

*Facultad de Ciencias y Tecnología.
ingeniería Civil.*

*“Análisis de datos de Auscultación Cuantitativa
del bloque E-06 de la Presa ITAIPU utilizando
Redes Neuronales Artificiales”*

UNIVERSIDAD NACIONAL DE CAAGUAZÚ

FACULTAD DE CIENCIAS Y TECNOLOGÍAS



CARRERA DE INGENIERÍA CIVIL

***“Análisis de datos de Auscultación Cuantitativa del
bloque E-06 de la Presa ITAIPU utilizando Redes
Neuronales Artificiales”***

CÉSAR ALBERTO BARRETO MELGAREJO

Coronel Oviedo – Paraguay

Año 2021

*Facultad de Ciencias y Tecnología.
ingeniería Civil.*

*“Análisis de datos de Auscultación Cuantitativa
del bloque E-06 de la Presa ITAIPU utilizando
Redes Neuronales Artificiales”*

***“Análisis de datos de Auscultación Cuantitativa del bloque
E-06 de la Presa ITAIPU utilizando Redes Neuronales
Artificiales”***

Elaborado por

César Alberto Barreto Melgarejo

Tutor

Ing. Fredy G. Ramírez Villanueva

Co - Tutor

Ing. Ricardo E. Miranda Rojas

Trabajo presentado a la Facultad de Ciencias y Tecnología
de la Universidad Nacional de Caaguazú, como requisito para
la obtención del título de Ingeniero Civil

UNIVERSIDAD NACIONAL DE CAAGUAZÚ

FACULTAD DE CIENCIAS Y TECNOLOGÍAS

CARRERA DE INGENIERÍA CIVIL

Coronel Oviedo - Paraguay

Año 2021

*Facultad de Ciencias y Tecnología.
ingeniería Civil.*

*“Análisis de datos de Auscultación Cuantitativa
del bloque E-06 de la Presa ITAIPU utilizando
Redes Neuronales Artificiales”*

PÁGINA DE APROBACIÓN

Trabajo de fin de grado para la obtención del Título de Ingeniero Civil aprobado en representación de la Facultad de Ciencias y Tecnología de la Universidad Nacional de Caaguazú, por el Tribunal Examinador constituido por los siguientes profesores.

Aprobado el _____

Banca Examinadora

AGRADECIMIENTOS

A Dios, por iluminar mis pasos y darme la oportunidad de cumplir este objetivo.

A mi amada madre Nelly, por su amor, comprensión y apoyo incondicional en todos mis proyectos.

A mi querido padre César (†), por enseñarme a crecer como persona con su ejemplo y consejos que llevo presente en todo momento.

A mi hermano Sergio, por la paciencia y calma en los momentos complicados.

A mis amigos y familia en general por el cariño y apoyo en todo momento.

A mis compañeros de trabajo y gerencia, por la excelente predisposición desde el momento que se empezó a desarrollar este trabajo.

A mis tutores, el Ing. Freddy Ramírez y el Ing. Ricardo Miranda, por el apoyo, paciencia y esfuerzo que han dedicado a este trabajo.

DEDICATORIA

Dedicado a una persona que me inculcó desde pequeño el valor del sacrificio y el cariño a la construcción e ingeniería. Le prometí ser Ingeniero civil. Sin prisa, pero sin pausa, primera promesa cumplida, papá.

RESUMEN

Las presas son obras hidráulicas de contención destinadas a diversas finalidades, entre las que se encuentran principalmente la generación de energía, abastecimiento y control de nivel de agua. Este tipo de obras de ingeniería genera gran impacto ambiental, económico y social, es por eso que se pone mucho empeño a la hora de proyectar, construir y monitorear este tipo de estructuras. Para contribuir al monitoreo de las estructuras deben ser instalados instrumentos que permitan evaluar el comportamiento de estas debido a la magnitud de estas y los riesgos que podría generar.

La Presa Itaipú posee más de 2700 instrumentos, 1701 instalados en el hormigón, 1045 en la fundación y 161 para geodesia. Además de 5268 drenes de los cuales 1018 son de hormigón y 4250 de fundación, que acompañan el desempeño de las estructuras y sus fundaciones.

Con una cantidad tan alta de instrumentos y que representan magnitudes distintas resulta complicado realizar un análisis global del comportamiento de la estructura. Este trabajo busca analizar los datos de instrumentos con diferentes funciones y unidades de medida en un bloque representativo de la presa utilizando redes neuronales artificiales de manera a encontrar alguna relación que pueda ayudar a realizar previsiones del comportamiento de un instrumento en base a otro relacionado.

*Facultad de Ciencias y Tecnología.
ingeniería Civil.*

*“Análisis de datos de Auscultación Cuantitativa
del bloque E-06 de la Presa ITAIPU utilizando
Redes Neuronales Artificiales”*

Palabras llave: Itaipú, presa, auscultación, monitoreo, instrumentación, bloque representativo, hormigón, fundación, Redes Neuronales.

ABSTRACT

The dams are hydraulic containment that can work for different purposes, among the main ones are power generation, supply and water level control. This kind of engineering structure are good for environmental, economic and social impact, this is why a lot of effort is put into projecting, building and monitoring dams.

In order to contribute in monitoring the structures, instruments must be installed to allow an evaluation of the behavior of the dam due to their magnitude and the risks that they could generate.

The ITAIPU Dam has approximately 2,700 instruments, 1,701 installed in the concrete, 1045 in the foundation and 161 for geodesy. In addition 5,268 drains of which 1018 are for concrete and 4,250 for foundation, which monitor the performance of the structures and their foundations.

With such a large number of instruments representing different magnitudes, it turns difficult to carry out a global analysis of the behavior of the entire structure. This work seeks to analyze the data of instruments with different functions and units of measurement in a representative block using artificial neural networks, in order to find a relationship that can help to make predictions of the behavior of an instrument based on another one related.

Keywords: ITAIPU, Dam, Auscultation, Monitoring, magnitude, representative block, concrete, foundation, Neural Network

INDICE

PÁGINA DE APROBACIÓN	3
AGRADECIMIENTOS	4
DEDICATORIA	5
RESUMEN	6
ABSTRACT	8
INDICE	13
LISTA DE FIGURAS	16
LISTA DE TABLAS	18
LISTA DE ABREVIATURAS Y SIGLAS	19
INTRODUCCIÓN	13
CAPÍTULO I	15
1.1 PLANTEAMIENTO Y FORMULACIÓN DEL PROBLEMA	15
1.2 OBJETIVOS	16
1.3 SOLUCIÓN PROPUESTA POR EL PROYECTO	17
1.4 JUSTIFICACIÓN	17
1.5 ALCANCE	18
1.6 ANTECEDENTES	19
1.7 METODOLOGÍA	21
1.7.1 Base de datos y análisis	21
1.7.2 Programación	21
CAPÍTULO 2	23
PRESAS	23
2.1 CONCEPTOS	23

2.1.1 Presa	23
2.1.2 Seguridad de Presas	26
2.1.3 Auscultación	26
2.1.4 Monitoreo e Instrumentación	27
2.1.5 Inspección Visual	27
2.2 PRESA DE ITAIPU	28
2.2.1 Disposición general	29
2.2.2 Instrumentos instalados en la presa ITAIPU.	30
2.2.3 Instrumentación del hormigón	31
2.2.4 Instrumentación de la fundación	32
2.2.5 Comportamiento de los bloques	33
2.2.6 Bloque E-06	36
CAPÍTULO 3	40
BASE DE DATOS	40
3.1 ORGANIZACIÓN Y LIMPIEZA DE DATOS	40
3.2 AGRUPACIÓN DE DATOS	41
CAPÍTULO 4	43
APRENDIZAJE AUTOMÁTICO	43
4.1 ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)	43
4.1.1 Definiciones	43
4.1.1.1 Esperanza matemática	43
4.1.1.2 Varianza	44
4.1.1.3 Desviación estándar	45
4.1.1.4 Covarianza	45

4.1.1.5 Coeficiente de correlación	46
4.1.2 El método PCA	47
4.2 REDES NEURONALES ARTIFICIALES	49
4.2.1 Perceptrón	49
4.2.2 Perceptrón multicapa	51
CAPÍTULO 5	57
ANÁLISIS DE DATOS	57
5.1 VISIÓN GENERAL DE LA BASE DE DATOS	57
5.2 ANÁLISIS DE COMPONENTES PRINCIPALES	60
5.2.1 Estandarización de variables	60
5.2.2 Matriz de correlaciones	61
5.2.3 Componentes Principales	65
5.2 DEFINICIÓN DE COMBINACIONES DE VARIABLES.	68
5.3 REDES NEURONALES ARTIFICIALES	70
CAPÍTULO 6	83
CONCLUSIONES	83
6.1 ANÁLISIS DE COMPONENTES PRINCIPALES	83
6.2 REDES NEURONALES	83
6.3 RECOMENDACIONES	84
REFERENCIAS	85
APÉNDICE	88

LISTA DE FIGURAS

Figura 1 - Sección geológica longitudinal de la fundación de la Presa principal de ITAIPU	29
Figura 2 - Disposición general de la Presa ITAIPU.....	30
Figura 3 - Instrumentación de un bloque llave.....	33
Figura 4 - Comportamiento de juntas de dilatación entre bloques.	34
Figura 5 - Descripción de comportamiento del bloque en invierno.	35
Figura 6 - Descripción de comportamiento del bloque en verano. .	36
Figura 7 - Sección transversal del bloque E-06.....	39
Figura 8 - Matriz de datos de entrada.....	42
Figura 9 - Arquitectura Básica de un PERCEPTRÓN.....	50
Figura 10 - Arquitectura del PERCEPTRÓN Multicapa.	51
Figura 11 - Gráfico de comportamiento histórico del PD-E-006.....	58
Figura 12 - Gráfico de comportamiento histórico EM-E-001.....	59
Figura 13 - Gráfico de comportamiento histórico EM-E-002.....	59
Figura 14 - Gráfico de comportamiento histórico. EM-E-003.....	59
Figura 15 - Gráfico de comportamiento histórico. EM-E-004.....	60
Figura 16 - Gráfico de Dispersión PD-E-006/R vs EM-E-001	64
Figura 17 - Gráfico de dispersión PD-E-006/R vs EM-E-002/	64
Figura 18 - Gráfico de dispersión EM-E-001 vs EM-E-002/3.....	65
Figura 19 - Desarrollo de códigos en Jupyter Notebook.....	71
Figura 20 - Predicción de X1 con las combinaciones aplicadas.	73
Figura 21 - Predicción de X3 con las combinaciones aplicadas.	74
Figura 22 - Predicción de X6 con las combinaciones aplicadas.	74

Figura 23 - Datos Predichos X1 vs Datos Reales X1- Total	75
Figura 24 - Datos Predichos X1 vs Datos Reales X1 - PCA 1	75
Figura 25 - Datos Predichos X1 vs Datos Reales X1 -Técnico1	76
Figura 26 - Datos Predichos X3 vs Datos Reales X3 - Total 2.....	76
Figura 27 - Datos Predichos X3 vs Datos Reales X3 - PCA 2.....	77
Figura 28 - Datos Predichos X3 vs Datos Reales X3 - Técnico 2...	77
Figura 29 - Datos Predichos X6 vs Datos Reales X6 - Total 3.....	78
Figura 30 - Datos Predichos X6 vs Datos Reales X6 - PCA 3.....	78
Figura 31 - Datos Predichos X6 vs Datos Reales X6 - Técnico 3...	79
Figura 32 - Datos Predichos X8 vs Datos Reales X8 - Total 4.....	79
Figura 33 - Datos Predichos X8 vs Datos Reales X8 - PCA 4.....	80
Figura 34 - Datos Predichos X8 vs Datos Reales X8- Técnico 4....	80
Figura 35 - Datos Predichos X10 vs Datos Reales X10 - Total 5....	81
Figura 36 - Datos Predichos X10 vs Datos Reales X10 - Total 5....	82
Figura 37 - Datos Predichos X10 vs Datos Reales X10 - Técnico 5.	82
Figura 38 - Bloque E-06. Fase aguas abajo.	88
Figura 39 - Códigos de desarrollo en Jupyter Notebook.	89
Figura 40 - Códigos de desarrollo en Jupyter Notebook.	89
Figura 41 - Códigos de desarrollo en Jupyter Notebook.	90
Figura 42 - Códigos de desarrollo en Jupyter Notebook.	90

LISTA DE TABLAS

Tabla 1 - Resumen de instrumentos para el análisis.	37
Tabla 2 - Instrumentos seleccionados para el análisis de datos.....	40
Tabla 3 - Datos estandarizados	61
Tabla 4 - Evaluación sistematizada del coeficiente de correlación lineal.	62
Tabla 5 - Matriz de correlaciones "R"	62
Tabla 6 - Porcentaje de variabilidad explicado de cada componente principal	65
Tabla 7 - Autovalores de las componentes principales correspondientes a cada variable.	67
Tabla 8 - Combinaciones utilizando PCA.....	68
Tabla 9 - Combinaciones técnicas.....	70
Tabla 10 - Coeficientes R2 de todas las combinaciones.	71
Tabla 11 - Coeficientes R2 de todas las combinaciones ordenadas por variable	72

LISTA DE ABREVIATURAS Y SIGLAS

Auscultación

EM	Extensómetro de barras múltiples
Hz	Hertz
JS	Base de elongómetro
MCA	Metros de columna de agua
PD	Péndulo Directo
PS	Piezómetro Standpipe
RD	Roseta de Deformímetro
RT	Roseta de Tensómetro
SGM	Sistema de Gestión de Mediciones
TS	Termómetro de Superficie
H	Altura
l/s	Litros por segundo
ρ	Coeficiente de correlación
°C	Grados Celsius
Ω	Ohm

Aprendizaje Automático

α	Factor de escala para el cálculo del número de neuronas
α_i	Multiplicadores de LaGrange
\bar{x}	Media muestral
β	Matriz de pesos de salida
ϵ	Vector error
Λ	Matriz diagonal de valores característicos

\hat{a}	VARIABLES DE ENTRADA NORMALIZADAS y
A	Matriz principal del sistema de ecuaciones lineales
$E[X]$	Esperanza matemática de la variable aleatoria X
H	Matriz de salida de la capa oculta
Q	Matriz ortogonal para descomposición espectral de la matriz de covarianzas
R	Matriz de correlaciones
S	Matriz de covarianzas
X	Vectores de variables aleatorias
Y	Vectores de valores de salida
Z	Vectores de componentes principales
μ	Media de la población
$\rho(X,Y)$	Coefficiente de correlación entre las variables aleatorias X e Y
σ_X	Desviación estándar de la variable aleatoria X
b	Bías
c	Índice de la capa de salida
g	Índice de función de activación
i, j	Índices
k, d	Índices
m	Número de variables de salida
N	Número de puntos. Tamaño del conjunto de entrenamiento.
n_0	Número de neuronas en la capa de entrada
n_h	Número de neuronas en la capa oculta
n_k	Número de neuronas en la capa k

$\rho_X(x)$ Probabilidad de que la variable aleatoria X asuma el valor x

x_j Variables independientes

y Variable dependiente

NE Tamaño del conjunto de entrenamiento

NV Tamaño del conjunto de verificación

INTRODUCCIÓN

ITAIPU Binacional es actualmente la mayor generadora de energía eléctrica del mundo, y es responsable de proveer el 86% de la energía consumida por Paraguay y 15% en Brasil.

Teniendo en cuenta la magnitud de esta obra de ingeniería, con 196 metros de altura máxima y un embalse de 29 millones de metros cúbicos, la vigilancia de la seguridad de presas es un aspecto fundamental ya que esta garantiza que sus estructuras están correctamente monitoreadas y permanecen estables. Este monitoreo se realiza a través de los instrumentos de auscultación sumados a las inspecciones visuales que son realizadas periódicamente por el equipo técnico.

La utilización de técnicas de observación y desenvolvimiento de sistemas que ayuden a interpretar y validar los datos provenientes de los instrumentos de auscultación han sido foco de varios trabajos e investigaciones en los últimos años, debido a la importancia que tienen estas estructuras. Por este motivo el presente trabajo busca aportar, a través de la utilización de redes neuronales artificiales, un método alternativo de validación de datos provenientes de la auscultación cuantitativa. Que ayude a observar desde otra perspectiva los resultados obtenidos en la instrumentación y nos ofrezca un panorama más amplio, ya que es un campo donde se pueden obtener bastantes beneficios. Varias disciplinas, no solo en el ámbito ingenieril ya están implementando esta

herramienta para optimizar diferentes análisis y avanzar en sus respectivas áreas. Las Redes Neuronales han demostrado ser eficiente en aprender el comportamiento de diversos fenómenos, pudiéndose considerar los resultados para datos que no fueron específicamente analizados de modo a obtener respuestas futuras. A pesar de ser conocido como un método eficiente para interpretar el comportamiento de fenómenos, identificar patrones y predecir valores, las redes neuronales necesitan datos de entrada representativos para que los valores se aproximen a la realidad.

CAPÍTULO I

1.1 PLANTEAMIENTO Y FORMULACIÓN DEL PROBLEMA

La colecta y el análisis de datos son de vital importancia para la Seguridad de Presas, ya que con ellos se realizan las evaluaciones del comportamiento de las estructuras. Se utiliza una gran cantidad de instrumentos que monitorean distintas magnitudes y regiones de las estructuras, pero a su vez forman parte de un mismo conjunto, razón por la cual nos cuestionamos si existe alguna relación en el comportamiento de estos, y si podrían colaborar en la optimización del proceso de validación de datos y evaluación del comportamiento de las estructuras disminuyendo el margen de error.

1.2 OBJETIVOS

1.2.1 OBJETIVO GENERAL

- Analizar los datos provenientes de la auscultación cuantitativa del bloque E-06 de la Presa ITAIPU utilizando redes neuronales artificiales.

1.2.2 OBJETIVOS ESPECÍFICOS

- Estudiar la relación entre el comportamiento de los diferentes instrumentos que monitorean el bloque E-06 utilizando Redes Neuronales Artificiales.
- Demostrar que las Redes Neuronales Artificiales pueden ser una herramienta de apoyo para validar los datos de auscultación cuantitativa.
- Exponer sobre la importancia de la utilización de nuevas herramientas en el monitoreo de presas.

1.3 SOLUCIÓN PROPUESTA POR EL PROYECTO

Utilizar la metodología de las Redes Neuronales Artificiales, como herramienta para encontrar posibles relaciones entre las mediciones de los instrumentos de auscultación y optimizar la colecta y el análisis de los datos.

1.4 JUSTIFICACIÓN

A lo largo de la historia han ocurrido varias fallas y accidentes en presas. A razón de esto surgieron varias organizaciones como el Comité Internacional de Seguridad de Presas (ICOLD), el Comité Nacional de Grandes Presas (SPANCOLD) y en nuestro país el Comité Paraguayo de Grande Presas (CPDP), entre muchas otras, con el objetivo de difundir información sobre las presas en todas las etapas del proyecto.

La Seguridad de presas toma un papel importante debido a que una falla en las estructuras puede causar un gran impacto económico, social y ambiental.

ITAIPU en conjunto con el Parque Tecnológico de ITAIPU (PTI) y el Centro de Estudios Avanzados en Seguridad de Presas (CEASB) constantemente realizan investigaciones y proyectos para implementar nuevas tecnologías y métodos para el monitoreo de las estructuras. Por esta razón se busca aportar una herramienta que podría ayudar a optimizar los procesos de colecta y análisis de datos.

