



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA CIVIL

**CONFECCIÓN DE MODELOS SUSTITUTOS PARA PROCESOS DE
SIMULACIÓN GEOTÉCNICOS Y CUANTIFICACIÓN DE INCERTIDUMBRE**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL

MAXIMILIANO RODOLFO MORALES NAVARRETE

PROFESOR GUÍA:
RAFAEL RUIZ GARCÍA

MIEMBROS DE LA COMISIÓN:
MARÍA ELENA VALENCIA VERA
YOLANDA ALBERTO HERNÁNDEZ

Este trabajo ha sido parcialmente financiado por:
ITASCA CHILE SPA

SANTIAGO DE CHILE
2021

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL
POR: MAXIMILIANO RODOLFO MORALES NAVARRETE
FECHA: 2021
PROFESOR GUÍA: RAFAEL RUIZ GARCÍA

CONFECCIÓN DE MODELOS SUSTITUTOS PARA PROCESOS DE SIMULACIÓN GEOTÉCNICOS Y CUANTIFICACIÓN DE INCERTIDUMBRE

El presente trabajo de título tiene el objetivo de desarrollar y presentar una metodología de confección de modelos sustitutos para procesos de modelamiento o simulación de carácter geotécnico, particularmente un sistema de depósito de relaves. Esto implica la realización de un modelo mediante el software de volúmenes finitos denominado FLAC3D, a partir del cual se obtienen resultados que conforman una base de datos, luego de ingresar un rango adecuado de valores de diferentes variables significativas a analizar. Esta base de datos permite el entrenamiento de un modelo sustituto mediante la aplicación del método Kriging o también conocido como regresión en procesos Gaussianos. De este modo es posible obtener un modelo sustituto capaz de predecir el comportamiento del modelo original, de forma mucho más ágil y eficiente. Finalmente, con las condiciones otorgadas por el modelamiento sustituto, se realiza un análisis de cuantificación de incertidumbre, aplicando los respectivos métodos referidos a esta área.

Los resultados esperados prácticos consisten en la propuesta de solución respecto a la problemática de eficiencia de estos procesos, para un sistema de monitoreo, de respuesta ágil, mejorada y capaz de percibir las incertezas respectivas, así como también la demostración de una metodología de modelamiento sustituto y cuantificación de incertidumbre para sistemas ingenieriles, el cual en este caso es un sistema de embalse de relaves.

Dedicado a mis seres queridos

Agradecimientos

Quiero agradecer a mi familia, a mis padres, a mi madre Verónica y a mi padre Rodolfo, por darme todo su apoyo, por darme la formación, educación y su ejemplo, que permitió convertirme en la persona que soy hoy. Agradezco a mi hermano Francisco, por contribuir con su juicio en este proceso. A mis abuelos, que siempre están presentes conmigo, y siempre han demostrado orgullo por mis logros y me han expresado todo su cariño, y a mi abuela Tita, que se qué está presenciando este momento en un lugar mejor.

A mis amigos de la universidad, a "la banda", con la que se vivieron tantos momentos buenos y malos, desafíos, traspasos, carretes, recuerdos imborrables, que me hicieron disfrutar mi tiempo en la universidad.

A mis amigos de vida, Nico, Felipe, Mati, por su amistad incondicional, por su apoyo y compañía, y por todos los momentos vividos.

A Geraldine, por estar presente, por todo lo entregado, por siempre apoyarme y estar en los mejores y peores momentos de mi vida.

Finalmente quiero agradecer a mi equipo de trabajo, al profesor Rafael, a María Elena y a la profesora Yolanda. Agradezco su disposición, su paciencia y la oportunidad entregada de trabajar con ustedes.

* El presente trabajo se realizó en colaboración con la empresa Itasca Chile SPA

Tabla de Contenido

1. Introducción	1
1.1. Introducción y motivación	1
1.2. Objetivos	4
1.2.1. Objetivo general	4
1.2.2. Objetivos específicos	4
1.3. Contenido del trabajo de memoria	4
2. Metodología de Trabajo	6
2.1. Diseño de sistema	6
2.2. Comunicación entre softwares y manipulación de datos	6
2.3. Confección de modelos sustitutos	6
2.4. Cuantificación de incertidumbre	7
3. Marco Teórico	8
3.1. Geotecnia: Embalse de relaves	8
3.1.1. Depósitos de relaves	8
3.1.1.1. Descripción de elementos	9
3.1.1.2. Tipos de depósitos	11
3.1.1.3. Métodos constructivos	11
3.1.1.4. Sistemas de control y monitoreo	12
3.1.1.5. Parámetros geotécnicos	12
3.1.1.6. Mecanismos de falla	14
3.1.2. Análisis de depósitos de relaves: Estabilidad física	15
3.1.2.1. Análisis estático	15
3.1.2.2. Análisis pseudoestático	15
3.1.2.3. Análisis dinámico	16
3.1.3. Métodos de análisis de estabilidad física	16
3.1.3.1. Método de equilibrio límite	16
3.1.3.2. Métodos finitos	17
3.1.4. Mecanismo de cálculo y modelamiento en FLAC3D	18
3.1.4.1. Descripción	18
3.1.4.2. Formulación matemática	18
3.1.4.3. Representación de elementos	19
3.1.4.4. Resolución de equilibrio y factor de seguridad	19
3.1.4.5. Argumento de uso de método	21
3.1.4.6. Modelamiento	21
3.1.5. Comentarios sobre revisión bibliográfica	21

3.2.	Modelamiento Sustituto	23
3.2.1.	Modelos sustitutos	23
3.2.2.	Métodos predictivos	23
3.2.3.	Método Kriging o regresión en procesos Gaussianos	25
3.2.3.1.	Conceptos previos	25
3.2.3.2.	Definición y descripción del método	26
3.2.3.3.	Preparación de base de datos	27
3.2.3.4.	Entrenamiento y optimización del modelo	28
3.2.3.5.	Predicción	30
3.2.3.6.	Validación de modelo	31
3.2.3.6.1.	Validación cruzada	31
3.2.3.6.2.	Validación de datos de prueba o test	33
3.2.3.7.	Características del método	34
3.2.4.	Comentarios sobre revisión bibliográfica	34
3.3.	Cuantificación de Incertidumbre	36
3.3.1.	Conceptos básicos	36
3.3.2.	Propagación de incertidumbre	37
3.3.3.	Probabilidad de falla	39
3.3.4.	Análisis de sensibilidad	40
3.3.5.	Comentarios sobre revisión bibliográfica	42
4.	Diseño de Modelo de Embalse de Relaves	43
4.1.	Bases de diseño	43
4.1.1.	Categorización de embalse	43
4.1.2.	Dimensiones	44
4.1.3.	Análisis de estabilidad	46
4.2.	Geometría	47
4.3.	Caracterización geotécnica	49
4.4.	Modelamiento del sistema	49
4.4.1.	Método numérico	49
4.4.2.	Modelo constitutivo	50
4.4.3.	Condiciones de borde	50
4.4.4.	Condiciones iniciales	50
4.4.5.	Nivel freático	50
4.4.6.	Resolución de equilibrio	51
4.4.7.	Cálculo de factor de seguridad	51
4.5.	Selección de parámetros a analizar	51
4.5.1.	Parámetros a analizar	52
4.5.2.	Rangos de parámetros	53
4.5.3.	Selección de valores de parámetros	53
4.6.	Resultados	56
5.	Confección de Modelo Sustituto	61
5.1.	Importación, filtro y manipulación de datos	62
5.1.1.	Importación de datos	62
5.1.2.	Filtro de datos	62
5.1.3.	Manipulación de datos	62

5.2.	Construcción de modelo sustituto	66
5.2.1.	Definición de características de método	66
5.2.1.1.	Función base	66
5.2.1.2.	Función de covarianza (Kernel)	66
5.2.2.	Entrenamiento del modelo	66
5.2.2.1.	Modelo de FS global	67
5.2.2.2.	Modelo de contorno de FS local	67
5.2.3.	Validación del modelo	67
5.3.	Resultados de modelo sustituto	68
5.3.1.	Modelo de FS global	68
5.3.1.1.	Tiempo de entrenamiento	68
5.3.1.2.	Predicción	68
5.3.1.3.	Validación	68
5.3.2.	Modelo de contorno de FS	71
5.3.2.1.	Tiempo de entrenamiento	71
5.3.2.2.	Predicción	71
5.3.2.3.	Validación	71
6.	Cuantificación de Incertidumbre	78
6.1.	Modelo de FS	78
6.1.1.	Planteamiento	78
6.1.2.	Propagación de incertidumbre	80
6.1.3.	Probabilidad de falla	81
6.1.4.	Análisis de sensibilidad	82
6.2.	Modelo de contorno de FS	82
6.2.1.	Planteamiento	83
6.2.2.	Propagación de incertidumbre	83
7.	Análisis y Discusión de Resultados	85
7.1.	Diseño de modelo de embalse de relaves	85
7.1.1.	Software de modelamiento FLAC3D	85
7.1.2.	Definición de modelo	86
7.1.3.	Modelo planteado	86
7.1.4.	Selección de parámetros y casos	87
7.1.5.	Resultados de FS y contorno de FS	88
7.2.	Confeción de modelo sustituto	89
7.2.1.	Manipulación de datos	89
7.2.2.	Construcción de modelo sustituto	90
7.2.3.	Resultados	91
7.3.	Cuantificación de incertidumbre	93
7.3.1.	Preliminares	93
7.3.2.	Propagación de incertidumbre	93
7.3.3.	Probabilidad de falla	94
7.3.4.	Análisis de sensibilidad	94
7.4.	Tiempos de procesamiento	94
7.4.1.	Resumen de tiempos	95
7.4.2.	Mejorías de tiempo	95

7.5.	Comparación	96
7.5.1.	Modelo con muestreo con LHS	96
7.5.2.	Modelo con menos cantidad de casos	96
7.5.3.	Modelo con distinta librería	97
7.6.	Aplicaciones	98
7.6.1.	Presentación y generalización de metodología	98
7.6.2.	Cuantificación de incertidumbre	99
7.6.3.	Sistema de monitoreo	99
8.	Conclusiones	103
	Bibliografía	105
Anexo A.	Conceptos de Probabilidad y Estadística	108
A.1.	Definiciones	108
A.2.	Parámetros básicos	108
A.3.	Distribuciones de probabilidad	109
A.4.	Propiedades generales	110
A.4.1.	Marginalidad	110
A.4.2.	Condicionabilidad	110
A.4.3.	Teorema de Bayes	110
A.5.	Procesos Gaussianos	111
A.5.1.	Definición	111
A.5.2.	Propiedades	111
A.5.2.1.	Distribución conjunta de vectores aleatorios	111
A.5.2.2.	Marginalidad en vectores aleatorios	111
A.5.2.3.	Condicionabilidad en vectores aleatorios	111
Anexo B.	Herramientas Python	112
B.1.	Librerías	112
Anexo C.	Procedimiento de selección de parámetros	113
C.1.	Parámetros a analizar	113
C.2.	Ejecución de casos	114
C.3.	Parámetros a analizar finales	116
Anexo D.	Distribución de casos a ejecutar	117
D.1.	Gráficos de distribución de casos	117
Anexo E.	Resultados de modelo de embalse de relaves en software de diseño	119
E.1.	Factor de seguridad global y contorno	119
E.1.1.	Caso más favorable	119
E.1.2.	Caso intermedio	120
E.1.3.	Caso más desfavorable	120
E.2.	Presiones de poros	121
E.2.1.	Caso más favorable	121
E.2.2.	Caso intermedio	121
E.2.3.	Caso más desfavorable	122

E.3. Desplazamientos	122
E.3.1. Caso más favorable	122
E.3.2. Caso intermedio	123
E.3.3. Caso más desfavorable	123
E.4. Esfuerzos	124
E.4.1. Caso más favorable	124
E.4.2. Caso intermedio	124
E.4.3. Caso más desfavorable	125
Anexo F. Desarrollo: Caso con muestreo con LHS	126
F.1. Selección de parámetros	126
F.2. Confección de modelo sustituto	126
Anexo G. Desarrollo: Caso con menor cantidad de datos	129
G.1. Selección de parámetros	129
G.2. Confección de modelo sustituto	130
Anexo H. Desarrollo: Caso con librería alternativa	132
H.1. Selección de parámetros	132
H.2. Confección de modelo sustituto	132

Índice de Tablas

3.1.	Variables a controlar y sistemas de control	12
3.2.	Parámetros geotécnicos	13
3.3.	Mecanismos de falla en embalses de relaves	14
3.4.	Métodos predictivos	24
4.1.	Categorización de muro según altura y capacidad	44
4.2.	Registro de principales embalses de Chile	45
4.3.	Rangos de dimensiones de muro	45
4.4.	Dimensiones del modelo	48
4.5.	Discretización del modelo	48
4.6.	Propiedades geotécnicas del modelo	49
4.7.	Parámetros a analizar	52
4.8.	Parámetros a analizar - Selección	52
4.9.	Rangos de parámetros	53
4.10.	Análisis de casos y tiempo de ejecución	53
4.11.	Casos obtenidos por método	54
4.12.	Selección de valores de parámetros por método	54
4.13.	Casos obtenidos por método	54
4.14.	Análisis de casos y tiempo de ejecución - Test	55
4.15.	Selección de valores de parámetros por método - Test	55
4.16.	Casos obtenidos por método - Test	55
4.17.	Parámetros de entrada y salida	56
4.18.	Formato datos a extraer	56
4.19.	Datos a extraer	57
4.20.	Datos de tensor de esfuerzos	57
4.21.	Datos de tensor de deformaciones	57
4.22.	Datos - Caso más favorable	58
4.23.	Datos - Caso intermedio	58
4.24.	Datos - Caso más desfavorable	59
5.1.	Formato de datos actualizado	62
5.2.	Formato de datos final	65
5.3.	Formato de datos - FSG	67
5.4.	Formato de datos - Contorno FS	67
5.5.	Errores - Datos de soporte - Modelo de FS global	69
5.6.	Errores - Valid. Cruzada - Modelo de FS global	69
5.7.	Errores - Modelo de FS global	70
5.8.	Errores - Test - Modelo de FS global	71
5.9.	Errores - Modelo de FS global	71
5.10.	Tiempo de ejecución de submodelos de contorno de FS	71

5.11.	Errores - Datos de soporte - Modelo de contorno de FS	72
5.12.	Continuación Tabla 5.11	72
5.13.	Errores residuales - Valid. Cruzada - Modelo de contorno de FS	72
5.14.	Continuación Tabla 5.13	72
5.15.	Errores RMSE y LOO - Valid. cruzada Modelo de contorno de FS	73
5.16.	Continuación Tabla 5.15	73
5.17.	Datos - Caso más favorable - Modelo sustituto	73
5.18.	Datos - Caso intermedio - Modelo sustituto	74
5.19.	Datos - Caso más desfavorable - Modelo sustituto	75
5.20.	Errores residuales - Test - Modelo de contorno de FS	75
5.21.	Continuación Tabla 5.20	75
5.22.	Errores RMSE y LOO - Test - Modelo de contorno de FS	76
5.23.	Continuación Tabla 5.22	76
5.24.	Datos - Test 1 - Modelo sustituto	76
5.25.	Datos - Test 2 - Modelo sustituto	77
6.1.	Información estadística de variables de entrada	79
6.2.	Análisis de sensibilidad - Índices de Sobol	82
6.3.	Información estadística de variables de entrada - Modelo contorno de FS	83
6.4.	Información estadística de variables de salida - Modelo contorno de FS	84
7.1.	Resumen de tiempos de procesos	95
7.2.	Mejorías en tiempo de respuesta	95
C.1.	Parámetros a analizar - Proceso de selección	113
C.2.	Parámetros fijos - Proceso de selección	114
C.3.	Variabilidad de parámetros a analizar	115
C.4.	Parámetros a analizar - Selección previa	115
C.5.	Parámetros a analizar - Selección filtrada	116
C.6.	Parámetros a analizar - Selección final	116
F.1.	Casos obtenidos por método	126
F.2.	Errores - Valid. Cruzada - Modelo de FS global	127
F.3.	Errores - Modelo de FS global - Anexo F	128
G.1.	Selección de valores de parámetros por método	129
G.2.	Casos obtenidos por método	129
G.3.	Errores - Valid. Cruzada - Modelo de FS global	131
G.4.	Errores - Modelo de FS global - Anexo G	131
G.5.	Comparación - Tiempo de ejecución	131
H.1.	Errores - Valid. Cruzada - Modelo de FS global	133
H.2.	Errores - Modelo de FS global - Anexo H	133
H.3.	Errores - Datos de soporte - Modelo de contorno de FS - Anexo H	133
H.4.	Continuación Tabla H.3	134
H.5.	Errores residuales - Valid. Cruzada - Modelo de contorno de FS	134
H.6.	Continuación Tabla H.5	134
H.7.	Errores RMSE y LOO - Modelo de contorno de FS	134
H.8.	Continuación Tabla H.7	135
H.9.	Tiempo de ejecución de submodelos de contorno de FS	135

Índice de Ilustraciones

2.1.	Metodología de Trabajo de Memoria	7
3.1.	Diagrama de depósito de relaves	9
3.2.	Método de aguas abajo	11
3.3.	Gráfico mecanismos de falla	15
3.4.	Elemento - Tetraedro	19
3.5.	Diagrama - Tensor de esfuerzos	20
3.6.	Proceso de entrenamiento de modelo sustituto	28
3.7.	Proceso de validación cruzada	31
4.1.	Embalse de relaves - Perspectiva 3D	47
4.2.	Geometría de embalse a diseñar	47
4.3.	Geometría de embalse - FLAC3D	48
4.4.	Nivel freático en embalse - FLAC3D	51
4.5.	FSG versus nivel freático	57
4.6.	FSG versus nivel freático 2	57
4.7.	FSG versus ángulo de fricción (ϕ)	58
4.8.	FSG versus cohesión	58
4.9.	Caso más favorable - Contorno FS - FLAC3D - FSG=1.88	58
4.10.	Caso intermedio - Contorno FS - FLAC3D - FSG=1.3813	59
4.11.	Caso más desfavorable - Contorno FS - FLAC3D - FSG=1.006	59
5.1.	Contorno de factor de seguridad - Python	63
5.2.	Puntos de borde de contorno de factor de seguridad - Python	63
5.3.	Parametrización de puntos	64
5.4.	Parametrización de puntos de borde de contorno de factor de seguridad - Python	64
5.5.	Errores residuales Modelo FS global	69
5.6.	FSG versus nivel freático - Modelo sustituto	70
5.7.	FSG versus nivel freático 2 - Modelo sustituto	70
5.8.	FSG versus ángulo de fricción (ϕ) - Modelo sustituto	70
5.9.	FSG versus cohesión - Modelo sustituto	70
5.10.	Errores residuales Modelo Contorno FS	73
5.11.	Comparación - Caso más favorable - Contorno FS - FSG=1.88	74
5.12.	Comparación - Caso Intermedio - Contorno FS - FSG=1.3813	74
5.13.	Comparación - Caso más desfavorable - Contorno FS - FSG=1.006	75
5.14.	Comparación - Test - Contorno FS - FSG=1.331	76
5.15.	Comparación - Test - Contorno FS - FSG=1.431	77
6.1.	Gráfica de PDFs de variables de entrada	79
6.2.	Histograma - FSG	80
6.3.	PDF - FSG	81
6.4.	Gráfica de Probabilidad de Falla	82

6.5.	Propagación de incertidumbre - Modelo de Contorno de FS - FS Local=1.4	83
7.1.	Diagrama referencial - Embalse de relaves	100
7.2.	Diagrama explicativo - Sistema con modelo original	100
7.3.	Diagrama explicativo - Sistema con modelo sustituto	101
D.1.	Distribución de muestras - LHS	117
D.2.	Distribución de muestras - Equiespaciamiento uniforme de variables	118
E.1.	Caso más favorable - Contorno FS - FSG=1.88	119
E.2.	Caso más desfavorable - Contorno FS - FSG=1.3813	120
E.3.	Caso más desfavorable - Contorno FS - FSG=1.006	120
E.4.	Caso más favorable - Presión de poros - FSG=1.88	121
E.5.	Caso intermedio - Presión de poros - FSG=1.3813	121
E.6.	Caso más desfavorable - Presión de poros - FSG=1.006	122
E.7.	Caso más favorable - Desplazamientos en XX - FSG=1.88	122
E.8.	Caso intermedio - Desplazamientos en XX - FSG=1.3813	123
E.9.	Caso más desfavorable - Desplazamientos en XX - FSG=1.006	123
E.10.	Caso más favorable - Esfuerzos en XX - FSG=1.88	124
E.11.	Caso intermedio - Esfuerzos en XX - FSG=1.3813	124
E.12.	Caso más desfavorable - Esfuerzos en XX - FSG=1.006	125
F.1.	FSG versus nivel freático - Modelo sustituto	127
F.2.	FSG versus nivel freático 2 - Modelo sustituto	127
F.3.	FSG versus ángulo de fricción (ϕ) - Modelo sustituto	127
F.4.	FSG versus cohesión - Modelo sustituto	127
G.1.	FSG versus nivel freático - Modelo sustituto	130
G.2.	FSG versus nivel freático 2 - Modelo sustituto	130
G.3.	FSG versus ángulo de fricción (ϕ) - Modelo sustituto	130
G.4.	FSG versus cohesión - Modelo sustituto	130
H.1.	FSG versus nivel freático - Modelo sustituto	132
H.2.	FSG versus nivel freático 2 - Modelo sustituto	132
H.3.	FSG versus ángulo de fricción (ϕ) - Modelo sustituto	133
H.4.	FSG versus cohesión - Modelo sustituto	133

Capítulo 1

Introducción

1.1. Introducción y motivación

El desarrollo de la ingeniería civil ha expresado avances con el paso del tiempo, demostrado en la evolución de las metodologías, procedimientos y herramientas a utilizar, con el objetivo de realizar una mejor ingeniería, un mejor diseño, cálculo, planificación y mejoras en general en las disciplinas que aborda el rubro. Se busca constantemente un perfeccionamiento en estos procedimientos al querer representar cada vez de forma más efectiva y adecuada los comportamientos, fenómenos, situaciones o sistemas que se desean modelar.

Los programas computacionales han sido una herramienta fundamental para la realización del trabajo y han otorgado la posibilidad de efectuar un análisis confiable y preciso en sistemas ingenieriles, principalmente por el procedimiento de modelamiento o simulación mediante estos programas, el cual se ha instaurado como una actividad clave en las disciplinas y en la labor del ingeniero. Sin embargo, existe una problemática, en la que se enfoca el presente trabajo, la cual es la gran carga computacional que implica el funcionamiento de los modelos en software complejos, lo que conlleva un gasto de tiempo en los procesamientos, que en ocasiones se vuelve inmanejable, como por ejemplo, frente a la situación en que se requiere ejecutar un mismo modelo para múltiples casos con diferentes variables. El software de modelamiento geotécnico FLAC3D, otorgado por la empresa Itasca para su uso en la presente memoria, y la mayoría de los software enfocados en el planteamiento de modelos de sistemas geotécnicos presentan principios matemáticos y físicos, fundamentos de elementos o volúmenes finitos y procedimientos iterativos que permiten la realización de cálculos asociados a distintos análisis que pueden presentar estos sistemas mencionados (análisis de estabilidad, hidráulico, sísmico, etc) e implican una carga computacional considerablemente alta.

Antes de mencionar las razones que hacen necesaria la búsqueda por obtener respuestas más rápidas y eficientes, se deben describir los embalses de relaves, los cuales son obras civiles y sistemas geotécnicos cuyo objetivo es el almacén y soporte de grandes cantidades de materia. Estas obras se caracterizan típicamente por ser de gran magnitud y de considerable riesgo, lo que implica un criterio cuidadoso en su análisis y diseño. Conlleva el análisis de múltiples parámetros geométricos, resistentes, hidráulicos, entre otros, y la comprensión de determinados mecanismos de falla, tales como la inestabilidad de taludes. El análisis de estos sistemas ha adquirido suma importancia, debido al aumento de su demanda y al aumento de la

dimensionalidad requerida de estas obras, esto a su vez dado por el crecimiento de la actividad minera. De esta manera se ha ido generando una cultura o práctica del análisis de estabilidad de estas obras. Esto es mencionado de forma muy completa y adecuada en el documento de Barrera y Campaña [1], en el cual se presenta la evolución de los métodos de análisis de estabilidad física. Debido a las características y a los antecedentes existentes respecto a estas obras, tales como fallas y condiciones desfavorables en las mismas, es que surgen los esfuerzos por evitar que se produzcan este tipo de situaciones, y se hace necesaria la búsqueda de más herramientas útiles y eficientes que otorguen mayor seguridad y confiabilidad para el desarrollo de una mejor y más responsable ingeniería.

Dicho esto, es posible mencionar las razones para la mejoría en la respuesta del sistema: en primer lugar, es requerida por la necesidad de realizar ejecuciones múltiples del modelo en un tiempo razonable y adecuado, y así tener la capacidad de evaluación para diferentes escenarios del sistema planteado y comprender el comportamiento y las variables asociadas, para así prevenir la falla y lograr desarrollar un mejor diseño; en segundo lugar, surge la necesidad por el planteamiento de un proyecto adjudicado a Itasca (a describir en una próxima sección), asociado a la respuesta rápida del cálculo de estabilidad de un embalse de relaves, a modo de establecer un sistema de monitoreo, por lo que una respuesta ágil es clave frente a estas circunstancias; y finalmente, es un requerimiento para la implementación del área de estudio referida a cuantificación de incertidumbre, ya que sus métodos involucran la aplicación de procedimientos de muestreo, es decir, se requiere la evaluación de una gran cantidad de muestras o casos (Método de Montecarlo).

Es por esto que la necesidad de mejorar y agilizar estos procesos adquiere suma importancia, y para resolver la problemática planteada en el presente trabajo, se propone la confección de modelos sustitutos. El desarrollo de modelos sustitutos tiene múltiples utilidades y esta área de estudio esta relacionada con el área referida a metodologías estadísticas de análisis predictivo y a fundamentos de inteligencia artificial y Machine Learning (o aprendizaje automático). En base al documento de Yaochu [2], se identifica y comprende la evolución de este tipo de modelos, que surgen debido a las necesidades y desafíos que van presentándose en el campo científico e ingenieril. El modelamiento sustituto se centra básicamente en la búsqueda de relaciones entre variables, basado en una base de datos previamente establecida, para realizar una predicción según un nuevo set de datos, es decir, entrando en mayor detalle, dichos modelos se crean a partir de una base de datos, en la cual se presentan diferentes casos para un mismo sistema diseñado, ejecutados por un software comercial (FLAC3D). Esta base sirve para aplicar en ella procesos regresivos, interpolantes, estableciendo así un proceso de entrenamiento del modelo sustituto, y de este modo otorgarle una capacidad predictiva evaluando nuevos valores y obtener nuevas respuestas de forma confiable y validada, es decir con errores que sean los más bajos posibles, respecto a la diferencia entre los valores reales y los predichos. Esta predicción saca ventaja respecto a la respuesta del modelo original (FLAC3D), ya que logra obtener datos intermedios y la razón principal es que el tiempo de ejecución de una respuesta del modelo sustituto es considerablemente menor, lo cual se evaluará y verificará en esta memoria. El método que se propone para su uso en el presente trabajo, que es el algoritmo que realiza el proceso regresivo, es el método Kriging o regresión en procesos Gaussianos, y su uso es justificado debido a sus características de suavidad y de adecuado ajuste, lo cual es descrito en mayor detalle en una próxima sección, además de su utilidad en los procedimientos de cuantificación de incertidumbre.

El estudio de las incertidumbres y su inclusión en las distintas aplicaciones de la ingeniería es de suma relevancia y otorga muchas utilidades en la comprensión del sistema a analizar. Los fundamentos probabilísticos y estadísticos del área de estudio de la cuantificación de incertidumbre dan solidez y respaldo a sus métodos, y debido a su planteamiento es posible su uso frente a diversos tipos de sistemas ingenieriles, científicos o incluso ámbitos externos a estas categorías. Esta área de estudio, al asociarla con los sistemas geotécnicos o mineros presenta variadas relaciones, tal y como es mencionado por Gúzman para Minería Chilena [3]. Algunas de estas son incertezas comprendidas por la geoestadística para las leyes y ubicaciones referidas a distintas materias primas mineras, o también la planificación de obra, referido a disponibilidad de equipo minero por ejemplo. Las incertezas están incluidas en los sistemas geotécnicos de un modo u otro, y los embalse de relaves tienen esta condición, viéndose reflejada en las propiedades características, geométricas y resistentes, tales como el talud, el ángulo de fricción, la cohesión, etc, las cuales presentan un grado de incerteza en su obtención o determinación y a su vez provocan una incertidumbre en las variables de salida del sistema, como el factor de seguridad. De esta manera, la aplicación de los métodos de cuantificación de incertidumbre, tales como propagación de incertidumbre, probabilidad de falla y análisis de sensibilidad, que serán explicados a lo largo del documento, presentan utilidades en la toma de decisiones, en el énfasis que se le debe dar a las diversas variables involucradas en el sistema, entre otras.

La memoria pretende entonces proponer y plantear un mejoramiento en el análisis, en la comprensión de incertezas y en la respuesta de un modelo geotécnico en términos de tiempo y carga computacional, a partir de los procesos de modelamiento sustituto y la cuantificación de incertidumbre. Estos procedimientos se realizan planteando una metodología que consiste en el modelamiento del sistema geotécnico referido a un embalse de relaves en el software FLAC3D, de modo de obtener una base de datos con capacidad de exportación a diferentes formatos y así importarlo a una nueva interfaz, con herramientas y posibilidades de manipulación de datos e implementación de algoritmos de modelamiento sustituto (Interfaz Python). Así se permite la realización del entrenamiento del modelo sustituto mediante el método Kriging o regresión en procesos Gaussianos, con su respectiva validación, y ya se establece el contexto adecuado para la aplicación de los métodos de cuantificación de incertidumbre.

1.2. Objetivos

El presente trabajo tiene los siguientes objetivos generales y específicos.

1.2.1. Objetivo general

Desarrollar una metodología de creación y confección de modelos sustitutos en base a un proceso de simulación de un sistema geotécnico, específicamente un embalse de relaves, demostrándose en un ejemplo referencial y típico del rubro, con el cual se busca realizar una mejora en el tiempo de procesamiento y disminuir la carga computacional que implica la respuesta de estos procesos, de modo de otorgar un contexto ideal para la realización de un proceso de análisis y comprensión de las incertezas del sistema, mediante cuantificación de incertidumbre.

1.2.2. Objetivos específicos

- Planteamiento de un sistema geotécnico, específicamente un embalse de relaves, para su modelación y posterior evaluación y análisis en el software de modelamiento FLAC3D.
- Creación de modelos sustitutos, mediante la implementación del método Kriging o regresión en procesos Gaussianos, para su uso en la predicción de la respuesta del sistema geotécnico presentado y otorgar el contexto adecuado para la aplicación de métodos de cuantificación de incertidumbre.
- Análisis de incertezas de un sistema de embalse de relaves mediante métodos de cuantificación de incertidumbre en el modelo sustituto, tales como propagación de incertidumbre, cálculo de probabilidad de falla y análisis de sensibilidad.
- Proposición de una metodología de modelamiento sustituto y cuantificación de incertidumbre.

1.3. Contenido del trabajo de memoria

El contenido del trabajo presenta la siguiente distribución, ordenada en capítulos:

- Capítulo 1: Introducción
Se presenta una explicación introductoria del trabajo, la motivación de su realización, los objetivos y el orden o capitulación.
- Capítulo 2: Metodología de trabajo
Se explica las diferentes etapas de la memoria, que conforman la estructura del trabajo a realizar. Esta sección describe uno de los objetivos de la memoria, el cual es la presentación de una metodología, junto con sus respectivas componentes que se irán desarrollando a lo largo del documento.
- Capítulo 3: Marco teórico
El presente documento involucra un marco teórico dividido en tres áreas: Geotecnia, Modelamiento Sustituto y Cuantificación de incertidumbre.

- **Capítulo 4: Diseño de modelo de embalse de relaves**
Este capítulo describe el modelo del sistema geotécnico que se busca diseñar, el cual es un embalse de relaves. Este se modela mediante el software de modelamiento o simulación FLAC3D.
- **Capítulo 5: Confección de modelo sustituto**
La confección del modelo sustituto se realiza en base a los datos extraídos y manipulados, aplicando métodos predictivos, específicamente el método Kriging (o regresión en procesos Gaussianos), en el mismo software de análisis.
- **Capítulo 6: Cuantificación de incertidumbre**
El área de estudio referido a cuantificación de incertidumbre, aplicada directamente al modelo sustituto en el software de análisis, permite realizar propagación de incertidumbre, cálculo de probabilidad de falla y análisis de sensibilidad.
- **Capítulo 7: Análisis y discusión de resultados**
Se analizan y discuten los resultados del trabajo realizado, tales como el modelo original diseñado, la precisión del modelo sustituto y el análisis de incertidumbre. Se argumenta y justifica las decisiones tomadas en el trabajo, así como también los software, las herramientas y los criterios ocupados.
- **Capítulo 8: Conclusiones**
Se presentan las observaciones, análisis y comentarios finales, que concluyen el trabajo de memoria.

Capítulo 2

Metodología de Trabajo

La metodología establecida para realizar este trabajo, la cual es en sí, el objetivo y resultado que se quiere plantear y presentar en esta memoria, se describe en la actual sección. Se menciona cada paso o actividad a realizar, y se tiene un mapa conceptual que resume esta metodología en la Figura 2.1, donde es posible observar una división en cuatro secciones importantes que se describen en los siguiente párrafos.

2.1. Diseño de sistema

En primer lugar se realiza el modelamiento del sistema geotécnico que se busca diseñar, el cual es un embalse de relaves. Este se modela mediante el software de modelamiento o simulación FLAC3D. Se describen las características del sistema, los pasos a seguir para su modelación en el software, su implementación y la obtención de resultados.

2.2. Comunicación entre softwares y manipulación de datos

Se establece la comunicación entre el software de modelamiento FLAC3D, con el software de análisis determinado por la interfaz de Python, el cual contiene herramientas y adecuadas para la manipulación de datos y aplicación de metodologías de modelamiento sustituto y cuantificación de incertidumbre. Se realiza estableciendo una extracción de los datos desde el primero en un formato común y universal (planilla), de modo de poder realizar la importación al segundo software. Una vez realizada la importación de datos se procede a filtrar y manipular los resultados para así preparar la base de datos a utilizar con un determinado formato adecuado para la confección del modelo sustituto.

2.3. Confección de modelos sustitutos

Se realiza la confección del modelo sustituto, mediante el uso de la base de datos obtenida previamente y la aplicación del método Kriging o regresión en procesos Gaussianos, con las respectivas características que lo definen (función base, kernel, entre otras) que se serán descritas en una próxima sección. De este modo se efectúa el proceso de entrenamiento y luego un proceso de validación, y el modelo queda preparado para realizar predicciones.

2.4. Cuantificación de incertidumbre

Con el modelo sustituto ya creado y calibrado, es posible aplicar los métodos referidos a cuantificación de incertidumbre, para obtener parámetros y estudios significativos, tales como propagación de incertidumbre, probabilidad de falla y análisis de sensibilidad.

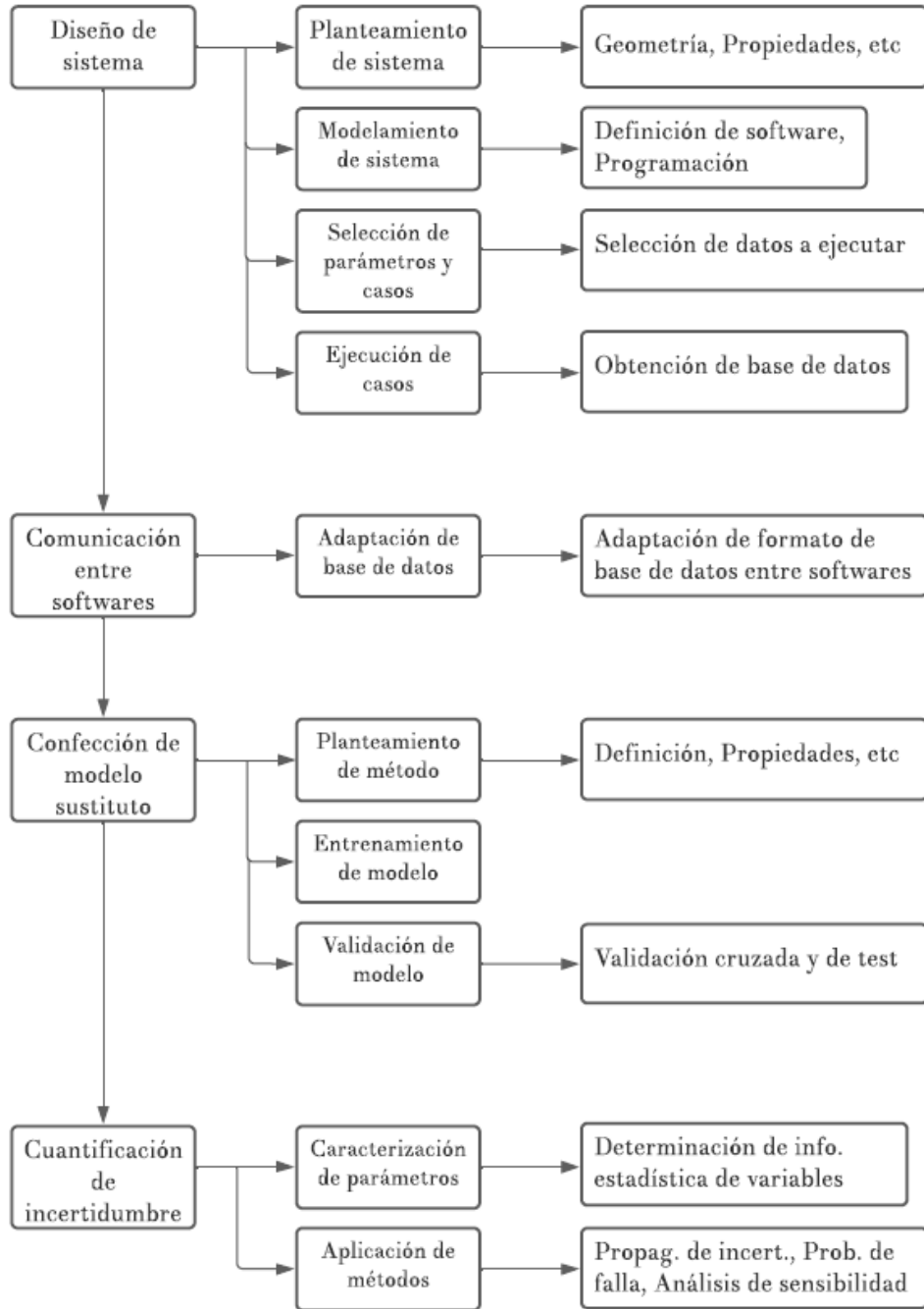


Figura 2.1: Metodología de Trabajo de Memoria

Capítulo 3

Marco Teórico

3.1. Geotecnia: Embalse de relaves

El análisis de los suelos, de sus propiedades y de su comportamiento mecánico, requieren de la integración y comprensión de las siguientes áreas de estudio: geología, geotecnia y geomecánica. El enfoque principal de la memoria es referido al área de geotecnia y geomecánica.

Los depósitos de relaves son obras que se enfocan en estas dos áreas de estudio. La identificación y obtención de los parámetros necesarios para caracterizar este tipo de sistema es abordado mediante conceptos geotécnicos, y el comportamiento y respuesta del mismo sistema es realizado en base a la geomecánica. Es posible considerar el área referida a la hidrogeología o hidráulica, debido a que un determinado conocimiento hidráulico es clave para el entendimiento del sistema mencionado, sin embargo, no se considera para el presente trabajo.

Para su comprensión y análisis se requiere entender las múltiples disciplinas que conlleva y se requiere la capacitación en el uso de determinados software, tales como FLAC3D, el cual es el que se implementa en el trabajo. La modelación requiere además conceptos de programación, y de métodos finitos.

En las siguientes secciones se describen los depósitos de relaves, sus características, los parámetros requeridos para su análisis y los métodos de cálculo y resolución.

3.1.1. Depósitos de relaves

Según el Decreto Supremo 248 [4], un depósito de relaves es toda obra estructurada en forma segura para contener los relaves provenientes de una planta de concentración húmeda de especies de minerales. Además, contempla sus obras anexas. Su función principal es la de servir como depósito, generalmente, definitivo de los materiales sólidos proveniente del relave transportado desde la planta, permitiendo así la recuperación, en gran medida, del agua que transporta dichos sólidos.

3.1.1.1. Descripción de elementos

Los depósitos de relaves, en general, están compuestos por los siguientes elementos o componentes, los cuales se pueden visualizar en la Figura 3.1 y se describen en base a la guía técnica del SERNAGEOMIN [5].

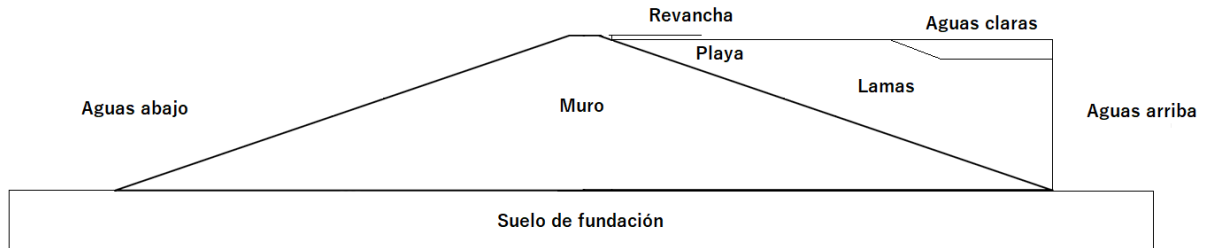


Figura 3.1: Diagrama de depósito de relaves

- i. Cubeta: Corresponde al volumen físico disponible donde se depositan las lamas y gran parte del agua. Es la componente más importante en relación con la vida útil del depósito. Esta a su vez es posible dividirla en tres zonas:
 - a. Zona de aguas claras: Corresponde a la laguna de aguas clarificadas que se forma en la cubeta debido a la sedimentación o decantación de las partículas sólidas. Esta poza permite la recuperación de aguas y al mismo tiempo la evacuación de estas desde la cubeta. Esta laguna debe mantenerse lo más alejada posible del muro y su evacuación debe hacerse siempre para no disminuir el grado de estabilidad del muro.
 - b. Zona de lamas: Zona donde quedan contenidos la parte sólida fina de los relaves (lamas).
 - c. Zona de playa: Zona donde se descargan los residuos a la cubeta. Se le llama zona de playa porque usualmente está seca en la superficie y se asemeja a una playa de arenas finas. Es la parte del depósito de relaves o lamas situada en las cercanías de la línea de vaciado.
- ii. Muro: Componente del depósito que delimita la cubeta y permite contener los residuos que en ella se descargan. Por lo tanto, el muro generalmente es la componente más importante en lo que dice relación con la estabilidad o grado de seguridad del depósito.
- iii. Sistema de Drenaje: Es el sistema utilizado para deprimir al máximo el nivel freático en el interior del cuerpo del muro, de modo de evitar el aumento excesivo de las presiones de poros y se produzca la colmatación del sistema.

