PhD Thesis

PhD Program in Structural Analysis

Department of Civil and Environmental Engineering
Barcelona School of Civil Engineering

Universitat Politècnica de Catalunya

A machine learning based methodology for anomaly detection in dam behaviour

Author:
Fernando Salazar González

Advisors:
Prof. Eugenio Oñate Ibáñez de Navarra
Prof. Miguel Ángel Toledo Municio

February 18, 2017

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola de Doctorat

Curso académico: 2016/2017

Acta de calificación de tesis doctoral

Nombre y apellidos

FERNANDO SALAZAR GONZÁLEZ

Programa de doctorado

ANÁLISIS ESTRUCTURAL

Unidad estructural responsable del programa

Departamento de Ingeniería Civil y Ambiental

Resolución del Tribunal

Reunido el Tribunal designado al efecto, el doctorando expone el tema de su tesis doctoral titulada:

A machine learning based methodology for anomaly detection in dam behaviour

Acabada la lectura y después de dar respuesta a las cuestiones formuladas por los miembros titulares del tribunal, éste otorga la calificación:

☐ NO APTO ☐ APROBADO ☐ NOTABLE ☐ EXCELENTE

(Nombre, apellidos y firma) Presidente		(Nombre, apellidos y firma) Secretario
(Nombre, apellidos y firma) Vocal	(Nombre, apellidos y firma) Vocal	(Nombre, apellidos y firma) Vocal

_____, _____ de _____ de _____

El resultado del escrutinio de los votos emitidos por los miembros titulares del tribunal, efectuado por la Comisión Permanente de la Escuela de Doctorado, otorga la MENCIÓN CUM LAUDE:

☐ SÍ ☐ NO

(Nombre, apellidos y firma) Presidente de la Comisión Permanente de la Escuela de Doctorado	(Nombre, apellidos y firma) Secretario/a de la Comisión Permanente de la Escuela de Doctorado

Barcelona, _____ de _____ de _____

Abstract

Dam behaviour is difficult to predict with high accuracy. Numerical models for structural calculation solve the equations of continuum mechanics, but are subject to considerable uncertainty as to the characterisation of materials, especially with regard to the foundation. As a result, these models are often incapable to calculate dam behaviour with sufficient precision. Thus, it is difficult to determine whether a given deviation between model results and monitoring data represent a relevant anomaly or incipient failure.

By contrast, there is a tendency towards automatising dam monitoring devices, which allows for increasing the reading frequency and results in a greater amount and variety of data available, such as displacements, leakage, or interstitial pressure, among others.

This increasing volume of dam monitoring data makes it interesting to study the ability of advanced tools to extract useful information from observed variables.

In particular, in the field of Machine Learning (ML), powerful algorithms have been developed to face problems where the amount of data is much larger or the underlying phenomena is much less understood.

In this thesis, the possibilities of machine learning techniques were analysed for application to dam structural analysis based on monitoring data. The typical characteristics of the data sets available in dam safety were taking into account, as regards their nature, quality and size.

A critical literature review was performed, from which the key issues to consider for implementation of these algorithms in dam safety were identified.

A comparative study of the accuracy of a set of algorithms for predicting dam behaviour was carried out, considering radial and tangential displacements and leakage flow in a 100-m high dam. The results suggested that the algorithm called "Boosted Regression Trees" (BRT) is the most suitable, being more accurate in general, while flexible and relatively easy to implement.

At a later stage, the possibilities of interpretation of the mentioned algorithm were evaluated, to identify the shape and intensity of the association between external variables and the dam response, as well as the effect of time. The tools were applied to the same test case, and allowed more accurate identification of the time effect than the traditional statistical method.

Finally, a methodology for the implementation of predictive models based on BRT for early detection of anomalies was developed and implemented in an interactive tool that provides information on dam behaviour, through a set of selected devices. It allows the user to easily verify whether the actual data for each of these devices are within a pre-defined normal operation interval.

Resumen

El comportamiento estructural de las presas de embalse es difícil de predecir con precisión. Los modelos numéricos para el cálculo estructural resuelven bien las ecuaciones de la mecánica de medios continuos, pero están sujetos a una gran incertidumbre en cuanto a la caracterización de los materiales, especialmente en lo que respecta a la cimentación. Como consecuencia, frecuentemente estos modelos no son capaces de calcular el comportamiento de las presas con suficiente precisión. Así, es difícil discernir si un estado que se aleja en cierta medida de la normalidad supone o no una situación de riesgo estructural.

Por el contrario, muchas de las presas en operación cuentan con un gran número de aparatos de auscultación, que registran la evolución de diversos indicadores como los movimientos, el caudal de filtración, o la presión intersticial, entre otros. Aunque hoy en día hay muchas presas con pocos datos observados, hay una tendencia clara hacia la instalación de un mayor número de aparatos que registran el comportamiento con mayor frecuencia.

Como consecuencia, se tiende a disponer de un volumen creciente de datos que reflejan el comportamiento de la presa, lo cual hace interesante estudiar la capacidad de herramientas desarrolladas en otros campos para extraer información útil a partir de variables observadas.

En particular, en el ámbito del aprendizaje automático (machine learning), se han desarrollado algoritmos muy potentes para entender fenómenos cuyo mecanismo es poco conocido, acerca de los cuales se dispone de grandes volúmenes de datos.

En la tesis se ha hecho un análisis de las posibilidades de las técnicas más recientes de aprendizaje automático para su aplicación al análisis estructural de presas basado en los datos de auscultación. Para ello se han tenido en cuenta las características habituales de las series de datos disponibles en las presas, en cuanto a su naturaleza, calidad y cantidad.

Se ha realizado una revisión crítica de la bibliografía existente, a partir de la cual se han identificado los aspectos clave a tener en cuenta para implementación de estos algoritmos en la seguridad de presas.

Se ha realizado un estudio comparativo de la precisión de un conjunto de algoritmos para la predicción del comportamiento de presas considerando desplazamientos radiales, tangenciales y filtraciones. Para ello se han utilizado datos reales de una presa bóveda. Los resultados sugieren que el algoritmo denominado “Boosted Regression Trees” (BRTs) es el más adecuado, por ser más preciso en general, además de flexible y relativamente fácil de

implementar.

En una etapa posterior, se han identificado las posibilidades de interpretación del citado algoritmo para extraer la forma e intensidad de la asociación entre las variables exteriores y la respuesta de la presa, así como el efecto del tiempo. Las herramientas empleadas se han aplicado al mismo caso piloto, y han permitido identificar el efecto del tiempo con más precisión que el método estadístico tradicional.

Finalmente, se ha desarrollado una metodología para la aplicación de modelos de predicción basados en BRTs en la detección de anomalías en tiempo real. Esta metodología se ha implementado en una herramienta informática interactiva que ofrece información sobre el comportamiento de la presa, a través de un conjunto de aparatos seleccionados. Permite comprobar a simple vista si los datos reales de cada uno de estos aparatos se encuentran dentro del rango de funcionamiento normal de la presa.

Acknowledgements

I want to express my deepest gratitude to my supervisors, Prof. Oñate and Prof. Toledo. They encouraged me to enrol in this mission that, fortunately, has come to fruition. Working with them is a privilege as engineer and researcher, though almost insignificant if compared to what I learned as a person.

I also want to thank my colleagues at CIMNE and UPM, who helped me so much, one way or another, to accomplish this task. In chronological order, thanks to those I met the very first day I visited CIMNE (Antonia, Miguel Ángel, Salva), who inoculated me the virus of research, even though none of us was aware at that time.

From that moment on, many people enriched my path by sharing their achievements and drawbacks. The list is long, but I don't want to forget Riccardo, José Manuel, Benjamín and Juan Miquel.

Special thanks to my colleagues at CIMNE Madrid, who provided interesting comments and suggestions, always useful. Thanks to Javier, Joaquín and David.

I appreciate the support of Rafa, who showed me the path, and that of the rest of the members of the Serpa group at UPM.

Also chronologically, I must finish by mentioning Javi, who encouraged me (and strongly contributed) to wrap up the application and host it in the cloud for secure remote access. He was the closest representative of the community of software developers willing to share their knowledge. I could'n have dreamed of doing anything like this work without people like Dean Attali and Max Woolf.

This research would have been impossible without the interest of Carlos Barbero (Catalan Water Agency), in charge of safety, maintenance and operation of La Baells dam, who provided real monitoring data.

The research has been partially supported by the Spanish Ministry of Economy and Competitiveness (*Ministerio de Economía y Competitividad*, MINECO) through the projects iComplex (IPT-2012-0813-390000) and AIDA (BIA2013-49018-C2-1-R and BIA2013-49018-C2-2-R).

To my family (in the broadest sense of the word), and to those who flew too high, too soon, so unfairly.

"If all a man has is a hammer, then every problem looks like a nail."

Abraham Maslow

"You were right about the stars. Each one is a setting sun"

Jeff Tweedy (Wilco), "Jesus etc.", Yankee Hotel Foxtrot, Nonesuch Records

"El mejor momento de las cosas es cuando no han pasado, porque luego todo lo que puede hacerse es comentarlo"

Manolo Martínez (Astrud), "La nostalgia es un arma", Mi fracaso personal, Virgin/Chewaka

Contents

List of Figures

List of Tables

Nomenclature

Mathematical Symbols

a	Magnitude of the artificial anomaly introduced
$a_1 \dots a_{10}$	Coefficients in the HST formula
d	Number of days since 1 January
$F(x_i)$	Output of predictive model for t_i
$F_n(X^j)$	Output of an ensemble model at iteration n
$f_m(X^j)$	Weak learner fitted at iteration m
F_2	Index of accuracy of anomaly detection models
h	Reservoir level
s	Argument of the trigonometrical functions in the HST model. $s = 2\pi d/365.25$
Mn	n Version of a BRT model
min	minimum
μ	Residuals average
N	Number of records in a period
ν	Regularisation parameter
p	Number of inputs
R	Model residuals (prediction-observation)
σ^2	Variance
S_m	Subsample of a training set used to fit $f_m(X^j)$
sd_{res}	Standard deviation of the residuals
t	time
t_{det}	Detection time
X^j	Input variable
x_i^j	Observed value for input X^j at time t_i
x_k^j	Equally-spaced values of X^j to be used in PDPs
Y	Measured response variable
$\hat{Y}$	Predicted response variable
y_i	Observed value of the output variable at time t_i
$\bar{y_i}$	Mean of y_i

Mathematical tools

ANFIS	Adaptive Neuro-Fuzzy System
ARX	Auto-Regressive Exogenous
BRT	Boosted regression Trees
FEM	Finite Element Method
GA	Genetic Algorithms
HST	Hydrostatic Season Time
HSTT	Hydrostatic Season Temperature Time
KNN	K-Nearest Neighbours
MARS	Multivariate Adaptive Regression Splines
ML	Machine Learning
MLP	Multi Layer Perceptron
MLR	Multilinear regression
NN	Neural Network
PCA	Principal Component Analysis
RF	Random Forest
SVM	Support Vector Machine

Other Acronyms

ACA	Agència Catalana de l'Aigua
ARV	Average Relative Variance
KDE	Kernel Density Estimation
MAE	Mean Absolute Error
MINECO	Ministerio de Economía y Competitividad
MSE	Mean Squared Error
OOR	Out of range
PDP	Partial dependence plot
RI	Relative Influence

1 Introduction and Objectives

1.1 Introduction

Dams play a key role in our society, since they provide essential services to our way of living, such as flood defence, water storage and power generation. Moreover, an eventual failure might have catastrophic consequences in terms of casualties, economic and environmental losses, as was unfortunately verified in the past [19].

As a consequence, safe dam operation needs to be ensured, and potentially anomalous performance shall be detected as early as possible, to avoid serious malfunctioning or failure. While the first objective is achieved by means of an appropriate maintenance program both for the structure and the hydro-electromechanical devices, failure prevention by early detection of anomalies is primarily based on surveillance tasks [32], [33].

In turn, surveillance is based on two main pillars [32]: a) visual inspection and b) monitoring of dam and foundation. Its main objective is to reduce the probability of failure [33].

Lombardi [40] formulated the objectives of dam and foundation monitoring in a concise way, by posing four questions to be answered:

1. Does the dam behave as expected/predicted?
2. Does the dam behave as in the past?
3. Does any trend exist which could impair its safety in the future?

4. Was any anomaly in the behaviour of the dam detected?

The answer to these questions requires the analysis of dam monitoring data two ways:

- In the short term (some times "on-line"), the measurements of some devices are compared to reference values, which correspond to the dam response to the concurring loads in "normal" or "safe" condition. These reference values and associated prediction intervals above and below them are obtained from some behaviour model, which accounts for the actual value of the acting loads. Those measurements outside the cited interval are considered as potential symptoms of anomalous behaviour, hence further verified.
- In the medium to long term, behaviour models and observed data are analysed to draw conclusions on the overall dam performance. In particular, the association between each load and output is observed, and the evolution over time is evaluated.

The result of this analysis is essential in dam safety assessment and decision making, together with the rest of available information about dam construction and operation, including visual inspection. Figure 1.1 shows schematically the monitoring data analysis process.

1.2 Motivation

Dam monitoring data analysis, and the answer to the above mentioned questions, require a behaviour model that provides an estimate on the response of the structure at a given time, taking into account the acting loads.

Existing models can be classified as follows [73]:

- Deterministic: typically based on the finite element method (FEM), these methods calculate the dam response on the basis of the physical governing laws.
- Statistical: exclusively based on dam monitoring data.
- Hybrid: deterministic models which parameters have been adjusted to fit the observed data.
- Mixed: comprised by a deterministic model to predict the dam response to hydrostatic pressure, and a statistical one to consider deformation due to thermal effects.

Numerical models based on the FEM provide useful estimates of dam displacements and stresses, but are subject to a significant degree of uncertainty in the characterisation of the materials, especially with respect to the structural behaviour of the foundation and the thermal evolution of the dam body in concrete (particularly arch) dams. Other assumptions

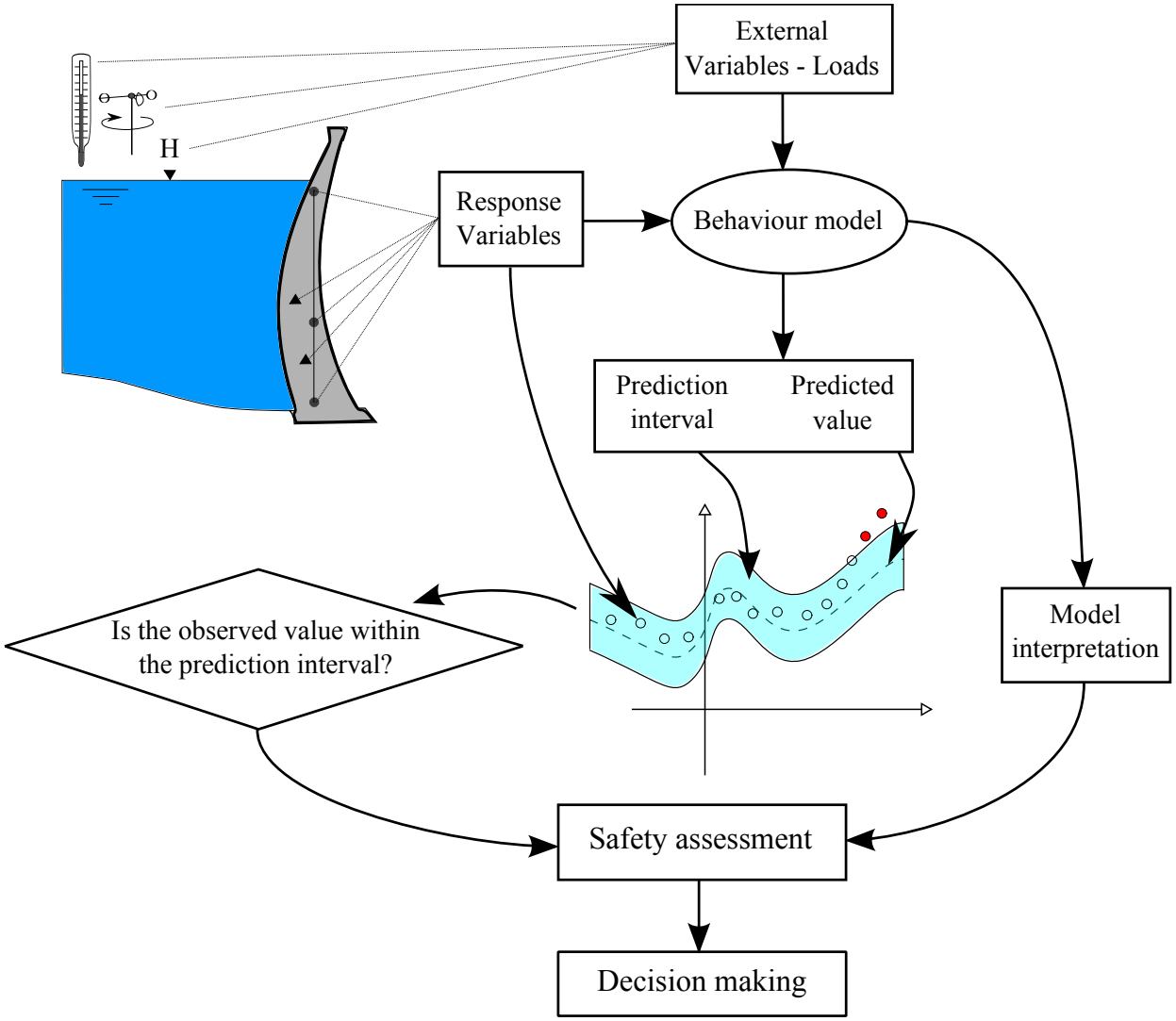


Figure 1.1: Flow diagram of dam monitoring data analysis.

and simplifications have to be made, regarding geometry and boundary conditions. These tools are essential during the initial stages of the life cycle of the structure, provided that not enough data are available to build data-based predictive models. However, their results are often not accurate enough for a precise assessment of dam safety.

This is more acute when dealing with determined variables such as leakage in concrete dams and their foundations, due to the intrinsic features of the physical process, which is often non-linear [14], and responds to threshold and delayed effects [71], [41]. Numerical analysis cannot deal with such a phenomenon, because comprehensive information about the location, geometry and permeability of each fracture would be needed. Other phenomena are also difficult to reproduce with numerical models, such as the beginning of failure by concrete plasticising or cracking, although tools have been developed for this purpose [49].

These drawbacks are shared by all approaches that make use of a FEM model: deterministic, hybrid and mixed.

Many of the dams in operation have a number of monitoring devices that record the evolution of various indicators such as displacements, leakage flow or pore water pressure, among others. Although there are still many dams with few observed data, there is a clear trend towards the installation of a larger number of devices with higher data acquisition frequency [33]. As a result, there is an increasing amount of information on dam performance.

Statistical tools employed in regular engineering practice for dam monitoring data analysis are relatively simple. They are frequently limited to graphical exploration of the time series of data [46], along with simple statistical models [33], [83]. The hydrostatic-season-time (HST) model [81] is the most widely applied, and the only generally accepted by practitioners.

HST is based on multiple linear regression considering the three most influential external variables: hydrostatic load, air temperature and time. It often provides useful estimations of displacements in concrete dams [75], and does not require air temperature time series data (it is assumed to follow a constant yearly cycle). Moreover, the resulting model is easily interpretable, since the contribution of each input is assumed to be cumulative.

Nonetheless, HST also features conceptual limitations that damage the prediction accuracy [75] and may lead to misinterpretation of the results [1]. For example, it is based on the assumption that the hydrostatic load and the temperature are independent, whereas it is well known that they are coupled, since the thermal field is influenced by the the water temperature in the upstream face [67]. On another note, it lacks flexibility, since the functions have to be defined beforehand, and thus may not represent the true behaviour of the structure [71]. Also, they are not well-suited to model non-linear interactions between input variables [14].

In the recent years, non-parametric techniques have emerged as an alternative to HST for building data-based behaviour models [61], e.g. support vector machines (SVN) [55], neural networks (NN) [42], adaptive neuro-fuzzy systems (ANFIS) [17], among others [61]. In general, these tools are more suitable to model non-linear cause-effect relations, as well as interaction among external variables, as that previously mentioned between hydrostatic load and temperature. On the contrary, they are typically more difficult to interpret, what led them to be termed as "black box" models (e.g. [3]). As a consequence, the vast majority of related works are limited to the verification of their prediction accuracy when estimating determined output variables (e.g. [56], [62], [38]).

Therefore, dam engineers face a dilemma: the HST model is widely known and used and easily interpretable. However, it is based on some incorrect assumptions, and its accuracy can be increased. On the other hand, more flexible and accurate models are available, but they are more difficult to implement and analyse.

The research aims at solving this issue by exploring the possibilities of machine learning

algorithms to improve dam monitoring data analysis and safety assessment.

1.3 Objectives

The main objective is the development of a methodology for dam behaviour analysis based on machine learning, efficient in early detection of anomalies. To achieve that goal, the following specific objectives need to be fulfilled:

1. **Literature review on data-based models for dam monitoring data analysis**, with focus on the following topics:
    - Critical analysis of relevant articles and conference proceedings.
    - Identification of areas to improve in the field of dam monitoring data interpretation.
    - Revision of the statistical and machine learning tools with potential for application to the problem to be solved.
    - Verification of the applicability of each tool to predict output variables of different nature.
    - Analysis of the key methodological issues as regards the implementation of predictive models in day-to-day practice.
    - Selection of a group of algorithms for a more detailed analysis.
2. **Algorithm selection**, in terms of accuracy, flexibility, robustness and ease of implementation.
3. **Analysis of the effect of the training set size**, to have an estimate on the time period required from the first filling before having the possibility of employing some data-based behaviour model.
4. **Identification of tools for interpretation of ML models**, i.e., analysis of the influence of each input on dam response and retrospective assessment of dam performance to detect potential changes in time.
5. **Implementation of the methodology in a software tool for anomaly detection**, with the following functionalities:
    - Accuracy: the better the prediction of the model fits the actual response of the dam, the more reliable the conclusions drawn from its interpretation [9]. Moreover, a more accurate model will result in a narrower prediction interval which in turn would allow earlier anomaly detection.

- Flexibility: each dam typology presents different characteristics in terms of the most influential loads, the strength and nature of their association with dam response, the most representative output variables and the potential failure modes, among other aspects. The behaviour model should ideally be able to adapt to highly different situations.
- Interpretability: model analysis should throw information on the nature and intensity of the association between each input and response, and in particular on the time effect, i.e., whether dam performance changed over time, and which way.
- Ability to detect anomalies: a criterion to determine a prediction interval around the model prediction is required, to classify upcoming observations of different output variables as "normal" or "potentially anomalous".
- Ability to identify extraordinary situations due to load combination.
- A graphical user interface for its practical application, including tools for data exploration, model fitting and anomaly detection.

1.4 Publications

This thesis is presented as a compendium of articles, previously published in indexed scientific journals. The list and the association with this document follows:

Chapter 2 contains a summary of the articles related to the literature review:

- Salazar, F., Morán, R., Toledo, M.Á., Oñate, E. Data-Based Models for the Prediction of Dam Behaviour: A Review and Some Methodological Considerations. *Archives of Computational Methods in Engineering* (2015). doi:10.1007/s11831-015-9157-9
- Salazar, F., Toledo, M.Á., Discussion on "Thermal displacements of concrete dams: Accounting for water temperature in statistical models", *Engineering Structures*, Available online 13 August 2015, ISSN 0141-0296, http://dx.doi.org/10.1016/j.engstruct.2015.08.001.

Chapter 3 is a summary of the article dealing with algorithm selection, based on a comparison of candidate techniques:

- Salazar, F., Toledo, M.Á., Oñate, E., Morán, R. An empirical comparison of machine learning techniques for dam behaviour modelling, *Structural Safety*, Volume 56, September 2015, Pages 9-17, ISSN 0167-4730, http://dx.doi.org/10.1016/j.strusafe.2015.05.001.

Chapter 4 focuses on model interpretation, and is associated with the fourth paper in the compendium:

- Salazar, F., Toledo, M.Á., Oñate, E., Suárez, B. Interpretation of dam deformation and leakage with boosted regression trees, *Engineering Structures*, Volume 119, 15 July 2016, Pages 230-251, ISSN 0141-0296, http://dx.doi.org/10.1016/j.engstruct.2016.04.012.

The overall methodology for anomaly detection is described in Chapter 5. It takes into account the conclusions of the precedent works, and is the subject of another article currently under review.

Finally, part of the work was presented in the following conferences:

Salazar, F., Oñate, E., Toledo, M.Á. *Posibilidades de la inteligencia artificial en el análisis de auscultación de presas.* III Jornadas de Ingeniería del Agua, Valencia (Spain), October 2013 (in Spanish).

Salazar, F., Morera, L., Toledo, M.Á., Morán, R., Oñate, E. *Avances en el tratamiento y análisis de datos de auscultación de presas.* X Jornadas Españolas de Presas, Spancold, Sevilla (Spain), February 2015 [1] (in Spanish).

Salazar, F., Oñate, E., Toledo, M.Á. *Nuevas técnicas para el análisis de datos de auscultación de presas y la definición de indicadores cuantitativos de su comportamiento*, IV Jornadas de Ingeniería del Agua, Córdoba (Spain), October 2015.

Salazar, F., González, J.M., Toledo, M.Á., Oñate, E. *A methodology for dam safety evaluation and anomaly detection based on boosted regression trees.* 8^{th} European Workshop on Structural Health Monitoring, Bilbao (Spain), July 2016.

A copy of the post-print version of the articles is included in Appendix A, while the works presented in conferences form Appendix B.

Therefore, Chapters 2, 3 and 4 include a summary of the methods and results of the correspondent articles, while Chapter 5 contains the final part of the research, in which the previous results were taken into account.

[1] Section 3 of this paper was carried out by León Morera, thus is not part of this thesis

2

State of the art review

2.1 Introduction

A literature review was performed on a selection of articles and conference proceedings featuring examples of application of data-based models in dam behaviour modelling. This chapter includes a summary of this analysis.

In what follows, $Y \in \mathbb{R}$ stands for some response variable (e.g. displacement, leakage flow, crack opening, etc.), which is estimated in terms of a set of inputs X^j: $Y \approx \hat{Y} = F(X^j)$. The observed values are denoted as $(x_i^j, y_i), i = 1, ..., N$, where N is the number of observations and $j = 1 \dots p$ refer to the dimensions of the input space.

2.2 Statistical and machine learning techniques used in dam monitoring analysis

2.2.1 Models based on linear regression

The Hydrostatic-Season-Time model (HST)

Linear regression is the simplest statistical technique, appropriate to reproduce certain phenomena. It is also the basis of the most popular data-based behaviour model in dam engineering: the Hydrostatic-Season-Time (HST). It was first proposed by Willm and Beaujoint in 1967 [81].

It is based on the assumption that the dam response is a linear combination of three effects:

$$\hat{Y} = F_1(h) + F_2(s) + F_3(t) \tag{2.1}$$

- A reversible effect of the hydrostatic load which is commonly considered in the form of a fourth-order polynomial of the reservoir level (h) ([73], [5], [71]):

$$F_1(h) = a_0 + a_1 h + a_2 h^2 + a_3 h^3 + a_4 h^4 \tag{2.2}$$

- A reversible influence of the air temperature, which is assumed to follow an annual cycle. Its effect is approximated by the first terms of the Fourier transform:

$$F_2(s) = a_5 cos(s) + a_6 sen(s) + a_7 sen^2(s) + a_8 sen(s) cos(s) \tag{2.3}$$

  where $s = 2\pi d/365.25$ and d is the number of days since 1 January.

- An irreversible term due to the evolution of the dam response over time. A combination of monotonic time-dependant functions is frequently considered. The original form is [81]:

$$F_3(t) = a_9 log(t) + a_{10} e^t \tag{2.4}$$

The model parameters $a_1 ... a_{10}$ are adjusted by the least squares method: the final model is based on the values which minimise the sum of the squared deviations between the model predictions and the observations.

The main advantages are:

- It frequently provides useful estimations of displacements in concrete dams [75].
- It is simple and thus easily interpretable: the effect of each external variable can be isolated in a straightforward manner, since they are assumed to be cumulative.
- Since the thermal effect is considered as a periodic function, the time series of air temperature are not required. This widens the possibilities of application, as only the reservoir level variation needs to be available to build an HST model.
- It is well known by practitioners and frequently applied in several countries [75].

It also features relevant limitations:

- The functions have to be defined beforehand, and thus may not represent the true behaviour of the structure [71].

- The governing variables are supposed to be independent, although some of them have been proven to be correlated [74].
- They are not well-suited to model non-linear interactions between input variables [14].

Several alternatives have been proposed to overcome these shortcomings. Penot *et al.* [50] introduced the HSTT method, in which the thermal periodic effect is corrected according to the actual air temperature.

Related approaches also based on linear regression were applied in dam safety, often by means of the addition of further input variables following some heuristics or after a trial-and-error process [42], [14], [71], [83], [10]. In all cases, the need to make a priori assumptions about the model remains, although variable selection procedures have also been proposed, such as Stojanovic *et al.* [72], who combined greedy MLR with variable selection by means of genetic algorithms (GA).

Consideration of delayed effects

It is well known that dams respond to certain loads with some delay [41]. The most typical examples are the change in pore pressure in an earth-fill dam due to reservoir level variation [6] and the influence of the air temperature in the thermal field in a concrete dam body [71].

Several alternatives have been proposed to account for these effects. The most popular is based on an enrichment of the linear regression by including moving averages or gradients of some explanatory variables in the set of predictors. Guedes and Coelho [27] predicted the leakage flow on the basis of the mean reservoir level over the course of a five-days period. Sánchez Caro [64] included the 30 and 60 days moving average of the reservoir level in the conventional HST formulation to predict the radial displacements of El Atazar Dam. Further examples are due to Popovici *et al.* [53] and Crépon and Lino [15].

A more formal alternative to conventional HST to account for delayed effects was proposed by Bonelli [7], [5]. It was intended to account for the delayed response of an arch dam in terms of the temperature field, with the final aim of predicting radial displacements. Lombardi *et al.* [40] suggested an equivalent formulation, also to compute the thermal response of the dam to changes in air temperature. Although the formulation differs from a multiple linear regression, its numerical integration leads to a predictive model which is a linear combination of:

- the value of the predictors at t_i and t_{i-1}.
- the value of the output variable at t_{i-1}.

which is the conventional form of a first order auto-regressive exogenous (ARX) model.

This is the most enriched version of multiple linear regression, where predictors of different types are combined. This gives greater flexibility to the algorithm to adapt to different

situations or response variables. By contrast, the number of potential inputs can become very large, which generally leads to the need for some variable selection procedure. For example, Piroddi and Spinelli [52] applied a specific algorithm for selecting 11 out of 40 predictors considered. Principal component analysis (PCA) was also employed for variable selection (e.g. [44], [13], [14]).

A further drawback of linear regression with many input variables is that model interpretation becomes difficult, since the contribution of each predictor is harder to isolate.

Moreover, the use of the previous (lagged) value of the output to calculate a prediction for current record may induce to question a) whether the observed previous value or the precedent prediction should be used, and b) whether the model parameters should be readjusted at every time step.

In addition, current and previous values of response variables different from the target variable (e.g. radial displacements or leakage) can be considered as inputs. They implicitly encompass information from unrecorded or unknown phenomena, so the resulting model will probably be more accurate. However, it can also "learn" the anomalous behaviour and consider it as normal, in which case it would be inappropriate to detect anomalies.

The higher accuracy obtained by increasing the information given to the model invites exploring the utility of this approach, keeping their limitations in mind.

2.2.2 Machine learning based models

Among the non-conventional data-based algorithms, neural networks (NNs) are by far the most popular in the field of dam monitoring data analysis. NN models are flexible, and allow modelling complex and highly non-linear phenomena. Most of the published works employ the conventional multi-layer perceptron (MLP) and some sigmoid as the activation function.

These models often result in greater accuracy than MLR, due to the higher flexibility. However, the results are highly dependent on some issues to be determined by the user:

1. The network architecture, i.e., number of layers and perceptrons in each layer, which is not known beforehand. Some authors focus on the definition of an efficient algorithm for determining an appropriate network architecture [66], whereas others use conventional cross-validation [42] or a simple trial and error procedure [76].

2. The training process, which may reach a local minimum of the error function. The probability of occurrence of this event can be reduced by introducing a learning rate parameter [76].

3. The stopping criterion, to avoid over-fitting. Various alternatives are suitable for solving this issue, such as *early stopping* and *regularisation* [28].

The fitting procedures greatly differ among authors. While Simon *et al.* [71] trained an MLP with three perceptrons in one hidden layer for 200,000 iterations, Tayfur *et al.* [76] used regularisation with 5 hidden neurons and 10,000 iterations. Neither of them followed any specific criterion to set the number of neurons. For his part, Mata [42] tested NN architectures with one hidden layer having 3 to 30 neurons on an independent test data set. He repeated the training of each NN model 5 times with different initialisation of the weights.

It can be concluded that NNs share some of the target features (flexibility, accuracy), but lack ease for implementation and robustness. Model interpretation is not straightforward, and the results depend on the initialisations, so several models need to be trained and their results averaged to increase robustness. Moreover, only numerical inputs can be considered, which need to be normalised for model fitting (and de-normalised afterwards).

Other ML approaches were also applied in dam safety, such as Adaptive neuro-fuzzy systems (ANFIS) ([56], [82]), Support Vector Machines (SVM) ([12], [55]), or K-nearest neighbours (KNN) ([68]). They mostly share the mentioned properties of NNs: greater flexibility and accuracy, more difficult interpretation and potential over-fitting.

2.3 Methodological issues

Although each algorithm has its peculiarities, they all need to face intrinsic aspects of the problem to be solved, which can be analysed independently of the selected technique. Some of them have been mentioned before as variable selection. Others are specific to data-based prediction tasks, and in particular to the dam behaviour problem.

2.3.1 Input selection

The vast majority of statistical and ML algorithms are highly dependent on the inputs considered, which results in a need for input variable selection. The issue has arisen in combination with the use of NN [18], [57], [35], [39], [48], ARX [52], MLR [72] and ANFIS models [56].

The selection of predictors can be useful to reduce the dimensionality of the problem (essential for ARX models), as well as to facilitate the interpretation of the results.

The criterion to be used depends on the type of data available, the main objective of the study (prediction or interpretation), and the characteristics of the phenomenon to be modelled. Engineering judgement is thus essential to make these decisions.

By contrast, some ML algorithms are insensitive to the presence of highly-correlated or uninformative predictors, such as those based on decision trees. Boosted regression trees (BRTs) and random forests (RFs) stand out among those included in this category, though they are relatively new and unknown for most dam engineers.

2.3.2 Model interpretation

There is an obvious interest in model interpretation to analyse the effect of each input on dam response, once the parameters have been fitted. This contributes to answer the first question posed by Lombardi [40]: does the dam behave as expected/predicted? For example, an arch dam is expected to move in the downstream direction in front of a combination of high hydrostatic load and low temperature.

The evolution over time is particularly relevant, since it is related to the second and third questions [40]:

- Does the dam behave as in the past?
- Does any trend exist which could impair its safety in the future?

The effect of time, hydrostatic load and temperature can be easily obtained from an HST model, since it is based on the assumption that they are additive. However, it was already mentioned that they are actually correlated. Paraphrasing Breiman [9], when a pre-defined model is fit to data, "the conclusions drawn are about the model's mechanism, not about nature's mechanism [1]. Moreover, "if the model is a poor emulation of nature, the conclusions may be wrong".

Therefore, the interpretation of a more accurate predictive model will offer more reliable conclusions. The price to be paid for the greater flexibility and accuracy is the more difficult interpretation.

The vast majority of published studies are limited to the analysis of model accuracy for the output variable under consideration, as compared to HST. Only a few come to deal with model interpretation, that is, to analyse the strength and nature of the contribution of each action to the dam response. They are often limited to cases where a low number of inputs are considered (e.g. [42], [65], [71], [53]).

2.3.3 Training and validation sets

Accuracy is the main (and most obvious) measure of model performance, i.e. how well the model predictions fit to the observed data. However, it is well known that an increase in the number of parameters results in models more susceptible to over-fit. The higher complexity of ML algorithms has a similar effect as regards over-fitting. Hence, model accuracy must be computed properly.

It has been proven that the prediction accuracy of a data-based model, measured on the training data, is an overestimation of its overall performance [2]. Therefore, part of the available data needs to be reserved for model accuracy estimation (validation set). In

[1] Breiman employs "nature" to denote any phenomenon partially understood, which associates the predictor variables to the outcome. In this research, "nature's mechanism" is homologous to "dam behaviour"

principle, any sub-setting of the available data into training and validation sets is acceptable, provided the data are independent and identically distributed (i.i.d.).

This is not the case in dam monitoring series, which are time-dependant in general. Moreover, the amount of available data is limited, what in turn limits the size of the training and validation sets. Ideally, both should cover all the range of variation of the most influential variables.

On another note, a minimum amount of data is necessary to build a predictive model with appropriate generalisation ability. Some authors estimate the minimum period to be 5 [73] to 10 years [16], though it is case-dependent.

A further problem for the application of data-based models is that transient phenomena take place during the first years of operation [40]. Therefore, data from that period should be analysed in detail, since it might not be representative of subsequent dam performance.

In spite of these issues, many authors use the training set for computing model generalisation capability, or use a small sample for validation. This raises doubts about the actual accuracy of these models, in particular of those more strictly data-based, such as NN or SVM.

The deviation between predictions and observations is essential for dam behaviour assessment [40]. Moreover, the prediction intervals are typically based on some multiple of the standard deviation of the residuals. Hence, the proper estimation of model accuracy, over an adequate validation set, is fundamental from a practical viewpoint.

This topic is covered in depth in Chapter 5.

2.3.4 Practical implementation

Despite the increasing amount of literature on the use of advanced data-based tools, very few examples described their practical integration in dam safety analysis. The vast majority were limited to the model accuracy assessment, by quantifying the model error with respect to the actual measured data.

The information provided by reliable automated systems, based on highly accurate models, can be a great support for decision making regarding dam safety [33], [32].

To achieve that goal, the outcome of the predictive model must be transformed into a set of rules that determine whether the system should issue a warning. The actions to be taken need to be defined on a case-by-case basis, taking into consideration the relevance of each device as regards the overall dam safety [40].

Actually, an overall analysis of the most representative instruments is recommended, to identify (and discard) any isolated reading error. Cheng and Zheng [12] proposed a procedure for calculating normal operating thresholds ("control limits"), and a qualitative classification of potential anomalies: a) extreme environmental variable values, b) global structure damage, c) instrument malfunctions and d) local structure damage.

A more accurate analysis could be based on the consideration of the major potential modes of failure to obtain the corresponding behaviour patterns and an estimate of how they would be reflected on the monitoring data. Mata *et al.* [43] employed this idea to develop a system that takes the measurements of several devices and classifies them as correspondent to normal or accidental situation. This scheme can be easily implemented in an automatic system, though requires a detailed analysis of the possible failure modes, and their numerical simulation to provide data with which to train the classifier.

2.4 Conclusions

There is a growing interest in the application of innovative tools in dam monitoring data analysis. Although only HST is fully implemented in engineering practice, the number of publications on the application of other methods has increased considerably in recent years, specially NN.

It seems clear that the models based on ML algorithms can offer more accurate estimates of the dam behaviour than the HST method in many cases. In general, they are more suitable to reproduce non-linear effects and complex interactions between input variables and dam response.

However, most of the published works refer to specific case studies, certain dam typologies or determined outputs. Many focus on radial displacements in arch dams, although this typology represents roughly 5% of dams in operation worldwide.

A useful data-based algorithm should be versatile to face the variety of situations presented in dam safety: different typologies, outputs, quality and volume of data available, among others. Data-based techniques should be capable of dealing with missing values and robust to reading errors.

These tools must be employed rigorously, given their relatively high number of parameters and flexibility, what makes them susceptible to over-fit the training data. It is thus essential to check their generalisation capability on an adequate validation data set, not used for fitting the model parameters.

The main limitation of these methods is their inability to extrapolate, i.e., to generate accurate predictions outside the range of variation of the training data. Therefore, before applying these models for predicting the dam response in a given situation, it should be checked whether the load combination under consideration lies within the values of the input variables in the training data set.

From a practical viewpoint, data-based models should also be user-friendly and easily understood by civil engineering practitioners, typically unfamiliar with computer science, who have the responsibility for decision making.

Finally, two overall conclusions can be drawn from the review:

- ML techniques can be highly valuable for dam safety analysis, though some issues remain unsolved.
- Regardless of the technique used, engineering judgement based on experience is critical for building the model, for interpreting the results, and for decision making with regard to dam safety.

3
Algorithm Selection

3.1 Introduction

In view of the conclusions of the literature review, a set of ML algorithms were selected for a detailed comparative analysis. The main features were already known, but there was a need for testing their appropriateness to build dam behaviour models.

A selection of algorithms were faced to a practical case study, and the results were compared. Specifically, the following techniques were considered: random forests (RF), boosted regression trees (BRT), support vector machines (SVM) and multivariate adaptive regression splines (MARS). Both HST and NN were also used for comparison purposes. Similar analyses had been previously performed in other fields of engineering, such as the prediction of urban water demand [29].

3.2 Case study

The data used for the study correspond to La Baells dam. It is a double curvature arch dam, with a height of 102 m, which entered into service in 1976. The monitoring system records the main indicators of the dam performance: displacement, temperature, stress, strain and leakage. The data were provided by the Catalan Water Agency (*Agència Catalana de l'Aigua*, ACA), the dam owner, for research purposes. Among the available records, the study focused on 14 variables: 10 correspond to displacements measured by pendulums (five radial and five

tangential), and four to leakage flow. Several variables of different types were considered in order to obtain more reliable conclusions. The details of the available data are included in the article, whereas the location of each monitoring device is depicted in Figure 3.1.

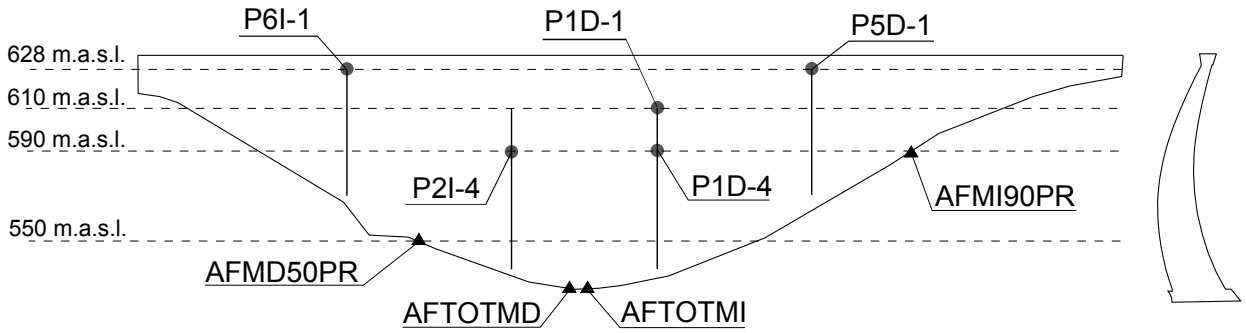


Figure 3.1: La Baells Dam geometry and location of the monitoring devices considered. Left: view from downstream. Right: highest cross-section.

The specific features of dam monitoring data analysis were taken into account to design the experiment. In all cases, approximately 40% of the records (from 1998 to 2008) were left out as the testing set. This is a large proportion compared with previous studies, which typically leave 10-20 % of the available data for testing [56], [42], [68]. A larger test set was selected in order to increase the reliability of the results.

On another note, it is well known that the early years of operation often correspond to a transient state, non-representative of the quasi-stationary response afterwards [40]. In such a scenario, using those years for training a predictive model would be inadvisable. This might lead to question the optimal size of the training set in achieving the best accuracy ([16], [14]). The available time series for La Baells dam span from 1979 to 2008. To analyse this issue, four different training sets were chosen to fit each model, spanning five, 10, 15 and 18 years of records. In all cases, the training data used correspond to the closest time period to the test set (e.g. periods 1993-1997, 1988-1997, 1983-1997, and 1979-1997, respectively).

The predictor set included inputs related to the environmental actions: air temperature and hydrostatic load. A time-dependent term was also added, to account for possible variations in dam behaviour over the period of analysis. Several variables derived from those actually measured at the dam site (reservoir level and the average daily temperature) were also included. They are listed in Table 3.2.

The variable selection was performed according to dam engineering practice. Both displacements and leakage are strongly dependant on hydrostatic load. Air temperature is well known to affect displacements, in the form of a delayed action. It may also influence leakage flow (as Seifart *et al.* reported for Itaipú Dam [70]), although it is uncertain (Simon *et al.* observed no dependency [71]). Both the air temperature and some moving averages were

Code	Group	Type	Period (days)
Level	Hydrostatic load	Original	-
Lev007	Hydrostatic load	Moving average	7
Lev014			14
Lev030			30
Lev060			60
Lev090			90
Lev180			180
Tair	Air temperature	Moving average	1
Tair007			7
Tair014			14
Tair030			30
Tair060			60
Tair090			90
Tair180			180
Rain	Rainfall	Accumulated	1
Rain030			30
Rain060			60
Rain090			90
Rain180			180
NDay	Time	Original	-
Year			-
Month	Season	Original	-
n010	Hydrostatic load	Rate of variation	10
n020			20
n030			30

Table 3.1: Predictor variables considered.

included in the analysis.

A relatively large set of predictors was used to capture every potential effect, overlooking the high correlation among some of them. The comparison sought to be as unbiased as possible, thus all the models were built using the same inputs[1] and data pre-process (only normalisation was performed when necessary). While it is acknowledged that this procedure may favour the techniques that better handle noisy or scarcely important variables, theoretically all learning algorithms should discard them automatically during the model

[1] with the exceptions of MARS and HST, as explained in the article

Type	Target	RF	BRT	NN	SVM	MARS	HST
Radial (mm)	P1DR1	1.70	0.93	**0.58**	0.75	2.32	1.35
	P1DR4	1.05	**0.71**	**0.68**	0.76	1.50	1.37
	P2IR4	0.94	0.97	1.02	1.05	**0.85**	1.12
	P5DR1	0.86	**0.70**	**0.64**	1.35	0.89	0.88
	P6IR1	1.47	0.69	0.72	**0.60**	1.67	0.91
Tangential (mm)	P1DT1	**0.24**	**0.25**	0.52	0.35	0.55	0.47
	P1DT4	**0.15**	**0.15**	0.18	0.19	0.22	0.20
	P2IT4	0.13	**0.11**	0.13	0.12	0.14	**0.10**
	P5DT1	0.40	0.22	**0.19**	0.38	0.47	**0.18**
	P6IT1	**0.28**	**0.27**	0.39	0.94	0.39	0.51
Leakage (l/min)	AFMD50PR	1.24	**0.90**	2.11	4.25	1.74	2.24
	AFMI90PR	0.18	0.15	**0.07**	0.33	0.25	0.28
	AFTOTMD	1.82	**1.60**	3.04	5.38	1.85	2.60
	AFTOTMI	0.91	**0.42**	0.83	1.49	1.49	1.11

Table 3.2: MAE for each output and model, fitted on the whole training set (18 years). The values within 10% from the minimum are highlighted in bold, and the minimum MAE are also underlined. The results correspond to the test set.

fitting.

3.3 Results and discussion

Table 3.2, contains the mean absolute error (MAE) for each target and model, computed as:

$$MAE = \frac{\sum_{i=1}^{N} \left| y_i - F\left(x_k^j\right) \right|}{N} \tag{3.1}$$

where N is the size of the training (or test) set, y_i are the observed outputs and $F(x_i)$ the predicted values.

It can be seen that models based on ML techniques mostly outperform the reference HST method. NN models yield the highest accuracy for radial displacements, whereas BRT models are better on average both for tangential displacements and leakage flow. It should be noted that the MAE for some tangential displacements is close to the measurement error of the device ($\pm 0.1mm$).

The effect of the training set size is depicted in Figure 3.2, where the model accuracy is

measured in terms of the average relative variance (ARV) [79]:

$$ARV = \frac{\sum_{i=1}^{N}\left(y_i - F\left(x_k^j\right)\right)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} = \frac{MSE}{\sigma^2} \tag{3.2}$$

where $\bar{y}$ is the output mean. Given that ARV denotes the ratio between the mean squared error (MSE) and the variance (σ^2), it accounts both for the magnitude and the deviation of the target variable. Furthermore, a model with ARV=1 is as accurate a prediction as the mean of the observed outputs.

Although the use of the whole training set is optimal for six out of 14 targets, significant improvements are reported in some cases by eliminating some of the early years. Surprisingly, for two of the outputs, the lower MAE corresponds to a model trained over five years, which in principle was assumed to be too small a training set. MARS is especially sensitive to the size of the training data. The MARS models trained on five years improve the accuracy for P1DR4 and P6IT1 by 13.3 % and 14.8 % respectively.

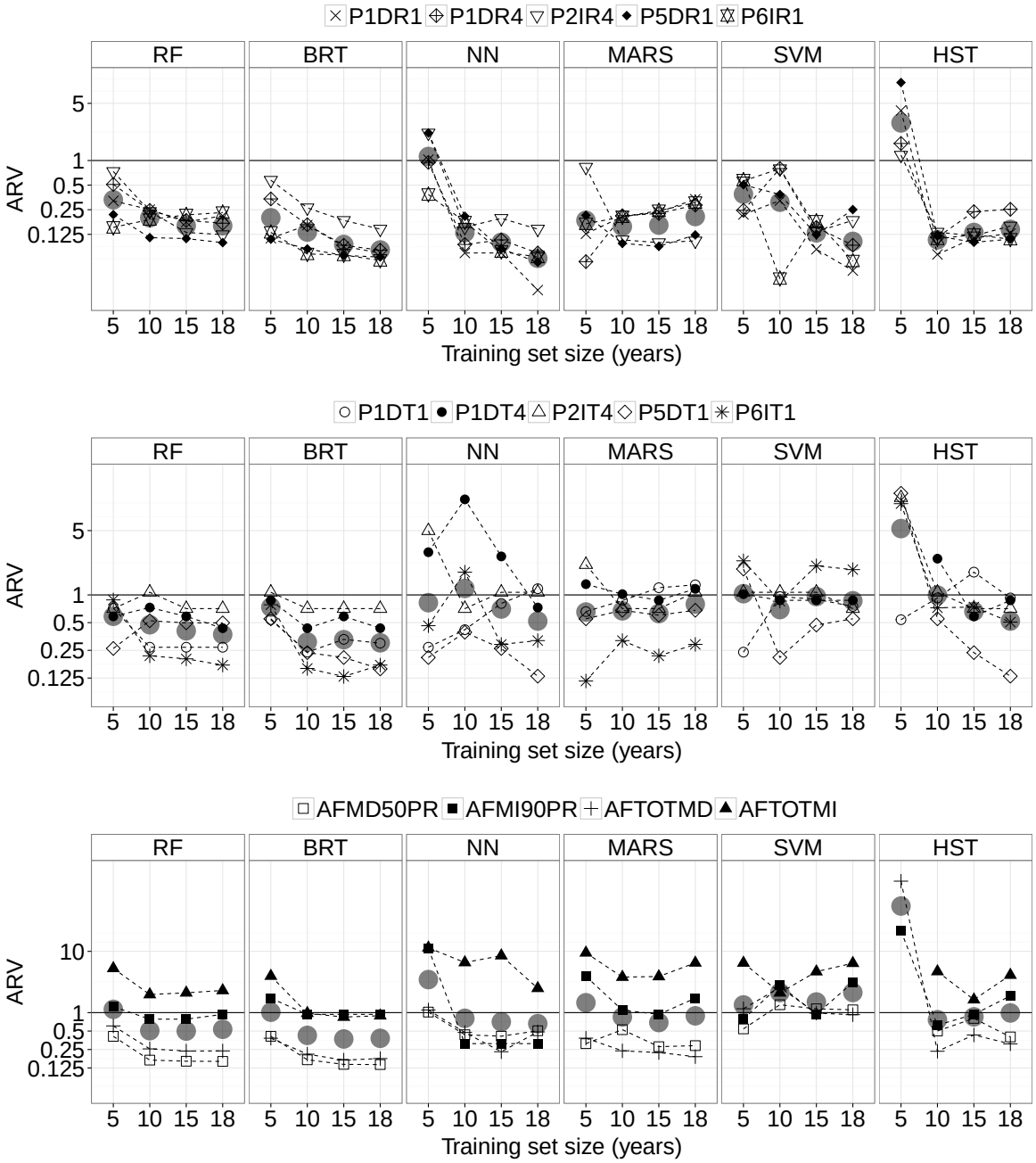


Figure 3.2: ARV for each model and training set size. Models with $ARV > 1.0$ are less accurate than the sample mean. The average values for each output, algorithm and training set size are plotted with black dots. Note the logarithmic scale of the vertical axis. Top: radial displacements. Middle: tangential displacements. Bottom: leakage flow. Some HST models trained over 5 years are out of the range of the vertical axis, thus highly inaccurate. The results correspond to the test set.

These results strongly suggest that it is advisable to select carefully the most appropriate training set size. This should be done by leaving an independent test set.

3.4 Conclusions

It was found that the accuracy of currently applied methods for predicting dam behaviour can be substantially improved by using ML techniques.

The sensitivity analysis to the training set size shows that removing the early years of dam life cycle can be beneficial. In this work, it has resulted in a decrease in MAE in some cases (up to 14.8%). Hence, the size of the training set should be considered as an extra parameter to be optimised during training.

Some of the techniques analysed (MARS, SVM, NN) are more susceptible to further tuning than others (RF, BRT), given that they have more hyper-parameters and are more sensitive to the presence of correlated or uninformative inputs. As a consequence, the former might have a larger margin for improvement than the latter.

However, both detailed tuning and careful variable selection increase the computational cost and complicate the analysis. Since the objective is the extension of these techniques for the prediction of a large number of variables of many dams, the simplicity of implementation is an aspect to be considered in model selection.

In this sense, BRT showed to be the best choice: it was the most accurate for five of the 14 targets; easy to implement; robust with respect to the training set size; able to consider any kind of input (numeric, categorical or discrete), and not sensitive to noisy and low relevant predictors.

4

Model Interpretation

4.1 Introduction

As a result of the comparative analysis, BRT was selected as the most appropriate tool to achieve the research objectives. In this stage, the possibilities of interpretation were investigated to:

1. Identify the effect of each external variable on the dam behaviour

2. Detect the temporal evolution of the dam response

3. Provide meaningful information to draw conclusions about dam safety

For this purpose, the same data from La Baells Dam were employed, though the analysis focused on 12 variables: 8 corresponded to radial displacements measured by pendulums (along the upstream-downstream direction), and four to leakage flow. The location of each monitoring device is depicted in Figure 4.1.

Since BRT models automatically discard those predictors not associated with the output [24], the initial model considered the same inputs as described in section 3. All the calculations were performed on a training set covering the period 1980-1997, and the model accuracy was assessed for a validation set correspondent to the years 1998-2008.

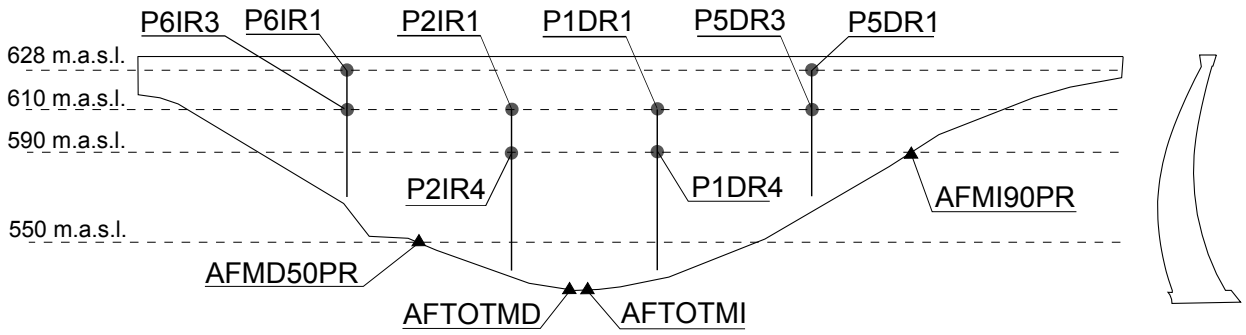


Figure 4.1: Geometry and location of the targets considered for model interpretation. Left: view from downstream. Right: highest cross-section.

4.2 Methods

4.2.1 Boosted regression trees

BRT models are built by combining two algorithms: a set of single models are fitted by means of decision trees [8], and their output is combined to compute the overall prediction using boosting [23]. For the sake of completeness, a short description of both techniques follow, although excellent introductions can be found in [58], [37], [21], [3].

Regression trees

Regression trees were first proposed as statistical models by Breiman *et al.* [8]. They are based on the recursive division of the training data in groups of "similar" cases. The output of a regression tree is the mean of the output variable for the observations within each group.

When more than one predictor is considered (as usual), the best split point for each is computed, and the one which results in greater error reduction is selected. As a consequence, non-relevant predictors are automatically discarded by the algorithm, as the error reduction for a split in a low relevant predictor will generally be lower than that in an informative one.

Other interesting properties of regression trees are:

- They are robust against outliers.
- They require little data pre-processing.
- They can handle numerical and categorical predictors.
- They are appropriate to model non-lineal relations, as well as interaction among predictors.

By contrast, regression trees are unstable, i. e., small variations in the training data lead to notably different results. Also, they are not appropriate for certain input-output relations, such as a straight 45° line [21].

Boosting

Boosting is a general scheme to build ensemble prediction models [23]. It is based on the generation of a (frequently high) number of simple models (also referred to as "weak learners") on altered versions of the training data. The overall prediction is computed as a weighted sum of the output of each model in the ensemble. The rationale behind the method is that the average of the prediction of many simple learners can outperform that from a complex one [69].

The idea is to fit each learner to the residual of the previous ensemble. The main steps of the original boosting algorithm when using regression trees and the squared-error loss function can be summarised as follows [45]:

1. Start predicting with the average of the observations (constant):

$$F_0\left(X^j\right) = f_0\left(X^j\right) = \bar{y}_i$$

2. For $m = 1$ to M

   (a) Compute the prediction error on the training set:

$$\tilde{y}_i = y_i - F_{m-1}\left(x_i^j\right)$$

   (b) Draw a random sub-sample of the training set (S_m)

   (c) Consider S_m and fit a new regression tree to the residuals of the previous ensemble:

$$\tilde{y}_i \approx f_m\left(X^j\right), i \in S_m$$

   (d) Update the ensemble:

$$F_m(X^j) \Leftarrow F_{m-1}(X^j) + f_m(X^j)$$

3. F_M is the final model

It is generally accepted that this procedure is prone to over-fitting, because the training error decreases with each iteration [45]. To overcome this problem, it is convenient to add a regularization parameter $\nu \in (0, 1)$, so that step (d) turns into:

$$F_m(X^j) \Leftarrow F_{m-1}(X^j) + \nu \cdot f_m(X^j)$$

Some empirical analyses showed that relatively low values of ν (below 0.1) greatly improve generalisation capability [23]. In practice, it is common to set the regularisation parameter and consider a number of trees such that the training error stabilises [58]. Subsequently, a certain number of terms are pruned using for example cross-validation. This is the approach employed in this work, with $\nu = 0.01$ and a maximum of 1,000 trees. It was verified that the training error reached the minimum before adding the maximum number of trees.

Five-fold cross-validation was applied to determine the amount of trees in the final ensemble. The process was repeated using trees of depth 1 and 2 (*interaction.depth*), and the most accurate for each target was selected. The rest of the parameters were set to their default values [25].

All the calculations were performed in the R environment [54].

Several procedures to interpret ML models, often termed "black box" models, can be found in the literature. In this work, the relative influence (RI) of each predictor and the partial dependence plots (PDP) were employed.

4.2.2 Relative influence (RI)

BRT models are robust against the presence of uninformative predictors, as they are discarded during the selection of the best split. Moreover, it seems reasonable to think that the most relevant predictors are more frequently selected during training. In other words, the relative influence (RI) of each input is proportional to the frequency with which they appear in the ensemble. Friedman [23] proposed a formulation to compute a measure of RI for BRT models based on this intuition. Both the relative presence and the error reduction achieved are considered in the computation. The results are normalised so that they add up to 100.

Based on this measurement, the most influential variables were identified for each output, and the results were interpreted in relation to dam behaviour. In order to facilitate the analysis, the RI was plotted as word clouds [36]. These plots resemble histograms, with the advantage of being more appropriate to visualise a greater set of variables. The code representing each predictor was displayed with a font size proportional to its relative influence with the library "wordcloud" [22].

Furthermore, two degrees of variable selection were applied, based on the RI of each predictor. First, a BRT model (M1) was trained with all the variables considered (section 5.2.3). Second, the inputs with $RI\left(X^j\right) > min\left(RI\left(X^j\right)\right) + sd\left(RI\left(X^j\right)\right)$ were selected to build a new model (M2). This criteria is heuristic and based on the *1-SE rule* proposed by Breiman *et al.* [8]. Finally, a model with three predictors was generated (M3), featuring the more relevant variables of each group: temperature, time and reservoir level for radial displacements, and rainfall, time and level for leakage flows.

These three versions were generated to analyse the effect of the presence of uninformative variables in the predictor set. Moreover, the simplest model facilitates the analysis, as the

effect of each action is concentrated in one single predictor.

In this sense, the temporal evolution is particularly relevant for dam safety evaluation, as it can help to identify a progressive deterioration of the dam or the foundation, which might result in a serious fault if not corrected.

4.2.3 Partial dependence plots

Multi-linear regression models and HST in particular are based on the assumption that the input variables are statistically independent, so the prediction is computed as the sum of their contributions. As a result, the effect of each predictor in the response can be easily identified, by plotting $F(X^j), \forall j = 1...p$.

This method is not appropriate for BRT models, as interactions among predictors are accounted for. While this results in more flexibility, it also implies that the identification of the relation between predictors and response is not straightforward.

Nonetheless, it is possible to examine the predictor-response relationship by means of the partial dependence plots [23]. This tool can be applied to any black box model, as it is based on the marginal effect of each predictor on the output, as learned by the model. Let X^j be the variable of interest. A set of equally spaced values are defined along its range: $X^j = x_k^j$. For each of those values, the average of the model predictions is computed:

$$\bar{F}\left(x_k^j\right) = \frac{1}{N}\sum_{i=1}^{N} F\left(x_i^j, x_i^{jc}\right) \tag{4.1}$$

where x_i^{jc} is the value for all inputs other than X^j for the observation i.

Similar plots can be obtained for interactions among inputs: the average prediction is computed for couples of fixed x_k^j, where j takes two different values. Hence, the results can be plotted as a three-dimensional surface (section 4.3.3). In this work, partial dependence plots were restricted to the simplest model, which considered three predictors. Therefore, three 3D plots allowed investigating the pairwise interactions among all the inputs considered in the simplified model.

4.2.4 Overall procedure

The complete process comprised the following steps:

1. Fit a BRT model on the training data with the variables in table 3.2 (M1).
2. Compute the RI and generate the word cloud.
3. Select the most relevant predictors with the 1-SE rule [8] and fit a new BRT model (M2).

4. Build a simple BRT model (M3) with the most influential variable of each group (temperature, level and time for displacements, and rainfall, level and time for leakage).

5. Generate the univariate and bivariate partial dependence plots for the simplest model.

6. Compute the goodness of fit for each model in both the training and the validation sets.

4.3 Results

4.3.1 Effect of input selection

Table 4.3.1 contains the error indices for each target. For those models with variable selection, the predictors are also listed. The results show that BRT efficiently discarded irrelevant inputs, since the fitting accuracy was similar for each version in most cases (i.e., the presence of uninformative predictors did not damage the fitting accuracy).

	Train		Validation		
Target	MAE	ARV	MAE	ARV	Inputs
	0.64	0.03	0.91	0.08	All
P1DR1	0.68	0.03	0.81	0.06	Tair090,Level,NDay,Lev007,Lev014
	0.69	**0.03**	**0.78**	**0.06**	**NDay,Tair090,Level**
	0.46	**0.03**	**0.65**	**0.08**	**All**
P1DR4	0.50	0.03	0.66	0.08	Level,Tair090,NDay,Lev007,Lev014,Lev030
	0.51	0.03	0.67	0.08	NDay,Tair090,Level
	0.66	0.03	1.03	0.09	All
P2IR1	0.85	0.05	1.09	0.09	Tair090,Level,Lev007,Lev014
	0.71	**0.04**	**0.98**	**0.08**	**NDay,Tair090,Level**
	0.48	**0.05**	**0.90**	**0.14**	**All**
P2IR4	0.61	0.06	0.93	0.14	Level,Tair090,Lev007,Lev014,Lev030
	0.53	0.06	0.94	0.16	NDay,Tair090,Level
	0.66	**0.05**	**0.82**	**0.08**	**All**
P5DR1	0.64	0.05	0.87	0.10	Tair060,Level,Tair030
	0.83	0.08	0.93	0.11	NDay,Tair060,Level
	0.25	**0.03**	**0.47**	**0.21**	**All**
P5DR3	0.33	0.05	0.55	0.22	Tair060,Level,Tair030
	0.31	0.04	0.52	0.24	NDay,Tair060,Level
	0.60	0.04	0.80	0.09	All
P6IR1	**0.65**	**0.05**	**0.78**	**0.08**	**Tair060,Tair030,Level,NDay**
	0.83	0.08	0.85	0.1	NDay,Tair060,Level
	0.23	**0.02**	**0.40**	**0.08**	**All**
P6IR3	0.37	0.05	0.67	0.17	Tair060,Level,Tair030
	0.29	0.03	0.43	0.09	NDay,Tair060,Level
	1.28	**0.16**	**0.93**	**0.19**	**All**
AFMD50PR	1.45	0.17	1.36	0.28	Level,Lev014,Lev007
	1.16	0.14	1.23	0.48	NDay,Rain090,Level
	0.08	0.09	0.15	0.51	All
AFMI90PR	0.08	0.10	0.12	0.45	Lev007,NDay,Level,Lev014,Lev030
	0.08	**0.10**	**0.12**	**0.46**	**NDay,Rain030,Lev007**
	1.64	**0.15**	**1.67**	**0.37**	**All**
AFTOTMD	1.87	0.19	1.73	0.45	Level,Lev007,Lev014
	1.69	0.18	1.97	0.52	NDay,Rain180,Level
	0.41	**0.11**	**0.44**	**0.40**	**All**
AFTOTMI	0.44	0.12	0.44	0.42	NDay,Lev060,Lev014,Lev007,Lev030,Lev180,Lev090,Level
	0.54	0.18	0.46	0.60	NDay,Rain180,Lev060

Table 4.1: Accuracy of each model and target for the training and validation sets. The results and inputs considered by the most accurate version are highlighted in bold.

4.3.2 Relative influence

The analysis of the wordclouds of RI allowed identifying some interesting features of La Baells dam behaviour. As for the radial displacements, (Figure 4.2), the thermal inertia was observed as higher RI for Tair060 and Tair090 than for Tair (which in fact resulted negligible). By contrast, the reservoir level at the date of the record was always more influential than all the moving averages, what reveals an immediate response of the dam to this load.

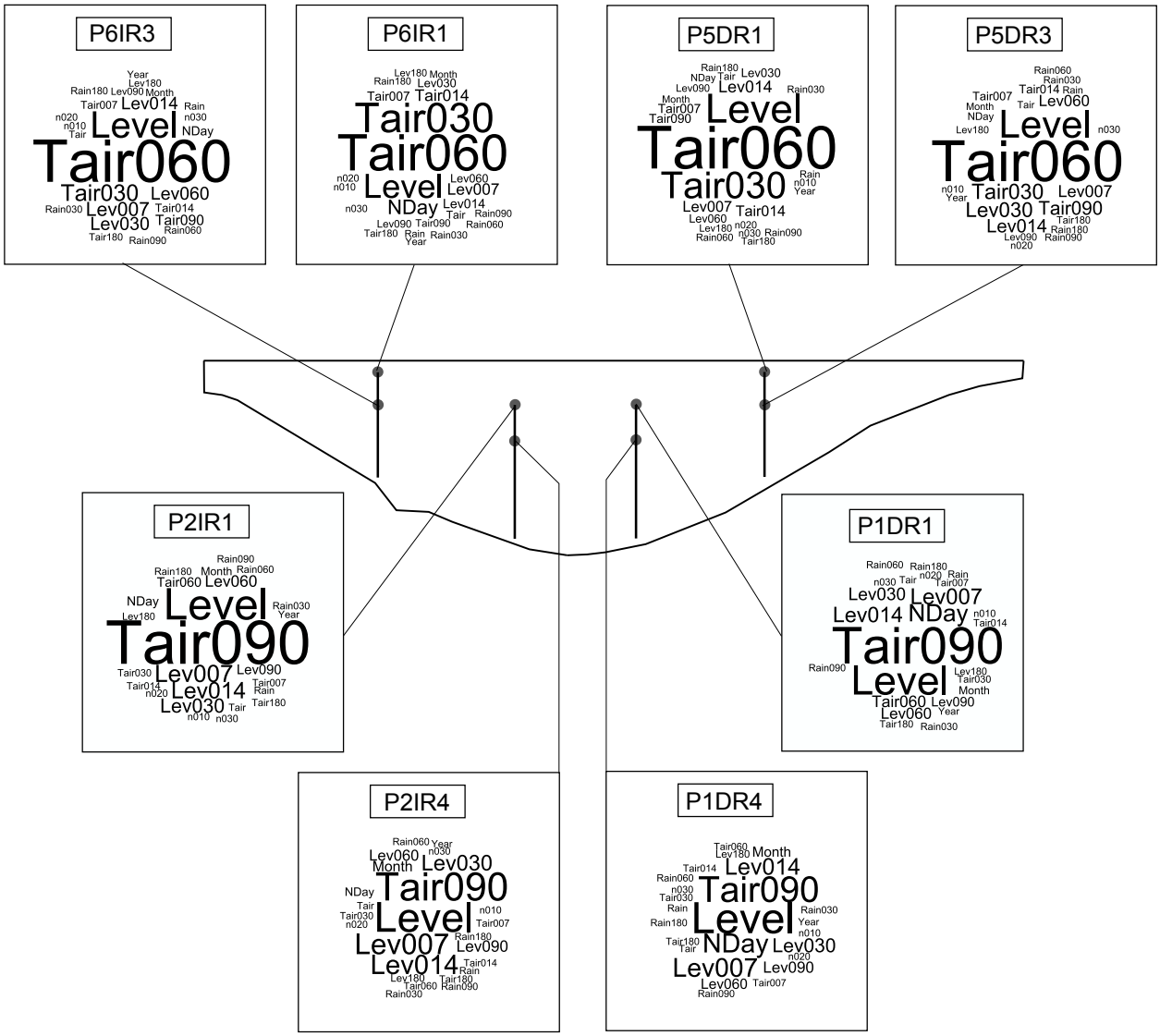


Figure 4.2: Word clouds for the radial displacements analysed.

Other conclusions derived from Figure 4.2 are:

- The thermal inertia was lower near the abutments.
- The RI of the temperature with respect to that of the hydrostatic load increased from the foundation towards the crown, and from the centre to the abutments.
- The dam behaviour is sensibly symmetrical.

The same analysis for the leakage flows revealed a clear different behaviour between the right (AFMD50PR and AFTOTMD) and the left margins (AFMI90PR and AFTOTMI). While the former responded mainly to the hydrostatic load, with little inertia, the latter

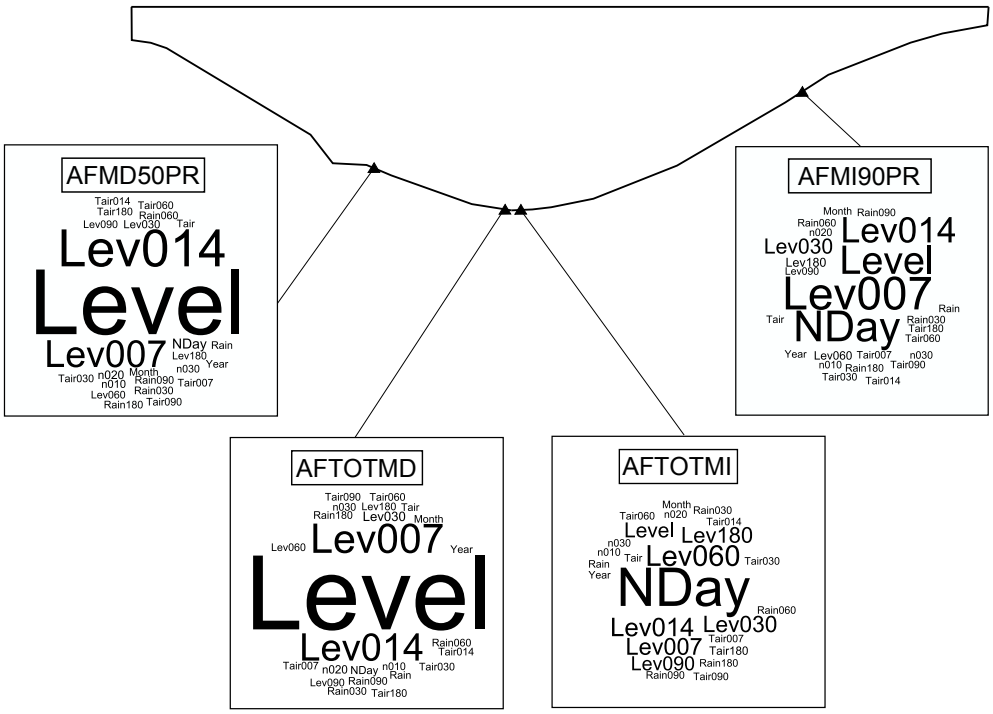


Figure 4.3: Word clouds for the leakage measurement locations analysed

showed a remarkable dependence on time, as well as a greater relevance of several rolling means of reservoir level. Figure 4.3 shows the word clouds for the leakage flows.

The low inertia with respect to the hydrostatic load suggests that most of the leakage flow comes from the reservoir, while the effect of rainfall is negligible.