1.5 ALCANCE

Este trabajo consistirá en analizar la relación entre los datos provenientes de los instrumentos de auscultación, específicamente el péndulo y los extensómetros instalados en el bloque llave E.06 de la Presa ITAIPU utilizando Análisis de Componentes Principales y Redes Neuronales Artificiales.

1.6 ANTECEDENTES

- D. LISKA (2019) presenta en su trabajo de posgraduación “*Análise do comportamento de um bloco chave da barragem de ITAIPU através da correlação da instrumentação.*” un análisis de la correlación existente entre los instrumentos del bloque llave D58 de la presa ITAIPU.

- E. LAZZAROTO (2016) propone en su tesis “*Obtenção de limites estatísticos de controle em gráficos de controle univariados e multivariados aplicados a dados de instrumentação de barragens*” la correcta interpretación y evaluación de la información de los datos durante el periodo operacional de la presa utilizando diferentes métodos estadísticos univariados y multivariados para crear límites estadísticos de los datos correspondientes a instrumentos la presa de ITAIPU.

- LOPES R. (2009) concluye en su publicación “*Emprego de modelo regressivo linear e rede neural artificial no controle da segurança de barragens de concreto*” que ambos modelos eran aceptables, siendo el de redes neuronales más correspondientes con los valores observados.

- B. MEDEIROS (2013) en su trabajo de posgraduación “*Análise por meio de Redes Neurais Artificiais dos dados do monitoramento dos piezômetros da barragem de concreto de ITAIPU*” utiliza las redes neuronales para analizar el comportamiento de los valores piezómetros en la fundación.

- J. LOPES (2011) en su disertación “*Aplicação de redes neuronais*

no planeamento de emergência em barragens de aterro” nos enseña un modelo de redes neuronales para previsión del nivel de alerta para la presa de Valtorno/Murao,

- J. CAIRES DA MATA (2008) en su tesis *“Aplicação de redes neuronais ao controle de segurança de barragens de Betão”* nos muestra la utilización de redes neuronales artificiales tipo Perceptrón Multicapa.

- J. DE SOUZA (2019) analiza en su tesis *“Estudo da correlação entre a variação do lago de ITAIPU e a piezometria da fundação do bloco E-06”* la correlación existente entre estas dos variables mencionadas.

1.7 METODOLOGÍA

1.7.1 Base de datos y análisis

Las bases de datos corresponden a las mediciones de los instrumentos del bloque llave E-06, desde su puesta en operación hasta la fecha de inicio del análisis de este trabajo. Los datos fueron modificados con una constante antes de su uso en el presente trabajo por cuestiones de confidencialidad. Esta modificación no altera los procesos ni resultados.

Para realizar el análisis se adopta la siguiente metodología:

- Organización de datos de cada instrumento.
- Limpieza y agrupación de datos por fecha de realización de lectura.
- Elaboración de hipótesis con bases estadísticas y técnicas para seleccionar las variables a ser relacionadas.
- Desarrollo de algoritmos de las Redes Neuronales Artificiales.
- Entrenamiento de las Redes Neuronales con los datos ordenados.
- Análisis de los resultados y conclusiones.

1.7.2 Programación

Los algoritmos computacionales se desarrollaron en Python 3 con sus respectivos paquetes científicos.

Según Gutttag (2013), Python es un lenguaje de programación de propósito general que puede ser utilizado efectivamente para construir casi cualquier clase de programas que no necesitan acceso directo al hardware

de computadora.

Entre las fortalezas de Python podemos destacar que viene junto con una librería estándar muy completa, además de miles de librerías de terceros que proporcionan utilidades más potentes y sofisticadas, como la librería numérica de Numpy y la librería simbólica Simpy (Summerfield,2010).

CAPÍTULO 2

PRESAS

2.1 CONCEPTOS

2.1.1 Presa

Una presa puede definirse como una barrera o una estructura colocada cruzando un curso de agua o de un río para retener el agua y así controlar el caudal (SPANCOLD, 2021).

Según el UNITED STATES DEPARTMENT OF THE INTERIOR BUREAU OF RECLAMATION WASHINGTON, D.C. (1966), las presas se pueden clasificar en un número de categorías diferentes que depende del objeto de la clasificación. Considera tres amplias clasificaciones de acuerdo con: El uso, el proyecto hidráulico, o los materiales que forman parte de la estructura.

Clasificación según el uso: las presas se pueden clasificar de acuerdo con la función más general que van a desempeñar, como de almacenamiento, de derivación o regulación.

Presas de almacenamiento, se construyen para embalsar el agua en los periodos en que sobra para utilizarla cuando escasea. Estos periodos pueden ser estacionales, anuales o más largos. Las presas de almacenamiento se pueden clasificar a su vez de acuerdo con el objeto del almacenamiento, como para abastecimiento de agua, para recreación, para la cría de peces y animales salvajes, para generación de energía

hidroeléctrica, irrigación, etc.

Las *presas de derivación* se construyen ordinariamente de manera a proporcionar la carga necesaria para desviar el agua hacia zanjas, canales u otros sistemas de conducción al lugar en que se van a usar. Se utilizan en los sistemas de riego, para la derivación de una corriente natural hacia un vaso de almacenamiento fuera del cauce natural de la corriente, para usos municipales e industriales, o para una combinación de estos.

Las *presas reguladoras* se construyen para retardar el escurrimiento de las avenidas y disminuir el efecto de las ocasionales. Las presas reguladoras se dividen en dos tipos. En uno de ellos, el agua se almacena temporalmente, y se deja salir por una obra de toma con un gasto que no exceda de la capacidad del cauce de aguas abajo. En el otro tipo, el agua se almacena tanto tiempo como sea posible y se deja infiltrar en las laderas del valle o por los estratos de grava de la cimentación, a este último se le llama algunas veces de distribución o dique, porque su principal objeto es recargar el acuífero. Las presas reguladoras también se construyen para detener los sedimentos. A menudo se les llama presas para arrastres.

Clasificación según su proyecto hidráulico. Las presas se pueden clasificar también como presas vertedoras o no vertedoras

Las *presas vertedoras* se proyectan para descargar sobre sus coronas. Deben estar hechas de materiales que no se erosionen con tales descargas.

Presas no vertedoras son las que se proyectan para que no rebase el agua por su corona. Este tipo de proyecto permite ampliar la elección de materiales incluyendo las presas de tierra y de enrocado.

Clasificación según los materiales. La clasificación más común que se usa en la discusión de los procedimientos de construcción se basa en los materiales que forman la estructura.

Según Novak, Moffat, & Nalluri (2001), existen diversos tipos de estructuras que son utilizadas para la construcción de una presa, cada una con sus particularidades. Dependiendo de los principales materiales de construcción utilizados, es posible hacer una amplia clasificación inicial en dos grupos:

Presas de relleno que se construyen con terraplenes de los suelos o enrocados. Las pendientes de los paramentos aguas arriba y aguas abajo son similares y con un ángulo moderado, lo que produce una sección ancha y con volumen de construcción grande con respecto a su altura.

Presas de hormigón que se construyen con hormigón macizo. Los taludes de los paramentos son diferentes, en general muy fuertes aguas abajo y casi verticales aguas arriba, estas presas tienen perfiles relativamente esbeltos según el tipo.

Esta clasificación no es absoluta, ya que se pueden utilizar de forma mixta como ya se han construido varias presas.

2.1.2 Seguridad de Presas

La seguridad de presas es el manejo coordinado de distintas disciplinas técnicas especializadas que intervienen en el diseño, cálculo, construcción, mantenimiento, operación y remoción de una presa, con el objeto de alcanzar y mantener el mayor nivel de seguridad factible. (Novak, Moffat, & Nalluri, 2001)

Según la Instrucción para el Proyecto, Construcción y Explotación de Grandes Presas (1967), en su Artículo 1 define a las Grandes Presas como todas las presas de más de 15 m de altura, o las presas entre 10 y 15 m de altura que respondan a una, al menos, de las condiciones siguientes:

- Capacidad de embalse superior a 100.000 m³.
- Características excepcionales de cimientos o cualquier otra

circunstancia que permita calificar la obra como importante para la seguridad o economía pública.

2.1.3 Auscultación

La auscultación es el conjunto de formas de observación del comportamiento de la represa y sus fundaciones, para controlar sus condiciones de seguridad, comprobar la validez de las hipótesis y de los métodos de cálculo utilizados en el proyecto y verificar la necesidad de medidas correctivas. (ITAIPU Binacional, 2021)

2.1.4 Monitoreo e Instrumentación

Según la Comisión Federal de Regulación de Energía (FERC, 1994), el monitoreo es la colecta, reducción, presentación y evaluación de los datos de instrumentación. Por lo tanto, el monitoreo y la instrumentación son herramientas que deben utilizarse con un programa de inspección atento para evaluar continuamente la seguridad de las presas.

El propósito de la instrumentación y el monitoreo es mantener y mejorar la seguridad de la presa mediante la provisión de información para:

- 1) evaluar si una presa está funcionando como se esperaba y
- 2) advertir de cambios que puedan poner en peligro la seguridad de esta.

La instrumentación y el monitoreo, combinados con la observación visual, pueden proporcionar alerta temprana de muchas condiciones que podrían contribuir a fallas e incidentes de represas.

2.1.5 Inspección Visual

Una forma de observar el comportamiento de la presa son las inspecciones visuales, que complementan los resultados de instrumentación y auscultación. (Itaipu Binacional, 2021)

2.2 PRESA DE ITAIPU

La Usina Hidroeléctrica de ITAIPU se encuentra sobre el Río Paraná, en la frontera entre Paraguay y Brasil, región constituida por derrames basálticos de la Cuenca del Paraná.

Según Silveira (2003) las principales características geológicas del local son las siguientes:

- Derrames basálticos esencialmente horizontales, con 20 a 70 metros de espesor.
- Camadas de brecha basáltica entre los derrames basálticos, con espesor variable entre 1 y 30 m, en general heterogéneas y normalmente menos resistentes y más deformables que el basalto.
- Discontinuidades en planos paralelos a los derrames basálticos generalmente localizadas en los contactos entre derrames o en la base de la zona de transición.
- Permeabilidad horizontal varias veces superior a la permeabilidad vertical.

En la fundación de las estructuras de hormigón de la presa de ITAIPU fueron investigados cinco derrames basálticos, que fueron nombrados en orden ascendente de A, B, C, D y E con espesores variables de 30 a 70 metros como se puede observar en la Figura 1.

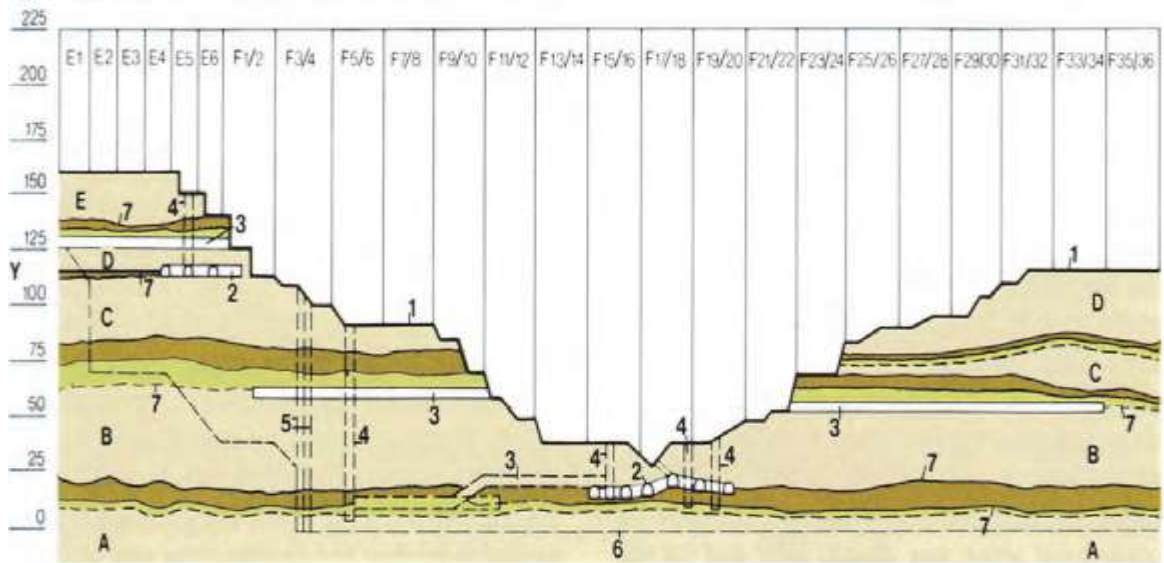


Figura 1 - Sección geológica longitudinal de la fundación de la Presa principal de ITAIPU. Fuente: ITAIPU Binacional

Y - Elevación (m)

1 - Perfil de excavación

2 - Chavetas

3 - Túneles de exploración

4 - Pozos de prospección

5 - Cortina de inyección transversal

6 - Límite de la cortina de inyección

7 - Discontinuidades

2.2.1 Disposición general

Como se muestra en la Figura 2, se asignó una letra a cada tramo correspondiente a los distintos tipos de estructuras. Los tramos correspondientes a estructuras de hormigón fueron construidos por

bloques. Y se enumeran de derecha a izquierda hidráulica. Mientras las estructuras de enrocado y tierra se trazan por progresiva.

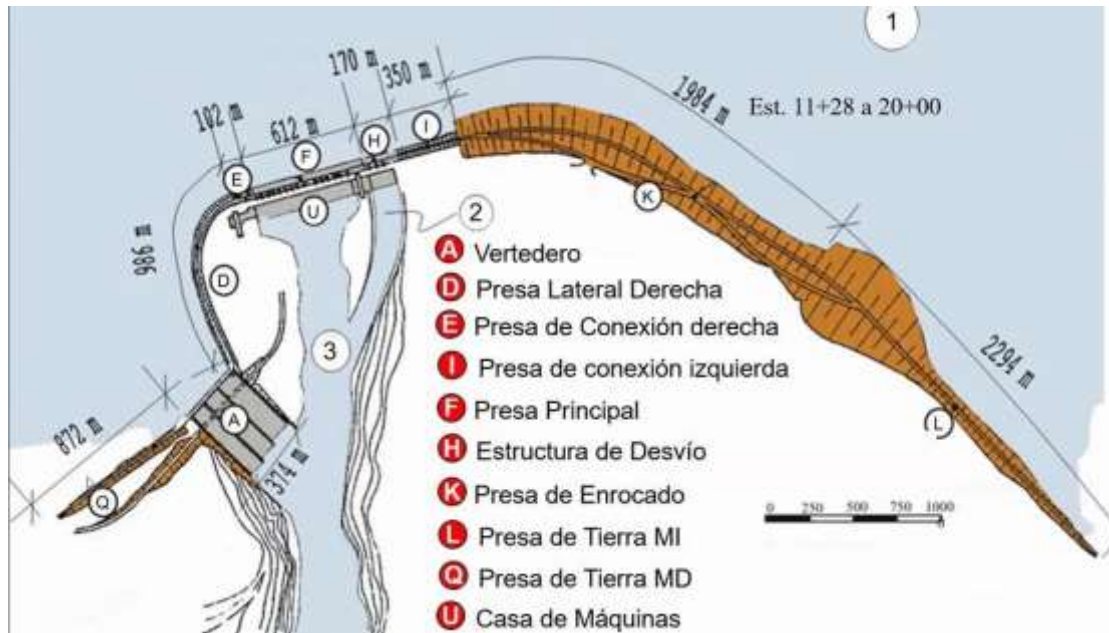


Figura 2 - Disposición general de la Presa ITAIPU. Fuente: ITAIPU Binacional

2.2.2 Instrumentos instalados en la presa ITAIPU.

La presa cuenta con aproximadamente 2700 instrumentos, 1701 instalados en el hormigón, 1045 en la fundación y 161 para geodesia. Además de 5268 drenes de los cuales 1018 son de hormigón y 4250 de fundación, que acompañan el desempeño de las estructuras y sus fundaciones.

A continuación, se presentan los instrumentos de auscultación instalados en las distintas estructuras de la presa. (ITAIPU Binacional, 2021)

2.2.3 Instrumentación del hormigón

- Caja selectora: Reúne los cables eléctricos provenientes de varios instrumentos en una caja central que, al ser conectada al aparato de lectura, proporciona los datos de estos instrumentos.
- Péndulo directo: Mide los desplazamientos horizontales de puntos situados en determinadas cotas de los bloques instrumentados de la presa con relación a la fundación de la estructura.
- Péndulo invertido: Mide los desplazamientos de la base de la presa con relación a un punto de la fundación suficientemente profundo como para ser considerado fijo.
- Medidor eléctrico de junta: Mide los desplazamientos de abertura y cierre en determinadas juntas de contracción de estructuras de hormigón.
- Base de elongómetro: Mide abertura, cierre, asentamiento y desplazamiento entre bloques, o juntas de monolitos.
- Deformímetro de armadura: Mide las tensiones en barras de armadura, en el interior de estructuras de hormigón.
- Tensómetro de hormigón: Mide la tensión en el interior de estructuras de hormigón.
- Deformímetro de hormigón: Mide la deformación del hormigón y, por

esta deformación, se obtiene la tensión que obra dentro de la estructura.

- Termómetro de resistencia: Mide la temperatura en el interior de la estructura de hormigón.

2.2.4 Instrumentación de la fundación

- Medidores de caudal: Mide los caudales de percolación a través de las estructuras y fundaciones de las obras de tierra y hormigón.
- Extensómetros de barras múltiples: Mide las deformaciones de la fundación con relación al punto de anclaje de su barra.
- Piezómetro Standpipe: Permite conocer la subpresión actuante en el lugar de su instalación.
- Piezómetro eléctrico: Permite conocer la subpresión actuante en el lugar de su instalación.
- Medidor de asentamiento IPT: Mide deformaciones verticales ocurridas en las presas de tierra.
- Célula de presión total: Mide las presiones totales actuantes en la zona de contacto suelo-hormigón
- Medidor Triortogonal: Mide los desplazamientos entre juntas de hormigón y zonas fracturadas en macizos rocosos
- Medidor de nivel del agua: Mide el nivel de agua presente en el

subsuelo (napa freática).

