



**UNIVERSIDAD DE CHILE**

**FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS**

**DEPARTAMENTO DE INGENIERÍA DE MINAS**

**SIMULACIÓN SECUENCIAL GAUSSIANA NO ESTACIONARIA DE LEYES**

**TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN MINERÍA**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL DE MINAS**

**YERKO ANDRÉ MARTÍNEZ FERNÁNDEZ**

**PROFESOR GUÍA:**

**JULIÁN ORTIZ CABRERA**

**MIEMBROS DE LA COMISIÓN**

**XAVIER EMERY**

**JUAN LUIS YARMUCH GUZMÁN**

**ALEJANDRO LEIVA RODRÍGUEZ**

**SANTIAGO DE CHILE**

**ENERO DE 2014**

## RESUMEN

El presente trabajo corresponde al desarrollo de una herramienta que permite simular valores de una variable regionalizada considerando que tales valores tienen una variación sistemática en el espacio. En este contexto, se desarrolla una nueva herramienta de simulación consistente en un algoritmo de simulación Gaussiana secuencial con rechazo considerando una deriva de referencia como input, bajo la hipótesis que esta herramienta permite respetar tal deriva, obteniendo resultados representativos de la base de datos en cuanto a sus estadísticos de orden 1 (histograma) y orden 2 (variograma). La metodología del algoritmo comienza definiendo la secuencia de visitas de nodos a simular de manera aleatoria. Se acepta o rechaza el nodo simulado en base a la deriva de referencia considerando un rechazo determinístico o probabilístico y una tolerancia dinámica. Para cada nodo se considera una vecindad de búsqueda de datos condicionantes para la simulación y una vecindad de búsqueda de datos para el cálculo de una media local simulada. El algoritmo permite ajustar el número aceptable de rechazos, el tamaño de la vecindad de búsqueda de la media local, la tolerancia y el tipo de rechazo.

Se presentan dos casos de estudio. El primero consiste en un ejemplo sintético de una coordenada con deriva lineal. En este primer caso se tiene que, a mayor tolerancia o mayor vecindad de búsqueda de la media local, los valores simulados se distribuyen con mayor dispersión en torno a la deriva de referencia. El segundo estudio de caso consiste en una zona de interés del yacimiento Compañía Minera Cerro Colorado donde se realiza el proceso de simulación en seis unidades de estimación considerando diecisiete sensibilizaciones de los parámetros del algoritmo más una simulación basada en Kriging Simple (SK) y otra basada en Kriging de residuos (BT). En el caso de presencia de deriva se obtiene en general mejores resultados con el algoritmo propuesto que con el SK o BT cuando la deriva se ve reflejada de manera clara en el variograma como en la unidad de estimación cuatro. Las estadísticas de validación en términos de desempeño de las simulaciones como estimación (coeficiente de determinación  $R^2$ , pendiente de la regresión de datos reales versus simulados y error medio) y en términos de cuantificación de la incertidumbre de los datos originales (accuracy plot) mejoran en relación al SK y BT. De esta manera, la herramienta desarrollada ofrece una alternativa flexible que mejora los estadísticos de validación en comparación al enfoque tradicional frente a un escenario de simulación con presencia de deriva clara en el variograma.

## ABSTRACT

This research study comprises the development of a computational tool which simulates values from a regionalized variable considering a systematic spatial variation, in other words, the presence of a drift. In this context, this new tool consists in a sequential Gaussian simulation algorithm with rejection based on a reference drift as input, on the hypothesis that this new algorithm obtains representative results of the original data base regarding to its first (histogram) and second (variogram) order statistics. The first step of the global algorithm methodology defines the random path to visit each node. Simulation is conducted and the resulting value is accepted or rejected based on the reference drift considering a deterministic or probabilistic rejection rule and a dynamic tolerance. Two different neighborhoods are defined to simulate each node: conditioning data neighborhood to simulate values in each node and local mean neighborhood to calculate the simulated local mean. The algorithm integrates flexibility by the possibility to adjust different parameters: number of rejections, size of local mean neighborhood, tolerance and rejection type.

Two case studies are undertaken. The first one is a synthetic one-coordinate example with lineal drift. This first case study indicates that the bigger the tolerance or size of local mean neighborhood, the bigger the dispersion of the simulated values around the reference drift. The second case study corresponds to an area of interest of the Compañía Minera Cerro Colorado orebody, in which the whole simulation process is conducted on six estimation units. This process is performed using simple Kriging (SK), Kriging of residuals (BT) and seventeen sensitization of the rejection algorithm. When drift is present, better results are obtained in comparison with SK and BT, moreover, when this drift is clearly seen in the experimental variogram as in estimation unit number four, validation statistics of simulations as estimation (coefficient of determination  $R^2$ , regression slope of original data versus simulated data, and mean error) and as uncertainty characterization (accuracy plot), improve their results in relation to SK and BT. This way, the developed tool offers a flexible alternative able to improve the validations statistics in comparison with traditional approach when a clearly drift-variogram scenario is simulated.

*Un viaje de mil millas comienza*

*con un solo paso*

**A mi familia**

## AGRADECIMIENTOS

Agradezco a mis padres, Flor y César, por inculcarme buenos valores desde pequeño y darme la base para hacer de mí la persona que hoy soy. Por enseñarme que independiente de las convicciones de cada uno, lo importante en la vida es obrar de manera justa y respetuosa, que para triunfar se debe ser perseverante y luchar por lo que uno quiere intentándolo una y otra vez hasta alcanzarlo. Agradezco en particular a mi madre por mantener siempre el empuje que la caracteriza y por nunca rendirse ante nada, por alentarme siempre a seguir y por apoyarme a lo largo de todo este camino de manera incondicional, sin duda sin ella no estaría en este momento donde estoy. Agradezco a mi mujer y compañera, Diana, por estar presente a lo largo de esta etapa de mi vida y tener siempre una sonrisa para regalarme o un hombro en que apoyarme cuando lo necesité, por estar junto a mí y alegrar mis días. A mi hermano, Igor, por todos los buenos momentos vividos y su compañerismo. A mis abuelos Raquel y Pepe, y tía Lorena por su apoyo en momentos necesarios.

Me siento agradecido de haber tenido una buena educación en mi hogar, colegio y universidad, de estudiar en lugares donde pude encontrar excelentes docentes y tener una visión integral y pluralista permitiéndome conocer variadas realidades. Agradezco en este sentido a los profesores Julián Ortiz y Xavier Emery por su buena disposición y orientación en el desarrollo de este trabajo, sin duda son referentes en materias técnicas y académicas. Al personal de secretaría docente del departamento de minería y en particular a Juanita por su eterna paciencia. A mis amigos que me han acompañado a lo largo de este camino universitario (Iván, Emilio, Camila, Valeria). A los mineros con los que compartí no sólo largas jornadas de trabajo y esfuerzo sino que también tristezas y alegrías. En particular a Sebastián Ávalos, Yerko Yávar y María Francisca Rubilar, compañeros incansables de trasnoches. A los amigos del colegio con los que se mantiene la amistad (Ángel, David y Sebastián). Y en general a todos quienes me han apoyado a lo largo de este camino.

Finalmente agradezco al laboratorio ALGES (Advanced Laboratory for Geostatistical Supercomputing), al Centro Avanzado de Tecnología para la Minería (AMTC) y a BHP Billiton por permitir el desarrollo del presente trabajo enmarcado en el proyecto CAM-G. A la Universidad de Chile por la Beca Universidad de Chile (BUCH), a la fundación Mellado por su beca empresa y a AMSA por la beca otorgada en mis últimos años de carrera.

## TABLA DE CONTENIDO

1	INTRODUCCIÓN.....	1
1.1	Motivación del tema .....	2
1.2	Objetivos .....	3
1.2.1	Objetivo general .....	3
1.2.2	Objetivos específicos.....	3
1.3	Alcances.....	4
2	ESTADO DEL ARTE .....	5
2.1	Estacionaridad.....	6
2.1.1	Momento de primer orden y segundo orden.....	7
2.1.2	Estacionaridad estricta.....	8
2.1.3	Estacionaridad de segundo orden .....	8
2.1.4	Hipótesis intrínseca .....	9
2.1.5	Cuasi-Estacionaridad .....	9
2.2	Estimación caso estacionario .....	9
2.2.1	Kriging simple .....	10
2.2.2	Kriging ordinario .....	11
2.3	Estimación caso no estacionario .....	12
2.3.1	Componentes de la deriva.....	13
2.3.2	Kriging en presencia de deriva .....	15
2.3.3	Kriging con deriva externa .....	17
2.3.4	Kriging de residuos.....	19
2.4	Simulación .....	19

2.4.1	Enfoque de simulación secuencial.....	21
2.4.2	Simulación secuencial Gaussiana.....	21
2.5	Librería GSLib.....	22
2.6	Comentarios sobre el estado del arte .....	23
3	METODOLOGÍA.....	26
3.1	Metodología de simulación del algoritmo de rechazo .....	27
4	MODIFICACIÓN CODIGO SGSIM.....	31
4.1	Cálculo de media local.....	32
4.2	Tolerancia dinámica.....	34
4.3	Mecanismo de rechazo.....	36
4.3.1	Rechazo determinista.....	36
4.3.2	Rechazo probabilístico .....	36
4.4	Resimulación .....	38
4.5	Metodología detallada de simulación del algoritmo de rechazo.....	39
5	CASO DE ESTUDIO I: EJEMPLO SINTETICO EN 1D .....	43
5.1	Variación en el mecanismo de rechazo.....	48
5.2	Variación de tolerancia dinámica .....	51
5.3	Variación de vecindad de búsqueda de media local .....	54
5.4	Variación de número aceptable de rechazos .....	57
5.5	Sensibilización con respecto al variograma modelado .....	61
5.6	Comentario sobre caso de estudio I .....	65
6	CASO DE ESTUDIO II: BASE DE DATOS 3D .....	66
6.1	Descripción del yacimiento .....	66
6.2	Descripción de la base de datos .....	69
6.3	Estudio exploratorio de los datos.....	73

6.3.1	Estadísticas básicas.....	73
6.3.2	Vistas isométricas.....	75
6.3.3	Nubes direccionales.....	77
6.4	Estudio exploratorio de los datos desagrupados.....	79
6.5	Modelo de bloques y de roca.....	81
6.6	Modelo de deriva de la media de CuT.....	82
6.7	Cálculo de residuos.....	89
6.8	Transformación Gaussiana.....	90
6.8.1	Estadísticas.....	90
6.8.2	Validación hipótesis multi Gaussiana.....	90
6.9	Análisis variográfico.....	91
6.9.1	Variogramas modelados compósitos originales.....	91
6.9.2	Variogramas modelados residuos.....	94
6.10	Selección de sondajes para simulación.....	96
6.11	Simulación.....	97
6.11.1	Plan de Kriging.....	98
6.12	Estadísticas y análisis de las simulaciones.....	101
6.12.1	Histogramas.....	102
6.12.2	Variogramas.....	104
6.12.3	Media local.....	106
6.12.4	Coefficiente de determinación ( $R^2$ ).....	109
6.12.5	Pendiente de la regresión.....	114
6.12.6	Error medio.....	118
6.12.7	Accuracy plot.....	122
6.12.8	Tiempos de simulación.....	129



6.12.9	Nodos rechazados .....	130
6.12.10	Nodos no simulados .....	132
6.13	Desempeño global Kriging con rechazo según estadísticas de validación .....	133
7	APLICACION GENERAL DEL ALGORITMO .....	137
8	ANALISIS Y CONCLUSIONES.....	139
9	BIBLIOGRAFÍA .....	148
	ANEXO A: Archivo de parámetros SGSIM .....	150
	ANEXO B: Descripción de variables creadas en implementación del algoritmo .....	153
	ANEXO C: Caso de estudio I – Sensibilización con respecto a modelo variográfico .....	156
	ANEXO D: Caso de estudio II – Base maps y análisis de contacto.....	175
	D.1 Base maps .....	175
	D.2 Análisis de contacto.....	178
	ANEXO E: Caso de estudio II – Histogramas .....	182
	ANEXO F: Caso de estudio II – Gráficos cuantil vs cuantil.....	184
	ANEXO G: Caso de estudio II – Desagrupamiento y estadísticas desagrupadas .....	187
	ANEXO H: Caso de estudio II – Validación hipótesis multigaussiana .....	193
	ANEXO I: Caso de estudio II – Mapas variográficos .....	199
	I.1 Compósitos originales.....	199
	I.2 Residuos .....	202
	ANEXO J: Caso de estudio II – Variogramas experimentales.....	204
	J.1 Compósitos originales.....	205
	J.1 Residuos .....	208
	ANEXO K: Caso de estudio II – Histograma por UE.....	211
	ANEXO L: Caso de estudio II – Variograma por UE.....	219
	ANEXO M: Caso de estudio II – Media local por UE.....	241

ANEXO N: Caso de estudio II – Coeficiente de determinación ( $R^2$ ) por UE.....	263
ANEXO O: Caso de estudio II – Pendiente de la regresión por UE .....	268
ANEXO P: Caso de estudio II – Error medio por UE.....	273
ANEXO Q: Caso de estudio II – Accuracy plot por UE.....	278
Q.1 Accuracy plot.....	278
Q.2 Pendiente de la regresión – Accuracy plot .....	285
ANEXO R: Caso de estudio II – Tiempo de simulación por UE.....	290
ANEXO S: Caso de estudio II – Nodos rechazados por UE.....	295
ANEXO T: Caso de estudio II – Nodos no simulados por UE .....	300

## ÍNDICE DE TABLAS

Tabla 1: Base – Estudio de caso 1. ....	43
Tabla 2: Parámetros generales de la simulación.....	44
Tabla 3: Grilla de simulación. ....	45
Tabla 4: Vecindad de búsqueda.....	45
Tabla 5: Variograma modelados – Caso de estudio 1. ....	46
Tabla 6: Parámetros adicionales – Variación mecanismo de rechazo.....	48
Tabla 7: Parámetros adicionales – Variación de tolerancia dinámica.....	52
Tabla 8: Parámetros adicionales – Variación de vecindad de búsqueda de media local.....	54
Tabla 9: Situación final de nodos por caso. Casos C9 al C12 considerando C9.a. ....	57
Tabla 10: Parámetros adicionales – Variación de número aceptable de rechazos. ....	57
Tabla 11: Situación final de nodos por caso. Casos C13 a C17. ....	59
Tabla 12: Grilla tipo de simulación.....	60
Tabla 13: Tiempos estimados de cómputo en grilla tipo de simulación. ....	61

Tabla 14: Coeficiente de determinación por modelo variográfico – Caso de estudio I. ....	64
Tabla 15: Volumen por tipo de mineral. ....	69
Tabla 16: Datos de ley de CuT por tipo de mineral.....	70
Tabla 17: Volumen por tipo de mineral – Recodificación. ....	71
Tabla 18: Datos de ley de CuT por tipo de mineral – Recodificación. ....	71
Tabla 19: Media muestras y media desagrupada.....	79
Tabla 20: Modelo de bloques 10x10x10 m <sup>3</sup> .....	81
Tabla 21: Número de bloques por UE.....	81
Tabla 22: Parámetros estimación de deriva por UE. ....	83
Tabla 23: Diferencia media desagrupada vs media deriva. ....	84
Tabla 24: Modelos variográficos.....	92
Tabla 25: Rotación modelos variográficos.....	92
Tabla 26: Modelos variográficos residuos.....	94
Tabla 27: Rotación modelos variográficos residuos. ....	94
Tabla 28: Número de datos por conjunto C30/C70.....	96
Tabla 29: Media y varianza por conjunto C30/C70 – Compósitos originales.....	97
Tabla 30: Media y varianza por conjunto C30/C70 – Residuos.....	97
Tabla 31: Radio de búsqueda por UE.....	98
Tabla 32: Tamaño matriz de covarianza por UE.....	99
Tabla 33: Cola superior por UE – SK/RDR. ....	99
Tabla 34: Cola inferior y superior por UE – BT.....	100
Tabla 35: Parámetros Kriging con rechazo. ....	100
Tabla 36: Desempeño RDR vs SK y BT – Coeficiente de determinación (R <sup>2</sup> ). ....	113
Tabla 37: Desempeño RDR vs SK y BT – Pendiente. ....	117
Tabla 38: Desempeño RDR vs SK y BT – Error medio.....	121

Tabla 39: Desempeño RDR vs SK y BT – Pendiente accuracy plot.....	128
Tabla 40: Desempeño RDR vs SK y BT – Validación simulación como estimación.....	134
Tabla 41: Desempeño promedio RDR vs SK y BT – Validación simulación como estimación.....	134
Tabla 42: Desempeño RDR vs SK y BT – Global.....	135
Tabla 43: Desempeño promedio RDR vs SK y BT – Global.....	135
Tabla 44: N° de pasos en Z por UE.....	199
Tabla 45: Parámetros variogramas experimentales.....	204
Tabla 46: Número y distancia del paso – Compósitos originales.....	205
Tabla 47: Número y distancia del paso - Residuos.....	205
Tabla 48: Coeficiente de determinación ( $R^2$ ) por UE.....	263
Tabla 49: Diferencia absoluta $R^2$ vs 100 [%] por UE.....	265
Tabla 50: Pendiente de la regresión por UE.....	268
Tabla 51: Diferencia absoluta pendiente vs 1.0 por UE.....	270
Tabla 52: Error medio por UE.....	273
Tabla 53: Diferencia absoluta error medio vs 0.0 por UE.....	275
Tabla 54: Pendiente de la regresión por UE – Accuracy plot.....	285
Tabla 55: Diferencia absoluta pendiente vs 1.0 por UE – Accuracy plot.....	287
Tabla 56: Tiempo de simulación – UE1.....	290
Tabla 57: Tiempo de simulación – UE2.....	290
Tabla 58: Tiempo de simulación – UE3.....	291
Tabla 59: Tiempo de simulación – UE4.....	291
Tabla 60: Tiempo de simulación – UE5.....	292
Tabla 61: Tiempo de simulación – UE6.....	292
Tabla 62: Nodos rechazados – UE 1.....	295
Tabla 63: Nodos rechazados – UE 2.....	295

Tabla 64: Nodos rechazados – UE 3. ....	296
Tabla 65: Nodos rechazados – UE 4. ....	296
Tabla 66: Nodos rechazados – UE5. ....	297
Tabla 67: Nodos rechazados – UE6. ....	297
Tabla 68: Nodos no simulados – UE 1. ....	300
Tabla 69: Nodos no simulados – UE 2. ....	300
Tabla 70: Nodos no simulados – UE 3. ....	301
Tabla 71: Nodos no simulados – UE 4. ....	301
Tabla 72: Nodos no simulados – UE5. ....	302
Tabla 73: Nodos no simulados – UE6. ....	302

## ÍNDICE DE FIGURAS

Figura 1: Esquema de procesos CAM-G. ....	2
Figura 2: Metodología de simulación – Paso 1. ....	28
Figura 3: Metodología de simulación – Paso 2. ....	28
Figura 4: Metodología de simulación – Paso 3.a-b. ....	29
Figura 5: Metodología de simulación – Paso 3.c. ....	29
Figura 6: Metodología de simulación – Paso 3.d. ....	30
Figura 7: Metodología de simulación – Paso 3.e (Rechazo determinista). ....	30
Figura 8: Parámetros de control en implementación de simulación con rechazo. ....	32
Figura 9: Ubicación CMCC. ....	66
Figura 10: Mapa geológico CMCC. ....	67
Figura 11: Vista isométrica CMCC (Agosto 1995) - Zonas principales. ....	68
Figura 12: Isométricos tipo de mineral. Dimensiones: 600x700x600 (X,Y,Z).....	72

Figura 13: Vista isométrica – Ley de CuT. ....	75
Figura 14: Vista isométrica – Tipo de mineral. ....	76
Figura 15: Proyección XY (Planta). ....	175
Figura 16: Proyección XZ (Perfil EW).....	176
Figura 17: Proyección YZ (Perfil NS).....	177
Figura 18: Mapa variográfico – UE1 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.....	199
Figura 19: Mapa variográfico – UE2 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.....	200
Figura 20: Mapa variográfico – UE3 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.....	200
Figura 21: Mapa variográfico – UE4 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.....	200
Figura 22: Mapa variográfico – UE5 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.....	201
Figura 23: Mapa variográfico – UE6 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.....	201
Figura 24: Mapa variográfico – UE1 (Izq: XY - Cen: XZ - Der: YZ) – Residuos. ....	202
Figura 25: Mapa variográfico – UE2 (Izq: XY - Cen: XZ - Der: YZ) – Residuos. ....	202
Figura 26: Mapa variográfico – UE3 (Izq: XY - Cen: XZ - Der: YZ) – Residuos. ....	202
Figura 27: Mapa variográfico – UE4 (Izq: XY - Cen: XZ - Der: YZ) – Residuos. ....	203
Figura 28: Mapa variográfico – UE5 (Izq: XY - Cen: XZ - Der: YZ) – Residuos. ....	203
Figura 29: Mapa variográfico – UE6 (Izq: XY - Cen: XZ - Der: YZ) – Residuos. ....	203

## ÍNDICE DE GRAFICOS

Gráfico 1: Tolerancia dinámica – Tipo de decaimiento. ....	35
Gráfico 2: Función de probabilidad de aceptación - $p(\epsilon) = e - C \cdot  \epsilon $ .....	37
Gráfico 3: Izquierda: Tolerancia 0.05 – Derecha: Tolerancia 0.95. (referror 1.0 en ambos). ....	38
Gráfico 4: Base – Estudio de caso 1. ....	44
Gráfico 5: Variograma omnidireccional – Datos Gaussianos. ....	45

Gráfico 6: Variograma modelados – Caso de estudio 1. ....	46
Gráfico 7: Simulación caso de estudio I – Kriging simple.....	47
Gráfico 8: Variación de mecanismo de rechazo – Probabilístico con error de referencia 0.01. ...	48
Gráfico 9: Variación de mecanismo de rechazo – Probabilístico con error de referencia 0.1. ....	49
Gráfico 10: Variación de mecanismo de rechazo – Probabilístico con error de referencia 1. ....	49
Gráfico 11: Variación de mecanismo de rechazo – Probabilístico con error de referencia 10. ....	49
Gráfico 12: Variación mecanismo de rechazo – Determinista. ....	50
Gráfico 13: Variación de tolerancia dinámica – Inicial: 0.999 – Final: 0.950. ....	52
Gráfico 14: Variación de tolerancia dinámica – Inicial: 0.005 – Final: 0.001. ....	52
Gráfico 15: Variación de tolerancia dinámica – Inicial: 0.999 – Final: 0.001. ....	53
Gráfico 16: Variación de vecindad de media local – Fracción de radio: 0.01. ....	54
Gráfico 17: Variación de vecindad de media local – Fracción de radio: 0.1. ....	55
Gráfico 18: Variación de vecindad de media local – Fracción de radio: 1. ....	55
Gráfico 19: Variación de vecindad de media local – Fracción de radio: 10. ....	55
Gráfico 20: C9 auxiliar – Tolerancia inicial: 0.005 – Tolerancia final: 0.001 ....	56
Gráfico 21: Variación de número aceptable de rechazos – Rechazos: 10.....	58
Gráfico 22: Variación de número aceptable de rechazos – Rechazos: 100.....	58
Gráfico 23: Variación de número aceptable de rechazos – Rechazos: 1,000.....	58
Gráfico 24: Variación de número aceptable de rechazos – Rechazos: 10,000.....	59
Gráfico 25: Variación de número aceptable de rechazos – Rechazos: 100,000.....	59
Gráfico 26: Porcentaje simulado de la grilla. Casos C13 a C17. ....	60
Gráfico 27: Variación modelo variográfico – Simulaciones – Kriging simple.....	62
Gráfico 28: Variación modelo variográfico – Simulaciones – Rechazos: 1,000.....	62
Gráfico 29: Variación modelo variográfico – Variograma simulaciones – Kriging simple. ....	63
Gráfico 30: Variación modelo variográfico – Variograma simulaciones – Rechazos 1,000. ....	63

Gráfico 31: Volumen por tipo de mineral – Recodificación. ....	71
Gráfico 32: Blox plot de CuT por unidad de estimación – Escala aritmética. ....	74
Gráfico 33: Blox plot de CuT por unidad de estimación – Escala logaritmica. ....	74
Gráfico 34: Nubes direccionales – UE 1. (Izq: X - Cen: Y - Der: Z).....	77
Gráfico 35: Nubes direccionales – UE 2. (Izq: X - Cen: Y - Der: Z).....	77
Gráfico 36: Nubes direccionales – UE 3. (Izq: X - Cen: Y - Der: Z).....	78
Gráfico 37: Nubes direccionales – UE 4. (Izq: X - Cen: Y - Der: Z).....	78
Gráfico 38: Nubes direccionales – UE 5. (Izq: X - Cen: Y - Der: Z).....	78
Gráfico 39: Nubes direccionales – UE 6. (Izq: X - Cen: Y - Der: Z).....	79
Gráfico 40: Blox plot de CuT desagrupado por unidad de estimación – Escala aritmética. ....	80
Gráfico 41: Blox plot de CuT desagrupado por unidad de estimación – Escala logaritmica. ....	80
Gráfico 42: Volumen por tipo de mineral – Modelo de roca. ....	82
Gráfico 43: Blox plot modelo de deriva – Escala aritmética.....	83
Gráfico 44: Blox plot modelo de deriva – Escala logarítmica. ....	84
Gráfico 45: Media local – Este (X) en cada UE. (Azul: Compósitos – Rojo: Deriva).....	86
Gráfico 46: Media local – Norte (Y) en cada UE. (Azul: Compósitos – Rojo: Deriva). ....	87
Gráfico 47: Media local – Elevación (Z) en cada UE. (Azul: Compósitos – Rojo: Deriva). ....	88
Gráfico 48: Blox plot residuos – Escala aritmética. ....	89
Gráfico 49: Variogramas modelados.....	93
Gráfico 50: Variogramas modelados residuos. ....	95
Gráfico 51: Histograma – Kriging simple UE4.....	102
Gráfico 52: Histograma – Kriging simple UE5.....	102
Gráfico 53: Histograma – Kriging de residuos UE4. ....	102
Gráfico 54: Histograma – Kriging de residuos UE5. ....	103
Gráfico 55: Histograma – Kriging con rechazo (caso base) UE4. ....	103



Gráfico 56: Histograma – Kriging con rechazo (caso base) UE5. ....	103
Gráfico 57: Variograma X-Y-Z – Kriging simple UE4. ....	104
Gráfico 58: Variograma X-Y-Z – Kriging simple UE5. ....	105
Gráfico 59: Variograma X-Y-Z – Kriging de residuos UE4. ....	105
Gráfico 60: Variograma X-Y-Z – Kriging de residuos UE5. ....	105
Gráfico 61: Variograma X-Y-Z – Kriging con rechazo (caso base) UE4. ....	105
Gráfico 62: Variograma X-Y-Z – Kriging con rechazo (caso base) UE5. ....	106
Gráfico 63: Media local X-Y-Z – Kriging simple UE4. ....	107
Gráfico 64: Media local X-Y-Z – Kriging simple UE5. ....	107
Gráfico 65: Media local X-Y-Z – Kriging de residuos UE4. ....	107
Gráfico 66: Media local X-Y-Z – Kriging de residuos UE5. ....	108
Gráfico 67: Media local X-Y-Z – Kriging con rechazo (caso base) UE4. ....	108
Gráfico 68: Media local X-Y-Z – Kriging con rechazo (caso base) UE5. ....	108
Gráfico 69: Coeficiente de determinación ( $R^2$ ) – Media UE4. ....	110
Gráfico 70: Coeficiente de determinación ( $R^2$ ) – Media UE5. ....	111
Gráfico 71: Diferencia absoluta $R^2$ vs 100 – Media UE4. ....	112
Gráfico 72: Diferencia absoluta $R^2$ vs 100 – Media UE5. ....	113
Gráfico 73: Desempeño RDR vs SK y BT – Coeficiente de determinación ( $R^2$ ). ....	114
Gráfico 74: Pendiente de la regresión – Media UE4. ....	115
Gráfico 75: Pendiente de la regresión – Media UE5. ....	115
Gráfico 76: Diferencia absoluta pendiente vs 1 – Media UE4. ....	116
Gráfico 77: Diferencia absoluta pendiente vs 1 – Media UE5. ....	116
Gráfico 78: Desempeño RDR vs SK y BT – Pendiente. ....	118
Gráfico 79: Error medio – Media UE4. ....	119
Gráfico 80: Error medio – Media UE5. ....	119

Gráfico 81: Diferencia absoluta error medio vs 0.0 – Media UE4.....	120
Gráfico 82: Diferencia absoluta error medio vs 0.0 – Media UE5.....	120
Gráfico 83: Desempeño RDR vs SK y BT – Error medio.....	121
Gráfico 84: Accuracy plot – Kriging simple – UE4.....	122
Gráfico 85: Accuracy plot – Kriging simple – UE5.....	123
Gráfico 86: Accuracy plot – Kriging de residuos – UE4. ....	123
Gráfico 87: Accuracy plot – Kriging de residuos – UE5. ....	123
Gráfico 88: Accuracy plot – Kriging con rechazo (caso base) UE4. ....	124
Gráfico 89: Accuracy plot – Kriging con rechazo (caso base) UE5. ....	124
Gráfico 90: Pendiente de la regresión – Accuracy plot UE4.....	126
Gráfico 91: Pendiente de la regresión – Accuracy plot UE5.....	126
Gráfico 92: Diferencia absoluta pendiente vs 1 – Accuracy plot UE4.....	127
Gráfico 93: Diferencia absoluta pendiente vs 1 – Accuracy plot UE5.....	127
Gráfico 94: Desempeño RDR vs SK y BT – Pendiente accuracy plot.....	128
Gráfico 95: Tiempo de simulación medio – UE4.....	129
Gráfico 96: Tiempo de simulación medio – UE5.....	130
Gráfico 97: Nodos rechazados medio – UE4. ....	131
Gráfico 98: Nodos rechazados medio – UE5. ....	131
Gráfico 99: Nodos no simulados medio – UE4.....	132
Gráfico 100: Nodos no simulados medio – UE5.....	133
Gráfico 101: Desempeño RDR vs SK – Global. ....	135
Gráfico 102: Desempeño RDR vs BT – Global. ....	136
Gráfico 103: Variación modelo variográfico – Kriging simple. ....	156
Gráfico 104: Variación modelo variográfico – Probabilístico con error de referencia: 0.01.....	157
Gráfico 105: Variación modelo variográfico – Probabilístico con error de referencia: 0.1.....	158

Gráfico 106: Variación modelo variográfico – Probabilístico con error de referencia: 1.....	159
Gráfico 107: Variación modelo variográfico – Probabilístico con error de referencia: 10.....	160
Gráfico 108: Variación modelo variográfico – Determinista.....	161
Gráfico 109: Variación modelo variográfico – Tolerancia alta.....	162
Gráfico 110: Variación modelo variográfico – Tolerancia baja.....	163
Gráfico 111: Variación modelo variográfico – Tolerancia de amplio rango. ....	164
Gráfico 112: Variación modelo variográfico – Fracción de radio: 0.01. ....	165
Gráfico 113: Variación modelo variográfico – Fracción de radio: 0.01 – Tolerancia baja.....	166
Gráfico 114: Variación modelo variográfico – Fracción de radio: 0.1. ....	167
Gráfico 115: Variación modelo variográfico – Fracción de radio: 1. ....	168
Gráfico 116: Variación modelo variográfico – Fracción de radio: 10. ....	169
Gráfico 117: Variación modelo variográfico – Rechazos: 10. ....	170
Gráfico 118: Variación modelo variográfico – Rechazos: 100. ....	171
Gráfico 119: Variación modelo variográfico – Rechazos: 1,000. ....	172
Gráfico 120: Variación modelo variográfico – Rechazos: 10,000. ....	173
Gráfico 121: Variación modelo variográfico – Rechazos: 100,000. ....	174
Gráfico 122: Correlogramas – UE1.....	178
Gráfico 123: Correlogramas – UE2.....	179
Gráfico 124: Correlogramas – UE3.....	179
Gráfico 125: Correlogramas – UE4.....	180
Gráfico 126: Correlogramas – UE5.....	180
Gráfico 127: Correlogramas – UE6.....	181
Gráfico 128: Histogramas.....	183
Gráfico 129: QQ-Plot – UE 1. (Izq: Gaussiano - Der: Lognormal). ....	184
Gráfico 130: QQ-Plot – UE 2. (Izq: Gaussiano - Der: Lognormal). ....	184

Gráfico 131: QQ-Plot – UE 3. (Izq: Gaussiano - Der: Lognormal). .....	185
Gráfico 132: QQ-Plot – UE 4. (Izq: Gaussiano - Der: Lognormal). .....	185
Gráfico 133: QQ-Plot – UE 5. (Izq: Gaussiano - Der: Lognormal). .....	185
Gráfico 134: QQ-Plot – UE 6. (Izq: Gaussiano - Der: Lognormal). .....	186
Gráfico 135: Media desagrupada por tamaño de celda. ....	188
Gráfico 136: Histogramas desagrupados. ....	189
Gráfico 137: Gráficos de probabilidad desagrupado – Lognormal. ....	190
Gráfico 138: Gráficos de probabilidad desagrupado de residuos – Normal.....	191
Gráfico 139: Gráficos de probabilidad desagrupado de residuos – Lognormal. ....	192
Gráfico 140: Box plot de CuT valores normalizados por unidad de estimación.....	193
Gráfico 141: Box plot de CuT valores normalizados de residuos por unidad de estimación.....	194
Gráfico 142: Histogramas valores Gaussianos – Compósitos originales. ....	195
Gráfico 143: Histogramas valores Gaussianos - Residuos. ....	196
Gráfico 144: Nube de correlación diferida–UE1 (Izq:25-50m - Cen:50-75m - Der:100-125m). 197	
Gráfico 145: Nube de correlación diferida–UE2 (Izq:25-50m - Cen:50-75m - Der:100-125m). 197	
Gráfico 146: Nube de correlación diferida–UE3 (Izq:25-50m - Cen:50-75m - Der:100-125m). 197	
Gráfico 147:Nube de correlación diferida–UE4 (Izq:25-50m - Cen:50-75m - Der:100-125m). 198	
Gráfico 148:Nube de correlación diferida–UE5 (Izq:25-50m - Cen:50-75m - Der:100-125m). 198	
Gráfico 149:Nube de correlación diferida–UE6 (Izq:25-50m - Cen:50-75m - Der:100-125m). 198	
Gráfico 150: Variogramas experimentales – UE1 – Compósitos. ....	205
Gráfico 151: Variogramas experimentales – UE2 – Compósitos. ....	205
Gráfico 152: Variogramas experimentales – UE3 – Compósitos. ....	206
Gráfico 153: Variogramas experimentales – UE4 – Compósitos. ....	206
Gráfico 154: Variogramas experimentales – UE5 – Compósitos. ....	206
Gráfico 155: Variogramas experimentales – UE6 – Compósitos. ....	206

Gráfico 156: Variogramas experimentales – UE1 – Residuos.....	208
Gráfico 157: Variogramas experimentales – UE2 – Residuos.....	208
Gráfico 158: Variogramas experimentales – UE3 – Residuos.....	208
Gráfico 159: Variogramas experimentales – UE4 – Residuos.....	209
Gráfico 160: Variogramas experimentales – UE5 – Residuos.....	209
Gráfico 161: Variogramas experimentales – UE6 – Residuos.....	209
Gráfico 162: Histograma acumulado – UE 1.....	212
Gráfico 163: Histograma acumulado – UE 2.....	213
Gráfico 164: Histograma acumulado – UE 3.....	214
Gráfico 165: Histograma acumulado – UE 4.....	215
Gráfico 166: Histograma acumulado – UE 5.....	216
Gráfico 167: Histograma acumulado – UE 6.....	217
Gráfico 168: Histograma acumulado por UE – Kriging de residuos.....	218
Gráfico 169: Variograma experimental – Dirección X – UE 1.....	220
Gráfico 170: Variograma experimental – Dirección X – UE 2.....	221
Gráfico 171: Variograma experimental – Dirección X – UE 3.....	222
Gráfico 172: Variograma experimental – Dirección X – UE 4.....	223
Gráfico 173: Variograma experimental – Dirección X – UE 5.....	224
Gráfico 174: Variograma experimental – Dirección X – UE 6.....	225
Gráfico 175: Variograma experimental por UE – Dirección X – Kriging de residuos.....	226
Gráfico 176: Variograma experimental – Dirección Y – UE 1.....	227
Gráfico 177: Variograma experimental – Dirección Y – UE 2.....	228
Gráfico 178: Variograma experimental – Dirección Y – UE 3.....	229
Gráfico 179: Variograma experimental – Dirección Y – UE 4.....	230
Gráfico 180: Variograma experimental – Dirección Y – UE 5.....	231

Gráfico 181: Variograma experimental – Dirección Y – UE 6.....	232
Gráfico 182: Variograma experimental por UE – Dirección Y – Kriging de residuos.....	233
Gráfico 183: Variograma experimental – Dirección Z – UE 1. ....	234
Gráfico 184: Variograma experimental – Dirección Z – UE 2. ....	235
Gráfico 185: Variograma experimental – Dirección Z – UE 3. ....	236
Gráfico 186: Variograma experimental – Dirección Z – UE 4. ....	237
Gráfico 187: Variograma experimental – Dirección Z – UE 5. ....	238
Gráfico 188: Variograma experimental – Dirección Z – UE 6. ....	239
Gráfico 189: Variograma experimental por UE – Dirección Z – Kriging de residuos. ....	240
Gráfico 190: Media local – Dirección X – UE 1.....	242
Gráfico 191: Media local – Dirección X – UE 2.....	243
Gráfico 192: Media local – Dirección X – UE 3.....	244
Gráfico 193: Media local – Dirección X – UE 4.....	245
Gráfico 194: Media local – Dirección X – UE 5.....	246
Gráfico 195: Media local – Dirección X – UE 6.....	247
Gráfico 196: Media local por UE – Dirección X – Kriging de residuos.....	248
Gráfico 197: Media local – Dirección Y – UE 1.....	249
Gráfico 198: Media local – Dirección Y – UE 2.....	250
Gráfico 199: Media local – Dirección Y – UE 3.....	251
Gráfico 200: Media local – Dirección Y – UE 4.....	252
Gráfico 201: Media local – Dirección Y – UE 5.....	253
Gráfico 202: Media local – Dirección Y – UE 6.....	254
Gráfico 203: Media local por UE – Dirección Y – Kriging de residuos.....	255
Gráfico 204: Media local – Dirección Z – UE 1. ....	256
Gráfico 205: Media local – Dirección Z – UE 2. ....	257

Gráfico 206: Media local – Dirección Z – UE 3. ....	258
Gráfico 207: Media local – Dirección Z – UE 4. ....	259
Gráfico 208: Media local – Dirección Z – UE 5. ....	260
Gráfico 209: Media local – Dirección Z – UE 6. ....	261
Gráfico 210: Media local por UE – Dirección Z – Kriging de residuos. ....	262
Gráfico 211: Coeficiente de determinación ( $R^2$ ) – Media UE1.....	263
Gráfico 212: Coeficiente de determinación ( $R^2$ ) – Media UE2.....	264
Gráfico 213: Coeficiente de determinación ( $R^2$ ) – Media UE3.....	264
Gráfico 214: Coeficiente de determinación ( $R^2$ ) – Media UE6.....	265
Gráfico 215: Diferencia Absoluta $R^2$ vs 100 [%] – Media UE1.....	266
Gráfico 216: Diferencia Absoluta $R^2$ vs 100 [%] – Media UE2.....	266
Gráfico 217: Diferencia Absoluta $R^2$ vs 100 [%] – Media UE3.....	267
Gráfico 218: Diferencia Absoluta $R^2$ vs 100 [%] – Media UE6.....	267
Gráfico 219: Pendiente de la regresión – Media UE1.....	268
Gráfico 220: Pendiente de la regresión – Media UE2.....	269
Gráfico 221: Pendiente de la regresión – Media UE3.....	269
Gráfico 222: Pendiente de la regresión – Media UE6.....	270
Gráfico 223: Diferencia absoluta pendiente vs 1.0 – Media UE1.....	271
Gráfico 224: Diferencia absoluta pendiente vs 1.0 – Media UE2.....	271
Gráfico 225: Diferencia absoluta pendiente vs 1.0 – Media UE3.....	272
Gráfico 226: Diferencia absoluta pendiente vs 1.0 – Media UE6.....	272
Gráfico 227: Error medio – Media UE1.....	273
Gráfico 228: Error medio – Media UE2.....	274
Gráfico 229: Error medio – Media UE3.....	274
Gráfico 230: Error medio – Media UE6.....	275

Gráfico 231: Diferencia absoluta error medio vs 0.0 – Media UE1.....	276
Gráfico 232: Diferencia absoluta error medio vs 0.0 – Media UE2.....	276
Gráfico 233: Diferencia absoluta error medio vs 0.0 – Media UE3.....	277
Gráfico 234: Diferencia absoluta error medio vs 0.0 – Media UE6.....	277
Gráfico 235: Accuracy plot – UE1.....	278
Gráfico 236: Accuracy plot – UE2.....	279
Gráfico 237: Accuracy plot – UE3.....	280
Gráfico 238: Accuracy plot – UE4.....	281
Gráfico 239: Accuracy plot – UE5.....	282
Gráfico 240: Accuracy plot – UE6.....	283
Gráfico 241: Accuracy plot – Kriging de residuos.....	284
Gráfico 242: Pendiente de la regresión – Accuracy plot UE1.....	285
Gráfico 243: Pendiente de la regresión – Accuracy plot UE2.....	286
Gráfico 244: Pendiente de la regresión – Accuracy plot UE3.....	286
Gráfico 245: Pendiente de la regresión – Accuracy plot UE6.....	287
Gráfico 246: Diferencia absoluta pendiente vs 1 – Accuracy plot UE1.....	288
Gráfico 247: Diferencia absoluta pendiente vs 1 – Accuracy plot UE2.....	288
Gráfico 248: Diferencia absoluta pendiente vs 1 – Accuracy plot UE3.....	289
Gráfico 249: Diferencia absoluta pendiente vs 1 – Accuracy plot UE6.....	289
Gráfico 250: Tiempo de simulación medio – UE1.....	293
Gráfico 251: Tiempo de simulación medio – UE2.....	293
Gráfico 252: Tiempo de simulación medio – UE3.....	294
Gráfico 253: Tiempo de simulación medio – UE6.....	294
Gráfico 254: Nodos rechazados medio – UE 1. ....	298
Gráfico 255: Nodos rechazados medio – UE 2. ....	298



Gráfico 256: Nodos rechazados medio – UE 3. ....	299
Gráfico 257: Nodos rechazados medio – UE6. ....	299
Gráfico 258: Nodos no simulados medio – UE 1. ....	303
Gráfico 259: Nodos no simulados medio – UE 2. ....	303
Gráfico 260: Nodos no simulados medio – UE 3. ....	304
Gráfico 261: Nodos no simulados medio – UE6. ....	304

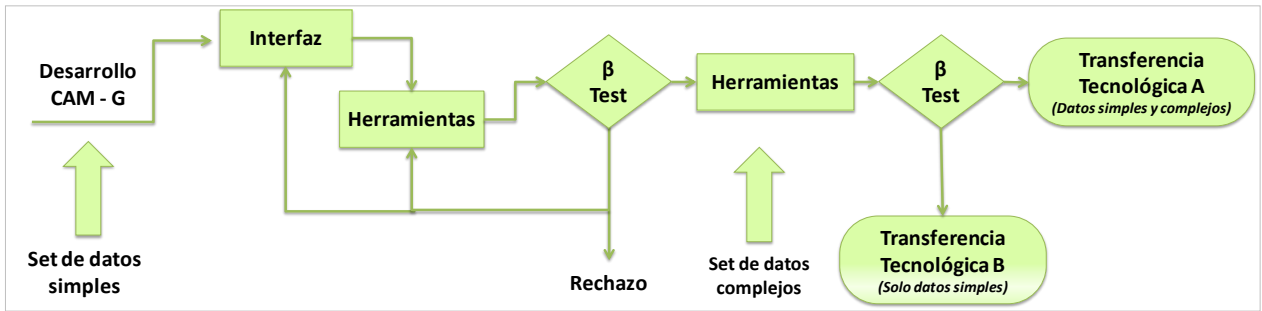
# 1 INTRODUCCIÓN

La geoestadística corresponde a un estudio estadístico de los datos integrando dentro de la información que compete al análisis, la disposición espacial de los valores a estudiar. De esta manera, toman importancia conceptos no representados en la estadística convencional como lo son el contexto geológico del espacio muestreado, la continuidad espacial de una cierta variable regionalizada y las redundancias de información debido a la disposición de las muestras.

Esta rama de la matemática está fuertemente ligada con el rubro minero y en particular con las etapas de estimación y categorización de recursos mediante el procesamiento de muestras de la especie de interés de manera geo-referenciada. Dada la importancia de esta primera etapa en la definición del negocio minero, la exactitud y precisión del análisis permite agregar valor al negocio. Más aún, la geoestadística continúa desarrollándose a lo largo de la vida de una mina en etapas de definiciones posteriores siendo un aporte en tales etapas.

El presente trabajo corresponde justamente al desarrollo de una herramienta que permite simular valores de una variable regionalizada considerando que tales valores tienen una variación sistemática en el espacio. Dicho trabajo se enmarca en el proyecto de investigación CAM-G (Computer Assisted Modeling of Geology) desarrollado por el laboratorio ALGES (Advanced Laboratory for Geostatistical Supercomputing), el cual es parte del Departamento de Minería de la Universidad de Chile (DIMIN) y del Centro Avanzado de Tecnología para la Minería (AMTC). El proyecto CAM-G tiene como metas realizar el modelamiento estocástico de información geológica dispersa de manera rápida y semi-automatizada con el fin de obtener sólidos estimados y/o simulados, los cuales sean rápidos de actualizar de una manera semi-automatizada en caso de agregarse diversas fuentes que vuelvan más denso el muestreo, como lo son los sondajes de aire reverso, los de diamantina y el mapeo geológico.

A modo de contexto, se presenta en la Figura 1 el esquema de procesos del proyecto CAM-G.



**Figura 1: Esquema de procesos CAM-G.**

### 1.1 Motivación del tema

En el campo de la estimación y categorización de recursos se debe tener en cuenta que usualmente la naturaleza de los cuerpos mineralizados y, en general, la de los recursos extraídos en minería, es bastante variada, encontrándose estos de manera masiva, en vetas, diseminados, estratificado, con direcciones preferenciales, entre otras muchas situaciones. Por otro lado, además de la naturaleza propia del estudio, hay que tener en cuenta el escenario de estimación, es decir, la cantidad de datos muestreados, la disposición de estos, posibles muestreos preferenciales (en la mayoría de los casos en la práctica), escenarios de interpolación o extrapolación, entre otros. Por todo esto, tener una herramienta adecuada para la estimación permite obtener mejores resultados en base a la representación de las estadísticas de los datos originales.

En general en este campo se trabaja con variables que se interpretan como una realización de una función aleatoria que usualmente se supone estacionaria. Sin embargo, en la realidad se tiene un set de datos que pueden presentar variaciones locales en sus estadísticas (media, varianza, continuidad espacial). El ignorar estas variaciones locales (el caso del modelo de función aleatoria estacionario) puede conducir a resultados poco satisfactorios debido al uso de una herramienta no apropiada para la estimación. Si bien, es conocido este hecho, no hay un consenso en la literatura con respecto al tema. De esta manera el rol y el significado de la estacionaridad parece ser una fuente de confusión en diversos trabajos siendo de gran relevancia en los resultados obtenidos.

Así, el presente trabajo se enfoca en la simulación secuencial Gaussiana no estacionaria de leyes, proponiendo un nuevo algoritmo que permite capturar dicha no estacionaridad de las leyes en el

modelo resultante. Tal algoritmo corresponde a una simulación Gaussiana secuencial con rechazo mediante una deriva de referencia con lo cual el modelador tiene la opción de definir una cierta tendencia a respetar por parte de los valores a simular bajo la hipótesis que esta herramienta desarrollada, en efecto, permite respetar tal deriva obteniendo resultados representativos de la base de datos en cuanto a sus estadísticos de orden 1 (histograma) y orden 2 (variograma). La idea principal que subyace en el algoritmo propuesto, sobre la cual se fundamenta la hipótesis propuesta, corresponde a un concepto teórico al considerar que el paso de rechazar valores simulados hasta obtener uno que se ajuste en base a ciertos criterios a la deriva propuesta es equivalente a quedarse con un set de simulaciones, subconjunto del universo de simulaciones, que respeta tal deriva. Dado que este concepto teórico no deriva de una formulación matemática que permita asegurar y generalizar validez, se presenta en este trabajo, además del estudio de casos en los cuales se implementa este algoritmo, una breve descripción de la aplicación general del algoritmo para un correcto uso en otros escenarios en los que se requiera su aplicación.

## **1.2 Objetivos**

### *1.2.1 Objetivo general*

El objetivo central apunta a generar una nueva herramienta de simulación de leyes capaz de imponer un modelo de derivas a través de un algoritmo de rechazo. El modelo debe ser semi automatizado y rápido de actualizar en presencia de nueva información.

### *1.2.2 Objetivos específicos*

Se desprenden del objetivo central del trabajo, los siguientes objetivos específicos:

1. Realizar un estudio crítico del estado del arte con el fin de considerar distintos enfoques utilizados en la búsqueda de una solución al problema.
2. Definir el concepto de estacionaridad y los supuestos que le dan el formalismo al estudio.
3. Realizar modificaciones al algoritmo de simulación Gaussiana secuencial con el fin de implementar el algoritmo propuesto.
4. Implementar mecanismos que permitan personalizar las condiciones de la simulación.
5. Verificar, mediante una etapa de validación, la validez de los resultados por medio de un ejemplo sintético en una dimensión, simple de presentar visualmente.

6. Realizar un análisis exploratorio de los datos, variografía y simulación mediante el algoritmo propuesto para un caso de estudio.
7. Validar resultados obtenidos mediante el caso de estudio.
8. Concluir sobre la verificación o rechazo de la hipótesis.

### **1.3 Alcances**

El alcance del presente trabajo corresponde a temas de investigación en el campo de la geoestadística.

Se puede resumir el alcance en los siguientes puntos:

1. El modelo de deriva se considera un input en el proceso de simulación por lo cual no se estudia la forma en que se construye este modelo.
2. El modelo de deriva corresponde a una deriva en la media. La deriva en la varianza no se analiza.
3. Implementación, estudio de resultados y ajustes.
4. Contraste con otros métodos de simulación de leyes con derivas.
5. Conclusiones y análisis sobre utilidad de algoritmo como nueva herramienta para abordar el problema.
6. No se trabaja en la optimización del código ni en la interfaz mediante la cual se despliega esta herramienta.
7. El rango de tolerancias en la implementación de tolerancias dinámicas para el algoritmo de rechazo sólo se sensibiliza, es decir, el rango aceptable no es optimizado.

## 2 ESTADO DEL ARTE

En geoestadística se suele usar la hipótesis de estacionaridad como supuesto necesario para inferir los estadísticos en la estimación y simulación de variables, como lo son el histograma y el variograma. Normalmente, sólo se dispone de un set de datos del cual se conocen ciertas propiedades de manera discreta en un dominio. A partir de estos datos y, en conjunto con el conocimiento geológico del dominio, se pretende inferir de ellos los parámetros de una función aleatoria que se supone estacionaria. Ahora bien, si el conjunto de los datos indica que existen tendencias locales, la hipótesis de estacionaridad pierde validez y se debe agregar al modelo este comportamiento mediante la inclusión de una deriva.

Existen diversos métodos para modelar leyes (variables continuas) considerando una deriva espacial los cuales se explican en la sección 2.3:

1. Simulación mediante Kriging en presencia de deriva (Kriging universal).
2. Simulación considerando Kriging con deriva externa.
3. Modelamiento de la deriva, cálculo de residuos y simulación de residuos.

La mayoría de estos modelos no son satisfactorios por alguna de las siguientes razones:

1. No reproducen el histograma global (se produce algún tipo de sesgo).
2. Los modelos de deriva son arbitrarios y pueden dejar los residuos sin estructura.
3. La inferencia del variograma y del modelo de deriva no es buena.

En el contexto en el que se enmarca el estudio es importante entender el concepto de derivas. Estas reflejan una tendencia sistemática de la distribución espacial de los valores de cierta variable regionalizada [5]. Las derivas tanto en la media como en la varianza de la variable regionalizada son importantes al momento de realizar el modelamiento de dicha variable. Los estimadores por Kriging han sido modificados por muchos autores para lograr hacerse cargo de la deriva en las medias. En general, los métodos se asemejan al Kriging universal en el sentido que la deriva es estimada por un set de funciones y sus coeficientes son iterativamente estimados junto con la inferencia del variograma de los residuos.

Ignorar las derivas y aplicar geoestadística convencional puede conducir a [3]:

1. Aumento de la varianza condicional debido al uso de Kriging ordinario.
2. Una definición incorrecta del correspondiente variograma modelado.
3. Compensación de la varianza condicional sobre diferentes regiones en el dominio, es decir, sobrestimación en ciertas regiones y subestimación en otras.

En ese sentido, es importante tener claro el concepto de estacionaridad y la formalidad en la cual se enmarca este concepto.

## 2.1 Estacionaridad

El concepto de estacionaridad permite desarrollar el formalismo probabilístico de función aleatoria con la información de la que se dispone con el fin de poder realizar la estimación de la variable en cuestión. En este sentido se debe recordar que la información de la que se dispone se considera como una realización de una función aleatoria y, por tanto, sólo se puede acceder a una sola de estas realizaciones. De esta manera, tal como lo hace notar Emery [5], son dos las razones que no permiten realizar geoestadística en su manera general, las cuales son, por un lado el hecho de poseer información de tan sólo una realización de la función aleatoria y, por otra parte, el de poseer esta información sólo de manera discreta en el dominio de estudio.

Siendo  $Z(\mathbf{u})$  una función aleatoria,  $k$  entero y un set cualquiera de puntos de soporte  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ , se define el vector aleatorio dado por  $\{Z(\mathbf{u}_1), \dots, Z(\mathbf{u}_k)\}$  el cual está caracterizado por su distribución conjunta definida en la ecuación 1.

$$F_{\mathbf{u}_1, \dots, \mathbf{u}_k}(z_1, \dots, z_k) = Prob\{Z(\mathbf{u}_1) < z_1, \dots, Z(\mathbf{u}_k) < z_k\} \quad (1)$$

El conjunto de todas las funciones de distribución (1) para cualquier entero  $k$  y cualquier set de puntos de soporte  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$  es lo que constituye la distribución espacial de la función aleatoria  $Z(\mathbf{u})$ . Ahora bien, en geoestadística y más precisamente en geoestadística lineal, son sólo los primeros dos momentos los utilizados [9].

### 2.1.1 Momento de primer orden y segundo orden

Sea  $Z(\mathbf{u})$  una variable aleatoria en un punto  $\mathbf{u}$  en el cual se asume que su momento de primer orden existe. Este valor corresponde a la esperanza matemática y en general, es función de la posición  $\mathbf{u}$ .

$$E\{Z(\mathbf{u})\} = m(\mathbf{u}) \quad (2)$$

En cuanto al momento de segundo orden, tres son los considerados en geoestadística [9].

1. **Varianza:** O, más precisamente, la varianza a priori de  $Z(\mathbf{u})$ . Cuando esta existe se define como el momento de segundo orden sobre la esperanza  $m(\mathbf{u})$  como se muestra en la ecuación 3.

$$Var\{Z(\mathbf{u})\} = E\{[Z(\mathbf{u}) - m(\mathbf{u})]^2\} \quad (3)$$

2. **Covarianza:** En caso que exista la varianza para las variables aleatorias  $Z(\mathbf{u}_1)$  y  $Z(\mathbf{u}_2)$  en los puntos  $\mathbf{u}_1$  y  $\mathbf{u}_2$ , entonces la covarianza entre estos puntos existe y queda definida según la ecuación 4.

$$Cov(\mathbf{u}_1, \mathbf{u}_2) = E\{[Z(\mathbf{u}_1) - m(\mathbf{u}_1)][Z(\mathbf{u}_2) - m(\mathbf{u}_2)]\} \quad (4)$$

3. **Variograma:** La función de variograma se define como la varianza de los incrementos, de tal manera que la ecuación 5 representa el semi-variograma.

$$\gamma(\mathbf{u}_1, \mathbf{u}_2) = \frac{1}{2} Var\{Z(\mathbf{u}_1) - Z(\mathbf{u}_2)\} \quad (5)$$

De las definiciones de varianza y covarianza se depende que estas dependen simultáneamente de  $\mathbf{u}_1$  y  $\mathbf{u}_2$ , por lo cual se necesitarían muchas realizaciones de  $Z(\mathbf{u}_1)$  y  $Z(\mathbf{u}_2)$  para poder realizar la inferencia estadística. Se definen los siguientes tipos de estacionaridad que permiten resolver la necesidad de repeticiones sobre la variable aleatoria por repeticiones en el espacio.



### 2.1.2 Estacionaridad estricta

Una función aleatoria se dice que es estacionaria estricta si es que su distribución espacial es invariante por traslación. En términos matemáticos, esto implica lo siguiente.

$$F_{\mathbf{u}_1, \dots, \mathbf{u}_k}(z_1, \dots, z_k) = F_{\mathbf{u}_1 + \mathbf{h}, \dots, \mathbf{u}_k + \mathbf{h}}(z_1, \dots, z_k)$$

Esta propiedad es bastante fuerte y en la práctica son sólo los dos primeros momentos los que se utilizan en la estimación.

### 2.1.3 Estacionaridad de segundo orden

Dado que es difícil testear la hipótesis de estacionaridad estricta y el hecho de que ésta no implica la existencia de momentos (media, varianza, covarianza), es usual que se utilice de manera más frecuente una forma más débil de estacionaridad.

Una función aleatoria es estacionaria de segundo orden si y sólo si cumple con:

- i.  $E\{Z(\mathbf{u})\}$  existe y se tiene que  $E\{Z(\mathbf{u})\} = m \forall \mathbf{u}$ , es decir, no depende de  $\mathbf{u}$ .
- ii.  $\forall \{Z(\mathbf{u}_1), Z(\mathbf{u}_2)\}$ , la covarianza existe y depende sólo del vector separación  $\mathbf{h} = \mathbf{u}_1 - \mathbf{u}_2$ .

$$Cov(Z(\mathbf{u}_1), Z(\mathbf{u}_2)) = Cov(\mathbf{u}_1 - \mathbf{u}_2) = Cov(\mathbf{h}) = E\{Z(\mathbf{u} + \mathbf{h})Z(\mathbf{u})\} - m^2$$

Se hace notar el hecho que la estacionaridad de la covarianza implica la estacionaridad de la varianza y del variograma, de hecho, (i) permite determinar la relación  $\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$ , la cual permite inferir el modelo de covarianza a partir del modelo de variograma y viceversa.

Bajo la estacionaridad de segundo orden se tiene entonces que la covarianza y el variograma funcionan como herramientas equivalentes para definir la correlación entre un par de variables aleatorias separadas a una distancia  $\mathbf{h}$ . Si bien esta estacionaridad de segundo orden asume la existencia de la covarianza (lo cual implica la existencia de la varianza a priori:  $C(\mathbf{0})$ ), existen muchos fenómenos físicos con capacidad infinita de dispersión donde no se puede definir ni la varianza ni la covarianza pero sí el variograma [9] [14], lo cual motiva a la definición de un tipo más débil de estacionaridad.

#### 2.1.4 Hipótesis intrínseca

La forma de estacionaridad que implica la hipótesis intrínseca es esencialmente una estacionaridad de segundo orden dado que ve las diferencias de primer orden  $Z(\mathbf{u}_1) - Z(\mathbf{u}_2)$ . De esta manera, una función aleatoria  $Z(\mathbf{u})$  se dice intrínseca si y sólo si cumple con:

- i.  $E\{Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})\} = 0 \forall \mathbf{u}$ , es decir, no depende de  $\mathbf{u}$ .
- ii.  $\forall \mathbf{h}, \{Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})\}$  tiene varianza finita y no depende de  $\mathbf{u}$ , es decir,

$$\text{Var}\{Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})\} = E\{[Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})]^2\} = 2\gamma(\mathbf{h})$$

Dado que la hipótesis (i) de la estacionaridad de segundo orden implica la hipótesis (i) de la hipótesis intrínseca, y que la existencia de la covarianza como función de  $\mathbf{h}$  implica la existencia del variograma como función de  $\mathbf{h}$ , se tiene que la estacionaridad de segundo orden implica la hipótesis intrínseca, pero la inversa no es necesariamente cierta.

#### 2.1.5 Cuasi-Estacionaridad

En la práctica interesa poder definir un modelo de covarianza o variograma dentro de una vecindad de estudio. Considerar que se cumple la hipótesis de estacionaridad de segundo orden (o hipótesis intrínseca) en un dominio acotado consiste en la cuasi-estacionaridad. El hecho es que siempre se puede definir una vecindad lo suficientemente pequeña para validar la estacionaridad aunque tal vez esto implique la no suficiencia de datos para inferir el primer y el segundo momento.

## 2.2 Estimación caso estacionario

La estimación por Kriging se realiza en el marco de la geoestadística lineal. De esta forma se define el estimador de la variable aleatoria  $Z(\mathbf{u})$  en un punto  $\mathbf{u}_o$  como  $Z_o^*$ , el cual, por construcción cumple con ser lineal, insesgado y óptimo.

- **Linealidad** :  $Z_0^* = \lambda_0 + \sum_{\alpha=1}^n \lambda_{\alpha} Z_{\alpha}$
- **Insesgo** :  $E\{Z_0 - Z_0^*\} = E\{Z_0\} - E\{Z_0^*\} = m_0 + \lambda_0 + \sum_{\alpha=1}^n \lambda_{\alpha} m_{\alpha} = 0$
- **Óptimo** :  $\{\lambda_{\alpha}\}, \alpha \in \{1, \dots, n\}$  es tal que  $Var\{Z_0 - Z_0^*\}$  es mínima.

### 2.2.1 Kriging simple

De la condición de insesgo se tiene la siguiente implicancia:

- **Insesgo** :  $E\{Z_0 - Z_0^*\} = 0 \Rightarrow \lambda_0 = m_0 - \sum_{\alpha=1}^n \lambda_{\alpha} m_{\alpha}$

Notar que se debe conocer la esperanza  $m_{\alpha}$  con  $\alpha \in \{1, \dots, n\}$  para asegurar la condición de insesgo. En la práctica, el Kriging simple asume media constante y conocida  $m_{\alpha} = m \forall \alpha \in \{1, \dots, n\}$ .

- **Óptimo** :  $\text{Min } Var\{Z_0 - Z_0^*\} = Var\{Z_0\} - 2Cov\{Z_0, Z_0^*\} + Var\{Z_0^*\}$

El desarrollo de cada uno de los términos permite obtener el siguiente resultado equivalente:

- $Var\{Z_0\} = Cov\{Z_0, Z_0\} = C_{00} = C(0)$
- $Cov\{Z_0, Z_0^*\} = Cov\{Z_0, \lambda_0 + \sum_{\alpha=1}^n \lambda_{\alpha} Z_{\alpha}\} = \sum_{\alpha=1}^n \lambda_{\alpha} Cov\{Z_0, Z_{\alpha}\}$
- $Var\{Z_0^*\} = Cov\{\lambda_0 + \sum_{\alpha=1}^n \lambda_{\alpha} Z_{\alpha}, \lambda_0 + \sum_{\alpha=1}^n \lambda_{\alpha} Z_{\alpha}\} = \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C_{\alpha\beta}$

Así, se busca minimizar.

$$\text{Min } C_{00} - 2 \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C_{\alpha\beta} \quad (6)$$

De esta manera la optimización (minimización de la varianza del error) de la estimación, correspondiente a igualar a cero las derivadas parciales con respecto a los ponderadores  $\lambda_{\alpha}$ , conduce al siguiente sistema.

$$\frac{\partial (C_{00} - 2 \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C_{\alpha\beta})}{\partial \lambda_{\alpha}} = 0 \Rightarrow C_{\alpha 0} = \sum_{\beta=1}^n \lambda_{\beta} C_{\alpha\beta} \quad \forall \alpha \in \{1, \dots, n\} \quad (7)$$

Que en su manera matricial se escribe como.

$$\begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} C_{10} \\ \vdots \\ C_{n0} \end{bmatrix}$$

Por su parte, la varianza del error de estimación por Kriging simple ( $\sigma_{KS}^2$ ) tiene la siguiente fórmula.

$$\sigma_{KS}^2 = Var\{Z_0 - Z_0^*\} = C_{00} - 2 \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} + \sum_{\alpha=1}^n \lambda_{\alpha} \left[ \sum_{\beta=1}^n \lambda_{\beta} C_{\alpha\beta} \right]_{=C_{\alpha 0}}$$

$$\sigma_{KS}^2 = C_{00} - \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} \quad (8)$$

Se hace notar que para la construcción del estimador se utiliza una media  $m$  conocida y constante para cualquier punto  $y$ , además, un modelo de covarianzas. De esta manera, basta asumir estacionaridad de segundo orden con el fin de tener el formalismo necesario que de sustento a la estimación.

### 2.2.2 Kriging ordinario

Si bien en el caso del Kriging simple se asume una media conocida  $m$  constante en el dominio de estimación, es decir,  $E\{Z(\mathbf{u}_{\alpha})\} = m \quad \forall \alpha \in \{1, \dots, n\}$ , en el caso del Kriging ordinario se evita la inferencia previa de la media y se obtiene ésta de manera implícita en el algoritmo de estimación. Así, las condiciones para su construcción plantean lo siguiente.

Asumiendo una media constante y desconocida, la única manera de asegurar la condición de insesgo consiste en imponer  $\lambda_0 = 0$ . De esta manera la condición de insesgo queda como.

- **Insesgo** :  $E\{Z_0 - Z_0^*\} = 0 \Rightarrow m(1 - \sum_{\alpha=1}^n \lambda_{\alpha}) = 0 \Rightarrow \sum_{\alpha=1}^n \lambda_{\alpha} = 1$

Es decir, que la suma de los ponderadores del Kriging sea igual a 1.

- **Óptimo** :  $\text{Min} \{C_{00} - 2 \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C_{\alpha\beta}\}$   
s. a.  $\sum_{\alpha=1}^n \lambda_{\alpha} - 1 = 0$

Es decir, se tiene el mismo problema de optimización dado por (6), sólo que en este caso está sujeto a la restricción de la sumatoria de los ponderadores de Kriging. Para resolver el problema de optimización se utiliza el método de multiplicadores de Lagrange, para lo cual se define la función  $L(\lambda_1, \dots, \lambda_n, 2\mu)$  obteniéndose el siguiente sistema.

$$L(\lambda_1, \dots, \lambda_n, 2\mu) = C_{00} - 2 \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C_{\alpha\beta} + 2\mu(\sum_{\alpha=1}^n \lambda_{\alpha} - 1)$$

$$\frac{\partial L(\lambda_1, \dots, \lambda_n, 2\mu)}{\partial \lambda_{\alpha}} = 0 \Rightarrow C_{\alpha 0} = \sum_{\beta=1}^n \lambda_{\beta} C_{\alpha\beta} + \mu \quad \forall \alpha \in \{1, \dots, n\}$$

$$\frac{\partial L(\lambda_1, \dots, \lambda_n, 2\mu)}{\partial \mu} = 0 \Rightarrow \sum_{\alpha=1}^n \lambda_{\alpha} = 1$$

Que en su manera matricial se escribe como:

$$\begin{bmatrix} C_{11} & \dots & C_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{n1} & \dots & C_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} C_{10} \\ \vdots \\ C_{n0} \\ 1 \end{bmatrix}$$

Por su parte, la varianza del error de estimación por Kriging ordinario ( $\sigma_{KO}^2$ ) considerando ahora el parámetro  $\mu$  queda de la siguiente forma.

$$\sigma_{KO}^2 = C_{00} - \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} - \mu \quad (9)$$

### 2.3 Estimación caso no estacionario

Hasta ahora se ha presentado la estimación para el caso estacionario de segundo orden. Sin embargo, es común encontrar fenómenos físicos que poseen una cierta tendencia en los datos a lo largo del dominio de estimación, lo cual se entiende como deriva y corresponde al caso no estacionario.

Por definición, la deriva corresponde a una no estacionaridad en el valor esperado de la función aleatoria ( $E\{Z(\mathbf{u})\} = m(\mathbf{u})$ ) que motiva a que sea expresada como la suma de su componente residual ( $R(\mathbf{u})$ , aleatoria) y su deriva ( $m(\mathbf{u})$ , determinística) [9].

$$Z(\mathbf{u}) = R(\mathbf{u}) + m(\mathbf{u}) \quad (10)$$

La función  $R(\mathbf{u})$  definida en (10) puede no ser estacionaria pero debe cumplir con tener esperanza cero ( $E\{R(\mathbf{u})\} = 0$ ).

Con respecto al momento de segundo orden se tiene lo siguiente.

- $Cov(Z(\mathbf{u}_1), Z(\mathbf{u}_2)) = E\{R(\mathbf{u}_1)R(\mathbf{u}_2)\} - E(R(\mathbf{u}_1))_{=0} E(R(\mathbf{u}_2))_{=0} = Cov(R(\mathbf{u}_1), R(\mathbf{u}_2))$

- $$2\gamma_Z(\mathbf{u}_1, \mathbf{u}_2) = \text{Var}\{Z(\mathbf{u}_1) - Z(\mathbf{u}_2)\} = E\{[R(\mathbf{u}_1) - R(\mathbf{u}_2)]^2\} + E^2\{R(\mathbf{u}_1) - R(\mathbf{u}_2)\}_{=0}$$

$$= 2\gamma_R(\mathbf{u}_1, \mathbf{u}_2)$$

Es decir, tanto la covarianza como el semi-variograma de la función aleatoria corresponden a la covarianza y semi-variograma de los residuos. Tal como se menciona en la literatura, la principal dificultad cuando se tiene una aparente no estacionaridad corresponde a la estimación del variograma. De hecho el variograma experimental, el cual es el estimador más comúnmente utilizado corresponde más bien a un estimador del valor esperado de las diferencias cuadráticas de primer orden más que de las varianzas [14], por lo cual el variograma experimental se encuentra sesgado. De esta manera es importante comprender los componentes de la no estacionaridad y la forma en que se presenta el sesgo en que se incurre.

