeters, 2000). Stabilization diagrams help distinguish between physical and numerical modes, identifying physical modes that consistently appear in multiple models.

Though the application of the SSI-Cov method is not part of the work developed in this thesis, it corresponds to a brief summary of how the natural frequencies were obtained. Fig. 20 presents the stabilization diagram obtained by applying the SSI-Cov method to a series of data collected from the instrumented dam (Pereira et al., 2018).

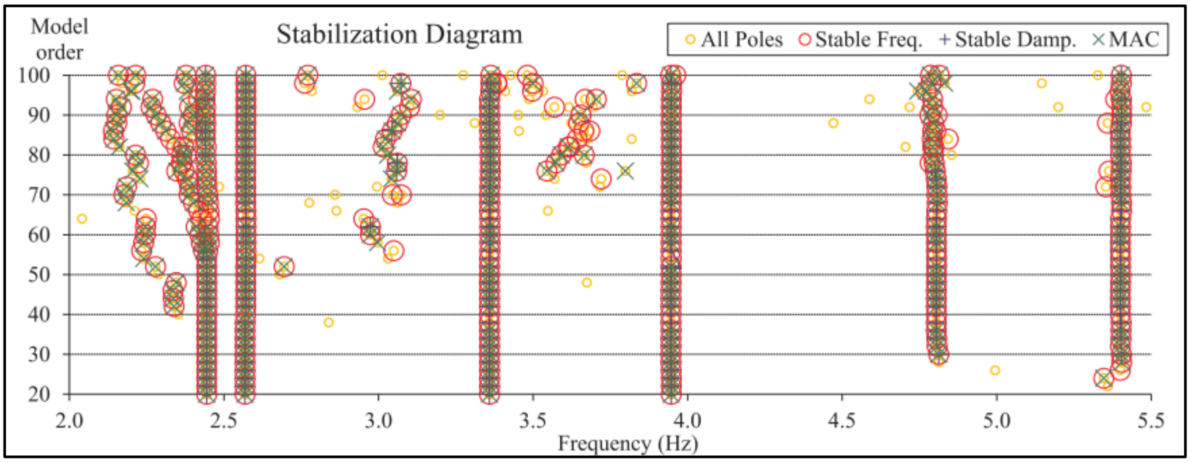


Fig. 20 - Stabilization diagram obtained for the analysis of the time series collected at 05/08/2016 13:00. (Pereira et al., 2018).

Automating the modal analysis is crucial for handling the large amount of data generated and obtaining real-time results. Clustering algorithms are used to group similar modes and remove outliers through outlier analysis. Final modal estimates are obtained by averaging the pole values included in each cluster (Magalhães et al., 2009). Fig. 21 shows modal estimates derived from the stabilization diagram, before and after outlier removal (Pereira et al., 2018).

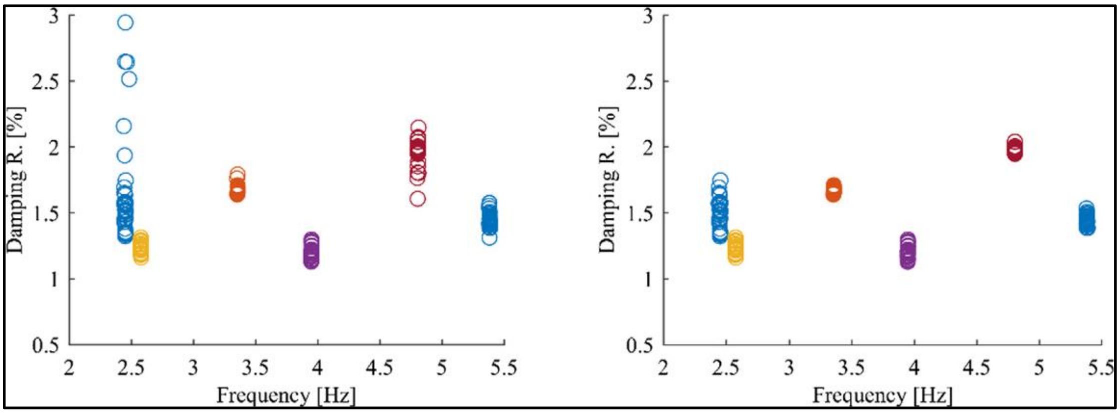


Fig. 21 - Damping vs Frequency diagram of the mode estimates that belong to the clusters with more elements, before (on the left) and after (on the right) the outlier removal (Pereira et al., 2018).

The identified modal parameters in each configuration are compared with previous references using the Modal Assurance Criterion (MAC). This ensures that the modal properties are consistent over time, allowing for the monitoring of the dam's dynamic evolution. The algorithm seeks estimates corresponding to true vibration modes, eliminating spurious modes and harmonics associated with turbine operation.

These methods were applied to the data recorded by the dynamic monitoring system, identifying natural frequencies, damping ratios, and mode shapes. In this work, the natural frequencies resulting from the prior application of identification methods are used to develop linear regression and machine learning models, with the aim of predicting future natural frequencies and evaluating the structural health of the dam, where the use of the first five vibration modes were already defined for the elaboration of this work.

Continuous Dynamic Monitoring is an essential tool for the assessment and monitoring of the structural health of concrete dams. By utilizing historical data and advanced modal analysis techniques, it is possible to make accurate predictions about future natural frequencies and take preventive measures to ensure the integrity and safety of the dam.

An overview of how the natural frequencies were obtained based on the accelerations and continuous monitoring of the Baixo Sabor dam were provided in this section. For the present thesis, the base information on the natural frequencies was supplied to ensure completely reliable and consistent data with which to develop the various models and perform deterministic analyses.

Fig. 22 illustrates the natural frequencies throughout continuous monitoring spanning three years. Data were recorded by devices installed along the dam, capturing frequencies every 60 minutes. This setup yielded 24 data records per day for each of the vibration modes. The dataset used for this thesis includes data monitored from December 1, 2015 (1/12/2015) at 00:00 hours until November 30, 2018 (30/11/2018) at 23:00 hours.

The extensive dataset, comprising numerous records over three years, allowed for the development of robust models to analyze the dam's behaviour under various conditions. The continuous recording of data every hour provided a detailed temporal resolution, facilitating a comprehensive understanding of how the natural frequencies evolved over time. This high-resolution data is essential for accurately capturing the dam's dynamic responses and for building reliable predictive models.

The goal of using such a detailed and prolonged dataset was to ensure that the developed models could accurately reflect the dam's behaviour and provide reliable predictions for future conditions. By working with such a comprehensive set of monitored data, the research aimed to achieve a high level of precision and confidence in the deterministic models, ensuring their applicability for effective structural health monitoring and maintenance planning for the Baixo Sabor dam.

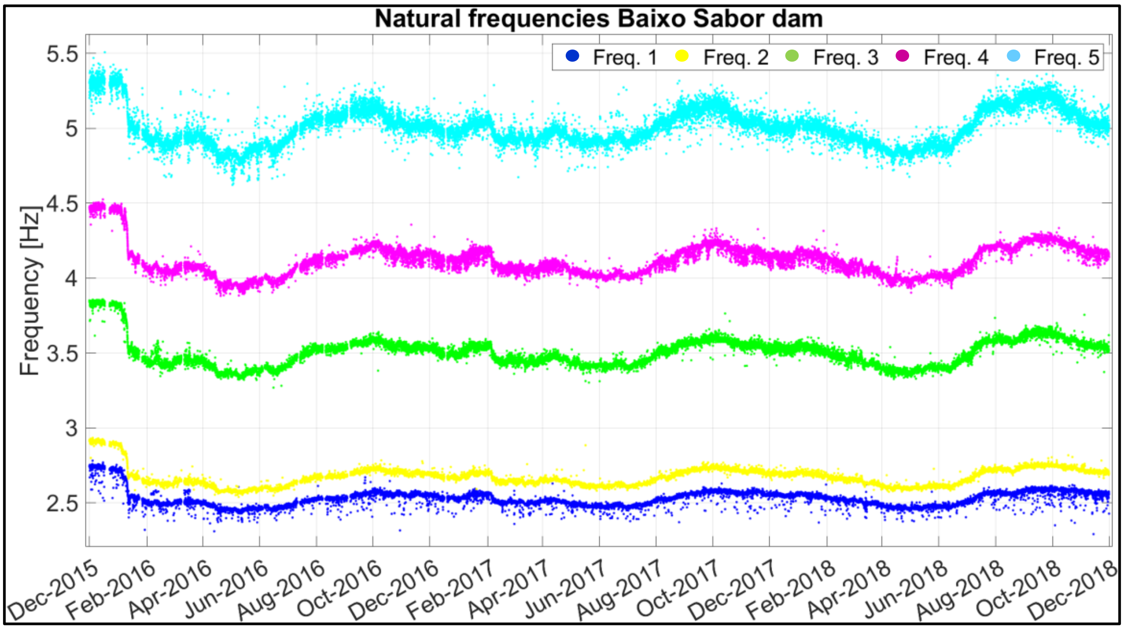


Fig. 22 - Evolution of natural frequencies through three years of monitoring.

Frequencies are typically expressed in Hertz [Hz] and correspond to the rate at which the dam structure naturally oscillates or vibrates under ambient conditions. In this case, the first frequency, ranging from approximately 2.2 to 2.8 hertz, represents the first mode of the dam-reservoir system. As the frequency index increases, indicating higher vibrational modes, the frequency values also increase accordingly. This progression signifies the varying degrees of stiffness and flexibility within the dam structure, as well as the interactions between the dam and its surrounding environment.

4.4. WATER LEVEL AND AIR TEMPERATURE VARIATIONS

Fig. 23 illustrates the water level observed over a period of three years. The x-axis represents the time, while the y-axis shows the corresponding water level. The sudden fall of natural frequency values in January 2016 is explained by the variation in the reservoir water level. Comparing Fig. 22 with Fig. 23, it is possible to notice that an inverse proportionality occurs between natural frequency values and reservoir water level. Natural vibration frequencies of the structure decreased considerably and continuously since the beginning of the monitoring, whereas the water level of the reservoir has been increasing throughout this period. This phenomenon is seen in January, when the first filling of the reservoir occurred, motivating a sudden rise in the reservoir water level and a sudden drop in the frequency values which would be expected, given that a large amount of mass was added to the dam reservoir system. Moreover, relatively low variations of the water level occurred between February and May 2016 that were accompanied by identical types of natural frequency change, thus, confirming the accuracy achieved with the obtained estimates (Pereira et al., 2018).

Additionally, data reveals variations in the water level over time due to the season, having more water in the reservoir during the spring and early summer, and having less during the autumn and early winter and these fluctuations show almost the same behaviour over time, providing insights into the dynamics of the reservoir. Serving as a crucial aspect of understanding the behaviour of the dam and its surrounding environment.

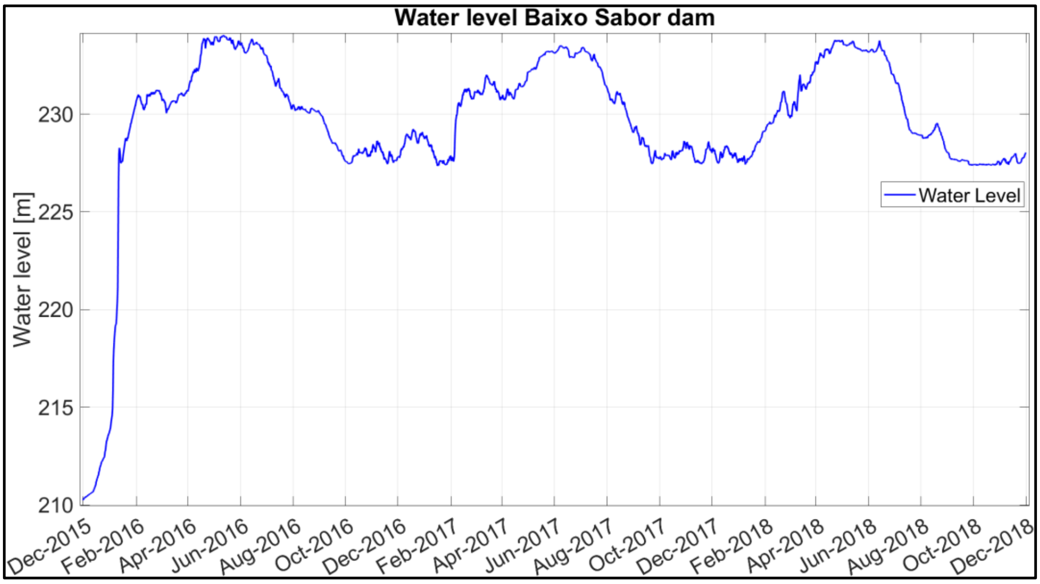


Fig. 23 - Water level records from 2015 to 2018.