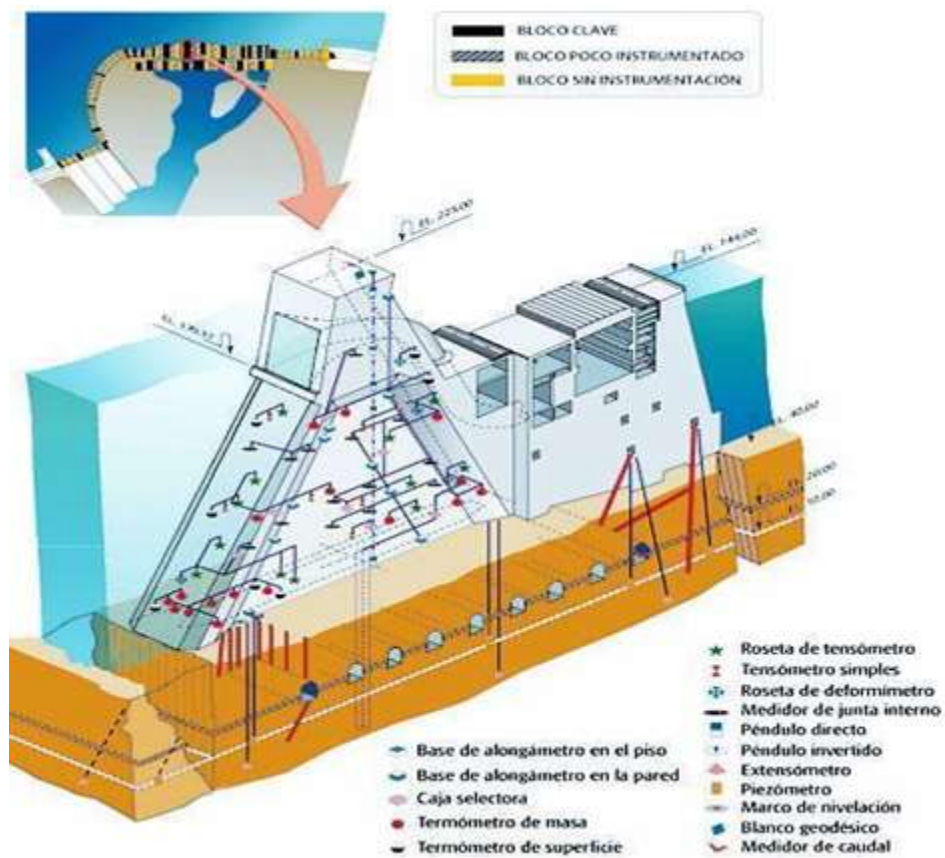


Figura 3 - Instrumentación de un bloque llave. Fuente: ITAIPU Binacional

2.2.5 Comportamiento de los bloques

En esta sección se presenta el comportamiento de un bloque de hormigón, con el objetivo de comprender mejor cómo esto se puede reflejar en las mediciones de los instrumentos.

Es bien sabido que el hormigón sufre variaciones de volumen por causa de la temperatura. En la figura 4 podemos visualizar el efecto de dilatación y contracción del hormigón en las juntas de los bloques, generando el cierre y abertura de estas. El monitoreo de las juntas se

realiza con los medidores triortogonales de juntas y las bases de elongómetros. Que miden los movimientos relativos entre bloques en tres ejes.

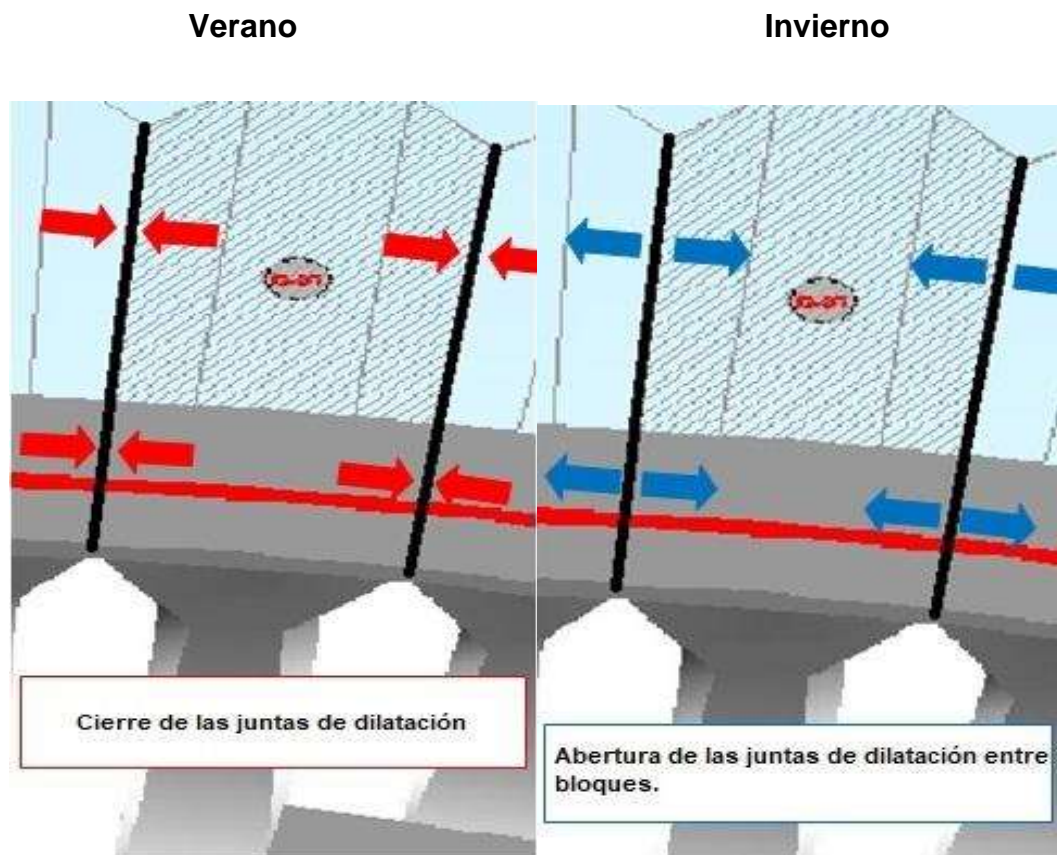


Figura 4 - Comportamiento de juntas de dilatación entre bloques.
Fuente: Adaptado de (Liska, 2019)

Las figuras 5 y 6 corresponden al corte transversal de un bloque de hormigón de la presa ITaipu, donde se puede observar el mismo fenómeno físico que en el párrafo anterior, pero en este caso la fase de aguas arriba se encuentra en contacto permanente con el agua, por tanto, la variación es menor en comparación a la fase de aguas abajo que está expuesto a la temperatura ambiente.

Entendiendo esto podemos comprender mejor la descripción que realizan las figuras. En la Figura 5 podemos ver como en temporadas de bajas temperaturas, el desplazamiento de la cresta de la presa tiende a dirigirse en dirección a aguas abajo debido a la mayor variación de temperatura que ocurre en la fase de aguas abajo y esto provoca una mayor contracción del hormigón en esta fase. Lo contrario ocurre en épocas de temperaturas altas donde el hormigón se dilata y la presa tiende a moverse hacia aguas arriba.

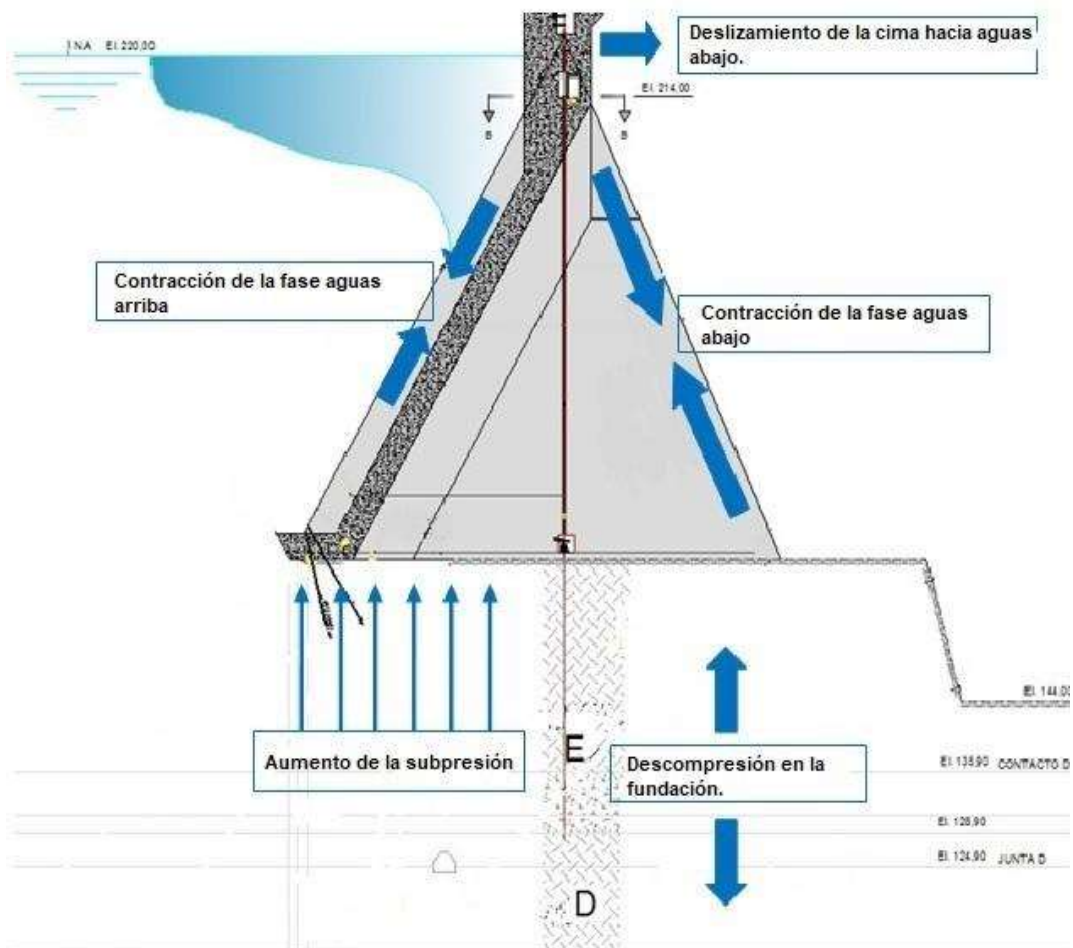


Figura 5 - Descripción de comportamiento del bloque en invierno.

Fuente. Adaptado de (Liska, 2019)

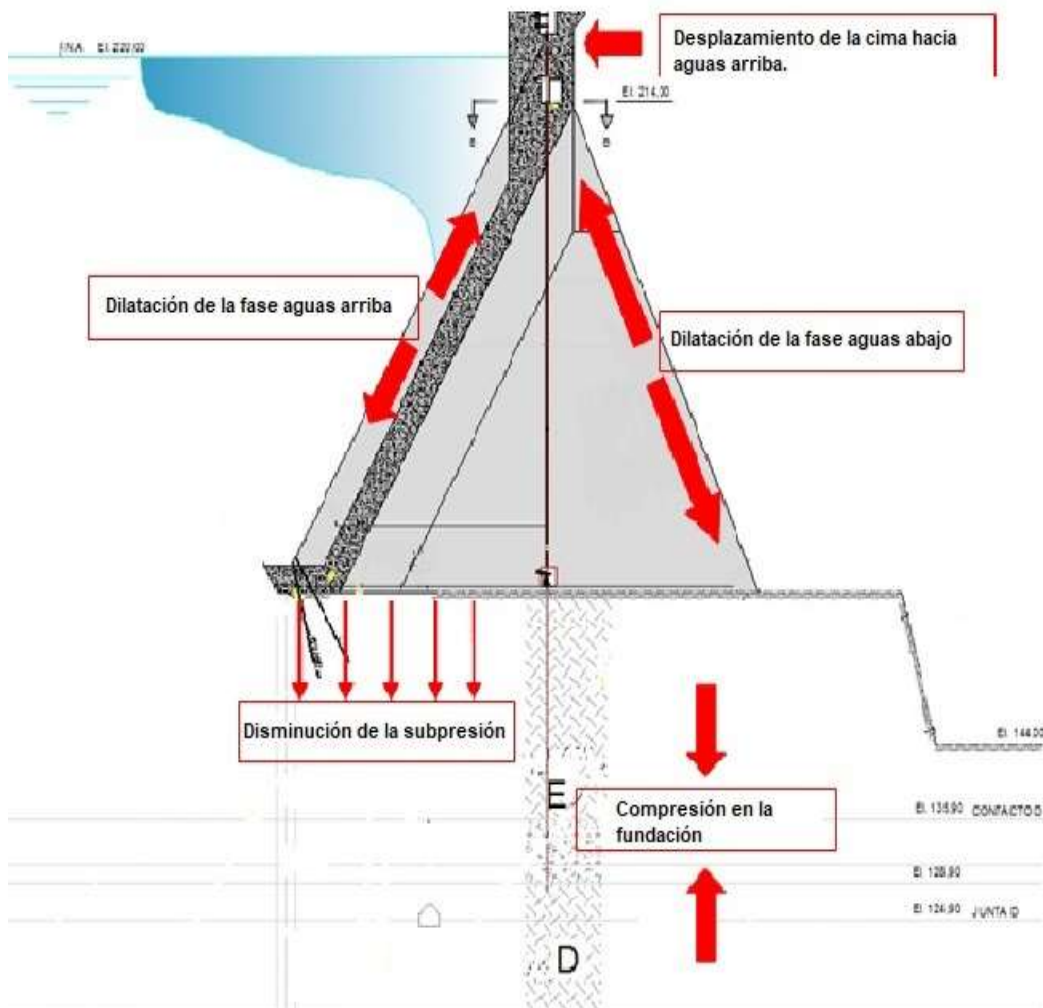


Figura 6 - Descripción de comportamiento del bloque en verano.
Fuente. Adaptado de (Liska, 2019)

2.2.6 Bloque E-06

El bloque E-06 está situado en la presa de Conexión Derecha de ITaipu. La altura de este bloque varía de 85 m aguas arriba y 100 m. aguas abajo, el ancho del bloque es de 15 m. En cuanto a su tipología, está clasificada como presa de contrafuerte.

El bloque E-06 es considerado un bloque llave, debido a la importancia

y las condiciones en la que fue construida, ya que se encuentra en la conexión con la Presa Principal y la fundación es de altura variable. Posee una amplia gama de instrumentos que monitorean distintas magnitudes.

Para este trabajo utilizaremos dos tipos de instrumentos, los cuales fueron el Péndulo Directo y los Extensómetros. En la Tabla 1 se presentan un resumen de los instrumentos que serán utilizados para el análisis del presente trabajo.

Tabla 1 - Resumen de instrumentos para el análisis.

Instrumento	Cantidad	Función del instrumento
Extensómetro de barras múltiples.	4	Conocer las deformaciones de la fundación con relación al punto de anclaje de la barra
Péndulo directo	1	Medir los desplazamientos horizontales de puntos situados en determinadas cotas de los bloques instrumentados de la presa con relación a la fundación de la estructura.

Son 12 puntos de medición en total los que serán utilizados para el análisis (Tabla 2). Estos están compuestos por 4 bases de extensómetros que miden las deformaciones en diferentes cotas y zonas, como así también el Péndulo directo que mide los desplazamientos horizontales en

las direcciones Radial y Tangencial en diferentes cotas del bloque.

La elección de estos instrumentos entre toda la gama existente se debe a que están estrechamente ligados, ya que los desplazamientos del bloque repercuten directamente en la fundación como podemos ver en la sección anterior donde se explica el comportamiento de los bloques con relación a la variación de temperatura.

Es importante mencionar que en este trabajo inicialmente fueron incluidos los piezómetros, pero estos tienen un tiempo de respuesta diferente en comparación a los instrumentos seleccionados, y no reflejaban buenos resultados por este motivo, razón por la cual fueron descartados del análisis.

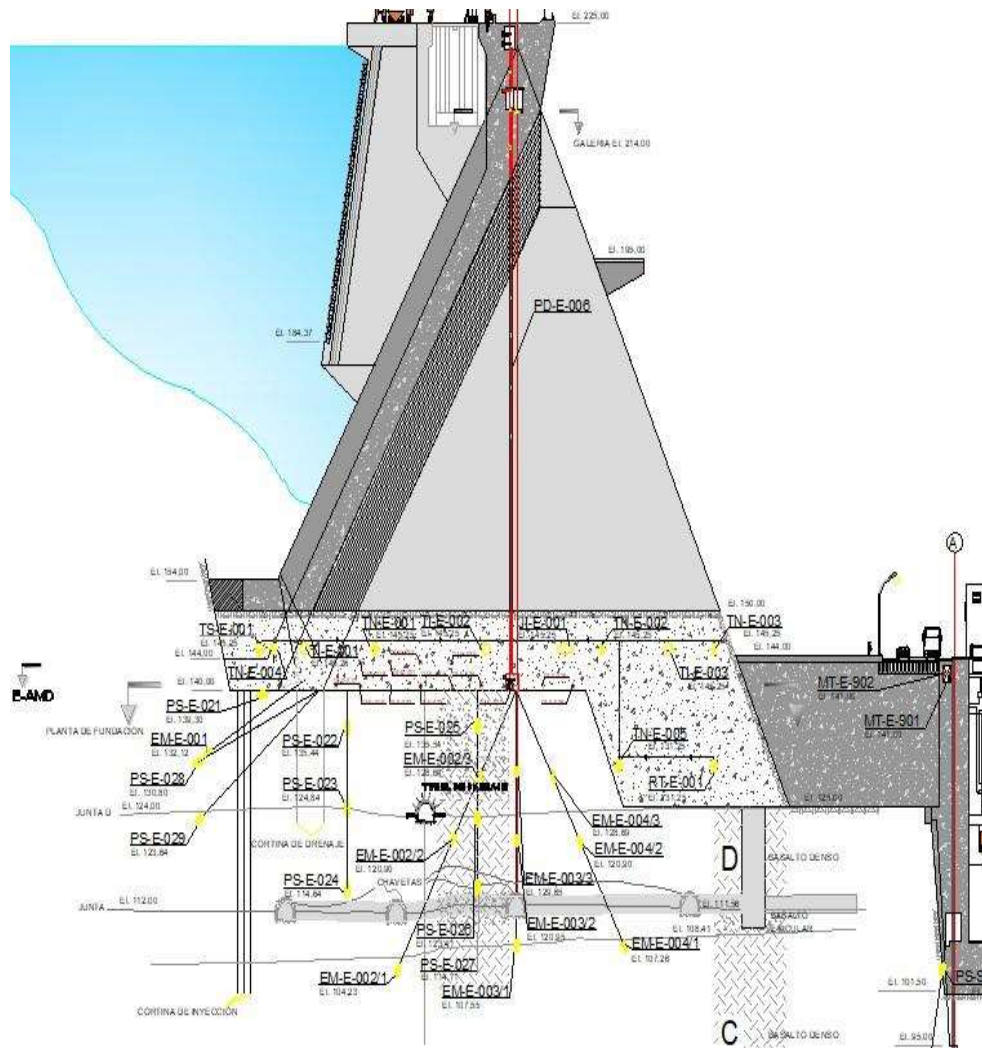


Figura 7 - Sección transversal del bloque E-06. Fuente: ITaipu Binacional

CAPÍTULO 3

BASE DE DATOS

3.1 ORGANIZACIÓN Y LIMPIEZA DE DATOS

Los datos de los instrumentos del bloque E-06 datan de la fecha de operación de los instrumentos de octubre de 1982. Para este estudio, fueron tenidos en cuenta 5 instrumentos (4 Extensómetros y un Péndulo Directo).

Teniendo en cuenta que las bases de extensómetro poseen hasta 3 barras de medición, y el péndulo es medido en dos direcciones, radial y tangencial, se obtienen así 12 fuentes de datos en total.

Cada instrumento posee una cantidad distinta de datos, debido a que durante el periodo de construcción y llenado del embalse la periodicidad era mayor, para monitorear el comportamiento de estos.

Tabla 2 - Instrumentos seleccionados para el análisis de datos.