### 2.3.1 Componentes de la deriva

El variograma se puede expresar como la varianza de los incrementos según (5).

$$\gamma(\mathbf{u}, \mathbf{u} + \mathbf{h}) = \frac{1}{2} \text{Var}\{Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})\}$$

Se pueden obtener dos tipos de variogramas, centrado (C) y no centrado (NC):

$$2\gamma_{NC}(\mathbf{h}) = E\{[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2\}$$

$$2\gamma_C(\mathbf{h}) = \text{Var}\{Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})\}$$

$$= E\{[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2\} - [E\{Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})\}]^2$$

$$= 2\gamma_{NC}(\mathbf{h}) - [E\{Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})\}]^2$$

En el caso de no poseer deriva se tiene que  $E\{Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})\} = 0$ , por lo que ambos variogramas, el centrado y el no centrado son iguales. El estimador del semivariograma experimental se puede reescribir como (considerando el variograma no centrado):

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h})]^2$$

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2}[\sigma_u^2 + \sigma_{u+h}^2] + \frac{1}{2}[m_u - m_{u+h}]^2 - \hat{C}(\mathbf{h})$$

Es importante recalcar que la componente de la deriva de las medias es debido a que se está utilizando un estimador del variograma NC. En caso de usarse el variograma centrado esta componente se anula. Más aún, se puede reescribir esta ecuación para dejarla en términos de las siguientes tres componentes: deriva de la varianza, deriva de la media y componente estacionaria [3].

$$\hat{\gamma}(\mathbf{h}) = \underbrace{\frac{1}{2}[\sigma_u - \sigma_{u+h}]^2}_{\text{Deriva de la varianza}} + \underbrace{\frac{1}{2}[m_u - m_{u+h}]^2}_{\text{Deriva de la media}} + \underbrace{\sigma_u \sigma_{u+h} - \hat{C}(\mathbf{h})}_{\text{Componente estacionaria}}$$

En caso de no tener deriva de la media ( $m_u \approx m_{u+h} = m$ ) y tampoco tener deriva en la varianza ( $\sigma_u^2 \approx \sigma_{u+h}^2 = \sigma^2$ ) se tiene la relación entre la covarianza y el semivariograma dada por la relación:

$$\hat{\gamma}(\mathbf{h}) = \hat{C}(\mathbf{0}) - \hat{C}(\mathbf{h})$$

En general se dice que cuando el semivariograma está por sobre la varianza de los datos (o meseta) es debido a una correlación negativa de los datos separados a una distancia  $\mathbf{h}$ . Sin embargo puede ser que la deriva en la media haga que el semivariograma suba por sobre la meseta. De esta manera es importante tener una muestra representativa del dominio dado que de no ser así la búsqueda de derivas o incluso la inferencia del variograma será poco confiable.

Incluso cuando se tiene un dominio densamente muestreado, se debería remover la deriva de medias previo a la modelación dado que el semivariograma en estos casos tiende a exagerar la continuidad espacial y aumentar la varianza condicional lo que implica un modelo más suavizado del que debería ser y con una incerteza aumentada. Se debe recordar que el semivariograma es un estadístico de dos puntos y como tal no puede capturar características complejas de la deriva como una deriva cóncava en las medias [3].

La descomposición del semivariograma en su parte estacionaria y no estacionaria es un tema ya tratado por otros autores, aunque por lo general, se centran en la deriva de la media buscando alguna dirección libre de deriva. Olea explica una metodología práctica para realizar el modelamiento variográfico [15]. En cuanto a variogramas con deriva se muestra el sesgo que se

produce al usar como estimador el variograma no centrado y se propone el uso de una dirección libre de deriva para calcular el variograma y luego asumir isotropía. La realización de un estudio de anisotropías se desarrolla mediante el ajuste de distintos modelos anisótropos los cuales se comparan y analizan en términos de su validez en función de una validación cruzada. Otros métodos más complejos incluyen un modelamiento iterativo de la deriva y del variograma y la remoción de la deriva por medio de filtros mediante el cálculo de incrementos [2].

### 2.3.2 Kriging en presencia de deriva

En el caso del Kriging ordinario se está en presencia de un estimador que permite filtrar una media desconocida y conservar el criterio de insesgo. El caso del Kriging con deriva, o también llamado por algunos autores como Kriging universal, corresponde a una extensión de este concepto aplicado en el Kriging ordinario que permite filtrar formas más complejas de deriva ( $m(\mathbf{u})$ ) que dependen de la ubicación de los puntos a estimar. Siguiendo con el mismo concepto, se asume una deriva dada por lo siguiente.

$$m(\mathbf{u}) = \sum_{l=0}^L a_l f_l(\mathbf{u}) \quad (11)$$

Donde  $f_l(\mathbf{u})$  y  $a_l$  con  $l \in \{0, \dots, L\}$  son funciones conocidas y parámetros a estimar respectivamente. Se puede notar que el Kriging ordinario es un caso particular del Kriging en presencia de deriva donde  $L = 0$ ,  $a_0 = m$  y  $f_0(\mathbf{u}) = 1$ .

El algoritmo de estimación en base a los criterios de estimación viene dado por lo siguiente.

- **Insesgo** :  $E\{Z_0 - Z_0^*\} = 0 \Rightarrow \sum_{l=0}^L a_l (f_l(\mathbf{u}_0) - \sum_{\alpha=1}^n \lambda_\alpha f_l(\mathbf{u}_\alpha)) = 0 \Rightarrow$   
 $\sum_{\alpha=1}^n \lambda_\alpha f_l(\mathbf{u}_\alpha) = f_l(\mathbf{u}_0) \quad \forall l \in \{0, \dots, L\} \quad (12)$

Notar que la condición (12) impone  $L+1$  condiciones en el sistema de ecuaciones a resolver, con lo cual se obtiene un sistema de  $n+L+1$  ecuaciones con  $n$  ponderadores  $\lambda_\alpha$  y  $L+1$  parámetros de Lagrange  $\mu_l$ . Al igual que en el Kriging ordinario se impone  $\lambda_0 = 0$  para asegurar insesgo.

- **Óptimo** : Procediendo de manera análoga al Kriging ordinario se obtiene lo siguiente.



$$L(\lambda_1, \dots, \lambda_n, 2\mu_0, \dots, 2\mu_L) = C_{00} - 2 \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C_{\alpha\beta} + 2 \sum_{l=0}^L \mu_l (f_l(u_0) - \sum_{\alpha=1}^n \lambda_{\alpha} f_l(\mathbf{u}_{\alpha}))$$

$$\frac{\partial L(\lambda_1, \dots, \lambda_n, 2\mu_0, \dots, 2\mu_L)}{\partial \lambda_{\alpha}} = 0 \Rightarrow C_{\alpha 0} = \sum_{\beta=1}^n \lambda_{\beta} C_{\alpha\beta} + \sum_{l=0}^L \mu_l f_l(u_{\alpha}) \quad \forall \alpha \in \{1, \dots, n\}$$

$$\frac{\partial L(\lambda_1, \dots, \lambda_n, 2\mu_0, \dots, 2\mu_L)}{\partial \mu_l} = 0 \Rightarrow \sum_{\alpha=1}^n \lambda_{\alpha} f_l(u_{\alpha}) = f_l(u_0) \quad \forall l \in \{0, \dots, L\}$$

Lo cual matricialmente es equivalente a lo siguiente.

$$\begin{bmatrix} C_{11} & \dots & C_{1n} & f_0(u_1) & \dots & f_L(u_1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ C_{n1} & \dots & C_{nn} & f_0(u_n) & \dots & f_L(u_n) \\ f_0(u_1) & \dots & f_0(u_n) & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_L(u_1) & \dots & f_L(u_n) & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu_1 \\ \vdots \\ \mu_L \end{bmatrix} = \begin{bmatrix} C_{10} \\ \vdots \\ C_{n0} \\ f_0(u_0) \\ \vdots \\ f_L(u_0) \end{bmatrix}$$

La varianza del error de estimación del Kriging en presencia de deriva ( $\sigma_{KT}^2$ ) se calcula de forma análoga al Kriging ordinario y se obtiene como sigue:

$$\sigma_{KT}^2 = C_{00} - \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} - \sum_{l=0}^L \mu_l f_l(u_0) \quad (13)$$

Es interesante notar que el adjetivo de Kriging universal proviene de las L+1 condiciones de insesgo, sin embargo, tal como hace notar Journel, esta universalidad no es intrínseca al sistema de estimación dado que depende de la arbitrariedad en la elección de las funciones que componen la deriva ( $f_l(\mathbf{u})$ 's) [8].

El modelo de covarianza que se utiliza corresponde al de los residuos ( $C_R(\mathbf{h})$  donde se tiene  $R(\mathbf{u}) = Z(\mathbf{u}) - m(\mathbf{u})$ ). Dado que en las ubicaciones de las muestras no se conoce su deriva ( $m_{\alpha}$ ), el valor residual también es desconocido y, por lo tanto, el modelo de covarianza también lo es al no poder ser inferido directamente. Una posible solución a esto presenta Deutsch, realizando un filtro de la deriva por medio de combinaciones lineales de los datos  $z_{\alpha}$  que realizan la función de filtro [4]. Como ejemplo se tienen las diferencias de primer orden ( $Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})$ ) que filtran cualquier tendencia de orden 0 como  $m(\mathbf{u}) = m$ . Si bien se ha desarrollado todo un formalismo correspondiente a las IRF-k (Funciones Intrínsecas de Orden k) para la inferencia directa de  $C_R(\mathbf{h})$  a partir de los datos originales, la inferencia de la covarianza generalizada para órdenes  $k > 0$  posee severos problemas dado que las diferencias de

orden  $k + 1$  no son fáciles de obtener por un set de datos que no se encuentren en una grilla regular. Es por eso que en la versión de GSLIB con la cual se realiza el estudio, no se incluye ni covarianza generalizada ni sus correspondientes funciones aleatorias intrínsecas [4].

En la práctica, lo que se suele hacer es buscar direcciones preferenciales a sub-zonas del dominio donde la deriva puede ser ignorada y por tanto se tenga que  $Z(\mathbf{u}) \cong R(\mathbf{u})$  mediante lo cual se puede inferir el modelo de covarianzas como se presenta en la sección 2.3.1.

### 2.3.3 Kriging con deriva externa

El Kriging con deriva externa considera un modelo de deriva definido en función de una variable secundaria externa. Se define de esta forma la deriva  $m(\mathbf{u}) = a_0 + a_1s(\mathbf{u})$  con lo que el valor esperado de  $Z(\mathbf{u})$  queda de la siguiente forma:

$$E\{Z(\mathbf{u})\} = m(\mathbf{u}) = a_0 + a_1s(\mathbf{u}) \quad (14)$$

Así, este método asume que esta variable secundaria refleja la deriva de la variable primaria luego de un re-escalamiento lineal dado por  $a_0$  y  $a_1$ , lo cual puede ser interpretado como un proceso de regresión a través de la variable secundaria con el fin de definir la deriva que guía la estimación de la variable primaria [4].

Como se puede apreciar, este método conduce al mismo desarrollo que el Kriging en presencia de deriva considerando  $K = 1$ ,  $f_0(\mathbf{u}) = 1$  (por convención) y  $f_1(\mathbf{u}) = s(\mathbf{u})$ .

De esta manera la condición de insesgo implica dos restricciones a cumplir sobre los ponderadores dados por la ecuación (12).

- **Insesgo** : De la ecuación (12):
  - 1)  $\sum_{\alpha=1}^n \lambda_{\alpha} = 1$
  - 2)  $\sum_{\alpha=1}^n \lambda_{\alpha} s(\mathbf{u}_{\alpha}) = s(\mathbf{u}_0)$
- **Óptimo** : Del sistema de Kriging con deriva externa se tiene:

$$L(\lambda_1, \dots, \lambda_n, 2\mu_0, 2\mu_1) = C_{00} - 2\sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha} \lambda_{\beta} C_{\alpha\beta} + 2\mu_0(\sum_{\alpha=1}^n \lambda_{\alpha} - 1) + 2\mu_1(\sum_{\alpha=1}^n \lambda_{\alpha} s(u_{\alpha}) - s(u_0))$$

$$\frac{\partial L(\lambda_1, \dots, \lambda_n, 2\mu_0, 2\mu_1)}{\partial \lambda_{\alpha}} = 0 \Rightarrow C_{\alpha 0} = \sum_{\beta=1}^n \lambda_{\beta} C_{\alpha\beta} + \mu_0 + \mu_1 s(u_{\alpha}) \quad \forall \alpha \in \{1, \dots, n\}$$

$$\frac{\partial L(\lambda_1, \dots, \lambda_n, 2\mu_0, 2\mu_1)}{\partial \mu_1} = 0 \Rightarrow \sum_{\alpha=1}^n \lambda_{\alpha} = 1$$

$$\frac{\partial L(\lambda_1, \dots, \lambda_n, 2\mu_0, 2\mu_1)}{\partial \mu_1} = 0 \Rightarrow \sum_{\alpha=1}^n \lambda_{\alpha} s(u_{\alpha}) = s(u_0)$$

Lo cual matricialmente es equivalente a lo siguiente.

$$\begin{bmatrix} C_{11} & \dots & C_{1n} & 1 & s(u_1) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ C_{n1} & \dots & C_{nn} & 1 & s(u_n) \\ 1 & \dots & 1 & 0 & 0 \\ s(u_1) & \dots & s(u_n) & 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu_0 \\ \mu_1 \end{bmatrix} = \begin{bmatrix} C_{10} \\ \vdots \\ C_{n0} \\ 1 \\ s(u_0) \end{bmatrix}$$

La varianza del error de estimación del Kriging con deriva externa ( $\sigma_{KED}^2$ ) se obtiene de la ecuación (13).

$$\sigma_{KED}^2 = C_{00} - \sum_{\alpha=1}^n \lambda_{\alpha} C_{\alpha 0} - \mu_0 - \mu_1 s(u_0)$$

La ventaja del Kriging con deriva externa radica en tener una integración rápida y directa de una variable secundaria dado que tiene tiempos de cálculos similares al Kriging de una variable y, a su vez, no necesita un modelo de varianzas/covarianzas del atributo secundario. Al igual que el Kriging en presencia de deriva, tiene la dificultad de inferir el modelo de covarianza de los residuos además de no haber manera de calibrar el efecto de la variable secundaria ya que se asume una correlación perfecta entre las dos variables. Se debe considerar también que la variable secundaria debe tener una variación suave, si no el sistema de Kriging se vuelve inestable en el sentido que se obtendrán estimaciones extremadamente bajas o altas.

Para poder aplicar este método se debe conocer el valor de la variable secundaria tanto en los puntos con el set de datos como en cada uno de los puntos de la grilla donde se pretende estimar la variable primaria. Es importante que esta variable secundaria guarde una relación física con la variable primaria que haga sentido al uso de ésta como deriva de la primera. Un ejemplo de aplicación de este método es la estimación de un modelo de profundidad utilizando como deriva externa el tiempo de viaje de ondas sísmicas.

### 2.3.4 Kriging de residuos

Dado un modelo de deriva  $m(\mathbf{u})$ , lo que usualmente realizan los algoritmos de simulación consiste en calcular los residuos en las posiciones con muestras y luego generar realizaciones de los residuos para posteriormente agregar estos al modelo de deriva. Para esto se presume que una función aleatoria posee una componente estructurada y otra componente estocástica dada por (10). El enfoque consiste por tanto en:

1. Calcular los residuos  $r(\mathbf{u}_\alpha)$  en los puntos muestreados  $\mathbf{u}_\alpha$ .
2. Generar múltiples realizaciones  $R(\mathbf{u})$  para la componente residual usando geoestadística.
3. Agregar  $R(\mathbf{u})$  al modelo de deriva  $m(\mathbf{u})$ .

Sin embargo, el hecho de estimar la deriva y posteriormente los residuos implica que los residuos estimados se asuman como los residuos verdaderos (los cuales se desconocen). Esto tiene implicancias en un sesgo en el variograma de los residuos, dado que los residuos estimados (a diferencia de los verdaderos) tienen menor varianza y menor estructura espacial.

## 2.4 Simulación

El proceso de estimación entrega tan sólo una única respuesta de valores estimados  $\{z^*(\mathbf{u})\}$  de una distribución espacial desconocida  $\{z(\mathbf{u})\}$ , lo cual no permite medir el grado de certeza de la estimación. Mediante simulaciones estocásticas<sup>1</sup> se realizan modelos equiprobables  $\{z^{(t)}(\mathbf{u})\}$  de esta distribución espacial que se desconoce. Estas simulaciones se considerarán condicionales cuando el valor simulado coincide con el valor de los datos para cada simulación y para cada punto con valor conocido.

$$z^{(t)}(\mathbf{u}_\alpha) = z(\mathbf{u}_\alpha) \quad \forall \alpha, \forall t \quad (15)$$

Entre las propiedades del Kriging se cuenta con la característica de realizar una interpolación exacta  $\{z^*(\mathbf{u}_\alpha) = z(\mathbf{u}_\alpha) \forall \alpha\}$  y, a su vez, con el suavizamiento de los valores originales, es decir,

---

<sup>1</sup> Se denomina estocásticos a aquellos procesos cuyo comportamiento son no determinísticos.

los valores estimados son menos dispersos que los reales. Como hace notar Emery, esta última propiedad tiene varias implicancias en el estudio de la dispersión de la variable en estudio [6].

- El variograma de los valores estimados no posee efecto pepa, mientras que el de los valores reales en general sí tiene.
- Se tiene una variabilidad no uniforme en los valores estimados, siendo ésta menor mientras menos datos se tengan en la localidad del punto estimado.
- Se tiene un rango menor en el caso de los valores estimados que en el caso de los valores originales, lo que conduce a una incorrecta predicción de valores extremos.

Esto muestra que la estimación por Kriging no tiene las mismas propiedades que los valores reales, lo cual motiva al desarrollo de una herramienta que pueda capturar estas propiedades a las que se hace referencia. De esta manera la diferencia principal entre el proceso de estimación y el de simulación radica en el objetivo para el cual se realiza. Por su parte, la estimación mediante Kriging busca encontrar el mejor estimador  $z^*(\mathbf{u}) \forall \mathbf{u} \in \mathcal{D}$  en el sentido que sea lo más próximo posible al valor real desconocido  $z(\mathbf{u})$ , tomando como criterios el insesgo del valor estimado y la minimización de la varianza del error de estimación. Esto no asegura, sin embargo, la reproducción de la variabilidad de  $Z(\mathbf{u})$ , y, en efecto, los criterios utilizados producen un suavizamiento en los valores estimados. Por otra parte, el resultado de una simulación  $z_s(\mathbf{u})$ , o mejor aún, de una simulación condicional  $z_{sc}(\mathbf{u})$ , tiene los mismos primeros dos momentos que los valores reales por lo que captura la dispersión. Sin embargo, el valor simulado en cada punto  $z_{sc}(\mathbf{u})$  no corresponde al mejor estimador. Así se tiene que una estimación será preferible cuando se quiere localizar y estimar recursos, mientras que se optará por una simulación cuando se quiera determinar las características de la dispersión de dichos recursos [9].

### 2.4.1 Enfoque de simulación secuencial

El enfoque de simulación secuencial permite reducir un problema multivariable de conocer la distribución acumulada condicional conjunta (ccdf) dado  $n$  valores de  $Z(\mathbf{u})$  conocidos (16), a un problema univariable de conocer  $N$  ccdf's univariables.

$$F_{(N)}(z_1, \dots, z_N | (n)) = Prob\{Z_i \leq z_i, i = 1, \dots, N | (n)\} \quad (16)$$

Más precisamente, el problema se reduce a saber las siguientes  $N$  ccdf's univariables.

1.  $Prob\{Z_1 \leq z_1 | (n)\}$
- $\vdots$
2.  $Prob\{Z_N \leq z_N | (n + N - 1)\}$

Este enfoque es independiente del algoritmo utilizado para definir la secuencia. Si bien en muchos casos se desconocen la forma de las ccdf's univariables, si uno busca sólo la reproducción de la covarianza (o del variograma) mediante el set de valores simulados  $\{z^{(t)}(\mathbf{u}_i), i = 1, \dots, N\} \forall t$ , es suficiente que cada una de las  $N$  ccdf univariables en la secuencia se identifique por la esperanza y varianza de Kriging simple  $(\mu_{KS}^*, \sigma_{KS}^{*2})$  dados los  $(n + i - 1)$  datos condicionantes [4].

### 2.4.2 Simulación secuencial Gaussiana

El enfoque secuencial es típicamente aplicado a funciones aleatorias Gaussianas dada la exactitud en términos de reproducir la distribución y la covarianza. La simulación secuencial Gaussiana utiliza este enfoque de simulación secuencial para realizar una simulación condicional a una variable continua  $z(\mathbf{u})$  la cual se modela mediante una variable aleatoria estacionaria  $Z(\mathbf{u})$  conforme la siguiente metodología [4]:

1. Determinar  $F_Z(z)$  global (cdf) y no sólo de las zonas muestreadas. En caso de ser necesario, desagrupar y suavizar mediante extrapolación.
2. Mediante  $F_Z(z)$  realizar una transformación Gaussiana  $\varphi$  (anamorfosis) de  $z$  en  $y$  con cdf  $(F_Y(y))$  normal estándar  $(N(0,1))$ .

3. Chequear hipótesis multi-Gaussiana de la variable  $y$ . En la práctica se chequea normalidad uni y bi-variable, asumiendo normalidad de orden mayor.
4. Proceder con la simulación secuencial.
  - a. Definir un camino aleatorio para visitar cada nodo de la grilla una vez. En cada nodo  $\mathbf{u}$  retener un número específico de valores condicionantes incluyendo tanto datos originales  $y$  como datos previamente simulados.
  - b. Usar SK con el variograma de los valores Gaussianos  $y$  para determinar los parámetros  $(\mu_{KS}^*, \sigma_{KS}^{*2})$  de la ccdf de la variable aleatoria  $Y(\mathbf{u}_i)$  en el punto  $\mathbf{u}_i$ .
  - c. Determinar un valor  $y^{(t)}(\mathbf{u}_i)$  de la ccdf determinada en b mediante simulación de Monte Carlo, y agregar a la base de datos.
  - d. Seguir con el siguiente nodo.
5. Destransformar los valores Gaussianos  $\{y^{(t)}(\mathbf{u}), \mathbf{u} \in \mathcal{D}\}$  en valores simulados de la variable original  $\{z^{(t)}(\mathbf{u}) = \varphi^{-1}(y^{(t)}(\mathbf{u})), \mathbf{u} \in \mathcal{D}\}$ . A menudo se realiza interpolación dentro del rango y extrapolación de las colas.

Se debe tener en cuenta que cualquier que sea el enfoque elegido para la simulación, el modelo final debe ser chequeado en el sentido que reproduzca los datos originales y sus estadísticas, específicamente su histograma y variograma.

## 2.5 Librería GSLib

GSLib (Geostatistical Software Library) consiste en un conjunto de rutinas que permiten realizar cada una de las etapas de un estudio geoestadístico de los datos hasta la estimación y/o simulación de valores de la variable estudiada.

Si bien los códigos se desarrollaron en la Universidad de Stanford en UNIX, estos no tienen restricción de sistema operativo. Los códigos pueden ser compilados mediante un compilador de Fortran 77 en adelante.

En particular, para la estimación y simulación secuencial Gaussiana de variables continuas, GSLib cuenta con los códigos *kt3d* y *sgsim* respectivamente. El primero consiste en un código que permite realizar un Kriging puntual o de bloques en un dominio tridimensional ya sea

mediante Kriging ordinario, simple o en presencia de deriva mientras que el segundo permite realizar una simulación secuencial Gaussiana con la metodología descrita en 2.4.2.

Otros códigos útiles en el proceso de estimación/simulación vienen dados por *nscore*, que permite realizar una transformación Gaussiana de los valores originales; *gamv*, que permite determinar el variograma experimental de los datos a estimar/simular; *vmodel*, que permite ajustar un modelo variográfico al variograma experimental previamente determinado; y una serie de códigos auxiliares para realizar el estudio exploratorio de los datos.

## 2.6 Comentarios sobre el estado del arte

La estacionaridad es una propiedad del modelo y no de los datos, es decir, dependiendo de la escala del problema o de la disponibilidad de datos puede cambiar la decisión sobre la estacionaridad. Un ejemplo clásico de esto corresponde al estudio de la elevación a lo largo de una dirección en un cierto dominio: en caso de considerar una escala local para el estudio, las diferencias de elevación a lo largo de una dirección pueden ser consideradas ciertamente como una variación sistemática a lo largo de tal dirección, y por ende, se puede modelar como una deriva. Ahora bien, al mirar el problema en una escala global, la dimensión de las montañas son insignificantes en comparación con la superficie de la Tierra por lo cual el problema puede ser ahora considerado como estacionario. He aquí la paradoja: si bien la estacionaridad es una propiedad de la función aleatoria y no de los datos, estos son la única información disponible para testear la estacionaridad y por lo tanto se vuelve intesteable [14]. Asumir que todas las aplicaciones son estacionarias o ignorar las ventajas de algunas formas de estacionaridad es inadecuado.

En el sentido de la importancia de la no estacionaridad en el modelo, es interesante tener en cuenta los resultados presentados por Journel, donde se muestra que en casos en los cuales se implemente una estimación por vecindad de búsqueda en situaciones de interpolación, la componente de deriva global y la componente residual otorgada por el KO y por el KT son prácticamente iguales [10]. Es decir, bajo estas condiciones, la estimación final no se ve afectada por especificaciones de la deriva de orden mayor. Esto sucede aún cuando los datos posean una evidente deriva o un muestreo preferencial por lo cual se recomienda utilizar KO en situaciones



de interpolación por su simpleza, a menos que la física de la distribución espacial sugiera una separación en distintas componentes de la deriva. Esto presenta por lo tanto una aparente paradoja al usar KO en casos de evidente no estacionaridad, para lo cual se debe notar que cualquier tipo de estimación mediante Kriging que utilice una vecindad móvil puede ser visto como un modelo no estacionario de una función aleatoria que en la práctica implica utilizar una deriva representada como función escalón en cada nodo a simular dada por una media constante pero desconocida que se estima implícitamente en cada vecindad. Si esta media es estimada como un solo término o como una deriva más compleja no importa mayormente en casos de interpolación dado que se obtienen valores similares para ambos tipos de Kriging. Ahora bien, en casos de extrapolación la decisión del tipo de Kriging y por ende del modelo de deriva cobra mayor importancia. De esta manera, en el contexto de la investigación realizada, se debe tener en cuenta el desarrollo del algoritmo en situaciones con pocos y muchos datos.

En el desarrollo de un proceso de simulación que incorpore la deriva la principal dificultad radica en la estimación e inferencia del variograma de los residuos. Esto se debe básicamente a que los residuos no son reales, si no que son un supuesto asumido a partir de la descomposición de la variable según (10), por lo que el variograma teóricamente correcto no es conocido.

En la literatura, diversos autores han desarrollado metodologías de simulación para hacerse cargo de la deriva de los datos. McLennan utiliza una modificación al algoritmo de simulación Gaussiana secuencial para capturar la deriva [13]. En este caso la modificación se encuentra en la transformación Gaussiana de los datos y en la inferencia de la distribución espacial (que en la práctica se traduce en la inferencia del variograma). En lugar de utilizar una anamorfosis global se realiza una transformación y destransformación Gaussiana local (o no estacionaria). Manchuk presenta un programa formato GSLib que corresponde a un algoritmo secuencial flexible en el sentido que recoge distintas modificaciones y enfoques con el fin de enfrentar distintos escenarios y problemas [12]. Las opciones varían según se trate de sólo una variable primaria con o sin estacionaridad aparente, de la presencia de variables secundarias, del tipo de correlación existente entre las múltiples variables y de la cantidad de información que se tenga de cada variable, es decir, una variable primaria muestreada de igual manera que una secundaria, o una variable primaria muestreada de manera menos intensa que una secundaria. Todas estas componentes definen distintos escenarios para los cuales existen enfoques y variantes de simulación secuencial

que se pueden usar y de los cuales se trata de hacer cargo el programa en cuestión. Otro tema relevante a la hora de realizar las simulaciones corresponde a la posible generación de valores implausibles como leyes negativas, para lo cual Leuangthong presenta una simulación de valores con una metodología de transformación de residuos condicionados a una deriva que permite solucionar este problema [11]. También se han realizado enfoques de simulación espectral para capturar la deriva mediante el método de Fourier, el cual permite trabajar en el espacio de frecuencias [16]. A través de esta metodología se logra evitar tener artefactos entre las fronteras de las sub-zonas definidas para los diferentes modelos de variogramas, aunque el método espectral tiene como limitación el requerimiento de una grilla uniformemente espaciada debido a la estacionaridad global del espectro de densidad.

Es interesante destacar, que si bien en la literatura existen trabajos enfocados a comparar metodologías que involucran métodos en presencia de deriva, por lo general estas comparaciones se enfocan en ver la utilidad del método en presencia de deriva versus otro método que no la considera. Como ejemplos de dichas comparaciones se tiene: Kriging simple o Kriging ordinario versus Kriging universal; co-Kriging o regresiones lineales simples versus Kriging con deriva externa; Kriging con múltiples derivas externas, entre otras.

En síntesis, en la literatura existen distintas metodologías y enfoques para enfrentar el tema de capturar la deriva en la simulación de valores de una variable regionalizada, los cuales tendrán condicionada la utilidad de su aplicación en función de la naturaleza del problema, es decir, qué tipo de variable se esté modelando y con qué objetivo, y del escenario propio del caso de estudio, es decir, cuantas variables se tengan, qué tipo de correlación posean, de qué forma se presenta la información (en grilla regular, datos dispersos), entre otras cosas. De esta manera, no existe una herramienta única que solucione el problema de representar de manera correcta la deriva en la simulación de valores, debiéndose tener en consideración tanto la naturaleza del problema como el escenario de estimación al momento de pretender capturar la deriva en los resultados en el sentido de obtener una buena reproducción de las estadísticas de los datos originales, llámese histograma y variograma, en los valores simulados.

### 3 METODOLOGÍA

El presente estudio considera una metodología de trabajo que está dividida básicamente en dos etapas. La primera de ellas considera el estudio y desarrollo del algoritmo, mientras que la segunda, guarda relación con la aplicación de dicho algoritmo en un estudio de caso real y el posterior análisis de resultados en contraste con metodologías ya existentes. La primera etapa tiene por objetivo determinar un algoritmo base sobre el cual se trabaje actualmente y, mediante una revisión exhaustiva, aplicar las modificaciones necesarias que permitan implementar una simulación secuencial Gaussiana con rechazo mediante una deriva de referencia. Para esto, se define cada una de las siguientes sub-etapas.

1. Definir algoritmo base a modificar.
2. Definir lenguaje de programación sobre el cual trabajar.
3. Implementar algoritmo de rechazo mediante modificaciones al código base y chequear la integridad del código desarrollado.
4. Validar resultados mediante un estudio de caso sintético.

La segunda etapa tiene por objetivo aplicar el algoritmo desarrollado en un estudio de caso real, contrastando su desempeño frente a dos metodologías de simulación ya existentes: simulación secuencial Gaussiana con Kriging simple (SK) y simulación secuencial Gaussiana de residuos (BT<sup>2</sup>). De esta manera, se definen las siguientes sub-etapas:

1. Obtención de base de datos para estudio de caso real.
2. Estudio exploratorio y definición de unidades de estimación.
3. Transformación Gaussiana y análisis variográfico.
4. Generación del modelo de bloques, de roca y de deriva.
5. Cálculo de residuos, transformación Gaussiana y análisis variográfico.
6. Simulación mediante Kriging Simple, Kriging con rechazo y Kriging de residuos.
7. Análisis de resultados y contraste de algoritmos.

---

<sup>2</sup> Referido al proceso de Back-Trending, el cual corresponde a la suma de la deriva a los residuos simulados.

A continuación se presenta la metodología de simulación del algoritmo desarrollado.

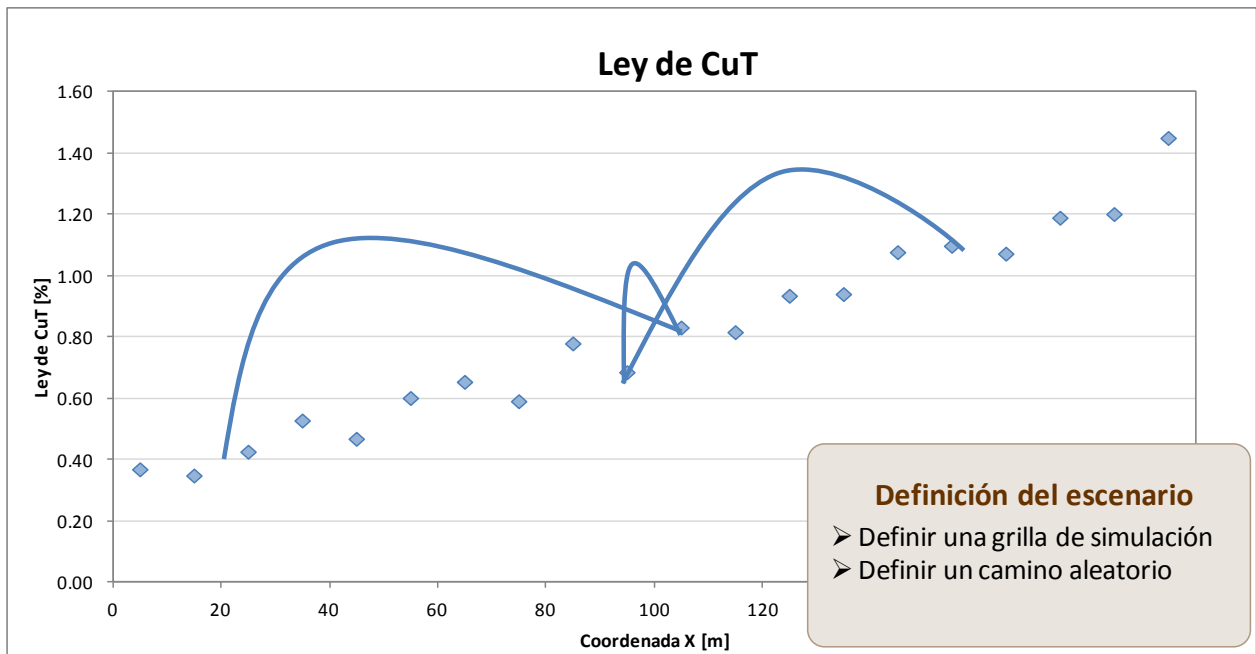
### 3.1 Metodología de simulación del algoritmo de rechazo

Respecto al algoritmo de rechazo, su metodología de simulación global corresponde a:

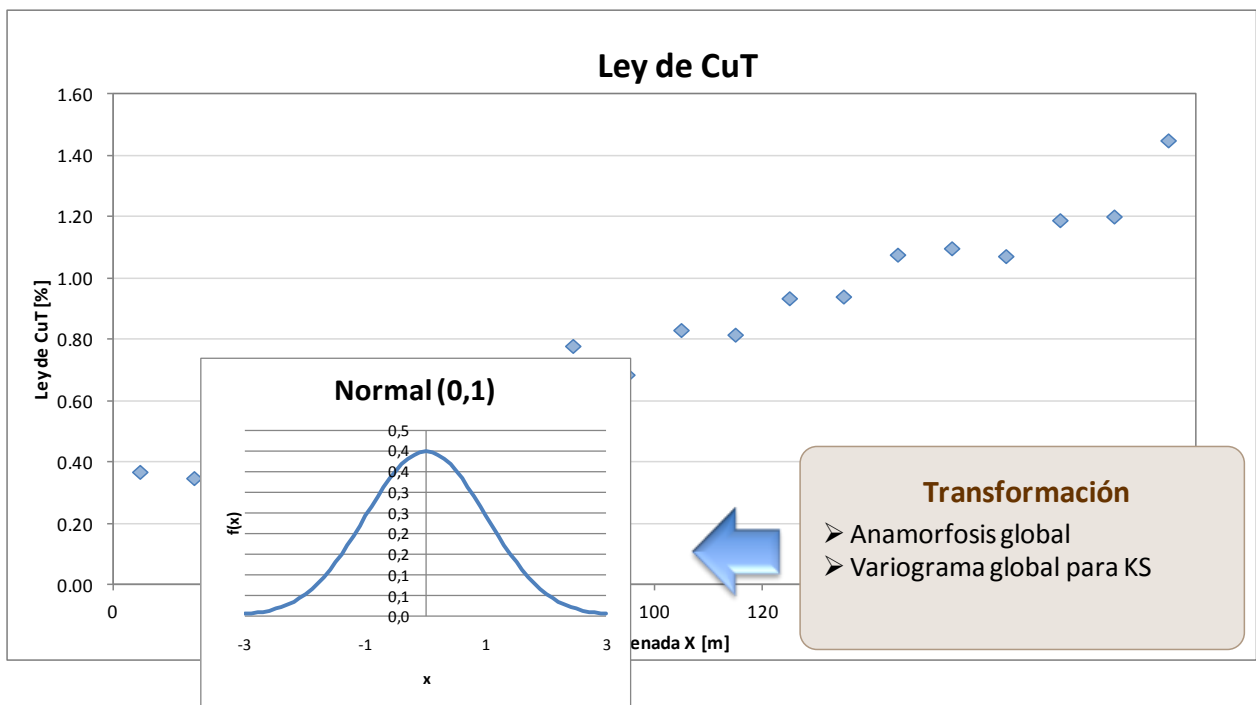
1. Se define una grilla de simulación  $G$  con  $K$  nodos y un camino aleatorio  $P_t$  en la grilla de simulación, con  $t \in \{1, \dots, T\}$  donde  $T$  representa el número de simulaciones a realizar. Se tiene que  $P_t(j) \in \{1, \dots, K\}$  y que  $\forall i \neq j, P_t(i) \neq P_t(j)$ .
2. Hacer una transformación Gaussiana global y calcular un variograma global de los datos transformados para la estimación por SK.
3. Para cada nodo  $k$ :
  - a. Tomar una vecindad en torno al punto  $\mathbf{u}_k$  y realizar un SK. En la vecindad del punto ( $D_{\mathbf{u}_k}$ ) se tendrán los valores de  $y(\mathbf{u}_i)$  con  $i \in D_{\mathbf{u}_k}$  que condicionan la distribución de  $y(\mathbf{u}_k)$ .
  - b. Se toma la estimación sobre  $\mathbf{u}_k$  junto con su varianza como los parámetros para la construcción de una cdf Gaussiana ( $y^*(\mathbf{u}_k), \sigma_{KS}^2(\mathbf{u}_k)$ ).
  - c. Simular mediante Monte Carlo un valor entre  $[0,1]$  y mediante la cdf se obtiene el valor  $y'(\mathbf{u}_k)$  simulado.
  - d. Se define una vecindad de búsqueda  $V_{\mathbf{u}_k}$  en torno al punto  $\mathbf{u}_k$  para el cálculo de la media local, la cual tiene la misma anisotropía y rotación que  $D_{\mathbf{u}_k}$  difiriendo tan sólo en los radios de búsqueda. Se calcula la media local en esta vecindad de búsqueda  $V_{\mathbf{u}_k}$  y luego se considera el valor simulado para el cálculo de una nueva media local lo cual se define como media local simulada ya que utiliza para su cálculo el valor simulado.
  - e. Se utiliza una función de control que rechaza o acepta este valor simulado considerando una deriva de referencia  $m(\mathbf{u}_k)$ , la cual se considera como un input para la metodología planteada. Esta función de control puede ser tanto determinística como probabilística. Cuando se acepta la simulación, se agrega este valor  $y'(\mathbf{u}_k)$  al conjunto de mediciones  $SD$ .

Realizar este procedimiento para cada nodo  $k$  en el orden dado por el camino aleatorio  $P_t$ , y realizar la simulación para cada camino aleatorio  $t$ . La metodología extendida del algoritmo de

simulación planteado se presenta en sección 4.5. Se muestra esta metodología de manera esquemática para su mejor comprensión.



**Figura 2: Metodología de simulación – Paso 1.**



**Figura 3: Metodología de simulación – Paso 2.**

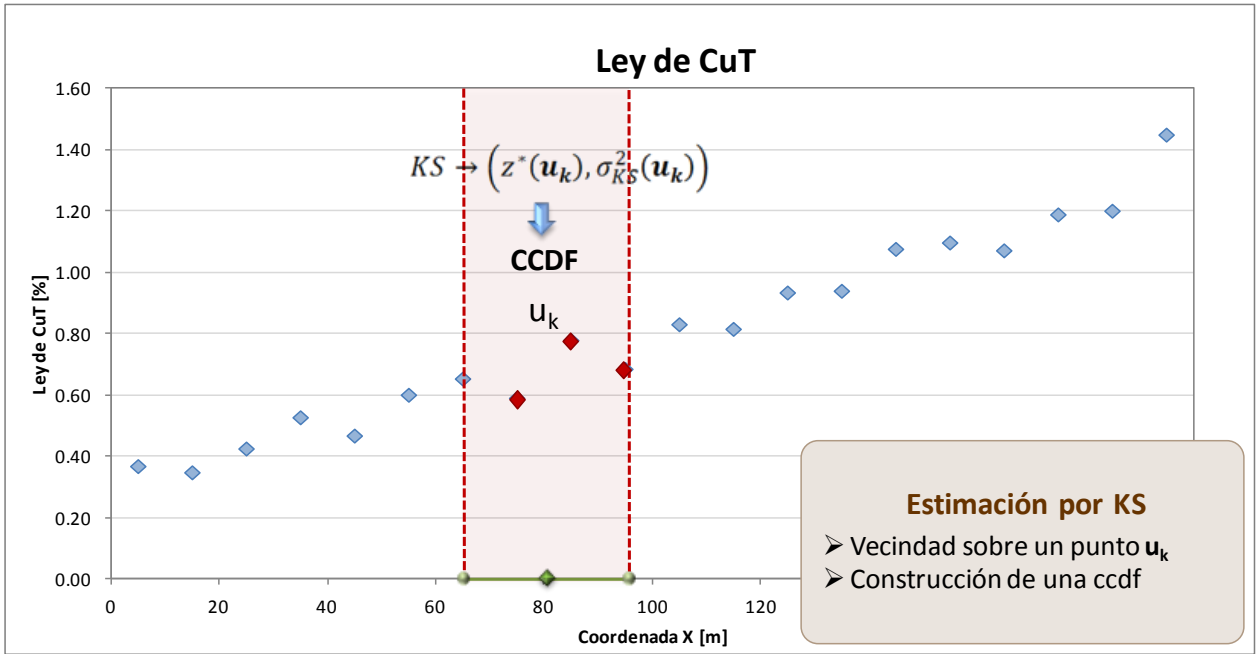


Figura 4: Metodología de simulación – Paso 3.a-b.

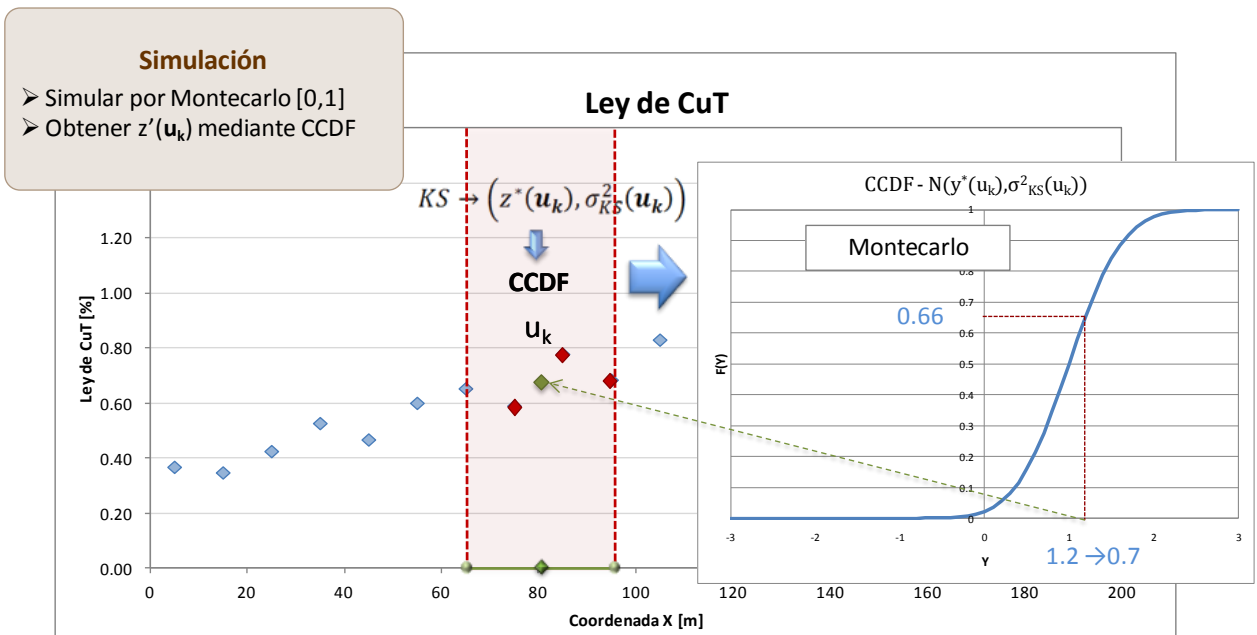


Figura 5: Metodología de simulación – Paso 3.c.

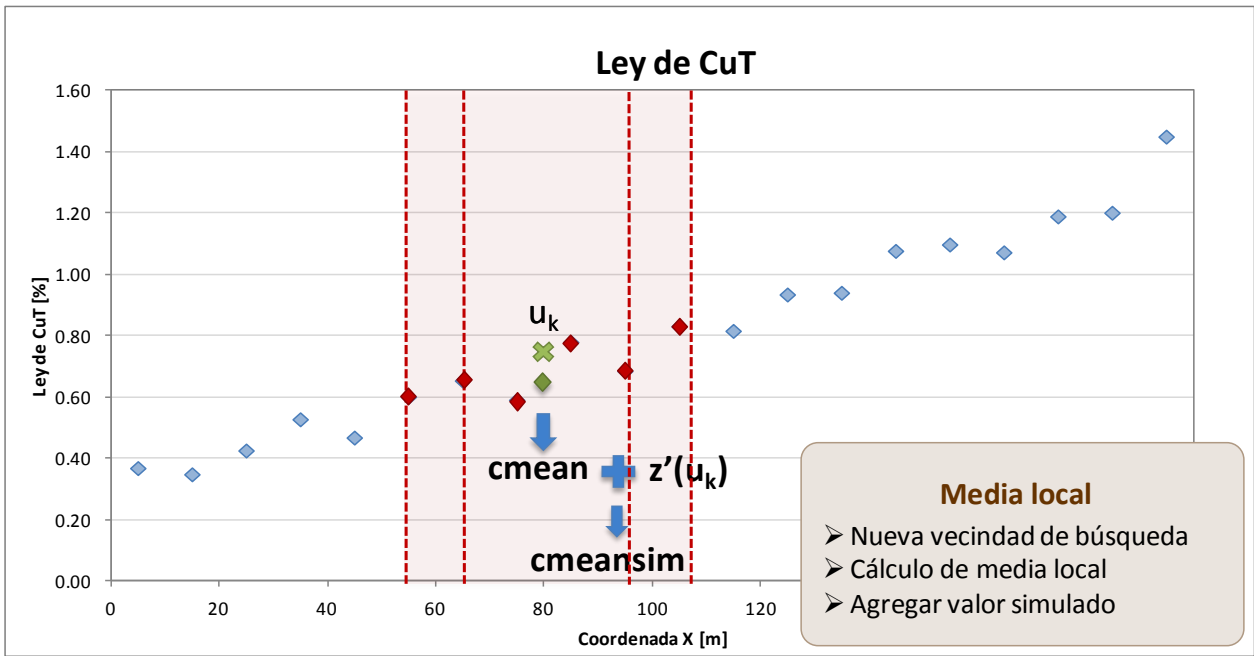


Figura 6: Metodología de simulación – Paso 3.d.

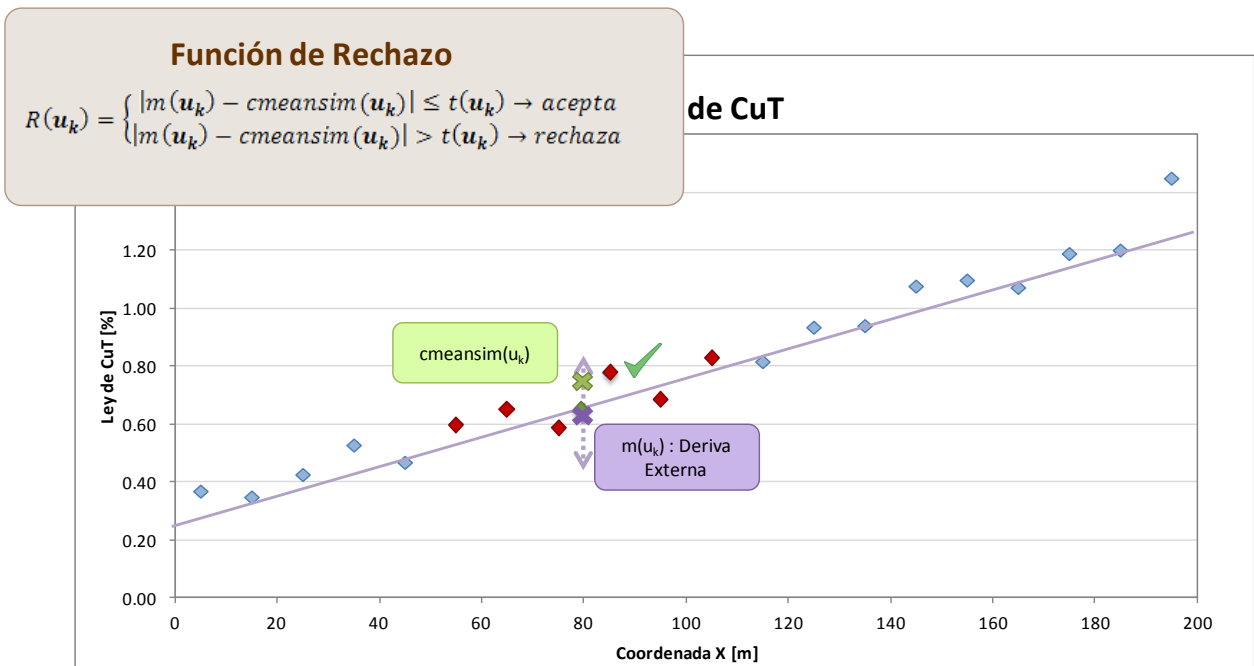


Figura 7: Metodología de simulación – Paso 3.e (Rechazo determinista).

## 4 MODIFICACIÓN CODIGO SGSIM

Dado que la metodología presentada en 3.1 corresponde a una simulación Gaussiana secuencial en la que se implementa un algoritmo de rechazo, se utiliza como base el código `sgsim.f` (extensión de archivo en código Fortran) perteneciente a la librería de herramientas para geoestadística GSLib. La compilación de dicho código entrega un archivo `sgsim.exe` (ejecutable), el cual utiliza un archivo de parámetros (`sgsim.par`) donde se entregan los parámetros correspondientes a la simulación (Anexo A).

Para realizar la modificación del código primeramente se debe definir de manera concreta los puntos 3.d y 3.e de la metodología de simulación que hacen relación a la definición del mecanismo de rechazo y aceptación de valores simulados. De esta manera, la implementación de la metodología de simulación se puede estructurar en cuatro etapas principales, las cuales corresponden a; cálculo de una media local simulada, cálculo de una tolerancia dinámica, definición del mecanismo de rechazo y resimulación de nodos en caso de ser necesario. Si bien estas cuatro etapas muestran la estructura global de la implementación del código, existe una serie de modificaciones e implementaciones menores que se deben realizar en la extensión de dicho código con el fin de dar sustento a estas cuatro etapas, las cuales corresponden, básicamente, a la definición de diversas variables que permiten apoyar los algoritmos respectivos. Para una mejor comprensión del código modificado se presentan en Anexo B cada una de las variables creadas en el desarrollo de este trabajo, así como su naturaleza y el uso que se les da.

La implementación de la metodología de simulación hace necesario agregar un nuevo tipo de Kriging al archivo de parámetros, el cual se define como tipo de Kriging 5-Kriging con Deriva de Referencia (RDR), (siendo los otros cuatro: 0-Kriging Simple (SK), 1-Kriging Ordinario (OK), 2-Kriging con Medias Locales (LVM), 3-Kriging con Deriva Externa (EXDR) y 4-Kriging Colocalizado (COLC)) Con el fin de tener un mejor manejo en el algoritmo de simulación por parte del usuario y por lo tanto una mayor flexibilidad en la implementación del código, se agregan 7 nuevos parámetros en el archivo `sgsim.par` que permiten influir en las cuatro etapas mencionadas anteriormente: tolerancia inicial y tolerancia final para la deriva de referencia, tipo de decaimiento de la tolerancia dinámica, fracción de la vecindad de búsqueda por Kriging a considerar en la vecindad de búsqueda para el cálculo de la media local, tipo de rechazo



(probabilístico o determinístico), error de referencia para el mecanismo de rechazo en caso de utilizarse un rechazo probabilístico y número aceptable de rechazos para un valor en un nodo a simular. Estos nuevos parámetros se revisan con mayor detención en las siguientes secciones sobre las etapas en las cuales se detalla la implementación del algoritmo.

```

0.5  0.2      - Initial and final tolerance for the imposed drift
1      - Type of tolerance decay (1:lin,2:exp)
1.5      - Fraction of search radii to search radii tol
1      1      - Probabilistic Rejection? (0=no, 1=yes), reference error to tolerance criterion
1000     - Acceptable number of rejected sim value

```

**Figura 8: Parámetros de control en implementación de simulación con rechazo.**

#### 4.1 Cálculo de media local

El cálculo de la media local viene a responder a la necesidad de tener un mecanismo capaz de realizar una integración de los datos condicionantes en el mecanismo de rechazo. Si bien se pudiese realizar tal mecanismo de rechazo considerando tan sólo el valor en el punto a simular y la deriva de referencia, el hecho de utilizar una media local permite integrar de una manera flexible los datos condicionantes en el rechazo o aceptación del valor simulado. Tal flexibilidad se logra al utilizar una vecindad de búsqueda para el cálculo de la media local de dimensiones distintas a la vecindad de búsqueda para la estimación por Kriging. Para esto se utiliza el parámetro de control ‘-Fraction of search radii to search radii tol’ el cual corresponde a la fracción a la cual se aumenta o disminuye la vecindad de búsqueda por Kriging y se guarda en la variable *ftol*.

Previo a la explicación del cálculo de la media local, se debe hacer referencia al mecanismo de búsqueda del algoritmo sgsim. Existen dos fuentes de datos que se pueden utilizar como condicionantes; por un lado está la base de datos con la que se está trabajando mientras que por otro lado se tienen los nodos simulados que se agregan a la base de datos una vez aceptados. Cada uno de estos dos conjuntos de datos tiene un mínimo y un máximo de datos a usar en el condicionamiento del Kriging a la vez que tienen distintos mecanismos de búsqueda. Para los valores de la base de datos original se utiliza a priori una búsqueda por súper bloques mientras que para los nodos previamente simulados se utiliza una búsqueda en espiral. Sin embargo existe

la opción de asignar los datos originales a los nodos más cercanos, lo que radicará en una búsqueda en espiral para todos los datos condicionantes y un máximo de estos dado sólo por las especificaciones del máximo de nodos a utilizar. En el caso de la búsqueda por súper bloques, el algoritmo utiliza dos subrutinas externas a ser modificadas: *'srchsupr'*, la cual realiza la búsqueda propiamente tal; y *'picksupr'*, la cual define los súper bloques para la búsqueda por *srchsupr*. Por su parte, la búsqueda en espiral se realiza mediante una subrutina interna a ser modificada *'srchnd'*. También se debe modificar la definición de la tabla de covarianzas que se calcula previamente mediante la subrutina interna *'ctable'*, dado que ésta se define en función de las dimensiones de la vecindad de búsqueda.

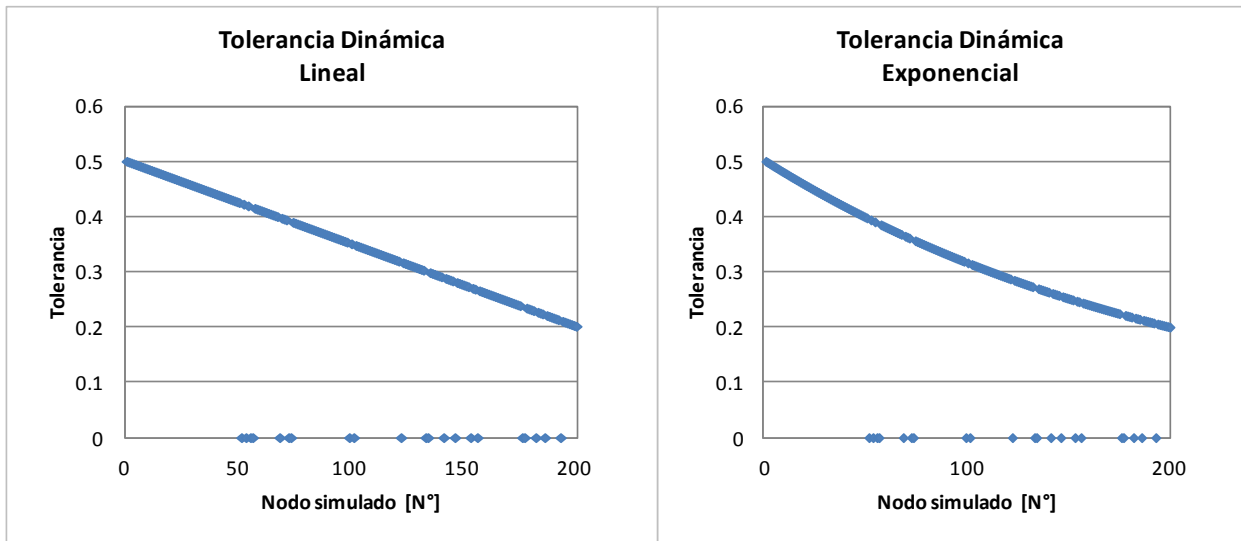
En esta etapa se crean variables indicadoras del tipo de búsqueda (*'srindy'*, *'srctype'* y *'supsrctype'*) que determinan si se está buscando datos para condicionamiento por Kriging o para cálculo de media local. Para un mayor detalle de éstas y el resto de las variables utilizadas en esta etapa revisar Anexo B.

De esta forma, para el cálculo de la media local se realiza una búsqueda de datos ya sea en súper bloques, en espiral o en ambos, dependiendo del caso, considerando una vecindad de búsqueda que tiene la misma rotación y anisotropía que la vecindad de búsqueda para el condicionamiento del Kriging. Dada esta restricción sólo basta considerar un factor (en este caso el parámetro *ftol*) que modifique proporcionalmente el radio de búsqueda *hmax* para tener totalmente definida la nueva vecindad. De esta manera un valor 2 para *ftol* corresponde a una vecindad de dimensiones equivalentes al doble de la original mientras que uno de 0.5 corresponde a una equivalente a la mitad, siempre con la misma rotación y anisotropía. En el caso extremo de considerar un factor cercano a 0, se tendrá una vecindad que sólo contendrá el nodo a simular y que, por lo tanto, tendrá una media local equivalente al nodo simulado. Es importante destacar que la función de rechazo, en cualquiera de los dos mecanismos de rechazo, se realiza en el espacio no Gaussiano, por lo cual es necesario destransformar cada uno de los valores en la vecindad de búsqueda para la media local, previo el cálculo de ésta, con el fin de obtener una media local simulada en el espacio no Gaussiano.

## 4.2 Tolerancia dinámica

De la etapa anterior (4.1), se obtiene una variable que según se haya definido mediante el parámetro *ftol*, considerará en la etapa de rechazo pocos o muchos datos condicionantes además del valor simulado. Dicha etapa sintetiza en una sola variable lo que se quiere ajustar a la deriva de referencia, pudiendo ser representada tan sólo por el valor simulado como por la tendencia en una localidad incluso mayor a la utilizada para el condicionamiento del Kriging. Ahora bien, para la etapa de rechazo se debe tener además de un criterio de rechazo, una cierta tolerancia que determine que tan ajustado se quiere la deriva de referencia a la variable a contrastar (media local).

La implementación de una tolerancia dinámica permite tener distintos criterios para el ajuste a medida que la simulación progresa en la grilla. De esta manera se puede definir una tolerancia mayor al inicio de la simulación, de tal manera de permitir un ajuste menos estricto al inicio considerando que se tienen menos datos condicionantes, y una tolerancia menor ya hacia el final de los nodos a simular dado que se tiene una gran cantidad de nodos condicionantes. Los parámetros sobre los cuales se tiene control en el archivo *sgsim.par* para esta etapa consisten en *'Initial and final tolerance for the imposed drift'*, los cuales corresponden a la tolerancia inicial siendo ésta la que se utiliza en el primer nodo a simular, y la tolerancia final correspondiente al último nodo a simular, cuyos valores se guardan en las variables *dtoli* y *dtolf* respectivamente. Se permite controlar también el tipo de decaimiento de la tolerancia mediante el parámetro *'Type of tolerance decay (1:lin, 2:exp)'* el cual se guarda en la variable *dtoltype* y determinará si se trata de un decaimiento lineal (1) o exponencial (2). En el Gráfico 1 se muestra un ejemplo utilizando los dos tipos de decaimiento para una tolerancia inicial de 0.5 y final de 0.2. El eje x representa el número en la secuencia aleatoria del nodo, por lo tanto, el nodo de menor valor en este eje corresponde al primero en ser simulado, mientras que el de mayor valor corresponde al último nodo en ser simulado. Los puntos con valor 0 corresponden a los nodos donde ya se tiene un valor gracias a la base de datos original y que, por lo tanto, no utilizan la tolerancia dinámica dejándola en su valor por defecto equivalente a 0.



**Gráfico 1: Tolerancia dinámica – Tipo de decaimiento.**

El decremento de la tolerancia dinámica, dado por  $mtol$  dependerá del tipo de decaimiento. Considerando un nodo a simular  $u_i$  y una cantidad  $N$  de nodos, se tiene que la tolerancia  $tol(u_i)$  se obtiene como:

$$\begin{aligned}
 tol(u_k) &= dtoli + mtol \cdot (k - 1) && 1: \text{Lineal} \\
 tol(u_k) &= dtoli \cdot e^{\left(\frac{k-1}{mtol}\right)} && 2: \text{Exponencial}
 \end{aligned}$$

Donde la variable  $mtol$  viene dada por:

$$\begin{aligned}
 mtol &= \frac{dtoli - dtolf}{(N-1)} && 1: \text{Lineal} \\
 mtol &= \frac{N-1}{\ln(dtolf) - \ln(dtoli)} && 2: \text{Exponencial}
 \end{aligned}$$

El rango de valores de la tolerancia dependerá del tipo de mecanismo de rechazo que se esté utilizando. En la sección 4.3 se explica este punto.

### 4.3 Mecanismo de rechazo

Tal como se explica previamente, el algoritmo permite realizar dos tipos distintos de rechazo de la media local simulada en función de la deriva de referencia: un rechazo determinista o uno probabilístico. Esta opción se presenta como ‘-*Probabilistic Rejection?* (0=no, 1=yes), *reference error to tolerance criterion*’, que guarda los parámetros *preject* y *referror*. El primero determina el tipo de mecanismo de rechazo, siendo 1 un rechazo probabilístico y 0 un rechazo determinista. El segundo se utiliza sólo en caso de un rechazo probabilístico y guarda relación con el rango de valores de la tolerancia.

#### 4.3.1 Rechazo determinista

El mecanismo de rechazo determinista consiste en calcular la diferencia absoluta entre la media local simulada y la deriva de referencia. En caso de que tal diferencia sea menor a la tolerancia dinámica para el punto a simular, el valor se acepta, en caso contrario, se rechaza y se simula nuevamente un valor para el nodo que determinará una nueva media local simulada. Para este tipo de mecanismo, la función de rechazo ( $R(\mathbf{u}_k)$ ) queda dada por:

$$R(\mathbf{u}_k) = \begin{cases} |m(\mathbf{u}_k) - cmsim(\mathbf{u}_k)| \leq tol(\mathbf{u}_k) \rightarrow \text{acepta} \\ |m(\mathbf{u}_k) - cmsim(\mathbf{u}_k)| > tol(\mathbf{u}_k) \rightarrow \text{rechaza} \end{cases}$$

De esta manera, el rango de valores que adopta la tolerancia dinámica tiene relación con el rango de valores que tiene la variable en estudio pues tanto la deriva de referencia  $m(\mathbf{u}_k)$  como la media local simulada  $cmsim(\mathbf{u}_k)$  se encuentran en este rango, del cual se hace hincapié en que corresponde a un espacio no Gaussiano.

#### 4.3.2 Rechazo probabilístico

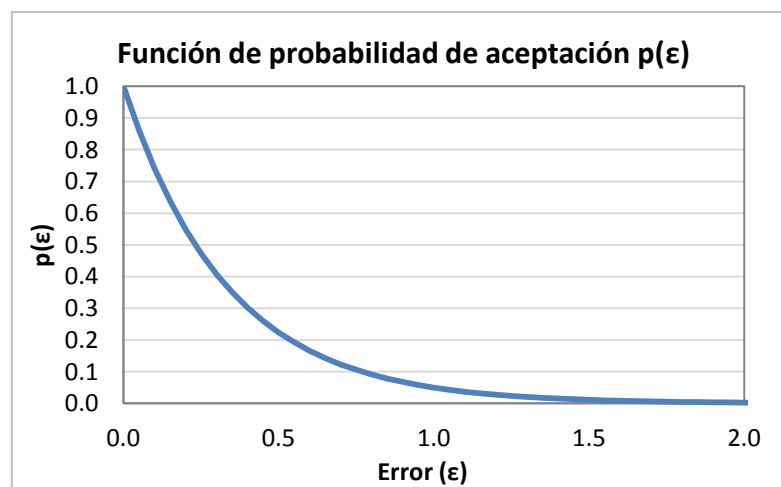
En el caso del mecanismo de rechazo probabilístico existe una cierta probabilidad de aceptar el error de la simulación con respecto a la deriva de referencia, la cual viene definida por una función de probabilidad que se ajusta en base a la tolerancia dinámica y al parámetro *referror*. La metodología de este mecanismo tiene los siguientes puntos:

1. Definir un número aleatorio  $p$  entre  $[0,1]$ .
2. Determinar el error de  $\varepsilon(\mathbf{u}_k)$  como la diferencia absoluta entre la media local simulada y la deriva de referencia  $|m(\mathbf{u}_k) - cmsim(\mathbf{u}_k)|$ .
3. Calcular la probabilidad de aceptar dicho error mediante la función de probabilidad  $p(\varepsilon(\mathbf{u}_k))$ .
4. Si  $p(\varepsilon(\mathbf{u}_k)) \geq p$  aceptar valor simulado, en caso contrario resimular nodo y obtener nueva media local simulada.

Por lo tanto, para este tipo de mecanismo la función de rechazo ( $R(\mathbf{u}_k)$ ) queda dada por:

$$R(\mathbf{u}_k) = \begin{cases} p(\varepsilon(\mathbf{u}_k)) \geq p \rightarrow \text{acepta} \\ p(\varepsilon(\mathbf{u}_k)) < p \rightarrow \text{rechaza} \end{cases}$$

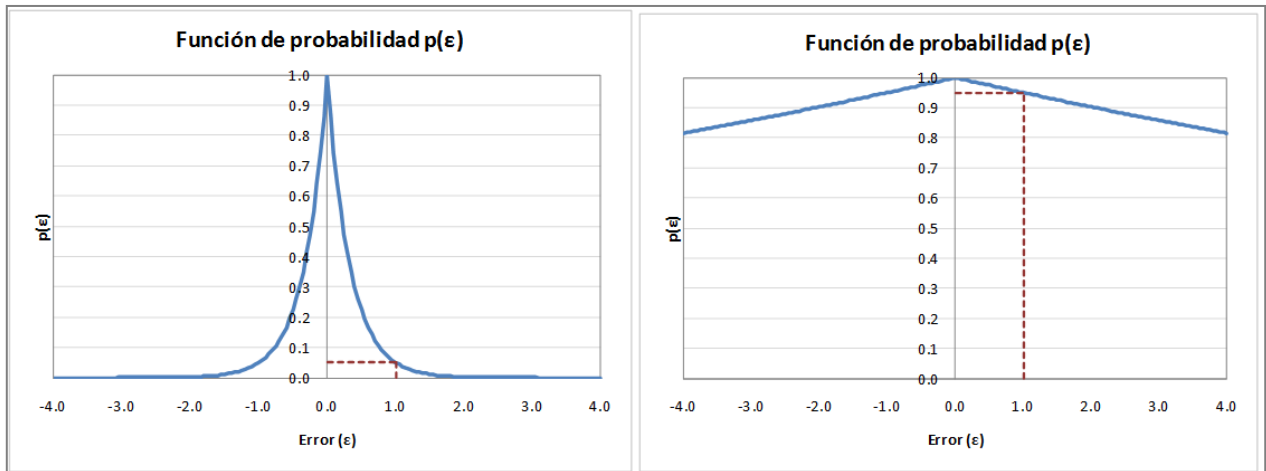
La función de probabilidad descrita previamente viene dada por:



**Gráfico 2: Función de probabilidad de aceptación -  $p(\varepsilon) = e^{-C \cdot |\varepsilon|}$ .**

Tal como se aprecia del Gráfico 2, en caso de tener un error igual a 0 (lo que es equivalente a que la media local simulada sea igual a la deriva de referencia en el nodo), la probabilidad de aceptar el valor simulado es igual a 1, es decir, siempre se aceptará este valor, y a medida que el error aumenta, disminuye esta probabilidad. El parámetro  $C$  corresponde a una estandarización de la tolerancia dinámica de tal manera de darle un sentido lógico al rango de valores de ésta. Bajo dicha estandarización, el rango de valor de la tolerancia dinámica va desde 1 como máximo para la tolerancia inicial, hasta 0 como mínimo para la tolerancia final. En este nuevo rango, un valor

para la tolerancia dinámica de 0.95 significa que la probabilidad de aceptar el parámetro *refererror* es de un 95%, mientras que una tolerancia dinámica de 0.05 significa que la probabilidad de aceptar el mismo parámetro es tan sólo de 5%. En el Gráfico 3 se muestra esta situación considerando un *refererror* de 1.0.



**Gráfico 3: Izquierda: Tolerancia 0.05 – Derecha: Tolerancia 0.95. (refererror 1.0 en ambos).**

#### 4.4 Resimulación

Una vez definida qué variable se quiere contrastar (media local) con la deriva de referencia, qué tan buen ajuste se desea (tolerancia dinámica) y de qué manera se realizará la aceptación o rechazo de la simulación, se debe definir qué tantas veces se considerará aceptable el rechazo de un nodo a simular dado que, en caso contrario, se puede generar un ciclo del cual no se termine de calcular en una cantidad considerable de tiempo. Para esto se define en el archivo *sgsim.par*, el parámetro '*Acceptable number of rejected sim value*', el cual guarda la variable *anorsv* que determina el máximo número de veces que éste se simulará antes de pasar al siguiente nodo. En caso que el número de rechazos en determinado nodo sea mayor que *anorsv*, dicho nodo quedará como no simulado (-99) y se procederá a simular el siguiente nodo.