- iv. Sistema de Impermeabilización:
 - a. En cubeta: Disposición de un sistema de impermeabilización del piso y paredes de la cubeta, usualmente construido en base a una serie de estratos impermeables y "permeables-drenantes". Esto con el objetivo de evitar la infiltración de aguas al subsuelo.
 - b. En muro: Disposición de sistema de impermeabilización en talud aguas arriba del muro, de modo de limitar las filtraciones provenientes desde las lamas de la cubeta. Generalmente para este sistema se dispone una capa de hormigón en el talud mencionado.
- v. Sistema de descarga, clasificación y selección de relaves: El sistema de descarga de residuos mineros permite depositar estos en la cubeta. Sistema se utiliza también para clasificar y seleccionar los residuos, de modo de utilizar parte de los mismos como material para la construcción del muro.
- vi. Sistema de Recuperación de Aguas: Sistema que permite devolver a la planta, las aguas claras que se han recuperado desde la poza o laguna de aguas clara, mediante bombeo desde balsas y/o descargas de torres de evacuación o decantación.
- vii. Canales de Desvío: Zanjas construidas o túneles construidos en las laderas para captar y desviar las escorrentías superficiales, impidiendo su ingreso a la cubeta del tranque o embalse de relaves.
- viii. Coronamiento: Zona de cota máxima del muro del depósito.
- ix. Revancha: Es la diferencia de cota entre la línea de coronamiento y la superficie inmediatamente vecina de la fracción aguas adentro de la cubeta. Se denomina como revancha mínima al desnivel entre el coronamiento del muro el punto más alto de las lamas; y se denomina revancha máxima al desnivel entre el coronamiento y la superficie de la poza de decantación.
- x. Nivel Freático: Es la cota de los puntos en que el agua de poros tiene presión nula.
- xi. Muro de Partida (o Inicial): Muro construido con material grueso al inicio del proceso de construcción del depósito de relaves.
- xii. Piezometría: Es el sistema de control de las presiones hidrostáticas en el interior del muro para detectar la presencia de sectores saturados. Los instrumentos utilizados se llaman piezómetros y con ellos se detecta el nivel freático del subsuelo.
- xiii. Vertederos de emergencia: Son vertederos de umbral variable destinados a evacuar el exceso de aguas acumuladas en la cubeta de un tranque al crecer la laguna de decantación por lluvias intensas.
- xiv. Empréstito: Zona de recolección de material granular de relleno para construcción de muro en embalses de relaves.
- xv. Suelo de fundación: Componente del depósito ubicado en la base del muro y que da soporte al mismo. Generalmente está compuesto por roca.

3.1.1.2. Tipos de depósitos

Los tipos de depósitos, según el SERNAGEOMIN [5], son los siguientes: Tranque de relaves, Embalse de relaves, Depósitos de relaves espesados, Depósito de relaves filtrados, Depósitos de relaves en pasta y depósitos de relaves especiales cuyo desarrollo es factible en otros países, no así en Chile debido a las restricciones de la respectiva normativa. Cabe definir los primeros dos sistemas, en primer lugar el tranque de relaves, el cual es un depósito de relaves compuesto por muro resistente construido con la fracción más gruesa del relave, correspondiente a las arenas, mientras que la parte fina (lamas), se almacena en la cubeta. Por otro lado, se tiene los embalses de relaves, que se describen como depósito de relaves compuesto por muro resistente hecho totalmente de material de empréstito, compactado e impermeabilizando el talud interior del muro y también parte o todo su coronamiento; los relaves se depositan completos en la cubeta sin necesidad de clasificación.

3.1.1.3. Métodos constructivos

La construcción de este tipo de obras se basan en tres métodos constructivos: el método aguas arriba, el cual consiste en un muro inicial construido y posterior a este se construye uno nuevo en la zona superior al anterior y así sucesivamente, en dirección hacia la cubeta. Si bien este método es el que requiere un menor volumen de material arenoso, es el que produce el tipo de muro menos resistente frente a oscilaciones sísmicas, es por ello que en la actualidad en Chile, no está contemplado aprobar este tipo de proyecto de tranques de relaves en la legislación vigente [5]. En segundo lugar se tiene el método de eje, que se inicia con un muro de partida para luego construir nuevos muros sobre el de partida, siguiendo el mismo plano vertical inicial de la berma de coronamiento del muro de partida, lo que permite lograr un muro de arenas cuyo eje se mantiene en el mismo plano vertical. Por último se tiene el método de aguas abajo, cuya construcción se inicia también con un muro de partida de material compactado y a medida que aumenta la cantidad de relaves necesaria a depositar en cubeta se va aumentando el volumen del muro en dirección aguas abajo, tal como indica la Figura 3.2. Este método de aguas abajo requiere disponer de un gran volumen de arenas y permite lograr muros resistentes más estables del punto de vista de la resistencia sísmica [5].

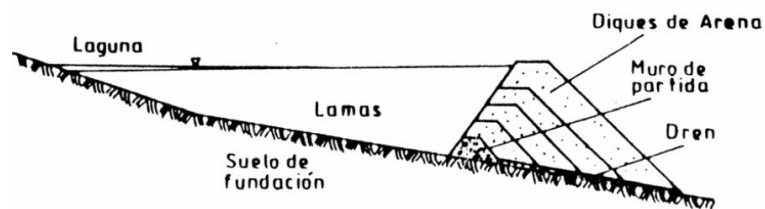


Figura 3.2: Método de aguas abajo

3.1.1.4. Sistemas de control y monitoreo

Los depósitos de relaves requieren de sistemas que regulen, controlen y monitoreen el comportamiento y las respuestas del mismo. Esto de modo de poder enfrentar posibles anomalías y tener alertas respecto al funcionamiento del depósito así como lograr la prevención de los mecanismos de falla al controlar ciertas variables de importancia. En la siguiente Tabla se muestran las variables significativas a considerar y que pueden variar durante la vida útil del depósito. Dichas variables presentan incertidumbre en su obtención, debido a las condiciones que se enfrenta cada variable, como por ejemplo en el nivel freático que depende de condiciones climáticas y situaciones de drenaje, en las aceleraciones del muro por la sismicidad, en la geometría del muro enfocada en el talud por inestabilidad, etc. Esta variabilidad se puede describir mediante mediciones y obteniendo registros estadísticos.

La siguiente lista se realiza en base a documentación referenciada, poniendo énfasis en el documento de Carvajal [6].

Tabla 3.1: Variables a controlar y sistemas de control

Variables	Sistema de control
i. Compactación, densidad	Densímetro nuclear, Densidad relativa
ii. Nivel freático	Piezómetro, Inspección visual
iii. Presión de poros	Piezómetro
iv. Geometría de muro	Inspección visual, Topografía, Fotogrametría
v. Desplazamientos	Celdas de asentamiento, monolitos
vi. Aceleraciones	Acelerómetro
vii. Esfuerzos internos, rigidez	Celdas de presión, SPT, CPT
viii. Sistemas de drenaje y desvío	Inspección visual, análisis hidráulico

3.1.1.5. Parámetros geotécnicos

Los depósitos de relaves son obras en las que el área subde geotecnia cumple un rol primordial, por lo cual para el modelamiento y el diseño de estos se requiere un conocimiento y comprensión de los parámetros geotécnicos involucrados, característicos de este tipo de sistemas, y en general de los sistemas geotécnicos, los cuales se enlistan en la Tabla 3.2, junto con su forma de obtención. Estos también pueden ser seleccionados, teniendo un criterio menos conservador, o más bien poder hacerse una referencia de valores mediante bibliografía referida a correlaciones y fórmulas geotécnicas, sin embargo el criterio empírico-experimental predomina por sobre todo. [7] [8] [9]. Estas variables tienen un grado de incerteza determinado por una variabilidad espacial, así como también por diferencias en la medición y en el uso de instrumentación para la obtención de estos parámetros. Realizando histogramas y distribuciones a partir de registros sobre las variables es posible obtener información útil para describir su comportamiento en mayor detalle y aplicar los métodos de cuantificación de incertidumbre.

Tabla 3.2: Parámetros geotécnicos

Parámetro	Fuente de obtención
i. Densidad, Peso unitario	Densímetro nuclear, Ensayo densidad relativa, Proctor, Otros
ii. Ángulo de fricción interna	Ensayos de laboratorio (Corte directo, triaxial, otros)
iii. Cohesión	Ensayos de laboratorio (Corte directo, triaxial, otros)
iv. Permeabilidad	Sondaje, Ensayos in situ de filtraciones, Otros
v. Plasticidad	Ensayos de laboratorio (Método de cucharas, otros)
vi. Compresibilidad	Ensayo de lab. (Ensayo de consolidación, Ensayo placa de carga.
vii. Módulo de elasticidad	Ensayo placa de carga
viii. Coeficiente de balasto	Ensayo placa de carga
ix. Granulometría	Análisis visual, análisis granulométrico
x. Resistencia al corte no drenada	Ensayos de lab. (Corte directo, triaxial, otros), Sondaje
xi. Coeficiente de presión lateral	Ensayos de lab. (Corte directo, triaxial, otros), Ensayo in situ de presión

3.1.1.6. Mecanismos de falla

Los mecanismos de falla de este tipo de obra son los mencionados a continuación [10] [11].

Tabla 3.3: Mecanismos de falla en embalses de relaves

Mecanismo de falla	Descripción	Causas
i. Inestabilidad de taludes	Falla por exceso de esfuerzos cortantes en taludes	- Control inadecuado de presiones de poros, esfuerzos internos y/o deformaciones
ii. Erosión	Desgaste, degradación o desplazamiento de zonas del material granular	- Inadecuado control de filtraciones - Protección insuficiente de taludes - Condición climática extrema (Lluvia intensa)
iii. Overtopping (Rebalse)	Incremento en el nivel freático hasta sobre el nivel de coronamiento, provocando una revancha nula	- Diseño hidráulico inadecuado - Asentamiento de coronamiento - Condición climática extrema (Lluvia intensa)
iv. Licuefacción	Comportamiento de material granular como líquido, debido a la falta de adherencia o cohesión de las partículas y el movimiento del cuerpo de masa formado por dicho material	- Diseño dinámico inadecuado - Condiciones extremas (Terremoto)
v. Falla de fundación	Falla en elementos de fundación	- Diseño estructural deficiente - Diseño insuficiente de condiciones de suelos

La frecuencia de los mecanismos de falla se verifica en la siguiente figura obtenida del artículo de Oldecop y Rodriguez [12].

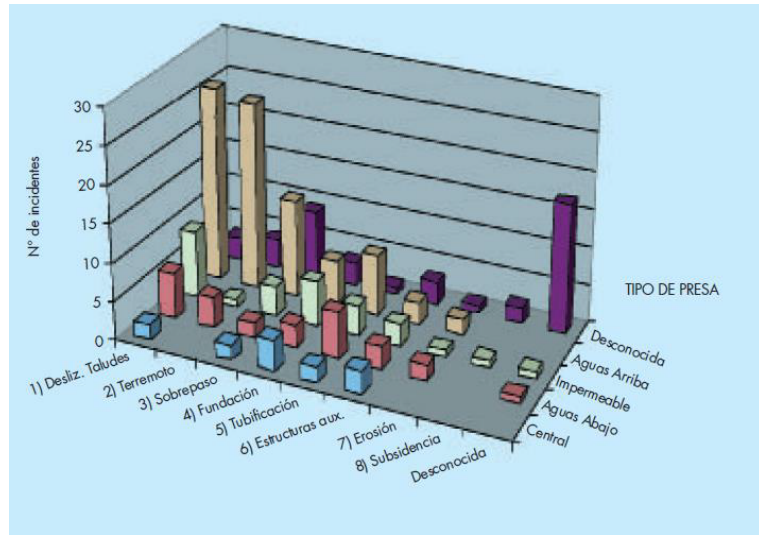


Figura 3.3: Gráfico mecanismos de falla

Mediante la visualización de la figura 3.3 es posible mencionar el mecanismo de falla más típico referido a la estabilidad de taludes y es por esto que los métodos de análisis de un depósito de relaves van enfocados en este mecanismo.

3.1.2. Análisis de depósitos de relaves: Estabilidad física

Existen diferentes estudios para este tipo de obras civiles, los cuales involucran un análisis determinado según el área que corresponda. Estos pueden ser: Análisis de estabilidad física, Análisis hidráulico, Análisis de filtraciones, Análisis sísmico, Análisis químico, Análisis de sistemas de drenaje, Análisis de sistemas de impermeabilización, Análisis estructural, entre otros.

A continuación se describen los tipos de análisis referidos a la estabilidad física de depósitos de relaves, tal como se menciona en el documento de Barrera y Campaña [1].

3.1.2.1. Análisis estático

Este análisis es el más básico e implica un cálculo de estabilidad de taludes del sistema estando en condiciones estáticas, es decir, las solicitaciones están únicamente descritas por el peso propio del cuerpo de masa, provocado a su vez por la gravedad.

3.1.2.2. Análisis pseudoestático

Este análisis conlleva la inclusión de una contribución por solicitaciones sísmicas. El camino más simple para incluir los efectos del sismo es considerar que las fuerzas inducidas por éste son incorporadas como fuerzas horizontales en el análisis. La magnitud de la fuerza horizontal F_h es tomada igual al peso de la masa que desliza, multiplicada por un coeficiente k_h que “representa” al sismo, transformándolo en un problema pseudoestático. Aunque en la realidad el sismo impone fuerzas horizontales y verticales, diversos autores recomiendan no incluir estas últimas en el análisis.

El coeficiente sísmico k_h , es un valor empírico que depende entre otros factores, de la aceleración horizontal máxima de campo libre, las características del sismo (magnitud, profundidad, contenido de frecuencias, etc.), dimensiones de la obra (volumen potencial masa deslizando y condiciones de borde), existencia de poblados, instalaciones y equipos que puedan verse afectados. Diversos autores han propuesto valores y expresiones para determinar el valor de k_h .

3.1.2.3. Análisis dinámico

El análisis dinámico se presenta como el análisis generalmente más complejo y que implica una mayor carga computacional en la ejecución de los procesos de cálculo. Este análisis permite obtener la respuesta del sistema frente a sollicitaciones dinámicas, lo cual involucra la determinación de aceleraciones, velocidades y desplazamientos en función del tiempo, así como la definición de parámetros de amortiguamiento, de un registro de ondas sísmicas y condiciones de borde e iniciales específicas del sistema [13] [11].

3.1.3. Métodos de análisis de estabilidad física

3.1.3.1. Método de equilibrio límite

Según Barrera y Campaña [1] se menciona lo siguiente: estos métodos son los más utilizados en la práctica para el cálculo de estabilidad de taludes y se basan fundamentalmente en una consideración de equilibrio plástico límite. Se han desarrollado una gran variedad de métodos, los cuales en su mayoría se basan en dividir la potencial masa de deslizamiento en un cierto número de dovelas o rebanadas.

Las fuerzas que actúan sobre cada dovela deben encontrarse en equilibrio, tanto para cada dovela, como para la totalidad de la masa deslizando. Algunos métodos consideran sólo equilibrio de fuerzas, otros sólo de momento y otros satisfacen ambas fuerzas.

Las hipótesis del Método de Equilibrio Límite son:

- El suelo sigue una ley de comportamiento Mohr-Coulomb, tal como se indica en la siguiente fórmula y la resistencia se moviliza total y simultáneamente a lo largo de la superficie de deslizamiento

$$\tau = c + \sigma \cdot \tan(\phi') \quad (3.1)$$

- El Factor de Seguridad (FS) de la componente cohesiva y friccionante del material es igual para todos los suelos involucrados en la superficie de deslizamiento y es el mismo para todas las dovelas.
- Se debe verificar el equilibrio de cada dovela como para toda la masa de suelo potencialmente deslizando.

Los métodos de equilibrio límite más comunes son: Felleniuos, Bishop Simplificado, Janbu Simplificado, Janbu Generalizado, Spencer, Morgenstern-Price, GLE, Corps of Engineers y Lowe-Karafiath. De los métodos indicados, el más preciso es el de Morgenstern-Price, pero su comprensión y aplicación es laboriosa.

3.1.3.2. Métodos finitos

Los métodos finitos implican una discretización del sistema u objeto del estudio en el cual la naturaleza continua y de infinitos grados de libertad del mismo es representada por un conjunto finito de elementos cada uno con un número limitado de grados de libertad. La representación discreta del problema implica una descripción precisa de la geometría del sistema, las propiedades del mismo y las condiciones a las cuales está sometido, para lo cual se utilizan elementos geométricos tales como puntos, líneas, polígonos y volúmenes. La preparación del problema tiene como resultado un sistema de ecuaciones que sirve como insumo de entrada para el procesamiento o solución del sistema, el cual se puede realizar por métodos directos o iterativos [14].

Estos métodos se basan en la inversión de una matriz, donde la ecuación matricial puede ser escrita como $[K]D=R$, donde D es un vector de cantidades desconocidas, tales como los valores en los nodos. $[K]$, por otra parte, es un arreglo de valores constantes y conocidos, y R corresponde a un vector de cargas conocidas [15].

- i. Método de diferencias finitas: El método de diferencias finitas se basa en asignar puntos o conjuntos de puntos discreto, espaciados de manera apropiada sobre el dominio y aplicar las ecuaciones constitutivas correspondientes del fenómeno a cada punto [14] [16].
- ii. Método de elementos finitos: En este método se discretiza el dominio en elementos finitos y calcula las propiedades en cada nodo. Las funciones de forma o funciones básicas se utilizan para interpolar, dentro del elemento finito, las soluciones obtenidas en estos nodos [14].
- iii. Método de volúmenes finitos: Este método consiste en dividir el dominio en volúmenes de control y aplicar un balance de conservación a cada volumen. Las propiedades se calculan para cada celda. Está basado en la forma integral de las leyes de conservación y puede manejar discontinuidades en las soluciones.

Realizando una comparación entre el método de elementos finitos y el de volúmenes finitos se puede mencionar que ambos métodos traducen un conjunto de ecuaciones diferenciales en ecuaciones matriciales para cada elemento, relacionando las fuerzas en los nodos con los desplazamientos en los nodos [17]. Sin embargo hay ciertos aspectos en los que difieren, los cuales justifican el uso de este método en FLAC3D, y se describen en la sección 3.1.11.5.

3.1.4. Mecanismo de cálculo y modelamiento en FLAC3D

Se establece esta sección para describir en forma más detallada el mecanismo de cálculo y modelamiento del software a ocupar para la realización del modelo [13].

3.1.4.1. Descripción

El software FLAC3D es un programa tridimensional de enfoque Lagrangiano y explícito de volúmenes finitos para mecánicas computacionales ingenieriles.

3.1.4.2. Formulación matemática

FLAC3D es un programa explícito de volumen finito para estudiar, numéricamente, el comportamiento mecánico de un medio tridimensional continuo a medida que alcanza el equilibrio o el flujo plástico estable.

La mecánica del medio se deriva de principios generales (definición de deformación, leyes de movimiento) y el uso de ecuaciones constitutivas que definen el material idealizado. La expresión matemática resultante es un conjunto de ecuaciones diferenciales parciales, que relacionan variables mecánicas (tensión) y cinemáticas (tasa de deformación, velocidad), que deben resolverse para geometrías y propiedades particulares, dados los límites específicos y las condiciones iniciales.

Un aspecto importante del modelo es la inclusión de las ecuaciones de movimiento, aunque FLAC3D se ocupa principalmente del estado de tensión y deformación del medio cerca del estado de equilibrio. La sección de implementación numérica muestra que los términos inerciales se utilizan como un medio para alcanzar el estado de equilibrio de una manera numéricamente estable. El método de solución en FLAC3D se caracteriza por los siguientes tres enfoques:

- i. Enfoque de volumen finito (las derivadas de espacio y tiempo de primer orden de una variable se aproximan mediante volúmenes finitos asumiendo variaciones lineales de la variable en intervalos de tiempo y espacio finitos, respectivamente)
- ii. Enfoque de modelo discreto (el medio continuo se reemplaza por un equivalente discreto, uno en el que todas las fuerzas involucradas se concentran en los nodos de una malla tridimensional utilizada en la representación del medio)
- iii. Enfoque de solución dinámica (Los términos inerciales en las ecuaciones de movimiento se utilizan como medios numéricos para alcanzar el estado de equilibrio del sistema en consideración)

Las leyes del movimiento para el continuo se transforman, mediante estos enfoques, en formas discretas de la ley de Newton en los nodos. El sistema resultante de ecuaciones diferenciales ordinarias se resuelve numéricamente utilizando un enfoque explícito de diferencias finitas en el tiempo.

Las derivadas espaciales involucradas en la derivación del medio equivalente son las que aparecen en la definición de tasas de deformación en términos de velocidades. Con el fin de

definir las variaciones de velocidad y los intervalos espaciales correspondientes, el medio se discretiza en elementos de tasa de deformación constante de forma tetraédrica cuyos vértices son los nodos de la malla mencionada anteriormente.

3.1.4.3. Representación de elementos

Los materiales están representados por elementos poliédricos dentro de una cuadrícula tridimensional que el usuario ajusta para adaptarse a la forma del objeto a modelar. Cada elemento se comporta de acuerdo con una ley de tensión / deformación lineal o no lineal prescrita en respuesta a las fuerzas aplicadas o restricciones en los límites. El material puede ceder y fluir, y la rejilla puede deformarse (en modo de gran deformación) y moverse con el material que se representa. Un ejemplo de un elemento con forma de tetraedro se presenta en la siguiente Figura:

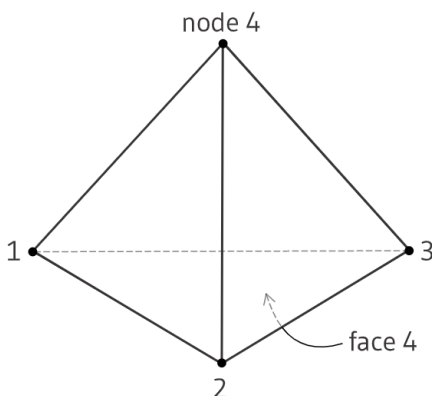


Figura 3.4: Elemento - Tetraedro

En este elemento se puede identificar algunos componentes, como las caras, los nodos o gridpoints y las zonas. Estas últimas representan el elemento en sí. El elemento puede adquirir también forma de hexaedro, el cual es el que se aplica para este trabajo.

3.1.4.4. Resolución de equilibrio y factor de seguridad

FLAC3D usa un procedimiento iterativo informado por la física para alcanzar el equilibrio estático. Cuando nos referimos a estar en equilibrio, nos referimos a lo mismo que estar completamente convergido a una solución estática. FLAC3D proporciona varias formas diferentes de medir y visualizar qué tan cerca se está de converger o haber alcanzado un estado de equilibrio en su modelo. Generalmente el criterio utilizado es en base al equilibrio de fuerzas. Al realizar estos procedimientos, los resultados quedan descritos en cada elemento mediante los tensores de esfuerzos y deformaciones, que se denotan como:

$$\sigma_{i,j} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} \quad \varepsilon_{i,j} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix} \quad (3.2)$$

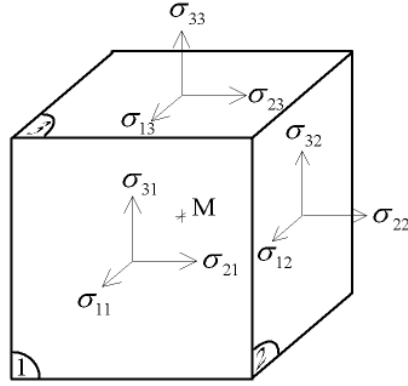


Figura 3.5: Diagrama - Tensor de esfuerzos

El factor de seguridad se calcula con el método de reducción de esfuerzos (strength reduction technique), mediante la reducción progresiva de la resistencia al corte del material para llevar el talud a un estado de equilibrio límite. El método se aplica comúnmente con el criterio de falla de Mohr-Coulomb. En este caso, el factor de seguridad se define de acuerdo con las ecuaciones:

$$c^{trial} = \frac{c}{FS^{trial}} \quad (3.3)$$

$$\phi^{trial} = \arctan\left(\frac{\tan(\phi)}{FS^{trial}}\right) \quad (3.4)$$

Se realizan una serie de simulaciones utilizando valores de prueba de FS^{trial} , variando según la discretización que se seleccione, para reducir la cohesión c y el ángulo de fricción ϕ hasta que ocurra la falla del talud y así establecer el factor de seguridad en ese paso.

El procedimiento conlleva además el trazado de los contornos del factor de seguridad, lo cual permite representar las zonas asociadas a un respectivo factor de seguridad. Estos son asociados a una determinada zona, luego de realizar la resolución del equilibrio frente a ese respectivo factor de seguridad y habiendo reducido los parámetros resistentes que se definieron previamente. De este modo las zonas que cumplan con el equilibrio quedan etiquetadas con el factor de seguridad utilizado, y así sucesivamente.

Se establecen las siguientes definiciones respecto a los factores de seguridad, para mayor facilidad en la explicación. Se denomina factor de seguridad 'global' (FSG) al mínimo factor de seguridad que provoca la estabilidad, luego del desarrollo del proceso de reducción de esfuerzos. Por otro lado, los factores de seguridad 'locales' (FSL) representan los factores de seguridad asociados a cada etapa de reducción de esfuerzos.

3.1.4.5. Argumento de uso de método

Se utiliza el método de volúmenes finitos y a continuación se menciona el argumento de su uso, establecido en la documentación oficial de software de Itasca, FLAC3D [13].

- El esquema de “discretización mixta” (Marti y Cundall. 1982) se utiliza para modelar con precisión las cargas de colapso plástico y el flujo de plástico. Se cree que este esquema es físicamente más justificable que el esquema de “integración reducida” comúnmente utilizado con elementos finitos.
- Se utilizan las ecuaciones dinámicas de movimiento completas, incluso cuando se modelan sistemas que son esencialmente estáticos. Esto permite que FLAC3D siga procesos físicamente inestables sin problemas numéricos. El enfoque para proporcionar una solución estática en el tiempo se analiza en la definición de solución estática.
- Se utiliza un esquema de solución explícito (en contraste con los métodos implícitos más habituales). Los esquemas explícitos pueden seguir la no linealidad arbitraria en las leyes de tensión / deformación casi al mismo tiempo que las leyes lineales, mientras que las soluciones implícitas pueden tardar mucho más en resolver problemas no lineales.

Además, no es necesario almacenar matrices, lo que significa que: 1) se puede modelar un gran número de elementos con un modesto requisito de memoria; y 2) una simulación de deformación grande apenas requiere más tiempo que una ejecución de deformación pequeña, porque no hay una matriz de rigidez que actualizar.

- FLAC3D es robusto en el sentido de que puede manejar cualquier modelo constitutivo sin ajuste al algoritmo de la solución; muchos códigos de elementos finitos necesitan diferentes técnicas de solución para diferentes modelos constitutivos.
- FLAC3D utiliza una formulación incremental en casi todas las áreas. Esto significa que los desplazamientos no están directamente relacionados con las tensiones y se pueden cambiar en cualquier momento, y que las propiedades del material pueden cambiar sin afectar el estado de tensión actual.

3.1.4.6. Modelamiento

La etapa de modelamiento, enfocado en el planteamiento de un sistema geotécnico del tipo embalse de relaves se describe en el capítulo 4 de forma aplicada en la realización del modelo.

3.1.5. Comentarios sobre revisión bibliográfica

Los análisis de los depósitos de relaves han adquirido suma importancia a lo largo del tiempo, esto es mencionado de forma bastante completa y adecuada en el documento de Barrera y Campaña [1], en el cual se presenta la evolución de los métodos de análisis de estabilidad física. Según se menciona en el documento descrito, y como se menciona en Marco Teórico, los métodos de cálculo de estabilidad física han sido el método de equilibrio límite y métodos finitos, siendo estos últimos los que han predominado en la actualidad. Es por

esto que su aplicación es justificada en el desarrollo de este documento, y el uso de FLAC3D, motivado por la empresa colaborante, involucra este tipo de procedimientos de cálculo.

Respecto a los trabajos realizados en los que se emplea el uso de software de modelamiento, que permiten modelar sistemas de embalse de relaves, se presentan metodologías que involucran el uso de FLAC3D, Plaxis, RS2, entre otros. Estos conllevan un análisis de deformaciones, esfuerzos, presiones de poros y el cálculo de factores de seguridad, el cual es el enfoque del presente trabajo. Cabe mencionar que también el enfoque está puesto en los contornos de factor de seguridad, los cuales se puede decir que es uno de los aspectos innovadores, ya que no es identificado el análisis de este concepto en la revisión bibliográfica. Por otro lado, se presentan análisis estáticos, pseudoestáticos y dinámicos, y en el presente documento se realiza un enfoque en el análisis estático básicamente, esto ya que se plantea un caso de modelamiento básico, estándar y referencial para utilizar en la obtención de una base de datos que servirá para los procedimiento consecutivos que se verán a lo largo de la memoria. Sin embargo, se presenta esta metodología con capacidad de adaptabilidad frente a los demás análisis mencionados.

3.2. Modelamiento Sustituto

3.2.1. Modelos sustitutos

El análisis de problemas complejos, tal como el sistema geotécnico presentado en este informe, son resueltos mediante procedimientos numéricos de un modo bastante preciso y confiable, debido a la solidez de la teoría matemática aplicada. Sin embargo, el costo implicado es muy alto, referido principalmente al computacional y temporal, lo que se traduce en una respuesta lenta y no eficiente. Es debido a esto que la búsqueda y ejecución de procedimientos de análisis alternativos es de suma importancia. De este modo es que surgen los métodos sustitutos o meta modelos, para solucionar esta problemática. En base al documento de Yaochu [2], se identifica y comprende la evolución de este tipo de modelos, que surgen debido a las necesidades y desafíos que van presentándose en el campo científico e ingenieril.

El modelo sustituto se crea mediante la implementación de métodos estadísticos o de Machine Learning. Para el presente caso se utiliza el método Kriging o regresión en procesos Gaussianos, el cual se describe a lo largo de esta sección. El modelo recurre a un proceso de entrenamiento, que consiste en la utilización de una base de datos otorgada por un registro de ejecuciones o casos planteados en el modelo original (modelo realizado en FLAC3D). A partir de este se realiza un proceso regresivo y los datos de esta base se interpolan al ingresarse en el algoritmo. De esta forma se puede obtener un modelo que realice predicciones del modelo original, al realizar interpolaciones de los puntos de soporte (base de datos) y obteniendo resultados para datos intermedios, lo que permite una generalización de la respuesta del modelo original.

El modelo sustituto logra establecer el contexto adecuado para la implementación del área de estudio de la cuantificación de incertidumbre. Una respuesta rápida y ágil se hace necesaria para lograr el objetivo de implementar los métodos de cuantificación de incertidumbre, donde se requiere una cantidad considerablemente alta de ejecuciones del modelo, debido principalmente a los fundamentos del método de Montecarlo, el que es usado para el uso de los métodos mencionados y será definido en una próxima sección.

Es necesario entender conceptos como ciencia de datos, minería de datos, métodos estadísticos o de Machine Learning, análisis predictivo, así como conceptos básicos de probabilidad y estadística que se irán describiendo a lo largo de este informe.

3.2.2. Métodos predictivos

Se describen a continuación los métodos estadísticos y de machine learning más comunes (Tabla 3.4) [18][19]. Estos pueden ser de tres tipos:

- i. Regresión: Tipo de algoritmo cuyo objetivo es establecer un método para la relación entre un cierto número de características y una variable objetivo continua. En términos de machine learning consiste en un algoritmo de aprendizaje supervisado.
- ii. Clasificación: Tipo de algoritmo que se usa cuando el resultado a obtener es una etiqueta discreta. En términos de machine learning consiste en un algoritmo de aprendizaje

supervisado.

- iii. Clustering (Agrupamiento): Algoritmo de agrupación cuyo objetivo es clasificar un conjunto de datos en un clúster (o grupo) específico. En términos de machine learning consiste en un algoritmo de aprendizaje no supervisado.

Tabla 3.4: Métodos predictivos

Método	Descripción	Categoría
i. Regresión lineal	- Ajuste de datos con tendencia lineal	Regresión
ii. Regresión polinomial	Ajuste de datos con tendencia no lineal	Regresión
iii. Soporte de vectores	Ajuste de datos mediante la creación de hiperplano	Regresión, Clasificación
iv. Árbol de decisión	Algoritmo que mediante nodos y condiciones caracteriza un determinado sistema	Regresión, Clasificación
v. Bosques aleatorios	Conjunto de árboles de decisión	Regresión, Clasificación
vi. Regresión logística	Algoritmo basado en correlación con puntos cercanos	Regresión, Clasificación
vii. K-vecinos más próximos	Algoritmo basado en correlación con puntos cercanos	Clasificación
viii. Agrupamiento de K-medias (K-mean clustering)	Algoritmo de agrupación mediante condición de K	Clustering (Agrupamiento)
ix. Agrupamiento jerárquico (Hierarchical clustering)	Algoritmo de agrupación mediante métodos aglomerativos	Clustering (Agrupamiento)
x. Redes Neuronales	Algoritmo que crea un sistema de interconexión en capas de neuronas artificiales que colaboran para procesar datos de entrada y generar salidas	Regresión, Clasificación
xi. Kriging (o regresión en procesos Gaussianos)	Algoritmo que implementa procesos Gaussianos para regresión	Regresión

3.2.3. Método Kriging o regresión en procesos Gaussianos

En la presente sección se describe el método a utilizar en el trabajo, así como sus principios, características y propiedades.

3.2.3.1. Conceptos previos

Se procede a describir algunos conceptos previos respecto al método Kriging o regresión en procesos Gaussianos [19] [20] [21] [22]

- i. Procesos Gaussianos: Un proceso Gaussiano es una colección de variables aleatorias, que cumplen que cualquier subconjunto finito de la colección, tiene una distribución Gaussiana conjunta (Ver Anexo A.1).
- ii. Función base: Funciones que permiten proyectar los datos de entrada en un espacio de determinada dimensión. Por ejemplo, una entrada escalar x podría proyectarse en el espacio de potencias de x : $x = (1, x, x^2, x^3, \dots)$ para implementar la regresión polinomial, y de este modo ajustar una regresión no lineal, adaptándola y asimilándola hacia una función lineal. Estas funciones se ponderan con los pesos.
- iii. Pesos: Parámetros que ponderan la función base. Estos son obtenidos mediante la resolución de un problema de optimización en base a mínimos cuadrados, lo cual se describe más adelante.
- iv. Función de covarianza (Kernel): Función que indica el grado de variación conjunta de dos variables aleatorias respecto a sus medias. Se expresa como matriz al haber un grado mayor de dimensiones. Contiene los hiperparámetros, los cuales son los parámetros a calibrar mediante el proceso de optimización. Induce las propiedades del proceso de regresión, tales como la suavidad, la escala de los datos, la periodicidad, entre otros. Mayor información sobre las distintas funciones de covarianza o kernel se presentan en el documento de Rasmussen y Williams [21].
- v. Hiperparámetros (θ): Parámetros que definen la función de covarianza o kernel. Se deben calibrar mediante un proceso de optimización.
- vi. Optimizador: Sistema que realiza la optimización de una función en particular (búsqueda de mínimos y máximos). En el caso a analizar, la optimización se realiza con el objetivo de calibrar los hiperparámetros. Por otro lado también se realiza un proceso de optimización para la obtención de los pesos por mínimos cuadrados, sin embargo el optimizador de hiperparámetros es el que se busca establecer de modo más detallado.

3.2.3.2. Definición y descripción del método

Kriging (también conocida como regresión del proceso Gaussiano) es un método bayesiano que tiene como objetivo general el aproximar funciones capaces de predecir determinados valores, o dicho de otro modo realizar una regresión para obtener un predictor, a partir de procesos Gaussianos y una base de datos preestablecida [20] [21] [23].

Su formulación se presenta como sigue.

$$y = f(x) + e \quad (3.5)$$

$$e \sim GP(0, k(x, x')) \quad (3.6)$$

donde $f(x)$ es la función a determinar y e es el error o ruido que tiene una distribución de tipo Gaussiana con media cero y kernel a determinar. Para el presente problema se determina un error igual a 0, debido a que este no se considera en este punto, principalmente por las características de la base de datos.

Dicho esto la función $f(x)$ se describe como:

$$f(x) = h(x)^T \cdot \beta + g(x) \quad (3.7)$$

$$g(x) \sim GP(0, k(x, x')) \quad (3.8)$$

donde $h(x)$ es la función base, β es la representación de los pesos y $g(x)$ es un proceso Gaussiano con media cero y función de covarianza o kernel a determinar.

De este modo la función $f(x)$ es un proceso Gaussiano expresado a continuación [24] [25]:

$$f(x) \sim GP(m(\cdot), k(\cdot, \cdot)) \quad (3.9)$$

donde m es la media y k es el kernel, y esta ecuación equivale a

$$f(x) \sim GP(F \cdot B, \sigma^2 \cdot R) \quad (3.10)$$

donde F es igual al conjunto de funciones base o matriz de regresión, B el vector de los pesos y R es la matriz de covarianza determinada por la función de covarianza o kernel, con σ^2 como la respectiva varianza.

La función mencionada entonces requiere de la definición de una función base y una función de covarianza, para luego entrenar el modelo a partir de una base de datos preestablecida.

3.2.3.3. Preparación de base de datos

La base de datos que se usan para entrenamiento del modelo puede provenir de diversas fuentes, tales como mediciones, datos experimentales, registros estadísticos, o de software capaces de realizar modelaciones o simulaciones. Se hará énfasis en esta última fuente, ya que será por la cual se obtenga la base de datos del presente trabajo.

La obtención a partir de esta fuente se realiza de modo de realizar ejecuciones del programa de un modelo en particular preestablecido. Este modelo recibe datos de entrada de cierta dimensión y entrega otra cantidad de datos de salida. Estos datos de entrada y salida establecen la base de datos a utilizar para la confección del modelo sustituto.

Se debe profundizar en ciertos aspectos, como lo es la selección de los datos de entrada. Estos deben ser elegidos de forma adecuada, para que se pueda abarcar la mayor cantidad de casos y una mayor representatividad, y así lograr una mejor efectividad del modelo sustituto. Para su selección se establece entonces el método de "Latin Hypercube Sampling".

Latin Hypercube Sampling (LHS) es un método de muestreo que permite cubrir mejor el dominio de variaciones de las variables de entrada, gracias a una estrategia de muestreo estratificado. Este método es aplicable en el caso de variables de entrada independientes. El procedimiento de muestreo se basa en dividir el rango de cada variable en varios intervalos de igual probabilidad. El muestreo se realiza de la siguiente manera:

- i. El rango de cada variable de entrada se estratifica en celdas isoprobabilísticas
- ii. Una celda se elige uniformemente entre todas las celdas disponibles
- iii. El número aleatorio se obtiene invirtiendo la función de densidad acumulada localmente en la celda elegida
- iv. Todas las celdas que tienen un estrato común con la celda anterior se separan de la lista de celdas disponibles

Por otro lado, en ciertas ocasiones es necesario realizar un proceso de normalización de los datos de modo de evitar complejidades numéricas. Esta se plantea modificando los datos según las siguientes fórmulas.

$$x_{i,d} = \frac{(x_{i,d} - \mu_{x,d})}{\sigma_{x,d}} \quad y_{i,d} = \frac{(y_{i,d} - \mu_{y,d})}{\sigma_{y,d}} \quad (3.11)$$

donde i es el número de muestra, d el indicador de la dimensión que corresponda, μ_d es el promedio por dimensión y σ_d es la desviación estándar por dimensión.

3.2.3.4. Entrenamiento y optimización del modelo

El proceso de entrenamiento consiste en la utilización de datos previamente establecidos, denominados datos de entrenamiento, definido como:

$$D = \{(x_i, y_i) \mid i=1, \dots, n\}$$

Estos son utilizados en la función $f(x)$ de modo de que pasan por la función base y la función de covarianza preestablecidos, es decir, se ingresan estos valores de entrenamiento y se calculan los pesos mediante un proceso de regresión y optimización y se calculan los coeficientes asociados a la función de covarianza que conforma una matriz, llamada matriz de covarianza, en la que se relacionan todos los valores del conjunto X de la base de datos de entrenamiento. Este último proceso conlleva al mismo tiempo un proceso de optimización de los parámetros característicos o hiperparámetros, referidos a la respectiva función de covarianza o kernel. La representación de estos procesos se pueden ver en mayor detalle en los documentos de Morales [26] y Parra [27], donde se visualiza los procesos para un caso univariable en datos de entrada y salida, multivariable en datos de entrada y multivariable en datos de salida.

Estos procesos pueden ordenarse según el siguiente formato, representado en el diagrama de la Figura 3.6.

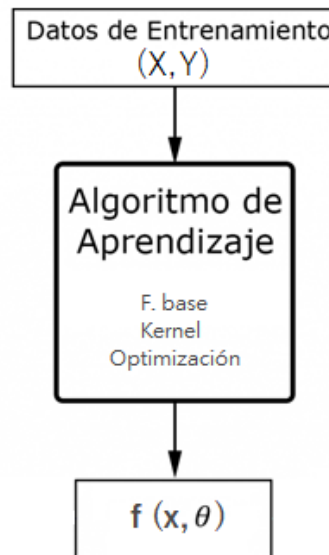


Figura 3.6: Proceso de entrenamiento de modelo sustituto

El proceso de obtención de los pesos se basa en el planteamiento de la ponderación y optimización en base a mínimos cuadrados. Se plantea un problema de optimización en el cual se adapta el de mínimos cuadrados, de modo de realizar una maximización de likelihood, con el objetivo de obtener los pesos, tal como lo indica el documento de Dubourg [24]. El proceso de maximización de likelihood está relacionada con la reducción del error entre el

valor real y el predicho, y esta expresión queda definida como la distribución de probabilidad de las observaciones o datos de entrenamiento, dados los valores β , σ y θ , siguiendo una distribución Gaussiana multivariada.

Así se obtiene el valor de β óptimo:

$$\hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} y \quad (3.12)$$

donde y es el conjunto de valores de salida de entrenamiento.

Además se obtiene σ^2 óptimo ($\hat{\sigma}^2$), en función de m , que es número de muestras o largo de datos de entrenamiento, que servirá más adelante. Detalles en documento de Dubourg [24].