Instrumento	Código del instrumento	Puntos de medición.	Número de datos	Variable asignada
Péndulo Directo	PD-E-006	Radial	742	X1
		Tangencial	742	X2
	EM-E-001	1	901	X3
		1	873	X4

Extensómetro de barra múltiple.	EM-E-002	2	873	X5
		3	873	X6
	EM-E-003	1	864	X7
		2	864	X8
		3	864	X9
	EM-E-004	1	876	X10
		2	876	X11
		3	876	X12

3.2 AGRUPACIÓN DE DATOS

Una vez ordenados estos datos, se procedió a la agrupación de estos, teniendo en cuenta las fechas de lectura de los instrumentos, obteniendo así una matriz de datos de 12x 670. Totalizando 8052 lecturas en la base de datos que serán procesados.

Teniendo en cuenta que cada instrumento mide diferentes magnitudes, es importante seleccionar los instrumentos adecuados para realizar un relacionamiento entre estos, utilizando redes neuronales. Para este fin se adoptarán 3 configuraciones de datos con el objetivo de analizar cuál es más efectiva.

La primera será utilizando todas las variables, de manera a analizar si es que todos los instrumentos tienen alguna correlación en sus

comportamientos.

En la segunda combinación se aplicará el Análisis de Componentes Principales (PCA, por sus siglas en inglés) con el fin de identificar las componentes principales de la base de datos y cuáles son las variables con más influencia en cada componente principal de manera a relacionarlas entre sí.

Por último, se seleccionarán las variables teniendo en cuenta aspectos técnicos y de funcionamiento de los instrumentos. Por ejemplo, elegir variables correspondientes a instrumentos que miden la misma discontinuidad. Esta combinación se explicará con más detalle en el Capítulo 5.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
0	14.896145	-3.885575	7.54210	1.44275	0.79600	0.52735	0.05970	-0.11940	-0.10945	-1.34325	-1.15420	-0.70645
1	15.215440	-3.813934	7.54210	1.51240	0.83580	0.61690	0.06965	-0.12935	-0.11940	-1.31340	-1.11440	-0.67660
2	15.368173	-3.532250	7.29335	1.55220	0.89550	0.64675	0.10945	-0.08955	-0.07960	-1.27360	-1.06465	-0.64675
3	14.451380	-1.891694	6.69635	1.47260	0.80595	0.59700	0.07960	-0.11940	-0.11940	-1.29350	-1.08455	-0.65670
4	14.298747	-1.530111	6.55705	1.44275	0.78605	0.54725	0.08955	-0.10945	-0.10945	-1.30345	-1.10445	-0.68655
...
665	-1.492500	-0.298500	0.31840	0.12935	0.07960	0.12935	-0.19900	-0.16915	-0.11940	-0.37810	-0.39800	-0.24875
666	-1.393000	-0.497500	0.08955	0.10945	0.06965	0.11940	-0.16915	-0.14925	-0.09950	-0.27860	-0.31840	-0.19900
667	-0.796000	-0.497500	-0.12935	0.07960	0.05970	0.09950	-0.17910	-0.11940	-0.09950	-0.08955	-0.20895	-0.11940
668	-0.199000	-0.199000	-0.25870	0.08955	0.05970	0.10945	-0.17910	-0.11940	-0.06965	-0.06965	-0.17910	-0.09950
669	0.000000	0.000000	-0.25870	0.05970	0.03980	0.07960	-0.13930	-0.11940	-0.06965	-0.06965	-0.16915	-0.08955

670 rows × 12 columns

Figura 8 - Matriz de datos de entrada. Fuente: Elaboración propia

CAPÍTULO 4

APRENDIZAJE AUTOMÁTICO

4.1 ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

Cuando recogemos información de un conjunto de datos, lo ideal es tomar el mayor número de variables posibles. No obstante, si tomamos demasiadas variables, es difícil visualizar las relaciones entre estas variables. Además de esto, las fuertes correlaciones que se presentan entre algunas variables indican otro problema, ya que normalmente están relacionadas o miden lo mismo desde distintos puntos de vista.

Entonces se hace necesario reducir el número de variables. Resaltando que el concepto de mayor información se relaciona con el de mayor variabilidad de datos o varianza. Por tanto, cuanto mayor sea la variabilidad de los datos, se considera que existe mayor información (CALLEGARI, 2003), este concepto está relacionado con la entropía.

4.1.1 Definiciones

4.1.1.1 Esperanza matemática

Algunas definiciones destacadas, tomadas de (Tsitsiklis & Bertsekas, 2002)

$$E[X] := \sum_x x \rho_X(x) \quad (4.1)$$

Donde:

E Esperanza Matemática.

X Variable aleatoria.

x Valor que asume la variable aleatoria X .

$\rho_X(x)$ Probabilidad que la variable aleatoria X asuma el valor x .

Si $\rho_X(x)$ es la misma para todos los valores x , entonces la esperanza matemática es igual a la media:

$$E[X] = \frac{1}{N} \sum_{d=1}^N x_d = \mu \quad (4.2)$$

Para muestras, la media se denota con \underline{x}

4.1.1.2 Varianza

$$var(X) := E[(X - E[X])^2] \quad (4.3)$$

En los casos en que se cumple 4.2, se tiene:

$$\begin{aligned} var(X) &= \frac{1}{N} \sum_{d=1}^N (x_d - \underline{x})^2 = \frac{1}{N} \|x - \mu \mathbf{1}_v\|^2 \\ &= \frac{1}{N} (x - \mu \mathbf{1}_v)^T (x - \mu \mathbf{1}_v) \end{aligned} \quad (4.4)$$

Donde 1_v es un vector en R^N de unos:

$$1_v := [1 \dots 1]_{N \times 1} \quad (4.5)$$

Para muestras, la varianza es:

$$s^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (4.6)$$

4.1.1.3 Desviación estándar

$$\sigma_X := \sqrt{\text{var}(X)} \quad (4.7)$$

La desviación estándar muestral es:

$$s := \sigma_X := \sqrt{s^2} \quad (4.8)$$

4.1.1.4 Covarianza

$$\text{cov}(X, Y) := E [(X - E[X]) (Y - E[Y])] = E [XY] - E[X] E[Y] \quad (4.9)$$

Notar que si $E[X] = E[Y] = 0 \therefore$

$$\text{cov}(X, Y) = E [XY] = \frac{1}{N} x^T y \quad (4.10)$$

Si se tienen $X_i (i = 1, \dots, n)$ variables aleatorias, la siguiente es la matriz de covarianzas:

$$S = [E [X_1 Y_1] \cdots E [X_1 Y_n] \vdots \vdots E [X_n Y_1] \cdots E [X_n Y_n]] \quad (4.11)$$

Se observa que si $E[X_i] = 0$ de (4.9):

$$S = [E [X_1 Y_1] \cdots E [X_1 Y_n] \vdots \vdots E [X_n Y_1] \cdots E [X_n Y_n]] \frac{1}{N} X X^T \quad (4.12)$$

Siendo $X = X_1, \dots, X_n$ el vector de variables aleatorias.

4.1.1.5 Coeficiente de correlación

$$\rho(X, Y) = \frac{cov(X, Y)}{\sqrt{var(X)var(Y)}} \quad (4.13)$$

Para varias variables, la siguiente es la matriz de correlaciones:

$$R = [\rho [X_1 Y_1] \cdots \rho [X_1 Y_n] \vdots \vdots \rho [X_n Y_1] \cdots \rho [X_n Y_n]] \quad (4.14)$$

Si las variables aleatorias, por ejemplo Y_i , tienen varianzas unitarias, lo cual se logra con la transformación de estandarización (Tsitsiklis & Bertsekas, 2002)

$$Y_i = \frac{X_i - \mu_i}{\sigma_n} \quad (4.15)$$

Entonces:

$$S = R \quad (4.16)$$

4.1.2 El método PCA

El método de análisis de componentes principales o PCA por sus siglas en inglés, se presenta a continuación.

Si se tienen N observaciones de las variables X_1, \dots, X_n , la matriz de datos muestrales es:

$$X = [x_{11} \ \dots \ x_{1n} \ \vdots \ \vdots \ x_{N1} \ \dots \ x_{Nn}]_{NXn} \quad (4.17)$$

En adelante se supone que X , es una matriz centrada, esto es, los promedios de cada variable son igual a cero, si no fuera así, la transformación HX , donde:

$$H = I - \frac{1}{n} \mathbf{1}_v \mathbf{1}_v^T \quad (4.18)$$

logra el centrado deseado (Grané Chávez, 2010)

De esta manera, si $X = [X_1, \dots, X_n]^T$, es el vector de variables aleatorias con medias iguales a cero. Se desea encontrar una matriz ortogonal Q tal que $Z = Q^T X$ (Maiti, 2014) tenga una matriz de covarianzas diagonal, entonces:

$$\begin{aligned} \text{cov}(Q^T X) &= E[(Q^T X)(Q^T X)^T] \\ &= E[Q^T X X^T Q] \\ &= Q^T E[XX^T] Q \\ &= Q^T \text{cov}(X) \\ &= Q^T S Q \\ &= \Lambda \end{aligned} \tag{4.19}$$

Como S es simétrica real, con valores no negativos, entonces S es una matriz semidefinida positiva, con lo cual $Q^T = Q^{-1}$ es una matriz ortogonal y Λ es una matriz diagonal. (Strang, 2006).

Por lo tanto:

$$S = Q \Lambda Q^T \tag{4.20}$$

Es la descomposición espectral de la matriz de la matriz de covarianzas. En las columnas de $Q = [q_1, \dots, q_n]$ están los vectores

característicos de S y en la diagonal de $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ sus valores característicos.

De esta manera, las componentes principales son:

$$Z = Q^T X \quad (4.21)$$

Por tanto, para cada componente principal sería:

$$Z_i = q_i^T X \quad (4.22)$$

4.2 REDES NEURONALES ARTIFICIALES

Según (Ramírez, 2019), básicamente una neurona tiene como misión recibir información de otras neuronas o de receptores, procesa estas informaciones mediante un código de activación interno y luego transmite nuevos impulsos a las neuronas siguientes u otras células.

4.2.1 Perceptrón

Es la unidad básica en el campo de las redes neuronales artificiales, la cual recibe unos valores de entradas ponderados, los procesa y emite una salida.

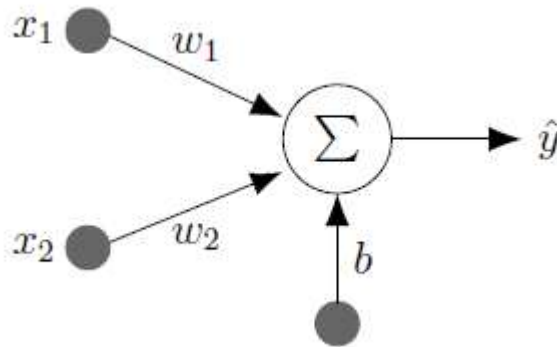


Figura 9 - Arquitectura Básica de un PERCEPTRÓN. Fuente: (Ramírez, 2019)

Conforme a la Figura 9 los valores de entrada son x_i , la salida es \hat{y} y existe un parámetro adicional b que es el umbral o bias.

$$a = w^T x + b = \sum_{i=1}^n w_i x_i + b \quad (4.23)$$

En donde: $x = [x_1 \dots x_n]$ son los valores de entrada y: $w = [w_1 \dots w_n]$ el vector de pesos.

La función:

$$g(a) = \hat{y} \quad (4.24)$$

es la denominada función de activación.

4.2.2 Perceptrón multicapa

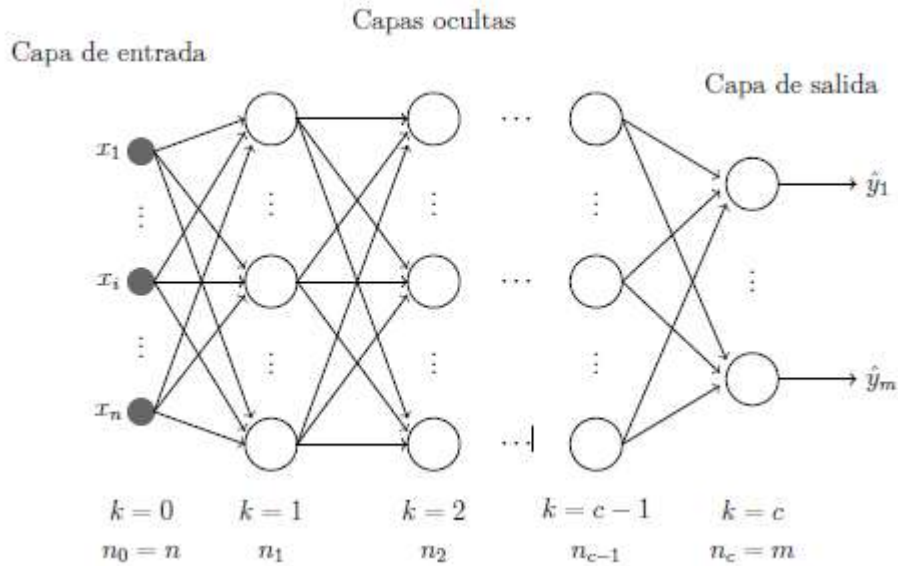


Figura 10 - Arquitectura del PERCEPTRÓN Multicapa. (Ramírez, 2019)

Analizando la Figura 10

x_i : valores de entrada $i = 1, \dots, n$;

n_k : número de neuronas, capa $k = 0 \dots, c$;

$c + 1$: número total de capas, la capa de entrada (capa 0), las ocultas (capas $k = 1 \dots, c - 1$) y la capa de salida "c", el número de capas ocultas es $c - 1$

\hat{y}_j : valores de salida $j = 1, \dots, m$.

Además, se utiliza la siguiente nomenclatura:

$w_{ij}^k, k > 0$: peso de conexión entre el nudo i y el nudo j de las capas

$k - 1$ y k respectivamente:

b_i^k : bias del nodo i , capa $k > 0$;

g : función de activación.

Formalmente, se tiene lo siguiente (McGonagle, 2019)

Conjunto de entrenamiento: Es un conjunto de pares de vectores, de entrada x_d , y salida y_d , de tamaño N denotado por:

$$X = \{(x_1, y_1), \dots, (x_N, y_N)\} \quad (4.25)$$

Red neuronal: como la Figura 10, cuyos parámetros son denotados por θ de manera colectiva. Tiene pesos w_{ij}^k de conexión entre el nudo i de la capa $k - 1$ al nudo j de la capa k , y b_j^k el bias de este último nudo. No hay conexiones entre neuronas de la misma capa, sin embargo, todas las de una capa están conectadas a todas las de la capa siguiente.

Función de error: $E(X, \theta)$, representa el error promedio entre la salida deseada y , respecto a la calculada \hat{y}^t en la iteración t . Es promedio de los errores de cada par (x_d, y_d) del conjunto de entrenamiento X para un particular conjunto de valores θ .

Los parámetros θ se actualizan en cada iteración utilizando la técnica descenso estocástico del gradiente:

$$\theta^{t+1} = \theta^t - \alpha \frac{\partial E(X, \theta)}{\partial \theta} \quad (4.26)$$

Donde θ^t es el valor del parámetro en la iteración t y α es la razón de

aprendizaje.

Según este proceso se realiza con el objetivo de minimizar el error hasta un valor aceptable. El error adoptado es el de mínimos cuadrados definido como sigue:

$$\varepsilon_d := \frac{1}{2} \|y_d - \hat{y}_d\|^2 \quad (4.27)$$

Que corresponde al par $(x_d - y_d)$. El error promedio para todo el conjunto de entrenamiento es:

$$E(X, \theta) := \frac{1}{N} \sum_{d=1}^N \varepsilon_d = \frac{1}{N} \sum_{d=1}^N \|y_d - \hat{y}_d\|^2 \quad (4.28)$$

Su derivada respecto a los pesos w_{ij}^k es:

$$\frac{\partial E(X, \theta)}{\partial w_{ij}^k} = \frac{1}{N} \sum_{d=1}^N \frac{\partial \varepsilon_d}{\partial w_{ij}^k} \quad (4.29)$$

Simplificando, si $b_j^k = w_{0j}^k$ y $o_0^{k-1} = 1$ para el nudo 0 de la capa $k - 1$, entonces la ecuación (4.29) pasa a ser:

$$a_j^k := \sum_{i=0}^N w_{ij}^k o_i^{k-1} \quad (4.30)$$

Con los cuales la salida del nudo i de la capa k , según se presentó en (4.24)

$$o_i^k = g(a_i^k) \quad (4.31)$$

Para la capa de entrada $k = 0$

$$o_i^0 = x_i \quad (4.32)$$

La derivada del error ε es:

$$\frac{\partial \varepsilon}{\partial w_{ij}^k} = \frac{\partial \varepsilon}{\partial a_j^k} \frac{\partial a_j^k}{\partial w_{ij}^k} \quad (4.33)$$

Se adopta la notación:

$$\delta_j^k := \frac{\partial \varepsilon}{\partial a_j^k} \quad (4.34)$$

Además, se puede observar que:

$$\frac{\partial a_j^k}{\partial w_{ij}^k} = \frac{\partial}{\partial w_{ij}^k} \left(\sum_{i=0}^{n_k} w_{ij}^k o_i^{k-1} \right) = o_i^{k-1} \quad (4.35)$$

Entonces la ecuación (4.33) se puede escribir:

$$\frac{\partial \varepsilon}{\partial w_{ij}^k} = \delta_j^k o_i^{k-1} \quad (4.36)$$

De esta manera, la derivada del error ε para una muestra dada respecto a w_{ij}^k , es el producto del término δ_j^k de la neurona j de la capa k multiplicada por la salida o_i^{k-1} de la neurona i de la capa interior $k - 1$, lo cual tiene sentido considerando que w_{ij}^k es el peso de conexión entre la neurona i (de salida) y de la neurona j (de entrada) de capas $k - 1$ y k respectivamente.

Capa de salida

En la capa de salida ($k = c$) de acuerdo con la ecuación (4.27):

$$\varepsilon = \frac{1}{2} \sum_{i=1}^m (y_i - \hat{y}_i)^2 = \frac{1}{2} \sum_{i=1}^m (y_i - g(a_i^c))^2 \quad (4.37)$$

Con base en esto, se puede aplicar la ecuación (4.34)

$$\delta_j^c := \frac{\partial \varepsilon}{\partial a_j^c} = - \sum_{i=1}^m (y_i - g(a_i^c)) g'(a_i^c) \frac{\partial a_i^c}{\partial a_j^c} \quad (4.38)$$

Como

$$\frac{\partial a_i^c}{\partial a_j^c} = \{1 \text{ si } i = j; 0 \text{ si } i \neq j\} \quad (4.39)$$

$$\delta_j^c = - g'(a_i^c) (y_i - g(a_i^c)) \quad (4.40)$$

Por lo tanto, de 4.36 y 4.40 obtenemos la siguiente derivada de error

ε :

$$\frac{\partial \varepsilon}{\partial w_{ij}^c} = - g'(a_i^c) o_i^{c-1} (y_i - g(a_i^c)) \quad (4.41)$$

Capas ocultas

Conforme a (4.36), para $1 \leq k < c$ es posible escribir:

$$\delta_j^k := \frac{\partial \varepsilon}{\partial a_j^k} = \sum_{i=1}^{n_{k+1}} \frac{\partial \varepsilon}{\partial a_i^{k+1}} \frac{\partial a_i^{k+1}}{\partial a_j^k} = \sum_{i=1}^{n_{k+1}} \delta_j^{k+1} \frac{\partial a_i^{k+1}}{\partial a_j^k} \quad (4.42)$$

Observando la definición dada en (4.30), se obtiene la derivada siguiente:

$$\frac{\partial a_i^{k+1}}{\partial a_j^k} = w_{ji}^{k+1} g'(a_j^k) \quad (4.42)$$

con la cual se obtienen la fórmula que da el nombre de retro propagación a este método:

$$\delta_j^k = g'(a_j^k) \sum_{i=1}^{n_{k+1}} w_{ji}^{k+1} \delta_i^{k+1} \quad (4.42)$$

ya que el valor δ_j^k de la cara k depende de los valores δ_i^{k+1} de la capa siguiente.