Con el fin de minimizar el número de puntos sin simular, se implementa en el algoritmo un proceso de resimulación. Para esto se guarda el índice de cada uno de los nodos en los cuales el número de rechazos superó el valor aceptable y una vez terminada la simulación se realiza una nueva simulación en la que sólo se consideran estos índices. En caso de que se vuelva a superar el

número máximo aceptable de rechazos en un punto, este queda con un valor asignado de -99. Esta implementación considera tan sólo una etapa más de simulación de nodos.

#### **4.5 Metodología detallada de simulación del algoritmo de rechazo**

Cada una de las secciones anteriores permite tener un mejor entendimiento de la implementación del código en cada una de las etapas en que se desestructura éste. De esta manera, se presenta una metodología detallada de la simulación Gaussiana con rechazo considerando una deriva de referencia.

#### **I Definiciones previas**

1. Definir las matrices de anisotropía y rotación para la búsqueda de datos y el cálculo del variograma. (*'setror'*)
2. Definir la búsqueda por súper bloques. Esto sólo se realiza en caso que la base de datos se busque de manera aparte a los nodos simulados. (*'setsupr'*)
3. Definir la tabla de covarianza y la búsqueda en espiral. (*'ctable'*)

#### **II Loop principal sobre todas las simulaciones (DO isim=1,nsim)**

1. Definir un camino aleatorio para la realización.
2. Realizar una búsqueda en “multigrilla” en caso que esta opción esté activada.
3. Inicializar la simulación: nodos con valor por defecto equivalente a -99.
4. Asignar datos a los nodos en caso que esta opción esté activada.
5. Inicializar tolerancia.
6. Definir si se está realizando simulación o re simulación.



### III Loop principal sobre todos los nodos (DO IN=1, nxyz)

1. Determinar ubicación del nodo y asegurarse que este no tenga ya un valor asignado.
  - a. Índice del nodo será determinado por el orden dado en la definición del camino aleatorio ( $\text{index}=\text{order}(\text{IN})$ ) o por los nodos a re simular ( $\text{index}=\text{resim}(\text{IN})$ ) en caso que se encuentre en etapa de re simulación.
  - b. Revisar si valor simulado en el índice determinado ( $\text{sim}(\text{index})$ ) fue o no ya visitado y asignado.
  - c. En caso de no tener un valor asignado (-99), determinar coordenadas (xx,yy,zz) e índices (ix,iy,iz).
2. Simular el nodo (ix,iy,iz)
  - a. Definir que la vecindad de búsqueda es para condicionamiento del Kriging ( $\text{supsrctype}=\text{srctype}=0$ ).
    - i. Si la vecindad de búsqueda es por súper bloques (para base de datos), se toman los  $nclose$  datos más cercanos, tomando como máximo los  $ndmax$  más cercanos definido en el archivo de parámetros.
    - ii. Si la búsqueda es en espiral (para nodos), se toman los  $ncnode$  nodos simulados más cercanos, considerando un máximo de  $nodmax$  nodos, definido en el archivo de parámetros.
  - b. Calcular la media y desviación condicional. Si hay datos se realiza el cálculo mediante Kriging, sino, se usa la media y desviación global.
    - i. Media global  $gmean=0$
    - ii. Si el número de datos ( $nclose+ncnode$ ) es menor a 1, entonces la media condicional  $cmean=0$  y la desviación estándar condicional  $cstdev=1$
    - iii. En caso contrario se usa Kriging simple para obtener  $cmean$  y  $cstdev$ .
  - c. Definir que la vecindad de Kriging ahora es para el cálculo de la media local ( $\text{supsrctype}=\text{srctype}=1$ ).
    - i. Se realiza búsqueda de datos mediante súper bloques al igual que en el punto a.i. Destransformar estos datos para volver al espacio no Gaussiano. Calcular la suma de los  $nclose$  datos:  $sumnclose$

- ii. Se realiza búsqueda de datos mediante búsqueda en espiral al igual que en el punto a.ii. Destransformar estos datos para volver al espacio no Guassiano. Calcular la suma de los  $ncnode$  nodos:  $sumncnode$ .
  - iii. Cálculo de la media local considerando los datos de la nueva vecindad de búsqueda.
- d. Actualizar tolerancia
- i. Si  $dtoltype=1$  (decaimiento lineal), entonces la actualización es:  

$$updtol = dtoli + mtol \cdot (k - 1)$$
  - ii. Si  $dtoltype=2$  (decaimiento exponencial), entonces la actualización es:  

$$updtol = dtoli \cdot e^{\left(\frac{k-1}{mtol}\right)}$$
  - iii. Guardar el valor actualizado en el arreglo  $ddtol(IN)$ , para tener un registro de la tolerancia dinámica.
- e. Determinar un valor simulado
- i. Definir un contador de rechazos  $rcont$
  - ii. Definir un número aleatorio entre  $[0,1]$ :  $p=acorni^3(idum)$  y llevarlo al espacio Gaussiano  $N(0,1)$  para obtener un valor aleatorio Gaussiano normal estándar:  $xp$ .
  - iii. El valor simulado será equivalente a:  $sim(index)=xp*cstdev+cmean$ .
  - iv. Des transformar el valor simulado Gaussiano.
- f. Mecanismo de rechazo
- i. Calcular la media local considerando el valor simulado mediante la ecuación:  

$$cmeansim = \frac{sumnclose + sumncnode + sim(index)}{nclose + ncnode + 1}$$
  - ii. Calcular la diferencia absoluta entre la media local simulada y la deriva de referencia:  

$$diff=abs(cmeansim-lvm(index))$$
  - iii. Definir el mecanismo de rechazo. En caso de tener un rechazo probabilístico ( $preject=1$ ): definir un número aleatorio  $p$  y determinar probabilidad de aceptación  $acp$  del valor simulado mediante la función de

---

<sup>3</sup> acorni(idum): Subrutina para generar un valor aleatorio entre  $[0,1]$ .

probabilidad de aceptación  $p(\varepsilon)$  considerando la tolerancia dinámica  $uptol$  y el error de referencia  $referror$  para obtener esta función.

- iv. Si ( $diff > uptol$ ) en caso de rechazo determinista o si ( $p > p(\varepsilon)$ ) en caso de rechazo probabilístico, el valor simulado se rechaza.
  1. Revisar si supera el máximo de rechazos aceptables  $anorsv$ . De ser así se guarda el índice del nodo para una posterior resimulación y en caso de encontrarse en etapa de resimulación ya, se deja el nodo como no simulado -99.
  2. En caso no tener más de  $anorsv$  se vuelve al paso 2.e.
- v. Si ( $diff \leq uptol$ ) en caso de rechazo determinista o si ( $p \leq p(\varepsilon)$ ) en caso de rechazo probabilístico, el valor simulado se acepta.

### **Fin del loop principal sobre todos los nodos (III)**

7. Chequear si se debe realizar una etapa de resimulación.
8. Reasignar los datos a los nodos, en caso de que se haya realizado una búsqueda por super bloques para la base de datos.
9. Destransformar cada valor y escribir los resultados.

### **Fin del loop principal sobre todas las simulaciones (II)**

Una vez terminado el loop principal sobre todas las realizaciones se obtiene como resultado un archivo de donde se guardan los resultados.

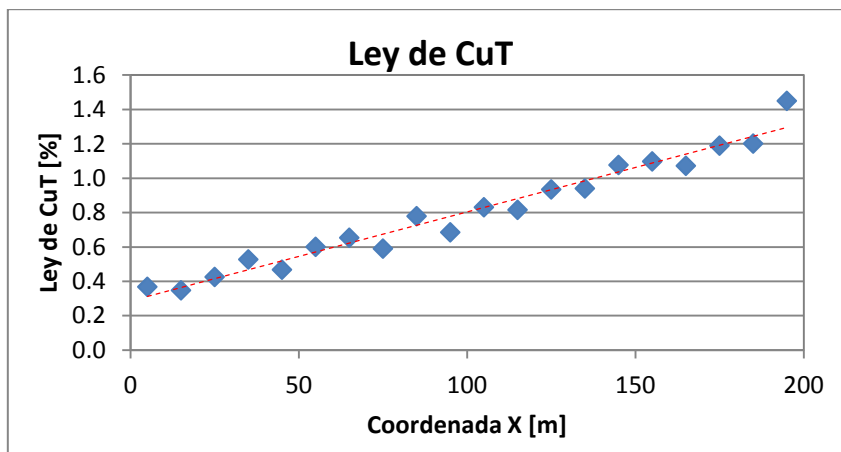
## 5 CASO DE ESTUDIO I: EJEMPLO SINTETICO EN 1D

Como primer estudio de caso se genera un ejemplo sintético en una dimensión con la finalidad de obtener resultados rápidos, de manera de lograr una visualización de los efectos de los parámetros implementados en la simulación Gaussiana con rechazo. El objetivo de este primer caso de estudio, tal como se explica en la sección de metodología, es validar los resultados obtenidos con el algoritmo de rechazo en el sentido de obtener resultados coherentes al sensibilizar los parámetros de entrada. Por ende, el foco de esta sección no está en el contraste de resultados obtenidos mediante el algoritmo de rechazo con respecto a la simulación mediante Kriging simple o Kriging de residuos ya que esto se estudia en detalle en el caso de estudio 3D presentado en la sección 6. Para efectos de chequear la integridad del código, se realiza una serie de simulaciones con el código original *sgsim* utilizando la opción de SK y las mismas simulaciones con el algoritmo desarrollado utilizando la opción de SK obteniéndose resultados idénticos.

El caso de estudio sintético consiste en muestreos de ley de cobre total (CuT) a lo largo de la coordenada X entre los rangos 0 a 200 metros, el cual se realiza cada 10 metros partiendo en la coordenada 5 metros, tal como se aprecia en el Gráfico 4.

Ley de Cu			
X	Cu	X	Cu
5	0.37	105	0.83
15	0.35	115	0.82
25	0.43	125	0.94
35	0.53	135	0.94
45	0.47	145	1.08
55	0.60	155	1.10
65	0.65	165	1.07
75	0.59	175	1.19
85	0.78	185	1.20
95	0.69	195	1.45

Tabla 1: Base – Estudio de caso 1.



**Gráfico 4: Base – Estudio de caso 1.**

Los datos presentan, a la escala del estudio, una evidente tendencia de la ley de cobre a tomar valores más altos a medida que se avanza de manera creciente a lo largo de la coordenada X. El ejemplo creado posee esta no estacionaridad en los datos por construcción y a su vez tiene una forma sencilla de crecimiento lineal sin presentar estructuras más complejas, lo cual, en una dimensión, es sencillo de visualizar, permitiendo realizar un rápido chequeo de los resultados.

Para la simulación de este estudio de caso se utilizan los siguientes parámetros.

Parámetros Generales	
Nº simulaciones	1
Semilla	69069
Transformar data?	Si
Min data	0
Max data	8
Nº Nodos simulados a usar	8
Búsqueda multigrilla?	Si
Asignar data a los nodos?	Si
Número de multigrilla <sup>4</sup>	3
zmin (extrapolación de colas)	0.0
zmax (extrapolación de colas)	5.0
Búsqueda por octante?	No

**Tabla 2: Parámetros generales de la simulación.**

<sup>4</sup> Número de refinamientos de multigrilla a utilizar. Una búsqueda un multigrilla es una técnica que utiliza distintos niveles de búsqueda que parten desde grillas más gruesas hasta grillas más densas.

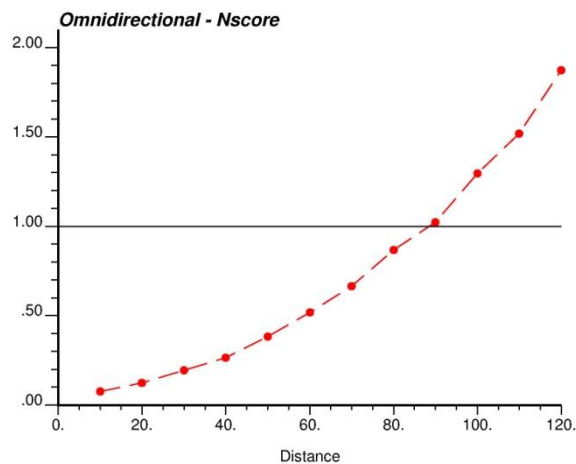
Grilla			
Parámetro	x	y	z
Nodos	201	1	1
Mínimo	0	0	0
Tamaño	1	1	1

**Tabla 3: Grilla de simulación.**

Vecindad de Búsqueda			
Parámetro	x	y	z
Radio de búsqueda	20	20	20
Cov. Lookout Table <sup>5</sup>	130	130	130
Parámetro	1	2	3
Ángulo de rotación	0	0	0

**Tabla 4: Vecindad de búsqueda.**

Dado que el proceso corresponde a una simulación secuencial Gaussiana, se presenta el variograma experimental de los datos Gaussianos correspondiente a un variograma omnidireccional con 20 pasos de 10 metros y tolerancia 5 metros.



**Gráfico 5: Variograma omnidireccional – Datos Gaussianos.**

Tal como se puede ver del Gráfico 5, el variograma posee claramente un comportamiento en que hay presencia de deriva en los datos, es decir, no alcanza la meseta teórica de valor 1, si no que aumenta indefinidamente por sobre esta. Con el fin de chequear la sensibilidad del algoritmo

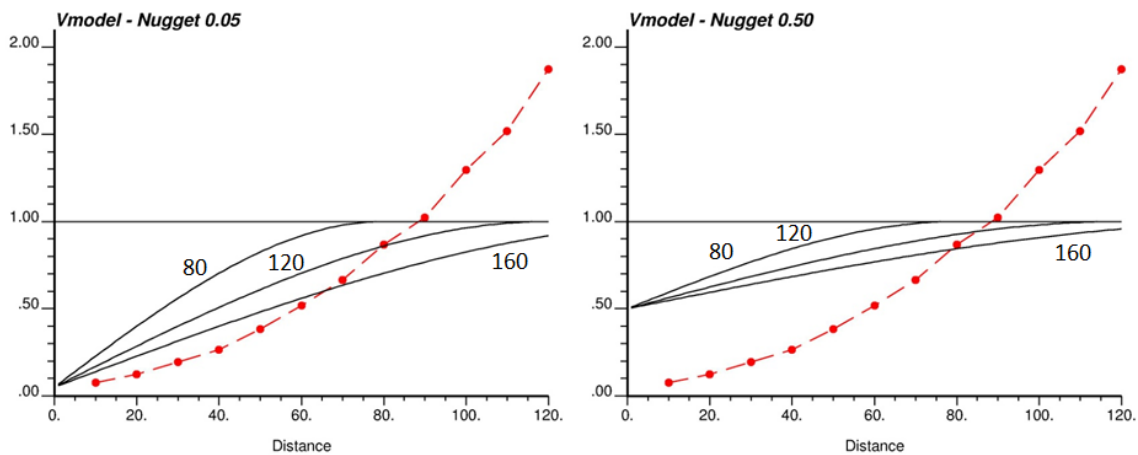
---

<sup>5</sup> Matriz de covarianzas pre-computada: Corresponde a una red 3D que contiene el valor de la covarianza en una vecindad definida en X, Y y Z. La finalidad de esta matriz es ahorrar tiempos de cómputo dentro del algoritmo.

frente al variograma modelado, se utilizan 6 modelos esféricos distintos en cuanto a alcance (80, 120 y 160 metros) y efecto pepa (0.05 y 0.50), los cuales se presentan en la Tabla 5.

N°	Pepa	Alcance	Modelo
1	0.05	80	$\gamma(\mathbf{h}) = 0.05Pepa(\mathbf{h}) + 0.95Esférico(\mathbf{h})_{(80,80,80)}$
2	0.05	120	$\gamma(\mathbf{h}) = 0.05Pepa(\mathbf{h}) + 0.95Esférico(\mathbf{h})_{(120,120,120)}$
3	0.05	160	$\gamma(\mathbf{h}) = 0.05Pepa(\mathbf{h}) + 0.95Esférico(\mathbf{h})_{(160,160,160)}$
4	0.50	80	$\gamma(\mathbf{h}) = 0.50Pepa(\mathbf{h}) + 0.50Esférico(\mathbf{h})_{(80,80,80)}$
5	0.50	120	$\gamma(\mathbf{h}) = 0.50Pepa(\mathbf{h}) + 0.50Esférico(\mathbf{h})_{(120,120,120)}$
6	0.50	160	$\gamma(\mathbf{h}) = 0.50Pepa(\mathbf{h}) + 0.50Esférico(\mathbf{h})_{(160,160,160)}$

**Tabla 5: Variograma modelados – Caso de estudio 1.**



**Gráfico 6: Variograma modelados – Caso de estudio 1.**

La deriva de referencia debe encontrarse a la resolución de la grilla de estimación, por lo cual para su cálculo se utiliza la siguiente ecuación.

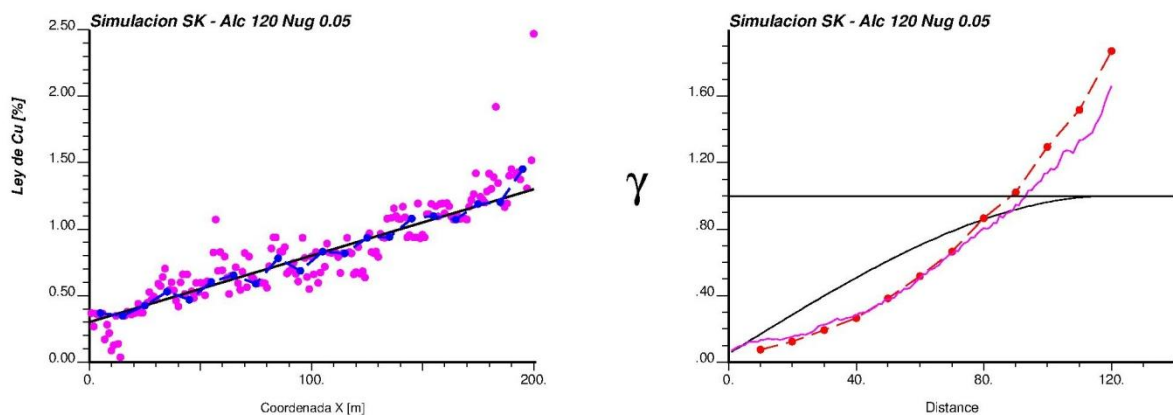
$$m(x) = 0.3 + 0.005x$$

Los parámetros presentados en la Tabla 2, Tabla 3 y Tabla 4 corresponden a los parámetros comunes utilizados en los distintos chequeos de los resultados sobre el caso de estudio I. Por su

parte los parámetros a estudiar corresponden a: mecanismo de rechazo, valor de la tolerancia dinámica, tamaño de la vecindad de búsqueda para cálculo de media local y número aceptable de rechazos. Se debe notar que el mecanismo de rechazo posee dos componentes: por un lado está el hecho de si se trata de un rechazo determinista o probabilista, y por otro el error de referencia que se utiliza en el rechazo probabilístico, lo cual incide básicamente en la tolerancia utilizada dado que mientras mayor sea el error de referencia, más tolerante será la simulación bajo el mismo rango de tolerancia dinámica. Si bien se dispone de 6 modelos variográficos, primeramente se muestra un análisis sobre un modelo base para posteriormente enfocarse en la sensibilización de los resultados al cambio de modelo variográfico. Se utiliza el modelo número 2 como base (pepa 0.05 y alcance 120 metros).

Con el fin de tener una medida de ajuste de la simulación con respecto a la deriva se calcula el valor del coeficiente de determinación ( $R^2$ ) de cada uno de los casos presentados: a mayor  $R^2$ , mayor ajuste de la simulación en torno a la deriva de referencia.

Los resultados son presentados mediante dos gráficos. El primer gráfico (izquierdo), contrasta los datos (puntos azules), la deriva (línea continua negra) y el valor de una simulación (puntos morados). En caso de haber datos no estimados en la simulación, estos no son mostrados en los gráficos respectivos. El segundo gráfico (derecho), contrasta el variograma modelado (línea continua negra), el variograma experimental (línea punteada roja) y el variograma de la simulación (línea continua morada). La meseta teórica de valor 1 también es presentada en este gráfico.



**Gráfico 7: Simulación caso de estudio I – Kriging simple.**



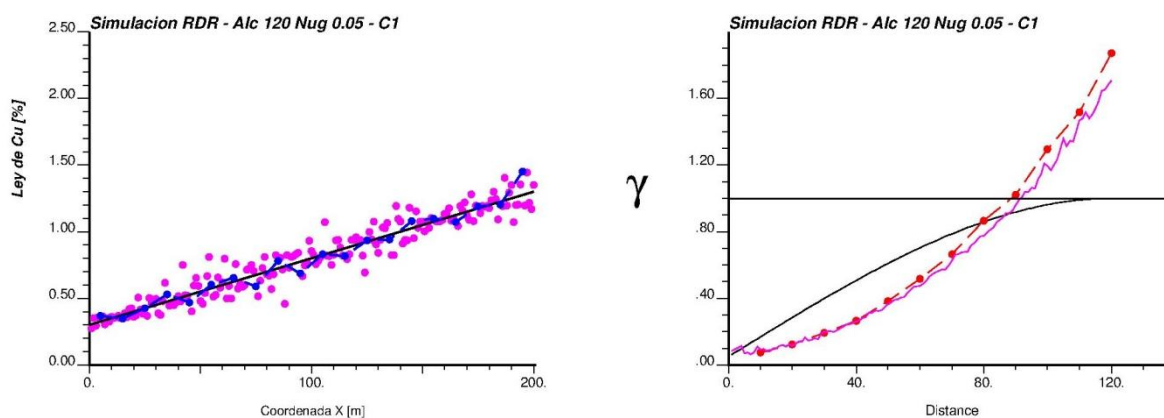
El Gráfico 7 presenta una simulación base correspondiente a una simulación mediante Kriging simple, la cual tiene un valor de  $R^2$  de 65.4%. En la siguiente sección se muestra cómo el cambio de método de simulación desde Kriging simple a Kriging con rechazo afecta los resultados en cuanto a reproducción de la deriva.

### 5.1 Variación en el mecanismo de rechazo

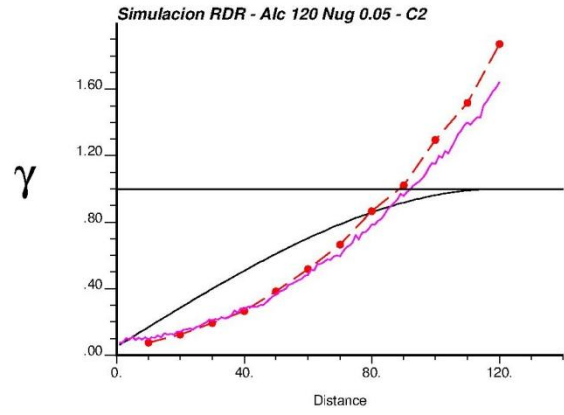
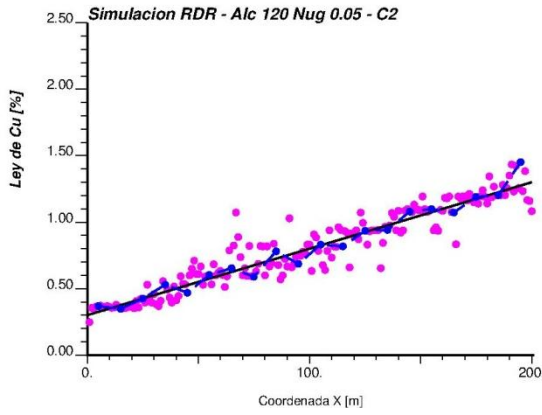
Se realiza un chequeo sobre el efecto de variar el error de referencia en el mecanismo de rechazo probabilístico. Para esto se dejan fijos los otros parámetros según la Tabla 6.

Parámetros Adicionales	
Tolerancia inicial	0.10
Tolerancia final	0.05
Tipo de decaimiento tolerancia	EXP
Fracción del radio de búsqueda	0.2
Rechazo probabilístico?	Si
<b>Error de referencia</b>	<b>0.01 – 0.1 – 1 – 10</b>
Número de rechazos	1,000

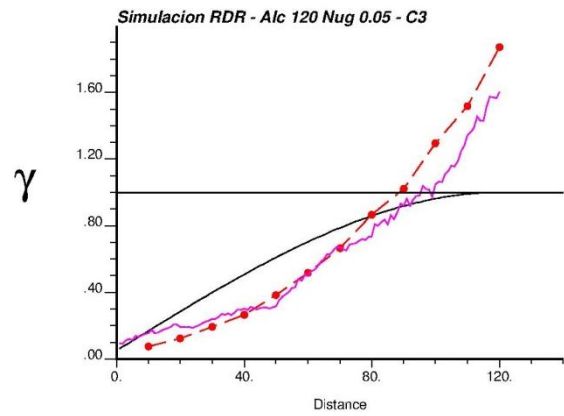
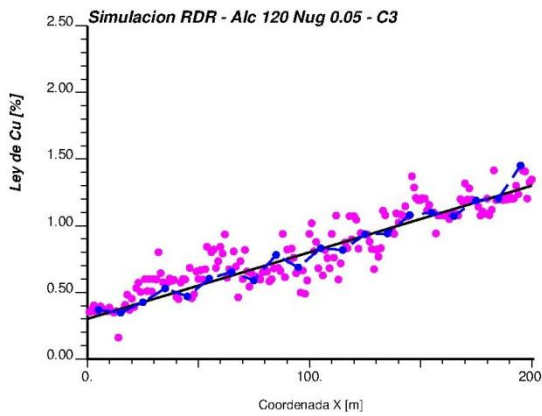
**Tabla 6: Parámetros adicionales – Variación mecanismo de rechazo.**



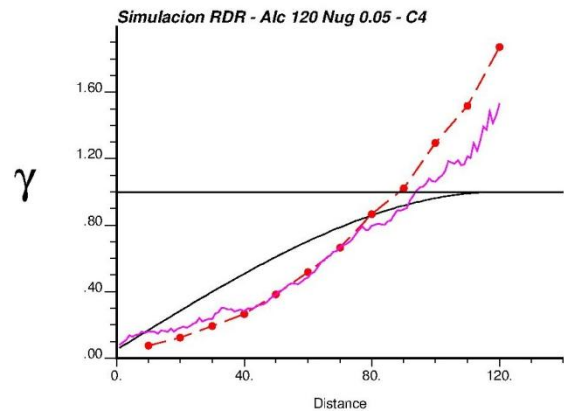
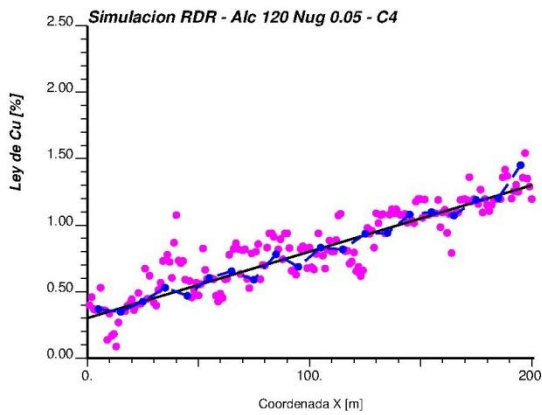
**Gráfico 8: Variación de mecanismo de rechazo – Probabilístico con error de referencia 0.01.**



**Gráfico 9: Variación de mecanismo de rechazo – Probabilístico con error de referencia 0.1.**



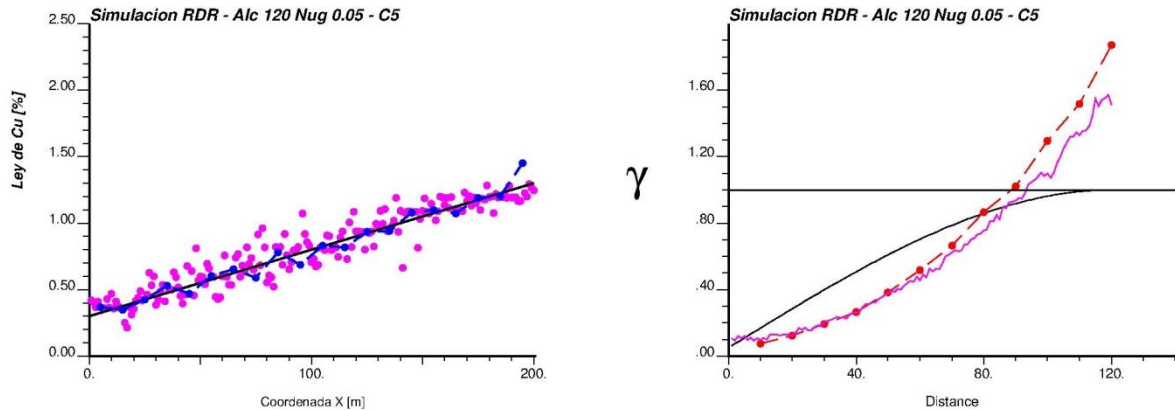
**Gráfico 10: Variación de mecanismo de rechazo – Probabilístico con error de referencia 1.**



**Gráfico 11: Variación de mecanismo de rechazo – Probabilístico con error de referencia 10.**

Se aprecia el comportamiento esperado en los resultados: a medida que se aumenta el error de referencia se obtiene una mayor dispersión de los valores simulados en torno a la deriva de referencia debido a una mayor tolerancia dinámica. De hecho los valores de  $R^2$  para los casos 1 a 4 son 92.9%, 90.8%, 85.4% y 74.4% respectivamente, lo que implica un aumento de dispersión desde el caso 1 al 4. Con respecto a la reproducción del variograma, se aprecia una menor reproducción del variograma experimental a medida que aumenta el error de referencia.

En el caso de utilizar un rechazo determinista, el parámetro de error de referencia no se utiliza, por lo cual sólo se cambia la opción del tipo de rechazo obteniéndose tan sólo un resultado al dejar los mismos parámetros fijos que para el caso probabilístico. En este caso se obtiene un valor de  $R^2$  de 90.4%, similar al caso 2, en el cual se utiliza un error de referencia de valor 0.1. El variograma obtenido también es similar al obtenido en el caso 2.



**Gráfico 12: Variación mecanismo de rechazo – Determinista.**

Se debe notar que el resultado de este caso no es directamente comparable con los resultados probabilísticos dado que, pese a tener el mismo rango de tolerancia (tolerancia inicial de 0.10 y final de 0.05), éstas no significan lo mismo: en un rechazo probabilista ésta corresponde a la probabilidad de aceptación de la diferencia absoluta calculada a través del error de referencia mientras que en el rechazo determinista, ésta corresponde directamente a la diferencia absoluta aceptable, por lo cual el hecho de decir que la tolerancia es alta o baja dependerá directamente del rango de valores de la especie de interés. En el ejemplo en estudio, el promedio de las variaciones entre un valor y su vecino es de 0.086, lo cual quiere decir que el rango de tolerancia en este caso es moderado, partiendo con una tolerancia mayor al promedio y terminando con una menor. Para comparar con el caso probabilista se debe tener en consideración tanto el rango de tolerancias

como el error de referencia. En el C1 se comienza aceptando un 10% de las veces una diferencia absoluta de 0.01 lo que significa que se acepta una cantidad aún menor de veces la diferencia de 0.086 equivalentes a la diferencia promedio de nodos vecinos. Luego esta tolerancia va disminuyendo aún más por lo que este caso es de baja tolerancia obteniéndose valores muy cercanos a la deriva de referencia. En el C2 se comienza aceptando un 10% de las veces una diferencia de 0.1 por lo que la diferencia promedio de 0.086 se acepta una cantidad aún mayor de veces, específicamente un 13.7% de las veces, bajando al último nodo a un 7.6%. De esta manera se puede considerar este caso como de tolerancia moderada, lo cual se condice con lo obtenido mediante el método determinista calificado como tolerancia moderada: ambos resultados (C2 y C5) son similares. Por último, para el C3 y el C4, dado que se tiene un valor de referencia muy alto, se tiene que la diferencia promedio de 0,086 se acepta desde un 82.0% a un 77.2% de las veces y desde un 98.0% a un 97.5% respectivamente. Estos últimos dos casos corresponden a altas tolerancias dada la alta probabilidad de aceptación de valores.

Otro hecho importante a recalcar es que sólo en el caso 1 existieron nodos a resimular (1 nodo), quedando en los otros 4 casos, todos los nodos simulados en una primera fase de simulación. De esta manera, el caso C1 tuvo tiempos de cómputo levemente mayores a los casos restantes dada su baja tolerancia y por ende su alto nivel de rechazo de valores en términos comparativos.

Para el estudio de la variación de los demás parámetros a chequear se utiliza como base el caso C2, es decir, la simulación Gaussiana con rechazo probabilístico que bajo el escenario propuesto resultó tener tolerancia moderada.

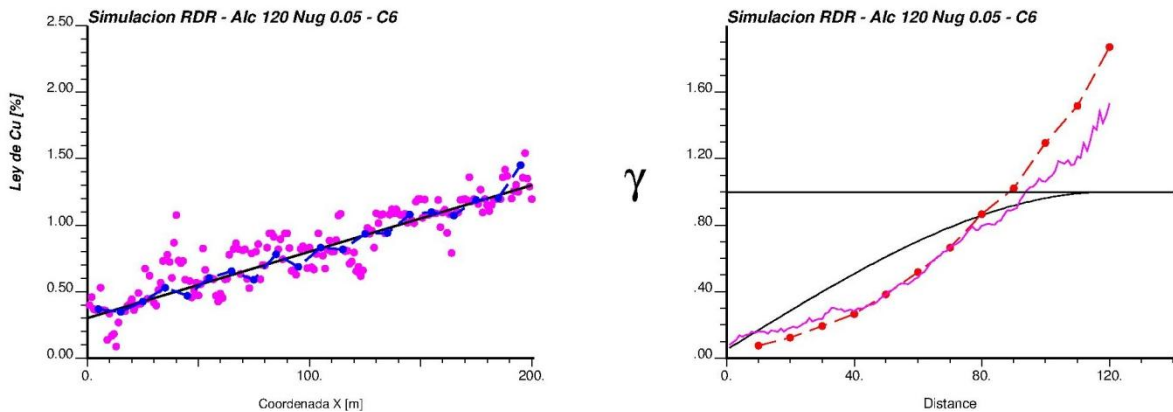
## **5.2 Variación de tolerancia dinámica**

Dado que se utiliza como base el caso C2, se mantienen los mismos parámetros adicionales que en la Tabla 6 dejando fijo el error de referencia en 0.1. Las combinaciones de tolerancia a utilizar corresponden a los pares dados por la Tabla 7.

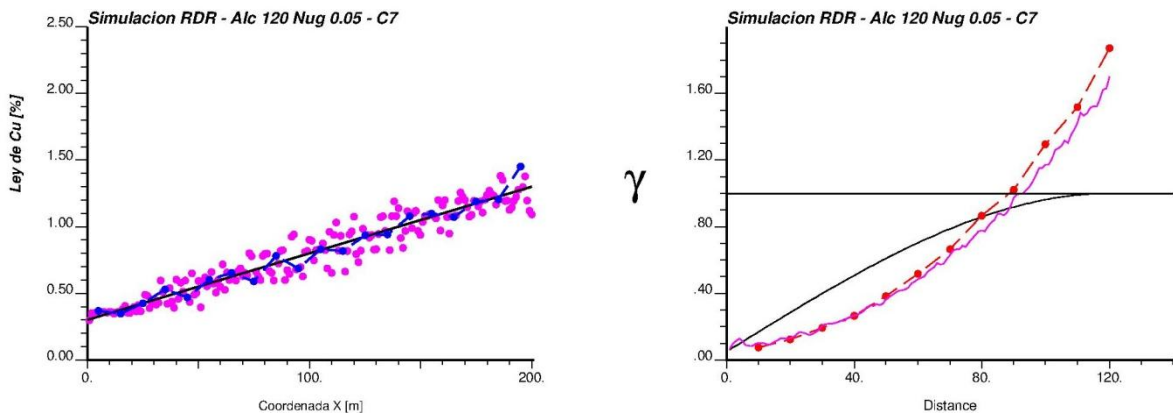
Parámetros Adicionales	
Tolerancia inicial – final	(0.999-0.95), (0.005-0.001), (0.999-0.001)
Tipo de decaimiento tolerancia	EXP
Fracción del radio de búsqueda	0.2
Rechazo probabilístico?	Si
Error de referencia	0.1
Número de rechazos	1,000

**Tabla 7: Parámetros adicionales – Variación de tolerancia dinámica.**

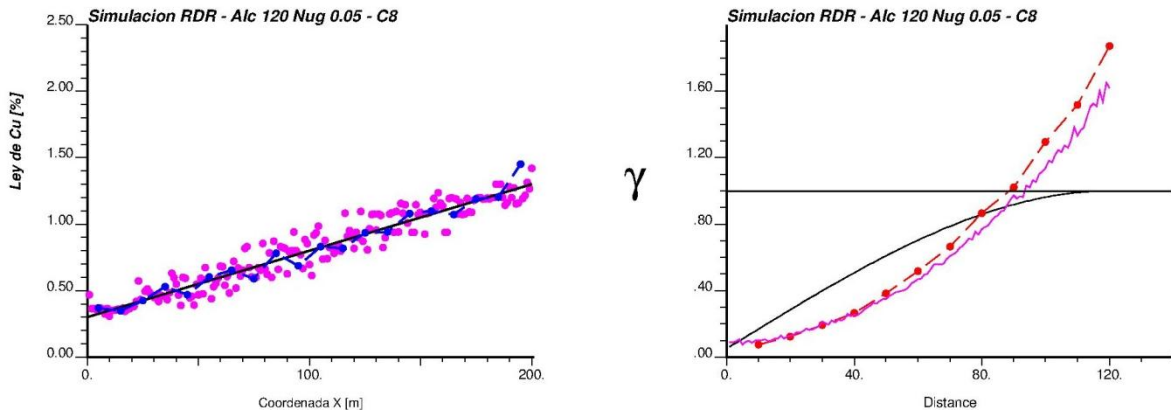
Los tres pares de tolerancia inicial-final tienen por finalidad estudiar el efecto de utilizar una alta tolerancia (0.999-0.950), una baja tolerancia (0.005-0.001) y una tolerancia dinámica de amplio rango (0.999-0.001).



**Gráfico 13: Variación de tolerancia dinámica – Inicial: 0.999 – Final: 0.950.**



**Gráfico 14: Variación de tolerancia dinámica – Inicial: 0.005 – Final: 0.001.**



**Gráfico 15: Variación de tolerancia dinámica – Inicial: 0.999 – Final: 0.001.**

Se aprecia de los resultados presentados en los gráficos de alta tolerancia (C6) y baja tolerancia (C7) una mayor dispersión para el primero y una menor para el segundo con valores de  $R^2$  de 74.4% y 92.6% respectivamente. Esto se condice con lo esperado dado que al aumentar la tolerancia se espera un aumento de la dispersión. El resultado C8 correspondiente a una tolerancia dinámica de amplio rango muestra un caso bastante similar al de baja tolerancia (C7) en términos de ajuste a la deriva con un  $R^2$  de 92.8%. En el caso de una simulación de alta tolerancia es evidente que el número de nodos a resimular será más bajo que el de una de amplio rango, sin embargo este efecto no lo es tanto al comparar una de baja tolerancia con la de amplio rango. Al utilizar una baja tolerancia inicial se está realizando un fuerte ajuste desde el inicio por lo cual los datos condicionantes utilizados en los últimos nodos a simular se encuentran ya fuertemente ajustados a la deriva de referencia. Ahora bien, si se comienza con una alta tolerancia inicial, los datos condicionantes para los últimos nodos a simular serán más dispersos, lo cual, en un nuevo escenario de simulación de baja tolerancia, implica un mayor rechazo de nodos. En términos de reproducción del variograma, se observa un mejor desempeño en los casos de baja tolerancia y tolerancia de amplio rango, no ocurriendo lo mismo en el caso de alta tolerancia donde se observa una dificultad para ajustar el variograma para distancias grandes. Para el estudio de la variación de los demás parámetros a chequear se utiliza como base el caso C8, es decir, la simulación Gaussiana con tolerancia dinámica de amplio rango.

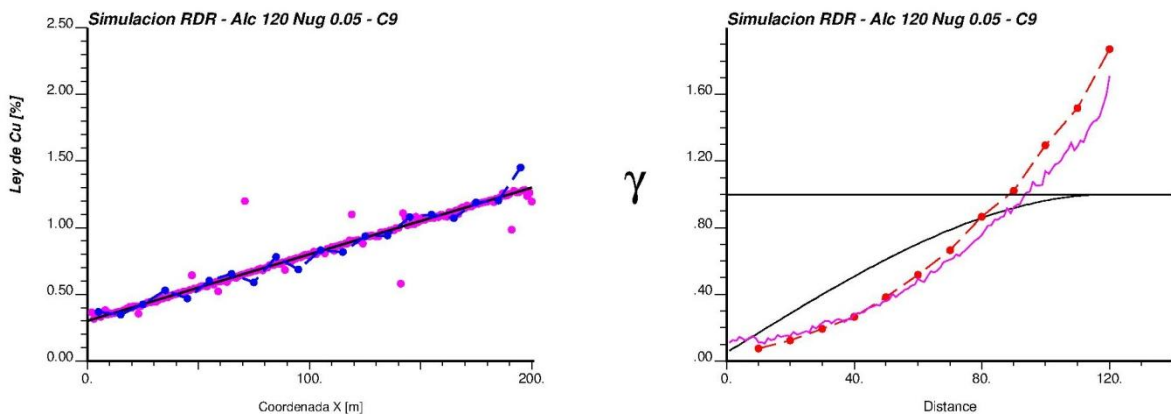
### 5.3 Variación de vecindad de búsqueda de media local

Dado que se utiliza como base el caso C8, se mantienen los mismos parámetros adicionales que en la Tabla 7 dejando fijo el par tolerancia inicial-final en (0.999-0.001) y cambiando en este caso el error de referencia por un valor 0.01 con el fin de obtener un mayor ajuste en torno a la deriva dado que se está utilizando como base un caso de tolerancia de amplio rango. Las fracciones del radio de búsqueda para el cálculo de la media local se presentan en la Tabla 8.

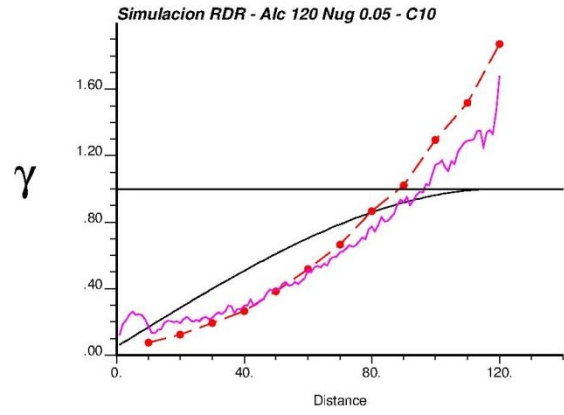
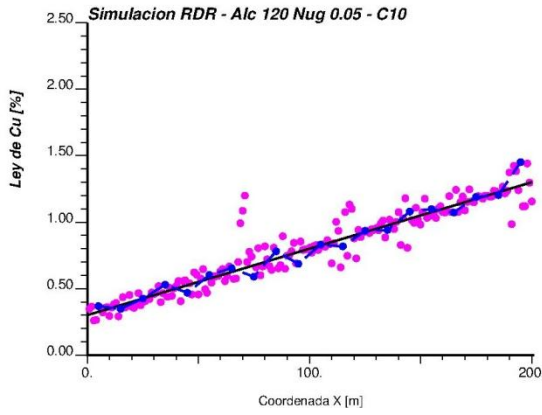
Parámetros Adicionales	
Tolerancia inicial – final	(0.999 - 0.001)
Tipo de decaimiento tolerancia	EXP
Fracción del radio de búsqueda	<b>0.01-0.1-1-10</b>
Rechazo probabilístico?	Si
Error de referencia	0.01
Número de rechazos	1,000

**Tabla 8: Parámetros adicionales – Variación de vecindad de búsqueda de media local.**

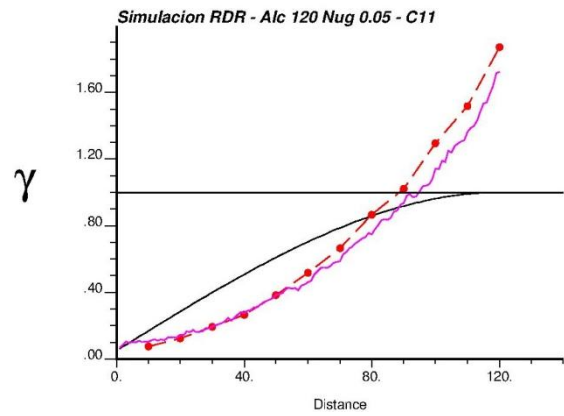
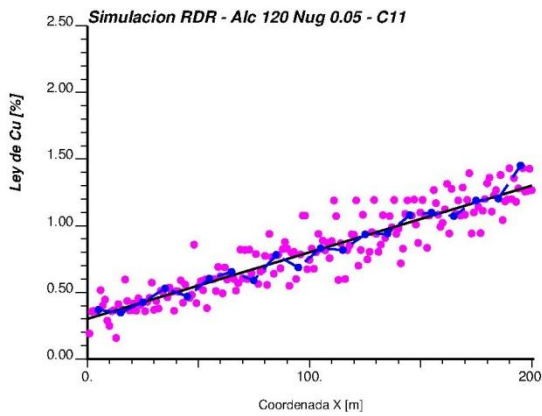
Se tiene una vecindad de búsqueda de datos condicionantes del Kriging de (20x20x20), por lo que la variación de la fracción del radio de búsqueda en los valores 0.01, 0.1, 1 y 10, implica vecindades de búsquedas para el cálculo de la media local de dimensiones (0.2x0.2x0.2), (2x2x2), (20x20x20) y (200,200,200) respectivamente.



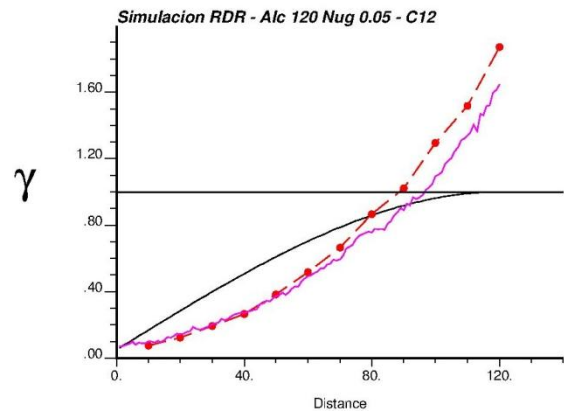
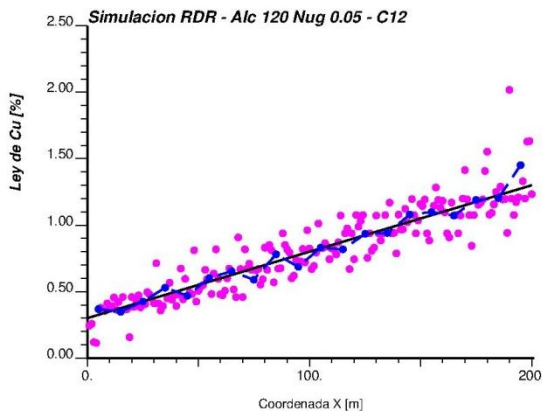
**Gráfico 16: Variación de vecindad de media local – Fracción de radio: 0.01.**



**Gráfico 17: Variación de vecindad de media local – Fracción de radio: 0.1.**



**Gráfico 18: Variación de vecindad de media local – Fracción de radio: 1.**

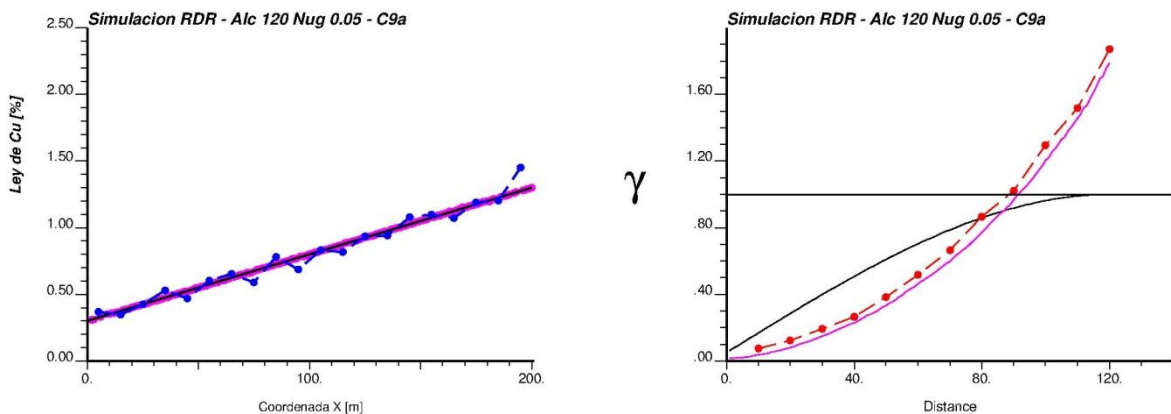


**Gráfico 19: Variación de vecindad de media local – Fracción de radio: 10.**



Dado que la resolución de la grilla es de 1 metro se tiene que en el C9 la búsqueda de datos para el cálculo de la media local encuentra tan sólo el valor simulado y por ende se debe obtener una simulación prácticamente igual a la deriva impuesta, sin embargo se obtienen valores muy alejados de la media de la simulación debido al uso de un mecanismo de rechazo probabilístico con tolerancia de amplio rango. Se realiza por tanto un nuevo caso auxiliar al C9 que permite chequear que estos valores muy distintos a la deriva son en efecto producto de la tolerancia de amplio rango, utilizando en este nuevo caso un valor para la tolerancia inicial y final de 0.005 y 0.001 respectivamente (Gráfico 20) con lo cual se obtiene un valor de  $R^2$  de 99.5%.

Los casos C10, C11 y C12 muestran una dispersión similar en torno a la deriva con valores de  $R^2$  de 90.9%, 86.4% y 84.7.0% respectivamente. Se debe notar que, a medida que aumenta el tamaño de la vecindad de búsqueda hasta abarcar la totalidad del dominio (200x200x200), aumenta también la dificultad de ajustar la media local simulada con la deriva. Esto permite decir que para el caso de estudio en cuestión, una mayor integración horizontal incide en una mayor dificultad de aceptación de valores simulados y por ende en un aumento de valores resimulados y no simulados, tal como se aprecia en la Tabla 9.



**Gráfico 20: C9 auxiliar – Tolerancia inicial: 0.005 – Tolerancia final: 0.001**

Se presenta en la Tabla 9 el número de datos resimulados y el número de datos no simulados para cada caso.

Situación	Caso				
	C9	C10	C11	C12	C9.a
N° de nodos resimulados	17	10	23	33	35

**Tabla 9: Situación final de nodos por caso. Casos C9 al C12 considerando C9.a.**

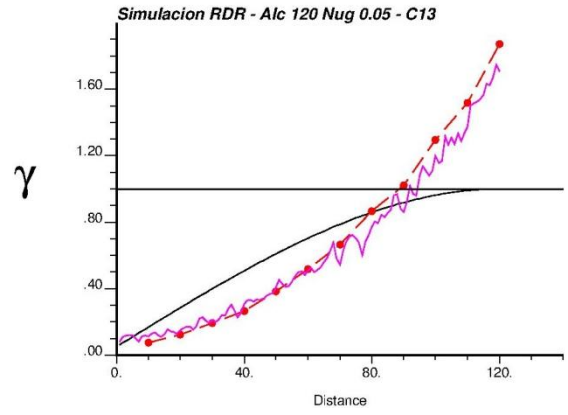
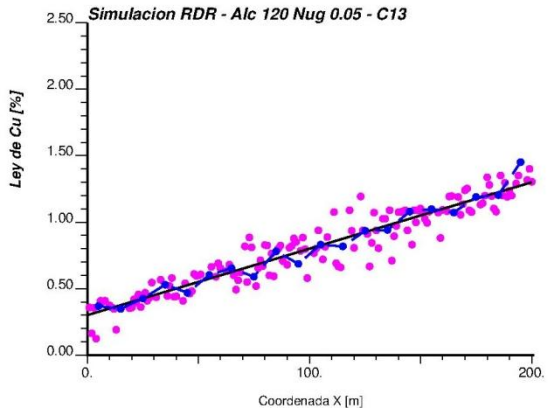
#### 5.4 Variación de número aceptable de rechazos

El último chequeo corresponde a una sensibilización en el número aceptable de rechazo. Dado que un aumento en el número aceptable de rechazos debiese disminuir el número de nodos a resimular y el número de nodos no simulados, se utiliza como base el C11 que mantiene los mismos parámetros adicionales que la Tabla 8, dejando fijo la fracción del radio de búsqueda para el cálculo de la media local en 1. El número aceptable de rechazo utilizado en cada caso se presenta en la Tabla 10.

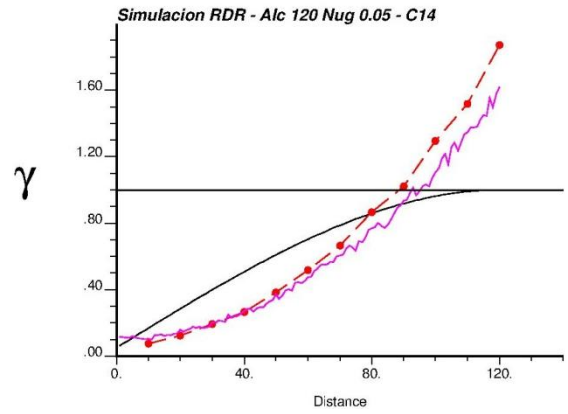
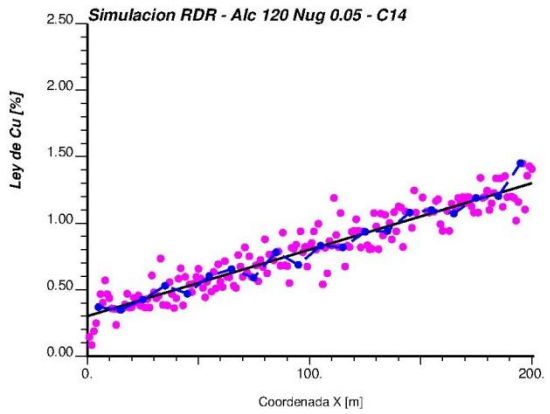
Parámetros Adicionales	
Tolerancia inicial – final	(0.999 - 0.001)
Tipo de decaimiento tolerancia	EXP
Fracción del radio de búsqueda	1
Rechazo probabilístico?	Si
Error de referencia	0.01
Número de rechazos	<b>10 – 100 – 1,000 – 10,000 – 100,000</b>

**Tabla 10: Parámetros adicionales – Variación de número aceptable de rechazos.**

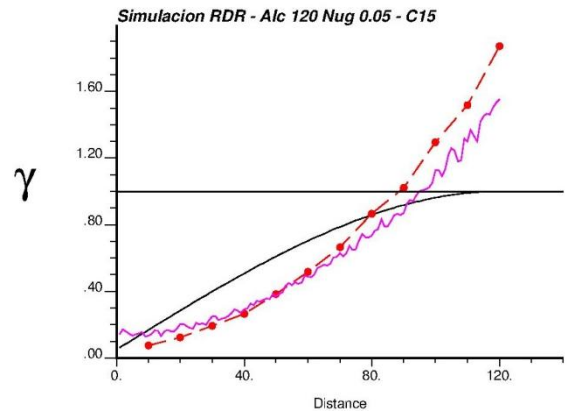
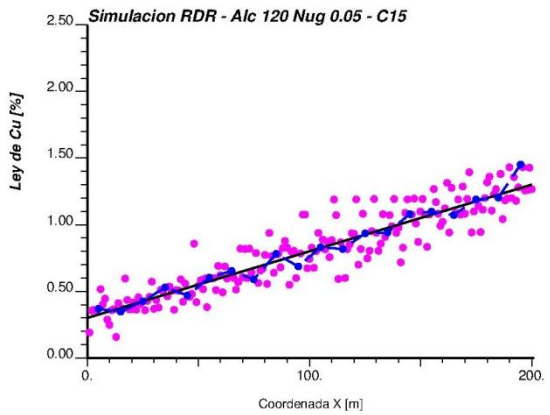
Se utiliza un contador de tiempo de cómputo cuyo tiempo inicial está al inicio del código principal y final luego de escribir los resultados de la simulación. Para cada uno de los cinco casos se obtienen simulaciones y variogramas de las simulaciones similares considerando los nodos simulados.



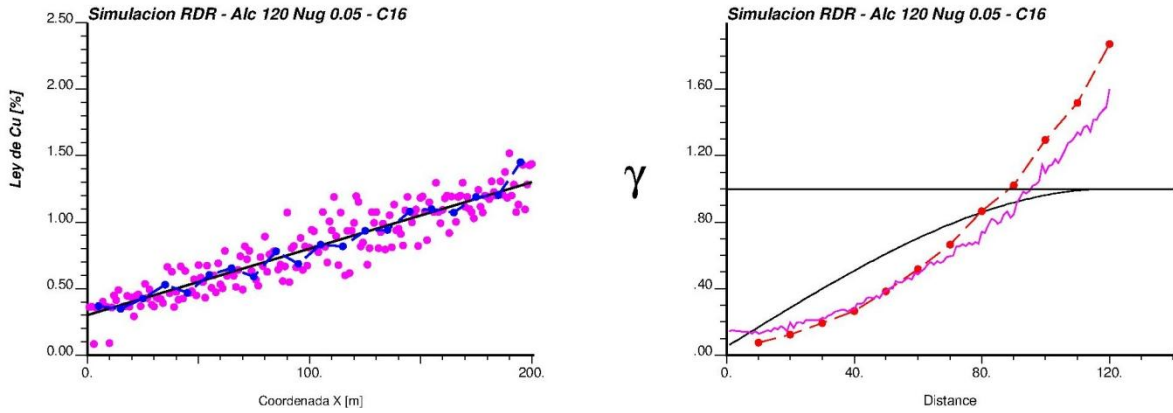
**Gráfico 21: Variación de número aceptable de rechazos – Rechazos: 10.**



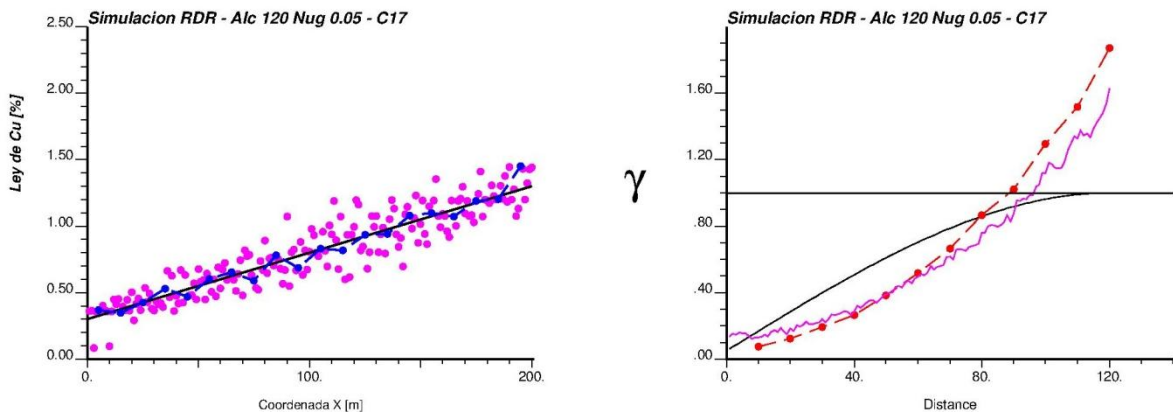
**Gráfico 22: Variación de número aceptable de rechazos – Rechazos: 100.**



**Gráfico 23: Variación de número aceptable de rechazos – Rechazos: 1,000.**



**Gráfico 24: Variación de número aceptable de rechazos – Rechazos: 10,000.**



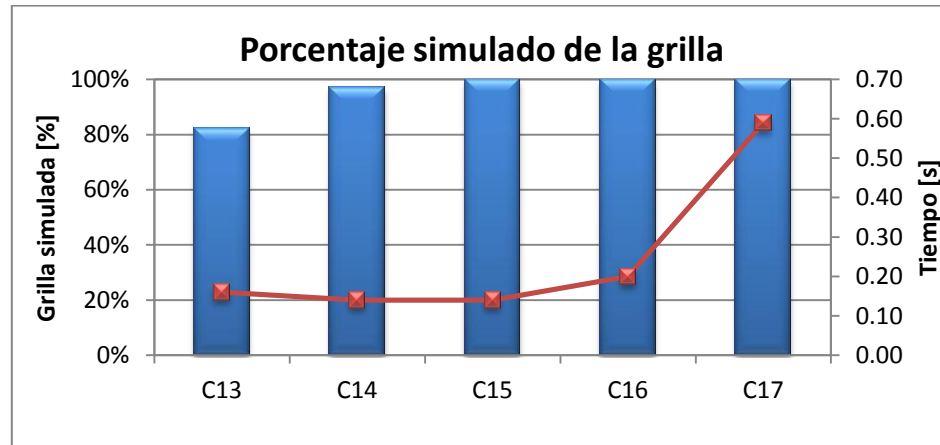
**Gráfico 25: Variación de número aceptable de rechazos – Rechazos: 100,000.**

Situación	Caso				
	C13	C14	C15	C16	C17
Nº aceptable de rechazos	10	100	1,000	10,000	100,000
Nº de datos resimulados	99	48	23	8	7
Nº de datos no simulados	36	6	0	0	0
Tiempo de cómputo [seg]	<b>0.16</b>	<b>0.14</b>	<b>0.14</b>	<b>0.20</b>	<b>0.59</b>

**Tabla 11: Situación final de nodos por caso. Casos C13 a C17.**

Si bien los resultados de los cinco casos son todos similares, se obtiene un parámetro importante como resultado de esta sección: el tiempo de cómputo, el cual se puede apreciar en la Tabla 11.

Considerando que el número de nodos de la grilla es de 201 se obtiene el siguiente gráfico que muestra el porcentaje de nodos simulados en la grilla y el tiempo de simulación.



**Gráfico 26: Porcentaje simulado de la grilla. Casos C13 a C17.**

Con los resultados sobre tiempos de cómputos se puede extrapolar el estudio de caso a una grilla tipo de dimensiones reales para estimar a grandes rasgos el tiempo de simulación en el que se incurre para cada uno de los valores de número aceptable de rechazos.

Grilla Tipo			
Parámetro	x	y	z
Dimensiones [m]	400	400	400
Tamaño de bloque [m]	10	10	10
N° de nodos	40	40	40

**Tabla 12: Grilla tipo de simulación.**

Considérese la grilla tipo de simulación dada por la Tabla 12. En esta se tienen 64,000 nodos a simular, es decir, 318 veces más nodos que el caso de estudio. Si se realiza una estimación del tiempo de cómputo de manera proporcional se obtienen los resultados de la Tabla 13. Se debe tener en cuenta que los resultados presentados siempre corresponden a tan sólo una realización y por ende se debe amplificar el tiempo de cómputo por el número de simulaciones que se piense realizar en el estudio, por lo cual se presentan los resultados de tiempo en términos de una realización y también de 100 realizaciones.

De la tabla se aprecia que al considerar 1,000 rechazos como número aceptable se obtendrían tiempos de cómputos del orden de 0.74 minutos por realización lo que implica considerar

simulaciones de 1.2 horas en caso de querer tener 100 realizaciones. Es por esto que se considera, a través de este chequeo, que 1,000 es un número aceptable de rechazos que permite obtener tiempos de cómputos razonables con una alto porcentaje de la grilla simulada (en este caso 100%). Se debe tener en cuenta que estos tiempos son sólo referenciales y dependerán de la naturaleza y condiciones de simulación del caso de estudio.

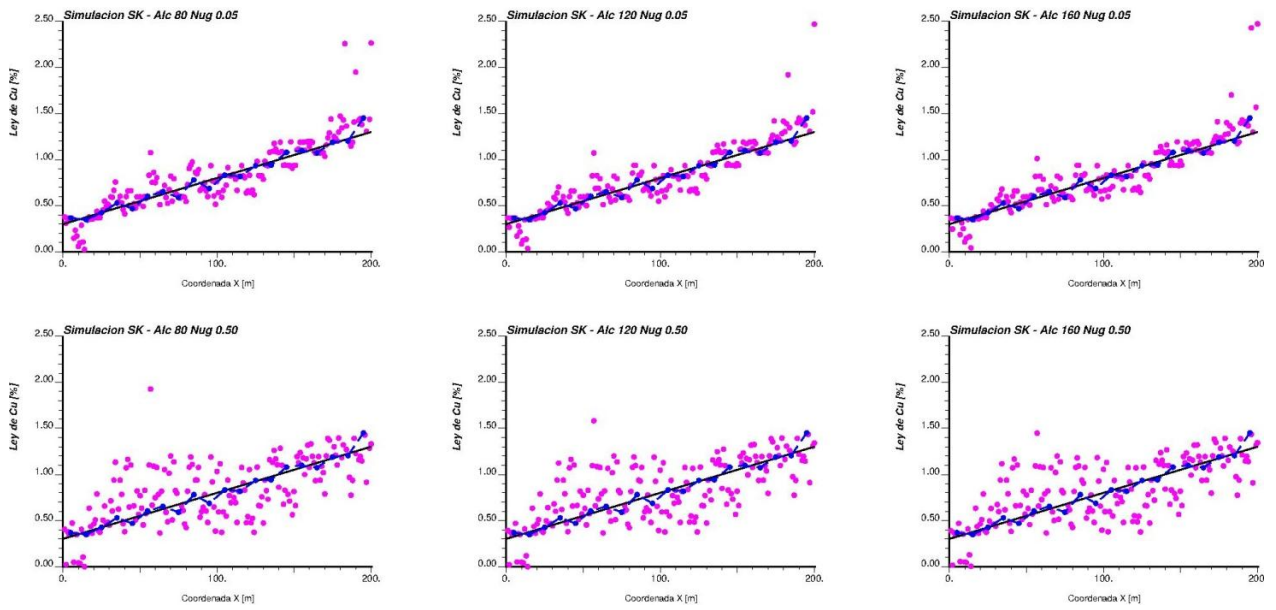
Situación	Caso				
	C13	C14	C15	C16	C17
N° aceptable de rechazos	10	100	1,000	10,000	100,000
Tiempo de cómputo [min]	0.85	0.74	0.74	1.06	3.13
Tiempo de cómputo [h]	0.01	0.01	0.01	0.02	0.05
Considerando 100 realizaciones					
Tiempo de cómputo [min]	85	74	74	106	313
Tiempo de cómputo [h]	1.4	1.2	1.2	1.8	5.2

**Tabla 13: Tiempos estimados de cómputo en grilla tipo de simulación.**

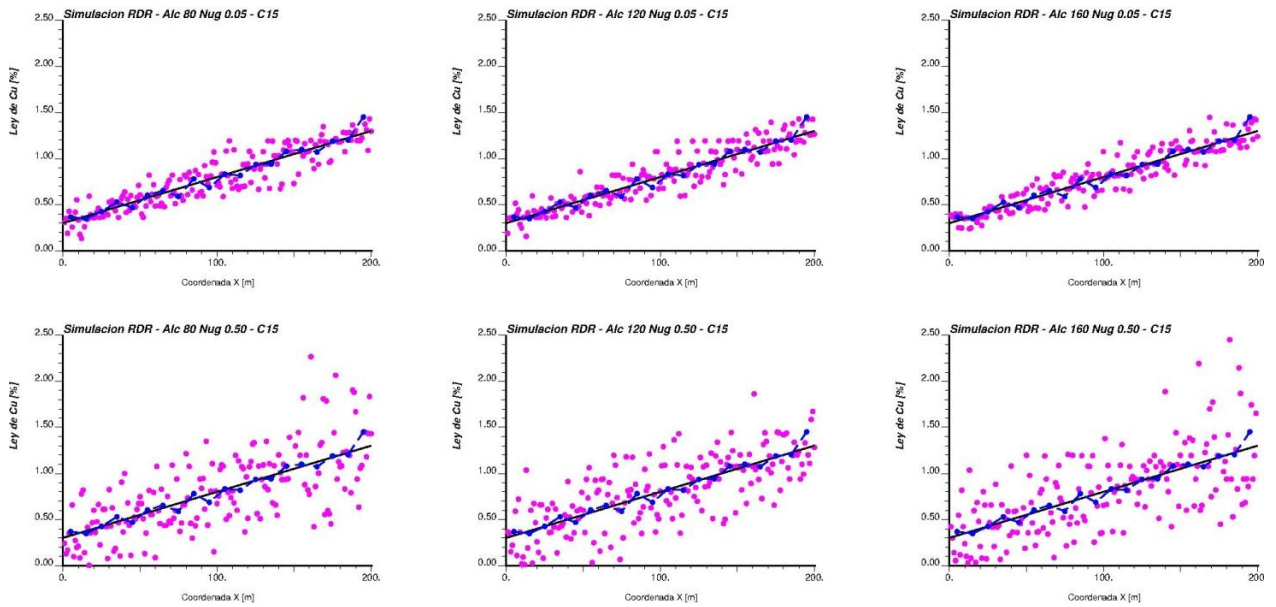
## 5.5 Sensibilización con respecto al variograma modelado

Se realizan los mismos 18 casos para cada uno de los otros 5 modelos variográficos y se grafican tanto las simulaciones como los variogramas de las simulaciones. Se presenta en esta sección los resultados de la simulación con Kriging simple y el caso C15 correspondiente a la variación del número aceptable de rechazos con 1,000 rechazos aceptables.