The monitored period corresponds to the dam first-filling, thus producing unique results that will be helpful to better understand the effect of the water level on the structure's dynamic behaviour, and will consequently allow the calibration and updating of numerical models.

Fig. 24 displays the air temperature recorded during a three-year monitoring period. Similar to the water level graph, the x-axis represents the monitoring time, and the y-axis indicates the records of the air temperature. The variations in temperature over time offer valuable information about seasonal variations and climatic conditions in the region. Analyzing these patterns can help in assessing the impact of the air temperature on various aspects of the dam's pattern behaviour.

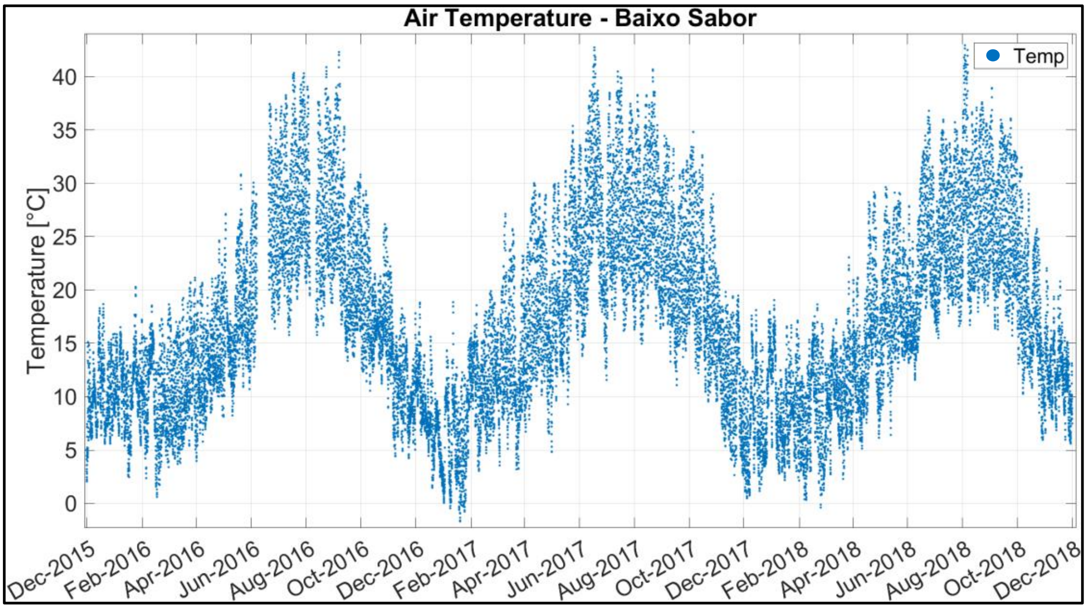


Fig. 24 - Air temperature records from 2015 to 2018.

Fig. 24 details a consistent pattern behaviour due to the seasonal variations, where higher temperatures are indeed observed during the spring-summer seasons, and consequently, lower temperatures occur during the autumn-winter seasons. Although these air temperature variations may not seem significant, they do affect the behaviour of the dam structure. They can cause dimensional changes that generate stresses in the dam structure, and the natural frequencies of the structure. It's essential to consider these environmental factors when assessing the structural dynamic behaviour of the dam over time.

4.5. MULTIPLE LINEAR REGRESSION MODELS

4.5.1. INTRODUCTION

The HST approach has primarily been used to analyze quasi-static physical quantities, such as horizontal displacements measured using the pendulum method. One of the recent innovations in this approach is the general characterization of the natural frequency pattern through data-based models, which allows the transposition of the HST approach to characterize dynamic behaviour. Since there are enough records of natural frequency available for the case study dam, the Baixo Sabor dam, with high-quality data recorded at hourly intervals, the characterization of the spectral dynamic behaviour could be performed.

In many applications of the HST approach, it's not necessary to use the complete polynomial to represent the hydrostatic effect. For the development of the thesis, only the first term of the full polynomial (h^4) will be considered, which represents the water level in the reservoir and is mentioned in equation (2).

In addition, for the case of the sinusoidal thermal effect the use of the variable day of the year is going to be used in the function of sine and cosine with a period of one year, which was shown in eq. (4).

The characterization of the variation of the dam's natural frequencies presented in the previous section indicates that, though other factors may play an important role, these are essentially influenced by reservoir water level and temperature. In turn, this section aims to establish a model to minimize the effects of these environmental and operational factors on the experimental natural frequencies.

4.5.2. DEVELOPMENT OF THE MLR MODEL BASED ON THE HST APPROACH

In the present application, a multiple linear regression model was adopted as the first test under the HST approach. The natural frequencies collected during the three years of continuous monitoring mentioned above were used, whereas Fig. 25 illustrates the two years of monitoring, from 01/12/2015 to 30/11/2017, used to train the model and to establish the base model. Furthermore, the remaining data which refers to the third year of monitoring, from 01/12/2017 to 30/11/2018 were used to validate the quality of the forecasts provided by the regression model having a trained model based on the training data with the two years of monitoring.

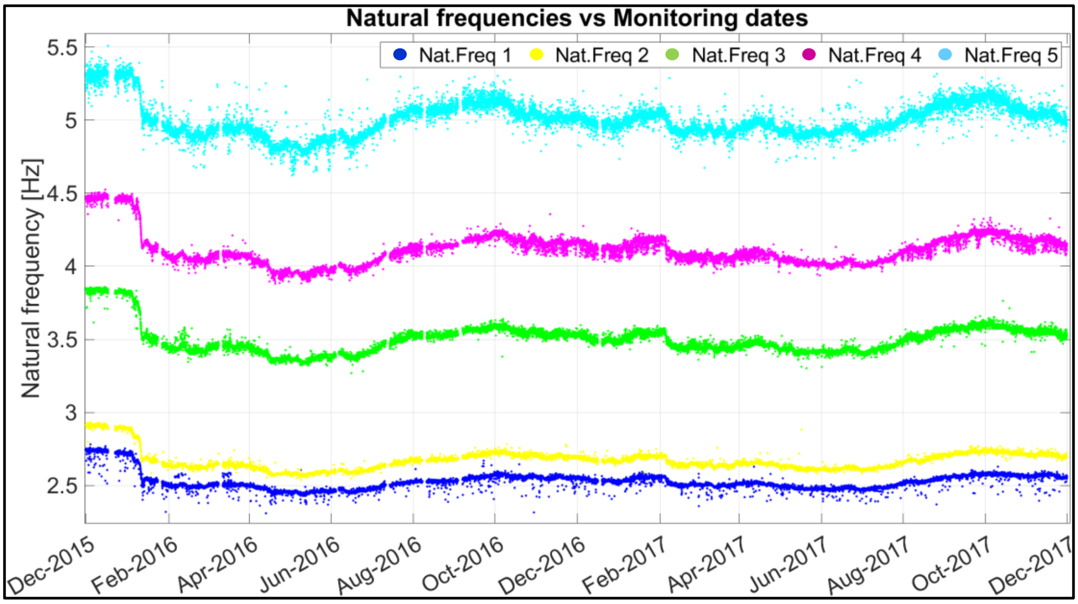


Fig. 25 - Evolution of natural frequencies from 2015 to 2017 (data adopted as a training set).

This division is important for ensuring the robustness and accuracy of the predictive models. By using data from the first two years for training, the models learn and capture the underlying patterns, trends, and relationships present in the dataset. This extended period of training enables the models to develop a comprehensive understanding of the system dynamics, including the complex relations between variables such as water level, and natural frequencies.

Subsequently, the third year's data, which the models have not encountered during training, serves as an independent evaluation set for testing the models' predictive performance. By evaluating the models on unseen data their ability to generalize and make accurate predictions beyond the training dataset can be assessed. This process helps to validate the models' effectiveness and identify any potential shortcomings or areas for improvement.

Overall, the utilization of separate training and testing datasets enables to develop reliable and effective predictive models that can provide valuable insights into the behaviour and performance of the dam system over time.

The whole procedure was performed under the software (Matlab R2022a) which, thanks to its wide use and magnitude at a software level, this project was developed.

The model was developed using the HST approach, where three inputs were considered. One input to represent the effect of the water height variations and two inputs to represent the effect of the annual variation of the temperature. Furthermore, the natural frequencies (one at a time) are the outputs of the model.

The main terms of the model adopted in this work are presented in Equation 6:

$$\delta_{HST} = \beta_0 + \beta_1 * h^4 + \beta_2 * \sin(d) + \beta_3 * \cos(d) \tag{6}$$

where:

$\beta_1 * H^4$ – Hydrostatic effect.

$\beta_2 * \sin(d) + \beta_3 * \cos(d)$ – Seasonal thermal effect.

$\beta_1, \beta_2, \beta_3$ – Unknown coefficients

$h^4, sin(d), cos(d)$ – Independent variables

The training period was from 01/12/2015 to 30/11/2017. Table 1 shows the results obtained from the training model based on the HST approach. This table allows to identify the coefficients $\beta_1, \beta_2, \beta_3$. These coefficients were used to conduct testing in the third year, which is dedicated to testing.

Table 1- Regression coefficients for the MLR

Frequency mode	β0	β1	β2	β3
Freq. 1	2.9543	$-2.2235\text{x}10^{-9}$	-0.010095	-0.0040453
Freq. 2	3.1742	$-2.6178\ \text{x}10^{-9}$	-0.14286	-0.008275
Freq. 3	4.2382	$-3.8372\ \text{x}10^{-9}$	-0.026121	-0.018613
Freq. 4	4.8933	$-4.0449\ \text{x}10^{-9}$	-0.027739	-0.0081896
Freq. 5	5.7554	$-3.9158\ \text{x}10^{-9}$	-0.050674	-0.034902

With all these input data, a linear regression model was constructed using the information extracted from each predictor variable and the corresponding coefficients obtained through the input variables. Subsequently, a prediction obtained is shown in Fig. 26 using the trained model as a basis to verify the reliability of the data.

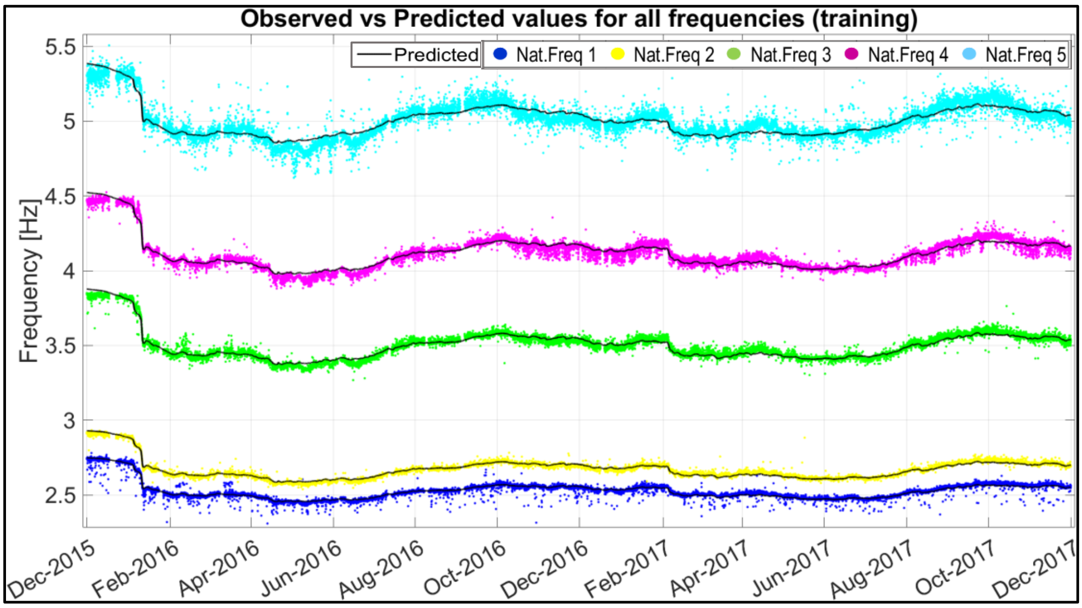


Fig. 26 MLR predicted values and observed frequency modes values for the training set.