4.3.3 Partial dependence plots (PDPs)

The resulting PDPs allowed verifying that the dam "behaved as expected", in terms of the first question posed by Lombardi. Figure 4.4 contains the univariate PDP for P1DR1, which shows that higher hydrostatic load and lower air temperature are associated with displacement towards downstream and vice-versa.

Similar plots can be generated in 3D, which allow investigating the pairwise interactions for all the inputs considered (Figure 4.5).

The analysis of the leakage flows (Figure 4.6) confirmed that the time effect was irrelevant in the right abutment, except by certain erratic behaviour in the first two years and in the last three. On the contrary, a sharp decrease in leakage flow was revealed around 1983 for both locations in the left abutment, and a lower decrease in later years.

The shape of the effect of the hydrostatic load is sensibly exponential, with low influence for reservoir level below 610 m.a.s.l.

The PDPs also provide information to answer the second and third questions, by means of analysing the partial dependence on time. In the particular case of P1DR1, these plots

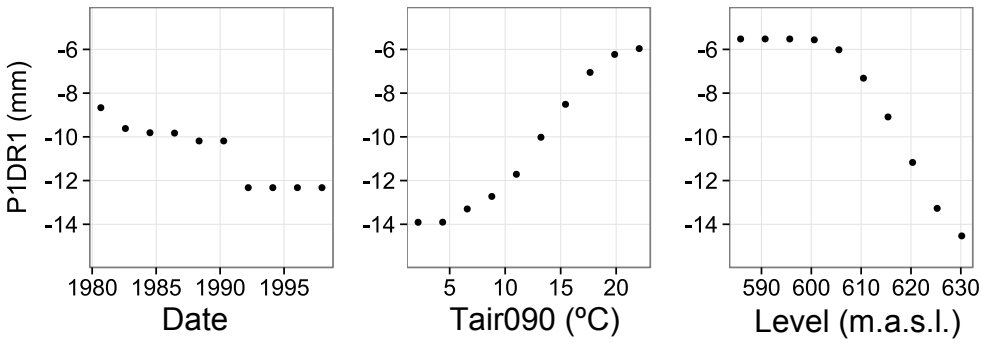


Figure 4.4: Partial dependence plot for P1DR1. Movement towards downstream correspond to lower values in the vertical axis, and vice-versa.

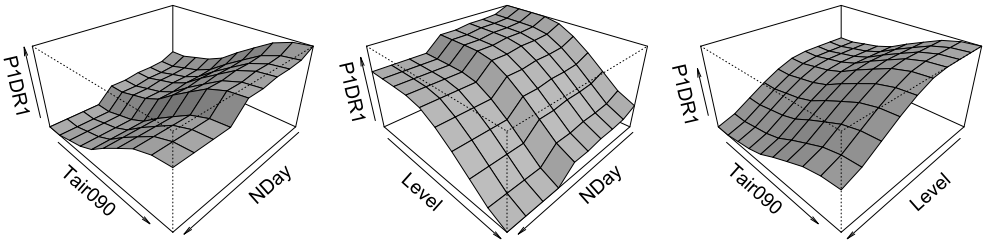


Figure 4.5: 3D PDPs for the main acting loads and P1DR1.

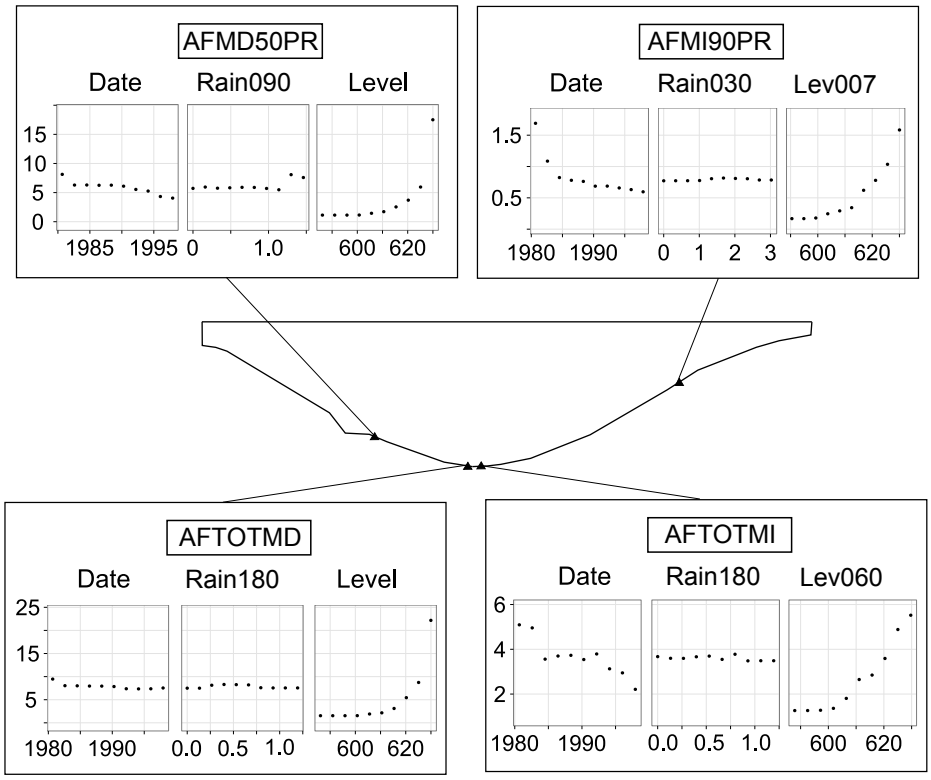


Figure 4.6: Partial dependence plot for leakage flows.

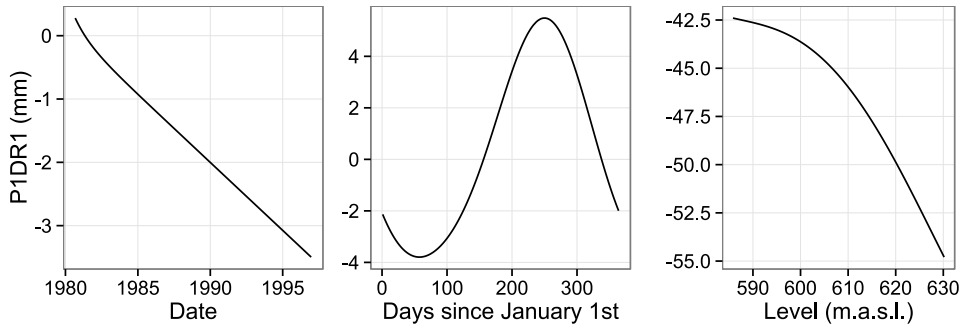


Figure 4.7: Contribution of time, temperature and hydrostatic load on P1DR1, as derived from the interpretation of HST.

show a step around 1991-1992 for the whole ranges of level and temperature, which might represent some change in dam response (Figures 4.4 and 4.5. This issue was object of further verification.

First, an HST model was fitted and similarly interpreted (Figure 4.7). The time effect was a linear trend towards downstream, in contrast with the step suggested by the BRT model.

On another note, the average reservoir level in the period 1991-1997 was significantly higher than before 1991, and might be the cause of the step registered in Figures 4.4 and 4.5: it represents a greater displacement towards downstream in the most recent period, which is consistent with the higher average hydrostatic load.

To clarify the divergence in the results, a new BRT model was fitted to artificial data generated by plugging actual time series of reservoir level into the HST model, while removing the time-dependent terms:

$$\begin{aligned} P1\hat{D}R1_{mod} &= a_1 h + a_2 h^2 + a_3 h^3 + a_4 h^4 + a_5 h^5 \\ &\quad + a_8 cos(s) + a_9 sin(s) \\ &\quad + a_{10} sin^2(s) + a_{11} sin(s) cos(s) \end{aligned} \tag{4.2}$$

The artificial time series data maintains the original reservoir level variation, and thus the higher load in the 1991-1997 period. Figure 4.8 contains the partial dependence plot for this BRT model, which clearly shows that the independence of the artificial data with respect to time was correctly captured. This result confirms that the step in the time dependence captured by BRT is not a consequence of the higher hydrostatic load in 1991-1997.

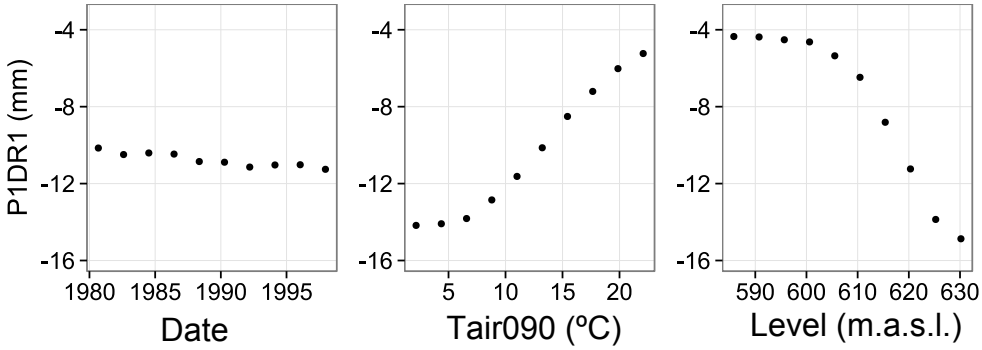


Figure 4.8: Partial dependence plot for the artificial time-independent data. P1DR1. It should be noted that time influence is negligible.

4.4 Conclusions

The interpretation of BRT models resulted in meaningful information on dam behaviour and the effect of each input variable. It allowed verifying that the dam response was in agreement with intuition (e.g. higher hydrostatic load generated displacement towards downstream), and isolating the evolution over time.

The observation of the relative influence of each predictor allowed detecting the thermal inertia of the dam, its symmetrical behaviour, as well as the high variation over time for the leakage flows in the left abutment.

Moreover, the analysis of the time effect suggested that partial dependence plots based on BRT models are more effective to identify performance changes, as they are not coerced by the shape of the regression functions that need to be defined a priori for HST.

5 Anomaly detection

5.1 Introduction

In the precedent sections, the first three questions posed in Chapter 1 were answered: BRT models allowed to study the dam response to the main loads, the relevance of each of the potential inputs, and the evolution over time. The high accuracy of BRTs imply that the conclusions drawn from the model interpretation are reliable.

However, the main objective of dam safety is to prevent failures, for which anomalies need to be detected at early stage. This refers to the fourth question: "was any anomaly in the behaviour of the dam detected?" [40]. The capability of predictive models to identify anomalies has been much less frequently studied than their accuracy. Mata *et al.* [43] developed a model based on linear discriminant analysis for the early detection of developing failure scenarios. This methodology belongs to the Type 2 among those defined by Hodge and Austin [30]: the system is trained with both normal and abnormal behaviour data, and classifies new inputs as belonging to one of those categories. The drawback of this approach is that the failure mode must be defined beforehand and simulated with sufficient accuracy to provide the training data. Hence, the system is specific for the failure mode considered.

Jung *et al.* [34] used a similar approach: abnormal situations were defined based on the discrepancy between model predictions and observed data. They focused on embankment dam piezometer data, and only the reservoir level was considered as external variable (although they acknowledge that the rainfall can also be influential). It is not clear whether this methodology could be applied to other dam typologies or response variables.

Cheng and Zeng [12] presented a methodology based on the definition of some control limits, which depend on the prediction error of a regression model. In addition, they proposed a classification of anomalies based on the trend of the deviation and on how the overall deviance is distributed among the devices considered. It has the advantage of being simultaneously applied to a set of devices, although the case study presented is simple and the test period considered very short (30 days), as compared to the available data (1,555 days).

Other examples of application of advanced tools together with prediction intervals have been published by Gamse and Oberguggenberger [26], who employed the procedure of probabilistic quality control, Yu *et al.* [83], based on principal component analysis (PCA), Kao and Loh [35], who used PCA together with neural networks (NN), Li *et al.* [38], who considered the autocorrelation of the residuals and Loh *et al.* [39], who presented models for short and long term prediction.

Most of these works follow a conceptually similar methodology: a prediction model is built, the density function of the residuals is calculated and used to define the prediction intervals, which are applied to detect anomalies. In all cases, the efficiency is verified by means of its application to a short period of records. As an exception, Jung *et al.* [34] and Mata *et al.* [43] used abnormal data obtained from finite element models (FEM).

In this Chapter, the results of the previous stages are implemented in a methodology for early detection of anomalies, with the following innovative features:

- The prediction model is based on boosted regression trees (BRTs), which showed to be more accurate than other machine learning and statistical tools in previous works [62].
- Causal, non-causal and auto-regressive models are considered and jointly analysed.
- Artificially-generated data are taken as reference. They were obtained from a FEM model considering the coupling between thermal and hydrostatic loads. This allows to identify normal and abnormal behaviour, as observed by some authors ([34], [43]). In this work, the FEM results are compared to actually observed data to verify their reliability.
- A methodology is proposed to neglect false anomalies due to the occurrence of extraordinary loads. It is based on the values of the two main actions (thermal and hydrostatic).
- Three types of anomalies are considered, affecting both to isolated devices and to the whole structure.
- Although radial displacements in an arch dam were selected for the case study, the method can be applied to other dam typologies and response variables. Moreover, it

adapts well to different amount and type of input variables, due to the great flexibility and robustness of BRTs.

The outputs considered correspond to the same radial displacements employed in Chapter 4 (Figure 4.1).

5.2 Methods

5.2.1 Prediction intervals

As mentioned above, most of the published works on the application of data-based models in dam monitoring are limited to the assessment of the model accuracy. However, the main practical utility of these models is the early detection of anomalies, for which it is necessary to compare the predictions with monitoring readings, and verify whether they fall within a predefined range. If the residual density function follows a normal distribution, that range can be defined in terms of the standard deviation of the residuals. For example, Kao and Loh [35] presented the 99% prediction intervals for models based on neural networks, while Jung *et al.* [34] tested 1, 2 and 3 standard deviations of the residuals as the width of the prediction interval.

Based on the results of a preliminary study [60], the prediction interval was set to $[\mu - 2\ sd_{res}, \mu + 2\ sd_{res}]$, being μ and sd_{res} the mean and the standard deviation of the residuals, respectively. Special attention was paid to the determination of a realistic residual distribution. It is well known that the accuracy of a machine learning prediction model must be calculated from a data set not used for model fitting [31] (validation set). In the case of time series, this validation set should be more recent in time than the training data, since in practice the model is used for predicting a time period subsequent to the training data [2].

The hold-out cross-validation method meets this requirement, with the most recent data in the hold-out set (Figure 5.1).

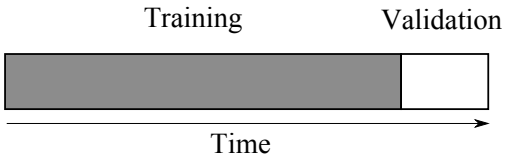


Figure 5.1: Hold-out cross-validation scheme.

However, this implies discarding the most recent data for the model fit, which are generally the most useful, since they represent the most similar behaviour to that to be predicted (assuming there may be a gradual change in behaviour over time). Moreover, the validation data may be biased, if they correspond, for instance, to a especially warm (or cold) period.

To overcome these drawbacks while maintaining good estimate of the prediction error, an approach based on the hold-out cross validation method suggested by Arlot and Celisse [2] for non-stationary time series data was employed.

The proposed method takes into account the following specific aspects of dam behaviour: a) changes in the dam-foundation system are generally gradual, and b) dam behaviour models are typically revised annually, coinciding with the update of safety reports.

Let us consider that a behaviour model is to be fitted at the beginning of year Z_i, to be applied for anomaly detection during that year. The available data corresponds to the years $Z_1 \dots Z_{i-1}$, with Z_1 being the initial year of dam operation. With the simple hold-out method, a model is fitted with data in years $Z_1 \dots Z_{i-2}$, whose accuracy is evaluated on data in Z_{i-1}.

In this work, a minimum training period of 5 years was considered. This value was chosen in view of the results of previous studies [62], and the evolution of model accuracy on the reference data, as described in section 5.3.2. Then, an iterative process was followed to reduce potential bias in the loads during Z_{i-1}. A set of predictions is generated as follows:

- For $k = 5 \dots i-2$
- Fit a model M_k trained with the period $Z_1 \dots Z_k$.
- Compute R_k as the residuals of M_k when predicting year Z_{k+1}.
- Compute the mean (μ_k) and standard deviation ($sd_{res,k}$) of R_k

At the end of the process, residuals for a set of models $M_k, k = 5 \cdots i-2$ are obtained, with the particularity that they are computed over different time periods, always subsequent to the training set ($Z_6 \cdots Z_{i-1}$). That is, the amount of observations in the training sample increases, and are used to predict the following year. The potential bias of some abnormal loads for one year is compensated by averaging, while a realistic prediction error is achieved, since it is always based on precedent data. A similar approach was employed by Herrera *et al.* to estimate demand in water supply networks, who employed the term *growing window strategy* [29].

Additionally, since the model accuracy typically increases as the training data grows, the actual model accuracy for the application period (year Z_i) will be more similar to that obtained for Z_{i-1}. Hence, R_{i-2} is more representative of the expected model performance for Z_i. To account for this issue, the prediction intervals are based on a weighted average of μ_k and $sd_{res,k}$. In particular, the weights for each year decrease geometrically from the most recent to the first available. A schematic representation of the procedure is included in Figure 5.2.

Finally, to take advantage of all the available data, a model is fitted with the entire period $Z_1 \dots Z_{i-1}$, with which the predictions for the following year (Z_i) are computed.

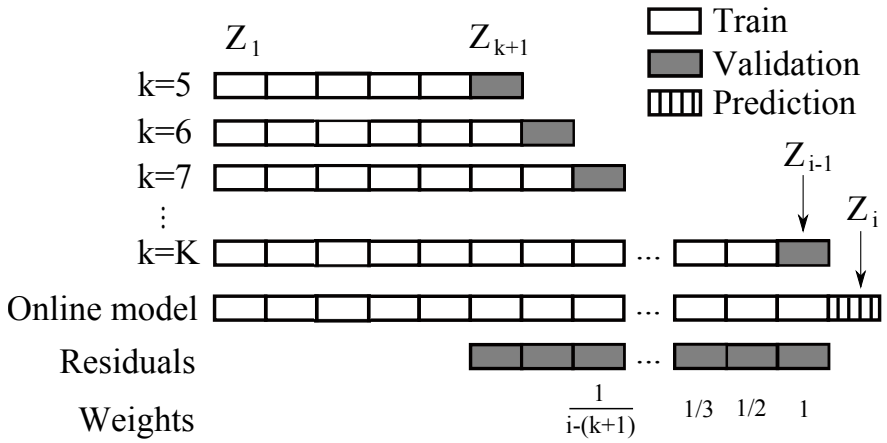


Figure 5.2: Graphical representation of the weighted growing-window cross-validation procedure. The prediction interval is estimated as a function of the weighted average of the standard deviation of the residuals for previous years, each one is computed from a model trained with a different training set.

Since the test set becomes part of the validation period in the subsequent years, the residuals generated during the application of the model in the test period can be added to those computed for previous years, so that there is no need to repeat the whole process: the previous residuals can be employed to obtain the new prediction interval, after updating the correspondent weights.

5.2.2 Causal and non-causal models

BRT models are robust against the presence of uninformative or highly correlated predictors [23], [63]. Hence, variable selection is much less influential for tree-based methods than for other machine learning tools [24]. This property was employed to build BRT models of three types.

The first is a causal model, as that described in section 3.2, which considers as predictors those inputs related to air temperature, hydrostatic load and time (Table 3.2). A priori, a model of this type is expected to detect reading errors and changes in dam behaviour. However, its accuracy might be improved, since the response of the dam may depend on variables not considered, such as the maximum and minimum daily temperatures, or the solar radiation.

The second version is the Non-Causal model. In addition to the predictors described above, dam response variables were also considered as inputs. This means that each radial displacement is included in the input set to predict other radial displacements. This version will in principle give greater precision, since the record from a neighbouring device (e.g. another station of the same pendulum) implicitly contains the effect of external variables not considered in the causal version. By contrast, this model might not be able to detect

anomalies affecting several devices. For example, a slide in a block of a concrete gravity dam will be reflected in all stations of the correspondent plumb line; therefore, the relation between the hydrostatic load and the displacement would be abnormal, while the relationship between several readings of the same pendulum could be normal.

Finally, an auto-regressive with exogenous inputs (ARX) [47] model was also fitted for each output, where the lagged values of all radial displacements were added to the Non-Causal model input set[1] . Specifically, the response at time t_i is estimated based on the readings at t_{i-1} and t_{i-2}, both for the variable to predict and other response variables.

One of the objectives of this work is to test the ability of all three models to detect various types of abnormalities, and draw conclusions for practical purposes.

5.2.3 Case study

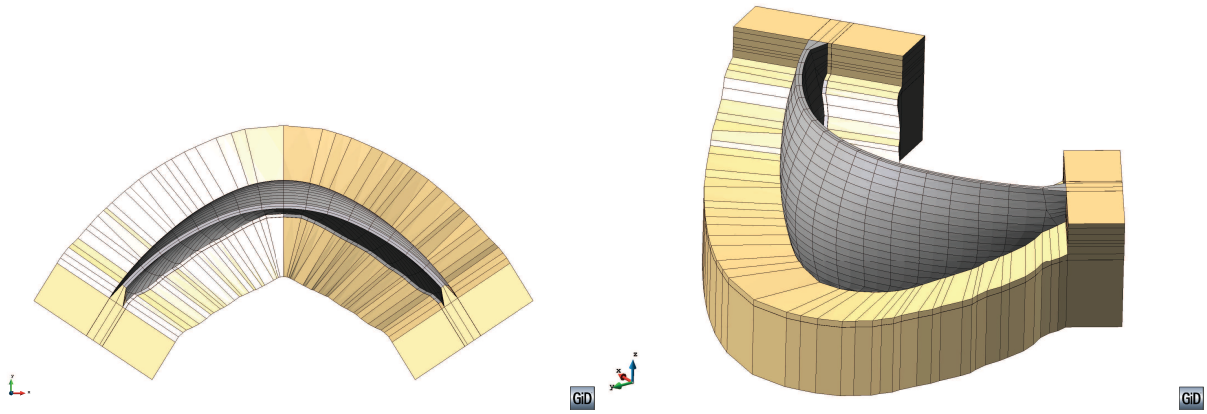
As in previous analyses, La Baells arch dam was also selected as the case study (Section 3.2). In this case, the air temperature and the reservoir level time series were considered as inputs to a FEM model. The results of this model in terms of radial displacements at the location of the pendulums were extracted and compared to the actual measurements. The objective was to check that the FEM model could provide realistic data to generate reference time series of dam behaviour. These artificial data are free from any temporal variation (the reference numerical model does not vary with time; only environmental loads do).

The dam was considered as a three-dimensional solid discretised in hexahedral serendipity 27-node elements. A portion of the foundation was also included, resulting in a total of 13,029 nodes and 2,530 elements. The thermal and mechanical problems were solved separately on the resulting finite element mesh (Figure 5.3). The material properties are shown in table 5.2.3.

Property	Dam	Foundation
Young modulus ($N \cdot m^{-2}$)	$4.76 \cdot 10^{10}$	$3.10 \cdot 10^{10}$
Poisson ratio	0.25	0.25
Density ($kg \cdot m^{-3}$)	2,400	3,000
Thermal conductivity ($W \cdot^{\circ} K^{-1} \cdot m^{-1}$)	2.4	2.2
Thermal expansion coefficient	10^{-5}	10^{-5}
Specific heat ($J \cdot kg^{-1} \cdot^{\circ} K^{-1}$)	982	950

Table 5.1: Material properties considered in the FEM model

[1]The ARX model is also non-causal, in the sense that variables with non-causal relation with the outputs are included as predictors. The acronym ARX was employed to distinguish both models when necessary, although they are occasionally jointly referred to as "non-causal models". For the sake of clarity, the capitalised version ("Non-Causal") is used to specifically refer to the second model, excluding the ARX.



(a) Plan view

(b) Perspective from upstream

Figure 5.3: FEM model.

For the thermal problem, a transient computation was run over the 1981-2008 period with time step of 30 days. The temperature was imposed in both dam faces, with different values for the wet and dry areas. For the boundaries below the reservoir level, the temperature was considered as equal to that of the water, which in turn was estimated by means of the Bofang formula [4]. Although it allows accounting for the temperature variation with depth, a unique value was considered in this work for all the wetted boundaries, equal to that obtained for 50% depth. For the dry faces, the 30-days moving average of air temperature was imposed, to take into account the thermal inertia. The result was increased by 2 degrees to account for the solar radiation, following the approach proposed by Pérez and Martínez for Spanish dams in the North-East region [51]. The temperature evolution for the first year was repeated 4 times to ensure that the result was not influenced by the initial conditions.

The mechanical response was assumed to be elastic and instantaneous (without inertia), hence for each time step, the hydrostatic load correspondent to the actual reservoir level was applied.

The results of both models (thermal and mechanical) were added, and the displacement evolution at the location of the monitoring devices were extracted. The model results, which are generated in global axes, were later transformed to the local axes correspondent to the radial displacements, as measured by the monitoring devices.

Finally, weekly values were obtained via interpolation, according to the average reading frequency for the available data.

In addition to radial displacements, also the temperature evolution in the dam body was compared to observed data from several thermometers embedded in the dam body.

The goodness of fit of the FE model was computed in terms of the mean absolute error (MAE) (equation 3.1).

5.2.4 Anomalies

As described in the previous section, the reference time series were those obtained with the FEM model for the 1980-2008 period, where the boundary conditions and loads correspond to the reservoir level and air temperature actually measured in the dam site. Three different types of anomalies were later introduced to modify those data:

- Scenario 1: Progressive breakdown of an isolated device. An increasing value was added to the reference series, with constant rate (a mm $\cdot$ year^{-1}).
- Scenario 2: The same as scenario 1, though the magnitude of the deviation is constant (a mm)
- Scenario 3: Imposed displacement of the left abutment. The data for this scenario were obtained from a modified FEM model representing a hypothetical sliding of the left abutment. For that purpose, the boundary condition at that region was set to a mm both in x and y axes (instead of null displacement, as for the reference case).

It is important to note that the anomaly of scenario 3 affects differently to each of the devices analysed. Since a displacement in the left abutment was imposed, the results in the left half of the dam body are anomalous. However, those in the right half are not affected. This can be observed in Figure 5.4, which depicts the displacement field in the dam body generated by the imposed anomaly with $a = 2mm$.

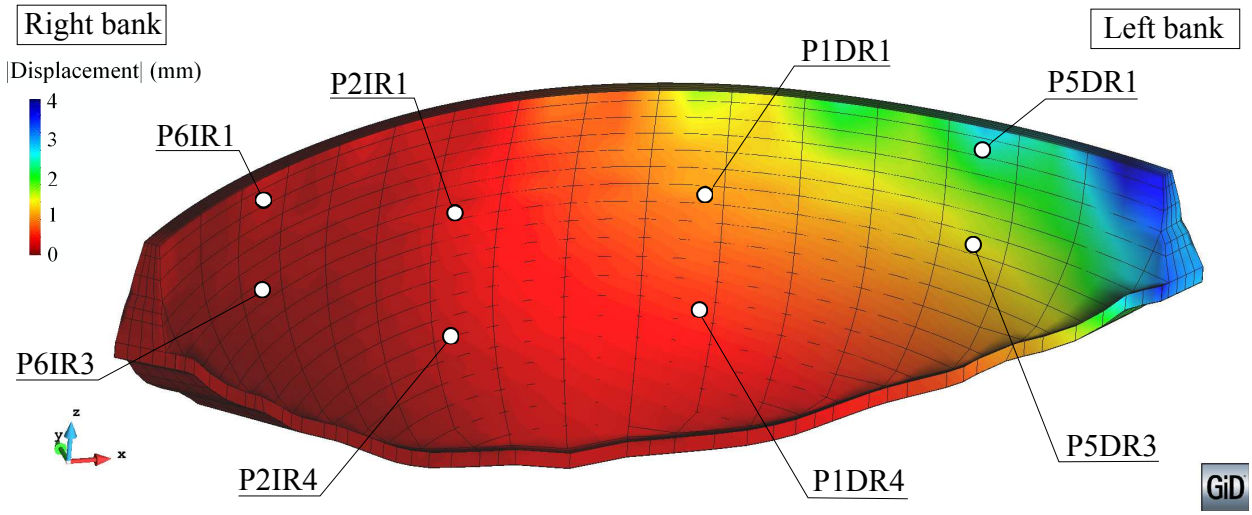


Figure 5.4: Displacement field resulting from the anomaly in scenario 3. View from downstream.

Table 5.2.4 contains the mean absolute deviation between the reference and the anomalous time series for each device for $a = 2mm$. Since the anomaly in scenario 3 does not affect to some devices, those values considered as abnormal by the system will be false positives.

Device	MAE (mm)	Device	MAE (mm)
P1DR1	0.61	P5DR1	1.42
P1DR4	0.52	P5DR3	1.05
P2IR1	0.10	P6IR1	0.02
P2IR4	0.13	P6IR3	0.01

Table 5.2: Discrepancy between the normal displacements, as computed with the FEM model, and those imposed in scenario 3 for $a = 2mm$. Mean absolute error (mm)

For each scenario, the performance of the three models considered (causal, Non-Causal and auto-regressive) was analysed. 4,000 anomalous cases were generated, where the following parameters were randomly selected:

- Initial date of abnormal period
- Anomaly scenario
- Output variable (Scenarios 1 and 2)
- Magnitude: 0.5, 1.0 or 2.0 mm $\cdot$ year $^{-1}$ for scenario 1; 0.5, 1.0 or 2.0 mm for scenario 2; 1.0 or 2.0 mm for scenario 3.

Each anomalous case was presented to all three models to compare their ability for anomaly detection. This was computed in terms of the *detection time* (t_{det}), defined as the elapsed time from the start of the anomaly until the first observation considered anomalous by each model, measured in days (Figure 5.5). Since the abnormal period was limited to 1 year, the models which did not detect any anomaly were assigned a t_{det} value of 365 days.

Moreover, the effectiveness of an anomaly detection system also depends on the number of false positives (observations considered abnormal by the model, which are actually normal) and false negatives (abnormal values not detected as such by the model). The two most commonly used metrics to account for these are precision (equation 5.1) and recall (equation 5.2). The comparison was mainly based on the F_2 index 5.3 [34], which jointly considers precision and recall, giving more importance to the latter.

$$precision = \frac{true\ positives}{true\ positives + false\ positives} \tag{5.1}$$

$$recall = \frac{true\ positives}{true\ positives + false\ negatives} \tag{5.2}$$

$$F_2 = (1 + 2^2)\frac{precision \cdot recall}{4 \cdot precision + recall} \tag{5.3}$$

However, these indexes are not useful for model performance assessment when analysing the unaffected devices in scenario 3. In these cases, there are not true positives (all records are normal, since these devices are not affected by the anomaly). Hence, both precision and recall equal zero. Nonetheless, it is highly relevant to know whether the proposed models correctly identify these records within the prediction interval. For that purpose, scenario 3 was analysed by means of the amount of false positives, whose computation depends on the device. For those in the left half of the dam body (as viewed from upstream), which are actually anomalous, the observations above the upper limit of the prediction interval are considered as false positives, since they would imply a deviation towards upstream (while the actual anomaly corresponds to a displacement in the downstream direction). By contrast, for the unaffected devices, every record outside the prediction interval is a false positive, both above the upper limit and below the lower limit of the interval.

5.2.5 Load combination verification

In general, model accuracy is dependent on the values of the input variables. The more input data available for similar situations to that to be predicted, the more accuracy is to be expected. In dam behaviour, it will depend on the thermal and hydrostatic loads.

This effect is more important when input values are out of the training data range [20]. In particular, the accuracy of data-based models as BRTs may decrease dramatically when extrapolating.

Cheng et *al.* [12] defined a possible abnormal state of the dam (State 3), that "may be caused by extreme environmental values variables". In this work, this issue was explicitly verified, and out-of-range (OOR) instances were considered as potential false positives.

This verification was carried out following an original procedure, specifically designed for the dam behaviour problem, where there are three main loads: thermal, mechanical (hydrostatic head) and temporal.

If the behaviour of the dam does not change over time, the importance of time variable is negligible. This was checked when fitting BRT models to the reference data, which correspond to time-independent dam behaviour. The inclusion of these variables is useful for retrospective analysis, as confirmed in Chapter 4. In practice, a previously trained model is employed to predict future values. Hence, it is obvious that the model prediction is an extrapolation in time axis and thus does not need to be verified.

As for the other two loads (thermal and hydrostatic), the simplest approach would be to check whether their values for the test period are greater (lower) than the maximum (minimum) within the training data set. However, that would not consider that both effects are coupled: the water temperature is different to that of the air, hence the water surface elevation affects the boundary condition in the upstream dam face and, as a result, conditions the thermal response of the dam [74].

Moreover, there is not a widely accepted agreement on what extrapolation is and how to handle it [20]. In dam behaviour modelling, it seems obvious that a hydrostatic load above the maximum in the training set is out-of-range. However, a more detailed definition seems appropriate to account for the "empty space phenomenon" [78], i.e., the existence of areas without training samples within the range of the inputs.

To account for this issue, a procedure that takes into consideration the combination of both loads is proposed:

1. The training data are plotted in the (Reservoir level, Air temperature) plane.

2. A two-dimensional density function is computed by means of the kernel density estimation (KDE) method.

3. The training instance with lower density value is localised, and the corresponding isoline is plotted.

4. The input values for the new data are plotted on the same plane. Those falling outside the isoline are considered as OOR.

With this procedure, it is taken into account that the predictive accuracy can be poor for a load combination not previously presented, even though their values, if considered separately, are within the training range. An example of this issue is presented in Figure 5.5.

5.3 Results and discussion

5.3.1 FEM model accuracy

Figure 5.6 shows the comparison between the observed radial displacements for P1DR1 and those obtained with the FE model for the period 1994-2008. Results for other outputs are similar (Table 5.3.1). The FEM model accuracy is comparable to that obtained in previous Chapters with data-based models 3.2.

As regards the temperature, Figure 5.7 shows the numerical results and the observed data for 4 thermometers and the January 2007 - June 2008 period. Both the devices and the time period correspond to the results published by Santillán *et al.* [67], who employed a highly detailed thermal model for the same case study.

Since predicting the thermal response is not the main objective of this analysis, relevant simplifications were employed to generate the reference data (neglecting the variation in water temperature with depth, using a relatively large time step). Nonetheless, the temperature within the dam body was well captured.

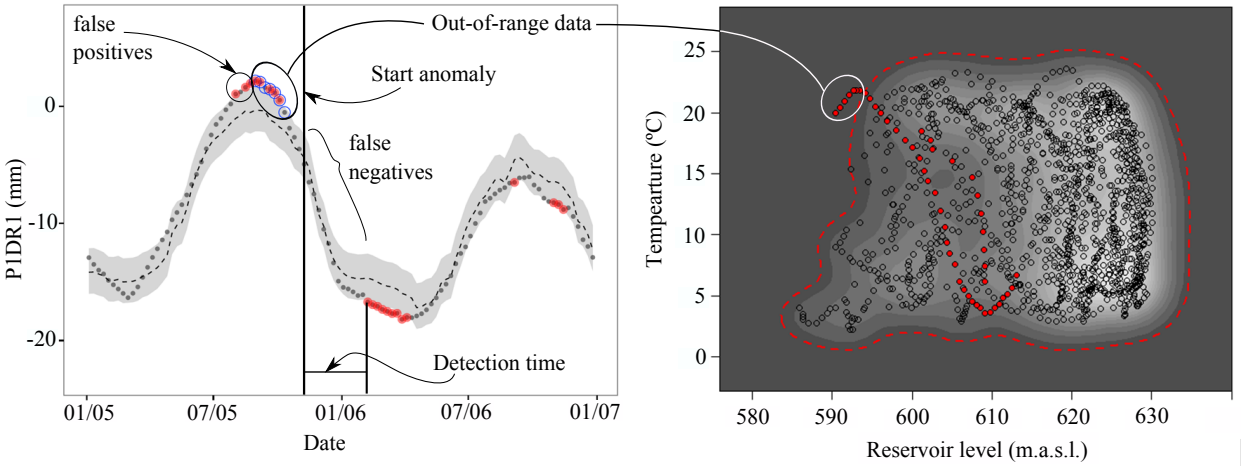


Figure 5.5: Model performance indicators. Left: typical output plot, with the observations (circles), the predictions (dotted line), and the prediction interval (shaded area). Before the start of anomaly, some data fall outside the prediction interval (in red). Of those, some are false positives, whereas others correspond to out-of-range inputs (blue circles), since they fall in a low-density region in the 2D density plot (right). In this case, a combination of high temperature and low reservoir level was presented for the first time in dam history.

This, together with the results for displacements, confirm that the resulting data series mostly reproduce the dam response to the main loads. Therefore, they are representative of the normal behaviour of the dam and useful to evaluate the ability of the methodology to detect anomalies.

5.3.2 Prediction accuracy

The performance of all models on the reference data (without anomalies) was first assessed. The objectives are:

1. Verify the evolution of the prediction accuracy over time

Output	MAE (mm)	Output	MAE (mm)
P1DR1	0.70	P5DR1	0.81
P1DR4	0.65	P5DR3	1.01
P2IR1	1.08	P6IR1	0.96
P2IR4	0.98	P6IR3	0.58

Table 5.3: Deviation between the radial displacements as computed with the FEM and the actual records for the 1994-2008 period. Mean absolute error

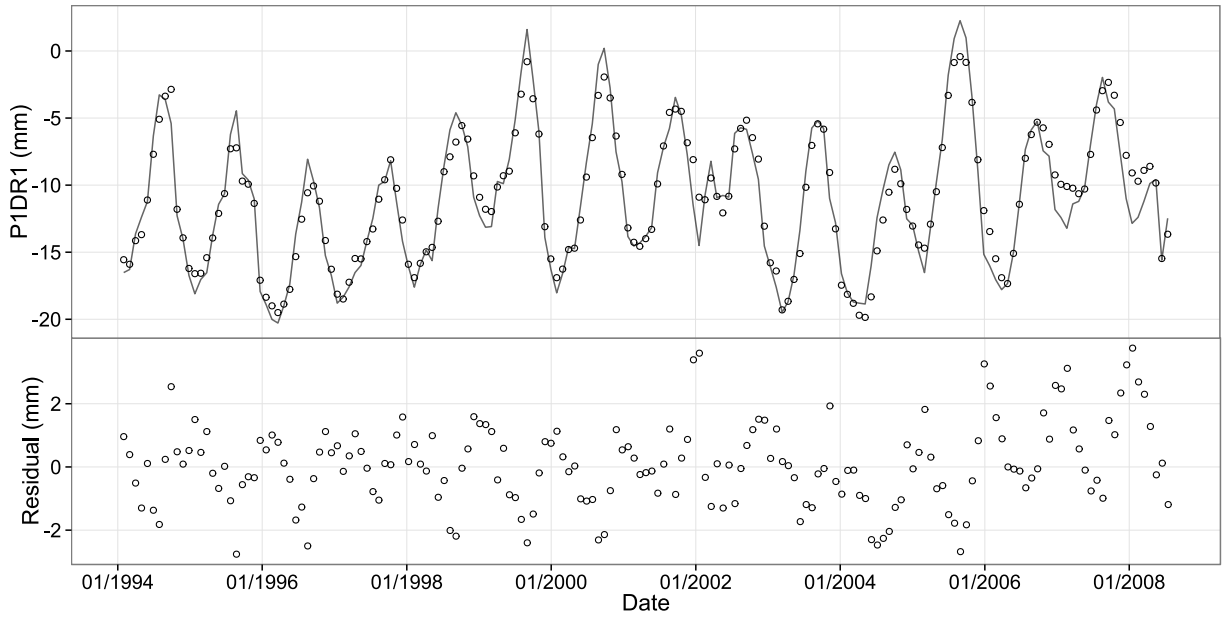


Figure 5.6: FEM results versus observations for P1DR1

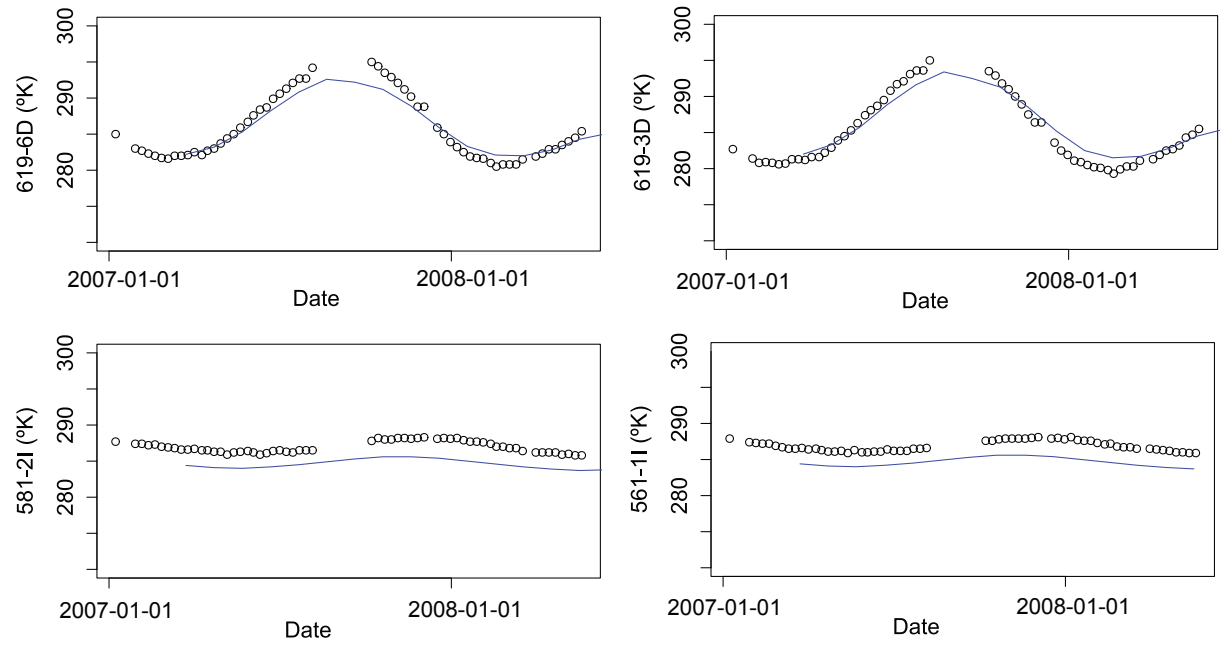


Figure 5.7: Comparison between numerical and measured temperature in 4 locations within the dam body

2. Check the effect of averaging the standard deviation

3. Compare all models in terms of false positives

4. Evaluate the efficiency of the criterion to detect out-of-range data

For that purpose, the iterative process described in section 5.2.1 was followed, i.e., each model is re-fitted yearly over an increasing training set, and the prediction interval is updated as a function of the actualised value of the weighted average of the residual standard deviation. Since the dam-foundation behaviour is time-independent for the reference case, the variation in model accuracy is due to the increase of training data.

Figure 5.8 shows the evolution of both the raw and the weighted average of the residual standard deviation for all devices and models. Some conclusions can be drawn:

- As expected, the accuracy of the Non-Causal and ARX models model is higher, since the non-causal inputs implicitly contain information regarding external variables not considered in the causal version.

- The inclusion of lagged variables in the ARX model is not relevant, as compared to the Non-Causal one.

- The raw values show high variance, especially for the causal model, which is eliminated by averaging

- The time evolution of the weighted standard deviation of the residuals is similar for all models: a sharp decrease in the first years, followed by quasi-constant behaviour. Nonetheless, the causal model requires more data to reach the low-slope part of the curve.

Table 5.3.2 contains the amount of false positives for all targets and models, as well as those correspondent to out-of-range inputs. Although the prediction interval for the causal model is wider (due to the higher residual standard deviation), it also generates a greater quantity of false positives. However, the average amount is low in all cases, as compared to the total amount of records (1,464). Moreover, the procedure to identify out-of-range inputs reduces the false positives by 27 % for the causal model and by 45% for both the non-causal and the ARX. As a result, the mean percentage of false positives is 8.0, 2.8 and 2.6 % respectively. It should be noticed that the results for the non-causal and ARX models are lower than the theoretical percentage of values outside the interval within 2 times the standard deviation in a normal distribution (5%).

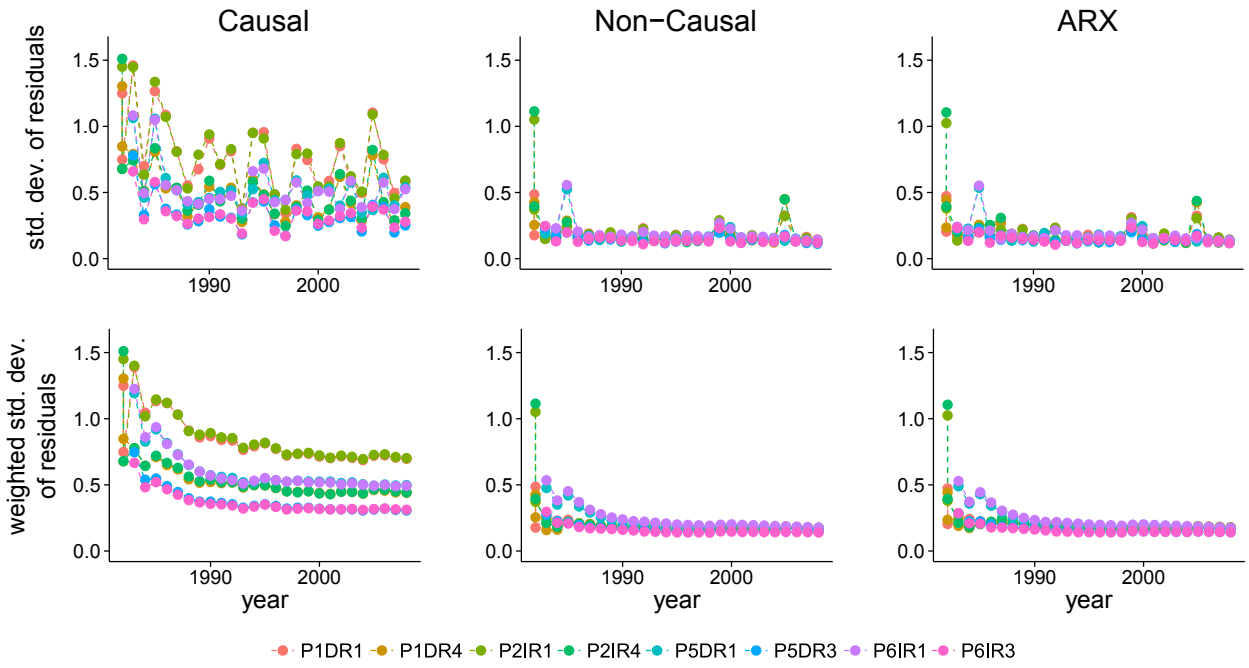


Figure 5.8: Time evolution of the prediction accuracy for all models and outputs. Top: standard deviation of residuals per year. Bottom: weighted average.

Model	Causal		Non-Causal		ARX	
Target	# False pos.	# OOR	# False pos.	# OOR	# False pos.	# OOR
P1DR1	179	53	91	40	82	35
P1DR4	178	54	89	42	75	38
P2IR1	184	54	89	41	85	35
P2IR4	198	54	95	50	75	38
P5DR1	125	31	50	21	51	21
P5DR3	164	49	72	31	68	30
P6IR1	129	31	51	21	50	21
P6IR3	171	42	63	27	65	28
Mean	166	46	75	34	69	31

Table 5.4: Amount of false positives

5.3.3 Anomaly detection

Figure 5.9 (a) shows the F_2 results as a function of the model and the anomaly magnitude a for scenarios 1 and 2. As expected, the larger anomalies were more easily detected in all cases. As for the input variables, Non-Causal model performed better on average, especially for small anomalies and as compared to the causal model. Again, the inclusion of lagged variables generated a minor effect, in this case towards slightly poorer performance.

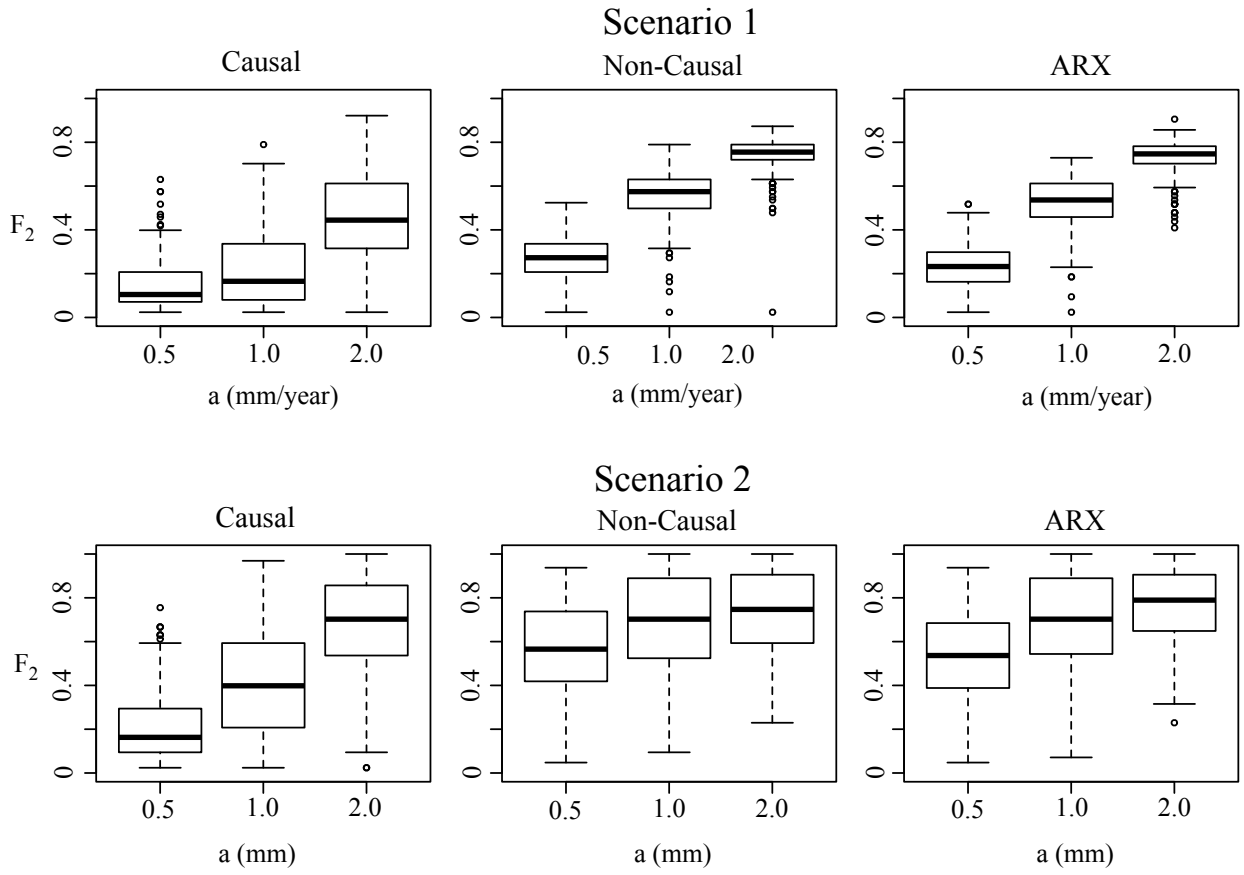


Figure 5.9: F_2 index for scenarios 1 and 2.

The results for Scenario 3 are more interesting to analyse, since they correspond to a realistic anomaly affecting the overall dam behaviour. Since the effect of this anomaly is different on each output, the results are presented in terms of the true detection time t_d per device, i. e., the elapsed time until the first record identified as a deviation towards downstream. Figure 5.10 shows the results.

A perfect model would feature null detection time for the affected devices (P1DR1, P1DR4, P5DR1 and P5DR3), and 365 days for the remaining (P2IR1, P2IR4, P6IR1 and P6IR3). Both the Non-Causal and the ARX models showed almost perfect performance. As regards the causal model, the anomaly in the most affected devices (P5DR1 and P5DR3) is

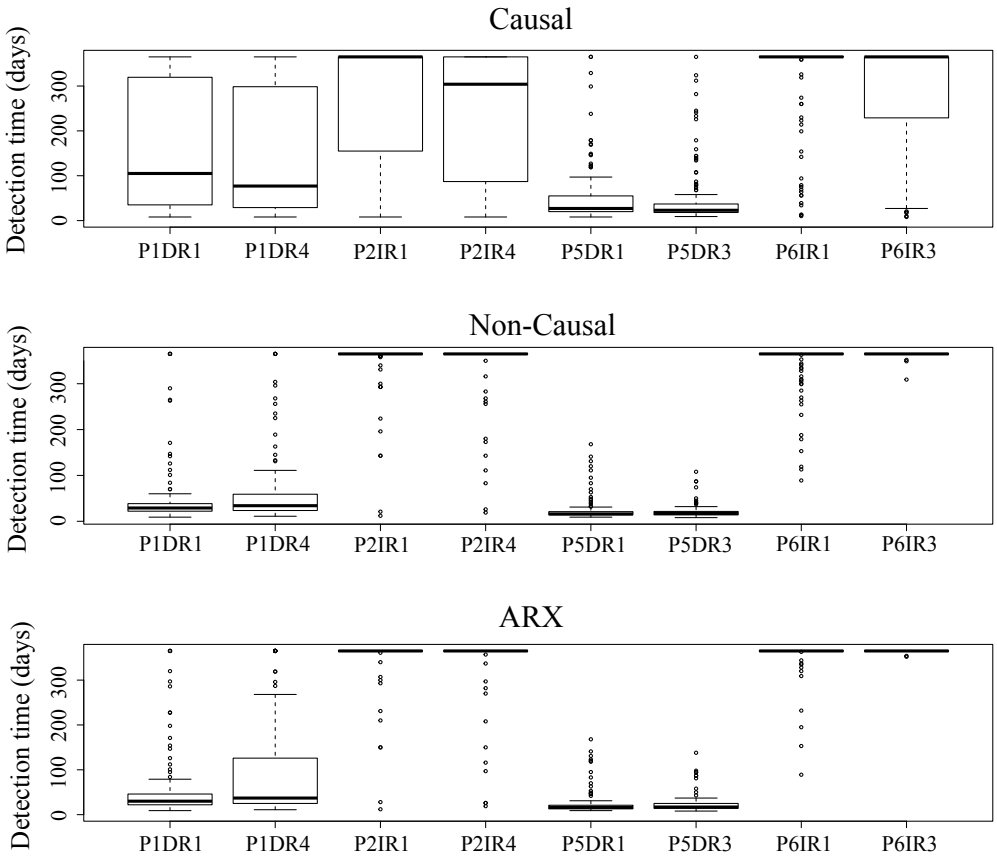


Figure 5.10: Detection time (days) per target and model for scenario 3.

detected almost instantly, but is less effective for P1DR1 and P1DR4, whose deviation from the reference behaviour is low (see Table 5.3.1). The detection time for P1DR1 and P1DR4 is around two months, with high variation up to 300 days.

A complete assessment of the model performance requires analysing the amount of false positives. They correspond to any value outside the prediction interval for the targets in the right half of the dam body, and to anomalies correspondent to deviations towards upstream for those in the left region. Figure 5.11 shows these results.

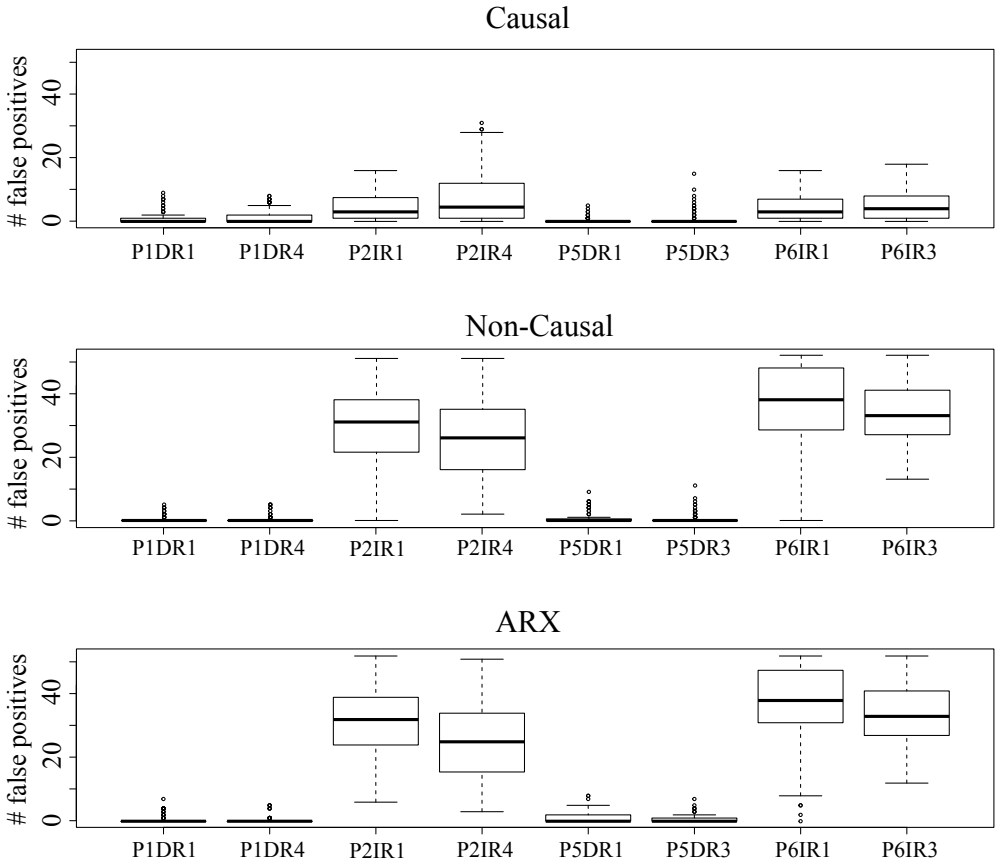


Figure 5.11: False positives per target and model for scenario 3.

It can be observed that the causal model is clearly more effective in this regard: both the Non-Causal and the ARX models classify around half of the observations for the unaffected devices as abnormal (there are 52 observations in the period of analysis). This result is due to the nature of the inputs for each model. For example, the Non-Causal model generates a prediction for P6IR1 based on the value of P5DR1 (among other inputs, but this is particularly important for being symmetrical within the dam body). In scenario 3, P5DR1 deviates towards downstream with respect to the reference (training) period. Since that input is anomalous, the resulting prediction is also wrong. In this case, the model interprets that the value of P6IR1 falls in the upstream side of the prediction interval.

This issue is highly relevant, since the final aim of the system is not only to detect a

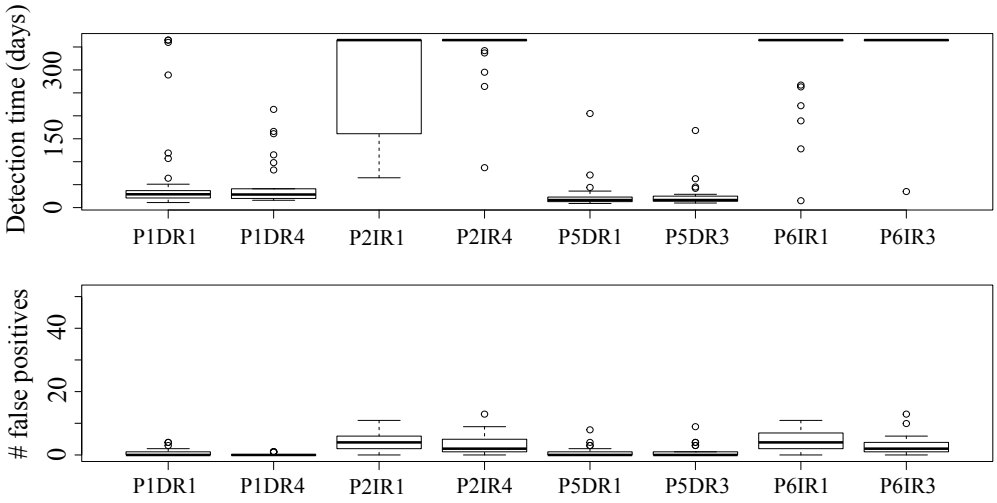


Figure 5.12: Detection time and false positives per target for scenario 3 and the Non-Causal model, once the anomalous variables are removed from the input set.

potentially anomalous behaviour, but also to support the correct identification of the cause, and then the decision making. In fact, similar results would have been obtained had the devices been analysed jointly in scenarios 1 and 2: a real deviation towards downstream in some device is (in general) correctly identified by the non-causal models, but that same value would generate an incorrect prediction for other devices, of opposite sign.

Causal models do not give these spurious results, since they predict the dam response only based on the external variables, at the cost of a generally higher detection time.

A straightforward option to avoid this behaviour is to discard non-causal models. However, their good performance for detecting true anomalies suggests that they can be useful overall.

As an alternative, the outputs whose value is identified as anomalous by the Non-Causal model can be removed from the input set. This requires re-training, but it can still offer accurate results, thanks to the flexibility of BRTs.

A new set of 240 cases was run for scenario 3 and the Non-Causal model. The results shown in Figure 5.12 confirm that the removal of abnormal variables is effective against false positives, while maintaining the ability for anomaly detection. The model performance is only poorer for P2IR1 (unaffected by the anomaly in scenario 3): the detection time is lower than 365 days, which indicates the existence of false positives. Nonetheless, the average detection time is still 270 days, and the total amount of false positives is lower than 10 %.

This approach was implemented in a new interactive tool, which was developed to present the results for all devices involved. It is based on the shiny library [11], and includes two plots for each model (Figure 5.13).

First, each device is plotted on its actual location within the dam body, with a symbol that is a function of the deviation between prediction and observation for the date under

Figure 5.13: Interface of the dam monitoring data analysis tool for a case from scenario 3. The imposed displacement in the left abutment is correctly identified

consideration. Then, the evolution of observations and predictions for the most recent period is plotted for two devices selected by the user. Figure 5.13 shows the application interface for one of the anomalies from scenario 3. It can be observed that the anomaly is correctly localised.

With this tool, the user jointly receives the overall information on all devices under consideration, and a more detailed plot of the selected output, where the value of the deviation, as well as the trend, can be observed. In this version, devices whose residuals are lower than two times the standard deviation are plotted in green; those between two and three times are depicted in yellow, and those above three times are shown in red. The shapes correspond to the direction of the deviation (upstream or downstream), as interpreted by each model. This criterion can be tailored to the user preferences.

5.4 Summary and conclusions

A methodology for early detection of anomalies in dam behaviour was presented, which includes a prediction model based on BRT, a criterion for detecting anomalies based on the residual density function, and a procedure for realistic estimation of the prediction interval. Also, extraordinary loads are identified by jointly considering the two most important external loads (hydrostatic load and temperature).

Causal models (which only consider external variables) and non-causal (including both

internal and lagged variables as predictors) were compared in terms of detection time for three different anomaly scenarios. The results showed that non-causal models are more effective for the detection of anomalies, both affecting to isolated devices (Scenarios 1 and 2), and those resulting from an overall malfunction of the dam (Scenario 3).

In the case study considered, the inclusion of lagged variables had minor effect both in the model accuracy and the detection time. This suggests that the Non-Causal model (without lagged variables) might be a better choice due to its higher simplicity.

Causal models were more robust as regards the precision (when accounting for false positives). In abnormal periods, the prediction of non-causal models for unaffected devices is often wrong because it is partially based on anomalous data (that from the devices actually affected by the anomaly). This type of behaviour is a consequence of the nature of the model itself, and is the price to pay in exchange for a greater ability for early detection of anomalies.

However, an updated version of the Non-Causal model, where the anomalous variables are removed from the input set, avoided the above-mentioned issue, and showed to be as effective for anomaly detection as the original Non-Causal, and even more robust against false positives than the causal model. Hence, this approach is the best option to provide useful information to the dam safety managers. To that end, it was implemented in an interactive on-line tool, which shows the devices whose behaviour is interpreted as potentially abnormal by the predictive model, together with the plot of the evolution of predictions and observations for all relevant outputs.

This tool can be used as a support for decision making, since it facilitates the identification of a potential deviation from normal behaviour. Thus, it can be used as an indicator to generate a warning which might lead to intensify the dam safety monitoring activity.

6

Achievements, Conclusions and Future Research Lines

6.1 Achievements

A comprehensive literature review on data-based models for dam behaviour estimation was performed. A selection of articles was analysed, paying attention to the essential aspects of model building and assessment. The weaknesses of the published works were highlighted, and conclusions were drawn on criteria for building data-based dam behaviour models.

The possibilities of 5 state-of-the-art machine learning algorithms for dam behaviour modelling were analysed. Two of them had seldom been applied in this field before (neural networks and support vector machines), while the rest (random forests, boosted regression trees and multi-adaptive regression splines) had, to the best of my knowledge, never been used in dam safety to date. Their prediction accuracy was computed for 14 output variables of three different types (radial and tangential displacements, and leakage), correspondent to a real 100-m high arch dam. Issues related to the training algorithms and criteria to determine the value of the meta-parameters were addressed.

As a result of the previous analysis, BRT models were selected for further assessment. Based on the same case study, the effectiveness of the available tools (partial dependence plots and variable importance measure) for BRT model interpretation was verified. The results of the variable importance measure were presented in an innovative way: as wordclouds. This

kind of plots are well known and often employed in other fields, and showed to be useful for agile interpretation of the results.

The effect of the inclusion of non-causal inputs was assessed, leading to the non-causal models. They showed to be even more accurate, though new issues arose regarding their implementation in dam safety assessment. Criteria for overcoming them were proposed, as well as for the practical implementation of data-based predictive models for the early detection of anomalies:

- A methodology to neglect false anomalies due to the occurrence of extraordinary loads.
- An innovative approach to obtain a realistic estimate of the model accuracy.
- A residual-based criterion to determine the prediction interval (range of safe operation).

These criteria were applied to develop an interactive tool for dam monitoring data analysis and anomaly detection that allows on-line control of dam performance at a glance. Both the code of the application and images of the user interface are included in Section C.3.

A second interactive tool was also developed, which makes use of the "shiny" [11] and "ggplot2" [80] libraries within RStudio [59]. It has the following functionalities:

1. **Data Import** is designed to load time series data to be analysed and used to build predictive models. Alternatively, a previously fitted model can be loaded for its analysis.

2. **Data Exploration** allows pseudo-4D representation of dam monitoring data. Time series for all installed devices (both external and response variables) can be plotted. The user can select which variable to plot in the horizontal and vertical axes. The values are depicted with shape and colour dependent on two extra variables, also selected by the user. In this same section, time series of several outputs can be jointly plotted, together with some external variable in a secondary y-axis. This plot is based on the library "dygraphs" [77], which is highly interactive.

3. **Model Fitting** is designed to build BRT models to estimate different output variables for users unfamiliar with RStudio. The following parameters can be tuned:

    - Output variable to predict
    - Inputs to consider (the resulting model can thus be causal or non-causal)
    - Training parameters (number of trees, shrinkage, interaction depth and bag fraction)
    - Training and validation periods

The output of the application is a plot with predictions and observations, together with the residuals and the MAE for the training and prediction sets.

4. **Interpretation** includes plots showing the input variable importance: a bar chart with the 5 most important variables and a wordcloud with all inputs. Also, partial dependence plots for three predictors selected by the user.

6.2 Conclusions

The main conclusions of the research can be summarised as follows:

- Machine-learning and other advanced data-based tools are becoming familiar in the field of dam safety. The amount of published papers on the field increased in the recent years, most of which showed that the accuracy of deterministic or statistical models can be increased. However, most of them referred to specific case studies, certain dam typologies or determined outputs, and did not deal with model interpretation. As a result, these tools are far from being fully implemented in day-to-day practice.

- ML models typically feature a relatively high amount of parameters. This makes them flexible, but also susceptible to over-fit the training data. Hence, it is essential to check their generalisation capability on an adequate validation data set, not used for fitting the model parameters.

- Among the ML tools analysed, Boosted Regression Trees resulted to be advantageous from an overall viewpoint, since they showed to be more accurate on average for different type of output variables, easy to implement, robust with respect to the training set size, able to consider any kind of input (numeric, categorical or discrete), and low sensitive to noisy and low relevant predictors.

- Nonetheless, other ML algorithms such as Neural Networks, Support Vector Machines or Multi-Adaptive Regression Splines, produced more accurate predictions for some response variables. Moreover, some of them allow further tuning (e.g. variable selection). Therefore, if the main objective is to achieve the best possible fit, the analysis should not be limited to a single technique.

- The accuracy of data-based models as BRTs may decrease dramatically when extrapolating, so the conclusions drawn from their interpretation should be analysed carefully when those situations arise. In this sense, a load combination that presents for first time in dam history is an extraordinary situation, even though the values of the loads are within the historical range, when considered separately. As an example, it was found that the model predictions were unreliable in situations with a combination of

low hydrostatic load and air temperature, although both were higher than the respective historical minimum.

- The application of BRT models to make predictions for a more recent period than that used for training, involves extrapolation over time (provided that some time dependent predictor is considered). Hence, results should be analysed carefully, in particular if the time effect seems relevant. This applies to any data-based model considering time as input, including HST.

- The removal of the early years of dam life cycle from the training set can be beneficial, though results suggest that its influence depends on the algorithm. While it resulted in a decrease in MAE above 10 % for some response variables, BRT accuracy showed lower dependency. Nonetheless, the size of the training set should be considered as an extra parameter to be optimised during training.

- The minimum required training period to obtain a model with reasonable accuracy can be estimated in 5 years, although this value is highly case-dependent. The aspects that influence this minimum are:

  - The load combinations acting during the first years of operation: for example, if the reservoir level remains low, a data-based model will be highly inaccurate when estimating dam response in front of the design flood.
  - The behaviour of dam and foundation during the first filling and the subsequent months. Although transient phenomena are frequent, their magnitude can differ greatly from case to case.
  - The algorithm used to generate the model, and the input variables considered. In particular, non-causal models can be highly accurate with a shorter training period.

- Non-causal models (which include both external and response variables as inputs) are more accurate than causal ones for dam behaviour modelling, and more effective for early detection of anomalies. The reason is that the response variables implicitly provide information that is not included in the causal variables. However, it should be noted that if an anomaly affects several response variables, models that include them as inputs will probably give spurious results. This effect was actually observed in the case study, although they were still advantageous after removing the anomalous variables from the input set.

- BRT models can be efficiently interpreted as regards the relevant questions to be solved in dam safety assessment. Partial dependence plots show the contribution of each input to the output under consideration, as well as the performance evolution over

time. Variable importance measures allow identifying the thermal inertia, as well as the relative influence of each acting load. The results are objective and reliable, since no a priori assumptions need to be made on the shape and intensity of the association between each input and the dam response.

- In spite of the observed advantages of ML algorithms, their results should be checked, when possible, against those provided by other means, such as deterministic models. Also, all available information about the dam behaviour should be taken into account, especially that obtained by visual inspection. Ultimately, engineering judgement based on experience is critical for building the model, interpreting the results, and making decisions with regard to dam safety.

6.3 Future research lines

Future research lines can be drawn from the results of the work, as well as from identified open issues:

- The work focused on BRTs because a robust and highly adaptive algorithm was looked for. However, other tools may be equally or more convenient in certain cases, depending on the variable to predict, the available information, and the characteristics of the dam. As an example, MARS provided greater accuracy in 3 of the 14 variables analysed in the comparative study, and always with a shorter training period. More sophisticated approaches such as the committees of experts (which can be of different nature) could also throw more accurate predictions. A more detailed discussion of this and similar algorithms might determine in which conditions they can be more effective.

- Data-based models obviously require a minimum amount of data to be generated. This means that they cannot be employed during the initial stage of dam life cycle, and in particular during the first filling. In this period, only numerical models are available, thought they also require real data for calibration. Interesting information might be obtained from the application of data-based models on numerically generated data, to narrow prediction intervals in the initial years of dam operation.

- The joint application and analysis of numerical and data-based models can also be advantageous in subsequent stages of dam life cycle, when enough monitoring data is available to build predictive models. Numerical models can be employed to estimate dam response in front of extraordinary loads, to enlarge and enrich the training data. Also, they can be modified to simulate potential anomalies or modes of failure, to generate response data to feed the data-based model. Research on this topic might reveal further possibilities.

- This research was based on the assumption that the time series data available were accurate and complete. Actually, data for the case study presented a low amount of missing values, which were simply interpolated. In the general case, it is highly frequent that long periods of data for determined sensors are missing. This prevents the inclusion of such variables among the inputs or output set, unless the missing values are imputed. Research is necessary to formulate criteria for missing value imputation. They should be dependent, at least, on the type of variable and the length of the missing period. Linear interpolation is appropriate for some variables (e.g. weekly mean temperature) if the missing period is short, but that is not the case in general.

- Many of the dams in operation were built decades ago, and their monitoring data is heterogeneous, incomplete or hand-written. In some cases, the lack of information might make impossible to apply any data-based model. A general picture of the quality of monitoring data would allow to develop tools and criteria to import data into an appropriate format and take full advantage of the available information.

- Flexibility was one of the premises throughout the research. BRTs were chosen because of their accuracy, but also because they automatically adapt to a variety of situations in terms of input variables availability and strength and shape of input-output association. Nonetheless, application of the tool and methodology to a set or real dams of different typologies would reveal specific issues to be solved.

- When an anomaly that affects several response variables occur, the non-causal models that rely on such variables as inputs give false positives on the unaffected devices. In the implementation developed, this problem is avoided by simply eliminating all variables considered anomalous in a first iteration. A more detailed study of this issue could allow developing a general criterion for identifying variables that are in fact abnormal, taking full advantage of all available information.

- The application developed displays the observations of the selected devices with different colour, depending on whether the system considered them as normal or abnormal. These colours are drawn over a front view of the dam, with each device in its actual location. In case of incipient failure, it could be useful to identify the potential causes, taking into account the dam typology, and the number and location of devices whose measure is identified as anomalous. A more detailed study, would allow defining colour patterns associated to potential failure modes.