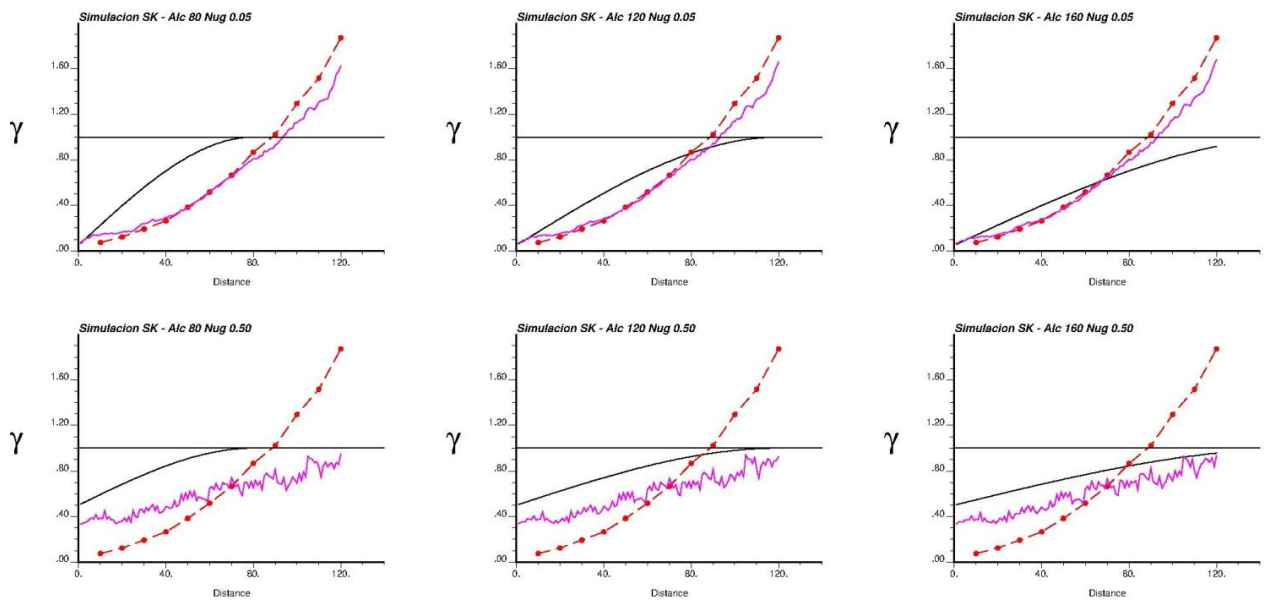
Los resultados se presentan en grupos de 6 tanto para las simulaciones como para los variogramas de las simulaciones. En la fila superior se encuentran los resultados correspondientes a los modelos variográficos con efecto pepa igual a 0.05, mientras que en la fila inferior se encuentran los resultados correspondientes a los modelos variográficos con efecto pepa igual a 0.50. De izquierda a derecha el alcance del variograma tiene valores de 80, 120 y 160 metros.



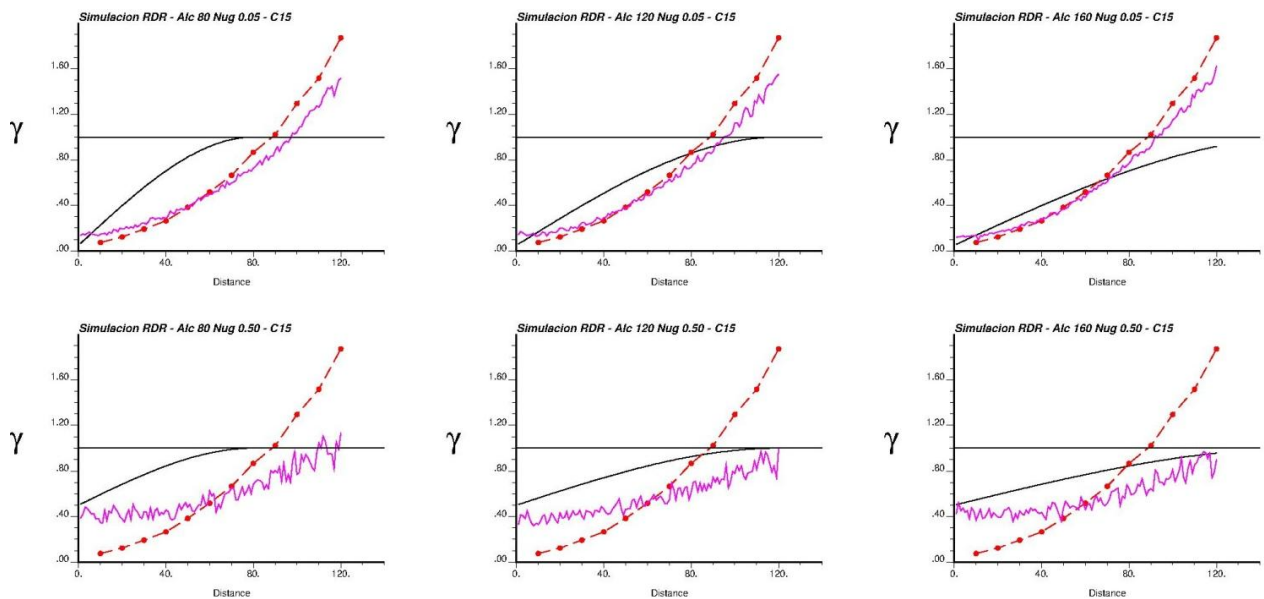
**Gráfico 27: Variación modelo variográfico – Simulaciones – Kriging simple.**



**Gráfico 28: Variación modelo variográfico – Simulaciones – Rechazos: 1,000.**



**Gráfico 29: Variación modelo variográfico – Variograma simulaciones – Kriging simple.**



**Gráfico 30: Variación modelo variográfico – Variograma simulaciones – Rechazos 1,000.**

Es posible apreciar del Gráfico 27 y el Gráfico 28 que la variación en el alcance del modelo variográfico no produce un cambio importante en los resultados de las simulaciones en lo que respecta a la dispersión de estas en torno a la deriva. Por otra parte, queda claro de estos gráficos, que el cambio en el valor del efecto pepa si produce un cambio importante en la dispersión de los resultados en torno a la deriva al menos en este caso de estudio. Se puede notar en el Gráfico 29 y



el Gráfico 30, que un cambio en el alcance del variograma modelado no tiene mayor incidencia en el variograma de la simulación. Sin embargo el cambio del efecto pepa desde 0.05 a 0.50 implica una pérdida en el buen ajuste entre el variograma experimental y el variograma de las simulaciones. Si bien se puede pensar que esto es debido a la dispersión de los valores en torno a la deriva, se debe tener en cuenta que el caso de Kriging simple con efecto pepa 0.05 y alcance 120 tiene una  $R^2$  de 65.4% mientras que el caso C8 con efecto pepa 0.50 y alcance 120 tiene un  $R^2$  de 65.5%. Sin embargo el ajuste del modelo logrado por el caso CSK usando dicho modelo variográfico es bastante mejor que el ajuste logrado por el caso C8 como es posible apreciar en el Gráfico 111 del Anexo C. Al respecto, la Tabla 14 muestra un resumen del coeficiente de determinación de cada uno de los casos para cada uno de los modelos variográficos destacando en gris el modelo número 2 correspondiente al modelo base. Es posible apreciar en esta tabla el mismo efecto que en el caso de Kriging simple y el caso 15 al variar el alcance y el valor del efecto pepa del modelo variográfico en cada unos de los otros casos, los cuales se encuentran en el Anexo C.

Coeficiente de determinación (R2 [%]) - Deriva vs Simulación						
Caso	Pepa 0.05			Pepa 0.50		
	Alc 80	Alc 120	Alc 160	Alc 80	Alc 120	Alc 160
CSK	63.1	<b>65.4</b>	67.4	29.9	30.7	31.0
C1	94.1	<b>92.9</b>	94.0	92.2	91.9	90.6
C2	91.4	<b>90.8</b>	93.5	74.4	73.8	79.2
C3	79.8	<b>85.4</b>	87.9	28.4	30.9	31.1
C4	72.8	<b>74.4</b>	75.5	10.8	11.4	11.7
C5	92.6	<b>90.4</b>	92.4	71.6	75.5	77.5
C6	72.8	<b>74.4</b>	75.5	15.2	15.6	15.7
C7	93.7	<b>92.6</b>	94.3	83.2	80.2	78.9
C8	93.3	<b>92.8</b>	91.3	69.5	65.5	63.3
C9	56.9	<b>56.9</b>	56.9	56.9	56.9	56.9
C9a	99.5	<b>99.5</b>	99.5	99.5	99.5	99.5
C10	90.4	<b>90.9</b>	92.1	89.0	86.5	88.4
C11	85.4	<b>86.4</b>	89.1	50.1	55.3	43.5
C12	77.2	<b>84.7</b>	86.1	47.4	40.6	43.2
C13	90.0	<b>90.4</b>	92.4	71.1	68.8	67.3
C14	87.7	<b>89.4</b>	90.6	65.4	61.8	58.2
C15	85.4	<b>86.4</b>	89.1	50.1	55.3	43.5
C16	85.0	<b>86.9</b>	85.9	38.2	57.2	41.2
C17	81.2	<b>86.6</b>	85.7	36.5	57.3	38.5

**Tabla 14: Coeficiente de determinación por modelo variográfico – Caso de estudio I.**

## 5.6 Comentario sobre caso de estudio I

El caso de estudio I permite visualizar de manera sencilla los efectos de la variación de cada parámetro del algoritmo de simulación con rechazo así como su resultado en términos comparativos con el Kriging simple. Es posible apreciar que el algoritmo desarrollado es consistente en cuanto obtiene los resultados esperados: mayor dispersión en torno a la deriva de referencia para tolerancias altas, errores de referencias altos y vecindad de búsqueda para el cálculo de la media local de mayores dimensiones. Por su parte el parámetro de número aceptable de rechazos también arroja resultados esperables en cuanto a medida que aumenta este parámetro, aumenta el porcentaje simulado de la grilla y los tiempos de cómputo. Es importante tener en consideración que el algoritmo de simulación con rechazo permite agregar flexibilidad a los resultados dado que la simulación mediante Kriging simple sólo obtiene un único resultado, el cual es comparable con los casos C4 (error de referencia 10), C6 (alta tolerancia) y C12 (fracción del radio de búsqueda 10), todos los cuales son los de resultados de mayor dispersión en torno a la deriva de referencia. En este sentido, el algoritmo con rechazo permite obtener una mayor gama de resultados, desde valores muy ajustados a la deriva de referencia hasta valores de mayor dispersión en torno a tal deriva, asemejándose estos últimos a los resultados obtenidos mediante Kriging simple.

Con respecto a la sensibilización del variograma modelado, se debe tener en cuenta que los resultados no están fuertemente condicionados por el alcance del modelo variográfico, no así por el efecto pepa utilizado, el cual tiene una fuerte incidencia en la dispersión de los simulaciones en torno a la deriva de referencia y en la reproducción del variograma experimental a través del variograma de las simulaciones. Finalmente, es posible observar de este estudio de caso que el parámetro de mayor incidencia en los resultados corresponde al tamaño de la vecindad de búsqueda de la media local seguido en importancia por los parámetros de tolerancia y error de referencia. Por su parte, el parámetro de número aceptable tiene relevancia en el porcentaje de la grilla simulado y en los tiempos de simulación, no así en los resultados mismos de la simulación.

## 6 CASO DE ESTUDIO II: BASE DE DATOS 3D

El siguiente caso de estudio corresponde a la implementación del algoritmo propuesto en una base de datos real proporcionada por Compañía Minera Cerro Colorado (CMCC) con el fin de ser usado en una determinada zona de interés para ambas partes: a saber, una zona de óxidos en donde se presenta una tendencia sistemática a la baja de los valores de las leyes de cobre a medida que se disminuye en la cota. Además de esta zona de interés, se realiza el estudio sobre cada una de las distintas unidades de estimación dentro del modelo de bloques a simular.

### 6.1 Descripción del yacimiento

Compañía Minera Cerro Colorado es una empresa de BHP Billiton, la cual produce alrededor de 125 mil toneladas de cátodos de cobre de alta pureza anuales. El yacimiento está ubicado en la región de Tarapacá, a 120 km al Este de la ciudad de Iquique, en la comuna de Pozo Almonte, primera región de Chile. Se encuentra localizado geográficamente en la precordillera a una altura promedio de 2,600 metros sobre el nivel del mar.

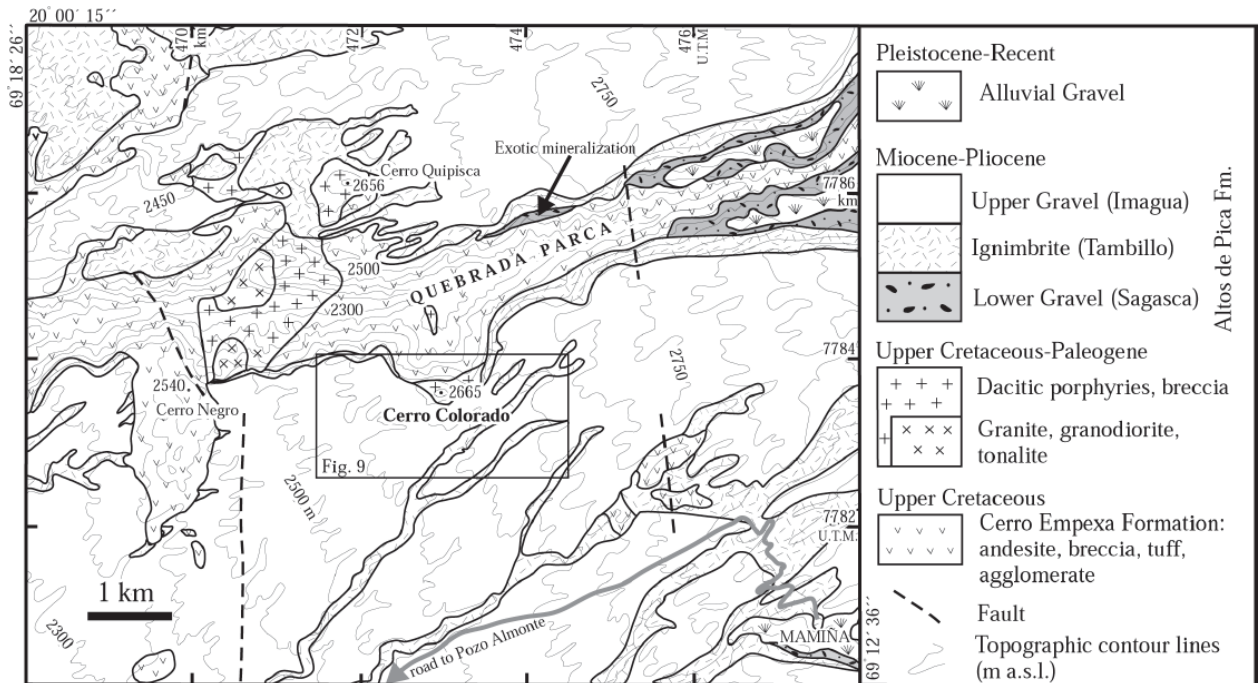


**Figura 9: Ubicación CMCC<sup>6</sup>.**

<sup>6</sup> Referencias [1].

La geología local comprende una litología compuesta por: andesitas, las cuales se distribuyen aledañas al rajo, pórfidos dioríticos a andesíticos, pórfidos dacíticos, ignimbritas, las cuales corresponden a un depósito post-mineral compuesto de toba litocristalina y toba de ceniza lítica, y gravas correspondientes a depósitos aluviales post-mineral formados por clastos subredondeados de tamaños variables y composiciones principalmente andesíticas. Bajo las ignimbritas aparece una brecha coluvial también post-mineral que es codificada como grava.

En cuanto a las alteraciones, se encuentra presencia de alteración potásica (biotita-feldespato potásico), alteración clorítica-fílica verde (cuarzo-sericita-clorita), la cual se interpreta como una alteración transicional entre potásica y fílica, alteración fílica (cuarzo-sericita-pirita), presente de manera sobreimpuesta a la alteración potásica anterior, alteración propilítica (clorita-epidota-calcita), evidenciada en afloramientos de Quebrada Parca y en algunos sondajes en los márgenes del depósito, alteración argílica intermedia (caolín-smectita), ubicada en los niveles superiores del depósito asociada a zonas de fractura, y alteración argílica avanzada (cuarzo-alunita-arcillas-pirofilita-andalucita).



**Figura 10: Mapa geológico CMCC<sup>7</sup>.**

<sup>7</sup> Referencias [1].

El yacimiento posee las siguientes zonas minerales principales:

1. **Zona lixiviada:** Formada principalmente por óxidos de fierro y afectada por fuerte alteración supérgena de tipo argílica sobreimpuesta a la alteración hidrotermal fílica.
2. **Zona de óxidos de cobre:** Se encuentra afectada por fuerte alteración fílica y menos alteración argílica sobreimpuesta. Representa el 65% de los recursos supérgenos.
3. **Zona de sulfuros supérgenos de cobre:** Ubicada bajo la zona de óxidos de cobre. Se encuentra afectada por fuerte alteración fílica. Representa el 35% de los recursos supérgenos.
4. **Zona de transición “MSH”:** Transición entre sulfuros supérgenos y sulfuros primarios de Cu. Afectada por alteración hidrotermal fílica predominante que se sobreimpone a las alteraciones tardimagmáticas hipógenas.
5. **Zona de sulfuros primarios hipógenos:** Mineralización de sulfuros de cobnre, afectada por alteraciones tardimagmáticas predominantes con débil a moderada sobreimposición de alteración fílica.



**Figura 11: Vista isométrica CMCC (Agosto 1995) - Zonas principales<sup>8</sup>.**

<sup>8</sup> Referencias [1].

## 6.2 Descripción de la base de datos

La base de datos consiste en información de sondajes de diamantina conteniendo información acerca de las leyes de las diversas especies mineralógicas contenidas en el yacimiento, así como del tipo de mineral, abarcando la totalidad del yacimiento. Asimismo se dispone de los sólidos correspondientes al tipo de mineral (TMIN), encontrándose 8 categorías: LIX (Lixiviado), OXI (Oxidado), MSH (Zona de transición), SULF\_SEC (Sulfuros secundarios), COVER (Capa superior), HYP1A, HYP1B y HYP2 (Zonas de hipógenos).

La modelación de la categorización por tipo de mineral se considera como input base para la simulación de leyes según reunión con personal de CMCC. A su vez, se determina el elemento y la zona de interés a simular siendo esta la simulación de cobre en la bajada de leyes de la zona de óxidos presente entre las coordenadas 73,300 a 73,900 (Este); 83,200 a 83,900 (Norte) y 2,000 a 2,700 (Elevación), abarcando un área de 420,000 m<sup>2</sup> y un volumen de 294 Mm<sup>3</sup>. Si bien el interés recae en la zona de óxidos, los alcances de la simulación abarcan un volumen que contiene los 8 tipos de mineral, por lo que la simulación se realizará en cada una de estas unidades por separado ya que se consideran como poblaciones distintas. Sin embargo se debe realizar un análisis por unidad primeramente para validar el uso de los 8 tipos de mineral o fusionar ciertos tipos en caso de ser necesario.

N°	TMIN	VOLUMEN [m3]	VOLUMEN [%]	VOLUMEN <sup>9</sup> [%]
1	cover	10,549,857	3.6%	4.0%
2	hyp1a	32,793,711	11.2%	12.5%
3	hyp1b	88,413,655	30.1%	33.7%
4	hyp2	180,457	0.1%	0.1%
5	lix	53,905,316	18.3%	20.6%
6	msh	11,791,827	4.0%	4.5%
7	oxi	39,223,453	13.3%	15.0%
8	sulf_sec	25,370,670	8.6%	9.7%
9	outside	31,771,053	10.8%	-
	<b>TOTAL</b>	<b>294,000,000</b>	<b>100.0%</b>	<b>100.0%</b>

**Tabla 15: Volumen por tipo de mineral.**

---

<sup>9</sup> No considera volumen de aire.

La Tabla 15 muestra el volumen de cada tipo de mineral. Es posible apreciar que el hyp2 corresponde sólo a un 0.1% de los datos, y más aún, su ubicación es en el vértice inferior del cubo y no distribuido a lo largo del sector del yacimiento a simular, tal como se puede apreciar en la Figura 12. Por esta razón no se considera este tipo de mineral en el proceso de simulación.

Una vez importado los sólidos en el software GEMS se preparan las correspondientes tablas de sondajes considerando como elemento de interés el cobre total y como unidad de estimación el tipo de mineral. De esta manera se importan un total de 2,135 sondajes de diamantina, de los cuales 324 de ellos se seleccionan como sondajes dentro del volumen de simulación utilizando como criterio que el collar del sondaje esté dentro del área comprendida entre las coordenadas 73,300-73,900 Este y 83,200-83,900 Norte.

Dado que los sondajes poseen soporte variable, se realiza un proceso de compositación a 4 metros con el fin de poseer muestras a igual soporte. Esto genera una base de datos de 324 sondajes conteniendo un total de 8,433 puntos con ley de cobre total informada, los cuales están distribuidos en las distintas unidades de estimación según lo muestra la Tabla 16.

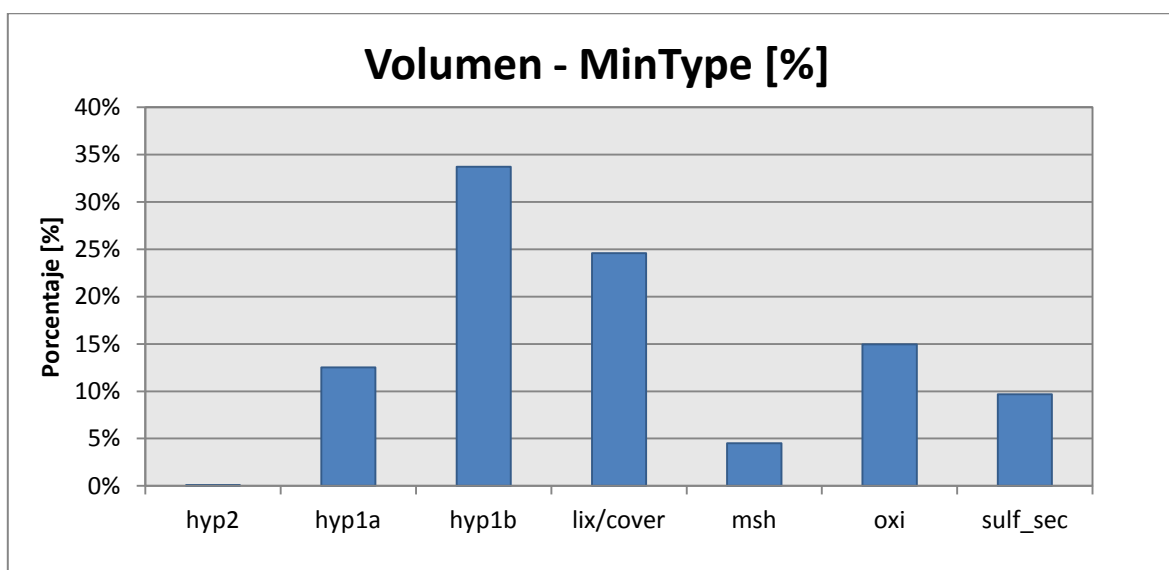
N°	Tipo de mineral	Datos
1	cover	8
2	hyp1a	953
3	hyp1b	813
4	hyp2	0
5	lix	1,567
6	msh	1,064
7	oxi	2,326
8	sulf_sec	1,702
9	outside	0
	<b>TOTAL</b>	<b>8,433</b>

**Tabla 16: Datos de ley de CuT por tipo de mineral.**

Es posible apreciar de esta tabla no sólo el hecho de la imposibilidad de modelar la unidad de estimación correspondiente a HYP2 debido a que no está informada, sino que además, la casi nula presencia de datos en el tipo de mineral COVER. De esta manera, y para efectos del estudio, se adiciona la unidad de estimación COVER a la unidad de estimación LIX, por lo que las nuevas tablas tanto de volumen de tipo de mineral como de compósitos quedan de la siguiente manera.

N°	TMIN	VOLUMEN [m3]	VOLUMEN MinType [%]	VOLUMEN [Mm3]	TONELAJE [Mton]
0	hyp2	180,457	0.1%	0.2	0.5
1	hyp1a	32,793,711	12.5%	32.8	88.5
2	hyp1b	88,413,655	33.7%	88.4	238.7
3	lix/cover	64,455,173	24.6%	64.5	174.0
4	msh	11,791,827	4.5%	11.8	31.8
5	oxi	39,223,453	15.0%	39.2	105.9
6	sulf_sec	25,370,670	9.7%	25.4	68.5

**Tabla 17: Volumen por tipo de mineral – Recodificación.**



**Gráfico 31: Volumen por tipo de mineral – Recodificación.**

N°	TMIN	N° DATA
0	hyp2	0
1	hyp1a	953
2	hyp1b	813
3	lix/cover	1,575
4	msh/cc018	1,064
5	oxi	2,326
6	sulf_sec	1,702
	<b>TOTAL</b>	<b>8,433</b>

**Tabla 18: Datos de ley de CuT por tipo de mineral – Recodificación.**



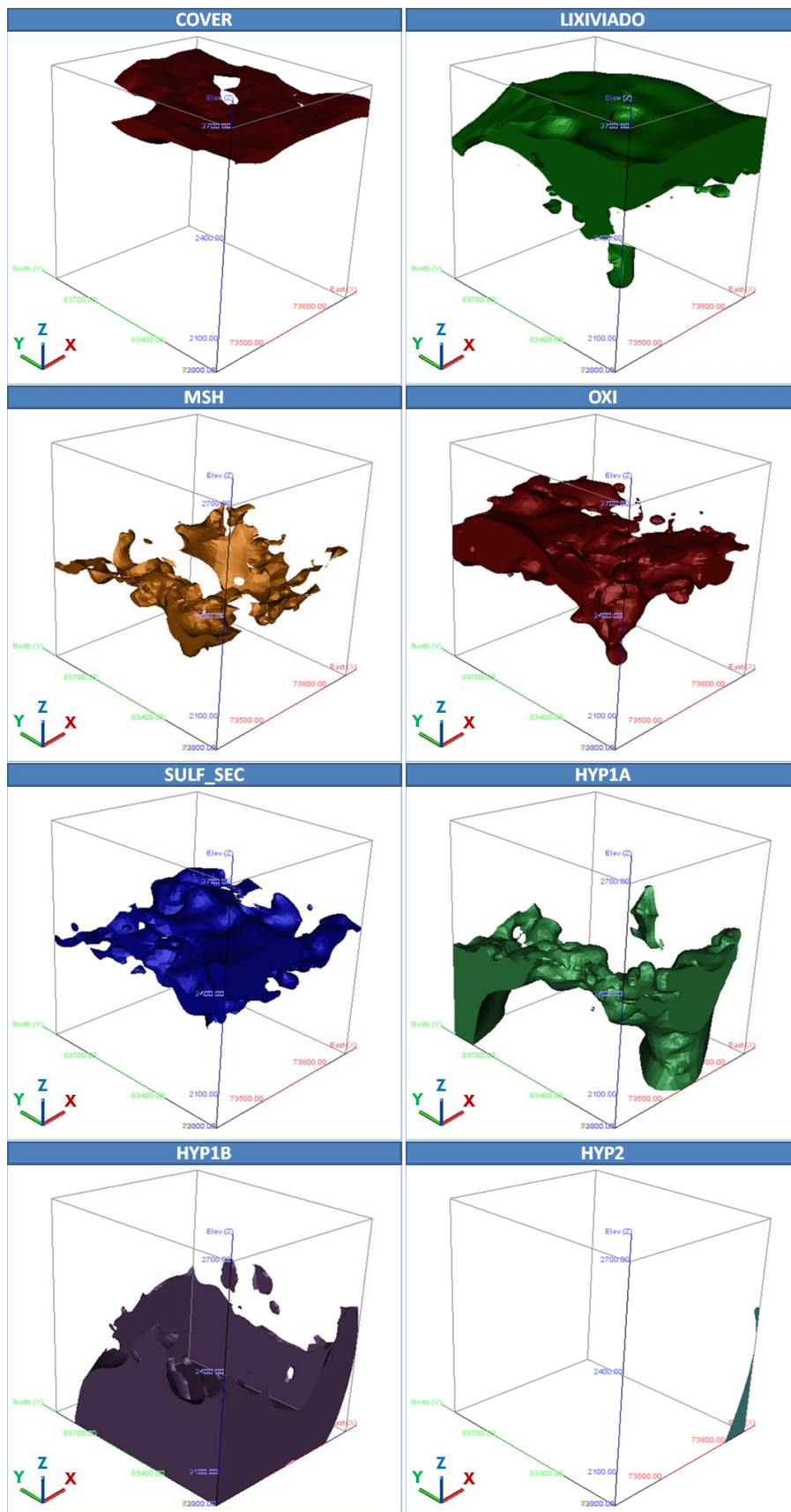


Figura 12: Isométricos tipo de mineral. Dimensiones: 600x700x600 (X,Y,Z)

En la Tabla 17 se agrega además de los volúmenes recodificados, una columna con un tonelaje referencial utilizando una densidad de 2.7 [ton/m<sup>3</sup>]. Es posible apreciar que el tipo de mineral predominante corresponde a HYP1B con un 33.7% en volumen, seguido por la fusión LIX-COVER (que en adelante sólo se llamará LIX) con un 24.6% y en tercer lugar la unidad de OXI con un 15%. Juntas, estas tres unidades de estimación corresponden al 73.3% del volumen a estimar. De esta manera se define una base de datos de seis unidades de estimación, 324 sondajes y 8,433 datos compositados a 4 metros sobre la cual se realiza el estudio exploratorio de los datos. Un análisis de contacto presentado en Anexo D descarta la presencia de contactos blando por lo cual se consideran sólo los datos presentes dentro de cada unidad de estimación como datos condicionantes para la simulación de las mismas.

### **6.3 Estudio exploratorio de los datos**

Cada uno de los estadísticos y análisis presentados se realiza sobre cada unidad de estimación, las cuales serán mostradas en adelante por su código de identificación, a saber: 1-HYP1A., 2-HYP1B, 3-LIX, 4-MSH, 5-OXI y 6-SULF\_SEC. Se hablará de unidades de estimación para referirse a la categorización por tipo de mineral.

Se presenta a continuación, en cada una de las secciones respectivas, el estudio exploratorio de los datos, abarcando sus estadísticas básicas, vistas isométricas y nubes direccionales (el resto de las estadísticas se presenta en Anexo D, E y F).

Una vez realizado este primer análisis, se realiza el desagrupamiento de los datos y se vuelven a presentar sus estadísticos calculados ahora sobre los datos desagrupados (el resto de las estadísticas se presentan en Anexo G).

#### *6.3.1 Estadísticas básicas*

El Gráfico 32 muestra las estadísticas básicas de cada una de las unidades de estimación. Es posible apreciar de este gráfico que todos los valores de ley de cobre total están dentro de un rango normal de valores (entre 0.0% y 6.7%), es decir, a primera vista no se sospecha de datos aberrantes, sin embargo se hará revisión de los histogramas para corroborar esto.

### CuT by Tmin - Arithmetic

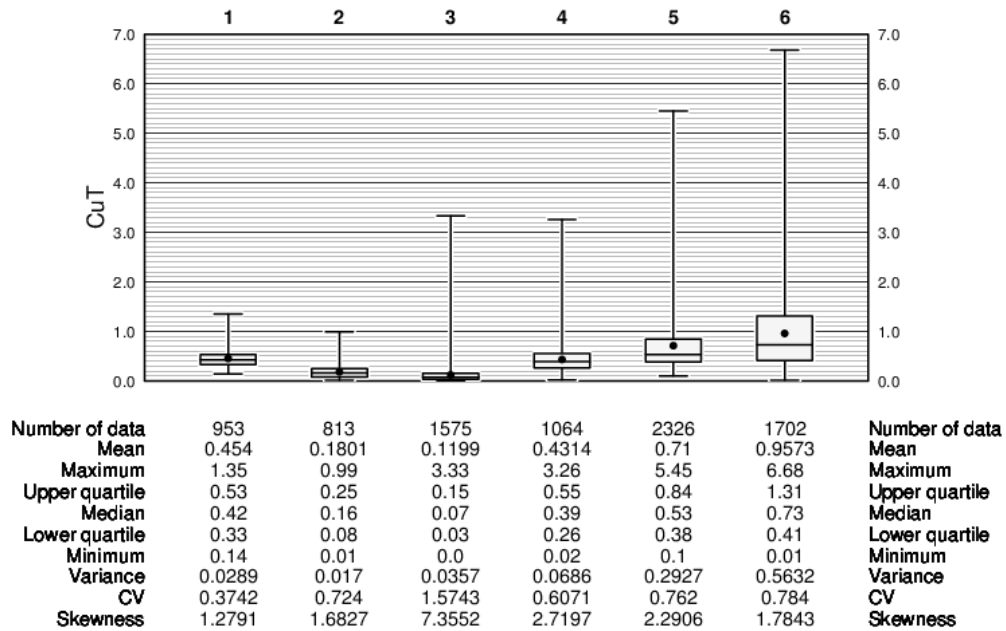


Gráfico 32: Blox plot de CuT por unidad de estimación – Escala aritmética.

### CuT by Tmin - Log Scale

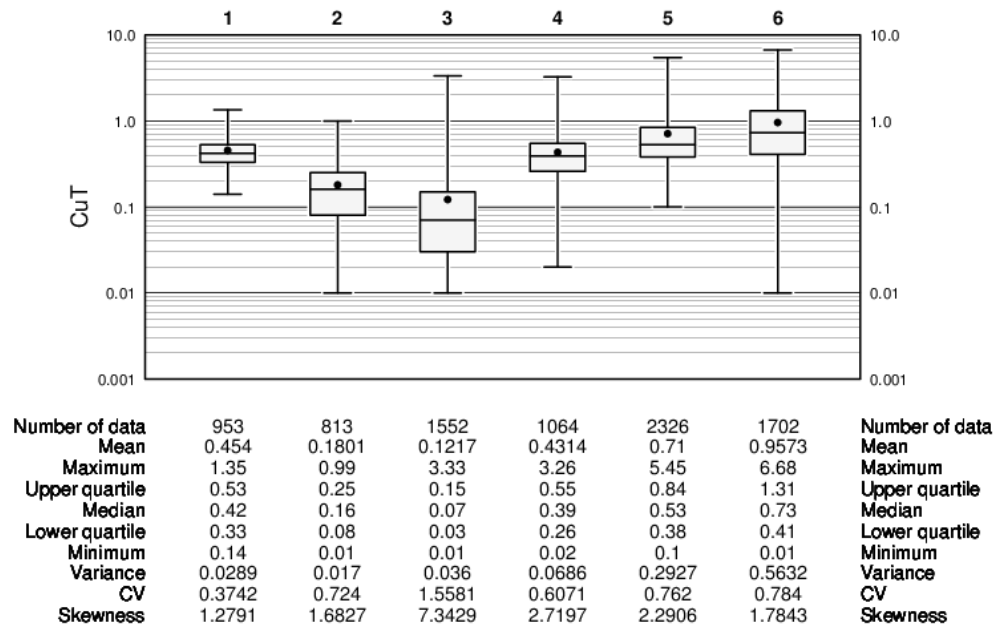


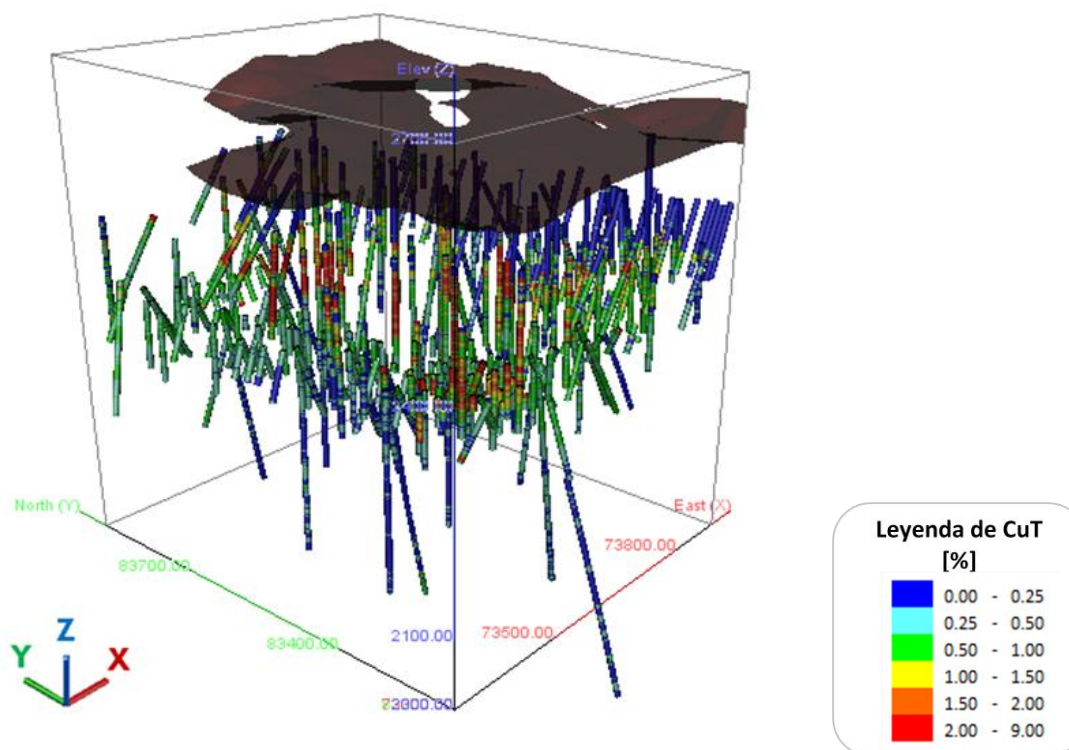
Gráfico 33: Blox plot de CuT por unidad de estimación – Escala logarítmica.

El orden de las unidades de estimación en cuanto a su media de CuT de manera ascendente es: 3-LIX, 2-HYP1B, 4-MSH, 1-HYP1A, 5-OXI y 6-SULF\_SEC. Considerando ahora la varianza de CuT de manera ascendente el orden cambia a: 2-HYP1B, 1-HYP1A, 3-LIX, 4-MSH, 5-OXI y 6-SULF\_SEC. Se aprecia de esta manera que el orden no cambia en las últimas tres unidades de estimación, no así en las primeras tres. Se definen de esta forma 3 tipos de unidades de estimación en cuanto a la media de sus leyes: unidades de baja ley, correspondientes a 3-LIX y 2-HYP1B; unidades de ley media correspondientes a 1-HYP1A y 4-MSH; y unidades de alta ley correspondientes a 5-OXI y 6-SULF\_SEC.

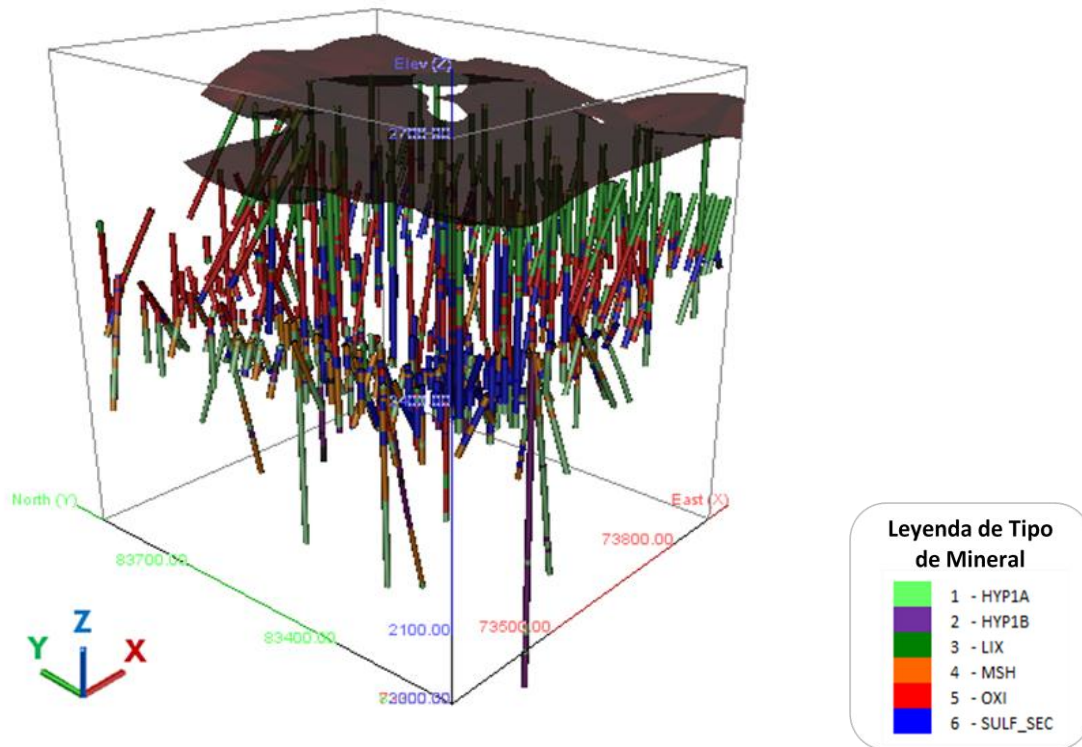
Si bien los boxplot de ambos hipógenos (unidades 1:HYP1A y 2:HYP1B) son relativamente similares entre ellos, estas unidades no pueden ser tratadas como la misma población dada su notoria diferencia en las medias de CuT.

### 6.3.2 Vistas isométricas

Se presenta en la Figura 13 y en la Figura 14 las vistas isométricas de los sondajes compositados a 4 metros mostrando la ley de cobre total y el tipo de mineral respectivamente.



**Figura 13: Vista isométrica – Ley de CuT.**

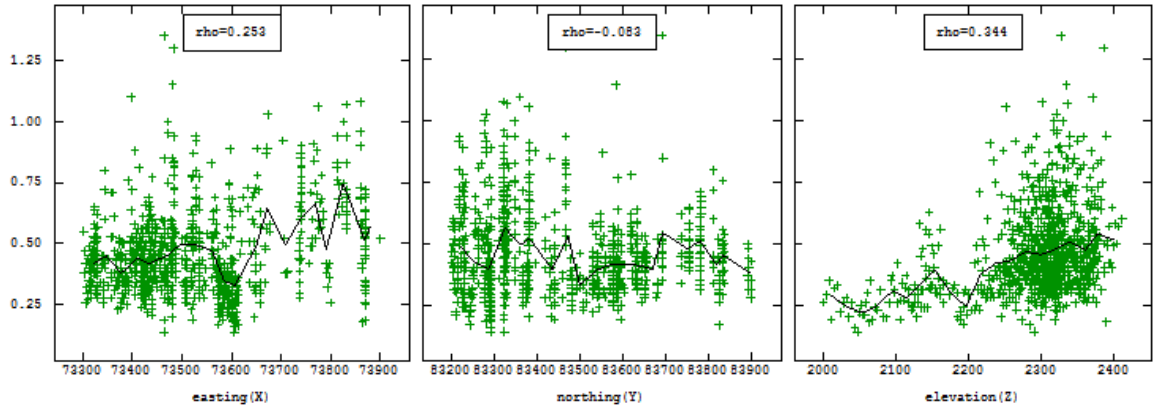


**Figura 14: Vista isométrica – Tipo de mineral.**

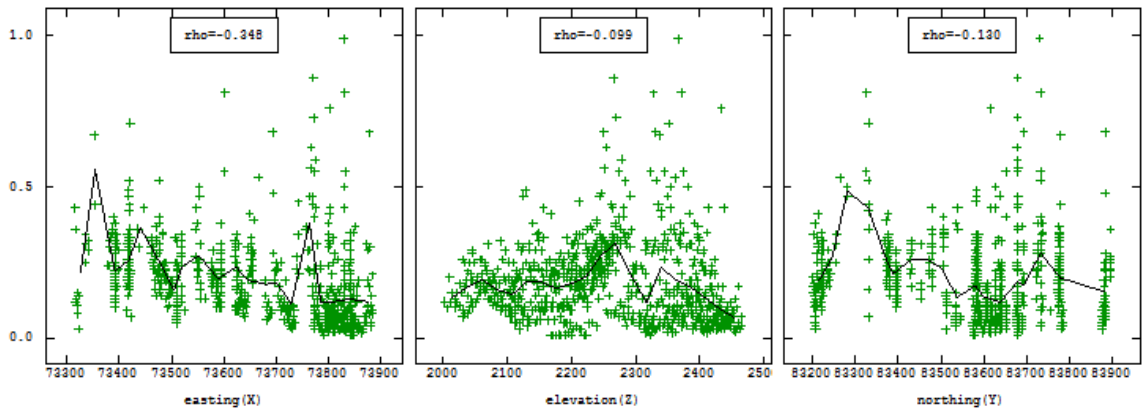
En el Anexo D se presenta la proyección de los sondajes en los planos XY (planos), XZ (perfil EO) e YZ (perfil NS) donde cada uno de los mapas posee una escala local respecto a la unidad de estimación. En la proyección en planta (Figura 15 Anexo D) no se muestra ninguna tendencia sistemática de las leyes de cobre a lo largo de las coordenadas X e Y, aunque es posible apreciar un muestreo preferencial en ciertas zonas, lo que refuerza la necesidad de un posterior desagrupamiento. En las vistas en perfil (Figura 16 y Figura 17 Anexo D) se muestra una cierta tendencia de la ley de cobre total a la baja a medida que se profundiza en la cota, lo cual es apreciable de manera más evidente en las unidades 1, 4 y 5. La unidad 6 presenta tal tendencia pero de manera local, específicamente en el sector SW.

### 6.3.3 Nubes direccionales

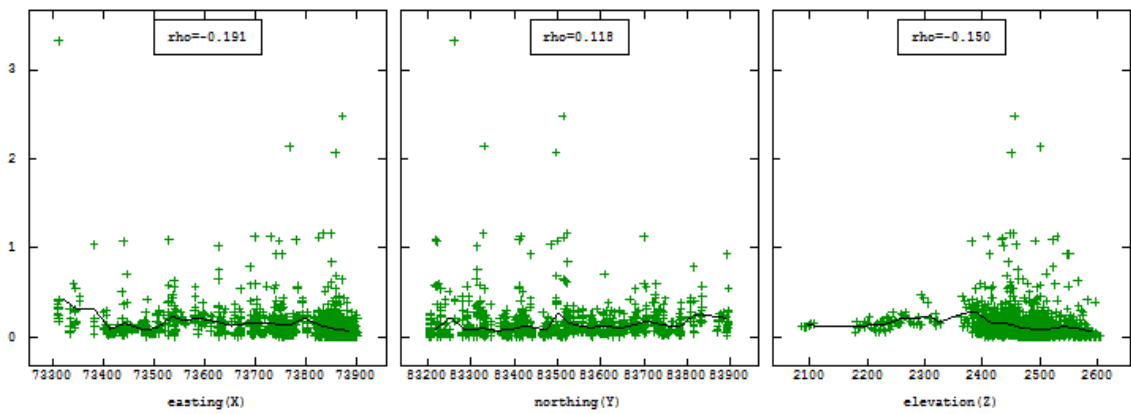
Las nubes direccionales permiten apreciar la existencia de deriva en cada una de las tres direcciones. De esta manera es posible corroborar de los siguientes gráficos una tendencia sistemática de los valores de la ley de cobre total al alza a medida que se aumenta en la coordenada Z en las unidades de estimación 1, 4, 5 y 6. Además es posible notar la no existencia de tendencia a lo largo de las coordenadas X e Y para ninguna de las 6 unidades de estimación.



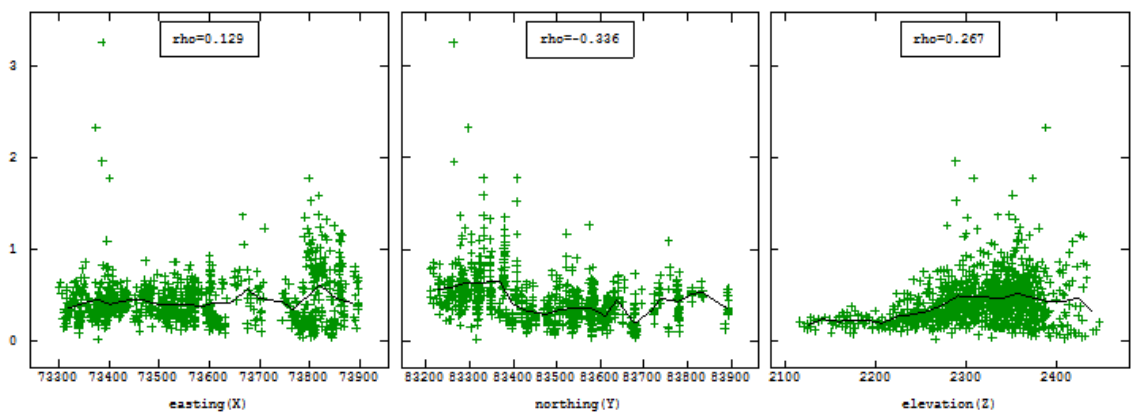
**Gráfico 34: Nubes direccionales – UE 1. (Izq: X - Cen: Y - Der: Z).**



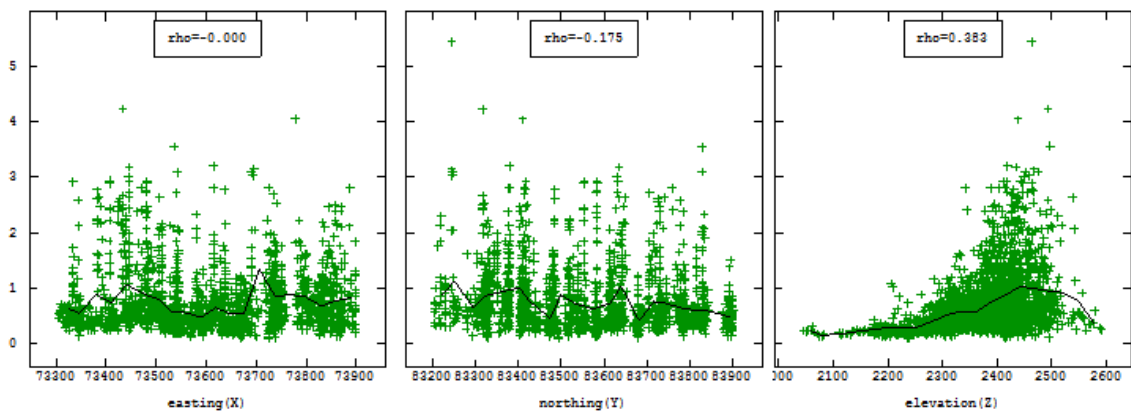
**Gráfico 35: Nubes direccionales – UE 2. (Izq: X - Cen: Y - Der: Z).**



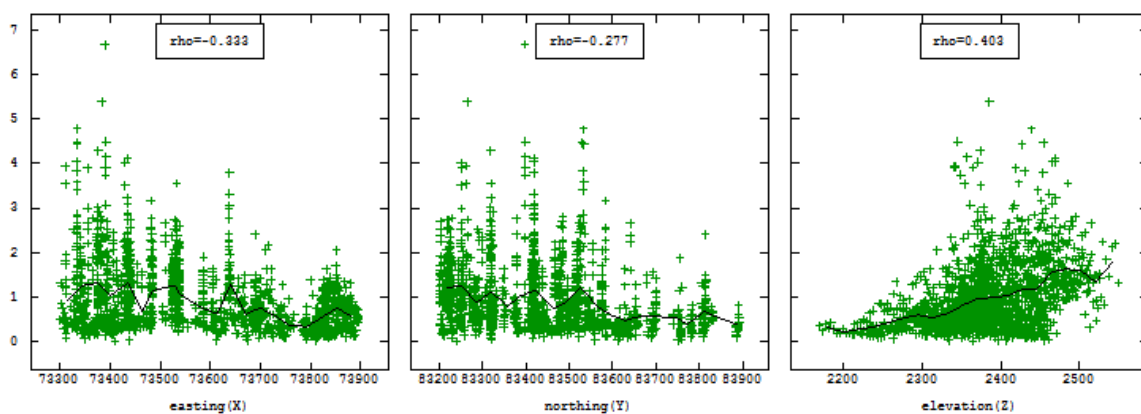
**Gráfico 36: Nubes direccionales – UE 3. (Izq: X - Cen: Y - Der: Z).**



**Gráfico 37: Nubes direccionales – UE 4. (Izq: X - Cen: Y - Der: Z).**



**Gráfico 38: Nubes direccionales – UE 5. (Izq: X - Cen: Y - Der: Z).**



**Gráfico 39: Nubes direccionales – UE 6. (Izq: X - Cen: Y - Der: Z).**

Con todos estos antecedentes se procede a la siguiente etapa del estudio consistente en un desagrupamiento de los datos debido a un muestreo irregular.

#### 6.4 Estudio exploratorio de los datos desagrupados

Para la etapa de desagrupamiento se utiliza el método de celdas, el cual subdivide el dominio en celdas de igual tamaño, todas las cuales tienen un mismo ponderador. El tamaño de celda definido es de  $100 \times 100 \times 25 \text{ m}^3$ , el cual se determina en Anexo G.

Se presentan a continuación las principales estadísticas de los datos desagrupados (histogramas y gráficos de probabilidad lognormal se presentan en Anexo G) con el fin de destacar las diferencias entre las mismas estadísticas de las muestras sin desagrupar. Mediante el Gráfico 40, que muestra las estadísticas básicas de los datos desagrupados, es posible calcular la diferencia de las medias con respecto a la media de las muestras sin desagrupar y su diferencia porcentual tomando como base estas últimas. La mayor variación ocurre en las unidades de estimación 2 y 3 con un 13.4% y 38.6% de variación respectivamente. Esto se debe básicamente a que son unidades de baja ley media por lo cual pequeñas variaciones absolutas implican grandes variaciones porcentuales. En términos de diferencia absolutas, estas no superan el 0,05% en ninguna de las seis unidades de estimación.

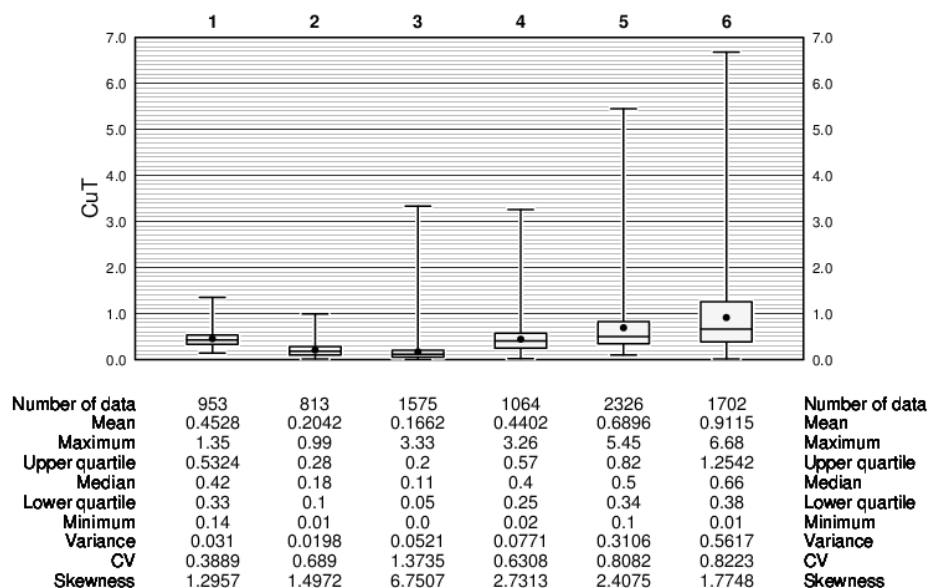
MEDIA	UNIDAD	UE-1	UE-2	UE-3	UE-4	UE-5	UE-6
Muestras	CuT	0.454	0.180	0.120	0.431	0.710	0.957
Desagrupado	CuT	0.453	0.204	0.166	0.440	0.690	0.912
Diferencia Absoluta	CuT	0.00	0.02	0.05	0.01	0.02	0.05
Diferencia Relativa	[%]	0.26	13.38	38.62	2.04	2.87	4.78

**Tabla 19: Media muestras y media desagrupada.**



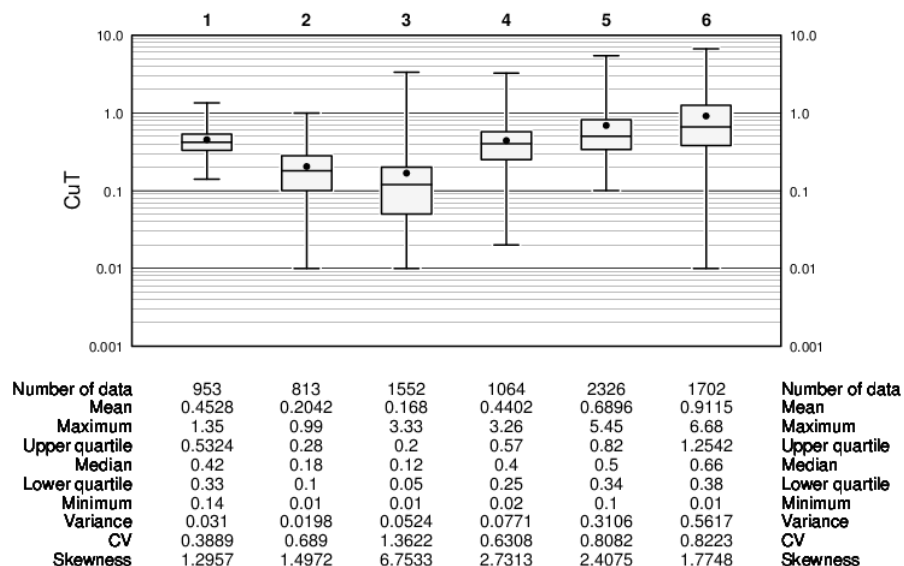
En general es posible apreciar del Gráfico 40, Tabla 19 y gráficos en Anexo G que las diferencias tanto en estadísticas básicas como en la distribución de los datos no cambian significativamente.

**CuT by Tmin - Arithmetic - Declustered**



**Gráfico 40: Blox plot de CuT desagrupado por unidad de estimación – Escala aritmética.**

**CuT by Tmin - Log Scale - Declustered**



**Gráfico 41: Blox plot de CuT desagrupado por unidad de estimación – Escala logarítmica.**

## 6.5 Modelo de bloques y de roca

Una vez determinado el volumen a simular y las estadísticas básicas, se procede a definir el modelo de bloques y roca que define tanto el soporte como las unidades de estimación a simular.

Coordenada	Tamaño de bloque	Número de bloques	Mínimo (Centroide)	Máximo (Centroide)
Este (X)	10	60	73,305	73,895
Norte (Y)	10	70	83,205	83,895
Elevación (Z)	10	66	2,005	2,655

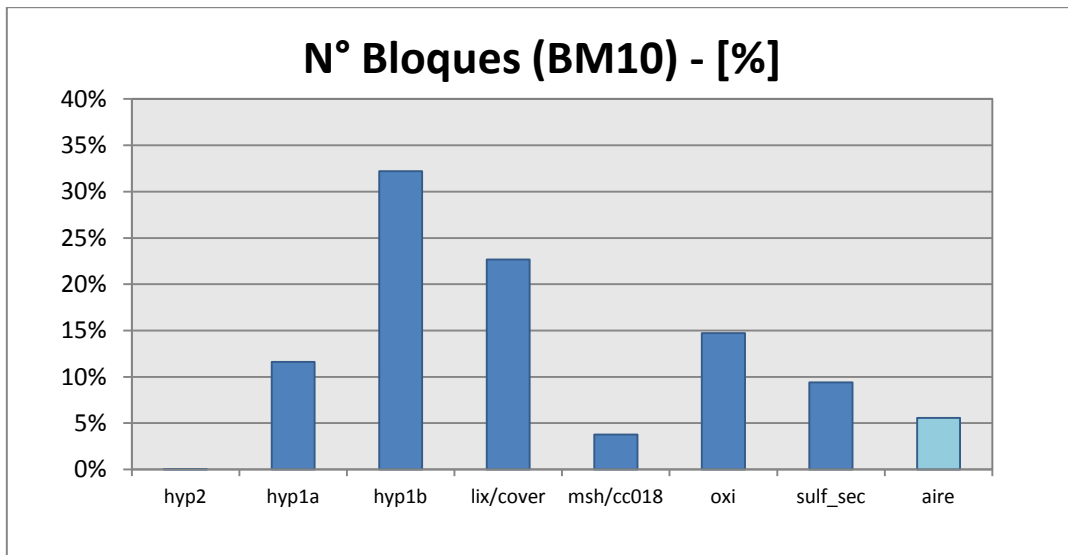
**Tabla 20: Modelo de bloques 10x10x10 m<sup>3</sup>.**

Se define un modelo de bloques de 10x10x10 m<sup>3</sup> cuyos parámetros se presentan en la Tabla 20. De esta manera el modelo de bloques consta de 60 bloques en la coordenada X, 70 en la coordenada Y y 66 en la coordenada Z, lo cual implica 277,200 bloques a simular involucrando un volumen de 600x700x660 m<sup>3</sup>. Se debe notar que se realizó una reducción del alcance en el eje Z dado que sobre la cota 2,660 el modelo corresponde a bloques de aire.

Una vez definido el modelo de bloques se inicializa el modelo de rocas (unidades de estimación) sobre dicho soporte considerando el siguiente orden de procedencia: Aire (0) – HYP2 (7) – LIX (3) – MSH (4) – HYP1A (1) – HYP1B (2) – SULFSEC (6) – OXI (5). Este orden tiene relevancia en el caso de bloques en los contactos ya que la última UE inicializada con código de roca tendrá preferencia en estos casos. Se considera un 33% como mínimo porcentaje de integración del bloque dentro de la correspondiente UE para la inicialización.

UE	Número de bloques	[%]
0	15,438	5.6
1	32,166	11.6
2	89,244	32.2
3	62,860	22.7
4	10,447	3.8
5	40,825	14.7
6	26,067	9.4
7	153	0.1
<b>Total</b>	<b>277,200</b>	<b>100.00</b>

**Tabla 21: Número de bloques por UE.**



**Gráfico 42: Volumen por tipo de mineral – Modelo de roca.**

En la Tabla 21 y en el Gráfico 42 se muestra el número de bloques por UE que definen el modelo de rocas y el volumen correspondiente.

## 6.6 Modelo de deriva de la media de CuT

Tal como se explica en la sección de alcances, no se estudiará la creación del modelo de deriva. Sin embargo se debe generar este para utilizar como input en el algoritmo generado, el cual debe estar totalmente informado dentro del volumen a simular, es decir, cada uno de los 277,200 bloques debe tener un valor para la deriva.

Para generar el modelo de deriva se utiliza el ejecutable 'KT3D' de GSLib con la opción de estimar tendencia. Lo que realiza esta función es un tipo de Kriging que calcula la tendencia en cada punto cuando se utiliza un variograma de gran alcance y alto efecto pepa. De esta manera se utiliza el siguiente modelo variográfico para cada una de las UE con el fin de cumplir con las condiciones de un variograma de gran alcance y con alto efecto pepa. Si bien esta elección es arbitraria, se realiza un proceso de validación de la deriva generada con este modelo para efectos de continuar con las siguientes etapas del estudio.

$$\gamma(\mathbf{h}) = 0.4 P(\mathbf{h}) + 0.6 Esf_{(350,300,AVert)}(\mathbf{h})$$

Este modelo considera un efecto pepa igual a 0.4, con una meseta en 1.0 y un alcance de 350 m y 300 m en X e Y respectivamente, variando el alcance en la vertical por UE según la Tabla 22.

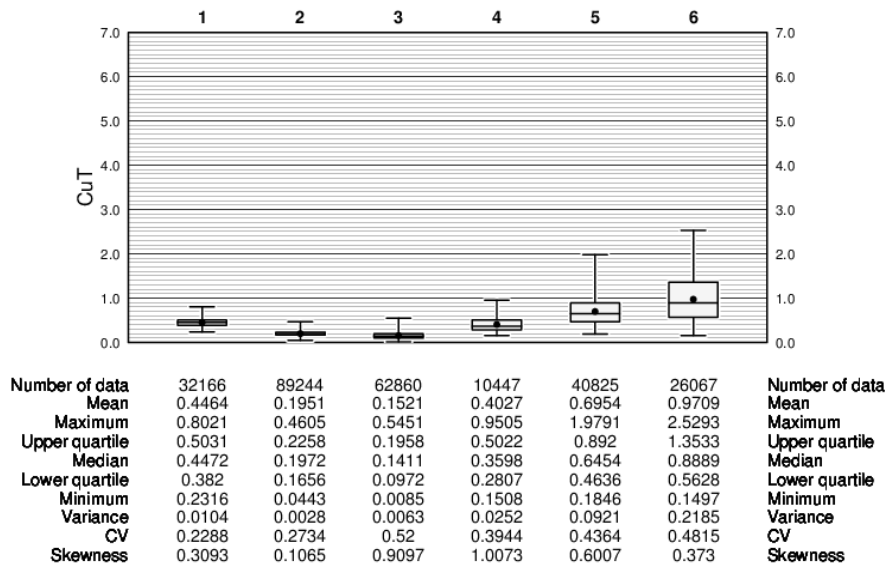
UE	Alcance Vertical	Radio de Búsqueda	Máximo de datos
1	200	350	40
2	200	350	40
3	250	300	56
4	150	300	56
5	250	300	40
6	200	350	56

**Tabla 22: Parámetros estimación de deriva por UE.**

Se presenta además en esta tabla, otros parámetros que varían por UE en la estimación de la deriva como lo es el radio de búsqueda (igual en todas direcciones) y el número máximo de datos. Dado que no está dentro de los alcances el estudio de la generación de esta deriva, basta con validar el modelo para poder utilizarse como input para el algoritmo de simulación con rechazo aún cuando los parámetros de la Tabla 22 pudiesen ser en cierta medida arbitrarios.

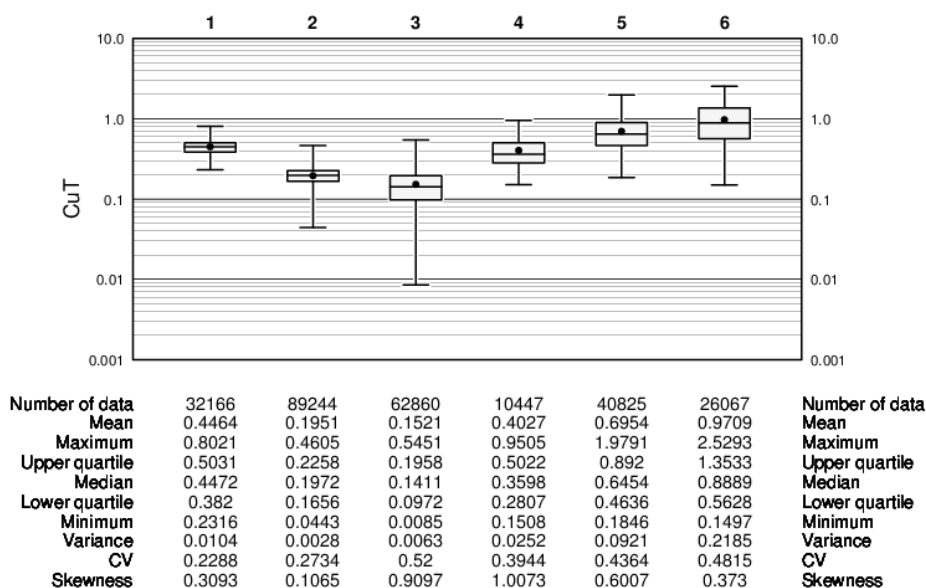
Para efectos de validación se chequea por un lado que la media global por UE del modelo de deriva sea aproximadamente igual a la media global de los datos desagrupados por UE, y por otro lado que la media local del modelo de deriva por coordenada siga una tendencia dada por la misma media local de los datos originales.

**CuT by Tmin - Drift - Arithmetic**



**Gráfico 43: Blox plot modelo de deriva – Escala aritmética.**

### CuT by Tmin - Drift - Log Scale



**Gráfico 44: Blox plot modelo de deriva – Escala logarítmica.**

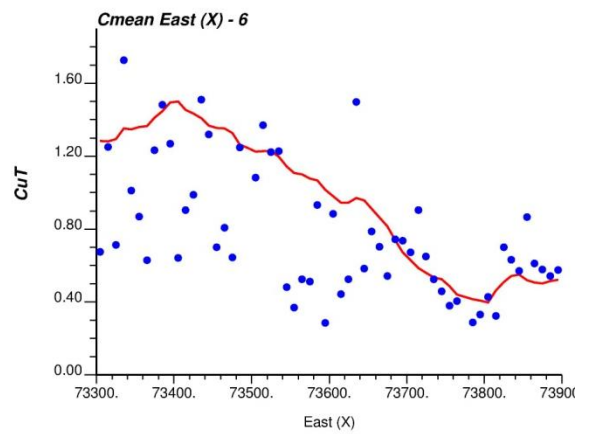
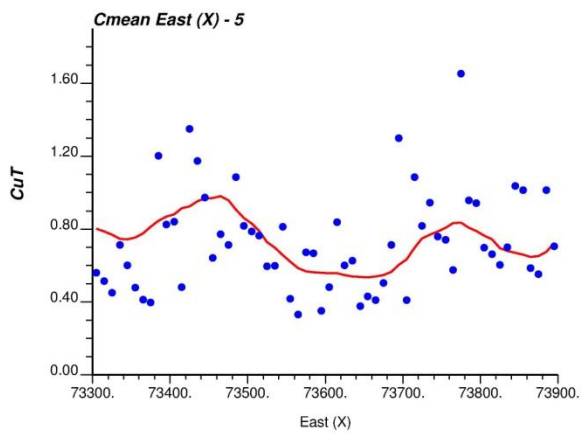
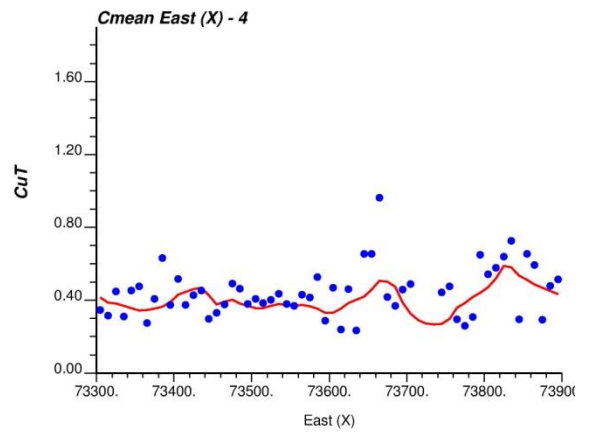
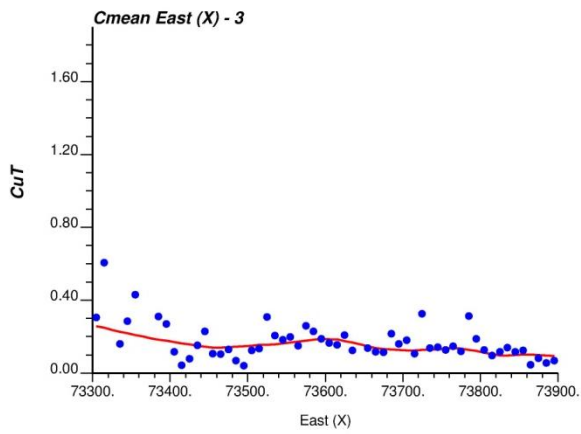
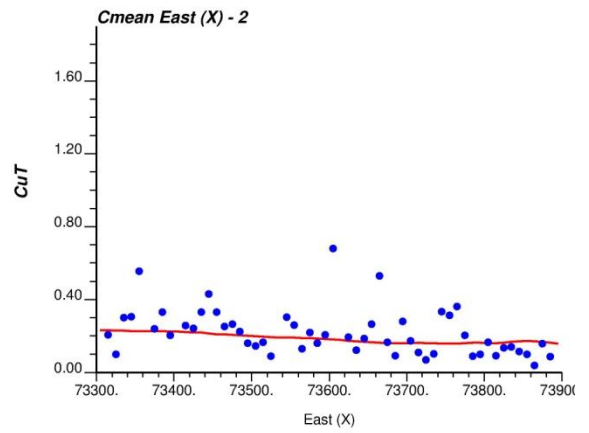
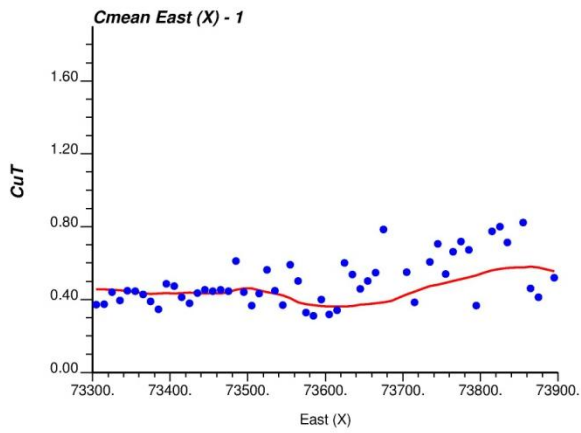
El Gráfico 43 muestra las estadísticas básicas del modelo de deriva en cada UE. Es posible apreciar una clara disminución de la varianza de los datos, así como una reducción del máximo y un aumento del mínimo, en síntesis, un estrechamiento del histograma de los datos. Mediante dichas estadísticas se calcula la diferencia de la media de los datos desagrupados y la media del modelo de deriva lo cual se muestra en la Tabla 23.