El proceso de optimización de los hiperparámetros sigue la siguiente estructura basada en la maximización del likelihood. La formulación y desarrollo del problema se presenta a continuación [24].

El likelihood mencionado se escribe como sigue:

$$L(y|\beta, \sigma^2\theta) = \frac{1}{((2\pi\sigma^2)^m [\det R(\theta)])^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(y - F\beta)^T R(\theta)^{-1}(y - F\beta)\right] \quad (3.13)$$

Y se presenta el logaritmo opuesto de esta expresión, denominado log-likelihood opuesto, lo cual transforma este problema de optimización en uno de minimización:

$$-\log(L(y|\beta, \sigma^2\theta)) = \frac{1}{2\sigma^2}(y - F\beta)^T R(\theta)^{-1}(y - F\beta) + \frac{m}{2}\log(2\pi) + \frac{m}{2}\log(\sigma^2) + \frac{1}{2}\log([\det R(\theta)]) \quad (3.14)$$

En desarrollos anteriores se obtuvo $\hat{\beta}$ y $\hat{\sigma}$, los cuales son valores obtenidos de la optimización por mínimos cuadrados. Estos valores se reemplazan en la ecuación del likelihood y de este modo la expresión queda dependiendo solo de los hiperparámetros θ [24].

$$-\log(L(y|\beta, \sigma^2\theta)) = \frac{m}{2}\log(\mathcal{L}(\theta)) + \frac{m}{2}(\log(2\pi) + 1) \quad (3.15)$$

$$\mathcal{L}(\theta) = \hat{\sigma}^2(\theta)[\det R(\theta)]^{1/m} \quad (3.16)$$

Donde $\mathcal{L}(\theta)$ es la función de likelihood reducido (reduced likelihood function) la cual se debe minimizar, y es equivalente a minimizar la expresión general del log-likelihood negativo, ya que lo demás es constante. Se busca el valor que logra este cometido, denotado por:

$$\hat{\theta} = \operatorname{argmin}_{\theta}(\mathcal{L}(\theta)) \quad (3.17)$$

Este valor óptimo representa el estimador de máxima verosimilitud (MLE, maximum likelihood estimate), que se centra en maximizar el likelihood [28].

Existen diferentes métodos de optimización, los cuales se presentan en mayor detalle en el documento de Rasmussen & Williams [21].

3.2.3.5. Predicción

El proceso de predicción se realiza en base a la siguiente formulación del método Kriging o regresión en procesos Gaussianos [20] [21] [24] [25], el cual es caracterizado como el mejor predictor lineal insesgado.

Si la función f es un proceso Gaussiano, entonces presenta una distribución Gaussiana, y se asume que el valor f^* a predecir también es parte de dicho proceso Gaussiano, por lo que presenta la respectiva distribución normal. La función f se denota en la siguiente ecuación.

$$f \sim N(m(\cdot), k(\cdot, \cdot)) \quad (3.18)$$

$$f \sim N(F(X)\hat{\beta}, \hat{\sigma}^2 R(X, \hat{\theta})) \quad (3.19)$$

Entonces es posible expresar la siguiente distribución compuesta de tipo Gaussiana (Anexo A.1):

$$\begin{Bmatrix} f \\ f^* \end{Bmatrix} \sim N \left(\begin{Bmatrix} F(X)\hat{\beta} \\ h(x)^T \hat{\beta} \end{Bmatrix}, \hat{\sigma}^2 \begin{Bmatrix} R(X, \hat{\theta}) & r(x, X, \hat{\theta}) \\ r(x, X, \hat{\theta}) & 1 \end{Bmatrix} \right) \quad (3.20)$$

donde $r(x)$ es el vector de correlación entre la variables de entrada a predecir y de las observaciones, $\hat{\beta}$ es en este caso el vector de los pesos óptimos (también puede ser denotado como \hat{B}) y $\hat{\sigma}^2$ es el valor de la varianza óptimo.

La propiedad de condicionalidad (Anexo A), permite la obtención de la siguiente formulación ('y' representa los valores Y de la base de datos):

$$f^* | X, Y, x, \hat{\theta} \sim N(\mu^*, \sigma^{2*}) \quad (3.21)$$

$$\mu^* = h(x)^T \hat{\beta} + r^T \gamma \quad (3.22)$$

$$\sigma^{2*} = \hat{\sigma}^2 [1 - r^T R^{-1} r + u^T (F^T R^{-1} F)^{-1} u] \quad (3.23)$$

$$\gamma = R^{-1} (y - F \hat{\beta}) \quad (3.24)$$

$$u = F^T R^{-1} r - h(x) \quad (3.25)$$

Esta expresión permite realizar la predicción y obtener el valor de μ y σ^2 de f^* en base al conocimiento de un conjunto de datos de entrenamiento (X,Y) , un nuevo conjunto de datos de entrada (x) y los hiperparámetros optimizados $(\hat{\theta})$.

3.2.3.6. Validación de modelo

3.2.3.6.1. Validación cruzada

Es necesario validar la precisión del modelo, es decir verificar que las predicciones que se realizarán se acerquen a los resultados reales.

El método a utilizar se denomina validación cruzada [29]. Se puede visualizar el diagrama de la Figura 3.7 que representa su estructura.



Figura 3.7: Proceso de validación cruzada

Este actúa de modo que una vez realizado el proceso de entrenamiento y la optimización, el modelo queda definido por los parámetros característicos (hiperparámetros), y ahora a partir de este, se reemplazan los datos de entrenamiento menos uno, sin realizar la optimización, para luego evaluar el dato de entrenamiento excluido y así obtener un valor predicho y compararlo con el valor real de ese mismo dato. De este modo se obtiene las siguientes expresiones que definen los errores entre el valor real y la predicción [29]:

$$Error\ residual_i = \Delta_i = f(x)_i - \hat{f}(x)_i \quad (3.26)$$

El cual generalmente se utiliza en su valor absoluto, obteniendo el error residual absoluto:

$$Error\ residual\ absoluto_i = \Delta_{abs_i} = |f(x)_i - \hat{f}(x)_i| \quad (3.27)$$

Los que permiten obtener un promedio, denominado error residual medio absoluto o simplemente error medio absoluto:

$$Error\ absoluto\ promedio = \frac{1}{N} \sum_i^N \Delta_{abs_i} \quad (3.28)$$

Es posible obtener también el error cuadrático:

$$Error\ residual\ cuadrático_i = \Delta_i^2 = (f(x)_i - \hat{f}(x)_i)^2 \quad (3.29)$$

y un parámetro denominado raíz del error residual cuadrático medio o simplemente raíz del error cuadrático medio (RMSE):

$$Error\ cuadrático\ promedio = \sqrt{\frac{1}{N} \sum_i^N \Delta_i^2} \quad (3.30)$$

Por otra parte, se tiene el valor del coeficiente de determinación, o mejor conocido como R cuadrado o R^2 :

$$R^2 = \frac{\frac{1}{N} \sum_i^N \Delta_i^2}{Var(Y)} \quad (3.31)$$

Además, se tiene un error de validación cruzada, junto con su respectiva normalización [20] [29]:

$$Error_{LOO} = \frac{1}{N} \sum_i^N \Delta_i^2 \quad (3.32)$$

$$Error\ normalizado_{LOO} = \varepsilon_{LOO} = \frac{Error_{LOO}}{Cov(Y)} \quad (3.33)$$

donde

$$Cov(Y) = \frac{1}{N-1} \sum_i^N (y_i - \hat{Y})^2 \quad (3.34)$$

$$\hat{Y} = \frac{1}{N} \sum_i^N y_i \quad (3.35)$$

donde $f(x)$ representa el valor real, $\hat{f}(x)$ el valor predicho, \hat{Y} es el promedio de Y, es decir, los valores de la variable de salida en espacio físico, y los valores de 'y' son los del mismo vector. El $error_{LOO}$, es el error por leave one out, es decir el método de validación cruzada por dejar uno fuera, el cual es equivalente al mencionado en el principio de esta sección.

Cabe mencionar que para el caso que se normalicen los valores previamente al proceso de entrenamiento, entonces el error será respecto a esos nuevos valores normalizados, por lo que

para obtener el error (residual) de los valores originales, o en espacio físico, se debe hacer el siguiente ajuste:

$$Error\ nuevo_i = \Delta^* \approx \Delta_i * \sigma \quad (3.36)$$

esto dado por:

$$\Delta_i = f(x)_i - \hat{f}(x)_i \quad (3.37)$$

haciendo el proceso inverso a la normalización:

$$(f(x)_i * \sigma + \mu) - (\hat{f}(x)_i * \sigma + \mu) \quad (3.38)$$

$$f(x)_i * \sigma - \hat{f}(x)_i * \sigma \quad (3.39)$$

$$(f(x)_i * -\hat{f}(x)_i) * \sigma \quad (3.40)$$

$$Error\ nuevo_i = \Delta^* \approx \Delta_i * \sigma \quad (3.41)$$

Este procedimiento se realiza para cada dato de entrenamiento y se obtiene una lista de errores, así como un promedio general de estos, para cada salida involucrada en el sistema. Es importante recalcar que este se realiza sin considerar los datos de soporte que tienen valores extremos del rango de las variables, sin embargo para este trabajo si se consideran, lo cual será discutido en la sección de análisis y discusión de resultados. Por último, μ y σ son el promedio y desviación estándar de los datos en espacio físico, respectivamente.

3.2.3.6.2. Validación de datos de prueba o test

Existe un proceso de validación adicional, el cual consiste en la prueba de datos no considerados en el modelo, es decir, datos que no pertenecan a la base de datos de entrenamiento. Estos deben encontrarse dentro del rango considerado o espacio en el que viven las variables o debe existir un criterio razonable para la determinación de sus valores. Son nombrados como valores de prueba o test. Su error se establece del siguiente modo:

$$Error\ test_i = \Delta\ test^* \approx \Delta\ test_i * \sigma \quad (3.42)$$

esto dado por:

$$\Delta\ test_i = f(x)_i - \hat{f}(x)_i \quad (3.43)$$

Los demás errores previamente descritos para la validación cruzada, son aplicables de forma análoga a la formulación referida a los errores de prueba o test.

3.2.3.7. Características del método

En esta sección se presentan las características del modelo que reflejan la justificación de su uso, basados en la documentación de scikit-learn [19] y el documento de Adapt [24].

- i. Interpolante: La predicción interpola las observaciones.
- ii. Predicción es probabilística: Esto implica que uno puede calcular los intervalos de confianza empíricos y decidir basándose en ellos si debe reajustar la predicción.
- iii. Adaptabilidad y versatilidad: Es posible ajustar y determinar las funciones base y funciones de covarianza o kernels.
- iv. Uso para cantidades limitadas de datos: Su uso es eficiente y presenta un adecuado comportamiento con una cantidad limitada de datos. Por otro lado se debe tener cuidado con la dimensionalidad de estos, ya que con su aumento se produce una pérdida de eficiencia.
- v. Mejor predictor lineal insesgado: El documento de Dubourg [24] define de este modo al predictor, con una descripción de subcaracterísticas aplicadas a un contexto lineal.
- vi. Suavizador lineal: El método y en general los procesos Gaussianos cumplen un rol adicional como suavizador lineal de funciones [21] [23].
- vii. Precisión en puntos de soporte (o de entrenamiento): El método permite otorgar alta precisión en los puntos de soporte, es decir, se tiene un error bajo referido al valor real y predicho de un punto que es parte de la base de datos utilizada en el proceso de entrenamiento. Esto funciona de forma adecuada cuando se tiene una dimensionalidad baja en la salida (para múltiples salidas disminuye esta precisión).

3.2.4. Comentarios sobre revisión bibliográfica

El desarrollo de modelos sustitutos tiene múltiples utilidades y esta área de estudio está relacionada con el área referida a metodologías estadísticas de análisis predictivo y a fundamentos de inteligencia artificial y Machine Learning (o aprendizaje automático), tal como se ha mencionado en la presente sección. El proceso de validación de este tipo de modelos es uno de los más importantes para su confección, ya que define la fidelidad y los errores asociados a su respuesta en comparación al modelo original, por lo que se debe poner énfasis en ello y considerar el costo computacional en su procesamiento, el cual conlleva cerca de un 80 % del tiempo total de la confección.

Debido a la gran magnitud de las áreas de estudio involucradas, es que se ha reducido el espectro de trabajos, realizando un enfoque en los documentos referidos a la aplicación de métodos estadísticos o predictivos en sistemas de carácter geotécnicos, y siendo más específicos, que aborden sistemas de embalses de relaves. Se identifican documentos que utilizan diferentes métodos, tales como equilibrio límite [30], el cual en términos simples implica una menor carga computacional y otro nivel de precisión, generalmente de menor grado que los

métodos finitos, para las cuales existe documentación referida. Estos trabajos se presentan con objetivos como: aplicar una variabilidad espacial de propiedades en las zonas que conforman el embalse de relaves en base a un método predictivo que permita interrelacionar dichas características a lo largo de la geometría del embalse [31] [32]; también la determinación de probabilidad de falla, la cual se establece en la sección de Cuantificación de Incertidumbre del presente trabajo; el cálculo de factores de seguridad, entre otros. En estos trabajos se utilizan distintos métodos estadísticos o predictivos, tales como las redes neuronales [33], el método de superficie de respuesta y procesos Gaussianos [30]. Se identifica el uso del método Kriging o regresión en procesos Gaussianos en sistemas geotécnicos en múltiples fuentes, con distintos usos. Su uso es fundamentado verificando sus características ya mencionadas en Marco Teórico, específicamente en la sección 3.2.3.7.

Si se analizan las fuentes, el documento de Sung [33] demuestra un procedimiento similar al planteado para el presente trabajo de memoria. En dicho texto se establece el análisis de un modelo de embalse de relaves en FLAC, junto con el planteamiento de una metodología predictiva basada en redes neuronales y la aplicación de métodos probabilísticos asociados a incertidumbre. Las diferencias asociadas al presente trabajo son variadas: en primer lugar, el método predictivo que se plantea es diferente, dándose el caso de establecer el método Kriging o regresión en proceso Gaussianos; en segundo lugar, el método de muestreo es distinto, estableciéndose un método de casos equiespaciados y LHS; en tercer lugar, se determinan distintos criterios para la aplicación de métodos de cuantificación de incertidumbre; y por último, el modelo del sistema físico y las variables consideradas, ya que si bien es similar, al verificar que es un embalse de relaves, las variables de entrada y salida son diferentes, y como diferencia fundamental, estas últimas presentan una mayor dimensionalidad, lo cual aumenta la complejidad del problema y presenta desafíos en la precisión del entrenamiento del modelo sustituto.

3.3. Cuantificación de Incertidumbre

La cuantificación de incertidumbre es un área de estudio que tiene el objetivo de analizar e interpretar la información aparentemente impredecible, por efectos aleatorios de un determinado sistema. Su análisis es de suma importancia para lograr una mayor comprensión de las respuestas de un sistema, y da mayor confiabilidad respecto a decisiones a tomar respecto al mismo, mediante la entrega de ciertos indicadores probabilísticos referido a las componentes o variables que pueden estar contenidas en dicho sistema.

La confección de un embalse de relaves conlleva la comprensión y análisis de múltiples variables, las cuales a su vez presentan incertezas y comportamientos aleatorios. Los materiales granulares en sí, por su composición tienden a tener comportamientos complejos y de carácter no lineal, los cuales aumentan la imprecisión del conocimiento respecto a su respuesta. Por otro lado la caracterización de los parámetros que se requieren para modelar, es decir los valores de entrada para el sistema tienen a su vez una variabilidad significativa, debido a todos los factores que influyen, tales como la representatividad de las muestras de suelo a medir en laboratorio, las diferentes composiciones físicas y químicas de las muestras, las diversas granulometrías presentes, entre otras consideraciones. Es por ello que la sensibilidad del sistema frente a tal configuración de variables puede ser considerablemente alta y se debe realizar una búsqueda por obtener indicadores al respecto de los parámetros de salida (respuesta), y como estos se ven afectados por las variables involucradas. Es por esto que es importante un análisis de sensibilidad, en el que se identifica el grado en que afectan las variables a la respuesta del sistema, el cual a su vez es sometido a un análisis de propagación de incertidumbre, basándose en la probabilidad de ocurrencia de situaciones extremas respecto a las mismas variables. La propagación de incertidumbre, permite transferir las incertezas de las variables al sistema en general, y de este modo entender el comportamiento y su respuesta.

Los embales de relaves son obras complejas y a su vez deben presentar una gran confiabilidad en su diseño y modelamiento debido a la magnitud de los mismos y al riesgo que estos conllevan, tanto físico, químico, social y ambiental. Es por ello que se hace tan importante el estudio de la incertidumbre de estos sistemas.

3.3.1. Conceptos básicos

La notación de una función de un determinado sistema puede ser como sigue:

$$z = G(u, \theta) \quad (3.44)$$

donde z es la respuesta del sistema, G es la función del sistema, u las variables externas al sistema y θ las variables internas del sistema.

Es posible considerar un error en la obtención de la respuesta, teniendo la expresión:

$$y = z + e \quad (3.45)$$

$$e \sim N(0, \sigma_e) \quad (3.46)$$

$$y \sim N(z, \sigma_e) \quad (3.47)$$

donde e es un error que sigue una distribución Gaussiana y por ende ahora se tiene que y sigue también una distribución Gaussiana.

Por simplificación en la notación se establecerá que:

$$z = g(\theta) \quad (3.48)$$

considerando que θ abarca todas las variables de entrada que ingresan a la función z .

3.3.2. Propagación de incertidumbre

El proceso de propagación de incertidumbre [34] permite expresar las incertezas o la variabilidad, determinadas por las PDF de las variables de entrada, y propagarlas de modo que se expresen en la respuesta o valores de salida del sistema.

Para realizar este procedimiento se establece el uso del método de Monte Carlo [35], el cual permite resolver integrales complejas de forma numérica. Este método se basa en el Teorema Central del Límite (TCL) [35], que expresa lo siguiente:

Si se tienen muestras $\theta = x_i, i = 1, \dots, K$, generados a partir de cualquier distribución de probabilidad $p(\theta)$, entonces se establece la siguiente ecuación:

$$g(\hat{\theta}) = \frac{1}{K} \sum_i g(\theta_i) \quad (3.49)$$

donde $g(\hat{\theta})$ es el promedio de la función, y esta se convierte en un variable aleatoria que seguirá una distribución normal $N(\mu, \frac{\sigma^2}{K})$. Estos datos estadísticos se definen como:

$$E[g(\hat{\theta})] = \frac{1}{K} \sum_i E[g(\theta_i)] = \mu \quad (3.50)$$

$$Var[g(\hat{\theta})] = \frac{1}{K} \frac{1}{K} \sum_i Var[g(\theta_i)] = \frac{\sigma^2}{K} \quad (3.51)$$

Luego si las muestras K tienden a infinito, entonces la varianza es muy pequeña y el valor esperado de $g(\hat{\theta})$ es el valor de la integral que se quiere resolver.

Dicho esto el método busca resolver la siguiente expresión

$$H = E[g(\theta)] = \int_{\theta_1} \dots \int_{\theta_n} g(\theta) p(\theta_1) \dots p(\theta_n) d\theta_1 \dots d\theta_n \quad (3.52)$$

que representa el valor esperado de la respuesta del sistema, considerando las incertezas.

Gracias a los principios del método es posible obtener los siguientes valores, que representan la aproximación del valor de la integral mencionada. En primer lugar se tiene el estimador de H :

$$\hat{H} = \frac{1}{K} \sum_i^K g(\theta_i) \quad (3.53)$$

El cual por el TCL, sigue una distribución Gaussiana con media y varianza:

$$E[\hat{H}] = \frac{1}{K} \sum_i E[g(\theta_i)] \approx \frac{1}{K} \sum_i g(\theta_i) \quad (3.54)$$

$$Var[\hat{H}] = \frac{1}{K} \frac{1}{K} \sum_i Var[g(\theta_i)] \approx Var[g(\theta)] \quad (3.55)$$

Estos cálculos (esperanza y varianza) se realizan un determinado número de veces (similar a K), para distintos set de combinaciones de variables de entrada, determinados por el número de muestras y su distribución conjunta. De este modo se tiene que a una gran cantidad de veces que se realiza, el estimador tiene una aproximación mucho más precisa. Para el presente caso se usa la aproximación dada en las ecuaciones anteriores, en las que se presenta la esperanza y varianza de los procesos a partir de un solo set de combinaciones de variables.

El procedimiento a realizar para ejecutar el método y aplicar estas ecuaciones junto al respectivo estimador, tiene tres variantes: Monte Carlo Directo, Monte Carlo con Estándar Gaussiano e Importance Sampling.

Se describe la metodología del caso de Monte Carlo con Estándar Gaussiano, debido a que es el que se ocupa en el desarrollo del presente documento. Mayor información sobre las demás metodologías en el documento de Robert y Casella [35].

El método comienza estableciendo los datos estadísticos de las variables de entrada, esto es, la media y desviación estándar principalmente. Se realiza una distribución conjunta de estas, las cuales pueden ser del tipo que se requiera o que se obtenga según los registros de las variables. Para el caso de una distribución conjunta Gaussiana se denota como sigue:

$$\theta \sim N \left(\begin{bmatrix} \mu_i \\ \vdots \\ \mu_n \end{bmatrix}, \begin{bmatrix} \sigma_{i,i}^2 & \dots & \sigma_{i,n}^2 \\ \vdots & \cdot & \vdots \\ \sigma_{n,i}^2 & \dots & \sigma_{n,n}^2 \end{bmatrix} \right) \quad (3.56)$$

El siguiente paso consiste en obtener K muestras de una distribución estándar Gaussiano (K a determinar).

$$\theta_G \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right) \quad (3.57)$$

Luego en base a la distribución en la ecuación 3.57, se realiza una descomposición de Cholesky inferior 'C', la cual cumple con la siguiente igualdad $\Sigma_\theta = CC^T$

Se realiza la transformación a espacio físico en base a la siguiente ecuación:

$$\theta_{nuevo} = \mu_\theta + C \cdot \theta_G \quad (3.58)$$

Se reemplaza en la función y se obtienen las muestras de salida. Estas pueden ser representadas como histogramas y como PDF.

Se define un parámetro que mide la calidad del método Monte Carlo, dado por:

$$c.o.v.MC = COV_{MC} = \frac{COV_g}{\sqrt{K}} \quad (3.59)$$

donde:

$$c.o.v.g = COV_g = \frac{\sqrt{Var[g(\theta)]}}{E[g(\theta)]} \quad (3.60)$$

3.3.3. Probabilidad de falla

El procedimiento de cálculo de probabilidad de falla es necesario para determinar la probabilidad de que el modelo exceda un valor límite preestablecido [35].

Se debe agregar un paso en el procedimiento de propagación de incertidumbre, para convertirlo a un proceso de cálculo de probabilidad de falla. Al momento de obtener las muestras de salida del modelo $g(x)$, se debe crear un nuevo set de valores a partir de una condición sobre dichas muestras, de modo que no excedan cierto valor.

La formulación se expresa como sigue:

$$P_F = \int_{\theta_1} \dots \int_{\theta_n} I_F(\theta) p(\theta_1) \dots p(\theta_n) d\theta_1 \dots d\theta_n \quad (3.61)$$

lo cual se aproxima:

$$E[P_F] \approx E[I_F(\theta)] \approx \frac{1}{K} \sum_i I_F(\theta_i) \quad (3.62)$$

$$Var[P_F] \approx Var(I_F(\theta)) \approx E[I_F(\theta)] - E[I_F(\theta)]^2 \quad (3.63)$$

donde la función $I_F(\theta)$ se define como sigue, donde se evalúa el valor límite a analizar:

$$I_F = \begin{cases} 0, & \text{si } g(\theta) < \text{valor lim} \\ 1, & \text{si } g(\theta) \geq \text{valor lim} \end{cases} \quad (3.64)$$

Es posible graficar los valores de la función $I_F(\theta)$, ordenando este set de menor a mayor, y graficandolo versus un vector definido por valores de 0 a K dividiendo cada uno en K . De este modo se obtiene un gráfico que permite visualizar la probabilidad de falla de forma general, es decir la asociada a cada valor que puede tomar la variable aleatoria a analizar.

Al igual que en la propagación de incertidumbre se define un parámetro que mide la calidad del método Monte Carlo:

$$COV_{MC} = \frac{COV_{I_F}}{\sqrt{K}} = \sqrt{\frac{1 - \hat{P}_F}{K \hat{P}_F}} \quad (3.65)$$

3.3.4. Análisis de sensibilidad

La realización de un análisis de sensibilidad permite identificar el impacto de los parámetros de entrada del modelo sobre la respuesta o salida del mismo. En otras palabras, permite visualizar en que grado afecta la variabilidad de cada parámetro en la variabilidad de la respuesta o salida.

Un procedimiento común para cuantificar la sensibilidad de los parámetros del modelo es la estimación de los índices de Sobol [36] [37]. Este se realiza a partir de lo que se conoce como el análisis de descomposición de varianzas que se basa en las siguientes ecuaciones [38]

Se define las siguientes ecuaciones:

$$\mu_h = E[H(\theta)] = \int_{\theta} H(\theta)p(\theta)d\theta \quad (3.66)$$

$$V_h = E[(H(\theta) - \mu_h)^2] = E[H(\theta)^2] - \mu_h^2 = \int_{\theta} H(\theta)^2 p(\theta)d\theta - \mu_h^2 \quad (3.67)$$

Por otro lado se tiene la descomposición de varianzas:

$$V_h = \sum_{i=1}^N V_i + \sum_{i=1}^{N-1} \sum_{j=1}^N V_{i,j} + \dots \quad (3.68)$$

V_i y $V_{i,j}$ se definen como:

$$V_i = E_i[\psi_i(\theta_i)^2] \quad (3.69)$$

$$V_{i,j} = E_{i,j}[\psi_{i,j}(\theta_i, \theta_j)^2] \quad (3.70)$$

donde:

$$\psi_i(\theta_i) = \xi_i(\theta_i) - \mu_h \quad (3.71)$$

$$\psi_{i,j}(\theta_i, \theta_j) = \xi_{i,j}(\theta_i, \theta_j) - \mu_h - \psi_i(\theta_i) - \psi_j(\theta_j) \quad (3.72)$$

y:

$$\xi_i(\theta_i) = E_{\sim i}[H(\theta)|\theta_i] = \int_{\theta_{\sim i}} H(\theta_{\sim i}, \theta_i) p(\theta_{\sim i}) d\theta_{\sim i} \quad (3.73)$$

$$\xi_{i,j}(\theta_i, \theta_j) = E_{\sim i,j}[H(\theta)|\theta_i, \theta_j] = \int_{\theta_{\sim i,j}} H(\theta_{\sim i,j}, \theta_i, \theta_j) p(\theta_{\sim i,j}) d\theta_{\sim i,j} \quad (3.74)$$

Es posible entonces definir el valor de V_i y $V_{i,j}$ como:

$$V_i = \int_{\theta_i} \left[\int_{\theta_{\sim i}} H(\theta) p(\theta_{\sim i}) d\theta_{\sim i} \right]^2 p(\theta_i) d\theta_i \quad (3.75)$$

$$V_{i,j} = \int_{\theta_{i,j}} \left[\int_{\theta_{\sim i,j}} H_i(\theta) p(\theta_{\sim i,j}) d\theta_{\sim i,j} \right]^2 p(\theta_{i,j}) d\theta_{i,j} - V_i - V_j - \mu_h \quad (3.76)$$

V_i es la varianza de primer orden y $V_{i,j}$ es la varianza de segundo orden. La notación $\theta_{\sim i}$ representa todos los valores de esa variable denotada excepto el valor i de la misma.

Los índices de Sobol, S_i , que representa la contribución de cada factor de entrada (θ_i) a la varianza total de la salida (V_h), y $S_{i,j}$, que representa la parte de la varianza de H explicada por la interacción de θ_i y θ_j . se definen como:

$$S_i = \frac{V_i}{V_h} \quad S_{i,j} = \frac{V_{i,j}}{V_h} \quad (3.77)$$

Finalmente se define un índice de Sobol de orden total S_{Ti} , que representa la contribución total de la salida H debido al factor θ_i , esto es, su efecto de primer orden más todos los efectos de orden superior producto de las interacciones de este factor con los demás. Se obtiene condicionando con respecto a todos los factores a excepción de θ_i .

$$S_{Ti} = 1 - S_{\sim i} \quad (3.78)$$

$$S_{\sim i} = \frac{E_{\sim i}[\xi_{\sim i}(\theta_{\sim i})^2] - \mu_h^2}{V_h} \quad (3.79)$$

Se debe mencionar que el cálculo de estas expresiones conlleva una carga computacional alta, por lo que se crean estimadores que buscan resolver este problema, en base al documento de Saltelli [39]

3.3.5. Comentarios sobre revisión bibliográfica

El estudio de las incertidumbres y su inclusión en las distintas aplicaciones de la ingeniería es de suma relevancia y otorga muchas utilidades, tal como se ha mencionado en la presente sección. Su planteamiento es utilizable en diversos tipos de sistemas ingenieriles, científicos o incluso ámbitos externos a estas categorías, siempre y cuando se cumpla el formato característico del problema. Este formato está asociado al establecimiento de un sistema asemejado a una función o caja negra, la cual integra variables de entrada e internas y que provocan una variable de salida, tal como se indica en la sección 3.3.1.

Esta área de estudio aplicada a los sistemas geotécnicos o mineros presenta variadas aplicaciones, tal y como es mencionado por Gúzman para Minería Chilena [3]. Las incertezas están incluidas en los sistemas geotécnicos de un modo u otro, y los embalse de relaves tienen esta condición.

Para el caso que acontece en esta memoria, el documento de Sung [33] y el trabajo de Ashley y Ali [31] se asemejan al obtener resultados asociados a la cuantificación de incertidumbre, tales como los asociados a la obtención de un histograma o una distribución de probabilidad del factor de seguridad, procedimiento relacionado a la propagación de incertidumbre, o la probabilidad de falla determinada para el mismo parámetro, así como también un análisis de sensibilidad. Estos procedimientos se aplican con una metodología diferente en el presente trabajo, pero sin alejarse de los resultados esperados.

Los principios y fundamentos de la cuantificación de incertidumbre están descritos de forma adecuada en el documento de Robert y Casella [35].

Capítulo 4

Diseño de Modelo de Embalse de Relaves

El sistema geotécnico a modelar consiste en un embalse de relaves, el cual presenta las características que se describen en este capítulo.

El modelo fue desarrollado en FLAC3D, software de modelamiento numérico otorgado por empresa colaborante (Itasca Chile SPA) en el presente trabajo de memoria. Este consiste en realizar el análisis de estabilidad física del embalse mencionado, modelando la sección considerada más crítica, de modo de establecer el caso más desfavorable.

Los datos que caracterizan el modelo (dimensiones geométricas, parámetros geotécnicos, entre otros), son referidos a un proyecto de embalse de relaves adjudicado a Itasca Chile SPA, para el cual se desea hacer un mejoramiento de los procesos de cálculo del modelo, tal como se indica en secciones anteriores.

Se busca ejecutar el modelo para distintos casos, referidos a diferentes combinaciones de parámetros de entrada, que definen el marco de acción y respuesta del sistema establecido. Esto permite obtener una base de datos, lo cual da el paso a la siguiente etapa, que se menciona en el próximo capítulo.

4.1. Bases de diseño

En esta sección se presentan las bases de diseño del modelo a realizar basándose en la normativa existente referida y enfocada a la estabilidad física de este tipo de obras civiles.

4.1.1. Categorización de embalse

Se establece un embalse de relaves, el cual está compuesto por material de empréstito. De este modo es posible seleccionar el material conformador del muro y sus propiedades según se solicite.

En base al artículo 12 del decreto 50 del MOP [40], es posible categorizar los embalses según su altura y capacidad de embalsamiento, tal como indica la Tabla 4.1, y al establecer la altura del muro de 40 metros se identifica que consiste en un embalse del tipo C.

Tabla 4.1: Categorización de muro según altura y capacidad

Categoría	Tamaño	Altura de muro máxima [m]	Capacidad [m ³]
A	Pequeños	$5 \leq h < 15$	$50.000 \leq h < 1.500.000$
B	Medianos	$15 \leq h < 30$	$1.500.000 \leq h < 60.000.000$
C	Grandes	$30 \leq h$	$60.000.000 \leq h$

4.1.2. Dimensiones

Las dimensiones del embalse de relave deben tener los siguientes requerimientos mínimos dados por la normativa vigente:

- i. Altura y Talud: Si bien no se presenta normativa referente a restricciones estrictas en dimensiones de muros, como la altura y talud, en la Tabla 4.2 se presenta la geometría de los depósitos de relaves con muro de empréstito conocidos del país, principalmente de las compañías mineras que son parte del consejo minero. Esta es obtenida del documento de Barrios [11]. Con esto es posible considerar los rangos descritos en la Tabla 4.3 como referencia.
- ii. Ancho de coronamiento: El artículo 50 del D.S. N°248 [4] establece un ancho de coronamiento que asegure la estabilidad, y define como mínimo, un ancho de 2 metros.
- iii. Revancha: Según el artículo 49 del D.S. N°248 del Ministerio de Minería [4] y del artículo 36 del D.S. N°50 del MOP [40], se establece que la revancha en los depósitos de relaves debe ser, como mínimo, de un 1 metro.

Tabla 4.2: Registro de principales embalses de Chile

Embalse	Compañía	Altura de muro [m]	Talud aguas abajo [H:V]	Talud aguas arriba [H:V]	Ancho de coron. [m]
Los Leones	Codelco	128	-	-	-
Carmen de Andacollo	Teck	151	-	-	-
Pampa Austral	Codelco	45	2:1	2.5:1	7
Pampa Pa-bellón	Collahuasi	63	1.8:1	1.6:1	-
Alto Norte	Glencore	31	2.5:1	2.5:1	6
Rekaves III	Glencore	41	2:1	2:1	7
Centinela (Muro sur y oeste)	AMSA	41	2.5:1	2:1	25
Laguna Seca	BHP	107	2:1	1.8:1	15
Los Diques	Minera Candelaria	156	-	-	-
El Rome-ral	CAP	50	-	-	-
Cerro Negro	CAP	73	-	-	-
Carén	Codelco	137	1.6:1	1.6:1	10
Talabre	Codelco	46	2.5:1	2:1	12

Tabla 4.3: Rangos de dimensiones de muro

Altura de muro [m]	Talud aguas abajo [H:V]	Talud aguas arriba [H:V]	Ancho de coronamiento [m]
< 150	Min 1.6:1 Max 2.5:1	Min 1.6:1 Max 2.5:1	6 - 25

4.1.3. Análisis de estabilidad

Se considera un criterio de estabilidad referido a un análisis de estabilidad físico basado en el factor de seguridad. Este se calcula mediante el software FLAC3D, basado en métodos finitos, específicamente el método de volúmenes finitos.

De forma preliminar se establece un análisis estático, con proyección a la realización de un análisis seudoestático y dinámico.

El artículo 31 del D.S. N°50 del Ministerio de Obras Públicas [40], establece los análisis de estabilidad mínimos exigidos para cada una de las categorías de embalses:

- i. Embalse Categoría A: Estático y pseudoestático, teniendo en consideración el estudio sísmológico aplicable a esta categoría.
- ii. Embalse Categoría B: Se elaborarán, al menos, los análisis estáticos y seudo estáticos, para los Sismos Máximo Creíble y de Diseño. Dependiendo de la sismicidad de la zona en que se encuentren las obras, del tipo de presa, de su fundación y de otras particularidades de interés para este tipo de análisis, aspectos que tienen que estar debidamente justificados en el proyecto, se deberá demostrar que no es necesario incluir un análisis dinámico para esta categoría de embalse.
- iii. Embalse Categoría C: Estático y dinámico, para los Sismos Máximo Creíble y de Diseño. En la verificación de la estabilidad de los muros, se deben obtener coeficientes o factores de seguridad mínimos, según lo que se especifica a continuación:
 - Caso estático: $FS \geq 1,4$
 - Caso seudoestático: $FS \geq 1,2$

Para embalses de Categorías B y C se deberá comprobar la estabilidad en condiciones post sísmicas referidas al Sismo Máximo Creíble, justificando adecuadamente los parámetros utilizados para este análisis. En este caso el factor de seguridad mínimo deberá ser mayor a 1.0.

El mismo D.S: N°50 describe aspectos relacionados al estudio sísmológico e hidráulico, los cuales no se analizarán en detalle en el presente documento, sin embargo se menciona y proyecta su estudio.

4.2. Geometría

La geometría del embalse de relaves a diseñar se presenta en la Figura 4.1 y en la Figura 4.2 con sus respectivas dimensiones, así como en la Tabla 4.4, respaldada por la normativa y referencias mencionadas en las bases de diseño. Este diseño se ha realizado de forma referencial, con bases en proyectos adjudicados a empresa Itasca. Como se puede ver, se establece una sección, la cual es considerada la más crítica o desfavorable (Corte B - B' en este caso). Es posible diferenciar tres zonas características: el muro que representa la obra en sí que permite sostener el material depositado de relaves o lamas; la base, que representa el suelo de fundación sobre el que se construirá el muro, el cual generalmente consiste en roca; y la zona de lamas, que representa el material a depositar y que otorga una sollicitación y confinamiento en esa localización.

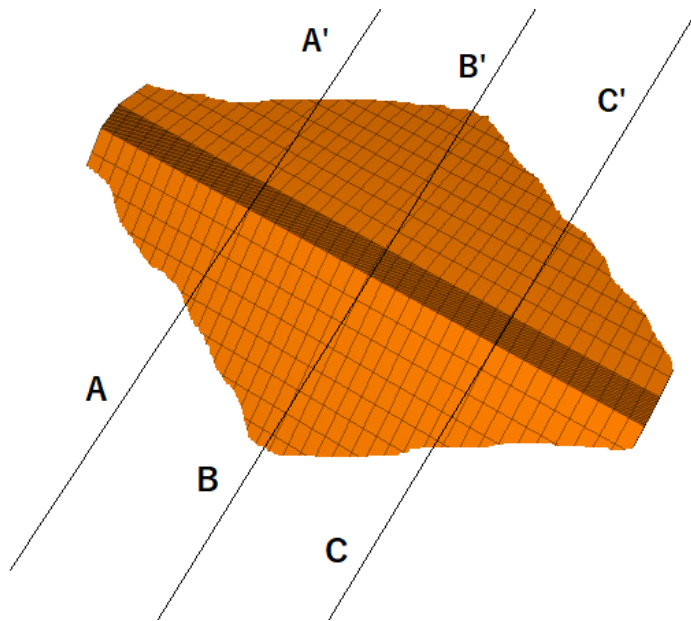


Figura 4.1: Embalse de relaves - Perspectiva 3D

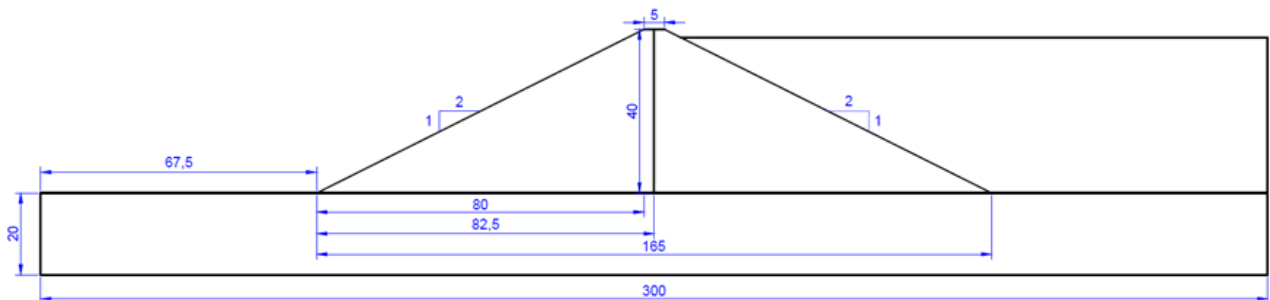


Figura 4.2: Geometría de embalse a diseñar

Tabla 4.4: Dimensiones del modelo

Parámetro	Valor	Unidad
Altura de muro (h_m)	40	[m]
Ancho de base de muro (B_m)	165	[m]
Ancho de coronamiento (b_m)	5	[m]
Altura de base (h_b)	20	[m]
Ancho de base (B_b)	300	[m]
Revancha	3	[m]
Talud	2:1	[-]

La representación de la geometría en el programa FLAC3D se visualiza en la siguiente Figura:

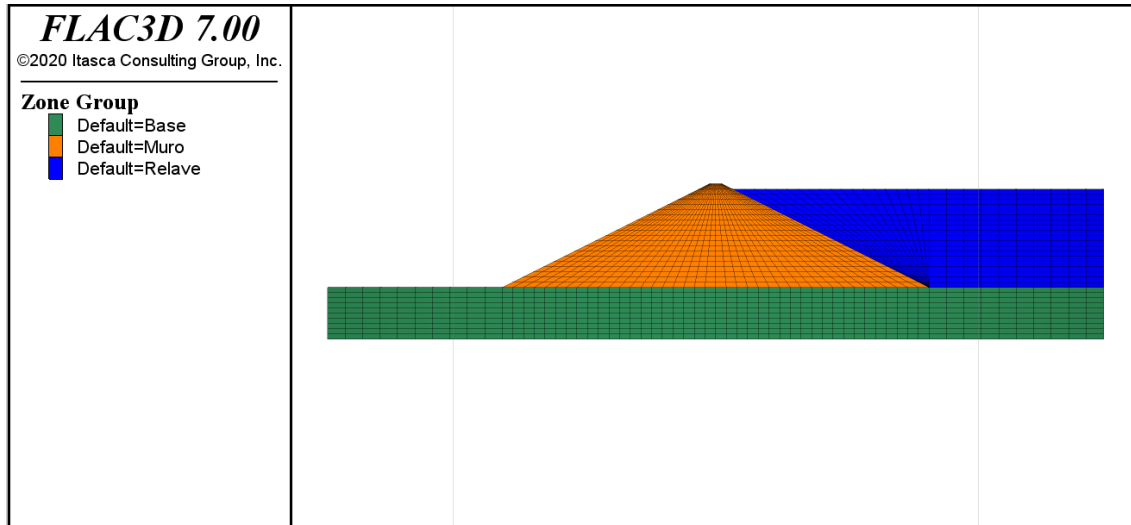


Figura 4.3: Geometría de embalse - FLAC3D

Se toman las siguientes consideraciones respecto al mallado de la figura, para establecer los elementos diferenciales:

Tabla 4.5: Discretización del modelo

Zona	Discretización	Valor
Muro	Vertical	20
	Horizontal	20
Base	Vertical	10
	Horizontal	10+40+10
Lamas	Vertical	19
	Horizontal	18+10

4.3. Caracterización geotécnica

Las propiedades de las zonas descritas previamente se determinan en esta sección.