This was done by evaluating the residuals and the determination coefficients (R^2), which indicate how much variation in the dependent variable (response variable) is explained by the variation in the independent variable (predictor variable) in the regression model as can be appreciated in Table 2. This analysis allowed for the assessment of the quality and accuracy of the developed model and its ability to predict the behaviour of the dam based on the variables considered.

Table 2 - Performance parameters of the MLR model based on the training set.

Frequency mode [Hz]	**Freq. 1**	**Freq. 2**	**Freq. 3**	**Freq. 4**	**Freq. 5**
Determination coefficient [%]	92.4	96.9	95.3	93.5	81.6
Standard deviation of residuals [Hz]	0.0162	0.0116	0.0212	0.0277	0.0498
Max. abs. value of residuals [Hz]	0.245	0.274	0.256	0.183	0.362

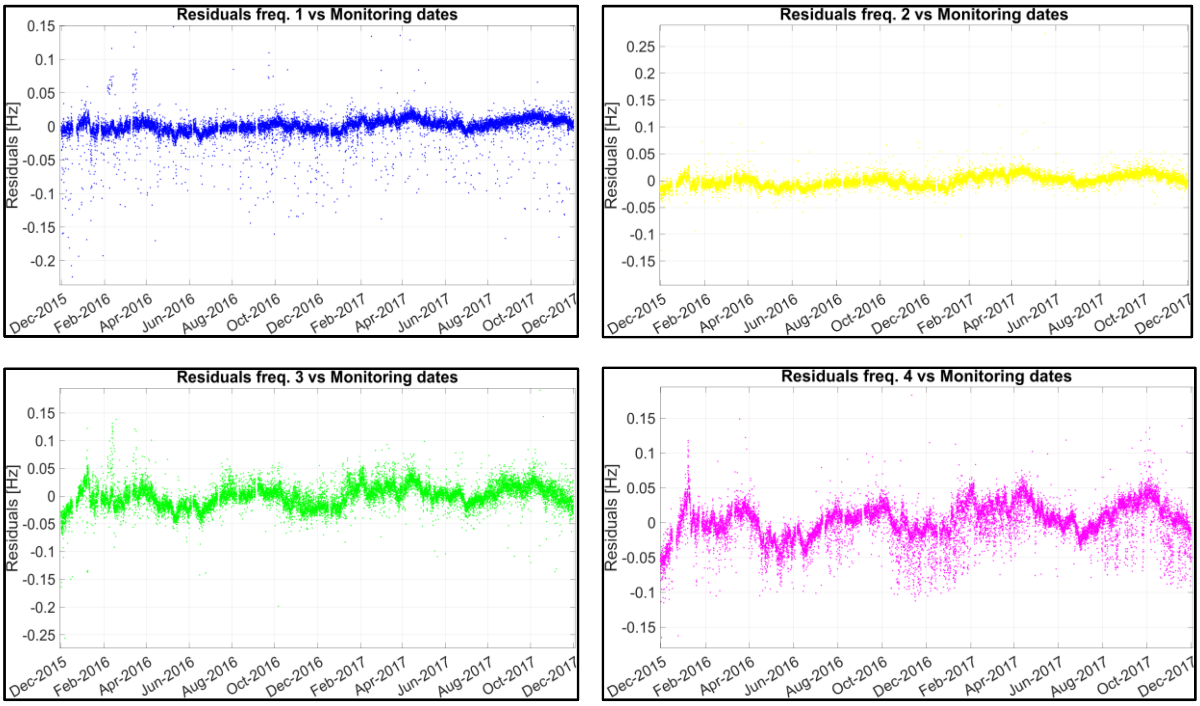
Fig. 26 visually compares the observed data (real-world measurements) with the predictions made by the model for each frequency over time. Each colour line on the graph represents a different frequency mode.

A good model fit is indicated when the lines of "observed (dots) + predicted (line)" closely follow the actual behaviour of the vibration mode shape. If the predicted lines closely match the observed data, it suggests that the model is effectively capturing the variability in the data. According to the observed results, it can be appreciated that the predicted data follow a trend within the variability range of the monitored vibration modes, which is also supported by Table 2 where it can be seen the determination coefficient which indicates how much the prediction (response variable) is explained by the model (predictor variables).

Discrepancies between the "observed + predicted frequencies" lines and the actual data may indicate conditions where the model is not adequately capturing the variation in the data. These discrepancies could result from errors in the model formulation, missing important data, or the presence of unaccounted factors.

Ideally, the residuals should be randomly spread around zero (0), indicating that the model's predictions are fair and accurate. If the residuals show a consistent pattern or trend over time, it suggests that the model is missing some aspect of the data.

In the analysis of residuals for the training set, it is observed that the residuals associated with each vibration mode exhibit a balanced distribution around zero. Upon visually inspecting the residual plots for each natural frequency, it is evident that both positive and negative residuals are present in similar quantities, indicating a lack of systematic bias in the model predictions. Fig. 27 provides the maximum absolute value of the residuals for each natural frequency ranges between 0.183 and 0.362 [Hz], suggesting that the residuals do not deviate significantly from zero and that the model tends to make predictions with acceptable accuracy for the training set.



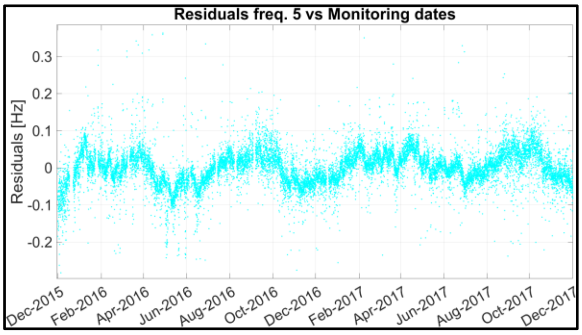


Fig. 27 - Evolution of MLR residuals for each frequency mode of the training set.

The residuals observed in the training period with the regression model show results where the residuals for lower frequencies are close to zero and present a lower standard deviation of the residuals. On the other hand, for higher frequencies, the residuals start experiencing a higher dispersion, which is still not significant due to the small differences between low and high frequencies. In addition, there is a balance between positive and negative residuals. This indicates that the model neither overestimates nor underestimates the observed values.

4.5.3. PERFORMANCE OF THE MLR MODEL FOR THE TEST SET

After completing the training for the base model, where the model learned patterns and sequences from the provided data collected over the first two years of monitoring, the regression model was used to predict data for the third year of analysis. This predicted data is going to be analyzed to evaluate the accuracy of the predictions when applied to unseen data.

Following this, the predictions are going to be examined for the third year which is the testing section to assess the model's performance and effectiveness. Below Fig. 28 shows a closer view of the results predicted by the regression models for the third year (testing set).

The testing phase of our model plays a critical role in validating its performance, assessing its predictive accuracy, and identifying areas for improvement. By rigorously evaluating the model's performance on unseen data, we gain confidence in its ability to make accurate predictions and inform decision-making processes effectively.

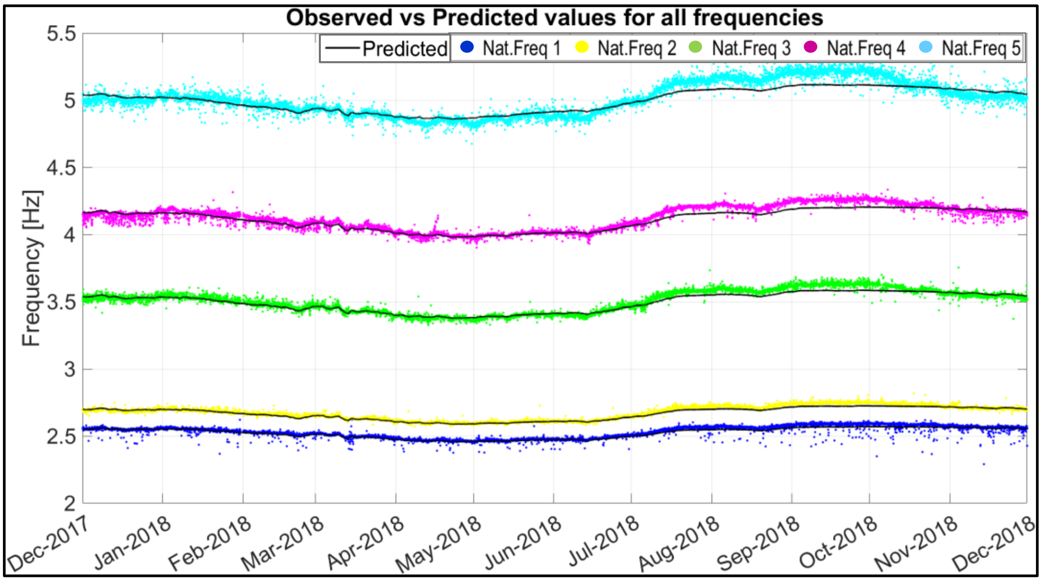


Fig. 28 - MLR predicted values and observed frequency modes values for the test set.

After training the model on historical data, the testing phase allows to evaluate how well the model performs on new, unseen data. This is crucial for assessing the generalization ability of the model and ensuring that it can make accurate predictions on data it has not encountered before. By comparing the model's predictions with the actual observed values in the testing dataset, it can assess the model's performance metrics, such as determination coefficient, and other relevant metrics, which provide insight into how well the model captures the underlying patterns and variability in the data. A summary table of the maximum absolute values of the residuals is also provided in Table 3.

Table 3 - Performance parameters of the MLR model based on the test set.

Frequency mode [Hz]	Freq. 1	Freq. 2	Freq. 3	Freq. 4	Freq. 5
Determination coefficient [%]	83.4	91.8	89.6	84.6	78.5
Standard deviation of residuals [Hz]	0.0163	0.0116	0.0247	0.0313	0.055
Max. Abs. Value residuals [Hz]	0.269	0.12	0.191	0.194	0.353

When analyzing the results of the testing phase, aspects such as the distribution of residuals, standard deviations, maximum absolute value of residuals, and the overall fit of the model were examined. Considering whether the model's predictions are within acceptable bounds and whether any patterns or trends exist in the residuals.

Fig. 29 shows the testing model and prediction in red, showing a strong alignment between the base model and the testing one. For each frequency, the model fits almost perfectly, with minimal margin for error. In addition, Fig. 29 shows the combination of both models for the training period and also for the testing period, where the experimental data observed during the testing period is represented in red for all the five vibration modes.

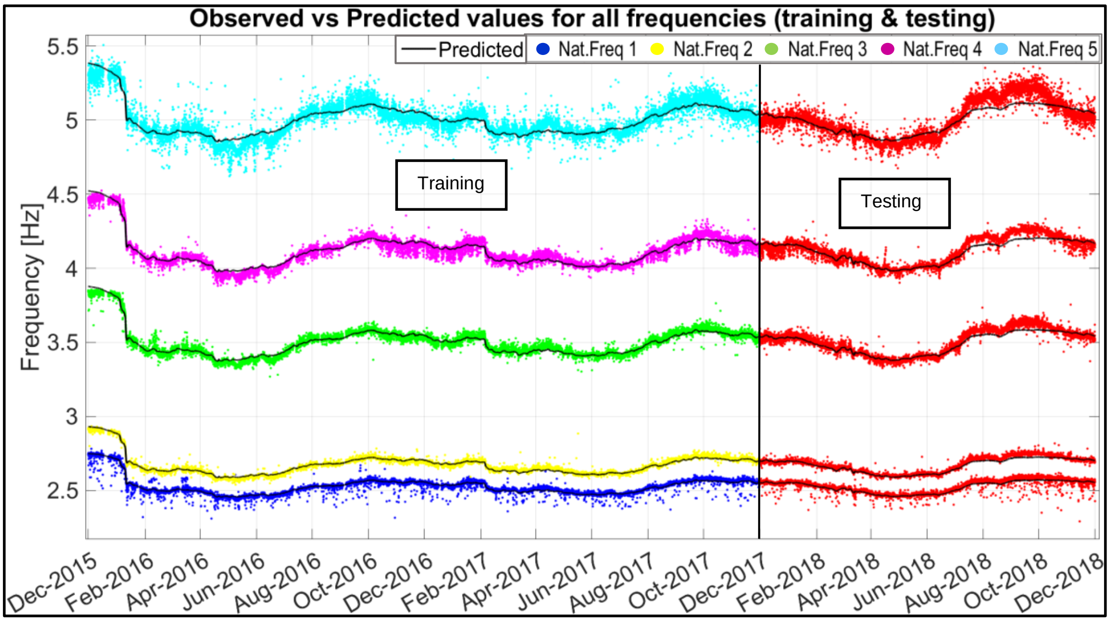


Fig. 29 - MLR predicted values and observed frequency modes values from 2015 to 2018 (training and test sets).