Bibliography

[1] F. Amberg. Interpretative models for concrete dam displacements. In *23th ICOLD Congress*, 2009. Q91-R43.

[2] S. Arlot, A. Celisse, et al. A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79, 2010.

[3] L. Auret and C. Aldrich. Empirical comparison of tree ensemble variable importance measures. *Chemometrics and Intelligent Laboratory Systems*, 105(2):157–170, 2011.

[4] Z. Bofang. Prediction of water temperature in deep reservoirs. *Dam Engineering*, 8:13–26, 1997.

[5] S. Bonelli and H. Félix. Interpretation of measurement results, delayed response analysis of temperature effect. In *Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams*, Salzburg, Austria, Oct. 2001.

[6] S. Bonelli and K. Radzicki. Impulse response function analysis of pore pressure in earthdams. *European Journal of Environmental and Civil Engineering*, 12(3):243–262, 2008.

[7] S. Bonelli and P. Royet. Delayed response analysis of dam monitoring data. In *Proceedings of the Fifth ICOLD European Symposium on Dams in a European Context*, Geiranger, Norway, 2001.

[8] L. Breiman. *Classification and regression trees*. Chapman & Hall/CRC, 1984.

[9] L. Breiman et al. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3):199–231, 2001.

[10] A. Carrère and C. Noret-Duchêne. Interpretation of an arch dam behaviour using enhanced statistical models. In *Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams*, Salzburg, Austria, 2001.

[11] W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson. *shiny: Web Application Framework for R*, 2016. R package version 0.13.2.

[12] L. Cheng and D. Zheng. Two online dam safety monitoring models based on the process of extracting environmental effect. *Advances in Engineering Software*, 57:48–56, 2013.

[13] L. Chouinard, D. Bennett, and N. Feknous. Statistical analysis of monitoring data for concrete arch dams. *Journal of Performance of Constructed Facilities*, 9(4):286–301, 1995.

[14] L. Chouinard and V. Roy. Performance of statistical models for dam monitoring data. In *Joint International Conference on Computing and Decision Making in Civil and Building Engineering, Montreal, June*, pages 14–16, 2006.

[15] O. Crépon and M. Lino. An analytical approach to monitoring. *Water Power and Dam Construction*, 1999.

[16] A. De Sortis and P. Paoliani. Statistical analysis and structural identification in concrete dam monitoring. *Engineering Structures*, 29(1):110–120, 2007.

[17] S. Demirkaya. Deformation analysis of an arch dam using ANFIS. In *Proceedings of the second international workshop on application of artificial intelligence and innovations in engineering geodesy. Braunschweig, Germany*, page 2131, 2010.

[18] S. Demirkaya and M. Balcilar. The contribution of soft computing techniques for the interpretation of dam deformation. In *Proceedings of the FIG working week*, Rome, Italy, 2012.

[19] P. Duffaut. The traps behind the failure of malpasset arch dam, france, in 1959. *Journal of Rock Mechanics and Geotechnical Engineering*, 5(5):335–341, 2013.

[20] T. Ebert, J. Belz, and O. Nelles. Interpolation and extrapolation: Comparison of definitions and survey of algorithms for convex and concave hulls. In *Computational Intelligence and Data Mining (CIDM), 2014 IEEE Symposium on*, pages 310–314. IEEE, 2014.

[21] J. Elith, J. R. Leathwick, and T. Hastie. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4):802–813, 2008.

[22] I. Fellows. *wordcloud: Word Clouds*, 2014. R package version 2.5.

[23] J. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, pages 1189 – 1232, 2001.

[24] J. H. Friedman and J. J. Meulman. Multiple additive regression trees with application in epidemiology. *Statistics in medicine*, 22(9):1365–1381, 2003.

[25] G. Ridgeway with contributions from others. *gbm: Generalized Boosted Regression Models*, 2013. R package version 2.1.

[26] S. Gamse and M. Oberguggenberger. Assessment of long-term coordinate time series using hydrostatic-season-time model for rock-fill embankment dam. *Structural Control and Health Monitoring*, 2016.

[27] Q. Guedes and P. Coelho. Statistical behaviour model of dams. In *15th ICOLD Congres*, pages Q56–R16, 319–334, 1985.

[28] T. Hastie, R. Tibshirani, and J. Firedman. *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*. Springer, New York, 2 edition, 2009.

[29] M. Herrera, L. Torgo, J. Izquierdo, and R. Pérez-García. Predictive models for forecasting hourly urban water demand. *Journal of Hydrology*, 387(1):141–150, 2010.

[30] V. J. Hodge and J. Austin. A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22(2):85–126, 2004.

[31] R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2014.

[32] International Commission on Large Dams. Automated dam monitoring systems. guidelines and case histories. Technical Report B-118, ICOLD, 2000.

[33] International Commission on Large Dams. Dam surveillance guide. Technical Report B-158, ICOLD, 2012.

[34] I.-S. Jung, M. Berges, J. H. Garrett, and B. Poczos. Exploration and evaluation of ar, mpca and kl anomaly detection techniques to embankment dam piezometer data. *Advanced Engineering Informatics*, 29(4):902–917, 2015.

[35] C.-Y. Kao and C.-H. Loh. Monitoring of long-term static deformation data of feitsui arch dam using artificial neural network-based approaches. *Structural Control and Health Monitoring*, 20(3):282–303, 2013.

[36] O. Kaser and D. Lemire. Tag-cloud drawing: Algorithms for cloud visualization. *arXiv preprint cs/0703109*, 2007.

[37] J. Leathwick, J. Elith, M. Francis, T. Hastie, and P. Taylor. Variation in demersal fish species richness in the oceans surrounding new zealand: an analysis using boosted regression trees. *Marine Ecology Progress Series*, 321:267–281, 2006.

[38] F. Li, Z. Wang, and G. Liu. Towards an error correction model for dam monitoring data analysis based on cointegration theory. *Structural Safety*, 43:1220, 2013.

[39] C.-H. Loh, C.-H. Chen, and T.-Y. Hsu. Application of advanced statistical methods for extracting long-term trends in static monitoring data from an arch dam. *Structural Health Monitoring*, 10(6):587–601, 2011.

[40] G. Lombardi. Advanced data interpretation for diagnosis of concrete dams. Technical report, CISM, 2004.

[41] G. Lombardi, F. Amberg, and G. Darbre. Algorithm for he prediction of functional delays in the behaviour of concrete dams. *Hydropower and Dams*, (3):111–116, 2008.

[42] J. Mata. Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Engineering Structures*, 3(3):03 – 910, 2011.

[43] J. Mata, N. S. Leitão, A. T. de Castro, and J. S. da Costa. Construction of decision rules for early detection of a developing concrete arch dam failure scenario. a discriminant approach. *Computers & Structures*, 142:45–53, 2014.

[44] J. Mata, A. Tavares de Castro, and J. Sá da Costa. Constructing statistical models for arch dam deformation. *Structural Control Health Monitoring*, 21(3):423–437, Mar. 2014.

[45] A. Michelis. Traditional versus non-traditional boosting algorithms. Master's thesis, University of Manchester, 2012.

[46] B. Myers and D. Scofield. Providing improved dam safety monitoring using existing staff resources: Fern Ridge Dam case study. In *Proceedings of 28th Annual USSD Conference*, 2008.

[47] P. Palumbo, L. Piroddi, S. Lancini, and F. Lozza. NARX modeling of radial crest displacements of the Schlegeis arch dam. In *Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams*, Salzburg, Austria, 2001.

[48] A. Panizzo and A. Petaccia. Analysis of monitoring data for the safety control of dams using neural networks. In *New Trends in Fluid Mechanics Research*, page 344347. Springer, 2009.

[49] M. Papadrakakis, V. Papadopoulos, N. D. Lagaros, J. Oliver, A. E. Huespe, and P. Sánchez. Vulnerability analysis of large concrete dams using the continuum strong discontinuity approach and neural networks. *Structural Safety*, 30(3):217–235, 2008.

[50] I. Penot, B. Daumas, and J. Fabre. Monitoring behaviour. *Water Power and Dam Construction*, 2005.

[51] J. Pérez and E. Martínez. La acción térmica del medio ambiente como solicitación de diseño en proyectos de presas españolas. *Rev Obras Públicas*, 3349:79–90, 1995.

[52] L. Piroddi and W. Spinelli. Long-range nonlinear prediction: a case study. In *42nd IEEE Conference on Decision and Control*, volume 4, pages 3984–3989. IEEE, 2003.

[53] A. Popovici, C. Ilinca, and T. Ayvaz. The performance of the neural networks to model some response parameters of a buttress dam to environment actions. In *Proceedings of the 9th ICOLD European Club Symposium*, Venice, Italy, 2013.

[54] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[55] V. Ranković, N. Grujović, D. Divac, and N. Milivojević. Development of support vector regression identification model for prediction of dam structural behaviour. *Structural Safety*, 48:33–39, 2014.

[56] V. Ranković, N. Grujović, D. Divac, N. Milivojević, and A. Novaković. Modelling of dam behaviour based on neuro-fuzzy identification. *Engineering Structures*, 35:107–113, 2012.

[57] V. Ranković, A. Novaković, N. Grujović, D. Divac, and N. Milivojević. Predicting piezometric water level in dams via artificial neural networks. *Neural Computing and Applications*, 24(5):1115–1121, 2014.

[58] G. Ridgeway. *Generalized Boosted Models: A guide to the gbm package*, 2007. R package vignette.

[59] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2015.

[60] F. Salazar, J. González, M. Toledo, and E. Oñate. A methodology for dam safety evaluation and anomaly detection based on boosted regression trees. In *Proceedings of the 8th European Workshop on Structural Health Monitoring*, Bilbao, Spain, 2016.

[61] F. Salazar, R. Morán, M. Á. Toledo, and E. Oñate. Data-based models for the prediction of dam behaviour: A review and some methodological considerations. *Archives of Computational Methods in Engineering*, pages 1–21, 2015.

[62] F. Salazar, M. Toledo, E. Oñate, and R. Morán. An empirical comparison of machine learning techniques for dam behaviour modelling. *Structural Safety*, 56:9–17, 2015.

[63] F. Salazar, M. Á. Toledo, E. Oñate, and B. Suárez. Interpretation of dam deformation and leakage with boosted regression trees. *Engineering Structures*, 119:230–251, 2016.

[64] F. J. Sánchez Caro. *Dam safety: contributions to the deformation analysis and monitoring as an element of prevention of pathologies of geotechnical origin.* PhD thesis, UPM, 2007. [In Spanish].

[65] D. Santillán, J. Fraile-Ardanuy, and M. Toledo. Dam seepage analysis based on artificial neural networks: The hysteresis phenomenon. In *Neural Networks (IJCNN), The 2013 International Joint Conference on*, pages 1–8. IEEE, 2013.

[66] D. Santillán, J. Fraile-Ardanuy, and M. Toledo. Seepage prediction in arch dams by means of artificial neural networks. *Water Technology and Science*, V(3), 2014.

[67] D. Santillán, E. Salete, D. Vicente, and M. Toledo. Treatment of solar radiation by spatial and temporal discretization for modeling the thermal response of arch dams. *Journal of Engineering Mechanics*, 140(11), 2014.

[68] V. Saouma, E. Hansen, and B. Rajagopalan. Statistical and 3d nonlinear finite element analysis of schlegeis dam. In *Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams*, pages 17–19, 2001.

[69] R. E. Schapire. The boosting approach to machine learning: An overview. In *Nonlinear estimation and classification*, pages 149–171. Springer, 2003.

[70] L. Seifard, A. Szpilman, and C. Piasentin. Itaipu structures. evaluation of their performance. In *15th ICOLD Congress*, pages 287–317, 1985. Q56-R15.

[71] A. Simon, M. Royer, F. Mauris, and J. Fabre. Analysis and interpretation of dam measurements using artificial neural networks. In *Proceedings of the 9th ICOLD European Club Symposium*, Venice, Italy, 2013.

[72] B. Stojanovic, M. Milivojevic, M. Ivanovic, N. Milivojevic, and D. Divac. Adaptive system for dam behavior modeling based on linear regression and genetic algorithms. *Advances in Engineering Software*, 65:182190, 2013.

[73] Swiss Committee on Dams. Methods of analysis for the prediction and the verification of dam behaviour. Technical report, ICOLD, 2003.

[74] M. Tatin, M. Briffaut, F. Dufour, A. Simon, and J. Fabre. Thermal displacements of concrete dams: Finite element and statistical modelling. In *9th ICOLD European Club Symposium*, Venice, Italy, 2013.

[75] M. Tatin, M. Briffaut, F. Dufour, A. Simon, and J.-P. Fabre. Thermal displacements of concrete dams: Accounting for water temperature in statistical models. *Engineering Structures*, 91:26–39, 2015.

[76] G. Tayfur, D. Swiatek, A. Wita, and V. P. Singh. Case study: Finite element method and artificial neural network models for flow through Jeziorsko earthfill dam in Poland. *Journal of Hydraulic Engineering*, 131(6):431440, 2005.

[77] D. Vanderkam, J. Allaire, J. Owen, D. Gromer, P. Shevtsov, and B. Thieurmel. *dygraphs: Interface to 'Dygraphs' Interactive Time Series Charting Library*, 2016. R package version 1.1.1.3.

[78] M. Verleysen et al. Learning high-dimensional data. *Nato Science Series Sub Series III Computer And Systems Sciences*, 186:141–162, 2003.

[79] A. S. Weigend, B. A. Huberman, and D. E. Rumelhart. Predicting sunspots and exchange rates with connectionist networks. In S. Eubank and M. Casdagli, editors, *Proc. of the 1990 NATO Workshop on Nonlinear Modeling and Forecasting (Santa Fe, NM)*, volume 12, pages 395–432. Addison-Wesley, Redwood, CA, 1992.

[80] H. Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009.

[81] G. Willm and N. Beaujoint. Les méthodes de surveillance des barrages au service de la production hydraulique d'Electricité de France-Problèmes ancients et solutions nouvelles. In *9th ICOLD Congres*, pages 529–550, 1967. Q34-R30. [in French].

[82] H. Xu and X. Li. Inferring rules for adverse load combinations to crack in concrete dam from monitoring data using adaptive neuro-fuzzy inference system. *Science China Technological Sciences*, 55(1):136141, 2012.

[83] H. Yu, Z. Wu, T. Bao, and L. Zhang. Multivariate analysis in dam monitoring data with PCA. *Science China Technological Sciences*, 53(4):1088–1097, 2010.

Appendices

A

Articles in the compilation

A.1 Data-based models for the prediction of dam behaviour. A review and some methodological considerations

Title: Data-based models for the prediction of dam behaviour. A review and some methodological considerations

First Author: Fernando Salazar González. CIMNE - International Center for Numerical Methods in Engineering

Second Author: Rafael Morán Moya. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Third Author: Miguel Á. Toledo Municio. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Fourth Author: Eugenio Oñate Ibáñez de Navarra. CIMNE - International Center for Numerical Methods in Engineering

Journal: Archives of Computational Methods in Engineering

D.O.I. 10.1007/s11831-015-9157-9

Impact Factor 4.214

Archives of Computational Methods in Engineering manuscript No.
(will be inserted by the editor)

Data-based models for the prediction of dam behaviour

A review and some methodological considerations

Fernando Salazar · Rafael Morán · Miguel Á. Toledo · Eugenio Oñate

doi:10.1007/s11831-015-9157-9

Abstract Predictive models are an important element in dam safety analysis. They provide an estimate of the dam response faced with a given load combination, which can be compared with the actual measurements to draw conclusions about dam safety. In addition to numerical finite element models, statistical models based on monitoring data have been used for decades for this purpose. In particular, the hydrostatic-season-time method is fully implemented in engineering practice, although some limitations have been pointed out. In other fields of science, powerful tools such as neural networks and support vector machines have been developed, which make use of observed data for interpreting complex systems. This paper contains a review of statistical and machine-learning data-based predictive models, which have been applied to dam safety analysis. Some aspects to take into account when developing analyses of this kind, such as the selection of the input variables, its division into training and validation sets, and the error analysis, are discussed. Most of the papers reviewed deal with one specific output variable of a given dam typology and the majority also lack enough validation data. As a consequence, although results are promising, there is a need for further validation and assessment of generalisation capability. Future research should also focus on the development of criteria for data preprocessing and model application.

Keywords Dam monitoring · Dam safety · Data analysis · Machine learning · Statistical models · Behaviour models

F. Salazar
CIMNE Centre Internacional de Metodes Numerics en Enginyeria
Tel.: +34-93-401-74-95
E-mail: fsalazar@cimne.upc.edu

R. Morán
Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment. Profesor Aranguren s/n, 28040, Madrid, Spain
Tel.: +34-91-336-67-09
E-mail: rmoran@caminos.upm.es

Miguel Á. Toledo
Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment. Profesor Aranguren s/n, 28040, Madrid, Spain
Tel.: +34-91-336-67-09
E-mail: matoledo@caminos.upm.es

Eugenio Oñate
CIMNE Centre Internacional de Metodes Numerics en Enginyeria
Universitat Politècnica de Catalunya (UPC). Barcelona, Spain
Tel.: +34-93-401-74-95
E-mail: onate@cimne.upc.edu

1 Introduction

Behaviour models are a fundamental component of dam safety systems, both for the daily operation and for long-term behaviour evaluation. They are built to calculate the dam response under safe conditions for a given load combination, which is compared to actual measurements of dam performance [71]. The result is an essential ingredient for dam safety assessment, together with visual inspection and engineering judgement [27].

Numerical models based on the finite element method (FEM) are widely used to predict dam response, in terms of displacements, strains and stresses. They are based on the physical laws governing the involved phenomena, which gives them some interesting features: a) they are useful for the design and,

more importantly, for dam safety assessment during the first filling, and b) they can be conveniently interpreted, provided that their parameters have physical meaning.

On the contrary, some relevant indicators of dam safety, such as uplift pressure and leakage flow in concrete dams, cannot be predicted accurately enough with numerical models [38], [39]. In addition, the knowledge of the stress-strain properties of the dam and foundation materials is always limited [75], and so is the prediction accuracy of FEM models [27].

These limitations, together with the availability of monitoring data, have fostered the application of statistical models to predict dam response. They have been used in dam safety analyses for decades as a complement to visual inspection and numerical models, to support decision making.

In recent years, there is a tendency towards automatising dam monitoring devices [27], which allows for increasing the reading frequency and results in a greater amount of data available. Although it encourages extraction of as much information as possible in relation with dam safety conditions [57], it has revealed certain limitations of traditional statistical tools to manage dam monitoring data [58].

On another note, advanced tools have been developed in the machine learning (ML) community to build data-based predictive models. They have been applied in various fields of science and engineering, where similar problems have emerged more dramatically, provided that the amount of data is much larger or the underlying phenomena is much less understood. This is the case, for example, of medicine, e-commerce, smartphone applications, econometrics or business intelligence, among others. Most of these tools exclusively rely on data to build predictive models, i.e. no prior assumptions on the physics of the phenomenon have to be made beforehand [25].

The limitations of traditional statistical tools and the availability of these advanced learning algorithms have motivated dam engineers to search the possibilities of the latter for building dam behaviour models, as well as for analysing dam behaviour.

This paper reports a review on dam behaviour models based on monitoring data. The work focuses on prediction accuracy, although it also refers to model suitability for interpreting dam performance. The most popular techniques are dealt with in section 2, whereas some common issues in building data-based models and evaluating their results are analysed in section 3. The analysis is performed on the basis of the review of 40 papers on the field.

2 Statistical and machine learning techniques used in dam monitoring analysis

The aim of these models is to predict the value of a given variable $Y \in \mathbb{R}$ (e.g. displacement, leakage flow, crack opening, etc.), in terms of a set of inputs[1] $X \in \mathbb{R}^d$:

$$Y = \hat{Y} + \varepsilon = F(X) + \varepsilon \tag{1}$$

ε is an error term, which encompasses the measurement error, the model error, and the deviation of the dam response from the expected behaviour [71]. This term is important, given that it is frequently used to define safety margins and warning thresholds [27].

The models are fitted on the basis of a set of observed input data x_i, and the correspondent registered outputs y_i, where $i = 1, ..., N$ and N is the number of observations. Note that each x_i is a vector of d components, being d the number of inputs.

The inputs may be of different nature, depending on the method:

- Raw data recorded by the monitoring system, which in turn can be:
  - External variables: reservoir level (h), air temperature (T), etc.
  - Internal variables: temperature in the dam body, stresses, displacements, etc.
- Variables derived from observed data. For example:
  - Polynomials
  - Moving averages
  - Derivatives

2.1 Hydrostatic-seasonal-time (HST) model

The most popular data-based approach for dam monitoring analysis is the hydrostatic-seasonal-time (HST) model. It was first proposed by Willm and Beaujoint in 1967 [76] to predict displacements in concrete dams, and has been widely applied ever since. It is based on the assumption that the dam response is a linear combination of three effects:

$$\hat{Y} = F_1(h) + F_2(s) + F_3(t) \tag{2}$$

[1] Traditionally, the statistical models applied in dam monitoring analysis were based on causal variables, e.g., hydrostatic load and temperature, which are often termed "independent variables". On the contrary, other algorithms make use of transformed variables (such as gradients or moving averages), and non-causal observations (e.g. the previous value of the output). This has led to the use of various terms to refer to the model inputs, such as "predictors", "covariates", and "features". In this paper they are used indistinctly.

– A reversible effect of the hydrostatic load which is commonly considered in the form of a fourth-order polynomial of the reservoir level (h) [71], [4], [67]:

$$F_1(h) = a_0 + a_1 h + a_2 h^2 + a_3 h^3 + a_4 h^4 \quad (3)$$

– A reversible influence of the air temperature, which is assumed to follow an annual cycle. Its effect is approximated by the first terms of the Fourier transform:

$$F_2(s) = a_5 cos(s) + a_6 sen(s) + a_7 sen^2(s) + a_8 sen(s) cos(s) \quad (4)$$

where $s = 2\pi d/365.25$ and d is the number of days since 1 January.

– An irreversible term due to the evolution of the dam response over time. A combination of monotonic time-dependant functions is frequently considered. The original form is [76]:

$$F_3(t) = a_9 log(t) + a_{10} e^t \quad (5)$$

The model parameters $a_1 ... a_{10}$ are adjusted by the least squares method: the final model is based on the values which minimise the sum of the squared deviations between the model predictions and the observations.

Some authors used variations of the original HST model, by using some heuristics or after a trial-and-error process. Mata [40] considered the irreversible effect by means of $F_3(t) = a_9 t + a_{10} e^{-t}$. Chouinard and Roy [12] used a linear term in t and a third-order polynomial of h. Simon *et al.* [67] chose $F_3(t) = a_9 e^{-t} + a_{10} t + a_{11} t^2 + a_{12} t^3 + a_{13} t^4$, whereas Yu *et al.* [80] used $F_3(t) = a_9 t + a_{10} t^2 + a_{11} t^3$. Carrère applied a variation of HST in which the possibility of a sudden change in the dam response at a certain time is considered by adding a step function to the irreversible term [9].

The method makes use of strong assumptions on the response of the dam, which might not be fulfilled in general. In particular, the three effects are considered as independent, although it is well known that certain collinearity exists. The reservoir level affects the thermal response of the dam, provided that the air and water temperatures differ [73]. In some cases, the reservoir operation follows an annual cycle due to the evolution of the water demand, so there is a strong correlation between h and the air temperature [38], [66], [33], [13]. Collinearity may lead to poor prediction accuracy and, more importantly, to misinterpretation of the results [1].

Another limitation of the original form of HST model is that the actual air temperature is not considered. On one hand, this makes it more flexible, because it can be applied in dams where air temperature measurements are not available. On the other hand, it reduces its prediction accuracy for particularly warm or cold years [73], [66].

Several alternatives have been proposed to overcome this shortcoming. Penot *et al.* [50] introduced the HSTT method, in which the thermal periodic effect is corrected according to the actual air temperature. This procedure has been applied at Electricité de France (EDF) [73], [20] with higher accuracy than HST, especially during the 2003 European heat wave. Although the proposal of this method has been frequently attributed to Penot *et al.*, Breitenstein *et al.* [8] applied a similar scheme 20 years earlier.

Tatin *et al* [73], [74] proposed further corrections of HSTT. The HST-Grad model takes into account both the mean and the gradient of the temperature in the dam body, considered as a one-dimensional domain. They are estimated from the air temperature in the downstream face, and from a weighted average of the air and water temperatures in the upstream one. A similar and more detailed approach was applied by the same authors, called the SLICE model [73]. It considers different thermal conditions for the portion of the dam body located below the pool level to that situated above, which is not affected by the water temperature.

Other common choice is to replace the periodic function of the thermal component by the actual temperature in the dam body, resulting in the hydrostatic-thermal-time (HTT) method. One difficulty of this approach is how to select the appropriate thermometers among those available. In arch dams, some authors only consider the thermometers in the central cantilever, assuming that it represents the thermal equilibrium between cantilevers in the right and left margins [66]. Mata *et al.* [42] solved this issue by applying principal component analysis (PCA), while other authors [33] considered all the available instruments. Li *et al.* [34] proposed an error correction model (ECM), featuring a term which depends on the error in the estimation of previous output values.

Although HST was originally devised for the prediction of displacements in concrete dams, it has also been applied to predict other variables. Simon *et al.* [67] estimated uplifts and leakage with HST, although they obtained more accurate results with neural networks (NN). Guedes and Coelho [24] built a model for the prediction of leakage in Itaipú Dam with the form $a_1 h_{6,11}^2 + a_2 t + a_3 t^2 + a_4 log(1+t)$, where $h_{6,11}$ is the average reservoir level between 6 and 11 days before the measurement. Breitenstein *et*

al. [8] also studied leakage, although they discarded both the seasonal and the temporal terms. Yu *et al.* [80] combined HST with PCA to predict the opening of a longitudinal crack in Chencun Dam.

A common feature to HST and its variations is that the output is computed as a linear combination of the inputs. Hence, they are all multi-lineal regression models (MLR), so their coefficients can be fitted by least squares. Other approaches based on MLR have been applied in dam safety, considering a larger set of inputs (e.g. [69], [19]).

2.2 Models to account for delayed effects

It is well known that dams respond to certain loads with some delay [39]. The most typical examples are:

- The change in pore pressure in an earth-fill dam due to reservoir level variation [6].
- The influence of the air temperature in the thermal field in a concrete dam body [67].

Other phenomena have been identified which are governed by similar processes. For example, Lombardi [38] noticed that the structural response of an arch dam to hydrostatic load comprised both elastic and viscous components. Hence, the displacements not only depended on the instantaneous reservoir level, but also on the past values. Simon *et al.* [67] reported that leakage flow at Bissorte Dam responded to rainfall and snow melt with certain delay.

Several approaches have been proposed to account for these effects. The most popular consists of including moving averages or gradients of some explanatory variables in the set of predictors. In the above mentioned study, Guedes and Coelho [24] predicted the leakage flow on the basis of the mean reservoir level over the course of a five-days period. Sánchez Caro [62] included the 30 and 60 days moving average of the reservoir level in the conventional HST formulation to predict the radial displacements of El Atazar Dam. Popovici *et al.* [53] used moving averages of 3, 10 and 30 days of the air temperature, together with the pool level in the previous 3 days to the measurement in order to predict displacements in a buttress dam with neural networks (NN). Crépon and Lino [15] reported significant improvement in the prediction of piezometric levels and leakage flows by considering the accumulated rainfall and the derivative of the hydrostatic load as predictors.

This approach requires a criterion to determine which moving averages and gradients should be considered for each particular case. Demirkaya and Balcilar [19] performed a sensitivity analysis to select the number of past values to include both in an MLR and in a NN model. They used the same period for the external and internal temperatures, as well as for the reservoir level, and found that the most accurate results were obtained with an MLR model considering data from 30 previous days. Although their results compared well to those proposed by the participants in the 6th ICOLD Benchmark Workshop[2] [81], they lacked physical meaning: they would imply that the dam responded with the same delay to the water level, the air temperature, and the internal temperature field.

Santillán *et al.* [64] proposed a methodology to select the optimal set of predictors among various gradients of air temperature and reservoir level. They used the gradients instead of the moving averages to ensure independence among predictors (moving averages are correlated with the original correspondent variables). They combined it with NN to predict leakage flow in an arch dam.

A more formal alternative to conventional HST to account for delayed effects was proposed by Bonelli and Royet [7]. It is based on the hypothesis that the delayed effect depends on the convolution integral of the impulse response function (IRF) and the loadings:

$$\hat{Y} = \alpha \frac{1}{t_0} \int_0^t e^{-\left(\frac{t \cdot t'}{t_0}\right)} h\left(t'\right) \partial t' \tag{6}$$

where α is a damping coefficient, t_0 is the characteristic time, which depends on the phenomenon, and $h\left(t'\right)$ is the reservoir level at time t'. Although the analytical integration of this function is cumbersome, it can be solved by means of numerical approximation. The advantage of this approach is that the coefficients have physical meaning: the characteristic time provides insight into the lag with which the dam reacts to a variation in the input variable, whereas the damping reflects the relation between the amplitude of the reservoir level variation and that of the pore pressure in the location considered within the dam body.

A similar approach was followed by the same author in the frame of the above mentioned 6th ICOLD Benchmark Workshop [4]. In this case, it was intended to account for the delayed response of the

[2] In the 6th ICOLD Benchmark Workshop, the participants were asked to provide a data-based model for predicting the radial displacement of Schlegeiss arch dam for the period 1999-2000. Time histories of water level, air temperature and concrete temperatures at various locations were provided for the period 1992-2000, as well as the observed values of the target variable for the period 1992-1998.

dam in terms of the temperature field, with the final aim of predicting radial displacements.

Lombardi *et al.* [38] suggested an equivalent formulation, also to compute the thermal response of the dam to changes in air temperature. Although the development was slightly different, the numerical approximation to the integral is equivalent. Lombardi arrived at the following expression [39]:

$$\hat{Y}(t) = \alpha \cdot Y(t - \Delta t) + \left(1 + \frac{\alpha}{\beta} - \frac{1}{\beta}\right) X(t) + \left(\frac{1}{\beta} - \frac{\alpha}{\beta} - \alpha\right) X(t - \Delta t) \quad (7)$$

where $\alpha = e^{\frac{-\Delta t}{t_0}}$, $\beta = \frac{\Delta t}{t_0}$, and Δt is the measurement interval. It should be noted that the numerical integration of (6) by means of (7) leads to a predictive model which is a linear combination of:

- the value of the predictors at t and $t - \Delta t$
- the value of the output variable at $t - \Delta t$

This is the conventional form of a first order auto-regressive exogenous (ARX) model. In general, these models require specific algorithms to determine the appropriate order of the model for a given case, i.e., the amount of past values to consider for the output and each of the input variables. The next section is devoted to this aspect and to auto-regressive models.

In practice, an input transformed by equation (6) is similar to a weighted moving average (WMA) [39]. Figure 1 shows the comparison between both transformations of 4 inputs: a) a sinusoidal, b) a random variable, c) a cyclic variable with random noise and d) an isolated pulse. It can be seen that the transformed sinusoidal can be accurately modelled with an appropriate moving average. The difference between IRF and WMA is greater for random inputs, and the discrepancy increases as the signal-to-noise ratio decreases.

IRF has the advantage of its physical meaning, and has offered accurate results for determined outputs. Nonetheless, given that it makes a strong assumption on the characteristics of the phenomenon, it is restricted to specific processes. Even when applied to a similar phenomenon, such as the effect of precipitation on the pore pressure on an earth-fill dam, the accuracy decreases [7]. Moreover, the coefficients lose their physical meaning in this case.

2.3 Auto-regressive (AR) models

The use of the previous (lagged) value of the output to calculate a prediction for current record may induce to question a) whether the observed previous value or the precedent prediction should be used, and b) whether the model parameters should be readjusted at every time step.

In general, using the actual previous value and refitting the model should provide better prediction accuracy, but such a model would not be able to detect gradual anomalies [79]: it would learn the abnormal behaviour and treat it as ordinary [38]. Riquelme *et al.* [59] improved the accuracy of a NN model by several orders of magnitude by applying this approach.

The opposite alternative is to fit the model to data gathered for a given time period, and make long-term predictions on a step-by-step basis [48], i.e., predict the output at $t + 1$, and use it (the prediction; not the observation) to estimate the value at $t + 2$. This procedure may fail in error propagation [10], but in principle should be appropriate to unveil gradual anomalies.

An intermediate choice is to use the actual measurement of the output variable, without readjusting the model parameters. In this case, the coefficients obtained on the basis of a period of normal behaviour are applied to future observations, hence the model could detect changes in the relation between current and next values of the output.

Although several authors built predictive models based on lagged output values, most of them did not mention which of the described approaches applied. Palumbo *et al.* [48], should have used the previous prediction, given that they presented a solution to the 6th ICOLD Benchmark Workshop, and the observed values of the output were not provided to the participants beforehand.

If the possibility of including past values of the variables is considered, a criterion to select some of the available shall be defined. Otherwise, the amount of predictors is quite high. For example, Piroddi and Spinelli [52] considered the most general form of a non-linear autoregressive exogenous model (NARX), which depended on current and previous values of the input variables, on precedent values of the output, as well as on linear and non-linear combinations of them. They applied a specific algorithm for selecting 11 predictors in the final model.

In general, these models prioritise prediction accuracy over explanatory capability. The greater the number of variables in the model, the harder it is to interpret and to isolate the effect of each component. Nonetheless, some procedures have been proposed to interpret models whose parameters do not have physical meaning, as described in section 3.2.

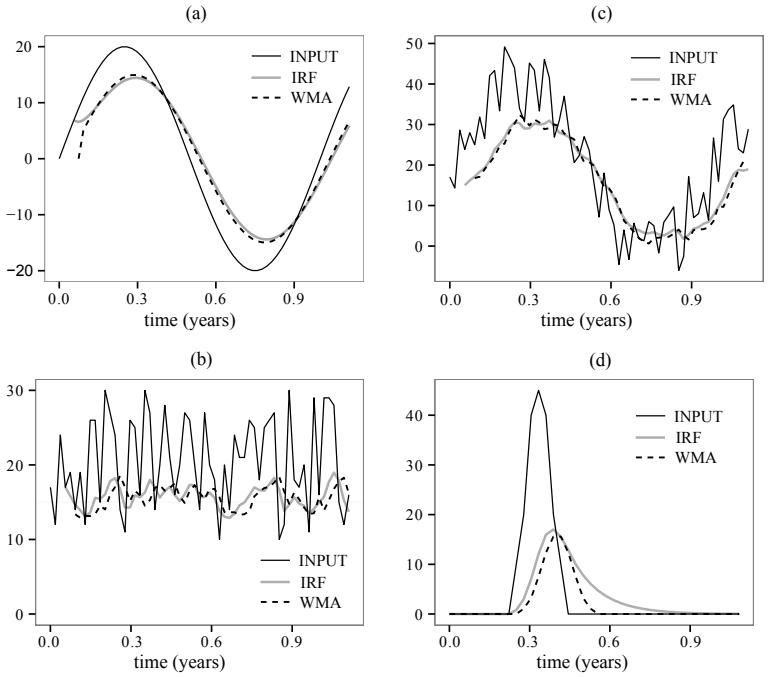


Fig. 1 Comparison between impulse response function (IRF) and weighted moving average (WMA) for various inputs: a) sinusoidal, b) random, c) sinusoidal with random noise and d) impulse.

2.4 Neural networks (NN)

Linear models are not well suited to reproduce non-linear behaviour, even though some actions are considered in the form of high order polynomials [12]. On the contrary, NN models are flexible, and allow modelling complex and highly non-linear phenomena. Although there are various types of NN models [3], the vast majority of applications for dam monitoring data analysis are based on the multilayer perceptron (MLP). Such models, as their name suggests, are comprised by a number of perceptrons (also called "units", or "neurons") organised in different layers: input, hidden, and output (Figure 2). In principle, several hidden layers can be used (see section 2.6), but one is mostly adopted in practice [3].

The input of each unit U_l is a linear combination of the predictors X^j:

$$c_l = \sum_{j=1}^{d} X^j \cdot w_l^j + b_l \tag{8}$$

which is later transformed by an activation function g to compute the neuron's output:

$$z_l = g(c_l) \tag{9}$$

Several forms of g can be chosen (non-linear in general), although sigmoid functions are often employed, such as the logistic (10) and the hyperbolic tangent (11) (Figure 3). As an exception, Su and Wu [70] selected Mexico-hat wavelet functions (12) to obtain a wavelet neural network (WNN) model, otherwise similar to conventional NN models described in this section.

$$g(c_l) = \frac{1}{1 + e^{-c_l}} \tag{10}$$

$$g(c_l) = \frac{e^{c_l} - e^{-c_l}}{e^{c_l} + e^{-c_l}} \tag{11}$$

$$g\left(c_l\right) = \left(1 - {c_l}^2\right) \cdot e^{\left(1 - \frac{{c_l}^2}{2}\right)} \tag{12}$$

The output layer may be composed of one of the described neurons, although a linear transform is frequently chosen, so that the overall model output is computed as:

$$\hat{Y} = \sum_{l=1}^{L} w_{out}^l \cdot g\left(\sum_{j=1}^{d} X^j w_l^j + b_l\right) + b_{out} \tag{13}$$

NN models can be thought of as an extension of MLR, which output c_l is expanded by the perceptron through a non-linear transformation g [25]. It should be noted (Figure 3) that the sigmoid functions have

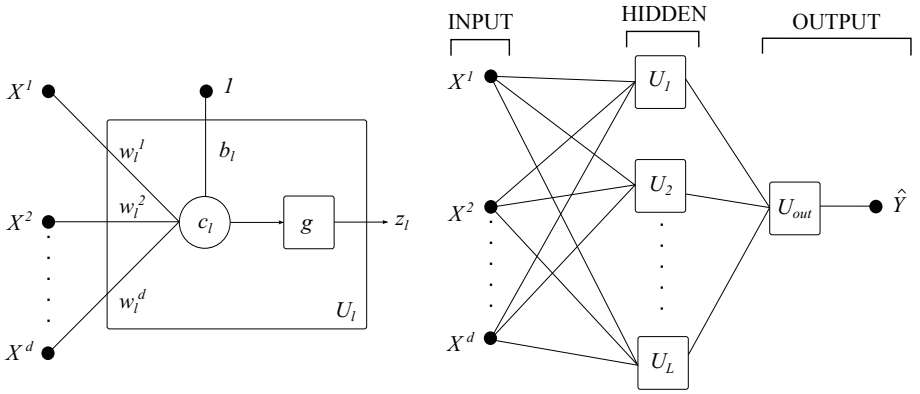


Fig. 2 Left: schematic model of a perceptron U_l. Right: Multilayer Perceptron formed by L units, $U_1 ... U_L$

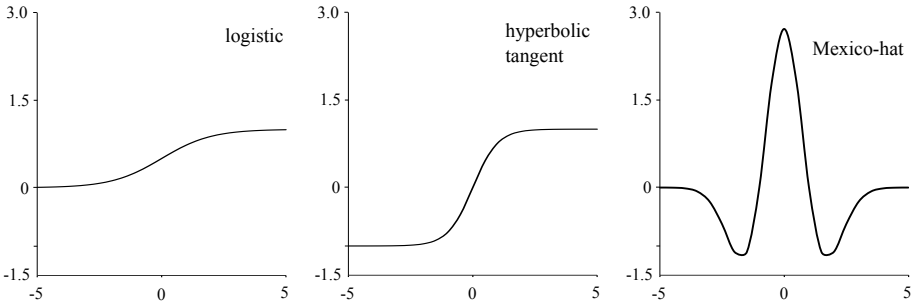


Fig. 3 Common activation functions in NN models.

a linear interval, thus an unit with small weights performs a linear transformation. On the contrary, they have horizontal asymptotes, which may cause numerical problems. While it is widely acknowledged that the variables shall be normalised before fitting an NN model, some authors restrict them to the range [0.1, 0.9] to avoid the above mentioned problems [75], [23], [56].

The most common learning algorithm is called back-propagation: NN model parameters $\{w_l^j, b_l, w_{out}^l, b_{out}\}$ are randomly initialised, and iteratively updated to minimise a cost function (typically the sum of the squared errors), by means of the gradient descent method [25].

The issues to be considered for building an NN model are the following:

1. The best network architecture, i.e., number of layers and perceptrons in each layer, is not known beforehand. Some authors focus on the definition of an efficient algorithm for determining an appropriate network architecture [64], whereas others use conventional cross-validation [40] or a simple trial and error procedure [75].
2. The training process may reach a local minimum of the error function. The probability of occurrence of this event can be reduced by introducing a learning rate parameter [75].
3. The NN models are prone to over-fitting. Various alternatives are suitable for solving this issue, such as *early stopping* and *regularisation* [25].

The fitting procedures greatly differ among authors. While Simon *et al.* [67] trained an MLP with three perceptrons in one hidden layer for 200,000 iterations, Tayfur *et al.* [75] used regularisation with 5 hidden neurons and 10,000 iterations. Neither of them followed any specific criterion to set the number of neurons. For his part, Mata [40] tested NN architectures with one hidden layer having 3 to 30 neurons on an independent test data set. He repeated the training of each NN model 5 times with different initialisation of the weights.

Kao and Loh [30] proposed a two-step procedure: first, the number of neurons was fixed whereas the optimal amount of iterations was computed. Second, NN models with different numbers of hidden nodes were trained with the selected amount of iterations, and the final architecture was chosen as the one which provided the lowest error in a validation set.

The results of the different studies are not comparable, due to the specific features of each case. Nonetheless, the lack of agreement on the training process suggests that similar results can be obtained with different criteria, provided enough care is taken to avoid over-fitting. This is in accordance with Hastie

et al. [25], who stated that in general it is enough to set the architecture and compute the appropriate regularisation parameter, or vice versa.

NN models have been used regularly in dam monitoring in recent years. There is an increasing number of published studies, both in academic and professional journals. The most recent ICOLD bulletin on dam surveillance [27] mentions NN as an alternative to HST and deterministic models, although it terms the tool as a "possible future alternative" to be developed, which suggests that it is far from being implemented in the daily practice.

2.5 Adaptive neuro-fuzzy systems (ANFIS)

Fuzzy logic allows inclusion of prior knowledge of the phenomenon, as opposed to the NN, who "learn" from the data. ANFIS models bring together the flexibility and ability to learn of the NN with the feasibility of interpretation of fuzzy logic. In fact, ANFIS can be considered a class of NN [60]. They are meant for highly non-linear, complex phenomena which vary with time [28].

Among the different types of ANFIS schemes, most previous references in dam monitoring used Takagi-Sukeno (T-S) type, whose singularity is that its output is a combination of linear functions [72]. As an exception, Opyrchal [47] used fuzzy logic to qualitatively locate seepage paths in Tresna and Dobczyce dams.

Fuzzy logic is based on the concept of membership functions (MF). Each continuous variable X^j is decomposed into K^j classes (for example, the reservoir level, which is continuous, can be transformed into "low", "medium" and "high"; see Figure 4). The particularity of fuzzy logic is that these classes have certain overlapping. Thus, a given reservoir level will generally have a different degree of membership (DOM), between zero and one, for more than one class. For Gaussian MFs:

$$DOM_{jk}\left(X^j\right) = \frac{1}{1+\left[\left(\frac{X^j-\nu_{jk}}{\lambda_{jk}}\right)^2\right]^{\mu_{jk}}} \qquad j=1,...,d; k=1,...,K^j \tag{14}$$

The number of classes for each input (K^j, which can be different among inputs X^j) are prescribed by the modeller, whereas the shape and position of their membership functions are determined by the *premise* parameters ν, λ and μ (Eq. 14), to be determined during training.

The other essential component in an ANFIS model is a set of rules, which take the form:

$$\begin{aligned} R_1 : if\ X^1 \in MF_{11} \wedge X^2 \in MF_{21} \wedge ... \wedge X^d \in MF_{d1} \Rightarrow \\ f_1 = p_{10} + p_{11}X^1 + p_{12}X^2 + ... + p_{1d}X^d \\ R_2 : if\ X^1 \in MF_{11} \wedge X^2 \in MF_{21} \wedge ... \wedge X^d \in MF_{d2} \Rightarrow \\ f_r = p_{20} + p_{21}X^1 + p_{22}X^2 + ... + p_{2d}X^d \\ ... \\ R_R : if\ X^1 \in MF_{1K^1} \wedge X^2 \in MF_{2K^2} \wedge ... \\ ... \wedge X^d \in MF_{dK^d} \Rightarrow \\ f_R = p_{R0} + p_{R1}X^1 + p_{R2}X^2 + ... + p_{Rd}X^d \end{aligned} \tag{15}$$

where $p_{10}...p_{Rd}$ are the *consequent* parameters, to be adjusted during model training. It should be noted that there can be up to $\prod_{j=1}^{d} K^j$ rules.

The model output is computed by means of 5 steps:

1. Compute the DOM of every input to each fuzzy category (14).
2. Compute the product of the correspondent DOM_{jk}, in accordance with the rules. In ANFIS terminology, these terms are referred to as the *firing strengths* ($w_r; r = 1...R$) for each rule:

$$\begin{aligned} w_1 &= DOM_{11} \cdot DOM_{21} \cdot ... \cdot DOM_{d1} \\ w_2 &= DOM_{11} \cdot DOM_{21} \cdot ... \cdot DOM_{d2} \\ &... \\ w_R &= DOM_{1K^1} \cdot DOM_{2K^2} \cdot ... \cdot DOM_{dK^d} \end{aligned} \tag{16}$$

3. Normalise the firing strengths:

$$\overline{w_r} = \frac{w_r}{\sum w_r} \tag{17}$$

4. Compute the output of each rule, as a linear function of the consequent parameters:

$$O_r = \overline{w_r} f_r = \overline{w_r}\left(p_{r0} + p_{r1}X^1 + p_{r2}X^2 + ... + p_{rd}X^d\right) \qquad r = 1...R \tag{18}$$

5. Combine the outputs of each rule to compute the overall output of the ANFIS model:

$$\hat{Y} = \sum_{r=1}^{R} O_r \tag{19}$$

The final result is a combination of linear functions of the input variables. The non-linearity is modelled in the membership functions, which are typically Gaussian, as shown in the example of Figure

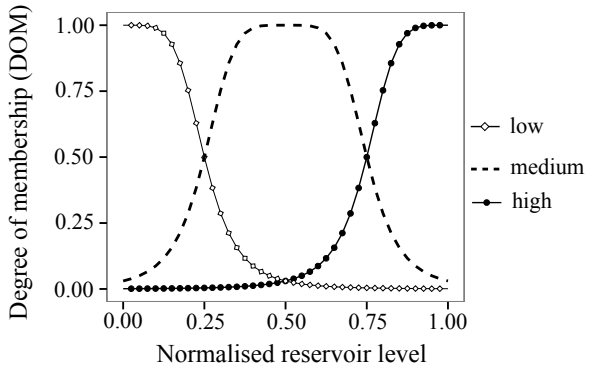


Fig. 4 Possible transformation of the normalised reservoir level into three fuzzy sets with Gaussian form: "low", "medium" and "high".

4. Each membership function is determined on the basis of 3 *premise* parameters, fitted with a hybrid method, in which the following steps are alternated:

1. The membership functions are fixed, and the consequent parameters are adjusted by least squares.
2. The premise parameters are modified by means of the gradient descent method.

The criterion of the user is more important for building ANFIS than for other kinds of models. Both the prediction accuracy and the possibility of interpreting the results may vary greatly according to the number of inputs (d), membership functions (K^j) and rules (R). It should be noted that the number of parameters in a first order T-S ANFIS model is:

$$3 \cdot \sum_{j=1}^{d} K^j + (d+1) \cdot \prod_{j=1}^{d} K^j \tag{20}$$

Rankovic *et al.* [54] prioritised prediction accuracy over model interpretation, by considering lagged values of both the input and output variables as predictors, resulting in an ANFIS model with $d = 5$, $K^j = 2$, $\forall j$ and $R = 32$. They used a zero-order T-S model, in which $f_r = p_{r0}$, $\forall r \in [1, R]$, and two-sided Gaussian membership functions, defined by 4 parameters each. No attempt was made to interpret the 32 rules.

On the contrary, Xu and Li [78] considered only 9 rules and could identify the worst environmental conditions for crack opening in the Chencun Dam.

For his part, Demirkaya [18] chose $d = 5$ and $K = 4$. Although he limited the number of rules to 4, the final model had 84 parameters.

ANFIS models can be as flexible and accurate as NN, while allowing for introducing engineering knowledge to some extent. If the amount of rules and membership functions is low, the resultant model can be interpreted. Furthermore, an ANFIS model can be used for qualitatively describing dam behaviour, especially if the output is "fuzzyfied" into linguistic variables [78].

On the contrary, they may comprise a high number of parameters, even with a few rules, which results in a high risk of over-fitting and low interpretability.

2.6 Principal component analysis (PCA) and dimensionality reduction

PCA is a well known technique in statistics. It was devised to transform a set of partially dependent variables into independent features called principal components (PCs), which are linear combination of the original variables. It is acknowledged that the first PCs contain the relevant information, whereas the less influential correspond to the signal noise. It has been used in dam monitoring for various purposes.

Mata *et al.* [42] used PCA to select the most useful thermometers to predict radial displacements in an arch dam. They pointed out the potentiality of this tool to select a group of sensors to be automatised in a given dam.

Yu *et al.* [80] applied PCA to a group of sensors to measure the opening of a longitudinal crack in an arch dam. They reported that PCA was useful for reducing the dimensionality of the problem, as well as to separate the signal from the noise. They also defined alarm thresholds as a function of the first PCs. Cheng and Zheng [11] followed a similar procedure: they analysed the covariance matrix of the outputs to separate the effect of the environmental variables from the signal noise.

Similar applications were due to Chouinard *et al.* [13], and Chouinard and Roy [12], who extracted PCs from a set of outputs (radial displacements at

pendulums) to better understand the behaviour of the structure. They focused on the model interpretation, rather than on the prediction accuracy. In this line, Nedushan [44] extracted PCs from a group of sensors to analyse them jointly, as well as to identify the correlations by means of stepwise linear regression. He defined a set of predictors (reservoir level, temperature and time), and built linear regression models by adding the most relevant one by one.

A limitation of PCA is that only linear relations between variables are considered. If the dependency is non-linear, it may lead to misinterpretation of the results. Non-linear principal component analysis (NPCA) can be an alternative, as showed by Loh *et al.* [37] and Kao and Loh [30], who applied it by means of auto-associative neural networks (AANN) to predict radial displacements in an arch dam.

AANN are a special kind of NN models, formed by 5 layers (Figure 5), which can be viewed as two NN models put in series. The intermediate (bottleneck) layer has fewer neurons than the number of model inputs, and the target outputs equal the inputs. Thus, the first part of the model reduces its dimensionality, computing some sort of non-linear PCA. The right-hand-side of the AANN is a conventional NN whose inputs are the non-linear PCs.

Jung *et al.* [29] developed a methodology to identify anomalies in piezometric readings in an earth-fill dam by means of moving PCA (MPCA), which is conventional PCA applied to different time periods. The goal was to detect significant variations in the PCs over time, which would reveal a change in dam behaviour.

PCA is mostly applied to input or output variable selection. The first option may increase the prediction accuracy, whereas the second can be useful for managing very large dams with a large amount of devices. For example, more than 8,000 instruments were installed to control the behaviour of the Three Gorges Dam [80].

2.7 Other ML techniques

There is a wide variety of ML algorithms which can be useful for dam monitoring data analysis. Their accuracy depends on the specific features of every prediction task. Given that research on ML is a highly active field, the algorithms are constantly improved and new practical applications are reported each year. Some of them have been applied to dam monitoring analysis. They are considered in this section more briefly than others, in accordance with their lower popularity in dam engineering so far. This does not mean that they can not offer advantages over the methods described previously.

Support Vector Machines (SVM) stand among the most popular ML algorithms nowadays. They combine a non-linear transformation of the predictor variables to a higher dimensional space, a linear regression on the transformed variables, and an ε-insensitive error function that neglects errors below a given threshold [68]. Cheng and Zheng [11] used SVM in combination with PCA for short-term prediction of the response of the Minhuatan gravity dam. Although the results were highly accurate, the computational time was high. Rankovic *et al.* [55] built a behaviour model based on SVM for predicting tangential displacements.

K-nearest neighbours (KNN) is a non-parametric method which requires no assumptions to be made about the physics of the problem; it is solely based on the observed data. The KNN method basically consists on estimating the value of the target variable as the weighted average of observed outputs in similar conditions within the training set. The similarity between observed values is measured as the Euclidean distance in the d-dimensional space defined by the input variables.

A clear disadvantage of this type of model is that if the Euclidean distance is used as a measure of similarity, all the predictors are given the same relevance. Hence, including a low relevant variable may result in a model with poor generalisation capability. As a consequence, variable selection is a critical aspect for fitting a KNN model.

Saouma *et al.* [65] presented a solution to the 6th ICOLD Benchmark Workshop based on KNN. To determine the similarity of observations, they used only two significant predictors (the reservoir level and a thermometer in the dam body) among the eight available. This selection of variables was performed by trial and error, although other criteria exist, as described in the next section.

Stojanovic *et al.* [69] combined greedy MLR with variable selection by means of genetic algorithms (GA). Unlike HST, they considered all the observed variables in various forms (e.g. $h, h^2, h^3, \sqrt{h}$, etc.). They defined a methodology to select the best set of predictors which could be useful to update the predictive model in case of missing variables. A similar approach was followed by Xu *et al.* [77], though with a smaller set of potential inputs.

Salazar *et al.* [61] performed a comparative study among various statistical and ML methods, including HST, NN, and others which had never been used before in dam monitoring, such as random forests (RF) or boosted regression trees (BRT). It was re-

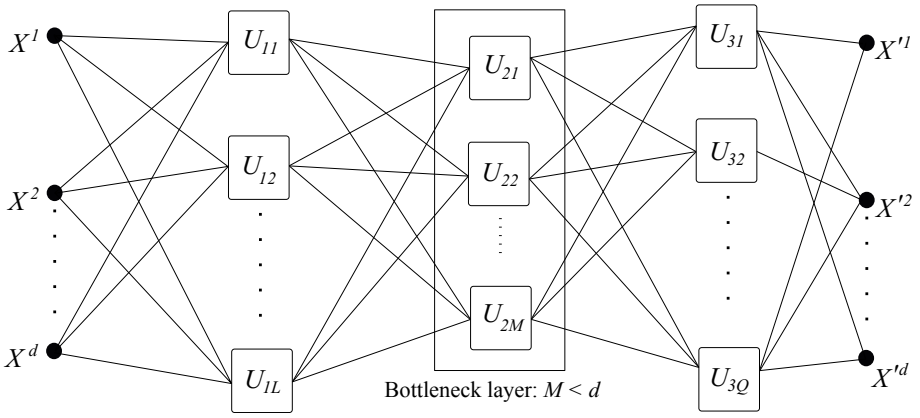


Fig. 5 Architecture of an auto-associative neural network. There are 3 hidden layers between the inputs and the output. The central one is called "bottleneck" layer, and shall have fewer nodes than model inputs, so that each one can be considered a non-linear principal component of the inputs.

ported that innovative ML algorithms offered the most accurate results, although no one performed better for all 14 outputs analysed, which corresponded to radial and tangential displacements and leakage flow in an arch dam.

3 Methodological considerations for building behaviour models

While each model has specific issues to take into account, there are also common aspects to consider when developing a prediction model, regardless of the technique. They are discussed in this section, in relation with a selection of 59 studies corresponding to 40 papers presented at conferences and in scientific journals. It is not an exhaustive review: the studies were selected on the basis of their relevance and interest, following the authors' criterion.

The Tables 2 and 3 summarise the main characteristics of the studies reviewed. It was found that most of them (38/59) considered radial displacements, especially in arch dams (31/59). This reflects the greater concern of dam engineers for this variable and dam typology, although other indicators such as leakage or uplift are acknowledged as equally relevant for dam safety [39]. The lower frequency with which the latter are chosen as target variables may be partly due to their more complex behaviour, which makes them harder to reproduce and interpret [39]. The HST and MLR methods, which have been the only ones available for a long time, are not suitable to model them [67], although some references exist [8], [24].

3.1 Input selection

In previous sections, it was pointed out that the model performance depended on the predictor variables considered. The range of options for variable selection is wide. Most of the papers reviewed do not use any specific method for variable selection, apart from user criterion (e.g. [49]) or "a priori knowledge" (e.g. [54]).

This issue has arisen in combination with the use of NN [19], [56], [30], [37], [49], NARX [52], [37], MLR [69] and ANFIS models [54].

First, the selection is limited by the available data. While the reservoir level and the temperature are usually measured at the dam site, other potentially influential variables, such as precipitation, are frequently not available. One of the advantages of the HST method is that only the reservoir level is required.

Second, it must be decided whether or not to use the lagged values of the target variable for prediction. The consequences of making predictions from the output itself have already been mentioned, regardless whether the observed or the estimated previous value is used. It can be concluded that the AR models prioritise prediction accuracy over model interpretation.

Third, the possibility of adding derived variables (and which ones), such as moving averages and gradients, can be considered. They can be set beforehand, on the basis of engineering judgement, or selected by means of some performance criterion from a wide set of variables.

Finally, consideration should be given to include non-causal variables in the model. For example, is it appropriate to base the prediction of radial displacements at a given location on the displacement

recorded at another point of the dam? Will it improve the model accuracy? What consequences would it have in the interpretation of the results?

Some models like the HST are often used with a set of specific predictors, and therefore variable selection is restricted to the order of the polynomial of the reservoir level, and the shape of the time dependent functions. The opposite case is the NARX method, which can be used with a high amount of predictor variables.

Hence, the criterion to be used depends on the type of data available, the main objective of the study (prediction or interpretation), and the characteristics of the phenomenon to be modelled. Again, engineering judgement is essential to make these decisions.

The selection of predictors may be useful to reduce the dimensionality of the problem (essential for NARX models), as well as to facilitate the interpretation of the results. PCA can be used for this purpose [42], as well as AANN [37]. Some specific methods for variable selection in dam monitoring analysis have been proposed, by means of backward elimination [**?**] genetic algorithms (GA) [69], and singular spectrum analysis (SSA) [37], although the vast majority of authors applied trial and error or engineering judgement.

3.2 Model interpretation

The main interest of this work focuses on model accuracy: a more accurate predictive model allows defining narrower thresholds, and therefore reducing the number of false anomalies. Nonetheless, once a value above (or below, if appropriate) the warning threshold is registered, an engineering analysis of the situation is needed to assess its seriousness. The ability of the model to interpret dam behaviour may be useful for this purpose.

The HST method has been traditionally used to identify the effect on the response of the dam of each considered action: hydrostatic load, temperature and time (e.g. [40]). However, it is clear that this analysis is only valid if the predictor variables are independent, which is not generally true [38], [66].

On the contrary, the ability of NN and similar models for interpreting dam behaviour is often neglected. They are frequently termed "black box" models, in reference to its lack of interpretability.

It turns out that NN models are well suited to capture complex interactions among inputs, as well as non-linear input-output relations. If an NN model offers a much better accuracy than the HST for a given phenomenon, it is probable that it does not fulfil the hypothesis of HST (input independence, linearity). Hence, it would be more appropriate to extract information on the dam behaviour from the interpretation of the NN model.

The effect of each predictor can be analysed by means of *ceteris paribus* analysis [40]: the output is computed for the range of variation of the variable under consideration, while keeping the rest at constant values. They can be set either to the correspondent mean or to several other values, in order to gain more detailed information on the dam response. Analyses of this kind can be found in the pertinent literature: Mata [40] calculated the effect of the reservoir level on the radial displacements of an arch dam for each season of the year, and the effect of temperature when setting the pool level at several constant values. Similar studies are due to Santillán *et al.* [63], Simon *et al.* [67] and Popovici *et al.* [53].

More complex algorithms have been proposed in related fields to unveil the relevance of each input in NN models (see for example [14], [22] and [46]), which may be helpful in dam monitoring.

Therefore, even though NN and similar models must be interpreted with great care, their ability to extract information on the dam behaviour should not be underestimated.

3.3 Training and validation sets

It is common and convenient to divide the available data into two subsets: the training set is used to adjust the model parameters, whereas the validation set is solely used to measure the prediction accuracy[3]. In statistics, this need is well known, since it has been proven that the prediction accuracy of a predictive model, measured on the training data, is an overestimation of its overall performance [2]. Any subsetting of the available data into training and validation sets is acceptable, provided the data are independent and identically distributed (i.i.d.). This is not the case in dam monitoring series, which are time-dependant in general.

The amount of available data is limited, what in turn limits the size of the training and validation sets. Ideally, both should cover all the range of variation of the most influential variables. This is particularly relevant for the training set of the more complex models, as they are typically unable

[3] the terminology is not universal; the data which are not used to fit the model is sometimes called test or prediction set.

to produce accurate results beyond the range of the training data [21].

It is not infrequent that reservoir level follow a relatively constant yearly cycle by which situations from the lowest to the highest pool level are presented each year. Temperature, which is the second most influential variable on average, responds to a more defined annual cycle. As a consequence, many authors measure the size of the training and validation sets in years.

Moreover, dam behaviour models are used in practice to calculate the future response, on the basis of the observed, normal functioning, and draw conclusions about the safety state. Therefore, it seems reasonable to estimate the model accuracy with a similar scheme, i.e., to take the most recent data as the validation set. This is the procedure used in the vast majority of the reviewed papers (39/40), with the unique exception of Santillán *et al.* [64], who made a random division of the data.

Models based on the underlying physics of the phenomenon and those with fewer parameters (HST, IRF and MLR), are less prone to over-fitting. As a result, a higher value can be given to the training error. This is probably the reason why most studies do not consider a validation set, but rather use all the data for the model fit e.g. [42], [7] (Figure 6 (a)).

When a validation set is used, 10% of the available data is reserved for that purpose on average. The higher frequency observed around 20% corresponds to the papers dealing with the data from the 6th ICOLD Benchmark Workshop, where the splitting criterion was fixed by the organisers.

Tayfur *et al.* [75] reserved only one year for training, but explicitly mentioned that it contained all the range of variation of the reservoir level. Some authors proposed to set a minimum of 5 to 10 observations per model parameter to estimate [71].

A fundamental premise for the successful implementation of any prediction model is that the training data correspond to a period in which the dam has not undergone significant changes in its behaviour. In practice, it is not easy to ensure that this condition is fulfilled. While the history of major repairs and events is usually available, it is well known that the behaviour in the first years of operation usually corresponds to a transient state, which may not be representative of its response in normal operation afterwards [38]. Therefore, the use of data corresponding to the first period to adjust the model parameters may lead to an increase in prediction error. Lombardi [38] estimated that 12 years from dam construction are required for a data-based model to be effective.

This issue can be checked by analysing the training error: ideally, errors shall be independent, with zero mean and constant variance [71]. Some authors compute some of these values for evaluating the goodness of fit (e.g. [67], [34], [30]).

On another note, a minimum amount of data is necessary to build a predictive model with appropriate generalisation ability. De Sortis and Paoliani [17] run a sensitivity analysis of the prediction error as a function of the training set size. They concluded that 10 years were necessary for obtaining stable results. For their part, Chouinard and Roy [12] performed a similar work on a dam set. Provided that most of them were run-of-the-river small dams, which remained full most of the time, the thermal effect was the preponderant variable. As this is almost constant every year, 5 years of data were enough for most cases to achieve high accuracy.

According to the Swiss Comittee on Dams [71], a minimum of "5 yearly cycles" should be available, which suggests that they refer to filling-emptying cycles throughout a year (to account for the thermal variation). On the contrary, ICOLD [27] recommended to set thresholds as a function of the prediction error along "2 or 3 years of normal operation".

Salazar *et al.* performed a similar analysis for 14 instruments in an arch dam [61], and reported that the prediction accuracy was higher in some cases for models trained over the most recent 5 years of data (the maximum training set length was 18 years).

The size of the validation set ranges from 1 to 25 years (Figure 6 (b)), and depends on the amount of data available, rather than on the type of model.

Such verifications regarding the training and testing data sets are not performed in general in dam monitoring analyses, probably due to a) the number of data available at a given time cannot be arbitrarily increased, and b) the validation data shall be the most recent. In practice, there is not agreement on the appropriate criterion to define training and validation sets. Consequently, the comparison between models which predict different variables has limited reliability, although it was sometimes considered [69], [56].

Again, engineering judgement is essential to assess the appropriateness of the train and validation sets, as well as the model performance.

3.4 Missing values

There are several potential sources of data incompleteness, such as insufficient measurement frequency [16], [42] or fault in the data acquisition system [41],

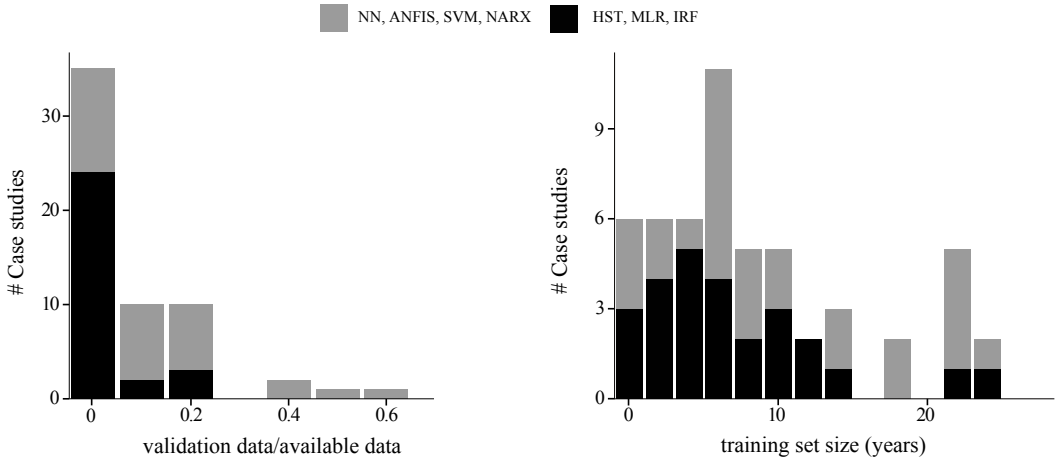


Fig. 6 Training and validation sets in the papers reviewed. Left: ratio of validation data with respect to available data. Right: training set size (years)

[69]. Although there is a tendency towards increasing the quality of measurements and the frequency of reading, there are many dams in operation with long and low-quality monitoring data series to be analysed. According to Lombardi [38], only a small minority of the world population of dams feature adequate, properly-interpreted monitoring records. Curt and Gervais [16] showed the importance of controlling the quality of the data on which the dam safety studies are based, although they focused on proposing future corrective measures rather than on how to improve imperfect time series.

However, the vast majority of published articles overlooked this issue. They limited to the selection of some specific time period for which complete data series were available. For example, Mata *et al.* [42] only considered the period 1998-2002 for their analysis of the Alto Lindoso dam, due to the absence of simultaneous readings of displacements and temperatures in subsequent periods. In general, the need for simultaneous data of both the external variables and the dam response reduces the amount of data available for model fitting and limits the prediction accuracy.

If the missing values correspond to one of the predictors, these models are inapplicable, which limits their use in practice. If lagged variables are considered, there is also a need for equally time spaced readings. The above mentioned adaptive system proposed by Stojanovic *et al.* [69] can be applied in the event of failure of one or several devices.

Faults in the data acquisition process can also result in erroneous readings [36] which should be identified and eventually discarded or corrected. During model fitting, this would improve the model accuracy and increase its ability to interpret the dam response. Once a behaviour model is built, it can be used for that purpose [11].

Numerous statistical techniques have been developed to impute missing values. Their review is beyond the scope of this work, as they were not employed in the papers analysed. Moreover, their application should be tailored to the specific features of the problem, as well as to the nature of the variable in question. For example, missing values of air temperature can be reasonably filled from the average historical temperature for the period, or interpolated from available data [64]. By contrast, daily rainfall may change largely between consecutive readings, so that one missing value cannot be imputed with similar confidence.

3.5 Prediction accuracy measurement

It is important to appropriately estimate the prediction error of a model, since a) it provides insight into its accuracy, b) it allows comparison of different models, and c) it is used to define warning thresholds.

There are various error measures to assess how well a model matches the observed data, among which the most commonly used are included in Table 1.

The result of using any of these indexes is frequently equivalent when referred to a given prediction task: the more accurate model will have a smaller $RMSE$ value, but also the lowest MSE, and higher r and R^2. However, they also present differences which can be relevant, and are often not considered.

Provided that $MSE = (RMSE)^2$, they can be used indistinctly for model comparison. The only difference is that $RMSE$ can be compared to the target variable, given that both are measured in the same

Table 1 Measures of accuracy. p = number of parameters of the model. $\bar{\cdot}$ = mean

Error metric	Formula
Mean squared error	$MSE = \frac{\sum_{i=1}^{N}(y_i - F(x_i))^2}{N}$
Root mean squared error	$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - F(x_i))^2}{N}} = \sqrt{MSE}$
Mean absolute error	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert y_i - F(x_i)\rvert$
Correlation coefficient	$r = \frac{\sum_{i=1}^{N}(y_i-\overline{y})\left(F(x_i)-\overline{F(x_i)}\right)}{\left(\sum_{i=1}^{N}(y_i-\overline{y})^2\right)^{0.5}\left(\sum_{i=1}^{N}\left(F(x_i)-\overline{F(x_i)}\right)\right)^{0.5}}$
Coefficient of determination	$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - F(x_i))^2}{\sum_{i=1}^{N}(y_i-\overline{y})^2}$
Standard error of estimate	$\sigma_\varepsilon = \sqrt{\frac{\sum (y_i - f(x_i))^2}{N}}$
Mean absolute percentage error	$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left\lvert\frac{y_i - F(x)}{y_i}\right\rvert$
Maximum absolute error	$MaxAE = \lvert\varepsilon\rvert_{max}$
Adjusted R^2	$R^2_{adj} = R^2 - (1-R^2)\frac{p}{N-p-1}$
Sum of squared error	$SSE = \sum_{i=1}^{N}(y_i - F(x_i))^2$
Average relative variance	$ARV = \frac{\sum_{i=1}^{N}(y_i - F(x_i))^2}{\sum_{i=1}^{N}(y_i-\overline{y})^2} = 1 - R^2$
Mean error	$ME = \frac{1}{N}\sum_{i=1}^{N}(y_i - F(x_i))$

units. It should be noted that they are computed on the basis of the squared residuals, therefore they are sensitive to the presence of outliers, i.e., a few large prediction errors. In this sense, MAE could be considered a better choice, provided that it shares the advantage of $RMSE$ (it is measured in the same units as the output), and not its drawback. Mindful of this fact, both can be used interchangeably, if the analysis is complemented with a graphical exploration of the model fit, or other error measures.

The drawback to both MSE and $RMSE$ is that they are not suitable for comparing models fitting different variables, provided that they do not consider neither the mean nor the deviation of the output.

This limitation can be overcome by using the correlation coefficient r, since $r \in [-1, 1]$. On the contrary, it is not exactly an error rate, but rather an index of the strength of the linear relationship between observations and predictions. In other words, it indicates to what extent one variable increases as the other does, and vice versa. It can be checked that the value of r for a prediction calculated as $\hat{Y} = AY + B$ is equal to 1 for $A \neq 0$, while the error can be very large and will generally be non-zero (unless $A = 1$ and $B = 0$) [32]. As an example, Rankovic *et al.* [56] considered r and r^2, as well as MAE and MSE. While the results were similar for the training and validation sets in terms of r and r^2, both MAE and MSE were much greater in the validation set (as much as 7 times greater). These results may reflect some degree of over-fitting.

If r is used as a measure of goodness of fit, its value always increases with increasing number of model parameters (except in the highly unlikely event that the functions are completely independent of output). The R_{adj} coefficient can be used (e.g. [69], [34]) to account for the number of parameters of each model.

As an alternative, R^2, or its equivalent ARV can be chosen. They have the advantage over the correlation coefficient of being sensitive to differences in the means and variances of observations and predictions, while maintaining the ability to compare models fitted to different data [61].

Finally, it should be noted that the reading error of the devices (ε_r) may be relevant when predictions of variables of different nature are compared, although it is often ignored. It cannot be expected to obtain a model with an error below the measurement resolution [80]. Popovici *et al.* [53] reported that the overall accuracy of NN models was lower for tangential than for radial displacements, and attributed it to the lower range of variation of the for-

mer. It is possible that the reading error (which in principle should be the same for tangential and radial displacements) were relevant in the first case and negligible in the second.

Salazar *et al.* found that models with relatively high ARV corresponded with very low MAE, close to ε_r [61].

Reading error should always be considered for evaluating model accuracy. One possibility would be to neglect the errors below that value before computing the prediction accuracy, by means of substituting $(y_i - F(x_i))$ by $|y_i - F(x_i)| - \varepsilon_r$, in the calculation of MSE, $RMSE$, r and R^2. Similarly, MAE could be computed as:

$$MAE^* = \frac{1}{N}\sum_{i=1}^{N}\left(|y_i - F(x_i)| - \varepsilon_r\right) \tag{21}$$

It is convenient to compute more than one error rate, especially if the aim is to compare models predicting variables of different kind. In addition, a graphical analysis of the error is highly advisable.

3.6 Practical application

Despite the increasing amount of literature on the use of advanced data-based tools, very few examples described their practical integration in dam safety analysis. The vast majority were limited to the model accuracy assessment, by quantifying the model error with respect to the actual measured data. Only a few cases dealt with the interpretation of dam behaviour, by identifying the effect of each of the external variables on the dam response (e.g. [40], [17], [35]).

A detailed analysis of the results is always convenient [26], especially when complex models are employed. However, improvements in instrumentation and data acquisition systems allow the implementation of automatic warning generation schemes. The information provided by reliable automated systems, based on highly accurate models, can be a great support for decision making regarding dam safety [27], [31].