La caracterización geotécnica se obtiene a partir de la base de datos de Itasca Chile SPA. Estos son datos gruesos obtenidos directamente de ensayos realizados para el proyecto mencionado previamente.

Tabla 4.6: Propiedades geotécnicas del modelo

Parámetro	Muro	Base	Lamas	Unidad
Material	Arena	Roca	Lamas	[-]
Densidad	1950	2900	2100	[kg/m ³]
Densidad saturada	2000	2900	2400	[kg/m ³]
Ángulo de fricción (ϕ)	[30 - 37]	50	-	[°]
Cohesión	[0 - 15000]	3.2e6	-	[N/m ²]
Módulo de Young (E)	5e7	2e8	6e6	[N/m ²]
Coefficiente de Poisson (ν)	0.3	0.3	0.49	[-]
Resist. corte no drenada (S_u)	-	-	0.1*esfzz	[N/m ²]
Coefficiente de presión lateral (K_0)	1-sen(ϕ)	1	-	[-] [m]

4.4. Modelamiento del sistema

Las características del modelamiento del embalse de relaves son mencionadas a continuación, basadas en la documentación de Itasca acerca de FLAC3D [13]:

4.4.1. Método numérico

El método numérico utilizado por FLAC3D es el método de volúmenes finitos. Mayor detalle en Marco Teórico.

4.4.2. Modelo constitutivo

Se debe definir el modelo constitutivo del material que se modela, el cual es de tipo granular. Existen distintos modelos constitutivos que se pueden elegir para caracterizar el material, los cuales son mencionados en la documentación de FLAC3D de Itasca [13].

Se selecciona un comportamiento definido por Mohr-Coulomb. Su uso queda argumentado en la sección de Análisis y Discusión.

4.4.3. Condiciones de borde

Las condiciones de borde del problema deben fijar y limitar los desplazamientos, velocidades o aceleraciones de las zonas de borde del modelo.

Para el actual modelo se establecen del siguiente modo:

- Debido a que se analiza una sección plana, en dos dimensionaes, es que se fija el grado de libertad perpendicular al plano de dicha sección, el cual en este caso es la dirección en y .
- La zona más baja, en la zona inferior de la base se asume un empotramiento, por lo que se restringen todos los grados de libertad, en las tres direcciones x , y , z .
- En las zonas laterales de la zona de la base, se restringen adicionalmente los grados de libertad en x .
- En las zonas laterales de la zona del muro no se restringen los grados de libertad, excepto en la dirección y , como se menciona en el primer punto.

4.4.4. Condiciones iniciales

Las condiciones iniciales del problema se establecen tomando las siguientes consideraciones:

- Se considera como condición inicial, la fuerza de gravedad aplicada al modelo, siendo la única sollicitación modelada.
- Se establece el comportamiento del material granular, respecto a los esfuerzos laterales, estableciendo el valor del coeficiente de presión lateral K_0 .

4.4.5. Nivel freático

El nivel freático se modela de modo que sea posible establecer su trazado en base a puntos ubicados dentro del muro, los cuales representan la medición dada por piezómetros dispuestos en dicho sitio. Para el problema presentado se plantea la ubicación de dos niveles freáticos, los cuales se ingresan en el modelo y como respuesta se obtiene el trazado del nivel freático.

Este nivel impone una condición de saturación en el medio, de modo que por debajo de su delimitación el material adquiere una densidad saturada. Por otro lado se impone e inicializa la condición de presión de poros en la misma zona.

Es posible ver el nivel freático representado en FLAC3D, en la siguiente figura:

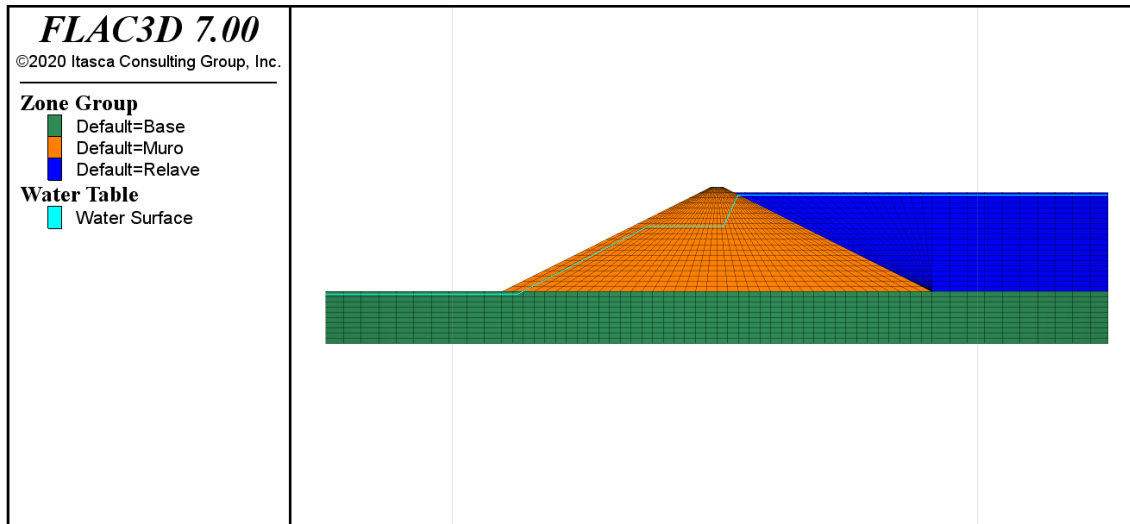


Figura 4.4: Nivel freático en embalse - FLAC3D

4.4.6. Resolución de equilibrio

La resolución del equilibrio se determina con un límite basado en la razón local de convergencia (local convergence ratio). Este límite hace referencia a un indicador del valor resultante neto dado por el equilibrio de fuerzas. Según recomendaciones de documentación de FLAC3D el valor de esta razón es igual a $1e-4$, sin embargo por recomendación de empresa Itasca, se utiliza una razón mucho más estricta igual a $2.5e-6$.

4.4.7. Cálculo de factor de seguridad

El cálculo del factor de seguridad se describe en detalle en el capítulo de marco teórico. Se establece tomando el criterio de razón local de convergencia (local convergence ratio) igual a $2.5e-6$, por experiencia de empresa Itasca. Adicionalmente se establece un intervalo de iteración del factor de seguridad de 0.1 y un rango a utilizar de valores mínimos y máximos igual a 1.0 y 2.0 respectivamente. Este procedimiento permite además trazar los contornos de factor de seguridad, que representan las zonas con un determinado factor de seguridad.

4.5. Selección de parámetros a analizar

Se debe realizar una selección de los parámetros que se analizarán, tal que que provoquen una variabilidad considerable en el modelo, dentro de rangos que permitan representar de forma completa el problema, y así obtener una cantidad considerable de casos o combinaciones distribuidas entre sí de forma heterogénea y abarcar la mayor cantidad de espacios o casos intermedios.

4.5.1. Parámetros a analizar

Los parámetros a analizar y que compondrán los casos o combinaciones a evaluar son seleccionados según su incidencia en la respuesta final. Esto se comprueba en primer lugar ajustando los rangos de acción, mediante el conocimiento físico del problema, la revisión bibliográfica y la experiencia, para luego evaluar los casos con las variables y sus valores extremos según ese rango mencionado (Procedimiento en Anexo C).

Los parámetros posibles para analizar y que se ingresan como variables de entrada en el modelo son los siguientes:

Tabla 4.7: Parámetros a analizar

Parámetro	Unidad
Densidad	[kg/m ³]
Densidad saturada	[kg/m ³]
Ángulo de fricción (ϕ)	[°]
Cohesión	[N/m ²]
Módulo de Young (E)	[N/m ²]
Coefficiente de Poisson (ν)	[-]
Resist. corte no drenada (S_u)	[N/m ²]
Coefficiente de presión lateral (K_0)	[-]
Nivel freático	[m]
Nivel freático 2	[m]
Altura de muro	[m]
Ancho de coronamiento	[m]
Altura de base	[m]
Ancho de base	[m]
Revancha	[m]
Talud	[-]
Discretización	[-]

Según el procedimiento anexado, se consideran las siguientes:

Tabla 4.8: Parámetros a analizar - Selección

Parámetro	Unidad
Nivel freático	[m]
Nivel freático 2	[m]
Ángulo de fricción	[°]
Cohesión	[N/m ²]

Cabe mencionar que dicho procedimiento exige un tiempo de ejecución (calculado en Anexo C) de 2.3 días.

4.5.2. Rangos de parámetros

Se determinan los rangos de valores de las variables elegidas en base a datos de proyecto mencionado, a bibliografía referida enfocada en correlaciones geotécnicas [7] [9] [8] y recomendaciones de empresa Itasca dadas por la experiencia y prácticas habituales. Se busca definir estos rangos en base a un esquema de entrenamiento global, es decir, se consideran todos los rangos de las variables donde se adquieran valores prácticamente viables o posibles según las fuentes ya mencionadas.

Tabla 4.9: Rangos de parámetros

Parámetro	Rango	Unidad
Nivel freático	[0 - 37]	[m]
Nivel freático 2	[0 - 37]	[m]
Ángulo de fricción	[30 - 37]	[°]
Cohesión	[0 - 15e3]	[N/m ²]

4.5.3. Selección de valores de parámetros

Se elige una cantidad de variables de modo que se logre una optimización respecto al tiempo de ejecución de los modelos y considerando valores que logren una representatividad adecuada del modelo. De esta manera se establece la Tabla 4.10 que describe las opciones a considerar. Cabe mencionar que el número de casos totales se determina realizando una permutación de la cantidad de valores por variable, elevando este al número de variables establecido.

Tabla 4.10: Análisis de casos y tiempo de ejecución

Opción	Valores por var.	Casos totales	T. por caso [hrs]	T. total [días]	T. total ef. [días]
1	6	1296	3	162	32.4
2	5	625	3	78.2	15.63
3	4	256	3	32	6.4
4	3	81	3	10.13	2.03

La Tabla descrita explica las opciones posibles con distintos números de casos totales y tiempo total de la ejecución de los mismos. Se evalúa un tiempo total efectivo, que considera la ejecución del total de casos divididos en grupos de casos, de modo de ejecutar esos casos simultáneamente en determinado número de "subejecuciones", los cuales en este caso son 5, por lo que el tiempo total se ve disminuido por la misma división.

Visualizando las opciones, se considera que una cantidad adecuada de casos y un tiempo razonable es la opción 2. Esto tomando en cuenta la utilización de un único computador, ejecutando los casos totales con sus respectivas subejecuciones.

Una vez obtenido el número de casos por opción y el tiempo asociado a su procesamiento

en el software de modelamiento (FLAC3D), así como también los rangos de los parámetros, se debe realizar el modo de selección de los casos dentro de estas condiciones.

Se realizan dos métodos de obtención de muestras. Estas muestras tienen el siguiente formato:

$$X = \{[NF, NF2, \phi, c]_i / i = 1, \dots, n\}$$

1. Latin Hypercube Sampling: casos se establecen implementando el algoritmo de LHS de la librería OpenTURNS, determinando una distribución uniforme para cada variable en base al respectivo rango de cada una y evaluando en el algoritmo el número de casos o muestras a obtener. Se obtiene la siguiente Tabla:

Tabla 4.11: Casos obtenidos por método

Caso	Combinación por caso
1	[5.03, 15.72, 34.12, 4347,8]
2	[23.57, 10.79, 32.93, 1176.58]
...	...
n	[19.12, 9.02, 32.58, 10878.02]

En Anexo D es posible visualizar los gráficos bidimensionales de las variables involucradas representadas entre sí, para verificar que la distribución de los casos a evaluar es adecuado para representar en gran parte el espectro total de casos. Cada punto en el gráfico representa un caso.

2. Casos equiespaciados por variable: casos establecidos por equiespaciamiento de los valores en el rango de cada variable con posterior permutación de éstos entre sí.

Tabla 4.12: Selección de valores de parámetros por método

Parámetro	Combinación por variable
Nivel freático	[0, 9.25, 18.5, 27.75, 37]
Nivel freático 2	[0, 9.25, 18.5, 27.75, 37]
Ángulo de fricción	[30, 31.75, 33.5, 35.25, 37]
Cohesión	[0, 3750, 7500, 11250, 15e3]

Tabla 4.13: Casos obtenidos por método

Caso	Combinación por caso
1	[0, 0, 30, 0]
2	[0, 0, 30, 3750]
...	...
n	[37, 37, 37, 15e3]

Se puede visualizar la distribución de los casos en Anexo D.

La Tabla 4.11 y la Tabla 4.13, representa los casos con distintas combinaciones de las variables de entrada a utilizar en el software FLAC3D.

La selección finalmente se realiza para la opción 2 de la Tabla 4.10 y para una distribución de casos equiespaciados por variables, y los datos presentados a lo largo del desarrollo son en base a esta decisión. Es importante señalar que se obtienen resultados basados en ejecuciones con el método de muestreo LHS y para una opción de la Tabla 4.10 que requiere menos casos, con el objetivo de realizar un análisis comparativo, además de realizar una comparación con una librería alternativa. Estos resultados y esta comparación se describen en Anexo F, G, H y en el capítulo Análisis y Discusión de Resultados.

Por otro lado, realizando un proceso adicional para una validación más respaldada en el proceso de confección del modelo sustituto, el cual se presenta en una próxima sección, se consume tiempo extra en la ejecución de casos de prueba, no considerados en la base principal relacionada al proceso de entrenamiento. Estos casos a calcular se establecen como sigue:

Tabla 4.14: Análisis de casos y tiempo de ejecución - Test

Opción	Valores por var.	Casos totales	T. por caso [hrs]	T. total [días]	T. total ef. [días]
Extra	2	16	3	2	0.4 (9.6 hrs.)

Tabla 4.15: Selección de valores de parámetros por método - Test

Parámetro	Combinación por variable
Nivel freático	[12.33, 24.66]
Nivel freático 2	[12.33, 24.66]
Ángulo de fricción	[32.33, 34.66]
Cohesión	[5e3, 1e4]

Tabla 4.16: Casos obtenidos por método - Test

Caso	Combinación por caso
1	[12.33, 12.33, 32.33, 5000]
2	[12.33, 12.33, 32.33, 10000]
...	...
n	[24.66, 24.66, 34.66, 10000]

4.6. Resultados

Los resultados obtenidos luego de la ejecución del modelo son bastante amplios y la información que entrega es muy valiosa para entender y analizar el comportamiento y la estabilidad del embalse de relaves. Algunos resultados que se pueden visualizar son:

- i. Esfuerzos
- ii. Deformaciones
- iii. Presiones de poros
- iv. Factores de seguridad
- v. Contorno de factores de seguridad

El enfoque en los resultados se hace en la información relacionada al factor de seguridad. Es por esto que se hace énfasis en el registro del factor de seguridad y el contorno de factores de seguridad, es decir el factor de seguridad asociado a una respectiva zona. Esto es:

Tabla 4.17: Parámetros de entrada y salida

Entrada	Salida
$X = \{[NF, NF2, \phi, c]_i / i = 1, \dots, n\}$	$Y = \{[FS, ContornoFS]_i / i = 1, \dots, n\}$

Específicamente los datos a extraer del modelo para cada caso son:

- i. Parámetros de entrada
- ii. Factor de seguridad (global)
- iii. Zonas
 - Número de zona
 - Grupo
 - Factor de seguridad (local)
 - Coordenadas
 - Esfuerzos
 - Deformaciones

Estos tienen el siguiente formato:

Tabla 4.18: Formato datos a extraer

nf1	nf2	ϕ	c	FS_g	Idz	Grz	FS_l	x	y	z	Esf.	Def.
$\{ \}_i$	$\{ \}_i$	$\{ \}_i$	$\{ \}_i$	$\{ \}_i$	$\{ \}_{i,j}$	$\{ \}_{i,j}$	$\{ \}_{i,j}$	$\{ \}_{i,j}$	$\{ \}_{i,j}$	$\{ \}_{i,j}$	$\{ \}_{i,j}$	$\{ \}_{i,j}$

Extrayendo un fragmento de la tabla de resultados se obtiene:

Tabla 4.19: Datos a extraer

nf1	nf2	ϕ	c	FS_g	Idz	Grz	FS_l	x	y	z	Esf.	Def.
0	0	30	0	1.044	1	1	1.6	71.51	0.5	21	*	**
0	0	30		1.044	2	1	1.65	75.54	0.5	21	*	**
...
37	37	37	15e3	1.569	1931	3	2	289.88	0.5	57	*	**
37	37	37	15e3	1.569	1932	3	2	296.63	0.5	57	*	**

donde los esfuerzos (*) y las deformaciones (**) queda descritos por los respectivos tensores.

Tabla 4.20: Datos de tensor de esfuerzos

Esfuerzos	σ_{xx}	σ_{xy}	σ_{xz}	σ_{yx}	σ_{yy}	σ_{yz}	σ_{zx}	σ_{zy}	σ_{zz}
-----------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------

Tabla 4.21: Datos de tensor de deformaciones

Deformaciones	ε_{xx}	ε_{xy}	ε_{xz}	ε_{yx}	ε_{yy}	ε_{yz}	ε_{zx}	ε_{zy}	ε_{zz}
---------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------

A continuación se presentan los resultados de forma gráfica de algunos de los casos ejecutados referidos al factor de seguridad y los respectivos contornos. Las siguientes gráficas representa el efecto de cada variable sobre la variable de salida u objetivo, fijando las demás (constantes) en un estado intermedio.

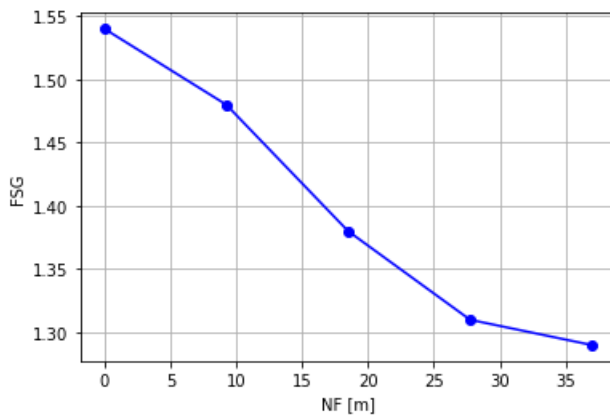


Figura 4.5: FSG versus nivel freático

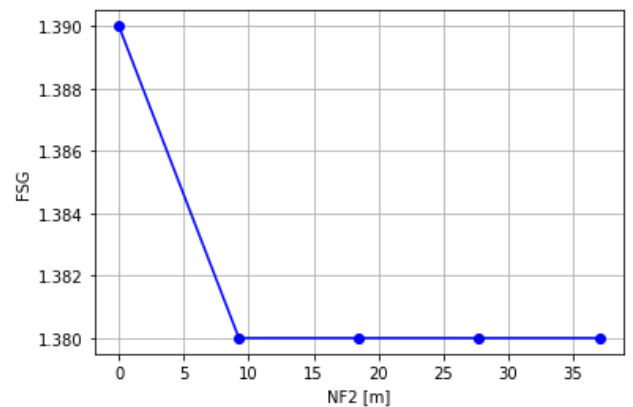


Figura 4.6: FSG versus nivel freático 2

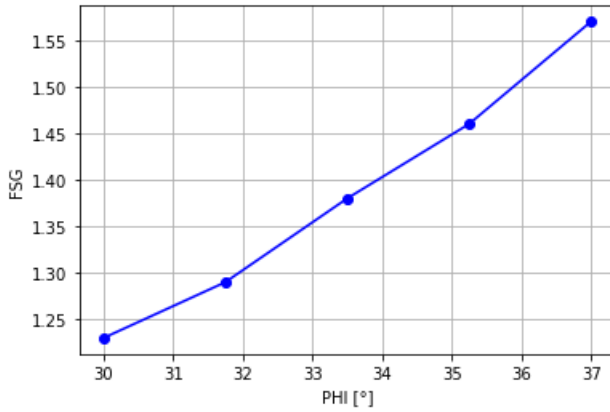


Figura 4.7: FSG versus ángulo de fricción (ϕ)

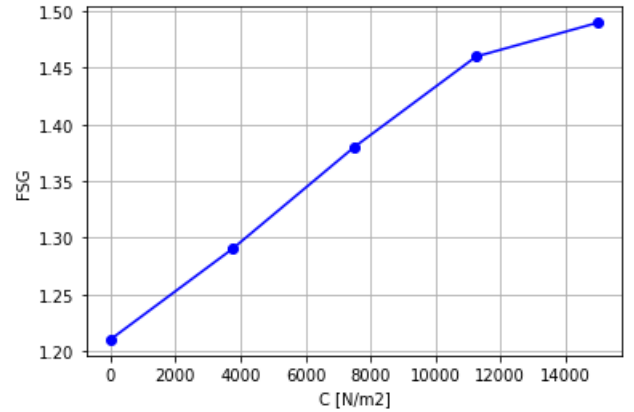


Figura 4.8: FSG versus cohesión

Los contornos del FS se presentan en la siguientes Figuras, donde se observan casos con distinto factor de seguridad 'global', el cual es el mínimo factor de seguridad que provoca la estabilidad, obtenido del proceso de reducción de esfuerzos. Por otro lado, los factores de seguridad 'locales' representan los factores de seguridad asociados a cada etapa de reducción de esfuerzos, y a su vez etiquetan una respectiva zona. Dicho esto las Figuras a continuación representan el caso más favorable, un caso intermedio y el más desfavorable respectivamente.

Tabla 4.22: Datos - Caso más favorable

nf1	nf2	ϕ	c	FS_g
0	0	37	15e3	1.88

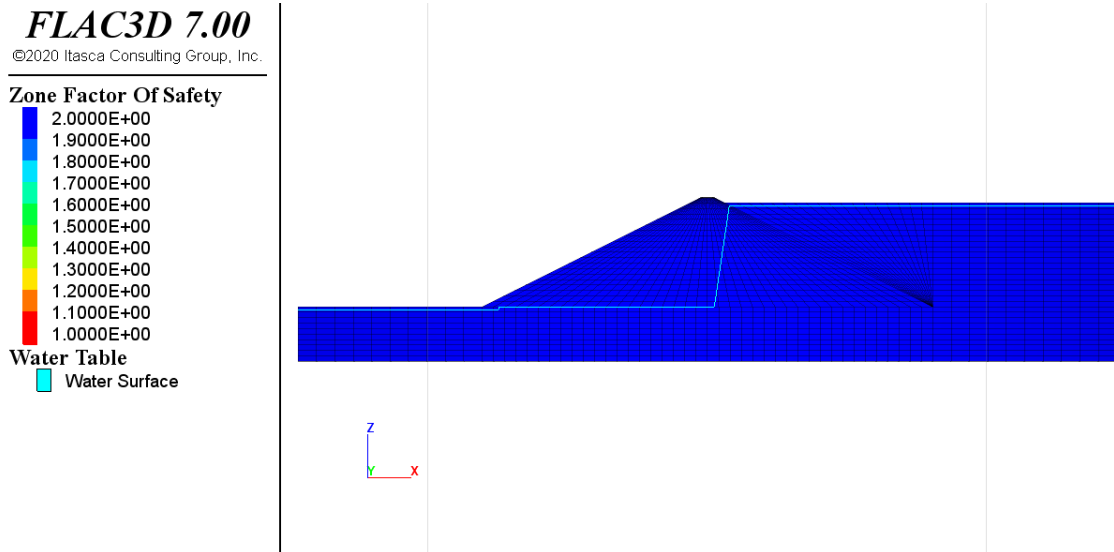


Figura 4.9: Caso más favorable - Contorno FS - FLAC3D - FSG=1.88

Tabla 4.23: Datos - Caso intermedio

nf1	nf2	ϕ	c	FS_g
18.5	18.5	33.5	7500	1.3813

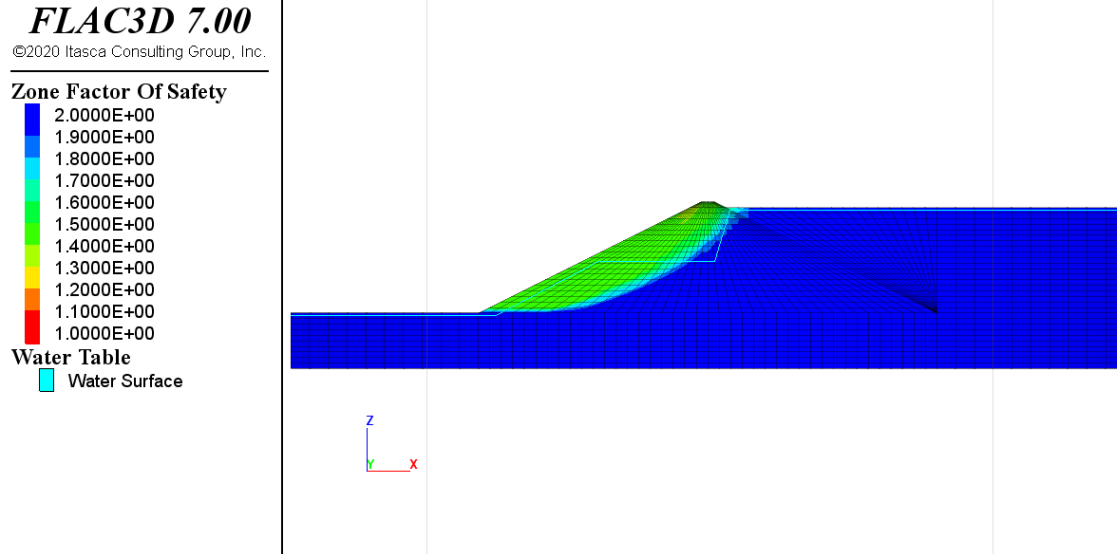


Figura 4.10: Caso intermedio - Contorno FS - FLAC3D - FSG=1.3813

Tabla 4.24: Datos - Caso más desfavorable

nf1	nf2	ϕ	c	FS_g
37	37	30	0	1.006

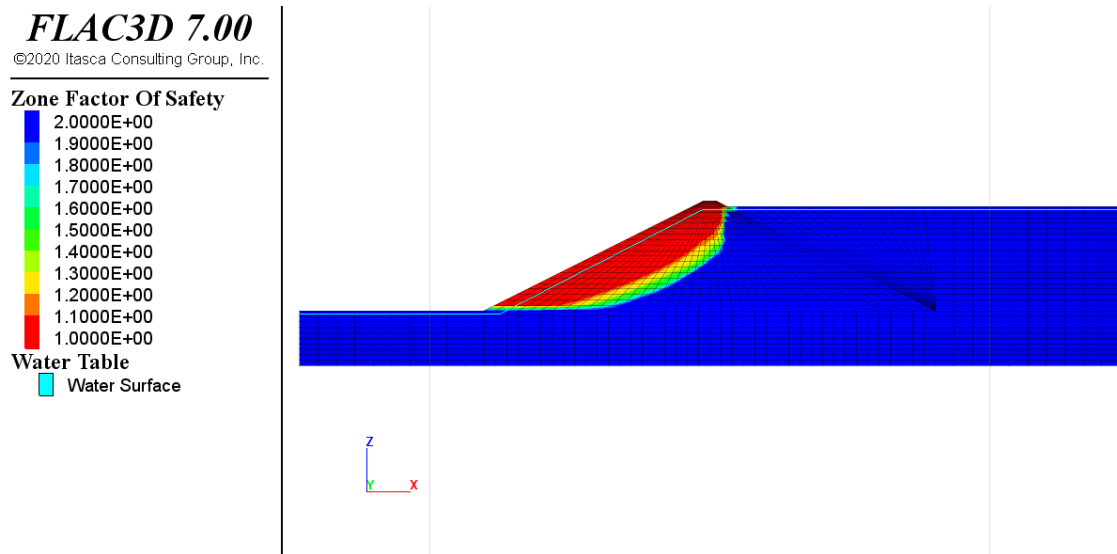


Figura 4.11: Caso más desfavorable - Contorno FS - FLAC3D - FSG=1.006

En Anexo E se presentan las Figuras con los resultados obtenidos, referidos a los esfuerzos, deformaciones, presiones de poros y contornos de factores de seguridad.

Capítulo 5

Confección de Modelo Sustituto

El modelamiento sustituto permite construir un modelo predictivo que pueda entregar una respuesta relacionada a un problema específico, a partir de un conjunto de variables de entrada, de forma ágil, rápida y precisa, en contraste con el modelo original diseñado en FLAC3D, el cual, si bien entrega una respuesta precisa, no presenta agilidad y rapidez en sus procesos, tal como hemos mencionado en apartados anteriores.

El modelo sustituto se construye a partir de una base de datos previamente establecida. Esta es la etapa de entrenamiento, la cual conlleva cierta carga computacional, muy inferior a la que conlleva la obtención de muestras, y además no es una problemática debido a que se debe realizar en una sola ocasión, y luego la predicción se realiza prácticamente de forma inmediata.

La base de datos debe ser seleccionada de forma adecuada para que se abarque la mayor cantidad de casos y que estos estén distribuidos de forma uniforme dentro del rango determinado de las variables. De este modo el modelo sustituto podrá predecir con mayor exactitud los datos intermedios, es decir, se podrá realizar una interpolación de forma más correcta y precisa. El proceso de predicción busca precisamente eso, que a partir de los datos de entrenamiento o de soporte sea posible ingresar nuevos valores de entrada y mediante el proceso de interpolación, se permita la obtención de la salida, que consiste en un valor intermedio de los datos de soporte, a partir de estos nuevos datos de entrada.

Mediante el procesamiento del modelo de embalse de relaves planteado en el capítulo anterior, es posible obtener la base de datos necesaria para ocupar en los siguientes procedimientos. Esta debe poder ser manipulada en un entorno de programación adecuado, el cual se ha decidido que será una interfaz de programación en base al lenguaje Python, debido a que es de código abierto, y permite el uso de una gran diversidad de herramientas, librerías y comandos relacionadas con el área de estadística, análisis predictivo, Machine Learning, entre otros.

5.1. Importación, filtro y manipulación de datos

La base de datos planteada, tiene el formato descrito en la Tabla 4.21. Esta en primer lugar debe importarse para poder ser manipulable, para luego poder hacer un filtro de los y enfocar las zonas de mayor interés para el análisis, y finalmente implementar los métodos predictivos para calibrar un modelo sustituto.

5.1.1. Importación de datos

Los datos han sido extraídos del programa FLAC3D en un formato del tipo texto (.txt), el cual es de uso común para la mayoría de programas. Así los datos en este formato son importados a la interfaz de Python para adaptarlos y convertirlos a un nuevo formato del tipo dataframe, elemento básico característico de la librería Pandas [41] en Python.

5.1.2. Filtro de datos

Una vez importados los datos, se debe hacer un filtro de los mismos, de modo de obtener los requeridos principalmente para el análisis.

Siguiendo el formato determinado en la Tabla 4.21, es posible visualizar que las zonas del modelo (para cada caso), se subdividen según el grupo de zona correspondiente. Como el interés esta puesto en analizar el comportamiento en el muro, es que se seleccionan solamente los datos de zonas que tienen este respectivo grupo, descartando los de las demás zonas (Base y Lamas).

Por otro lado las columnas con los valores de esfuerzos y deformaciones, si bien contienen información valiosa respecto al comportamiento del modelo, no se utilizan para realizar el modelo sustituto.

Con los datos ya filtrados, la base de datos se adapta y se obtiene el siguiente nuevo formato:

Tabla 5.1: Formato de datos actualizado

nf1	nf2	ϕ	c	FS_g	Idz	FS_l	x	y	z
$\{\}_i$	$\{\}_i$	$\{\}_i$	$\{\}_i$	$\{\}_i$	$\{\}_{i,j}$	$\{\}_{i,j}$	$\{\}_{i,j}$	$\{\}_{i,j}$	$\{\}_{i,j}$

5.1.3. Manipulación de datos

Hasta el momento se tiene un formato de datos como el que indica la Tabla 5.1. Se tiene que los datos de entrada son los cuatro parámetros dados por los dos niveles freáticos, el ángulo de fricción y la cohesión. Los datos de salida son el factor de seguridad (global) y el contorno del factor de seguridad, sobre el cual se abordará en mayor detalle a continuación, ya que su representación en los datos debe ser adaptada previo a su implementación en el método predictivo.

El contorno de factor de seguridad representa un conjunto de zonas que tienen asociadas el mínimo factor de seguridad luego de realizado el método de reducción de esfuerzos (Marco Teórico). Es posible decir que todas estas zonas son extraídas con una etiqueta o indicador asignado, referido al factor de seguridad respectivo. De este modo es posible graficar estas zonas en Python con la librería Matplotlib (ver Anexo B), y se obtiene un diagrama similar al de la Figura 5.1:

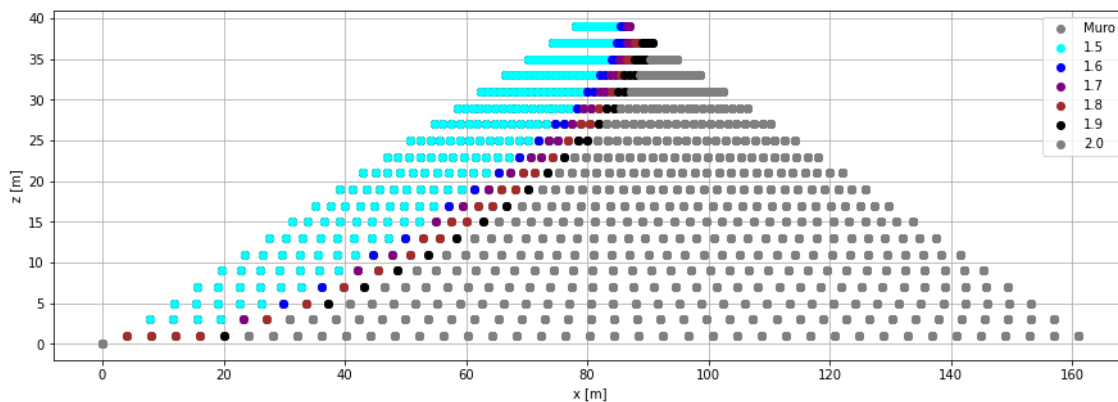


Figura 5.1: Contorno de factor de seguridad - Python

Se necesita reducir la cantidad de puntos a representar, es decir, la salida de los datos de entrenamiento debe tener una dimensión menor. Es por esto que de la Figura 5.1 se extraen los puntos de borde y se obtiene la Figura 5.3.

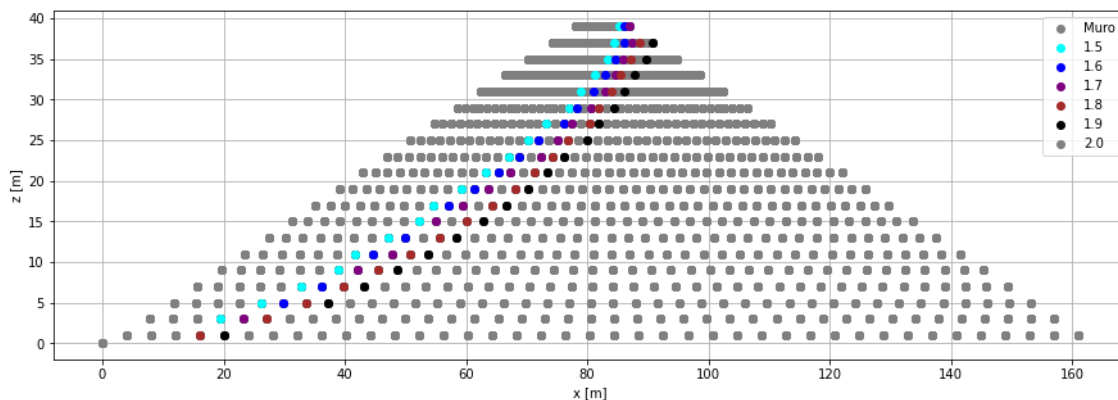


Figura 5.2: Puntos de borde de contorno de factor de seguridad - Python

Estos puntos caracterizan de forma adecuada el contorno a representar, sin embargo es posible reducir aún más la cantidad de puntos, realizando un ajuste de regresión de estos puntos del tipo polinomial de grado 3, utilizando la librería de scikit-learn (Ver Anexo B). Así cada conjunto de puntos de borde del contorno se puede parametrizar y queda definida por cuatro coeficientes. Esta idea puede verse reflejada en la Figura 5.3 y en la Ecuación 5.1.

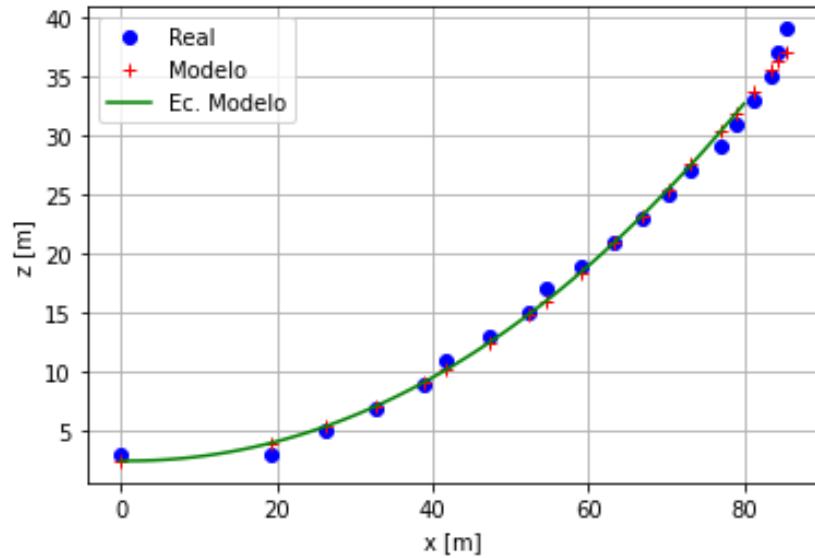


Figura 5.3: Parametrización de puntos

$$Curva = C0 + C1 \cdot x + C2 \cdot x^2 + C3 \cdot x^3 \quad (5.1)$$

En la Figura 5.3 se puede observar los datos reales en color azul y los datos en cruz rojos son los datos predecidos por el ajuste equivalentes a la ubicación en x de los puntos reales, y la curva en verda presenta el ajuste de forma continua.

De este modo la Figura 5.3 queda como sigue:

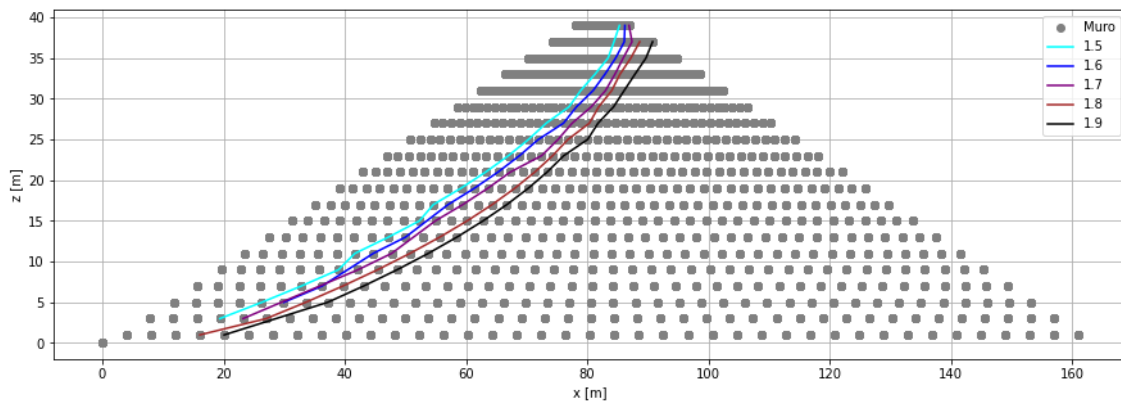


Figura 5.4: Parametrización de puntos de borde de contorno de factor de seguridad - Python

Finalmente es posible adaptar nuevamente el formato de datos de la Tabla 5.1 como sigue:

Tabla 5.2: Formato de datos final

nf1	nf2	ϕ	c	FS_g	FS_l	C0	C1	C2	C3
$\{\}_i$	$\{\}_i$	$\{\}_i$	$\{\}_i$	$\{\}_i$	$\{\}_{i,j}$	$\{\}_{i,j}$	$\{\}_{i,j}$	$\{\}_{i,j}$	$\{\}_{i,j}$

Cabe mencionar que se realiza un ajuste de dimensiones del muro del depósito, las cuales vienen con una configuración previa dada en el modelo efectuado en FLAC3D. Esto se realiza ubicando el punto inferior izquierdo del mismo en el punto (0,0), y así es posible obtener coordenadas de la figura que permitan una mejor manipulación y evitar errores al momento de entrenar el modelo sustituto.

Es necesario también adaptar los valores de los factores de seguridad, con el objetivo de lograr una discretización adecuada de estos. De este modo es posible etiquetar los casos en base a un determinado factor de seguridad local, lo cual permite realizar un entrenamiento adecuado para cada uno de estos conjuntos determinados pro esa respectiva etiqueta. Este proceso es necesario para la confección del modelo sustituto referido al modelo de contorno de FS local, que se describirá más adelante.

Por otro lado es importante recalcar que la calidad de los datos no es perfecta. Esto hace referencia a que los datos pueden presentar singularidades, lo cual implica que no haya una continuidad en los contornos. Esto puede presentar complicaciones al momento de realizar el entrenamiento, por lo que se deben imponer condiciones para adecuar estas singularidades. Estas complicaciones se mencionan a continuación, junto con su respectiva solución:

- i. Curvas de contorno de factor de seguridad irregulares: Estos casos provocan que el ajuste de los datos si bien es adecuado y logra obtener una buena relación entre los puntos, se presentan valores de los coeficientes que difieren de forma excesiva en comparación con la mayoría de los demás casos, por lo que esto se corrige agregando un punto estratégicamente, a una altura equivalente al punto más bajo del conjunto de puntos asociado a un respectivo factor de seguridad. De este modo se adecua principalmente el valor del coeficiente asociado al corte en el eje Y.
- ii. Casos que no presenten ciertas zonas de factor de seguridad local: Esta situación implica que no haya datos para ciertos factores de seguridad, es decir, no habría zonas caracterizadas por este valor, por lo que para resolver esto, se impone que esta curva de contorno asociada a un factor de seguridad local que no se presenta en un caso, está ubicada justo en el borde del muro. En términos sencillos, consiste en agregar datos donde no los hay por las condiciones del caso, sin falta de coherencia ni sentido físico. Esto permite ajustar de mejor modo el entrenamiento del modelo sustituto.

Finalmente este proceso de importación, filtro y manipulación, al involucrar una cantidad considerable de datos, conlleva un tiempo de procesamiento de 20 minutos.

5.2. Construcción de modelo sustituto

Como se ha mencionado, el método a ocupar es el método Kriging o regresión en procesos Gaussianos. En esta sección se definen las características del método, se describe el proceso de entrenamiento para finalmente realizar la validación del modelo.

El método se implementa en base a las herramientas y funciones otorgadas por la librería OpenTURNS [20].