During testing, an evaluation on how closely the model's predictions align with the observed values is considered. Ideally, the predicted values should closely match the observed values across the testing dataset. If there are discrepancies between the predicted and observed values, it could indicate areas where the model needs improvement or where there are limitations in the data.

In the same manner as for the training model, the residuals are analyzed in various aspects such as the balance between positive and negative residuals, how close the data points are to zero, and ensuring they do not follow any unusual patterns.

It is observed in Fig. 30 that the values are relatively low, as expected for a predicted model where the margin of error may be slightly higher than for the training model. The residuals associated with each vibration mode exhibit a balanced distribution around zero, with an equitable presence of positive and negative residuals. Additionally, the maximum absolute value of the residuals for each natural frequency varies between 0.120 and 0.353 Hz, indicating that the model maintains consistency in the accuracy of predictions for the test set, without significant deviations from zero.

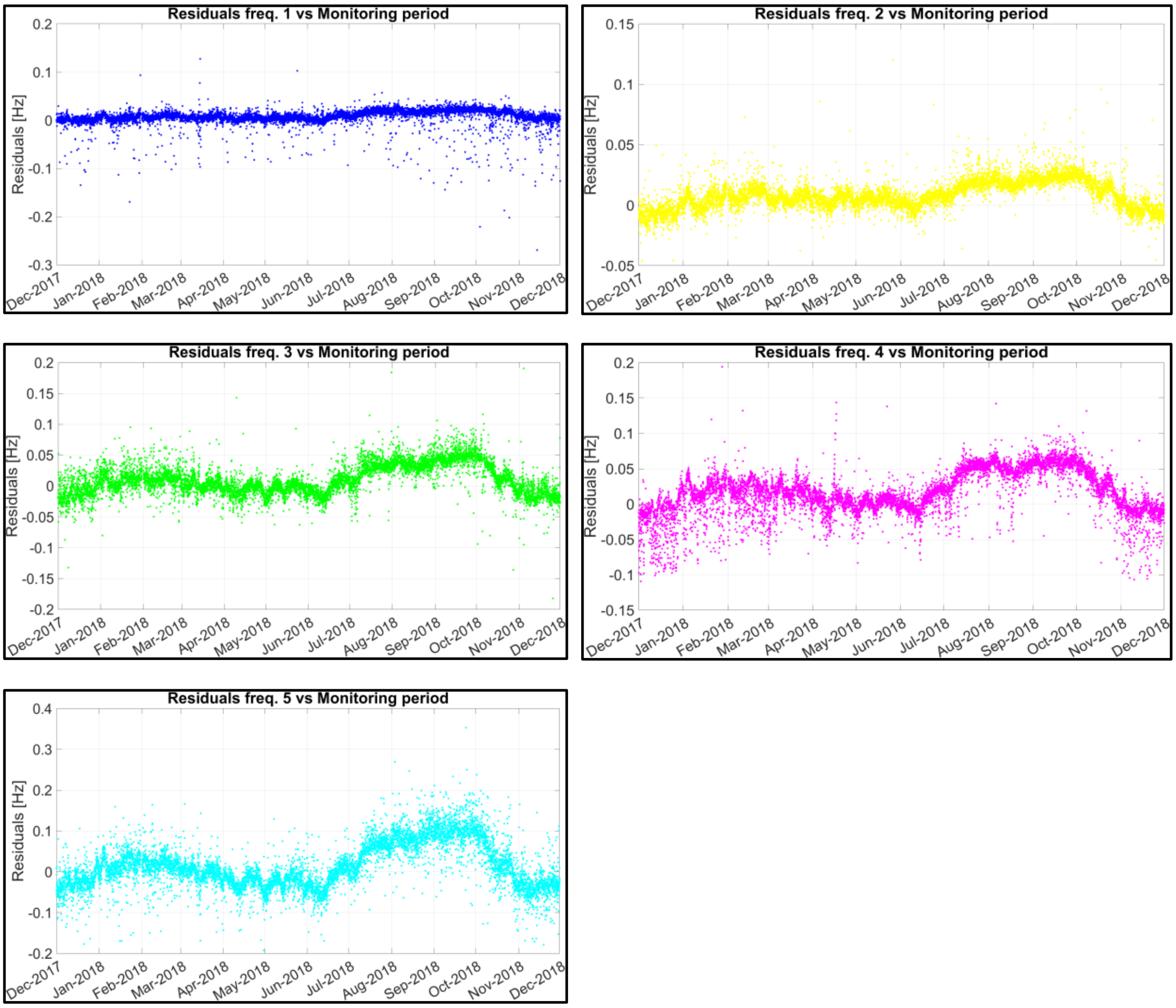


Fig. 30 - Evolution of MLR residuals for each frequency mode of the test set.

An analysis was also conducted on the variables $\beta_1, \beta_2,$ and β_3 where the first is related to water behaviour (hydrostatic effect), while β_2 and β_3 are associated with air temperature (seasonal thermal effect) as can be appreciated in Fig. 31. These environmental factors play a significant role in the dam's behaviour. On the other hand, the water effect is due to the hydrostatic pressures in the dam, while the temperature effect influences and/or alters the internal characteristics of the concrete such as the thermal expansion and contraction.

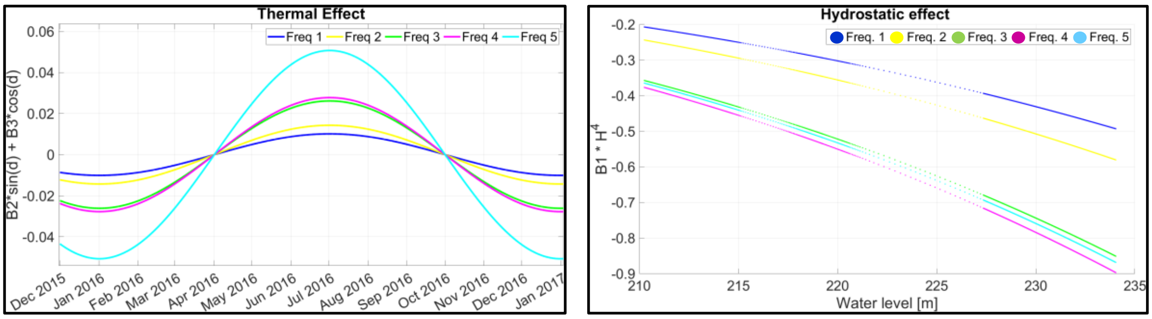


Fig. 31 - Thermal effect (left) and hydrostatic effect (right) for each frequency mode based on the MLR models.

Given the examination of natural frequencies, these environmental factors also contribute to the variations in frequencies, which should not be overlooked. Fig. 31 illustrates the effects of these environmental factors and how they vary over time in the case of the temperature due to the seasonal thermal effect for every season of the year, and the water level varies due to the level of the water in their reservoir.

4.5.4. INTERPRETATION OF RESULTS

After defining and understanding the concepts of standard deviation, residuals, and determination coefficients in previous chapters, the interpretation of the training and testing models was conducted. This analysis was based on the previously mentioned tables, which provided the necessary data for evaluating the models' performance.

In Table 2 and Table 3 presented earlier, the results obtained by the linear regression model were detailed, covering both the training and testing data. Both tables corresponding to the linear regression model show the results for each analyzed natural frequency where the evaluated model criteria was based on the determination coefficient, standard deviation, and the maximum absolute value of the residuals.

These results provide a basis for further analysis and estimation of the system's accuracy and data quality obtained, both for the training and testing models. It is observed that the determination coefficient exceeds 75% and 90%, indicating a good fit within the learning parameters.

4.6. MULTILAYER PERCEPTRON NEURAL NETWORK MODELS

As referred before, a MLP-NN model is adopted. The same dataset as the regression model was used. In addition, for the development of neural networks, the HST approach is also taken into consideration and an example of the architecture of these models based on neural networks is also detailed in Fig. 32.

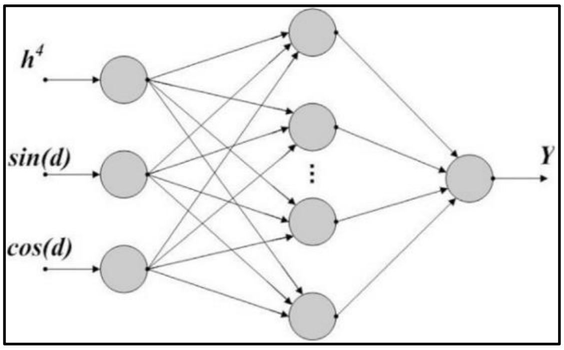


Fig. 32 - Architecture of an MLP-NN model based on the HST approach. (Mata et al., 2023).

The data collected during the two years of continuous dynamic monitoring, from 01/12/2015 to 30/11/2017, is used to establish the model and then, data collected during the third year of monitoring, from 01/12/2017 to 30/11/2018, is used to validate the quality of the forecasts provided by the neural network.

In the same way, as the multiple linear regression model was developed, the same independent variables are still considered. However, using the HST approach, this model's training proceeds differently.

Unlike the previous model, this approach does not require coefficients like $\beta_1, \beta_2, \beta_3$ due to its different architecture. The model was trained on all patterns, using the software Matlab in its R2022a version and the function fitrnet, which has the significant advantage of being able to run both multiple linear regression and neural network methods.

Using MATLAB software, the fitrnet function was employed for training and testing the model. This function was essential for evaluating the model, allowing the modification of some parameters to achieve better control of the model and consequently obtain good results.

The fitrnet function was adjusted to provide optimal results in terms of reliability, focusing on the standard deviation, the maximum value of the residuals, and the determination coefficient. The function captured the best results considering variables such as weights, hidden layers, number of neurons, and overfitting reduction. The model presented in this section was determined by the use of two hidden layers, with a range of 1 to 15 neurons per hidden layer.

In the adopted architecture, the weights are important parameters that determine the strength and direction of the influence that one neuron has on another in the neural network. During the training set, the neural network adjusted the weights of the connections between neurons in each layer in order to minimize the difference between the model's predictions and the real values. These weights were iteratively updated using optimization algorithms for each of the vibration modes, allowing the network to learn patterns from the data.

The initialization of the weights influenced how the model began the learning process. Good initialization helped the algorithm converge more quickly towards an optimal solution, avoiding problems such as getting stuck in local minima. For this reason, the initialization of the weights started randomly and close to zero to increase the chances of good convergence.

The inclusion of hidden layers in the neural network allowed the model to learn more abstract and complex representations of the data. Each hidden layer added a level of processing, enabling the network to extract deeper features from the inputs. In this specific case, different configurations of hidden layers were experimented with in order to identify the structure that best captured the non-linear relationships present in the data, resulting in a greater ability of the model to generalize and make accurate predictions.

The number of neurons in the hidden layers directly affected the model's ability to learn from the data. An appropriate number of neurons allowed the model to capture enough detail without overfitting the training data. A range of neurons from 1 to 15 was established, and the optimal configuration that balanced model complexity and generalization ability was selected. This adjustment helped maximize the accuracy of predictions in the test set.

The choice of the sigmoid activation function in the hidden layer allowed the introduction of non-linearities into the model, crucial for capturing complex relationships in the data. This function helped the neurons in the hidden layers activate appropriately in response to inputs, enabling the model to learn more detailed representations. As a result, there was an improvement in the model's ability to make accurate predictions about natural frequencies.