To achieve that goal, the outcome of the predictive model must be transformed into a set of rules that determine whether the system should issue a warning. In turn, these rules should be based on an overall analysis of the most representative instruments: a single value out of the normal-operation range will probably correspond to a reading error, if other instruments show no anomalies. However, the coincidence of out-of-range values in several devices may correspond to some abnormal behaviour. This is the idea behind the method proposed by Cheng and Zheng [11], which features a procedure for calculating normal operating thresholds ("control limits"), and a qualitative classification of potential anomalies: a) extreme environmental variable values, b) global structure damage, c) instrument malfunctions and d) local structure damage.

A more accurate analysis could be based on the consideration of the major potential modes of failure to obtain the corresponding behaviour patterns and an estimate of how they would be reflected on the monitoring data. Mata *et al.* [43] employed this idea to develop a methodology that includes the following steps:

- Identification of the most probable failure mode.
- Simulation of the structural response of the dam in normal and accidental situations (failure) by means of finite element models.
- Selection of the set of instruments that better identify the dam response during failure.
- Construction of a classification rule based on linear discriminant analysis (LDA) that labels a set of monitoring data as normal behaviour or incipient failure.

This scheme can be easily implemented in an automatic system. By contrast, it requires a detailed analysis of the possible failure modes, and their numerical simulation to provide data with which to train the classifier. Moreover, the finite element model must be able to accurately represent the actual behaviour of the dam, which is frequently hard to achieve.

4 Conclusions

There is a growing interest in the application of innovative tools in dam monitoring data analysis. Although only HST is fully implemented in engineering practice, the number of publications on the application of other methods has increased considerably in recent years, specially NN.

It seems clear that the models based on ML algorithms can offer more accurate estimates of the dam behaviour than the HST method in many cases. In general, they are more suitable to reproduce nonlinear effects and complex interactions between input variables and dam response.

However, most of the papers analysed refer to specific case studies, certain dam typologies or determined outputs. More than a half of them focus on radial displacements in arch dams, although this

typology represents roughly 5% of dams in operation worldwide.

Moreover, the vast majority of articles overlooked the data pre-process. It is implicitly assumed that the monitoring data are free of reading errors and missing values, whereas that is not the case in practice. The development of criteria to fix imperfect data would allow to take advantage of a large amount of stored dam monitoring data.

An useful data-based algorithm should be versatile to face the variety of situations presented in dam safety: different typologies, outputs, quality and volume of data available, etc. Data-based techniques should be capable of dealing with missing values and robust to reading errors.

These tools must be employed rigorously, given their relatively high number of parameters and flexibility, what makes them susceptible to over-fit the training data. It is thus essential to check their generalisation capability on an adequate validation data set, not used for fitting the model parameters.

In this sense, most of the studies reviewed did not include an evaluation of the predictive model on an independent data set, and there are very few examples that used more than 20% of the data for validation. This raises doubts about the generalisation capability of these models, in particular of those more strictly data-based, such as NN or SVM. It should be reminded that the main limitation of these methods is their inability to extrapolate, i.e., to generate accurate predictions outside the range of variation of the training data.

Before applying these models for predicting the dam response in a given situation, it should be checked whether the load combination under consideration lies within the values of the input variables in the training data set. Verifications of this kind were not reported in the reviewed papers, although they would provide insight into the reliability of the predictions.

From a practical viewpoint, data-based models should also be user-friendly and easily understood by civil engineering practitioners, typically unfamiliar with computer science, who have the responsibility for decision making.

Finally, two overall conclusions can be drawn from the review:

- ML techniques can be highly valuable for dam safety analysis, though some issues remain unsolved.
- Regardless of the technique used, engineering judgement based on experience is critical for building the model, for interpreting the results, and for decision making with regard to dam safety.

References

1. Amberg F (2009) Interpretative models for concrete dam displacements. In: 23th ICOLD Congress, q91-R43
2. Arlot S, Celisse A, et al (2010) A survey of cross-validation procedures for model selection. Statistics surveys 4:40–79
3. Bishop CM (1995) Neural networks for pattern recognition. Oxford university press
4. Bonelli S, Félix H (2001) Delayed response analysis of temperature effect. In: Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams, Salzburg, Austria
5. Bonelli S, Radzicki K (2007) The impulse response function analysis of pore pressures monitoring data. In: 5th International Conference on Dam Engineering, Lisbon, Portugal
6. Bonelli S, Radzicki K (2008) Impulse response function analysis of pore pressure in earthdams. European Journal of Environmental and Civil Engineering 12(3):243–262
7. Bonelli S, Royet P (2001) Delayed response analysis of dam monitoring data. In: Proceedings of the Fifth ICOLD European Symposium on Dams in a European Context, Geiranger, Norway
8. Breitenstein F, Klher W, Widman R (1985) Safety control of the dams of the Glockner-Kaprun hydro-electric development. In: 15th ICOLD Congress, pp 1121–1134, q56-R59
9. Carrère A, Noret-Duchêne C (2001) Interpretation of an arch dam behaviour using enhanced statistical models. In: Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams, Salzburg, Austria
10. Chen BJ, Chang MW, et al (2004) Load forecasting using support vector machines: A study on EUNITE competition 2001. Power Systems, IEEE Transactions on 19(4):1821–1830
11. Cheng L, Zheng D (2013) Two online dam safety monitoring models based on the process of extracting environmental effect. Advances in Engineering Software 57:4856
12. Chouinard L, Roy V (2006) Performance of statistical models for dam monitoring data. In: Joint International Conference on Computing and Decision Making in Civil and Building Engineering, Montreal, pp 14–16
13. Chouinard L, Bennett D, Feknous N (1995) Statistical analysis of monitoring data for concrete arch dams. Journal of Performance of Constructed Facilities 9(4):286–301

14. Cortez P, Embrechts MJ (2011) Opening black box data mining models using sensitivity analysis. In: Computational Intelligence and Data Mining (CIDM), 2011 IEEE Symposium on, IEEE, pp 341–348
15. Crépon O, Lino M (1999) An analytical approach to monitoring. Water Power and Dam Construction
16. Curt C, Gervais R (2014) Approach to improving the quality of data used to analyse dams–illustrations by two methods. European Journal of Environmental and Civil Engineering 18(1):87–105
17. De Sortis A, Paoliani P (2007) Statistical analysis and structural identification in concrete dam monitoring. Engineering structures 29(1):110–120
18. Demirkaya S (2010) Deformation analysis of an arch dam using ANFIS. In: Proceedings of the second international workshop on application of artificial intelligence and innovations in engineering geodesy. Braunschweig, Germany, p 2131
19. Demirkaya S, Balcilar M (2012) The contribution of soft computing techniques for the interpretation of dam deformation. In: Proceedings of the FIG working week, Rome, Italy
20. Fabre J, Geffraye G (2013) Study and control of thermal displacements of Gage II dam (France) through the contribution of special heating and cooling devices. In: Proceedings of the Seventh Argentinian Conference on Dams, San Juan, Argentina, [in Spanish]
21. Flood I, Kartam N (1994) Neural networks in civil engineering. i: Principles and understanding. Journal of computing in civil engineering 8(2):131–148
22. Gevrey M, Dimopoulos I, Lek S (2003) Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecological Modelling 160(3):249–264
23. Govindaraju RS (2000) Artificial neural networks in hydrology II: hydrologic applications. Journal of Hydrologic Engineering 5(2):124–137
24. Guedes Q, Coelho P (1985) Statistical behaviour model of dams. In: 15th ICOLD Congress, pp 319–334, q56-R16
25. Hastie T, Tibshirani R, Firedman J (2009) The Elements of Statistical Learning - Data Mining, Inference, and Prediction, Second Edition, 2nd edn. Springer
26. Hill C, Sundaram M (2013) Instrumentation data collection, management and analysis. Tech. rep., United States Society on Dams (USSD) Committee on Monitoring of Dams and Their Foundations
27. International Commission on Large Dams (2012) Dam surveillance guide. Tech. Rep. B-158, ICOLD
28. Jang JS (1993) ANFIS: adaptive-network-based fuzzy inference system. Systems, Man and Cybernetics, IEEE Transactions on 23(3):665–685
29. Jung IS, Berges M, Garrett JH, Kelly CJ (2013) Interpreting the dynamics of embankment dams through a time-series analysis of piezometer data using a non-parametric spectral estimation method. In: Computing in Civil Engineering (2013), ASCE, p 2532, URL `http://ascelibrary.org/doi/abs/10.1061/9780784413029.004`
30. Kao C, Loh C (2013) Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural network-based approaches. Structural Control and Health Monitoring 20(3):282–303
31. on Large Dams IC (2000) Automated dam monitoring systems. guidelines and case histories. Tech. Rep. B-118, ICOLD
32. Legates DR, McCabe GJ (1999) Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water resources research 35(1):233–241
33. Léger P, Leclerc M (2007) Hydrostatic, temperature, time-displacement model for concrete dams. Journal of Engineering Mechanics 133(3):267277
34. Li F, Wang Z, Liu G (2013) Towards an error correction model for dam monitoring data analysis based on cointegration theory. Structural Safety 43:1220
35. Li F, Wang Z, Liu G, Fu C, Wang J (2014) Hydrostatic seasonal state model for monitoring data analysis of concrete dams. Structure and Infrastructure Engineering (ahead-of-print):1–16
36. Ljunggren M, Logan T, Campbell P (2013) Is your dam as safe as your data suggest? In: Proceedings of the NZSOLD/ANCOLD Conference, Rotorua, New Zealand
37. Loh CH, Chen CH, Hsu TY (2011) Application of advanced statistical methods for extracting long-term trends in static monitoring data from an arch dam. Structural Health Monitoring 10(6):587–601
38. Lombardi G (2004) Advanced data interpretation for diagnosis of concrete dams. Tech. rep., CISM

39. Lombardi G, Amberg F, Darbre G (2008) Algorithm for he prediction of functional delays in the behaviour of concrete dams. Hydropower and Dams (3):111–116
40. Mata J (2011) Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. Engineering Structures 3(3):03 – 910, DOI 10.1016/j.engstruct.2010.12.011
41. Mata J, de Castro AT, da Costa JS (2013) Time–frequency analysis for concrete dam safety control: Correlation between the daily variation of structural response and air temperature. Engineering Structures 48:658–665
42. Mata J, Tavares de Castro A, Sá da Costa J (2014) Constructing statistical models for arch dam deformation. Structural Control Health Monitoring 21(3):423–437, DOI 10.1002/stc.1575
43. Mata J, Leitão NS, de Castro AT, da Costa JS (2014) Construction of decision rules for early detection of a developing concrete arch dam failure scenario. a discriminant approach. Computers & Structures 142:45–53
44. Nedushan B (2002) Multivariable statistical analysis of monitoring data for concrete dams. PhD thesis, McGill University
45. Nourani V, Babakhani A (2012) Integration of artificial neural networks with radial basis function interpolation in earthfill dam seepage modeling. Journal of Computing in Civil Engineering 27(2):183–195
46. Olden JD, Jackson DA (2002) Illuminating the black box: a randomization approach for understanding variable contributions in artificial neural networks. Ecological modelling 154(1):135–150
47. Opyrchal L (2003) Application of fuzzy sets method to identify seepage path through dams. Journal of Hydraulic Engineering 129(7):546–548
48. Palumbo P, Piroddi L, Lancini S, Lozza F (2001) NARX modeling of radial crest displacements of the Schlegeis arch dam. In: Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams, Salzburg, Austria
49. Panizzo A, Petaccia A (2009) Analysis of monitoring data for the safety control of dams using neural networks. In: New Trends in Fluid Mechanics Research, Springer, p 344347
50. Penot I, Daumas B, Fabre J (2005) Monitoring behaviour. Water Power and Dam Construction
51. Perner F, Obernhuber P (2010) Analysis of arch dam deformations. Frontiers of Architecture and Civil Engineering in China 4(1):102–108
52. Piroddi L, Spinelli W (2003) Long-range nonlinear prediction: a case study. In: Decision and Control, 2003. Proceedings. 42nd IEEE Conference on, IEEE, vol 4, pp 3984–3989
53. Popovici A, Ilinca C, Ayvaz T (2013) The performance of the neural networks to model some response parameters of a buttress dam to environment actions. In: Proceedings of the 9th ICOLD European Club Symposium, Venice, Italy
54. Ranković V, Grujović N, Divac D, Milivojević N, Novaković A (2012) Modelling of dam behaviour based on neuro-fuzzy identification. Engineering Structures 35:107113, DOI 10.1016/j.engstruct.2011.11.011
55. Ranković V, Grujović N, Divac D, Milivojević N (2014) Development of support vector regression identification model for prediction of dam structural behaviour. Structural Safety 48:33–39
56. Ranković V, Novaković A, Grujović N, Divac D, Milivojević N (2014) Predicting piezometric water level in dams via artificial neural networks. Neural Computing and Applications 24(5):1115–1121
57. Restelli F (2010) Systemic evaluation of dam monitoring using PCA. In: Proceedings of the Sixth Argentinian Conference on Dams, Neuquén, Argentina, [in Spanish]
58. Restelli F (2013) Systemic evaluation of the response of large dams instrumentation. Application at El Chocón dam. In: Proceedings of the 9th ICOLD European Club Symposium, Venice, Italy
59. Riquelme F, Fraile J, Santillán D, Morán R, Toledo M (2011) Application of artificial neural network models to determine movements in an arch dam. In: Proceedings of the 2nd International Congress on Dam Maintenance and Rehabilitation, Zaragoza, Spain, pp 117–123
60. Ruiz H (2013) Fisher networks: a principled approach to retrieval-based classification. PhD thesis, Liverpool John Moores University
61. Salazar F, Toledo M, Oñate E, Morán R (2015) An empirical comparison of machine learning techniques for dam behaviour modelling. Structural Safety 56(0):9 – 17, DOI http://dx.doi.org/10.1016/j.strusafe.2015.05.001
62. Sánchez Caro FJ (2007) Dam safety: contributions to the deformation analysis and monitoring as an element of prevention of pathologies of geotechnical origin. PhD thesis, UPM, [In Spanish]
63. Santillán D, Fraile-Ardanuy J, Toledo M (2013) Dam seepage analysis based on artificial neural

networks: The hysteresis phenomenon. In: Neural Networks (IJCNN), The 2013 International Joint Conference on, IEEE, pp 1–8
64. Santillán D, Fraile-Ardanuy J, Toledo M (2014) Seepage prediction in arch dams by means of artificial neural networks. Water Technology and Science V(3), [in Spanish]
65. Saouma V, Hansen E, Rajagopalan B (2001) Statistical and 3d nonlinear finite element analysis of Schlegeis dam. In: Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams, pp 17–19
66. Silva Gomes AF, Silva Matos D (1985) Quantitative analysis of dam monitoring results. State of the art, applications and prospects. In: 15th ICOLD Congres, pp 319–334, q56-R39
67. Simon A, Royer M, Mauris F, Fabre J (2013) Analysis and interpretation of dam measurements using artificial neural networks. In: Proceedings of the 9th ICOLD European Club Symposium, Venice, Italy
68. Smola AJ, Schlkopf B (2004) A tutorial on support vector regression. Statistics and computing 14(3):199222
69. Stojanovic B, Milivojevic M, Ivanovic M, Milivojevic N, Divac D (2013) Adaptive system for dam behavior modeling based on linear regression and genetic algorithms. Advances in Engineering Software 65:182190
70. Su H, Wu Z, Wen Z (2007) Identification model for dam behavior based on wavelet network. Computer-Aided Civil and Infrastructure Engineering 22(6):438–448
71. Swiss Committee on Dams (2003) Methods of analysis for the prediction and the verification of dam behaviour. Tech. rep., ICOLD
72. Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. Systems, Man and Cybernetics, IEEE Transactions on (1):116–132
73. Tatin M, Briffaut M, Dufour F, Simon A, Fabre J (2013) Thermal displacements of concrete dams: Finite element and statistical modelling. In: Proceedings of the 9th ICOLD European Club Symposium, Venice, Italy
74. Tatin M, Briffaut M, Dufour F, Simon A, Fabre JP (2015) Thermal displacements of concrete dams: Accounting for water temperature in statistical models. Engineering Structures 91:26–39
75. Tayfur G, Swiatek D, Wita A, Singh VP (2005) Case study: Finite element method and artificial neural network models for flow through Jeziorsko earthfill dam in Poland. Journal of Hydraulic Engineering 131(6):431440
76. Willm G, Beaujoint N (1967) Les méthodes de surveillance des barrages au service de la production hydraulique d'Electricité de France-Problèmes ancients et solutions nouvelles. In: 9th ICOLD Congres, pp 529–550, q34-R30. [in French]
77. Xu C, Yue D, Deng C (2012) Hybrid GA/SIMPLS as alternative regression model in dam deformation analysis. Engineering Applications of Artificial Intelligence 25(3):468475
78. Xu H, Li X (2012) Inferring rules for adverse load combinations to crack in concrete dam from monitoring data using adaptive neuro-fuzzy inference system. Science China Technological Sciences 55(1):136141
79. Yao Y, Sharma A, Golubchik L, Govindan R (2010) Online anomaly detection for sensor systems: A simple and efficient approach. Performance Evaluation 67(11):1059–1075
80. Yu H, Wu Z, Bao T, Zhang L (2010) Multivariate analysis in dam monitoring data with PCA. Science China Technological Sciences 53(4):1088–1097, DOI 10.1007/s11431-010-0060-1
81. Zenz G, Obernhuber P (2001) ICOLD benchmark workshops on dam safety. Hydropower and Dams (2):75–78

Table 2: Review summary. Case studies

Id	Author	Year	Dam(s)	Country	Typology	Output	# Outputs
1	Breitenstein [8]	1985	Limberg, Mooser, Drossen	Switzerland	ARC, GRA, ARC	RAD	7
2	Breitenstein [8]	1985	Limberg, Mooser, Drossen	Switzerland	ARC, GRA, ARC	LEAK	6
3	Guedes [24]	1985	São Simão	Brazil	EF+GRA	RAD	1
4	Guedes [24]	1985	Água Vermelha	Brazil	EF+GRA	RAD	1
5	Guedes [24]	1985	Funil	Brazil	ARC	PIEZ	1
6	Guedes [24]	1985	Sobradinho	Brazil	EF+GRA	JOINT	1
7	Guedes [24]	1985	Itaipú	Brazil	GRA	LEAK	1
8	Bonelli [7]	2001	Alzitone, Chamboux, La Verne	France	EF	PIEZ	9,6,4
9	Bonelli [4]	2001	Schelegeis	Austria	ARC	RAD	1
10	Carrere [9]	2001	Schelegeis	Austria	ARC	RAD	1
11	Saouma [65]	2001	Schelegeis	Austria	ARC	RAD	1
12	Palumbo [48]	2001	Shclegeis	Austria	ARC	RAD	1
13	Nedushan [44]	2002	Chute-à-Caron	Canada	GRA	RAD, TAN, VERT	1,1,1
14	Piroddi [52]	2003	Schelegeis	Austria	ARC	RAD	1
15	Tayfur [75]	2005	Jeziorsko	Poland	CFRD	PIEZ	4
16	De Sortis [17]	2006	Ancipa	Italy	BUT	RAD	5
17	De Sortis [17]	2006	Sabbione	Italy	BUT	RAD	3
18	De Sortis [17]	2006	Malga Bissina	Italy	BUT	RAD	5
19	S. Caro [62]	2007	El Atazar	Spain	ARC	RAD	46
20	Léger [33]	2007	Schelegeis	Austria	ARC	RAD	1
21	Su [70]	2007	?	China	AG	VERT	1
22	Panizzo [49]	2007	Pieve di Cadore	Italy	AG	RAD	1
23	Lombardi [39]	2008	?	?	ARC	RAD	1
24	Lombardi [39]	2008	?	?	ARC	LEAK	1
25	Bonelli [5]	2007	?	?	EF	PIEZ	14
26	Bonelli [6]	2008	?	?	EF	PIEZ	16
27	Yu [80]	2010	Chencun	China	AG	CRACK	5
28	Perner [51]	2010	Zillergruendl	Austria	ARC	RAD	2
29	Demirkaya [18]	2010	Schelegeis	Austria	ARC	RAD	1
30	Riquelme [59]	2011	La Baells	Spain	ARC	RAD	1
31	Mata [40]	2011	Alto Rabagão	Portugal	ARC	RAD	1
32	Rankoviĉ [54]	2012	Bocac	Bosnia Herzegovina	ARC	RAD	2
33	Xu [77]	2012	Chencun	China	AG	CRACK	1
34	Demirkaya [19]	2012	Schelegeis	Austria	ARC	RAD	1
35	Demirkaya [19]	2012	Schelegeis	Austria	ARC	RAD	1
36	Cheng [11]	2013	Mianhuatan	China	GRA	RAD	12
37	Cheng [11]	2013	Mianhuatan	China	GRA	UP	16
38	Popovici [53]	2013	Gura Râului	Romania	BUT	RAD, TAN, ROCK	2, 2, 3
39	Tatin [73]	2013	Castelnau	France	GRA	RAD	1
40	Tatin [73]	2013	Castelnau	France	GRA	RAD	1
41	Li [34]	2013	Wanfu	China	ARC	RAD	4
42	Li [34]	2013	Wanfu	China	ARC	RAD	4
43	Simon [67]	2013	Pareloup	France	ARC	PIEZ	1
44	Simon [67]	2013	Bissorte	France	GRA	LEAK	4
45	Simon [67]	2013	Monteynard	France	ARC	RAD	1
46	Simon [67]	2013	Monteynard	France	ARC	RAD	1
47	Nourani [45]	2013	Sahand	Iran	EF	PIEZ	4
48	Kao [30]	2013	Fei-Tsui	Taiwan	ARC	RAD	13
49	Kao [30]	2013	Fei-Tsui	Taiwan	ARC	RAD	13
50	Kao [30]	2013	Fei-Tsui	Taiwan	ARC	RAD	13
51	Mata [42]	2013	Alto Lindoso	Portugal	ARC	RAD	5
52	Jung [29]	2013	?	USA	EF	PIEZ	1
53	Stojanovic [69]	2013	Bocac	Bosnia Herzegovina	ARC	RAD	1
54	Rankoviĉ [56]	2014	Iron Gate 2	Serbia/ Romania	EF+GRA	PIEZ	2
55	Rankoviĉ [56]	2014	Iron Gate 2	Serbia/ Romania	EF+GRA	PIEZ	2
56	Santillán [64]	2014	La Baells	Spain	ARC	LEAK	1
57	Salazar [61]	2014	La Baells	Spain	ARC	RAD, TAN, LEAK	5, 5, 4
58	Rankoviĉ [55]	2014	Iron Gate 2	Serbia/ Romania	EF+GRA	TAN	2
59	Tatin [74]	2015	Eguzon, Izourt, Roselend, Tignes, Vouglans, Bissorte, Gittaz, Sarrans	France	GRA,ARC, BUT	RAD	1

Continued on next page

Table 2 – continued from previous page

Id	Author	Year	Dam	Country	Typology	Output	# Outputs

Typology: ARC = arch; GRA = gravity; EF = earth-fill; AG = arch-gravity; BUT = buttress; CFRD = concrete-faced rockfill dam; ? = Not specified. Outputs: RAD= radial displacements; LEAK = leakage flow; PIEZ = pore pressure; JOINT = joint opening; TAN = tangential displacements; VERT = vertical displacements; UP = uplift pressure; ROCK = rockmeter displacements; CRACK = crack opening.

Table 3: Review summary. Methods

Id	Model	Inputs	Training set (years/ # samples)	Validation set (years/ # samples)	%Validation data	Error metric (see Table 1)
1	MLR	$h, S, t, T_{air}, \partial(T_{air}), \partial(h)$	10/14600	0.0/	0%	R^2
2	MLR	h	10/14600	0.0/	0%	R^2
3	MLR	$h, mav(T_{air})$	1/103	0.0/	0%	r
4	MLR	$h, mav(T_{air})$	0.5/63	0.0/	0%	r
5	MLR	$t, mav(h)$	2/230	0.0/	0%	r
6	MLR	$t, mav(T_c)$	2.5/66	0.0/	0%	r
7	MLR	$t, mav(h)$	0.5/86	0.0/	0%	r
8	IRF	$h, lag(h), P, lag(P), t$	Var/Var	0.0/	0%	-
9	IRF	h, T_{air}	7/2557	2.0/730	22%	-
10	HST	h, T_{air}	7/2557	2.0/730	22%	r, R^2, σ_ϵ
11	KNN	h, T_c	7/2557	2.0/730	22%	r, R^2, σ_ϵ
12	NARX	$h, T_{air}, T_c, lag(h), lag(T_c), lag(T_{air})$	7/2555	2.0/730	22%	$RMSE$
13	NN	T_c, t	1.5/548	1.5/548	50%	R^2
14	NARX	$h, T_{air}, T_c, lag(h), lag(T_c), lag(T_{air})$	7/2555	2.0/730	22%	MSE
15	NN	h	1/26	2.0/52	67%	$RMSE$, MAE, R^2
16	HST	h, S, t	2 to 15/730 to 5475	0.0/	0%	r, σ_ε, $\frac{\sigma_\varepsilon}{D/2}$
17	HST	h, S, t	5/1825	0.0/	0%	r, σ_ε, $\frac{\sigma_\varepsilon}{D/2}$
18	HST	h, S, t	9/3285	0.0/	0%	r, σ_ε, $\frac{\sigma_\varepsilon}{D/2}$
19	MLR	$h, mav(h), S, T_{air}, mav(T_{air})$	24.5/8943	0.0/	0%	σ_ε, MSE
20	HTT	h, T_c, t	5/1825	0.0/	0%	r
21	WNN	h, S, t	11/44	2.0/8	15%	MAE
22	NN	$h, lag(rad), T_{air}, T_c$	7/2555	0.0/	0%	R^2, $pdf(\varepsilon)$, MSE
23	IRF	$h, lag(h), lag(rad), T_{air}$	4/?	0.0/	0%	σ_ε
24	IRF	$h, lag(h), lag(out)$	5/?	0.0/	0%	-
25	IRF	$h, lag(P)$	3/167	0.0/	0%	R^2
26	IRF	$h, hd, lag(P)$	var/var	0.0/	0%	R^2
27	HST	h, S, t	10/1200	0.0/	0%	r
28	HYB	h, T_c, t	22/8030	0.0/	0%	-
29	ANFIS	h, T_{air}, T_c	6/2044	1.0/365	15%	r, $RMSE$, MAE
30	NN	$h, T, mav(T), lag(out)$	18/706	12.0/470	40%	$MAPE$
31	NN	h, S	23/914	1.8/69	7%	MAE, $MaxAE$, r
32	ANFIS	$lag(h), lag(S), lag(out)$	9/657	2.0/140	18%	r, MAE, $RMSE$
33	ANFIS	T_{air}, h	15/400	?/?	0%	$RMSE$
34	MLR	$h, T_{air}, T_c, lag(h), lag(T_{air}), lag(T_c)$	7/2555	2.0/730	22%	ME, σ_ε, R^2
35	NN	$h, T_{air}, T_c, lag(h), lag(T_{air}), lag(T_c)$	7/2555	2.0/730	22%	ME, σ_ε, R^2
36	PCA, SVM	h, T_{air}, P	4.2/1525	0.1/30	2%	-
37	PCA, SVM	h, T_{air}, P	3/900	0.2/56	6%	-
38	NN	$t, h, T_{air}, lag(h), mav(T_{air})$	14/?	2.0/?	13%	r, R^2, σ_ε
39	GRAD	$h, S, t, IRF, T_{air}, T_w$	12/?	0.0/	0%	σ_ε
40	SLICE	$h, S, t, IRF, T_{air}, T_w$	12/?	0.0/	0%	σ_ε
41	HTT	h, S, T_c	3.2/169	0.4/20	11%	R^2_{adj}, σ_ϵ, $pdf(\varepsilon)$
42	ECM	$h, S, T_c, \varepsilon(t-1)$	3.2/169	0.4/20	11%	R^2_{adj}, σ_ϵ, $pdf(\varepsilon)$
43	NN	h, S, t	?/429	0.0/	0%	σ_ε, MSE
44	IRF+NN	$h, S, t, T_{air}, IRF(P), IRF(M)$	?/?	0.0/	0%	R^2
45	NN	$h, S, t, IRF(T_{air})$	?/?	0.0/	0%	σ_ε
46	HSTT	$h, S, t, IRF(T_{air})$	?/?	0.0/	0%	σ_ε
47	NN	$h, h_d, lag(P)$	1.1/58	0.4/18	24%	R^2
48	NN	h, T_c	22/8120	0.3/62	1%	R^2, $pdf(\varepsilon)$, MSE
49	NARXNN	$h, lag(h), lag(out)$	22/8120	0.3/62	1%	R^2, $pdf(\varepsilon)$, MSE
50	AANN	$lag(rad)$	22/8120	0.3/62	1%	R^2, $pdf(\varepsilon)$, MSE
51	HTT	h, T_c	5/95	0.0/	0%	R^2_{adj}, σ_ε, ε_{max}, ε_{min}, SSE
52	MPCA	h	6/4380	0.0/	0%	-
53	MLR	h, T_c, T_{air}, P, t	6/2550	1.0/365	13%	R^2_{adj}, $RMSE$
54	NN	$h_d, lag(h_d)$	8/163	1.0/20	11%	r, r^2, MSE, MAE
55	MLR	$h_d, lag(h_d)$	8/163	1.0/20	11%	r, MSE, MAE
56	NN	$h, T_{air}, \partial(h), \partial(T_{air})$	25.5/918	3.0/103	10%	$RMSE$
57	NN, MARS, RF, BRT, SVM	$h, T_{air}, S, t, mav(h), mav(T_{air}), P, \partial(h)$	18/600	10.0/400	40%	MAE, ARV
58	SVM	$h, h_d, lag(h), lag(h_d), lag(out)$	11/573	3.0/156	21%	r, MAE, MSE

Continued on next page

Table 3 – continued from previous page

Id	Model	Inputs	Training set (years/ # samples)	Validation set (years/ # samples)	%Validation data	Error metric (see Table 1)
59	GRAD	$h, S, t, IRF, T_{air}, T_w$	8/?	2/?	20%	σ_ε

Models: MLR = multilineal regression; IRF = impulse response function; HST = hydrostatic seasonal time; KNN = k-nearest neighbours; NN = neural networks; WNN = wavelet neural networks; NARX = non-linear autoregressive exogenous; HTT = hydrostatic thermal time; HYB = hybrid; ANFIS = adaptive neuro-fuzzy system; PCA = principal component analysis; MPCA = moving PCA; SVM = support vector machine; ECM = error correction model; HSTT = hydrostatic seasonal thermal time; NARXNN = non-linear autoregressive exogenous neural network; AANN = auto-associative neural network; RR = robust regression; MARS = multivariate adaptive regression splines; RF = random forest; BRT = boosted regression trees; WNN = wavelet neuran networks; ECM = error correction method. Inputs: h = upstream pool level; S = season; t = time; $\partial(\cdot)$ = time derivative; T_c = concrete temperature; T_{air} = air temperature; T_w = water temperature; $IRF(\cdot)$ = impulse response function; $lag(\cdot)$ = lagged variable; P = precipitation; out = output; $mav(\cdot)$ = moving average; h_d = downstream pool level; M = snow melt; $pdf(\varepsilon)$ = probability density function of error.

A.2 Discussion on "Thermal displacements of concrete dams: Accounting for water temperature in statistical models"

Title: Discussion on "Thermal displacements of concrete dams: Accounting for water temperature in statistical models"

First Author: Fernando Salazar González. CIMNE International Center for Numerical Methods in Engineering

Second Author: Miguel Á. Toledo Municio. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Journal: Engineering Structures

D.O.I. 10.1016/j.engstruct.2015.08.001.

Impact Factor 1.893

Discussion on "Thermal displacements of concrete dams: Accounting for water temperature in statistical models"

F. Salazar[1,*], M.Á. Toledo[2]

Abstract

A discussion on the paper by Tatin et al. (2015) is presented. The paper described an innovative statistical model to interpret dam behaviour, which was validated with artificial data and then applied to seven dams in operation. This discussion provides several comments about the model performance evaluation, as well as suggestions for further analysis of the monitoring data.

Keywords: Concrete dams, Structural health monitoring, Thermal effects, Statistical analysis, Finite element method, Pendulum displacements

1. Discussion

Tatin et al. (hereinafter "the authors") presented an innovative statistical method to interpret dam behaviour. It is based on the traditional HST (Hydrostatic, Season, Time) [1]. Likewise the more recent Thermal HST (HSTT) [2], the new method considers the actual temperature measurements, but also the water temperature and the reservoir level variation in a simplified manner [3]. The result is the method called HST-Grad.

We agree with the authors in that the thermal effect is important in concrete dams, and particularly that caused by the presence of water. In this sense, the new tool constitutes an advance over HST and HSTT, largely maintaining the simplicity of both methods. Nonetheless, it is noteworthy that recent studies have revealed that other phenomena such as solar radiation, shading [4], [5] and night and evaporative cooling [6], are also relevant to the

*Corresponding author

Email addresses: fsalazar@cimne.upc.edu (F. Salazar), matoledo@caminos.upm.es (M.Á. Toledo)

URL: www.cimne.com (F. Salazar)

[1]CIMNE – Centre Internacional de Metodes Numerics en Enginyeria. Campus Norte UPC. Gran Capitán s/n. 08034. Barcelona, Spain

[2]Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment. Profesor Aranguren s/n, 28040, Madrid, Spain

Simulation	Boundary conditions		
	Air temperature	Water temperature	Reservoir level
1	Variable	-	Empty (constant)
2	Variable	Variable	Full (constant)
3	Variable	Variable	Variable

Table 1: Simulations performed by the authors for the heuristic case [3].

temperature field in the dam body and thus affect the displacement field.

The authors admit the inaccuracy of assuming that the thermal load is homogeneous in the upstream face. In our opinion, the situation in the downstream face may be similar in the general case, due to solar radiation and shading [6].

The new method was validated by means of its application to an artificially-generated time series of dam displacements. These data were obtained from a bi-dimensional finite element (FE) model of the Izourt Dam, a 44-m height gravity dam. Then the performance of the new method was assessed by considering actual monitoring data from seven dams in operation. In both cases ("heuristic" and "real" cases from here on, following the authors' terminology in [3]), the HST-Grad model was compared to HST and HSTT.

The HST-Grad model resulted in smaller residuals in most cases, and thus offered a more accurate identification of dam behaviour. The main advantage of these tools is their ease of interpretation, as opposed to others based on machine learning, which nonetheless proved to be highly accurate in recent studies [7], [8], [9].

For the heuristic case, three simulations were performed in [3], whose features are summarised in Table 1.

The result of this analysis showed a better fit of the HST-Grad model, in particular for the Simulation 3. This is coherent with the construction of the three models, and with the boundary conditions applied in each simulation.

The results for the heuristic case could be considered as an estimate of the lower boundary of the residuals that can be expected with each model in practice (Simulation 3). In other words, the error resulting from the application of HST-Grad model to a real case should increase with respect to that shown in Fig. 5 in [3] due to the simplifications introduced

in the FE model: the actual water temperature and its spatial and temporal variation, the actual air temperature, the unconsidered thermal effects, and the thermal inertia of the water mass (in the case presented, the time delay to account for heat transfer between air and water was chosen arbitrarily). Furthermore, the measurement error should be added.

The authors presented the results for the real cases as the standard deviation of the residuals (mm), unlike in the heuristic case, where they were shown "as the ratio (in %) between the standard deviation of the residuals and the amplitude of the displacement analysed". In our opinion, it is more convenient to present the results in relative terms, for three reasons:

- In practice, the relevance of the residual depends on its relationship with the displacement amplitude. The same applies to model comparison.
- A dimensionless residual allows comparison between different dams, whose behaviour depend on several factors, including the dam typology and height.
- The use of the same goodness-of-fit index would ease comparison between the results for the heuristic and the real cases correspondent to Izourt Dam.

We found particularly relevant the similarity of the results for the calibration and forecast periods (Fig. 12 in [3]). This indicates that the residual for the training period is a good indicator of the general model accuracy. For the same reason, a deeper analysis would be needed for the Izourt Dam, which was the only exception: the standard deviation of the residuals increased between the calibration (0.5 mm) and the forecast period (0.8 mm; Fig. 12 in [3]). The description provided for Izourt Dam does not suggest a potential explanation, given that it is the lowest (44 m) and simplest (rectilinear gravity dam) of the seven dams considered.

We analysed the results published by the authors by extracting comparable values from Fig. 5 and Fig. 12 in [3]. They are showed in Table 2.

It should be noted that the results largely differ between the heuristic and the real cases in terms of relative residual reduction (75 to 10%). However, the difference in absolute value is much less relevant (0.1 mm).

In our opinion, though the heuristic case is highly valuable for validation purposes, the HST-Grad model assessment should be mostly based on the results for the real cases. For

Case	Residuals standard deviation (mm)		HST-Grad residual reduction	
	HSTT	HST-Grad	Absolute (mm)	Relative (%)
Heuristic (Simulation 3)	0.2	0.05	0.15	75
Real	0.5	0.45	0.05	10

Table 2: Results of HST-Grad for Izourt Dam. Approximated values extracted from [3]. The displacements amplitude was supposed to be 5.0 mm (Fig. 8 in [3])

the seven dams considered, the residual standard deviation obtained with HST-Grad was around 10-15% lower than that of HSTT (Fig. 12 in [3]).

A more detailed analysis of the contribution of each source of error for the real cases would be highly interesting, as well as the influence of other specific dam features: typology, height, location, orientation, and reservoir operation. Regarding the latter, it would be helpful to know the reservoir level variation for the analysed dams, both in the calibration and forecast periods, as well as the reading frequency (the amount of data available). This could help the interpretation of the results of the comparative study, once the heuristic case confirmed that the relative performance of HSTT and HST-Grad was strongly influenced by the reservoir level variation.

We also consider that a four-way comparison between the FE model, the monitoring data, and the HSTT and HST-Grad estimates for Izourt Dam would be highly valuable. It might allow evaluation of the advantages and disadvantages of either model. FE models can be useful, especially for gravity dams that can be modelled in 2D. They have been previously applied for interpreting dam behaviour even for buttress dams, which are more complex and must be modelled in 3D [10].

2. Conclusion

The HST-Grad method presented by the authors constitutes an improvement over other currently used statistical methods to interpret dam behaviour. However, it is our belief that the reduction in the residuals deviation with respect to HSTT will generally be much closer to that obtained for the real cases (10-15%), than for the heuristic one (75%).

3. Acknowledgements

The research was supported by the Spanish Ministry of Economy and Competitiveness (*Ministerio de Economía y Competitividad*, MINECO) through the projects iComplex (IPT-2012-0813-390000) and AIDA (BIA2013-49018-C2-1-R and BIA2013- 49018-C2-2-R).

References

[1] Willm G, Beaujoint N. Les méthodes de surveillance des barrages au service de la production hydraulique d'Electricité de France-Problèmes ancients et solutions nouvelles. In: 9th ICOLD Congres. 1967, p. 529–50. Q34-R30. [in French].

[2] Penot I, Daumas B, Fabre J. Monitoring behaviour. Water Power and Dam Construction 2005;.

[3] Tatin M, Briffaut M, Dufour F, Simon A, Fabre JP. Thermal displacements of concrete dams: Accounting for water temperature in statistical models. Engineering Structures 2015;91:26–39.

[4] Jin F, Chen Z, Wang J, Yang J. Practical procedure for predicting non-uniform temperature on the exposed face of arch dams. Applied Thermal Engineering 2010;30(14):2146–56.

[5] Sheibany F, Ghaemian M. Effects of environmental action on thermal stress analysis of Karaj concrete arch dam. Journal of Engineering Mechanics 2006;132(5):532–44.

[6] Santillán D, Salete E, Vicente D, Toledo M. Treatment of solar radiation by spatial and temporal discretization for modeling the thermal response of arch dams. Journal of Engineering Mechanics 2014;140(11).

[7] Mata J. Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. Engineering Structures 2011;3(3):903 –10. doi: 10.1016/j.engstruct.2010.12.011.

[8] Ranković V, Grujović N, Divac D, Milivojević N, Novaković A. Modelling of dam behaviour based on neuro-fuzzy identification. Engineering Structures 2012;35:107–13.

[9] Salazar F, Toledo M, Oñate E, Morán R. An empirical comparison of machine learning techniques for dam behaviour modelling. Structural Safety 2015;56(0):9 – 17. doi: http://dx.doi.org/10.1016/j.strusafe.2015.05.001.

[10] De Sortis A, Paoliani P. Statistical analysis and structural identification in concrete dam monitoring. Engineering Structures 2007;29(1):110–20.

A.3 An empirical comparison of machine learning techniques for dam behaviour modelling

Title: An empirical comparison of machine learning techniques for dam behaviour modelling

First Author: Fernando Salazar González. CIMNE International Center for Numerical Methods in Engineering

Second Author: Miguel Á. Toledo Municio. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Third Author: Eugenio Oñate Ibáñez de Navarra. CIMNE - International Center for Numerical Methods in Engineering

Fourth Author: Rafael Morán Moya. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Journal: Structural Safety

D.O.I. 10.1016/j.strusafe.2015.05.001

Impact Factor 2.086

An empirical comparison of machine learning techniques for dam behaviour modelling

F. Salazar[1,*], M.A. Toledo[2,*], E. Oñate[1,*], R. Morán[2,*]

Abstract

Predictive models are essential in dam safety assessment. Both deterministic and statistical models applied in the day-to-day practice have demonstrated to be useful, although they show relevant limitations at the same time. On another note, powerful learning algorithms have been developed in the field of machine learning (ML), which have been applied to solve practical problems. The work aims at testing the prediction capability of some state-of-the-art algorithms to model dam behaviour, in terms of displacements and leakage. Models based on random forests (RF), boosted regression trees (BRT), neural networks (NN), support vector machines (SVM) and multivariate adaptive regression splines (MARS) are fitted to predict 14 target variables. Prediction accuracy is compared with the conventional statistical model, which shows poorer performance on average. BRT models stand out as the most accurate overall, followed by NN and RF. It was also verified that the model fit can be improved by removing the records of the first years of dam functioning from the training set.

Keywords: Dam monitoring, Dam safety, Data analysis, Boosted regression trees, Neural networks, Random forests, MARS, Support vector machines, Leakage flow

1. Introduction and background

Dam safety assessment is a complex task due to the uniqueness of each of such structures and their foundations. It is commonly based on three main pillars: visual inspection, engineering knowledge and a behaviour model. The actual response of the dam is compared with the predictions of the model, with the aim of detecting anomalies and preventing failures. Current predictive methods can be classified as follows [1]:

- Deterministic: typically based on the finite element method (FEM), these methods calculate the dam response on the basis of the physical governing laws.
- Statistical: exclusively based on dam monitoring data.

*Corresponding author

Email addresses: fsalazar@cimne.upc.edu (F. Salazar), matoledo@caminos.upm.es (M.A. Toledo), onate@cimne.upc.edu (E. Oñate), rmoran@caminos.upm.es (R. Morán)

URL: www.cimne.com (F. Salazar)

[1]International Center for Numerical Methods in Engineering (CIMNE). Campus Norte UPC. Gran Capitán s/n. 08034. Barcelona, Spain

[2]Technical University of Madrid (UPM). Civil Engineering Department: Hydraulics and Energy. Profesor Aranguren s/n, 28040, Madrid, Spain

- Hybrid: deterministic models which parameters have been adjusted to fit the observed data.

- Mixed: comprised by a deterministic model to predict the dam response to hydrostatic pressure, and a statistical one to consider deformation due to thermal effects.

It is difficult to predict dam behaviour with high accuracy. Numerical models based on the FEM provide useful estimates of dam movements and stresses, but are subject to a significant degree of uncertainty in the characterisation of the materials, especially with respect to the dam foundation. Other assumptions and simplifications have to be made, regarding geometry and boundary conditions. These tools are essential during the initial stages of the life cycle of the structure, provided that there are not enough data available to build data-based predictive models. However, their results are often not accurate enough for a precise assessment of dam safety.

This is more acute when dealing with leakage in concrete dams and their foundations, due to the intrinsic features of the physical process, which is often non-linear [2], and responds to threshold and delayed effects [3], [4]. Numerical analysis cannot deal with such a phenomenon, because comprehensive information about the location, geometry and permeability of each fracture would be needed. As a result, deterministic models are not used in practice for the prediction of leakage flow in concrete dams [1].

Many of the dams in operation have a large number of monitoring devices, recording the evolution of various indicators such as movements, leakage flow or the pore water pressure, among others. Although there are still many dams with few observed data, there is a clear trend towards the installation of a larger number of devices with higher data acquisition frequency [5]. As a result, there is an increasing amount of information on the dam performance, which makes it interesting to study the ability of machine learning (ML) tools to process them, build behaviour models and extract useful information [6].

The paper assesses the potential of some state-of-the-art ML techniques to build models for the prediction of dam behaviour. The results are compared with the conventional statistical method.

1.1. Statistical models

The most popular data-driven approach for the prediction of dam behaviour is the hydrostatic-seasonal-time (HST) method, characterised by taking into account three effects:

- A reversible effect of the hydrostatic load.

- A reversible seasonal thermal influence of the temperature.

- An irreversible term due to the evolution of the dam response over time.

This assumption is coherent with the observed behaviour of many concrete dams in terms of displacements. However, it has also been applied to other variables, such as uplifts and leakage [3]. Similar schemes have been used for rock-fill dams, although it is acknowledged that the temperature is not relevant, and that the irreversible effect of settlements prevails on the elastic response to hydrostatic load. Furthermore, rainfall may have a strong influence on leakage [3].

The main drawbacks of HST and other methods based on linear regression are the following:

- The functions have to be defined beforehand, and thus may not represent the true behaviour of the structure [3].

- The governing variables are supposed to be independent, although some of them have been proven to be correlated [7].

- They are not well-suited to model non-linear interactions between input variables [2].

1.2. Advanced data analysis in dam monitoring

The examples of application of innovative techniques to improve the results of HST are becoming more frequent in recent years. As an example, Bonelli and Radzicki [8] used an impulse-response function for predicting the pore pressure in the dam body. The method provided accurate results in the test cases, showing the hysteresis effect by which the pore pressure is lower during filling than it should be in a stationary state, and vice versa. Nonetheless, given that it makes a strong assumption on the characteristics of the phenomenon, it is restricted to specific processes.

Li et *al.* [9] proposed a method to improve HST based on cointegration theory. They tested the stationarity of the monitoring data series before fitting a multi-linear regression (MLR) model.

One obvious weakness of linear regression is that it cannot reproduce nonlinear relations between variables. This problem is typically overcome by introducing higher order terms of the covariates. Neural networks (NN) constitute a powerful alternative to solve this issue. Their flexibility and capability to adapt to highly complex interactions have made them popular in several fields of engineering, including dam monitoring (see for example [3], [10], [11], and [12]).

However, it should be noted that NN have some drawbacks:

- The result depends on the initialisation of the weights.

- The best network architecture (number of hidden layers and neurons in each layer) is not known beforehand.

- The model is prone to over-fitting.

- The training process may reach a local minimum of the error function.

Several techniques have been developed to overcome these shortcomings, which in general lead to an increase in the computational cost [13]. In spite of this, NN stand out as the most popular ML tool in dam engineering, and the results are promising [3]. Further models have been also applied to dam monitoring, such as ANFIS (adaptive network-based fuzzy inference system) models [14], principal component analysis [6], NARX (nonlinear autoregressive with exogenous input) models [15] or K-nearest neighbours [16]. However, these tools are rarely used in practice, where HST still prevails. Moreover, most of the previous studies are limited to one single variable of specific dams [11], [12]. Hence, the results are not generally applicable.

1.3. Objectives

The study aims to assess the prediction accuracy of some ML tools, most of which have been seldom used in dam engineering. Specifically, the algorithms selected are: random forests (RF), boosted regression trees (BRT), support vector machines (SVM) and multivariate adaptive regression splines (MARS). Both HST and NN were also used for comparison purposes. Similar analyses have been performed in other fields of engineering, such as the prediction of urban water demand [17].

A further singularity of dams is that the early years of operation often correspond to a transient state, non-representative of the quasi-stationary response afterwards [18]. In such a scenario, using those years for training a predictive model would be inadvisable. This might lead to question the optimal size of the training set in achieving the best accuracy. De Sortis [19] ran a sensitivity analysis and concluded that at least 10 years were needed to obtain acceptable predictions. However, his study was limited to the prediction of the radial displacement in one particular location of a specific dam by using HST. A similar work was performed by Chouinard and Roy [2]. This paper seeks to extend such studies to other learning algorithms and output variables.

2. Case study and variable selection

The data used for the study correspond to La Baells dam. It is a double curvature arch dam, with a height of 102 m, which entered into service in 1976. The monitoring system records the main indicators of the dam performance: displacement, temperature, stress, strain and leakage. The data were provided by the Catalan Water Agency (*Agència Catalana de l'Aigua*, ACA), the dam owner, for research purposes. Among the available records, the study focuses on 14 variables: 10 correspond to displacements measured by pendulums (five radial and five tangential), and four to leakage flow. Several variables of different types were considered in order to obtain more reliable conclusions. Table 1 shows some statistics of the target variables, whereas the location of each monitoring device is depicted in Figure 1.

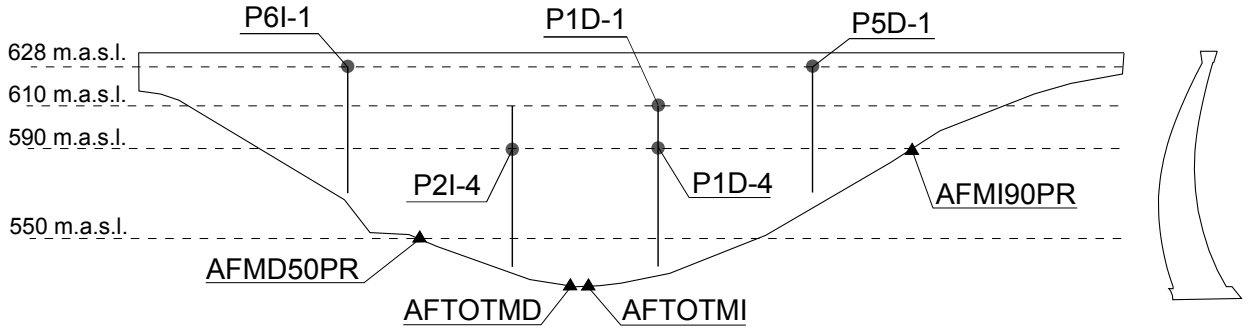


Figure 1: Geometry and location of the monitoring devices in La Baells Dam. Left: view from downstream. Right: highest cross-section.

The data acquisition frequency is of the order of one record per week. The measurement error of the devices is $\pm 0.1mm$ for displacements, and negligible for leakage flows (measured using the volumetric method). The series span from 1979 to 2008. In all cases, approximately 40% of the records (from 1998 to 2008) were left out as the testing set. This is a large

Target	# Observations	Type	units	Mean	Min.	Max.
P1DR1	1,194	Radial displ.	mm	-10.75	-20.6	2.1
P1DR4	1,194	Radial displ.	mm	-9.88	-16.8	0.0
P2IR4	1,191	Radial displ.	mm	-7.77	-17.5	1.3
P5DR1	1,193	Radial displ.	mm	-6.37	-14.8	1.9
P6IR1	1,198	Radial displ.	mm	-9.24	-17.5	0.1
P1DT1	1,194	Tangential displ.	mm	2.36	0.0	3.9
P1DT4	1,194	Tangential displ.	mm	-0.32	-1.5	0.3
P2IT4	1,191	Tangential displ.	mm	-1.56	-2.7	-1.1
P5DT1	1,193	Tangential displ.	mm	-0.09	-1.8	1.6
P6IT1	1,199	Tangential displ.	mm	-2.04	-4.2	0.1
AFMD50PR	1,016	Leakage	l/min	5.05	0.0	27.3
AFMI90PR	994	Leakage	l/min	0.63	0.0	3.0
AFTOTMD	1,064	Leakage	l/min	7.30	0.1	35.8
AFTOTMI	1,014	Leakage	l/min	2.89	0.1	12.4

Table 1: Target variables considered in the comparative study

proportion compared with previous studies, which typically leave 10-20 % of the available data for testing [14], [12], [16]. A larger test set was selected in order to increase the reliability of the results.

Four different training sets were chosen to fit each model, spanning five, 10, 15 and 18 years of records. In all cases, the training data used correspond to the closest time period to the test set (e.g. periods 1993-1997, 1988-1997, 1983-1997, and 1979-1997, respectively).

Other environmental variables are recorded at the dam site, and were considered as inputs: air temperature, reservoir level and rainfall (Figure 2). Although the latter has no influence on displacements, it has been included to test whether the methods analysed can handle noisy or non-influential inputs without decreasing prediction accuracy.

Some derived variables were also calculated, namely:

- Average velocity of reservoir level variation in different periods prior to the measurement (10, 20 and 30 days).
- Accumulated rainfall over various periods (30, 60, 90 and 180 days) prior to the reading.
- Moving averages of reservoir level and air temperature over seven, 14, 30, 60, 90 and 180 days before the record.

Finally, the year, number of days from the first record, and month were also considered as explanatory variables up to 25.

The variable selection was performed according to dam engineering practice. Both displacements and leakage are strongly dependant on hydrostatic load. Air temperature is well known to affect displacements, in the form of a delayed action. It is not clear whether it strongly influences leakage or not. Whereas Seifart et *al.* [20] reported that leakage in the Itaipú Dam follows a seasonal cycle "due clearly to the thermal effect on the opening and closure of joints", other studies showed no dependency [3]. Both the air temperature and some moving averages were included in the analysis.

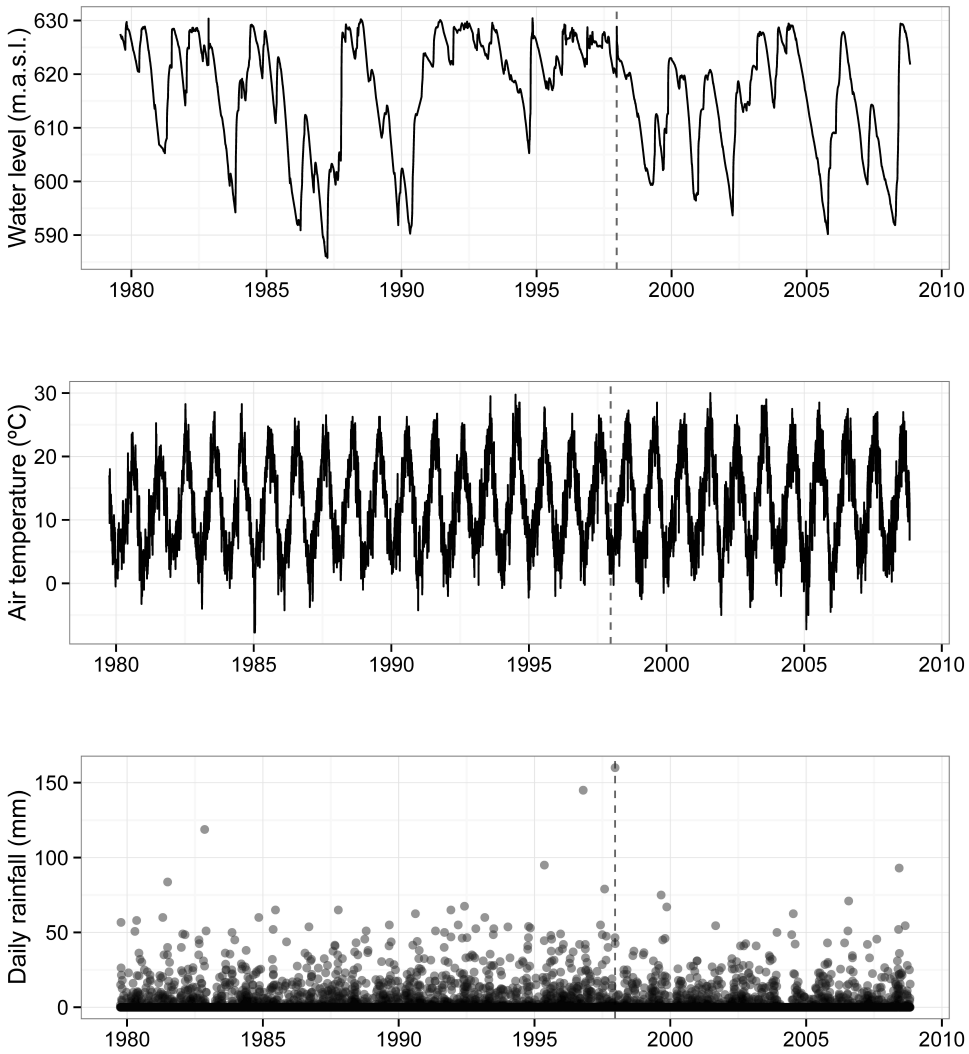


Figure 2: Time series of environmental variables at La Baells dam site. From top to bottom: water level, air temperature and daily rainfall. The vertical dashed line marks the division between training and test periods.

Hydrostatic load induces an almost immediate response of the dam, although some studies suggest that there may be also a delayed effect, specially for leakage [11], [21]. Moving averages of reservoir level were considered to capture it. The velocity of variation of reservoir level over different periods was also included, following studies that suggest the existence of an influence in dam displacements [22].

In order to account for the temporal drift of the structure, both the year and the number of days from the first record were also added.

A relatively large set of predictors was used to capture every potential effect, overlooking the high correlation among some of them. In addition, the comparison sought to be as unbiased as possible, thus all the models were built using the same inputs[3]. While it

[3]with the exceptions of MARS and HST, as explained in sections 3.5 and 3.6 respectively

is acknowledged that this procedure may favour the techniques that better handle noisy or scarcely important variables, theoretically all learning algorithms should discard them automatically during the model fitting.

3. Methods

In this section, the algorithms chosen to build the prediction models are briefly described. Although the detailed mathematical description is beyond the scope of the paper, a short description, the most relevant features, and some key references are included. All the models were built by using the language/programming environment R [23] and some of its packages, which are cited in each section.

The objective is to predict an output variable $Y \in \mathbb{R}$ based on the value of a set of predictors $X \in \mathbb{R}^p$, i.e. $Y \approx \hat{Y} = F(X)$. The observed values are denoted as $(x_i, y_i), i = 1, ..., N$, where N is the number of observations. Note that each x_i is a vector with p components, each of which is referred to as x_i^j, when necessary. Similarly, $X^j, j = 1, ..., p$ stands for each dimension of the input space.

3.1. Random forests (RF)

An RF model is a group of regression (or classification) trees [24], trained on altered versions of the training set. Given that its output is the average of the prediction of each individual tree, it is an *ensemble method*. Since RF were first introduced by Breiman [25], they have become highly popular as a method to build predictive models [26]. The training process has two random components:

- Only a random subset of the input variables is considered to perform each division of the input space.
- Each tree is built using a different training set, obtained from the original data via random sampling with replacement.

The aim of adding randomness is to generate substantially different trees, so that the ensemble captures as many patterns in the training set as possible. Other interesting features of RF are the following:

- They can easily handle continuous, discrete and categorical inputs, as well as missing values.
- They can naturally model non-linear interactions.
- They avoid the need to perform cross-validation, because an unbiased estimate of the generalisation error is computed during the training process (out-of-bag (OOB) error).

Two parameters can be tuned for building an RF model: the number of covariates to consider at each split (*mtry*), and the total number of trees in the forest. Neither has significant influence on the results, according to the majority of published authors (e.g. [25], [26]).

The default value of *mtry* for regression is $p/3$, with p being the number of covariates. An RF model was fitted with the default *mtry*, and then it was increased and decreased to find the value that gives the minimum OOB error. The function *tuneRF* of the R package "randomForest" [27] was used.

All RF models entailed 500 trees, with it being checked that the OOB error is stable with that size.

3.2. Boosted regression trees (BRT)

Boosting is a general scheme to build ensemble prediction models [28]. Although various methods can be selected to form the ensemble, regression trees are frequently chosen, and were used in this work.

The idea is to build an ensemble so that each single model (often referred to as *base learners*), is fitted to the residual of the previous ensemble. The overall prediction is calculated as the weighted sum of the outputs of the base learners (unlike RF, where the prediction of the ensemble is the average).

The algorithm includes two ingredients to avoid over-fitting:

- Each learner is trained on a random subset (without replacement) of the training set. This also decreases the computational cost.

- A regularisation (shrinkage) parameter $\nu \in (0, 1)$ is applied.

Some empirical analyses show that relatively low values of ν (below 0.1) greatly improve generalisation capability [28]. The optimal value depends on the number of base learners. In practice, it is common to set the regularisation parameter and calculate a number of trees such that the training error stabilises [29]. Subsequently, a certain number of terms are "pruned" by using for example cross-validation. The library used [29] allows choice of the number of trees by different methods. The value $\nu = 0.001$ was considered and the number of trees was selected by means of five-fold cross-validation. The process was repeated by using trees of depth 1 and 2 (*interaction.depth*), and the most accurate for each target selected.

3.3. Neural networks (NN)

NN models have been applied to solve a wide variety of problems. Among the different types of NN found in the literature [13], the multilayer perceptron (MLP) was selected for this work. An MLP is formed by a number of single units, called perceptrons, organised in different layers. The simplest architecture of an MLP was used, which involves three layers: input, hidden and output. Each perceptron in the hidden layer applies a nonlinear transformation to the inputs, and yields a response, which is later combined to compute the model output. Thus, NN are appropriate to model non-linear input-output relations.

NN stand out as one of the most popular machine learning techniques for civil engineers. Some previous applications to dam monitoring have been mentioned in section 1.2.

The package "nnet" [30] was used, which allows tuning the NN models by setting several parameters. The *size* (number of perceptrons in the hidden layer) and the *decay* (regularisation parameter) are the most influential in the results [31]. All the possible combinations of three, 12 and 25 perceptrons (*size*) with *decay* of 0.1, 0.01 and 0.001 were tried, and the pair of values which showed the minimum error via five-fold cross-validation was chosen. For each fold, 15 NN with different initialisations were fitted, and the average error was compared. Thus, the accuracy of every combination of parameters was computed on the basis of 75 NN.

The selected parameters were applied to train 20 NN models over the whole training set with the function *avNNet*, from the R package "caret" [32]. The final prediction was computed as the average of the 20 NN.

3.4. Support vector machines (SVM)

This learning algorithm is based on a non-linear transformation of the predictor variables to a higher dimensional space (often referred to as *feature space*), and a linear regression on these transformed variables. The mathematical development of the method is complex and beyond the scope of the paper. Detailed and rigorous descriptions can be found in [33] and [34], and a recent application in predicting dam behaviour is reported in [35]. The method uses an ε-insensitive error function that neglects errors below the threshold ε. The algorithm searches for a trade-off between the minimum error and the smoothness of the approximating function. The library "e1071" within the R environment [36] was used, which allows tuning the most important parameters [37] of an SVM model:

- The "cost" parameter, C. Values of 10, 100 and 500 were tested.
- The width of the ε-insensitive error function, ε. The default value (0.1) was chosen.
- The kernel function, which defines the mapping from the input to the feature space. A radial basis function was considered: $K(x_i, X) = e^{-\gamma|x_i - X|^2}$
- The γ parameter of the kernel. Values of 1, 0.1, 0.01, 0.001 and 0.0001 were tried.

The 15 possible combinations of C and γ were applied to fit SVM models on the training data. Four-fold cross-validation was performed to obtain the best values. Each fold and combination of parameters was repeated five times to account for randomness, and the one with the lowest average error was selected to train an SVM model over the whole training set.

3.5. Multivariate adaptive regression splines (MARS)

MARS is an adaptive algorithm originally proposed by Friedman [38]. It is based on the combination of elementary piecewise linear functions, which definition depends on the data. As an example, an input data $x_j^j = k$ defines a pair of basic functions of the form $(X^j - k)_+$ and $(k - X^j)_+$. The subscript "+" stands for the positive part, i.e.: $(X^j - k)_+ = X^j - k$ if $X^j > k$; 0, otherwise [31]. The algorithm starts with a constant value and adds pairs of functions as long as the training error decreases above a given threshold. This is the forward pass. At the end of this step, the resulting model typically over-fits the data. Then a "pruning" process is followed, during which some of the functions are eliminated according to the generalised cross validation (GCV) criterion. GCV is a modification of the residual sum of squares (RSS) which takes into account the number of parameters of the model [31]. In practice, the method searches for a trade-off between the reduction in the training error and the complexity of the model.

MARS models are well suited to non-linear problems, as well as easily interpretable. Furthermore, the algorithm implicitly performs variable selection, given that the functions in the final ensemble depend only on the most relevant predictors X^j.

The work was performed using the library "earth" [39], and the parameter tuned was the maximum number of terms in the final model (*nprune*). Five-fold cross-validation was run, repeated five times ($nfold = 5$ and $ncross = 5$ in the *earth* function). In principle, the model with the highest coefficient of determination (RSq) should be selected. However, the results of some preliminary tests showed that in most cases the RSq rose sharply after adding the first few terms, and remained almost constant afterwards. For the sake of model simplicity and

generalisation capability, the lower value of *nprune* with $RSq \geq mean(RSq) - SD(RSq)$ was selected, a criterion similar to the *1 SE rule* proposed by Breiman et *al.* [40]. The same tests also revealed that the models with one or more time-dependant functions in the final ensemble (i.e. considering the year and/or the day since the first record) had poor generalisation ability. Therefore, both inputs were removed from the set of predictors.

3.6. HST

A conventional HST model was also built, in order to compare the results with current engineering practice. The most typical form was chosen:

$$\begin{aligned}\hat{Y} = F(t,h,s) \;\; = \;\; & a_0 + a_1 h + a_2 h^2 + a_3 h^3 + a_4 h^4 + a_5 h^5 \\ & + a_5 e^{-t} + a_6 t + a_7 cos(s) + a_8 sen(s) \\ & + a_9 sen^2(s) + a_{10} sen(s) cos(s)\end{aligned}$$

$$s = \frac{d}{365,25} 2\pi$$

where d is the number of days since 1 January, t is the elapsed time (years), h is the reservoir level, and $a_1, a_2, ..., a_{10}$ are the coefficients to fit.

3.7. Measures of accuracy

The accuracy of regression models is frequently measured via the mean absolute error (MAE), computed as:

$$MAE = \frac{\sum_{i=1}^{N} |y_i - F(x_i)|}{N}$$

where N is the size of the training (or test) set, y_i are the observed outputs and $F(x_i)$ the predicted values. Given that MAE is measured in the same units as the target variable, it provides a useful indication of prediction accuracy. However, it takes into account neither the mean value of the output, nor its deviation. Moreover, it is not appropriate to compare results correspondent to outputs of a different nature (e.g. displacements vs flows). To overcome these drawbacks, the study is mostly based on the average relative variance (ARV) [41]:

$$ARV = \frac{\sum_{i=1}^{N} (y_i - F(x_i))^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} = \frac{MSE}{\sigma^2}$$

where $\bar{y}$ is the output mean. Given that ARV denotes the ratio between the mean squared error (MSE) and the variance (σ^2), it accounts both for the magnitude and the deviation of the target variable. Furthermore, a model with ARV=1 is as accurate a prediction as the mean of the observed outputs.

4. Results and discussion

Models for 14 targets, with six learning algorithms, trained over four different training sets were fitted, i.e. $14 \times 6 \times 4 = 336$ models. Due to space constraints, only one plot is presented in Figure 3 as an example. It shows the predictions of the BRT model trained over the whole training set, in comparison with the measured data for three targets of different kind (P1DR1, P1DT1 and AFMD50PR). It provides an intuition on the goodness of fit

achieved, and highlights how the ARV allows comparison of the accuracy between different targets. Although the highest MAE corresponds to P1DR1, it yields the lowest ARV at the same time, because of its higher variance.

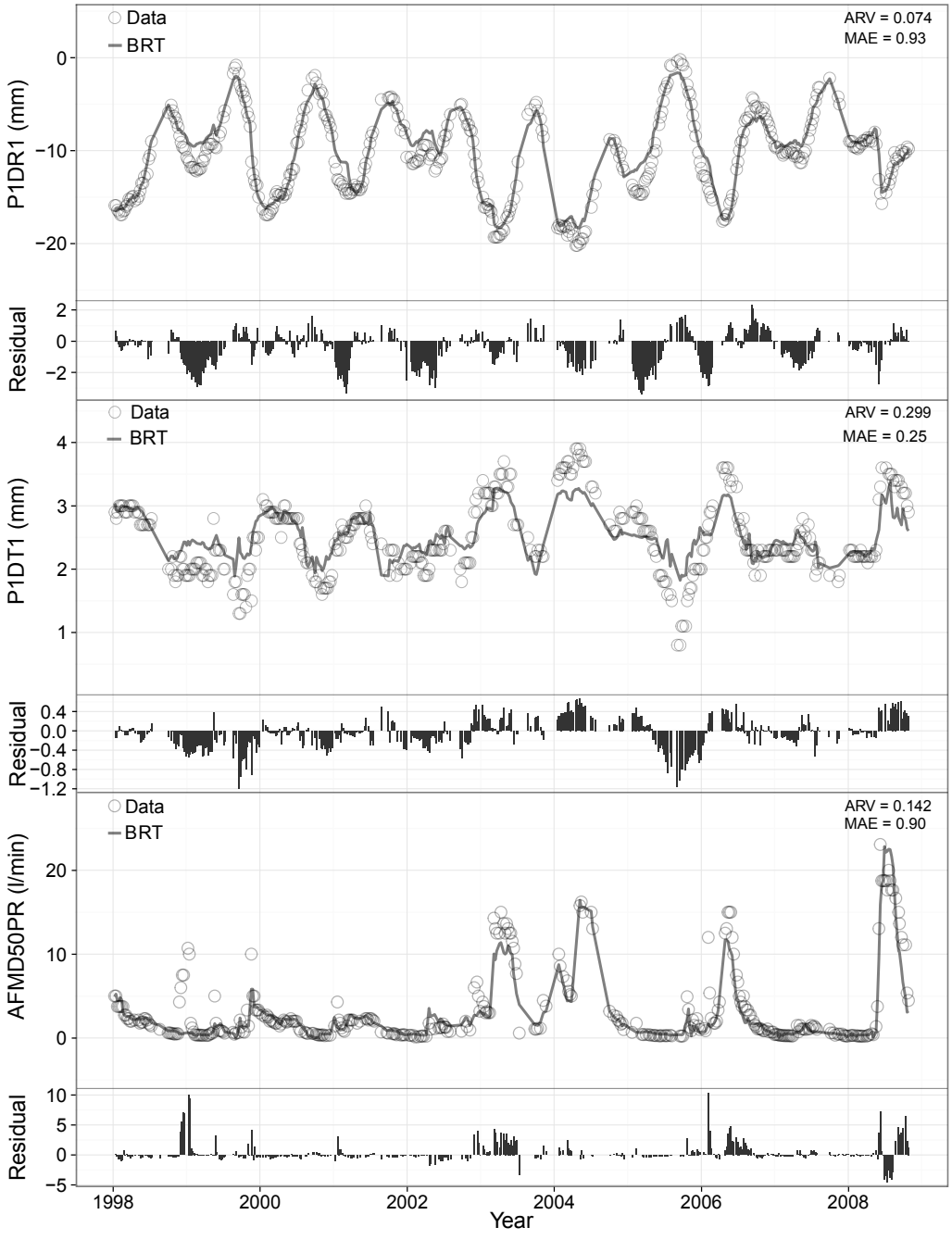


Figure 3: Measured data (circles) versus predictions of the BRT model (lines) for the test period. The residuals between them are included below each plot. From top to bottom: P1DR1, P1DT1 and AFMD50PR.

It is commonly accepted that increasing the amount of training data leads to a better model performance. Although this may not be the case of dams, in general the statistical models for dam monitoring are fitted using all the available data. Table 2 shows the MAE

Type	Target	RF	BRT	NN	SVM	MARS	HST
Radial (mm)	P1DR1	1.70	0.93	**0.58**	0.75	2.32	1.35
	P1DR4	1.05	**0.71**	**0.68**	0.76	1.50	1.37
	P2IR4	0.94	0.97	1.02	1.05	**0.85**	1.12
	P5DR1	0.86	**0.70**	**0.64**	1.35	0.89	0.88
	P6IR1	1.47	0.69	0.72	**0.60**	1.67	0.91
Tangential (mm)	P1DT1	**0.24**	**0.25**	0.52	0.35	0.55	0.47
	P1DT4	**0.15**	**0.15**	0.18	0.19	0.22	0.20
	P2IT4	0.13	**0.11**	0.13	0.12	0.14	**0.10**
	P5DT1	0.40	0.22	**0.19**	0.38	0.47	**0.18**
	P6IT1	**0.28**	**0.27**	0.39	0.94	0.39	0.51
Leakage (l/min)	AFMD50PR	1.24	**0.90**	2.11	4.25	1.74	2.24
	AFMI90PR	0.18	0.15	**0.07**	0.33	0.25	0.28
	AFTOTMD	1.82	**1.60**	3.04	5.38	1.85	2.60
	AFTOTMI	0.91	**0.42**	0.83	1.49	1.49	1.11

Table 2: MAE for each output and model, fitted on the whole training set (18 years). The values within 10% from the minimum are highlighted in bold, and the minimum MAE are also underlined. The results correspond to the test set.

for each target and model, fitted on the whole training set, i.e., 18 years.

It can be seen that models based on ML techniques mostly outperform the reference HST method. NN models yield the highest accuracy for radial displacements, whereas BRT models are better on average both for tangential displacements and leakage flow. It should be noted that the MAE for some tangential displacements is close to the measurement error of the device ($\pm 0.1mm$).

Figure 4 shows the results in terms of ARV for each model and type of output. It should be remembered that models with ARV > 1 can be considered as being of little use. The error is lower for radial displacements, whereas there is not a great difference between the ARV for leakage flow and tangential displacements. These results are in accordance with engineering knowledge: the prediction of tangential displacements is more difficult because the signal-to-noise ratio is lower than for radial displacements (while the measurement error is the same, the standard deviations are highly different, as shown in Table 1). The measurement error for leakage flow is negligible, but it is governed by a more complex physical process, which makes it harder to predict.