5.2.1. Definición de características de método

5.2.1.1. Función base

Se define una función base de tipo cuadrática como sigue:

$$h(x) = \{1, x_1, \dots, x_n, x_i x_j\} \quad (5.2)$$

Se selecciona este tipo debido a que se presenta un problema físico con características no lineales, además de su uso típico por la librería utilizada para el procedimiento de entrenamiento (OpenTURNS).

5.2.1.2. Función de covarianza (Kernel)

La función de covarianza o kernel es de tipo exponencial cuadrática:

$$k(x_i, x_j) = \sigma^2 \exp\left(-\frac{(x_i - x_j)^2}{2l^2}\right) \quad (5.3)$$

Su elección se basa en sus capacidades de otorgar propiedades en el modelo de suavidad, es decir correlacionar los puntos que están más cercanos entre sí, entregar propiedad de no periodicidad, y también su elección es respaldada por su uso típico con buenos resultados en diversas fuentes [21].

5.2.2. Entrenamiento del modelo

El entrenamiento del modelo se realiza implementando el algoritmo Kriging o regresión en procesos Gaussianos con el comando de la librería OpenTURNS "KrigingAlgorithm", el cual permite ingresar los siguientes argumentos:

i. Base de datos compuesta por:

- Datos de entrada (X)
- Datos de salida (Y)

ii. Función base

iii. Función de covarianza o Kernel

La función base y la función de covarianza o Kernel se establece con las ecuaciones 5.2 y 5.3 respectivamente. La base de datos está establecida en la Tabla 5.1.

Se establece la realización de dos modalidades de entrenamiento, de modo de obtener una mejor respuesta para cada variable de salida o predicción, sin tener una dimensionalidad muy alta que puede provocar problemas numéricos y de precisión en el método. Estos se presentan a continuación:

5.2.2.1. Modelo de FS global

El primer modelo sustituto recibe los cuatro datos de entrada y entrega el factor de seguridad global. Este modelo permite una única salida o respuesta con una alta precisión. Este recibe la siguiente base de datos:

Tabla 5.3: Formato de datos - FSG

nf1	nf2	ϕ	c	FS_g
$\{ \}_i$	$\{ \}_i$	$\{ \}_i$	$\{ \}_i$	$\{ \}_i$

5.2.2.2. Modelo de contorno de FS local

El segundo modelo sustituto recibe los cuatro datos de entrada y entrega el factor de seguridad global con los coeficientes de la curva a representar. Para cada índice j fijo, que representa un factor de seguridad local, se realiza un proceso de entrenamiento, de modo que se obtienen 10 modelos sustitutos, asociados a: FSL=[1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9]. Cada uno de estos modelos permite representar una curva asociada al contorno de un factor de seguridad local.

Cabe mencionar que se utiliza esta discretización (0.1), ya que esta permite obtener valores de interés, y un intervalo menor no entrega información considerablemente significativa para analizar. También se debe comentar que el factor de seguridad global no tiene la misma precisión que el primer modelo, debido a su mayor dimensionalidad en los parámetros de salida, y este valor de factor de seguridad global más bien sirve para lograr una compatibilidad entre los coeficientes.

Los datos de entrenamiento de estos "submodelos" tienen el siguiente formato:

Tabla 5.4: Formato de datos - Contorno FS

nf1	nf2	ϕ	c	FS_g	C0	C1	C2	C3
$\{ \}_i$	$\{ \}_i$	$\{ \}_i$	$\{ \}_i$	$\{ \}_i$	$\{ \}_{i,j}$	$\{ \}_{i,j}$	$\{ \}_{i,j}$	$\{ \}_{i,j}$

5.2.3. Validación del modelo

La validación del modelo se realiza mediante el método de validación cruzada (Marco Teórico). Dicho método se realiza para cada modelo sustituto, esto es para el modelo de FS global y para cada submodelo del modelo del contorno de FS.

5.3. Resultados de modelo sustituto

5.3.1. Modelo de FS global

5.3.1.1. Tiempo de entrenamiento

A continuación se presenta el tiempo de ejecución del primer modelo a analizar. Este tiempo incluye el tiempo de entrenamiento del modelo más el tiempo de validación del modelo.

Cabe mencionar que el procedimiento de validación conlleva cerca del 85 % del tiempo total de procesamiento de cada modelo sustituto (sea este modelo o submodelo).

El tiempo de ejecución de este modelo sustituto es de 1 minuto aproximadamente.

5.3.1.2. Predicción

Una vez realizados los procesos de entrenamiento y de validación del modelo sustituto, ya es posible realizar predicciones, ingresando nuevos valores de entrada no considerados en la base de datos de entrenamiento.

Los pasos que sigue este proceso son los siguientes:

- i. En primer lugar se determina el nuevo conjunto de valores de entrada
- ii. Estos valores se deben normalizar antes de entrar el modelo sustituto, debido a que este está entrenado con esta condición
- iii. Se ingresan este conjunto de nuevos valores normalizado al modelo sustituto
- iv. Se obtiene la respuesta normalizada
- v. La respuesta se debe transformar nuevamente a espacio físico, es decir se realiza el procedimiento inverso a la normalización.
- vi. Se obtiene respuesta de modelo sustituto referido a nuevo conjunto de valores de entrada

El tiempo de respuesta del sistema es aproximadamente de un segundo, es decir, la respuesta es prácticamente inmediata.

5.3.1.3. Validación

Es de suma importancia verificar el resultado de la validación del modelo sustituto, debido a que este procedimiento permite obtener la precisión del modelo, y por ende la viabilidad de su uso.

Cabe mencionar que el error residual dado por el entrenamiento del modelo, es decir, dado por los datos de soporte o base de datos y la predicción de los mismos, es prácticamente 0 para este modelo, tal y como se demuestra en la siguiente Tabla:

Tabla 5.5: Errores - Datos de soporte - Modelo de FS global

Caso	E. resid.
1	0
2	4e-17
...	...
n	3e-16
Promedio	8e-17

Es posible identificar en la siguiente Tabla, los errores de la validación cruzada asociados a la salida dada por el factor de seguridad global, y también representado gráficamente en la Figura 5.5. Los valores y promedio del error residual se toma a partir de los valores absolutos.

Tabla 5.6: Errores - Valid. Cruzada - Modelo de FS global

Caso	E. resid.	E. resid. cuad.
1	0.021	0.0005
2	0.018	0.0003
...
n	0.012	0.0002
Promedio	0.0102	0.0002

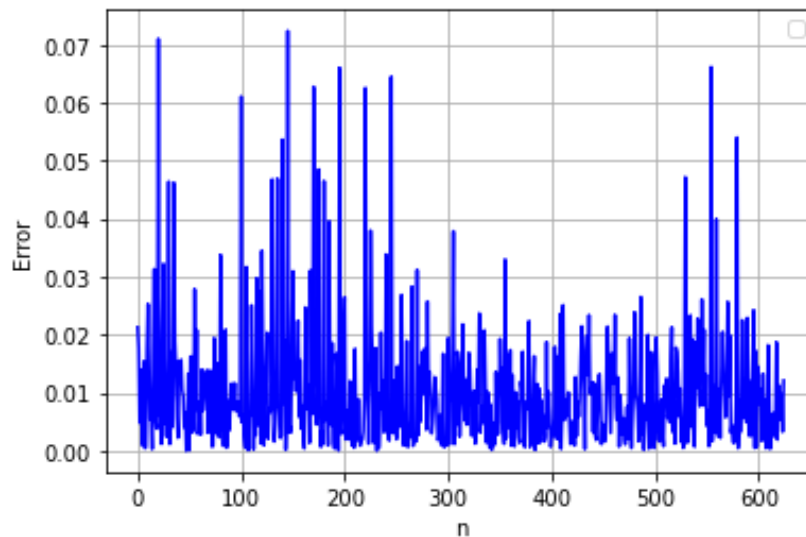


Figura 5.5: Errores residuales Modelo FS global

Se tiene entonces los siguientes errores representativos:

Tabla 5.7: Errores - Modelo de FS global

Error	Valor
Error absoluto medio	0.0102
RMSE	0.015
R^2	0.994
Err _{LOO}	0.0002

A continuación se evalúa el modelo sustituto, comparando las Figuras 4.5 a 4.8 presentadas en la sección 4.6.

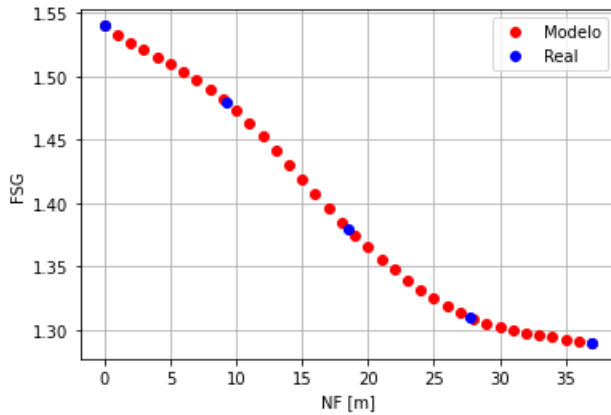


Figura 5.6: FSG versus nivel freático - Modelo sustituto

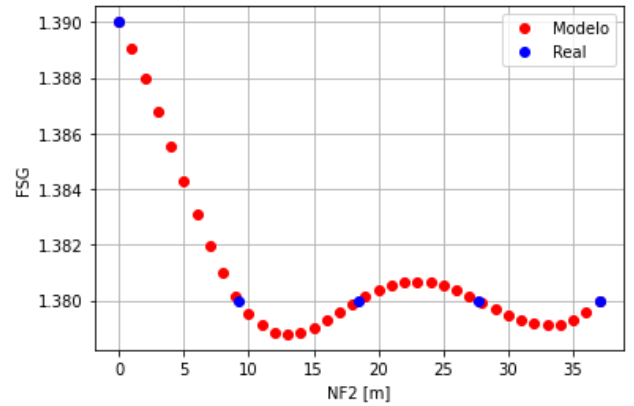


Figura 5.7: FSG versus nivel freático 2 - Modelo sustituto

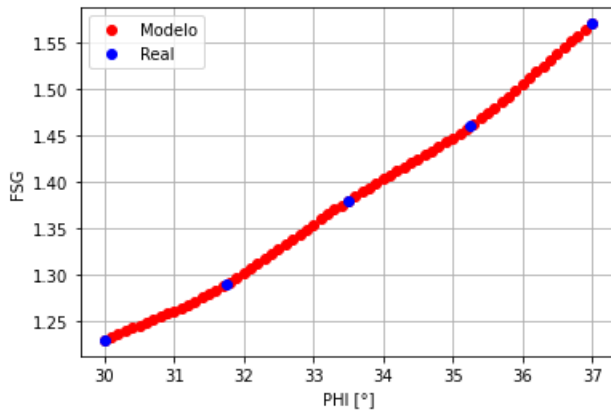


Figura 5.8: FSG versus ángulo de fricción (ϕ) - Modelo sustituto

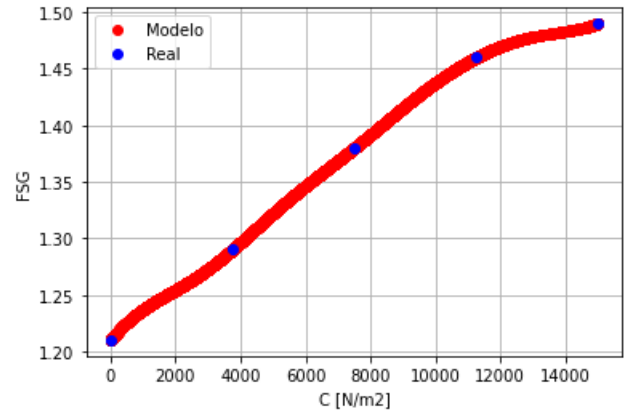


Figura 5.9: FSG versus cohesión - Modelo sustituto

Las Figuras anteriores muestra en color azul un conjunto de los puntos de los datos reales dados por la base de datos usadas para el entrenamiento, y en color rojo un conjunto de datos probados en el modelo sustituto ya confeccionado.

En base a los casos de prueba obtenidos, se establecen los errores de prueba o test absolutos en la siguiente Tabla:

Tabla 5.8: Errores - Test - Modelo de FS global

Caso	E. resid.	E. resid. cuad.
1	0.007	5e-5
2	0.008	6e-5
...
n	0.005	3e-5
Promedio	0.007	7e-5

Se tiene entonces los siguientes errores representativos:

Tabla 5.9: Errores - Modelo de FS global

Error	Valor
Error absoluto medio	0.007
RMSE	0.0081
R ²	0.998
Err _{LOO}	7e-5

5.3.2. Modelo de contorno de FS

5.3.2.1. Tiempo de entrenamiento

El tiempo de procesamiento de los submodelos se describe en la siguiente Tabla:

Tabla 5.10: Tiempo de ejecución de submodelos de contorno de FS

Submodelos	T. por submodelo [hrs]	T. total [hrs]	T. total [días]
10	3	30	1.25

5.3.2.2. Predicción

Para este modelo se realiza el mismo procedimiento para la predicción del modelo anterior. En este caso el tiempo de respuesta del sistema es aproximadamente de un minuto.

5.3.2.3. Validación

En primer lugar se verifican los errores residuales respecto a los datos de soporte o de entrenamiento. Estos como se puede ver no son considerablemente bajos, debido a la dimensionalidad aumentada en la salida para este modelo, sin embargo, siguen siendo aceptables, más si se verifica el sentido de cada parámetro.

Tabla 5.11: Errores - Datos de soporte - Modelo de contorno de FS

Error	1.0	1.1	1.2	1.3	1.4
FSG	5e-17	6e-17	3e-17	7e-17	4e-17
C0	3e-13	0.633	0.83	0.999	0.912
C1	3e-13	0.061	0.033	0.092	0.057
C2	4e-15	0.001	0.001	0.001	0.001
C3	3e-17	7e-6	5e-6	1e-5	9e-6

Tabla 5.12: Continuación Tabla 5.11

Error	1.5	1.6	1.7	1.8	1.9	Prom.
FSG	5e-17	6e-17	6e-17	8e-17	7e-17	5e-17
C0	1.039	0.977	1.13	0.942	0.676	0.814
C1	0.058	0.068	0.038	0.074	0.068	0.055
C2	0.001	0.001	0.001	0.002	0.002	0.001
C3	1e-5	1e-5	7e-6	1e-5	1e-5	7e-6

Se identifica en la siguiente tabla, los errores absolutos asociados a la salida dada por el FS global y los coeficientes de cada curva, dado por la validación cruzada. En primer lugar se describen los errores residuales para cada FS local. Para esta ocasión se denotan los promedios de los errores residuales para cada parámetro de salida.

Tabla 5.13: Errores residuales - Valid. Cruzada - Modelo de contorno de FS

Error	1.0	1.1	1.2	1.3	1.4
FSG	0.0116	0.0098	0.0112	0.0101	0.0111
C0	0.3482	0.7705	1.2101	1.3369	1.5077
C1	0.025	0.0419	0.2292	0.1216	0.2124
C2	0.0004	0.0011	0.0059	0.0031	0.0054
C3	4e-6	9e-6	4e-5	2e-5	3e-5

Tabla 5.14: Continuación Tabla 5.13

Error	1.5	1.6	1.7	1.8	1.9	Prom.
FSG	0.0098	0.0104	0.012	0.0105	0.0115	0,0108
C0	1.4457	1.4462	1.1508	0.5275	0.5238	1,0267
C1	0.2205	0.1518	0.1684	0.0824	0.0899	0,1343
C2	0.006	0.0043	0.0047	0.0026	0.0029	0,0036
C3	4e-5	3e-5	3e-5	2e-5	2e-5	2e-5

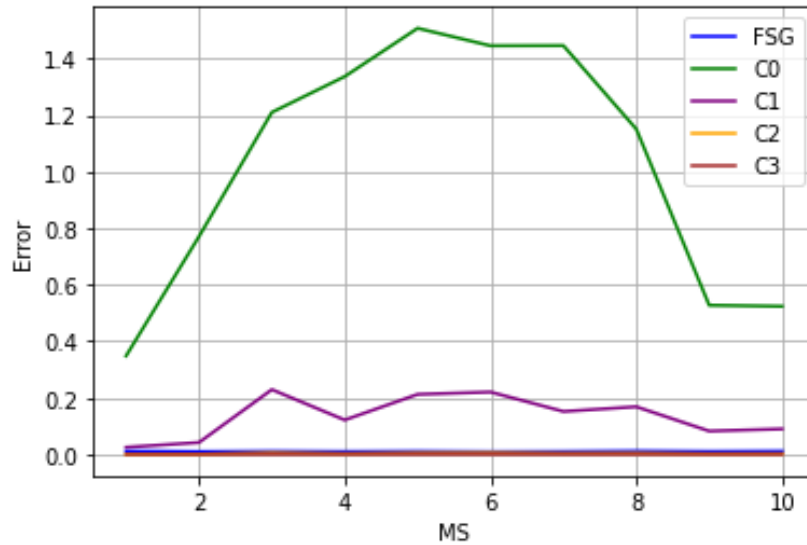


Figura 5.10: Errores residuales Modelo Contorno FS

Se tienen además los errores RMSE y se adiciona un promedio asociado a los errores LOO de cada submodelo.

Tabla 5.15: Errores RMSE y LOO - Valid. cruzada Modelo de contorno de FS

Errores	1.0	1.1	1.2	1.3	1.4
FSG	0.017	0.014	0.017	0.014	0.014
C0	1.509	2.866	3.194	2.787	2.851
C1	0.091	0.093	1.023	0.212	0.717
C2	0.003	0.003	0.028	0.006	0.02
C3	2e-5	2e-5	2e-4	5e-5	1e-4

Tabla 5.16: Continuación Tabla 5.15

Errores	1.5	1.6	1.7	1.8	1.9	RMSE	Err_{LOO}
FSG	0.014	0.014	0.017	0.014	0.017	0.015	0.0002
C0	2.623	2.944	2.113	0.961	1.224	2.307	5.9027
C1	0.611	0.336	0.836	0.152	0.219	0.429	0.2878
C2	0.017	0.009	0.022	0.005	0.006	0.012	2e-4
C3	1e-4	6e-5	1e-4	3e-5	5e-5	7e-5	8e-9

A continuación se evalúa el modelo sustituto, comparando las Figuras 5.11 a 5.13 presentadas en la sección 4.6.

Tabla 5.17: Datos - Caso más favorable - Modelo sustituto

nf1	nf2	ϕ	c	FS_g
0	0	37	15e3	1.88

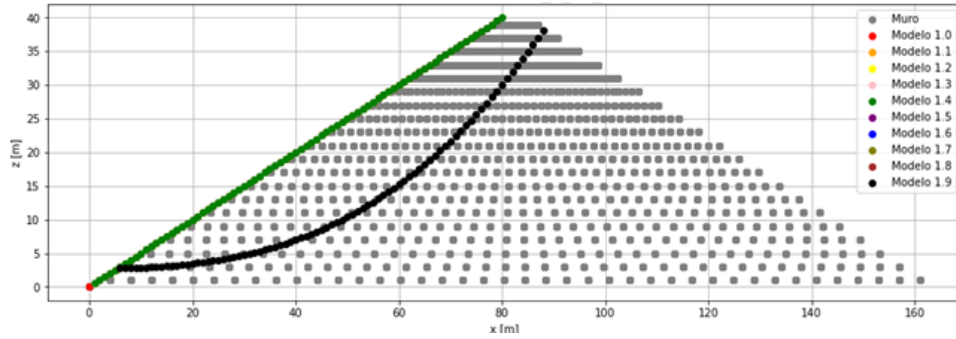
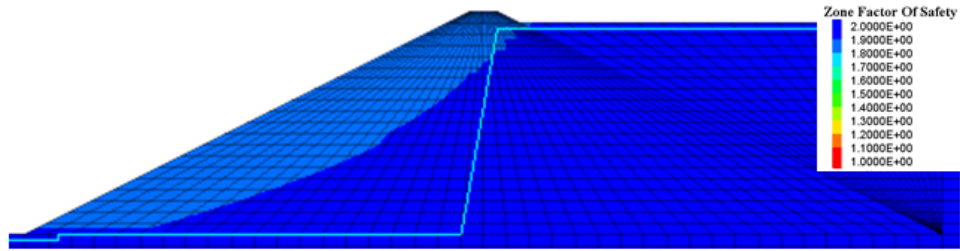


Figura 5.11: Comparación - Caso más favorable - Contorno FS - FSG=1.88

Tabla 5.18: Datos - Caso intermedio - Modelo sustituto

nf1	nf2	ϕ	c	FS_g
18.5	18.5	33.5	7500	1.3813

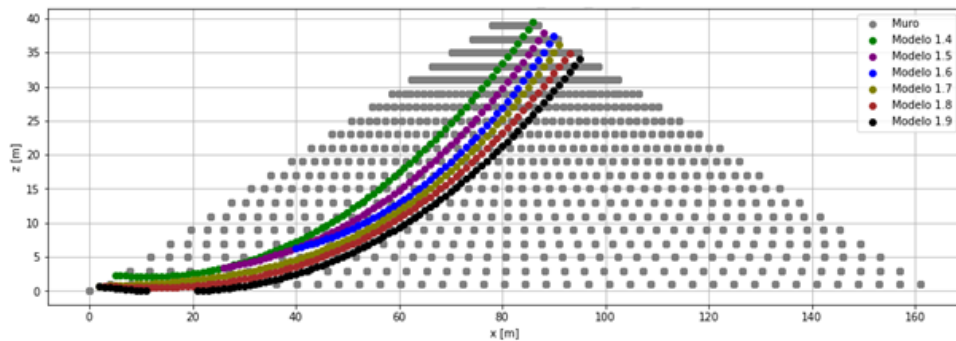
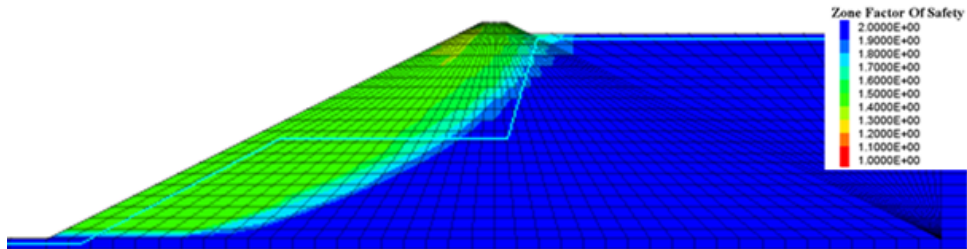


Figura 5.12: Comparación - Caso Intermedio - Contorno FS - FSG=1.3813

Tabla 5.19: Datos - Caso más desfavorable - Modelo sustituto

nf1	nf2	ϕ	c	FS_g
37	37	30	0	1.006

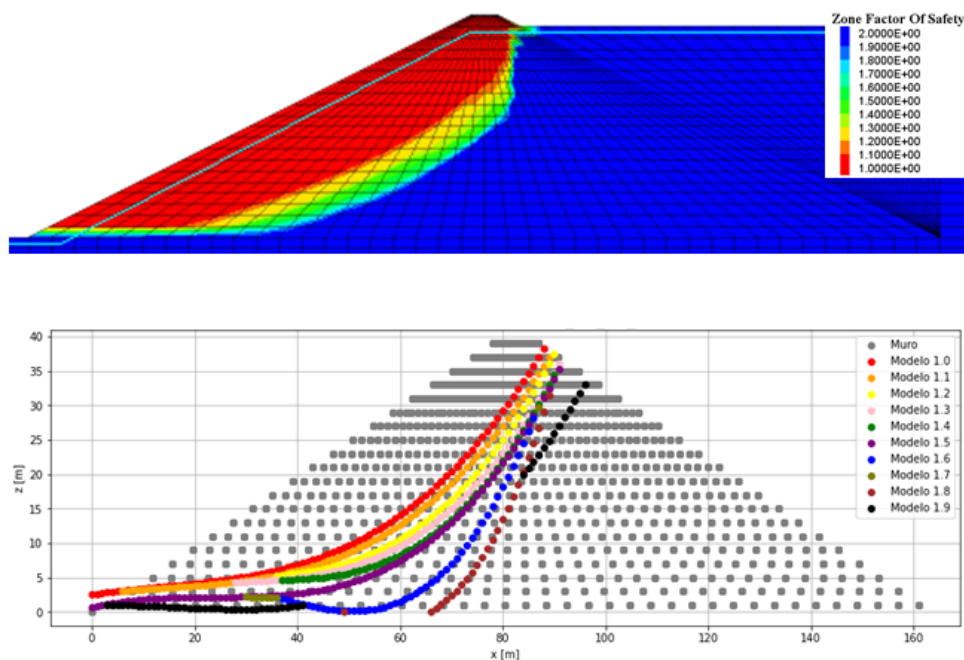


Figura 5.13: Comparación - Caso más desfavorable - Contorno FS - $FS_g=1.006$

Por otro lado, en base a los casos de prueba obtenidos, se establecen los errores de prueba o test absolutos en la siguiente Tabla:

Tabla 5.20: Errores residuales - Test - Modelo de contorno de FS

Error	1.0	1.1	1.2	1.3	1.4
FSG	0.016	0.014	0.018	0.007	0.018
C0	0.145	0.355	0.767	1.826	0.52
C1	0.013	0.038	0.147	0.22	0.21
C2	7e-5	0.0003	0.003	0.002	0.004
C3	2e-6	3e-06	2e-5	2e-5	3e-5

Tabla 5.21: Continuación Tabla 5.20

Error	1.5	1.6	1.7	1.8	1.9	Prom.
FSG	0.018	0.011	0.01	0.005	0.009	0.013
C0	1.274	2.609	1.22	0.39	0.966	1.007
C1	0.119	0.14	0.101	0.063	0.056	0.111
C2	0.003	0.004	0.002	0.002	0.002	0.002
C3	2e-5	3e-5	2e-5	2e-5	1e-5	2e-5

Se tienen los errores RMSE y el promedio asociado a los errores LOO de cada submodelo.

Tabla 5.22: Errores RMSE y LOO - Test - Modelo de contorno de FS

Errores	1.0	1.1	1.2	1.3	1.4
FSG	0.018	0.016	0.02	0.008	0.02
C0	0.155	0.6	0.947	2.08	0.679
C1	0.013	0.065	0.179	0.293	0.224
C2	7e-5	0.0004	0.003	0.003	0.005
C3	2e-6	5e-6	2e-5	3e-5	4e-5

Tabla 5.23: Continuación Tabla 5.22

Errores	1.5	1.6	1.7	1.8	1.9	RMSE	Err_{LOO}
FSG	0.02	0.012	0.011	0.007	0.01	0.014	0.0002
C0	1.443	4.734	1.563	0.554	1.567	1.432	3.577
C1	0.134	0.209	0.132	0.082	0.068	0.14	0.026
C2	0.004	0.007	0.003	0.003	0.002	0.003	1e-5
C3	3e-5	5e-5	2e-5	3e-5	2e-5	2e-5	7e-10

Finalmente, se puede visualizar la comparación en las Figuras 5.14 y 5.15.

Tabla 5.24: Datos - Test 1 - Modelo sustituto

nf1	nf2	ϕ	c	FS_g
12.33	12.33	32.33	5e3	1.331

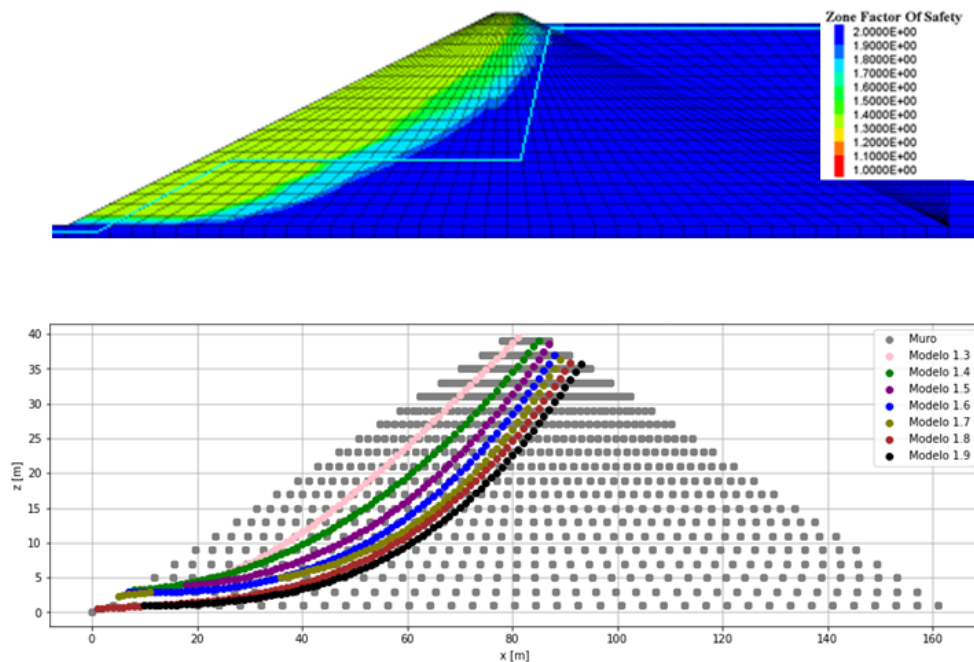


Figura 5.14: Comparación - Test - Contorno FS - FSG=1.331

Tabla 5.25: Datos - Test 2 - Modelo sustituto

nf1	nf2	ϕ	c	FS_g
24.66	24.66	34.66	1e4	1.431

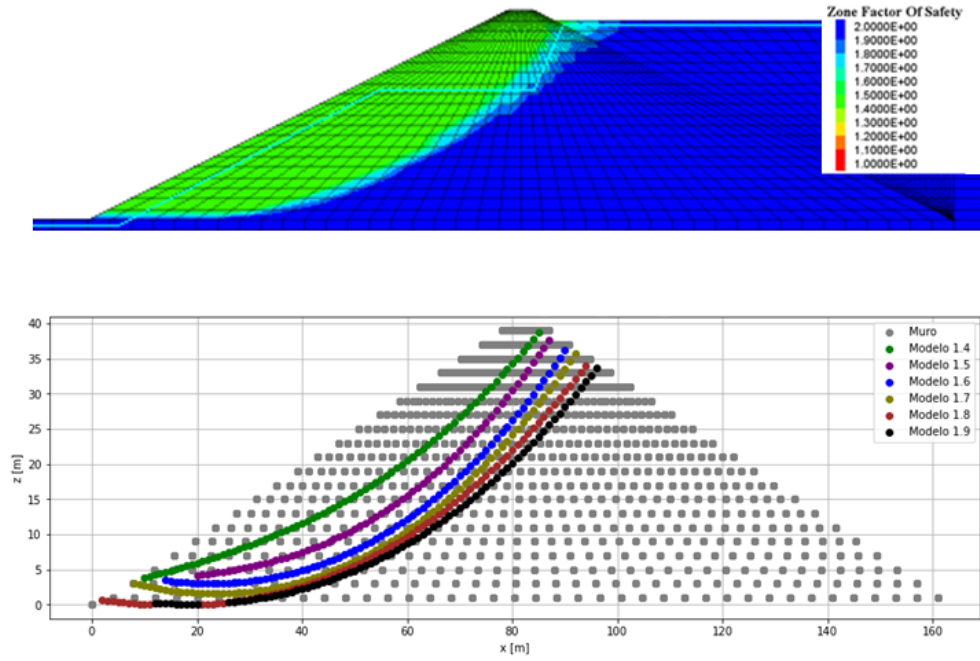


Figura 5.15: Comparación - Test - Contorno FS - FSG=1.431

Capítulo 6

Cuantificación de Incertidumbre

En el siguiente capítulo se explica la implementación de los métodos de cuantificación de incertidumbre en el problema a analizar del embalse de relaves. Estos procedimientos permiten denotar e interpretar las incertezas o la variabilidad que está presente en las variables involucradas, de modo de identificar como se ve influenciada la respuesta del sistema o salida, y presentar la variabilidad resultante respecto a esta última.

La implementación de estos procedimientos se realizan en base a los modelos sustitutos obtenidos en el capítulo anterior. Gracias a su formulación es posible la aplicación de los métodos de cuantificación de incertidumbre de un modo ágil y eficiente, lo cual no sería posible realizar de esta forma si se utiliza directamente el modelo desarrollado en el software de diseño FLAC3D.

6.1. Modelo de FS

Para este modelo que entrega el factor de seguridad como única variable de salida, es significativa la aplicación de los tres métodos de cuantificación de incertidumbre, tal y como se plantean a continuación.

6.1.1. Planteamiento

La función $M(\theta)$ representa el modelo sustituto confeccionado en el capítulo anterior, el cual recibe un conjunto de variables de entrada y entrega un conjunto valores de salida. Se utilizará el modelo sustituto referido al modelo de FS global.

El conjunto de valores de entrada está representado por X . La información estadística de las variables de entrada, que en este caso sería el nivel freático, nivel freático 2, ángulo de fricción y cohesión, esto es la distribución de probabilidad de las mismas, será determinada en base a las características del problema, lo cual implica la fuente de obtención, que puede ser por ejemplo por mediciones en terreno, datos experimentales, datos artificiales por softwares, entre otras. La información establecida para cada variable se presenta a continuación:

Tabla 6.1: Información estadística de variables de entrada

Variable	Distribución	Media	Desv. Est.	c.o.v.
Nivel freático	Normal	18.5	6.17	0.333
Nivel freático 2	Normal	18.5	6.17	0.333
Ángulo de fricción	Normal	33.5	1.675	0.05
Cohesión	Normal	7500	2500	0.333

Se considera un coeficiente de correlación igual a 0 para todos los casos, lo cual indica que las variables no presentan una relación directa entre sí o una correlación. Esta decisión es conservadora, y se comenta en el capítulo de Análisis y Discusión.

Es posible denotar la distribución conjunta de estas variables del siguiente modo.

$$\{f\} \sim N \left(\begin{bmatrix} \mu_{nf} \\ \mu_{nf2} \\ \mu_{\phi} \\ \mu_c \end{bmatrix}, \begin{bmatrix} \sigma_{nf}^2 & 0 & 0 & 0 \\ 0 & \sigma_{nf2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{\phi}^2 & 0 \\ 0 & 0 & 0 & \sigma_c^2 \end{bmatrix} \right) \quad (6.1)$$

Estas distribuciones se representan gráficamente como sigue:

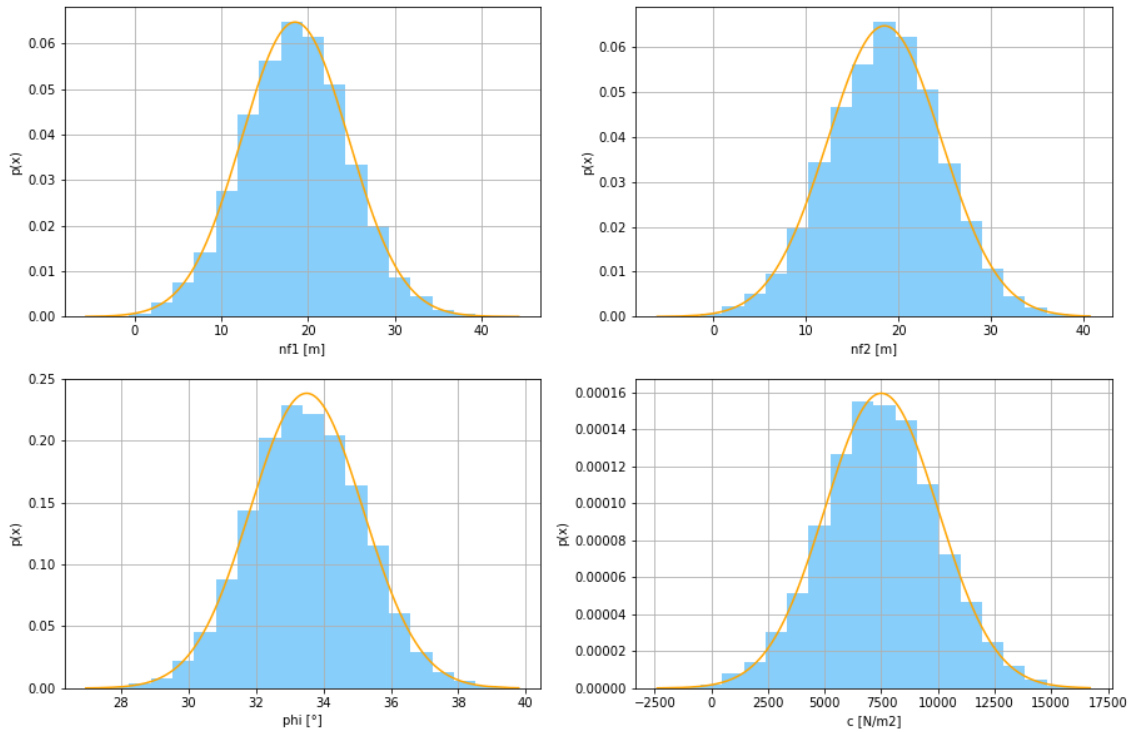


Figura 6.1: Gráfica de PDFs de variables de entrada

6.1.2. Propagación de incertidumbre

Para realizar este procedimiento se establece el uso del método de Monte Carlo con muestreo en estándar Gaussiano. Se plantea el problema en base a la siguiente expresión:

$$FSG = H = E[M(\theta)] = \int_{\theta_1} \dots \int_{\theta_n} M(\theta)p(\theta_1)\dots p(\theta_n)d\theta_1\dots d\theta_n \quad (6.2)$$

y mediante el método mencionado se tiene el estimador:

$$\hat{H} = \frac{1}{K} \sum_i M(\theta_i) \quad (6.3)$$

donde $\theta = \{\text{nf}, \text{nf2}, \phi, c\}$.

Se comienza obteniendo K muestras de una distribución estándar Gaussiano (K se establece como 10000). Luego en base a la distribución en la ecuación 6.1, se realiza una descomposición de Cholesky inferior C. Se realiza la transformación a espacio físico para finalmente las muestras de entrada se reemplazan en la función $M(\theta)$ y se obtiene las muestras de salida. Estas se representan a continuación en el histograma de la Figura 6.2, y como PDF en la Figura 6.3.

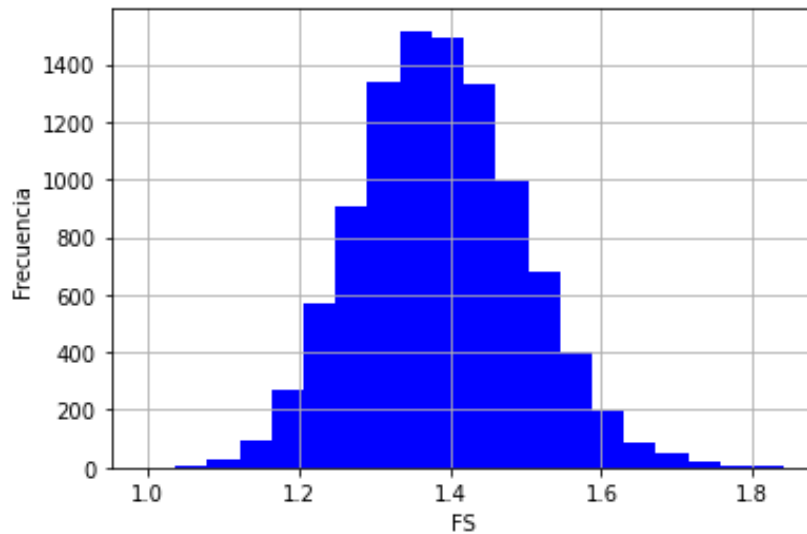


Figura 6.2: Histograma - FSG

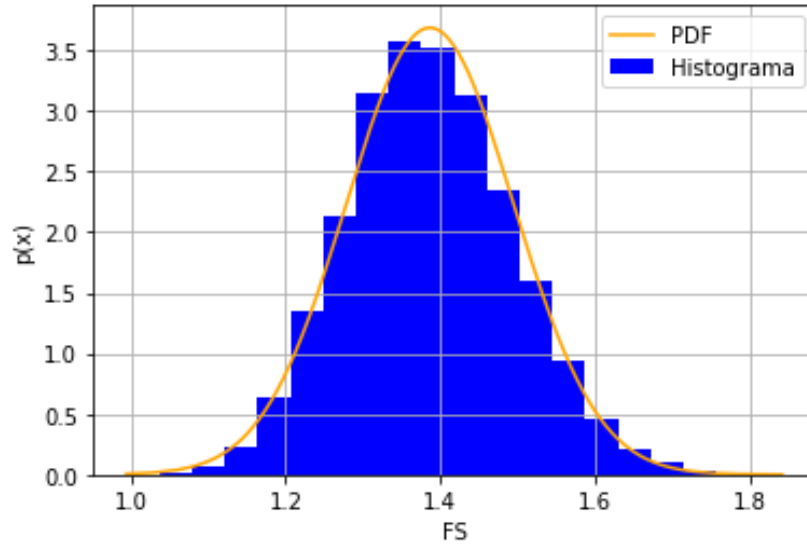


Figura 6.3: PDF - FSG

Es posible obtener los siguientes valores estadísticos a partir de las muestras de salida (Factor de seguridad global).

$$\mu_{FSG} = 1.385 \quad (6.4)$$

$$\sigma_{FSG} = 0.107 \quad (6.5)$$

$$c.o.v.MC = 0.078\% \quad (6.6)$$

6.1.3. Probabilidad de falla

El procedimiento de cálculo de probabilidad de falla es necesario para determinar la probabilidad de que el modelo no se reduzca a cierto valor límite, es decir que el factor de seguridad no sea menor a cierto valor.

La función $I_F(\theta)$ queda definida como sigue, en este caso adaptando la desigualdad debido a que se busca que la función sea 1 para valores menores al valor límite, el cual se establece como lo requiera el problema físico. En este caso, este valor se determina por las bases de diseño del embalse de relaves, donde se indica el valor límite permitido por normativa, el cual es 1.4 (normativa chilena).

$$I_F = \begin{cases} 0, & \text{si } g(\theta) > \text{valor lim} \\ 1, & \text{si } g(\theta) \leq \text{valor lim} \end{cases} \quad (6.7)$$

En la siguiente gráfica se visualiza la probabilidad de que la respuesta del sistema (factor de seguridad), adquiera valores bajo determinado valor límite de forma general.