Fig. 33, as shown, represents the result of the trained model with the five vibration modes identified by distinct colours. Each vibration mode is accompanied by its corresponding prediction. This approach ensures that the visual representation allows for an easy comparison between actual and predicted data points. To maintain consistency and reliability in the analysis, the model was trained using the same amount of data and identical datasets as the ones used in the linear regression model. The primary distinction between the two models lies in the method employed, highlighting the differences in predictive performance between the MLP-NN and MLR models.

The results obtained for the training period provided a determination coefficient greater than 90% for each of the frequencies, having them between 89 % and 98 %. In general terms, this indicates how much of the variation in the dependent variable (response variable) is explained by the independent variable (predictor variable) variation in the neural network model.

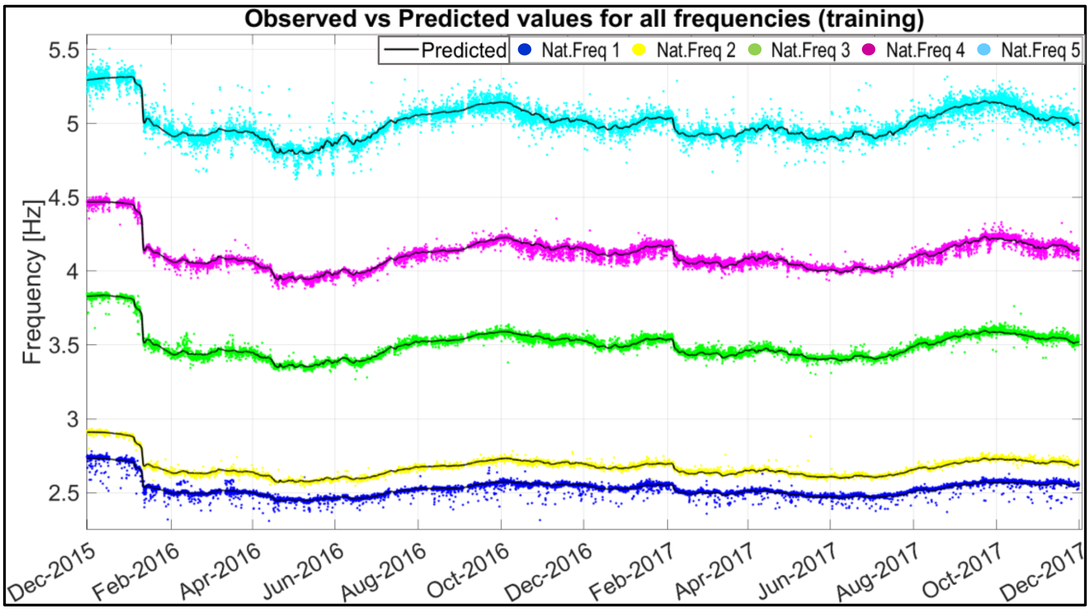


Fig. 33 - MLP-NN predicted values and observed frequency modes values for the training set.

A similar representation is used for Fig. 34, which presents the testing period. This figure is exclusively dedicated to the third year, providing a detailed view of the model's predictions during this period. The testing phase was crucial for evaluating the model's generalization capabilities and its accuracy in predicting new unseen data. By applying the same principles and data consistency as in the training phase, the model's predictions for each vibration mode during the third year were obtained and analyzed. The results obtained for the third year slightly decreased compared to the training model, with determination coefficient values ranging between 89% and 96%. This consistent approach ensures that the comparison between the training and testing results is valid and meaningful, offering insights into the model's robustness and reliability under different conditions.

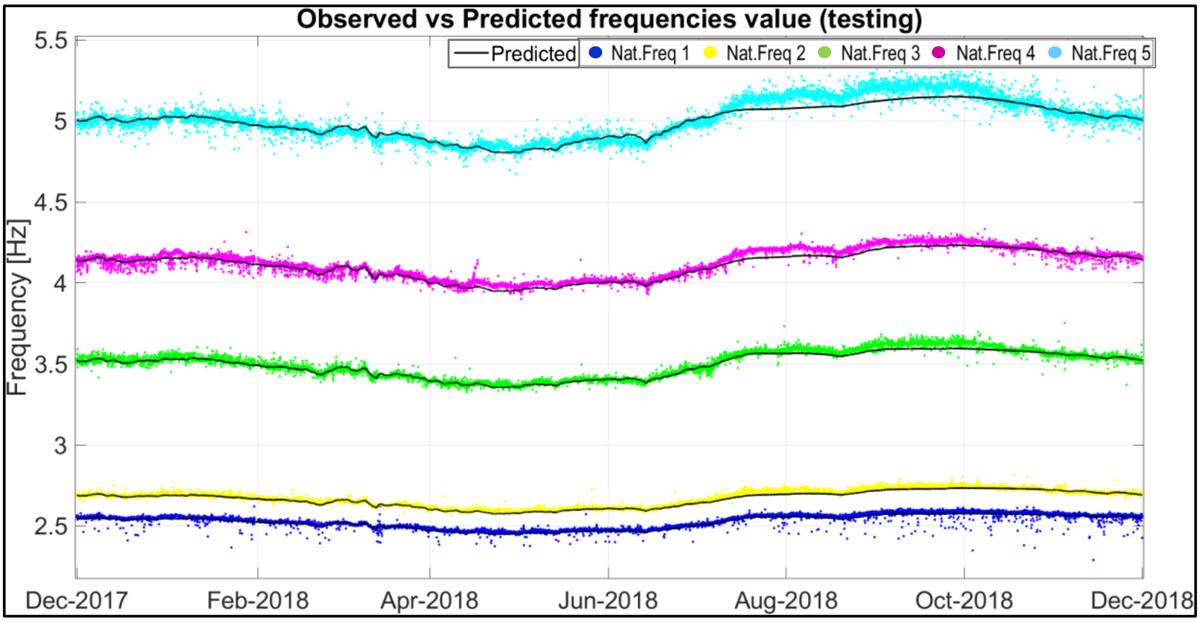


Fig. 34 - MLP-NN predicted values and observed frequency modes values for the test set.

In summary, Fig. 33 and Fig. 34 together provide a comprehensive view of the model's performance over both the training and testing periods, allowing for a thorough evaluation of its predictive accuracy and consistency across different datasets and timeframes.

As it was shown in Fig. 34, testing was conducted for the period from December 2017 to November 2018. At first glance, the prediction appears quite consistent and closely aligned with each of the natural frequencies. The complete Fig. 35 for a three-year period is presented, where the training model's frequencies are depicted in their respective colours. However, the tested year is represented by a red colour, with the prediction shown in black.

Fig. 35 illustrates the five natural frequencies, with the lowest magnitude frequencies at the bottom and the highest at the top of the graph. These frequencies are depicted over the monitoring time, where the first two years were used for training the model, and the last year represents the prediction of the trained model, i.e., the testing phase (shown in red for the natural frequencies and a black line for the prediction).

It is noteworthy that Fig. 35 demonstrated a significantly more adaptable and flexible behaviour in comparison to Fig. 29 (figure related to the MLR model), where it presents the outcomes of both the training and testing phases, providing a clear view of the model's performance across different stages. In contrast, Fig. 35 highlights a distinct pattern of flexibility that merits further analysis.

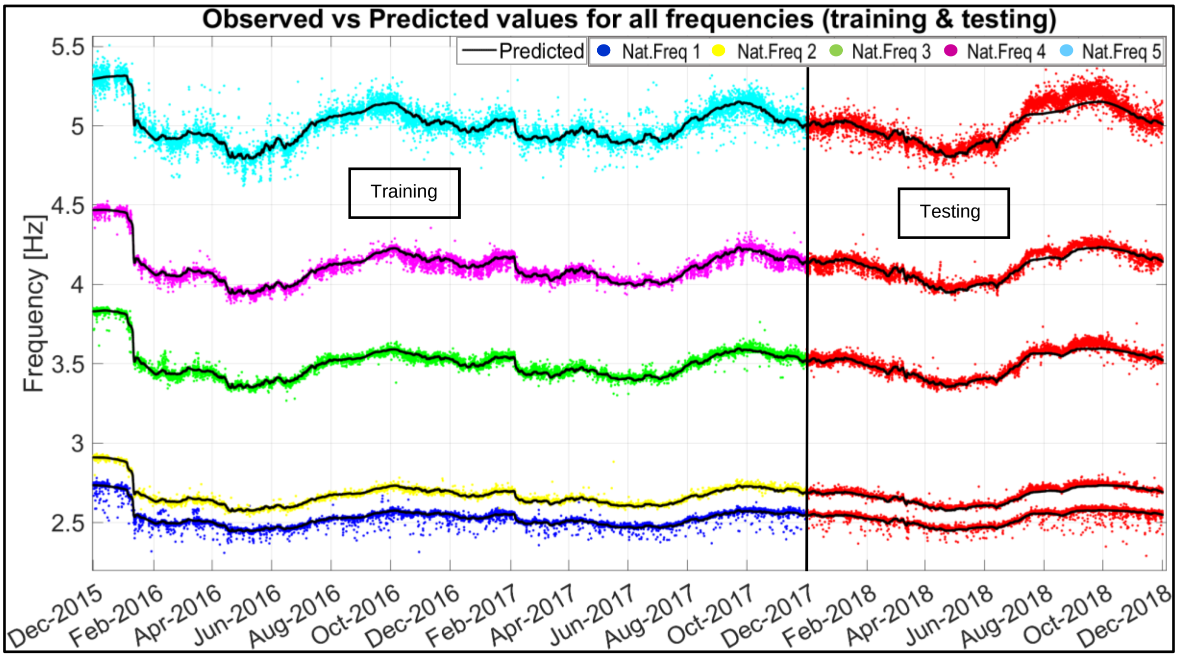


Fig. 35 - MLP-NN predicted values and observed frequency modes values for the training and test sets.

To thoroughly understand these differences, a detailed comparison between the MLR model (Fig. 29) and the MLP-NN model (Fig. 35) will be conducted forward in section (4.7). This comparison will delve into the specific aspects of each figure, examining the variations in behaviour and performance metrics observed during both the training and testing periods. By analyzing these figures side by side, we can gain deeper insights into the model's adaptability and accuracy, as well as its ability to generalize from the training data to new, unseen data.

This detailed comparative analysis will help identify the strengths and weaknesses of the model under different conditions, ultimately contributing to a more comprehensive understanding of its predictive capabilities and overall robustness.

4.7. COMPARISON BETWEEN REGRESSION AND NEURAL NETWORK MODEL

After successfully running both models through their training and testing phases, the results from the multiple regression and the neural network models were finally obtained. This allows to compare the performance of the two models, examining their advantages and disadvantages to help determine which model is better suited according to their purposes. The effectiveness was evaluated, the architecture that each model uses, the complexity involved, and how well they handle and process the given data.

Fig. 36 illustrates the results obtained by both models which represent the regression model's predictions and additionally are depicted in blue, while the neural network's predictions are shown in black. This distinction was made specifically for the testing set, which is the most representative and crucial part of the evaluation process. These graphs detail the first five natural frequencies of Baixo Sabor dam, with colours indicating the training period for both models. The training period was specified as 2 years, occurring from December 2015 to the end of November 2017. Subsequently, both models present their respective predictions, depicted in red, with the predicted results also shown as a black line, covering the time from December 2017 to the end of November 2018.