The study was repeated with each technique, using different training set sizes, namely five, 10 and 15 years. The results were compared to those obtained previously, with 18 years. The test set was the same as before (1998-2008). Figure 5 shows the results. An important decrease in error is observed in most cases between models trained on five and 10 years. This decrease is dramatic for HST (note that some of the results for HST and five years lie outside the vertical limit of the plots).

Although some previous studies offered similar results [19], in this case such effect may be more pronounced due to the fact that the reservoir level remained high in the 1993-1998 period (Fig. 2). Models fitted on those years have no information on the dam behaviour when the reservoir is at low levels, and therefore the prediction of the dam response in such situations may be highly inaccurate.

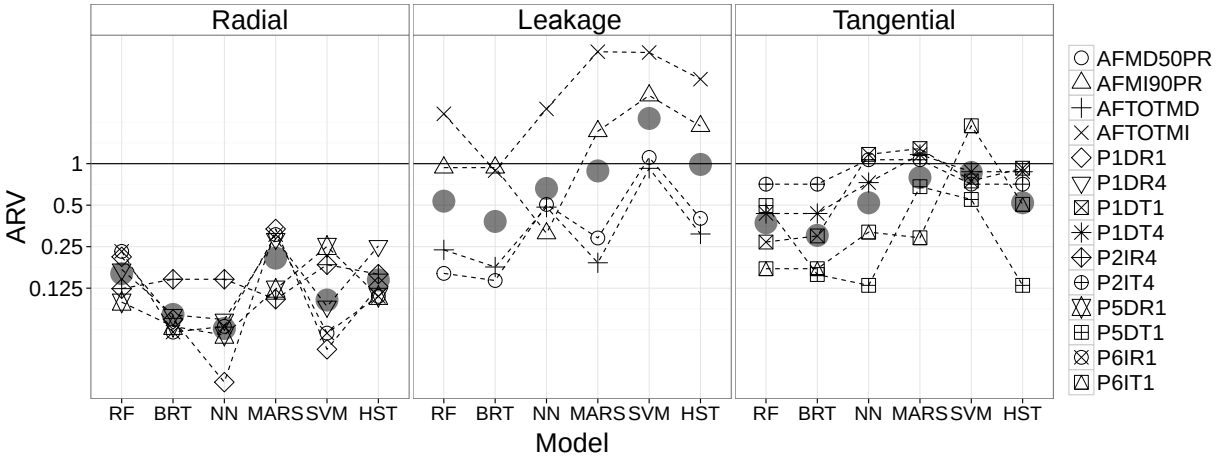


Figure 4: *ARV* for each target and model, fitted on the whole training set (18 years). Models with $ARV > 1.0$ are less accurate than the sample mean. The average values for each technique and type of output are plotted with black dots. Note the logarithmic scale of the vertical axis. The results correspond to the test set.

When increasing the training set up to 15 and 18 years, the variation is either negligible (i.e. BRT models for leakage, Figure 5, bottom), or there is a small decrease in error (i.e. NN models for radial displacements, Figure 5, top). In some cases, the error even increases, such as in HST models for radial displacements (Figure 5, top). Some techniques do not show a clear trend, such as MARS models for tangential displacements (Figure 5, bottom).

Table 3 compares the best models overall with those trained on the entire training set (18 years). Although the use of the whole training set is optimal for six out of 14 targets,

Target	Best model 18 years	MAE 18 years	Best model overall	MAE overall	Best training size (years)	MAE reduction (%)
P1DR1	NN	0.58	-	-	-	-
P1DR4	NN	0.68	MARS	0.60	5	13.3
P2IR4	MARS	0.85	MARS	0.81	15	4.7
P5DR1	NN	0.64	-	-	-	-
P6IR1	SVM	0.60	SVM	0.53	10	11.7
P1DT1	RF	0.24	BRT	0.22	10	8.3
P1DT4	RF/BRT	0.15	BRT	0.14	10	6.7
P2IT4	HST	0.10	-	-	-	-
P5DT1	HST	0.18	-	-	-	-
P6IT1	BRT	0.27	MARS	0.23	5	14.8
AFMD50PR	BRT	0.90	BRT	0.89	15	1.1
AFMI90PR	NN	0.07	-	-	-	-
AFTOTMD	BRT	1.60	BRT	1.57	15	1.9
AFTOTMI	BRT	0.42	-	-	-	-

Table 3: Comparison between the best models fitted using the whole training set and the best overall. Empty rows correspond to outputs for which no improvement is achieved by using a smaller training set. The results correspond to the test set.

significant improvements are reported in some cases by eliminating some of the early years.

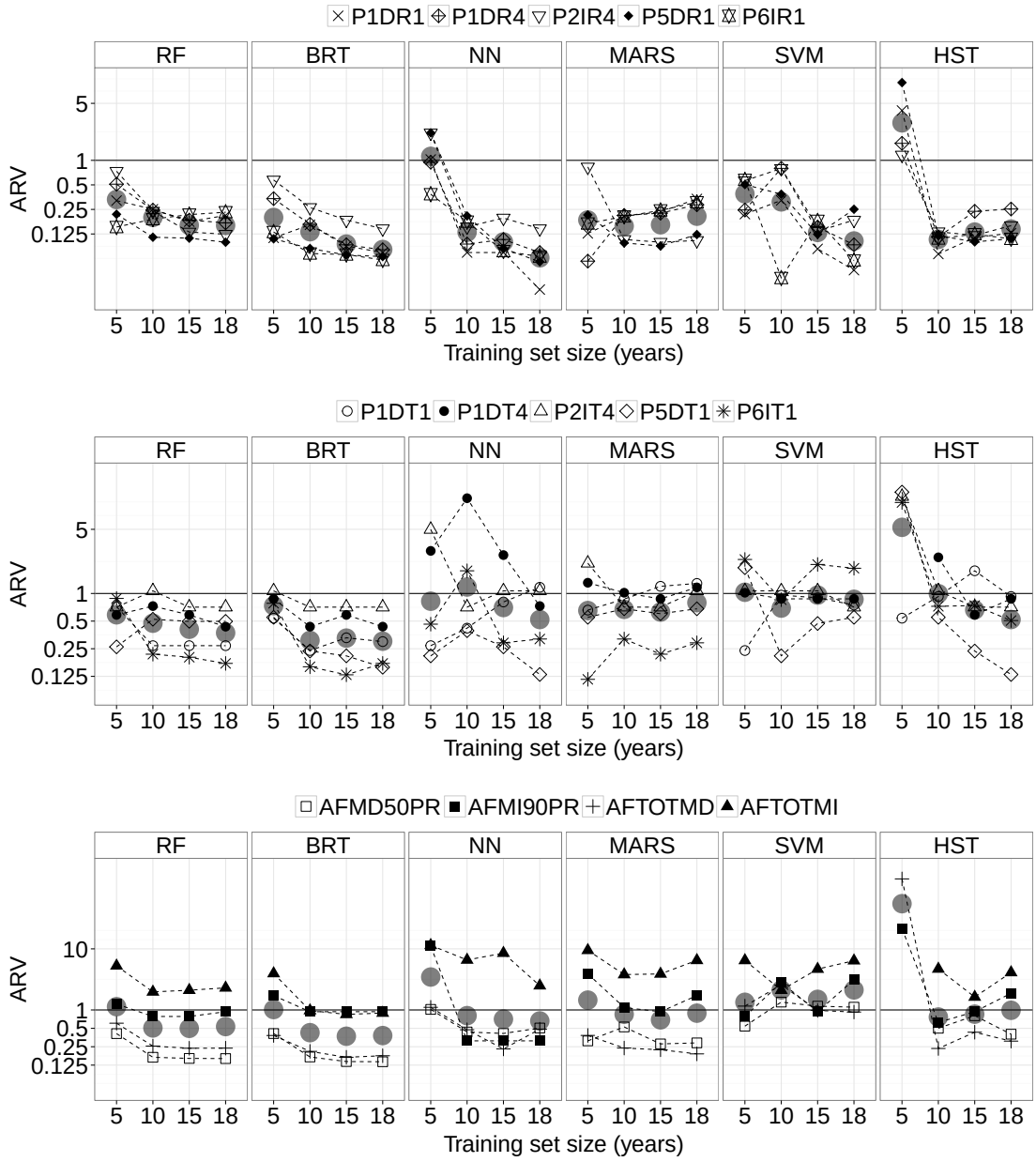


Figure 5: ARV for each model and training set size. Top: radial displacements. Middle: tangential displacements. Bottom: leakage flow. Some HST models trained over 5 years are out of the range of the vertical axis, thus highly inaccurate. The results correspond to the test set.

Surprisingly, for two of the outputs, the lower MAE corresponds to a model trained over five years, which in principle was assumed to be too small a training set. MARS is especially sensitive to the size of the training data. The MARS models trained on five years improve the accuracy for P1DR4 and P6IT1 by 13.3 % and 14.8 % respectively.

These results strongly suggest that it is advisable to select carefully the most appropriate training set size. This should be done by leaving an independent validation set.

5. Summary and conclusions

It was found that the accuracy of currently applied methods for predicting dam behaviour can be substantially improved by using ML techniques.

The sensitivity analysis to the training set size shows that removing the early years of dam life cycle can be beneficial. In this work, it has resulted in a decrease in MAE in some cases (up to 14.8%). Hence, the size of the training set should be considered as an extra parameter to be optimised during training.

Some of the techniques analysed (MARS, SVM, NN) are more susceptible to further tuning than others (RF, BRT), given that they have more hyper-parameters. As a consequence, the former might have a larger margin for improvement than the latter.

A more careful selection of variables could also improve the fit. It should be noted, though, that variable selection is an issue in itself, and will be the subject of further work. It may not only decrease the error, but also help to build more understandable models.

However, both detailed tuning and careful variable selection increase the computational cost and complicate the analysis. If the objective is the extension of these techniques for the prediction of a large number of variables of many dams, the simplicity of implementation is an aspect to be considered in model selection.

In this sense, BRT showed to be the best choice: it was the most accurate for five of the 14 targets; easy to implement; robust with respect to the training set size; able to consider any kind of input (numeric, categorical or discrete), and not sensitive to noisy and low relevant predictors.

However, none of the algorithms provides the highest accuracy in all cases. Therefore, if the main objective is to achieve the best possible fit, the analysis should not be limited to a single technique.

6. Acknowledgements

The authors thank the Catalan Water Agency, owner of La Baells dam, and the company Ofiteco for providing the monitoring data.
The research has been partially supported by the Spanish Ministry of Economy and Competitiveness (*Ministerio de Economía y Competitividad*, MINECO) through the project iComplex (IPT-2012-0813-390000).

[1] Swiss Committee on Dams, Methods of analysis for the prediction and the verification of dam behaviour, Tech. rep., ICOLD (2003).

[2] L. Chouinard, V. Roy, Performance of statistical models for dam monitoring data, in: Joint International Conference on Computing and Decision Making in Civil and Building Engineering, Montreal, June, 2006, pp. 199–207.

[3] A. Simon, M. Royer, F. Mauris, J. Fabre, Analysis and interpretation of dam measurements using artificial neural networks, in: 9th ICOLD European Club Symposium, Venice, Italy, 2013.

[4] G. Lombardi, F. Amberg, G. Darbre, Algorithm for he prediction of functional delays in the behaviour of concrete dams, Hydropower and Dams (3) (2008) 111–116.

[5] International Commission on Large Dams, Dam surveillance guide, Tech. Rep. B-158, ICOLD (2012).

[6] F. Restelli, Systemic evaluation of the response of large dams instrumentation. Application at El Chocón Dam, in: 9th ICOLD European Club Symposium, Venice, Italy, 2013.

[7] M. Tatin, M. Briffaut, F. Dufour, A. Simon, J. Fabre, Thermal displacements of concrete dams: Finite element and statistical modelling, in: 9th ICOLD European Club Symposium, Venice, Italy, 2013.

[8] S. Bonelli, K. Radzicki, Impulse response function analysis of pore pressure in earth-dams, European Journal of Environmental and Civil Engineering 12 (3) (2008) 243–262.

[9] F. Li, Z. Wang, G. Liu, Towards an error correction model for dam monitoring data analysis based on cointegration theory, Structural Safety 43 (2013) 1220. doi:10.1016/j.strusafe.2013.02.005.

[10] F. Riquelme, J. Fraile, D. Santillan, R. Moran, M. Toledo, Application of artificial neural network models to determine movements in an arch dam, in: 2nd International Congress on Dam Maintenance and Rehabilitation, Zaragoza, Spain, 2011, pp. 117–123.

[11] D. Santillán, J. Fraile-Ardanuy, M. Toledo, Seepage prediction in arch dams by means of artificial neural networks, Water Technology and Science V (3), accepted for publication. [In Spanish].

[12] J. Mata, Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models, Engineering Structures 3 (3) (2011) 903 – 910. doi:10.1016/j.engstruct.2010.12.011.

[13] C. M. Bishop, Neural networks for pattern recognition, Oxford University Press, 1995.

[14] V. Ranković, N. Grujović, D. Divac, N. Milivojević, A. Novaković, Modelling of dam behaviour based on neuro-fuzzy identification, Engineering Structures 35 (2012) 107–113. doi:10.1016/j.engstruct.2011.11.011.

[15] L. Piroddi, W. Spinelli, Long-range nonlinear prediction: a case study, in: 42nd IEEE Conference on Decision and Control, Vol. 4, IEEE, 2003, pp. 3984–3989.

[16] V. Saouma, E. Hansen, B. Rajagopalan, Statistical and 3d nonlinear finite element analysis of Schlegeis dam, Tech. rep., University of Colorado (2001).

[17] M. Herrera, L. Torgo, J. Izquierdo, R. Pérez-García, Predictive models for forecasting hourly urban water demand, Journal of Hydrology 387 (1) (2010) 141–150.

[18] G. Lombardi, Advanced data interpretation for diagnosis of concrete dams, Tech. rep., CISM (2004).

[19] A. De Sortis, P. Paoliani, Statistical analysis and structural identification in concrete dam monitoring, Engineering Structures 29 (1) (2007) 110–120.

[20] L. Seifard, A. Szpilman, C. Piasentin, Itaipú structures. Evaluation of their performance, in: 15th ICOLD Congress, 1985, pp. 287–317, Q56-R15.

[21] Q. Guedes, P. Coelho, Statistical behaviour model of dams, in: 15th ICOLD Congres, 1985, pp. 319–334, Q56-R16.

[22] F. J. Sánchez Caro, Dam safety: contributions to the deformation analysis and monitoring as an element of prevention of pathologies of geotechnical origin, Ph.D. thesis, UPM, [In Spanish] (2007).

[23] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2013).
URL http://www.R-project.org/

[24] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, Classification and regression trees, Wadsworth & Brooks, Monterrey, CA, 1984.

[25] L. Breiman, Random forests, Machine learning 45 (1) (2001) 05–32.

[26] R. Genuer, J. Poggi, C. Tuleau-Malot, Variable selection using random forests, Pattern Recognition Letters 31 (14) (2010) 225 – 236. doi:10.1016/j.patrec.2010.03.014.

[27] A. Liaw, M. Wiener, Classification and regression by randomForest, R news 2 (3) (2002) 18–22.

[28] J. Friedman, Greedy function approximation: a gradient boosting machine, Annals of Statistics (2001) 1189 – 1232.

[29] G. Ridgeway, Generalized Boosted Models: A guide to the gbm package, R package vignette (2007).
URL http://CRAN.R-project.org/package=gbm

[30] W. N. Venables, B. D. Ripley, Modern Applied Statistics with S, 4th Edition, Springer, New York, 2002, iSBN 0-387-95457-0.

[31] T. Hastie, R. Tibshirani, J. Firedman, The Elements of Statistical Learning - Data Mining, Inference, and Prediction, 2nd Edition, Springer, New York, 2009.

[32] M. Kuhn, Building predictive models in R using the caret package, Journal of Statistical Software 28 (5) (2008) 1–26.
URL http://www.jstatsoft.org/v28/i05/paper

[33] A. J. Smola, B. Schlkopf, A tutorial on support vector regression, Statistics and computing 14 (3) (2004) 199–222.

[34] J. M. Moguerza, A. Muñoz, Support vector machines with applications, Statistical Science (2006) 322–336.

[35] V. Ranković, N. Grujović, D. Divac, N. Milivojević, Development of support vector regression identification model for prediction of dam structural behaviour, Structural Safety 48 (2014) 33–39.

[36] D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, F. Leisch, e1071: Misc Functions of the Department of Statistics (e1071), TU Wien, r package version 1.6-1 (2012).

[37] B.-J. Chen, M.-W. Chang, et al., Load forecasting using support vector machines: A study on eunite competition 2001, IEEE Transactions on Power Systems 19 (4) (2004) 1821–1830.

[38] J. H. Friedman, Multivariate adaptive regression splines, The Annals of Statistics (1991) 1–67.

[39] S. Milborrow., earth: Multivariate Adaptive Regression Spline Models, R package version 3.2-6 (2013).
URL `http://CRAN.R-project.org/package=earth`

[40] L. Breiman, J. Friedman, R. Olshen, C. Stone, D. Steinberg, P. Colla, Classification and Regression Trees, Wadsworth: Belmont, CA, 1984.

[41] A. S. Weigend, B. A. Huberman, D. E. Rumelhart, Predicting sunspots and exchange rates with connectionist networks, in: S. Eubank, M. Casdagli (Eds.), Proc. of the 1990 NATO Workshop on Nonlinear Modeling and Forecasting (Santa Fe, NM), Vol. 12, Addison-Wesley, Redwood, CA, 1992, pp. 395–432.

A.4 Interpretation of dam deformation and leakage with boosted regression trees

Title: Interpretation of dam deformation and leakage with boosted regression trees

First Author: Fernando Salazar González. CIMNE - International Center for Numerical Methods in Engineering

Second Author: Miguel Á. Toledo Municio. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Third Author: Eugenio Oñate Ibáñez de Navarra. CIMNE - International Center for Numerical Methods in Engineering

Fourth Author: Benjamín Suárez Arroyo. CIMNE - International Center for Numerical Methods in Engineering

Journal: Engineering Structures

D.O.I. 10.1016/j.engstruct.2016.04.012

Impact Factor 1.893

Interpretation of dam deformation and leakage with boosted regression trees

Fernando Salazar[1], Miguel Á. Toledo[2], Eugenio Oñate[1], Benjamín Suárez[1]

Abstract

Predictive models are essential in dam safety assessment. They have been traditionally based on simple statistical tools such as the hydrostatic-season-time (HST) model. These tools are well known to have limitations in terms of accuracy and reliability. In the recent years, the examples of application of machine learning and related techniques are becoming more frequent as an alternative to HST. While they proved to feature higher flexibility and prediction accuracy, they are also more difficult to interpret. As a consequence, the vast majority of the research is limited to prediction accuracy estimation. In this work, one of the most popular machine learning techniques (boosted regression trees), was applied to model 8 radial displacements and 4 leakage flows at La Baells Dam. The possibilities of model interpretation were explored: the relative influence of each predictor was computed, and the partial dependence plots were obtained. Both results were analysed to draw conclusions on dam response to environmental variables, and its evolution over time. The results show that this technique can efficiently identify dam performance changes with higher flexibility and reliability than simple regression models.

Keywords:
machine learning, dam safety, dam monitoring, boosted regression trees

1. Introduction

Dam monitoring is essential to ensure its proper operation and its long-term safety [1]. One of the main tasks to be carried out is the comparison between the expected response and that registered by the monitoring system, to understand the dam behaviour and to detect potential anomalies. In this context, predictive models are necessary to estimate the dam response in a given situation.

Data-based tools allow building predictive models based on monitoring data, i.e., without explicitly considering the physical properties of the dam and the foundation. The hydrostatic-season-time (HST) model [2] is the most widely applied, and the only generally accepted by practitioners.

*Corresponding author: F. Salazar

Email addresses: fsalazar@cimne.upc.edu (Fernando Salazar), matoledo@caminos.upm.es (Miguel Á. Toledo), onate@cimne.upc.edu (Eugenio Oñate), benjamin.suarez@upc.edu (Benjamín Suárez)

[1]International Center for Numerical Methods in Engineering (CIMNE). Campus Norte UPC. Gran Capitán s/n. 08034. Barcelona, Spain

[2]Technical University of Madrid (UPM). Civil Engineering Department: Hydraulics, Energy and Environment. Profesor Aranguren s/n, 28040, Madrid, Spain

HST is based on multiple linear regression considering the three most influential external variables: hydrostatic load, air temperature and time. The main advantages of HST are:

1. It frequently provides useful estimations of displacements in concrete dams [3].
2. It is simple and thus easily interpretable: the effect of each external variable can be isolated in a straightforward manner, since they are cumulative.
3. Since the thermal effect is considered as a periodic function, the time series of air temperature are not required. This widens the possibilities of application, as only the reservoir level variation is needs to be available to build an HST model.
4. It is well known by practitioners and frequently applied in several countries [3].

Nonetheless, HST also features conceptual limitations that damage the prediction accuracy [3] and may lead to misinterpretation of the results [4]. For example, it is based on the assumption that the hydrostatic load and the temperature are independent, whereas it is obviously not the case: the thermal field in the dam body, especially in the vicinity of the water surface, is strongly dependant on the water temperature in the upstream face [5]. In turn, the thermal load influences the stress and displacement fields.

Several modifications to the original HST model have been proposed to overcome these drawbacks. They focus on improving the consideration of the thermal load, by taking into account the actual air temperature instead of the historical mean [6], or the effect of the water temperature on the upstream face [3], [7].

In the recent years, non-parametric techniques have emerged as an alternative to HST for building data-based behaviour models [8], e.g. support vector machines (SVN) [9], neural networks (NN) [10], adaptive neuro-fuzzy systems (ANFIS) [11], among others [8]. In general, these tools are more suitable to model non-linear cause-effect relations, as well as interaction among external variables, as that previously mentioned between hydrostatic load and temperature. On the contrary, they are typically more difficult to interpret, what led them to be termed as "black box" models (e.g. [12]).

Most of the published works focused on building predictive models whose accuracy was generally higher than that offered by HST (e.g. [10], [13], [14]). Since the resulting model was seldom analysed, little information was provided for dam safety assessment. Some exceptions worth mentioning, though simple, were due to Santillán *et al.* [15], Mata [10] and Cheng and Zheng [16].

Therefore, dam engineers face a dilemma: the HST model is widely known and used and easily interpretable. However, it is based on some incorrect assumptions, and its accuracy can be increased. On the other hand, more flexible and accurate models are available, but they are more difficult to implement and analyse. The same problem arose in the field of statistics [17].

The objective of this work is to investigate the possibilities of interpretation of one of these black box models to:

1. Identify the effect of each external variable on the dam behaviour
2. Detect the temporal evolution of the dam response
3. Provide meaningful information to draw conclusions about dam safety

Among the plethora of machine learning techniques available [18], a previous comparative study [13] showed boosted regression trees (BRT) as one of the more appropriate tools for

the prediction of dam response. In this paper, the technique was further explored, with focus on the interpretation of the results for dam behaviour identification. In particular, the partial dependence plots were examined to isolate the effect of each action, and the relative influence (RI) was computed to identify the strength of each input-output relation. Furthermore, the results were interpreted from an overall viewpoint to draw conclusions on the dam behaviour.

The method was applied to the analysis of La Baells Dam, as compared to the conventional HST model.

The rest of the paper is organised as follows. A brief introduction to BRT is presented, including the methods for interpretation. Then, the case study and the HST version taken as reference are described. The results are included and interpreted in terms of the dam behaviour, and the differences between both methods are discussed.

2. Methods

2.1. Boosted regression trees

The objective of a predictive model is to estimate the value of an output variable $Y \in \mathbb{R}$ (i.e. radial displacement or leakage), based on a set of predictors (reservoir level, air temperature, etc.) $X \in \mathbb{R}^p$, i.e. $Y \approx \hat{Y} = F(X)$. The observed values are denoted as $(x_i, y_i), i = 1, ..., N$, where N is the number of observations. Note that each x_i is a vector with p components, each of which is referred to as x_i^j, when necessary. Similarly, $X^j, j = 1, ..., p$ stands for each dimension of the input space.

BRT models are built by combining two algorithms: a set of single models are fitted by means of decision trees [19], and their output is combined to compute the overall prediction using boosting [20]. For the sake of completeness, a short description of both techniques follow, although excellent introductions can be found in [21], [22], [23], [12].

2.1.1. Regression trees

Regression trees were first proposed as statistical models by Breiman *et al.* [19]. They are based on the recursive division of the training data in groups of "similar" cases. The output of a regression tree is the mean of the output variable for the observations within each group.

When more than one predictor is considered (as usual), the best split point for each is computed, and the one which results in greater error reduction is selected. As a consequence, non-relevant predictors are automatically discarded by the algorithm, as the error reduction for a split in a low relevant predictor will generally be lower than that in an informative one.

Other interesting properties of regression trees are:

- They are robust against outliers.
- They require little data pre-processing.
- They can handle numerical and categorical predictors.
- They are appropriate to model non-lineal relations, as well as interaction among predictors.

By contrast, regression trees are unstable, i. e., small variations in the training data lead to notably different results. Also, they are not appropriate for certain input-output relations, such as a straight 45º line [23].

2.1.2. Boosting

Boosting is a general scheme to build ensemble prediction models [20]. It is based on the generation of a (frequently high) number of simple models (also referred to as "weak learners") on altered versions of the training data. The overall prediction is computed as a weighted sum of the output of each model in the ensemble. The rationale behind the method is that the average of the prediction of many simple learners can outperform that from a complex one [24].

The idea is to fit each learner to the residual of the previous ensemble. The main steps of the original boosting algorithm when using regression trees and the squared-error loss function can be summarised as follows [25]:

1. Start predicting with the average of the observations (constant):

$$F_0(X) = f_0(X) = \bar{y}_i$$

2. For $m = 1$ to M
   (a) Compute the prediction error on the training set:

$$\tilde{y}_i = y_i - F_{m-1}(x_i)$$

   (b) Draw a random sub-sample of the training set (S_m)
   (c) Consider S_m and fit a new regression tree to the residuals of the previous ensemble:

$$\tilde{y}_i \approx f_m(X), i \in S_m$$

   (d) Update the ensemble:

$$F_m(X) \Leftarrow F_{m-1}(X) + f_m(X)$$

3. F_M is the final model

It is generally accepted that this procedure is prone to over-fitting, because the training error decreases with each iteration [25]. To overcome this problem, it is convenient to add a regularization parameter $\nu \in (0, 1)$, so that step (d) turns into:

$$F_m(X) \Leftarrow F_{m-1}(X) + \nu \cdot f_m(X)$$

Some empirical analyses showed that relatively low values of ν (below 0.1) greatly improve generalisation capability [20]. In practice, it is common to set the regularisation parameter and consider a number of trees such that the training error stabilises [21]. Subsequently, a certain number of terms are pruned using for example cross-validation. This is the approach employed in this work, with $\nu = 0.001$ and a maximum of 10,000 trees. It was verified that the training error reached the minimum before adding the maximum number of trees.

Five-fold cross-validation was applied to determine the amount of trees in the final ensemble. The process was repeated using trees of depth 1 and 2 (*interaction.depth*), and the most accurate for each target was selected. The rest of the parameters were set to their default values [26].

All the calculations were performed in the R environment [27].

2.2. Model interpretation

2.2.1. Relative influence (RI)

BRT models are robust against the presence of uninformative predictors, as they are discarded during the selection of the best split. Moreover, it seems reasonable to think that the most relevant predictors are more frequently selected during training. In other words, the relative influence (RI) of each input is proportional to the frequency with which they appear in the ensemble. Friedman [20] proposed a formulation to compute a measure of RI for BRT models based on this intuition. Both the relative presence and the error reduction achieved are considered in the computation. The results are normalised so that they add up to 100.

Based on this measurement, the most influential variables were identified for each output, and the results were interpreted in relation to dam behaviour. In order to facilitate the analysis, the RI was plotted as word clouds [28]. These plots resemble histograms, with the advantage of being more appropriate to visualise a greater set of variables. The code representing each predictor was displayed with a font size proportional to its relative influence with the library "wordcloud" [29].

Furthermore, two degrees of variable selection were applied, based on the RI of each predictor. First, a BRT model (M1) was trained with all the variables considered (section 2.4). Second, the inputs with $RI\left(X^j\right) > min\left(RI\left(X^j\right)\right) + sd\left(RI\left(X^j\right)\right)$ were selected to build a new model (M2). This criteria is heuristic and based on the *1-SE rule* proposed by Breiman *et al.* [19]. Finally, a model with three predictors was generated (M3), featuring the more relevant variables of each group: temperature, time and reservoir level for radial displacements, and rainfall, time and level for leakage flows.

These three versions were generated to analyse the effect of the presence of uninformative variables in the predictor set. Moreover, the simplest model facilitates the analysis, as the effect of each action is concentrated in one single predictor.

In this sense, the temporal evolution is particularly relevant for dam safety evaluation, as it can help to identify a progressive deterioration of the dam or the foundation, which might result in a serious fault if not corrected.

2.2.2. Partial dependence plots

Multi-linear regression models and HST in particular are based on the assumption that the input variables are statistically independent, so the prediction is computed as the sum of their contributions. As a result, the effect of each predictor in the response can be easily identified, by plotting $f(X^j), \forall j = 1...p$.

This method is not appropriate for BRT models, as interactions among predictors are accounted for. While this results in more flexibility, it also implies that the identification of the relation between predictors and response is not straightforward.

Nonetheless, it is possible to examine the predictor-response relationship by means of the partial dependence plots [20]. This tool can be applied to any black box model, as it is based on the marginal effect of each predictor on the output, as learned by the model. Let X^j be the variable of interest. A set of equally spaced values are defined along its range: $X^j = x_k^j$. For each of those values, the average of the model predictions is computed:

$$\bar{F}\left(x_k^j\right) = \frac{1}{N}\sum_{i=1}^{N} F\left(x_k^j, x_i^{jc}\right) \tag{1}$$

where x_i^{jc} is the value for all inputs other than X^j for the observation i.

Similar plots can be obtained for interactions among inputs: the average prediction is computed for couples of fixed x_k^j, where j takes two different values. Hence, the results can be plotted as a three-dimensional surface (section 3.3). In this work, partial dependence plots were restricted to the simplest model, which considered three predictors. Therefore, three 3D plots allowed investigating the pairwise interactions for all the inputs considered in the simplified model.

2.3. HST model

A conventional HST model was fitted for comparison purposes:

$$\begin{aligned} \hat{Y} = F(t,h,s) &= a_0 + a_1 h + a_2 h^2 + a_3 h^3 \\ &\quad + a_4 h^4 + a_5 h^5 + a_6 e^{-t} \\ &\quad + a_7 t + a_8 cos(s) + a_9 sin(s) \\ &\quad + a_{10} sin^2(s) + a_{11} sin(s) cos(s) \end{aligned} \tag{2}$$

where

$$s = \frac{d}{365,25} 2\pi \tag{3}$$

where d is the number of days since 1 January, t is the elapsed time (years), h is the reservoir level, and $a_1, a_2, ..., a_{11}$ are the coefficients to fit.

The contribution of each action can be computed by adding the correspondent terms:

$$\hat{Y}_h = a_1 h + a_2 h^2 + a_3 h^3 + a_4 h^4 + a_5 h^5 \tag{4}$$

$$\begin{aligned} \hat{Y}_s &= a_8 cos(s) + a_9 sin(s) \\ &\quad + a_{10} sin^2(s) + a_{11} sin(s) cos(s) \end{aligned} \tag{5}$$

$$\hat{Y}_t = a_5 e^{-t} + a_6 t \tag{6}$$

This model was also employed to check the reliability of the temporal behaviour identified by BRT models for some devices. After an HST model was fitted to the training data, a modified version of the time series of the target variable was generated by removing the temporal term ($\hat{Y}_t$) and adding random noise of zero mean and a standard deviation equal to 0.5 (mm for displacements; l/min for leakage):

$$Y_{mod} = \hat{Y}_h + \hat{Y}_s + N(0, 0.5) \tag{7}$$

The result is a time series whose dependence from the temperature and the level approximates that of the actual displacement, while being totally time-independent.

2.4. Case study

The data used for the study correspond to La Baells Dam. It is a double curvature arch dam, with a height of 102 m, which entered into service in 1976. Among the available records, the study focused on 12 variables: 8 corresponded to radial displacements measured

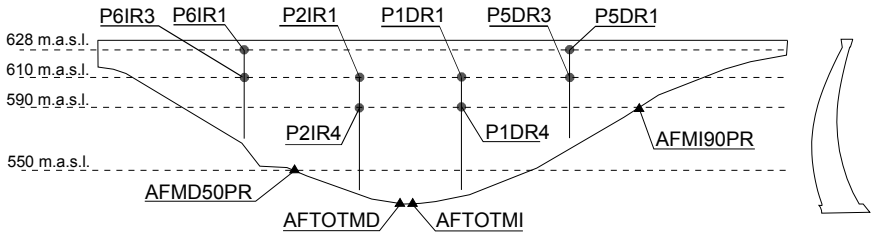


Figure 1: Geometry and location of the monitoring devices in La Baells Dam. Left: view from downstream. Right: highest cross-section.

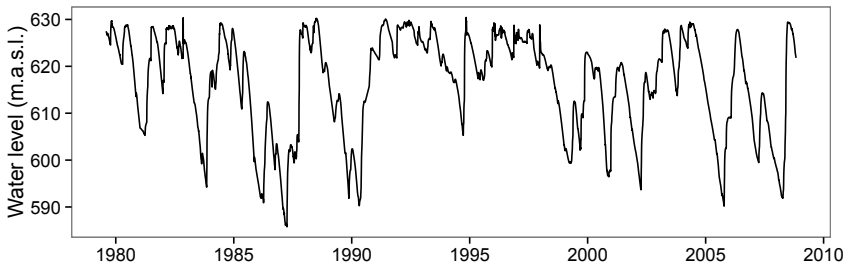


Figure 2: Time series of the reservoir level at La Baells Dam.

by pendulums (along the upstream-downstream direction), and four to leakage flow. The location of each monitoring device is depicted in Figure 1.

As for the environmental variables, the mean daily air temperature, the reservoir level and the daily rainfall were available. Figure 2 depicts the reservoir level variation in the period considered, whereas the other two are included in the Appendix (Figures A1 and A2).

Since BRT models automatically discard those predictors not associated with the output [30], the initial model considered a relatively large set of inputs. The objective was to test that property (by introducing obviously unimportant predictors), as well as to explore the rate of influence of several variables whose relevance was not so obvious (e. g. the rate of variation of the reservoir level). The complete list of predictors is included in table 1.

All the calculations were performed on a training set covering the period 1980-1997, when weekly records were available. The relative influence and the partial dependence were computed with this data set. The model accuracy was assessed for a validation set covering the period 1998-2008.

The goodness of fit was computed in terms of the mean absolute error (MAE):

$$MAE = \frac{\sum_{i=1}^{N} |y_i - F(x_i)|}{N} \tag{8}$$

where N is the size of the training (or validation) set, y_i are the observed outputs and $F(x_i)$ the predicted values. As MAE is measured in the same units as the variable to predict, it is an intuitive measure of accuracy. However, it is not appropriate to compare models for different targets, as it does not account for the standard deviation of the output.

To overcome this limitation, the average relative variance (ARV) [31] was also computed:

Table 1: Predictor variables considered for the initial BRT model (M1).

Code	Group	Type	Period (days)
Level	Hydrostatic load	Original	-
Lev007	Hydrostatic load	Moving average	7
Lev014			14
Lev030			30
Lev060			60
Lev090			90
Lev180			180
Tair	Air temperature	Moving average	1
Tair007			7
Tair014			14
Tair030			30
Tair060			60
Tair090			90
Tair180			180
Rain	Rainfall	Accumulated	1
Rain030			30
Rain060			60
Rain090			90
Rain180			180
NDay	Time	Original	-
Year			-
Month	Season	Original	-
n010	Hydrostatic load	Rate of variation	10
n020			20
n030			30

$$ARV = \frac{\sum_{i=1}^{N}(y_i - F(x_i))^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} = \frac{MSE}{\sigma^2} \tag{9}$$

where $\bar{y}$ is the output mean. Given that ARV denotes the ratio between the mean squared error (MSE) and the variance (σ^2), it accounts both for the magnitude and the deviation of the target variable.

2.5. Overall procedure

For each target, the complete process comprised the following steps:

1. Fit a BRT model on the training data with the variables in table 1 (M1).
2. Compute the RI and generate the word cloud.
3. Select the most relevant predictors with the 1-SE rule (see section 2.2.1) and fit a new BRT model (M2).
4. Build a simple BRT model (M3) with the most influential variable of each group (temperature, level and time for displacements, and rainfall, level and time for leakage).
5. Generate the univariate and bivariate partial dependence plots for the simplest model.
6. Compute the goodness of fit for each model in both the training and the validation sets.

3. Results and discussion

3.1. Model accuracy

Although the work focused on model interpretation and its implications on dam safety, the goodness of fit was also checked in order to a) observe the effect of variable selection, and b) check the prediction accuracy of the model used for interpretation (M3).

Table 2 contains the error indices for each target, while more detailed results are included in the Appendix. For those models with variable selection, the predictors are also listed. The results show that BRT efficiently discarded irrelevant inputs, since the fitting accuracy was similar for each version in most cases (i.e., the presence of uninformative predictors did not damage the fitting accuracy).

The residuals were higher for the validation period, what reveals some degree of overfitting. A probable reason is that time was considered as any other predictor, and thus extrapolation over time was required to calculate the response in a more recent period. It is well known that non-parametric models lose much of their accuracy when predictions are made outside the range of variation of the input variables [32]. The increase in prediction error is greater for those targets for which time influence is more important, as is the case of the leakage in the left margin (up to four times larger). In these cases (AFMI90PR and AFTOTMI), the usefulness of the ARV is clearly observed: while the MAE is similar for the training and validation periods, the ARV is notably greater in the latter case, because the variance is lower in the most recent period (leakage flow decreased significantly over time).

Table 2: Accuracy of each model and target for the training and validation sets. The results and inputs considered by the most accurate version are highlighted in bold.

	Train		Validation		
Target	MAE	ARV	MAE	ARV	Inputs
	0,64	0,03	0,91	0,08	All
P1DR1	0,68	0,03	0,81	0,06	Tair090,Level,NDay,Lev007,Lev014
	0,69	**0,03**	**0,78**	**0,06**	**NDay,Tair090,Level**
	0,46	**0,03**	**0,65**	**0,08**	**All**
P1DR4	0,50	0,03	0,66	0,08	Level,Tair090,NDay,Lev007,Lev014,Lev030
	0,51	0,03	0,67	0,08	NDay,Tair090,Level
	0,66	0,03	1,03	0,09	All
P2IR1	0,85	0,05	1,09	0,09	Tair090,Level,Lev007,Lev014
	0,71	**0,04**	**0,98**	**0,08**	**NDay,Tair090,Level**
	0,48	**0,05**	**0,90**	**0,14**	**All**
P2IR4	0,61	0,06	0,93	0,14	Level,Tair090,Lev007,Lev014,Lev030
	0,53	0,06	0,94	0,16	NDay,Tair090,Level
	0,66	**0,05**	**0,82**	**0,08**	**All**
P5DR1	0,64	0,05	0,87	0,10	Tair060,Level,Tair030
	0,83	0,08	0,93	0,11	NDay,Tair060,Level
	0,25	**0,03**	**0,47**	**0,21**	**All**
P5DR3	0,33	0,05	0,55	0,22	Tair060,Level,Tair030
	0,31	0,04	0,52	0,24	NDay,Tair060,Level
	0,60	0,04	0,80	0,09	All
P6IR1	**0,65**	**0,05**	**0,78**	**0,08**	**Tair060,Tair030,Level,NDay**
	0,83	0,08	0,85	0,1	NDay,Tair060,Level
	0,23	**0,02**	**0,40**	**0,08**	**All**
P6IR3	0,37	0,05	0,67	0,17	Tair060,Level,Tair030
	0,29	0,03	0,43	0,09	NDay,Tair060,Level
	1,28	**0,16**	**0,93**	**0,19**	**All**
AFMD50PR	1,45	0,17	1,36	0,28	Level,Lev014,Lev007
	1,16	0,14	1,23	0,48	NDay,Rain090,Level
	0,08	0,09	0,15	0,51	All
AFMI90PR	0,08	0,10	0,12	0,45	Lev007,NDay,Level,Lev014,Lev030
	0,08	**0,10**	**0,12**	**0,46**	**NDay,Rain030,Lev007**
	1,64	**0,15**	**1,67**	**0,37**	**All**
AFTOTMD	1,87	0,19	1,73	0,45	Level,Lev007,Lev014
	1,69	0,18	1,97	0,52	NDay,Rain180,Level
	0,41	**0,11**	**0,44**	**0,40**	**All**
AFTOTMI	0,44	0,12	0,44	0,42	NDay,Lev060,Lev014,Lev007,Lev030,Lev180,Lev090,Level
	0,54	0,18	0,46	0,60	NDay,Rain180,Lev060

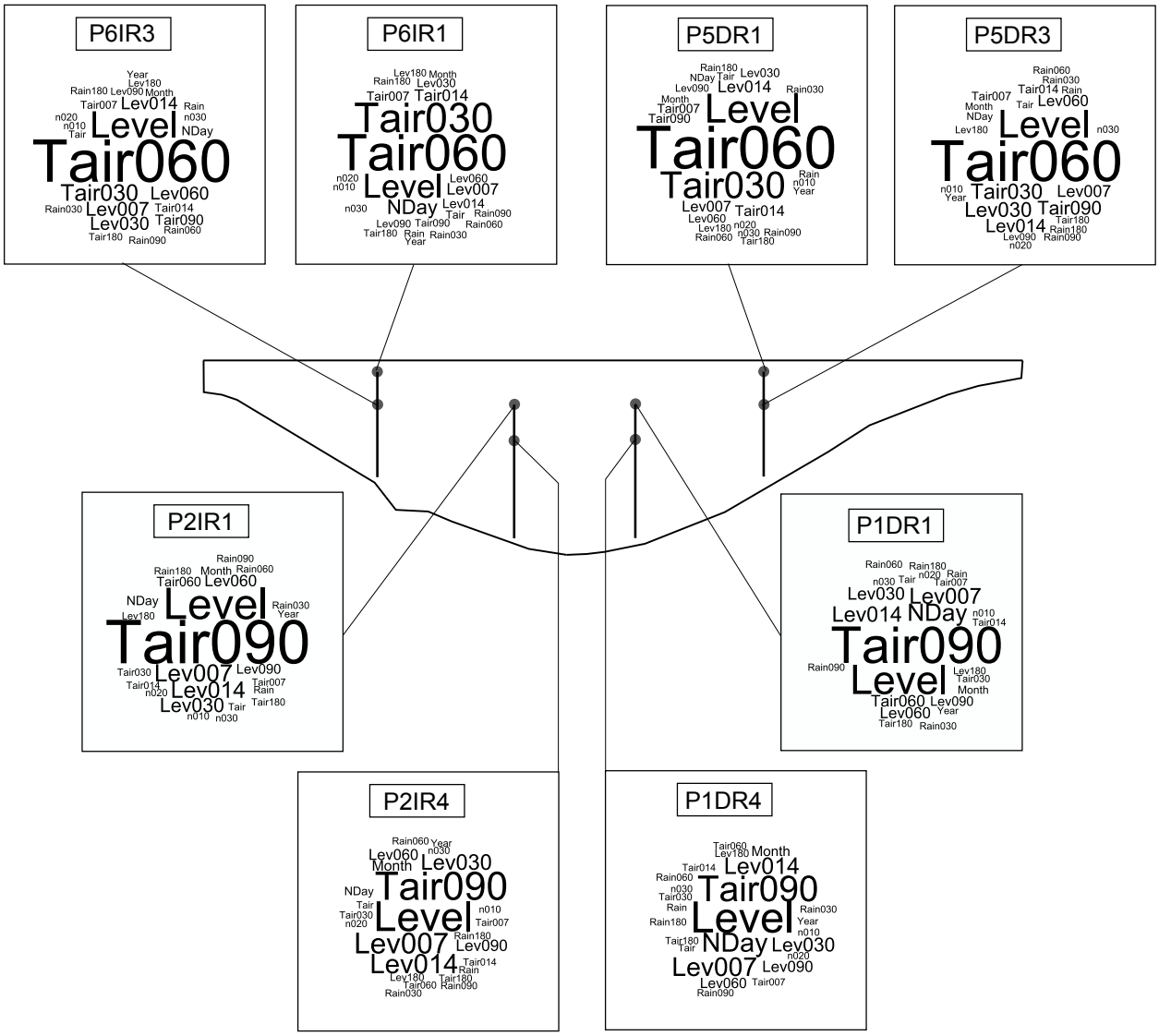


Figure 3: Word clouds for the radial displacements analysed.

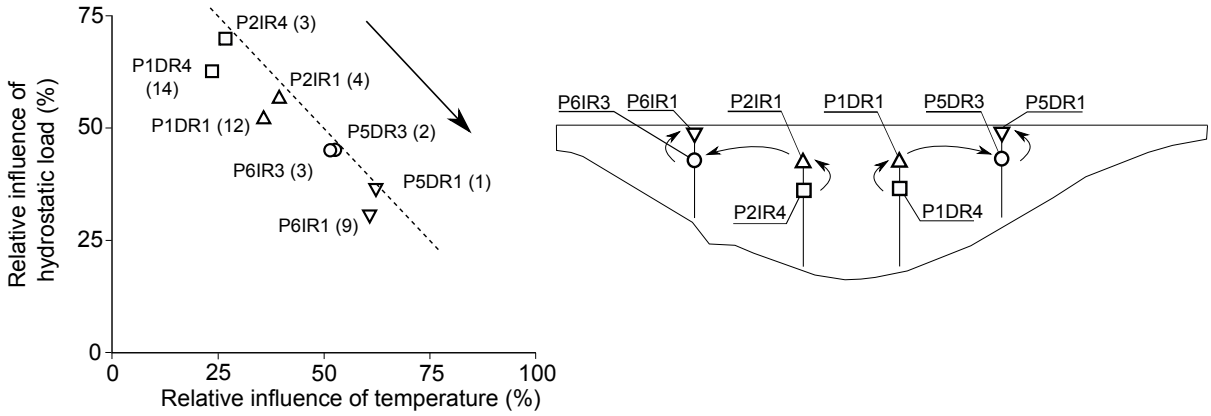


Figure 4: Relative influence of inputs in radial displacements, aggregated by type. The values in parenthesis correspond to the relative influence of time. Since the result is normalised, they sum 100 for each location. Hence, the distance to the $x + y = 100$ line (dashed) is proportional to the importance of the time effect. It should be noted that the devices in symmetrical locations with respect to the dam axis are grouped (i.e. P6IR3 and P5DR3). The arrows highlight the path of increasing influence of temperature in both plots, also symmetrical.

3.2. Variable importance

3.2.1. Radial displacements

Figure 3 depicts the RI of the predictors for each radial displacement considered. While Tair90 was the most relevant thermal input for the central sections (P1DR and P2IR), Tair060 took its place for those in the vicinity of the abutments (P5DR and P6IR). The higher thermal inertia of the central blocks might be due to their greater average thickness.

As for the hydrostatic load, the reservoir level at the date of the record was always more influential than all the moving averages, what reveals an immediate response of the dam to this load.

The RI of the rate of reservoir level variation was similar to that of rainfall, hence negligible.

From an overall viewpoint, a high degree of symmetry was observed, with the remarkable exception of the greater influence of NDay for P1DR1 and P1DR4 (Figure 3). This issue was further investigated by aggregating the relative influence of inputs by type: hydrostatic load, air temperature and time (Table 1). Figure 4 shows the result for each location considered. The symmetry is neatly observed, as well as the increasing RI of the temperature with respect to that of the hydrostatic load, from the foundation towards the crown, and from the centre to the abutments.

3.2.2. Leakage

The RI of the inputs for the leakage flows revealed a clear different behaviour between the right (AFMD50PR and AFTOTMD) and the left margins (AFMI90PR and AFTOTMI). While the former responded mainly to the hydrostatic load, with little inertia, the latter showed a remarkable dependence on time, as well as a greater relevance of several rolling means of reservoir level. Figure 5 shows the word clouds for the leakage flows.

The low inertia with respect to the hydrostatic load suggests that most of the leakage flow comes from the reservoir, while the effect of rainfall is negligible.

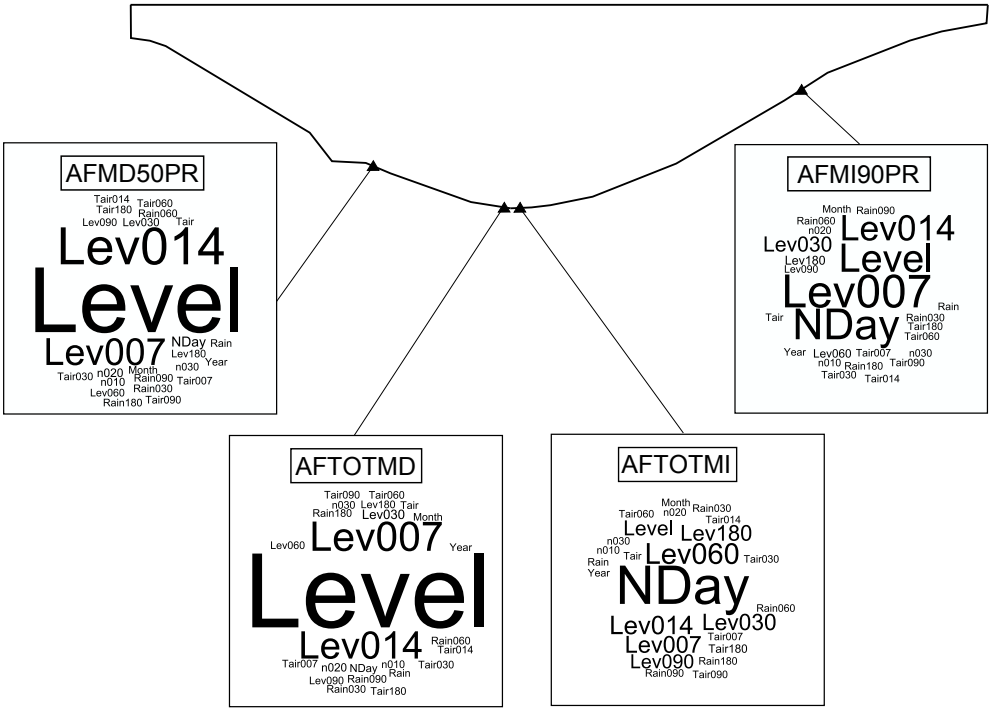


Figure 5: Word clouds for the leakage measurement locations analysed

Finally, it should be noted that temperature was irrelevant.

3.3. Partial dependence

3.3.1. Radial displacements

Figure 6 displays the univariate partial dependence plots for the radial displacements. The association between the air temperature and the reservoir level is intuitive: high temperature generates displacements towards upstream (positive direction in the vertical axis), due to concrete expansion, coerced by the abutments. The effect of the hydrostatic load is the opposite: high levels imply greater load and displacement towards downstream, and vice-versa.

These plots also show that the air temperature had greater influence than reservoir level for P5D and P6I (steeper curves with greater range of variation). The inverse effect was observed for P1DR4 and P2IR4, whereas both had similar relevance for P1DR1 and P2IR1. These results are coherent with the computed RI (Figures 3 and 4).

A deeper analysis was performed for P1DR1, which featured the greatest influence of time. The correspondent plot in Figure 6 shows a small step at the beginning of the period, followed by a sensibly constant behaviour until 1990, and a second larger step that stabilised afterwards. The bivariate plots for P1DR1 in Figure 7 show the step around 1991-1992 for the whole ranges of level and temperature.

This qualitative behaviour was observed for all displacements with lower magnitude, except for P5DR3. The effect was more clearly registered for P1DR1 and P1DR4, which in turn showed a greater influence of time (see Figure 3 and the Appendix).

Figure 8 shows the contribution of each action as obtained with the HST model. Unlike the previous case, the influence of time is linear, equivalent to a constant-rate displacement towards downstream.

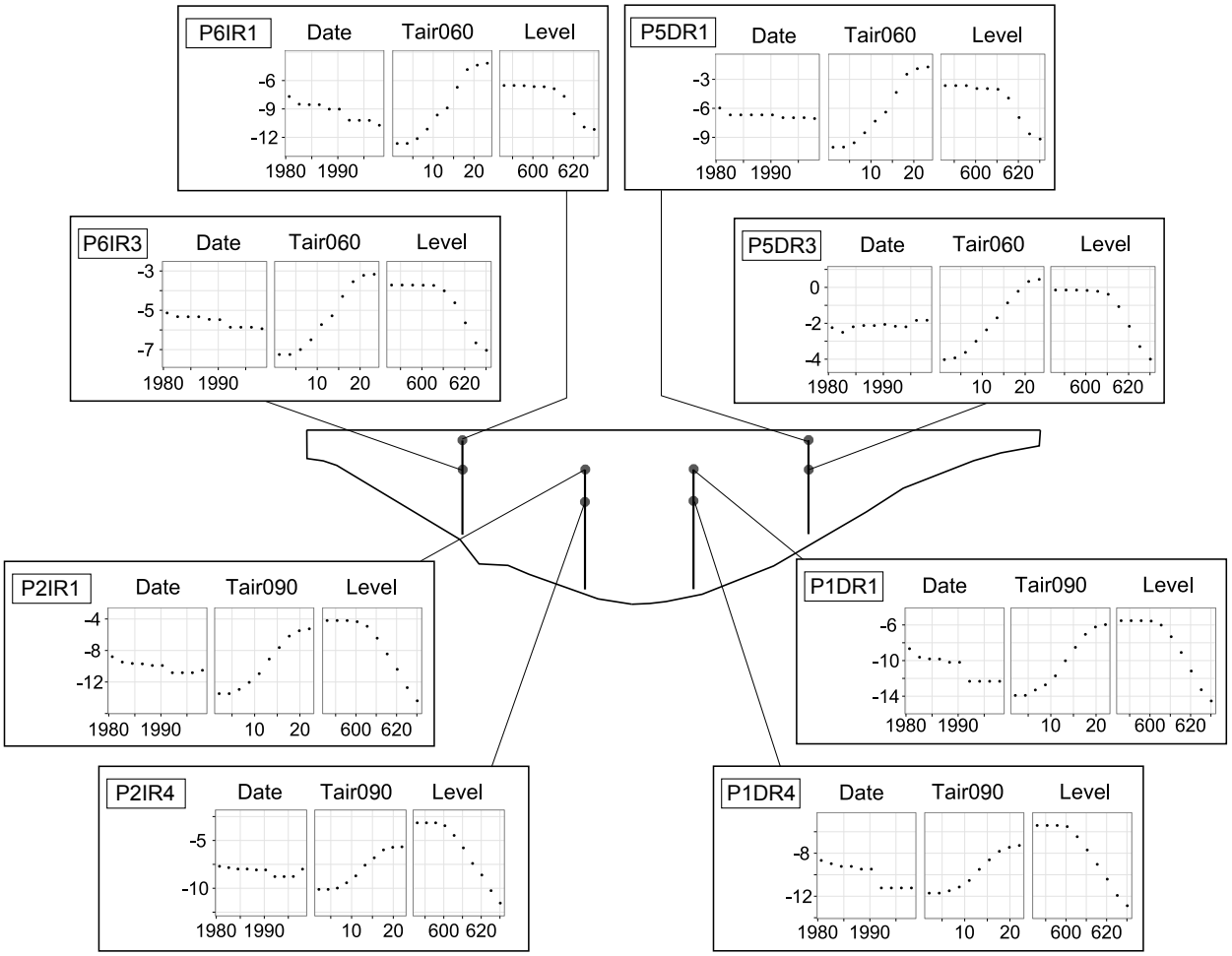


Figure 6: Partial dependence plot for the radial displacements analysed. Movement towards downstream correspond to lower values in the vertical axis, and vice-versa.

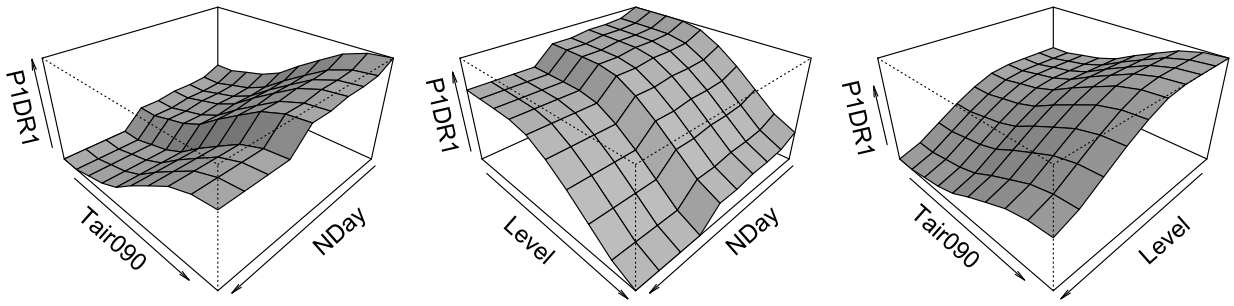


Figure 7: Interaction plots for P1DR1. It should be noted that the step along the temporal axis is observed for all the range of temperature and level.

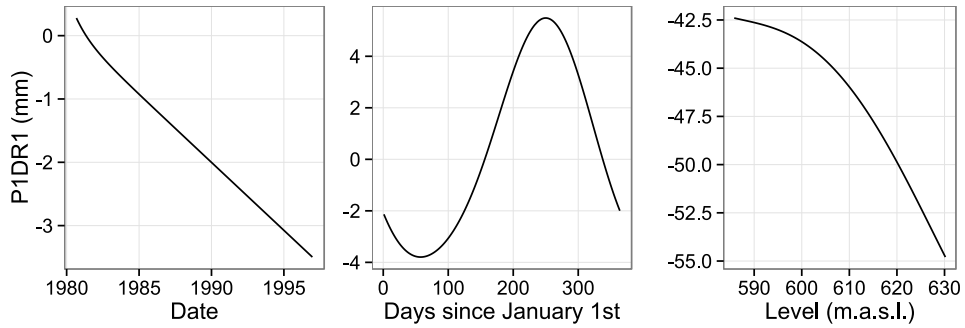


Figure 8: P1DR1. Contribution of the air temperature, the hydrostatic load and the time as drawn from the HST model.

This different model interpretation deserved a further verification. Not only because the results were substantially different, but also because the time effect is essential for the early detection of dam deterioration. In view of the temporal variation as captured by the BRT model, it could be concluded that some anomaly happened around 1991, which stabilised in the later years. On the contrary, the HST model interpreted a constant drift towards downstream of roughly 0.2 mm/year which might be serious in terms of the dam safety.

As mentioned above, the shape of the output-time dependency must be defined a priori for HST (in this case, a combination of exponential and negative linear functions was chosen), while in principle it can take any form for BRT. Therefore, it could be concluded that the actual behaviour of the dam was that showed by the BRT model, and that the result of the HST was due to the previously imposed restriction.

However, the average reservoir level in the period 1991-1997 was significantly higher than before 1991 (Figure 2), and might be the cause of the step registered in Figure 6: it represents a greater displacement towards downstream in the most recent period, which is consistent with the higher average hydrostatic load.

The verification was performed by fitting a new BRT model to the artificial data generated ($\hat{Y}_{mod}$) without time variation (eq. 7). It should be recalled that the artificial time series data maintains the original reservoir level variation, and thus the higher load in the 1991-1997 period. Figure 9 contains the partial dependence plot for this BRT model, which clearly shows that the independence of the artificial data with respect to time was correctly captured. This result confirms that the step in the time dependence captured by BRT is not a consequence of the higher hydrostatic load in 1991-1997.

As regards the HST model, it can be concluded that the linear trend is the best least squares fit that can be obtained to the observed behaviour (constant-step-constant) with a linear function. This might lead to a wrong interpretation of dam performance, not supported by the observed data.

It should be mentioned that more sophisticated versions of the HST model can be employed, and in particular a step can be considered, as Carrère and Noret-Duchêne showed in their analysis of the Schlegeiss Dam [33]. However, they only decided to try a step after observing a previous linear fit, where the sudden change in dam behaviour could be noted. In general, it can be difficult to identify a change in dam behaviour by simple data exploration, as is the case of La Baells Dam (see Figure A3). In this sense, the non-parametric nature of

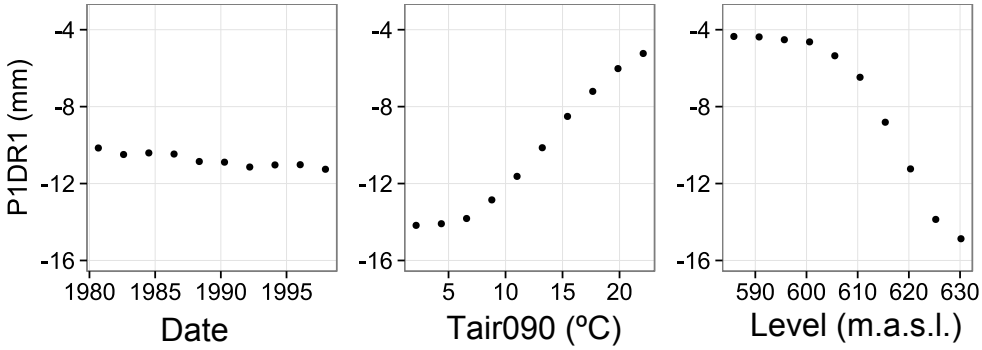


Figure 9: Partial dependence plot for the artificial time-independent data. P1DR1. It should be noted that time influence is negligible.

BRT models helps to identify performance changes of any type.

3.3.2. Leakage flows

Although the word clouds showed that neither rainfall nor temperature were influential on the leakage flow, partial dependence plots were generated as a further verification for the simplest model (M3 model; section 2.5).

Figure 10 contains the results, which confirm the conclusion of the word clouds: the time effect is irrelevant in the right abutment, except by certain erratic behaviour in the first two years and in the last three. On the contrary, a sharp decrease in leakage flow is revealed around 1983 for both locations in the left abutment. A lower decrease is observed in later years.

These results might be due to the colmatation of the cracking network in the left abutment, which would have led to lower permeability and leakage flow.

The shape of the effect of the hydrostatic load is sensibly exponential, with low influence for reservoir level below 610 m.a.s.l.

4. Summary and Conclusions

BRT models with different degree of variable selection were fitted to 8 radial displacements and 4 leakage flows at La Baells Dam. The relative influence of each input was computed and depicted via word clouds, which offered an efficient visualisation of the overall response of the dam. These graphs, together with the univariate and bivariate partial dependence plots, allowed interpretation of the BRT models: useful information regarding dam behaviour was obtained, such as the thermal inertia, the variation over time, and the performance of each area of the dam body.

The results showed a symmetrical behaviour of the dam in terms of displacements, as well as some interesting patterns, which will be the subject of future research:

- the thermal inertia was higher near the abutments.
- the RI of the temperature with respect to that of the hydrostatic load increased from the foundation towards the crown, and from the centre to the abutments.

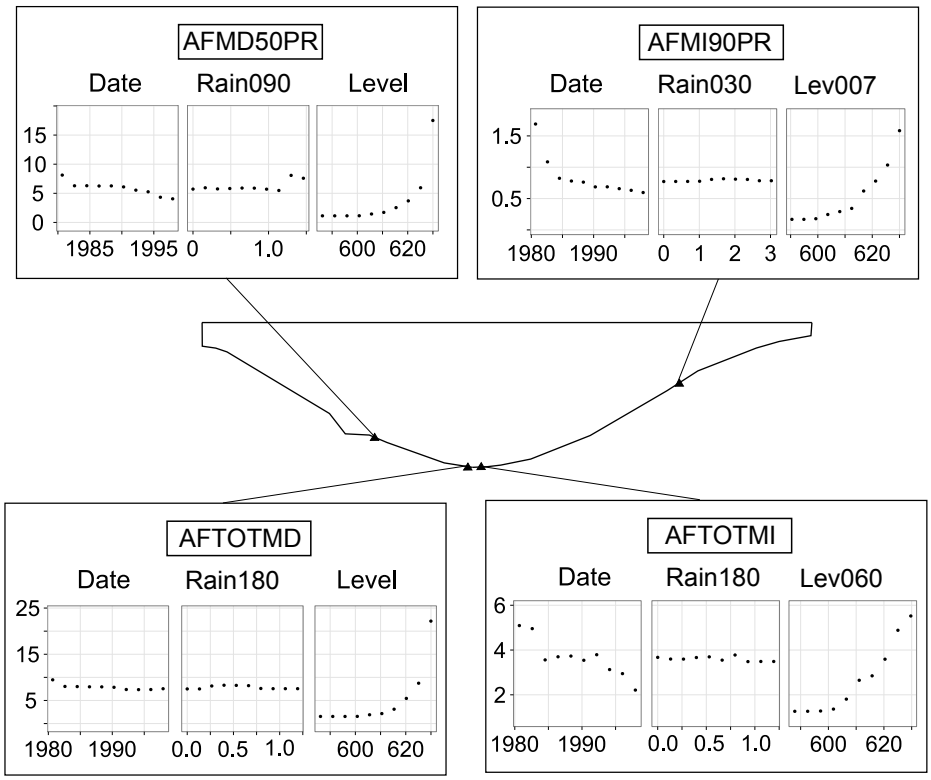


Figure 10: Partial dependence plot for leakage flows.

As regards the leakage flows, the different behaviour over time of each abutment was identified.

The amount of predictors considered in the BRT model did not significantly affect the prediction accuracy: the results confirm that the algorithm efficiently discard the less relevant inputs.

The application of BRT models to make predictions for a more recent period than that used for training involves extrapolation over time (provided that some time dependent predictor is considered). Hence, results should be analysed carefully, in particular if the time effect seems relevant. This applies to any data-based model considering time as input, including HST.

A sudden change in radial displacements was identified by the BRT model, especially for P1DR1. By contrast, the HST model suggested a constant-rate drift in the downstream direction. It was verified that the step towards downstream captured by the BRT model was not due to the higher average hydrostatic load actually registered for the 1991-1997 period. This suggests that partial dependence plots based on BRT models are more effective to identify performance changes, as they are not coerced by the shape of the regression functions that need to be defined a priori for HST.

The flexibility and robustness of BRT models make them suitable to model any output variable, as well as to identify changes in dam behaviour. Nevertheless, data-based models should never be the only source of information to make decisions on dam safety. Their results need to be checked against those provided by other means, such as deterministic models. Also, all available information about the dam behaviour should be taken into account, especially that obtained by visual inspection.

5. Acknowledgements

The authors thank Carlos Barbero, dam safety manager at the Catalan Water Agency, for providing the monitoring data.
The research was supported by the Spanish Ministry of Economy and Competitiveness (*Ministerio de Economía y Competitividad*, MINECO) through the projects iComplex (IPT-2012-0813-390000) and AIDA (BIA2013-49018-C2-1-R and BIA2013- 49018-C2-2-R).

References

[1] G. Lombardi, Advanced data interpretation for diagnosis of concrete dams, Tech. rep., CISM (2004).

[2] G. Willm, N. Beaujoint, Les méthodes de surveillance des barrages au service de la production hydraulique d'Electricité de France-Problèmes ancients et solutions nouvelles, in: 9th ICOLD Congres, 1967, pp. 529–550, q34-R30. [in French].

[3] M. Tatin, M. Briffaut, F. Dufour, A. Simon, J.-P. Fabre, Thermal displacements of concrete dams: Accounting for water temperature in statistical models, Engineering Structures 91 (2015) 26–39.

[4] F. Amberg, Interpretative models for concrete dam displacements, in: 23th ICOLD Congress, 2009, q91-R43.

[5] D. Santillán, E. Salete, D. Vicente, M. Toledo, Treatment of solar radiation by spatial and temporal discretization for modeling the thermal response of arch dams, Journal of Engineering Mechanics 140 (11).

[6] I. Penot, B. Daumas, J. Fabre, Monitoring behaviour, Water Power and Dam Construction.

[7] F. Salazar, M. Toledo, Discussion on "Thermal displacements of concrete dams: Accounting for water temperature in statistical models", Engineering Structures (2015) –doi:http://dx.doi.org/10.1016/j.engstruct.2015.08.001.

[8] F. Salazar, R. Morán, M. Á. Toledo, E. Oñate, Data-based models for the prediction of dam behaviour: A review and some methodological considerations, Archives of Computational Methods in Engineering (2015) 1–21.

[9] V. Ranković, N. Grujović, D. Divac, N. Milivojević, Development of support vector regression identification model for prediction of dam structural behaviour, Structural Safety 48 (2014) 33–39.

[10] J. Mata, Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models, Engineering Structures 3 (3) (2011) 903 – 910. doi:10.1016/j.engstruct.2010.12.011.

[11] S. Demirkaya, Deformation analysis of an arch dam using ANFIS, in: Proceedings of the second international workshop on application of artificial intelligence and innovations in engineering geodesy. Braunschweig, Germany, 2010, p. 21–31.

[12] L. Auret, C. Aldrich, Empirical comparison of tree ensemble variable importance measures, Chemometrics and Intelligent Laboratory Systems 105 (2) (2011) 157–170.

[13] F. Salazar, M. Toledo, E. Oñate, R. Morán, An empirical comparison of machine learning techniques for dam behaviour modelling, Structural Safety 56 (2015) 9–17.

[14] F. Li, Z. Wang, G. Liu, Towards an error correction model for dam monitoring data analysis based on cointegration theory, Structural Safety 43 (2013) 12–20.

[15] D. Santillán, J. Fraile-Ardanuy, M. Toledo, Seepage prediction in arch dams by means of artificial neural networks, Water Technology and Science V (3).

[16] L. Cheng, D. Zheng, Two online dam safety monitoring models based on the process of extracting environmental effect, Advances in Engineering Software 57 (2013) 48–56.

[17] L. Breiman, et al., Statistical modeling: The two cultures (with comments and a rejoinder by the author), Statistical Science 16 (3) (2001) 199–231.

[18] T. Hastie, R. Tibshirani, J. Firedman, The Elements of Statistical Learning - Data Mining, Inference, and Prediction, 2nd Edition, Springer, New York, 2009.

[19] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, Classification and regression trees, Wadsworth & Brooks, Monterrey, CA, 1984.

[20] J. Friedman, Greedy function approximation: a gradient boosting machine, Annals of Statistics (2001) 1189 – 1232.

[21] G. Ridgeway, Generalized Boosted Models: A guide to the gbm package, r package vignette (2007).
URL `http://CRAN.R-project.org/package=gbm`

[22] J. Leathwick, J. Elith, M. Francis, T. Hastie, P. Taylor, Variation in demersal fish species richness in the oceans surrounding new zealand: an analysis using boosted regression trees, Marine Ecology Progress Series 321 (2006) 267–281.

[23] J. Elith, J. R. Leathwick, T. Hastie, A working guide to boosted regression trees, Journal of Animal Ecology 77 (4) (2008) 802–813.

[24] R. E. Schapire, The boosting approach to machine learning: An overview, in: Nonlinear estimation and classification, Springer, 2003, pp. 149–171.

[25] A. Michelis, Traditional versus non-traditional boosting algorithms, Master's thesis, University of Manchester (2012).

[26] G. R. with contributions from others, gbm: Generalized Boosted Regression Models, r package version 2.1 (2013).

[27] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2013).
URL `http://www.R-project.org/`

[28] O. Kaser, D. Lemire, Tag-cloud drawing: Algorithms for cloud visualization, arXiv preprint cs/0703109.

[29] I. Fellows, wordcloud: Word Clouds, r package version 2.5 (2014).
URL `http://CRAN.R-project.org/package=wordcloud`

[30] J. H. Friedman, J. J. Meulman, Multiple additive regression trees with application in epidemiology, Statistics in medicine 22 (9) (2003) 1365–1381.

[31] A. S. Weigend, B. A. Huberman, D. E. Rumelhart, Predicting sunspots and exchange rates with connectionist networks, in: S. Eubank, M. Casdagli (Eds.), Proc. of the 1990 NATO Workshop on Nonlinear Modeling and Forecasting (Santa Fe, NM), Vol. 12, Addison-Wesley, Redwood, CA, 1992, pp. 395–432.

[32] G. Hooker, Diagnosing extrapolation: Tree-based density estimation, in: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2004, pp. 569–574.

[33] A. Carrère, C. Noret-Duchêne, Interpretation of an arch dam behaviour using enhanced statistical models, in: Proceedings of the Sixth ICOLD Benchmark Workshop on Numerical Analysis of Dams, Salzburg, Austria, 2001.

Appendix A.

All the plots generated during the analysis are included herein: the time series of mean air temperature and daily rainfall, and a set of plots for each target variable:

- The 2D and 3D partial dependence plots for BRT Model 3 fitted to the original data
- The location of each device within the dam body
- The 2D partial dependence plot for BRT Model 3 fitted on the altered version of the target (independent of time)
- The word cloud for Model 1
- Observations versus BRT model predictions for the training and validation sets, together with the model residuals.

The partial dependence for the artificial data was included to highlight that the BRT models correctly captured the time independence when it was imposed in the time series of the target variable.

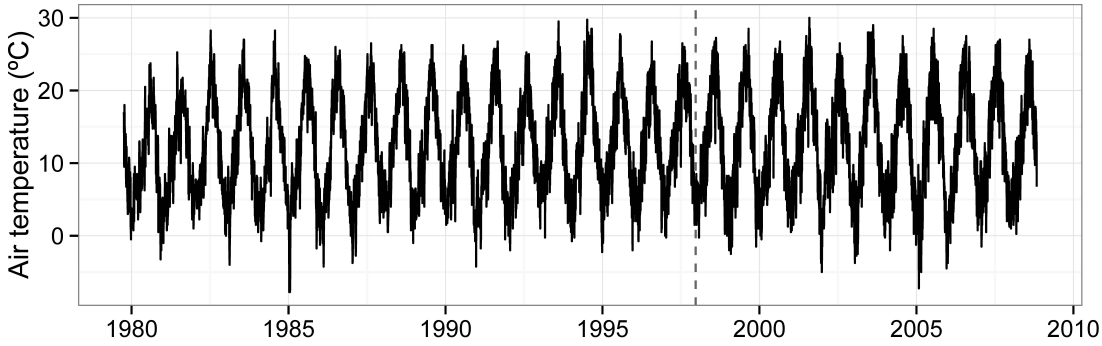


Figure A1: Time series of the mean air temperature at La Baells dam site.