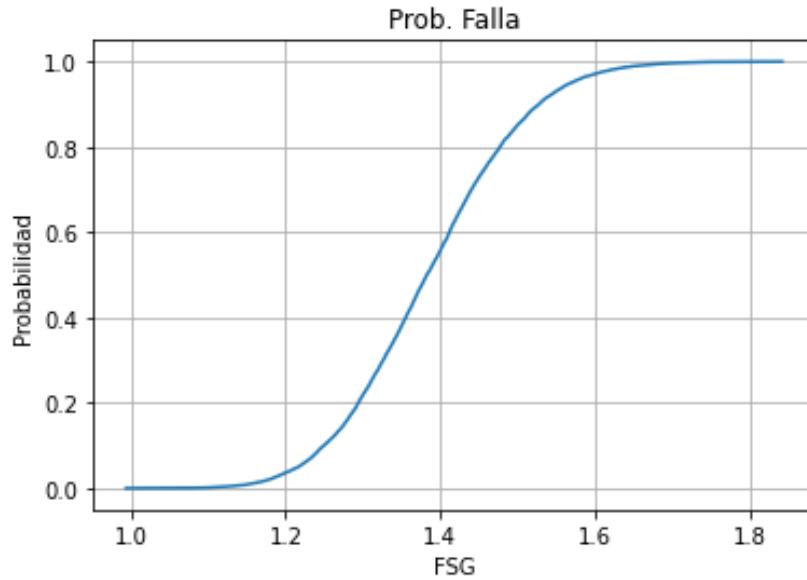


Figura 6.4: Gráfica de Probabilidad de Falla

6.1.4. Análisis de sensibilidad

Este análisis se realiza mediante el cálculo de los índices de Sobol, utilizando la librería OpenTURNS y su algoritmo respectivo, denominado SobolIndicesAlgorithm, y la estimación es realizada por el algoritmo de Saltelli, denominado SaltelliSensitivityAlgorithm. Este recibe muestras de entrada y salida, y el tamaño de las muestras, las cuales son previamente procesadas por el algoritmo SobolIndicesExperiment, que genera muestras con el formato adecuado. De este modo se obtienen los siguientes índices de Sobol:

Tabla 6.2: Análisis de sensibilidad - Índices de Sobol

Variable	Í. Sobol 1° Orden	Í. Sobol Orden Total
Nivel freático	0.213	0.222
Nivel freático 2	-0.007	0.002
Ángulo de fricción	0.524	0.555
Cohesión	0.223	0.236

6.2. Modelo de contorno de FS

Para este modelo, el análisis significativo es en base a la aplicación del método de propagación de incertidumbre, a partir del cual es posible ver la variabilidad de las curvas de contorno.

6.2.1. Planteamiento

En este caso, la información estadística se plantea de forma menos variable, es decir, existe menor dispersión de los datos. Esto para que se refleje de mejor manera la situación planteada y la representación del método de propagación incertidumbre aplicado, sin pérdida de generalidad.

Tabla 6.3: Información estadística de variables de entrada - Modelo contorno de FS

Variable	Distribución	Media	Desv. Est.	c.o.v.
Nivel freático	Normal	18.5	3.083	0.166
Nivel freático 2	Normal	18.5	3.083	0.166
Ángulo de fricción	Normal	33.5	0.838	0.025
Cohesión	Normal	7500	1250	0.166

6.2.2. Propagación de incertidumbre

Realizando la propagación de incertidumbre, en base a la misma formulación dada en la sección anterior, variando el número de muestras a 1000, se obtienen esta vez resultados gráficos, que representan la distribución de la salida, que en este caso son contornos y asociados a un respectiva curva con un determinado factor de seguridad 'local' etiquetado, tal como se presenta en la siguiente Figura:

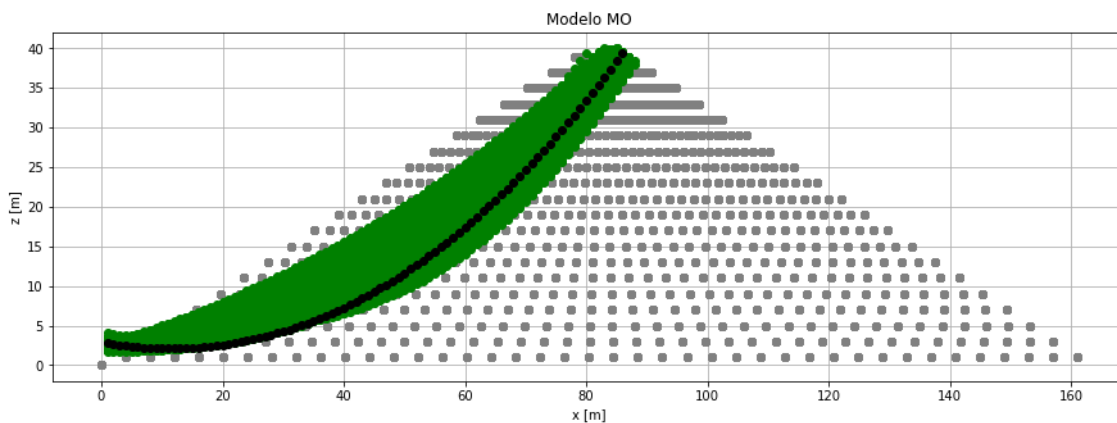


Figura 6.5: Propagación de incertidumbre - Modelo de Contorno de FS - FS Local=1.4

De este modo se refleja un determinado caso, y se realiza énfasis en una curva asociada a un factor de seguridad local en particular. Así es posible identificar el espacio o zona donde se puede ubicar esta curva de contorno, la cual en este caso está definida por la reducción de las muestras dentro de los percentiles 3 y 97, abarcando un rango máximo de aproximadamente 25 metros para el caso presentado. Por otro lado, la información estadística se establece para

cada variable de salida y se registra en la siguiente Tabla. Así verificando los valores de $c.o.v_{MC}$, se toma una referencia de la precisión del método de Monte Carlo y su validación.

Tabla 6.4: Información estadística de variables de salida - Modelo contorno de FS

Variable	Media	Desv. Est.	c.o.v	c.o.v._{MC}
FSG	1.39	0.053	0.0378	0.12
C0	2.619	0.38	0.145	0.459
C1	0.072	0.061	0.843	2.664
C2	0.002	0.001	0.704	2.228
C3	3e-5	1e-5	0.47	1.487

Capítulo 7

Análisis y Discusión de Resultados

En este capítulo se pretende analizar los resultados obtenidos en las tres etapas que conforman el trabajo de memoria, de modo de verificar, validar y respaldar su obtención, y así también verificar los argumentos y fuentes teóricas que presenta la metodología que se desarrolla a lo largo del documento. Se busca realizar una descripción más detallada de los datos, añadiendo juicios de valor, realizando una discusión crítica de los mismos, evaluando la calidad de estos y efectuando análisis comparativos.

Se pretende explicar las aplicaciones del presente trabajo, sus alcances, limitaciones, además de posibles generalizaciones del trabajo y proyecciones a evaluar dentro del mismo o fuera del campo de acción que tiene esta memoria.

7.1. Diseño de modelo de embalse de relaves

7.1.1. Software de modelamiento FLAC3D

El software seleccionado para realizar el modelamiento, FLAC3D, presenta múltiples argumentos para su uso. En primer lugar, la empresa Itasca, colaborante del presente trabajo y creadora del software FLAC3D, verifica la factibilidad de su uso, mediante aplicaciones en el rubro, en el cual se presenta menciones constantes en diferentes documentos de la academia, así como en el ámbito laboral y empresarial. En términos técnicos, el software presenta fundamentos teóricos adecuados, que permiten que su uso en este tipo de problemas, como lo es el análisis de estabilidad física o estabilidad de taludes en embalse de relaves sea eficiente. En base a múltiples documentos se hace referencia al software FLAC3D, así como también al software Plaxis y Abaqus, cuyo análisis y prueba es interesante para la comparación y factibilidad de la diversificación de estos procedimientos y la metodología impuesta en este trabajo. Existen diferencias entre los softwares de modelamiento, donde se presenta FLAC3D, aún cuando el uso y su interfaz conlleva mayor complejidad que otros en primera instancia, como una alternativa sobresaliente para el modelamiento de embalse de relaves, por la variabilidad de opciones, comandos, manipulabilidad de los parámetros del diseño, mejoras en la convergencia, entre otros comentarios, que son mencionados en el rubro a favor de este software.

7.1.2. Definición de modelo

El modelo descrito consiste en el análisis de una sección de un embalse de relaves compuesto por un muro de arena y base de roca. El hecho de analizar una sección, en vez del modelo completo en tres dimensiones (opción disponible en el software FLAC3D), es debido al alcance computacional que esto conlleva, ya que el tiempo de ejecución del modelo para una sección es mucho menor y viable. Esto de cierto modo puede quitar generalidad para la comprensión del comportamiento del modelo, lo cual busca resolverse al seleccionar la sección más crítica del embalse, de manera que se esté en el caso más desfavorable y con mayor posibilidad de falla. Esta condición implica que el modelo adquiera un carácter sumamente conservador, lo cual puede involucrar un costo económico, pero conlleva una seguridad considerable en el diseño.

El planteamiento de un modelo en tres dimensiones o un modelo que incluya otras área de cálculo, requiere tomar ciertas consideraciones: en primer lugar se debe considerar el tiempo que conlleva la ejecución de un determinado caso y a partir de este plantear las variables significativas involucradas para así determinar la cantidad y el formato de los casos o combinaciones de variables de entrada que se utilizarían. En caso de ser un tiempo de ejecución considerablemente alto, entonces se puede proponer como solución, el optar por reducir la cantidad de combinaciones y también la utilización de una mayor cantidad de servidores o equipos para el procesamiento de estos casos.

La metodología planteada en este trabajo permite adaptarse frente a un distinto diseño del embalse de relaves presentado y aplicar la misma para la confección del modelo sustituto y aplicación de la cuantificación de incertidumbre. Es importante aquí comprobar de forma efectiva el modelo que se utilizará, haciendo énfasis en la verificación de la convergencia del modelo para los casos extremos más desfavorables, es decir, delimitar, abarcar y tener claridad del espacio muestral, donde se encuentran las variables involucradas, así como también las variables de salida o respuestas que entrega el modelo.

Se establece la realización de un análisis estático para el sistema, debido a que es el proceso estándar y permite comprender la respuesta del sistema sin estar presto a factores que estén fuera del enfoque del trabajo, el cual es explicativo y con objetivo a la confección de modelos sustitutos. La normativa referida describe aspectos relacionados al estudio sísmológico e hidráulico, los cuales como se mencionó anteriormente no se analizarán en detalle en el presente documento. Sin embargo su inclusión en la metodología es de sumo interés e importancia. Es posible entonces proyectar su aplicación en áreas asociadas a: un estudio o análisis pseudoestático, estudio o análisis dinámico, estudio hidráulico, entre otros.

7.1.3. Modelo planteado

El modelo particular realizado en el presente trabajo se plantea de forma referencial y basándose en información de bibliografía referida y en una base de datos asociada a Itasca, de modo que el modelo sea representativo y lo más realista posible.

La definición del nivel freático es una decisión significativa del sistema, ya que verificando el objetivo del caso práctico a aplicarse, el cual consiste en el sistema automatizado de monitoreo

del embalse, el cual recibe información piezométrica, indica que se debe poner énfasis en este parámetro. Entonces se establece que el nivel freático debe ser trazado mediante dos o más puntos modelados dentro del muro, que a su vez quedan definidos por ubicaciones de los instrumentos de medición, los cuales se definen a su vez de forma referencial. Por ende, se consideran dos puntos dentro del nivel freático, los cuales permiten trazar el mismo, y el próximo objetivo es verificar la influencia de este, o también evaluar si la diferencia entre estos niveles es significativa o no.

Por otro lado, las condiciones de borde, plantean el contexto adecuado para un análisis de estabilidad de taludes, dejando libres los grados de libertad en x y z , asociados al talud que se quiere examinar, el cual es el ubicado aguas abajo. Las condiciones iniciales establecen la condición estática, teniendo como sollicitación únicamente la gravedad.

Respecto a la ley constitutiva, se define un comportamiento Mohr-Coulomb, el cual está implementado en la mayoría de los softwares comerciales disponibles para la simulación mediante el método de los elementos finitos, necesitando determinados parámetros como datos de entrada, los cuales pueden ser determinados a través de los ensayos convencionales [42]. Esto otorga simplicidad en su uso, al ser determinado principalmente por el ángulo de fricción y la cohesión, y además muestra predicciones conservadoras de la resistencia cortante del suelo [43]. Es de interés la comparación en el uso de distintas leyes constitutivas, sin embargo no es el enfoque del presente trabajo, por lo que el comportamiento Mohr-Coulomb es adecuado para este fin, siendo simple y utilizado típicamente por la empresa Itasca y por el rubro.

Los criterios de convergencia, como bien se ha mencionado, son conservadores, al establecer una razón local de convergencia bastante menor al recomendado por la documentación de FLAC3D [13], lo cual otorga mayor seguridad y confiabilidad en los resultados.

Respecto al cálculo del factor de seguridad, el planteamiento del método de reducción de esfuerzos es adecuado para la determinación de los denominados factores de seguridad locales, los cuales son establecidos por cada paso de reducción de esfuerzos dada por el respectivo factor de seguridad que se itera. De este modo es posible trazar los contornos de factor de seguridad, que establecen las zonas más comprometidas o críticas, lo cual en términos de diseño y análisis de estabilidad es de suma importancia, ya que si existe mayor zona asociada a un factor de seguridad bajo, entonces se tiene una condición más inestable o insegura.

7.1.4. Selección de parámetros y casos

Se seleccionan parámetros que son adecuados e incidentes en el problema, y se selecciona una cantidad de variables y valores de variables basándose en un tiempo razonable y acorde a las necesidades de plazo del proyecto (en este caso, el proyecto es el presente trabajo de memoria, por lo que el tiempo es el contemplado en la planificación del mismo). Este procedimiento está detallado en el capítulo 5. Es importante realizar este cálculo de tiempos, ya que indica la viabilidad de la ejecución de los casos a modelar. Para ello se debe ser consciente de los equipos disponibles para uso y la capacidad técnica de los respectivos.

El método de muestreo dado por LHS, es bastante adecuado ya que como se visualiza en Anexo (C.), los gráficos para este método reflejan una distribución con una uniformidad y

dispersión correctas dentro del rango de acción de las variables, de modo que las muestras, puntos o casos a utilizar abarcan una gran cantidad de casos posibles, eso sí, no asegurando los casos extremos o de borde. Para el caso de valores equiespaciados en su rango no se presenta esta heterogeneidad de los valores dentro del rango, sin embargo, con esta configuración se asegura la selección de puntos de borde o extremos, y se pierde precisión, no significativamente, en datos intermedios. Debido a la importancia que presentan los casos extremos en el problema físico (Factores de seguridad bajos principalmente), se prioriza el método de muestreo dado por el equiespaciamiento en los rangos de valores de las variables, para asegurar la inclusión de estos casos en el modelo sustituto que se entrena.

Un análisis comparativo entre el rendimiento de un modelo sustituto con cierto número de valores por variable y otro con un número menor, es de sumo interés, ya que permite verificar si es posible realizar un ahorro considerable de tiempo en la ejecución de los casos. Este es realizado en Anexo y discutido en una próxima sección.

Es posible establecer un mayor número de variables y/o un mayor número de valores por variable, siempre y cuando exista la capacidad de servidores o computadores, para realizar las ejecuciones, ya que a mayor cantidad de estos aspectos, implica una mayor carga computacional, en términos de tiempo, por lo que al distribuir esta operación en mayor número de equipos, el tiempo reduce. Esto puede implicar un mayor costo económico en un proyecto, sin embargo la inversión en estos procesos, puede otorgar ahorros a largo plazo considerables, en los mismos términos económicos, por supuesto en tiempo de ejecución, y en términos técnicos otorgaría mayor precisión al modelo. Se insiste en este tema, ya que si bien la ejecución de los casos o combinaciones dado por un gran número de variables puede implicar un tiempo de demora considerable, este se puede reducir con mayor cantidad de servidores, y lo importante es que este tiempo y carga computacional se exigiría en una sola instancia, por que luego que se haya realizado, el programa queda listo para realizar cálculos y predicciones prácticamente inmediatas.

7.1.5. Resultados de FS y contorno de FS

Los resultados visualizados en las Figuras 4.5 a la 4.8, responden correctamente a la relación que tienen las variables de entrada con el FS global y el contorno de FS, de modo que una situación más inestable, se traduce en un mayor nivel freático, menor valor del ángulo de fricción (ϕ) y menor valor de cohesión y por ende un menor FS y zonas más críticas para el contorno de FS. Respecto a los contornos de FS, se visualizan zonas bien diferenciadas entre un factor de seguridad, denominado como 'local', y otro, lo que permite ver las zonas con mayor riesgo o inestabilidad. El factor de seguridad denominado como 'global' es el factor de seguridad menor obtenido luego del procedimiento de reducción de esfuerzos tal que exista la falla (fuerzas solicitantes mayores a las resistentes) y se empiezan a mostrar zonas diferenciadas en la geometría del embalse dados por la misma, es decir, el hecho de que se tenga un factor de seguridad 'global' de 1.4, implica que es el valor desde el cual se presenta la falla en determinadas zonas y este es el valor mínimo que se muestra en los contornos.

Es importante mencionar las singularidades, las cuales se representan en los contornos de FS, como zonas que no siguen el patrón de la mayoría de las demás zonas. Estas pueden presentarse debido a cambios bruscos en ciertos parámetros o variables de entrada respecto

a las demás, también frente a casos extremos al combinar las mismas variables de entrada y además por selección de discretizaciones muy o poco detalladas. Como se menciona en secciones previas (5.1.3), se determina que para el ajuste de estas singularidades, se agregan puntos estratégicamente, de modo de evitar que los valores de coeficientes de las curvas ajustadas cambien bruscamente. Por otro lado, para el caso de que no existan zonas de FS 'local', entonces se establece agregar una curva directamente en el talud del embalse.

7.2. Confección de modelo sustituto

7.2.1. Manipulación de datos

El paso previo a la confección del modelo sustituto consiste en la importación, filtración y manipulación de la base de datos obtenida de las ejecuciones del modelo en FLAC3D, el cual se detalla en capítulos anteriores. Respecto al modelo de FS, la base de datos a utilizar paracel entrenamiento es directa, no así para el modelo de contorno de FS, por lo que para este último cabe resaltar una discusión.

La manipulación referida al contorno de FS, consiste en que a partir del gráfico dado por los puntos que definen la geometría en la base de datos, se identifican los puntos de borde asociados a cada factor de seguridad 'local', los cuales permiten representar los bordes de cada zona referida al factor de seguridad mencionado, sin pérdida de generalidad, ya que el límite o interfaz de estos contornos son los que permiten identificar la ubicación de una zona u otra. Esto como se mencionó en secciones previas, se realiza para disminuir la dimensionalidad en la salida, lo cual en este paso aún no es suficiente, por lo que se realiza un ajuste de los puntos obtenidos por cada curva, para reducir aún más la dimensión mencionada, llegando a tenerse 5 variables de salida.

El ajuste realizado para la obtención de las curvas del contorno de FS es una aproximación determinada por el polinomio que se selecciona, el cual en este caso es de grado 3. La elevación del grado puede permitir una mayor precisión en el ajuste, sin embargo esto provocaría un aumento en los coeficientes de parametrización de la curva, lo que se traduce en una mayor dimensionalidad del proceso de entrenamiento, lo cual reduciría la precisión, por lo que su selección debe ser adecuada y tomando en consideración estos aspectos. Esta etapa define la parametrización del problema y conlleva una complejidad asociada al tipo de zonas que se presente. Para otros tipos de zonas, se debe barajar la opción de plantear un mayor grado en el polinomio de ajuste o utilizar una parametrización diferente, con ayuda de las herramientas de programación (identificador de contornos en Python por ejemplo) que permita describir la geometría de la forma más adecuada. Por otro lado, es importante señalar que la manipulación de los datos conlleva un proceso de adaptación de los mismos, de modo que entre las curvas no difieran entre ellas de forma drástica, evitando que las variables asociadas al valor de C_0 , por ejemplo, y adquieran valores que se escapan de forma extrema de su rango y que provoquen curvas sin sentido físico. Esto se logra agregando puntos estratégicamente, sin pérdida de representatividad, tal como se menciona en el capítulo 6, al momento de hablar de singularidades. El modelamiento de singularidades y también la falta de zonas con un determinado factor de seguridad es presentado en la parte final de la sección 6.1.3. Las consideraciones aplicadas pueden considerarse conservadoras.

7.2.2. Construcción de modelo sustituto

Las características del modelo, esto es, la función base y la función de covarianza se determinan como una función cuadrática y una función cuadrática exponencial respectivamente. La función base es adecuada, ya que como se mencionó, permite ajustar un problema físico con características no lineales, además de su uso típico por la librería utilizada para el procedimiento de entrenamiento (OpenTURNS). La función de covarianza elegida es frecuentemente utilizada por diversos autores, obteniendo buenos resultados y es posible aprovechar sus propiedades que aseguran una suavidad en el ajuste.

Al momento de realizar el entrenamiento del modelo es importante tomar en consideración la dimensionalidad del problema o de la base de datos, tanto en los datos de entrada como de salida. Una mayor cantidad de dimensiones no presenta tantas complicaciones en los datos de entrada, como lo es en la dimensionalidad de los datos de salida. Se debe adecuar una cantidad necesaria y no excesiva de dimensiones, de modo de que el entrenamiento no presente problemas numéricos y de precisión.

Se toma en consideración dos modelos ya mencionados. El primer modelo, que consiste en los parámetros de entrada versus la salida definida por el FS global, tiene mayor simplicidad en su entrenamiento y conlleva menor carga computacional, en contraste con el segundo modelo, que consiste en los contornos de FS, que tiene los mismos parámetros de entrada y en la salida tiene el FS global junto a los coeficientes de las curvas parametrizadas. Este último conlleva un mayor tiempo de procesamiento debido al aumento de la dimensionalidad en la salida. Se hace una distinción entre estos dos procesos, ya que el primero permite una mayor precisión por tenerse una única variable de salida.

Respecto al modelo de contorno de FS, se establece la creación de 10 modelos sustitutos distintos, cada uno asociado a un factor de seguridad "local", y por ende, a una determinada curva. Se define así este procedimiento por varias razones: en primer lugar, debido a la capacidad computacional de la librería de entrenamiento OpenTURNS, la cual limita la opción de tener una gran dimensionalidad en el modelo, con una gran cantidad de casos (base de datos muy extensa); en segundo lugar, es por las características de la base de datos, ya que esta, para cada caso o combinación de parámetros de entrada, define todas las curvas asociadas a cada uno de los factores de seguridad locales, por lo que cada caso tendría asociado 10 curvas que a su vez tiene asociada cada una 4 coeficientes, por lo que la dimensionalidad sería intratable. Es por esto que hay que buscar otro enfoque, el cual es elaborar un modelo sustituto asociado a cada factor de seguridad local, de modo que cada uno de estos modelos trace la respectiva curva. El tema importante a tratar ahora respecto a estos submodelos sustitutos, es la coherencia y correlación entre ellos, lo cual se logra si el modelo presenta un adecuado ajuste en el entrenamiento, y realizando un procedimiento importante en la confección de los gráficos de los contornos, el cual consiste en el filtro de puntos dentro de los mismos, con el objetivo de evitar su intersección. Esto se logra, sin quitar el sentido y la coherencia del sistema.

Es importante señalar la capacidad de retroalimentación del modelo sustituto, a partir de nuevos datos, por lo que a medida que pase el tiempo y que se tenga la disponibilidad y opción de ejecutar mayor cantidad de casos, no incluidos en la base de datos original, es

posible agregar estos datos de modo de realizar un entrenamiento nuevo y así actualizar el modelo sustituto, aumentando su precisión. Este tiempo de entrenamiento como se ha visto y se denotará en determinadas secciones del presente trabajo, no es excesivo, sin embargo debe evaluarse para verificar la factibilidad de su desarrollo. Por otro lado, siguiendo esta línea de búsqueda de mejoras en la implementación del modelo, es de sumo interés la aplicación de nuevas librerías o programas para el entrenamiento del modelo, de manera que se pueda realizar una mejora en el procesamiento de los datos. Se evalúa esto en una siguiente sección.

7.2.3. Resultados

Los resultados entregados por la validación cruzada para el modelo de FS, indican un promedio de los errores residuales igual a 0.0102. Este parámetro indica que el valor que se predice mediante el modelo sustituto tiene un rango de confianza de ± 0.0102 . Este valor, en el contexto del problema físico, y verificando el carácter de la variable objetivo, demuestra una variabilidad baja en la predicción, ya que los valores de este parámetro, o los significativos para análisis, generalmente adquieren cifras que se diferencian entre sí con una discretización de 0.1, por lo que una variabilidad de 0.0102 representa cerca de un 10 % de dicha discretización, lo que implica una alta precisión en la predicción del valor real. Respecto a los valores de prueba, como se puede visualizar demuestran errores más bajos que los mencionados, por lo que el criterio a aplicar es el mismo, y se tiene una validación mucho más respaldada.

Para el modelo de contorno FS, se tiene un promedio de errores residuales para cada submodelo sustituto y para cada parámetro de ese respectivo submodelo. De este modo, es posible ver en base al mismo criterio aplicado en el párrafo anterior, que los errores para los parámetros FSG, C0 se ajustan de forma adecuada, es decir, son pequeños respecto a la discretización característica de estas variables objetivo o de salida y no representan una variabilidad muy excesiva en estas, ya que, por un lado se tiene el valor de FS que ya se describió en el párrafo anterior y se verificó que es adecuado el error, y el valor de C0, que representa el corte en el eje Y de las curvas ajustadas, presenta una variabilidad de 1 unidad, lo cual no se traduce en cambios determinantes gráficamente. Para el caso de los valores de C1, C2, C3, se presentan errores que se exceden dentro de su rango o de su discretización permitida, es decir, presentan cambios más bruscos, y estos se reflejan al momento de graficar las curvas, lo cual implica un comportamiento bastante particular del modelo. Por ejemplo, una variabilidad en la pendiente de las curvas, que se representa como C1, igual a 0.1343 supone cambios significativos gráficamente en las curvas, en algunos casos provocando una diferencia en la coordenada Y de 5 metros, por lo que se entiende que el error no es lo suficientemente bajo, y puede provocar ciertas irregularidades en la predicción. De todos modos, las tendencias del gráfico dado por un aumento o disminución de los parámetros son consistentes, es decir, que el aumento del nivel freático o disminución de parámetros resistentes, lo cual se considera desfavorable, provoca un factor de seguridad 'global' menor y un aumento en las zonas de menor factor de seguridad, y viceversa. Por otro lado, verificando los errores asociados a los datos de prueba se tiene un comportamiento similar, por lo que el criterio es el mismo.

Ahora es importante mencionar que la consideración de estos errores, tal como se menciona en la sección Marco Teórico, se realiza tomando en cuenta todos los datos y no filtrando los datos de soporte extremos respecto a los rangos de cada variable. Esto se realiza por

varias razones: en primer lugar las características del sistema, demostrado en los gráficos que confrontan las variables y el valor de salida, se muestra un comportamiento con indicios lineales, lo cual permitiría agregar estos puntos y los errores se ajustarían de forma adecuada y no serían altos al seguir esta tendencia lineal. En segundo lugar, la base de datos presente no es abundante, y se utiliza una gran cantidad de puntos que adquieren valores de los extremos de los rangos de las variables, por lo que se hace necesario utilizar estos, ya que se necesita la mayor cantidad de datos posibles y así lograr que el método de validación cruzada sea representativo. Por último, al realizar esto estamos frente a una situación desfavorable, lo cual permite tener un criterio conservador en la obtención de estos errores por validación cruzada.

La predicción se realiza para ambos modelos con un tiempo de respuesta prácticamente inmediato, lo cual se denota en mayor detalle en una próxima sección. Se visualiza en las Figuras 5.11 a 5.13 una comparación entre el modelo sustituto referido al modelo de contorno de FS. A partir de estas, es posible identificar que en el modelo sustituto existen zonas de factor de seguridad 'local' no visibles en el modelo FLAC3D. Esto es debido a que el modelo sustituto entrega una representación más directa respecto al apartado visual y también este modelo entrega una respuesta para todo factor de seguridad local, descartando únicamente las cruvas asociadas a factores de seguridad locales que son menores al factor de seguridad global. De todos modos, esto no presenta un problema, ya que se realiza un ajuste de modo que las curvas no influyan entre ellas, y el modelo responde bien frente a estos casos, ya que cuando se tiene un contorno de un determinando factor de seguridad 'local' que no se ve directamente en el modelo original, entonces el modelo por su configuración y entrenamiento logra adecuar estas curvas y las asimila correctamente, trazandolas en las cercanías de las líneas de contorno anteriores y que si se presentan visualmente.

Se puede concluir sobre los modelos sustitutos confeccionados, en primer lugar, respecto al modelo de FS, el cual presenta una sola variable de salida, que presenta una gran confiabilidad y permite entregar una respuesta precisa, por lo que su uso es recomendado para realizar la predicción, además de que a partir de este modelo es posible realizar la cuantificación de incertidumbre, y es el modelo para el cual existe interés de analizar desde el enfoque de esta área, debido a que se tiene como salida el valor del factor de seguridad, cuya importancia es fundamental para el análisis que involucra el presente trabajo referido a la estabilidad física del embalse de relaves. Respecto al modelo de contorno de FS, se recomienda su uso como aproximación del modelo original, pero con una precisión menor que el modelo de FS, debido a que existe una mayor sensibilidad en la representación gráfica, y estas variaciones se presentan de forma más brusca en la geometría planteada. Además, se presentan cambios bruscos en ciertas situaciones o casos, debido al trazado de determinadas curvas, y esto se produce por la naturaleza del problema analizado, ya que estos contornos presentan un comportamiento discontinuo frente a cambios en las variables de entrada y con una tendencia no lineal. La precisión de este último modelo depende del detalle del procedimiento que se realice para ajustar las curvas, haciendo un tratamiento de los datos adecuado y filtrando los datos que provocan un escape de los valores de los parámetros dentro de su rango característico, y realizando ajustes que logren suplir las discontinuidades del problema, es decir, la prueba de distintas formulaciones, ajustes y adición de puntos estratégicos en la base de datos, adaptando de forma adecuada los contornos, permitirá que estos adquieran una mayor precisión en la predicción, y sino, entonces buscar las configuraciones que presenten condi-

ciones menos favorables, aplicando así un criterio conservador. Específicamente esto se puede conseguir agregando y fijando puntos en zonas bajas y con pendientes más pronunciadas, de modo que se reproduzca un caso más desfavorable. Cabe mencionar nuevamente que el trazado de los gráficos es consistente a la variabilidad de las variables de entrada, es decir, un aumento de variables resistentes se traduce zonas de factor de seguridad mayores y viceversa y es importante que estas tendencias se mantengan para el funcionamiento adecuado del modelo. En una próxima sección, se describe una propuesta de uso de estos modelos de forma más específica.

7.3. Cuantificación de incertidumbre

7.3.1. Preliminares

Los preliminares de la cuantificación de incertidumbre, consisten en la determinación de la información estadística de las variables de entrada, esto es, el promedio y la desviación estándar. Esta información puede provenir de diversas fuentes, tales como, mediciones en terreno, datos experimentales, datos artificiales por softwares, entre otras. En este caso se utiliza información referencial, tomando el valor intermedio del rango seleccionado para cada variable y una desviación estándar que abarque al menos un valor de dicho rango en el modelo de FS. Para el modelo de contorno de FS, se toma valores menos conservadores, para reflejar de forma adecuada el método.

7.3.2. Propagación de incertidumbre

La propagación de incertidumbre permite determinar la distribución de probabilidad de la variable de salida, la cual en este caso es el factor de seguridad. Se observa una media igual a 1.385, con una desviación de 0.107, lo que implica que se pueden tener valores muy cercano al límite dado por la normativa chilena (1.4), por lo que para esta situación es conveniente mejorar las propiedades resistentes o realizar un control hidráulico o de filtraciones más efectivo, para de este modo ajustar y mejorar los parámetros de entrada, lo que conlleva mejoras en el parámetro de salida, el factor de seguridad. El procedimiento se ha realizado en base a la utilización de 10,000 muestras, la cual se definió por iteración, de modo que el valor de cov_{MC} sea adecuado, es decir, que sea menor a 10 %, según la práctica habitual en estos procesos, lo cual se consigue con gran holgura.

Respecto al modelo de contorno de FS, se utilizan 1000 muestras, que entrega resultados de $c.ov_{MC}$ adecuados para cada una de las variables de salida (menor a 10 %). En la Figura 6.5, se presenta la curva asociada al factor de seguridad local 1.4, con un intervalo o espacio donde se muestra sus posibles ubicaciones dentro del intervalo seleccionado, el cual es entre el percentil 3 y 97, de modo que se tiene un 94 % de posibilidad que la curva se ubique en la zona descrita, dada por la propagación. Se establecen estos percentiles por ser un criterio común y general, y dado por las características de las distribuciones normales. Este análisis es posible verificarse para la curva que se requiera, sin embargo el énfasis se realiza para la situación considerada más crítica, que es asociado al factor de seguridad global calculado, por lo tanto, la curva que se analiza es la que tiene el valor de factor de seguridad local igual o cercano.

La noción que entrega el modelo permite mayor confiabilidad en el análisis del sistema, ya que se puede identificar cual es la máxima área que puede abarcar asociada a un determinado factor de seguridad local a partir de la información estadística de las variables de entrada. Es decir, si se tiene que el factor de seguridad global es 1.4, el cual es el límite según la normativa chilena para un caso estático, entonces se analiza la curva de factor de seguridad local igual a ese valor y esta indica las zonas críticas y su rango posible.

7.3.3. Probabilidad de falla

La probabilidad de falla establece la probabilidad de que la variable objetivo sea mayor o menor que determinado valor. Mediante esta método es posible analizar el valor mencionado previamente, es decir cuando el valor del factor de seguridad es menor a 1.4, y verificar la probabilidad, la cual es igual a 60 %. Esto permite comprender si las variables y su información, es adecuada, si es conservadora o no y si se está en una situación de riesgo.

El procedimiento se ha realizado con las mismas condiciones mencionadas en el proceso de propagación de incertidumbre.

7.3.4. Análisis de sensibilidad

Mediante este análisis se logra identificar las variables que influyen en mayor o menor grado en la variable objetivo o respuesta del sistema. Los resultados indican que la variable asociada al ángulo de fricción es la que tiene mayor influencia, basado en el valor del índice de Sobol que es cercano a 0.5, lo que equivale aproximadamente a un 50 %. Esta información puede indicar el enfoque que se debe realizar en esta variable y en la precisión que se requiera en su determinación o medición para el respectivo diseño en el que se emplea. Por otro lado, se tienen porcentajes cercanos a 20 % asociados al nivel freático y a la cohesión, lo cual implica una influencia considerable dentro del sistema. Es interesante el resultado dado para la variable de nivel freático 2, ya que otorga un valor negativo para el índice de primer orden y bastante bajo para el índice de orden total, lo cual puede ser considerado prácticamente como 0 %, ya que el valor adquiere importancia cuando no es menor a 5 %. Es posible concluir con esto en el problema físico, que esta variable es posiblemente descartable en el análisis, debido a que no presenta gran influencia, lo cual apoyaría en gran medida en términos de carga computacional y tiempo de procesamiento, al reconsiderar el problema. Ahora, la baja influencia de esta variable puede deberse principalmente a la ubicación del piezómetro que entrega el valor del respectivo nivel freático, ya que se encuentra en el talud aguas arriba, alejado del talud aguas abajo, el cual presenta mayor interés para el análisis y donde se aprecia el mayor efecto de la inestabilidad del sistema.

7.4. Tiempos de procesamiento

En esta sección se presentan los tiempos necesarios para el desarrollo de cada uno de los procesos, tal como se indica en la Tabla 7.1, y en la Tabla 7.2, es posible ver el porcentaje de disminución del tiempo utilizado para la ejecución de un caso.

7.4.1. Resumen de tiempos

Tabla 7.1: Resumen de tiempos de procesos

Proceso	Tiempo	Unidad
Selección de parámetros	2.3	días
Distribución de casos	1	seg.
Ejecuciones de casos	15.7	días
Ejecución de casos de prueba	0.4	días
Manipulación de datos	20	min.
Entrenamiento de modelo FS	1	min.
Entrenamiento de modelo de contorno FS	2.5	días
Cuantif. de inc. de modelo FS	1	seg
Cuantif. de inc. de modelo de contorno FS	20	min
Total	21	días

Como se indica en la Tabla anterior, el tiempo total de los procesos es de 21 días aproximadamente, lo cual se puede ver como un tiempo extenso, sin embargo, este proceso es realizado una sola vez y es amplio debido principalmente a las ejecuciones de los casos en el modelo original de diseño FLAC3D. Una vez realizado todo este proceso, el modelo queda entrenado y contiene toda la información necesaria para realizar predicciones. Este tiempo es en base a la utilización de un solo servidor, equipo o computador con el programa FLAC3D, y utilizado a tiempo completo. Es posible disminuirlo mediante la consideración de menos casos en los valores o variables de entrada, lo cual puede implicar una menor precisión de la predicción, o también es posible mediante el uso de una mayor cantidad de servidores, tal y como se mencionó en secciones previas del presente capítulo. y se mencionará y evaluará en próximas secciones.

7.4.2. Mejoras de tiempo

Tabla 7.2: Mejoras en tiempo de respuesta

Proceso	Tiempo	Unidad
Modelo original	180	min.
Modelo sustituto	1	min.
Diferencia	179	min.
Diferencia en %	99.4	%

Como lo demuestra la Tabla, la reducción del tiempo en la ejecución de un determinado caso, utilizando el proceso de predicción del modelo sustituto, es considerablemente alta, dando como resultado un 99.4 %. Este tiempo conlleva la ejecución de ambos modelos simultáneamente, tanto el modelo de FS, como el de contorno de FS.

7.5. Comparación

En la presente sección se realiza un proceso comparativo, que busca contrastar el modelo base del trabajo de memoria, el cual es el que se ha realizado a lo largo del presente documento hasta ahora, con diferentes modelos que son configurados con cambios parciales en un determinado procedimiento. Esto con el objetivo de evaluar la influencia de estos procedimientos en el resultado final, y así también verificar cual es la opción que puede representar mejoras o no en la precisión del modelo, así como también en su procesamiento, considerando los tiempos y la carga computacional.

7.5.1. Modelo con muestreo con LHS

En primer lugar, se efectúa el desarrollo de un modelo en el que se modifica el método de obtención de muestras, esta vez utilizando el método LHS (Latin Hypercube Sampling). Este desarrollo se presenta en Anexo F, y se realiza para comparar los modelos en términos de precisión.

Como se puede ver en la Tabla F.2 los errores asociados a la validación cruzada son bastante bajos, incluso menores que los del modelo base, lo cual muestra que la distribución es adecuada para la interpolación de los casos, es decir, se tiene buena precisión para la predicción en los datos intermedios. Sin embargo, como se puede ver en las Figuras F.1 a F.4, las cuales comparan el comportamiento de la predicción (en color rojo) con los puntos reales o datos de soporte extraídos del modelo base (en color azul), la predicción cuando se requiere obtener datos extremos de cada variable, tiene una precisión no tan adecuada, llegando a haber un error en ciertos casos de 0.05, lo cual puede considerarse un valor alto tomando en consideración el contexto físico de los valores del factor de seguridad.

Es posible verificar entonces que este modelo mediante este método de muestreo, tiene una gran precisión para la predicción de valores intermedios, y presenta una menor precisión cuando se trata de predecir a partir de valores extremos o de borde de las variables de entrada involucradas, lo cual para este problema es de suma importancia, ya que esta situación de casos extremos generalmente representa el caso de falla.

Cabe mencionar que el tiempo de procesamiento para la realización de este modelo es el mismo que el modelo base, por lo tanto no hay mejorías en este aspecto.

7.5.2. Modelo con menos cantidad de casos

En este caso, se desarrolla un modelo que utilice una menor cantidad de casos para el proceso de entrenamiento. Este se muestra en Anexo G y se realiza para comparar los modelos en términos de precisión y tiempos de procesamiento.

Observando la Tabla G.3, se identifica que la validación cruzada presenta errores más altos para este modelo, respecto al modelo base, lo cual tiene sentido por el cambio de la cantidad de datos, lo que implica una variación en la capacidad de interpolación o predicción con datos intermedios. Aún así, este error no es considerablemente alto, tomando un valor cercano a 0.02. Comprendiendo el contexto físico del parámetro de salida a analizar, se puede verificar

esto. Por otro lado, visualizando las Figuras G.1 a G.4, donde se compara la predicción (en color rojo) con los datos reales basados en el modelo base (color azul), se demuestra que las predicciones para los puntos extremos o de borde son precisas y para datos intermedios se muestra un buen comportamiento, pero no lo suficiente para la variable de nivel freático, donde se ve que para los casos en los que no se tiene información de soporte o datos de entrenamiento, el error es mayor.

En términos de tiempo de procesamiento, este es considerablemente menor, tal como se ve en la Tabla G.5, reduciéndose en un 87% el tiempo del proceso de obtención de la base de datos, respecto al modelo base.

Se concluye con el desarrollo de este modelo que utilizar una menor cantidad de datos, provoca una disminución en la precisión lo cual es desventajoso, pero no de forma abrupta, y el tiempo de procesamiento es mucho menor, lo cual presenta ciertas ventajas. Esto incentiva el uso de este método, ya que el tiempo es un parámetro clave en el procesamiento de estos modelos, además de que se puede presentar una metodología alternativa, en la que estos casos que se utilizan no abarquen un rango de valores de entrada tan amplio y más bien enfocar la selección de casos a un rango de valores de salida más reducido pero con mayor utilidad en el análisis y así se lograría una mayor precisión en las predicciones, ya que como se dijo, se tiene esa misma cantidad de datos (la base de datos reducida respecto al modelo base) pero en un espacio más reducido. Por ejemplo, es posible reducir los casos para hacer énfasis en la obtención de factores de seguridad que adquieran valores cercanos al valor límite dado la normativa chilena para un caso estático (1.4) y de este modo enfocar el problema en este espacio muestral significativo. De todos modos, se tiene gran justificación en la elección de casos dado por el modelo base, ya que con esta se asegura una mayor precisión en desmedro de una mejoría en el tiempo por la consideración de mayor cantidad de casos y además se abarca un espectro mayor de escenarios, aunque como bien se menciona, depende del enfoque del problema que se esté planteando.

7.5.3. Modelo con distinta librería

Finalmente se realiza la comparación del modelo base con un nuevo modelo en el que se utiliza una librería alternativa. La librería es Pydace y está basada en la configuración de la librería Dace de Matlab. El desarrollo se presenta en Anexo H y se realiza para comparar los modelos en términos de precisión, tiempos de procesamiento y carga computacional.

Tal y como se indica en el desarrollo, se realiza un análisis más detallado. Primero se calculan los errores de validación cruzada asociados al modelo de FS en la Tabla H.1, los cuales muestran un promedio de 0.011, que es ligeramente mayor al entregado por el modelo base. Respecto a las Figuras H.1 a H.4 se verifica que presentan un ajuste adecuado a los puntos de soporte o de entrenamiento, bastante similar al que se muestra para el modelo base.