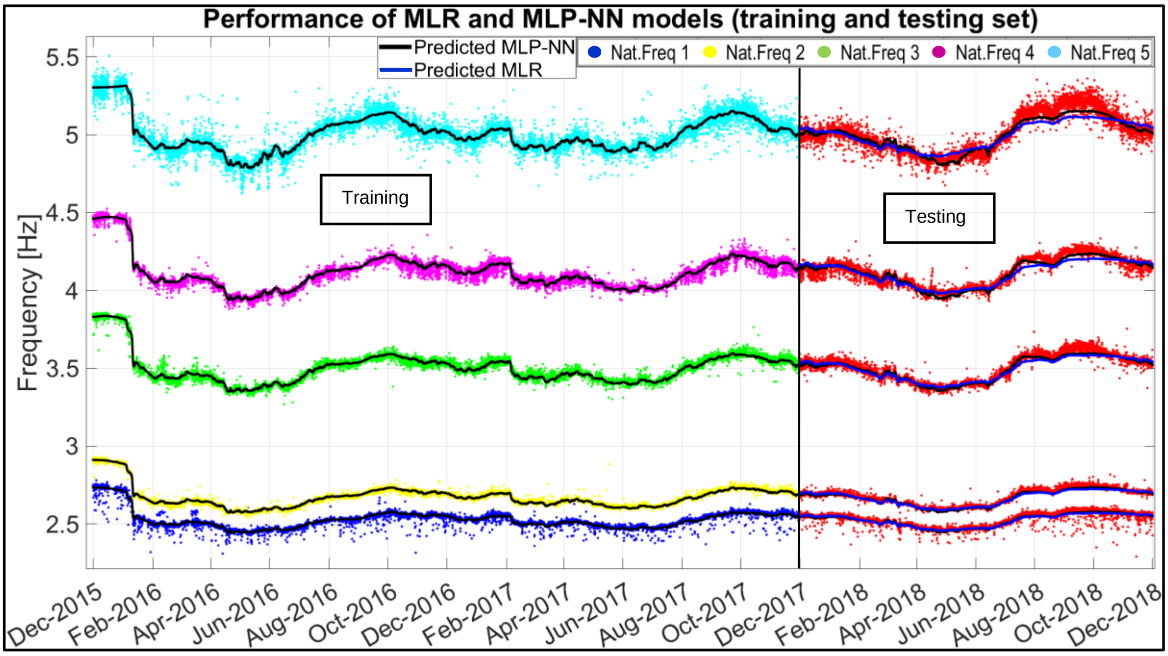


Fig. 36 - Comparison between MLR (blue line) and MLP-NN (black line) models for the training and testing sets.

At first glance, it is evident that both models have fulfilled their objective of predicting the natural frequencies, and both methods serve as reliable solutions. Additionally, the model created by the neural network slightly outperforms the regression model, providing a smoother prediction and tighter patterns. However, the results shown by the regression model are still considered valid and cannot be dismissed as a viable option.

The linear regression model, while effective, relies on a simpler approach that may not capture all the complexities within the data. On the other hand, the neural network model demonstrates superior learning capabilities due to its inherent flexibility. This flexibility allows it to learn complex relationships and patterns in the data, resulting in more accurate predictions.

Table 4 provides a detailed summary of the training and testing results obtained from the regression model MLR and the MLP-NN model. With its ability to adapt and improve through the learning process, the neural network model clearly outperformed the linear regression model in terms of accuracy and predictive power. This comparison underscores the advantages of using advanced machine learning techniques for modelling dynamic behaviours in complex systems.

Table 4 - Results of the training and testing set for MLR and MLP-NN.

Model	Frequency mode [Hz]	Set	Freq. 1	Freq. 2	Freq. 3	Freq. 4	Freq. 5
MLR	Determination coefficient [%]	Training	92,4	96,9	95,3	93,5	81,6
		Testing	83.4	91.8	89.6	84,6	78.5
	Standard deviation of residuals [Hz]	Training	0,0162	0,0116	0,0212	0,0277	0,0498
		Testing	0.0163	0.0116	0.0247	0.0313	0.055
	Max. abs. value of residuals [Hz]	Training	0,245	0,274	0,256	0,183	0,362
		Testing	0.269	0.12	0.191	0,194	0.353
MLP-NN	Determination coefficient [%]	Training	93.3	98.1	97.1	97.9	90.1
		Testing	89.4	96.3	95	93.4	89.3
	Standard deviation of residuals [Hz]	Training	0.0152	0.0091	0.0165	0.0202	0.0386
		Testing	0.0154	0.00858	0.0177	0.0224	0.0368
	Max. abs. value of residuals [Hz]	Training	0.249	0.274	0.218	0.199	0.414
		Testing	0.269	0.123	0.196	0.188	0.316
	Number of neurons in the hidden layer	Training	**[2 5]**	**[2 8]**	**[2 5]**	**[2 5]**	**[3 5]**
		Testing	**[2 5]**	**[2 8]**	**[2 5]**	**[2 5]**	**[3 5]**

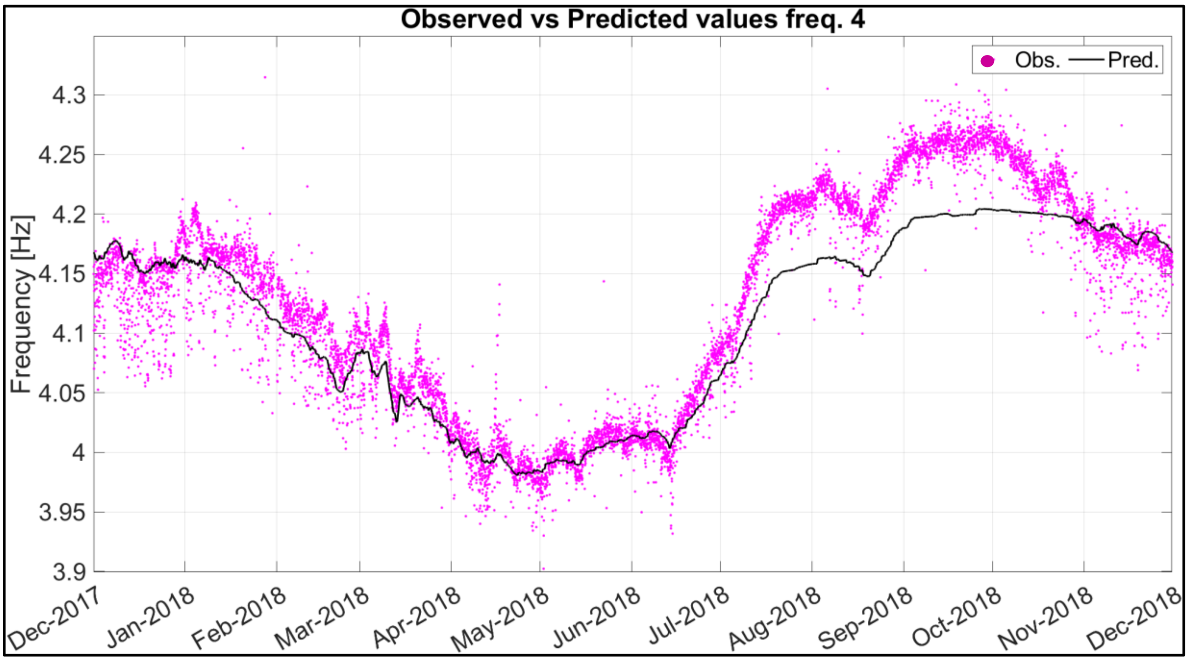
In order to verify the accuracy of the models several parameters were considered, such as the determination coefficient, standard deviation, and residuals. It can be observed that the MLP-NN showed better performance for both training and testing phases. However, it is important to note that both models successfully achieved the main objective. Even though the regression model should not be disregarded as it still provides acceptable results, counting as a viable model.

Additionally, for the neural network model, the number of hidden layers showed in the table are considered the best solution for each frequency. For instance, for the first frequency, the parameter [2 5] was obtained. This means that each column represents a hidden layer, and the value shown indicates the number of neurons each hidden layer contains. Explaining, for the first vibration mode, there are 2 hidden layers where the first hidden layer consists of 2 neurons, and the second hidden layer consists of 5 neurons. Using the same criteria, the number of neurons and hidden layers for each natural frequency is explained.

Fig. 37 compares the performance of the linear regression and neural network models focusing solely on the testing set. Additionally, frequency 4, ranging from 3.9 to 4.35 Hz, is highlighted. It is evident that both models successfully predicted the dynamic behaviour under the same conditions. This comparison reveals that each method has its unique approach to learn from the data and providing the most accurate results possible.

Furthermore, it was observed that both models fell short within the vibration range where neither could predict accurately, with the MLP-NN model showing closer predictions. It is noted that the training period occurred under certain conditions that could be considered ideal, whereas the testing set faced different conditions, leading to discrepancies in predicting vibration modes. Additionally, the coefficients of determination were slightly lower for the shown testing set.

Regression model (testing Freq. 4)



Neural network (testing Freq. 4)

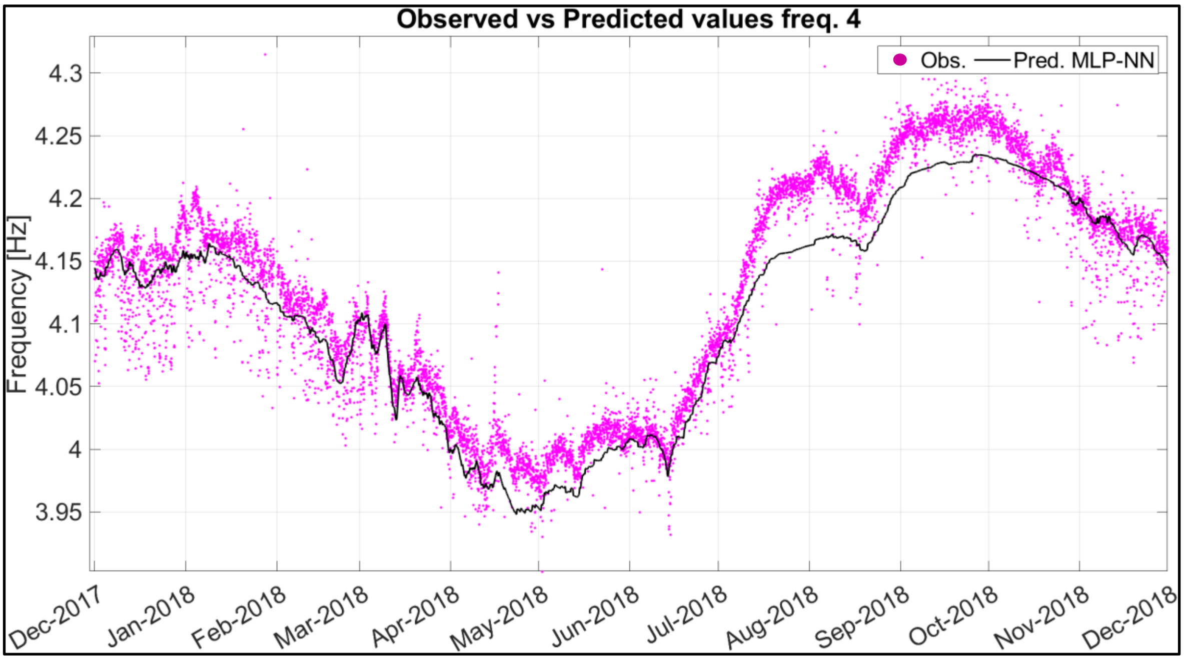
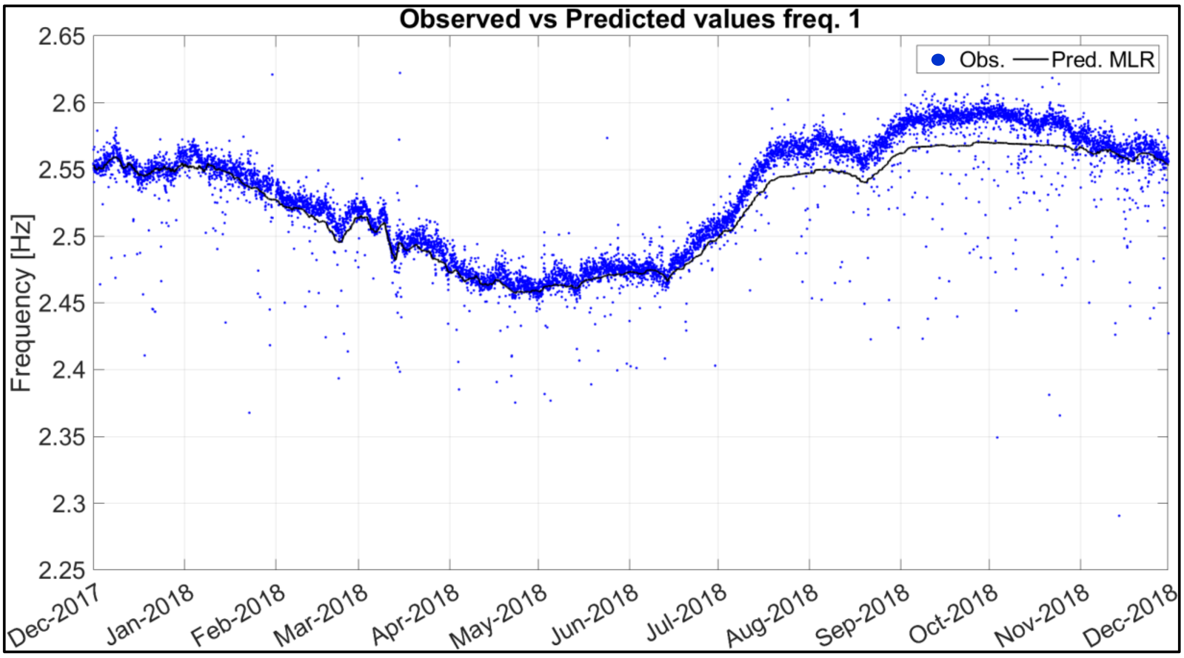


Fig. 37 - Comparison of the MLR and MLP-NN model predictions and the observed values for the frequency mode 4 in the test set.