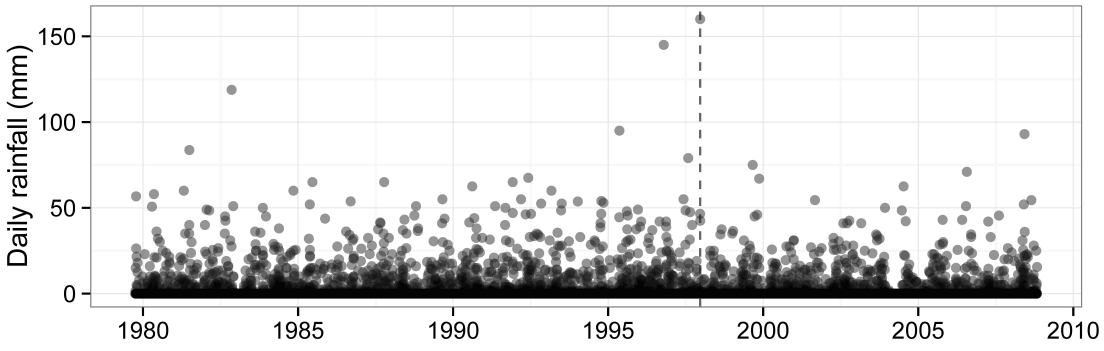


Figure A2: Time series of the daily rainfal at La Baells dam site.

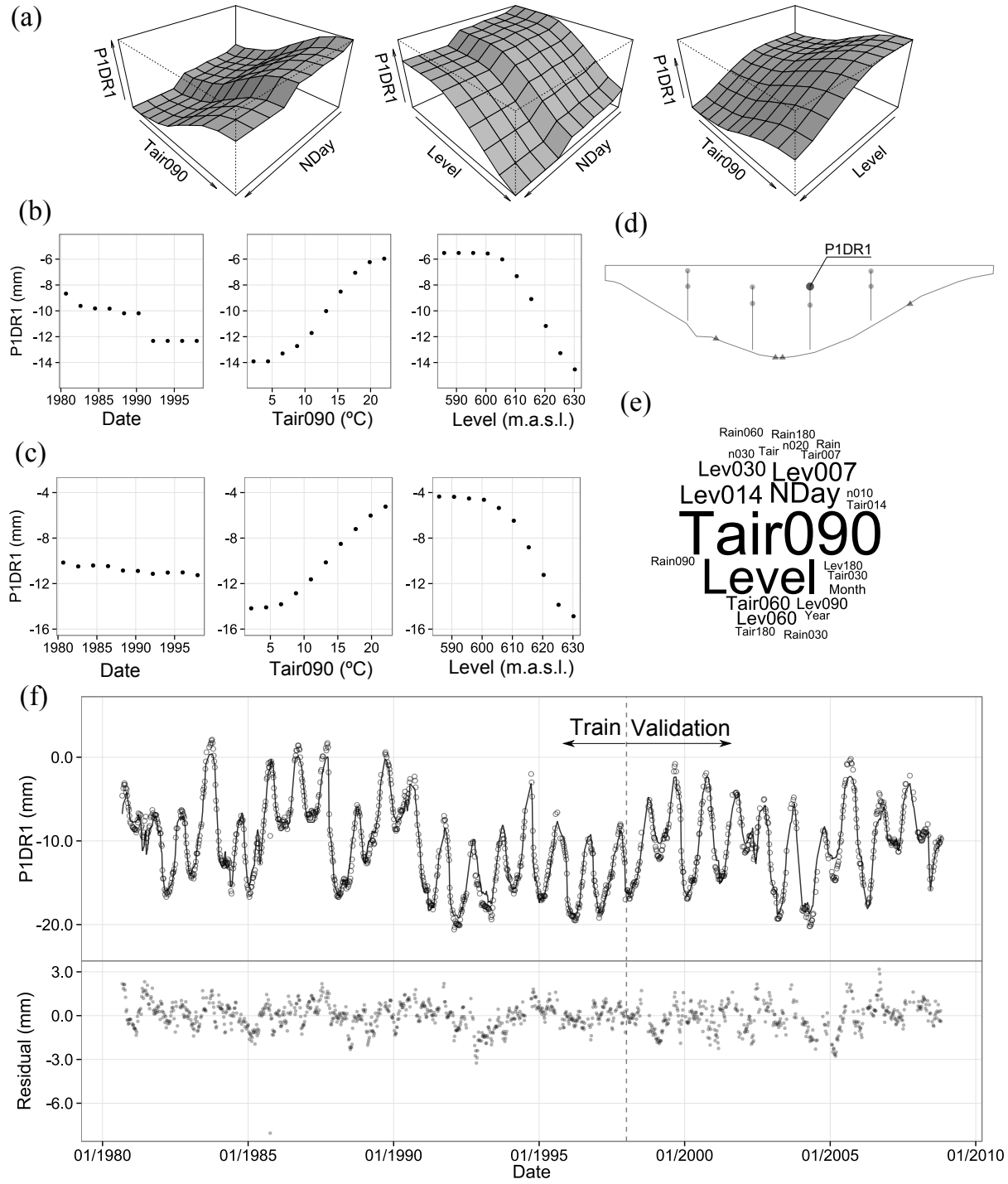


Figure A3: P1DR1. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

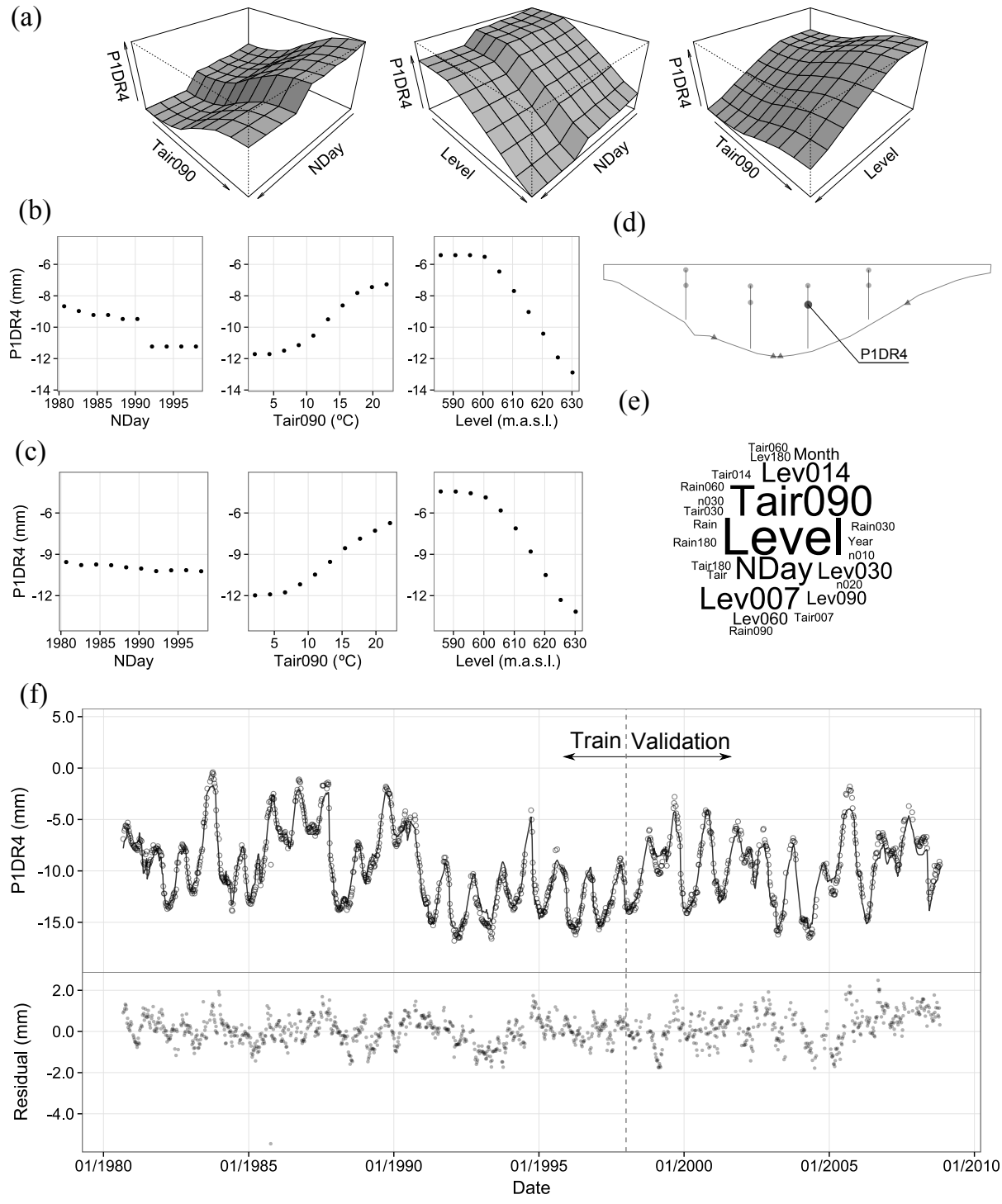


Figure A4: P1DR4. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

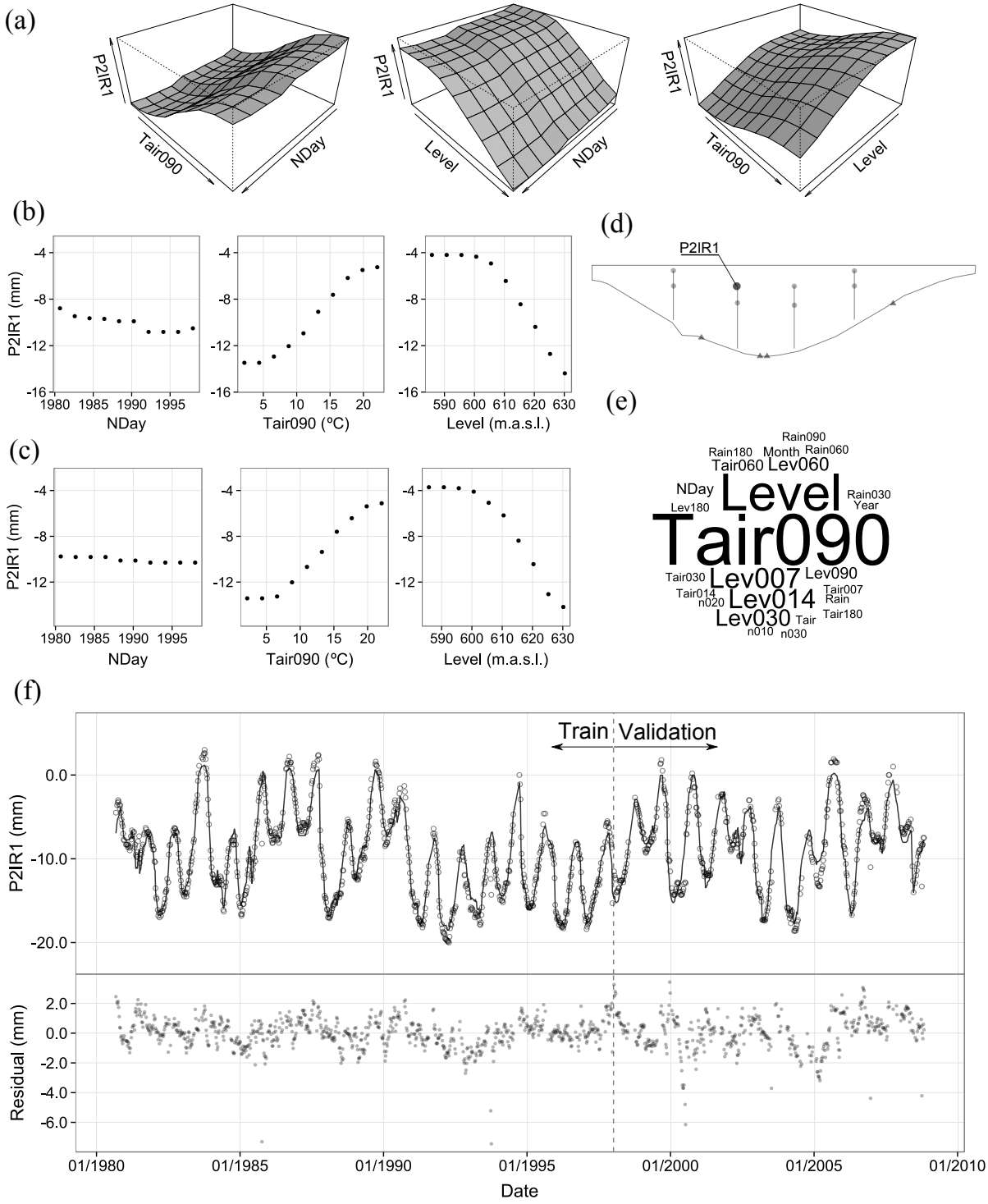


Figure A5: P2IR1. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

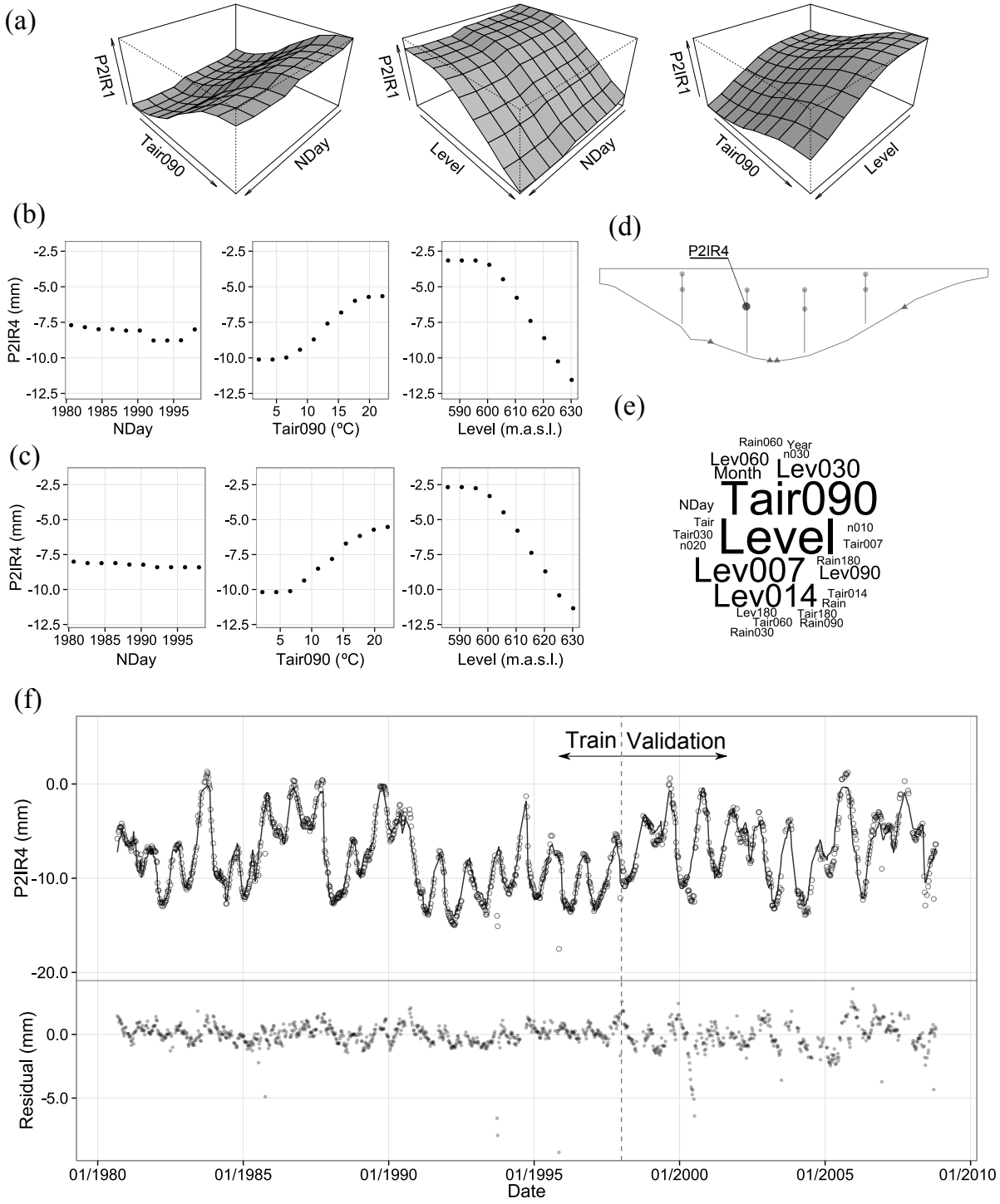


Figure A6: P2IR4. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

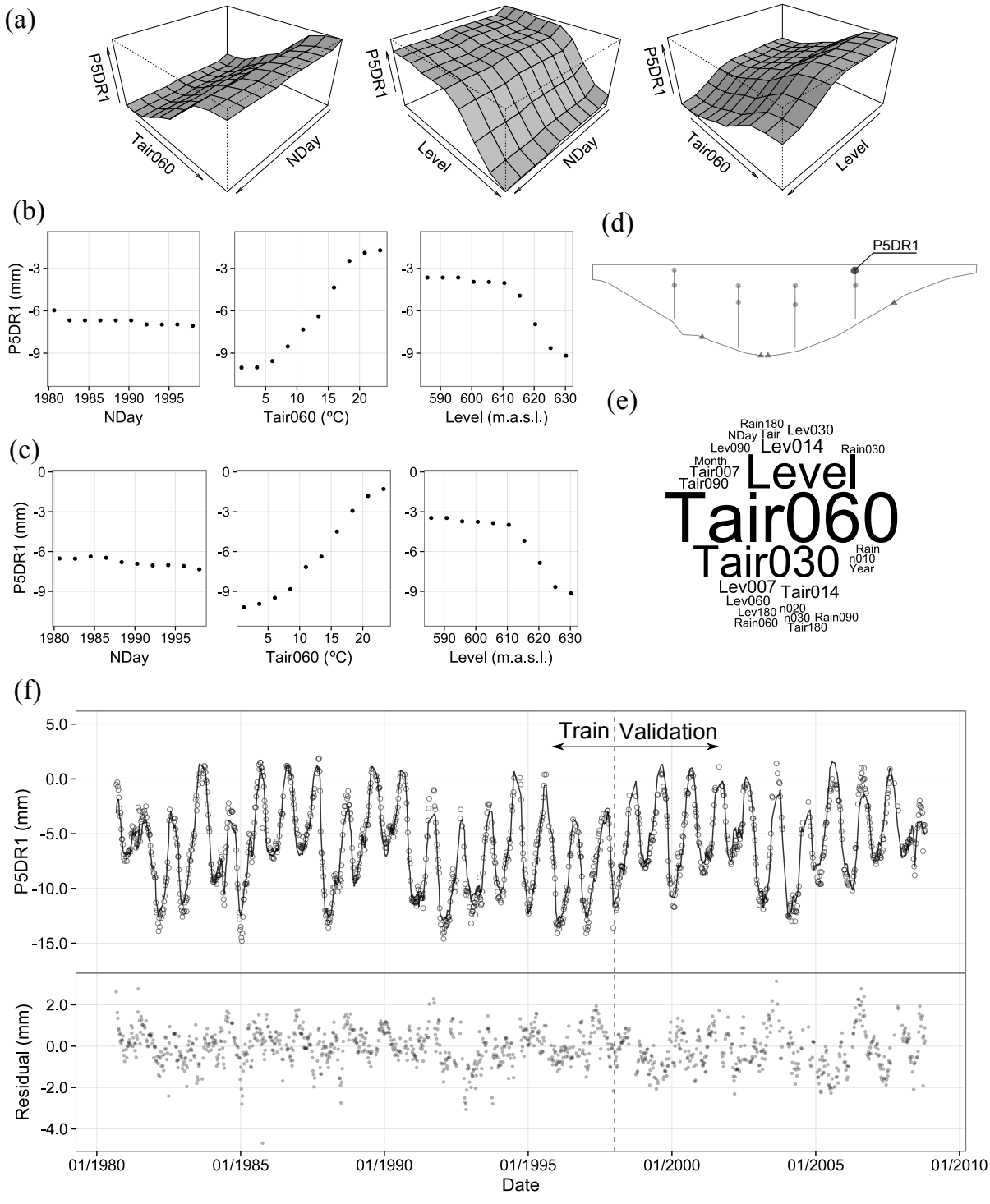


Figure A7: P5DR1. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

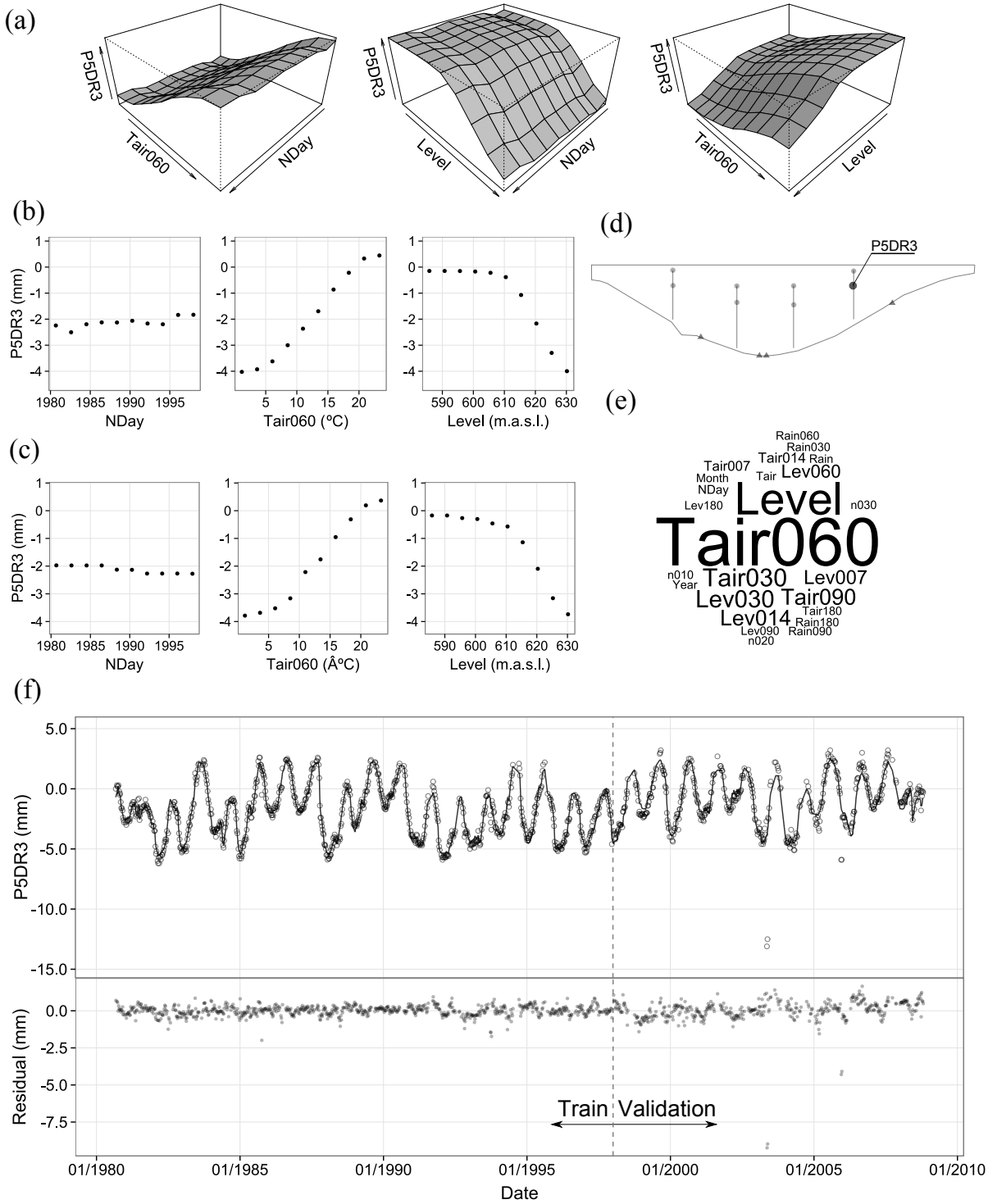


Figure A8: P5DR3. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

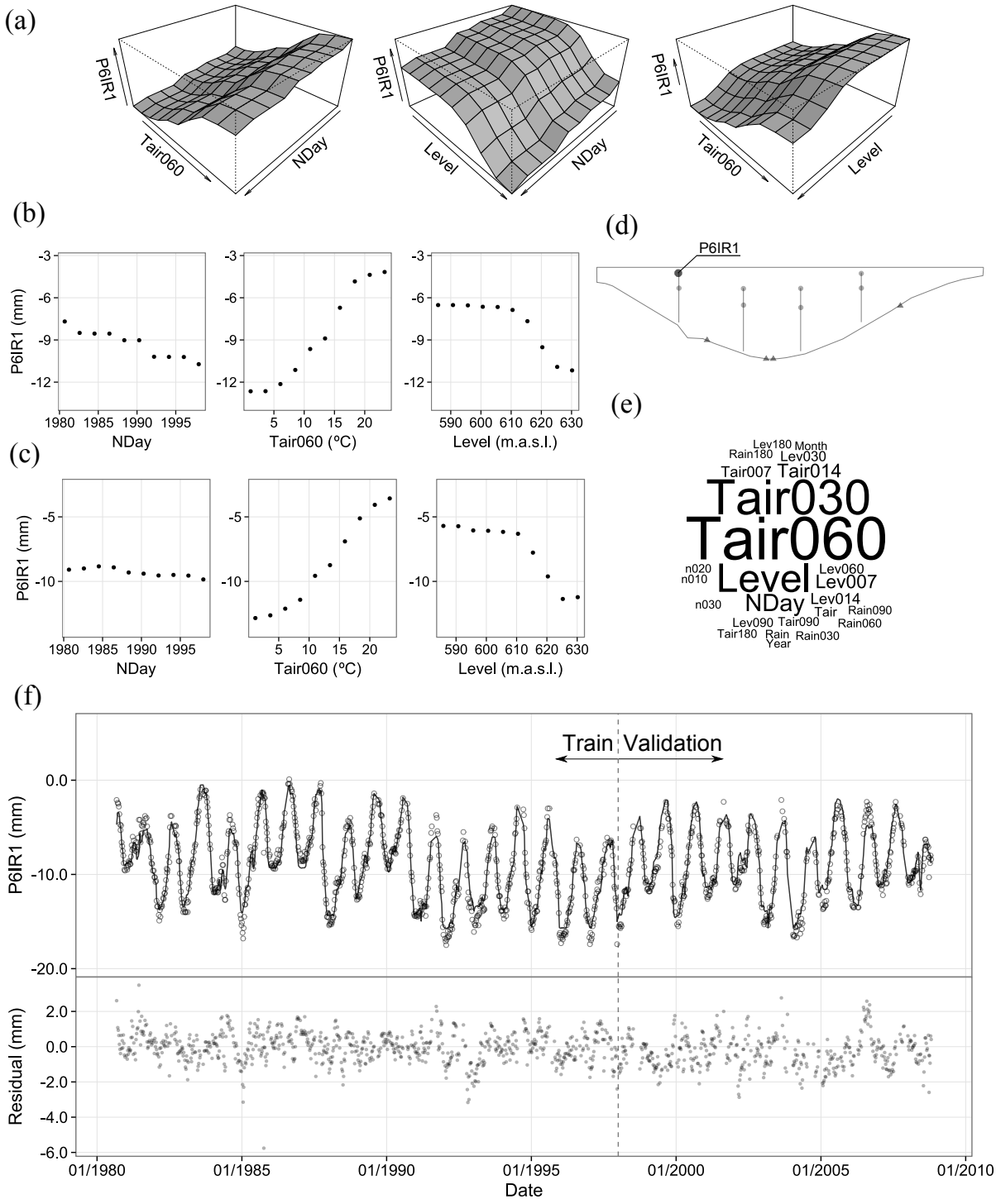


Figure A9: P6IR1. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

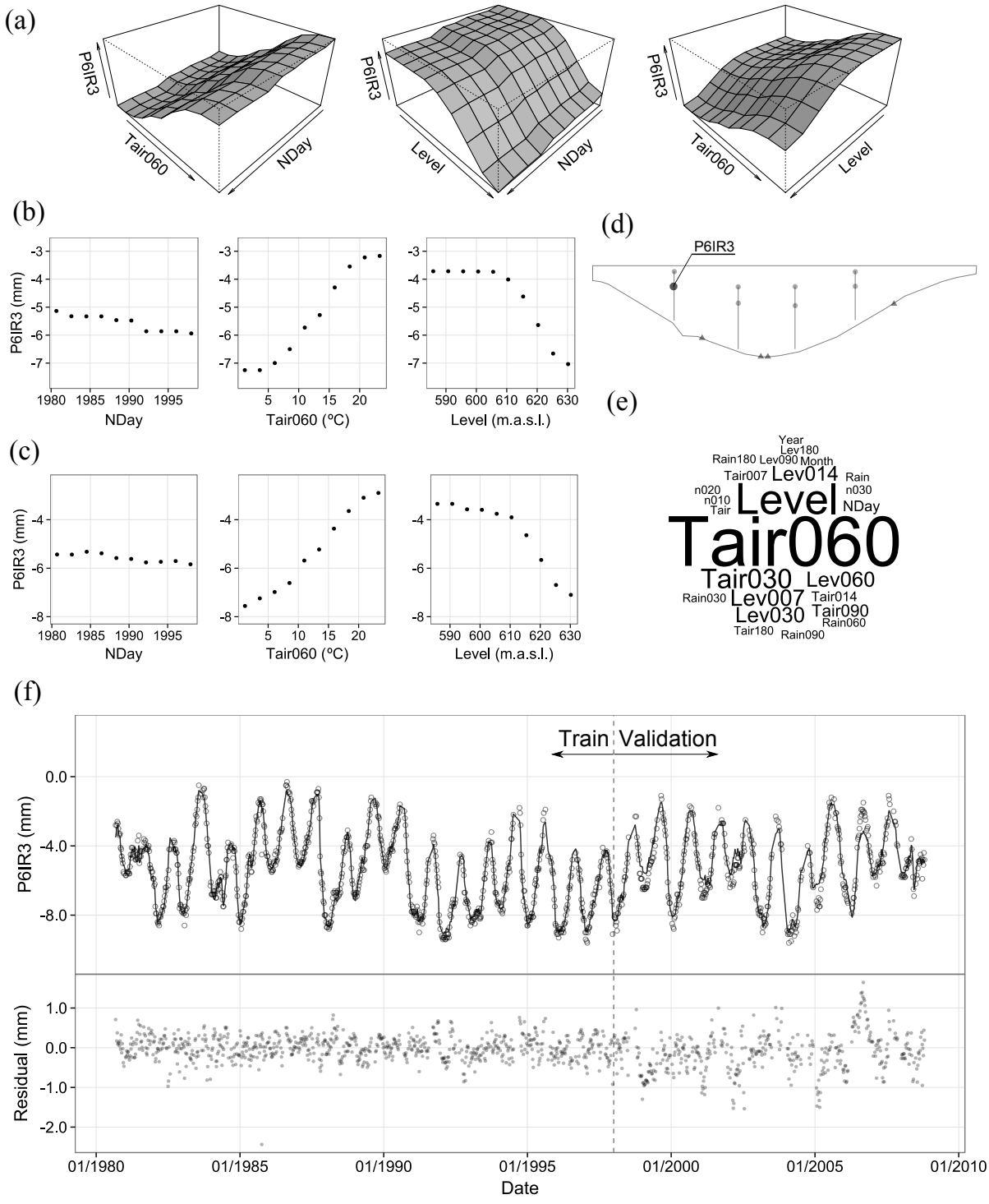


Figure A10: P6IR3. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

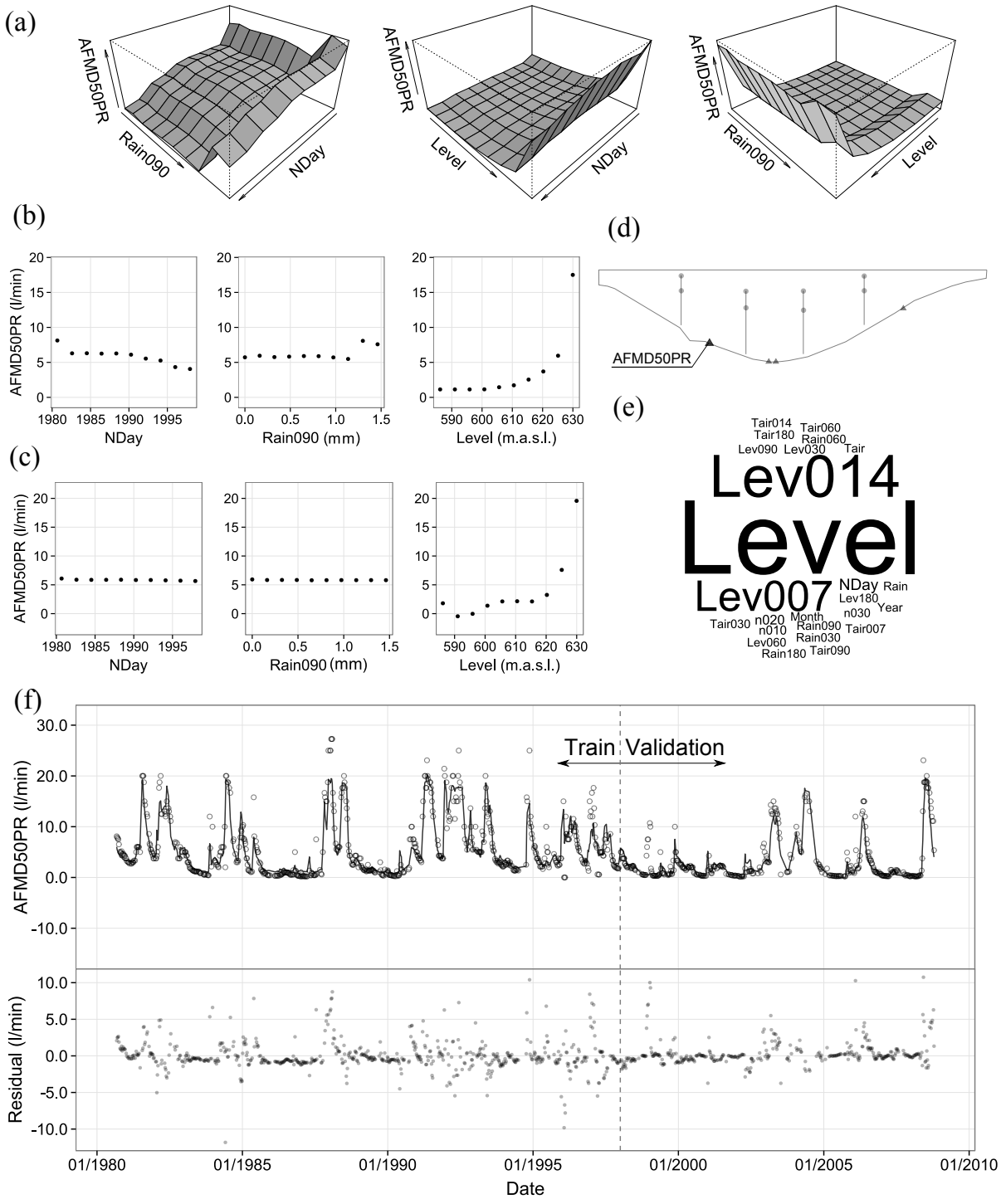


Figure A11: AFMD50PR. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

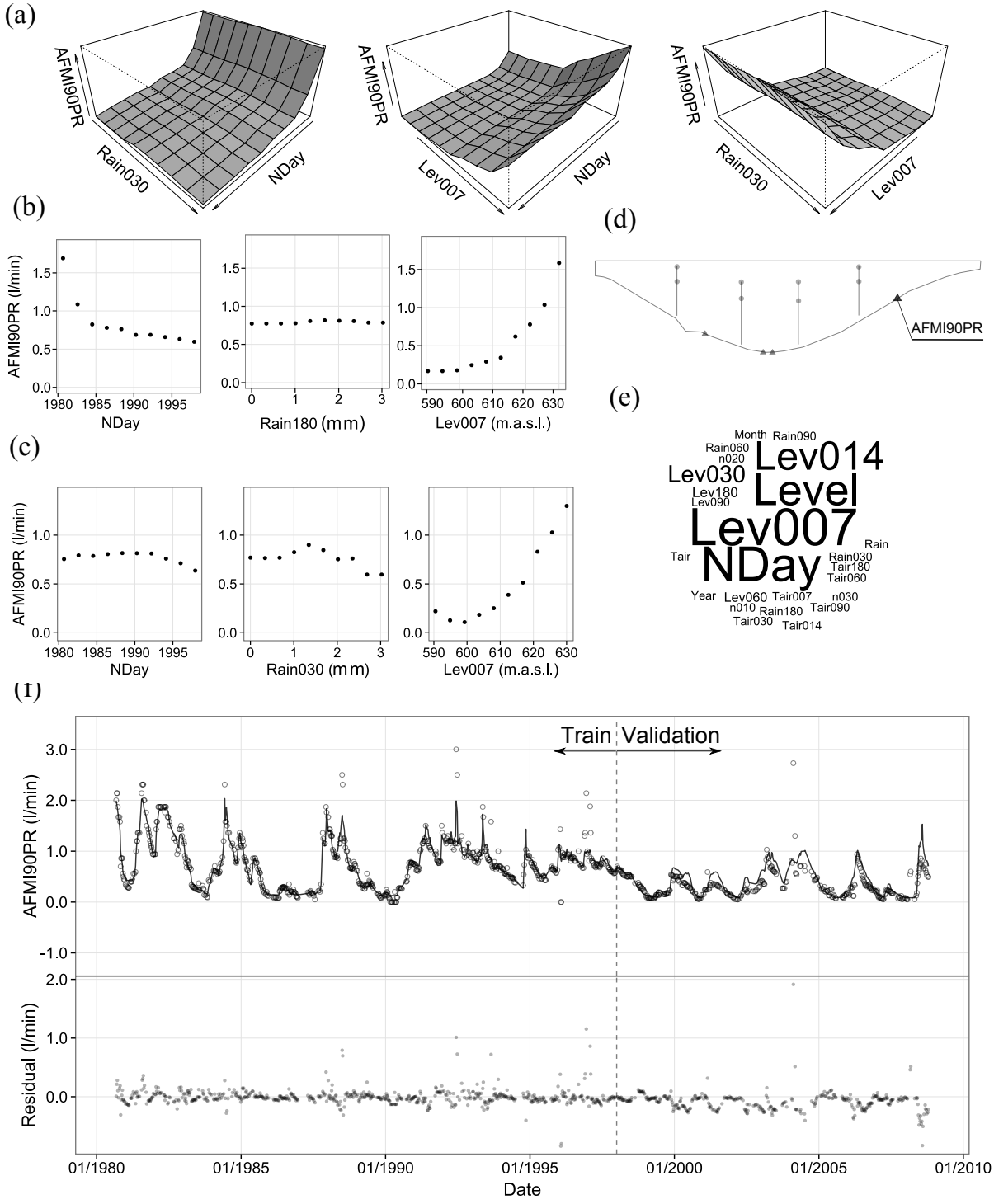


Figure A12: AFMI90PR. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

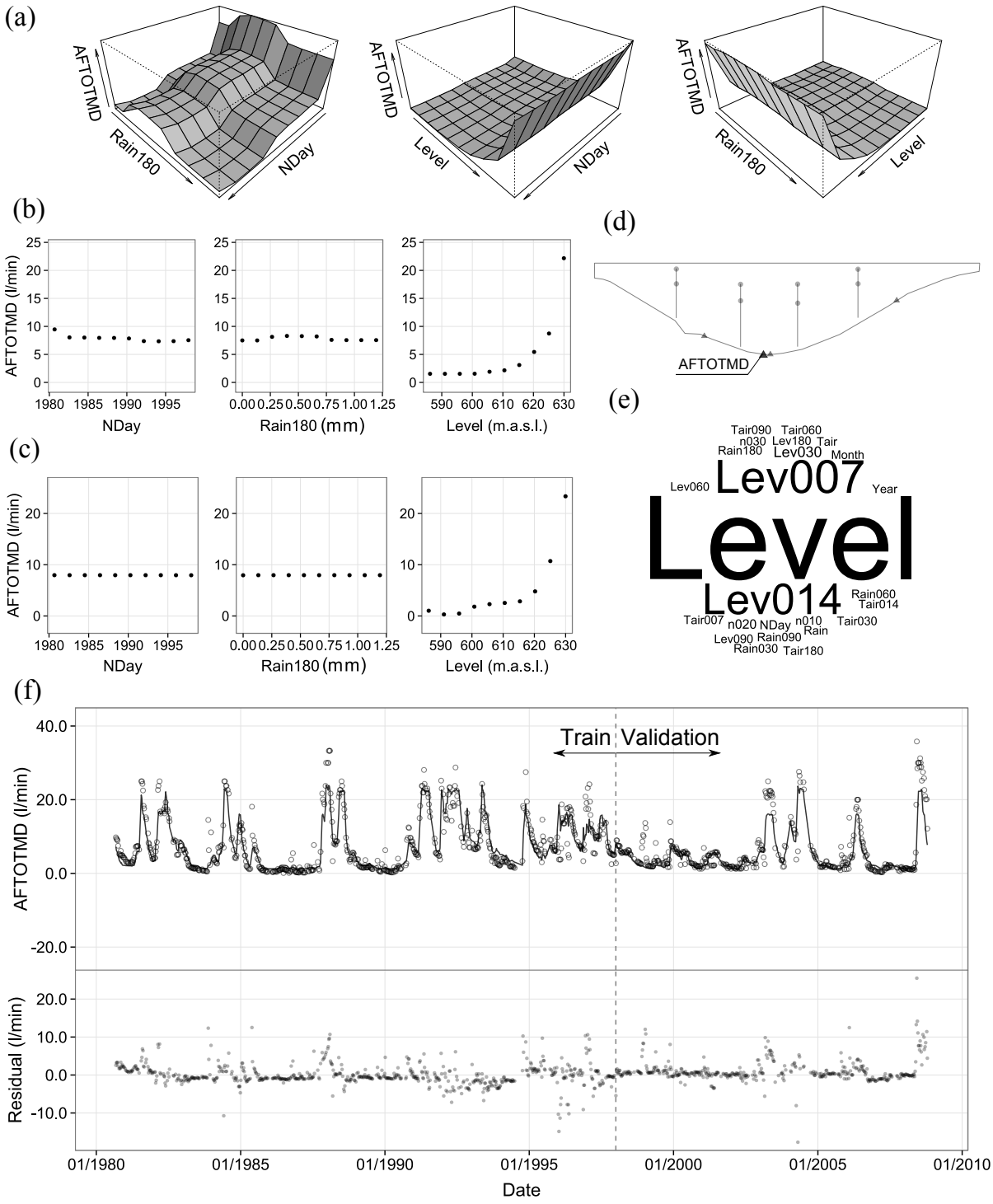


Figure A13: AFTOTMD. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

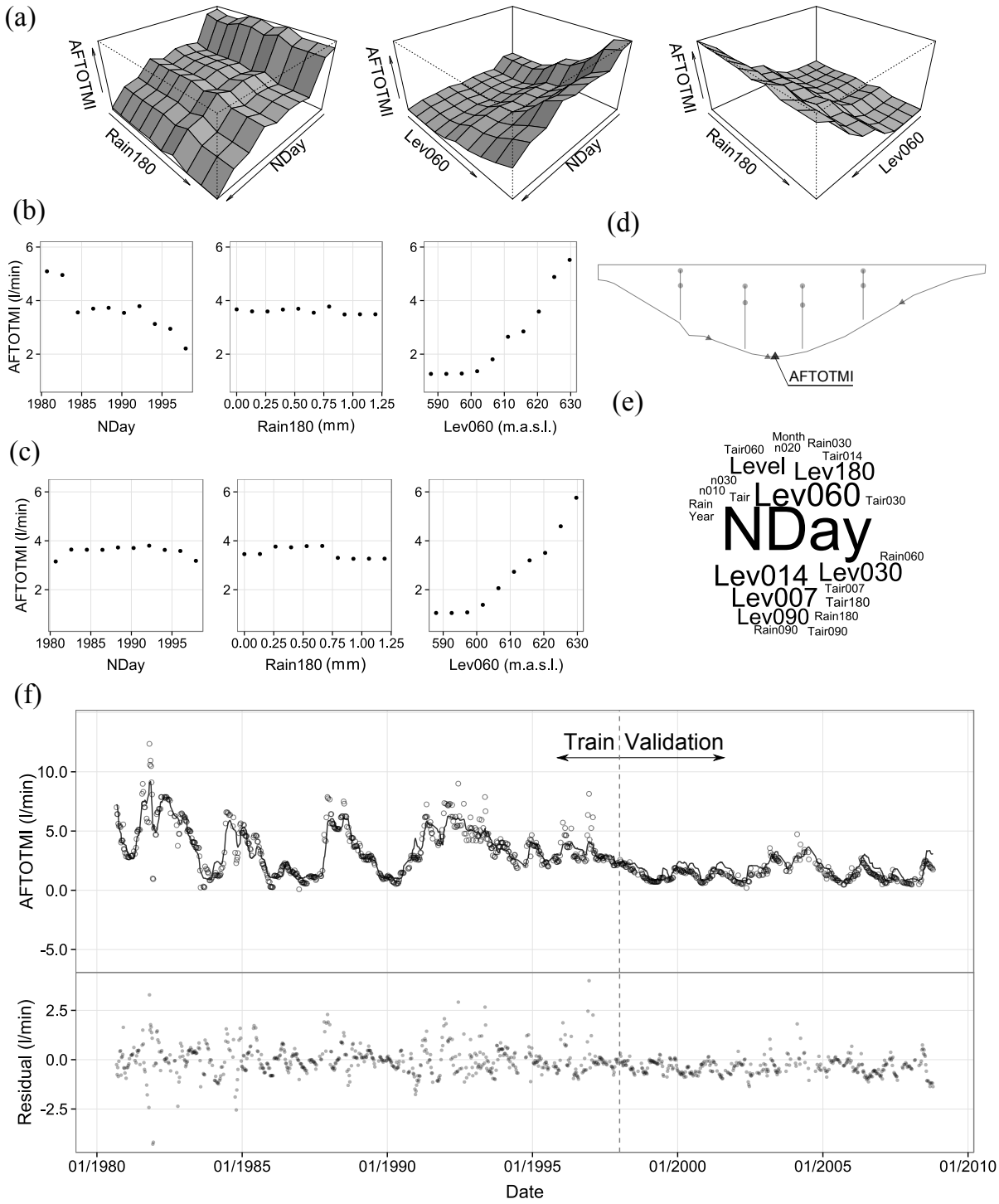


Figure A14: AFTOTMI. (a) 3D partial dependence plot; (b) 2D Partial dependence plot; (c) Idem for artificial data (time-independent); (d) Device location; (e) Word cloud of relative influence; (f) Model fit and residuals for the train and the validation sets.

B

Other publications

B.1 Posibilidades de la inteligencia artificial en el análisis de auscultación de presas

Title: Posibilidades de la inteligencia artificial en el análisis de auscultación de presas

First Author: Fernando Salazar González. CIMNE - International Center for Numerical Methods in Engineering

Second Author: Miguel Á. Toledo Municio. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Third Author: Eugenio Oñate Ibáñez de Navarra. CIMNE - International Center for Numerical Methods in Engineering

Conference: III Jornadas de Ingeniería del Agua. La protección contra los riesgos hídricos (JIA 2013)

Date-Location: October 2013 - Valencia (Spain)

ISBN 978-84-267-2070-2

III Jornadas de Ingeniería del Agua
La protección contra los riesgos hídricos

Posibilidades de la inteligencia artificial en el análisis de auscultación de presas

F. Salazar, E. Oñate
Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE)
Gran Capitán s/n. 08034 Barcelona.

M. Á. Toledo
Departamento de Ingeniería Civil. Hidráulica y Energética.
Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos
Universidad Politécnica de Madrid (UPM). Profesor Aranguren s/n. 28040 Madrid.

1. Introducción

El comportamiento estructural de las presas de embalse es difícil de predecir con precisión. Los modelos numéricos para el cálculo estructural resuelven bien las ecuaciones de la mecánica de medios continuos, pero están sujetos a una gran incertidumbre en cuanto a la caracterización de los materiales, especialmente en lo que respecta a la cimentación. Así, es difícil discernir si un estado que se aleja en cierta medida de la normalidad supone o no una situación de riesgo estructural.

Por el contrario, muchas de las presas en operación cuentan con un gran número de aparatos de auscultación, que registran la evolución de diversos indicadores como los movimientos, el caudal de filtración, o la presión intersticial, entre otros. Aunque hoy en día hay muchas presas con pocos datos observados, hay una tendencia clara hacia la instalación de un mayor número de aparatos que registran el comportamiento con mayor frecuencia (Restelli 2008)0. Como consecuencia, se tiende a disponer de un volumen creciente de datos que reflejan el comportamiento de la presa. En la actualidad, estos datos suelen tratarse con métodos estadísticos para extraer información acerca de la relación entre variables, detectar anomalías y establecer umbrales de emergencia.

El modelo general más común es el denominado HST (Hydrostatic-Season-Time), que calcula la predicción de una variable determinada a partir de una serie de funciones que tienen en cuenta los factores que teóricamente más influyen en la respuesta: la carga del embalse, el efecto térmico (en función de la época del año) y un término irreversible. El método fue desarrollado en 1967 (Simon et *al*. 2013) por ingenieros de Électricité de France (EDF) y se ha aplicado especialmente a la predicción de desplazamientos (Swiss Comittee on Dams, 2003).

Aunque se ha utilizado con éxito durante mucho tiempo, se han detectado algunas limitaciones del método, como el hecho de que asume que los tres efectos son independientes, y que las funciones deben definirse a priori. Esto permite un margen de mejora en caso de que no representen el efecto real en un caso concreto (Simon et *al*., 2013). Además, no permite reproducir la inercia de la presa en su respuesta frente a ciertas solicitaciones, como es el caso de la relación entre el nivel de embalse y la presión intersticial en presas de materiales sueltos (Bonelli y Radzicki 2008).

Puntualmente se han aplicado modelos más complejos para solventar las limitaciones mencionadas. En algunos casos se han introducido otras variables como la velocidad de variación del nivel de embalse (Sánchez Caro 2003), y en otros se han utilizado expresiones que se adaptan mejor al efecto de inercia, como la función impulso-respuesta (Bonelli y Radzicki 2008).

En otros campos de la ciencia, como la medicina o las telecomunicaciones, el volumen de datos es mucho mayor, lo que ha motivado el desarrollo de numerosas herramientas para su tratamiento y para la generación de modelos de predicción. Algunas de ellas, como las redes neuronales, ya han sido aplicadas al caso de la auscultación de presas (Santillán et *al*. 2010, Mata 2011, Simon et *al*. 2013) con resultados prometedores.

La aplicación de estas técnicas puede ayudar a mejorar la precisión de los modelos de predicción, y a entender mejor el comportamiento de la presa. Con esta idea se ha puesto en marcha el proyecto de investigación iComplex, en el que participan la empresa DACARTEC, la Universidad Politécnica de Madrid (UPM) y el Centro Internacional de Métodos Numéricos en Ingeniería (CIMNE).

Una de las tareas comprendidas en la primera fase del proyecto es la revisión de diversas herramientas de inteligencia artificial de cara a su aplicación a la predicción del comportamiento de presas: movimientos, tensiones, filtraciones, etc. Una de las herramientas que se está considerando como potencialmente útil está basada en los llamados bosques aleatorios. A continuación se describen someramente las bases de esta tecnología y se muestran resultados preliminares de su aplicación a un caso piloto.

2. Sobre los bosques aleatorios

Los bosques aleatorios (Breiman 2001) son modelos que permiten predecir el valor de una determinada variable (*variable objetivo*) a partir de una serie de *variables predictoras*, cuyo valor es conocido. Como los métodos estadísticos, requiere de unos datos de entrenamiento, a partir de los cuales se ajusta el modelo al caso de estudio. Un bosque aleatorio está formado por un conjunto de árboles de decisión. La predicción del bosque es el promedio de las predicciones de los árboles que lo forman. Por tanto, se trata de un *modelo de conjunto* (Martínez, 2006).

Los árboles de decisión se basan en la división sucesiva del conjunto de datos observados en grupos de casos “similares”. Suelen denominarse *árboles de regresión* aquellos cuya variable

objetivo es continua, y *árboles de clasificación* cuando es discreta o categórica. En adelante se utilizará por tanto el término árbol de regresión, ya que en auscultación de presas se trata de predecir variables continuas. La predicción un árbol de regresión es en general un valor constante para cada grupo, igual a la media de los valores observados. Para explicar el proceso de generación de un árbol de decisión, se utiliza un ejemplo sencillo relacionado con la auscultación de presas: supongamos que se desea ajustar un modelo para predecir el caudal de filtración en un determinado aforador a partir únicamente del nivel de embalse. La Figura 1 muestra la relación entre las variables de entrada (nivel de embalse) y objetivo (caudal de filtración).

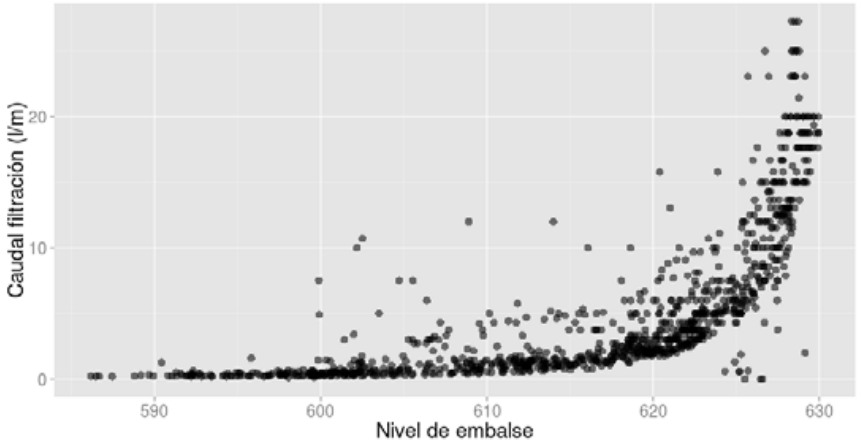


Figura 1. *Caudal de filtración en función del nivel de embalse en el caso de ejemplo*

En primer lugar, se dividen los datos en dos regiones según el nivel de embalse (en el ejemplo, según sea mayor o menor de 625,6; Figura 2). A continuación, una de las regiones creadas se subdivide a su vez en dos, y el proceso continúa hasta que se alcanza algún criterio de parada. En el ejemplo, el resultado final es la división de los casos en 7 grupos. La predicción del modelo es la media de los valores observados en cada grupo, y por tanto el resultado es una sucesión de escalones (Figura 3).

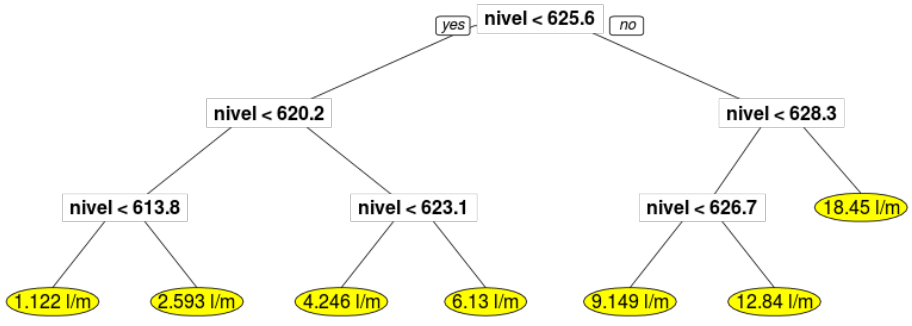


Figura 2. *Árbol de regresión para la predicción del caudal de filtración en función del nivel de embalse en el caso de ejemplo.*

El algoritmo de generación del árbol calcula la mejor división posible en cada paso como la que minimiza el error de predicción. Si hubiera más de una variable predictora, se calcula para cada una de ellas la división óptima y a continuación se selecciona la variable que produce una división más precisa (Hastie et *al.* 2009).

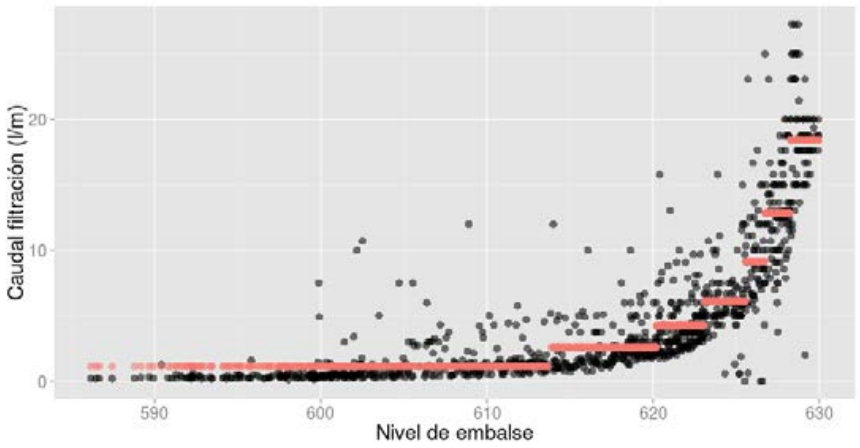


Figura 3. *Predicción del árbol de regresión (en rojo), en comparación con los datos registrados (en negro)*

Algunas de las propiedades más representativas de los árboles de decisión son las siguientes: a) su coste computacional es moderado; b) se adaptan bien a problemas no lineales; c) manejan sin problemas mezclas de variables continuas y discretas; además, las continuas pueden tener rangos muy diferentes lo que evita tener que transformarlas previamente, como ocurre con otros métodos; d) permiten considerar valores no medidos; e) los valores atípicos no modifican sustancialmente el resultado; f) no es necesario adoptar ninguna hipótesis a priori sobre la relación entre variables, ni sobre cuáles son más relevantes; g) son inestables, en cuanto a que una pequeña variación de los datos de entrenamiento puede producir una variación importante del resultado. Este problema puede convertirse en una ventaja si se utilizan métodos avanzados basados en árboles, como es el caso de los bosques aleatorios.

Los bosques aleatorios pertenecen a los denominados *modelos de conjuntos*, basados en la generación de un número (generalmente elevado) de modelos sobre una misma base de datos (o subconjuntos de ella). La predicción se calcula como la media de las predicciones de los modelos individuales. Los bosques aleatorios son un conjunto de árboles de decisión generados a partir de perturbaciones de los datos observados. El aspecto clave que caracteriza el método es que busca que los árboles sean independientes entre sí. Para ello, la diferencia principal con un árbol de decisión convencional es que en cada división, en lugar de considerar todas las variables predictoras disponibles para seleccionar la que minimiza el error, se analiza únicamente un **subconjunto aleatorio** de las mismas (Breiman 2001). De este modo, se aprovecha la propiedad de los árboles de decisión por la cual pequeñas perturbaciones en las primeras divisiones del espacio provocan resultados apreciablemente diferentes. Al introducir aleatoriedad en la construcción de cada árbol, se

consigue un conjunto de ellos sensiblemente independientes, de modo que se captura una proporción mayor de los patrones presentes en los datos de entrenamiento, y el resultado final mejora. Además, cada árbol se genera a partir de un conjunto de datos diferente, obtenido de los originales tomando una **muestra aleatoria con repetición**. Es decir, cada nuevo árbol se construye a partir de una muestra en la que aproximadamente un tercio de los datos originales aparece una vez, otro tercio aparece repetido, y el tercio restante no aparece.

Los bosques aleatorios han cobrado gran popularidad como método para generar modelos predictivos por su sencillez de programación y sus buenos resultados en diversas aplicaciones (Ganuer et *al*. 2008). Si bien suelen considerarse un modelo *de caja negra*, en cuanto a que no proporciona parámetros con interpretación física, se han desarrollado herramientas que permiten cuantificar cómo afecta cada variable al resultado final. En concreto, se define el índice de importancia de una variable como la variación del error de predicción que se produce al modificar aleatoriamente su valor, manteniendo el resto. Las variables más importantes provocarán un mayor aumento del error al ser permutadas.

3. Ejemplo de aplicación

Se ha elaborado un modelo de predicción en un caso de prueba basado en bosques aleatorios. Las variables que se pretende predecir son los caudales de filtración de una presa tomada como caso piloto. El periodo de datos disponible comprende desde la puesta en carga hasta el año 2008. Se han considerado los aforadores con un mayor número de datos registrados en ese periodo, que son los de la Tabla 1.

Además de los datos de aforo, del nivel de embalse y de otras magnitudes que no son objeto de este trabajo, se miden en la presa variables ambientales como precipitación y temperatura del aire.

Para esta primera prueba, las variables utilizadas para predecir el valor del caudal de filtración son: a) el número de día del registro, contado a partir de la puesta en carga de la presa; b) el año; c) el nivel de embalse medio el día de la lectura; d) la temperatura media ambiental; e) la precipitación acumulada en los 30 días anteriores a la lectura; f) la velocidad media de variación del nivel de embalse en los 10 días anteriores a la lectura; g) la media móvil de 60 días del nivel de embalse.

Aforador	Margen	Nº datos disponible
md50pr	Derecha	1023
md90pr	Derecha	748
totmd	Derecha	1071
mi50p	Izquierda	1066
mi90pr	Izquierda	1001
totmi	Izquierda	1021

Tabla 1. *Número de datos disponibles en cada aforador*

Se han dividido los datos disponibles en dos grupos. El primero se utiliza para ajustar los parámetros del modelo (datos de entrenamiento), y el segundo para comprobar la bondad del ajuste (datos de validación). La división se ha realizado de dos formas: a) el 60% de los datos más antiguos para entrenamiento, y el 40% más reciente para validación. Este es el criterio utilizado en el análisis de seguridad de la presa, así como en el trabajo de Santillán *et al.,* que utiliza redes neuronales (Santillán 2010); b) división aleatoria en todo el periodo registrado, con un 70% para entrenamiento y un 30% para validación.

En la 0 se muestra el error resultante en cada caso.

RMSE (l/min)				
División de los datos	60%-40% temporal		70%-30% aleatoria	
Aforador	Entrenamiento	Validación	Entrenamiento	Validación
md50pr	2,15	2,64	2,12	1,77
md90pr	0,40	1,58	0,55	0,47
totmd	2,56	4,42	2,41	2,64
mi50p	0,48	0,41	0,45	0,41
mi90pr	0,16	0,24	0,17	0,10
totmi	0,67	1,05	0,62	0,60

Tabla 2. *Errores de predicción del modelo (raíz del error cuadrático medio)*

Como ejemplo, en la Figura 4 muestra el resultado para el aforador "md50pr", que es al que corresponde también la Figura 3. Puede apreciarse la mejora de la predicción del bosque aleatorio respecto del árbol de regresión individual.

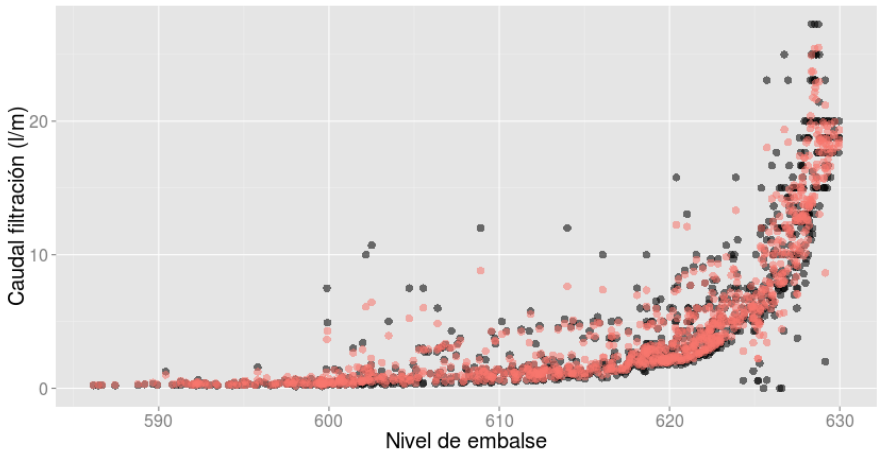


Figura 4. *Predicción del modelo basado en bosques aleatorios (en rojo), en comparación con los datos registrados (en negro). Aforador "md50pr".*

Se ha calculado también el índice de importancia de las variables, y llama la atención el hecho de que las relativas al periodo de medición (número de día desde la puesta en carga y

año de lectura) son siempre más relevantes que otras como la temperatura o la precipitación. En algunos casos, llegan a serlo más incluso que el nivel de embalse (Figura 5 izquierda). Para verificar estos resultados, se ha dibujando la relación entre el nivel de embalse y el caudal de filtración separando los registros por intervalos temporales. La Figura 6 muestra dos gráficos de este tipo, donde se observa claramente cómo en algunos casos (izquierda) el caudal de filtración depende en gran medida del periodo de la vida de la presa, lo cual no sucede en otros (derecha). Otro índice de la variación de la respuesta de la presa con el tiempo lo representa el hecho de que el error de entrenamiento es similar independientemente de cómo se dividan los datos, mientras que el de validación es sensiblemente inferior si se toman aleatoriamente. Si como se observa en la Figura 6 izquierda la respuesta de la presa varía con el tiempo, es lógico que un modelo entrenado con los datos de un periodo determinado se ajuste peor al aplicarlo a un periodo diferente.

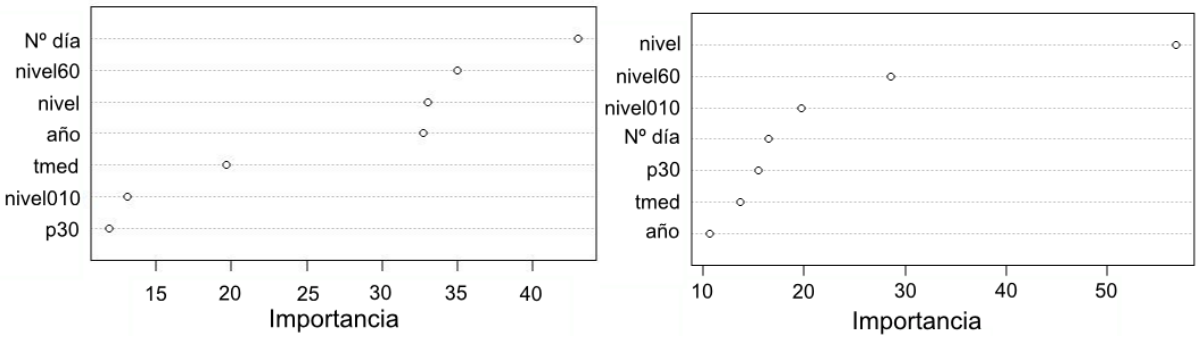


Figura 5. *Importancia de las variables predictoras en dos de los aforadores estudiados. Derecha: "totmi", donde el día del registro es la variable más importante, lo que denota una evolución temporal en el comportamiento. Izquierda: "md50pr", donde el nivel de embalse es claramente lo que más influye en la filtración.*

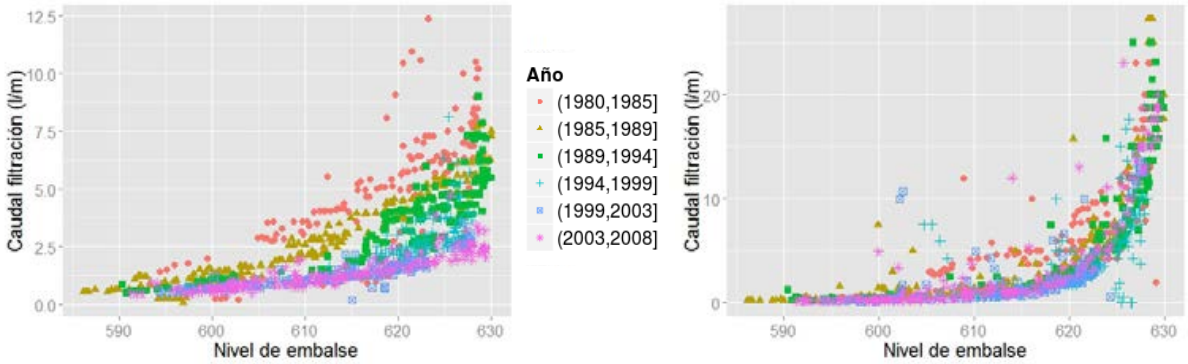


Figura 6. *Caudal de filtración "totmi" (izquierda) y "md50pr" (derecha) en función del nivel de embalse, separados por periodos de tiempo. Se observa que en el primer caso el comportamiento depende de manera importante del año de lectura, mientras que en el segundo depende más del nivel de embalse, lo cual es coherente con los índices de importancia calculados.*

Los resultados obtenidos hasta el momento con los bosques aleatorios sugieren que pueden ser una herramienta útil no solo como modelo de predicción, sino también para obtener información acerca del comportamiento de la presa, y del efecto de las variables de entorno.

En la actualidad se está trabajando para desarrollar criterios de selección de variables, así como para determinar cómo evoluciona la precisión del modelo en función del número de datos disponible.

Agradecimientos

Los autores quieren agradecer a la Agencia Catalana de l'Aigua la autorización para el uso de los datos de una de sus presas con fines de investigación, y a Ofiteco las gestiones realizadas para ello. También al Ministerio de Economía los medios facilitados para esta investigación dentro del proyecto "Desarrollo del software iComplex para el control y evaluación de la seguridad de infraestructuras críticas", del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2008-11; IPT-2012-0813-390000.

Referencias

Bonelli, S., & Radzicki, K. 2008. Impulse response function analysis of pore pressures in earthdams. *European Journal of Environmental and Civil Engineering*, *12*(3), 243-262.

Breiman, L. 2001. Random forests. *Machine learning*, 45(1), 5-32, 2001.

Ganuer, R., Poggi, J.M., Tuleau, C. 2008. Random Forests: some methodological insights. arXiv: 0811.3619.

Hastie, T., Tibshirani, R. y Friedman, J. 2009. The elements of statistical learning - Data mining, Inference and Prediction. Springer, 2ª edición.

Martínez, G. 2006. Clasificación mediante conjuntos. Tesis Doctoral. Universidad Autónoma de Madrid.

Mata, J. 2011. Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Engineering Structures*, 33(3), 903-910, 2011.

Restelli, F. 2008. Utilización de modelación conexionista para la determinación de patrones en respuestas pulsantes de sistemas automáticos. *V Congreso Argentino de Presas*, Tucumán, Argentina. 2008.

Sánchez Caro, F. 2007. Seguridad de presas: aportación al análisis y control de deformaciones como elemento de prevención de patologías de origen geotécnico. Tesis Doctoral. UPM.

Santillán, D., Morán, R., Fraile, J.J., Toledo, M.Á. 2010. Forecasting of dam flow-meter measurements using artificial neural networks. En Romeo García et *al.* (eds) Dam Maintenance and Rehabilitation II, CRC Press, Londres, 2010, pp 183-189.

Simon, A., Royer, M., Mauris, F. y Fabre, J.P. 2013. Analysis and Interpretation of Dam Measurements using Artificial Neural Networks. En *Proceedings of the 9th ICOLD European Club Symposium*, Venecia, 2013.

Swiss Committee on Dams. Methods of analysis for the prediction and the verification of dam behavior. En *21º Congress of the International Commission on Large Dams*, Montreal, 2003.

B.2 Avances en el tratamiento y análisis de datos de auscultación de presas

Title: Avances en el tratamiento y análisis de datos de auscultación de presas

First Author: Fernando Salazar González. CIMNE - International Center for Numerical Methods in Engineering

Second Author: León Morera. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Third Author: Miguel Á. Toledo Municio. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Fourth Author: Rafael Morán. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Fifth Author: Eugenio Oñate Ibáñez de Navarra. CIMNE International Center for Numerical Methods in Engineering

Conference: X Jornadas Españolas de Presas

Date-Location: February 2015 - Sevilla (Spain)

COMITÉ NACIONAL ESPAÑOL DE GRANDES PRESAS

AVANCES EN EL TRATAMIENTO Y ANÁLISIS DE DATOS DE AUSCULTACIÓN DE PRESAS

Fernando Salazar[1]

León Morera[2]

Miguel Ángel Toledo[3]

Rafael Morán[4]

Eugenio Oñate.[5]

RESUMEN: Los avances en los instrumentos de medida y en las técnicas de transmisión y almacenamiento de información han permitido aumentar el control de la seguridad de las presas, con medidas más fiables, precisas y frecuentes. Sin embargo, los métodos de tratamiento y análisis de los datos de auscultación no han evolucionado tanto, y con frecuencia se limitan a gráficos y modelos estadísticos sencillos.

Existen multitud de herramientas desarrolladas en diversos campos, generalmente alejados de la ingeniería civil, que facilitan el análisis y modelación de sistemas complejos: redes neuronales, redes complejas o bosques aleatorios son algunos ejemplos.

Estas técnicas han sido exploradas, implementadas y aplicadas a casos de ejemplo reales. Las estimaciones obtenidas son en general más precisas que las que resultan de aplicar los métodos convencionales, con lo que se puede definir mejor el rango de comportamiento normal de la presa. Además, permiten desvelar interacciones complejas entre variables de distinto tipo, más allá de la relación con el tiempo y el nivel de embalse. Por tanto, estos métodos, manejados e interpretados por expertos en ingeniería de presas, pueden ser de gran ayuda para conocer mejor el comportamiento de la presa y aumentar su seguridad.