UE	Media desagrupada	Media deriva	Diferencia absoluta	Diferencia relativa [%]
1	0.4528	0.4464	0.01	1%
2	0.2042	0.1951	0.01	4%
3	0.1662	0.1521	0.01	8%
4	0.4402	0.4027	0.04	9%
5	0.6896	0.6954	0.01	1%
6	0.9115	0.9709	0.06	7%

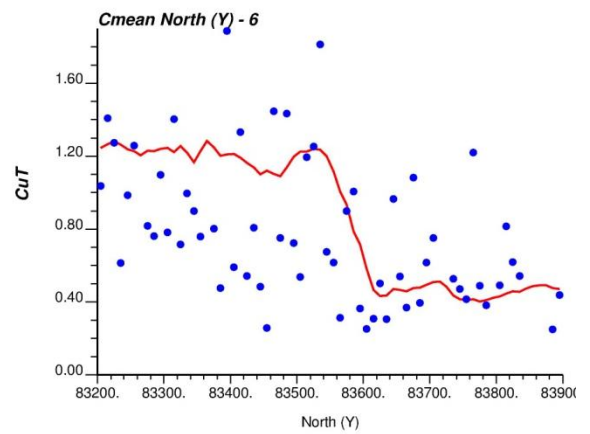
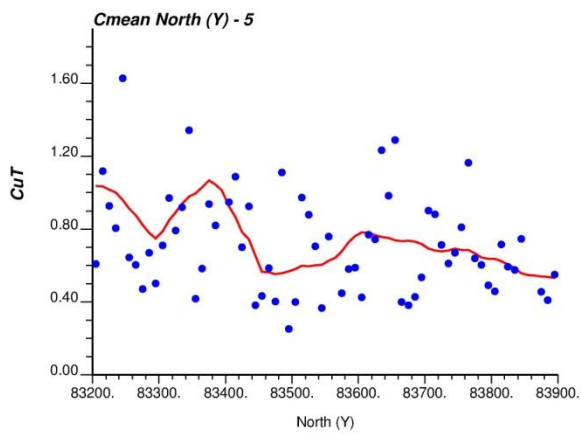
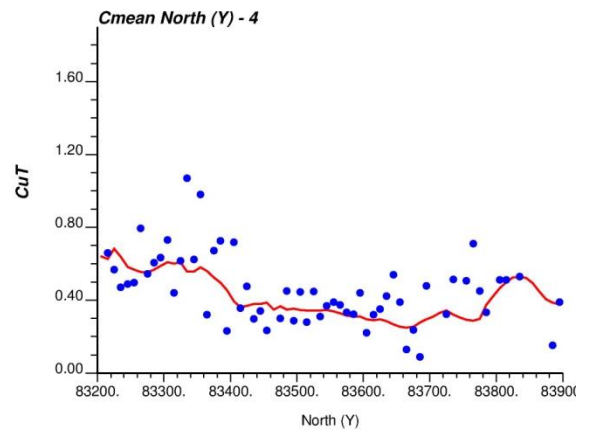
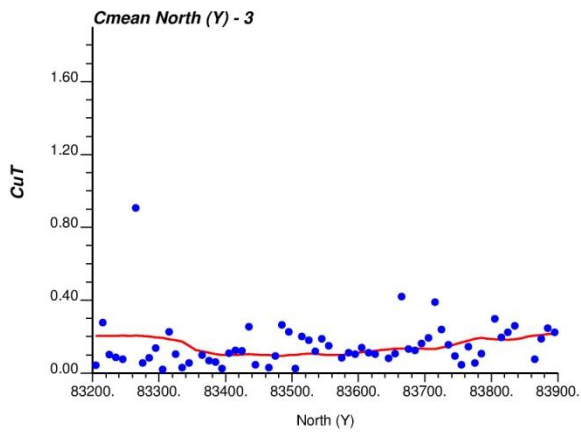
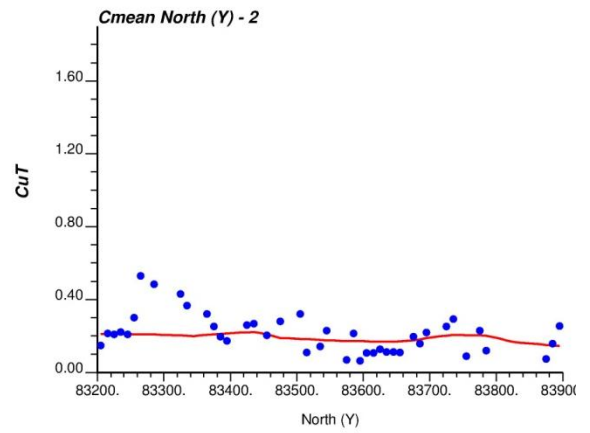
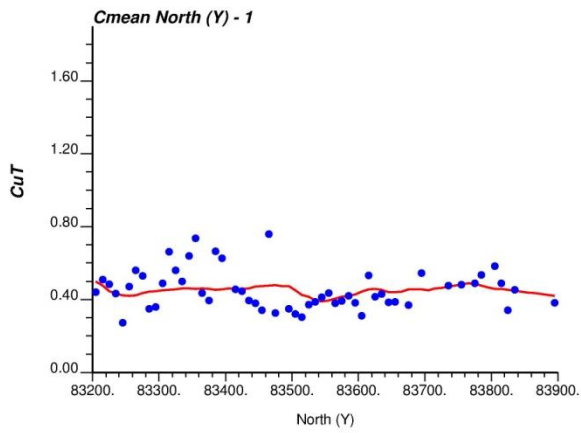
**Tabla 23: Diferencia media desagrupada vs media deriva.**

Se valida con esta tabla el primer criterio de validación dado que las diferencias de medias no superan el 0.06 en valor absoluto ni el 10% en valor porcentual con respecto a la media desagrupada.

Para revisar el segundo criterio de validación del modelo de deriva se calcula la media local con un paso de 10 metros dentro del modelo de bloques, es decir, 60 nodos en el eje X, 70 en el eje Y y 66 en el eje Z. Se realiza el cálculo tanto para los compósitos (puntos azules) como para la deriva (línea roja) para cada una de las 3 coordenadas mediante un ejecutable con formato GSLib (*'cmean'*) creado para tales efectos. Los resultados se muestran en el Gráfico 45, Gráfico 46 y Gráfico 47, los cuales validan el segundo criterio de validación en el sentido que la tendencia seguida por el modelo de deriva viene dada por la tendencia de los datos originales, siendo la tendencia del modelo más suave que la de los datos originales.

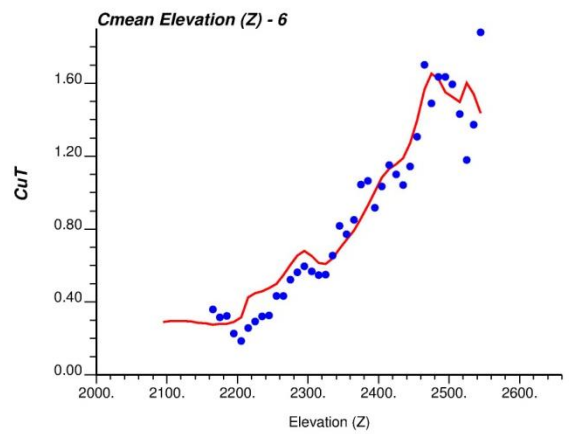
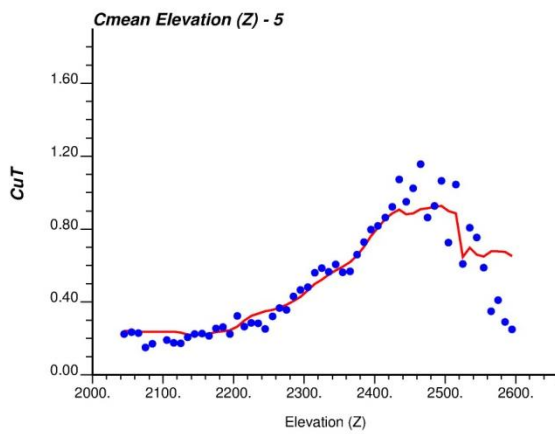
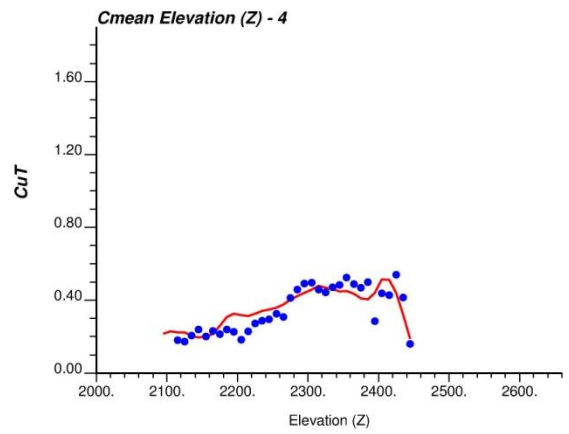
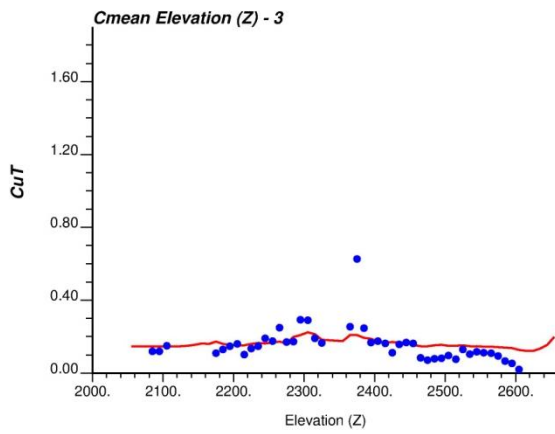
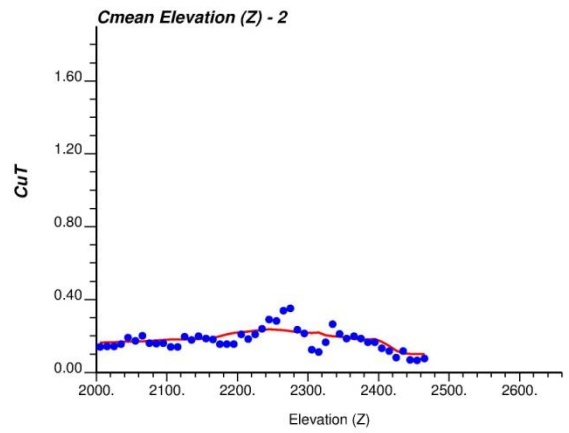
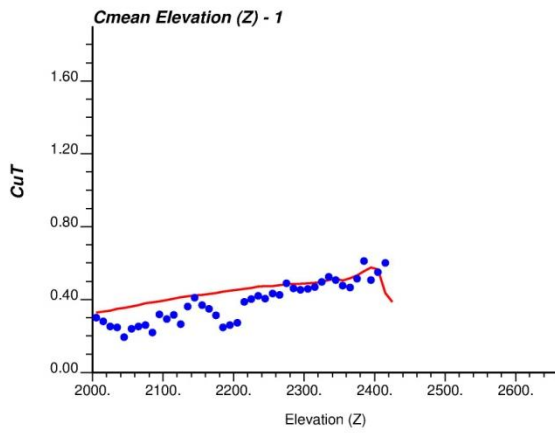


**Gráfico 45: Media local – Este (X) en cada UE. (Azul: Compósitos – Rojo: Deriva).**



**Gráfico 46: Media local – Norte (Y) en cada UE. (Azul: Compósitos – Rojo: Deriva).**

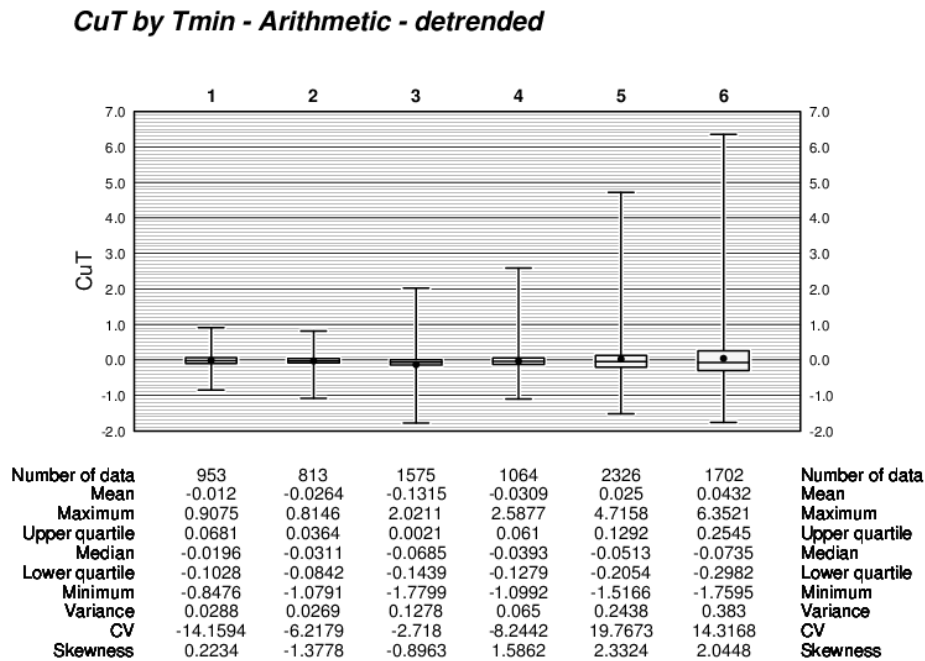




**Gráfico 47: Media local – Elevación (Z) en cada UE. (Azul: Compósitos – Rojo: Deriva).**

## 6.7 Cálculo de residuos

Una vez determinado el modelo de deriva es posible realizar el cálculo de residuos de los datos compositados realizando un proceso de sustracción de la deriva (*detrending*). Para dicho cálculo se utiliza la ecuación (10), la cual asume que la variable regionalizada puede ser descompuesta en su componente residual ( $R(\mathbf{u})$ , aleatoria) y su deriva ( $m(\mathbf{u})$ , determinística).



**Gráfico 48: Blox plot residuos – Escala aritmética.**

Las estadísticas de los residuos son presentadas en el Gráfico 48, donde se puede ver que los valores tienen media cercana a 0, lo cual es lo esperado al quitar su componente de deriva. Mediante estos valores es posible realizar el proceso de Kriging de residuos para lo cual primero se debe realizar el análisis variográfico respectivo.

Se debe considerar que tanto los datos de los compósitos originales como de los residuos utilizan los mismo ponderadores de desagrupamiento dado que estos guardan relación sólo con la distribución espacial de las muestras y no con el valor de ellas.

## 6.8 Transformación Gaussiana

El modelo de simulación Gaussiano secuencial necesita para su implementación que la distribución de la variable a simular posea una distribución Gaussiana, sin embargo, la distribución a la ley de cobre total en cada una de las unidades de estimación es lognormal como se puede apreciar en el Anexo F para el caso de los compósitos originales y en el Anexo G para el caso de los residuos, por lo cual se debe realizar un proceso de transformación Gaussiana previo al proceso de simulación.

### 6.8.1 Estadísticas

En el Gráfico 140 (Anexo H) se presentan las estadísticas de los valores normalizados de la ley desagrupada de cobre para cada una de las unidades de estimación tanto para los compósitos originales como para los residuos. Como se puede apreciar, tanto la media como la varianza toman un valor constante para cada unidad igual a 0 y a 1 respectivamente. Esto es lo esperado dado que se transforma la distribución original en una Gaussiana de media 0 y varianza 1.

### 6.8.2 Validación hipótesis multi Gaussiana

El modelo supone la hipótesis de multi-Gaussianidad, es decir, que las distribuciones de orden superior siguen siendo Gaussianas. En la práctica, la validación para distribuciones de orden mayor a dos se hace compleja por lo cual sólo se valida la hipótesis para la distribución de orden 1 y 2. Para la distribución univariable se revisa que el histograma tenga una distribución Gaussiana centrada en 0 y de varianza 1. Para la distribución bivivariable se revisa que la nube de correlación diferida dibuje elipses que tienden a la recta identidad para distancias pequeñas y a un círculo para distancias grandes.

En Anexo H se muestran los histogramas de los compósitos original y residuos, donde se aprecia en efecto una distribución normal (0,1) en cada una de las unidades de estimación, lo cual valida la hipótesis de Gaussianidad univariable. Los gráficos de correlación diferida de los compósitos originales también son presentados en Anexo H, los cuales se calculan de manera omnidireccional para tres distancias distintas: baja (de 25 m a 50 m), media (de 50 m a 75 m) y alta (de 100 m a 125 m). Dado que para cada una de las unidades de estimación se dibujan elipses que tienden a la recta identidad para distancias cortas y a un círculo para distancias grandes, se

valida la hipótesis Gaussiana bivariable. De esta manera se valida el modelo multi Gaussiano para distribuciones de orden 1 y 2 en el caso de los compósitos originales y para distribución de orden 1 en el caso de los residuos, asumiéndose su cumplimiento en distribuciones de orden mayor.

## **6.9 Análisis variográfico**

Una vez transformados los valores a su forma Gaussiana se realiza el estudio variográfico mediante el siguiente procedimiento:

1. Buscar posibles direcciones de anisotropía mediante los mapas variográficos.
2. Determinar los variogramas experimentales definiendo las direcciones en función de los resultados obtenidos de los mapas variográficos. Se chequean tres tipos de variogramas experimentales: omnidireccional, omnihorizontal y direccional.
3. Modelar el variograma experimental definido en el punto 2.

Esto permite obtener la información acerca de la relación existente entre pares de datos, es decir, información acerca de la relación bivariable de la distribución espacial de la función aleatoria, lo cual se utiliza en etapas posteriores para la implementación de la simulación Gaussiana. Se presenta en esta sección los variogramas modelados, mientras que los mapas variográficos y los variogramas experimentales se muestran en Anexo I y Anexo J respectivamente.

### *6.9.1 Variogramas modelados compósitos originales*

Para la modelación de los variogramas experimentales definidos en Anexo J se utilizan modelos anidados consistentes en un efecto pepa más uno o dos variogramas esféricos. Dado que se trata de valores con distribución normal (0,1), la meseta en cada uno de los modelos variográficos es equivalente a 1. Las ecuaciones presentadas en la Tabla 24 definen los variogramas modelados para cada UE, mientras que las rotaciones de cada dirección se muestran en la Tabla 25.

UE	Modelo variográfico
1	$0.20 P(\mathbf{h}) + 0.40 Esf_{(300,300,70)}(\mathbf{h}) + 0.40 Esf_{(300,300,220)}(\mathbf{h})$
2	$0.20 P(\mathbf{h}) + 0.80 Esf_{(105,105,105)}(\mathbf{h})$
3	$0.20 P(\mathbf{h}) + 0.40 Esf_{(350,160,130)}(\mathbf{h}) + 0.40 Esf_{(160,160,130)}(\mathbf{h})$
4	$0.20 P(\mathbf{h}) + 0.80 Esf_{(350,350,350)}(\mathbf{h})$
5	$0.20 P(\mathbf{h}) + 0.15 Esf_{(60,60,50)}(\mathbf{h}) + 0.65 Esf_{(600,600,195)}(\mathbf{h})$
6	$0.18 P(\mathbf{h}) + 0.52 Esf_{(100,100,100)}(\mathbf{h}) + 0.30 Esf_{(450,450,450)}(\mathbf{h})$

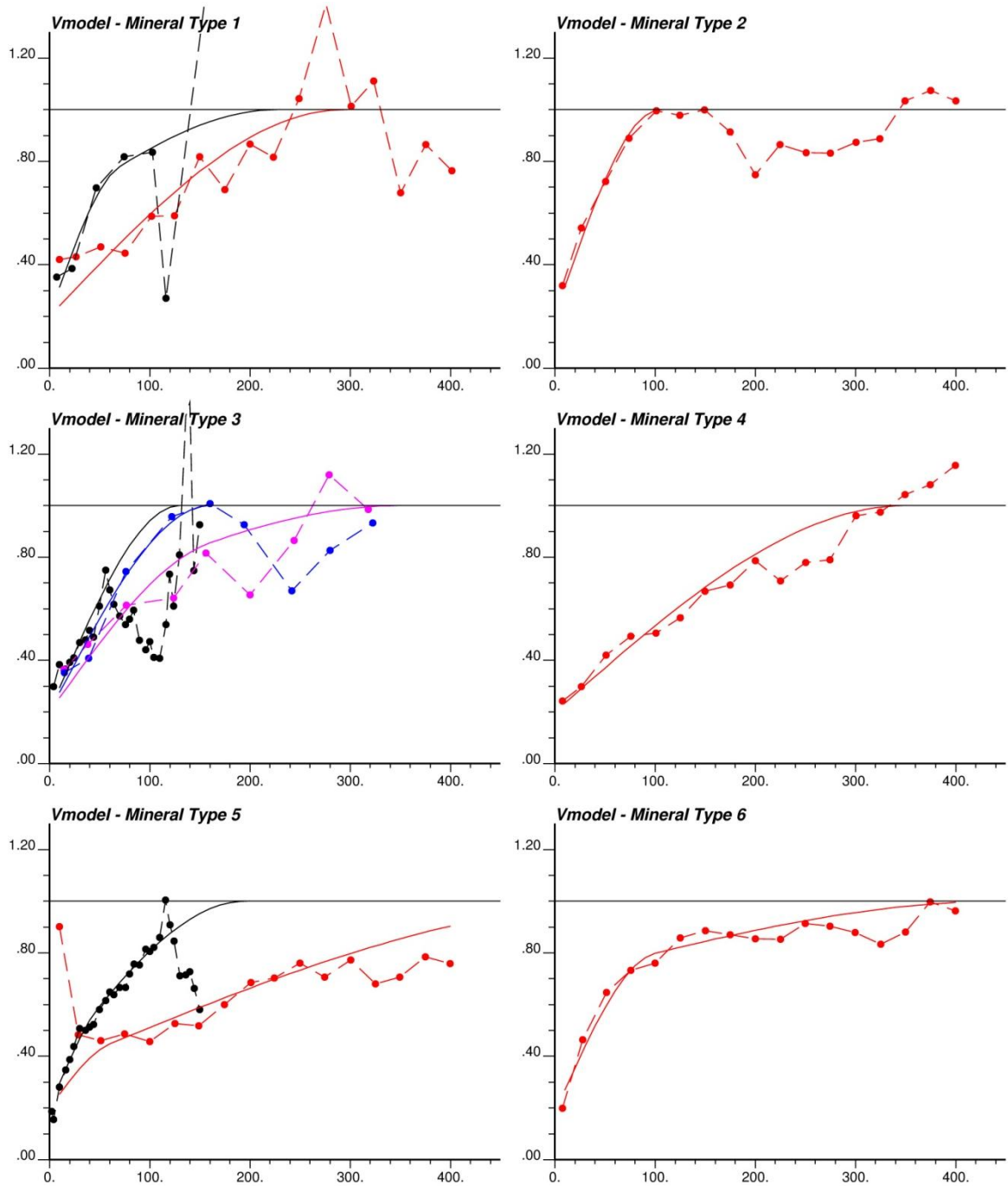
**Tabla 24: Modelos variográficos.**

UE	Dirección	Azimut [°]	Dip [°]
1	1	0	20
	2	90	20
	3	0	110
2	1	0	0
	2	90	0
	3	0	90
3	1	45	0
	2	135	0
	3	0	90
4	1	0	0
	2	90	0
	3	0	90
5	1	0	0
	2	90	0
	3	0	90
6	1	0	0
	2	90	0
	3	0	90

**Tabla 25: Rotación modelos variográficos.**

Se presenta en el Gráfico 49 el resultado del ajuste generado mediante los modelos variográficos junto con sus correspondiente variogramas experimentales. Se aprecia un buen ajuste de los modelos para cada una de las UE. Es importante tener en consideración el variograma modelado de la UE4, dado que esta unidad se estudiará en mayor detalle en la sección de resultados. Esto se debe básicamente a que, si bien la UE5 corresponde a la zona de interés definida inicialmente, es en la UE4 donde se aprecia de mejor manera un comportamiento que denota la presencia de deriva, es decir, un variograma experimental que aumenta linealmente por sobre la varianza de

los datos. En relación a este punto se debe notar que el modelo variográfico de la UE4 corresponde a un modelo omnidireccional por lo que el comportamiento no estacionario no se modela sólo en la dirección vertical sino que en todas direcciones. Una mayor discusión al respecto se presenta en el Anexo J (Gráfico 153).



**Gráfico 49: Variogramas modelados.**

### 6.9.2 Variogramas modelados residuos

Al igual que para los variogramas modelados de los compósitos originales, se modelan los variogramas experimentales definidos en Anexo J utilizando modelos anidados consistentes en un efecto pepa más uno, dos o tres variogramas esféricos. Dado que se trata de valores con distribución normal (0,1), la meseta en cada uno de los modelos variográficos es equivalente a 1. Las ecuaciones presentadas en la Tabla 26 definen los variogramas modelados para cada UE, mientras que las rotaciones de cada dirección se muestran en la Tabla 25.

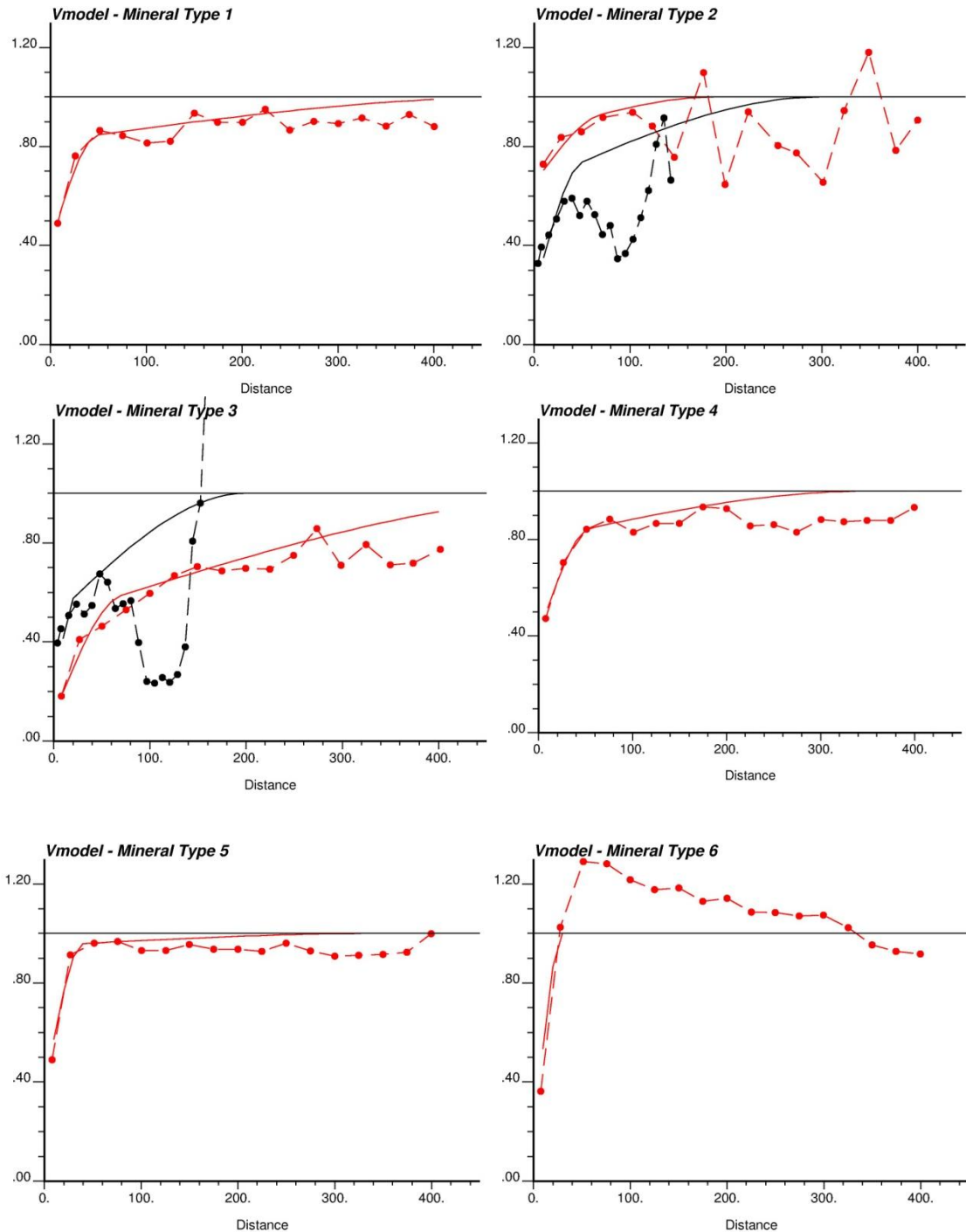
UE	Modelo variográfico
1	$0.40 P(\mathbf{h}) + 0.42 Esf_{(50,50,50)}(\mathbf{h}) + 0.18 Esf_{(500,500,500)}(\mathbf{h})$
2	$0.20 P(\mathbf{h}) + 0.45 Esf_{(10,10,50)}(\mathbf{h}) + 0.20 Esf_{(70,70,300)}(\mathbf{h}) + 0.15 Esf_{(190,190,300)}(\mathbf{h})$
3	$0.10 P(\mathbf{h}) + 0.40 Esf_{(70,70,20)}(\mathbf{h}) + 0.50 Esf_{(600,600,200)}(\mathbf{h})$
4	$0.40 P(\mathbf{h}) + 0.40 Esf_{(55,55,55)}(\mathbf{h}) + 0.20 Esf_{(350,350,350)}(\mathbf{h})$
5	$0.35 P(\mathbf{h}) + 0.60 Esf_{(40,40,40)}(\mathbf{h}) + 0.05 Esf_{(350,350,350)}(\mathbf{h})$
6	$0.10 P(\mathbf{h}) + 0.90 Esf_{(30,30,30)}(\mathbf{h})$

**Tabla 26: Modelos variográficos residuos.**

UE	Dirección	Azimut [°]	Dip [°]
1	1	0	0
	2	90	0
	3	0	90
2	1	0	0
	2	90	0
	3	0	90
3	1	0	0
	2	90	0
	3	0	90
4	1	0	0
	2	90	0
	3	0	90
5	1	0	0
	2	90	0
	3	0	90
6	1	0	0
	2	90	0
	3	0	90

**Tabla 27: Rotación modelos variográficos residuos.**

Se presenta en el Gráfico 50 el resultado del ajuste generado mediante los modelos variográficos junto con sus correspondiente variogramas experimentales. Se aprecia un buen ajuste de los modelos para cada una de las UE. Es interesante notar que en el caso de la variografía de residuos no se aprecia un comportamiento en el que hay presencia de deriva.



**Gráfico 50: Variogramas modelados residuos.**



## 6.10 Selección de sondajes para simulación

Previo al proceso de simulación se realiza una selección de los sondajes a simular de tal manera de separar la base de datos en dos conjuntos: el primero (C70) abarcando aproximadamente un 70% de los datos, se utiliza para realizar las simulaciones, y el segundo (C30) abarcando aproximadamente un 30% de los datos, se utiliza para una posterior validación de los resultados obtenidos.

Dado que esta selección de datos se realiza de manera aleatoria, la separación de conjuntos se realiza por sondaje en lugar de dato por dato. Se adopta esta forma de selección con el fin de realizar la validación de puntos con un valor conocido en sitios alejados de las muestras condicionantes (otros sondajes) que en sitios cercanos como sería el caso de sacar datos de manera aleatoria dentro de un mismo sondaje. Al definir un cierto porcentaje de datos en cada conjunto para cada UE y una selección por sondaje, el problema se vuelve no trivial, dado que cada sondaje tiene datos de distintas UE en distintas proporciones. Por lo tanto se crea el ejecutable con formato GSLib '*dhselect*' que realiza esta selección de sondajes de tal manera que se cumpla una proporción definida en cada UE dentro de un rango de tolerancia.

Se define una tolerancia de un 5%, lo cual da una separación de los 324 sondajes en un conjunto de 57 sondajes en el cual cada UE tiene entre un 25% y 35% de los datos y 267 sondajes en el cual cada UE tiene entre 65% y 75% de los datos. Los resultados de la separación de los datos en dos conjuntos en cuanto a número de datos y porcentaje que representa cada conjunto se muestran en la Tabla 28 mientras que los resultados en cuanto a media y varianza desagrupada para cada conjunto se presentan en la Tabla 29 mostrando los valores para los compósitos originales y en la Tabla 30 mostrando los valores para los residuos.

UE	N° Datos C30	N° Datos C70	C30 [%]	C70 [%]
1	312	641	32.7	67.3
2	281	532	34.6	65.4
3	478	1097	30.3	69.7
4	269	795	25.3	74.7
5	657	1669	28.2	71.8
6	531	1171	31.2	68.8

**Tabla 28: Número de datos por conjunto C30/C70.**

UE	Media Desagrupada C30	Media Desagrupada C70	Media Desagrupada UE	Varianza Desagrupada C30	Varianza Desagrupada C70	Varianza Desagrupada UE
1	0.4426	0.4588	<b>0.4528</b>	0.0328	0.0298	<b>0.0310</b>
2	0.2088	0.2017	<b>0.2042</b>	0.0202	0.0196	<b>0.0198</b>
3	0.1758	0.1620	<b>0.1662</b>	0.0419	0.0565	<b>0.0521</b>
4	0.4838	0.4243	<b>0.4402</b>	0.0751	0.0769	<b>0.0771</b>
5	0.7509	0.6644	<b>0.6896</b>	0.3410	0.2959	<b>0.3106</b>
6	1.0399	0.8531	<b>0.9115</b>	0.5392	0.5610	<b>0.5617</b>

**Tabla 29: Media y varianza por conjunto C30/C70 – Compósitos originales.**

UE	Media Desagrupada C30	Media Desagrupada C70	Media Desagrupada UE	Varianza Desagrupada C30	Varianza Desagrupada C70	Varianza Desagrupada UE
1	-0.0116	-0.0122	<b>-0.0120</b>	0.0193	0.0343	<b>0.0288</b>
2	-0.0360	-0.0211	<b>-0.0264</b>	0.0292	0.0256	<b>0.0269</b>
3	-0.1384	-0.1285	<b>-0.1315</b>	0.1805	0.1046	<b>0.1278</b>
4	-0.0147	-0.0368	<b>-0.0309</b>	0.0572	0.0677	<b>0.0650</b>
5	0.0400	0.0188	<b>0.0250</b>	0.3079	0.2173	<b>0.2438</b>
6	0.0653	0.0332	<b>0.0432</b>	0.4003	0.3748	<b>0.3830</b>

**Tabla 30: Media y varianza por conjunto C30/C70 – Residuos.**

## 6.11 Simulación

La etapa de simulación utiliza todos los elementos ya definidos: datos desagrupados, variogramas modelados, conjunto de sondajes que abarcan aproximadamente el 70% de las muestras para realizar las simulaciones respectivas, conjunto de sondajes que abarca aproximadamente el 30% de las muestras para validar los resultados, modelo de roca y modelo de deriva.

La simulación se realiza mediante el ejecutable ‘*sgsim*’, el cual es un ejecutable de GSLib que permite implementar una simulación secuencial Gaussiana a partir de una variable continua modelada como Gaussiana, por lo cual dicho algoritmo utiliza como input los datos desagrupados, los cuales pasan por un proceso de transformación Gaussiana. Es en esta etapa donde se utiliza el algoritmo desarrollado en el marco de este estudio, el cual consiste en el mismo ejecutable con la posibilidad de implementar un Kriging con rechazo mediante una deriva de referencia.

Para esta etapa de simulación, y con el fin de contrastar los resultados del algoritmo desarrollado con un algoritmo ya existente, se implementa una simulación secuencial Gaussiana utilizando Kriging simple (SK), Kriging de residuos (BT) y Kriging con rechazo (RDR). A su vez, se sensibilizan los resultados del Kriging con rechazo en cuanto al tamaño de la vecindad de búsqueda para la media local (Cmean), el número de rechazos (Nrej), la tolerancia (Tol), el tipo de rechazo (Rtype) y el error de referencia (Referror).

### 6.11.1 Plan de Kriging

Se utiliza el mismo plan de Kriging en cuanto a los parámetros comunes a cada método para Kriging simple y Kriging con rechazo, es decir, los mismos variogramas modelados y plan de búsqueda. El criterio para definir el radio de búsqueda en cada UE consiste en usar un radio de búsqueda tal que sea equivalente al primer múltiplo de 50 metros donde se cumpla con un 80% o más de la meseta en el correspondiente variograma modelado. Este criterio captura la posibilidad de una búsqueda anisótropa dependiendo de cómo se alcance la meseta en cada dirección para variogramas no omnidireccionales. Los radios de búsqueda por dirección se presentan en la Tabla 31. En el caso de la UE 1 el elipsoide de búsqueda está rotado con un dip de 20° en el plan de SK. En el caso de Kriging de residuos, los mismos radios de búsqueda son utilizados, sin embargo el variograma a utilizar corresponde a los variogramas de los residuos presentados en el Gráfico 50.

UE	Tipo de variograma	Radio de búsqueda [m]		
		X	Y	Z
1	Omnihorizontal	200	200	100
2	Omnidireccional	100	100	100
3	Direccional	100	100	150
4	Omnidireccional	200	200	200
5	Omnihorizontal	300	300	100
6	Omnidireccional	150	150	150

**Tabla 31: Radio de búsqueda por UE.**

En cada simulación se generan 100 realizaciones sobre la grilla dada por la Tabla 20 con una semilla para la definición del camino aleatorio igual a 69,069. Los datos se asignan a los nodos y se utiliza un máximo de 24 muestras para realizar la simulación en cada nodo. La búsqueda de

datos se realiza utilizando 3 multigrillas sin la opción de búsqueda por octantes. Para definir el tamaño de la matriz de covarianzas pre-computada se utiliza la siguiente fórmula:

$$Cov_i = \frac{4Radio_i}{siz_i} + 1$$

Donde *siz* es el tamaño del bloque en la dirección *i* (a saber 10 en cada dirección). De esta manera el tamaño de la matriz de covarianzas queda dado según la Tabla 32.

UE	Tamaño matriz de covarianza		
	X	Y	Z
1	81	81	41
2	41	41	41
3	41	41	61
4	81	81	81
5	121	121	41
6	61	61	61

**Tabla 32: Tamaño matriz de covarianza por UE.**

Se utiliza además el mismo criterio de extrapolación de colas en la destransformación de los valores a su distribución original tanto en SK como en RDR. En cada UE se realiza una extrapolación lineal fijando el mínimo en 0 y el máximo en un valor equivalente un 10% por sobre el máximo experimental de la respectiva UE, tal como se muestra en la Tabla 33.

UE	Máximo	Cola superior
1	1.35	1.49
2	0.99	1.09
3	3.33	3.66
4	3.26	3.59
5	5.45	6.00
6	6.68	7.35

**Tabla 33: Cola superior por UE – SK/RDR.**

En el caso de Kriging de residuos, el criterio de extrapolación de colas es levemente distinto. Dado que los residuos pueden tomar valores negativos, se realiza una extrapolación lineal fijando el mínimo y el máximo en un valor equivalente a un 10% por debajo y por sobre al valor mínimo y máximo experimental respectivamente, tal como lo muestra la Tabla 34.

UE	Mínimo	Cola Inferior	Máximo	Cola Superior
1	-0.85	<b>-0.93</b>	0.91	<b>1.00</b>
2	-1.08	<b>-1.19</b>	0.81	<b>0.90</b>
3	-1.78	<b>-1.96</b>	2.02	<b>2.22</b>
4	-1.10	<b>-1.21</b>	2.59	<b>2.85</b>
5	-1.52	<b>-1.67</b>	4.72	<b>5.19</b>
6	-1.76	<b>-1.94</b>	6.35	<b>6.99</b>

**Tabla 34: Cola inferior y superior por UE – BT.**

Con los parámetros hasta ahora fijados queda completamente definido el plan de Kriging para el caso de SK y BT, sin embargo en el caso de RDR se debe definir aún una serie de parámetros más los cuales además serán sensibilizados. En este sentido, se definen los siguientes parámetros bases sobre los cuales se realizará la sensibilización dependiendo del parámetro a estudiar.

Parámetros	Base	Sensibilización
Cmean	<b>1.0</b>	0.1 - 0.5 - 2.0
Nrej	<b>1,000</b>	10 - 100 - 10,000
Tol	<b>2</b>	1 - 3
Rtype	<b>1</b>	0
Refferror	<b>0.1</b>	0.01 - 1 - 10

**Tabla 35: Parámetros Kriging con rechazo.**

En cuanto a la tolerancia dinámica, esta es de decaimiento exponencial en todos los casos variando entre tolerancia alta (1) con una tolerancia inicial de 0.999 y final de 0.995; tolerancia de amplio rango (2) con una tolerancia inicial de 0.999 y final de 0.001; y una baja (3) con una tolerancia inicial de 0.050 y final de 0.001. Con respecto al tipo de rechazo, éste puede ser probabilístico (1) o determinístico (0).

## 6.12 Estadísticas y análisis de las simulaciones

Se analiza el resultado de las simulaciones basado en tres grupos de estadísticas. El primer grupo corresponde a estadísticas basadas solamente en los valores simulados: éstas corresponden a histograma acumulado, variograma experimental y media local, ambos en las direcciones X, Y y Z. Este grupo de estadísticas se compara además contra las mismas estadísticas de los datos originales (100% de los sondajes), presentando tanto las 100 simulaciones como los datos originales en el mismo gráfico.

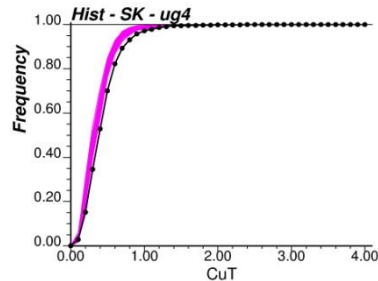
El segundo grupo corresponde a estadísticas de validación ya que se calculan mediante el uso del conjunto de sondajes que equivale a aproximadamente al 30% de los datos. En este grupo se presentan dos tipos de estadísticas. Por un lado se tienen estadísticas de validación del desempeño de la simulación como estimación (coeficiente de determinación  $R^2$ , pendiente de la regresión real versus simulado y error cuadrático medio) para lo cual se contrasta el valor de cada uno de los datos del conjunto C30, con el valor correspondiente según el resultado de la media de las simulaciones. Por otro lado se presentan estadísticas de validación del desempeño de la simulación en cuanto a la caracterización de la incertidumbre (accuracy plot) calculando qué tan bien caracteriza la simulación la incertidumbre de los datos originales.

Por último, el tercer grupo equivale a estadísticas específicas del Kriging con rechazo con la finalidad de testear su desempeño como nuevo algoritmo de simulación. En este grupo de estadísticas se cuenta con los tiempos de simulación, número de nodos rechazados a resimular y número de nodos finalmente no simulados. Las UE presentadas corresponden a la UE5 por ser la zona de interés a analizar y la UE4 ya que ésta presenta de manera más notoria un comportamiento que denota presencia de deriva en su variograma experimental. Para el primer grupo de estadísticas se presentan los resultados de la simulación por Kriging simple y Kriging de residuos, y los resultados de la simulación por Kriging con rechazo considerando sólo el caso base. Los resultados de las UE 1, 2, 3 y 6, así como los resultados de las sensibilizaciones (cmean, nrej, tol, rtype y referror) de cada una de las 6 unidades de estimación para el primer grupo de estadísticas, son presentados en los Anexos K (histogramas), L (variogramas), M (media local), N (coeficiente de determinación), O (pendiente de la regresión), P (error medio), Q (accuracy plot), R (tiempo de simulación), S (nodos rechazados) y T (nodos no simulados).

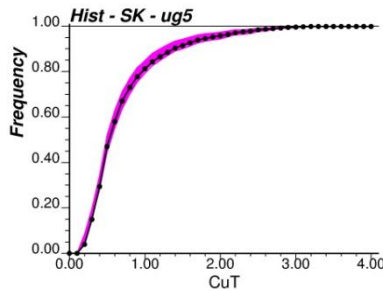
### 6.12.1 Histogramas

Se presentan los histogramas acumulados con una escala de 0% a 4% de ley de cobre total. Cada gráfico muestra los 100 histogramas acumulados de cada realización (líneas moradas) y el histograma acumulado de los datos originales desagrupados (línea negra).

#### 6.12.1.1 Kriging simple

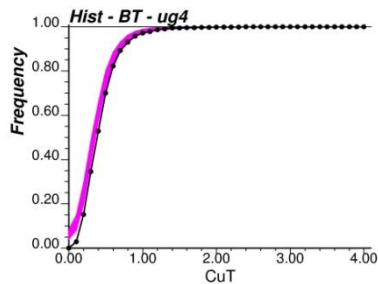


**Gráfico 51: Histograma – Kriging simple UE4.**

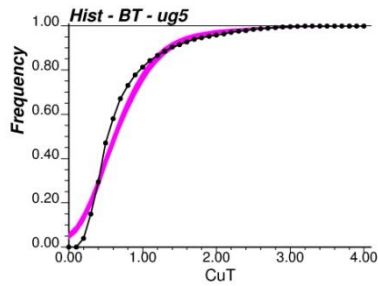


**Gráfico 52: Histograma – Kriging simple UE5.**

#### 6.12.1.2 Kriging de residuos

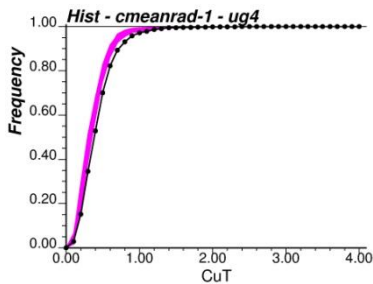


**Gráfico 53: Histograma – Kriging de residuos UE4.**

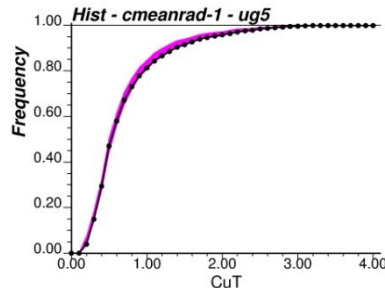


**Gráfico 54: Histograma – Kriging de residuos UE5.**

*6.12.1.3 Kriging con rechazo – Caso base*



**Gráfico 55: Histograma – Kriging con rechazo (caso base) UE4.**



**Gráfico 56: Histograma – Kriging con rechazo (caso base) UE5.**

A medida que aumenta el tamaño de la vecindad de búsqueda aumenta la dispersión de los histogramas acumulados de las simulaciones en torno al histograma acumulado de los datos originales, siendo todos menos dispersos que el caso de Kriging simple (SK). El cambio es considerable al pasar de una vecindad 0.1 a una 0.5, mientras que un aumento por sobre eso genera cambios prácticamente despreciables. El cambio en el número de rechazos no produce cambios significativos en los resultados siendo todos menos dispersos que el caso con SK. Al aumentar la tolerancia los valores de los histogramas se ajustan más al histograma de los datos originales. El caso de alta tolerancia (tolerancia 1) presenta una dispersión similar al SK. El tipo de rechazo determinístico presenta un sesgo con respecto a los datos originales mostrando una

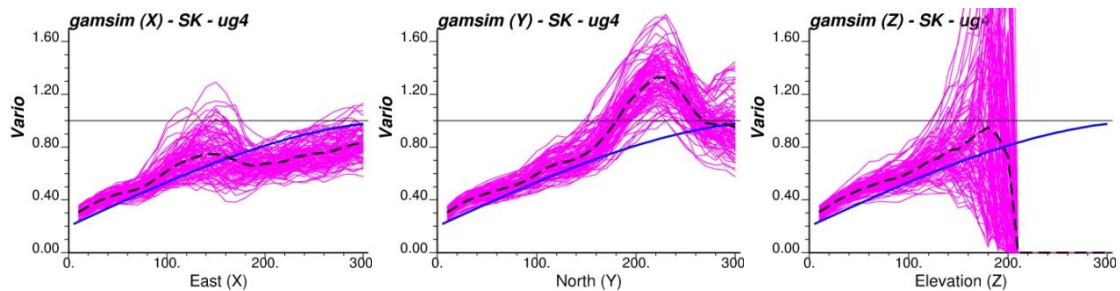


distribución un poco más amplia, no así el rechazo probabilístico, el cual presenta un buen ajuste siendo este menos disperso que el SK. El error de referencia muestra que para el menor de ellos (0.01) el ajuste al histograma original es casi exacto perdiéndose prácticamente la dispersión en torno a éste. Cada uno de los casos de este parámetro presenta una dispersión menor a la del SK, excepto por el error de referencia 10. Con respecto a los histogramas del BT, se puede apreciar en estos un sesgo con respecto al histograma de los datos originales. Esto queda aún más claro en la UE5 y UE6. Para el caso de la UE1 y UE4, este sesgo se hace presente para valores bajos de la ley. En este sentido, cada una de las sensibilizaciones realizadas con RDR logra un mejor ajuste al histograma de los datos originales en relación a los resultados obtenidos con BT en el caso de las UE1, UE4, UE5 y UE6.

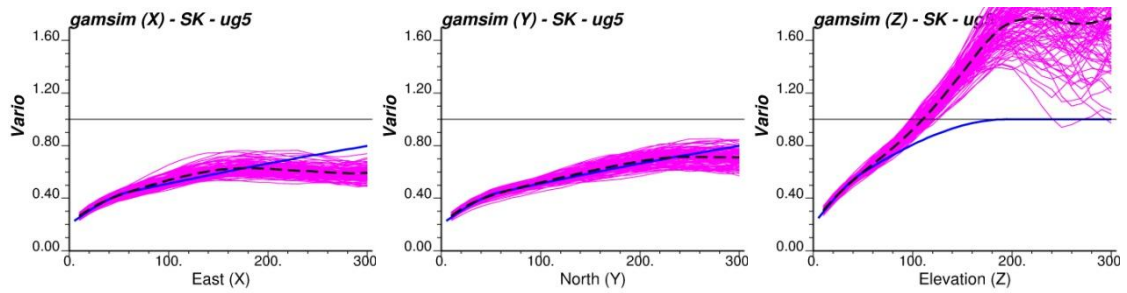
### 6.12.2 Variogramas

Se presentan los variogramas experimentales de los resultados de las simulaciones en el mismo formato en que se presentan los histogramas. Cada gráfico muestra, además de mostrar los 100 variogramas experimentales de cada simulación (líneas moradas), el variograma modelado correspondiente (línea azul) y el promedio de los 100 variogramas experimentales (línea negra segmentada). Para el cálculo de cada variograma experimental se utilizan 30 pasos de 10 metros, realizando dicho cálculo en las direcciones X, Y y Z.

#### 6.12.2.1 Kriging Simple

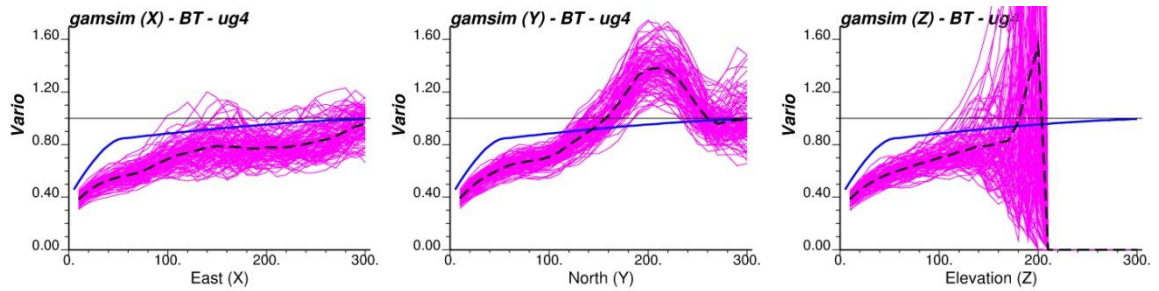


**Gráfico 57: Variograma X-Y-Z – Kriging simple UE4.**

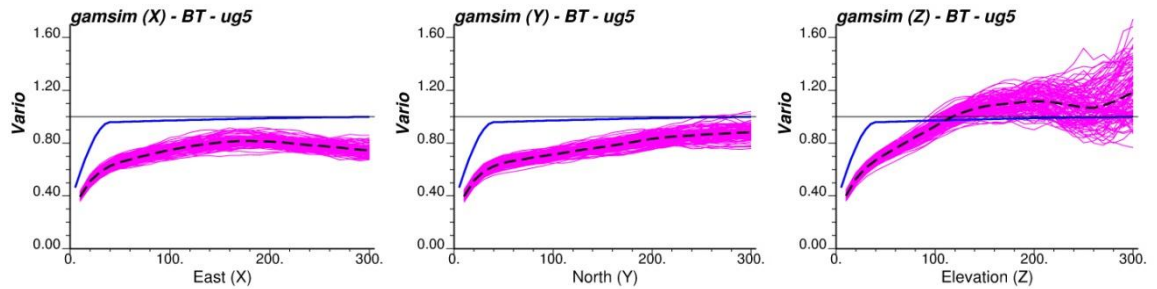


**Gráfico 58: Variograma X-Y-Z – Kriging simple UE5.**

6.12.2.2 Kriging de residuos

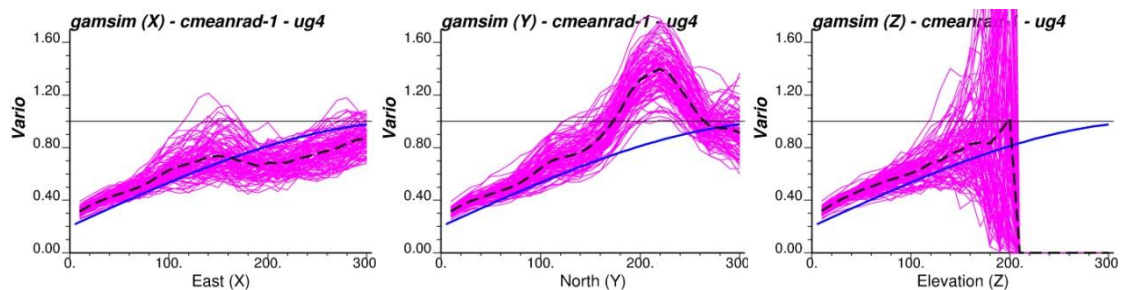


**Gráfico 59: Variograma X-Y-Z – Kriging de residuos UE4.**

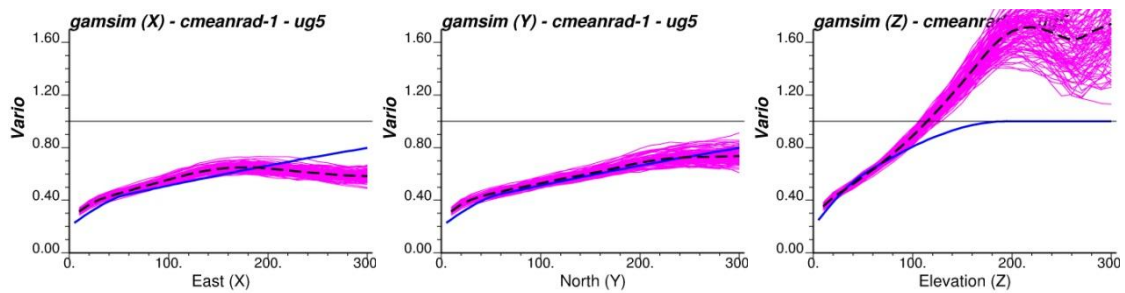


**Gráfico 60: Variograma X-Y-Z – Kriging de residuos UE5.**

6.12.2.3 Kriging con Rechazo – Caso base



**Gráfico 61: Variograma X-Y-Z – Kriging con rechazo (caso base) UE4.**



**Gráfico 62: Variograma X-Y-Z – Kriging con rechazo (caso base) UE5.**

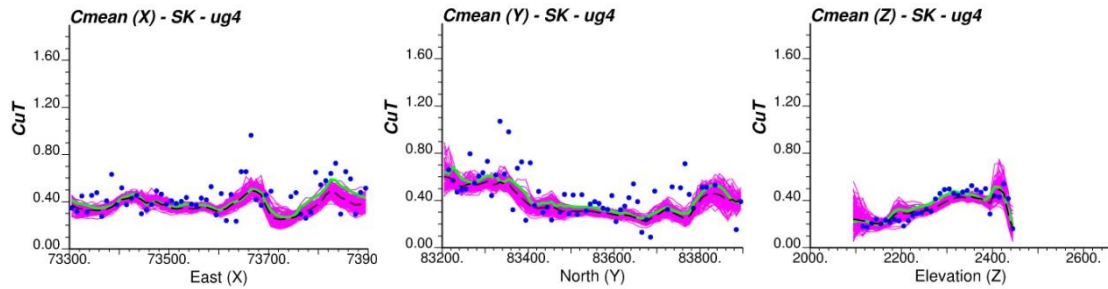
El tamaño de la vecindad de media condicional más pequeño (0.1) presenta un sesgo en cada una de las direcciones X, Y y Z. Los resultados para las vecindades 0.5, 1 y 2 presentan un buen ajuste, similar al obtenido por SK, tanto en dispersión en torno al variograma modelado como en el variograma medio de las simulaciones. La variación del número de rechazos no muestra diferencias significativas entre cada uno de los cuatro escenarios, siendo estos resultados similares a los de SK. La tolerancia a medida que aumenta provoca una disminución de la dispersión de los variogramas simulados en torno al variograma modelado. El caso de mayor ajuste corresponde a la alta tolerancia, siendo el caso de baja tolerancia el más sesgado en cuanto a reproducción del variograma modelado. El tipo de rechazo determinístico posee un mayor sesgo en la reproducción del variograma modelado que el tipo de rechazo probabilístico. En cuanto al error de referencia, este presenta un sesgo al utilizar un valor de 0.01, mientras que los resultados con valores de 0.1, 1 y 10 presentan resultados similares a los obtenidos por SK. Se puede apreciar que los resultados obtenidos con BT presentan, al igual que para el caso de los histogramas, un sesgo en la media de los variogramas de las simulaciones con respecto al variograma modelado. En este sentido se tiene que tener en consideración que el caso de BT términos generales no obtiene buenos ajustes, teniendo una buena reproducción del variograma modelado sólo en la UE2 y en las direcciones X e Y de la UE3.

### 6.12.3 Media local

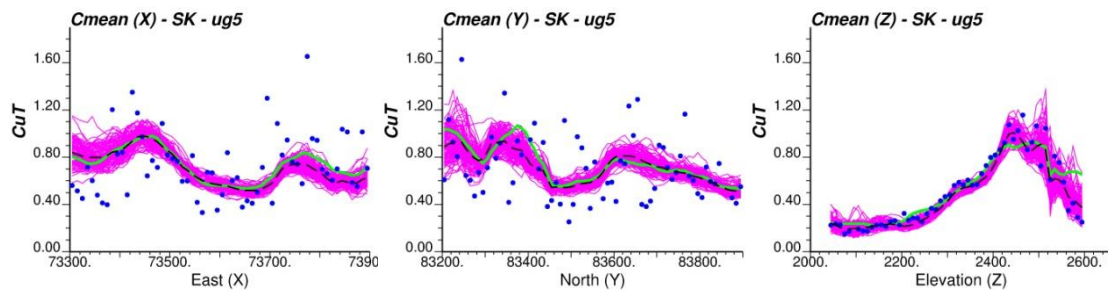
Se presenta la media local de los resultados de las simulaciones en el mismo formato en que se presentan los histogramas. Cada gráfico muestra las 100 medias locales de cada simulación (líneas moradas), el modelo de deriva (línea verde), la media de los compositos (puntos azules) y

el promedio de las simulaciones (línea negra segmentada). El cálculo de la media local de cada una de las 100 simulaciones se realiza considerando la grilla de simulación (Tabla 20).

### 6.12.3.1 Kriging Simple

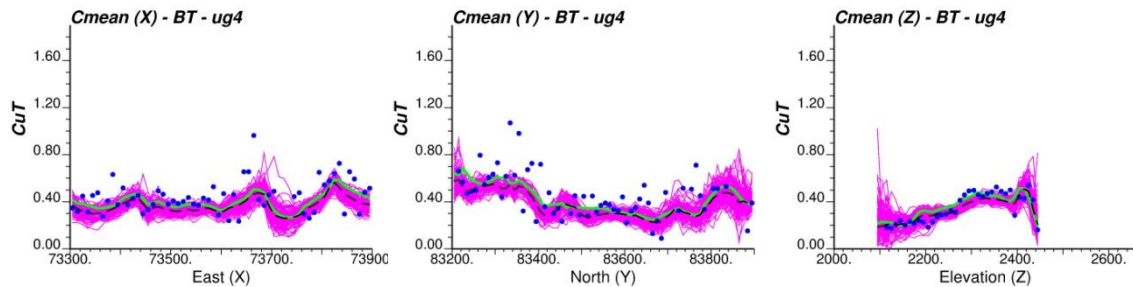


**Gráfico 63: Media local X-Y-Z – Kriging simple UE4.**

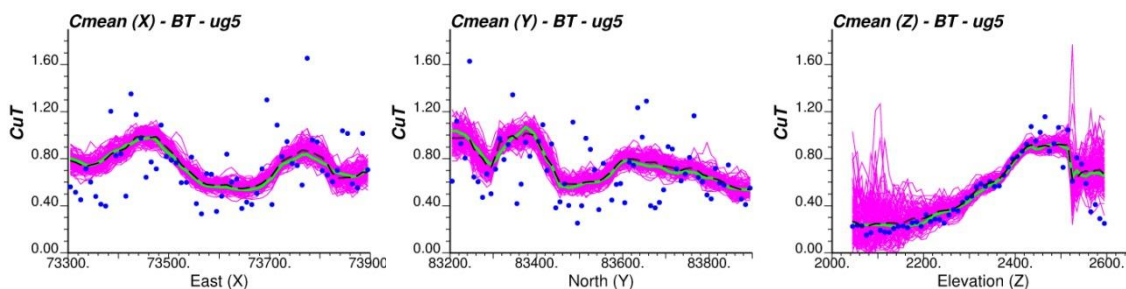


**Gráfico 64: Media local X-Y-Z – Kriging simple UE5.**

### 6.12.3.2 Kriging de residuos

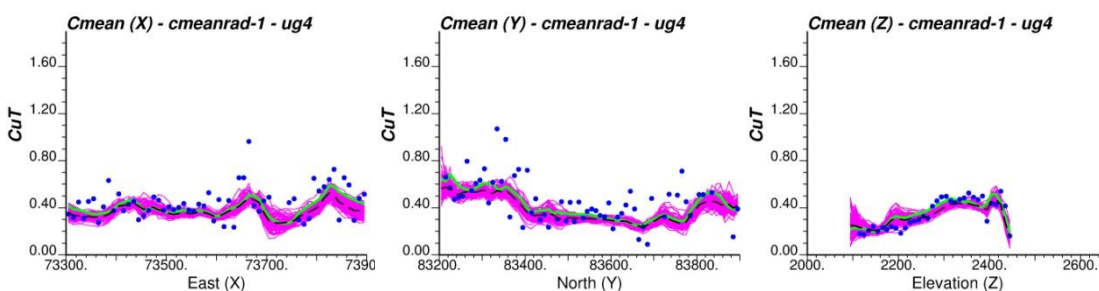


**Gráfico 65: Media local X-Y-Z – Kriging de residuos UE4.**

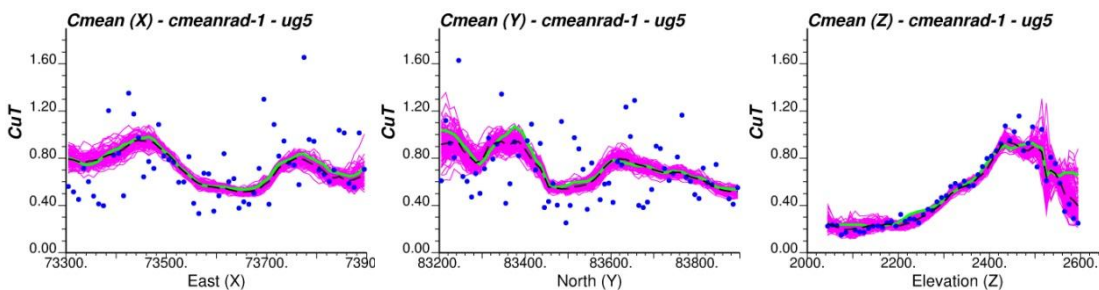


**Gráfico 66: Media local X-Y-Z – Kriging de residuos UE5.**

### 6.12.3.3 Kriging con Rechazo – Caso base



**Gráfico 67: Media local X-Y-Z – Kriging con rechazo (caso base) UE4.**



**Gráfico 68: Media local X-Y-Z – Kriging con rechazo (caso base) UE5.**

La media local como medida del ajuste a la deriva presenta distintos resultados para cada uno de los parámetros a sensibilizar. En el caso del tamaño de la vecindad, se presenta un aumento del ajuste a medida que esta disminuye de tamaño (0.01). Los casos de vecindad 0.5, 1, 2 y los resultados de SK son todos similares. Se aprecia una mayor dispersión y un menor ajuste a la deriva modelada en torno a la media local para las cotas extremas (piso y techo) en la coordenada Z. La variación del número de rechazos no tiene mayor influencia en los resultados. La disminución de la tolerancia (desde tolerancia alta a baja tolerancia) muestra mayor ajuste en torno a la deriva, aún cuando las cotas extremas siguen siendo poco ajustadas en la

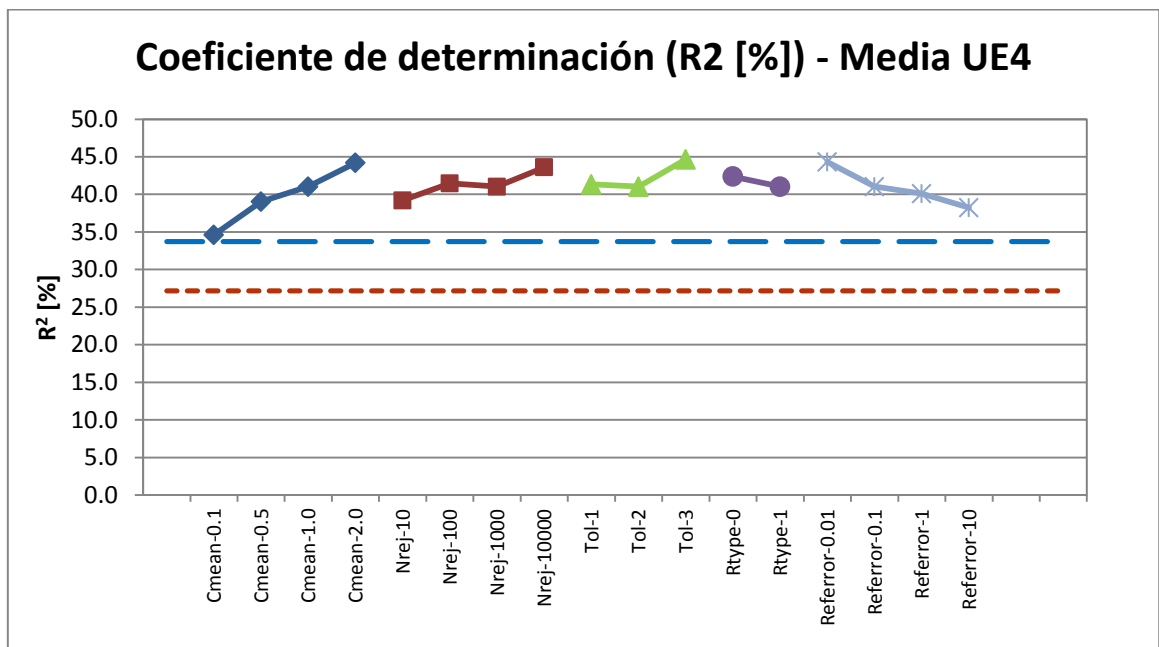
dirección Z. El tipo de rechazo determinista nuevamente muestra un sesgo en sus resultados en cuanto se obtienen medias locales por sobre la deriva a ajustar. El error de referencia mejora el ajuste a la deriva al tener un bajo valor (0.01) y a medida que aumenta, aumenta también la dispersión en torno a la deriva obteniéndose resultados similares al SK para los valores del error de referencia de 0.1, 1 y 10. El mejor ajuste a la deriva modelada se obtiene en el caso del BT el cual logra un buen ajuste en las cotas extremas de la coordenada Z que es justamente la zona con dificultades en el caso de SK y RDR. Los resultados de BT tienen un mejor ajuste en todas las UE salvo en la UE3. Si bien los resultados de BT son los que obtienen mejores ajustes, en el caso de la UE4, estos son bastante similares a los obtenidos con RDR.