Finalmente, la derivada del error ε , de acuerdo con (4.36) es:

$$\frac{\partial \varepsilon_d}{\partial w_{ij}^k} = g'(a_j^k) o_i^{k-1} \sum_{i=1}^{n_{k+1}} w_{ji}^{k+1} \delta_i^{k+1} \quad (4.43)$$

CAPÍTULO 5

ANÁLISIS DE DATOS

En este capítulo se explica el proceso del análisis de datos. Comenzando por una visión general de la base de datos, el análisis de correlación de los instrumentos para tener una visión preliminar de la relación entre los diferentes instrumentos que monitorean el bloque E-06.

Debido a la cantidad de variables, es conveniente seleccionar los instrumentos más adecuados para relacionar y realizar el análisis con las Redes Neuronales Artificiales. Para lograr esto se aplican las combinaciones mencionadas en el CAPÍTULO 3.

Posteriormente se describen las características de las redes neuronales desarrolladas para el presente trabajo y se realiza el análisis relacionando las variables propuestas. Por cada relación entre variables, se tendrá como resultado datos de predicción y un puntaje o “score” de la red neuronal con relación a los resultados obtenidos.

5.1 VISIÓN GENERAL DE LA BASE DE DATOS

Matriz de datos de entrada:

$$X = \begin{bmatrix} X1^{(1)} & X2^{(1)} & X3^{(1)} & X4^{(1)} & EX5^{(1)} & X6^{(1)} & X7^{(1)} & X8^{(1)} & X9^{(1)} & X10^{(1)} & X11^{(1)} & X12^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X1^{(N)} & X2^{(N)} & X3^{(N)} & X4^{(N)} & X5^{(N)} & X6^{(N)} & X7^{(N)} & X8^{(N)} & X9^{(N)} & X10^{(N)} & X11^{(N)} & X12^{(N)} \end{bmatrix}$$

A continuación, se presentan los gráficos de comportamiento de los instrumentos seleccionados desde su instalación.

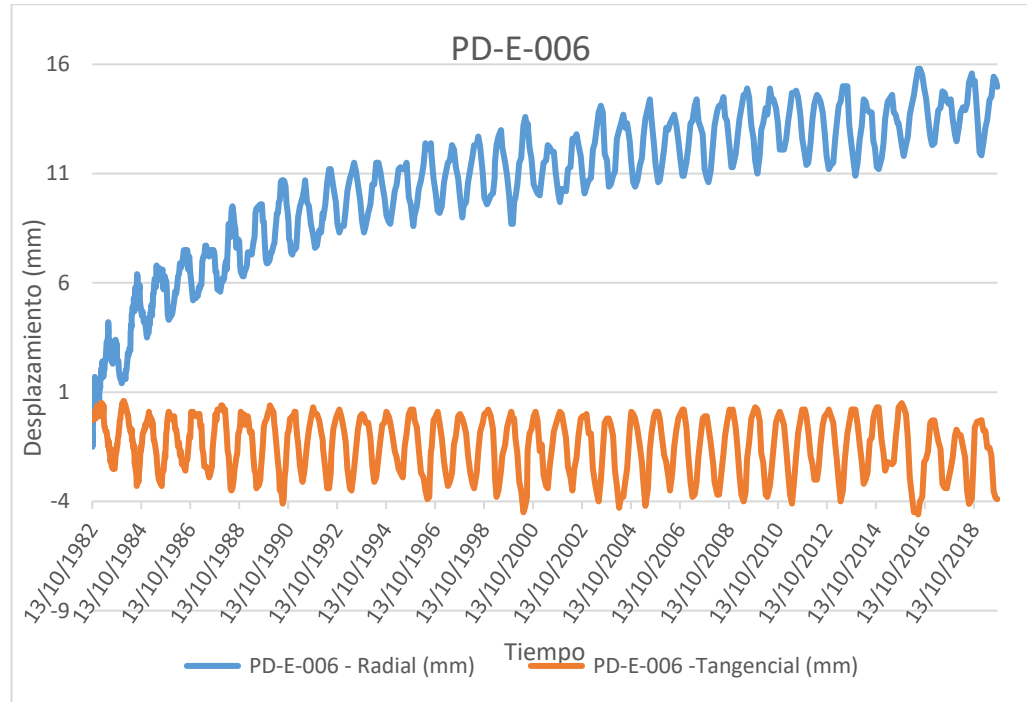


Figura 11 - Gráfico de comportamiento histórico del PD-E-006.

Fuente: Elaboración propia

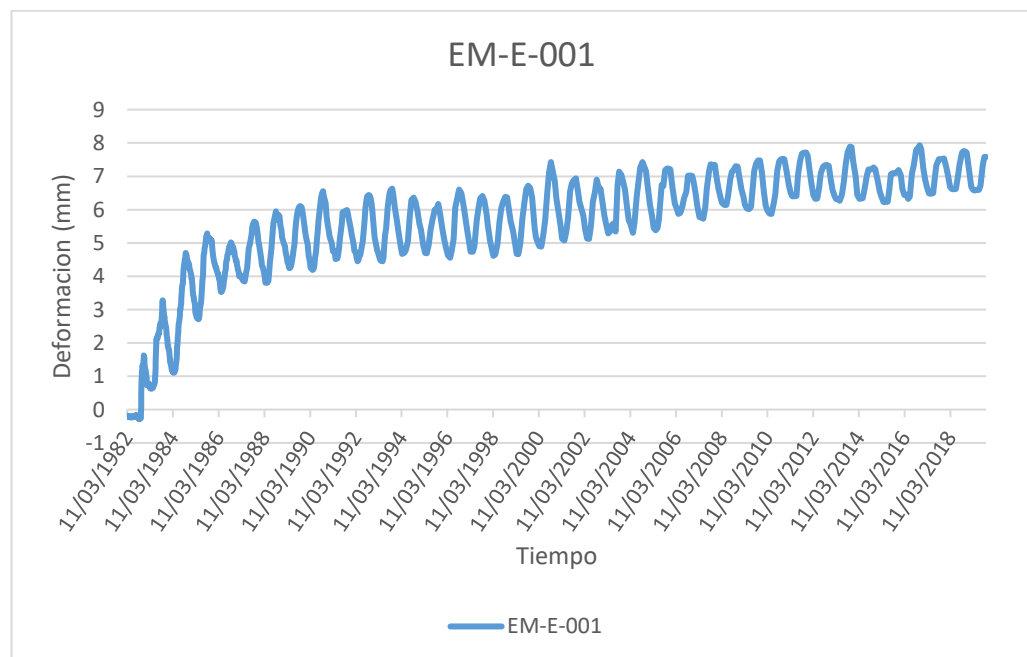


Figura 12 - Gráfico de comportamiento histórico EM-E-001

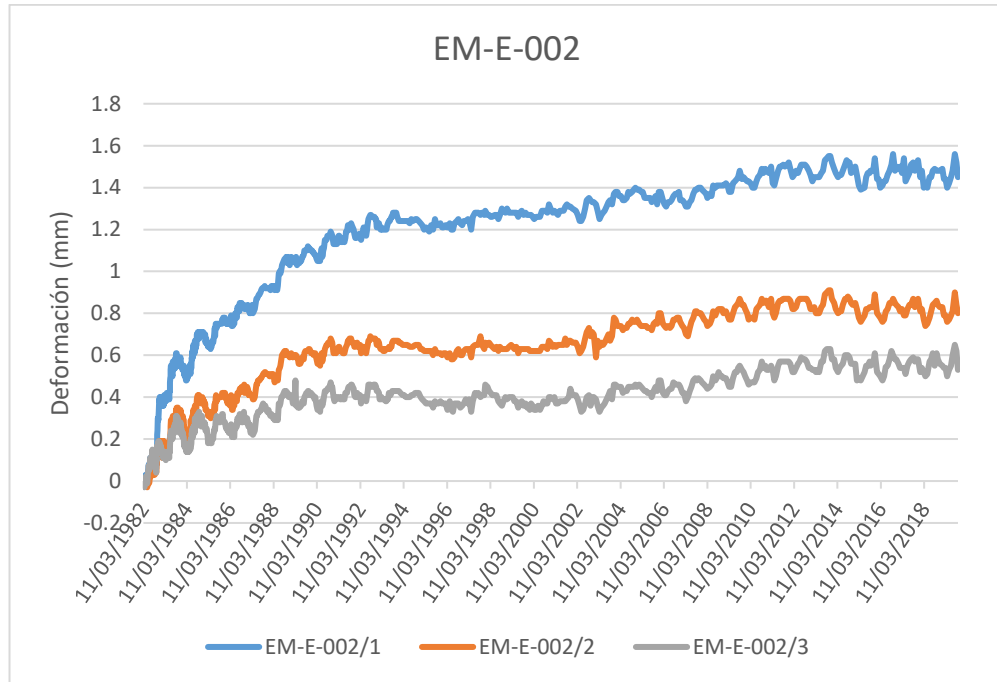


Figura 13 - Gráfico de comportamiento histórico EM-E-002

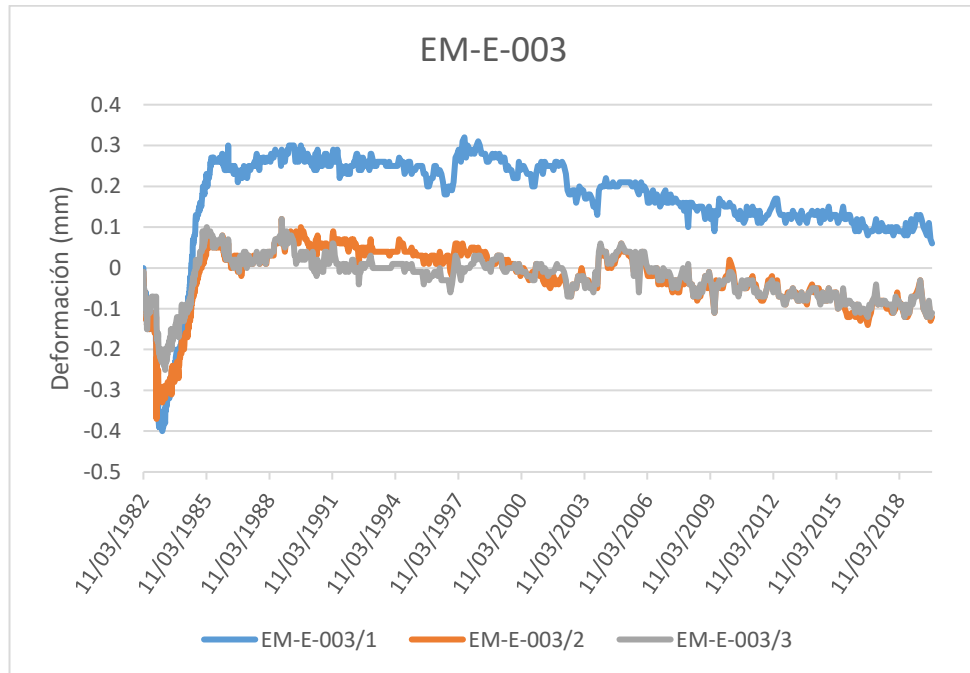


Figura 14 - Gráfico de comportamiento histórico. EM-E-003

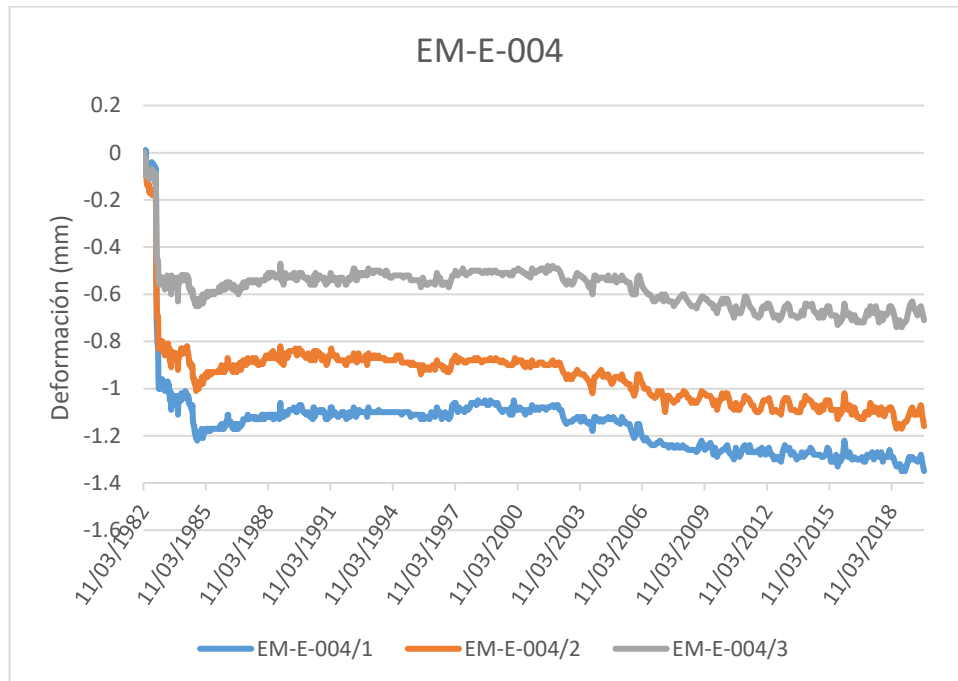


Figura 15 - Gráfico de comportamiento histórico. EM-E-004

Como podemos ver en los diferentes instrumentos, durante la época de llenado del embalse 1982, los instrumentos detectaron deformaciones en el caso de los extensómetros y desplazamientos en el caso del péndulo.

5.2 ANÁLISIS DE COMPONENTES PRINCIPALES

Para aplicar el análisis de componentes principales sigue la siguiente secuencia de pasos.

5.2.1 Estandarización de variables

Se obtiene restando a cada valor la media correspondiente y dividiendo éste entre su desviación estándar, para lograr, a partir de X la matriz con datos estandarizados X' , cada columna de X' tiene media igual a cero y desviación estándar igual a uno.

Tabla 3 - Datos estandarizados

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
0	1.519547	-2.067214	1.371418	-1.089306	1.105664	1.271922	-0.318103	-0.688807	-0.995945	-1.664541	-2.093585	-1.916214
1	1.595222	-2.010481	1.371418	1.275725	1.289026	1.964532	-0.265815	-0.779935	-1.125156	-1.437456	-1.734895	-1.506636
2	1.631420	-1.787409	1.244610	1.382251	1.564069	2.195402	-0.056659	-0.415422	-0.608313	-1.134677	-1.286533	-1.097058
3	1.414135	-0.488216	0.940269	1.169200	1.151505	1.810618	-0.213526	-0.688807	-1.125156	-1.286067	-1.465878	-1.233584
4	1.377960	-0.201871	0.869257	1.089306	1.059824	1.425835	-0.161237	-0.597679	-0.995945	-1.361761	-1.645223	-1.643162
...
665	-2.364660	0.773469	-2.311100	-2.426033	-2.194847	-1.806346	-1.677613	-1.144448	-1.125156	5.677859	4.721525	4.363983
666	-2.341078	0.615877	-2.427764	-2.479296	-2.240687	-1.883302	-1.520746	-0.962192	-0.866735	6.434807	5.438905	5.046613
667	-2.199585	0.615877	-2.539355	-2.559190	-2.286527	-2.037216	-1.573035	-0.688807	-0.866735	7.873009	6.425303	6.138821
668	-2.058093	0.852265	-2.605296	-2.532558	-2.286527	-1.960259	-1.573035	-0.688807	-0.479102	8.024398	6.694320	6.411873
669	-2.010929	1.009858	-2.605296	-2.612452	-2.378208	-2.191129	-1.363880	-0.688807	-0.479102	8.024398	6.783993	6.548399

670 rows x 12 columns

5.2.2 Matriz de correlaciones

Obtenemos la matriz de correlaciones conforme a la ecuación (4.14)

Con los valores estandarizados, la matriz de covarianzas es igual a la matriz de correlaciones (3.16).

La interpretación de los valores de la matriz R depende de los objetivos de la utilización por la cual es calculado. Según Callegari-Jacques (2003) el coeficiente de correlación puede ser evaluado cualitativamente de acuerdo con el siguiente cuadro.

Tabla 4 - Evaluación sistematizada del coeficiente de correlación lineal.

Coeficiente de correlación (ρ)	Intensidad de correlación
$0,00 < \rho < 0,30$	Débil correlación lineal
$0,30 \leq \rho < 0,60$	Moderada correlación lineal
$0,60 \leq \rho < 0,90$	Fuerte correlación lineal
$0,90 \leq \rho < 1,00$	Correlación lineal muy fuerte.

Fuente: (Callegari, 2003)

Tabla 5 - Matriz de correlaciones "R"

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1.000000	-0.467789	0.938762	0.967369	0.957769	0.916424	0.588423	0.451478	0.286298	-0.664636	-0.702833	-0.480741
X2	-0.467789	1.000000	-0.525927	-0.321669	-0.344665	-0.395982	-0.266376	-0.213017	-0.169620	0.278034	0.298186	0.247493
X3	0.938762	-0.525927	1.000000	0.941357	0.947374	0.926851	0.684879	0.568189	0.434569	-0.693646	-0.694496	-0.481029
X4	0.967369	-0.321669	0.941357	1.000000	0.989545	0.944068	0.623467	0.494909	0.327884	-0.660561	-0.676191	-0.437810
X5	0.957769	-0.344665	0.947374	0.989545	1.000000	0.972529	0.592203	0.468680	0.308918	-0.674181	-0.696158	-0.479153
X6	0.916424	-0.395982	0.926851	0.944068	0.972529	1.000000	0.519342	0.397386	0.242355	-0.671093	-0.705602	-0.532352
X7	0.588423	-0.266376	0.684879	0.623467	0.592203	0.519342	1.000000	0.966843	0.907434	-0.419296	-0.275196	-0.066853
X8	0.451478	-0.213017	0.568189	0.494909	0.468680	0.397386	0.966843	1.000000	0.952667	-0.305008	-0.143235	0.040462
X9	0.286298	-0.169620	0.434569	0.327884	0.308918	0.242355	0.907434	0.952667	1.000000	-0.242544	-0.076131	0.083860
X10	-0.664636	0.278034	-0.693646	-0.660561	-0.674181	-0.671093	-0.419296	-0.305008	-0.242544	1.000000	0.959526	0.892453
X11	-0.702833	0.298186	-0.694496	-0.676191	-0.696158	-0.705602	-0.275196	-0.143235	-0.076131	0.959526	1.000000	0.934370
X12	-0.480741	0.247493	-0.481029	-0.437810	-0.479153	-0.532352	-0.066853	0.040462	0.083860	0.892453	0.934370	1.000000

A partir de la matriz de correlaciones R, podemos hacer algunas observaciones.