Por otra parte, se tiene el modelo de contorno de FS, para el cual se calculan los errores referidos a los datos de soporte en la Tabla H.5, los cuales presentan promedios cuyos valores son levemente más bajos que los errores del modelo base. Luego se tienen los errores respecto

a la validación cruzada que siguen la misma tendencia al tener errores bajos y muy cercanos a los del modelo base.

En términos de tiempo de procesamiento, es importante resaltar que esta librería logra una mejoría, lo cual implica una reducción en el tiempo del proceso de entrenamiento del modelo sustituto de múltiples salidas, es decir, en el modelo de contorno de FS. Esta diferencia se muestra en la Tabla H.9. Como se puede ver, la reducción es considerable.

El modelo analizado permite efectuar el proceso de entrenamiento de una forma mucho más rápida, con una diferencia en los errores asociados a datos de soporte y validación cruzada, para los modelos de FS y de contorno de FS, considerablemente baja. Por otra parte, se debe mencionar que la carga computacional asociada, implica por un lado un determinado tiempo de procesamiento y también conlleva un aspecto referido a la capacidad de procesamiento del equipo o servidor, lo cual va relacionado con la memoria RAM. Para el caso de la librería OpenTURNS, el proceso de entrenamiento para un modelo de múltiples salidas implica un costo de este tipo demasiado alto, lo cual no ocurre con la librería Pydace. Se puede concluir entonces que esta última librería presenta ventajas en comparación con la librería del modelo base, principalmente por el tiempo de entrenamiento del modelo, por lo que su uso es una alternativa muy factible. El argumento de uso de la librería utilizada para el modelo base (OpenTURNS) es principalmente esta librería contiene una documentación mucho más detallada y extensa, lo cual permite reconocer de mejor manera la teoría aplicada y presenta también una mayor cantidad de herramientas para la manipulación y trabajo con los datos, tales como opciones de distribución de probabilidad conjuntas, diferentes métodos de entrenamientos, métodos de optimización, y una variedad interesante y potente de algoritmos asociados además a la cuantificación de incertidumbre. Por otro lado, su aplicación es adecuada para modelos con baja dimensionalidad en los parámetros de salida, por lo que el primer modelo de FS, no tiene problemas en este aspecto y los tiempos son adecuados, además de que los errores obtenidos son bajos para cada modelo, cada variable de salida y para cada situación de validación, y son levemente menores al dado por el modelo de la librería alternativa, es decir, en términos de precisión no hay grandes diferencias. Aún así, el tiempo de procesamiento es clave, por lo que es posible concluir que el uso de la librería del modelo base (OpenTURNS) es útil y completa para realizar modelos de baja dimensionalidad en la salida, como lo es el modelo de FS, y la librería alternativa (Pydace) es útil para modelos con mayor dimensionalidad en la salida, como lo es el modelo de contorno de FS, sin entrar en mayor detalle en su formulación, y usando procedimiento y herramientas estándar para el entrenamiento del mismo.

7.6. Aplicaciones

Esta sección tiene el objetivo de presentar de forma organizada las aplicaciones del presente trabajo de memoria.

7.6.1. Presentación y generalización de metodología

A lo largo del documento se ha propuesto una metodología de confección de modelos sustitutos enfocada a un sistema físico geotécnico, específicamente un embalse de relaves, el que generalmente se caracteriza por la limitación en la cantidad de datos que se pueden obtener

en el procesamiento y ejecución de los modelos. Es por esto que se propone el método Kriging o regresión en procesos Gaussianos, la cual cumple con la creación de un modelo sustituto bajo las condiciones mencionadas anteriormente. De este modo el trabajo de memoria cumple el cometido de presentar esta metodología enfocada, intentando abordar todos los puntos y pasos que esta conlleva, mediante bases teóricas bien respaldadas y aplicadas.

Se realiza un enfoque para embalses de relaves, debido al interés que se tiene por lograr una mejor representatividad, una noción de incertezas, mayor seguridad en el diseño y un procesamiento más ágil de los cálculos y datos, el cual en este caso se realiza en el software FLAC3D. La metodología está claramente conectada y dirigida a la utilización de los procedimientos de cálculo de este software de una forma alternativa a la forma clásica, con un enfoque en la mejoría de los procedimientos de obtención de datos a partir de este mismo programa. Cabe mencionar, tal y como se mencionó en secciones previas, que esta metodología es adaptable a diferentes fuentes de datos, estos es, otro software u otra forma de obtención, como datos experimentales, mediciones en terreno, etc.

Por otro lado, es posible adaptar esta metodología, básicamente estableciendo otro problema de diseño, frente a otro sistema físico (distinto a un embalse de relaves), tales como, por ejemplo, sistemas estructurales, sistemas hidraulicos, sistemas de transporte, incluso sistemas alejados de la ingeniería civil pura, siempre y cuando se establezca y asimile el problema planteado con el formato del problema que se evalúa en el presente trabajo de memoria. Esto permite una generalización de esta metodología para diferentes áreas de estudio.

7.6.2. Cuantificación de incertidumbre

Junto con la presentación de una metodología, se adhiere una metodología y aplicación de métodos de cuantificación de incertidumbre, los cuales permiten abarcar e incursionar en un estudio de incertezas significativo para la comprensión y confianza de los diferentes sistemas a modelar. Esta área de estudio tiene múltiples funcionalidades, tales como la obtención de intervalos de confianza de parámetros o variables de salida, probabilidades de falla de diversos sistemas, comprensión de influencia de variables sobre respuesta de sistema, toma de decisiones sobre énfasis en variables, y más. Se hace posible generalizar el uso de estos métodos para su aplicación en diversas áreas de estudio, como los que ya se han mencionado en este documento y en la sección anterior (sistemas de diversa índole, referidos a ingeniería civil y a otras ramas de ingeniería).

7.6.3. Sistema de monitoreo

Realizando un enfoque en la aplicación práctica directa del presente trabajo de memoria, incluyendo su énfasis en el sistema de embalse de relaves, se procede a describir el proyecto adjudicado a la empresa Itasca, el cual sirvió en parte, de motivación para este trabajo de memoria.

El proyecto parte de la base de un embalse de relaves contruido que posee piezómetros en determinadas zonas del muro, similar a la imagen referencial que se muestra en la siguiente Figura:

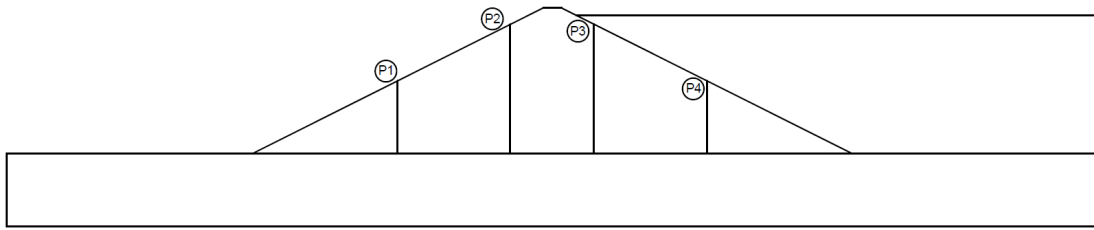


Figura 7.1: Diagrama referencial - Embalse de relaves

A partir de los registros de estos instrumentos, se plantea la confección de un sistema autónomo que almacene estos datos y los pueda utilizar para el desarrollo de otros procesos. Estos conllevan la determinación de información estadística de las variables involucradas (nivel freático), la cual es de utilidad para la realización de la cuantificación de incertidumbre, y el objetivo más importante, para el procesamiento de un cálculo mediante el software de modelamiento geotécnico FLAC3D, y así obtener resultados referidos al factor de seguridad (FS) y contornos de FS. El siguiente diagrama representa este sistema:

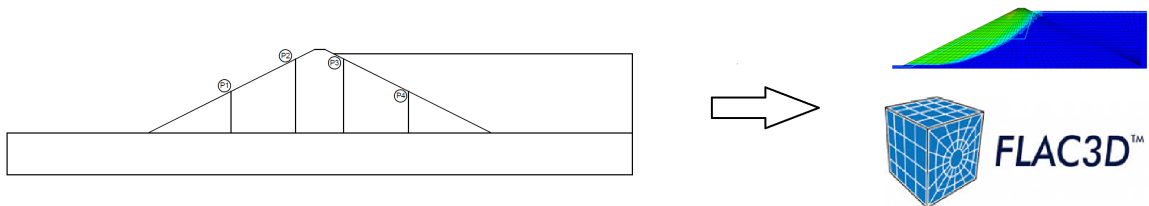


Figura 7.2: Diagrama explicativo - Sistema con modelo original

La problemática se presenta por el hecho de la falta de agilidad en el proceso de ejecución de los casos y por ende la respuesta del software. Es por ello que el modelo sustituto confeccionado en el presente trabajo de memoria soluciona esta situación. Esto se indica en la siguiente Figura:

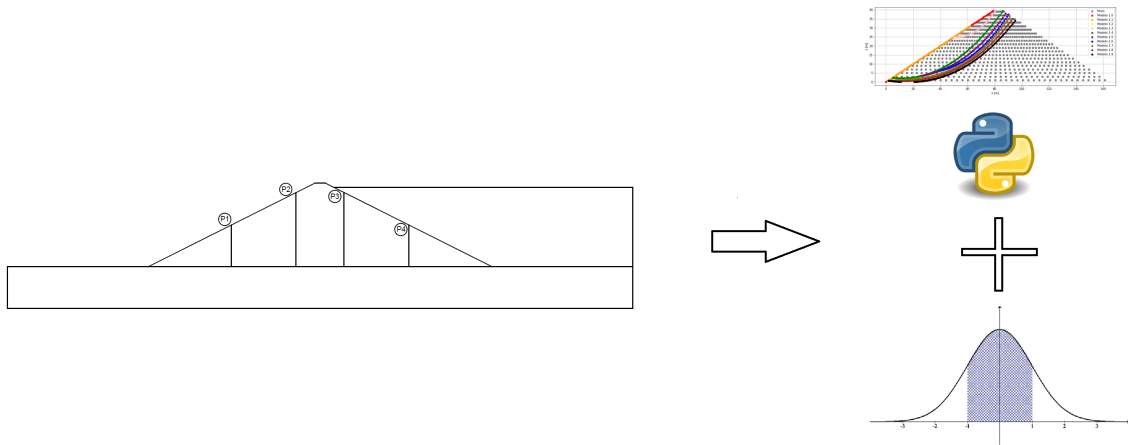


Figura 7.3: Diagrama explicativo - Sistema con modelo sustituto

De este modo, el modelo sustituto permite la confección de este sistema autónomo, el cual recibe valores de variables de entrada, a partir del registro in situ de las mediciones y con ello la obtención de información estadística, para luego calcular mediante el modelo sustituto el factor de seguridad, y también los contronos de FS, además de aplicar al mismo tiempo los métodos de cuantificación de incertidumbre. Cabe resaltar la capacidad de retroalimentación del modelo sustituto mencionada en una sección previa de este mismo capítulo, ya que mediante se realiza todo el proceso de registro de información y los respectivos cálculos, se puede realizar un procedimiento de cálculo de una mayor cantidad de casos externos a la base de datos original, para lograr nutrir de mejor manera el modelo sustituto y dotarlo de una mayor capacidad predictiva.

Dicho lo anterior, es posible utilizar y obtener conclusiones de los modelos del siguientes modo: primero se utiliza el modelo de FS para la obtención precisa del factor de seguridad global, que representa el menor valor obtenible de este parámetro para el respectivo caso que se esté evaluando, y en base a este valor es que se puede realizar una comprensión del sistema, eso sí, con un criterio muy conservador. Este criterio es así, debido a que se está considerando que el sistema completo del embalse se basa en este parámetro sin entrar en particularidades y sin tomar en cuenta la cantidad de área o el volumen asociado a esa respectivo factor de seguridad. De todos modos su uso no deja de ser un criterio seguro y responsable, lo cual es sumamente valorable. A partir de los métodos de cuantificación de incertidumbre se hace viable el análisis de este parámetro con mayor comprensión del mismo, identificando las incertezas y teniendo más información y criterio para su análisis. Al tratarse de un sistema de monitoreo, el objetivo de dotar de seguridad a este se cumple a cabalidad, otorgando una respuesta con un criterio bastante conservador y con una respuesta rápida y ágil. Por otro lado, el modelo de contorno de FS, al usarse luego del análisis del modelo de FS, permite obtener mayor información y comprensión de la respuesta del sistema, de modo que el criterio anterior se puede respaldar en mayor medida. Es decir, en esta etapa, al visualizar los contornos del determinado factor de seguridad obtenido, se verifica la zona involucrada y se establece si es significativa o no, teniendo una nueva noción del problema.

La validación obtenida para este último modelo, si bien no es igual de precisa y confiable que el modelo de FS, sirve como una aproximación, y en caso de precisar de información más detallada y se cuenta con un determinado tiempo adicional, se propone la ejecución de ese respectivo caso en el software original, esto claramente frente a condiciones que no tengan exigencia temporal. La mención de esta última propuesta no quita mérito a la funcionalidad de los modelos sustitutos confeccionados, ya que se cumple con sus objetivos, los cuales son entregar valores de factores de seguridad precisos y contornos aproximados del modelo, de modo que esta respuesta es ágil, permite generar alertas en el sistema y otorga las condiciones para una reacción y respuesta oportuna por parte del usuario.

Capítulo 8

Conclusiones

Como se ha descrito, analizado y discutido a lo largo del documento, el objetivo del trabajo de memoria es la confección de un modelo sustituto para un sistema geotécnico, específicamente un embalse de relaves, mediante la aplicación de un proceso predictivo, utilizando el método Kriging o regresión en procesos Gaussianos.

Mediante la realización de este objetivo se presenta una metodología sobre el desarrollo de este tipo de modelos, el cual conlleva un proceso de obtención de una base de datos, que a su vez depende del tipo de sistema que se está analizando y de la fuente de obtención, que en este caso es el software de modelamiento FLAC3D. Implica un proceso de entrenamiento y de validación a partir de esta base de datos, y la aplicación de métodos de cuantificación de incertidumbre, para la medición y noción de incertezas. De este modo se propone un procedimiento de trabajo claro y con múltiples utilidades, tal y como son mencionadas en la respectiva sección de Análisis y Discusión.

Los modelos sustitutos se presentan como una herramienta factible para el análisis de sistemas ingenieriles. Estos agilizan la respuesta del sistema, mediante un proceso de predicción, siempre y cuando se tenga una validación adecuada, la cual es la actividad clave e importante al momento de confeccionar estos modelos, tal y como se ha mencionado a lo largo del trabajo. Junto con esta utilidad, el modelo sustituto entrega las condiciones ideales para la aplicación de métodos de cuantificación de incertidumbre, los que entregan información valiosa para la toma de decisiones y en la comprensión y enfoque del problema. Esta área de estudio se muestra como una opción con bases teóricas bien desarrolladas, y con múltiples herramientas, con el objetivo de lograr una mejor representatividad y comprensión de las respuestas del sistema, obteniendo intervalos de confianza e información estadística y probabilística que complementa de gran manera a la determinista, lo cual aporta significativamente a la hora de analizar el respectivo sistema.

El análisis del tipo de sistemas geotécnicos evaluado en esta memoria, como lo son los embalses de relaves, debe realizarse asumiendo criterios conservadores y con estándares de seguridad adecuados, principalmente debido a la magnitud de las obras asociadas y al nivel de riesgo que conllevan, por lo que el presente trabajo busca aportar en el análisis de estos sistemas, otorgando una mejoría en la eficiencia de los procesos de cálculo, y otorgando herramientas para agilizar y detallar la respuesta de estos. Es posible ver este efecto en el caso práctico que se busca aplicar, al desarrollar una propuesta de solución de un sistema de

monitoreo en un embalse de relaves ya establecido, cuya respuesta es rápida y autónoma, lo cual no se podría efectuar de este modo sin el modelo sustituto, debido a las características del modelo original.

Es posible concluir en términos técnicos que se ha cumplido el objetivo de la memoria, con la obtención de un modelo sustituto que responde a determinadas características del modelo original, con errores, dados por la respectiva validación cruzada y validación de prueba, bastante adecuados para el modelo de FS, cercanos a 0.01, lo cual es aceptable para el sentido físico del parámetro, y aceptables para un fin de aproximación por parte del modelo de contorno de FS, con valores relativamente altos en consideración al sentido físico de las variables de salida, lo cual se presta para mejorías basadas en las propuestas de solución planteadas en la memoria. Se consigue una respuesta mejorada en términos de carga computacional y reduciendo el tiempo de procesamiento en aproximadamente un 99.4 %, con la posibilidad de aplicación de un área de estudio de amplia utilidad, como lo es la cuantificación de incertidumbre, mediante el desarrollo de métodos de propagación, probabilidad de falla y análisis de sensibilidad.

A lo largo de las diferentes secciones del capítulo de Análisis y Discusión se plantean propuestas referidas a la solución de dificultades o realización de mejoras para posibles problemas observables acociados a cada etapa de la metodología del trabajo, a modo de establecer recomendaciones frente a la realización de futuros trabajos, los cuales también se han propuesto, o frente a la implementación de la misma metodología planteada en esta memoria.

Finalmente, se espera que el presente trabajo contribuya al desarrollo de la ingeniería, en este caso para un contexto geotécnico, otorgando herramientas para aumentar la seguridad, confiabilidad y representatividad del análisis que conlleve, y de este modo realizar una mejor ingeniería y contribuir de forma responsable y con adecuado criterio en los desafíos que se presentan dentro del ámbito científico y social.

Bibliografía

- [1] Barrera. S., Campaña J., *Análisis de estabilidad de presas de relave-Práctica Chilena*. 2004.
- [2] Yaochu. J., *Surrogate-assisted evolutionary computation: Recent advances and future challenges*. 2011.
- [3] Guzmán. J., *Cuantificación de incertidumbres. Minería Chilena*. 2011.
- [4] Ministerio de Minería, *Decreto Supremo N°248. Reglamento para la aprobación de diseño, construcción, operación y cierre de los depósitos de relaves*. 2007.
- [5] SERNAGEOMIN, *Guía técnica de operación y control de depósitos de relaves*. 2007.
- [6] Carvajal. M., *Desarrollo de una metodología para análisis de estabilidad física de depósitos de relaves. Universidad de Chile*. 2018.
- [7] Bowles. J., *Foundation Analysis and Design. Fifth Edition*. 1996.
- [8] Budhu. M., *Soil Mechanics and Foundations. Third Edition*. 2011.
- [9] Look. B., *Handbook of Geotechnical Investigation and Design Tables. Second Edition*. 2014.
- [10] U.S. Environmental Protection Agency (EPA), *Design and evaluation of tailings dams. Technical Report*. 1994.
- [11] Barrios. D., *Análisis de deformaciones de depósitos de relaves con muros de material de empréstito sometidos a sismos chilenos. Universidad de Chile*. 2019.
- [12] Oldecop. L., Rodríguez R., *Mecanismos de fallas de las presas de residuos mineros*. 2007.
- [13] Itasca, *Software Documentation Itasca. FLAC3D*. 2019.
- [14] Bosch. J., *Métodos finitos. Instituto de energía y termodinámica Medellín. Universidad Pontificia Bolivariana*. 2006.
- [15] Cook R., *Finite element modeling for stress analysis. University of Wisconsin. Madison*. 1995.
- [16] *Modelación Numérica en Ingeniería Hidráulica y Ambiental. Universidad de Chile, Elementos Finitos*. 2008.
- [17] Cepe-eua, *Diferencias entre métodos finitos*. Desconocida.
- [18] González. L., *Introducción a Machine Learning. 1° Edición*. 2018.
- [19] Scikit-learn, *Documentation Scikit-learn*. 2021.

- [20] OpenTURNS, *OpenTURNS 1.17rc1 Documentation*. 2021.
- [21] Rasmussen. C., Williams K., *Gaussian Processes for Machine Learning*. 2006.
- [22] Tresp. V., *Basis Functions*. 2016.
- [23] De La Pompa. V., *Procesos Gaussianos para problemas de regresión y estimación de la incertidumbre*. 2018.
- [24] Dubourg. V., *Adaptive surrogate models for reliability analysis and reliability-based design optimization*. 2012.
- [25] N. W. Santner. T., Williams B., *The Design and Analysis of Computer Experiments*. 2003.
- [26] Morales. E., *Procesos Gaussianos*. Desconocida.
- [27] Parra G., *Procesos Gaussianos Multi-Output*. 2016.
- [28] Villalobos. A., *Procesos Gaussianos y análisis Bayesiano, aplicados a la detección de daños en placas de materiales compuestos*. 2020.
- [29] Blatman. G., *Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis*. 2009.
- [30] Bin. Z. Tetsuya. H. Huafu. P. Qing. Y., *Efficient reliability analysis of slopes integrating the random field method and a Gaussian process regression-based surrogate model*. 2020.
- [31] Ashley. D., Ali T., *Prediction and classification for finite element slope stability analysis by random field comparison*. 2019.
- [32] Kedong. T. Jialiang. W. Lielie. L., *A Prediction Method Based on Monte Carlo Simulations for Finite Element Analysis of Soil Medium considering Spatial Variability in Soil Parameters*. 2020.
- [33] Sung C., *Probabilistic stability analyses of slopes using the ANN-based response surface*. 2009.
- [34] Zhang. J., *Modern Monte Carlo Methods for Efficient Uncertainty Quantification and Propagation: A Survey*. 2020.
- [35] Robert. C., Casella G., *Monte Carlo statistical methods*. Springer Science. 2013.
- [36] Sobol. I.M., *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*. 2001.
- [37] Sobol. I. M., *Sensitivity Estimates for Nonlinear Mathematical Models*. 1990.
- [38] Jia. G., Taflanidis A., *Efficient Evaluation of Sobol' Indices Utilizing Samples from an Auxiliary Probability Density Function*. 2001.
- [39] Saltelli. A., *Making best use of model evaluations to compute sensitivity indices*. 2002.
- [40] Ministerio de Obras Públicas, *Decreto Supremo 50*. 2015.
- [41] Pandas development team, *Pandas Documentation*. 2021.
- [42] De La Rosa. A., *Modelos constitutivos empleados en la modelación de la respuesta mecánica de los suelos agrícolas*. 2018.
- [43] González. O. Herrera. M. Iglesias. C. López. C., *Análisis de los modelos constituti-*

vos empleados para simular la compactación del suelo mediante el método de elementos finitos. 2013.

- [44] Devore. J., *Probabilidad y Estadística para Ingeniería y Ciencias. Séptima Edición.* 2008.
- [45] NumPy developers, *Numpy Documentation.* 2021.
- [46] Matplotlib development team, *Matplotlib Documentation.* 2021.
- [47] SciPy developers, *SciPy Documentation.* 2021.
- [48] Souza F., *Documentación Pydace.* 2021.

Anexo A

Conceptos de Probabilidad y Estadística

En este apartado se describen los conceptos de probabilidad y estadística que se utilizan a lo largo del trabajo tanto de forma implícita como explícita [21] [44].

A.1. Definiciones

- i. Estadística descriptiva: rama de la estadística que se enfoca en la descripción y/o caracterización de un grupo de datos.
- ii. Estadística inferencial: rama de la estadística que se enfoca en la deducción y obtención de conclusiones acerca de situaciones generales mas allá del conjunto de datos obtenidos.
- iii. Variable aleatoria: función que asigna un valor, numérico, al resultado de un experimento aleatorio.
- iv. Vector aleatorio: vector formado por una o más variables aleatorias.
- v. Inferencia bayesiana: Tipo de inferencia estadística en la que las evidencias u observaciones se emplean para actualizar o inferir la probabilidad de que una hipótesis puedaser cierta.

A.2. Parámetros básicos

- i. Esperanza

$$E[x] = \int_x x \cdot p(x)dx = \mu_x \quad (\text{A.1})$$

$$E[f(x, y)] = \int_x \int_y f(x, y) \cdot p(x, y)dxdy = \mu_{f(x,y)} \quad (\text{A.2})$$

- ii. Varianza

$$VAR(x) = \int_x \int_y (x - \mu_x)^2 \cdot p(x) dx = \sigma_x^2 \quad (\text{A.3})$$

iii. Covarianza

$$COV(x, y) = \int_x \int_y (x - \mu_x)(y - \mu_y) \cdot p(x, y) dx dy \quad (\text{A.4})$$

iv. Desviación estandar

$$\sigma_x = \sqrt{VAR(x)} \quad (\text{A.5})$$

v. Coeficiente de correlación

$$Corr(x, y) = \frac{COV(x, y)}{\sigma_x \sigma_y} \quad (\text{A.6})$$

vi. Coeficiente de variación

$$c.o.v. = \frac{\sigma}{\mu} \quad (\text{A.7})$$

A.3. Distribuciones de probabilidad

Las distribuciones de probabilidad están definidas por funciones de densidad de probabilidad (PDF). Cuando la distribución de probabilidad depende de dos variables aleatorias es una distribución de probabilidad conjunta. Estas se denotan como $p(x, y)$.

También es posible definir la función de densidad acumulada (CDF):

$$P(x, y) = \int_x \int_y p(x, y) dx dy \quad (\text{A.8})$$

Algunos de los tipos más comunes de distribuciones de probabilidad están determinados por los siguientes PDF:

i. Distribución uniforme

$$x \sim U(\alpha, \beta) \quad (\text{A.9})$$

$$p(x) = \frac{1}{(\beta - \alpha)} \quad (\text{A.10})$$

ii. Distribución normal o Gaussiana

$$x \sim N(\mu, \sigma^2) \quad (\text{A.11})$$

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (\text{A.12})$$

iii. Distribución log-normal

$$\ln(x) \sim N(\mu, \sigma^2) \quad (\text{A.13})$$

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right) \quad (\text{A.14})$$

A.4. Propiedades generales

A.4.1. Marginalidad

Si se tiene que $p(x,y)$ es la probabilidad conjunta de x e y , entonces la probabilidad marginal de x es:

$$p(x) = \int_y p(x, y) dy \quad (\text{A.15})$$

A.4.2. Condicionalidad

Si se tiene que $p(x,y)$ es la probabilidad conjunta de x e y , entonces la probabilidad condicional de x es

$$p(x|y) = \frac{p(x, y)}{p(y)} \quad (\text{A.16})$$

A.4.3. Teorema de Bayes

A partir de la probabilidad condicional es posible obtener el teorema de Bayes:

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)} \quad (\text{A.17})$$

A.5. Procesos Gaussianos

A.5.1. Definición

Un proceso Gaussiano es una colección de variables aleatorias, que cumplen que cualquier subconjunto finito de la colección, tiene una distribución Gaussiana conjunta

A.5.2. Propiedades

A.5.2.1. Distribución conjunta de vectores aleatorios

$$\begin{Bmatrix} x \\ y \end{Bmatrix} \sim N \left(\begin{Bmatrix} \mu_x \\ \mu_y \end{Bmatrix}, \sigma^2 \begin{Bmatrix} A & C \\ C^T & B \end{Bmatrix} \right) \quad (\text{A.18})$$

A.5.2.2. Marginalidad en vectores aleatorios

Si se tiene la distribución A.18, entonces la distribución marginal de x es:

$$x \sim N(\mu_x, A) \quad (\text{A.19})$$

A.5.2.3. Condicionalidad en vectores aleatorios

Si se tiene la distribución A.18, entonces la distribución condicional de x es:

$$x|y = N(\mu_x + CB^{-1}(y - \mu_y), A - CB^{-1}C^T) \quad (\text{A.20})$$

Anexo B

Herramientas Python

B.1. Librerías

En el presente trabajo se utilizan las siguientes librerías de Python.

- i. Numpy [45]
- ii. Pandas [41]
- iii. Matplotlib [46]
- iv. Scikit-learn [19]
- v. OpenTURNS [20]
- vi. Scipy [47]
- vii. Pydace [48]

Anexo C

Procedimiento de selección de parámetros

C.1. Parámetros a analizar

Los parámetros posibles para analizar y que se ingresan como variables de entrada en el modelo son los siguientes:

Tabla C.1: Parámetros a analizar - Proceso de selección

Parámetro	Unidad
Densidad	[kg/m ³]
Ángulo de fricción (ϕ)	[°]
Cohesión	[N/m ²]
Módulo de Young (E)	[N/m ²]
Coefficiente de Poisson (ν)	[-]
Resist. corte no drenada (S_u)	[N/m ²]
Coefficiente de presión lateral (K_0)	[-]
Nivel freático	[m]
Nivel freático 2	[m]
Altura de muro	[m]
Ancho de coronamiento	[m]
Altura de base	[m]
Ancho de base	[m]
Revancha	[m]
Talud	[-]
Discretización	[-]

C.2. Ejecución de casos

En primer lugar se consideran las siguientes variables que se establecen como fijas. Estas son principalmente los parámetros geométricos, exceptuando el talud y la revancha, para los cuales se espera que no existan cambios en un sistema estable estática y dinámicamente. Estos son la altura de muro, de base, el ancho de base y de coronamiento. La discretización se considera fija también ya que es determinado una vez y no presenta variabilidad en los siguientes procesos. El talud se considera más adelante. Cabe mencionar que la revancha se considera y fija para el caso más desfavorable posible, es por esto que no se aplica variabilidad en este parámetro. Por otro lado la densidad saturada es dependiente de la densidad seca, por lo que se toma en consideración esta última simplemente. Esto sucede también con el coeficiente de presión lateral que depende del ángulo de fricción.

Para el análisis de las demás variables y su incidencia en el sistema, se consideran dos valores considerados extremos (máximo y mínimo) para cada variable, de modo de comparar la variabilidad del sistema frente a ella, específicamente la variabilidad del Factor de Seguridad global, en base a la siguiente ecuación:

$$Variabilidad = \frac{|FS_{m\acute{a}x} - FS_{m\acute{i}n}|}{\frac{V_{arm\acute{a}x} - V_{arm\acute{i}n}}{V_{arm\acute{a}x}}} \quad (C.1)$$

El sistema se fija para los siguientes valores (exceptuando por la variable que se considera en cada caso de sensibilidad):

Tabla C.2: Parámetros fijos - Proceso de selección

Parámetro	Valor	Unidad
Densidad	1960	[kg/m ³]
Densidad saturada	2000	[kg/m ³]
Ángulo de fricción (ϕ)	33.5	[°]
Cohesión	7500	[N/m ²]
Módulo de Young (E)	5e7	[N/m ²]
Coefficiente de Poisson (ν)	0.3	[-]
Resist. corte no drenada (S_u)	0.1*esfzz	[N/m ²]
Nivel freático	18.5	[m]
Nivel freático 2	18.5	[m]
Altura de muro	40	[m]
Ancho de coronamiento	5	[m]
Altura de base	20	[m]
Ancho de base	300	[m]
Revancha	3	[m]
Talud	2:1	[-]

En la siguiente Tabla se presenta la variabilidad mencionada para las variables en cuestión:

Tabla C.3: Variabilidad de parámetros a analizar

Parámetro	Unidad	Val. mín.	Val. máx.	FS mín.	FS máx.	Variab.
Densidad	[kg/m ³]	1500	2500	1.3688	1.3813	0.0312
Ángulo de fricción (ϕ)	[°]	30	37	1.2313	1.5688	1.7834
Cohesión	[N/m ²]	0	15000	1.2063	1.4938	0.2874
Modulo de Young (E)	[N/m ²]	5e6	5e8	1.3813	1.3813	0
Coefficiente de Poisson (ν)	[-]	0.2	0.4	1.3813	1.3813	0
Resist. corte no drenada (S_u)	[m]	0.05*esfzz	0.1*esfzz	1.3813	1.3813	0
Nivel freático	[m]	0	37	1.5438	1.2938	0.2479
Nivel freático 2	[m]	0	37	1.3938	1.3813	0.0111
Talud	[m]	1.5	2	1.0938	1.3813	1.15

De este modo se puede identificar que los parámetros que influirían de mayor forma en la respuesta del sistema (FS global), serían los indicados en la siguiente tabla.

Tabla C.4: Parámetros a analizar - Selección previa

Parámetro	Unidad
Nivel freático	[m]
Nivel freático 2	[m]
Ángulo de fricción	[°]
Cohesión	[N/m ²]
Densidad	[kg/m ³]
Talud	[-]

Sin embargo, es necesario realizar un mayor enfoque al problema físico y a la experiencia y bibliografía, de modo que es posible filtrar aún más estos parámetros. Esto se realiza debido a que si bien es posible usar todas estas variables para obtener una base de datos, para ello se requiere una mayor cantidad de equipos o servidores para realizar las ejecuciones en el programa FLAC3D, o conllevaría mayor tiempo y carga computacional en caso de usar menos servidores. Se discute en mayor detalle este tema en el capítulo de Análisis y Discusión de resultados.

Dicho esto, se considera entonces excluir del análisis la variable referida al talud, tomando como argumento el hecho de que la variabilidad en sus valores tiene una baja probabilidad, por lo que se considera fija. La densidad también se descarta para dar prioridad a la variable de nivel freático 2, la cual si bien provoca menor variabilidad, su análisis es significativo para este trabajo, principalmente por el problema aplicado, referido al sistema de monitoreo, que hace énfasis en estas variables asociado a la altura del nivel freático. De este modo se obtiene la siguiente Tabla:

Tabla C.5: Parámetros a analizar - Selección filtrada

Parámetro	Unidad
Nivel freático	[m]
Nivel freático 2	[m]
Ángulo de fricción	[°]
Cohesión	[N/m ²]

Por otro lado el tiempo de ejecución de estos modelos se define por la siguiente ecuación:

$$T = N * t_{caso}[\text{horas}] \quad (\text{C.2})$$

donde N es igual a 18 en este caso y t_{caso} es igual a 3 horas, lo cual da como resultado 54 horas, es decir 2.3 días aproximadamente.

C.3. Parámetros a analizar finales

Finalmente, realizado el procedimiento anterior, se seleccionan las siguientes variables para analizar y combinar para luego ingresar al sistema.

Tabla C.6: Parámetros a analizar - Selección final

Parámetro	Unidad
Nivel freático	[m]
Nivel freático 2	[m]
Ángulo de fricción	[°]
Cohesión	[N/m ²]

Anexo D

Distribución de casos a ejecutar

D.1. Gráficos de distribución de casos

A continuación se presentan los gráficos referidos a la distribución de los casos a ejecutar en el modelo original o de diseño en el software FLAC3D. En primer lugar se tiene la distribución por Latin Hypercube Sampling, en la cual es posible visualizar que los puntos que representan los casos o combinaciones están distribuidos de forma heterogénea dentro del rango de cada variable.

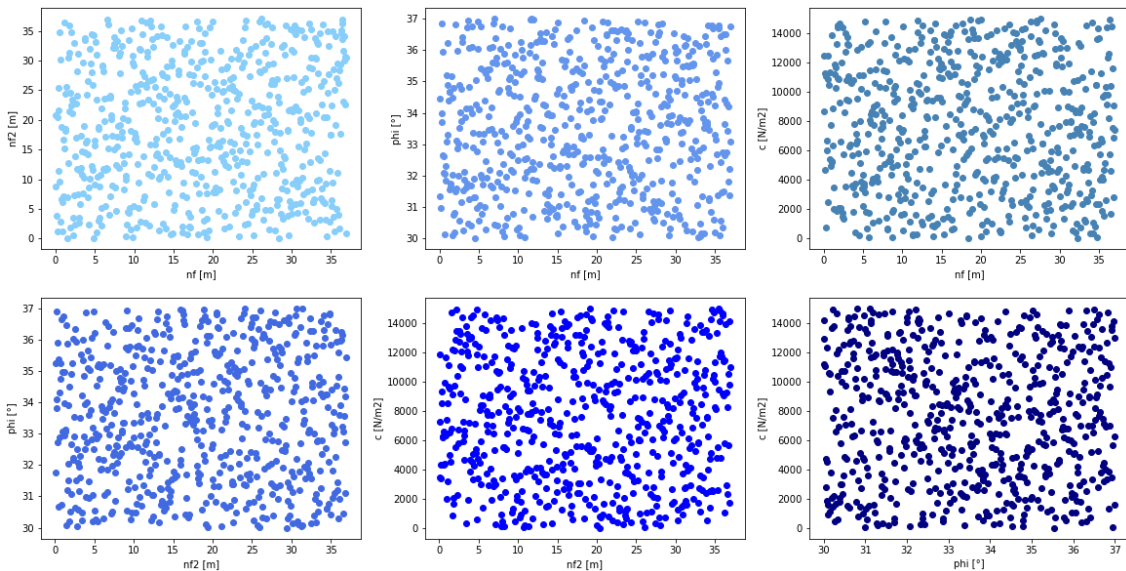


Figura D.1: Dsitribución de muestras - LHS

En segundo lugar se tiene una distribución realizando una división equiespaciada de los rangos de cada variable independientemente y aplicando uan permutación de estos valores. Se obtiene la siguiente Figura.

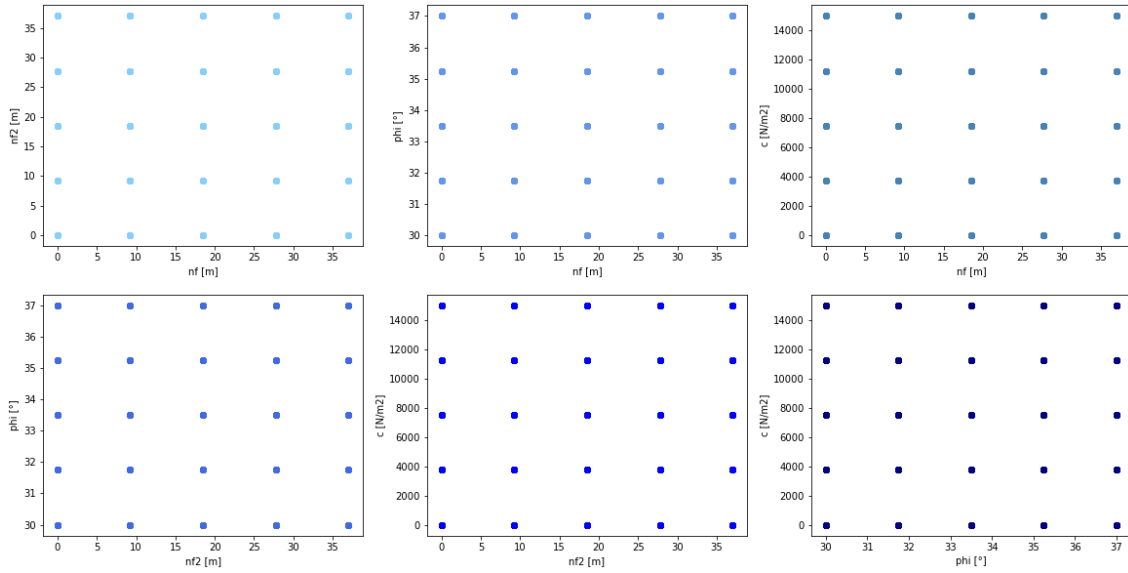


Figura D.2: Dsitrribución de muestras - Equiespaciamento uniforme de variables

Anexo E

Resultados de modelo de embalse de relaves en software de diseño

En esta sección se presentan algunos resultados del modelo de embalse de relaves diseñado en el software FLAC3D.