#### 6.12.4 Coeficiente de determinación ( $R^2$ )

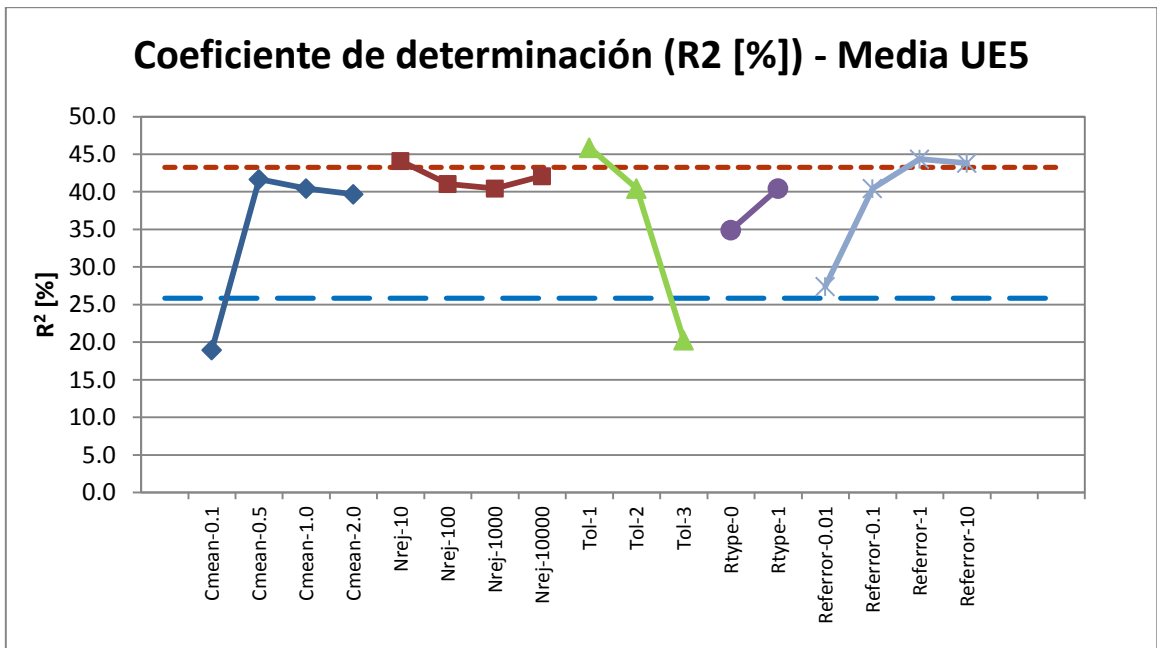
El coeficiente de determinación permite tener una medida de la proporción de la información de una variable que es posible predecir mediante otra variable, es decir, mide el poder predictivo. De esta manera, este estadístico permite revisar que tan bien ajusta la simulación utilizando el conjunto C30 para validar el modelo creado en cada simulación. Mientras más cercano a 1 el valor de  $R^2$ , mejor será el ajuste. El cálculo del coeficiente de determinación se realiza sobre el valor medio de las 100 simulaciones obteniéndose de esta manera un valor para cada una de las 17 sensibilizaciones de RDR y un valor para los casos de SK y BT contra los cuales se contrastan los resultados.

La UE5 muestra valores de  $R^2$  [%] de la media de las 100 simulaciones que van desde 19.0 para cmean-0.1, hasta 45.9 para tol-1, mientras que la UE4 muestra valores que van desde 27.2 para SK, hasta 44.6 para tol-3. El escenario con menor correlación no es el mismo para ambas UE. De hecho, tal como se puede ver en Anexo N, el caso con menor correlación es SK para la UE1 y UE4, referror-1 para la UE2, BT para la UE3, cmean-0.1 para la UE5 y tol-3 para la UE6. Es interesante notar que mientras para la UE5 el escenario tol-1 es el que tiene mayor correlación, el mismo escenario tiene una correlación muy cercana al mínimo en la UE2. Esto indica que la mejor elección de set de parámetros no está estandarizada y depende en efecto de la zona a simular considerando tanto el escenario de estimación como la naturaleza del problema. En este sentido tampoco existe una tendencia estricta del valor de  $R^2$  al variar un parámetro. El caso más claro ocurre en el parámetro de tolerancia: mientras en la UE5 el valor de  $R^2$  disminuye a medida

que se pasa de tol-1 a tol-3, en la UE2 el valor de  $R^2$  aumenta. Esto refuerza más la idea que el set de parámetros óptimos para la estimación depende fuertemente de las condiciones de los datos y la zona a simular. Se aprecia de los gráficos de  $R^2$  que a medida que aumenta el tamaño de la vecindad de la media, la correlación aumenta en la UE3, UE4, UE5 y UE6 y disminuye en la UE1 y UE2. En el caso de la tolerancia, el aumento de la correlación a medida que se pasa de tol-1 a tol-3 se produce en la UE1, UE2 y UE3 mientras que la disminución en la UE5 y UE6. Por su parte, al aumentar el error de referencia, el valor de la correlación aumenta en la UE5 y UE6 y disminuye en la UE1, UE2 y UE4. El número de rechazos no muestra una tendencia clara en su efecto en el valor de  $R^2$  mientras que en todas las UE se obtiene un mejor ajuste del modelo en el caso de rechazo probabilístico que el determinista salvo en la UE4 donde se obtienen valores muy similares para ambos casos. En relación a la comparación con el caso de SK y BT se presenta en el Gráfico 69 y Gráfico 70 los resultados de cada una de las 17 sensibilizaciones con RDR mientras que el valor de SK y BT se presentan con líneas punteadas naranja y azul respectivamente. Es posible ver del Gráfico 69 que cada una de las sensibilizaciones de RDR obtiene valores de  $R^2$  mayores que los casos de SK y BT. Lo mismo no sucede en la UE5 donde la mayor parte de las sensibilizaciones de RDR se encuentran entre los valores de BT como cota inferior y de SK como cota superior.



**Gráfico 69: Coeficiente de determinación ( $R^2$ ) – Media UE4.**



**Gráfico 70: Coeficiente de determinación ( $R^2$ ) – Media UE5.**

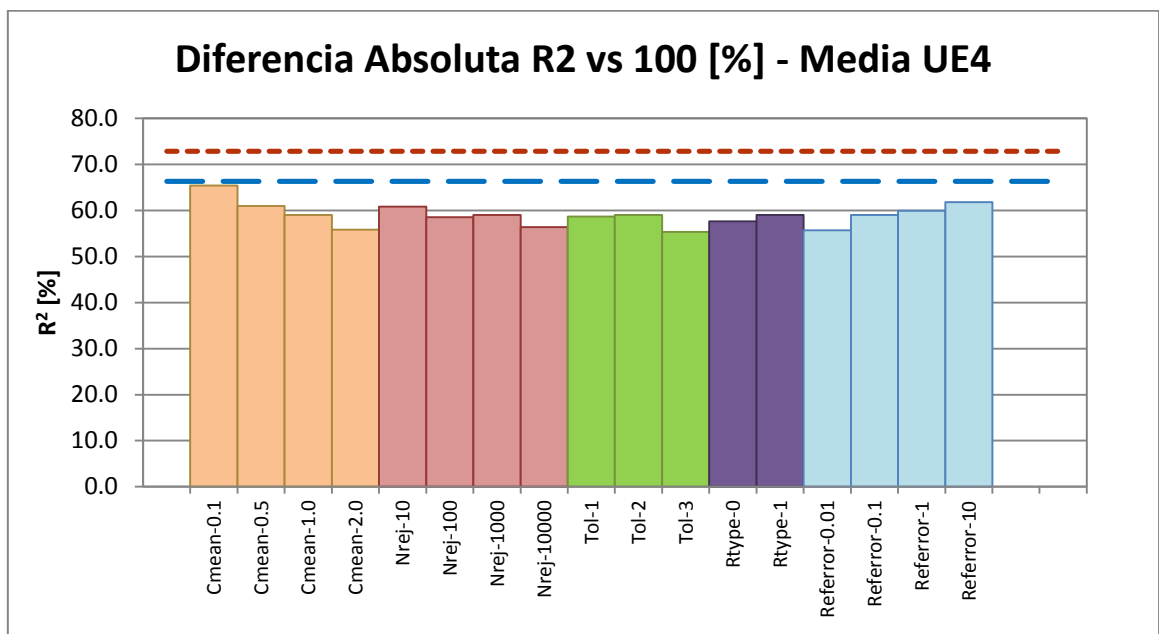
Si bien en estos gráficos se puede ver el desempeño del RDR versus los otros dos algoritmos de simulación, más interesante aún es tener un índice de desempeño que mida cuantitativamente y no cualitativamente el desempeño global del algoritmo considerando todas las sensibilizaciones. Para esto se define un valor objetivo a alcanzar el cual equivale a 100% para el caso del coeficiente de determinación. El Gráfico 71 y Gráfico 72 muestran la diferencia absoluta entre el valor de  $R^2$  de cada caso y el valor objetivo (100% en este caso). De esta manera, un valor cercano a 0 en estos gráficos indica un mejor alcance del valor objetivo. En base a esto se define el siguiente índice donde DA es equivalente a la diferencia absoluta entre el valor del subíndice correspondiente y el valor objetivo:

$$\text{Índice de SK: } I_{SK} = \sum_{i=1}^{17} \frac{D_{ASK} - D_{ARDR_i}}{|D_{ASK} - D_{ARDR_i}|}$$

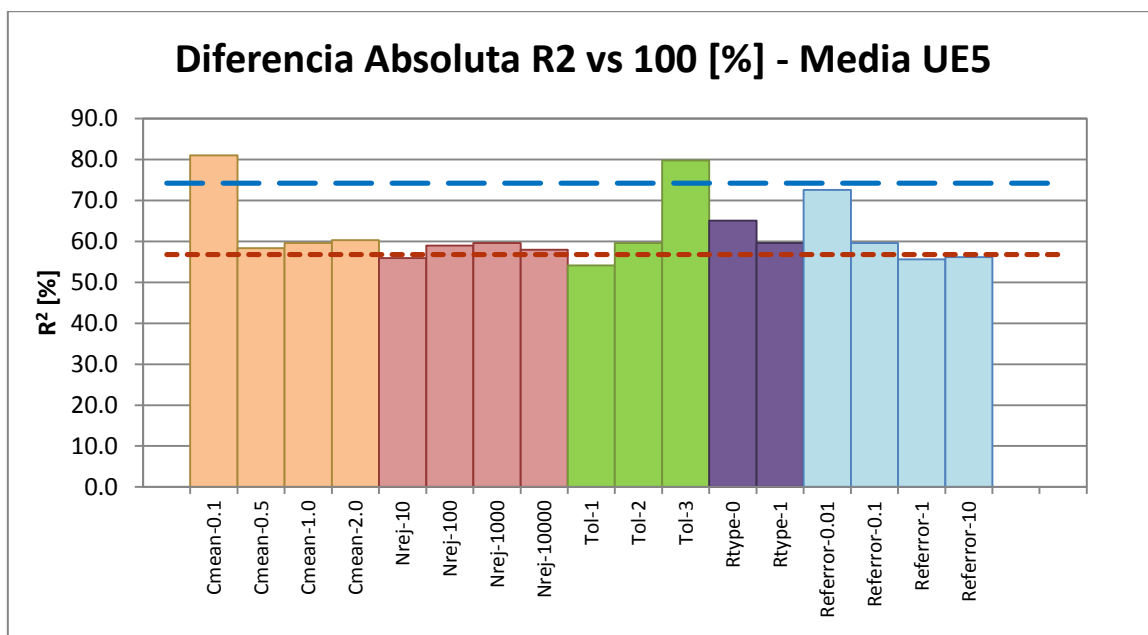
$$\text{Índice de BT: } I_{BT} = \sum_{i=1}^{17} \frac{D_{ABT} - D_{ARDR_i}}{|D_{ABT} - D_{ARDR_i}|}$$



Lo que hace este índice es sumar todas las diferencias entre la DA del algoritmo a comparar (SK o BT) y la DA de las sensibilizaciones de RDR y estandarizar estos sumandos para que el rango de valores esté entre -1 y 1. A modo de ejemplo, en el caso en que SK tenga un valor de 20% para  $R^2$ , su DA será de 80%. De esta manera si una sensibilización de RDR tiene un valor de 30% para  $R^2$ , su valor de DA será de 70% y por lo tanto en la fórmula del índice  $I_{SK}$  sumará un 10% (valor positivo). En el caso en que todas las sensibilizaciones presenten una mejora respecto al SK, todas sumarán un valor positivo y por lo tanto al momento de estandarizar el valor del índice será igual a 1. Por lo tanto este índice mide el desempeño global del RDR versus SK o BT obteniéndose un valor negativo cuando se trata de un peor desempeño y un valor positivo cuando se trata de un mejor desempeño. Los valores igual a -1, 0 y 1 indican un peor desempeño en todas las sensibilizaciones, un desempeño equivalente y un mejor desempeño en todas las sensibilizaciones respectivamente.



**Gráfico 71: Diferencia absoluta  $R^2$  vs 100 – Media UE4.**

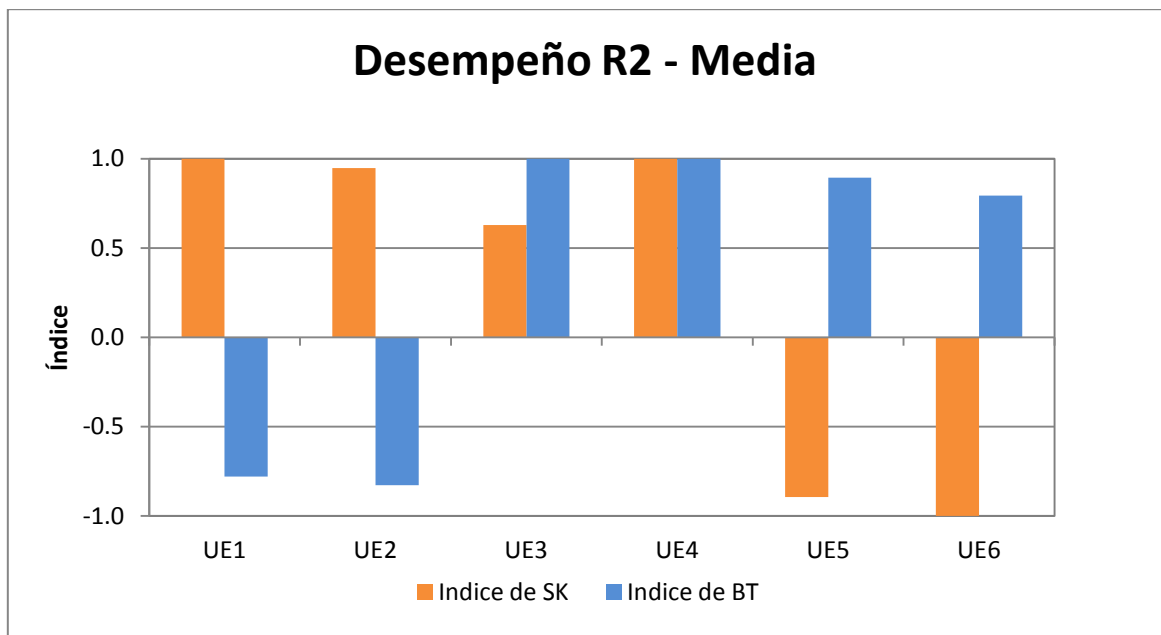


**Gráfico 72: Diferencia absoluta R<sup>2</sup> vs 100 – Media UE5.**

Tal como se aprecia en el Gráfico 71, el desempeño del RDR en la UE4 es mejor en cada una de las sensibilizaciones con respecto tanto al SK como al BT. En el caso de la UE5, el desempeño del RDR es mejor en la mayoría de los casos con respecto al BT pero no sucede lo mismo al comparar con SK. Esto se ve representado claramente por los índices respectivos en la Tabla 36. En el caso de la UE4 se observan mejoras en las 17 sensibilizaciones con respecto al SK y BT por lo cual el índice es equivalente a 1 en ambos casos. Por su parte la UE5 tiene un índice de 0.89 al comparar contra BT y un índice de -0.89 al comparar contra SK.

UE	RDR vs SK - Media			RDR vs BT - Media		
	N° de Mejoras	N° de Mejoras [%]	Índice	N° de Mejoras	N° de Mejoras [%]	Índice
UE1	17	100%	<b>1.00</b>	3	18%	<b>-0.78</b>
UE2	14	82%	<b>0.95</b>	3	18%	<b>-0.83</b>
UE3	13	76%	<b>0.63</b>	17	100%	<b>1.00</b>
UE4	17	100%	<b>1.00</b>	17	100%	<b>1.00</b>
UE5	4	24%	<b>-0.89</b>	15	88%	<b>0.89</b>
UE6	0	0%	<b>-1.00</b>	14	82%	<b>0.79</b>

**Tabla 36: Desempeño RDR vs SK y BT – Coeficiente de determinación (R<sup>2</sup>).**



**Gráfico 73: Desempeño RDR vs SK y BT – Coeficiente de determinación ( $R^2$ ).**

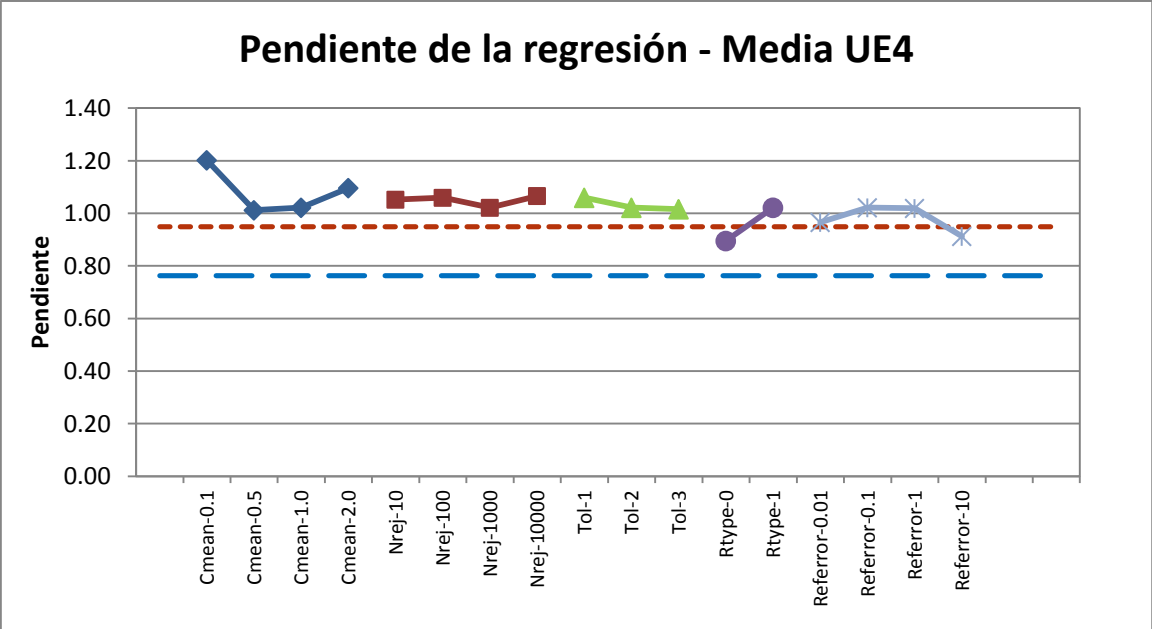
Se aprecia de la Tabla 36 que el RDR tiene un mejor desempeño para las UE1, UE2, UE3 y UE4 con respecto al SK mientras que el mejor desempeño ocurre en las UE3, UE4, UE5 y UE6 al comparar contra BT.

#### 6.12.5 Pendiente de la regresión

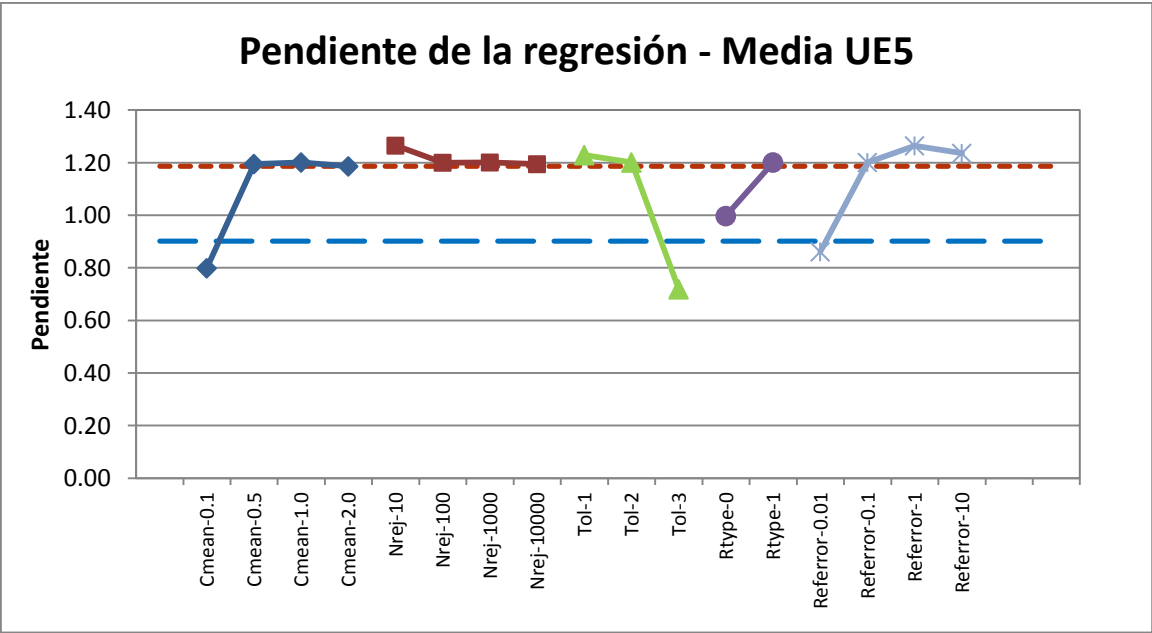
Este estadístico tiene el mismo carácter que el coeficiente de determinación, es decir, validar mediante el conjunto C30 los resultados. Se hace referencia a la regresión lineal entre la ley real versus la ley simulada, cuya pendiente mientras más cercana sea a 1 implicará un menor incesgo condicional. El cálculo de la pendiente de la regresión se realiza sobre el valor medio de las 100 simulaciones obteniéndose de esta manera un valor para cada una de las 17 sensibilizaciones de RDR y un valor para los casos SK y BT contra los cuales se contrastan los resultados.

La pendiente de la regresión como indicador del incesgo condicional toma valores desde 0.719 para tol-3 hasta 1.265 para nrej-10 en la UE5, mientras que en la UE4 toma valores que van desde 0.763 para BT hasta 1.202 para cmean-0.1. Es importante notar que si bien en el caso del coeficiente de determinación, un mejor desempeño con respecto al SK o BT implicaba un valor más alto, en el caso de la pendiente de la regresión no se cumple siempre lo mismo. Esto se debe

a que el valor objetivo en este estadístico es equivalente a 1.0 por lo tanto un valor más cercano a una pendiente igual a 1.0 implica un mejor desempeño. En el Gráfico 74 y Gráfico 75 se muestra el valor para la pendiente de la regresión de cada sensibilización así como el resultado obtenido mediante SK y BT presentados como líneas punteadas naranja y azul respectivamente.

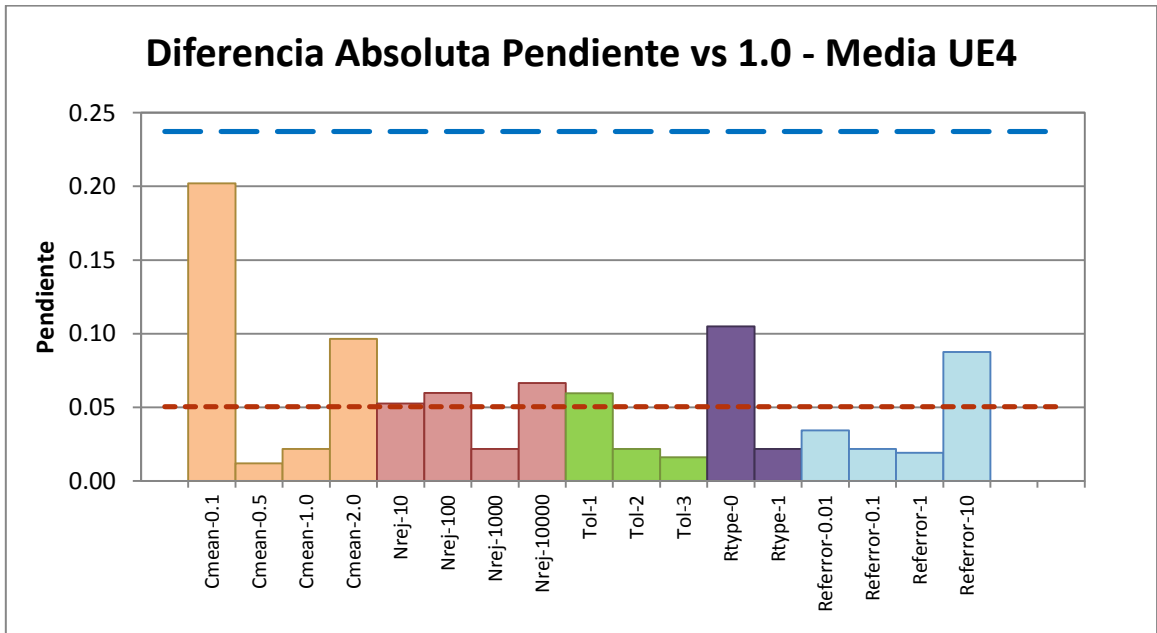


**Gráfico 74: Pendiente de la regresión – Media UE4.**

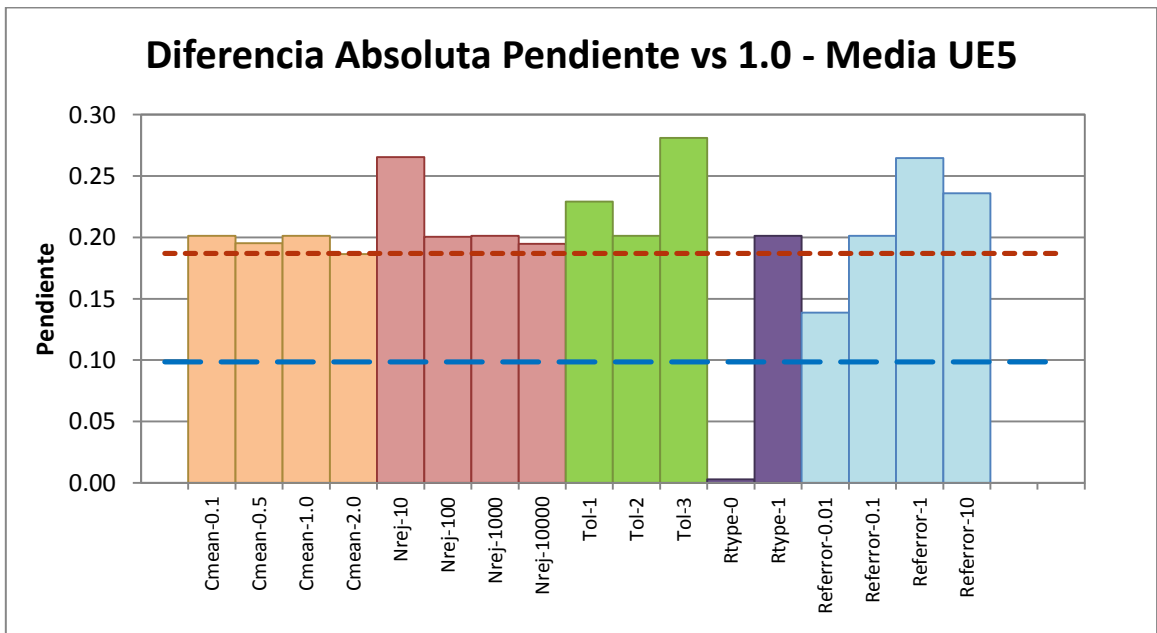


**Gráfico 75: Pendiente de la regresión – Media UE5.**

En el Gráfico 76 y Gráfico 77 se muestra la diferencia absoluta entre la pendiente de la regresión de SK, BT y cada una de las sensibilizaciones y el valor objetivo equivalente a 1.0 en este caso.



**Gráfico 76: Diferencia absoluta pendiente vs 1 – Media UE4.**



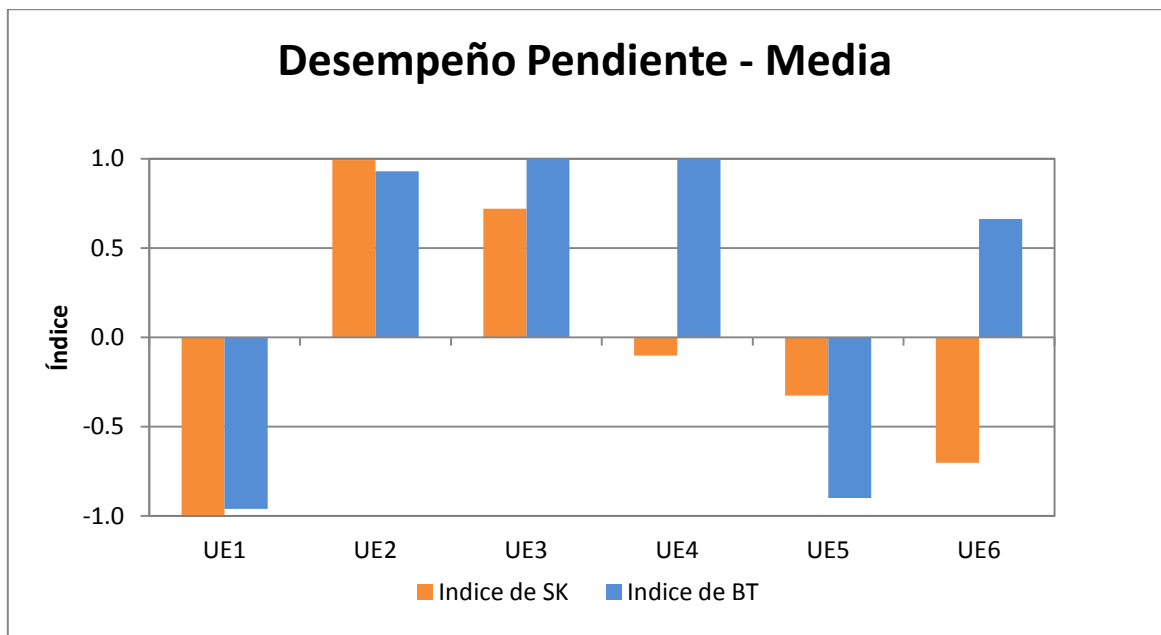
**Gráfico 77: Diferencia absoluta pendiente vs 1 – Media UE5.**

Es posible apreciar que en la UE4, el desempeño versus BT es siempre mejor, mientras que lo mismo no es tan claro al comparar contra SK. En este sentido es que el valor del índice mostrado en la Tabla 37 entrega una información cuantitativa más certera del desempeño del RDR en la UE4 obteniéndose un valor de 1.0 al comparar contra BT (un desempeño mejor en cada sensibilización) y un valor equivalente a -0.10 al comparar contra SK. Si bien en más de la mitad de los casos se obtienen mejoras con respecto al SK (9 de 17 sensibilizaciones) el valor de -0.10 del índice se debe a que en las sensibilizaciones en donde empeoró el algoritmo, el valor de la pendiente corresponde a un resultado muy alejado del valor objetivo, mostrando como la magnitud de la mejora o empeoramiento incide en el valor del índice. Aún cuando el índice de desempeño con respecto a SK es negativo e implica un peor desempeño, este valor es muy cercano a 0 por lo cual se encuentra cerca de tener un desempeño equivalente al SK. En la UE5 se aprecia que al comparar contra BT se tienen mejoras sólo en 1 de las 17 sensibilizaciones mientras que al comparar contra SK se tienen mejoras en 3 de las 17 sensibilizaciones. Esto se aprecia en los índices de BT y SK, obteniéndose un valor igual a -0.90 y -0.33 respectivamente, por lo cual el desempeño de RDR es peor que ambos algoritmos para la UE5.

UE	RDR vs SK – Media			RDR vs BT - Media		
	N° de Mejoras	N° de Mejoras [%]	Índice	N° de Mejoras	N° de Mejoras [%]	Índice
UE1	1	6%	<b>-1.00</b>	2	12%	<b>-0.96</b>
UE2	16	94%	<b>0.99</b>	14	82%	<b>0.93</b>
UE3	14	82%	<b>0.72</b>	17	100%	<b>1.00</b>
UE4	9	53%	<b>-0.10</b>	17	100%	<b>1.00</b>
UE5	3	18%	<b>-0.33</b>	1	6%	<b>-0.90</b>
UE6	9	53%	<b>-0.70</b>	14	82%	<b>0.66</b>

**Tabla 37: Desempeño RDR vs SK y BT – Pendiente.**

De la Tabla 37 es posible ver que se obtiene un mejor desempeño con respecto al SK en la UE2 y UE3, mientras que lo mismo sucede con respecto al BT en la UE2, UE3, UE4 y UE6.



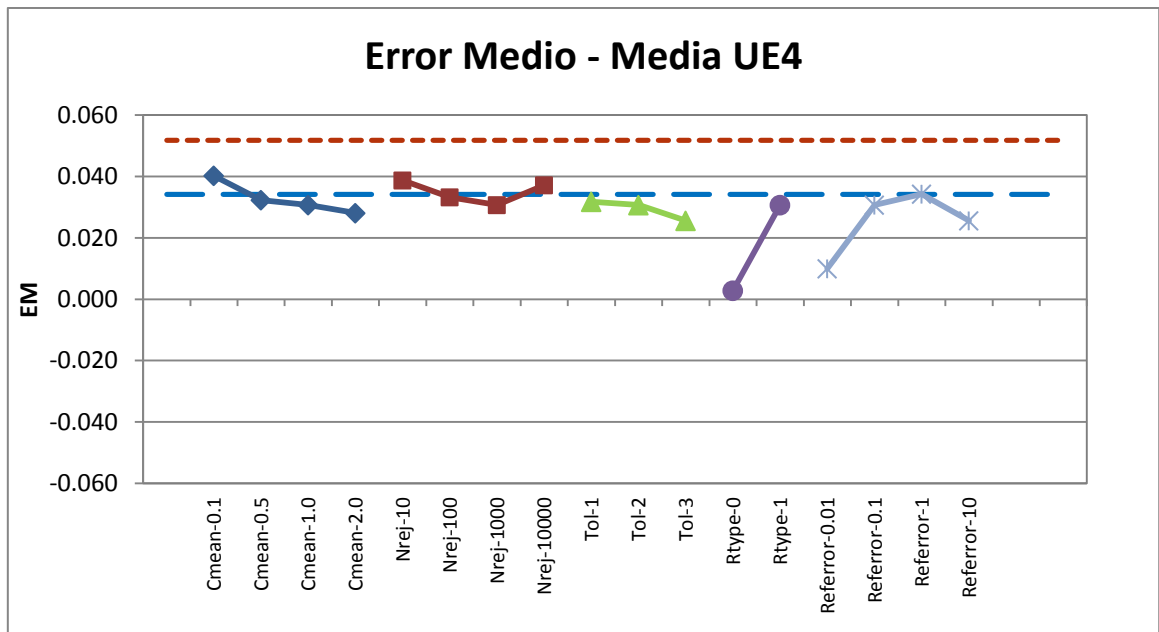
**Gráfico 78: Desempeño RDR vs SK y BT – Pendiente.**

#### 6.12.6 Error medio

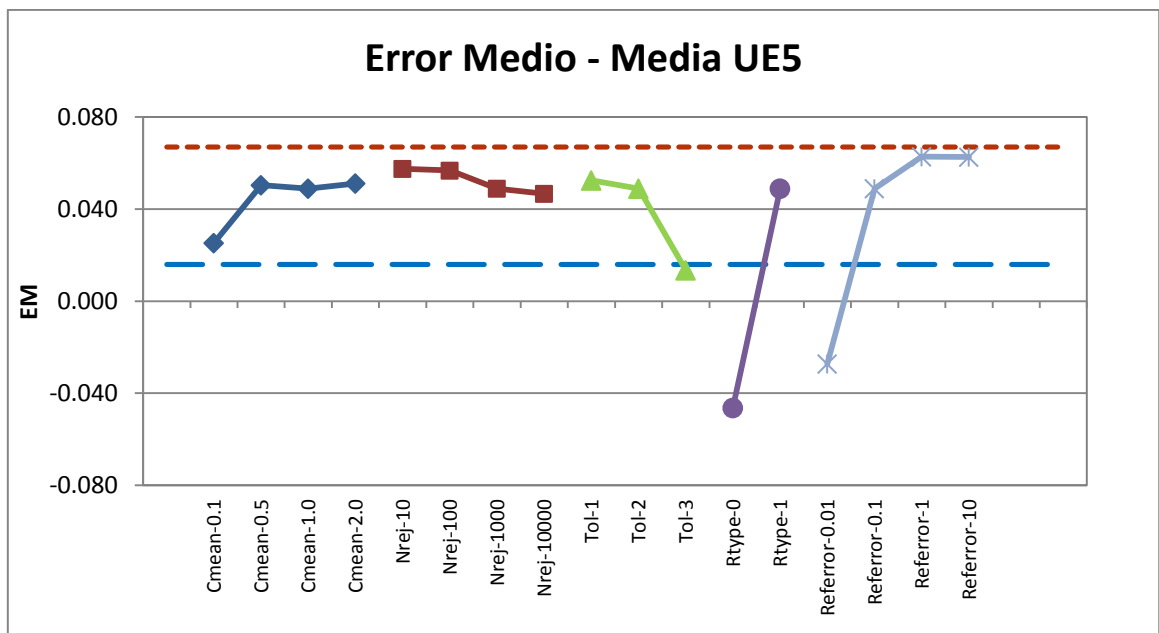
El error medio corresponde al tercero de los 3 estadísticos de validación de los resultados de la simulación como estimación. Se define como la media de las diferencias entre el valor real y el valor simulado considerando tanto valores negativos como positivos. De esta manera se mide el sesgo global y se espera un valor cercano a 0 con el fin de tener insesgo global. El cálculo del error medio se realiza sobre el valor medio de las 100 simulaciones obteniéndose de esta manera un valor para cada una de las 17 sensibilizaciones de RDR y un valor para los casos SK y BT contra los cuales se contrastan los resultados.

El error medio como indicador del insesgo global presenta valores que van desde -0.046 para rtype-0 hasta 0.067 para SK en la UE5 y desde 0.003 para rtype-0 hasta 0.052 para SK en la UE4. En el caso del estudio del error medio se busca que este sea lo más cercano a 0 cuando se chequea insesgo global. De esta forma tanto en la UE5 y la UE4 se tiene que el caso de SK es el que presenta peor desempeño en cuanto a insesgo global. Más aún, en la UE4 el desempeño del algoritmo es aún mejor dado que la brecha entre el SK y el escenario de mayor sesgo global de las sensibilizaciones es de 0.012, bastante más que los 0.004 de la misma brecha en la UE5. Se tiene como resultado un mejor desempeño del algoritmo con respecto al SK en todos los escenarios de las UE4, UE5 y UE6, las cuales presentan, según el Anexo D de base maps y 6.3.3

de nubes direccionales, cierta deriva en la coordenada Z. En el Gráfico 79 y Gráfico 80 se muestra el valor para error medio de la media de cada sensibilización así como el resultado obtenido mediante SK y BT presentados como líneas punteadas naranja y azul respectivamente. Es posible apreciar que el desempeño con respecto al BT no es tan claro como en el caso del SK.



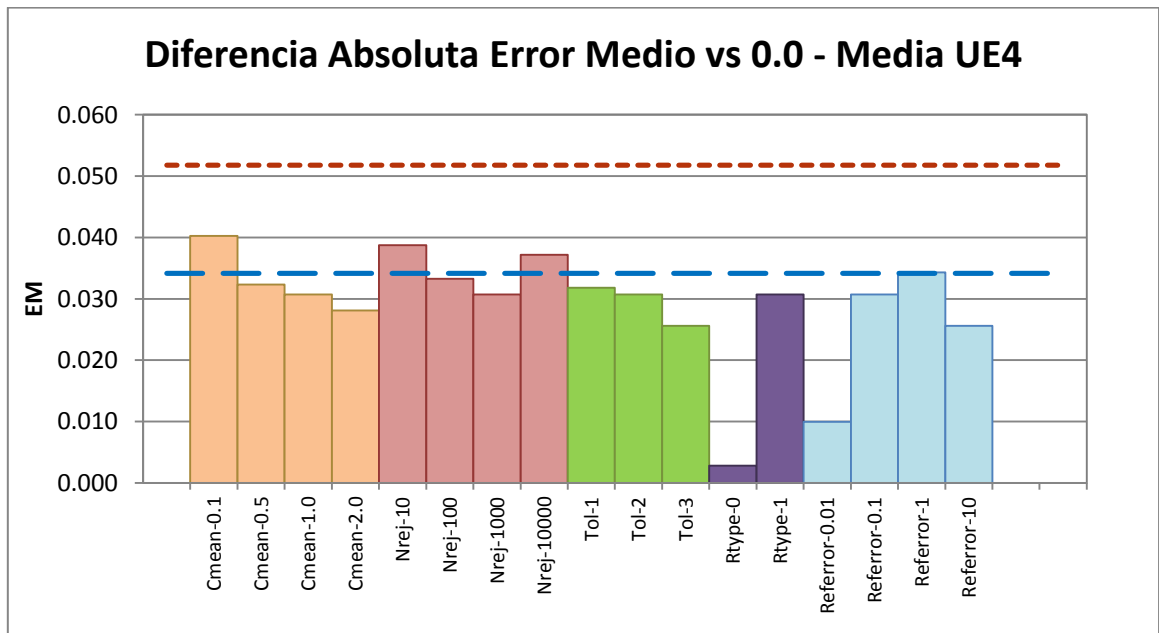
**Gráfico 79: Error medio – Media UE4.**



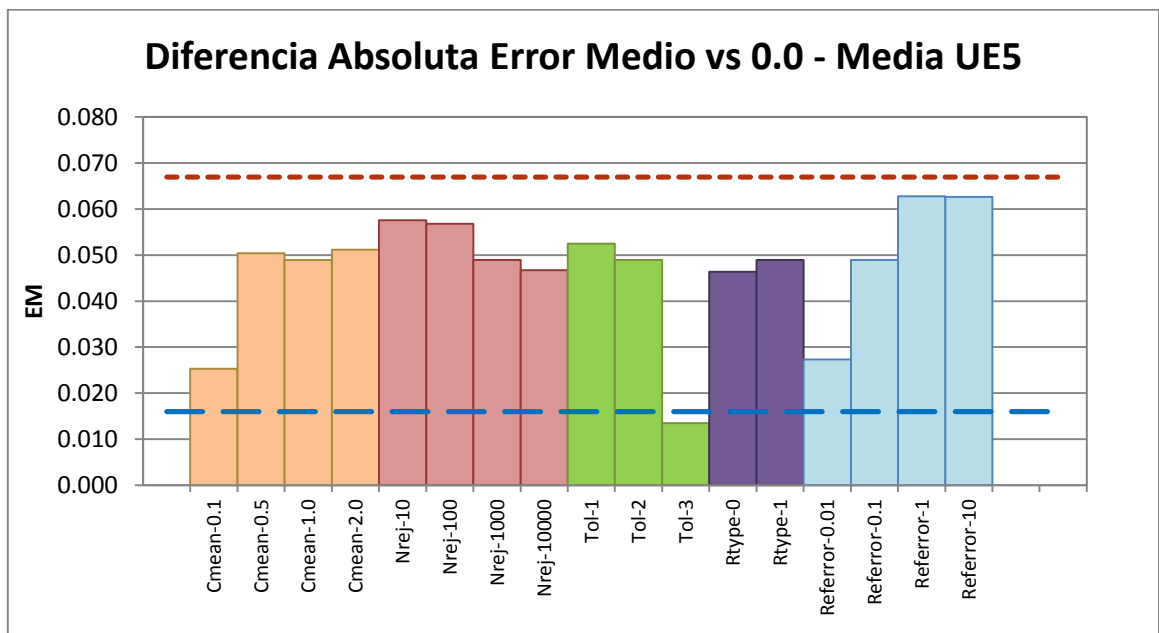
**Gráfico 80: Error medio – Media UE5.**



En el Gráfico 81 y Gráfico 82 se muestra la diferencia absoluta entre los resultados de error medio y el valor objetivo que en este caso corresponde a 0.0.



**Gráfico 81: Diferencia absoluta error medio vs 0.0 – Media UE4.**



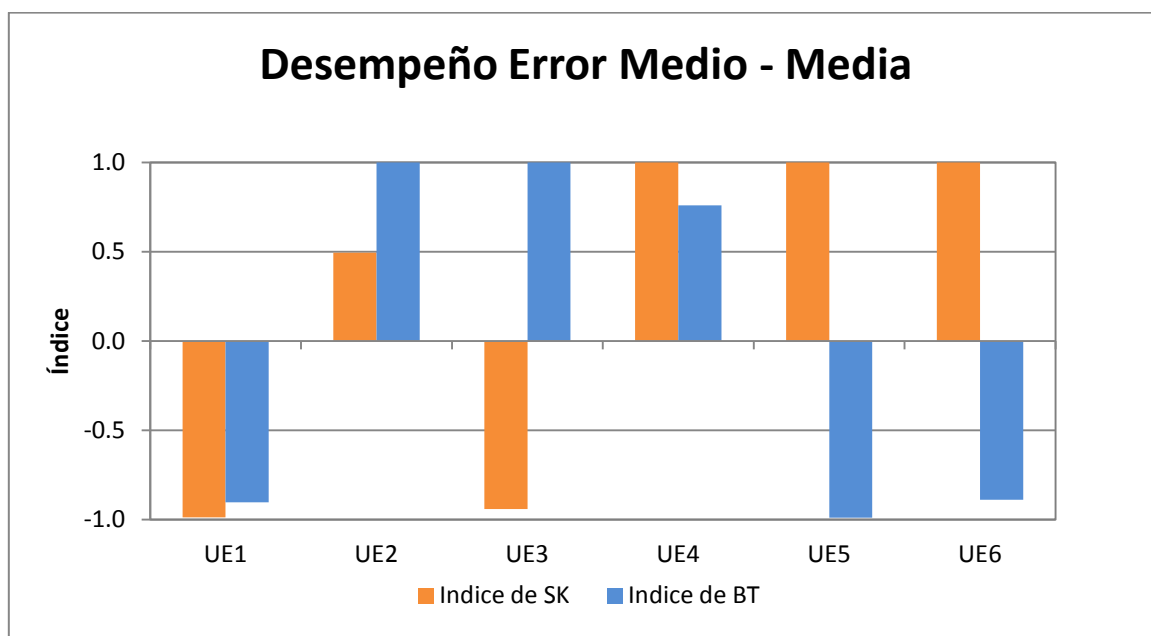
**Gráfico 82: Diferencia absoluta error medio vs 0.0 – Media UE5.**

Es posible apreciar que los resultados con respecto a SK con mejores en la UE4 y UE5 en cada una de las sensibilizaciones lo que queda representado por el índice de desempeño equivalente a 1 para SK en las UE4, UE5 y UE6 como se indica en la Tabla 38. Con respecto al desempeño de RDR versus BT, el índice respectivo indica un mejor desempeño para la UE4 con un valor de 0.76 y 13 de 17 sensibilizaciones con mejor resultados, mientras que para la UE5 se obtiene un peor desempeño con un índice de -0.99 y 1 de 17 sensibilizaciones con mejor resultado.

UE	RDR vs SK – Media			RDR vs BT - Media		
	Nº de Mejoras	Nº de Mejoras [%]	Índice	Nº de Mejoras	Nº de Mejoras [%]	Índice
UE1	1	6%	<b>-0.99</b>	3	18%	<b>-0.90</b>
UE2	12	71%	<b>0.49</b>	17	100%	<b>1.00</b>
UE3	3	18%	<b>-0.94</b>	17	100%	<b>1.00</b>
UE4	17	100%	<b>1.00</b>	13	76%	<b>0.76</b>
UE5	17	100%	<b>1.00</b>	1	6%	<b>-0.99</b>
UE6	17	100%	<b>1.00</b>	2	12%	<b>-0.89</b>

**Tabla 38: Desempeño RDR vs SK y BT – Error medio.**

De la Tabla 38 es posible ver que se obtiene un mejor desempeño con respecto al SK en la UE2, UE4, UE5 y UE6, mientras que lo mismo sucede con respecto al BT en la UE2, UE3 y UE4.



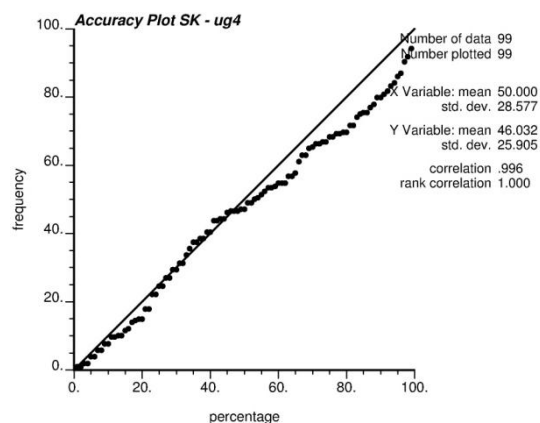
**Gráfico 83: Desempeño RDR vs SK y BT – Error medio.**

### 6.12.7 Accuracy plot

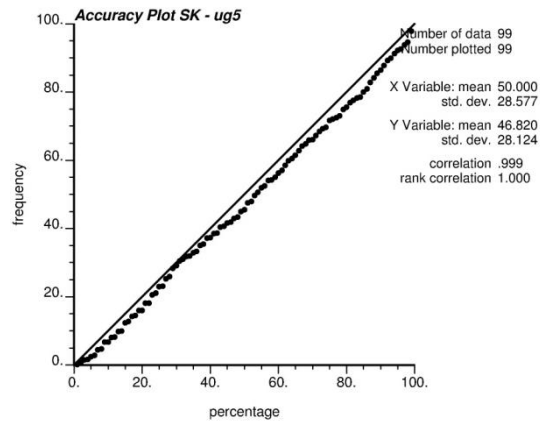
Un *accuracy plot* es un gráfico que permite determinar que tan bien caracterizada está la dispersión de los datos en los resultados de las simulaciones. Para esto, en cada nodo con dato de validación, se determina la distribución dada por los 100 valores simulados con lo cual se tiene una cantidad de distribuciones equivalente a la cantidad de datos del conjunto de validación C30. Luego, para un nivel de significancia dado, se determina el intervalo de confianza correspondiente en cada una de las distribuciones determinadas y se calcula que porcentaje de los datos validadores C30 se encuentran dentro del intervalo de confianza de su respectiva distribución. De esta manera, una simulación que caracteriza de buena manera la dispersión de los datos originales debiese dibujar una línea lo más cercana a la recta identidad al graficar la frecuencia de datos dentro del intervalo de confianza de su respectiva distribución versus el nivel de significancia del intervalo de confianza. Esto significa que para un nivel de significancia de un 95% (es decir, dentro del intervalo de confianza se encuentra el 95% de la distribución), el 95% de los datos validadores debiese estar dentro de sus respectivos intervalos de confianza.

Se presentan los *accuracy plot* de SK, BT y cada una de las sensibilizaciones de RDR. El eje X corresponde al porcentaje de la distribución dentro del intervalo de confianza respectivo mientras que el eje Y representa la frecuencia de datos validadores dentro de los intervalos de confianza respectivos.

#### 6.12.7.1 Kriging simple

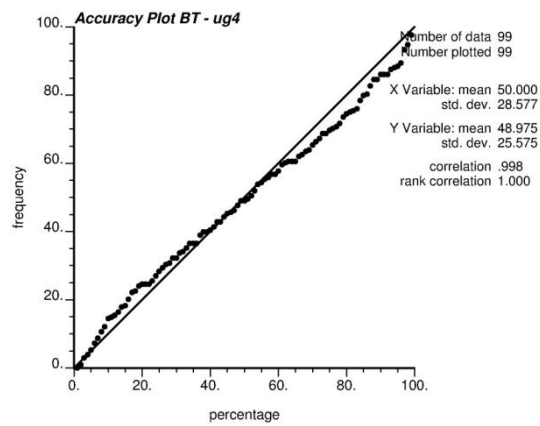


**Gráfico 84: Accuracy plot – Kriging simple – UE4.**

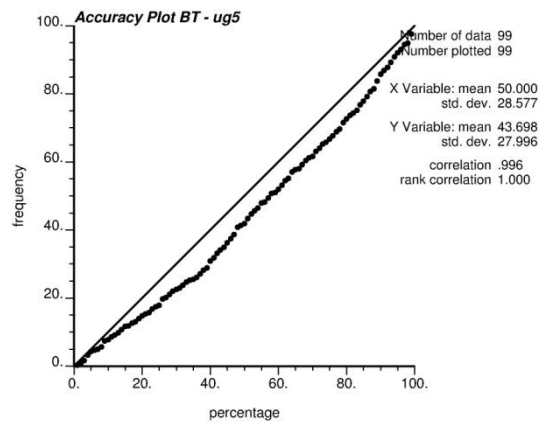


**Gráfico 85: Accuracy plot – Kriging simple – UE5.**

6.12.7.2 Kriging de residuos

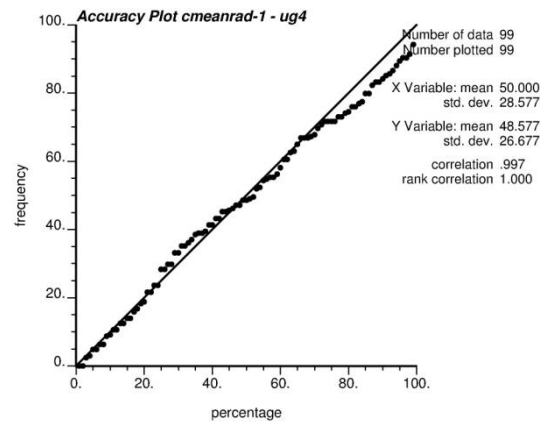


**Gráfico 86: Accuracy plot – Kriging de residuos – UE4.**

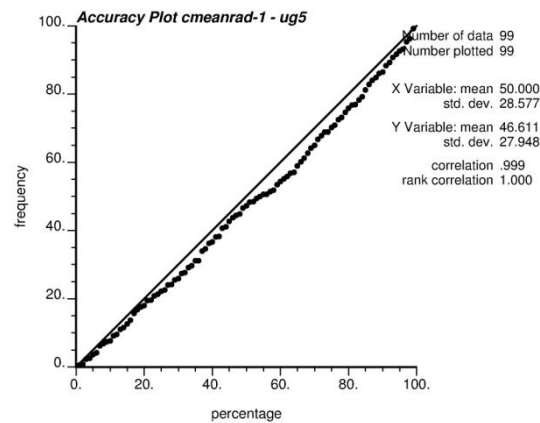


**Gráfico 87: Accuracy plot – Kriging de residuos – UE5.**

### 6.12.7.3 Kriging con rechazo – Caso base



**Gráfico 88: Accuracy plot – Kriging con rechazo (caso base) UE4.**



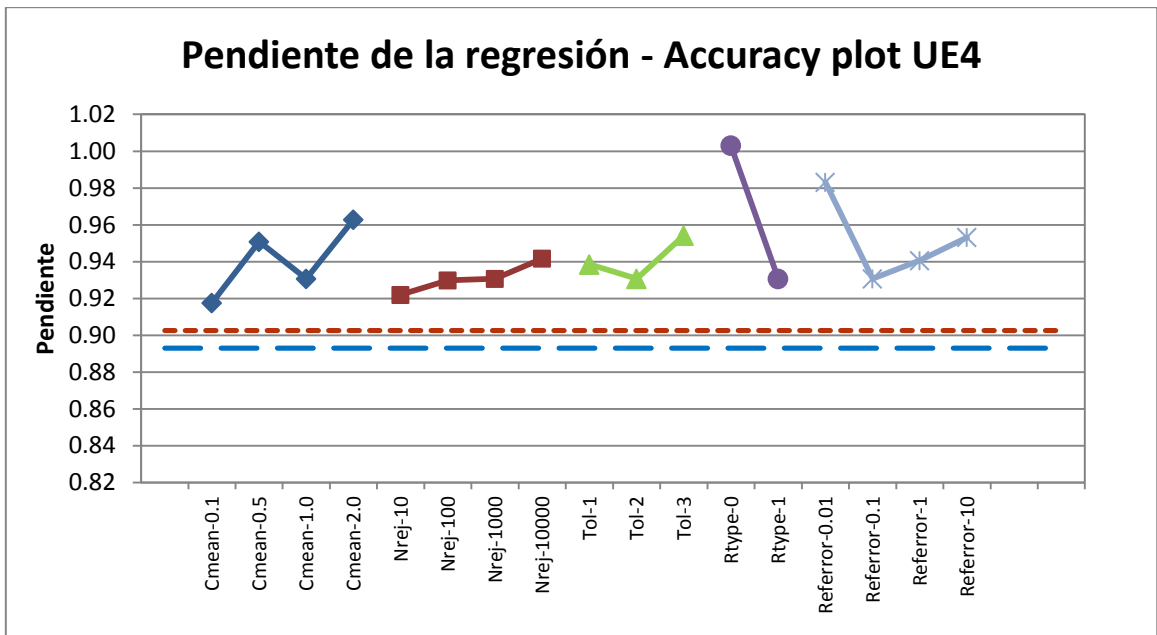
**Gráfico 89: Accuracy plot – Kriging con rechazo (caso base) UE5.**

Se puede apreciar un mejor ajuste a la recta identidad en el caso del RDR para la UE4 en comparación con el SK y con el BT. De hecho, en el caso del SK se aprecia un marcado sesgo para valores alto del porcentaje. En el caso de la UE5 se aprecia claramente un mejor desempeño del RDR con respecto al BT, pero la comparación con respecto al SK no es tan clara. De esta manera, se presenta en la sección siguiente, un análisis de desempeño considerando como estadístico de medición la pendiente de la regresión de la frecuencia versus el porcentaje, la cual, mientras más cercana a 1.0 indicará una mejor caracterización de la incertidumbre de los datos originales.

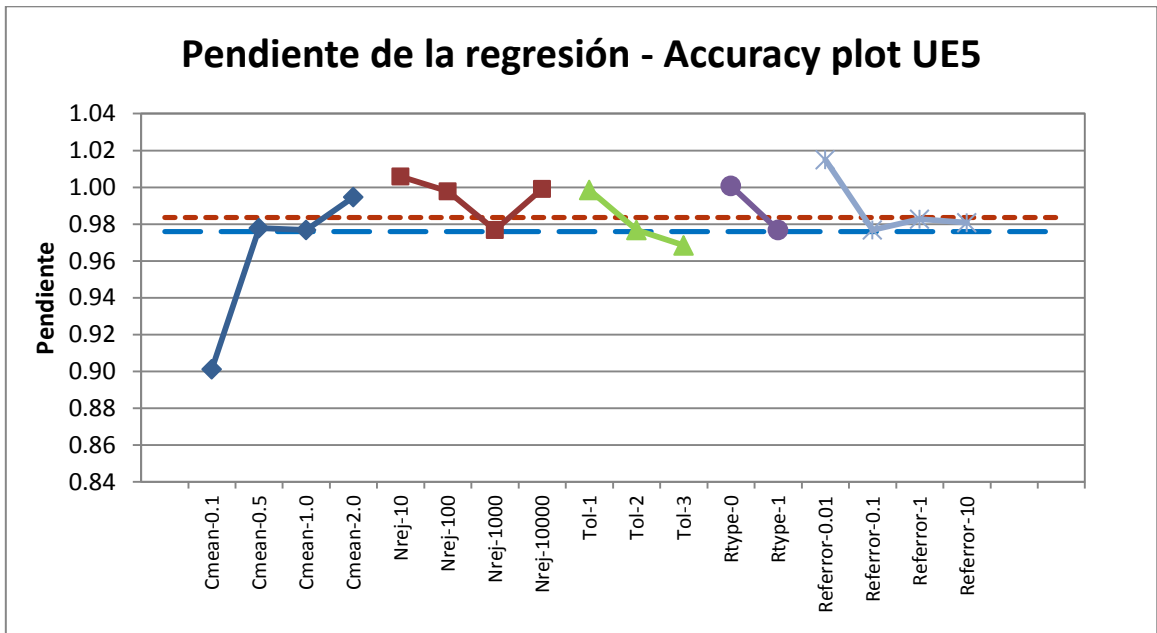
#### *6.12.7.4 Desempeño del Kriging con rechazo*

Para la determinación del desempeño del Kriging con rechazo en cuanto a caracterización de la incertidumbre de los datos originales se utiliza la pendiente de la regresión de cada accuracy plot. Se hace referencia a la regresión lineal entre el porcentaje de la distribución y la frecuencia de datos validadores dentro de los intervalos de confianza respectivos, cuya pendiente mientras más cercana sea a 1 implicará una mejor caracterización de la incertidumbre. De esta manera se obtiene un valor para cada una de las 17 sensibilizaciones de RDR y un valor para los casos SK y BT contra los cuales se contrastan los resultados.

La pendiente de la regresión de los accuracy plot toma valores desde 0.901 para cmean-0.1 hasta 1.015 para referror-0.01 en la UE5, mientras que en la UE4 toma valores que van desde 0.893 para BT hasta 1.003 para rtype-0. Tal como se explica en la sección 6.12.5, el valor objetivo en este estadístico es equivalente a 1.0 por lo tanto un valor más cercano a una pendiente igual a 1.0 implica un mejor desempeño. En el Gráfico 90 y Gráfico 91 se muestra el valor para la pendiente de la regresión de los accuracy plot de cada sensibilización así como el resultado obtenido mediante SK y BT presentados como líneas punteadas naranja y azul respectivamente. Se puede apreciar de estos gráficos que en la UE4 cada una de las sensibilizaciones de RDR tiene un valor más cercano a 1.0 para la pendiente de los accuracy plots en comparación con SK y BT. Sin embargo esto no sucede en la UE5 donde si bien el caso de SK está un poco por sobre BT, los resultados de las sensibilizaciones de RDR toman valores que van desde desempeños peores que BT hasta mejores que el SK. De esta manera se realiza el cálculo de la diferencia absoluta con respecto al valor objetivo de la pendiente de la regresión de los accuracy plots mostrada en el Gráfico 92 y el Gráfico 93. En el Anexo Q se presentan los resultados del resto de las UE. Es posible apreciar que los resultados del RDR en comparación con el BT empeoran en la UE1, UE2 y UE6, mientras que las mejoras ocurren claramente en la UE3 y UE4. Con respecto al desempeño del RDR con respecto al SK, se aprecia una mejora clara sólo en la UE4.



**Gráfico 90: Pendiente de la regresión – Accuracy plot UE4.**



**Gráfico 91: Pendiente de la regresión – Accuracy plot UE5.**

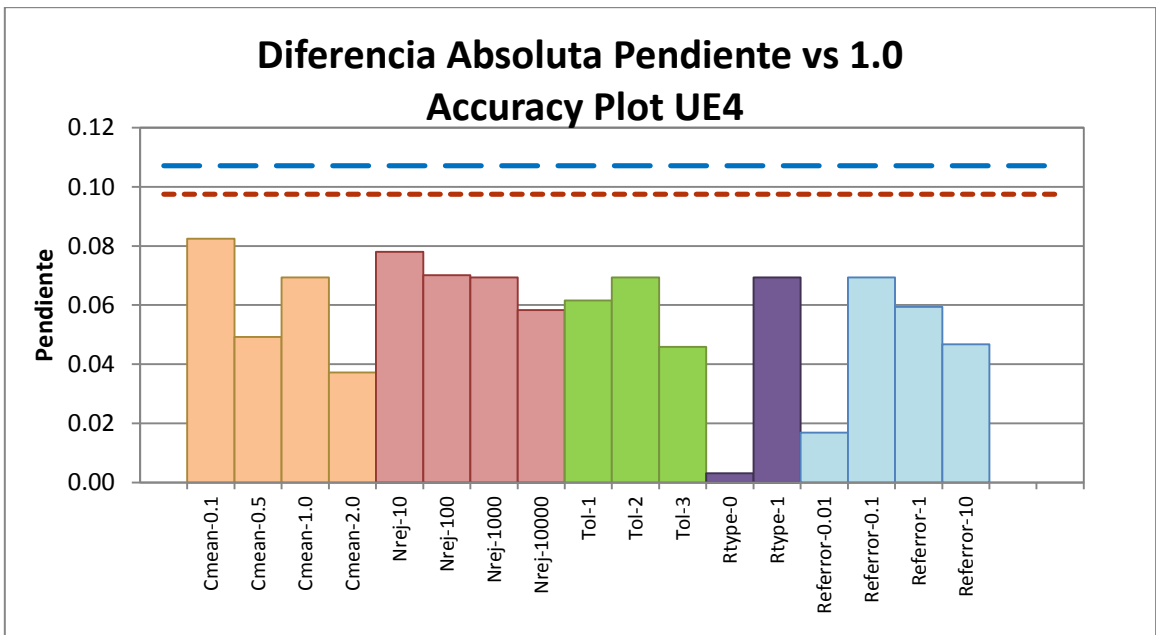


Gráfico 92: Diferencia absoluta pendiente vs 1 – Accuracy plot UE4.

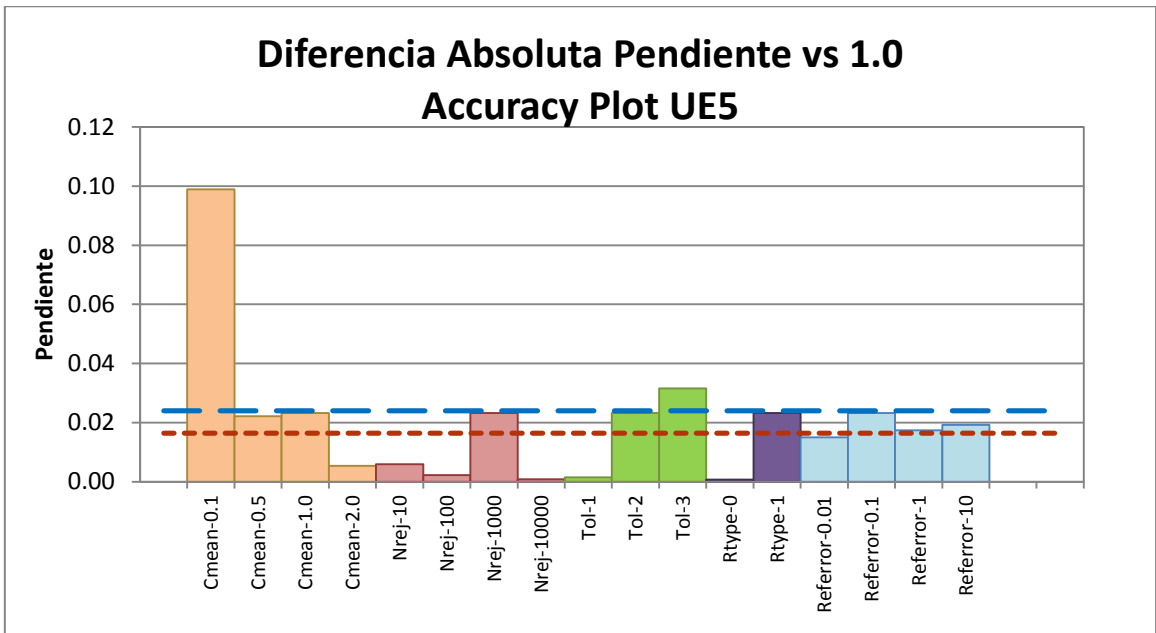


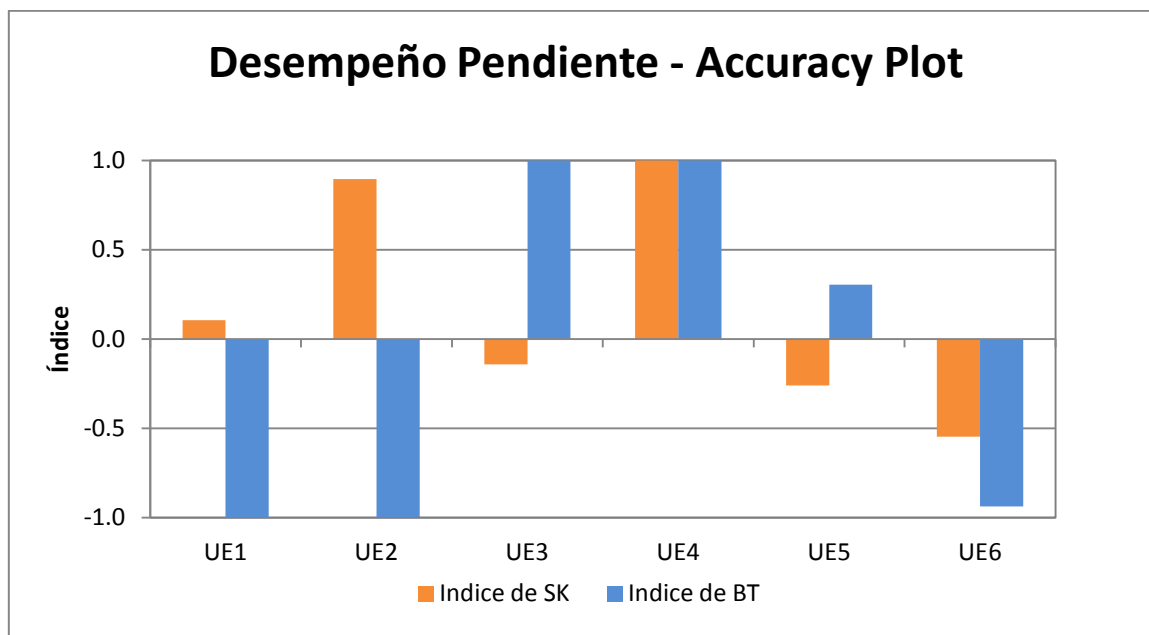
Gráfico 93: Diferencia absoluta pendiente vs 1 – Accuracy plot UE5.



Los resultados expresados por los gráficos quedan resumidos en la Tabla 39 donde se presentan los índices correspondientes a la comparación versus SK y BT de la pendiente de la regresión de los accuracy plots. Se puede apreciar que con respecto al BT, se obtienen mejoras en la UE3, UE4 y UE5 siendo las dos primeras mejoras en cada una de las 17 sensibilizaciones. Con respecto al índice de SK, las mejoras ocurren en la UE1, UE2 y UE4, siendo esta última la única que obtiene mejoras en cada una de las 17 sensibilizaciones. De esta manera se muestra que la UE4 es la única que tiene un mejor desempeño en cada una de las 17 sensibilizaciones ya sea comparando con el SK o BT.