- La variable X1 correspondiente al movimiento Radial del Péndulo, tiene una fuerte correlación lineal con las variables X10, X11, correspondiente a los instrumentos EM-E-004/1 y EM-E-004/2 una muy fuerte correlación con las variables X3, X4, X5 y X6 correspondientes a los extensómetros EM-E-001, EM-E-002/, EM-E-002/2 y EM-E-002/3.
- La variable X2 correspondiente al movimiento tangencial del péndulo PD posee una correlación moderada a débil con todas las variables.
- X8 y X9 correspondientes a los extensómetros EM-E-003/2 y EM-E-003/3 poseen correlación moderada o débil con los demás instrumentos distintos a los de su misma base.
- Las variables correspondientes a extensómetros de una misma base de extensómetros poseen una correlación fuerte o muy fuerte.
- Un aspecto interesante es la concordancia de los coeficientes de correlación entre las variables, ya que los altos valores de correlación, en su mayoría responden a la lógica del movimiento del bloque estudiado.

En las siguientes figuras podemos observar algunos ejemplos de gráficos de dispersión entre algunas variables.

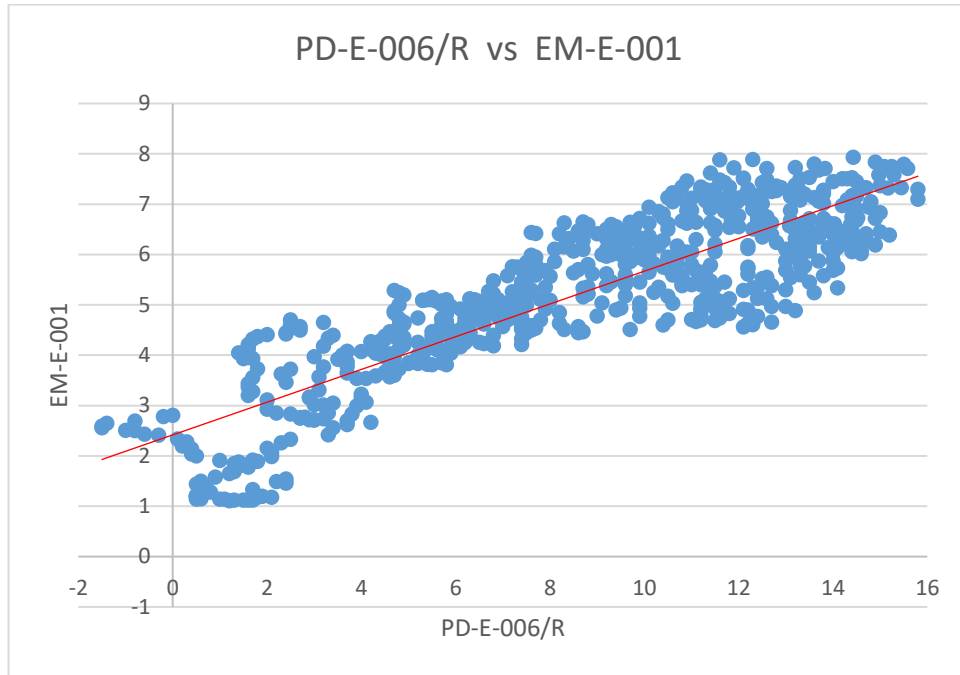


Figura 16 - Gráfico de Dispersión PD-E-006/R vs EM-E-001

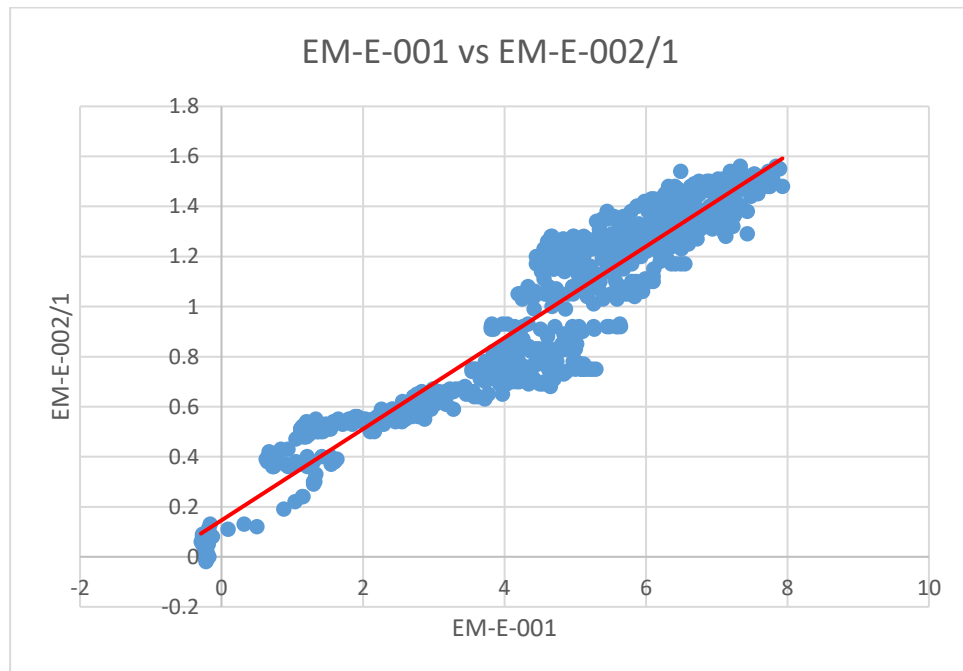


Figura 17 - Gráfico de dispersión PD-E-006/R vs EM-E-002/

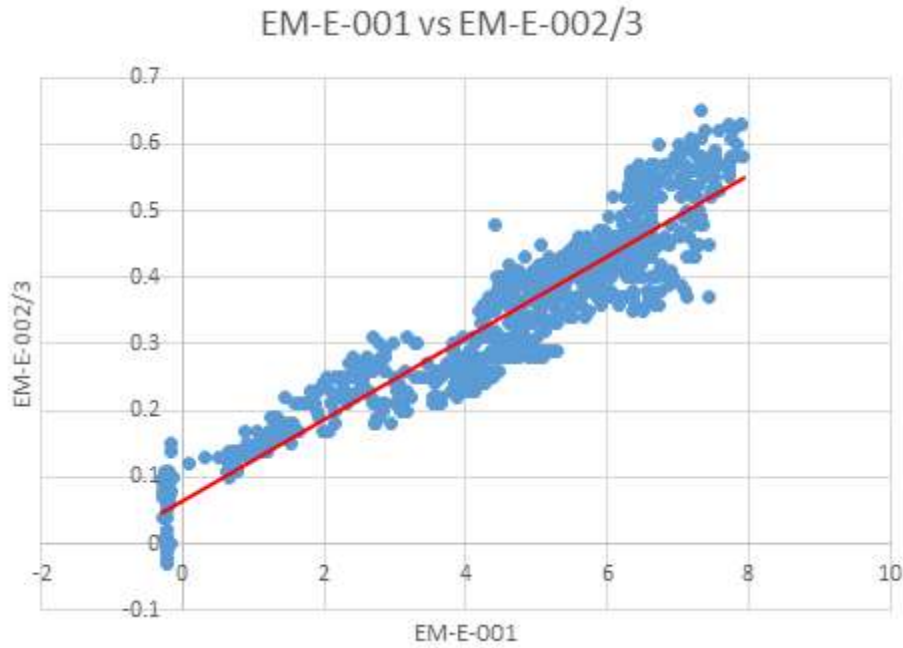


Figura 18 - Gráfico de dispersión EM-E-001 vs EM-E-002/3

A continuación, aplicaremos el análisis de componentes principales y seleccionaremos los instrumentos con mayor influencia en el sistema.

5.2.3 Componentes Principales

Tabla 6 - Porcentaje de variabilidad explicado de cada componente principal

Componente	Auto valor	%Explicado	%Acumulado
Z1	7,4036	61,60%	61,60%
Z2	2,5043	20,84%	82,44%
Z3	1,0350	8,61%	91,05%

Z4	0,8269	6,88%	97,93%
Z5	0,0992	0,83%	98,76%
Z6	0,0457	0,38%	99,14%
Z7	0,0377	0,31%	99,45%
Z8	0,0387	0,24%	99,69%
Z9	0,0183	0,15%	99,84%
Z10	0,0089	0,07%	99,91%
Z11	0,0070	0,06%	99,97%
Z12	0,0033	0,03%	100,00%

Con las componentes Z1, Z2 y Z3, el porcentaje de variabilidad explicada es del 91,05%.

En la siguiente tabla podemos observar los autovalores de las componentes Principales.

Estos valores corresponden a la correlación entre los valores reales y las Componentes principales.

Tabla 7 - Autovalores de las componentes principales correspondientes a cada variable.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
Z1	-0.34	0,16	-0.35	-0.34	-0.34	-0.33	-0,26	-0.21	-0.17	0.29	0.29	0.22
Z2	0.04	-0.01	-0.02	0.01	0.03	0.08	-0.42	-0.48	-0.50	-0.22	-0.32	-0,41
Z3	0.26	-0.33	0.17	0.22	0.22	0.24	-0.15	-0.19	-0.21	0.43	0.30	0.42

Podemos observar que no existen valores altos de intensidad de correlación de las variables con las componentes principales. pero si podemos identificar que en las 3 primeras componentes principales existen algunas variables que poseen una correlación con intensidad moderada.

Teniendo en cuenta que estos valores representan cuanta información aportan las variables al sistema.

Este será nuestro punto de partida para poder agrupar las variables en el análisis con las Redes Neuronales.

- Componente 1. Posee correlación moderada con las variables X1, X3, X4, X5 y X6.
- Componente 2. Posee correlación moderada con las variables X7, X8, X9, X11 y X12.
- Componente 3. Posee correlación moderada con las variables X2, X9, X10, X11 y X12.

5.2 DEFINICIÓN DE COMBINACIONES DE VARIABLES.

Seleccionamos de manera aleatoria 5 variables como ejemplos para ser relacionados con los instrumentos según el criterio de cada combinación.

Las variables seleccionadas son: X1, X3, X6, X8, X10.

Combinación con todas las variables

Cada variable será relacionada con todas las demás sin tener en cuenta algún criterio estadístico o técnico.

Combinación utilizando PCA

Tabla 8 - Combinaciones utilizando PCA.

Variable referencial	Variabes a relacionar
X1	X3, X4, X5, X6
X3	X1, X4, X5, X6
X6	X1, X3, X4, X5,
X8	X7, X9, X11, X12
X10	X2, X9, X11, X12

Combinación utilizando justificativas técnicas.

Para este criterio de combinación es necesario comprender el comportamiento del bloque y como está compuesta la fundación de este.

Sabiendo que los instrumentos están instalados a distintas cotas y monitoreando diferentes zonas. Inicialmente, tendremos en cuenta el movimiento hacia aguas arriba y aguas abajo que realiza el bloque, cuando ocurre esto se comprimen y descomprimen los extremos. Por tanto, utilizaremos la variable X1 correspondiente a la componente radial del péndulo, junto con la variable X3 del extensómetro que monitorea la fundación hacia aguas arriba, y las variables X5 y X11 pertenecientes a los extensómetros que monitorean la fundación en la zona central del bloque, pero cuyos perfiles de instalación tiene dirección hacia aguas arriba y aguas abajo. Además, están instalados posteriores a una discontinuidad de rocas con mayor deformabilidad que presentará mayores deformaciones ante el movimiento del bloque.

Teniendo en cuenta que la deformación en las rocas de fundación no ocurre instantáneamente ante cada movimiento del bloque, sino que ocurre lentamente, podemos decir que existe un tiempo de respuesta entre el movimiento del bloque y la deformación de la fundación. Entonces para esta combinación utilizaremos únicamente extensómetros y los relacionamos con sus pares instalados en la misma discontinuidad.

Tabla 9 - Combinaciones técnicas.

Variable referencial	Variables a relacionar
X1	X3, X5, X11
X3	X1, X5, X11
X6	X3, X4, X5, X9, X12
X8	X5, X7, X9, X11
X10	X4, X7, X11, X12

5.3 REDES NEURONALES ARTIFICIALES

Se utiliza el método sklearn. neural_network. MLPRegressor (), se trata de un regresor de perceptrón multicapa que minimiza el error cuadrático usando el descenso estocástico del gradiente entre otros métodos.

Se optó por los siguientes parámetros con los cuales se obtuvieron buenos resultados utilizando como parámetro el coeficiente de determinación R^2 .

- Cantidad de capas ocultas: 2
- Número de neuronas en la capa oculta=225
- Función de activación: relu
- Método de optimización: 'lbfgs' algoritmo de Broyden-Fletcher-Gold

farb.

- Razón de aprendizaje $\alpha = 0.001$ (ec. 4.26)

```
En [86]: X_train, X_test, y_train, y_test=train_test_split(X,y, test_size=0.2)

En [102]: mlr= MLPRegressor (solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(2,225), random_state=20,activation = "relu",max_iter=5000)
mlr.fit(X_train,y_train)
```

Figura 19 - Desarrollo de códigos en Jupyter Notebook.

Tabla 10 - Coeficientes R2 de todas las combinaciones.

Combinación	Variable referencial	Variables relacionadas	Coef. de Determinación R ²
Total-1	X1	Todas las demás variables	0.9860
Total-2	X3		0.9809
Total-3	X6		0.9814
Total-4	X8		0.9476
Total-5	X10		0.9810
PCA-1	X1	X3, X4, X5, X6	0.9606
PCA-2	X3	X1, X4, X5, X6	0.9601

PCA-3	X6	X1, X3, X4, X5,	0.9547
PCA-4	X8	X7, X9, X11, X12	0.9712
PCA-5	X10	X2, X9, X11, X12	0,9840
Técnico-1	X1	X3, X5, X11	0.9604
Técnico-2	X3	X1, X5, X11	0.9607
Técnico-3	X6	X3, X4, X5, X9, X12	0,9691
Técnico-4	X8	X5, X7, X9, X11	0.9715
Técnico-5	X10	X4, X7, X11, X12	0.9836

Tabla 11 - Coeficientes R2 de todas las combinaciones ordenadas por variable

Variable	Combinación	Coeficiente de Determinación R ²
X1	Total	0.9860
	PCA	0.9522
	Técnico	0.9604
X3	Total	0.9809
	PCA	0.9601
	Técnico	0.9607
X6	Total	0.9814
	PCA	0.9547
	Técnico	0,9691

X8	Total	0.9476
	PCA	0.9712
	Técnico	0.9715
X10	Total	0.9810
	PCA	0,9840
	Técnico	0.9836

A continuación, se presentan ejemplos utilizando las variables X1, X3, X6 de predicción de datos utilizando las combinaciones mencionadas anteriormente. Donde los datos de las columnas con título “Pred.” corresponden a los datos predichos y las columnas de título “Real” son los datos extraídos de la base de datos.

X1 - Total 1	X1 - PCA 1	X1 - Emp. 1
Pred. Real	Pred. Real	Pred. Real
13.23 12.64	9.80 9.78	6.44 6.27
4.08 4.18	9.54 9.21	14.23 14.54
9.77 9.98	11.97 13.68	-0.08 0.28
11.73 11.25	7.04 6.08	11.01 12.35
10.91 11.21	3.93 4.56	9.69 9.59
...
4.72 4.27	9.93 9.40	3.36 2.18
14.40 14.53	6.76 8.84	3.46 2.94
7.89 7.22	7.44 7.22	10.58 9.98
10.58 10.64	12.61 12.80	8.10 10.16
5.49 5.13	7.76 7.32	8.10 9.02

Figura 20 - Predicción de X1 con las combinaciones aplicadas.

X3 - Total 2		X3 - PCA 2		X3 - Emp. 3	
Pred. Real		Pred. Real		Pred. Real	
3.97	3.84	2.58	2.67	4.60	4.18
4.01	4.05	6.76	6.59	6.15	5.83
5.95	6.16	4.50	4.49	6.97	6.95
0.97	1.15	1.85	1.89	3.77	4.00
4.63	4.46	3.42	3.87	1.86	2.16
...
5.00	5.52	1.11	1.09	1.05	1.15
6.70	6.70	4.95	4.86	6.69	6.15
5.36	5.57	4.19	4.21	5.21	5.11
0.87	0.69	-0.06	0.09	5.19	5.57
5.29	5.08	4.44	4.20	5.35	5.98

Figura 21 - Predicción de X3 con las combinaciones aplicadas.

X6 - Total 3		X6 - PCA 3		X6 - Emp 3	
Pred. Real		Pred. Real		Pred. Real	
0.27	0.24	0.44	0.44	0.55	0.55
0.41	0.39	0.32	0.29	0.28	0.22
0.39	0.40	0.36	0.41	0.53	0.49
0.47	0.51	0.46	0.47	0.40	0.40
0.44	0.37	0.38	0.35	0.29	0.33
...
0.27	0.28	0.36	0.37	0.52	0.51
0.51	0.50	0.53	0.53	0.20	0.21
0.26	0.28	0.46	0.46	0.34	0.36
0.50	0.52	0.28	0.27	0.40	0.43
0.29	0.28	0.55	0.54	0.40	0.35

Figura 22 - Predicción de X6 con las combinaciones aplicadas.

A continuación, se presentan los gráficos de los Datos predichos vs Datos reales de todas las combinaciones realizadas.