E.1. Factor de seguridad global y contorno

E.1.1. Caso más favorable

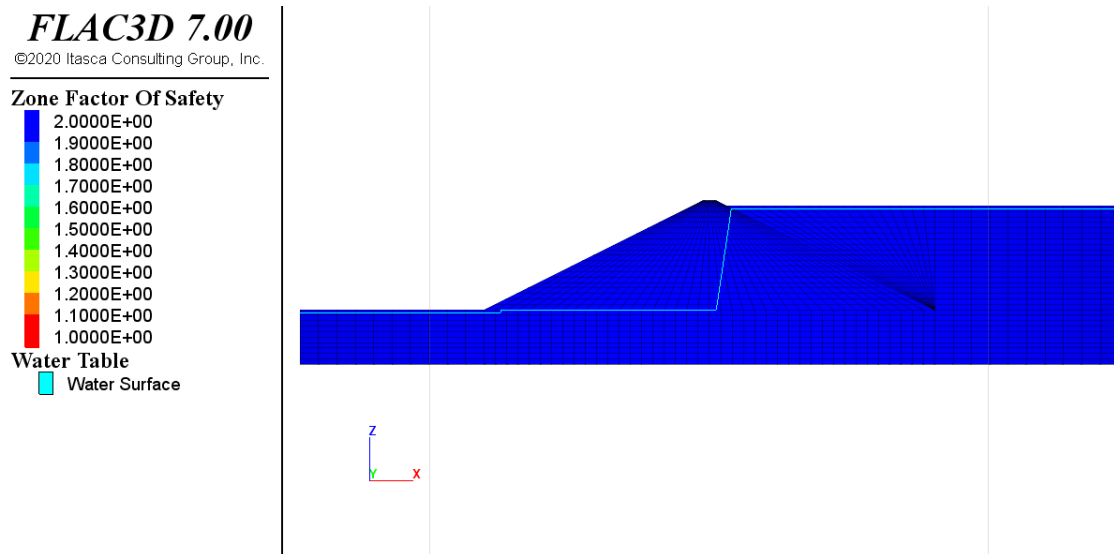


Figura E.1: Caso más favorable - Contorno FS - FSG=1.88

E.1.2. Caso intermedio

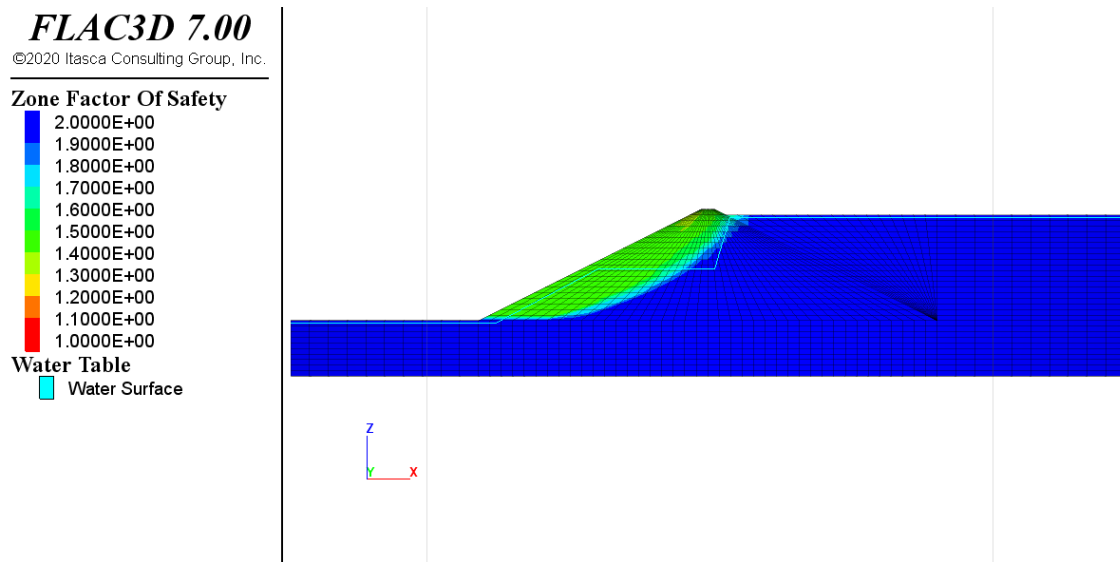


Figura E.2: Caso más desfavorable - Contorno FS - FSG=1.3813

E.1.3. Caso más desfavorable

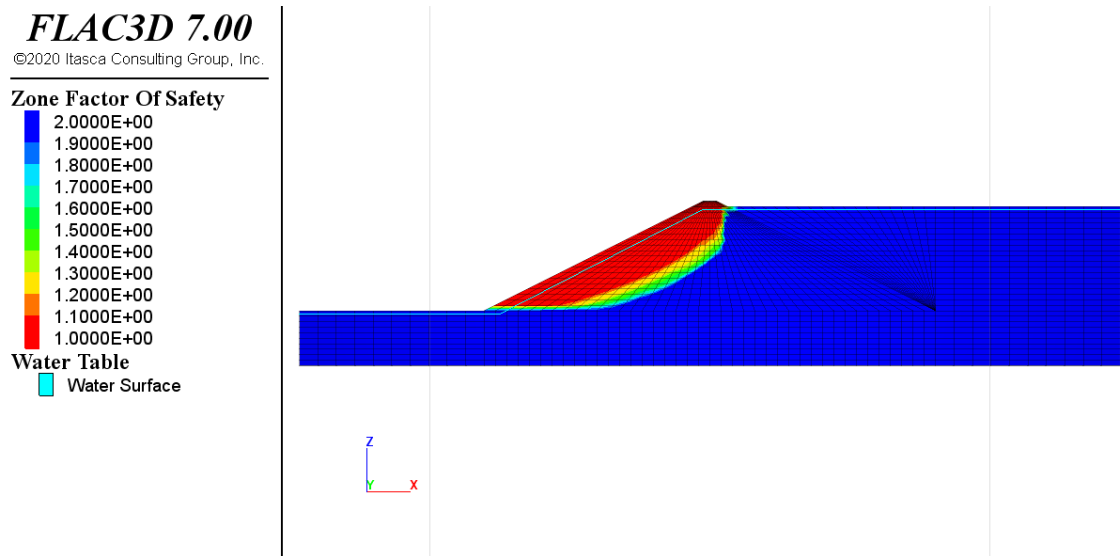


Figura E.3: Caso más desfavorable - Contorno FS - FSG=1.006

E.2. Presiones de poros

E.2.1. Caso más favorable

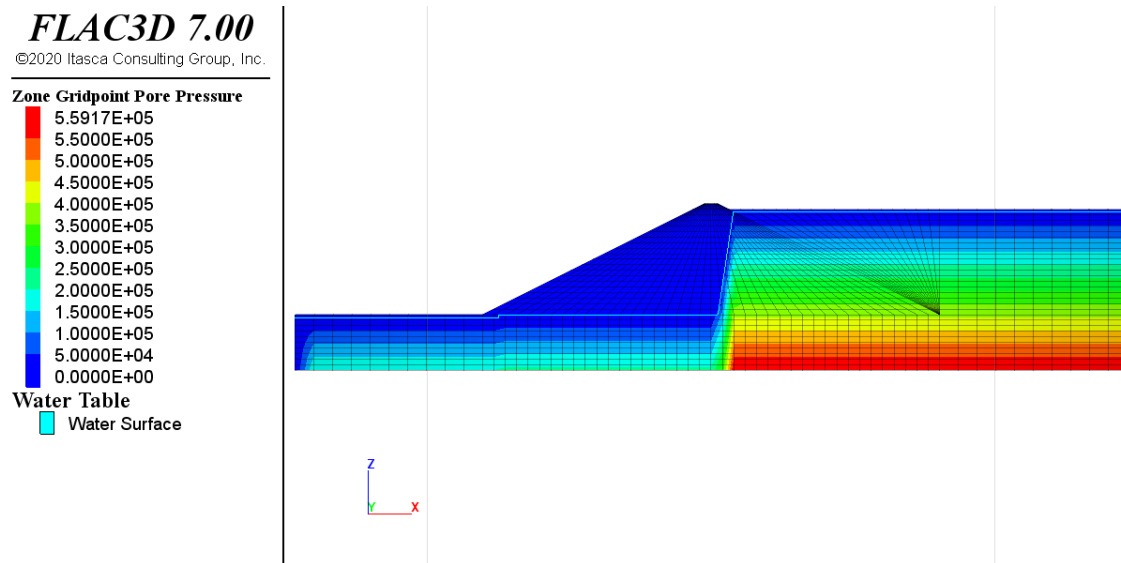


Figura E.4: Caso más favorable - Presión de poros - FSG=1.88

E.2.2. Caso intermedio

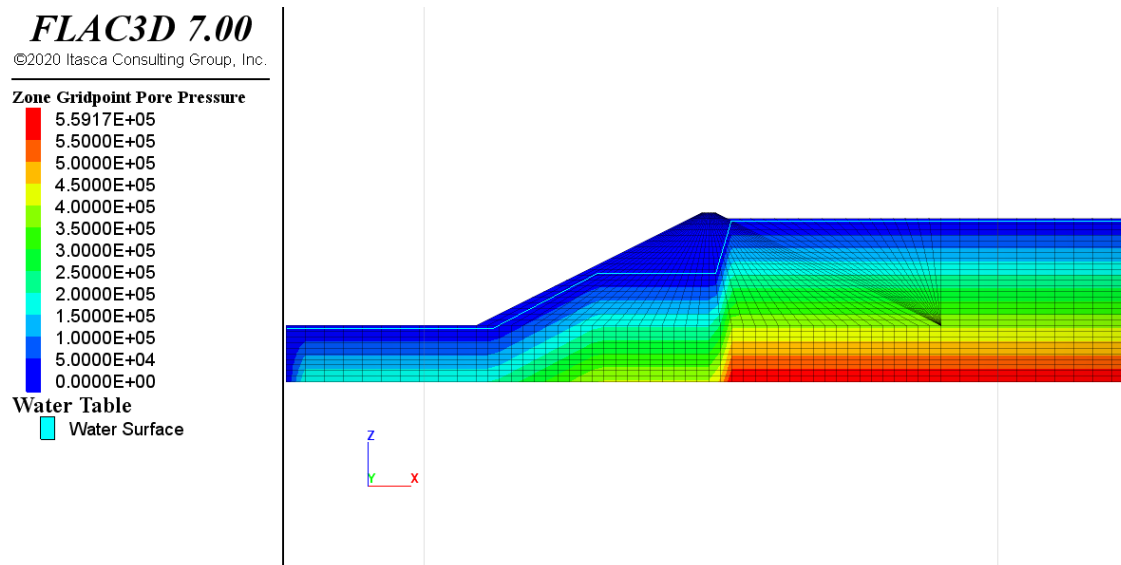


Figura E.5: Caso intermedio - Presión de poros - FSG=1.3813

E.2.3. Caso más desfavorable

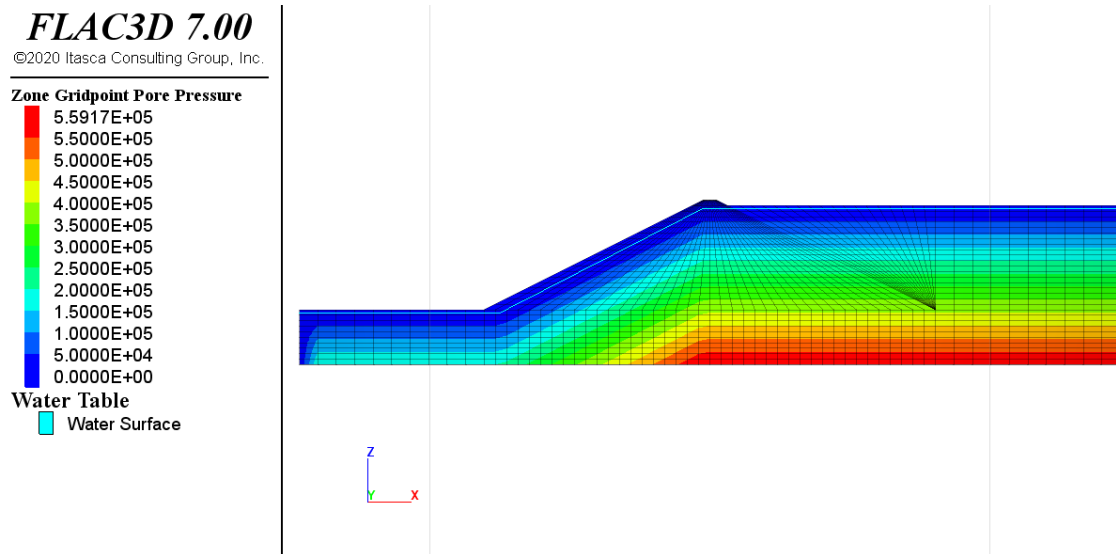


Figura E.6: Caso más desfavorable - Presión de poros - FSG=1.006

E.3. Desplazamientos

E.3.1. Caso más favorable

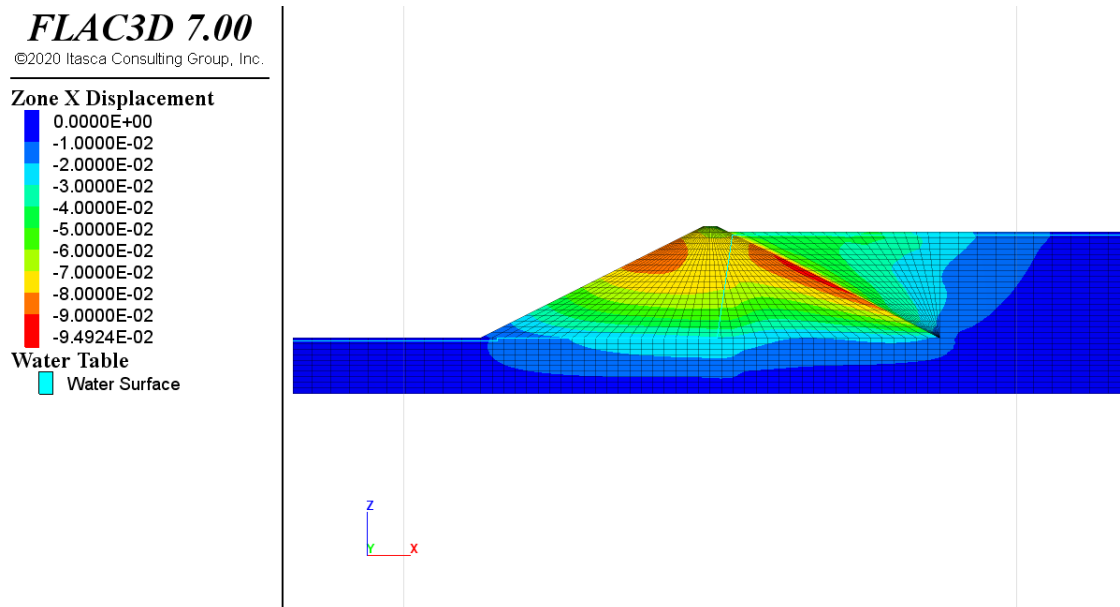


Figura E.7: Caso más favorable - Desplazamientos en XX - FSG=1.88

E.3.2. Caso intermedio

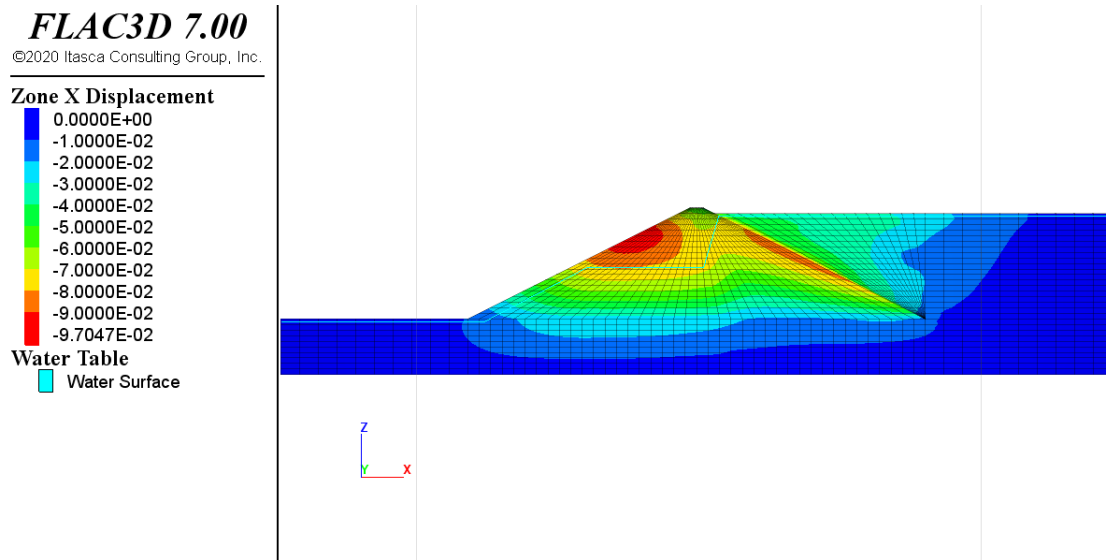


Figura E.8: Caso intermedio - Desplazamientos en XX - FSG=1.3813

E.3.3. Caso más desfavorable

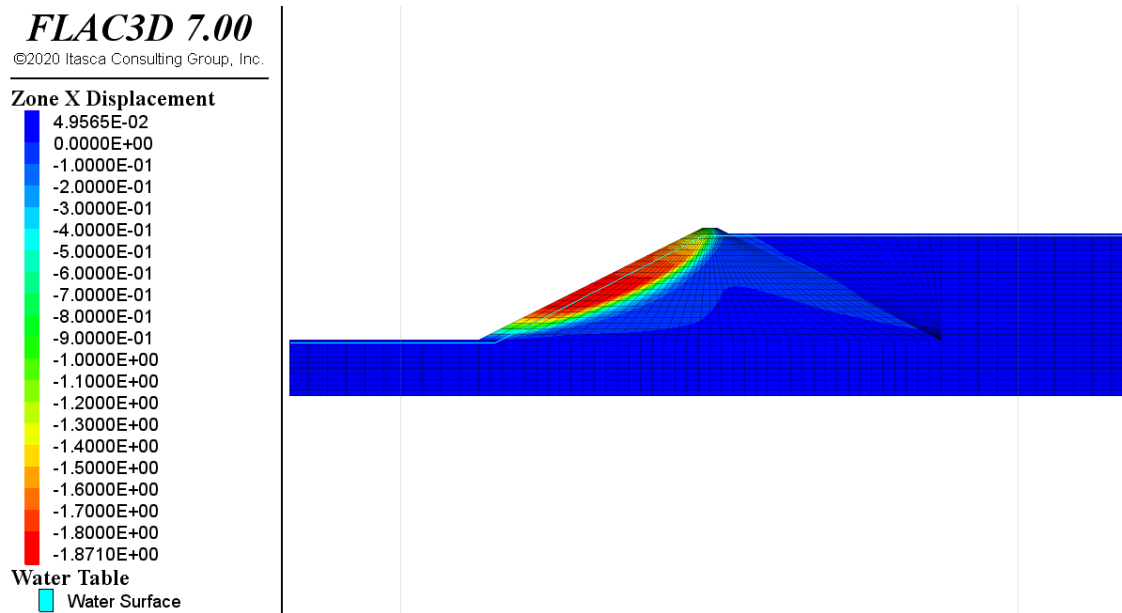


Figura E.9: Caso más desfavorable - Desplazamientos en XX - FSG=1.006

E.4. Esfuerzos

E.4.1. Caso más favorable

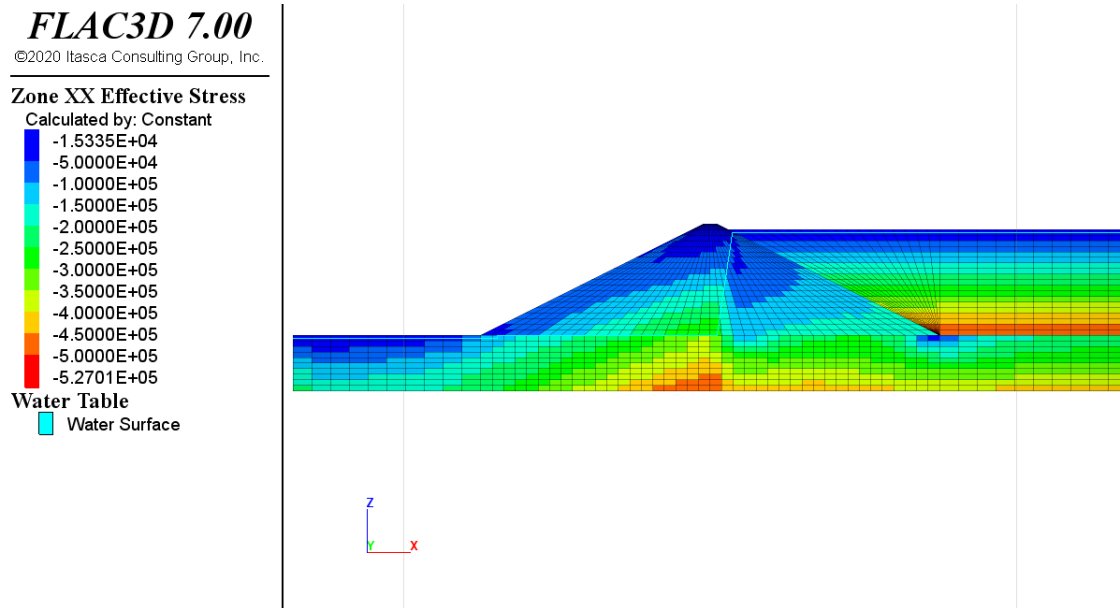


Figura E.10: Caso más favorable - Esfuerzos en XX - FSG=1.88

E.4.2. Caso intermedio

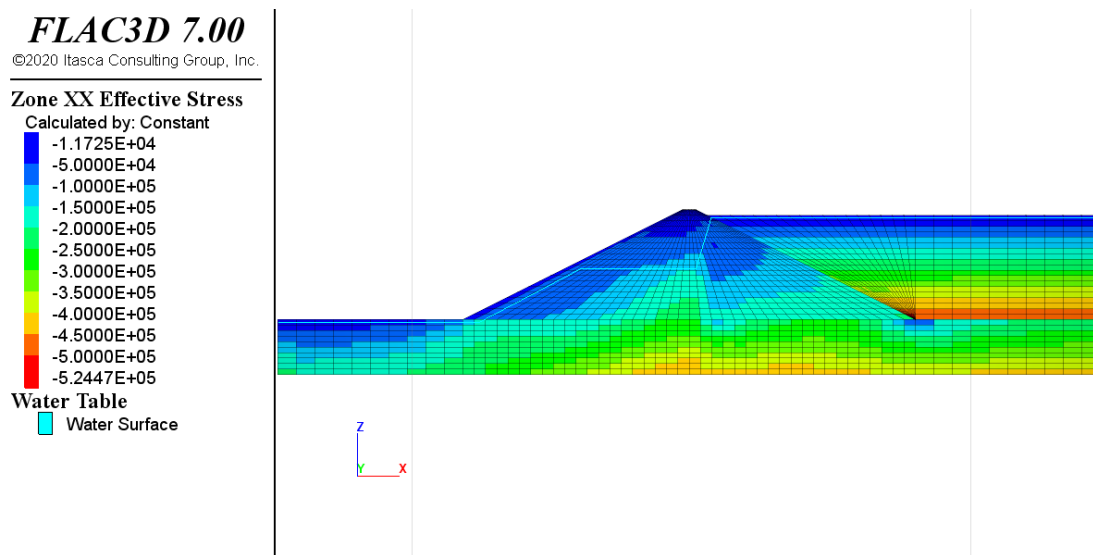


Figura E.11: Caso intermedio - Esfuerzos en XX - FSG=1.3813

E.4.3. Caso más desfavorable

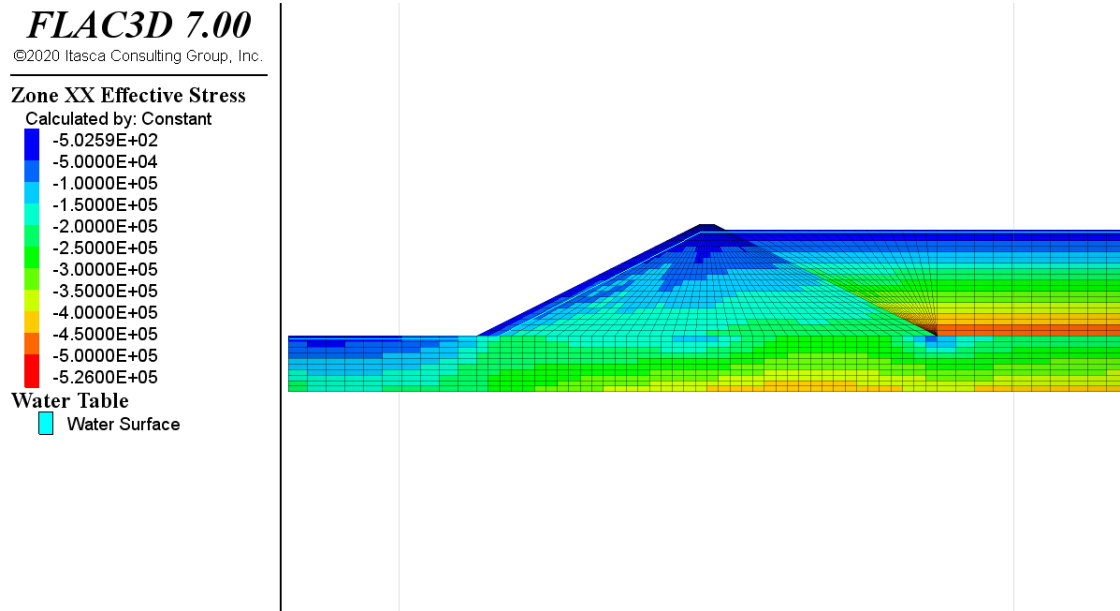


Figura E.12: Caso más desfavorable - Esfuerzos en XX - FSG=1.006

Anexo F

Desarrollo: Caso con muestreo con LHS

En este anexo se desarrolla el proceso de confección presentado en el trabajo, pero haciendo una diferencia en el proceso de distribución de casos.

F.1. Selección de parámetros

En esta etapa se seleccionan los mismos parámetros, el rango también es el mismo, y cambia la distribución de los casos dentro del rango, esta vez dada por el método de muestreo LHS (Latin Hypercube Sampling) y entonces se obtiene la siguiente Tabla:

Tabla F.1: Casos obtenidos por método

Caso	Combinación por caso
1	[5.03, 15.72, 34.12, 4347,8]
2	[23.57, 10.79, 32.93, 1176.58]
...	...
n	[19.12, 9.02, 32.58, 10878.02]

Una vez obtenidas las combinaciones de casos, se procede a ejecutarlos en el software FLAC3D y se obtienen los resultados. Cabe mencionar que el diseño del embalse de relaves en el software, es el mismo.

F.2. Confección de modelo sustituto

Esta etapa conlleva el mismo procedimiento, estableciendo las mismas propiedades en el método Kriging. Es importante visualizar los resultados obtenidos por este entrenamiento, por lo que se presentan los gráficos referidos al comportamiento del modelo sustituto, asociado al modelo de FS, en comparación con el modelo original de FLAC3D. Se utiliza como datos del modelo original, los datos de soporte utilizados en el modelo presentado en primera instancia (modelo base del trabajo de memoria).

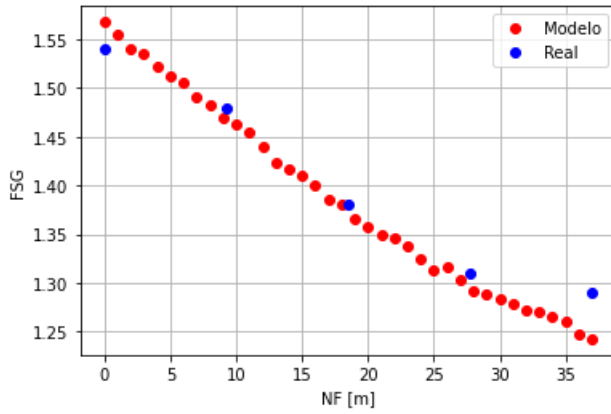


Figura F.1: FSG versus nivel freático - Modelo sustituto

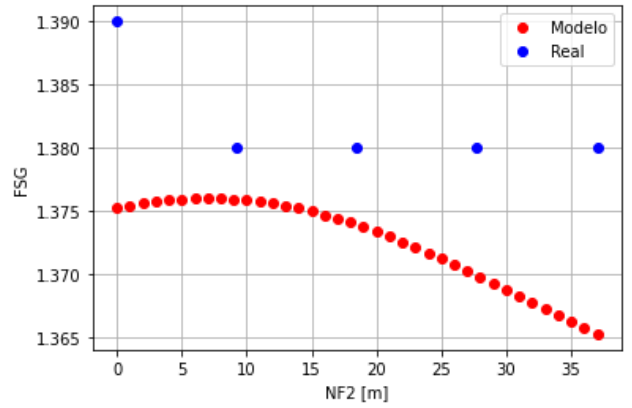


Figura F.2: FSG versus nivel freático 2 - Modelo sustituto

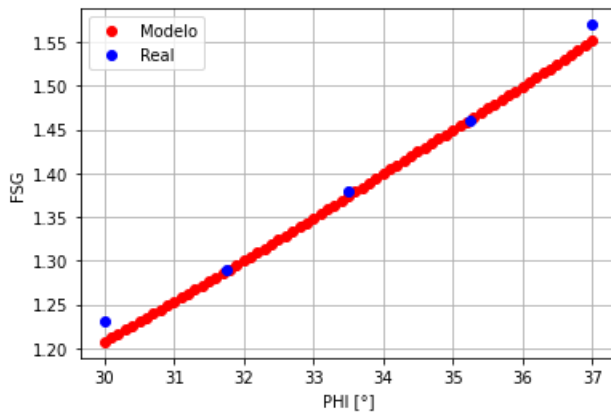


Figura F.3: FSG versus ángulo de fricción (ϕ) - Modelo sustituto

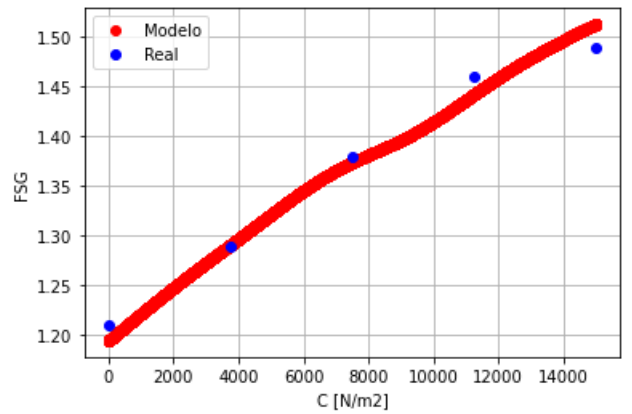


Figura F.4: FSG versus cohesión - Modelo sustituto

Por otro lado, se obtiene los errores de la validación cruzada para el mismo modelo de FS.

Tabla F.2: Errores - Valid. Cruzada - Modelo de FS global

Caso	E. resid.	E. resid. cuad.
1	0.004	1e-5
2	0.012	0.0002
...
n	0.0003	6e-8
Promedio	0.006	6e-5

Se tiene entonces los siguientes errores representativos:

Tabla F.3: Errores - Modelo de FS global - Anexo F

Error	Valor
Error absoluto medio	0.006
RMSE	0.0077
R^2	0.997

Este desarrollo se realiza para comparar los modelos en términos de precisión.

Anexo G

Desarrollo: Caso con menor cantidad de datos

Se presenta en este anexo el desarrollo del modelo para una cantidad menor de datos o valores dentro del rango de cada variable de entrada.

G.1. Selección de parámetros

Observando la Tabla ??, se opta por la opción 4 en este caso, donde se seleccionan 3 valores por variable, y el tiempo total de ejecución de casos se reduce considerablemente. Entonces los valores de las variables, manteniendo el mismo rango por cada una, queda establecido del siguiente modo:

Tabla G.1: Selección de valores de parámetros por método

Parámetro	Combinación por variable
Nivel freático	[0, 18.5, 37]
Nivel freático 2	[0, 18.5, 37]
Ángulo de fricción	[30, 33.5, 37]
Cohesión	[0, 7500, 15e3]

Tabla G.2: Casos obtenidos por método

Caso	Combinación por caso
1	[0, 0, 30, 0]
2	[0, 0, 30, 7500]
...	...
n	[37, 37, 37, 15e3]

G.2. Confección de modelo sustituto

Esta etapa conlleva el mismo procedimiento, estableciendo las mismas propiedades en el método Kriging. Es importante visualizar los resultados obtenidos por este entrenamiento, por lo que se presentan las gráficas referidas al comportamiento del modelo sustituto, asociado al modelo de FS, en comparación con el modelo original de FLAC3D. Se utiliza como datos del modelo original, los datos de soporte utilizados en el modelo presentado en primera instancia (modelo base del trabajo de memoria).

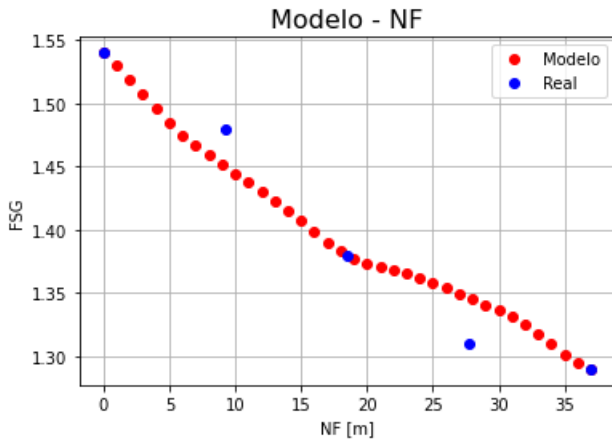


Figura G.1: FSG versus nivel freático - Modelo sustituto

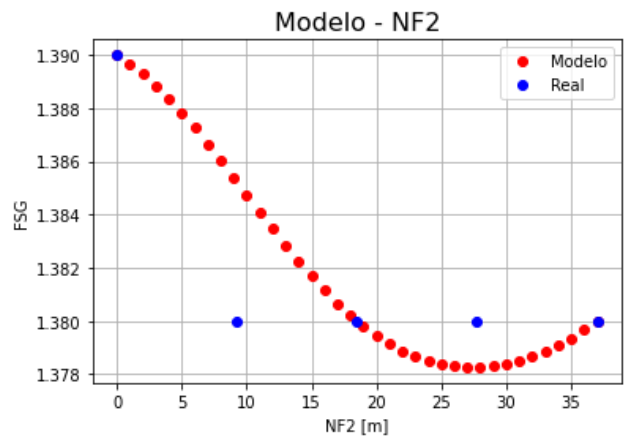


Figura G.2: FSG versus nivel freático 2 - Modelo sustituto

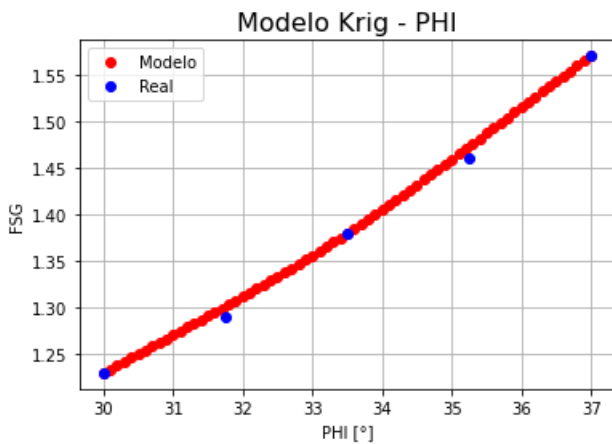


Figura G.3: FSG versus ángulo de fricción (ϕ) - Modelo sustituto

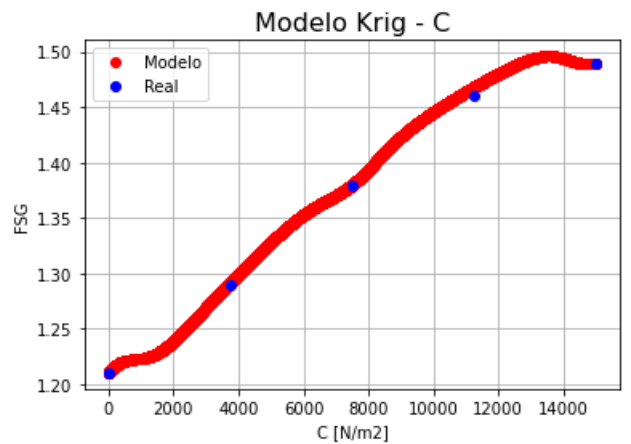


Figura G.4: FSG versus cohesión - Modelo sustituto

Por otro lado, se obtiene los errores de la validación cruzada para el mismo modelo de FS.

Tabla G.3: Errores - Valid. Cruzada - Modelo de FS global

Caso	E. resid.	E. resid. cuad.
1	0.013	0.0002
2	0.024	0.0006
...
n	0.039	0.002
Promedio	0.015	0.0004

Se tiene entonces los siguientes errores representativos:

Tabla G.4: Errores - Modelo de FS global - Anexo G

Error	Valor
Error absoluto medio	0.015
RMSE	0.021
R ²	0.991

Finalmente se realiza una comparación respecto al tiempo, donde la opción 1 representa la opción con el modelo base, y la opción 2 representa la opción con el nuevo modelo presentado (menos casos).

Tabla G.5: Comparación - Tiempo de ejecución

Opción	Valores por var.	Casos totales	T. por caso [hrs]	T. total [días]	T. total ef. [días]
1	5	625	3	78.2	15.63
2	3	81	3	10.13	2.03

Lo cual demuestra una considerable disminución en el proceso de ejecución de casos, de aproximadamente un 87 % que implica una gran reducción del tiempo total de procesamiento, al ser este proceso el que conlleva mayor tiempo y carga computacional.

Este desarrollo se realiza para comparar los modelos en términos de precisión y tiempos de procesamiento.

Anexo H

Desarrollo: Caso con librería alternativa

Se presenta en este anexo el desarrollo del modelo utilizando una librería distinta. En vez de ocupar la librería de Python OpenTURNS, se utiliza la librería de Pydace, la cual es una adaptación a Python de la librería original DACE disponible en Matlab.

H.1. Selección de parámetros

El proceso de selección no varía, ni la distribución de los casos, respecto al modelo base del trabajo de memoria.

H.2. Confección de modelo sustituto

Se utiliza la misma base de datos ya filtrada y transformada, y se ingresa en el algoritmo de la librería Pydace, bajo los mismos criterios y propiedades relacionadas a este método, ocupadas en el modelo base. Los resultados obtenidos se presentan en la siguientes Figuras:

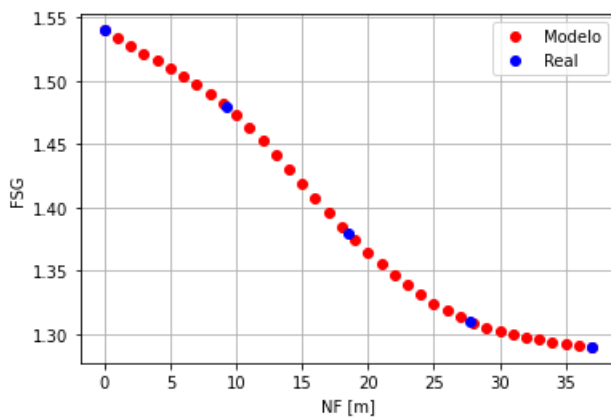


Figura H.1: FSG versus nivel freático - Modelo sustituto

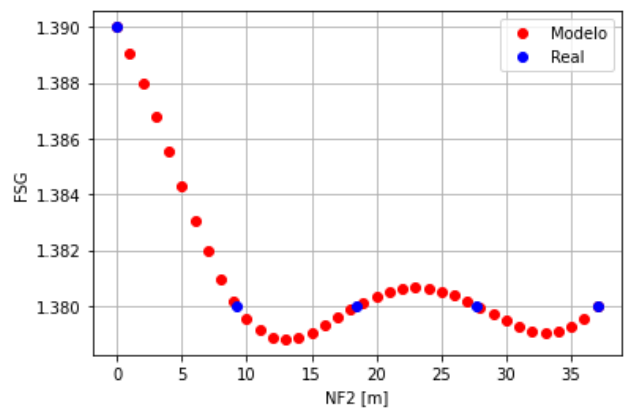


Figura H.2: FSG versus nivel freático 2 - Modelo sustituto

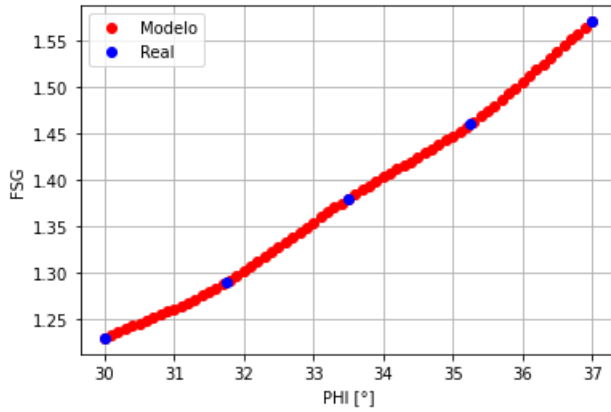


Figura H.3: FSG versus ángulo de fricción (ϕ) - Modelo sustituto

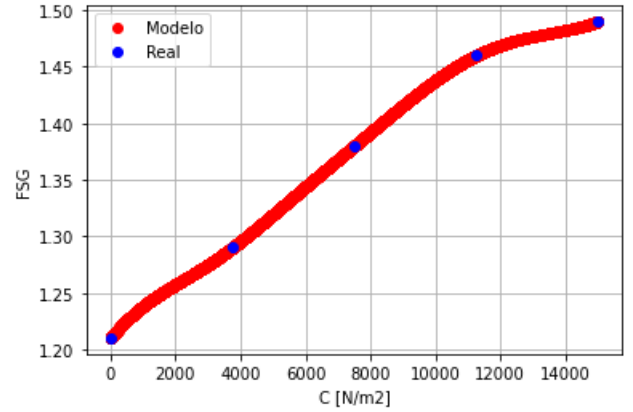


Figura H.4: FSG versus cohesión - Modelo sustituto

Se obtiene los errores de la validación cruzada para el mismo modelo de FS.

Tabla H.1: Errores - Valid. Cruzada - Modelo de FS global

Caso	E. resid.	E. resid. cuad.
1	0.024	0.0006
2	0.02	0.0004
...
n	0.013	0.0002
Promedio	0.011	0.0002

Se tiene entonces los siguientes errores representativos:

Tabla H.2: Errores - Modelo de FS global - Anexo H

Error	Valor
Error absoluto medio	0.011
RMSE	0.0154
R^2	0.994

Por otro lado, se realiza un análisis más detallado en este caso, efectuando el procesamiento del modelo de contorno FS. En primer lugar, se tienen los siguientes errores residuales de los datos de soporte o de entrenamiento:

Tabla H.3: Errores - Datos de soporte - Modelo de contorno de FS - Anexo H

Error	1.0	1.1	1.2	1.3	1.4
FSG	2e-15	2e-15	2e-15	2e-15	3e-15
C0	0.565	0.718	0.811	0.962	0.929
C1	0.191	0.281	0.202	0.304	0.203
C2	0.006	0.008	0.006	0.008	0.006
C3	4e-5	5e-5	4e-5	5e-5	4e-5

Tabla H.4: Continuación Tabla H.3

Error	1.5	1.6	1.7	1.8	1.9	Prom.
FSG	2e-15	2e-15	2e-15	2e-15	2e-15	2e-15
C0	0.973	0.918	0.99	0.867	7e-14	0.773
C1	0.184	0.167	0.128	0.112	1e-14	0.177
C2	0.005	0.005	0.004	0.003	4e-16	0.005
C3	4e-5	4e-5	3e-5	3e-5	3e-18	4e-5

En segundo lugar se obtienen los errores residuales dados por la validación cruzada.

Tabla H.5: Errores residuales - Valid. Cruzada - Modelo de contorno de FS

Error	1.0	1.1	1.2	1.3	1.4
FSG	0.018	0.011	0.018	0.011	0.018
C0	0.403	0.723	1.573	1.329	1.711
C1	0.028	0.046	0.248	0.129	0.233
C2	0.001	0.001	0.006	0.003	0.006
C3	5e-6	9e-6	4e-5	2e-5	4e-5

Tabla H.6: Continuación Tabla H.5

Error	1.5	1.6	1.7	1.8	1.9	Prom.
FSG	0.016	0.018	0.015	0.015	0.017	0.015
C0	1.6	1.633	1.27	0.483	0.546	1.13
C1	0.252	0.175	0.163	0.082	0.092	0.145
C2	0.007	0.005	0.005	0.003	0.003	0.004
C3	5e-5	3e-5	2e-5	2e-5	2e-5	2e-5

Se tienen además los errores RMSE y se adiciona un promedio asociado a los errores LOO de cada submodelo.

Tabla H.7: Errores RMSE y LOO - Modelo de contorno de FS

Errores	1.0	1.1	1.2	1.3	1.4
FSG	0.032	0.032	0.032	0.032	0.032
C0	1.5	2.634	3.722	2.8	3.187
C1	0.089	0.095	0.978	0.207	0.68
C2	0.003	0.003	0.032	0.006	0.032
C3	2e-5	2e-5	2e-4	5e-5	1e-4

Tabla H.8: Continuación Tabla H.7

Errores	1.5	1.6	1.7	1.8	1.9	RMSE	Err.<i>LOO</i>
FSG	0.032	0.032	0.032	0.032	0.032	0.032	0.001
C0	2.611	2.982	2.187	0.915	1.226	2.376	6.386
C1	0.636	0.349	0.794	0.17	0.215	0.421	0.271
C2	0.032	0.01	0.032	0.006	0.006	0.016	4e-4
C3	1e-4	7e-5	1e-4	5e-5	5e-5	8e-5	8e-9

Tabla H.9: Tiempo de ejecución de submodelos de contorno de FS

Opción	Submodelos	T. por submodelo [hrs]	T. total [hrs]
1	10	3	30
2	10	0.02 (1 min)	0.2

Este desarrollo se realiza para comparar los modelos en términos de precisión, tiempos de procesamiento y carga computacional.