UE	RDR vs SK - Media			RDR vs BT - Media		
	N° de Mejoras	N° de Mejoras [%]	Índice	N° de Mejoras	N° de Mejoras [%]	Índice
UE1	12	71%	<b>0.10</b>	0	0%	<b>-1.00</b>
UE2	11	65%	<b>0.90</b>	1	6%	<b>-1.00</b>
UE3	7	41%	<b>-0.14</b>	17	100%	<b>1.00</b>
UE4	17	100%	<b>1.00</b>	17	100%	<b>1.00</b>
UE5	7	41%	<b>-0.26</b>	15	88%	<b>0.30</b>
UE6	5	29%	<b>-0.55</b>	2	12%	<b>-0.94</b>

**Tabla 39: Desempeño RDR vs SK y BT – Pendiente accuracy plot.**

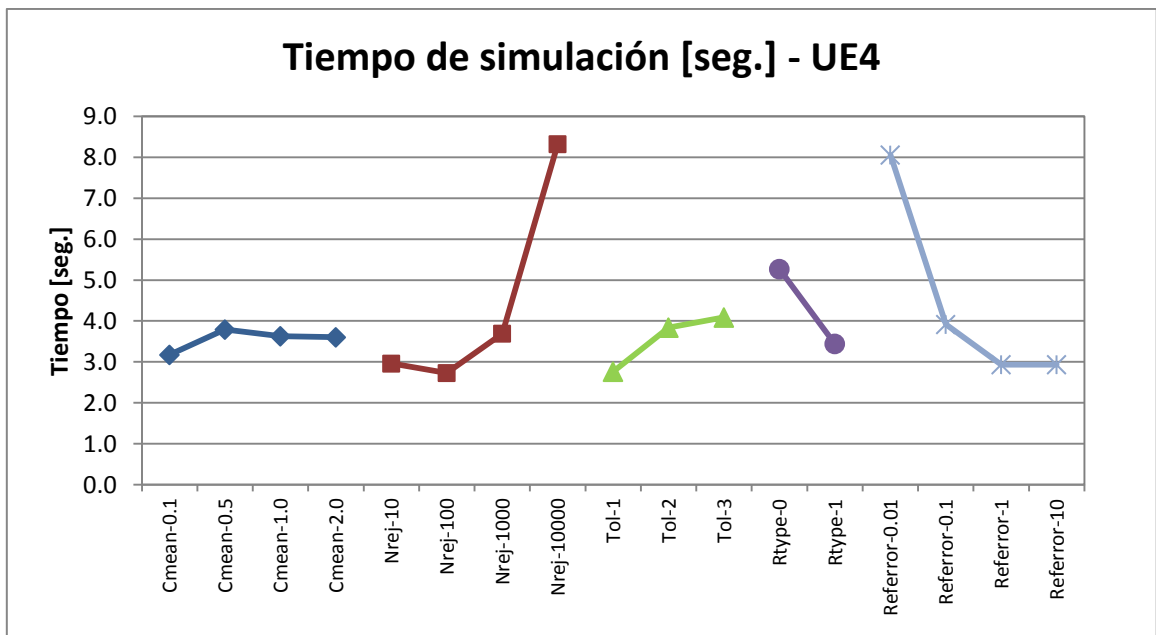


**Gráfico 94: Desempeño RDR vs SK y BT – Pendiente accuracy plot.**

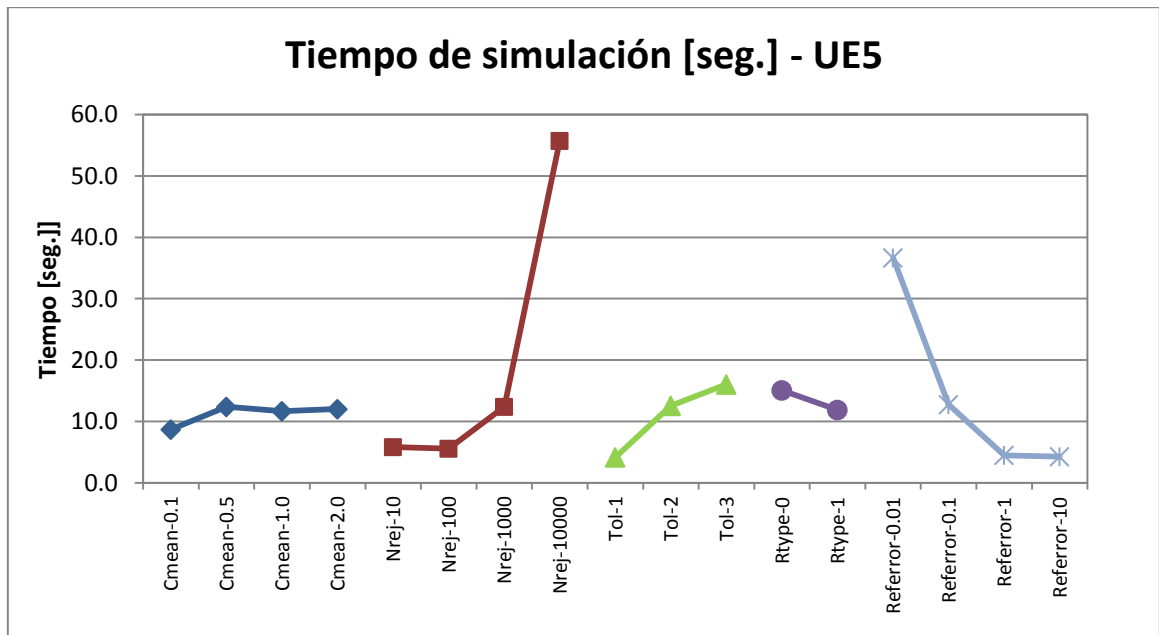
### 6.12.8 Tiempos de simulación

Los tiempos de simulación permiten estudiar el desempeño del algoritmo de rechazo frente a la utilización de distintos parámetros mediante las sensibilizaciones. Se presenta el valor mínimo, máximo, medio y la varianza de las 100 simulaciones en el SK y en cada una de las 17 sensibilizaciones.

Los tiempos de simulación (Anexo R) tienen la misma tendencia en todas las UE. De esta manera se aprecia que aumenta el tiempo de simulación cuando: aumenta el tamaño de la vecindad de búsqueda de la media dado que se integran más datos para los cálculos, aumenta el número de rechazos, la tolerancia se vuelve más estricta o el error de referencia es más pequeño dado que se realiza un mayor número de rechazos y por ende un mayor número de cálculos. También se aprecia que el tiempo de simulación baja en todas las UE cuando se usa un rechazo probabilístico a cuando se utiliza uno determinista, sin embargo esto no es generalizable dado el desempeño del rechazo determinista depende del rango de valores de la variable a simular.



**Gráfico 95: Tiempo de simulación medio – UE4.**

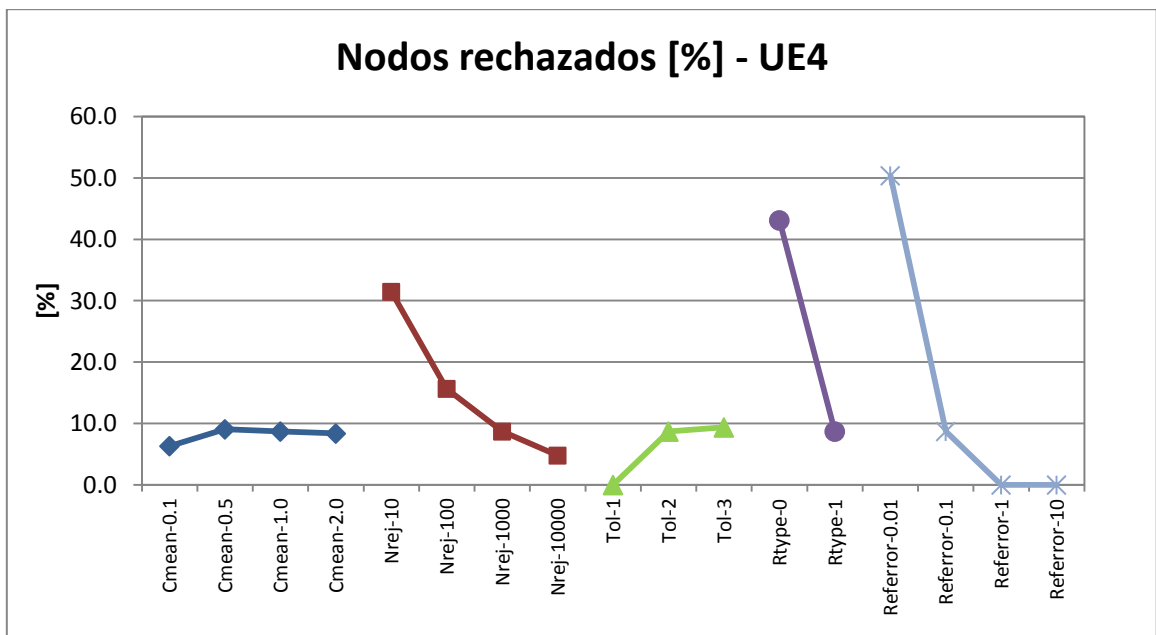


**Gráfico 96: Tiempo de simulación medio – UE5.**

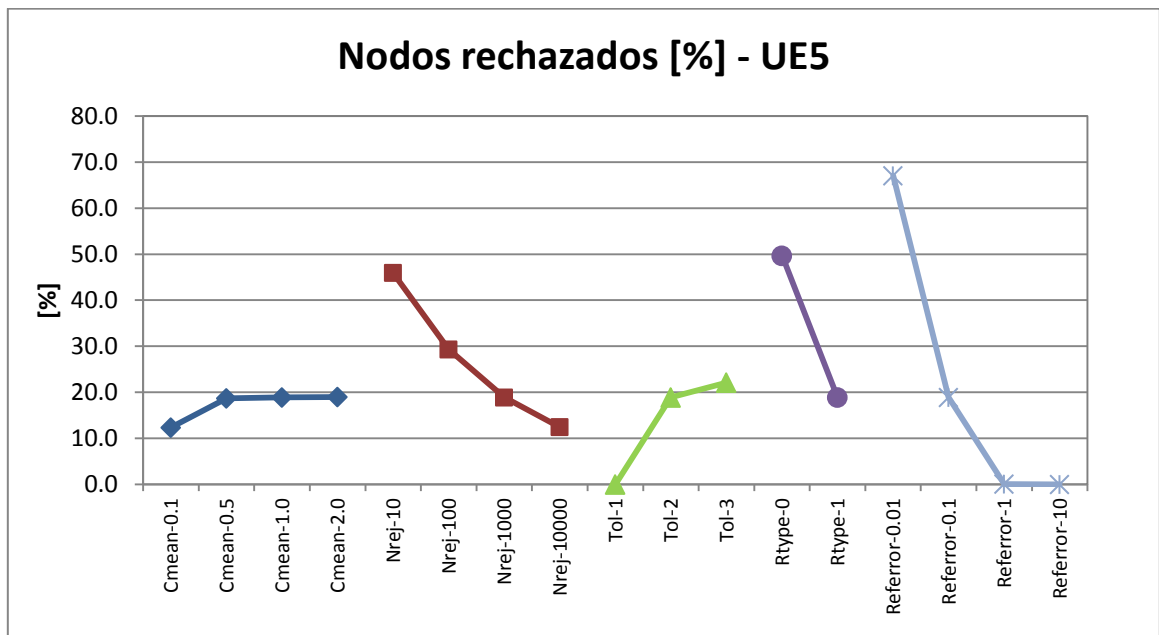
### 6.12.9 Nodos rechazados

El número de nodos rechazados corresponde al porcentaje de nodos que superan el número máximo aceptable de rechazos y que deben ser resimulados. Se presenta el valor mínimo, máximo, medio y la varianza de las 100 simulaciones en el SK y en cada una de las 17 sensibilizaciones.

Al igual que los tiempos de simulación, el porcentaje de los nodos rechazados tienen la misma tendencia en cada una de las UE. De hecho, se tiene la misma tendencia que los tiempos de simulación para cada parámetro excepto para el número aceptable de rechazos dado que a menor número aceptable, mayor es el número de nodos rechazados que se deben resimular. El mayor porcentaje de rechazos se produce en los escenarios de error de referencia 0.01, número aceptable de rechazos 10 y rechazo determinista, mientras que el menor porcentaje se produce para los casos de baja tolerancia y tamaño de la vecindad de búsqueda para la media de 0.1.



**Gráfico 97: Nodos rechazados medio – UE4.**

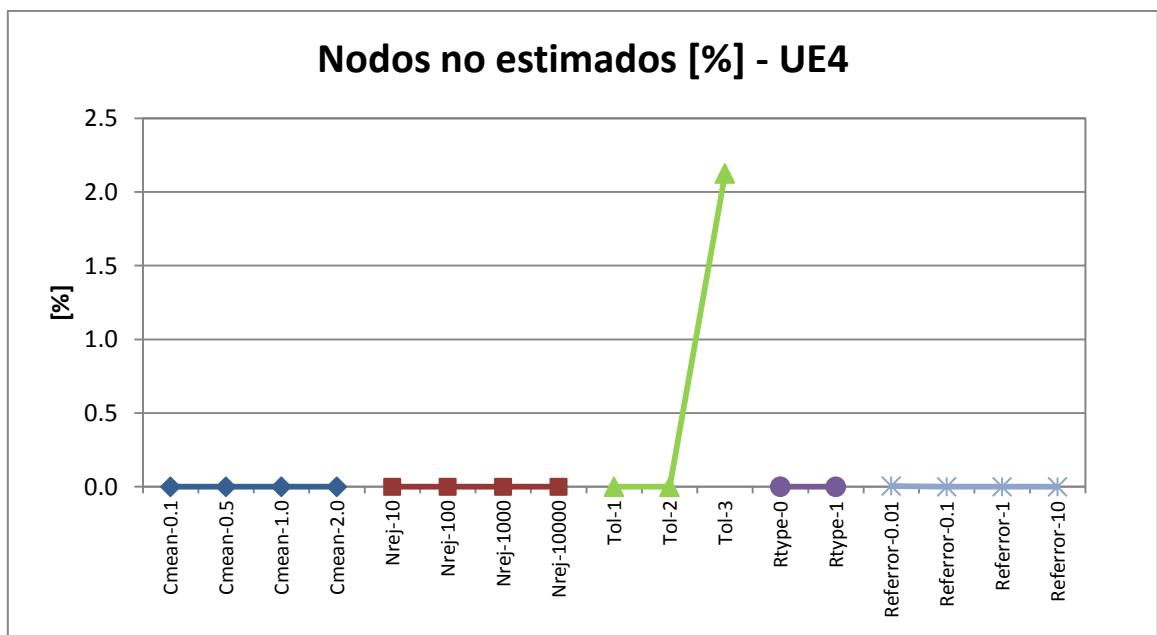


**Gráfico 98: Nodos rechazados medio – UE5.**

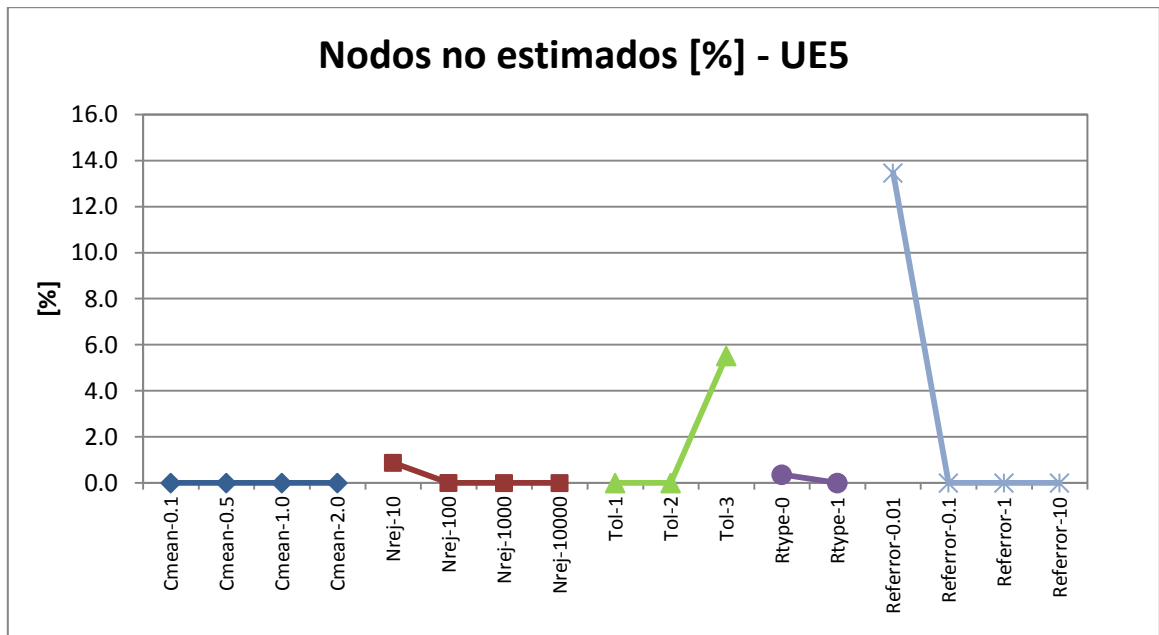
### 6.12.10 Nodos no simulados

El número de nodos no simulados corresponde al porcentaje de nodos que en una etapa de resimulación, aún superan el número máximo aceptable de rechazos. Se presenta el valor mínimo, máximo, medio y la varianza de las 100 simulaciones en el SK y en cada una de las 17 sensibilizaciones.

Para los nodos no simulados se tiene la misma tendencia que para el caso de los nodos rechazados, lo cual es intuitivo ya que esto no es más que una segunda etapa de aceptación/rechazo. En cada una de las UE en la mayoría de los escenarios se tiene que el porcentaje de nodos no simulados no es mayor al 1%. De hecho para todas las sensibilizaciones de tamaño de la vecindad de búsqueda de la media (0.1, 0.5, 1 y 2), las tolerancias 1 y 2 (alta y de amplio rango), el rechazo probabilístico, los números aceptables de rechazos de 1,000 y 10,000 y los errores de referencia 0.1, 1 y 10, se tiene que el número de nodos no simulados es de 0% en ambas UE. Es decir, en al menos 12 de los 17 escenarios la grilla queda totalmente simulada.



**Gráfico 99: Nodos no simulados medio – UE4.**



**Gráfico 100: Nodos no simulados medio – UE5.**

### 6.13 Desempeño global Kriging con rechazo según estadísticas de validación

Dado que en términos de estadísticas de validación, ya sea evaluando las simulaciones en su desempeño como estimación o como herramienta para caracterizar la incertidumbre, los resultados no son siempre los mismos, es importante tener una visión global del desempeño del algoritmo de rechazo frente a los algoritmos de contraste. De esta manera se presenta en la Tabla 40 el desempeño global de las estadísticas de validación de la simulación como estimación. Para esto se considera el total de mejoras que se realizan en las tres estadísticas (coeficiente de determinación, pendiente de la regresión y error medio) y el promedio de los índices respectivos ya sea para comparar versus SK o BT. La tabla muestra que al comparar el desempeño del RDR versus SK considerando el desempeño de las simulaciones como estimación, se encuentran mejoras en las UE2, UE3 y UE4, siendo la UE2 la de mayor mejora con un índice de 0.81, seguida de cerca por la UE4 con un índice de 0.63. Por otra parte, al comparar el RDR versus el BT se tiene que las mejoras se presentan en las mismas UE que al comparar contra el SK agregando ahora mejoras en la UE6. En este caso el mayor índice se obtiene en a UE2 con un índice de 1.0 (lo que indica mejoras en todas las sensibilizaciones en las tres estadísticas), seguida de cerca por la UE4 con un índice de 0.92. Por último se presenta en la Tabla 41 el promedio de

los índices globales de SK y BT con lo cual se obtiene un único valor de desempeño del algoritmo que permite tener una relación de orden entre las UE. De esta manera, considerando el desempeño de las simulaciones como estimación, se obtienen mejoras al usar el algoritmo de rechazo en la UE2, UE3 y UE4, siendo la de mayor mejora la UE4 con un índice de 0.78, seguida por la UE2 y UE3 con índices de 0.59 y 0.57 respectivamente.

UE	Global RDR vs SK			Global RDR vs BT		
	N° de Mejoras	N° de Mejoras [%]	Índice	N° de Mejoras	N° de Mejoras [%]	Índice
UE1	19	37%	<b>-0.33</b>	8	16%	<b>-0.88</b>
UE2	27	53%	<b>0.81</b>	34	67%	<b>0.37</b>
UE3	17	33%	<b>0.13</b>	51	100%	<b>1.00</b>
UE4	35	69%	<b>0.63</b>	47	92%	<b>0.92</b>
UE5	22	43%	<b>-0.07</b>	17	33%	<b>-0.33</b>
UE6	18	35%	<b>-0.33</b>	30	59%	<b>0.19</b>

**Tabla 40: Desempeño RDR vs SK y BT – Validación simulación como estimación.**

UE	Promedio Global RDR vs SK y BT		
	N° Mejoras	N° Mejoras [%]	Índice
UE1	<b>13.5</b>	<b>26%</b>	<b>-0.60</b>
UE2	<b>30.5</b>	<b>60%</b>	<b>0.59</b>
UE3	<b>34.0</b>	<b>67%</b>	<b>0.57</b>
UE4	<b>41.0</b>	<b>80%</b>	<b>0.78</b>
UE5	<b>19.5</b>	<b>38%</b>	<b>-0.20</b>
UE6	<b>24.0</b>	<b>47%</b>	<b>-0.07</b>

**Tabla 41: Desempeño promedio RDR vs SK y BT – Validación simulación como estimación.**

Dado que se ha analizado ya el desempeño de las simulaciones como estimación y como herramienta para caracterizar la incertidumbre y en ambas el mejor resultado se obtiene en la UE4 al usar el algoritmo de rechazo, se presenta un último ejercicio que consiste en determinar un índice de desempeño global considerando los cuatro estadísticos utilizados. De esta manera se obtiene la Tabla 42 y Tabla 43 las cuales tienen valores levemente diferentes a la Tabla 40 y Tabla 41. De hecho se sigue la misma tendencia (mejora o empeoramiento) en cada una de las UE al comparar versus SK o BT excepto en la comparación contra BT de la UE6 en la cual se pasa de una mejora a un empeoramiento leve con un índice de -0.09. De estas últimas tablas es posible determinar que ya sea al comparar contra SK o BT las mejoras en términos globales se producen en la UE2, UE3 y UE4 obteniéndose el mejor desempeño del algoritmo en la UE4.

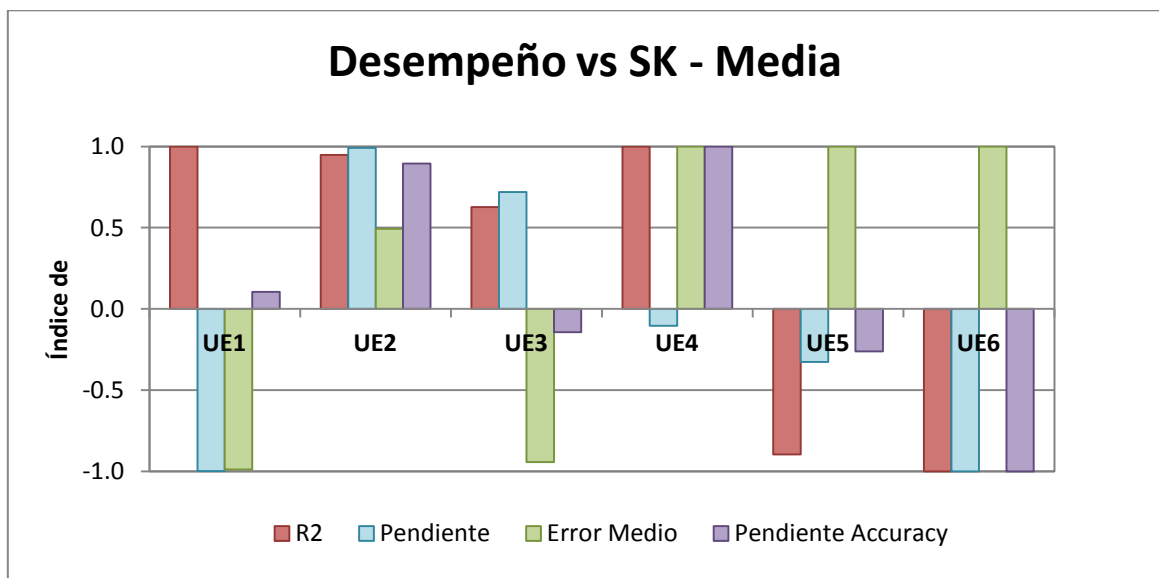
UE	Global RDR vs SK			Global RDR vs BT		
	N° de Mejoras	N° de Mejoras [%]	Índice	N° de Mejoras	N° de Mejoras [%]	Índice
UE1	31	46%	<b>-0.22</b>	8	12%	<b>-0.91</b>
UE2	38	56%	<b>0.83</b>	35	51%	<b>0.03</b>
UE3	24	35%	<b>0.07</b>	68	100%	<b>1.00</b>
UE4	52	76%	<b>0.72</b>	64	94%	<b>0.94</b>
UE5	29	43%	<b>-0.12</b>	32	47%	<b>-0.17</b>
UE6	23	34%	<b>-0.31</b>	32	47%	<b>-0.09</b>

**Tabla 42: Desempeño RDR vs SK y BT – Global.**

UE	Promedio Global RDR vs SK y BT		
	N° Mejoras	N° Mejoras [%]	Índice
UE1	19.5	29%	<b>-0.57</b>
UE2	36.5	54%	<b>0.43</b>
UE3	46.0	68%	<b>0.53</b>
UE4	58.0	85%	<b>0.83</b>
UE5	30.5	45%	<b>-0.15</b>
UE6	27.5	40%	<b>-0.20</b>

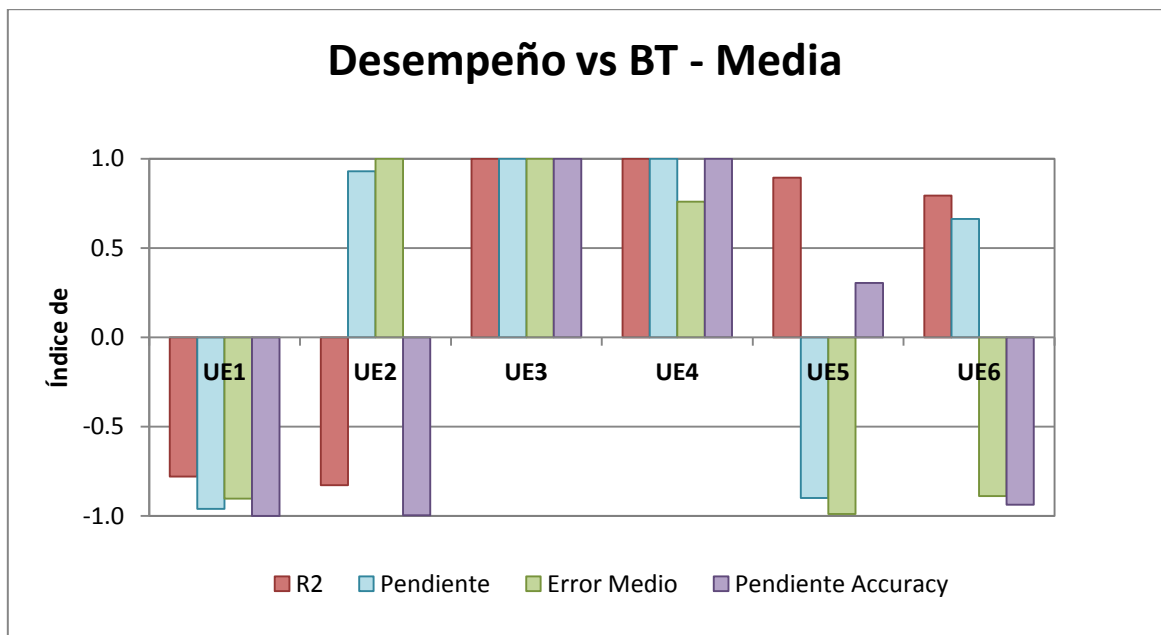
**Tabla 43: Desempeño promedio RDR vs SK y BT – Global.**

El Gráfico 101 y el Gráfico 102 muestran el desempeño global del algoritmo de rechazo considerando su comparación versus SK y BT respectivamente.



**Gráfico 101: Desempeño RDR vs SK – Global.**





**Gráfico 102: Desempeño RDR vs BT – Global.**

## 7 APLICACION GENERAL DEL ALGORITMO

El algoritmo desarrollado posee una serie de parámetros propios de su aplicación, por lo cual se presenta en esta sección, una breve descripción de los parámetros más relevantes así como recomendaciones en cuanto a valores recomendables de cada parámetro y aplicación general del algoritmo.

Dado que la hipótesis que subyace a la construcción del algoritmo no proviene de una derivación matemática, es recomendable, mas no obligatorio, realizar un proceso de validación del uso del algoritmo previo a su aplicación con el fin de corroborar que este efectivamente obtiene buenos resultados. Para esto se recomienda considerar un porcentaje de los datos condicionantes mucho menor a la base de datos total (del orden de un 10 a 30%), seleccionados de manera aleatoria dentro del espacio a simular y aplicar el algoritmo de rechazo considerando esta base de datos reducida. Con estos resultados chequear la reproducción del histograma y variograma de los datos originales. Esta base de datos reducida puede ser utilizada para posteriores procesos de validación. Si bien, cada set de datos a simular tiene sus particularidades, se tienen las siguientes recomendaciones de manera general en cuanto a los parámetros propios del algoritmo:

- *Vecindad de búsqueda de la media local:* El tamaño de esta vecindad es uno de los parámetros que puede tener mayor influencia en los resultados. Esto sucede cuando se toma un tamaño de la vecindad de búsqueda de la media local muy pequeño en comparación con la vecindad de búsqueda de datos condicionantes. Se recomienda no utilizar valores muy alejados de la vecindad de búsqueda de datos condicionantes, es decir, considerar un rango de valores entre 0.5 a 1.5. Para valores del orden de 0.1, si bien las simulaciones se ajustan de manera más estricta a la deriva impuesta, el variograma de las simulaciones no reproduce de buena manera el variograma de los datos originales.
- *Número de rechazos:* El número de rechazos no tiene mayor influencia en los resultados. Su efecto está en los tiempos de simulación y en el porcentaje simulados de la grilla, por lo cual se debe utilizar un parámetro que mantenga un balance entre estos resultados. Para un set de datos donde la deriva impuesta es coherente con el comportamiento de los datos, valores del orden de 10 a 100 obtienen bajos tiempos de simulación pero poco porcentaje de la grilla simulada. Valores del orden de 10,000 hacia arriba obtienen un

gran porcentaje de la grilla simulada, pero muy altos tiempos de simulación. Se recomienda utilizar un valor de 1,000 rechazos aceptables. Para yacimientos con deriva impuesta que no guarden relación con los datos, es recomendable aumentar este valor dado que el ajuste de las simulaciones a la deriva implica un mayor número de rechazos y por ende baja el porcentaje simulado de la grilla.

- *Tolerancia dinámica:* Este parámetro, junto con el error de referencia en el caso de un mecanismo de rechazo probabilístico, tiene igual importancia que el tamaño de la vecindad de búsqueda de la media local. Su valor se fija según los requerimientos del modelador: en caso de querer simulaciones ajustadas a la deriva impuesta debe usarse valores bajos de tolerancia y en caso contrario, valores altos de tolerancia son aceptables. Para un rechazo determinista, el rango de valores de la tolerancia corresponde al rango de valores de la variable a simular por lo cual se deben fijar valores de tolerancia dependiendo de dicha variable. Para un rechazo probabilístico, el rango de valores de la tolerancia va desde 0 a 1 (sin incluir los extremos), por lo cual para una baja tolerancia considerar rangos desde 0.050 a 0.001 y para una alta tolerancia considerar rangos de 0.999 a 0.950. En caso encontrarse en una postura indiferente a la tolerancia, un rango de valores de amplio rango es recomendado (valores desde 0.999 a 0.001).
- *Mecanismo de rechazo:* Si bien el mecanismo de rechazo no es tan influyente como el tamaño de la vecindad, la tolerancia dinámica o el error de referencia, es importante tener en consideración que a la luz de los resultados de ambos casos de estudios, un rechazo determinista es más propenso a presentar un sesgo en la reproducción de la deriva. Más aún, un rechazo probabilístico, si bien es un tanto más complejo de implementar dado que se debe definir el error de referencia, otorga mayor flexibilidad en los resultados. Por esto se recomienda un rechazo probabilístico el cual se define utilizando un valor 1 para este parámetro.
- *Error de referencia:* Es un tipo de tolerancia necesario de definir cuando se implementa un rechazo probabilístico, por lo cual las consideraciones al respecto son similares a la tolerancia dinámica. En este caso valores bajos del error de referencia (del orden de 0.01 a 0.1) deben ser utilizados en caso de querer simulaciones ajustadas a la deriva impuesta mientras valores más altos (del orden de 1 en adelante) se utilizan en caso contrario. No es recomendable utilizar valores muy bajos (0.01) de error de referencia dado que los resultados no reproducen de buena manera el variograma de los datos originales.

## 8 ANALISIS Y CONCLUSIONES

### *Sobre el estado del arte*

Se desprende la revisión bibliográfica que si bien existen diversas metodologías para enfrentar el problema de una simulación no estacionaria, éstas tendrán condicionada su utilidad a diversos aspectos del escenario de estimación y a la naturaleza del problema. De esta manera los resultados obtenidos y por ende la utilidad del algoritmo como nueva herramienta de simulación están condicionados por las características propias del caso de estudio.

### *Sobre el caso de estudio I*

El caso de estudio I permite apreciar de forma simple y rápida los efectos sobre los resultados de la sensibilización de los parámetros propios del algoritmo desarrollado, es decir, tamaño de la vecindad de búsqueda de la media condicional, número de rechazos, tolerancia y mecanismo de rechazo. El mecanismo de rechazo hace referencia al tipo de rechazo, ya sea determinista o probabilístico, y en el caso de este último, al error de referencia como parámetro de referencia para el grado de tolerancia. Un menor error de referencia implica una menor tolerancia a valores muy disimiles de la deriva de referencia en el correspondiente nodo, y por ende se tiene un mayor número de rechazo, mayor tiempo de simulación y resultados más ajustados a la deriva. Esto se aprecia claramente en los resultados donde a medida que aumenta el error de referencia desde 0.01 hasta 10, la dispersión de los resultados en torno a la deriva de referencia también aumenta, lo cual se ve reflejado en el coeficiente de determinación de la deriva impuesta versus los resultados, tomando valores desde 92.9% hasta 74.4%. En cuanto a la reproducción del variograma, los resultados son más cercanos al variograma de los datos originales para errores de referencia bajos. En el caso del rechazo determinista, se hace énfasis en que en este caso, el rango de valores para la tolerancia dinámica depende de los valores de la variable a simular, a diferencia del rechazo probabilístico en el cual el rango de valores corresponde a un porcentaje de aceptación del error entra la deriva y el valor simulado. Esto último condiciona la elección del rango de tolerancias a la distribución de los valores a simular pudiéndose obtener tan sólo un escenario en cada rango. En cambio, un rechazo probabilístico actúa con tolerancias en términos porcentuales permitiendo además la aceptación de errores fuera del rango aceptable en virtud del carácter probabilístico del método. Se permite de esta manera un mayor grado de calibración de los resultados por medio del parámetro de error de referencia pudiéndose obtener más de un

escenario en cada caso con este método de rechazo. El rango de tolerancias muestra resultados esperables en cuanto a que para tolerancias bajas, los resultados se ajustan estrechamente entorno a la deriva ( $R^2$  de 92.6%) mientras que a medida que se aumenta la tolerancia, la dispersión de los valores en torno a la deriva también aumenta ( $R^2$  de 74.4%). En cuanto a la reproducción del variograma, se observa un mejor desempeño en la baja tolerancia y la tolerancia de amplio rango. Con respecto al tamaño de la vecindad de búsqueda de la media condicional, es posible apreciar que a medida que aumenta esta, es decir, a medida que se aumenta la integración horizontal, la dispersión de los resultados en torno a la deriva aumenta. En el caso extremo de una vecindad de búsqueda tan pequeña que sólo encuentre el mismo nodo a simular dentro de la vecindad, los resultados ajustarán de manera prácticamente exacta a la deriva (alcanzando valores de  $R^2$  de 99.5%) salvo por la componente probabilística del rechazo. Por último, el número de nodos a rechazar no muestra variaciones considerables en cuanto al ajuste a la deriva de los resultados, sin embargo presenta diferencias en cuanto a los tiempos de simulación, es decir, a mayor número de rechazos, mayor tiempo de simulación. Lo interesante a revisar es el porcentaje de nodos con valor simulado en función del tiempo lo cual muestra que con un número de rechazos de 1,000 ya se tiene un 100% de la grilla con valor simulado, por lo cual un aumento desde 1,000 a 10,000 rechazos no realiza mejoras en el porcentaje de grilla con valor simulado mientras cuadruplica el tiempo de cómputo. De esta manera, el caso de estudio I además de mostrar de forma rápida y fácil de visualizar los resultados, valida el algoritmo desarrollado en la medida que éste obtiene resultados acordes a los esperados en cuanto a la variación de sus parámetros propios, es decir, un mejor ajuste a la deriva cuando la vecindad de búsqueda para la media condicional es pequeña, cuando la tolerancia es baja y cuando el error de referencia es bajo, además de un mayor porcentaje de la grilla con valor simulado y un mayor tiempo de computo a medida que aumenta el número aceptable de rechazos.

Finalmente, con respecto a la sensibilización del variograma modelado se aprecia que los resultados están fuertemente condicionados por el efecto pepa utilizado, mientras que una variación en el alcance del modelo no tiene una fuerte incidencia en los resultados en cuanto a dispersión de las simulaciones en torno a la deriva impuesta y reproducción del variograma de los datos.

## *Sobre el caso de estudio II*

El caso de estudio II permite analizar los resultados obtenidos mediante el algoritmo versus simulaciones con Kriging simple (SK) y simulaciones con Kriging de residuos (BT), en una base de datos real consistente en 6 unidades de estimación (UE) con distintas características donde la especie de interés en este caso corresponde al cobre total medido sobre sondajes de diamantina compositados a 4 metros. Si bien el caso de estudio posee 6 unidades de estimación, el foco del análisis se centra en las unidades de estimación 5 (óxidos) y 4 (transición). El análisis estadístico no muestra presencia de datos aberrantes pero sí un muestreo irregular el cual es compensado con el correspondiente desagrupamiento. El análisis variográfico determina tres modelos omnidireccionales (UE 2-4-6), dos omnihorizontales (UE 1-5) y uno direccional (UE3) para el caso de los compósitos originales, mientras que para el caso de los residuos se determinan 4 modelos omnidireccionales (UE 1-4-5-6) y dos modelos omnihorizontales (UE 2-3). La generación del modelo de rocas se realiza sobre el volumen a simular con un tamaño de bloque de 10x10x10 y 277,200 nodos sobre la base de las UE. El modelo de derivas por su parte se realiza mediante la estimación de tendencia local a través del ejecutable 'KT3D' de GSLib. Si bien este modelo es arbitrario, no se precisa dentro de los alcances estudiar la generación del modelo de deriva por lo cual este punto consiste en una posibilidad de extensión del estudio en cuanto se puede chequear cómo varían los resultados con la imposición de distintos modelos de derivas los cuales no necesariamente deben estar bien ajustados a los datos condicionantes. Los resultados obtenidos se presentan en base a tres grupos de tres estadísticas cada uno: estadísticas propias de las simulaciones contrastadas versus los datos originales (histograma, variograma y media local); estadísticas de validación de la simulación en su desempeño como estimación (coeficiente de determinación, pendiente de la regresión y error medio) y en su desempeño en cuanto a caracterización de la incertidumbre (accuracy plot), considerando sondajes no condicionantes de la simulación como datos validadores; y estadísticas acerca del desempeño del algoritmo mismo (tiempo de simulación, número de nodos rechazados y número de nodos no simulados). Al respecto de cada estadística se tiene:

- *Histogramas:* Se aprecia un buen ajuste de los histogramas de las simulaciones en cada una de las sensibilizaciones para la UE 5, a excepción del rechazo determinista. Además, cada una de las sensibilizaciones presenta una dispersión de los histogramas de las

simulaciones en torno a los valores originales menor que la del SK a excepción de la alta tolerancia y el error de referencia 10. En cuanto a la UE4, se aprecia que los histogramas de las simulaciones presentan leyes levemente menores que el histograma de los datos originales. Si bien esto no indica un buen ajuste, se debe tener en consideración que los resultados son prácticamente iguales tanto en el SK como en el resto de las sensibilizaciones, excepto por  $c_{mean}=0.1$  y  $tol=3$ . Esto indica que el uso del algoritmo de rechazo si bien no mejora los resultados en términos comparativos con el SK en la UE4, tampoco los empeora. Esta última afirmación es extensible a cada una de las 6 unidades de estimación en estudio. Con respecto a los histogramas de BT, estos presentan un sesgo al comparar contra el histograma de los datos originales, el cual se hace claramente presente en la UE5 mientras que en la UE4 se aprecia para valores bajos de la ley. En este sentido, el algoritmo de rechazo presenta mejores resultados en cuanto a la reproducción del histograma de los datos en comparación con BT para las UE1, UE4, UE5 y UE6.

- *Variogramas:* El comportamiento de los resultados en la UE5 y la UE4 al variar los parámetros a sensibilizar tiene la misma tendencia dado por el caso de estudio I. Respecto a esto, mientras mayor es el ajuste de la deriva (bajo tamaño vecindad de búsqueda, baja tolerancia o bajo error de referencia), menor es la dispersión de los histogramas en torno al histograma de los datos originales. Sin embargo en el caso de los variogramas, mientras mayor es el ajuste, se produce un sesgo con respecto al variograma modelado, lo cual sucede en los casos de vecindad 0.1, error de referencia 0.01 y tolerancia 3 (baja tolerancia). Si bien este sesgo se presenta de manera notoria en la UE5, no sucede lo mismo en la UE4, donde el variograma no se ve notoriamente afectado en las sensibilizaciones de mayor sesgo de la UE5. Esto último indica que para la UE4 el sesgo incurrido en usar un conjunto de parámetros no óptimo para el caso de estudio no tiene gran impacto, no así en el resto de las unidades de estimación (Ver Anexo L). La dispersión de los resultados en torno al variograma modelado es siempre menor o igual al del SK en cada uno de los escenarios estudiados. Al respecto del BT, al igual que para el caso de los histogramas, existe un sesgo ahora en la media de los variogramas de las simulaciones con respecto al variograma modelado. En este sentido el algoritmo de rechazo presenta mejores resultado que el BT para las UE4 y UE5 en cuanto a la

reproducción del variograma modelado excepto para distancias altas de la dirección Z de la UE4.

- *Media local:* El comportamiento de los resultados en cada una de las sensibilizaciones es el mismo que en el caso de los histogramas, es decir, se tiene un buen ajuste de la deriva para la UE5 y la UE4 en todos los escenarios excepto en el rechazo determinista en la UE5. Además los resultados presentan una dispersión en torno a la deriva a ajustar menor o igual a los resultados obtenidos por SK. En este caso, el BT obtiene mejores resultados dado que posee un mejor ajuste en las cotas extremas de la coordenada Z que es justamente la zona con dificultades del SK y el algoritmo de rechazo. De hecho, el BT obtiene mejores resultados en todas las UE (Ver Anexo M) salvo en la UE3. Se debe tener en cuenta que pese a los mejores resultados del BT, para el caso de la UE4, dichos resultados son muy similares a los obtenidos con el algoritmo de rechazo.
- *Coefficiente de determinación:* Vale recordar en este punto los variograma experimentales y modelados presentados en el Gráfico 49. Se puede apreciar acá que en efecto es la UE4 la que posee una deriva más clara en el variograma experimental y variograma modelado. En este sentido se valida la utilidad del algoritmo como nueva herramienta de simulación dado que en el caso de presencia más notoria de deriva en el variograma, el algoritmo obtiene mejores resultados en cuanto al ajuste del modelo que el SK y el BT. De hecho, considerando sus índices de desempeño contra el SK ( $I_{SK}$ ) y contra el BT ( $I_{BT}$ ), la UE4 es la única que obtiene valores 1 para ambos índices, lo cual indica mejoras en cada una de las 17 sensibilizaciones con respecto tanto al SK como el BT. En el caso de la UE5, se obtienen mejores resultados sólo con respecto al BT (con un índice de 0.89) mientras que al comparar contra el SK los resultados empeoran (con un índice de -0.89).
- *Pendiente de la regresión:* En el caso de la pendiente de la regresión, se tiene un mejor desempeño del algoritmo de rechazo con respecto sólo al BT, obteniéndose un índice de desempeño igual 1 lo que implica mejoras en cada una de las 17 sensibilizaciones. Ahora bien con respecto al SK el índice no indica mejoras pues se obtiene un valor igual a -0.1. Si bien esto indica que en términos globales el desempeño es peor, el número de mejoras es igual a 9 y el valor del índice es muy cercano a 0 lo que indica un desempeño muy



similar al SK con un leve empeoramiento. En este sentido se puede decir que el algoritmo desarrollado tiene un desempeño similar o mejor en cuanto a la pendiente de la regresión y por tanto insesgo condicional, con respecto al SK y BT justamente para el caso que posee una deriva más clara en el variograma experimental y variograma modelado (UE4). Para el caso de la UE5, se obtiene un peor desempeño del algoritmo de rechazo con respecto al SK (índice de -0.33) y al BT (índice de -0.90).

- *Error medio:* Se tiene que con respecto al criterio de insesgo global el algoritmo desarrollado tiene un mejor desempeño que el SK y BT con índices de desempeño de 1 y 0.76 para la UE4. En el caso de la UE5 el mejor desempeño sólo se hace presente con respecto al SK con un índice de desempeño de 1 mientras que al comparar contra el BT, se obtiene un peor desempeño con un índice de -0.99.
- *Accuracy plot:* Los accuracy plots son los gráficos que permiten tener una medida de que tan bien está caracterizada la incertidumbre de los datos originales en las simulaciones. En este sentido, al igual que en el caso del coeficiente de determinación, la UE4 obtiene mejores desempeños que el SK y el BT con valores de 1 para ambos casos, lo cual significa mejoras en cada una de las 17 sensibilizaciones al comparar ya sea contra el SK o contra el BT. Esto permite validar la utilidad del algoritmo como nueva herramienta de simulación en cuanto a su caracterización de la dispersión de los datos originales en las simulaciones ya que justamente en la UE donde se presenta una deriva clara en el variograma experimental y modelado obtiene mejores resultados que los otros dos algoritmos en todas las sensibilizaciones. En el caso de la UE5 se obtiene un mejor desempeño con respecto al BT (con un índice de 0.30) y un peor desempeño con respecto al SK (con un índice de -0.26).
- *Tiempo de simulación:* Es interesante notar que los menores tiempos de simulación se obtienen en la UE4, si bien se podría atribuir esto a la cantidad de datos, esta UE es la tercera en menor cantidad de datos con 1,064, mientras que las UE1 y UE2 tienen 953 y 813 respectivamente. Esto implica que el algoritmo de rechazo tiene un desempeño más eficiente en un caso en el cual exista presencia de deriva que en aquellos donde no.

- *Nodos rechazados*: En cuanto al desempeño del algoritmo, el porcentaje de rechazos es mayor en la UE5 y la UE6, mientras que en las UE1, UE2, UE3 y UE4 son menor y prácticamente igual entre ellas.
- *Nodos no simulados*: Los mayores porcentajes de nodos no simulados se obtienen en los escenarios de error de referencia 0.01 y de tolerancia baja. En particular se tienen que los menores porcentajes de nodos no simulados se obtienen en las UE1 y UE4. Esto muestra que el desempeño del algoritmo en cuanto a completitud de la grilla a estimar es mejor en el escenario con presencia de deriva como lo es la UE4.

### *Sobre el cumplimiento de la hipótesis*

La hipótesis del estudio defiende que el algoritmo de simulación Gaussiana secuencial con rechazo como nueva herramienta de simulación bajo un escenario no estacionario mediante una deriva de referencia permite respetar tal deriva obteniendo resultados representativos de los datos originales en cuanto a sus estadísticos de orden 1 y 2. En este sentido se cumple la hipótesis de estudio en base a los resultados obtenidos en ambos casos de estudio. En el caso de estudio I se aprecia claramente el buen ajuste a la deriva de los resultados, sin embargo no se chequea de manera completa si las estadísticas de los resultados son representativas. Esto si se realiza en el caso de estudio II donde se estudian a cabalidad las estadísticas de los resultados obteniéndose estadísticas representativas de los datos originales para los histogramas, variogramas por dirección y media local por dirección. Estos tres estadísticos muestran un comportamiento bastante similar en la mayoría de los escenarios dados por la sensibilización de los parámetros del algoritmo de rechazo, variando en general la dispersión en torno al mismo estadístico de los datos originales pero no su buen ajuste. Queda claro en este punto que el SK es comparable a un caso de mayor dispersión de las simulaciones en torno al estadístico a ajustar y que además estos distintos escenarios agregan flexibilidad al modelo resultante. Con respecto al BT se aprecia que este presenta estadísticas de las simulaciones sesgadas con respecto a las estadísticas de los datos originales en cuanto al histograma y variograma, por lo cual el algoritmo de rechazo mejora dichas estadísticas. Los estadísticos de validación contra datos originales permiten mostrar que el desempeño y utilidad del algoritmo está condicionado fuertemente tanto por la naturaleza como por el escenario de estimación, mas es importante destacar que en el caso de presencia de deriva

se obtiene en general mejores resultados que el SK. Es más, cuando esta deriva se ve reflejada de manera clara en el variograma como es el caso de la UE4, las estadísticas de validación de las simulaciones como estimación (coeficiente de determinación, pendiente de la regresión de datos estimados versus simulados y error medio) mejoran en relación al SK y BT en términos globales, indicando mejoras en el ajuste del modelo, insesgo condicional e insesgo global. Más aún, el desempeño de las simulaciones en cuanto a su capacidad de caracterizar la incertidumbre de los datos originales (accuracy plot) también mejora en relación al SK y BT. Esto queda claro en la tabla de desempeño global de algoritmo de rechazo, la cual considera las cuatro estadísticas de validación y la comparación con respecto a SK y BT, obteniendo un único índice que permite tener una relación de orden de desempeño en cada UE. De esta manera se aprecia que se producen mejoras en las UE2, UE3 y UE4 siendo la UE4 la de mejor desempeño. Se debe notar que si bien la UE1, UE5 y UE6 tienen un peor desempeño con el algoritmo de rechazo, en el caso de la UE5 y UE6 estas bajas en el desempeño tienen índices de -0.15 y -0.20 respectivamente, lo cual no es una baja considerable y, en efecto, para el caso de la UE5 su índice se encuentra bastante cercano a 0 lo cual indica un desempeño similar a los otros algoritmos. De hecho, si se toma el promedio de los 6 índices globales de las UE como índice para resumir en un solo número el desempeño del algoritmo del presente caso de estudio, se obtiene un valor de 0.15 lo cual indica mejoras con respecto a los otros dos algoritmos. Por último, es importante destacar que el algoritmo tiene una tendencia clara en su desempeño al variar los parámetros de las simulación en cuanto a tiempos de simulación nodos rechazados y nodos no simulados, notándose que se obtiene un comportamiento más eficiente al encontrarse en el caso para el que ha sido desarrollado el algoritmo, es decir, en presencia de deriva. Esto último queda demostrado ya que los tiempos de simulación bajan y el porcentaje de la grilla estimada aumenta en las UE4, la cual es el caso ideal para la implementación del algoritmo desarrollado.

#### *Sobre el impacto y aporte al rubro minero*

Si bien el algoritmo introduce una herramienta con mejoras tanto en el uso de la simulación como estimación y en la caracterización de la incertidumbre, es en este último punto donde se encuentra el mayor potencial dado que es justamente esto el principal uso de las simulaciones. En este sentido existe un potencial real en la planificación a largo plazo y valorización del negocio minero en términos de cuantificación de riesgo. Dado que esta herramienta permite caracterizar

de mejor manera la incertidumbre en un caso no estacionario, es que se puede pensar en realizar una planificación en cada una de las realizaciones procurando no tener simulaciones muy dispersas en torno a la deriva con el fin de no tener planificaciones con cambios importante en los criterios de diseño de la mina si no que más bien sensibilizaciones en términos de planes de producción. Con esto se puede calcular el valor actual neto (VAN) de cada una de las realizaciones y obtener una medida de cuantificación del riesgo calculando el valor en riesgo (VaR) respectivo, el cual mide la exposición al riesgo respecto del valor esperado del VAN a un determinado nivel de confianza. Esto es en efecto un potencial uso del algoritmo dado que la sensibilización que se realiza comúnmente en la valorización del negocio minero corresponde a simulaciones en torno a variables críticas (inversión, ley de cabeza de la variable, precio de la variable, recuperación, costo operación), para lo cual se utilizan distribuciones triangulares en general y se obtiene un VAN para cada escenario. En este sentido, esta metodología de valorizar cada una de las simulaciones, las cuales para un caso no estacionario debiesen caracterizar mejor la incertidumbre que simulaciones realizadas con SK o BT, permite incorporar de mejor manera la incertidumbre de los datos y por ende tener una mejor caracterización del riesgo del negocio en términos de variaciones en los planes de producción.

Por último es importante notar un potencial uso del algoritmo en zonas de bajas leyes donde por lo general se tiene un muestro menos exhaustivo y por ende se tienen menos datos condicionantes. En este sentido el algoritmo desarrollado permite entregar una ley de extrapolación que controla la presencia de datos de alta ley producto de la escasez de información y la aleatoriedad del método en zonas de baja ley.

## 9 BIBLIOGRAFÍA

- [1] Bouzari F., Clark A., 2002. *Anatomy, Evolution, and Metallogenic Significance of the Supergene Orebody of the Cerro Colorado Porphyry Copper Deposit, I Región, Northern Chile*. Economic Geology, Vol. 97 (2002), pp 1701-1740.
- [2] Chilès J-P, Delfiner P. 1999. *Geostatistics-modeling spatial uncertainty*. Wiley, New York.
- [3] Cuba, M., Leuangthong, O., Ortiz, J. 2012. *Detecting and quantifying sources of non-stationarity via experimental semivariogram modeling*. Stochastic Environmental Research and Risk Assessment (2012), pp 247-260.
- [4] Deutsch, C. and Journel, A.G. 1998. *GSLIB Geostatistical Software Library and User's Guide*. Applied Geostatistics Series. Oxford University Press.
- [5] Emery, X. 2011. *Geoestadística*. Departamento de Ingeniería en Minas, Universidad de Chile, Santiago.
- [6] Emery, X. 2012. *Simulación estocástica y geoestadística no lineal*. Departamento de Ingeniería en Minas, Universidad de Chile, Santiago.
- [7] Gerencia geología BHP Billiton Pampa Norte - Cerro Colorado, 2013. *GEO LP Geología CMCC*.
- [8] Journel, A.G. 1989. *Fundamentals of Geostatistics in Five Lessons*. Short Course of Geology: Volume 8. American Geophysical Union, Washington, D.C.
- [9] Journel, A.G. and Huijbregts, Ch. J. 1978. *Mining Geostatistics*. London: Academic.
- [10] Journel, A.G., Rossi, M.E., 1989. *When do we need a trend model in Kriging?* Mathematical Geology, Vol. 21, No. 7 (1989).
- [11] Leuangthong, O., Deutsch, C.V., 2003. *Stepwise conditional transformation for simulation of multiple variables*. Mathematical Geology, Vol. 35, No. 2 (2003).

- [12] Manchuk, J.G., Deutsch, C.V., 2012. *A flexible sequential Gaussian simulation program: USGSIM*. Computer & Geosciences 41 (2012), pp 208-216.
- [13] McLennan, J.A., Deutsch, C.V., 2008. *A new approach to SGS with a trend: A non-stationary Gaussian transformation*. Geostats (2008), pp 419-428.
- [14] Myers, D. E. 1989. *To be or not to be ... Stationary? That is the Question*. Mathematical Geology, Vol. 21, No. 3 (1989).
- [15] Olea, R.A., 2006. *A six-step practical approach to semivariogram modeling*. Stochastic Environmental Research and Risk Assessment (2006), pp 307-218.
- [16] Yao, T., Calvert, C., Jones, T., Bishop, G., Ma, Y., Foreman, Lincoln, 2006. *Spectral simulation and conditioning to local continuity trends in geologic modeling* Mathematical Geology, Vol. 38, No. 1 (2006).

## ANEXO A: Archivo de parámetros SGSIM

Se presentan a continuación el archivo de parámetros sgsim.par del ejecutable de GSLib sgsim.exe que permite realizar simulación Gaussiana secuencial e implementar el algoritmo propuesto de simulación con rechazo considerando una deriva de referencia. Aquellos parámetros destacados corresponden a implementaciones del algoritmo propuesto.

### Parameters for SGSIM

\*\*\*\*\*

#### START OF PARAMETERS:

nodata	- file with data
1 2 3 0 0 0	- columns for X,Y,Z,vr,wt,sec.var.
-1.0 1.0e21	- trimming limits
0	- transform the data (0=no, 1=yes)
sgsim.trn	- file for output trans table
0	- consider ref. dist (0=no, 1=yes)
../histsmth/histsmth.out	- file with ref. dist distribution
1 2	- columns for vr and wt
0.0 15.0	- zmin,zmax(tail extrapolation)
1 0.0	- lower tail option, parameter
1 15.0	- upper tail option, parameter
0	- debugging level: 0,1,2,3
sgsim.dbg	- file for debugging output
sgsim.out	- file for simulation output
1	- number of realizations to generate
50 1 5.0	- nx,xmn,xsiz
50 1 5.0	- ny,ymn,ysiz
1 1 5.0	- nz,zmn,zsiz
69069	- random number seed
0 8	- min and max original data for sim

12	- number of simulated nodes to use
1	- assign data to nodes (0=no, 1=yes)
1 3	- multiple grid search (0=no, 1=yes),num
0	- maximum data per octant (0=not used)
10 10 10	- maximum search radii (hmax,hmin,vert)
0.0 0.0 0.0	- angles for search ellipsoid
71 61 17	- size of covariance lookup table
1 0.60 1.0	- ktype: 0=SK,1=OK,2=LVM,3=EXDR,4=COLC, <b>5=RDR</b>
../data/cluster.dat	- file with LVM, EXDR, or COLC variable
1	- column for secondary variable
2 0.4	- nst, nugget effect
1 0.25 0.0 0.0 0.0	- it,cc,ang1,ang2,ang3
50.0 50.0 50.0	- a_hmax, a_hmin, a_vert
1 0.35 0.0 0.0 0.0	- it,cc,ang1,ang2,ang3
1000 1000 1000	- a_hmax, a_hmin, a_vert
<b>0.1 0.05</b>	<b>- initial and final tolerance for the imposed drift</b>
<b>1</b>	<b>- Type of tolerance decay (1:lin,2:exp)</b>
<b>0.2</b>	<b>- Fraction of search radii to search radii tol</b>
<b>0 0.01</b>	<b>- Probabilistic Rejection?(0=no, 1=yes), reference error to tolerance criterion</b>
<b>1000</b>	<b>- Acceptable number of rejected sim values</b>

Las implementaciones en el archivo sgsim.par corresponden a 7 nuevos parámetros y una extensión en opciones en uno de los parámetros ya existentes. A saber, estos 7 parámetros corresponden a:

1. **Initial tolerance:** Tolerancia inicial para el cálculo de la tolerancia dinámica.
2. **Final tolerance:** Tolerancia final para el cálculo de la tolerancia dinámica.
3. **Type of tolerance decay:** Tipo de decaimiento en tolerancia dinámica. Puede ser lineal como puede ser exponencial. Permite calcular la tolerancia dinámica a partir de las tolerancias inicial y final.



4. **Fraction of search radii to search radii tol:** Factor de proporcionalidad entre la dimensión de la vecindad de búsqueda para condicionamiento por Kriging y la vecindad de búsqueda para cálculo de media condicional.
5. **Probabilistic rejection:** Definición del mecanismo de rechazo, este puede ser tanto determinista como probabilístico.
6. **Reference error to tolerance criterion:** Error de referencia para la estandarización del rango de tolerancia dinámica en el caso de rechazo probabilístico.
7. **Acceptable number of rejected sim value:** Número máximo aceptable de rechazos para un nodo a simular antes de dejarlo para una posterior resimulación o como un dato no simulado dependiendo de la etapa del algoritmo.
8. **Ktype: 5=RDR:** Extensión de tipo de Kriging a utilizar. La opción 5 corresponde a la implementación del algoritmo desarrollado siendo denominado como Kriging con deriva de referencia (RDR: Reference Drift Kriging).

## ANEXO B: Descripción de variables creadas en implementación del algoritmo

En esta sección se detalla y se describe cada una de las variables creadas en la implementación del algoritmo desarrollado.

1. **anorsv**: Entero. Número máximo de veces que se puede rechazar un nodo antes de considerarse para una posterior resimulación. Se utiliza el mismo valor tanto para etapa de simulación como para etapa de resimulación.
2. **asim**: Real. Temporal que guarda el valor simulado destransformado para poder usarse en el mecanismo de rechazo.
3. **auxdrift**: Boolean. True: Si tipo de Kriging 3 o 5. Copia LVM (arreglo con la deriva de referencia) en SEC (arreglo con atributo secundario).
4. **cmeansim**: Real. Media condicional de la vecindad de búsqueda respectiva, considerando el valor simulado.
5. **cr1**: Real. Variable 1 en el criterio de rechazo. Para un rechazo determinista es equivalente a la diferencia absoluta entre la media condicional simulada y la deriva de referencia (diff). Para un rechazo probabilístico es equivalente al número aleatorio p utilizado como probabilidad para el criterio de rechazo.
6. **cr2**: Real. Variable 2 en el criterio de rechazo. Para un rechazo determinista es equivalente a la tolerancia dinámica actualizada. Para un rechazo probabilístico es equivalente a la probabilidad de aceptación de la diferencia absoluta diff calculada por medio de la función de probabilidad de aceptación en función de la tolerancia dinámica actualizada.
7. **diff**: Real. Diferencia absoluta entre la media condicional simulada y la deriva de referencia.
8. **ddtol**: Arreglo real de dimensión equivalente al número de nodos en la grilla. Mantiene un registro de la tolerancia dinámica en la simulación.
9. **dtoli**: Real. Tolerancia en nodo inicial.
10. **dtolf**: Real. Tolerancia en nodo final.
11. **dtoltype**: Entero que indica el tipo de decaimiento de la tolerancia dinámica. 1: Decaimiento lineal. 2: Decaimiento exponencial.

12. **finaltime**: Tiempo de cpu que se calcula al final de la ejecución del código para determinar el tiempo de simulación mediante  $\text{finaltime} - \text{starttime}$ .
13. **ftol**: Real. Factor de proporcionalidad entre dimensión de vecindad de búsqueda de Kriging y vecindad de búsqueda de media condicional. Se utiliza para pasar desde el radio cuadrático de la vecindad de Kriging ( $\text{radsqd}$ ) al radio cuadrático de la vecindad de búsqueda para cálculo de media condicional ( $\text{radsqdtol}$ ).  $\text{radsqdtol} = \text{radsqd} \cdot \text{ftol}^2$
14. **lvmg**: Real. Valor Gaussiano de la variable del arreglo LVM (deriva) utilizado para actualizar ésta a su valor Gaussiano.
15. **lvmr**: Real. Temporal para transformar la variable del arreglo LVM (deriva) a su valor Gaussiano.
16. **mtol**: Real. Variable que indica el decremento de la tolerancia dinámica y que depende del mecanismo de decaimiento que se esté utilizando (Ver sección 4.2).
17. **nlooku**: Entero. Define el tamaño de la matriz de covarianzas en función de la vecindad de búsqueda para el cálculo de la media condicional.
18. **nrd**: Entero. Número de nodos rechazados en la simulación respectiva. Se actualiza para cada simulación. Se utiliza para determinar el número de nodos a resimular en una posterior etapa de resimulación.
19. **nsbtosrtol**: Entero. Número de súper bloques a buscar para la búsqueda por súper bloques en caso de búsqueda de datos para el cálculo de la media condicional. Se utiliza en la subrutinas externas 'pícSKup' y 'srchsupr'.
20. **preject**: Binario que indica el tipo de mecanismo de rechazo. 0: Rechazo determinista. 1: Rechazo probabilístico.
21. **radsqdtol**: Real. Radio cuadrático de la vecindad de búsqueda para cálculo de media condicional.
22. **rcont**: Entero. Contador del número de rechazos en el nodo respectivo. En caso de ser superior al número máximo aceptable de rechazos ( $\text{anorsv}$ ), se inicializa en 0 y se guarda el nodo para una posterior etapa de resimulación.
23. **renxyz**: Entero. Indica el límite superior del loop principal sobre los nodos. Este será equivalente al total de datos ( $\text{nxyz}$ ) en caso de estar en etapa de simulación, o al número de nodos rechazados ( $\text{nrd}$ ) en caso de estar en etapa de resimulación.

24. **resim:** Arreglo que guarda los índices de los nodos que han tenido un número de rechazos superior al máximo aceptable (anorsv) y que por ende pasan a una posterior etapa de resimulación.
25. **resimulating:** Binario que indica si se está en etapa de simulación o de resimulación. 0: Simulación. 1: Resimulación.
26. **srctype:** Binario con el tipo de búsqueda en búsqueda espiral de nodos. 0: datos para Kriging. 1: datos para media condicional.
27. **srindy:** Arreglo entero binario con el tipo de búsqueda. 0: para búsqueda de datos condicionantes en la estimación por Kriging. 1: para búsqueda de datos para cálculo de media condicional.
28. **starttime:** Tiempo de cpu que se calcula al inicio de la ejecución del código para determinar el tiempo de simulación mediante  $\text{finaltime} - \text{starttime}$ .
29. **sumnclose:** Real. Suma del valor de los n datos para el cálculo de la media condicional buscados mediante la rutina 'srchsup'  $\sum_{i=1}^n y_i$ , es decir, sólo se usa en caso que la opción 'Assign data to nodes' no esté activada.
30. **sumncnode:** Real. Suma del valor de los n datos para el cálculo de la media condicional buscados mediante la rutina 'srchnd'  $\sum_{i=1}^n y_i$ . Corresponde a una búsqueda en espiral.
31. **supsrctype:** Binario con el tipo de búsqueda en búsqueda de súper bloques. 0: datos para Kriging. 1: datos para media condicional.
32. **updtol:** Real. Valor de la tolerancia actualizada para cada nodo. Este es el valor que se utiliza en el mecanismo de rechazo.

## ANEXO C: Caso de estudio I – Sensibilización con respecto a modelo variográfico

Se presenta en este anexo como varían los resultados de las simulaciones así como el variograma de las simulaciones para cada uno de los 17 set de parámetros para el Kriging con rechazo analizados en el caso de estudio 1 así como para el caso de simulación con Kriging simple.