On the other hand, Fig. 38 shows the comparison between both models used; the linear regression model and the neural network model, specifically focusing on the first mode of vibration, which ranges from 2.3 to 2.65 Hz. It can be observed that both models successfully achieved their goal of predicting dynamic behaviour under the same conditions.

The neural network model exhibited good learning capabilities, resulting in accurate predictions almost on par with those of the regression model.

Regression model (testing Freq. 1)



Neural network (testing Freq. 1)

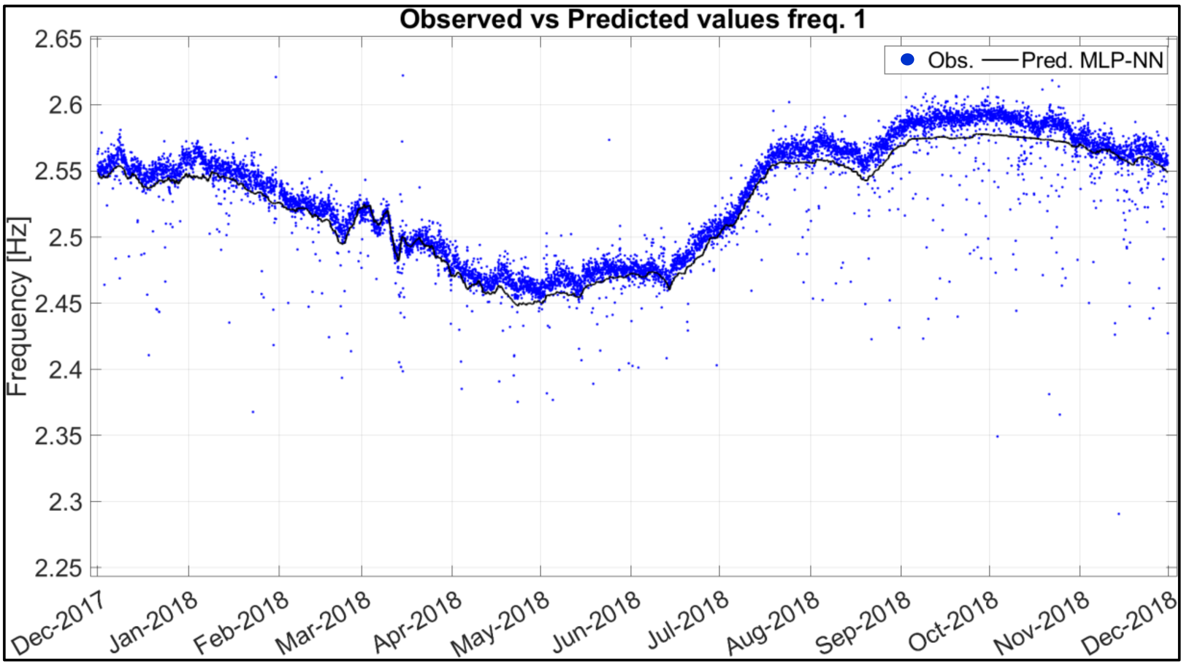


Fig. 38 - Comparison of the MLR and MLP-NN model predictions and the observed values for the frequency mode 1 in the test set.

It is also worth noting that as frequency increases, the level of noise tends to increase as well. Therefore, special care must be taken when working with higher frequencies, as predictions may vary slightly compared to working with lower frequencies. This variability can impact the accuracy of the models, necessitating more sophisticated approaches to handle higher-frequency data.

5.

FINAL CONSIDERATIONS

5.1. WORK ACCOMPLISHED

The developed work aimed to dynamically monitor a concrete dam through MLR and MLP-NN models by tracking the evolution of its dynamic behaviour, ensuring that it stays within expected variations. This activity relies on observation through on-site instrumentation and behaviour models. These models are characterized by pattern recognition, which allows the characterization of a structure's normal behaviour without explicitly relying on principles of mechanics.

The primary goal was to construct, calibrate, and compare statistical and machine learning models to support the interpretation of observed dynamic behaviour in a functioning concrete dam. This goal was reached, as the observed patterns and behaviour of the structure under everyday events were successfully analyzed.

As a result, a greater understanding and knowledge were achieved across the various chapters outlined at the beginning of this thesis. The following sections provide a detailed explanation of the findings and insights gained from this work:

- In Chapter 1, the work focused on establishing a comprehensive foundation by contextualizing the importance of monitoring concrete dam safety, outlining the research objectives and contributions, and providing a structured thesis outline. By emphasizing dynamic monitoring and the integration of statistical and machine learning models, this chapter set the stage for the detailed exploration and analysis that followed, aiming to enhance the understanding and management of concrete dam behaviour.
- In Chapter 2, the focus was on the monitoring and observation of concrete dams, introducing the concept of structural safety monitoring activities and emphasizing its importance in preventing structural failures and potential disasters. The chapter also provided a historical overview of Portugal's dams, highlighting their key characteristics. This included a look at existing dams in

Portugal and the most significant ones in terms of dimensions and storage capacity. As concrete dams are more stable over time, they are increasingly used. Specific activities related to dam safety control were studied, such as periodic inspections, analysis of historical data, and preventive measures to avoid structural problems. The regulations controlling dams were also discussed, addressing their classification, safety, and size, with a safety control cycle encompassing environmental safety, structural safety, and operational safety. An analysis of the main actions impacting a concrete dam and the expected structural behaviour was also conducted, focusing on the effects caused by thermal and seasonal variations.

- Chapter 3, the different approaches were explored to interpreting observed behaviour in concrete dams, emphasizing the importance of understanding how structures respond to varying conditions. Two main approaches for linear regression models and machine learning to interpret structural behaviour were described: the HST approach (Hydrostatic, Seasonal, Temperature) and the HTT approach. The primary task was to base the work on the existing data and opt for the HST approach. Both models were analyzed under this approach, using temperature and reservoir water level data from three years of continuous monitoring. Linear regression and neural networks worked successfully to identify the patterns in the models and the consistent behaviour of the Baixo Sabor dam, allowing for predictions with unseen data based on this extensive knowledge.
- Chapter 4 focused on a specific case study: the Baixo Sabor dam. The work began by determining the variables to use according to the HST approach. For both the MLR and MLP-NN methods, the first two years as the training set were used, with the third year for testing. Once both models were created, various graphs and tables were presented, showing the results from both models to better understand and visualize the outcomes. This helped identify which model had better performance for the case study. The chapter also discussed the different approaches and detailed the advantages and disadvantages of each model.

5.2. RESULTS DISCUSSION

Both methods fulfilled their function of capturing and predicting information for an unseen season, with both models considered functional based on performance, although selecting an appropriate method can be complex and dependent on various factors. Each project is unique and should be carefully analyzed.

For the Baixo Sabor dam case, based on the obtained results, the following considerations can be made regarding which criterion or model is necessary:

Overall Performance: Both methods performed well in developing data, suggesting that both approaches are suitable for the problem at hand.

Prediction Accuracy: The neural network slightly outperformed the regression model in terms of prediction accuracy. This suggests that the neural network was better able to capture the relationship between input features and the target variable compared to the regression model.

Model Flexibility: The neural network can capture complex relationships between input features and the target variable, explaining its better performance on a dataset. On the other hand, the linear regression model is more dependent on the input parameters' shape.

Model Interpretation: The linear regression model is easier to interpret than the neural network, as relationships between input features and the target variable are linear and easily interpreted. However, the neural network might be considered as a black-box model by some users, making it more challenging to understand its behaviour. Nevertheless, once the model is understood, it becomes more flexible and provides more control in analyzing and making sense of the data.

In summary, both linear regression and neural networks are applicable methods for any project, depending on the complexity of the data and the relevance of the results. For this particular project, focusing on natural frequencies, and environmental effects, and taking into account the magnitude of the project as the Baixo Sabor dam is, the neural network is a solid option due to its ability to capture complex relationships between input features and the target variable.

5.3. FUTURE WORK

Within the scope of this work and for future considerations, it is recommended to expand this model to different types of dams. The presented model provides a solid foundation for replicating this work to other dams and analyzing each of them to have a wide knowledge of the behaviour of dams.

Another suggestion is to apply the model to the same dam studied in this project with the aim of obtaining updated data. This would help to determine if the findings from this project remain accurate over the years and understand the impact of the dam's ageing process on the data.

For future analysis, given that the testing set exhibited different behaviour compared to the trained model, it is suggested to apply an HTT model. This could help to understand whether temperature is one of the main factors influencing the behaviour of dams.

Developing this model considering higher modes of vibration is also an interesting topic to explore. The current model has proven to be reliable, and expanding the range of vibration modes could provide a better interpretation of results, especially when dealing with higher modes of vibration. This could lead to more comprehensive insights into the dynamic behaviour of dams.

BIBLIOGRAPHIC REFERENCES

Agência Portuguesa do Ambiente, I.P. A.P.A. (2018). *Documentos técnicos de apoio ao regulamento de segurança de barragens (RSB)* (In Portuguese). (1ª Edition). Comissão dos Regulamentos de Barragens. Lisbon, Portugal.

Agência Portuguesa do Ambiente, A.P.A. (2021). *Barragens de Portugal* (in Portuguese). Retrieved from { HYPERLINK "https://apambiente.pt/prevencao-e-gestao-de-riscos/barragens-de-portugal" }. Website visited 22/May/2024.

Allemang, R. J. and Brown, D. L. (1982). *A Correlation Coefficient for Modal Vector Analysis*. Paper presented at Proceedings of the 1st International Modal Analysis Conference, Orlando, 8-10 November 1982, pp. 110-116.

Allemang, R. J. (2003). *The Modal Assurance Criterion (MAC): Twenty Years of Use and Abuse*. Structural Dynamics Research Laboratory. University of Cincinnati, Ohio.

Ayodele, T. (2010). *Types of machine learning algorithms*. New advances in machine learning, (pp. 19-48), Yagang Zhang, In Tech, Rijeka.

Banoula, M. *An Overview on Multilayer Perceptron (MLP)*. AI & Machine Learning, 10/August/2023, Retrieved from { HYPERLINK "https://www.simplilearn.com/tutorials/deep-learning-tutorial/multilayer-perceptron" }. Website visited 24/May/2024.

Batista, A. L. (1998). *Análise do Comportamento ao Longo do Tempo de Barragens Abóbada* (in Portuguese). Doctoral thesis in Civil Engineering, Technical University of Lisbon, Lisbon.

Batista, A. L (2022). *Deterioration processes of dams affected by concrete swelling reactions. The Portuguese experience in monitoring and rehabilitation*. Paper presented at 16th International Conference on Alkali-Aggregate Reaction in Concrete, Lisbon.

Bendat, J. S., & Piersol, A. G. (1980). *Engineering Applications of Correlation and Spectral Analysis*. New York, USA: John Wiley & Sons.

Brincker, R., (2014). *Some Elements of Operational Modal Analysis*. Shock and Vibration. 2014. 1-11. 10.1155/2014/325839. Department of Engineering, Aarhus University, Denmark. http://dx.doi.org/10.1155/2014/325839

Castro, A. T. (1998). *Métodos de Retroanálise na interpretação do comportamento de barragens de betão* (in Portuguese). Doctoral thesis in Civil Engineering, Instituto Superior Técnico IST, Lisbon.