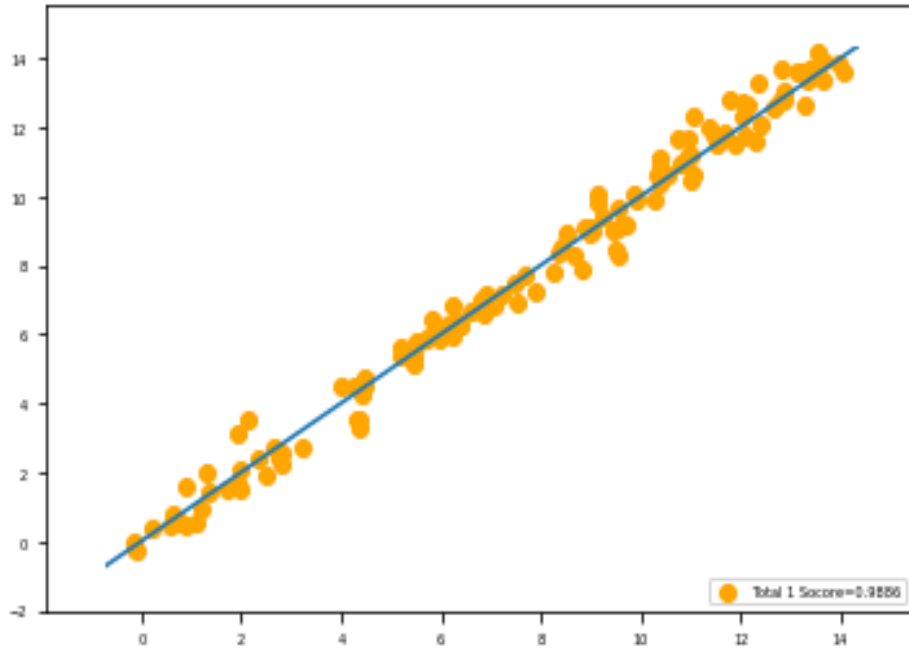


Figura 23 - Datos Predichos X1 vs Datos Reales X1- Total

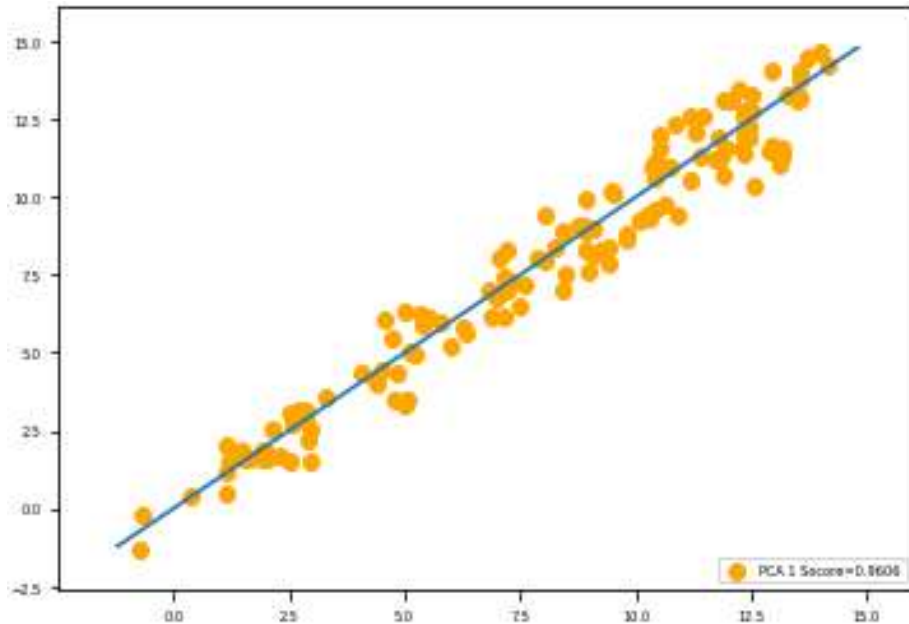


Figura 24 - Datos Predichos X1 vs Datos Reales X1 - PCA 1

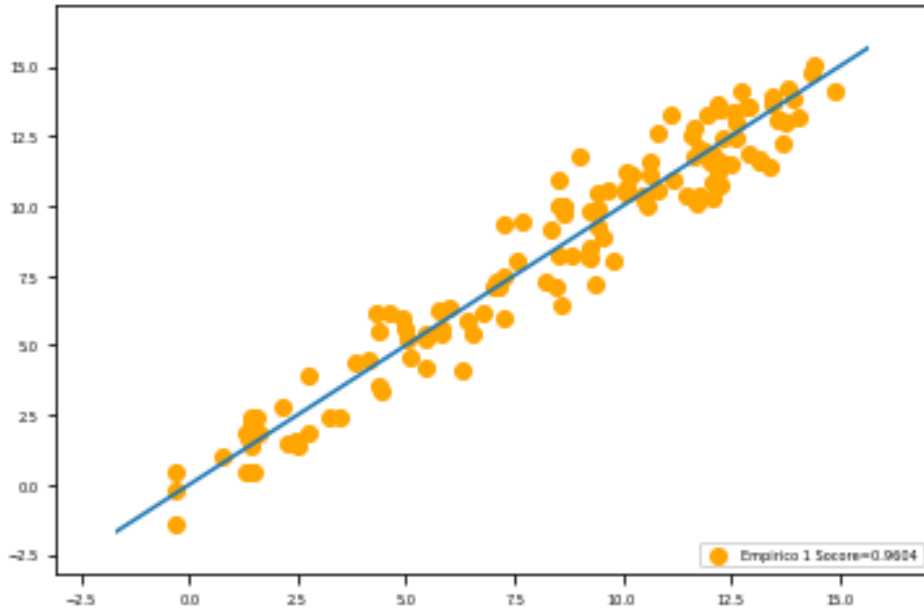


Figura 25 - Datos Predichos X1 vs Datos Reales X1 -Técnico1

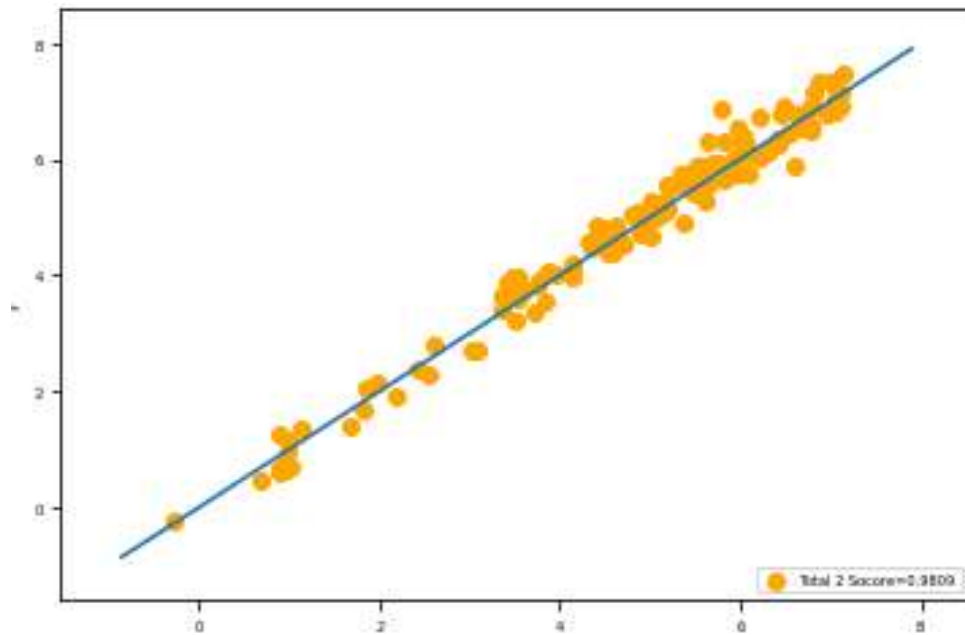


Figura 26 - Datos Predichos X3 vs Datos Reales X3 - Total 2.