[1] Centro Internacional de Métodos Numéricos en Ingeniería (CIMNE)
[2] Universidad Politécnica de Madrid (UPM)
[3] Universidad Politécnica de Madrid (UPM)
[4] Universidad Politécnica de Madrid (UPM)
[5] Centro Internacional de Métodos Numéricos en Ingeniería (CIMNE)

1. INTRODUCCIÓN

El objetivo principal de los sistemas de auscultación de presas es la detección temprana de anomalías, de modo que sea posible tomar las medidas adecuadas para su corrección, y en última instancia, para evitar averías graves y la rotura.

La normativa requiere definir unos valores máximos (también mínimos en algunos casos) admisibles para determinados aparatos de auscultación, de modo que al verse superados deben activarse unos determinados protocolos de actuación.

Este rango de funcionamiento "normal" se basa en la estimación del comportamiento de la presa en unas determinadas condiciones de operación (principalmente nivel de embalse y temperatura). Para el cálculo de la respuesta en situación normal y la determinación del rango admisible se utilizan modelos de predicción de diversa naturaleza.

Los modelos determinísticos se basan en las leyes de la física, y generalmente consisten en un modelo estructural de elementos finitos, que considera con cierto grado de detalle las acciones sobre la presa. Son fundamentales en la fase de proyecto y el inicio de la explotación de la presa, si bien requieren adoptar simplificaciones importantes con respecto a las acciones [1] y a la respuesta estructural de la presa y el cimiento [2], que en la inmensa mayoría de los casos se consideran medios continuos elásticos lineales. Esto lleva a discrepancias entre sus resultados y la respuesta real de la presa, medida con los aparatos de auscultación.

Los modelos estadísticos utilizan los datos realmente medidos por el sistema de auscultación de la presa durante un periodo determinado para predecir su respuesta en un periodo posterior. Ello implica que no pueden aplicarse durante la fase inicial de la vida de la presa, hasta que no se han recopilado datos suficientes para ajustar los parámetros del modelo (el periodo necesario depende de cada caso, habiendo estudios que lo cifran en 5 [3], 10 [4] y 12 años [5]). Los métodos estadísticos convencionales presentan limitaciones importantes, como se ha puesto de manifiesto en trabajos recientes [9].

En los últimos años, la UPM y el CIMNE han puesto en marcha una investigación conjunta que pretende mejorar la seguridad de presas extrayendo la máxima información de los datos de auscultación. Se basa en dos ideas fundamentales:

- El conjunto presa-cimiento es un sistema complejo cuyos elementos son las series temporales de registro (de variables externas e internas). A partir del análisis de dichas series de registros pueden definirse conexiones o relaciones entre los elementos del sistema.

- Existen herramientas de inteligencia artificial, redes complejas y sistemas expertos que pueden ser útiles para analizar el sistema en conjunto, estudiar las relaciones entre sus elementos, y generar modelos de predicción de las variables de respuesta.

En la comunicación se incluyen algunos de los resultados obtenidos hasta el momento.

2. EXPLORACIÓN DE DATOS

La primera operación a realizar sobre los datos de auscultación, una vez recibidos y almacenados, es el análisis gráfico. Es muy frecuente dibujar la evolución de los registros a lo largo del tiempo, así como en relación con las variables externas más influyentes, que generalmente son el nivel de embalse y la temperatura.

A primera vista, estos gráficos permiten tener una idea de los rangos de variación de las variables, el volumen de datos disponible, y las lagunas en las series de datos. Si además se lleva a cabo por un técnico experto y cualificado, pueden detectarse cambios claros de tendencia en las medidas, y tener una primera idea de si el comportamiento responde a lo esperado.

Las herramientas utilizadas para la generación de estos gráficos suelen ser las ofimáticas convencionales. En ocasiones se han desarrollado también herramientas específicas que incorporan ciertas funcionalidades, como la indicación de determinadas incidencias, ya sean del sensor o generales de la presa [6].

Los avances informáticos permiten de una manera fácil generar gráficos muy flexibles e interactivos. En el curso de la investigación, se han integrado diversas tecnologías para generar una herramienta de exploración de datos interactiva y adaptable, que permite:

1. Seleccionar las variables que se quieren visualizar en cada eje del gráfico.
2. Seleccionar qué variables utilizar para determinar el color y el tamaño de la visualización.
3. Navegar de forma dinámica sobre el gráfico, ampliando las regiones de interés en cada uno de los ejes.
4. Acceder al gráfico e interactuar con él desde cualquier dispositivo con acceso a internet, en cualquier momento.

La Figura 1 muestra una imagen de la aplicación. En el menú de la izquierda se pueden seleccionar las variables que se quieren representar en los ejes, así como las utilizadas para definir el tamaño y el color de los puntos.

En la Figura 2 se incluye un ejemplo de visualización dinámica de series temporales, con dos niveles de aumento diferentes.

Estas herramientas gráficas permiten además tomar decisiones para la generación de modelos de predicción. Como ejemplo, en la Figura 1 se observa la disminución del caudal de filtración (eje de ordenadas) con el tiempo (mapa de color) y su variación cuasi lineal con el nivel de embalse (eje de abscisas) para los registros más recientes. A la vista de este gráfico parece razonable restringir la selección a los datos más recientes para después ajustar una regresión lineal dependiente del nivel de embalse.

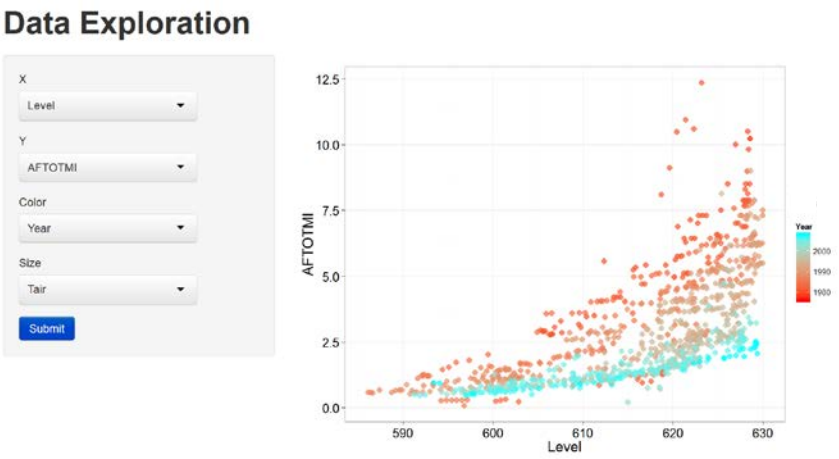


Figura 1. Aforo de filtración en función del nivel de embalse. Los colores se corresponden con el año de registro. Se observa que el caudal ha disminuido con el tiempo, y que su relación con el nivel de embalse es sensiblemente lineal en el periodo más reciente.

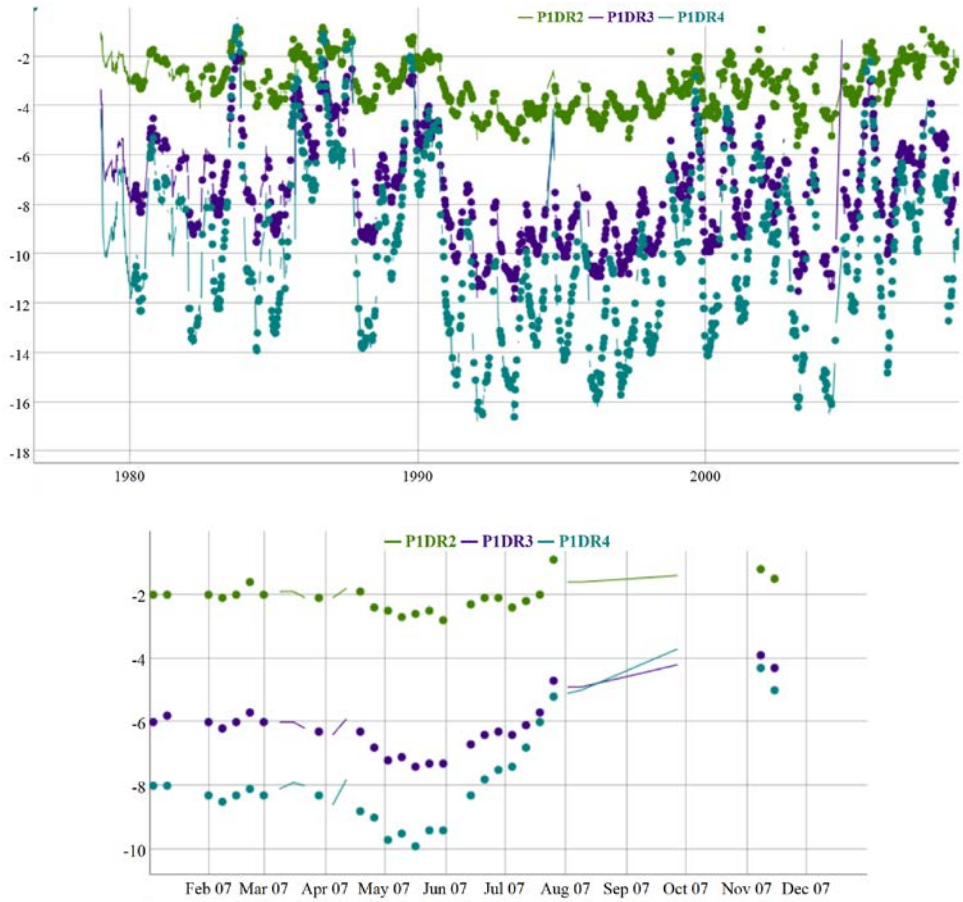


Figura 2. Interfaz de la aplicación para exploración dinámica de datos. Se muestra la evolución temporal de tres movimientos en péndulos. Arriba: series completas. Abajo: ampliación de los datos del año 2007.

3. INTERPRETACIÓN DEL COMPORTAMIENTO CONJUNTO. LAS REDES COMPLEJAS

La aplicación de la Teoría de Redes Complejas al análisis de los registros de auscultación se plantea como una posibilidad adicional dentro de la exploración de datos permitiendo el alcance de un punto de vista global o sistémico, a diferencia de otros procedimientos exploratorios más específicos (como los anteriormente descritos). Es posible definir la red compleja a partir de las variables o las series registradas en cada aparato (que serán los elementos o nudos de la red), mientras que las posibles conexiones entre las parejas de nudos se obtendrán a través de la comparación de los datos contenidos en cada una de las series, mediante la fijación de un determinado criterio. En las primeras aplicaciones [7], el criterio escogido para valorar la intensidad de la relación entre cada pareja de series fue el valor del mejor coeficiente de determinación resultante del ajuste de polinomios de segundo grado a los datos coincidentes en fecha entre ambos aparatos (se recuerda que el citado coeficiente establece una medida normalizada de la bondad del ajuste obtenido).

Una vez establecida la red compleja, la Teoría comprende diferentes técnicas que permiten efectuar una descriptiva del sistema:

- Medidas de centralidad para valorar la importancia de cada nudo o serie dentro del conjunto conforme a diferentes criterios.
- Algoritmos de partición o de clasificación para la detección de conjuntos homogéneos de aparatos.
- Caracterización de la red, de forma que puedan establecerse paralelismos entre fenómenos de diversos campos cientifico-técnicos.

Otro de los aspectos esenciales de las redes complejas es su representación en forma de grafos (conjuntos de nudos y lazos), donde la posición de los elementos puede definirse mediante diferentes algoritmos de distribución, y en los que es posible establecer mapas de colores y distribuciones de tamaños asociados a los nudos o a los lazos, aumentando considerablemente la información que se visualiza en una sola imagen. También es posible representar la red compleja tomando como base la ubicación real de los aparatos sobre la presa. Todas estas posibilidades de representación pueden combinarse con un enfoque dinámico o evolutivo de la red, para detectar así pautas de comportamiento cronológico o derivas temporales, así como anomalías.

De este modo, las redes complejas y sus representaciones en forma de grafo pueden servir como un soporte de visualización global del sistema presa de gran elocuencia, que además puede ser descompuesto o sometido a diferentes criterios de filtrado para la realización de análisis pormenorizados por parte del ingeniero experto. (En relación a la cantidad de información representada, se recuerda aquí que, si el número de nudos del sistema es de orden n, el número de lazos o conexiones representadas será de orden $n^2/2$).

Actualmente se está explorando la posibilidad de utilizar las redes complejas así obtenidas como soporte para la selección de variables de entrada para modelos de predicción, gracias a que los diferentes algoritmos de partición pueden ayudar a evitar la multi-colinealidad así como a enriquecer la información de entrada empleada para la predicción.

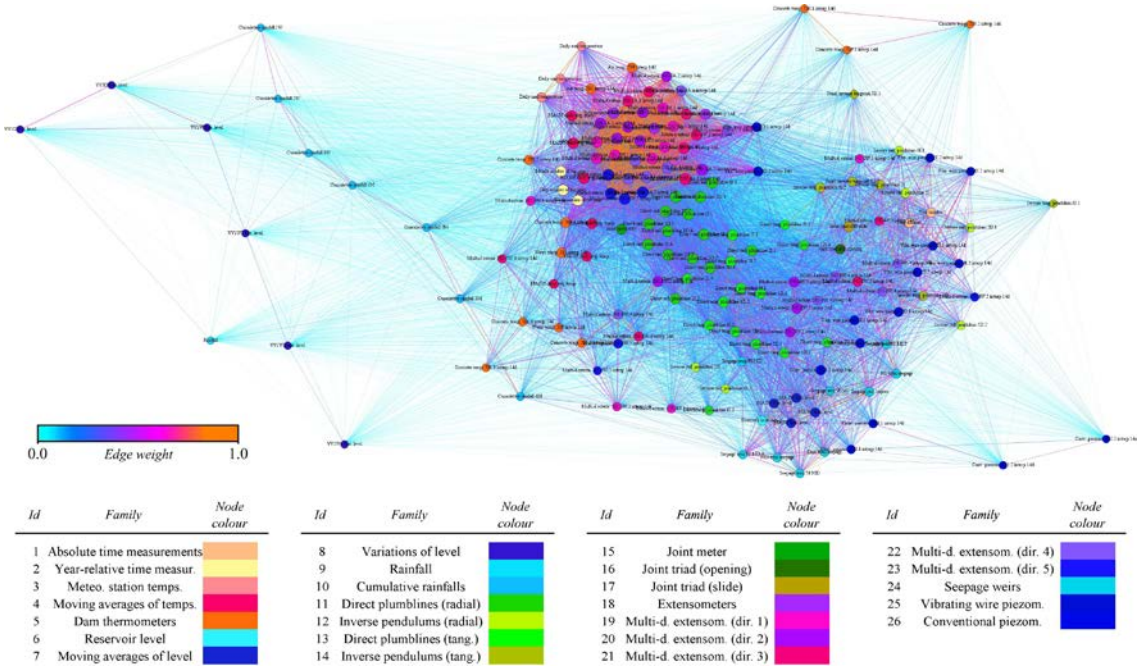


Figura 3. Ejemplo de representación de una red compleja obtenida a partir del análisis de los datos procedentes de una presa bóveda. Los colores de los nudos representan en este caso el tipo de aparato o variable. Los colores de las conexiones dan una idea de la intensidad de la relación entre los nudos que une. El algoritmo de distribución sobre el plano tiene en cuenta los grupos de aparatos densamente interconectados.

4. MODELOS DE PREDICCIÓN NO CONVENCIONALES

4.1. INTRODUCCIÓN

Los sistemas de aviso son claves en el sistema de auscultación. Con frecuencia, se implementan herramientas que lanzan un aviso cuando se registra un valor que queda fuera de un rango determinado previamente. Dicho rango de comportamiento "normal" se determina a partir de un modelo de predicción, que estima el valor más probable de la variable en cuestión en función de las variables exteriores de la presa: fundamentalmente, temperatura, nivel de embalse, y edad de la estructura.

Estos modelos suelen basarse en relaciones estadísticas sencillas de las variables mencionadas, siendo el método "HST" el más común. Se trata de una regresión lineal múltiple, de modo que la variable objetivo se calcula como una combinación lineal de:

- Varias potencias del nivel de embalse
- La temperatura ambiente, que suele considerarse como una función senoidal.
- Una función del tiempo, que considera los efectos no reversibles.

Este modelo se desarrolló originalmente para calcular el movimiento en presas bóveda, si bien se ha aplicado con algunas variantes para otro tipo de variables, como, por ejemplo, el aforo de filtraciones [8].

El método presenta limitaciones, según han puesto de manifiesto investigaciones recientes [5], [9]. Las principales son:

1. Asume que las variables son independientes, lo cual no siempre se cumple.

2. Asume que las relaciones entre las variables exteriores y la respuesta de la presa son lineales, lo cual tampoco es cierto con carácter general.

Para solventar estas limitaciones, se han comenzado a utilizar otras técnicas de generación de modelos, en general más flexibles, y que por tanto se permiten modelar relaciones no lineales y variables dependientes. Las más comunes son las redes neuronales, como muestra la cantidad de estudios presentados recientemente (por ejemplo, [10]).

Estos modelos no hacen suposiciones a priori sobre las relaciones entre variables. Por el contrario, pueden adaptarse y "descubrir" las interacciones existentes entre las series de datos. Como contrapartida, deben utilizarse con precaución, para evitar el efecto denominado "sobreajuste", por el cual el modelo aproxima muy bien los datos de entrenamiento (los utilizados para ajustar los parámetros del modelo), pero no tienen buena capacidad de generalización. Es decir, el error de predicción aumenta considerablemente al aplicarlo a un caso diferente de los de entrenamiento. Un procedimiento aplicable siempre, de gran utilidad para evitar este problema, consiste en reservar una parte de los datos disponibles (datos de test), que no se utilizan para ajustar los parámetros del modelo. Al contrario, se compara la predicción del modelo en el periodo de test con los valores observados. Un aumento del error en este periodo con respecto al registrado en los datos de entrenamiento denota sobreajuste.

En el curso de la investigación, se ha realizado un estudio comparativo entre algunas de estas herramientas, enmarcadas en el campo de la inteligencia artificial [11]. Los resultados muestran que las redes neuronales ofrecen en general mayor precisión que el modelo HST, y también que otras técnicas permiten obtener una precisión incluso mayor en muchos casos.

La idea es que modelos más precisos permitirán definir con mayor precisión los umbrales de comportamiento normal, y por tanto unos niveles de aviso más realistas.

4.2. Implementación

Se ha desarrollado una aplicación que permite generar modelos de predicción basados en distintas técnicas (incluso las convencionales y las redes neuronales), con diferentes valores de los parámetros que las definen.

Como ejemplo, para construir una red neuronal deben definirse el número de neuronas y los parámetros que controlan el proceso entrenamiento. La aplicación permite ajustar un modelo de este tipo, bien con unos parámetros por defecto (que suelen funcionar aceptablemente bien en el caso general), para que el usuario no iniciado pueda hacer pruebas, bien particularizándolos, para comprobar cómo afectan los cambios a la precisión del modelo.

Se incluye también documentación explicativa, en forma resumida, sobre a) las bases de funcionamiento del modelo, b) criterios para la definición de los

parámetros, y c) enlaces a textos explicativos de la base matemática de cada herramienta.

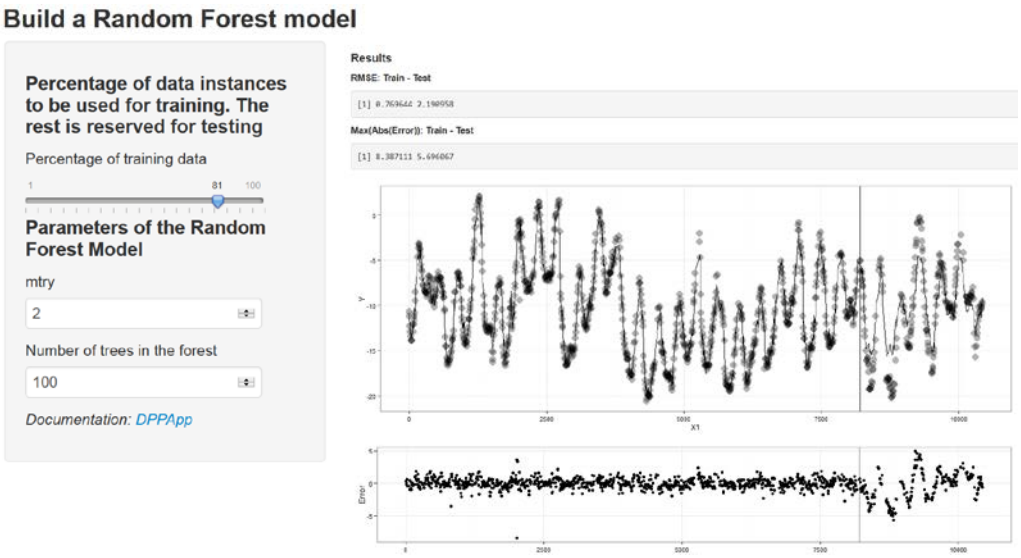


Figura 4. Aplicación para el cálculo de un modelo de predicción basado en bosques aleatorios. Los controles de la izquierda permiten particularizar la cantidad de datos utilizados para el entrenamiento, y los parámetros que definen el modelo (en este caso, el número de árboles y el parámetro "ntry"). Se incluye también un enlace a una documentación resumida, con criterios a seguir y enlaces a textos clave para entender la base matemática y la aplicación de cada modelo. En el ejemplo mostrado se observa sobreajuste del modelo, como indica el aumento claro del error entre los datos de entrenamiento (a la izquierda de la línea vertical) y los de test (no empleados para el ajuste del modelo).

Como en el ejemplo anterior, estas aplicaciones se han diseñado para ser accesibles a través de internet, desde cualquier dispositivo. La figura 4 muestra una imagen de la herramienta para la generación de modelos basados en bosques aleatorios [12].

5. CONCLUSIONES Y LÍNEAS DE INVESTIGACIÓN FUTURAS

Existen multitud de técnicas desarrolladas en diversos campos, generalmente alejados de la ingeniería civil, que han demostrado su utilidad para la interpretación y predicción del comportamiento de sistemas complejos. Su aplicación a sistemas de ingeniería civil ha sido relativamente escasa, muy escasa en el caso de las presas.

Es necesario considerar la complejidad del comportamiento del conjunto presa-terreno, y admitir las posibilidades que ofrecen los métodos no convencionales para comprender mejor la respuesta de la estructura. Se trata de combinar las técnicas habituales con otras nuevas en el campo de la ingeniería de presas, pero ya maduras gracias a su aplicación a diversos ámbitos del conocimiento.

Las posibilidades que ofrecen las técnicas de inteligencia artificial son enormes. Algunas de las que se han identificado, y que se prevé explorar en el futuro, son las siguientes:

- Selección de variables flexible: en algunos casos, se añaden al modelo variables derivadas de las exteriores, como por ejemplo la media móvil del nivel de embalse, o su velocidad de variación [13], en base a la intuición del modelador. Las herramientas de inteligencia artificial permiten seleccionar de forma automática las variables más útiles, y descartar las poco importantes [10].

- Generación de modelos no causales: pueden utilizarse algunas de las variables de respuesta de la presa para predecir otras (por ejemplo, calcular el movimiento en un péndulo a partir de otro). Ello puede reducir la utilidad del modelo para explicar el comportamiento, pero aumentar su aptitud para detectar determinadas anomalías.

- Modelos autorregresivos: son los que basan la predicción en el valor registrado en instantes de tiempo previos. También se llaman modelos de corto plazo, porque suelen utilizarse para predicción "paso a paso". Su precisión disminuye en general para predecir a largo plazo, porque el error se propaga hacia adelante en el tiempo. Pueden ser útiles sin embargo en combinación con otro tipo de modelos.

- Interpretación del comportamiento de la presa. Muchos de estos modelos permiten calcular un índice de importancia de las variables, que normalmente se basa en criterios empíricos (miden cuánto aumenta el error al eliminar cada una de las variables por separado). La interpretación de estos resultados no es tan directa como la de un modelo lineal, donde los coeficientes correspondientes a cada variable pueden compararse directamente y representan la variación de la variable respuesta al modificarse cada una de las variables predictoras. Sin embargo, pueden ser de gran utilidad como apoyo para el encargado de seguridad de la presa, en combinación con su conocimiento previo de la estructura y su comportamiento.

Finalmente hay que destacar dos ideas fundamentales:

1. la inspección visual es irreemplazable y debe siempre formar parte de un sistema de auscultación

2. la aplicación de estas herramientas y la interpretación de sus resultados debe llevarse a cabo por profesionales con formación y experiencia suficientes.

4. REFERENCIAS

[1] Santillán, D. (2014). Mejora de los modelos térmicos de las presas bóveda en explotación: aplicación al análisis del efecto del cambio climático. Tesis Doctoral, E.T.S.I. Caminos, Canales y Puertos (UPM).

[2] International Commission on Large Dams (2012). Dam surveillance guide. Boletín 158, ICOLD.

[3] Chouinard, L. y Roy, V. (2006). Performance of statistical models for dam monitoring data. In Joint International Conference on Computing and Decision Making in Civil and Building Engineering, Montreal. 2006 14–16.

[4] De Sortis, A. y Paoliani, P. (2007). Statistical analysis and structural identification in concrete dam monitoring. Engineering structures. 29(1):110–120.

[5] Lombardi, G. (2004). Advanced data interpretation for diagnosis of concrete dams. Technical report, CISM, 2004.

[6] Mora, J., López, E. y de Cea, J.C. (2008). Estandarización en la gestión de datos de auscultación e informe anual en presas de titularidad estatal. VIII Jornadas Españolas de Presas

[7] Morera, L. (2014). A systemic approach to dam data analysis using Complex Networks Theory. Tesina Fin de Máster Universitario en Sistemas de Ingeniería Civil. Universidad Politécnica de Madrid, España.

[8] Simon, A., Royer, M., Mauris, F. y Fabre, J. (2013). Analysis and interpretation of dam measurements using artificial neural networks. 9th ICOLD European Club Symposium. Venecia, Italia.

[9] Tatin, M., Briaut, M., Dufour, F., Simon, A. y Fabre, J. (2013). Thermal displacements of concrete dams: Finite element and statistical modelling, in: 9th ICOLD European Club Symposium, Venecia, Italia.

[10] Santillán D., Fraile-Ardanuy J., Toledo M. Á. (2014). Predicción de lecturas de aforos de filtraciones de presas bóveda mediante redes neuronales artificiales. Tecnología y Ciencias del Agua, vol. V, núm. 3, mayo-junio, 2014, pp. 81-96

[11] Toledo, M. Á., Salazar, F., Morán, R., Morera, L., Roa, A. y Santillán, D. (2013). Interpretación de los datos de auscultación de presas por métodos no convencionales. Jornada Técnica sobre Avances en investigación aplicada en seguridad hidráulica de presas. CEDEX, junio de 2013.

[12] Salazar, F., Toledo M. Á., Oñate, E. y Morán, R. (2015). An empirical comparison of machine learning techniques for dam behaviour modelling. Structural Safety. Enviado.

[13] Sánchez Caro, F. J. (2007). Seguridad de presas: aportación al análisis y control de deformaciones como elemento de prevención de patologías de origen geotécnico. Tesis Doctoral, E.T.S.I. Caminos, Canales y Puertos (UPM).

B.3 Nuevas técnicas para el análisis de datos de auscultación de presas y la definición de indicadores cuantitativos de su comportamiento

Title: Nuevas técnicas para el análisis de datos de auscultación de presas y la definición de indicadores cuantitativos de su comportamiento

First Author: Fernando Salazar González. CIMNE - International Center for Numerical Methods in Engineering

Second Author: Miguel Á. Toledo Municio. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Third Author: Eugenio Oñate Ibáñez de Navarra. CIMNE - International Center for Numerical Methods in Engineering

Fourth Author: León Morera. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Fifth Author: Rafael Morán. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Conference: III Jornadas de Ingeniería del Agua. La precipitación y los procesos erosivos (JIA 2015)

Date-Location: October 2013 - Córdoba (Spain)

IV Jornadas de Ingeniería del Agua
La precipitación y los procesos erosivos
Córdoba, 21 y 22 de Octubre 2015

Nuevas técnicas para el análisis de datos de auscultación de presas y la definición de indicadores cuantitativos de su comportamiento

F. Salazar, E. Oñate
Centro Internacional de Métodos Numéricos en Ingeniería. Campus Norte UPC. Ed. C-1. Barcelona 08034

L. Morera, R. Morán y M.Á. Toledo
Universidad Politécnica de Madrid

1. Introducción

El control de la seguridad de presas es un aspecto fundamental de su explotación. El elemento clave es el sistema de auscultación, considerado como el conjunto de a) los aparatos instalados en la presa que registran determinadas variables, b) el sistema de transmisión de datos y c) una metodología para evaluar el estado de seguridad en función de los datos registrados. En la práctica, el sistema debe determinar si los datos registrados se corresponden con una situación normal, o si existe un riesgo de avería.

Para ello, se seleccionan los aparatos más representativos del comportamiento general de la estructura, y se definen unos umbrales que se corresponden con el estado de seguridad de la presa (situación normal o algún grado de emergencia), que se incluyen en el Plan de Emergencia de la Presa como indicadores cuantitativos del estado de seguridad (Ministerio de Medio Ambiente 2001).

La propia Guía Técnica para Elaboración de Planes de Emergencia de Presas incluye una clasificación de los métodos que pueden emplearse para definir los indicadores, señalando algunos de los pros y los contras de cada uno de ellos:

- Estadísticos, que establecen la relación entre las variables externas e internas a partir de las series de observaciones reales. Solo pueden aplicarse a presas en servicio “durante un periodo largo de tiempo”.
- Deterministas, que se basan en la modelización del comportamiento de la presa (típicamente con elementos finitos). Pueden aplicarse a presas de nueva construcción, pero solo si es posible “caracterizar con fiabilidad la realidad estructural de la presa y su cimiento y su comportamiento y, por tanto, los parámetros que lo condicionan”.

- Mixtos, que combinan los dos anteriores.

La Guía recomienda también (y los autores lo suscriben) aplicar el juicio ingenieril antes de declarar cualquier escenario de emergencia. Sin embargo, del mismo modo establece que "cualquier cambio significativo que pueda tener un efecto negativo sobre la seguridad y no sea explicable directamente en función de otros parámetros (nivel de embalse, temperatura, etc.) debe considerarse causa suficiente para la declaración del Escenario 0 de control de la seguridad".

En la práctica, es relativamente frecuente que ocurran errores de lectura que producen registros fuera de rango de situación normal. El responsable de explotación puede aplicar su juicio para establecer que dicha medición no constituye una anomalía y por tanto no requiere la declaración de un escenario de emergencia. Sin embargo, sería útil disponer de herramientas capaces de ayudar a tomar una decisión que en algunos casos puede ser comprometida.

Es difícil establecer umbrales a partir de modelos deterministas, porque la información disponible de la presa y el cimiento es muchas veces insuficiente para obtener resultados precisos. Por ello los indicadores cuantitativos suelen basarse en modelos estadísticos. Las herramientas de este tipo que se utilizan más habitualmente son relativamente simples, limitándose frecuentemente a la regresión lineal múltiple. El modelo más utilizado es el Hydrostatic-Season-Time (HST), basado en una combinación lineal del nivel de embalse, la estación del año y el tiempo de vida de la presa. En ocasiones se añaden variables derivadas de las registradas para aumentar la complejidad del modelo y hacerlo más flexible con el objetivo de aumentar la precisión de las predicciones.

En otros campos de la ciencia en que el volumen de datos disponible es mucho mayor, se han desarrollado herramientas para su tratamiento y para la generación de modelos de predicción. Algunas de ellas, como las redes neuronales, ya han sido aplicadas al caso de la auscultación de presas (Santillán *et al*. 2014, Mata, 2011, Simon *et al*. 2013, Salazar *et al.,* 2015a, Salazar *et al.,* 2015b) con resultados prometedores.

La aplicación de estas técnicas puede ayudar a mejorar la precisión de los modelos de predicción, y a entender mejor el comportamiento de la presa (Toledo *et al*., 2013). Un modelo más preciso puede ayudar además a definir umbrales de emergencia más fiables.

El principal inconveniente de estos modelos es que son más difíciles de interpretar que el HST. Mientras que en éste puede extraerse directamente la contribución de cada una de las acciones (nivel de embalse, temperatura, tiempo) en la respuesta de la presa, en otros modelos más complejos los efectos no son aditivos, suelen ser no-lineales, y basarse en un conjunto de variables mayor. Todo ello ha provocado que sean calificados con frecuencia como modelos de "caja negra" (p.e. Olden y Jackson, 2002). La mayor parte de los trabajos publicados sobre la aplicación de estas técnicas en seguridad de presas se limitan a evaluar el error de predicción. No se interpreta el modelo obtenido y por tanto no se

pueden extraer conclusiones sobre el estado de seguridad de la presa. Como excepción, cabe citar los trabajos de Santillán *et al* (2014) y Mata (2011).

Nos encontramos por tanto ante una disyuntiva: el modelo HST es ampliamente conocido y utilizado y fácilmente interpretable. Sin embargo, se basa en unas hipótesis que no son ciertas, como que la temperatura es independiente del nivel de embalse (Tatin *et al*., 2013), y su precisión es limitada. Por otra parte, se dispone de unas técnicas más flexibles y precisas, pero más difíciles de implementar y analizar.

Ante esta situación, el presente trabajo tiene por objetivo proponer un método para interpretar un modelo de predicción del comportamiento de presas basado en bosques aleatorios (Breiman, 2001). El resultado de dicha interpretación se compara con el obtenido con el método HST.

2. Metodología

2.1. Caso de estudio

Se dispone de datos de auscultación de la presa de La Baells en el periodo 1.980-2.008. La figura 1 contiene el alzado esquemático y la sección de la presa, con la situación de algunos de los aparatos de auscultación. La descripción completa de los datos disponibles puede encontrarse en (Salazar *et al*., 2015a).

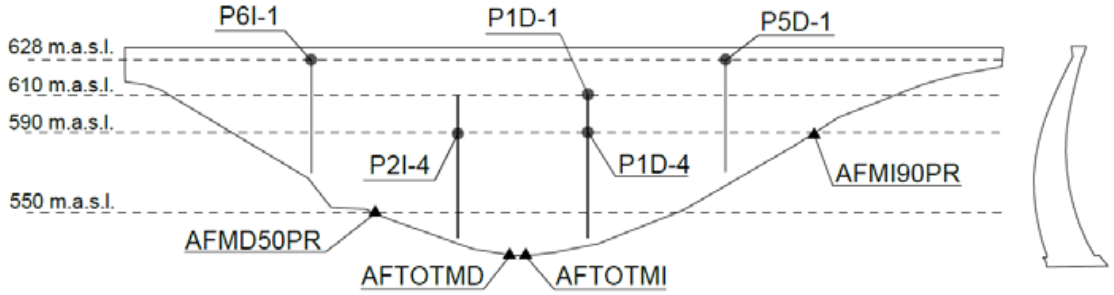


Figura 1. *Alzado esquemático y sección de la presa de La Baells. Sobre el primero se señala la situación de algunos de los aparatos de auscultación.*

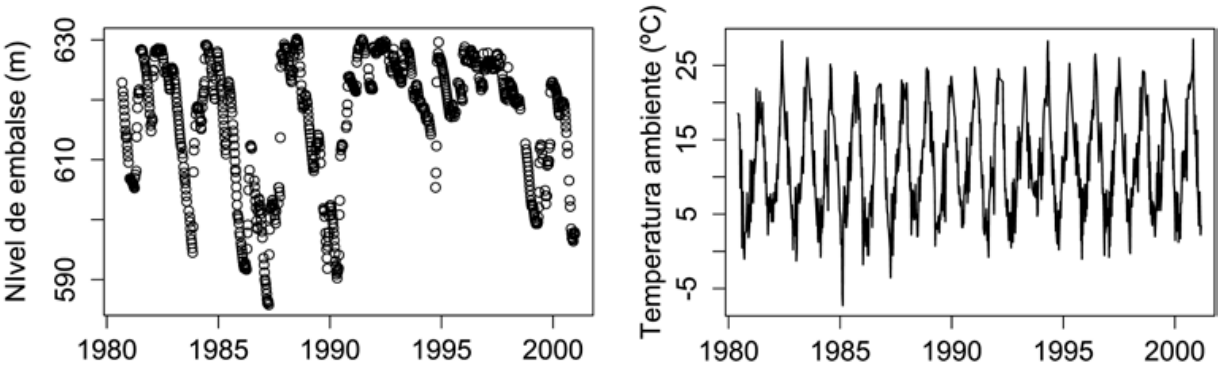


Figura 2. *Nivel de embalse y temperatura media diaria en el periodo de análisis*

Para el presente análisis se han seleccionado los datos de nivel de embalse y temperatura, para analizar los desplazamientos radiales registrados en el péndulo P1D1 (Figura 1). Se considera el periodo 1.980 – 2.000, durante el que se dispone de registros con frecuencia aproximadamente semanal (figura 2).

2.2. Método HST

La fórmula más habitual del modelo HST, que es la que se ha utilizado en este trabajo, es:

$$\hat{f}(h,s,t) = a_1 \cdot h + a_{21} \cdot h^2 + a_3 \cdot h^3 + a_4 \cdot h^4 + a_5 \cdot e^{-t} + a_6 \cdot t + a_7 \cdot \cos(s) + a_8 \cdot sen(s) + a_9 \cdot sen(2s) + a_{10} sen(s)\cos(s) \quad [1]$$

donde $\hat{f}$ es la función que define el comportamiento de la variable objetivo que se pretende evaluar, *h* es el nivel de embalse, *t* es el tiempo en días desde la puesta en carga de la presa, y *s* se calcula como:

$$s = \frac{d}{365{,}25} \cdot 2\pi \quad [2]$$

siendo *d* el número de días transcurridos desde el 1 de enero del año correspondiente.

Para identificar el efecto de cada una de las variables exteriores (temperatura, carga hidrostática y tiempo), basta con extraer los términos de [1] que dependen de cada una de ellas y dibujar su contribución parcial a la variable objetivo. Es decir, dibujar por separado *f(h)*, *f(s)* y *f(t)*.

2.3. Bosques aleatorios

Los bosques aleatorios (BA) pertenecen a un tipo de modelos basados en datos denominados “no paramétricos”. El motivo es que no se hace ninguna suposición a priori sobre el tipo de relación entre las variables exteriores (en nuestro caso, nivel, temperatura y tiempo) y la respuesta del sistema (desplazamiento radial). Esta característica contrasta con el método HST, en el que debe fijarse por ejemplo el orden del polinomio de la variable nivel de embalse, así como el tipo de función dependiente del tiempo.

Los BA aproximan el valor de la variable objetivo a partir del promedio de la predicción de un gran número (generalmente varios centenares) de modelos sencillos del tipo árbol de decisión. Son por tanto un *modelo de conjunto*. La descripción de la base teórica puede encontrarse en diversas fuentes (p.e. Hastie *et al*., 2009), así como en el artículo seminal de Breiman (2001). Se han publicado ejemplos de aplicación a seguridad de presas (Salazar *et al*, 2013).

Aunque su interpretación no es fácil, existen herramientas que permiten extraer información útil sobre el funcionamiento del sistema. La herramienta principal de interpretación de los modelos basados en BA es el índice de importancia de las variables.

Una vez construido el modelo, se permuta por separado cada una de las variables predictoras y se calcula el incremento en el error de predicción que supone. Se basa en la idea de que si una variable no afecta a la respuesta, la precisión dependerá poco de si se utiliza la serie original o la permutada. Al revés, el error aumenta más cuanto mayor es la importancia de la variable permutada. Esta medida se puede utilizar para seleccionar variables, si bien se ha demostrado que tiene sesgo en determinadas circunstancias (Strobl *et al*, 2008).

También puede extraerse información del modelo mediante los gráficos de dependencia parcial (Friedman 2000). En estos gráficos, para cada variable predictora, se selecciona un conjunto de valores distribuidos uniformemente a lo largo de su rango. Para cada uno de esos valores, se calcula la media de la predicción del modelo considerando los valores reales del resto de predictoras:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} f(x, x_{iC}) \qquad [3]$$

donde x es la variable cuyo efecto se quiere evaluar, y x_{iC} representan el resto de variables. Por ejemplo, si x es el nivel de embalse, se define un conjunto de valores equiespaciados a lo largo de la carrera de embalse $x_1, \ldots x_p$. Se sustituye la serie original del nivel de embalse por un valor constante $x=x_1$. Se calcula la predicción del modelo manteniendo el resto de variables con su valor original (x_{iC}), y se extrae la media de esas predicciones. Se repite el proceso para cada uno de los valores de $x=x_p$. Con ello se obtiene una serie de puntos que reflejan el efecto promedio de la variación del nivel de embalse en la predicción del modelo.

Los BA son la base de los denominados "quantile regression forests" (Meinshausen 2006), que permiten calcular la función de densidad de la variable objetivo, a partir de la cual pueden definirse intervalos de confianza de la predicción, que pueden ser útiles para definir umbrales de emergencia.

3. Resultados y discusión

3.1. Interpretación del comportamiento

La figura 3 contiene la contribución parcial al desplazamiento radial P1DR1 del tiempo, la temperatura y la carga hidrostática según el modelo HST. Se observa la relación no-lineal con el nivel de embalse, con desplazamiento hacia aguas abajo mayor con embalse lleno (sentido decreciente del eje *y*). Lo mismo ocurre en los meses fríos, como se comprueba en el gráfico parcial de la temperatura. Por último, el modelo identifica una deriva temporal hacia aguas abajo sensiblemente lineal de aproximadamente 0,2 mm/año.

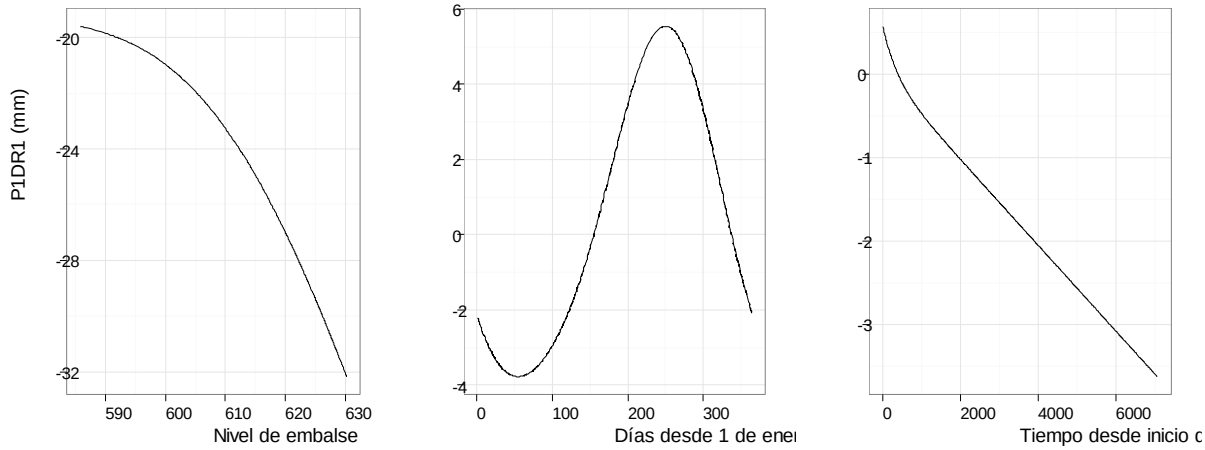


Figura 3. *Contribución parcial de las variables exteriores al desplazamiento radial según el modelo HST.*

Los gráficos de dependencia parcial que se obtienen del modelo BA se incluyen en la figura 4. Se observan tendencias muy similares en cuanto al efecto de la temperatura y el nivel de embalse. Sin embargo, la influencia del tiempo es cualitativamente diferente. En este caso, se registra una variación brusca sobre el año 1992, y un comportamiento sensiblemente constante entre ese momento y el año 2000.

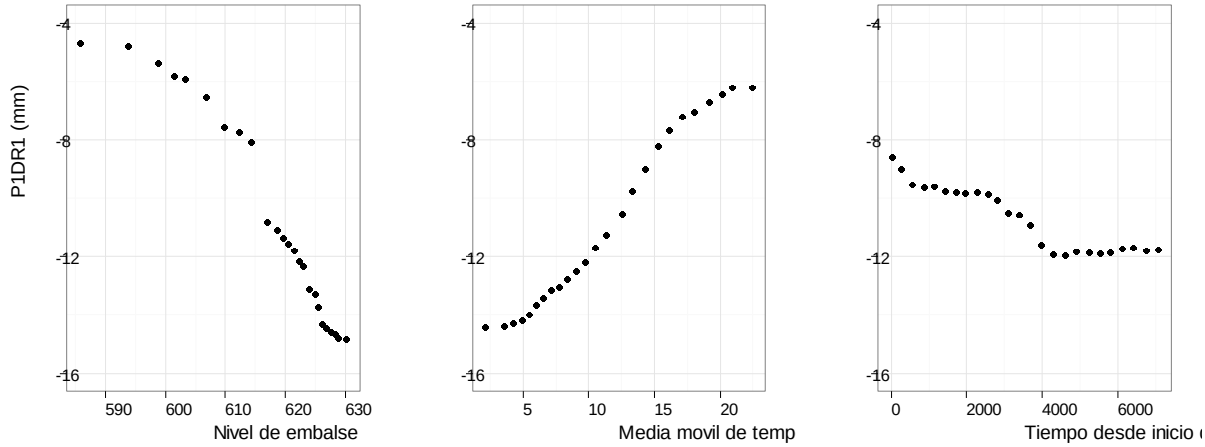


Figura 4. *Gráficos de dependencia parcial del desplazamiento radial en P1DR1 (datos registrados) calculados con el modelo BA*

La interpretación de estos resultados y su implicación con respecto del comportamiento de la presa son muy importantes y deben hacerse a partir de la máxima información sobre las incidencias en la estructura y en su sistema de auscultación.

Ya se ha mencionado que en el modelo HST debe definirse a priori la forma de la función dependiente del tiempo (en este caso, se ha optado por una combinación de exponencial negativa y lineal), mientras que en el modelo BA puede tener en principio cualquier forma.

Por lo tanto, podría pensarse que el comportamiento real de la presa es el que muestra el modelo BA, y que el resultado del HST se debe a la restricción impuesta a priori. En ese caso, la curva de contribución parcial del tiempo de la figura 3 sería el mejor ajuste del modelo HST al comportamiento real (escalón) que muestra la figura 4.

No obstante, cabe cuestionarse esta conclusión, ya que los gráficos de dependencia parcial no son equivalentes a los de la figura 3. Además, se da la circunstancia de que el nivel de embalse promedio en el periodo 1992-2000 fue sensiblemente mayor que en el periodo 1980-1992 (figura 2). Esta podría ser la causa del escalón registrado en la figura 4, que denota un desplazamiento mayor hacia aguas abajo en el periodo más reciente (coherente con el mayor nivel de embalse).

Para comprobar esta circunstancia, se ha realizado un experimento adicional. Se ha generado una serie de datos modificada del desplazamiento en P1DR1. Se ha obtenido introduciendo los valores reales de tiempo y nivel en [1], con los coeficientes calculados durante el ajuste del modelo HST, eliminando los términos dependientes del tiempo. Se obtiene por tanto una serie artificial, que representa cualitativamente la contribución del nivel de embalse y la temperatura, pero que es independiente del tiempo. Sí mantiene la particularidad de que la carga hidrostática en el periodo más reciente es mayor que en el periodo inicial. Se ha añadido al resultado un término aleatorio de valor medio nulo y desviación típica igual a 0,5 mm.

Se ha ajustado un modelo basado en BA para ajustar la serie artificial, y se incluyen los gráficos de dependencia parcial en las figura 5.

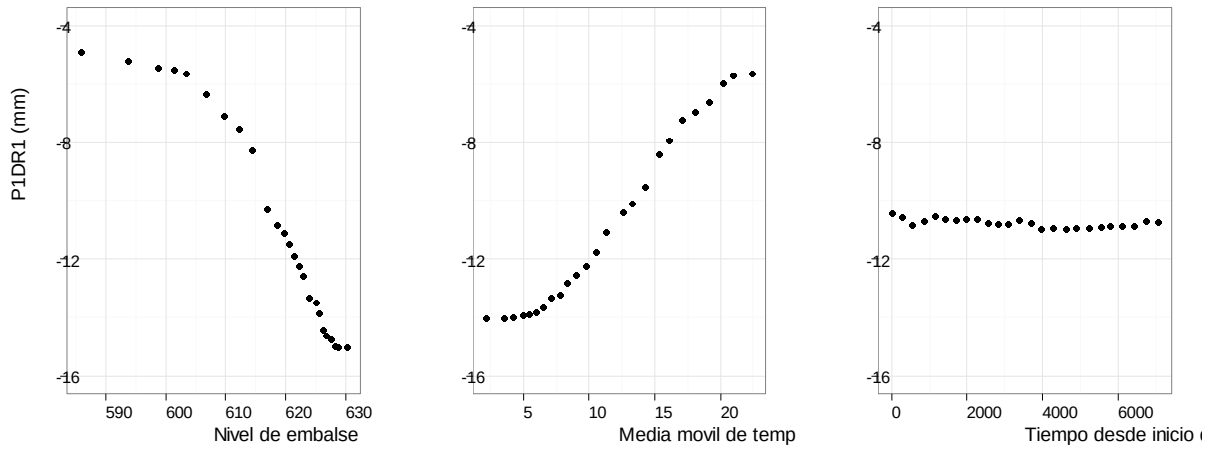


Figura 5. *Gráficos de dependencia parcial del desplazamiento radial $P1DR1_{HS}$ (serie generada artificialmente eliminando la influencia del tiempo), calculados con el modelo BA*

Se comprueba que la respuesta es prácticamente independiente del tiempo. Estos resultados confirman que el escalón registrado en el gráfico de dependencia parcial del desplazamiento con respecto del tiempo (figura 4) no se debe a que el nivel de embalse

promedio es mayor en el periodo 1992-2000. Ello demuestra la utilidad de los gráficos de dependencia parcial de modelos BA para interpretar el comportamiento de presas.

3.2. Definición de umbrales

La figura 6 muestra la predicción del modelo BA para el año 2000 con los datos reales registrados de nivel y temperatura (línea), junto con el rango del 95% de la función de densidad en cada punto (zona sombreada). Los datos registrados se representan por puntos. Se observa que los valores caen dentro del umbral así definido.

Si bien la declaración de un escenario de emergencia debe corresponder al responsable de seguridad de la presa, este tipo de gráficos pueden ser útiles para tal efecto. En particular, podría definirse una alerta que se activase al registrarse un valor fuera del rango de comportamiento normal del 95%.

Se trata de una prueba preliminar, pero muestra que el rango de confianza de la predicción del modelo varía dependiendo del valor de las variables exteriores.

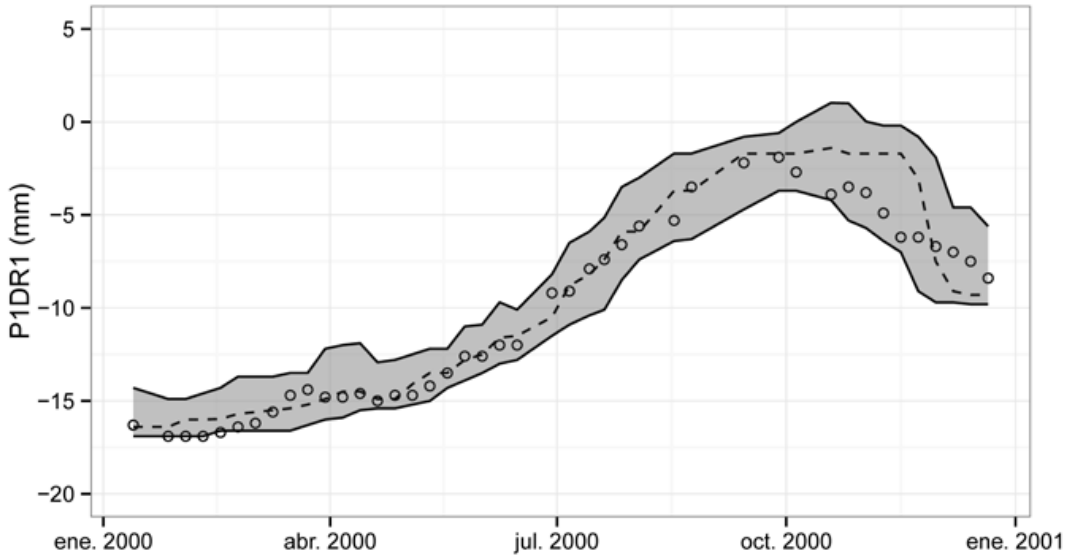


Figura 6. *Predicción del modelo BA (línea discontinua) y su rango de confianza del 95% (zona sombreada). Los datos registrados se muestran con círculos.*

4. Resumen y conclusiones

Los BA tienen gran flexibilidad para aproximar relaciones no lineales entre variables predictoras y respuesta. Aunque su complejidad dificulta su interpretación, los gráficos de dependencia parcial son una herramienta útil para identificar cambios en el comportamiento del sistema.

Su flexibilidad permite capturar variaciones de comportamiento de cualquier tipo. Esto supone una gran ventaja con respecto del modelo HST, en el que debe definirse a priori el

tipo de influencia entre cada variable predictora y la respuesta. El ejemplo presentado ilustra la ventaja que supone en la práctica. El análisis del modelo HST podría llevar a concluir que existe una deriva de la presa que se traduce en un desplazamiento hacia aguas abajo lineal con el tiempo y que dicha deriva no muestra síntomas de atenuación. Por el contrario, el modelo BA muestra que la variación temporal se produjo bruscamente, y que se ha estabilizado en los últimos años.

Otra conclusión importante es la posibilidad de asociar los umbrales de emergencia con funciones de densidad dependientes de la incertidumbre de las variables predictoras. Sobre esta línea se mantiene una investigación en marcha que pretende establecer criterios para la definición de dichos umbrales mediante funciones de probabilidad asociadas al modelo de comportamiento de la presa.

Agradecimientos

El trabajo ha sido financiado por el Ministerio de Economía y Competitividad de España (MINECO), a través de los proyectos iComplex (IPT-2012-0813-390000) y AIDA (BIA2013-49018-C2-1-R y BIA2013- 49018-C2-2-R).

Los autores desean expresar su agradecimiento a Carlos Barbero, de la Agencia Catalana del Agua, por facilitar los datos de auscultación de la presa de La Baells.

Referencias

Breiman, L. (2001). Random forests, *Machine learning* 45(1) (2001) 05-32.

Friedman, J.H. (2000). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29, 1189-1232.

Hastie, T., Tibshirani, R. y Friedman, J. 2009. The elements of statistical learning - Data mining, Inference and Prediction. Springer, 2ª edición.

Olden, J. D., Jackson, D. A. (2002). Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks. *Ecol. Model.* 154(1), 135-150.

Mata, J. (2011). Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Engineering Structures*, 33(3), 903-910.

Meinshausen, N. (2006). Quantile regression forests. *The Journal of Machine Learning Research*, *7*, 983-999.

Ministerio de Medio Ambiente. 2001. Guía Técnica para Elaboración de Planes de Emergencia de Presas.

Salazar, F., Oñate, E. y Toledo, M.Á. (2013). Posibilidades de la inteligencia artificial en el análisis de auscultación de presas. III Jornadas de Ingeniería del Agua.

Salazar, F., Toledo, M. A., Oñate, E., & Morán, R. (2015a). An empirical comparison of machine learning techniques for dam behaviour modelling. *Structural Safety*, 56, 9-17.

Salazar, F., Morán, R., Toledo, M.Á., Oñate, E. (2015b). Data-based models for the prediction of dam behaviour. A review and some methodological considerations. *Archives of Computational Methods in Engineering*. Doi: 10.1007/s11831-015-9157-9.

Santillán, D., Fraile-Ardanuy, J., Toledo, M.Á. (2014). Predicción de lecturas de aforos de filtraciones de presas bóveda mediante redes neuronales artificiales. *Tecnología y Ciencias del Agua*. Vol. V, núm. 3, mayo-junio de 2014, pp. 81-96.

Strobl, C., Boulesteix, A.L., Kneib, T., Augustion, T, Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics* 2008, 9:307

Tatin, M., Briffaut, M., Dufour, F., Simon, A., Fabre, J. (2013). Thermal displacements of concrete dams: Finite element and statistical modelling, 9th ICOLD European Club Symposium, Venice, Italy, 2013.

Toledo, M.Á., Salazar, F., Morera, L., Roa, A., Santillán, D. y Morán, R. 2013. Interpretación de los datos de auscultación de presas por métodos no convencionales. Jornada técnica sobre avances en investigación aplicada en seguridad hidráulica de presas. Madrid, junio de 2013. http://oa.upm.es/30175/1/INVE_MEM_2013_150013.pdf

B.4 A methodology for dam safety evaluation and anomaly detection based on boosted regression trees

Title: A methodology for dam safety evaluation and anomaly detection based on boosted regression trees

First Author: Fernando Salazar González. CIMNE - International Center for Numerical Methods in Engineering

Second Author: José M. González. CIMNE - International Center for Numerical Methods in Engineering

Third Author: Miguel Á. Toledo Municio. Technical University of Madrid (UPM). Department of Civil Engineering: Hydraulics, Energy and Environment.

Fourth Author: Eugenio Oñate Ibáñez de Navarra. CIMNE - International Center for Numerical Methods in Engineering

Conference: 8^{th} European Workshop on Structural Health Monitoring

Date-Location: July 2016 - Bilbao (Spain)

A methodology for dam safety evaluation and anomaly detection based on boosted regression trees

Fernando SALAZAR[1], José M. GONZÁLEZ[1], Miguel Á. TOLEDO[2], Eugenio OÑATE[1]

[1] Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE), UPC. Campus Nord UPC. Gran Capitán s/n, Barcelona 08034, SPAIN fsalazar@cimne.upc.edu, josem.gonzalez@cimne.upc.edu, onate@cimne.upc.edu

[2] Civil Engineering Department: Hydraulics, Energy and Environment, Technical University of Madrid (UPM) Profesor Aranguren s/n, Madrid 28040, SPAIN matoledo@caminos.upm.es

Key words: Statistical analysis, Dam safety, Machine learning, Anomaly detection.

Abstract

Many countries are implementing new dam safety regulations that often include more restrictive standards. This, together with the increasing average age of dams, results in a greater need for dam control and maintenance works. The advances in information and communications technologies improved the performance of dam monitoring systems, so a large amount of information on the dam behaviour can be collected. This has led to the use of more powerful tools for its analysis, many of which were first developed in the field of machine learning (e. g. neural networks). They offer some advantages over the conventional statistical methods. However, their capacity for early detection of anomalies has seldom been studied. As a result, they are far from being fully accepted by practitioners, whose analyses are often restricted to the interpretation of simple plots of time series data, together with basic statistical models. The present work describes a methodology for anomaly detection in dam behaviour, with the following features: a) The prediction model is based on boosted regression trees (BRTs). b) Causal and auto-regressive models are combined to detect different types of anomalies. c) It is checked whether the values of the external variables fall within the range of the training data. The performance of the proposed methodology was assessed through its application to a test case corresponding to an actual 100-m height arch dam, in operation since 1980. Artificial data were generated by means of a finite element model. Different anomalies were later added in order to test the anomaly detection capability. The method can be applied to other response variables and dam typologies, due to the great flexibility of BRTs, which automatically select the most relevant inputs.

1 INTRODUCTION

The advances in information and communication technologies have improved the performance of dam monitoring systems in terms of accuracy, reliability and reading frequency. This results in more comprehensive information about the behaviour of the structure [1].

The increase in the amount of information available led to the application of advanced tools for data analysis, most of which provide from the machine learning community, e.g. neural networks [2], support vector machines [3], the adaptive neuro-fuzzy inference systems

(ANFIS) [4], among others [5], [6].

Nonetheless, these tools have not been introduced among practitioners, who typically employ graphical data exploration [7], together with simple statistical models [1].

The vast majority of the examples of application of advanced tools focus on the development of a behaviour model that predicts the value of a given response variable (e. g. radial displacement). The prediction is compared with the actually observed data and some error index is extracted. In many cases the results are compared with those obtained by conventional statistical methods (e.g. [2]).

These advanced tools offer some advantages in terms of greater accuracy, flexibility, or ability to interpret the dam behaviour [8]. However, an accurate predictive model is just one of the necessary ingredients of an anomaly detection system. Some criterion needs to be developed to determine whether a given discrepancy between prediction and observation shall be considered as anomalous. This aspect was seldom considered, with a few exceptions for particular cases [9], [10], [11], which nonetheless were evaluated over a short validation period. Jung et al [10] and Mata et al [9] also generated artificial data with numerical models representative of abnormal situations.

In these work, a similar procedure is employed: a predictive model is built, and prediction intervals are derived from the standard deviation of the prediction error. This is the main ingredient of a methodology for anomaly detection in dam behaviour, with the following innovative features:

- The predictive model is based on boosted regression tress (BRTs from now on). This tool offered higher accuracy than other conventional and advanced tools in previous works [6].
- Both causal and auto-regressive models are assessed and the correspondent efficiency is compared in terms of anomaly detection capability.
- Artificial data are taken as reference, obtained with a numerical finite element (FE) model. It represents an actual dam currently in operation. The numerical results were compared to monitoring data to verify that they represent the actual dam behaviour.
- The value of the external variables was compared to the range of training data: during model application, abnormal values correspondent to extraordinary loads (e.g. reservoir level) are considered as due to lack of training data. This contributes to reduce the amount of false positives.

2 METHODS

2.1 Boosted regression trees

BRT models are built by combining two algorithms: a set of models are fitted by means of simple decision trees [12], whose output is combined to compute the overall prediction using boosting [13]. This tool was employed in previous works [8], where its main features were described. A more comprehensive introduction can be found in [14].

The main properties of BRTs are:

- They are robust against outliers.
- They require little data pre-processing.
- They can handle numerical and categorical predictors.
- They are appropriate to model non-lineal relations, as well as interaction among predictors.

All the calculations were performed in the R environment [15].

2.2. Prediction intervals

As mentioned above, a method for anomaly detection requires determining which magnitude of the discrepancy between prediction and observation is considered abnormal. In this work, the density function of error was computed and the normal interval defined as

$$[\hat{y}+\bar{e}+2\cdot sd_e,\ \hat{y}+\bar{e}-2\cdot sd_e] \quad (1)$$

where $\hat{y}$ is the model prediction, $\bar{e}$ is the mean error and sd_e is the error standard deviation. If the error density function follows a normal distribution, this margin contains 95% of the normal values. This criterion is heuristic, and was determined after some preliminary tests, considering both the true anomalies detected (true positives) and the normal situations considered abnormal (false positives).

However, a more relevant issue is the proper computation of the model prediction error. Since BRTs are non-parametric and typically feature a large number of parameters, they are susceptible to over fit the training data. It is well known that the training prediction error for machine learning tools results in an optimistic estimate of the actual model generalisation capability.

Cross validation is a conventional method to overcome this drawback. However, it cannot be directly applied to dam monitoring data, since they are time series: the training period shall be precedent to the validation interval. Moreover, the dam and its foundation feature behaviour changes over time in the general case.

We propose a method based on the hold-out cross validation described by Arlot and Celisse [16] for non-stationary time-series data. It comprises an iterative procedure:

- Take a minimum training period of 5 years. Build a BRT model, and compute the prediction errors for the sixth year.
- Build a new BRT model with 5+1 years of training data. Compute prediction errors for 7th year, and aggregate them to those obtained in the previous step.
- Repeat step 2 until a model is built with all the available data but the most recent year, and aggregate the prediction errors.
- Compute the density function of the aggregated-error and its statistics ($\bar{e}$ and sd_e). They are employed to define the interval for normal behaviour with Eq. (1).
- Build a BRT model with all the available data. Generate predictions and normal intervals for new (validation) data and compare to observations for anomaly detection.

2.3. Training range verification

Machine learning models typically produce highly inaccurate results when extrapolating, i.e., when new data falls outside the range of the training data set. In the case of dam monitoring, this situation corresponds with external loads above (or below) the maximum (minimum) value registered during its service life. Since dam response depends on several actions, also a combination of them may result in an “out-of-range” situation, even if none of the values are out of range when considered separately. To account for this issue, we chose the two principal environmental loads, i.e., hydrostatic (reservoir level) and thermal (air temperature), and build a two-dimensional density function via kernel density estimation [17]. The training sample with the lower density was computed, and its correspondent iso-

line plotted. New inputs falling outside this line are classified as out-of-range. In practice, they are not considered as anomalies even though the deviation between prediction and observation fell outside the normal interval. Figure 1 is an example of the density plots generated, with the training data, the iso-line and the validation data, part of which are out of range according to the described criterion.

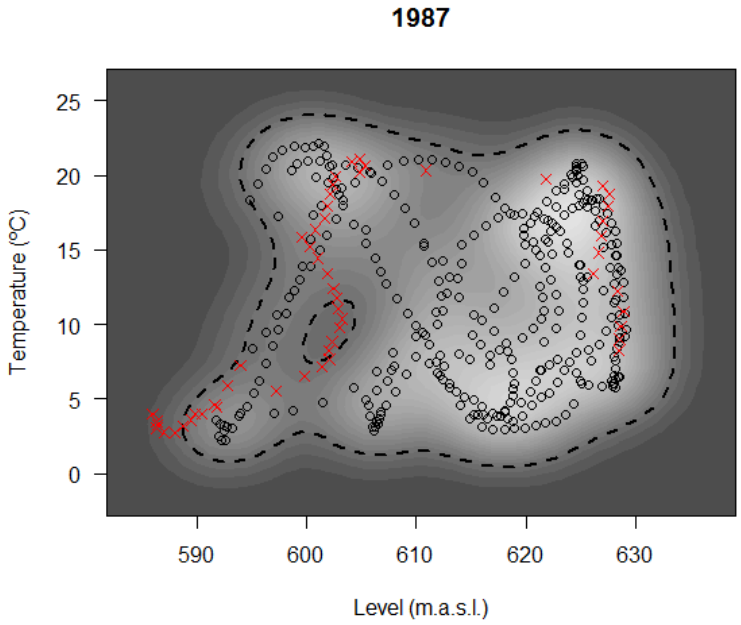


Figure 1: Example of density function for the main environmental loads. The circles represent the training set, and the red crosses the validation set. The dotted line is the iso-line with the lower density for the training set. It should be noted that some of the validation data are considered as out of range because they fall in low-density areas, even though they do not correspond to extraordinary high (or low) hydrostatic load (or air temperature).

2.4 Case study

La Baells dam is a double-curvature arch dam located in the Llobregat river, in the Barcelona region (Spain). The crest length is 403 m, whereas the maximum height above foundation is 102 m. Monitoring data were provided by the Catalan Water Agency for the period 1981-2008. These data correspond both to environmental and response variables. In this work, the air temperature and the reservoir level time series were considered as inputs to a finite element (FE) model. The results of this model in terms of radial displacements at the location of the pendulums were extracted and compared to the actual measurements. The objective was to check that the FE model could provide realistic data to generate reference time series of dam behaviour. This procedure allows obtaining data corresponding to the "ground truth", i.e., input-output time series without anomalies.

Time series data for hydrostatic load and air temperature were available for the period 1980-2008. Some derived variables were computed and considered as inputs, as described in previous works [5], [6] and summarised in Table 1. These are the inputs for the causal models.

The radial displacements measured at eight locations within the dam body were considered as outputs (Figure 2).

Moreover, a non-causal model was built for each output, taking as inputs both the environmental variables and all the outputs except that to be predicted in each case. Also the lagged output variables were included as inputs. This is, to predict the radial displacement $R(t_k,x_i)$ (at time t_k, location x_i), the following variables were added to those considered for the

causal model:

- $R(t_{k-1}, x_i)$
- $R(t_{k-2}, x_i)$
- $R(t_k, x_j; j \neq i)$
- $R(t_{k-1}, x_j; j \neq i)$
- $R(t_{k-2}, x_j; j \neq i)$

Hence, both a causal and a non-causal model are built for each output variable.

Id	Type	Period (days)
Level	Hydrostatic load-original	-
Lev007	Hydrostatic load-moving average	7
Lev014	Hydrostatic load-moving average	14
Lev030	Hydrostatic load-moving average	30
Lev060	Hydrostatic load-moving average	60
Lev090	Hydrostatic load-moving average	90
Lev180	Hydrostatic load-moving average	180
Tair	Air Temperature-original	-
Tair007	Air Temperature-moving average	7
Tair014	Air Temperature-moving average	14
Tair030	Air Temperature-moving average	30
Tair060	Air Temperature-moving average	60
Tair090	Air Temperature-moving average	90
Tair180	Air Temperature-moving average	180
NDay	Time	-
Month	Season	-

Table 1: External variables considered

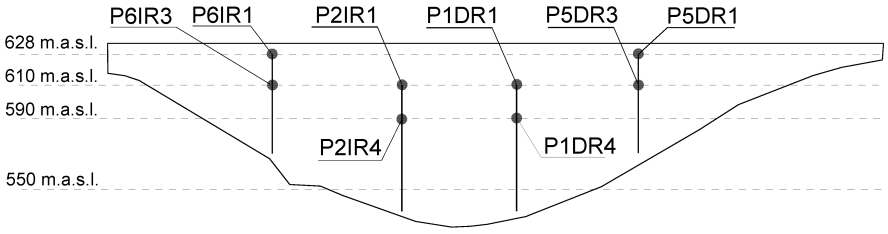


Figure 2: Location of the devices analysed within the dam body (view from downstream)

2.5 Anomalies

The results of the numerical model were artificially modified to simulate anomalies of two types and variable magnitude, which were introduced at a certain date:

- Offset, equivalent to adding a constant value (0.5, 1.0 or 1.5 mm) to the FEM model result.
- Incremental drift, where the added value is variable, growing linearly along time (0.5, 1.0 or 1.5 mm/year).

The abnormal period started at some random date between January 1st 1986 (5 years of minimum training period) and 2007 September 10th (one year before the end of the available period).

These criteria were applied to generate 1,000 abnormal time series, each one with random values of a) the target variable (among those depicted in Figure 2), b) the type and magnitude of anomaly, and c) the initial date of the abnormal period.

Each test case was presented both to the causal and the non-causal models. The time lapse since the initiation of the abnormal behaviour to the first observation identified as anomaly by each model was registered as "detection time". The test period was limited to two years of abnormal behaviour. If the correspondent model did not identify any observation as abnormal during that period, the detection time was set to 730 days (2 years) as regards results analysis.

Also the amount of false positives (regular values considered anomalous) were computed for each model and output. It should be reminded that the out-of-range instances (according to the criterion described in section 2.3) were not considered as anomalies.

3. RESULTS AND DISCUSSION

Figure 3 shows an example of application of the methodology. The vertical line was plotted at the initial date of anomaly. The shaded area represents the normal range (Eq. 1), and the anomalous values are depicted with red asterisks. The blue hollow circles represent out-of-range input data.

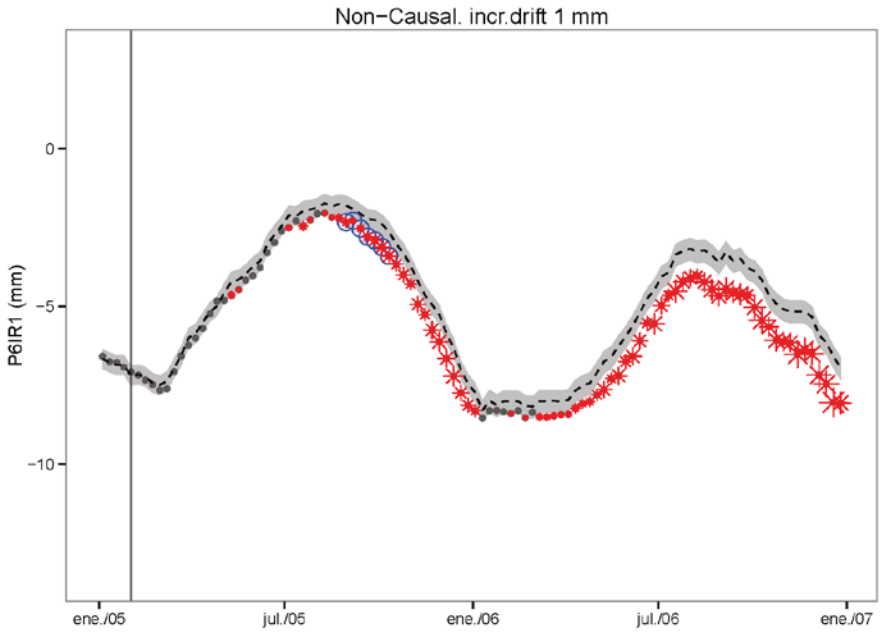


Figure 3: Anomaly detection. Example of application. Blue hollow circles depict out-of-range values. Red asterisks show observations out of the normal area (shaded). The vertical line is located at the starting of the abnormal interval.

An increasing deviation between observations and predictions can be clearly identified, correspondent to the artificial anomaly introduced. The model is re-trained at the beginning of each year, and the observations in the precedent year are added to the training data set. In this case, the model considers them as normal response, and partially adapts to the abnormal

behaviour. This is the reason why the observations in January 2006 fall within the normal interval. However, the anomalous behaviour is again detected in February 2006. In practice, the re-training at January 2006 should be modified accounting for the abnormal behaviour previously identified.

Table 2 shows some statistics of the performance of both models across the 2,000 cases analysed. The non-causal model outperforms the causal one for all the indicators considered.