Catarino, A. J. (2016). *Modelação do comportamento dinâmico de sistemas barragem-fundação-albufeira: análise sísmica de uma barragem de abóbadas múltiplas* (in Portuguese). Master's thesis in Civil Engineering, Technical University of Lisbon, Lisbon.

Chen, SH. (2015). *Actions on Hydraulic Structures and Their Effect Combinations. In: Hydraulic Structures*. Springer, Berlin, Heidelberg. doi.org/10.1007/978-3-662-47331-3_4

Cunha, A., & Caetano, E. (2006). *Experimental modal analysis of civil engineering structures*. Sound and Vibration, 40(6), 12-20.

Elmenshawy, M. R. (2015). *Static And Dynamic Analysis of Concrete Gravity Dams*. Master's degree in Civil Engineering (Irrigation and Hydraulics Engineering), Faculty of Engineering– Tanta University. doi: 10.13140/RG.2.2.19941.86242

Faria, J., (2022). *Baixo Sabor dam site description*.[PowerPoint slides]. Faculdade de Engenharia da Universidade do Porto, Porto, Portugal. Presentation retrieved from https://drive.google.com/file/d/1a36no8JSy60BMn4Y2L346FVfOnVVPT7b/view?usp=sharing.

Gomes, A. F. S. (1986). *A observação no controlo de segurança das barragens de betão portuguesas* (in Portuguese). Relatório 302/86 – NO, Laboratório Nacional de Engenharia Civil (LNEC), Lisbon.

Hastie, T., Tibshirani, R., Friedman, J. (2009). *The Elements of Statistical Learning. Basis Expansions and Regularization.* 2nd Edition, (pp. 139-189). New York, USA.

DOI: 10.1007/b94608

Haykin, S. (2009). *Neural Networks and Learning Machines*, 3rd Edition, Pearson Education Inc., Nova Jersey.

Howard, T. H. Retrieved from { HYPERLINK "https://tataandhoward.com/a-history-of-dams-from-ancient-times-to-today/" }. Website visited 14/5/2024.

ICOLD European Club, (2020). *Dam Legislation Report - 2020*. (pp. 13-14). Lisbon.

Jurafsky, D., Martin, J. (2024). Chapter 7 - Neural Networks and Neural Language Models. (pp. 16-25), Standford University. California.

Keim, R. *How to Train a Multilayer Perceptron Neural Network*. All about circuits, 26/December/2016, Retrieved from https://www.allaboutcircuits.com/technical-articles/how-to-train-a-multilayer-perceptron-neural-network/. Website visited 20/ May / 2024.

Lee, I.; Shin, Y. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. Business Horizons, Volume 63, 2nd Edition, March-April of 2020, pág. 157-170, Indiana.

Leung, K. (2008). *Introduction to artificial neural networks.* Department of Computer and Information Science, Polytechnic University, New York.

LNEC, FEUP, & Ambisig. (2015). *Barragem de montante do aproveitamento hidroelétrico do Baixo Sabor. Caracterização do comportamento dinâmico da barragem através de monitorização em contínuo* (in Portuguese). Installation report.

Lusa, (2023, 02 November). *Más de 30 presas están a más del 80% de su capacidad* (in Portuguese). The Portugal News. { HYPERLINK "https://www.theportugalnews.com/es/noticias/2023-11-02/mas-de-30-presas-estan-a-mas-del-80-de-su-capacidad/82904" }.

Lynch, A. *Case Study: Malpasset Dam (France, 1959).* Association of State Dam Safety Officials, Lesson Learned. Retrieved from https://damfailures.org/case-study/malpasset-dam-france-1959/. Website visited 14/May/2024.

Magalhães, F., Cunha, Á., & Caetano, E. (2008). *Dynamic monitoring of a long span arch bridge.* Engineering Structures, 30(11), 3034-3044. doi:10.1016/j.engstruct.2008.04.020

Magalhães, F., Cunha, Á., Caetano, E. (2009). *Online automatic identification of the modal parameters of a long span arch bridge.* Mechanical Systems and Signal Processing, 14/ May/2008. Elsevier. doi:10.1016/j.ymssp.2008.05.003

Magalhães, F. (2010). *Operational Modal Analysis for Testing and Monitoring of Bridges and Special Structures*. (PhD), Faculty of Engineering, University of Porto, Porto.

Marcelino, J., Ramos, J. M., & Santos, L. O. (2008). *Observação do comportamento estrutural de barragens e de pontes. Critérios e métodos* (in Portuguese). 5º Congresso Luso-Moçambicano de Engenharia, 2º Congresso de Engenharia de Moçambique, Maputo, September 2nd to 4th, 2008.

Martins, L., Mata, J., & Ribeiro, A. (2011a). *A Qualidade das Medições de Deslocamento Radial e Tangencial pelo Método do Fio-de-Prumo no Controlo da Segurança Estrutural de Barragens de Betão*. Medições E Ensaios (in Portuguese). CONFMET, 4/November/2010, Lisbon.

Martins, L., Mata, J., Ribeiro, A. (2012b). *A Qualidade das Medições de Deslocamento Radial e Tangencial pelo Método do Fio-de-Prumo no Controlo da Segurança Estrutural de Barragens de Betão*. Medições e ensayos (in Portuguese), March/2012, Vol. 32, Filipe, E., Ribeiro, A., Girão, P., Godinho, I., Pellegrino, O., Lisbon.

Mata, J. (2007). *Aplicação de redes neuronais ao controlo de segurança de barragens de betão* (in Portuguese). Master's thesis in Structural Engineering, Instituto Superior Técnico.

Mata, J., Tavares de Castro, A., & Sá da Costa, J. (2014). *Constructing statistical models for arch dam deformation*. Structural Control and Health Monitoring, Vol. 21, 3rd Edition, pag. 423-437. https://doi.org/10.1002/stc.1575

Mathworks. (2022). Matlab. R2022a.

McNeill. S. (2013). *A Modal Identification Algorithm Combining Blind Source Separation and State Space Realization.* Journal of Signal and Information Processing, 9/March/2013, Scientific Research. http://dx.doi.org/10.4236/jsip.2013.42025

Mendes, J. M. (2023). *Identificação de novidades no comportamento observado em barragens de betão com recurso a modelos de machine learning* (in Portuguese). Master's thesis in Civil Engineering, Faculty of Engineering, University of Porto, Porto.

Mikkelsen, P. E. (2003). *Piezometers in Fully Grouted Boreholes*. Symposium on Field Measurements in Geomechanics, FMGM 2003, Oslo, Norway, September.

Miranda, R. J. (2017). *Análise do Comportamento Sísmico de uma Barragem Gravidade Tipo* (in Portuguese). Master's thesis in Civil Engineering, Higher Institute of Engineering of Lisbon, Lisbon.

Moore, J. D., Henderson, A., Fletcher, J. S., Lockyer, N. P., & Vickerman, J. C. (2013). *Peak picking as a pre-processing technique for imaging time of flight secondary ion mass spectrometry*. Surface and Interface Analysis, 45(1), 461-465. Manchester. https://doi.org/10.1002/sia.5062

Morais, R., S. (2017) *Monitorização de deslocamentos em grandes barragens utilizando GNSS: aplicação à barragem do Cabril* (in Portuguese). Master's thesis in Civil Engineering, Higher Institute of Engineering of Lisbon, Lisbon.

Moura, G. M. G. (2005). *Determinação de tensões a partir de extensões observadas em barragens de betão* (in Portuguese). Master's thesis in Civil Engineering, Faculty of Engineering, University of Porto, Porto.

Novak, P., Moffat, A. I. B., & Nalluri, C. (2007). *Hydraulics in Civil and Environmental Engineering*, (T. Francis Ed. Fourth ed.). London and New York.

Oliveira, S., Ferreira, I., Berberan, A., Mendes, P., Boavida, J., Baptista, B. (2010). *Monitoring the structural integrity of large concrete dams: the case of Cabril dam, Portugal*. Paper presented at Hydropower 2010, Lisbon.

Pedro, J. O. (2001). Segurança e funcionalidade das barragens (in Portuguese). Memória N.º 824, LNEC, Lisbon.

Peeters, B. (2000). *System Identification and Damage Detection in Civil Engineering*. (PhD thesis), Katholieke Universiteit Leuven.

Pereira, S. (2019). *Structural condition assessment of dams based on continuous dynamic monitoring.* Doctoral thesis in Civil Engineering, Faculty of Engineering, University of Porto, Porto.

Pereira, S., Magalhães, F., Gomes, J., Cunha, Á., Lemos, J. (2018). *Dynamic monitoring of a concrete arch dam during the first filling of the reservoir*. Engineering Structures, 3/August/2018, Elsevier. https://doi.org/10.1016/j.engstruct.2018.07.076

Pereira Gomes, J., Magalhães, F., Monteiro, G., Palma, J., Pereira, S., & Silva Matos, D. (2018). *Seismic Monitoring System of Baixo Sabor Scheme for Structural Dynamic Behaviour Monitoring and Risk Management*. Paper presented at the ICOLD 2018, Vienna, Austria.

Posada, M., Grande, M., Blancat, C. (2017). *Manual de salud y seguridad en el trabajo. Hormigón elaborado Plantas* (in Spanish). Unión Obrera de la Construcción de la República Argentina UOCRA, Buenos Aires, Argentina.

Power Technology. *Power plant profile: Baixo Sabor, Portugal*. 31/January/2024, Retrieved from { HYPERLINK "https://www.power-technology.com/marketdata/power-plant-profile-baixo-sabor-portugal/" }. Website visited 7/May/2024.

Sysoi. (2017, September 6). *Branches of Artificial Intelligence.* [Post]. Facebook. https://www.facebook.com/SysoiTechnologies/photos/a.114131395356785/1071949122908336/?type=3

Ramírez, A. I. *La seguridad de presas desde la perspectiva hidrológica* (in Spanish). Paper presented at Conferencia Enzo Levi 2010, Mexico.

Shiksha Online. *What is a multilayer perceptron (MLP) neural network*. 23/January/2024. { HYPERLINK "https://www.shiksha.com/online-courses/articles/understanding-multilayer-perceptron-mlp-neural-networks/" }. Website visited 21/05/2024.

Silva, C. *Central da barragem do Alto Tâmega em funcionamento em março de 2024* (in Portuguese). 16/August/2023. { HYPERLINK "https://www.industriaeambiente.pt/noticias/central-barragem-alto-tamega-funcionamento-marco-2024/" }. Website visited 19/05/2024.

Silva, J., (2022). *Desenvolvimento de modelos de Machine learning baseados em Dados de monitorização contínua de Barragens de betão para Interpretação do comportamiento Estrutural observado* (in Portuguese). Master's thesis in Civil Engineering, Faculty of Engineering, University of Porto, Porto.

Silva, J. A, Mata, J., Pereira, S., & Cunha, A. (2022). *Desenvolvimento de modelos de interpretação quantitativa do tipo HTT com recurso a medições da temperatura do ar. Aplicação aos deslocamentos horizontais observados na Barragem do Baixo Sabor* (in Portuguese). JPEE2022-6, 6ª Jornadas Portuguesas de Engenharia de Estruturas, 2022.

SNIRH, Sistema Nacional de Informação de Recursos Hídricos. *Boletim de Armazenamento nas Albufeiras de Portugal Continental* (in Portuguese). 4/May/2024. Retrieved from { HYPERLINK "https://snirh.apambiente.pt/index.php?idMain=1&idItem=1.3" }. Website visited 19/05/2024.

Sobrevilla, M. (2020). *La seguridad de presas en el Perú* (in Spanish). https://www.cip.org.pe/publicaciones/2020/enero/la-seguridad-de-presas-en-el-peru.pdf. 23/May/2024.

Sundar, K.P., Shrimali, M.K., Bharti, S.D., Vibhute, A. (2023). *Seismic Response Analysis of Concrete Gravity Dam Using Response Spectrum Analysis and Response History Analysis*. In: Shrikhande, M., Agarwal, P., Kumar, P.C.A. (eds) Proceedings of 17th Symposium on Earthquake Engineering (Vol. 2). Lecture Notes in Civil Engineering, Vol 330. Springer, Singapore.

https://doi.org/10.1007/978-981-99-1604-7_16

Tanchev, L. (2014). *Dams and Appurtenant Hydraulic Structures (2nd edition)*. CRC Press/Balkema, Leiden.