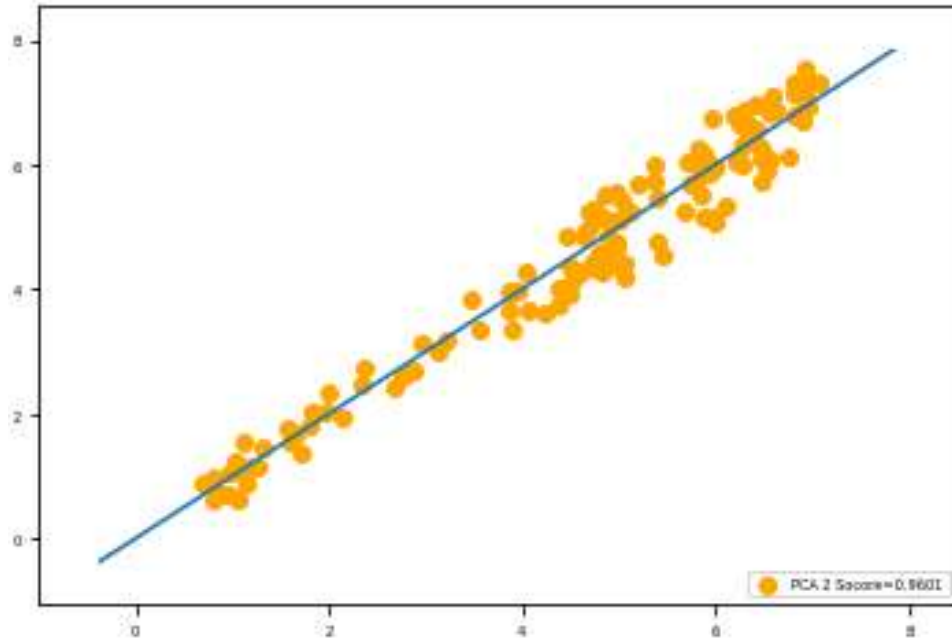


Figura 27 - Datos Predichos X3 vs Datos Reales X3 - PCA 2

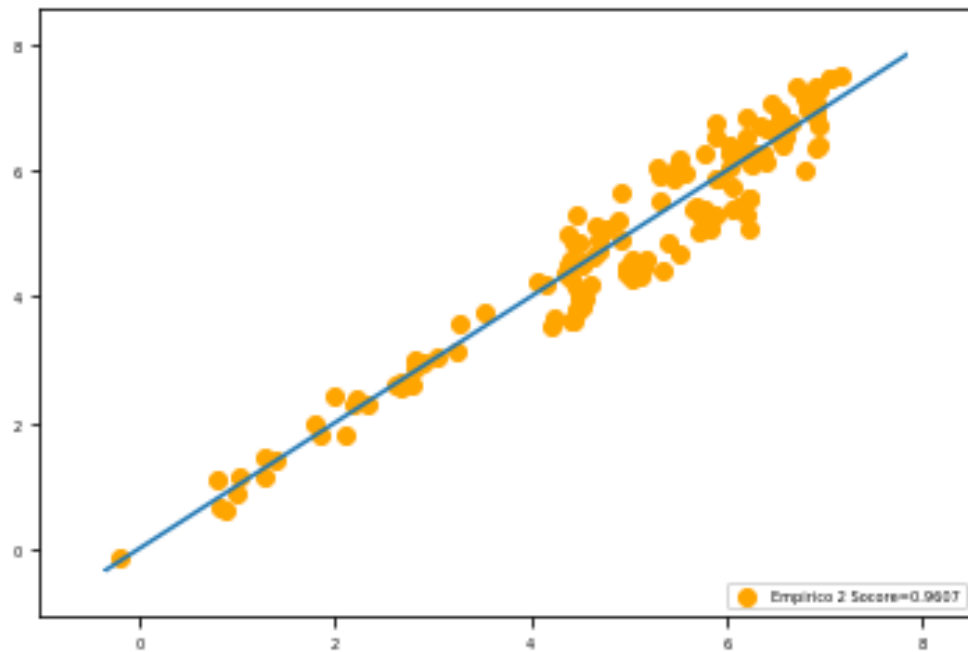


Figura 28 - Datos Predichos X3 vs Datos Reales X3 - Técnico 2

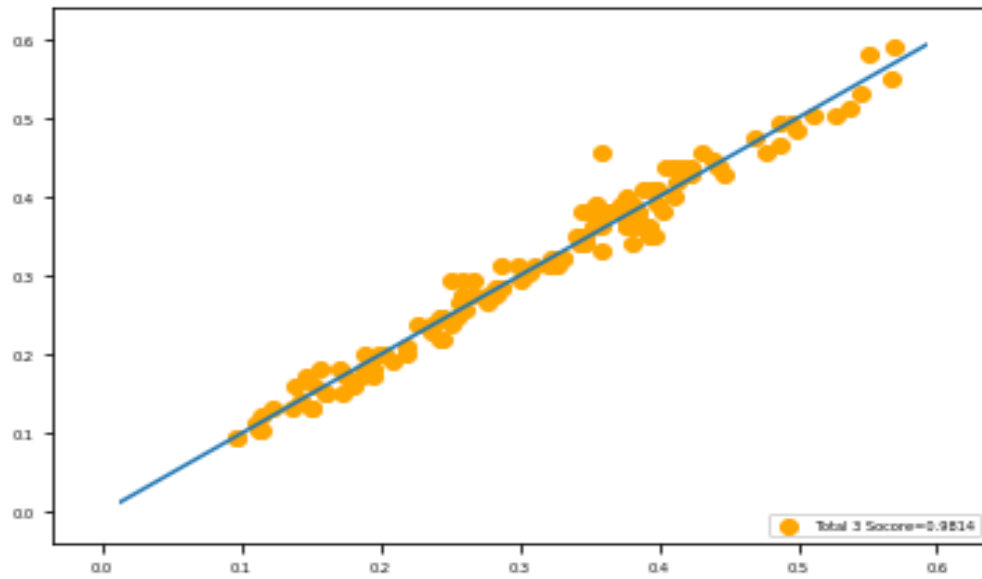


Figura 29 - Datos Predichos X6 vs Datos Reales X6 - Total 3

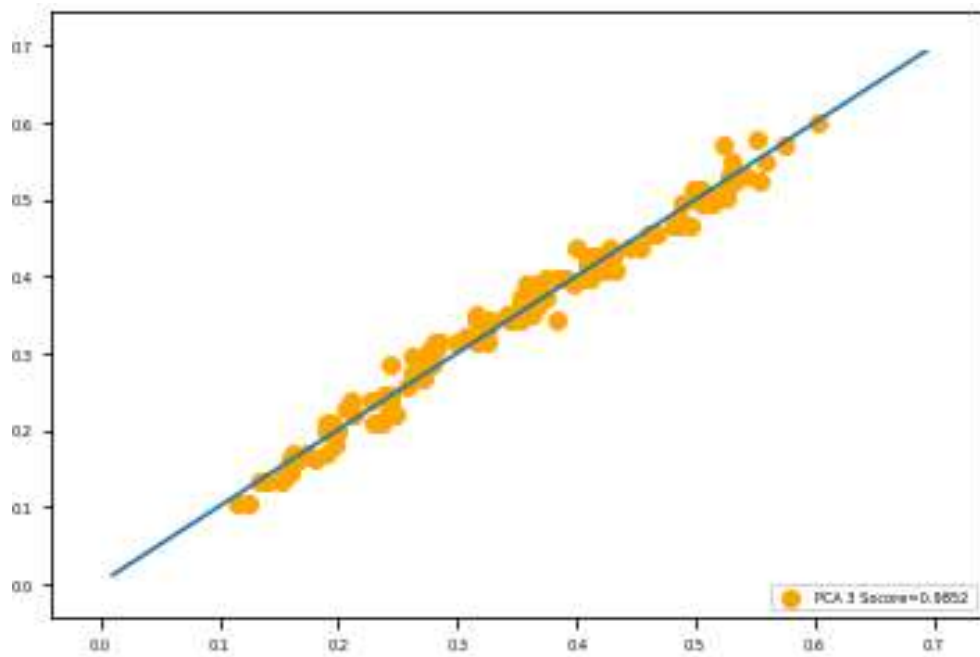


Figura 30 - Datos Predichos X6 vs Datos Reales X6 - PCA 3

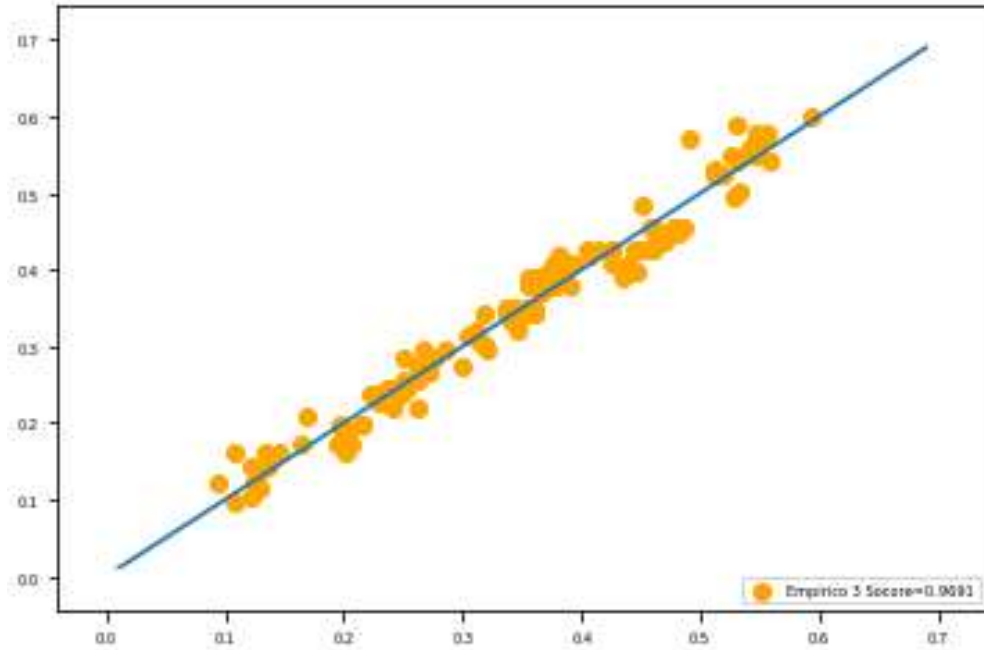


Figura 31 - Datos Predichos X6 vs Datos Reales X6 - Técnico 3

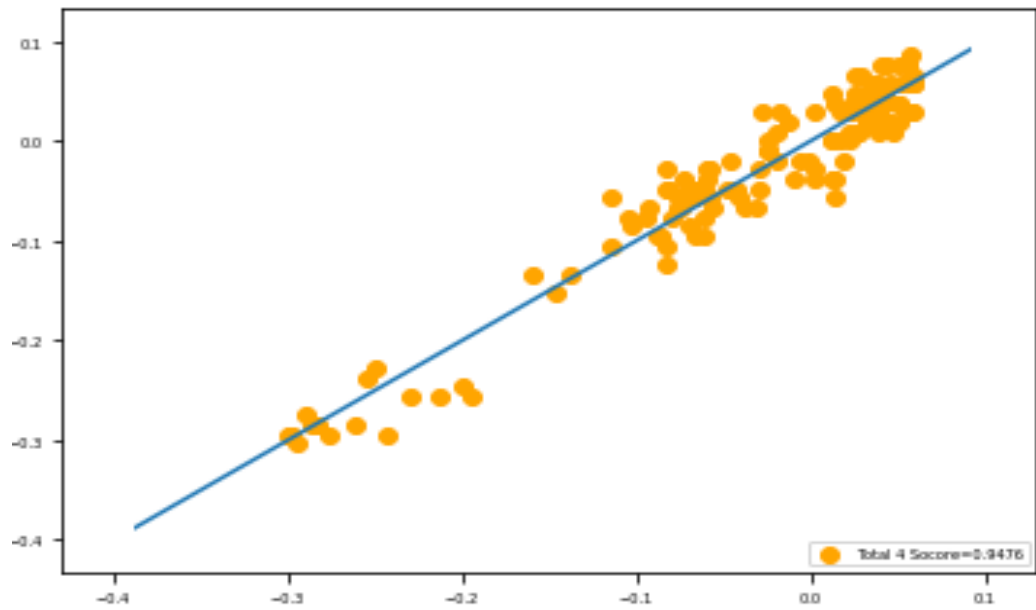


Figura 32 - Datos Predichos X8 vs Datos Reales X8 - Total 4

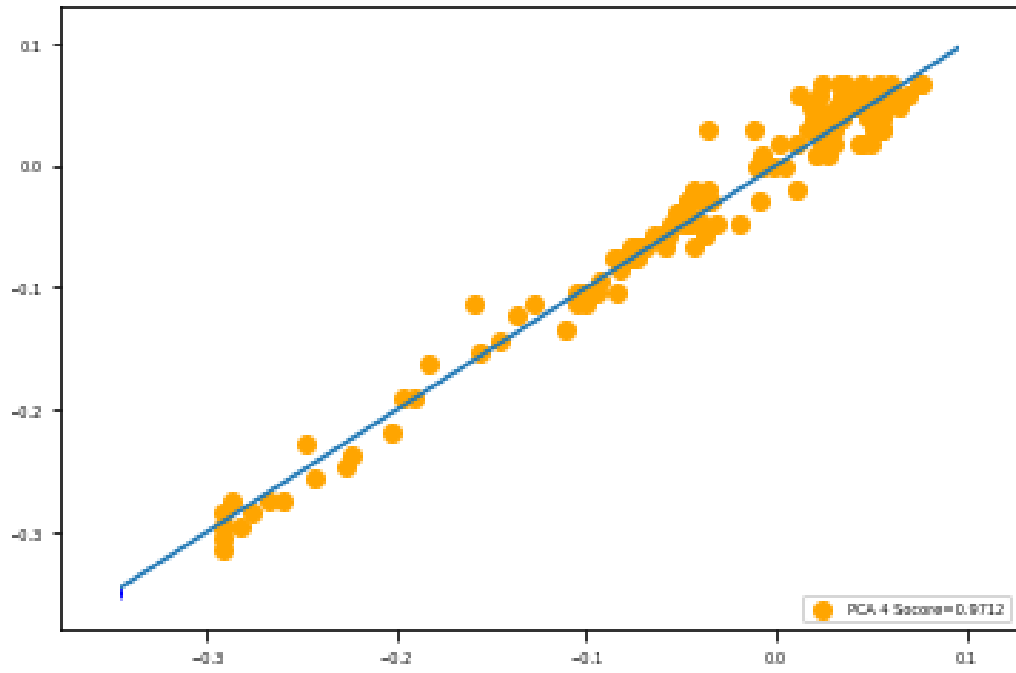


Figura 33 - Datos Predichos X8 vs Datos Reales X8 - PCA 4

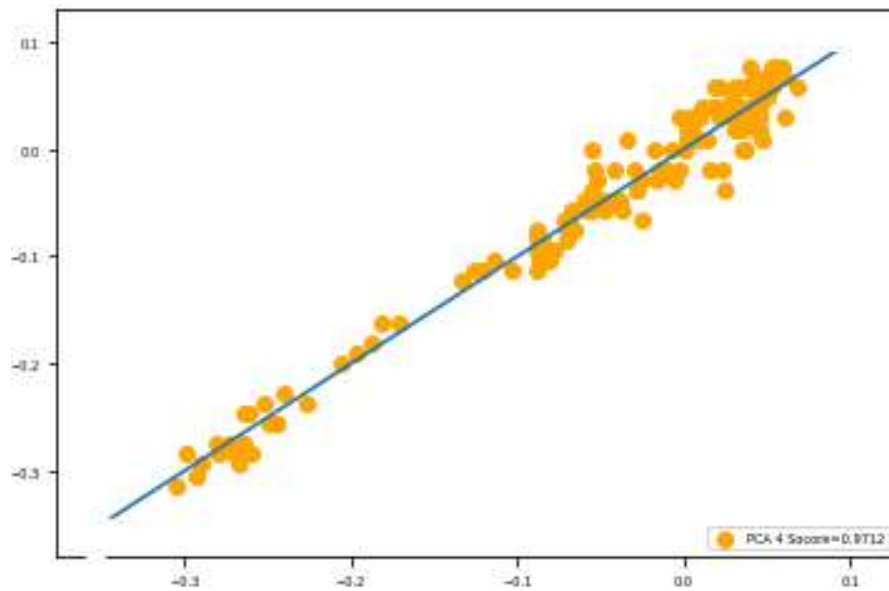


Figura 34 - Datos Predichos X8 vs Datos Reales X8- Técnico 4

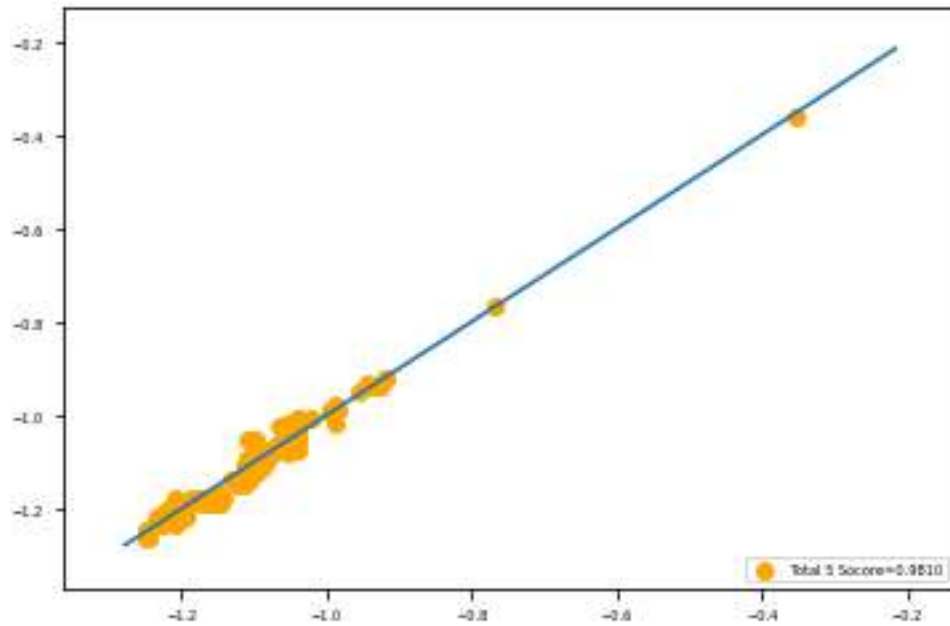


Figura 35 - Datos Predichos X10 vs Datos Reales X10 - Total 5

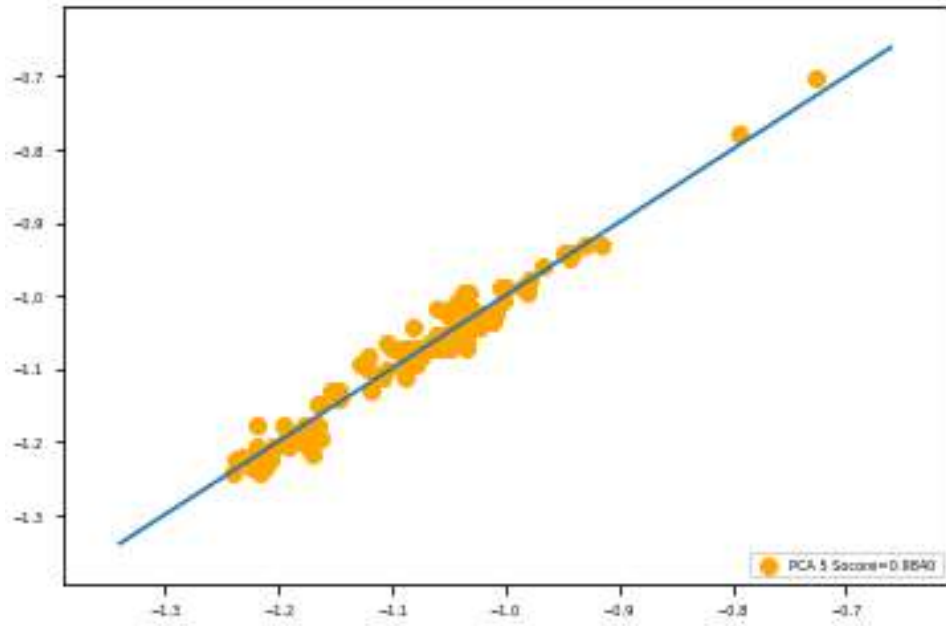


Figura 36 - Datos Predichos X10 vs Datos Reales X10 - Total 5

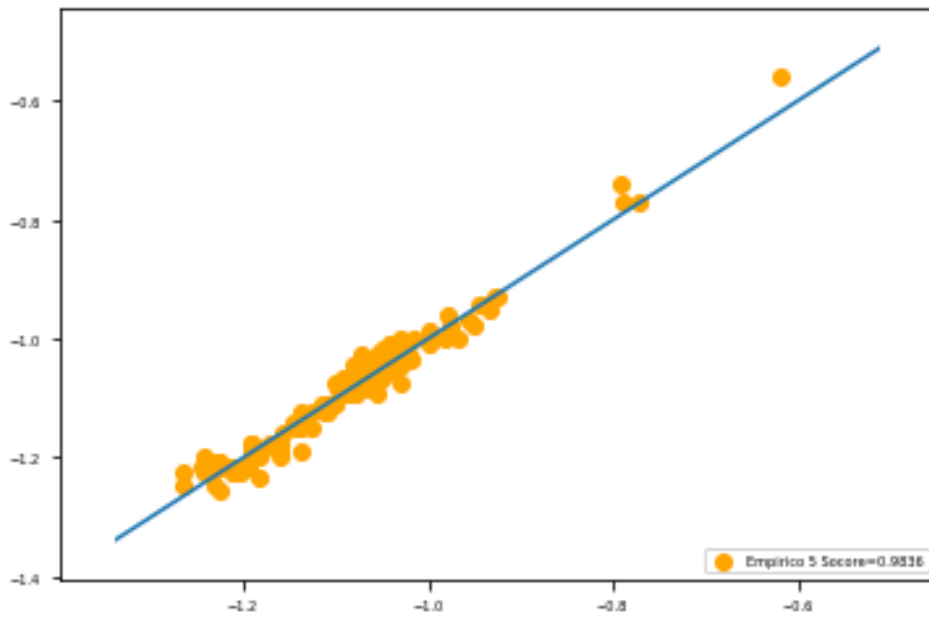


Figura 37 - Datos Predichos X10 vs Datos Reales X10 - Técnico 5.

CAPÍTULO 6

CONCLUSIONES

6.1 ANÁLISIS DE COMPONENTES PRINCIPALES

El análisis de componentes principales no produjo mayores ventajas a la hora de reducir la cantidad de variables para facilitar el procesamiento de datos. Esto se puede observar en la Tabla 8 donde en la componente principal Z1 que representa el 61,60% de la variabilidad explicada todas las variables poseen una correlación moderada a baja. Esto evidencia que todos los instrumentos aportan información importante al sistema por tanto no es posible realizar una separación en grupos.

6.2 REDES NEURONALES

Como se puede ver en la Tabla 11, los coeficientes de Determinación R^2 de las predicciones son en su mayoría muy buenas utilizando todas las combinaciones. En las figuras 17, 18 y 19 podemos observar el poder predictivo de las redes neuronales con las 3 combinaciones realizadas, demostrando alta efectividad y acercándose a los valores reales. Así también en los gráficos de los Datos predichos vs Datos reales de todas las variables se puede comprobar que los puntos se acercan a la línea a 45°, y mientras más cerca se encuentran de la línea, más precisa es la predicción.

Considerando que se trata de instrumentos de alta precisión, y además no se han tenido en cuenta parámetros como temperatura o desfase de tiempo de reacción de los instrumentos existe una diferencia en los resultados. Utilizando los parámetros mencionados y relacionando los instrumentos. Se podrían obtener mejores resultados.

Por lo tanto, podemos concluir que las redes neuronales son una herramienta potente, de fácil acceso que con los parámetros y selección de variables adecuadas es posible obtener predicciones y a partir de estas establecer mecanismos de validación.

6.3 RECOMENDACIONES

- Calcular el tiempo de respuesta de los instrumentos con respecto a las variaciones de temperatura.
- Utilizar parámetros como temperatura ambiente, temperatura interna de las estructuras, nivel del embalse u otros parámetros que contribuyan a comprender mejor el comportamiento de las estructuras.
- Analizar métodos para validar los datos provenientes de la auscultación cuantitativa.

REFERENCIAS

ASCE (American Society of Civil Engineers). (1975). *"Lessons from Dam Incidents, USA,"*.

ASCE (American Society of Civil Engineers). (1988). *"Lessons from Dam Incidents USA-II,"*.

CALLEGARI, J. (2003). *Bioestadística, principios y aplicaciones*.

Cybenko, G. (1989). *Mathematics of control, signals, and systems 2.4*.

FERC. (1994). *Federal Energy Regulatory Commission*. Obtenido de <https://www.ferc.gov/sites/default/files/2020-04/chap9.pdf>

Grané Chávez, A. (2010). *Análisis de Componentes Principales»*. En: *Publicaciones*. Universidad Carlos III de Madrid.

ICOLD. (2021). *ICOLD*. Obtenido de <https://www.icold-cigb.org/>

Instrucción para el Proyecto, Construcción y Explotación de Grandes Presas. (1967). Obtenido de Spancold: <https://www.spancold.org/wp-content/uploads/2016/09/instruccion.pdf>

Itaipu Binacional. (2021). *Itaipu Binacional*. Obtenido de <https://www.itaipu.gov.br/es/energia/auscultacion>

Jansen, R. (1980). *Dams and Public Safety*. "Water public Safety". Water Power and Resources Service (now U.S. Bureau of Reclamation).

- Liska, D. (2019). Análise do comportamento de um bloco chave da barragem da Itaipu através da instrumentação. Foz de Iguaçu.
- Maiti, J. (2014). Obtenido de <https://www.youtube.com/watch?v=hkCT-6KJAK0&>
- McGonagle, J. y. (2019). *Backpropagation*. Obtenido de <https://brilliant.org/wiki/backpropagation/>
- Novak, P., Moffat, A., & Nalluri, C. (2001). Estructuras Hidráulicas. En P. Novak, A. Moffat, & C. Nalluri, *Estructuras hidráulicas* (pág. 599).
- PTI, Ceasb. (2018). *Curso de Instrumentação e a Auscultação de Barragem*.
- Ramírez, F. (2019). *TÉCNICAS DE APRENDIZAJE AUTOMÁTICO PARA ANÁLISIS A FLEXIÓN EN VIGAS RECTANGULARES DE HORMIGÓN ARMADO*.
- Silveira, J. F. (2003). Instrumentação e comportamento de fundações de Barragens de Concreto. En J. F. Silveira.
- SPANCOLD. (2005). GUÍAS TÉCNICAS DE SEGURIDAD DE PRESAS. En C. D. COMITÉ NACIONAL ESPAÑOL DE GRANDES PRESAS, *SEGURIDAD DE PRESAS*.
- Tsitsiklis, J., & Bertsekas, D. (2002). *Introduction to probability. Vol. 1. Athena Scientific. MA*.

UNITED STATES DEPARTMENT OF THE INTERIOR BUREAU OF
RECLAMATION WASHINGTON, D.C. (1966). *DESIGN OF SMALL
DAMS*. (I. J. Lepe., Trans.) Washington, D.C.: Continental.

APÉNDICE



Figura 38 - Bloque E-06. Fase aguas abajo.

```
#librerías
from sklearn.decomposition import PCA
from sklearn import preprocessing
import sympy
import random # Permite obtener de distintos modos números aleatorios
import matplotlib.pyplot as plt

#Tratamiento de datos
# -----
import pandas as pd
import numpy as np

# Preprocesado y modelado
# -----
from scipy.stats import pearsonr
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error
import statsmodels.api as sm
import statsmodels.formula.api as smf

### Red neuronal
from sklearn.neural_network import MLPRegressor
from sklearn.model_selection import train_test_split
```

Figura 39 - Códigos de desarrollo en Jupyter Notebook.

```
En [700]: #lectura de la base de datos
df = pd.read_csv('TFG ULTIPO.csv')*0.95
df

Fuera [700]:
```

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
0	14.222460	-3.709845	7.2010	1.3775	0.7000	0.9035	0.9570	-0.1140	-0.1045	-1.2825	-1.1020	-0.0745
1	14.527305	-3.841445	7.2010	1.4440	0.7980	0.5880	0.9685	-0.1235	-0.1140	-1.2540	-1.0640	-0.0480
2	14.873130	-3.372500	8.9535	1.4820	0.8550	0.8175	0.1045	-0.0855	-0.0780	-1.2160	-1.0185	-0.6175
3	13.797800	-1.806540	0.3935	1.4950	0.7055	0.5700	0.0760	-0.1140	-0.1140	-1.2350	-1.0355	-0.6270
4	13.652070	-1.460910	0.2095	1.3775	0.7500	0.5225	0.0055	-0.1045	-0.1045	-1.2445	-1.0545	-0.5555
...
885	-1.425000	-0.205000	0.3040	0.1235	0.0780	0.1235	-0.1900	-0.1615	-0.1140	-0.3810	-0.3800	-0.2375
886	-1.330000	-0.475000	0.0855	0.1045	0.0955	0.1140	-0.1615	-0.1425	-0.0950	-0.2000	-0.3040	-0.1900
887	-0.780000	-0.475000	-0.1235	0.0780	0.0970	0.0850	-0.1710	-0.1140	-0.0950	-0.0855	-0.1995	-0.1140
888	-0.190000	-0.190000	-0.2470	0.0855	0.0570	0.1045	-0.1710	-0.1140	-0.0965	-0.0665	-0.1710	-0.0950
889	0.060000	0.000000	-0.2470	0.0570	0.0380	0.0760	-0.1330	-0.1140	-0.0985	-0.0985	-0.1615	-0.0855

670 filas x 12 columnas

Figura 40 - Códigos de desarrollo en Jupyter Notebook.

Combinaciones

```
Total 1 X = gl [['X2', 'X3', 'X4', 'X5', 'X6','X7', 'X8', 'X9', 'X10', '11', 'X12']] y = df [['X1']]
Total 2 X = df [['X1', 'X2', 'X4', 'X5', 'X6', 'X7', 'X8', 'X9', 'X10', 'X11', 'X12']] y = df [['X3']]
Total 3 X = gl [['X1', 'X2', 'X3', 'X4', 'X5', 'X7', 'X8', 'X9', 'X10', 'X11', 'X12']] y = df [['X6']]
Total 4 X = gl [['X1', 'X2', 'X3', 'X4', 'X5', 'X6', 'X7', 'X9', 'X10', 'X11', 'X12']] y = df [['X8']]
Total 5 X = df [['X1', 'X2', 'X3', 'X4', 'X5', 'X6', 'X7', 'X8', 'X9', 'X11', 'X12']] y = df [['X10']]
PCA 1 X = gl [['X3', 'X4', 'X5', 'X6']] y = gl [['X1']]
PCA 2 X = gl [['X1', 'X4', 'X5', 'X6']] y = gl [['X3']]
PCA 3 X = gl [['X1', 'X3', 'X5', 'X6']] y = gl [['X6']]
PCA 4 X = gl [['X7', 'X9', 'X11', 'X12']] y = gl [['X8']]
PCA 5 PCA 1 X = gl [['X2', 'X9', 'X11', 'X12']] y = gl [['X10']]
Empírico 1 X = gl [['X3', 'X5', 'X11']] y = gl [['X1']]
Empírico 2 X = gl [['X1', 'X5', 'X11']] y = gl [['X3']]
Empírico 3 X = gl [['X3', 'X9', 'X12']] y = gl [['X6']]
Empírico 4 X = gl [['X5', 'X11']] y = gl [['X8']]
Empírico 5 X = gl [['X4', 'X7', 'X12']] y = gl [['X10']]
```

Figura 41 - Códigos de desarrollo en Jupyter Notebook.

```
En [86]: X_train, X_test, y_train, y_test=train_test_split(X,y, test_size=0.2)
En [102]: mlr= MLPRegressor (solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(2,225), random_state=20,activation = "relu",max_iter=5000)
mlr.fit(X_train,y_train)
```

Figura 42 - Códigos de desarrollo en Jupyter Notebook.