Model	Undetected anomalies	Detection time. Mean/median (days)	False positives (average per year)
Causal	20%	276/186	0.85
Non causal	10%	163/72	0.56

Table 2: Results of the performance of causal and non-causal models

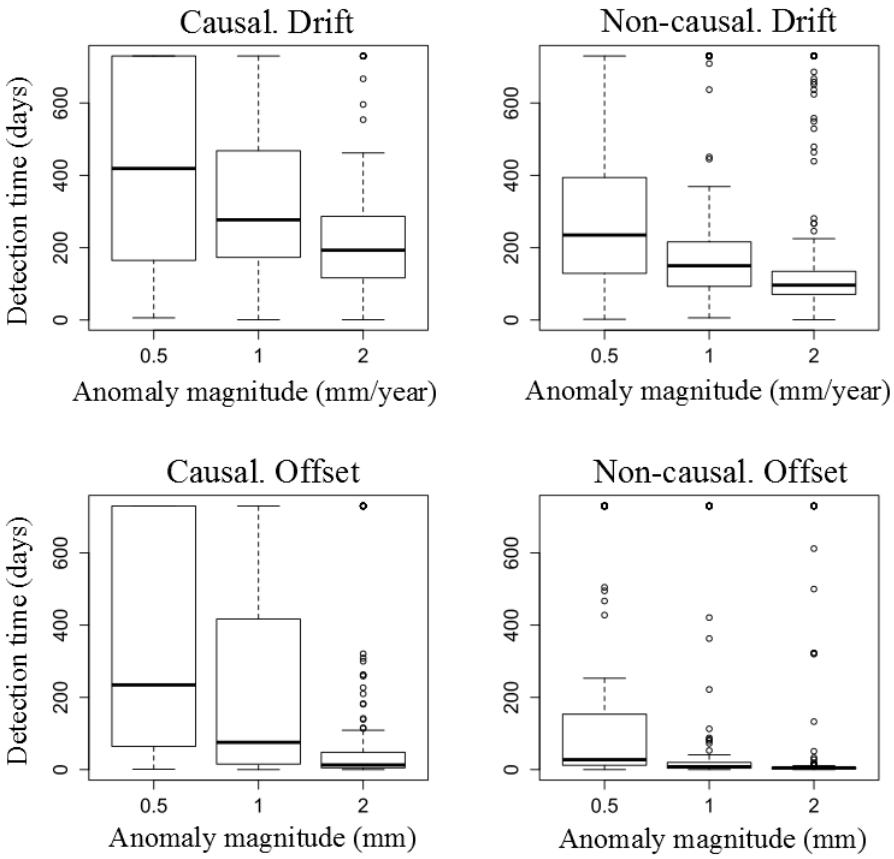


Figure 4: Detection time per type of anomaly, predictive model and magnitude. The non-causal model performs better, especially for the "offset" type.

The detection time per type of model and anomaly is depicted in Figure 3. It is lower for the non-causal model in all cases. As expected, the anomalies with lower magnitude are harder to detect, which results in a higher detection time (it should be reminded that the detection time is limited to 730 days). Also, the "incremental drift" anomalies require a longer time to be detected than the "offset".

Figure 4 shows the results per type of model and anomaly, as well as per year of initiation of the abnormal behaviour. The latter factor determines the size of the training set, which in turn affects the model accuracy and its ability to detect anomalies.

For the “offset” anomaly, the performance of both models neatly improves for later date of anomaly initiation. This effect is remarkable for the non-causal model, whose median detection time is close to zero for anomalies starting after 1995. The tendency is less clear for the “incremental drift” anomaly.

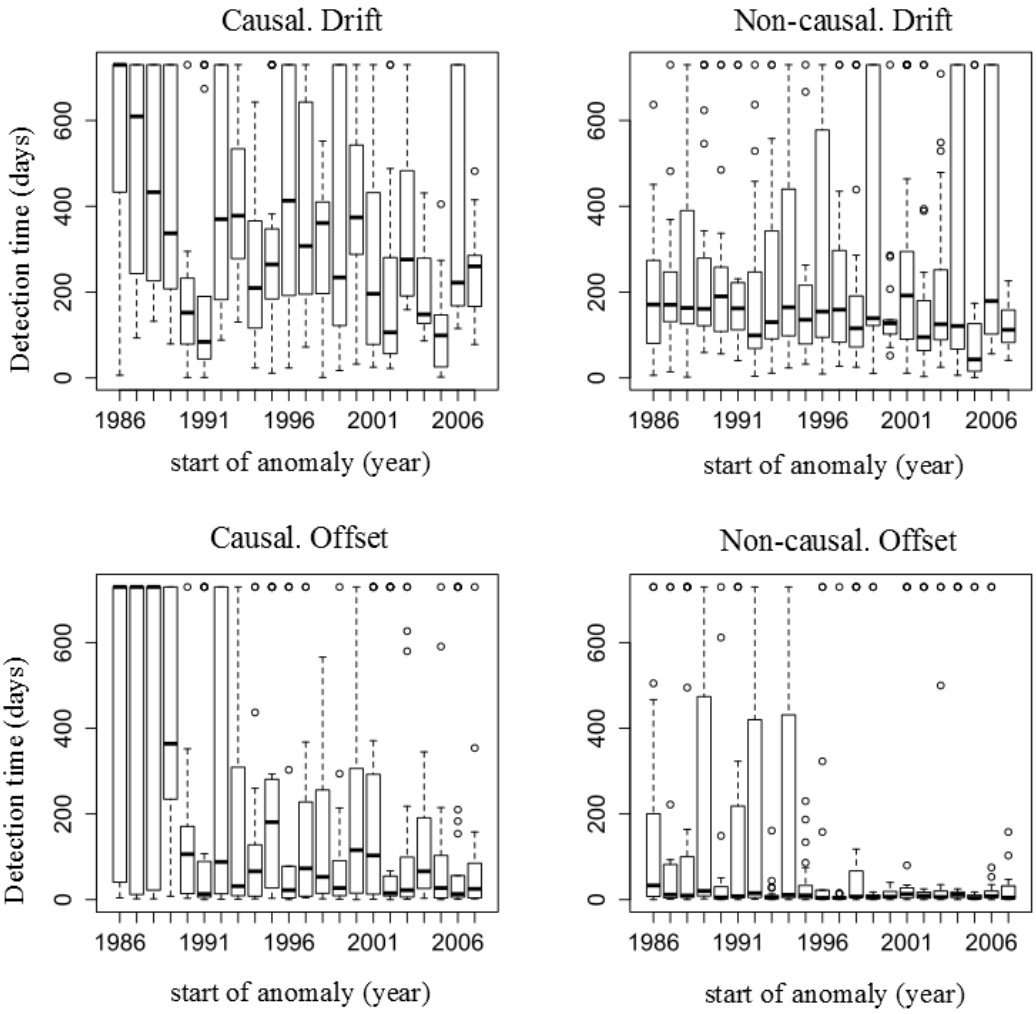


Figure 5: Detection time per year of anomaly initiation, type of model and anomaly. The performance is better along time for both models and “offset” anomalies, whereas the tendency is less clear for “incremental drift”.

9 SUMMARY, CONCLUSIONS AND FUTURE WORK

A methodology for anomaly detection in dam behaviour based on BRTs was presented. It is based on a criterion for defining a range of normal behaviour, based in turn on the model prediction and the statistics of the training error.

The occurrence of extraordinary loads is accounted for by computing a density function of the most relevant input variables (hydrostatic load and air temperature) via kernel density estimation. This reduces the amount of false positives due to lack of model accuracy for extrapolation.

Causal and non-causal models were compared, as regards their capability for detecting anomalies in radial displacements of an arch dam. Artificial anomalies were generated by adding certain values to the dam response, as computed by means of a FE model. The non-causal model showed better performance: fewer false positives, more anomalies detected and

lower detection time. This is due to its higher accuracy, which results in narrower intervals for normal behaviour.

However, non-causal models rely on response variables to predict each output, i.e., the predictions for the radial displacement at a given location and time is based on other displacements, as well as on the previous values of the variable to predict. This may lead to poor performance when the dam undergoes abnormal behaviour affecting several devices. We are currently working on this issue, by means of applying the same methodology to more realistic abnormal data: the boundary conditions of the numerical model are modified to reproduce hypothetical, though feasible, anomalies, which are reflected in several output variables.

In any case, these techniques should be used as a tool to provide detailed and accurate information to the dam safety managers, rather than as a totally automatic detection system. All relevant decisions influencing dam safety shall be made by an expert and capable engineer, based on the analysis of all the relevant information available. In this context, plots as that depicted in Figure 3 can be highly valuable: they allow visually identifying the occurrence of a deviation from normal behaviour (increasing along time in the example presented). Thus, the automatic system can be used as an indicator to generate a warning which leads to intensify the dam safety monitoring.

ACKNOWLEDGEMENTS

The authors thank Carlos Barbero, dam safety manager at the Catalan Water Agency (ACA), for providing the information regarding La Baells dam.

The research was supported by the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad, MINECO) through the projects iComplex (IPT-2012-0813-390000) and AIDA (BIA2013-49018-C2-1-R and BIA2013- 49018-C2-2-R).

REFERENCES

[1] International Commission on Large Dams (2012). Dam surveillance guide. Bulletin 158, ICOLD

[2] Mata, J. (2011). Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models. *Engineering Structures*, *33*(3), 903-910.

[3] Ranković, V., Grujović, N., Divac, D., & Milivojević, N. (2014). Development of support vector regression identification model for prediction of dam structural behaviour. *Structural Safety*, *48*, 33-39.

[4] Ranković, V., Grujović, N., Divac, D., Milivojević, N., & Novaković, A. (2012). Modelling of dam behaviour based on neuro-fuzzy identification. *Engineering Structures*, *35*, 107-113.

[5] Salazar, F., Morán, R., Toledo, M. Á., & Oñate, E. (2015a). Data-Based Models for the Prediction of Dam Behaviour: A Review and Some Methodological Considerations. *Archives of Computational Methods in Engineering*, 1-21.

[6] Salazar, F., Toledo, M. A., Oñate, E., & Morán, R. (2015b). An empirical comparison of machine learning techniques for dam behaviour modelling. *Structural Safety*, *56*, 9-17.

[7] S. D. Myers, B.K., Providing improved dam safety monitoring using existing staff resources: Fern Ridge Dam case study. Proceedings of 28th Annual USSD Conference, 2008.

[8] Salazar, F., Toledo, M. Á., Oñate, E. & Suárez, B. (2016). Interpretation of dam deformation and leakage with boosted regression trees. *Engineering Structures* doi: 10.1016/j.engstruct.2016.04.012

[9] Mata, J., Leitão, N. S., de Castro, A. T., & da Costa, J. S. (2014). Construction of decision rules for early detection of a developing concrete arch dam failure scenario. A discriminant approach. *Computers & Structures*, *142*, 45-53.

[10] Jung, I. S., Berges, M., Garrett, J. H., & Poczos, B. (2015). Exploration and evaluation of AR, MPCA and KL anomaly detection techniques to embankment dam piezometer data. *Advanced Engineering Informatics*, *29*(4), 902-917.

[11] Cheng, L., & Zheng, D. (2013). Two online dam safety monitoring models based on the process of extracting environmental effect. *Advances in Engineering Software*, *57*, 48-56.

[12] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

[13] Friedman, J.H. (2000). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29, 1189-1232.

[14] Leathwick, J. R., Elith, J., Francis, M. P., Hastie, T., & Taylor, P. (2006). Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Marine Ecology Progress Series*, *321*, 267-281.

[15] Team, R. C. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.

[16] Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*, *4*, 40-79.

[17] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, *27*(3), 832-837.

C
Code

C.1 Introduction

In this appendix, the code for the interactive tools is included. They all make use of the Shiny library and are formed by three files:

- global.R includes general instructions
- server.R contains the calculations
- ui.R controls the user interface

All files should be placed in the same directory, together with a *data* folder where the input data should be stored in an appropriate format to be read from global.R

C.2 Dam Monitoring App

C.2.1 User interface

Upload tab

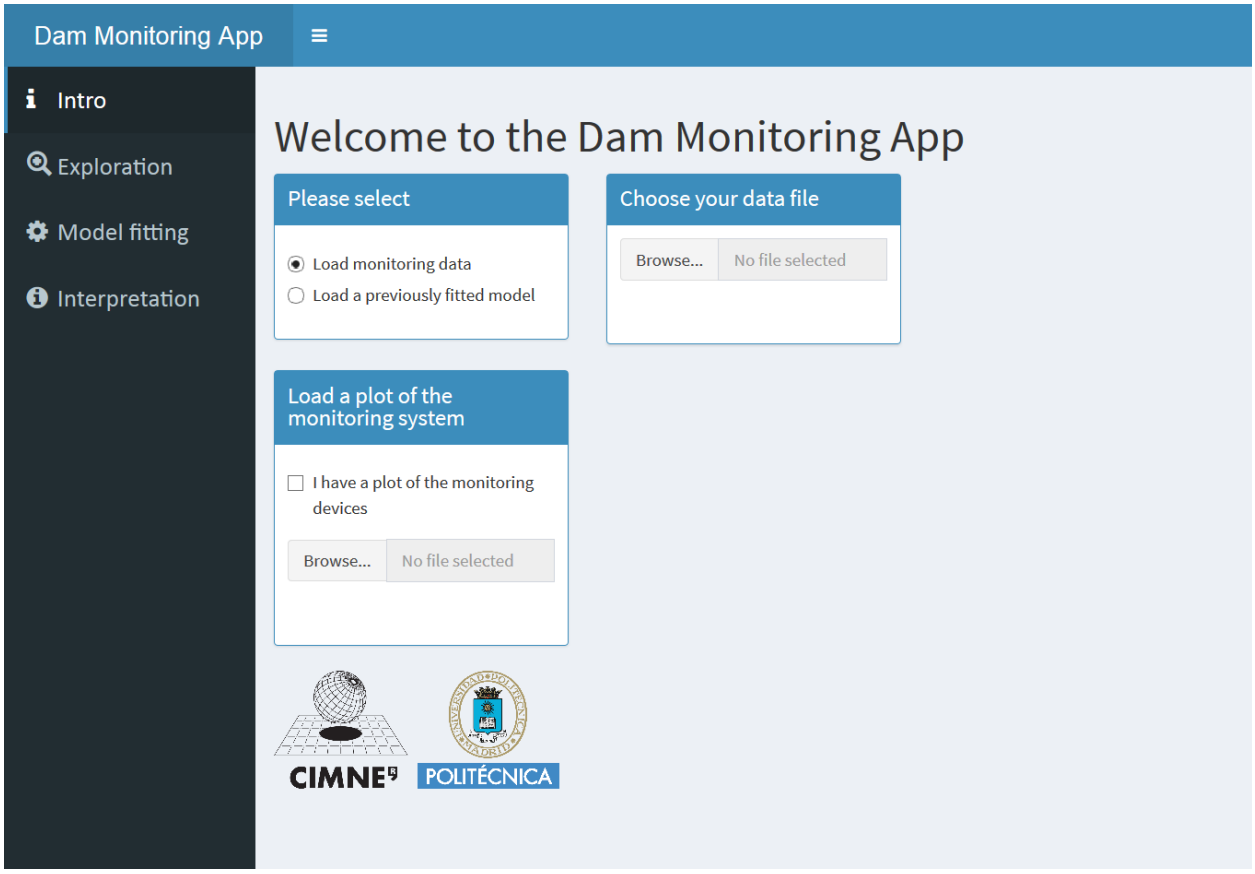


Figure C.1: Dam Monitoring App. Welcome tab. File upload.

Data exploration

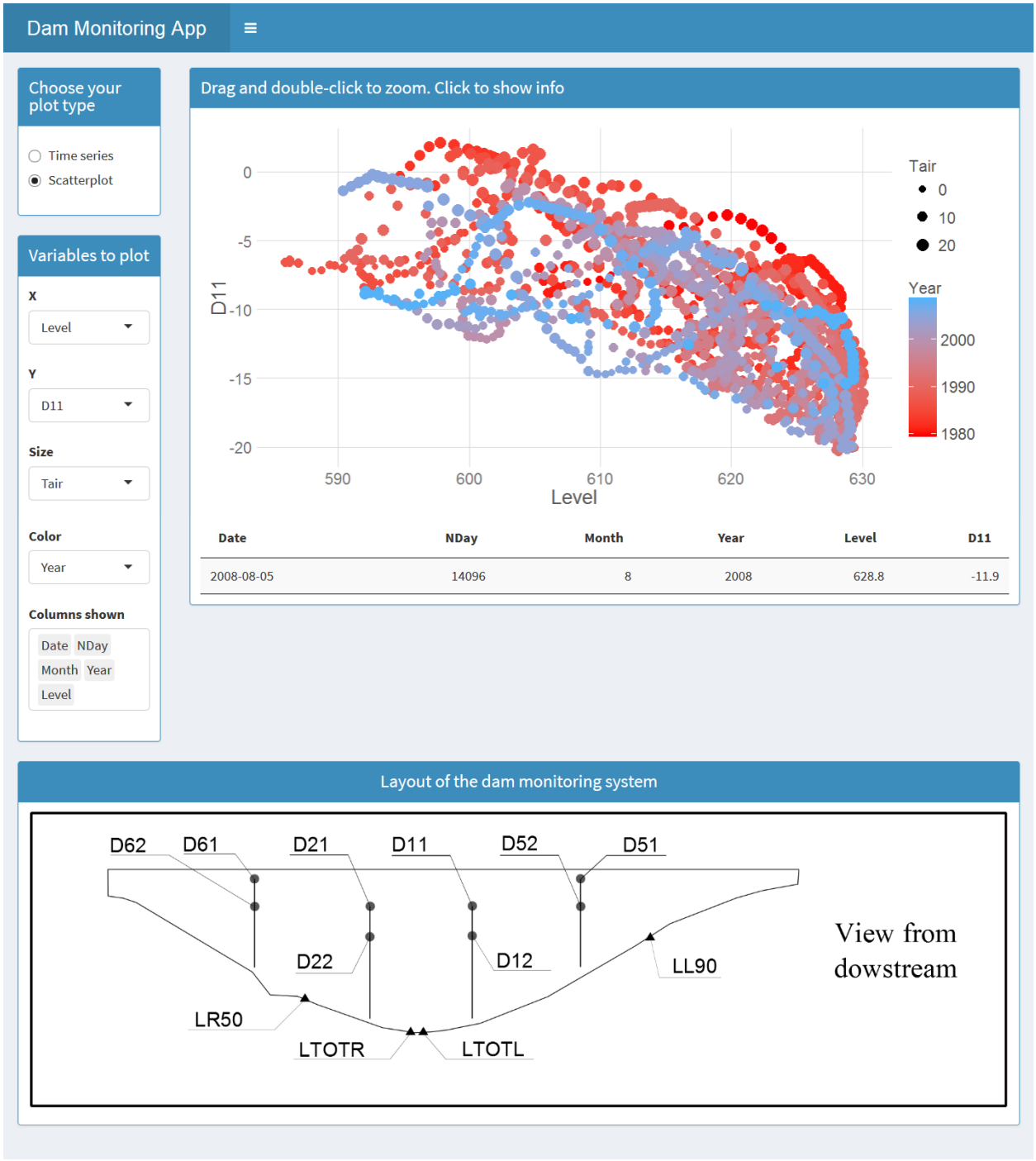


Figure C.2: Tab for data exploration. User interface for scatterplot.

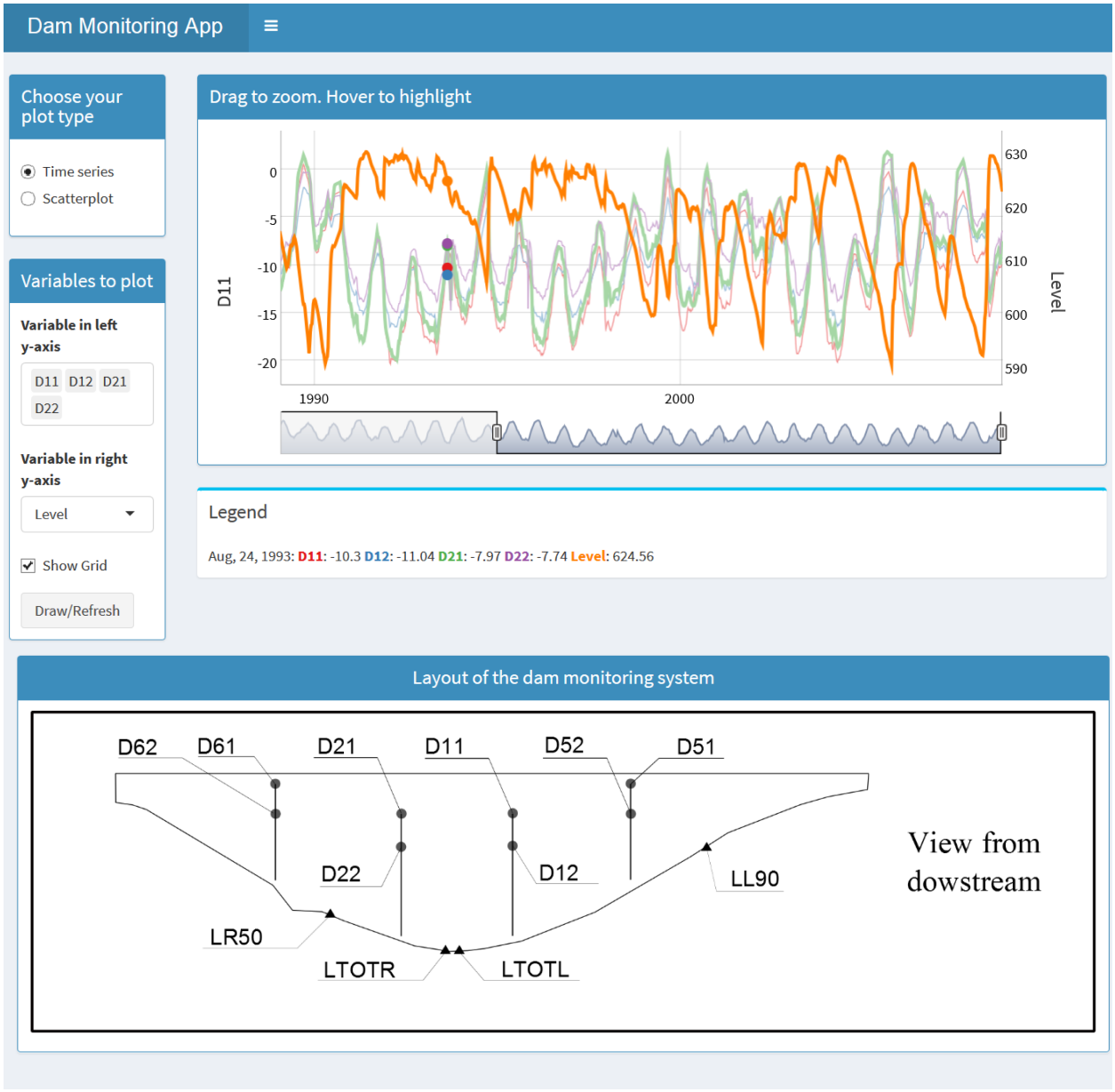


Figure C.3: Tab for data exploration. User interface for time series plot.

Model fitting

Figure C.4: Tab for model fitting. User interface.

Model interpretation

Figure C.5: Tab for model interpretation. User interface.

C.2.2 Code

global.R

```
#### Load libraries --------------------------------

library(lubridate)
library(forecast)
library(gbm)
library(shinydashboard)
library(wordcloud)
library('shiny')
library(scales)
library(grid)
library(RColorBrewer)
library('ggplot2')
library(DT)
library(xts)
library(dygraphs)


#### Function for partial dependence plot

part.plot <- function (model, xvar, points){
  pdp01 <- plot(model,
                i.var = match(xvar,model$var.names),
                n.trees = model$n.trees,
                continuous.resolution = points,
                return.grid = T)
  gplot1 <- ggplot(data=pdp01, aes(x=pdp01[,1], y=y))+
    geom_point(shape=21, fill="#F8766D", size=4)+
    geom_smooth(level=0.8, colour="black", span=1.5, linetype='dashed')
  gplot1 <- gplot1+fte_theme()+
    theme(axis.title.x = element_text(size=20, vjust=3))+
    labs(x = xvar, y= model$response.name)
  return(gplot1)
}

#### Function for bar chart in dygraphs

dyBarChart <- function(dygraph) {
  dyPlotter(dygraph = dygraph,
            name = "BarChart",
            path = system.file("examples/plotters/barchart.js",
                               package = "dygraphs"))
}

#### Pre-loaded theme function for ggplot2 plots (adapted from http://minimaxir.com/)
    ----------------------------

fte_theme <- function() {

  # Generate the colors for the chart procedurally with RColorBrewer

  palette <- brewer.pal("Greys", n=9)
  #color.background = palette[2]
  color.background = 'white'
  color.grid.major = palette[3]
```

```
54    color.axis.text = palette[6]
55    color.axis.title = palette[7]
56    color.title = palette[9]
57
58    # Begin construction of chart
59
60    theme_bw(base_size=9) +
61
62      # Set the entire chart region to a light gray color
63      theme(panel.background=element_rect(fill=color.background, color=color.background)) +
64      theme(plot.background=element_rect(fill=color.background, color=color.background)) +
65      theme(panel.border=element_rect(color=color.background)) +
66
67      # Format the grid
68      theme(panel.grid.major=element_line(color=color.grid.major,size=.25)) +
69      theme(panel.grid.minor=element_blank()) +
70      theme(axis.ticks=element_blank()) +
71
72      # Format the legend, but hide by default
73      #theme(legend.position="none") +
74      theme(legend.background = element_rect(fill=color.background)) +
75      theme(legend.text = element_text(size=16,color=color.axis.title)) +
76      theme(legend.title = element_text(size=16, color=color.axis.title))+
77
78      # Set title and axis labels, and format these and tick marks
79      theme(plot.title=element_text(color=color.title, size=10, vjust=1.25)) +
80      theme(axis.text.x=element_text(size=16,color=color.axis.text)) +
81      theme(axis.text.y=element_text(size=16,color=color.axis.text)) +
82      theme(axis.title.x=element_text(size=20,color=color.axis.title, vjust=0)) +
83      theme(axis.title.y=element_text(size=20,color=color.axis.title, vjust=1.25)) +
84
85
86      # Plot margins
87      theme(plot.margin = unit(c(0.35, 1, 0.3, 0.35), "cm"))
88  }
```

ui.R

```
1   #### User interface for BRT Load. --------------------------------
2
3   dashboardPage(
4     dashboardHeader(title="Dam␣Monitoring␣App"),
5     dashboardSidebar(
6       width = 200,
7         sidebarMenu(
8           tags$head(
9             includeCSS(path = "www/style.css")
10          ),
11          menuItem("␣Intro",
12                   tabName = "introTab",
13                   icon = icon("info")
14                   ),
15          menuItem("␣Exploration",
16                   tabName = "exploration",
17                   icon = icon("zoom-in", lib = "glyphicon","fa-0.5x")
18                   ),
19          menuItem("␣Model␣fitting",
20                   tabName = "fit",
21                   icon = icon("gear")
```

```
22                 ),
23         menuItem("␣Interpretation",
24                  tabName = "interpretation",
25                  icon = icon("info-circle")
26                  )
27       )                          # end of sidebarMenu
28   ),                             # end of sidebar
29   dashboardBody(
30     tabItems(
31       tabItem(tabName="introTab",                              # tabItem 1. Intro
32               h1("Welcome␣to␣the␣Dam␣Monitoring␣App"),
33               fluidRow(
34                 box(
35                   title = "Please␣select",
36                   width = 4, status = "primary", solidHeader = TRUE,
37                   radioButtons("radio", label = NULL,
38                                choices = list("Load␣monitoring␣data" = 1,
39                                "Load␣a␣previously␣fitted␣model" = 2),
40                                selected = character(0))
41                 ),                      # close the box model source
42                 conditionalPanel(
43                   condition = "input.radio␣==␣1",
44                     box(
45                       title = "Choose␣your␣data␣file",
46                       width=4,
47                       status="primary",
48                       solidHeader = TRUE,
49                       fileInput('file2',
50                         NULL,
51                         accept = '.rds')
52                     )               # close the box
53                 ),
54                 conditionalPanel(
55                   condition = "input.radio␣==␣2",
56                     box(
57                       title = "Choose␣the␣file␣with␣the␣saved␣model",
58                       width=4,
59                       status="primary",
60                       solidHeader = TRUE,
61                       fileInput('file1',
62                         NULL,
63                         accept = '.rds')
64                     )               # close the box
65                 )
66               ),          # end of fluidRow
67               fluidRow(
68                 box(
69                   title = "Load␣a␣plot␣of␣the␣monitoring␣system",
70                   width = 4, status = "primary", solidHeader = TRUE,
71                    checkboxInput("frontView", "I␣have␣a␣plot␣of␣the␣monitoring␣devices",
                          value = FALSE),
72                     fileInput('file3',
73                     NULL,
74                     accept = '.png')
75                 )               # close the box
76              ), # end of fluidRow
77              imageOutput("logos")
78               ),                                              # end of tabItem 1. Intro
79       tabItem(tabName = "exploration",                        # tabItem 2. Exploration
```

```
80              tags$style(type="text/css",
81              ".shiny-output-error␣{␣visibility:␣hidden;␣}",
82              ".shiny-output-error:before␣{␣visibility:␣hidden;␣}"),
83              fluidRow(
84                column(width = 2,      # start column 1 - data
85                  box(
86                    title = "Choose␣your␣plot␣type",
87                    width = NULL,
88                    solidHeader = TRUE,
89                    status = "primary",
90                      radioButtons("radioPlot", label = NULL,
91                        choices = list("Time␣series" = 1, "Scatterplot" = 2),
92                                 selected = 1)
93                    ),
94                  conditionalPanel(
95                    condition = "input.radioPlot␣==␣1",
96                    box(
97                      title = "Variables␣to␣plot",
98                      width = NULL,
99                      solidHeader = TRUE,
100                     status = "primary",
101                       uiOutput("tsVars"),
102                       uiOutput("tsVarsy2"),
103                       checkboxInput("showgrid", label = "Show␣Grid", value = TRUE),
104                       actionButton("refresh", "Draw/Refresh")
105                     )
106                 ),
107                 conditionalPanel(
108                   condition = "input.radioPlot␣==␣2",
109                   box(
110                     title = "Variables␣to␣plot",
111                     width = NULL,
112                     solidHeader = TRUE,
113                     status = "primary",
114                        uiOutput("xVar"),
115                        uiOutput("yVar"),
116                        uiOutput("size"),
117                        uiOutput("color"),
118                        uiOutput("tableVars")
119                     )
120                   )
121                 ),
122               column(width = 10,      # start column 2 - plot
123                 conditionalPanel(
124                   condition = "input.radioPlot␣==␣1",
125                   box(
126                     title = "Drag␣to␣zoom.␣Hover␣to␣highlight",
127                     width = NULL,
128                     solidHeader = TRUE,
129                     status = "primary",
130                          dygraphOutput("dy", height="300px")#,
131                     ),
132                   box(
133                     textOutput("legendDivID"),
134                     status = "info",
135                     solidHeader = FALSE,
136                     title = "Legend", width=NULL
137                     )
138                   ),
```

```
139                 conditionalPanel(
140                   condition = "input.radioPlot == 2",
141                   box(
142                     title = "Drag and double-click to zoom. Click to show info",
143                     width = NULL,
144                     solidHeader = TRUE,
145                     status = "primary",
146                        plotOutput("plotExp",
147                            click = "plot_click",
148                            dblclick = "plotExp_dblclick",
149                            brush = brushOpts(
150                            id = "plotExp_brush",
151                            resetOnNew = TRUE
152                            )
153                         ),
154                         DT::dataTableOutput("plot_clicked_points")
155                     )
156                   )
157                 )           # end of column
158               ),          # end of  first fluidRow
159               fluidRow(
160                 conditionalPanel(
161                   condition = "input.frontView == true",
162                   column(width = 12, align="center", #offset=2,
163                        box(
164                        title = "Layout of the dam monitoring system", width = NULL, status= "
                               primary",
165                        height = 370, #collapsible = TRUE, collapsed = TRUE,
166                        solidHeader = TRUE,
167                         imageOutput("FrontView1")
168                        )
169                      )              # end column
170                 )
171                )
172               ),                                          # end tabItem 2. Exploration
173       tabItem(tabName = "fit",                            # tabItem 3. Model fit
174               fluidRow(                                   # open row
175                 column(width = 3,                         # long column with parameters
176                        conditionalPanel(
177                          condition = "input.radio == 2",
178                          box(
179                            title = "Model info", width=NULL, status="primary",
180                            htmlOutput("mytarget", container=tags$h3),
181                            tags$hr(),
182                            htmlOutput("input.title", container=tags$h3),
183                            htmlOutput("myinputs", container=tags$h4),
184                            tags$hr(),
185                            htmlOutput("params.title", container=tags$h3),
186                            htmlOutput("myparams", container=tags$h4),
187                            htmlOutput("tr.title", container=tags$h3),
188                            htmlOutput("loadTrY", container=tags$h4),
189                            htmlOutput("te.title", container=tags$h3),
190                            htmlOutput("loadTeY", container=tags$h4)
191                          )              # close the box
192                        ),
193                        conditionalPanel(
194                          condition = "input.radio == 1",
195                          box(
196                          title = "Choose Model Parameters",
```

```
197                       width = NULL,
198                       status = "primary",
199                       solidHeader = TRUE,
200                       uiOutput("target"),
201                       uiOutput("inputs"),
202                        uiOutput("trainYears"),
203                        uiOutput("testYears"),
204                        numericInput('shrinkage', 'shrinkage', min=0.0001, max=0.9, value
                              =0.01, step=0.0001),
205                        numericInput('Int.depth', 'Int.depth', min=1, max=5, value=2, step
                              =2),
206                        numericInput('Bag.fraction', 'Bag.fraction', min=0.2, max=1, value
                              =0.5, step=0.1),
207                        numericInput('ntree', 'Number␣of␣trees', min=1000, max=10000,
                              value=1000, step=1000),
208                        p(em("Documentation:",a("BRT␣Model␣Info",href="Documentation.html"
                              )))
209                     )            # close the box
210                   ),                    # end of the conditional panel
211                   conditionalPanel(     # conditional panel for buils button
212                     condition = "input.radio␣==␣1",
213                     box(
214                       title = NULL,
215                       width = NULL, status = "primary",
216                       actionButton("build", "Build/Update")
217                     )
218                   ),
219                   conditionalPanel(
220                     condition = "input.radio␣==␣1",
221                     box(                    # box for save
222                       title = "Save␣model",
223                       width = NULL,
224                       status = "primary",
225                       textInput("modelName",
226                         label=NULL,
227                         value = "",
228                         width = NULL,
229                         placeholder = "File␣name"),
230                     actionButton('save', 'Save')#,
231                     )                       # close the box for save
232                   )
233                   ),                                      # end long column
234             column(width = 9,                           # column for results model fitting
235                   fluidRow(                             # top row for info boxes
236                    infoBoxOutput("MAE.Train",
237                      width=3),
238                    valueBoxOutput("R2.Train",
239                      width=3),
240                    valueBoxOutput("MAE.Test",
241                      width=3),
242                    valueBoxOutput("R2.Test",
243                      width=3)
244                    ),                                   # end of top row
245                   fluidRow(                             # bottom row for plots
246                     box(                                # open box for plot model fit
247                       title = "Model␣fitting",
248                       width = NULL,
249                       status = "primary",
250                       solidHeader = TRUE,
```

```
                        dygraphOutput("dyFit", height="300px"),
                        dygraphOutput("dyRes", height="300px")#,
                      )            # end of box for residuals
                     )                                  # end of bottom row
                    )                        # end column
           )                                 # end of row
         ),                                  # end of tabItem 3. Model fit
    tabItem(tabName = "interpretation",      # start tabItem 4. Interpretation
            fluidRow(
              column(width = 2,
                     box(
                       title = "Plot Controls", width = NULL, status = "primary",
                       solidHeader = TRUE,
                       uiOutput("xvarsP1"),
                       numericInput("pointsP1", "Points in Plot 1", value=10,
                                    min = 5, max = 20, step = NA,
                                    width = NULL),
                       uiOutput("xvarsP2"),
                       numericInput("pointsP2", "Points in Plot 2", value=10,
                                    min = 5, max = 20, step = NA,
                                    width = NULL),
                       uiOutput("xvarsP3"),
                       numericInput("pointsP3", "Points in Plot 3", value=10,
                                    min = 5, max = 20, step = NA,
                                    width = NULL)#,
                     )
              ),            # end column
              column(width = 10,
                     fluidRow(                          # top row for RI
                        box(
                          title = "Top 5 influential variables. Relative influence", width
                               = 6, status = "primary",
                          solidHeader = TRUE,
                          plotOutput("varimp")
                        ),
                        box(
                          title = "Relative influence", width = 5, status = "primary",
                          solidHeader = TRUE,
                          plotOutput("wordcloud")
                        )
                      ),                                 # end of top row
                     fluidRow(                          # bottom row for info boxes
                       box(
                         title = "P1. Partial Dependence", width = 4, status = "primary",
                         solidHeader = TRUE, height = 370,
                         plotOutput("pdp1")
                       ),
                       box(
                         title = "P2. Partial Dependence", width = 4, status = "primary",
                         solidHeader = TRUE, height = 370,
                         plotOutput("pdp2")
                       ),
                       box(
                         title = "P3. Partial Dependence", width = 4, status = "primary",
                         solidHeader = TRUE, height = 370,
                         plotOutput("pdp3")
                       )             # end of box for model fit
                     )               # end of bottom row
              )                      # end column
```

```
309                 )                                   # end of row
310         )                                           # end of tabItem 4. Interpretation
311       )                                             # end of tabItems
312     )                                           # end of dashboard body
313 )                                               # end of dashboardpage
```

server.R

```
# Allow larger files to upload.

options(shiny.maxRequestSize = 30*1024^2)

shinyServer(function(input, output, session) {

  ## Operations for tabItem #1. Intro ---------------------------------

  # Plot CIMNE and UPM logos

    output$logos <- renderImage({
      return(list(
           src = "www/logos.png",
           contentType = "image/png",
           alt = "Logos",
           width = 227, height = 94
      ))
    }, deleteFile = FALSE)

  # Upload rds file and assign to global variables.

  modelData <- reactive({
    if(input$radio==2){
    inFile <- input$file1       # file with model
    if (is.null(inFile))
      return(NULL)
    loadModelRes  <- readRDS(inFile$datapath)
    }
    loadModelRes
  })

  # Upload data file file (file2) and create 'dataset' data.frame.

  filedata <- reactive({
     validate(
       need(input$radio == 1, "Please load monitoring data")
     )
    inFile <- input$file2       # file with data
    if (is.null(inFile))
      return(NULL)
    if (exists("dataset"))   {rm(dataset)}
    dataset <<- readRDS(inFile$datapath)
    dataset
  })

  #### Operations for tabItem 2 Exploration ---------------------------------

  ## Input vars for plot  ---------------------------------

  output$xVar <- renderUI({
    df <-filedata()
    if (is.null(df) || input$radio==2 ) {
      return(HTML('Please load some data file'))
      }
    items=names(df)
    names(items)=items
    selectInput('x', 'X', items, items[5])
```

```
58    })
59    output$yVar <- renderUI({
60      df <-filedata()
61      if (is.null(df)) return(NULL)
62      items=names(df)
63      names(items)=items
64      selectInput('y', 'Y', items, items[11])
65    })
66    output$size <- renderUI({
67      df <-filedata()
68      if (is.null(df)) return(NULL)
69      items=names(df)
70      names(items)=items
71      selectInput('size', 'Size', names(items), names(items)[[6]])
72    })
73    output$color <- renderUI({
74      df <-filedata()
75      if (is.null(df)) return(NULL)
76      items=names(df)
77      names(items)=items
78      selectInput('color', 'Color', names(items), names(items)[[4]])
79    })
80
81    # Vars to show in bottom table
82
83    output$tableVars <- renderUI({
84      df <-filedata()
85      if (is.null(df)) return(NULL)
86      items=names(df)
87      names(items)=items
88      selectizeInput('tableVars', 'Columns␣shown',
89                     choices = items, multiple = TRUE,
90                     selected = items[1:5])
91    })
92
93    # Time series plot
94
95      output$tsVars <- renderUI({
96          df <-filedata()
97      if (is.null(df)) return(NULL)
98      items=names(df)
99      names(items)=items
100     selectizeInput('tsVars', '␣Variable␣in␣left␣y-axis␣',
101                    choices = items, multiple = TRUE,
102                    selected = names(items)[[3]])
103   })
104
105   output$tsVarsy2 <- renderUI({
106     df <-filedata()
107     if (is.null(df)) return(NULL)
108     items=c("None", names(df))
109     names(items)=items
110     selectizeInput('tsVarsy2', '␣Variable␣in␣right␣y-axis␣',
111                    choices = items, multiple = FALSE,
112                    selected = items[1]
113                    )
114   })
115
116   dyCols <- eventReactive(input$refresh, {
```

```
117     inFile <- input$file2       # file with data
118     mydat <- readRDS(inFile$datapath)
119     selVars <- c(input$tsVars, input$tsVarsy2)
120     dycols <- which(names(mydat)%in%selVars)
121     dycols
122   })
123
124    dy2 <- eventReactive(input$refresh, {
125      dy2 <- input$tsVarsy2
126      dy2
127   })
128
129   output$dy <- renderDygraph({
130     if (input$refresh == 0)
131       return()
132     input$refresh
133     inFile <- input$file2       # file with data
134     mydat <- readRDS(inFile$datapath)
135     selVars <- c(input$tsVars, dy2())
136     dyDframe <-  xts(mydat[,dyCols()], order.by = mydat[,1])
137     isolate({
138         if (identical(input$tsVarsy2,"None")){
139         dygraph(dyDframe, main = "")%>%
140           dyLegend(labelsDiv = "legendDivID", width = 1200, show = "onmouseover")%>%
141           dyOptions(drawGrid = input$showgrid, rightGap=20, colors = RColorBrewer::brewer.
                pal(9, "Set1"))%>%
142           dyAxis("y", label = input$tsVars[1], labelWidth=20)%>%
143           dyHighlight(highlightCircleSize = 5,
144                   highlightSeriesBackgroundAlpha = 0.5,
145                   hideOnMouseOut = FALSE,
146                   highlightSeriesOpts = list(strokeWidth = 3))%>%
147           dyRangeSelector()
148         } else {
149          dygraph(dyDframe, main = "")%>%
150          dySeries(dy2(), axis = 'y2')%>%
151          dyLegend(labelsDiv = "legendDivID", width = 400, show = "onmouseover")%>%
152          dyOptions(drawGrid = input$showgrid, rightGap=20,, colors = RColorBrewer::brewer.
               pal(9, "Set1"))%>%
153          dyAxis("y", label = input$tsVars[1], labelWidth=20)%>%
154          dyAxis("y2", label = dy2(), labelWidth=20,independentTicks = TRUE, drawGrid = FALSE
               )%>%
155          dyHighlight(highlightCircleSize = 5,
156                   highlightSeriesBackgroundAlpha = 0.5,
157                   hideOnMouseOut = TRUE,
158                   highlightSeriesOpts = list(strokeWidth = 3))%>%
159          dyRangeSelector()
160         }
161     })
162   })
163
164   # Plot Front view
165
166   output$FrontView1 <- renderImage({
167     inFile <- input$file3       # file with image
168     if (is.null(inFile))
169       return(NULL)
170     if (input$frontView == FALSE)
171       return(NULL)
172     return(list(
```

```
173             src = inFile$datapath,
174             contentType = "image/png",
175             alt = "FrontView",
176             width = 993, height = 300
177             ))
178   }, deleteFile = FALSE)
179
180   ## Exploration Plot  ------------------------------------------------
181
182   # Zoom brush
183
184   rangesExp <- reactiveValues(x = NULL, y = NULL)
185   observeEvent(input$plotExp_dblclick, {
186     brushExp <- input$plotExp_brush
187     if (!is.null(brushExp)) {
188       rangesExp$x <- c(brushExp$xmin, brushExp$xmax)
189       rangesExp$y <- c(brushExp$ymin, brushExp$ymax)
190     } else {
191       rangesExp$x <- NULL
192       rangesExp$y <- NULL
193     }
194   })
195
196   # Click Info on exploration plot  ----------------------------------------
197
198   output$click_info <- renderText({
199     if (is.null(input$plot_click$x))
200       return(NULL)
201     HTML(paste0(input$x, "␣=␣", round(input$plot_click$x, digits=1),
202             "<br/>", input$y, "␣=␣",
203             round(input$plot_click$y, digits=1),sep=''))
204   })
205
206   # Show data for clicked point.
207
208    output$plot_clicked_points <- DT::renderDataTable({
209      mydat <- filedata()
210      mydat <- cbind(Date=mydat[,1], round(mydat[,-1],digits=1))
211      selVars <- input$tableVars
212      myTableVars <- unique(c(input$tableVars, input$x, input$y))
213      mycols <- which(names(mydat)%in%myTableVars)
214      dat <- mydat[,mycols]
215      res <- nearPoints(dat, input$plot_click, xvar=input$x, yvar=input$y,
216           threshold = 10, maxpoints = 1,
217           addDist = FALSE)
218      datatable(res, options = list(
219                       paging = FALSE,
220                       searching = FALSE,
221                       info=FALSE,
222                       ordering=FALSE
223                       ),
224                       rownames= FALSE
225                 )
226   })
227
228   #  Exploration Plot  ----------------------------------------
229
230    output$plotExp <- renderPlot({
231     if (is.null(input$file2)) return(NULL)
```

```
232       validate(
233         need(input$radio == 1, "Please␣load␣monitoring␣data")
234       )
235      if(input$x != 'Date' && !is.null(input$x)){
236       ggplot(filedata(), aes_string(x=input$x, y=input$y,
237                   color = input$color)) +
238          geom_point() +
239        fte_theme()+
240        coord_cartesian(xlim = rangesExp$x, ylim = rangesExp$y)+
241         aes_string(color=input$color)+
242         aes_string(size=input$size)+
243        scale_colour_gradient(low="red")+
244        theme(legend.position="right")+
245        theme(legend.key.size = unit(1, "cm"))
246       } else {
247        ggplot(filedata(), aes_string(x=input$x, y=input$y)) + geom_point() +
248        fte_theme()+
249        coord_cartesian(xlim = rangesExp$x, ylim = rangesExp$y)+
250        aes_string(color=input$color)+
251        aes_string(size=input$size)+
252        scale_colour_gradient(low="red")+
253        scale_x_date()+
254        theme(legend.position="right")+
255        theme(legend.key.size = unit(1, "cm"))
256       }
257   })
258
259   ## Operations for tabItem 3. Model fit ---------------------------------
260
261   # Options for ui
262
263   output$target <- renderUI({
264     if (is.null(filedata())) return(NULL)
265     items=names(filedata())
266     names(items)=items
267     selectInput('target', 'Target', items, items[8])
268   })
269   output$inputs <- renderUI({
270     if (is.null(filedata())) return(NULL)
271     items=names(filedata())
272     selectizeInput('Inputs', 'Inputs',
273                    choices = items[-1], multiple = TRUE)
274   })
275   output$trainYears <- renderUI({
276     if (is.null(filedata())) return(NULL)
277     dateRangeInput("trainYears", label = h3("Training␣period"),
278                    start= min(filedata()$Date), end =max(filedata()$Date)-1825,
279                    min= min(filedata()$Date), max =max(filedata()$Date))
280   })
281   output$testYears <- renderUI({
282     if (is.null(filedata())) return(NULL)
283     dateRangeInput("testYears", label = h3("Test␣period"),
284                    start= max(filedata()$Date)-1824, end =max(filedata()$Date),
285                    min= min(filedata()$Date), max =max(filedata()$Date))
286   })
287
288   ##  Model load/build tabItem ------------------------------------------------
289
290   #  Main output: list with model, dataframe with (date, preds, obs, residual),
```

```
291   #              words, freqs, maes. ------------------------------------------
292
293   modelResFit <- eventReactive(input$build, {
294     traindata <- dataset[dataset$Date>=input$trainYears[1] & dataset$Date<=input$trainYears
          [2],]
295     testdata <- dataset[dataset$Date>input$testYears[1] & dataset$Date<=input$testYears[2]
          ,]
296     for (i in 1:length(input$Inputs)){
297           if (i == 1){
298             nam = input$Inputs[i]
299           } else {
300             nam <- paste(nam, '+', input$Inputs[i])
301           }
302         }
303     myform <- as.formula(paste(input$target, "~", nam))
304     withProgress(message = 'Training␣model', value = 0,{
305       brtModel <<- gbm(myform,
306                        data              = traindata,
307                        distribution      = "gaussian",
308                        n.trees           = input$ntree,
309                        shrinkage         = input$shrinkage,
310                        interaction.depth = input$Int.depth,
311                        bag.fraction      = input$Bag.fraction,
312                        train.fraction    = 1,
313                        n.minobsinnode    = 5,
314                        cv.folds          = 0,
315                        keep.data         = TRUE,
316                        verbose=F)
317     })
318   # compute predictions, extract observations, add to dataframe
319     predTest <- predict(brtModel, newdata=testdata, n.trees=brtModel$n.trees)
320     predTrain <- predict(brtModel, newdata=traindata, n.trees=brtModel$n.trees)
321     Prediction=c(predTrain, predTest)
322     Observation <- c(traindata[,brtModel$response.name], testdata[,brtModel$response.name])
323     Error=c(Prediction-Observation)
324     Date <- c(traindata[,'Date'], testdata[,'Date'])
325     dataOut <<- data.frame(Date, Observation, Prediction, Error)   # dataframe with date,
          preds, obs, residual
326     mae.train <<- round((accuracy(Prediction[1:nrow(traindata)],
327                                   Observation[1:nrow(traindata)]) [3]),
328                                   digits=2)
329     mae.test <<- round((accuracy(Prediction[(nrow(traindata)+1):nrow(dataOut)],
330                                  Observation[(nrow(traindata)+1):nrow(dataOut)])
                                      [3]),
331                                  digits=2)
332     R2.train <- round(1 - sum((traindata[,brtModel$response.name]-predTrain)^2)/sum((
          traindata[,brtModel$response.name]-mean(traindata[,brtModel$response.name]))^2),
333                   digits = 2)
334     R2.test <- round(1 - sum((testdata[,brtModel$response.name]-predTest)^2)/sum((testdata[,
          brtModel$response.name]-mean(testdata[,brtModel$response.name]))^2),
335                   digits = 2)
336     varImpB<-summary(brtModel, normalize=T)
337     words <<- row.names(varImpB)
338     freqs <<- varImpB[,2]/min(varImpB[varImpB[,2]!=0,2])+1
339     for (i in 1: nrow(varImpB)){ freqs[i] = max(varImpB[i,2], rep(1, nrow(varImpB))[i]  )}
340     modelRes <- list(model      = brtModel,
341                      dataout    = dataOut,
342                      words      = words,
343                      freqs      = freqs,
```

```
344                     mae.train   = mae.train,
345                     R2.train    = R2.train,
346                     mae.test    = mae.test,
347                     R2.test    = R2.test,
348                     trainY      = input$trainYears,
349                     testY       = input$testYears
350                           )
351     maeTest <<- mae.test
352     trainYears <<- input$trainYears
353     testYears  <<- input$testYears
354     modelResFit <- modelRes
355     modelResFit
356   })
357
358   #   If loaded model: Target variable ------------------------------------
359
360   output$mytarget <- renderPrint({
361     if(input$radio==2){
362       inFile <- input$file1
363       if (is.null(inFile))
364         return(HTML(""))
365       HTML(paste("Target: ", modelData()$model$response.name, sep=''))
366     } else {return(HTML(""))}
367   })
368
369   #   If loaded model: trainYears -----------------------------------------------
370
371   output$tr.title <-   renderPrint({ HTML("Train period ")})
372   output$loadTrY <- renderPrint({
373     if(input$radio==2){
374       inFile <- input$file1
375       if (is.null(inFile))
376         return(HTML(""))
377     HTML(paste(as.Date(modelData()$trainY[1]),"to",as.Date(modelData()$trainY[2]),  sep=' ')
         )
378     } else {return(HTML(""))}
379   })
380   output$te.title <-   renderPrint({ HTML("Test period ")})
381   output$loadTeY <- renderPrint({
382     if(input$radio==2){
383       inFile <- input$file1
384       if (is.null(inFile))
385         return(HTML(""))
386     HTML(paste(as.Date(modelData()$testY[1]),"to",as.Date(modelData()$testY[2]),  sep=' '))
387     } else {return(HTML(""))}
388   })
389
390   #   If loaded model: Input variables ------------------------------------------
391
392   output$params.title <-   renderPrint({ HTML("Model Parameters: ")})
393   output$myparams <- renderPrint({
394     if(input$radio==2){
395       inFile <- input$file1
396       if (is.null(inFile))
397         return(HTML(""))
398       params <- HTML(paste("n.trees:", modelData()$model$n.trees,'<br/>',
399                            "Shrinkage:", modelData()$model$shrinkage,'<br/>',
400                            "Bag fraction:", modelData()$model$bag.fraction,'<br/>',
401                            "Int. depth:", modelData()$model$interaction.depth,'<br/>'))
```

```
402       params
403     } else {return(HTML(""))}
404   })
405
406   #   If loaded model: Model parameters  ---------------------------------------------------
407
408   output$input.title <-   renderPrint({ HTML("Inputs:␣")})
409   output$myinputs <- renderPrint({
410     if(input$radio==2){
411       inFile <- input$file1
412       if (is.null(inFile))
413         return(HTML(""))
414       inputs <- modelData()$model$var.names[1]
415       for (i in 2: length(modelData()$model$var.names)){
416         inputs <- HTML(paste(inputs, modelData()$model$var.names[i], sep = '␣-␣'))
417       }
418       inputs
419     } else {return("")}
420   })
421
422   # Save the model  ----------------
423
424   observeEvent(input$save, {           # Only after fitting (input$radio == 1)
425     resList <- modelResFit()
426     saveRDS (resList, file = paste('data/models/',input$modelName,'.rds', sep=''))
427   })
428
429   # If build model: MAE Test -----------------
430
431   output$MAE.Test <- renderValueBox({
432     pdf(NULL)                          # to fix error when running on AWS
433     if(input$radio==1){
434       mae.test <- modelResFit()$mae.test
435     } else {
436       mae.test <- modelData()$mae.test
437     }
438     valueBox(
439           mae.test, "MAE␣Test", icon = icon("line-chart"),
440           color = "aqua"
441     )
442   })
443
444   output$R2.Test <- renderValueBox({
445     pdf(NULL)                        # to fix error when running on AWS
446     if(input$radio==1){
447       R2.test <- modelResFit()$R2.test
448     } else {
449       R2.test <- modelData()$R2.test
450     }
451     valueBox(
452           R2.test, "R2␣Test", icon = icon("line-chart"),
453           color = "aqua"
454     )
455   })
456
457   # If build model:  MAE Train -----------------
458
459   output$MAE.Train <- renderValueBox({
460     pdf(NULL)                        # to fix error when running on AWS
```

```
461     if(input$radio==1){
462         mae.train <- modelResFit()$mae.train
463     } else {
464         mae.train <- modelData()$mae.train
465     }
466     valueBox(
467          mae.train, "MAE␣Train", icon = icon("line-chart"),
468          color = "purple"
469     )
470   })
471
472    output$R2.Train <- renderValueBox({
473     pdf(NULL)                          # to fix error when running on AWS
474     if(input$radio==1){
475       R2.train <- modelResFit()$R2.train
476     } else {
477       R2.train <- modelData()$R2.train
478     }
479     valueBox(
480           R2.train, "R2␣Train", icon = icon("line-chart"),
481           color = "purple"
482     )
483   })
484
485   ## Fit plot
486
487   #  Dygraphs
488
489    output$dyFit <- renderDygraph({
490     pdf(NULL)                          # to fix error when running on AWS
491     if (input$radio == 1) {
492       dyData <- modelResFit()$dataout
493       endTrain <- input$trainYears[2]
494     } else {
495       dyData <- modelData()$dataout
496       endTrain <- modelData()$trainY[2]
497     }
498     dyDframe <-  xts(dyData[,2:3], order.by = dyData[,1])
499     dygraph(dyDframe, main = "", group = "fit")%>%
500       dySeries("Observation", label = "Observation", drawPoints = TRUE,strokeWidth= 0.0,
              pointSize= 3) %>%
501       dySeries("Prediction", label = "Prediction", strokePattern = "dashed", strokeWidth=
              2.0) %>%
502       dyLegend(show = "always", width = 400)%>%
503       dyOptions(, colors = RColorBrewer::brewer.pal(9, "Set1"))%>%
504       dyEvent(endTrain, "End␣of␣training", labelLoc = "bottom")%>%
505       dyHighlight(highlightCircleSize = 5,
506                   highlightSeriesBackgroundAlpha = 0.9,
507                   hideOnMouseOut = FALSE)%>%#,
508       dyRangeSelector()
509   })
510    output$dyRes <- renderDygraph({
511    pdf(NULL)                          # to fix error when running on AWS
512     if (input$radio == 1) {
513       dyData <- modelResFit()$dataout
514       endTrain <- input$trainYears[2]
515     } else {
516       dyData <- modelData()$dataout
517       endTrain <- modelData()$trainY[2]
```

```
518     }
519     names(dyData)[4] <- "Error"
520     dyDframe <-  xts(dyData[,4], order.by = dyData[,1])
521     dyDframe <- data.frame(Error=dyDframe)
522     dygraph(dyDframe, main = "", group = "fit")%>%
523       dyLegend(show = "onmouseover")%>%
524       dyAxis("y", label = "Residual", labelWidth=20)%>%
525       dyEvent(endTrain, "End␣of␣training", labelLoc = "bottom")%>%
526       dyHighlight(highlightCircleSize = 5,
527               hideOnMouseOut = TRUE)%>%
528       dyBarChart()
529   })
530
531   #  Brush reactive zoom  ----------------------------
532
533   ranges <- reactiveValues(x = NULL, y = NULL)
534   observeEvent(input$plot1_dblclick, {
535     brush <- input$plot1_brush
536     if (!is.null(brush)) {
537       ranges$x <- as.Date(c(brush$xmin, brush$xmax), origin='1970-01-10' )
538       ranges$y <- c(brush$ymin, brush$ymax)
539     } else {
540       ranges$x <- NULL
541       ranges$y <- NULL
542     }
543   })
544
545   #  Plot1. Model fitting  ----------------------------
546
547   output$plot1 <- renderPlot({
548     if (input$radio == 1) {
549       ggData <- modelResFit()$dataout
550       ggyLab <- modelResFit()$model$response.name
551     } else {
552       ggData <- modelData()$dataout
553       ggyLab <- modelData()$model$response.name
554     }
555     p <- ggplot(ggData) +
556         geom_point(aes(Date, Observation, colour="Observation"),
557                         shape=21, fill="#619CFF", size=4) +   # circles - observations
558         geom_line(aes(Date, Prediction, colour="Prediction"),
559                   size=1.2, linetype='dashed')+
560         scale_colour_manual(values = c("#619CFF", "orangered2"),
561                             guide = guide_legend(override.aes = list(
562                               linetype = c("blank", "dashed"),
563                             shape = c(21, NA))))+
564         coord_cartesian(xlim = ranges$x, ylim = ranges$y)
565     isolate({
566         if(input$radio==1){
567           xVertLine <- as.numeric(input$trainYears[2])
568         } else if (input$radio == 2) {
569           xVertLine <- as.numeric(modelData()$trainY[2])
570         }
571     })
572     p +
573       fte_theme()+
574       theme(legend.position=c(0.9, 0.9))+
575       theme(legend.title=element_blank())+
576       theme(legend.key = element_rect(colour = NA, fill = NA))+
```

```
577         theme(legend.key.size = unit(1.5, "cm"))+
578         geom_vline(xintercept = xVertLine,
579                                 color='red', linetype=2)+
580         labs(x = "Date", y= ggyLab)
581     })
582
583     #   Brush for bottom plot  ----------------------------
584
585     rangesx <- reactiveValues(x = NULL)
586     observe({
587       brush <- input$plot2_brush
588       if (!is.null(brush)) {
589         ranges$x <- as.Date(c(brush$xmin, brush$xmax), origin='1970-01-10' )
590       } else {
591         rangesx$x <- NULL
592       }
593     })
594
595     #  Residuals plot  ----------------------------
596
597     output$plot2 <- renderPlot({
598       if (input$radio == 1) {
599         ggResData <- modelResFit()$dataout
600       } else {
601         ggResData <- modelData()$dataout
602       }
603       p <- ggplot(ggResData, aes(x=Date, y=Error)) +
604          geom_point(shape=21, fill="#F8766D", size=4)+coord_cartesian(xlim = ranges$x)
605       isolate({
606           if(input$radio==1){
607             xVertLine <- as.numeric(input$trainYears[2])
608           } else {
609             xVertLine <- as.numeric(modelData()$trainY[2])
610           }
611       p + fte_theme()+geom_vline(xintercept = xVertLine,
612                                  color='red', linetype=2)+
613          labs(x = "Date", y= "Residual")
614       })
615     }, height=200)
616
617     ## Operations for tabItem 4 Interpretation ------------------------------
618
619     #  Vars for partial dep. plot #1 -----------------
620
621     output$xvarsP1 <- renderUI({
622       if (input$radio == 1) {
623         modVars <- modelResFit()$model$var.names
624       } else if (input$radio ==2){
625         modVars <- modelData()$model$var.names
626       }
627       selectizeInput(
628           'XvarsP1', 'Var␣in␣Plot␣1', choices = modVars, multiple = F,
629           selected = NULL
630       )
631     })
632
633     #  Vars for partial dep. plot #2 -----------------
634
635     output$xvarsP2 <- renderUI({
```

```
636     if (input$radio == 1) {
637       modVars <- modelResFit()$model$var.names
638     } else if (input$radio ==2){
639       modVars <- modelData()$model$var.names
640     }
641     selectizeInput(
642         'XvarsP2', 'Var in Plot 2', choices = modVars, multiple = F,
643         selected = NULL
644     )
645   })
646
647   #  Vars for partial dep. plot #3 -----------------
648
649   output$xvarsP3 <- renderUI({
650     if (input$radio == 1) {
651       modVars <- modelResFit()$model$var.names
652     } else if (input$radio ==2){
653       modVars <- modelData()$model$var.names
654     }
655     selectizeInput(
656         'XvarsP3', 'Var in Plot 3', choices = modVars, multiple = F,
657         selected = NULL
658     )
659   })
660
661   #  Variable importance plot. Top 5 vars. #1 -----------------
662
663    output$varimp <- renderPlot({
664     if (input$radio == 1) {
665       brtModel <- modelResFit()$model
666      } else if (input$radio ==2){
667       brtModel <- modelData()$model
668     }
669     varimp.data <- summary(brtModel, plotit=F, cBars=5, order=T)
670     varimp.data <- data.frame(var=rownames(varimp.data), rel.inf=varimp.data[,2])
671     vars2plot <- min(5, nrow(varimp.data))             # no more than 5 vars
672     vi.p <- ggplot(varimp.data[1:vars2plot,], aes(x = var, y = rel.inf)) +
673             geom_bar(stat = "identity")+ coord_flip()
674     vi.p+fte_theme()+
675          theme(axis.title.x = element_text(size=20, vjust=3))+
676          labs(x = "Variable", y= "Rel. Influence")
677   })
678
679   #   Partial Dependence Plot #1 -----------------
680
681   output$pdp1 <- renderPlot({
682      if (input$radio == 1) {
683       brtModel <- modelResFit()$model
684     } else if (input$radio ==2){
685       brtModel <- modelData()$model
686     }
687     pdp01 <- part.plot(brtModel, input$XvarsP1, input$pointsP1)
688     pdp01
689   }, height=300)
690
691   #   Partial Dependence Plot #2 -----------------
692
693   output$pdp2 <- renderPlot({
694      if (input$radio == 1) {
```

```
695       brtModel <- modelResFit()$model
696     } else if (input$radio ==2){
697       brtModel <- modelData()$model
698     }
699     pdp02 <- part.plot(brtModel, input$XvarsP2, input$pointsP2)
700     pdp02
701   }, height=300)
702
703   #   Partial Dependence Plot #3 -----------------
704
705   output$pdp3 <- renderPlot({
706      if (input$radio == 1) {
707       brtModel <- modelResFit()$model
708     } else if (input$radio ==2){
709       brtModel <- modelData()$model
710     }
711     pdp03 <- part.plot(brtModel, input$XvarsP3, input$pointsP3)
712     pdp03
713   }, height=300)
714
715 #  Wordcloud model -----------------
716
717   wordcloud_rep <- repeatable(wordcloud)
718   output$wordcloud <- renderPlot({
719     if (input$radio == 1) {
720       words <- modelResFit()$words
721       freqs <- modelResFit()$freqs
722     } else if (input$radio ==2){
723       words <- modelData()$words
724       freqs <- modelData()$freqs
725     }
726   wordcloud_rep(words,freqs,
727                 scale=c(6,0.8),
728                 min.freq=1,
729                 max.words=10,
730                 random.order=F,
731                 random.color=FALSE,
732                 rot.per=0,
733                 colors=brewer.pal(8, "Dark2"),
734                 ordered.colors=FALSE,
735                 use.r.layout=FALSE,
736                 fixed.asp=TRUE)
737   })
738 })                          # end of shinyServer
```

C.3 Anomaly Detection App

This application requires an image of the dam, also stored in the "data" folder.

C.3.1 User interface

Figure C.6: Anomaly detection application. User interface

C.3.2 global.R

```
library(shinydashboard)
library('shiny')
library('ggplot2')
library('grid')
library('png')
library('scales')
library(dygraphs)

img <- readPNG("data/frontView.png")
g <- rasterGrob(img, interpolate=TRUE)
mydf <- read.table("data/masteroutNoncVal2W.res")
mydf[,'date'] <- as.Date(mydf[,'date'])
myColors <- c("#00FF00",  "#FFFF00", "#FF0000")
names(myColors) <- c("green", "yellow", "red")
myShapes <- c(22,24,25)
names(myShapes) <- c("none", "upstream", "downstream")

############  Initialise data frame for the colors and shapes
mydate <- as.Date("2000-05-14")
res<- which(mydf$date>mydate)
res <- mydf[(res[1]-1):res[1],]      # current and last rows
res1 <- res[,9:16]                                                  # t and t-1
res2 <- t(res1)
state <- factor(x=rep('green', 8), levels =c('green', 'yellow', 'red'))
deviation <- factor(x=rep('none', 8), levels =c('none', 'upstream', 'downstream'))
pends <- data.frame(x=c(0.935, 0.935, 0.715, 0.715, 1.168, 1.168, 0.47, 0.47),
                    y= c(0.385, 0.324, 0.385, 0.324, 0.44, 0.385, 0.44, 0.385),
                    state = state, deviation = deviation)
res3 = data.frame(prev = res2[,1], curr = res2[,2], pends)
res3$state <- rep('green', 8)
for(i in 1:nrow(res3)){
  if (abs(res3[i,2])>3) {         # num.sd
    res3$state[i] <- 'red'
  }else if (abs(res3[i,2])>2){    #num.sd
    res3$state[i] <- 'yellow'
  } else
    res3$state[i] <- 'green'
}
for (j in 1:nrow(res3)){
  if (res3$state[j] != 'green') {                # yellow or red
    if (res3$curr[j] > 0){
      res3$deviation [j] <- 'upstream'
    } else res3$deviation[j] <- 'downstream'
  }
}
res4 <- res3
#################  End of initialising -> res3 and res4

# Function for plots

plot_sh <- function(d.frame, tg, max.date){
  tg.pr <- paste(tg, "_pr", sep="")
  date.in <- max.date-180    # start date
  date.end <- max.date +14 # final date
  d.frame <- d.frame[(d.frame$date > date.in),]        # data frame to plot
  d.frame <- d.frame [(d.frame$date < date.end),]
  d.frame[nrow(d.frame),] <- NA
```

```
58    tg.col <- which(names(d.frame) == tg)            # output to plot
59    tg.pr <- which(names(d.frame) == paste(tg, "_pr", sep=""))
60    tg.anom <- which(names(d.frame) == paste(tg, "_anom", sep=""))
61    tg.up <- which(names(d.frame) == paste(tg, "_up", sep=""))
62    tg.lw <- which(names(d.frame) == paste(tg, "_lw", sep=""))
63    names(d.frame)[tg.col] <- 'target'
64    names(d.frame)[tg.pr] <- 'pred'
65    names(d.frame)[tg.anom] <- 'anom'
66    names(d.frame)[tg.up] <- 'upr'
67    names(d.frame)[tg.lw] <- 'lwr'
68    myd.frame <- d.frame[,c(tg.col, tg.pr, tg.up, tg.lw)]
69    anom.vals <- d.frame$anom
70    for (i in 1:(length(anom.vals)-1)){
71      if (abs(anom.vals[i])>2) {
72        anom.vals[i] <- d.frame$target[i]
73      } else {
74        anom.vals[i] <- NA
75      }
76    }
77    myd.frame <- data.frame(myd.frame, anom = anom.vals)
78    yRange <- 10
79    yMax <- max(myd.frame$upr)+yRange
80    yMin <- min(myd.frame$lwr)-yRange
81    dygraph(myd.frame, main = "") %>%
82      dySeries("anom", label = NULL, drawPoints = TRUE, strokeWidth= 0.0,pointSize= 7) %>%
83      dySeries("target", label = "Observed", drawPoints = TRUE,strokeWidth= 0.0,pointSize= 3)
          %>%
84      dyAxis("y", label = tg, labelWidth=20)%>%
85      dySeries(c("lwr", "pred", "upr"), label = "Predicted", strokePattern = "dashed") %>%
86      dyOptions(axisLabelFontSize=16, rightGap=20,colors = RColorBrewer::brewer.pal(9, "Set1")
          )%>%
87      dyLegend(labelsDiv = "legendDivID")
88  }
89
90  # Function for plot with symbols on dam layout
91  semplot <- function(res3, g, myCols) {
92    base <- ggplot(res3, aes(x=x, y=y)) + xlim(0,1.733)+ylim(0,0.592)+
93      geom_blank()+
94      theme(axis.line=element_blank(),axis.text.x=element_blank(),
95            axis.text.y=element_blank(),axis.ticks=element_blank(),
96            axis.title.x=element_blank(),
97            axis.title.y=element_blank(),
98            panel.background=element_blank(),panel.border=element_blank(),panel.grid.major=
                element_blank(),
99            panel.grid.minor=element_blank(),plot.background=element_blank())
100
101   # Full panel annotation
102   pl <- base + annotation_custom(g,xmin = -0.05, xmax = Inf, ymin = -0.05, ymax = Inf)+
103     geom_point(aes(x=x, y=y, fill = state, shape=factor(deviation), show_legend = TRUE),
104                size=6 )+
105     scale_fill_manual(values = myCols, guide=FALSE)+
106     scale_shape_manual(name="Direction",values=myShapes)+
107     theme(legend.key = element_rect())
108   pl <- pl+coord_fixed(ratio = 1)+theme(legend.key.size = unit(1, "cm"))+
109     theme(legend.text = element_text(size = 20))+
110     theme(legend.title = element_text(size = 20))+
111     theme(legend.key = element_rect(colour = 'white', fill = 'white'))
112   pl
113 }
```

C.3.3 ui.R

```
dashboardPage(
   dashboardHeader(title="Dam Monitoring App"),
   dashboardSidebar(
     width = 300,
     dateInput('date',
               label = h3('Date'),
               value = as.Date("2000-05-15")
     ),
     selectInput('tg', label = h3("Select target in Plot 1"),
                 choices = list("P1DR1" = "P1DR1", "P1DR4" = "P1DR4", "P5DR1" = "P5DR1", "
                     P5DR3" = "P5DR3",
                                "P2IR1" = "P2IR1", "P2IR4" = "P2IR4", "P6IR1" = "P6IR1" , "
                                    P6IR3" = "P6IR3"),
                 selected = "P1DR1"),
     selectInput('tg.2', label = h3("Select target in Plot 2"),
                 choices = list("P1DR1" = "P1DR1", "P1DR4" = "P1DR4", "P5DR1" = "P5DR1", "
                     P5DR3" = "P5DR3",
                                "P2IR1" = "P2IR1", "P2IR4" = "P2IR4", "P6IR1" = "P6IR1" , "
                                    P6IR3" = "P6IR3"),
                 selected = "P2IR1")
   ),
   dashboardBody(
      fluidRow(
        column(10, align="center", offset=2,
            box(           # open the box
               title = "Dam layout",width = 10, status = "primary",
               solidHeader = TRUE, height = 370,
               plotOutput("plot3")
            )            # close the box
        )
      ),
      fluidRow(
            box(
              title = "Plot 1",width = 6, status = "primary",
              solidHeader = TRUE,
              dygraphOutput("dygraph01", height="300px")
            ),            # close the box
            box(
              title = "Plot 2",width = 6, status = "primary",
              solidHeader = TRUE, #height = 370,
              dygraphOutput("dygraph02", height="300px")
            )            # close the box
      ),
      fluidRow(
          column(12, align="center", offset=4,
                box(
                  textOutput("legendDivID"),
                  status = "primary",
                  solidHeader = TRUE,
                  title = "Legend", width=4
                )
          )
      )
   )
)
```

C.3.4 server.R

```
shinyServer(function(input, output) {
dataInput <- reactive({

############  Update data frame for the colors and shapes
    mydate <- input$date
    res2<- which(mydf$date>mydate)
    res2 <- mydf[(res2[1]-1):res2[1],]      # current and last rows
    res12 <- res2[,9:16]                                                  # t and t-1
    res22 <- t(res12)
    res32= data.frame(prev = res22[,1], curr = res22[,2], pends)
    res32$state <- rep('red', 8)
    res32$deviation <- rep('none', 8)
    for(i in 1:nrow(res32)){
      if (abs(res32[i,2])>3) {       # num.sd
        res32$state[i] <- 'red'
      }else if (abs(res32[i,2])>2){  # num.sd
        res32$state[i] <- 'yellow'
      } else
        res32$state[i] <- 'green'
    }
    for (j in 1:nrow(res32)){
      if (res32$state[j] != 'green') {                # yellow or red
        if (res32$curr[j] >0){
          res32$deviation [j]<- 'upstream'
        } else res32$deviation[j] <- 'downstream'
      }
    }
    res32
  })
#################  End of updating ->

  output$plot3 <- renderPlot({
    date.curr <- input$date
    semplot(dataInput(), g, myColors)
  }, height = 300)
  output$dygraph01 <- renderDygraph({
    tg <- input$tg
    max.date <- input$date
    plot_sh(mydf, tg, max.date)
  })
  output$dygraph02 <- renderDygraph({
    tg <- input$tg.2
    max.date <- input$date
    plot_sh(mydf, tg, max.date)
  })
})
```