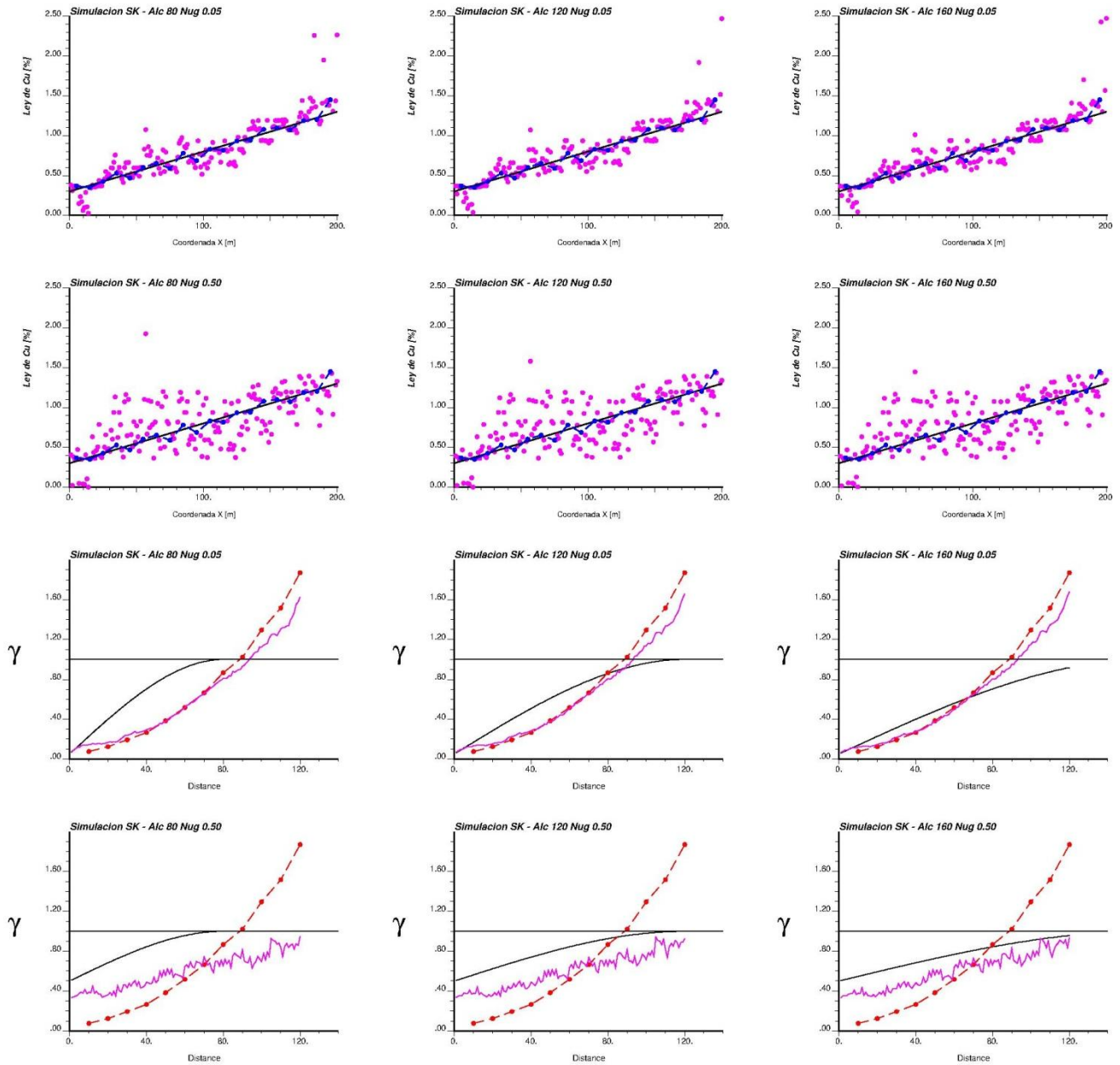
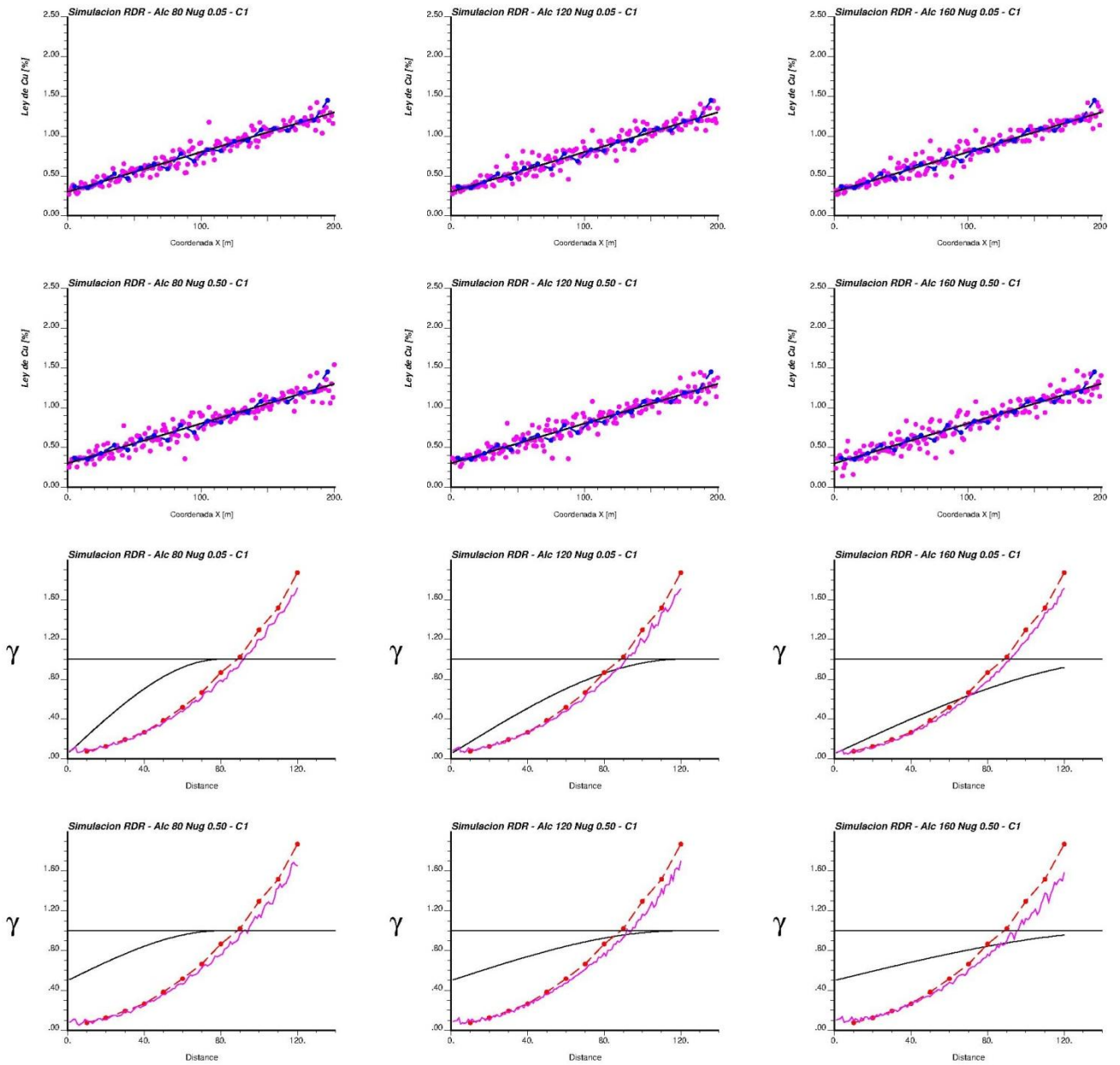
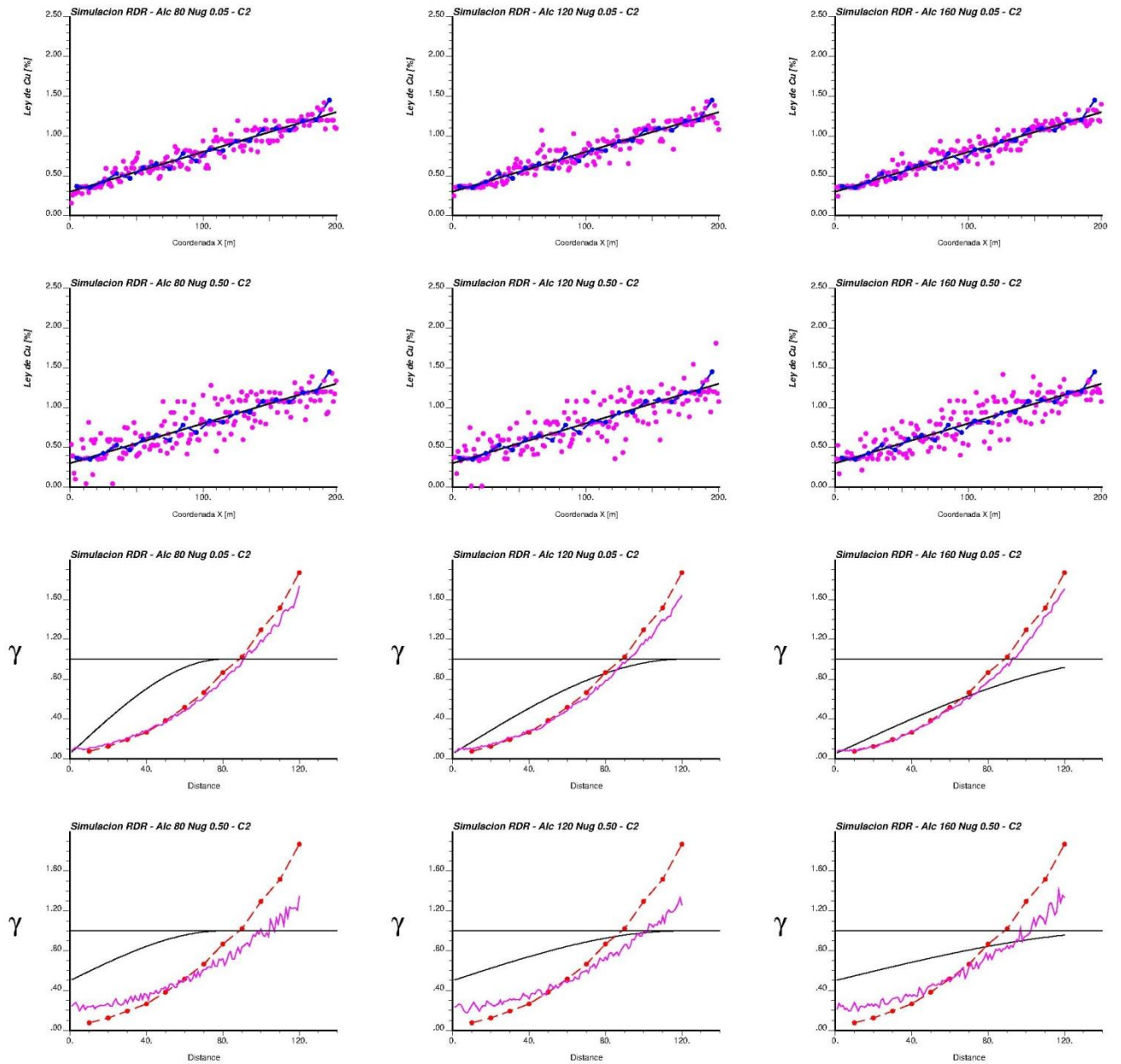


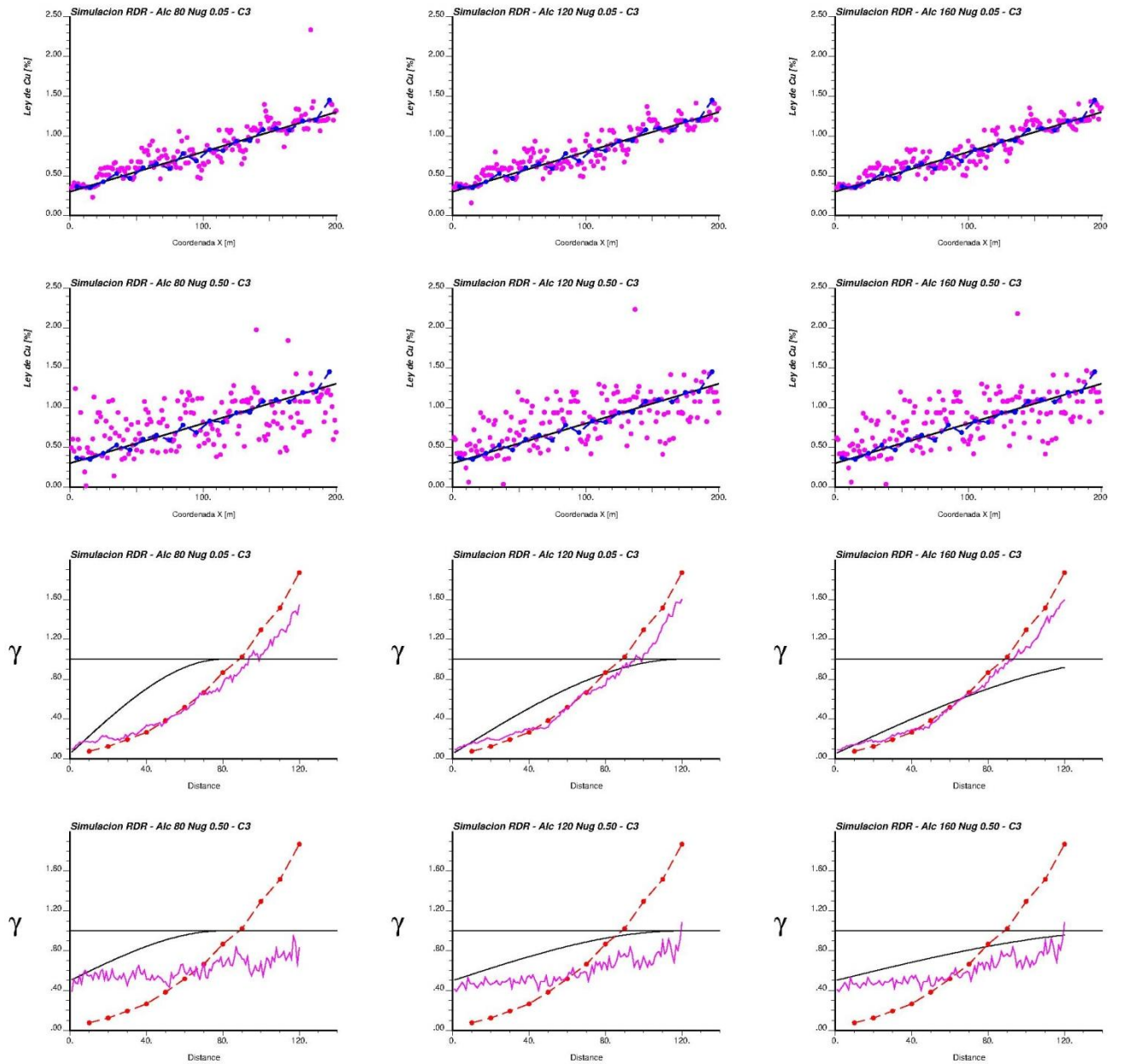
Gráfico 103: Variación modelo variográfico – Kriging simple.



**Gráfico 104: Variación modelo variográfico – Probabilístico con error de referencia: 0.01.**

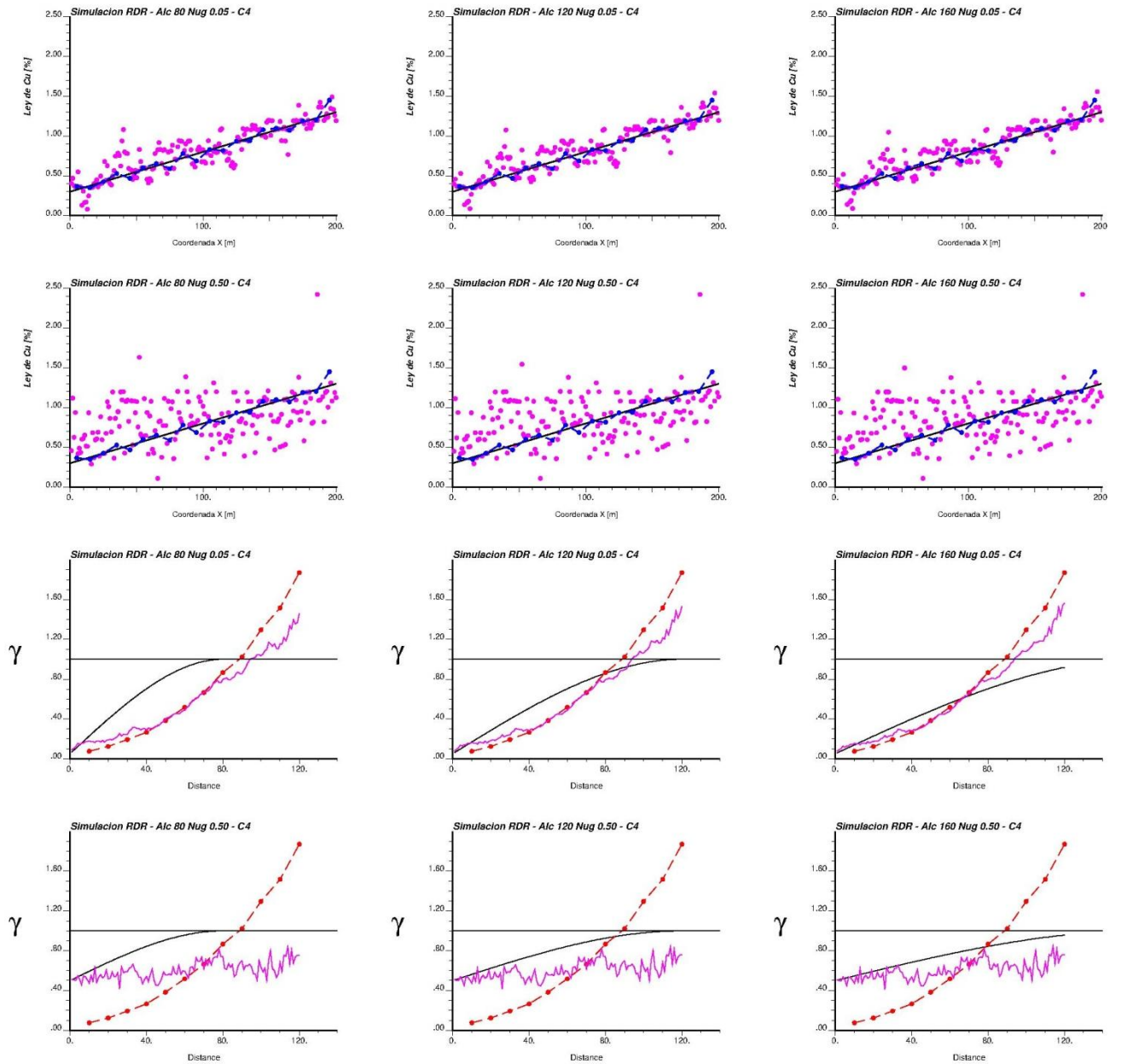


**Gráfico 105: Variación modelo variográfico – Probabilístico con error de referencia: 0.1.**

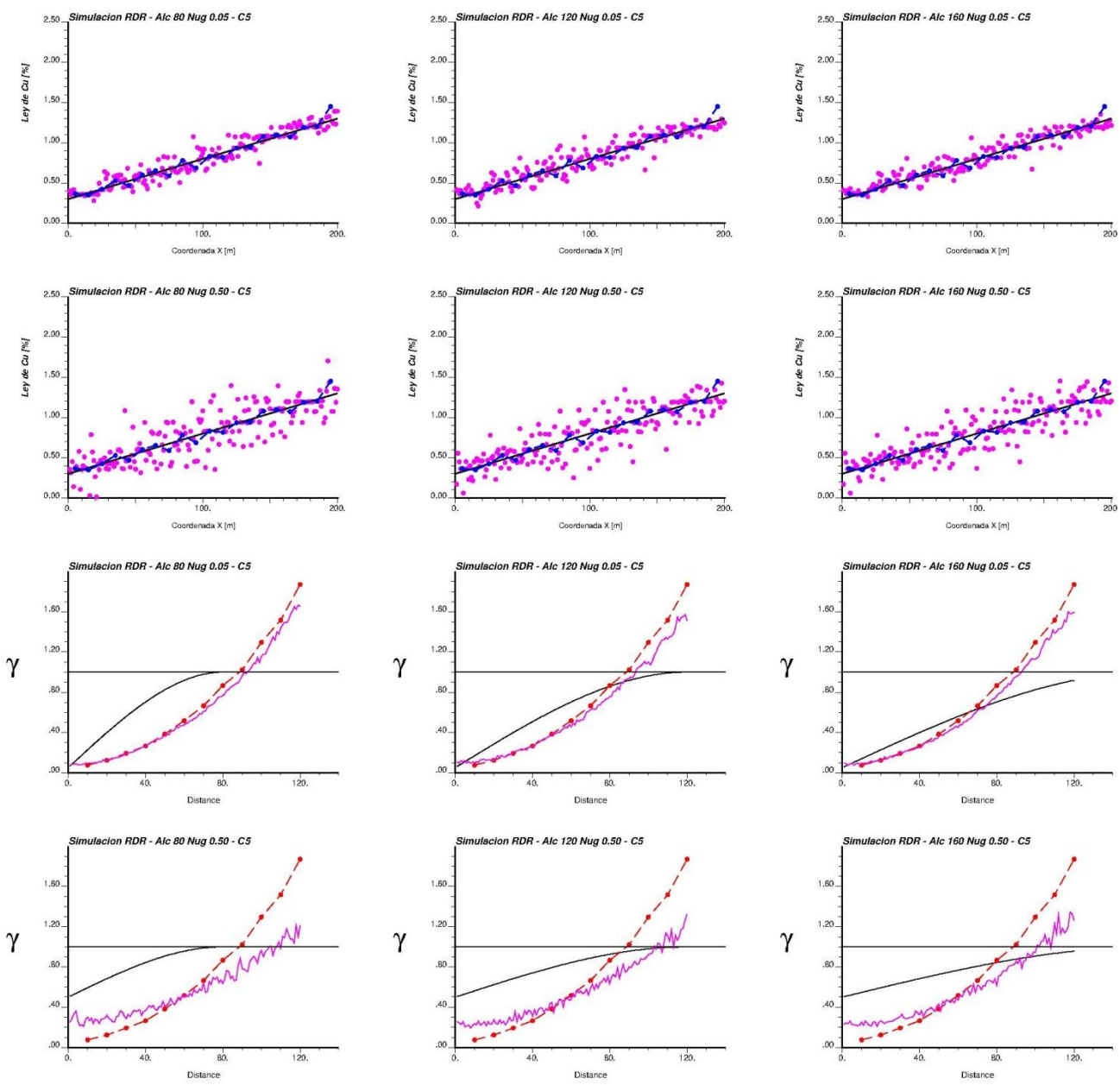


**Gráfico 106: Variación modelo variográfico – Probabilístico con error de referencia: 1.**

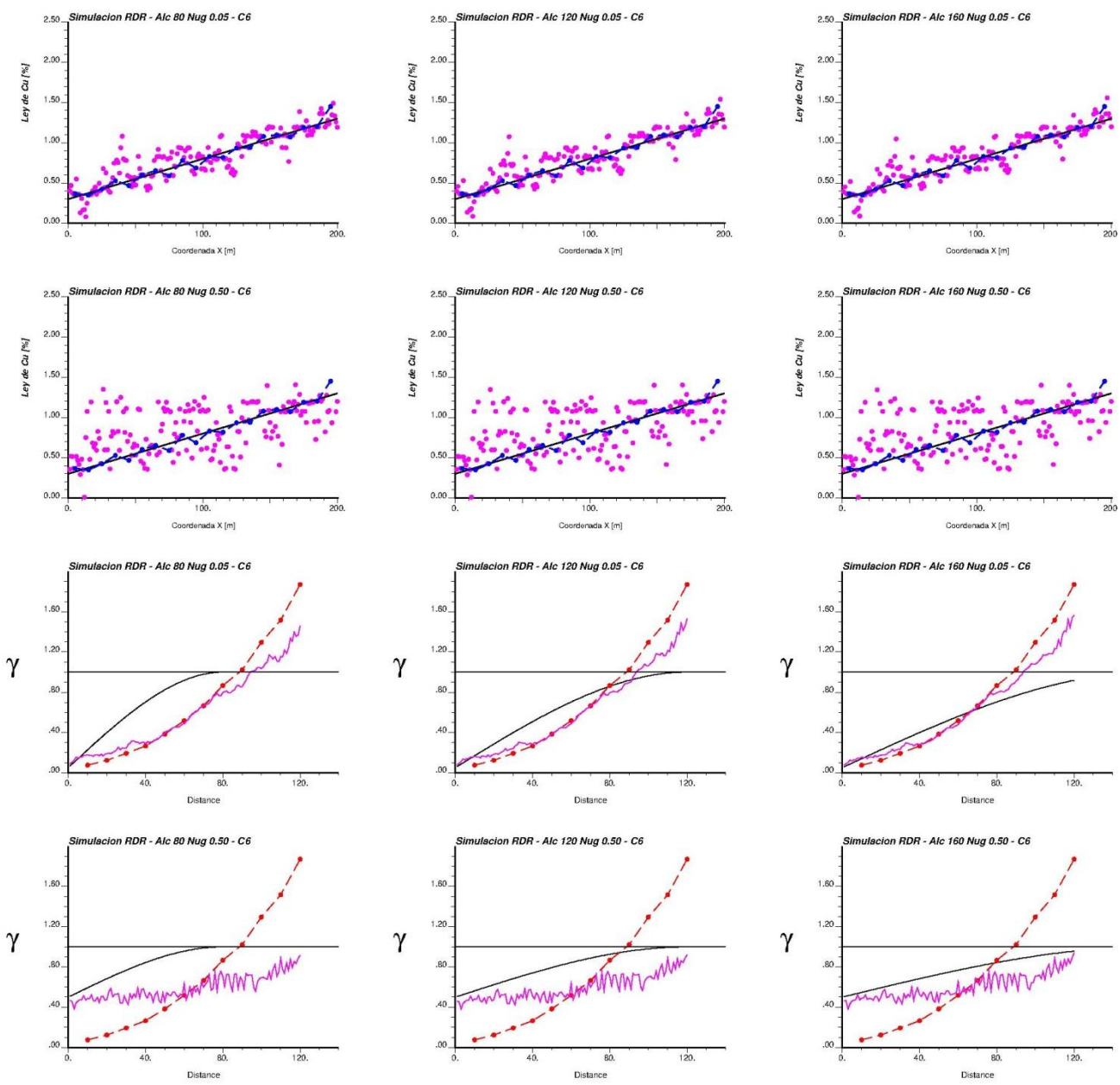




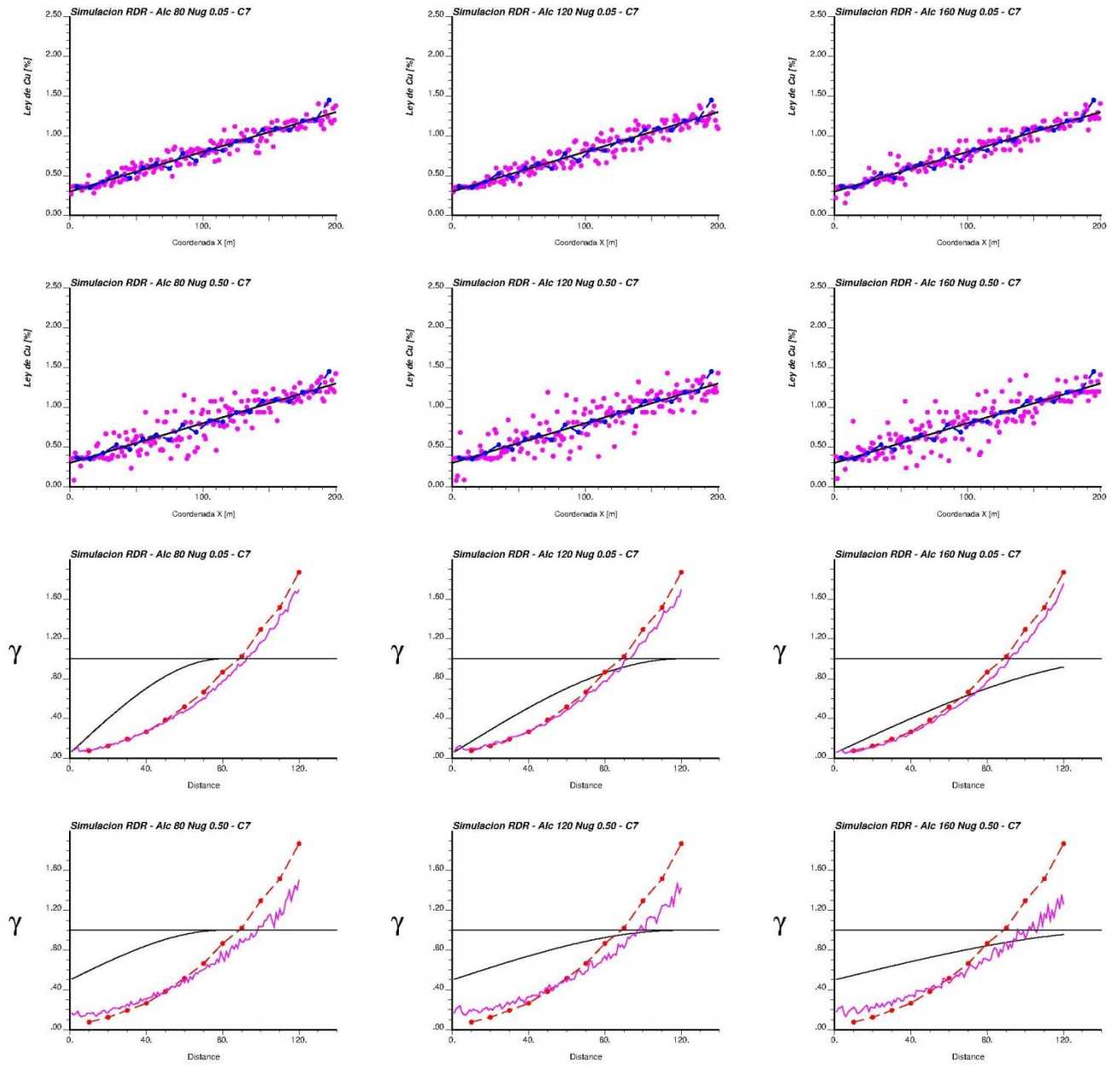
**Gráfico 107: Variación modelo variográfico – Probabilístico con error de referencia: 10.**



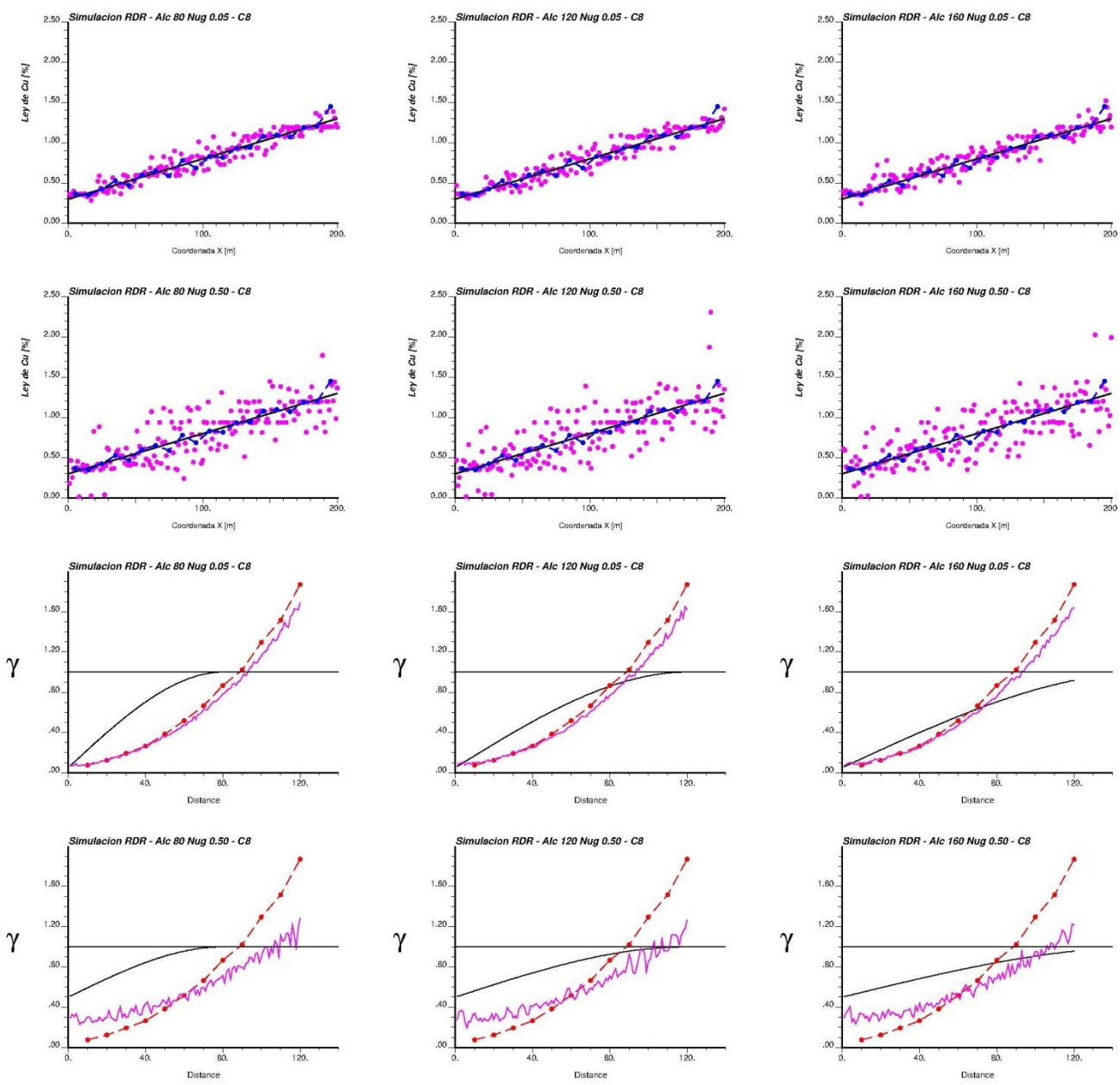
**Gráfico 108: Variación modelo variográfico – Determinista.**



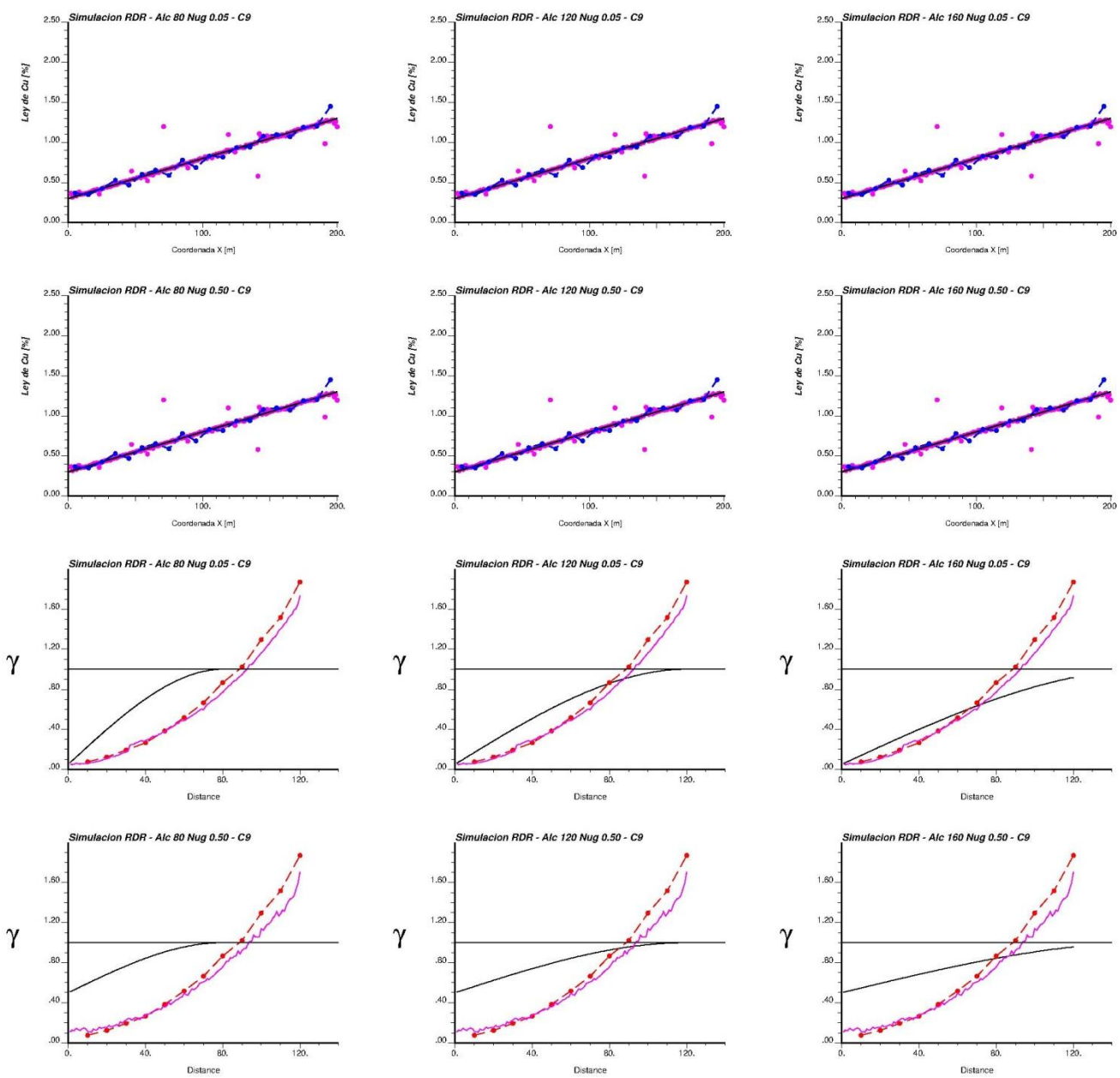
**Gráfico 109: Variación modelo variográfico – Tolerancia alta.**



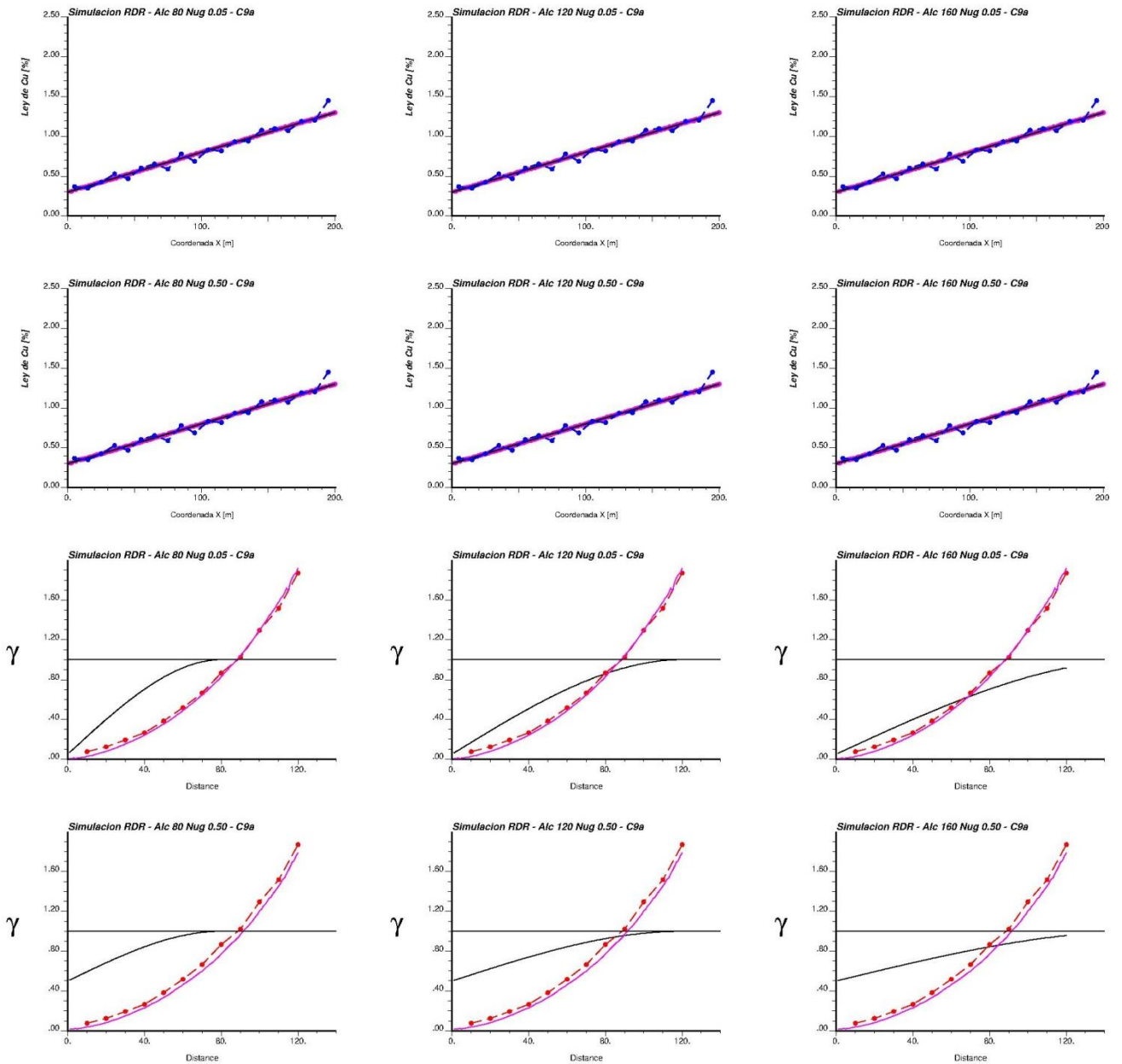
**Gráfico 110: Variación modelo variográfico – Tolerancia baja.**



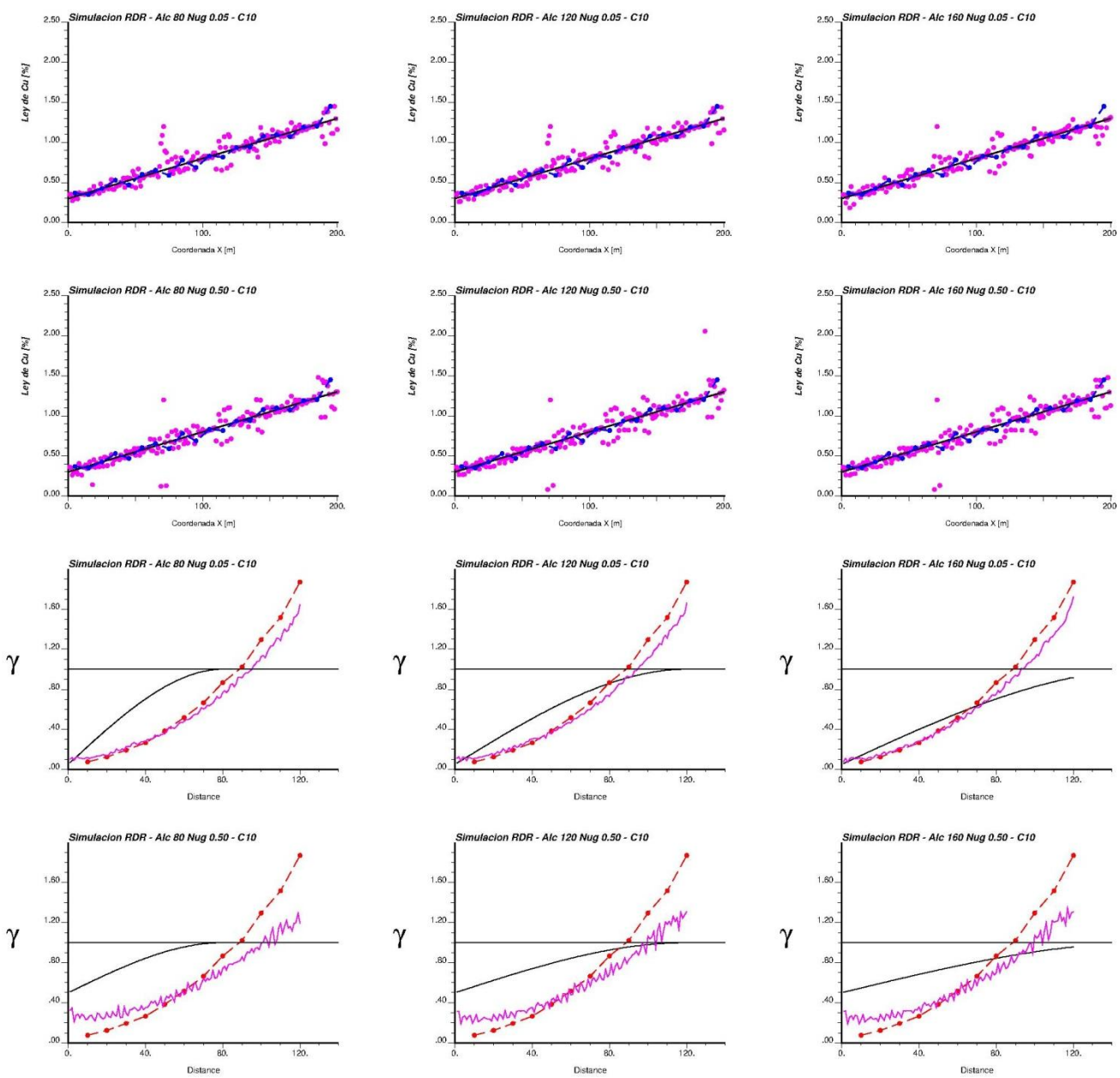
**Gráfico 111: Variación modelo variográfico – Tolerancia de amplio rango.**



**Gráfico 112: Variación modelo variográfico – Fracción de radio: 0.01.**

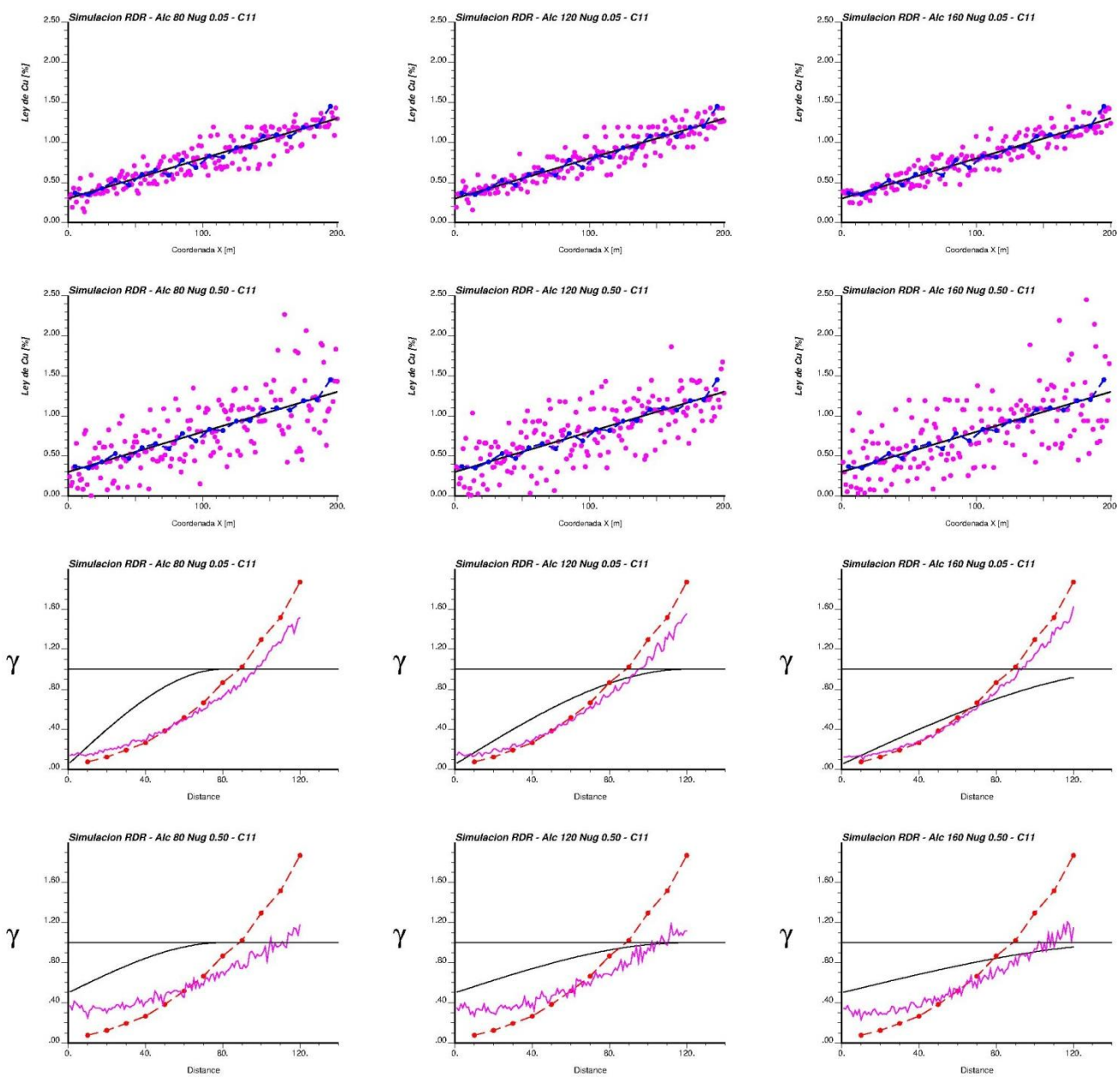


**Gráfico 113: Variación modelo variográfico – Fracción de radio: 0.01 – Tolerancia baja.**

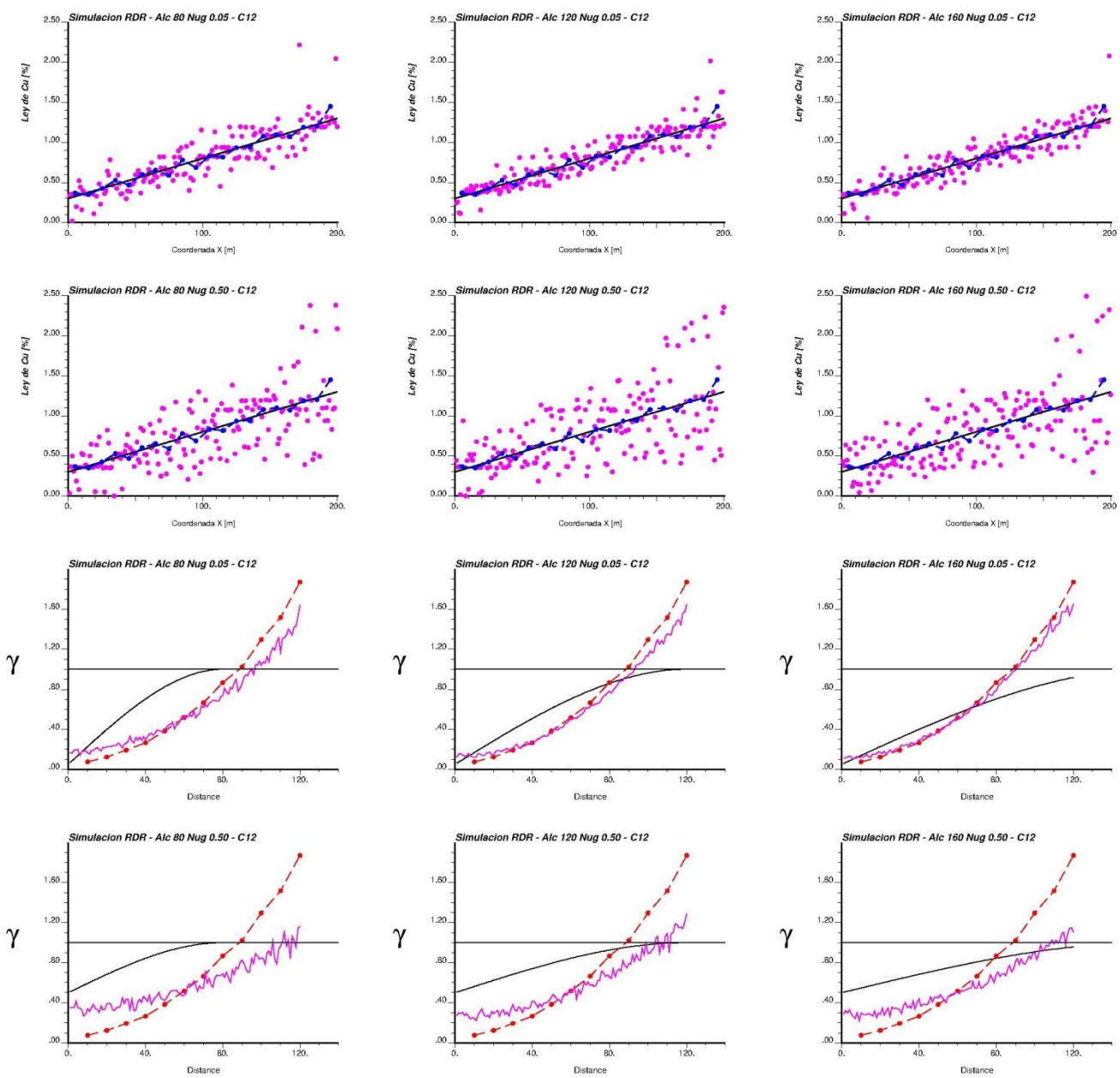


**Gráfico 114: Variación modelo variográfico – Fracción de radio: 0.1.**

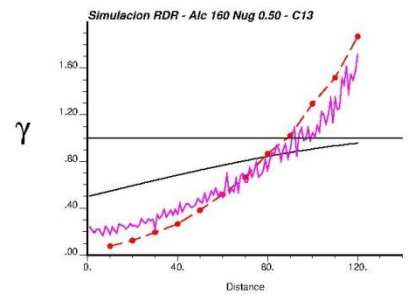
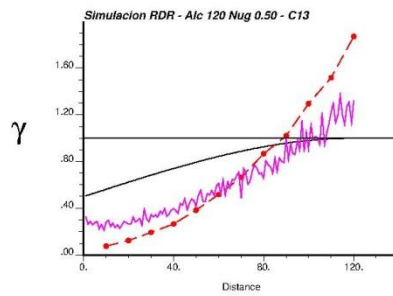
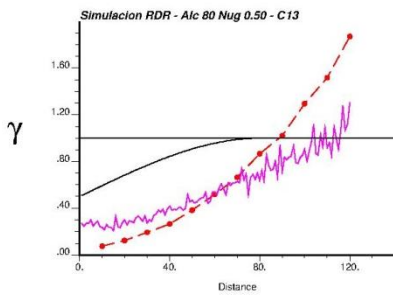
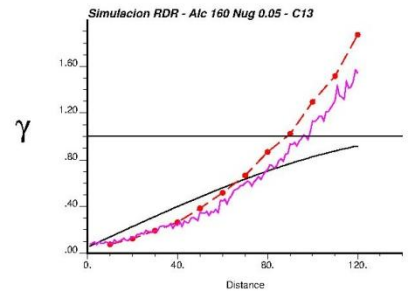
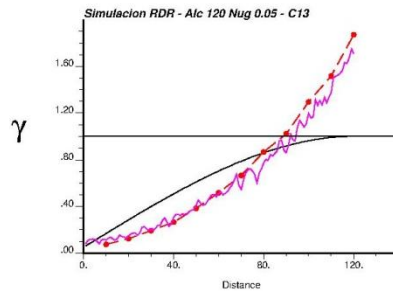
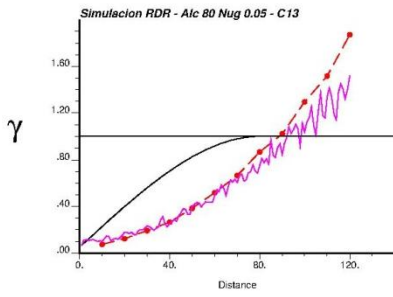
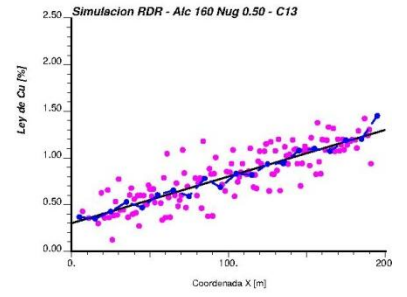
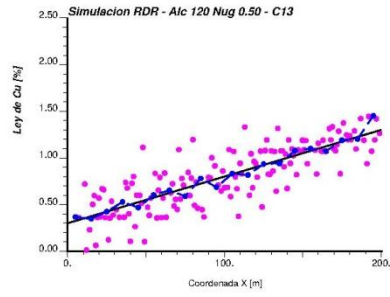
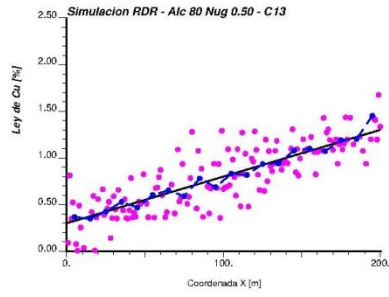
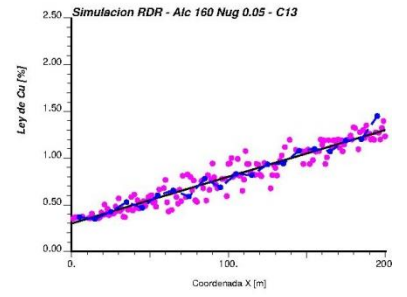
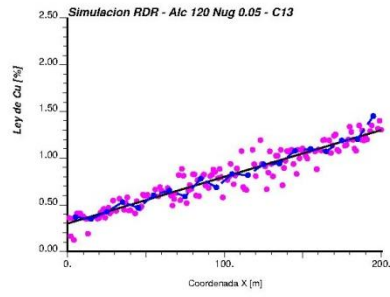
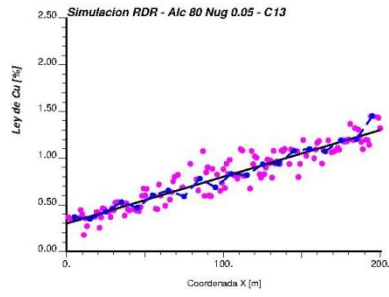




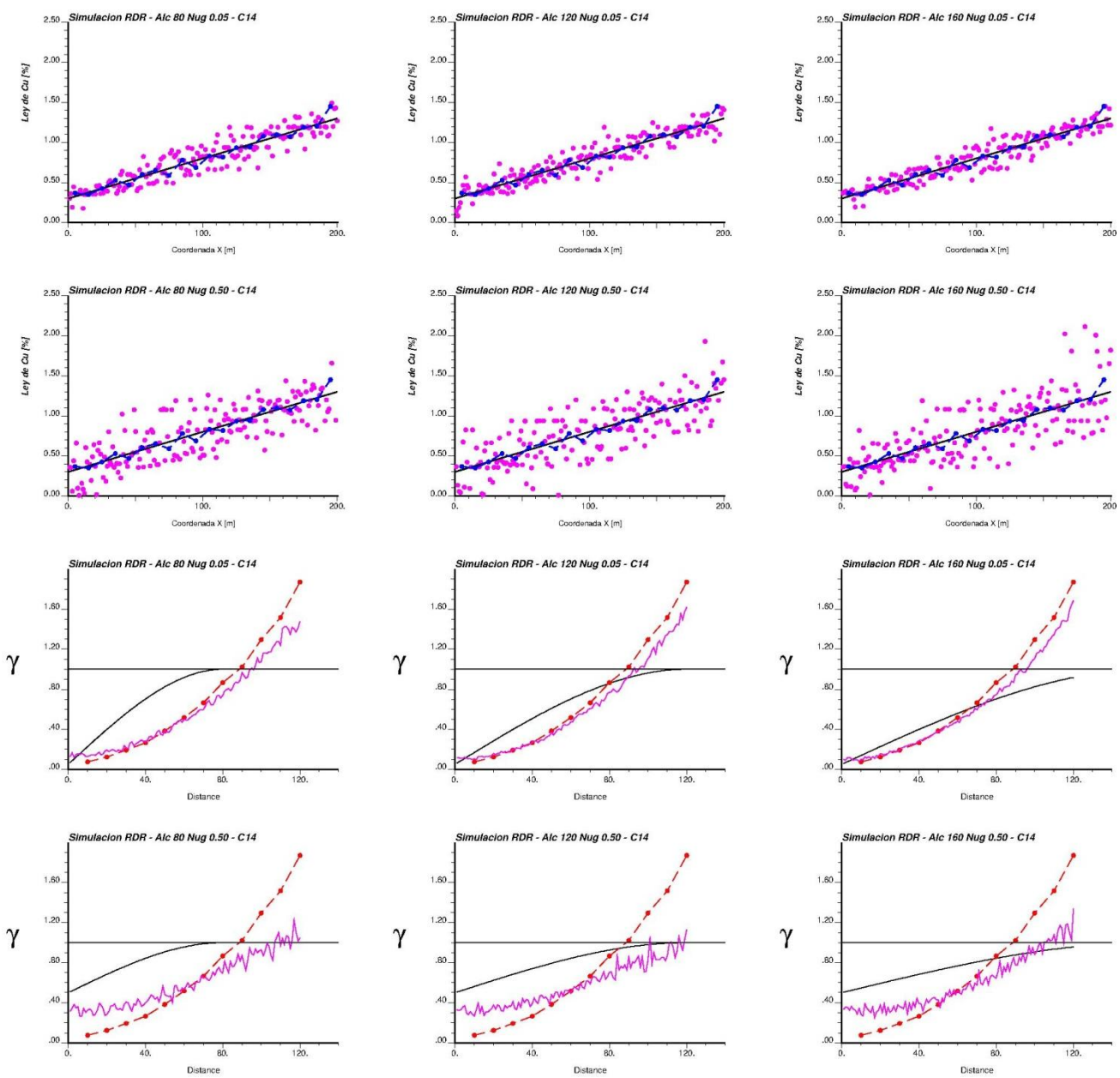
**Gráfico 115: Variación modelo variográfico – Fracción de radio: 1.**



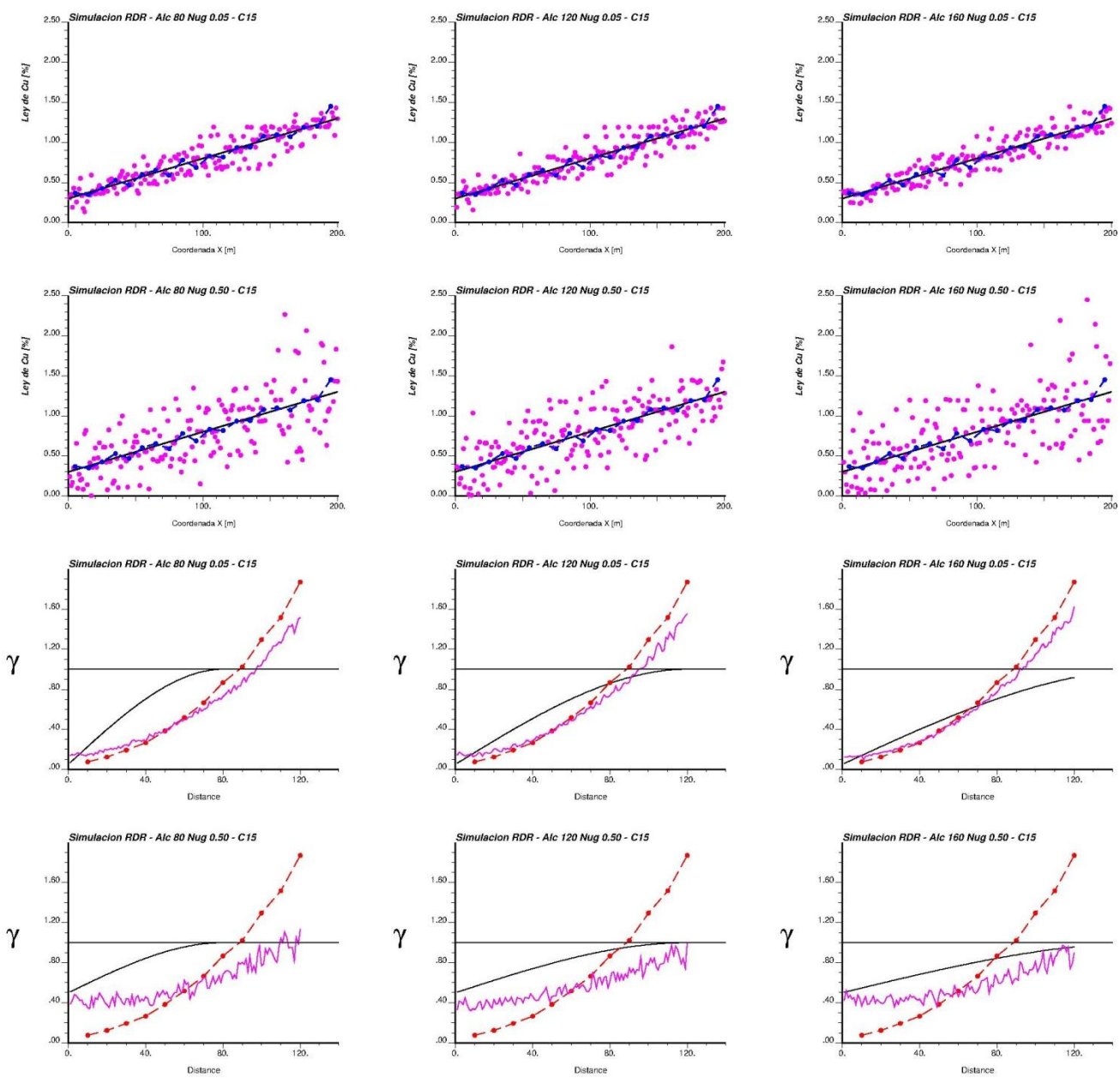
**Gráfico 116: Variación modelo variográfico – Fracción de radio: 10.**



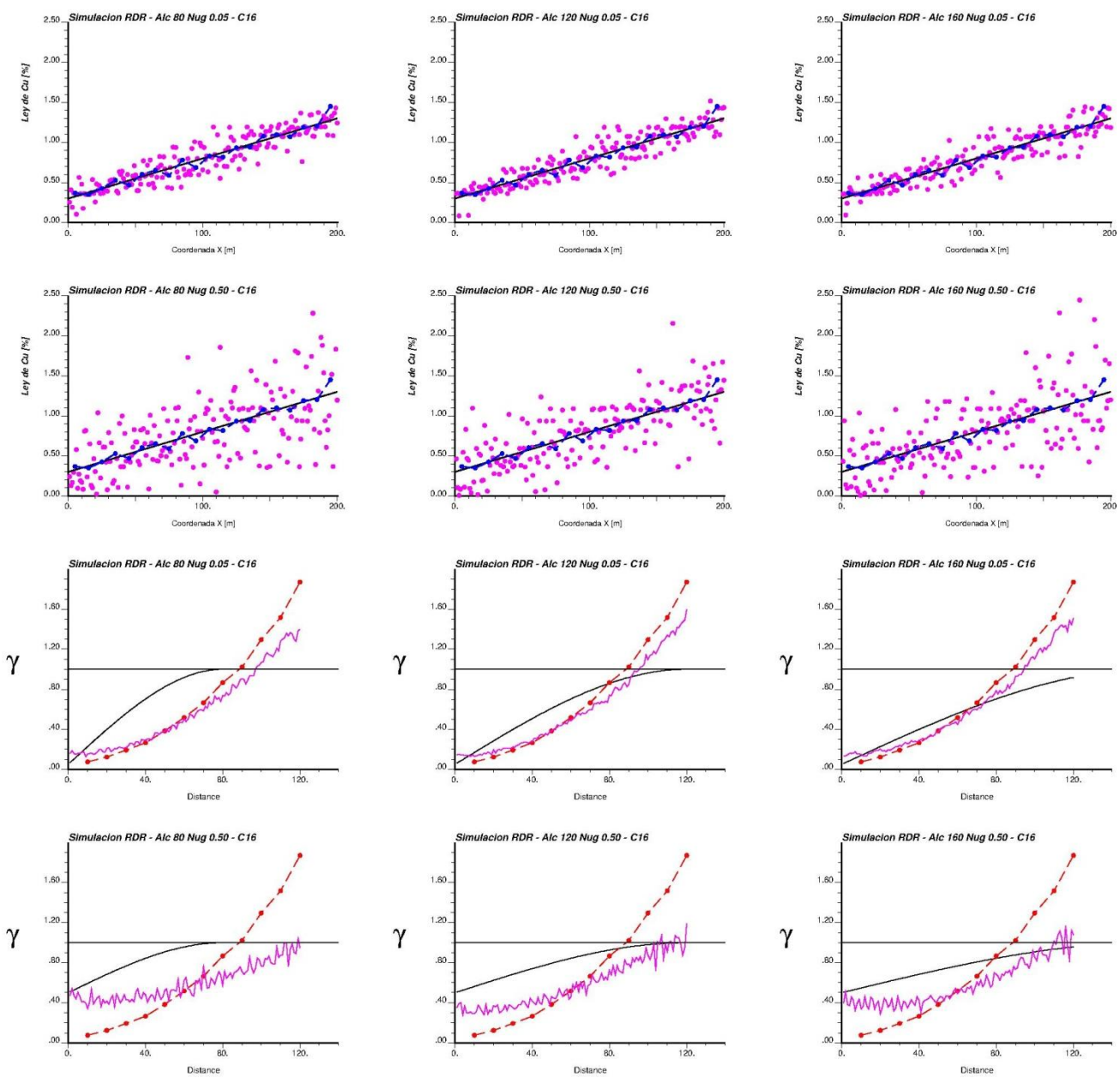
**Gráfico 117: Variación modelo variográfico – Rechazos: 10.**



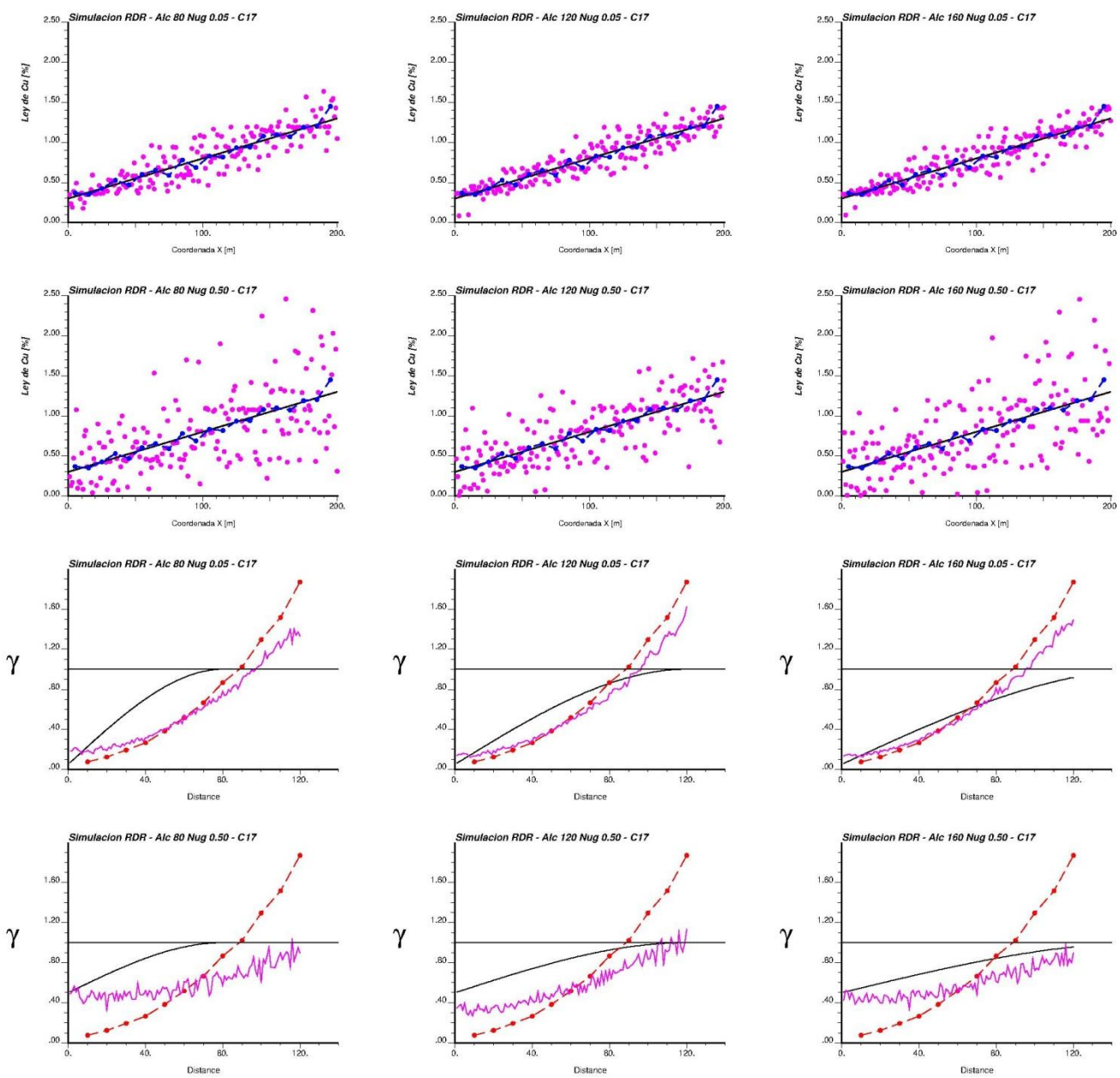
**Gráfico 118: Variación modelo variográfico – Rechazos: 100.**



**Gráfico 119: Variación modelo variográfico – Rechazos: 1,000.**



**Gráfico 120: Variación modelo variográfico – Rechazos: 10,000.**



**Gráfico 121: Variación modelo variográfico – Rechazos: 100,000.**

## ANEXO D: Caso de estudio II – Base maps y análisis de contacto

### D.1 Base maps

Se presenta en la Figura 15, Figura 16 y Figura 17, la proyección de los sondajes en los planos XY (planta), XZ (perfil EO) e YZ (perfil NS) respectivamente.

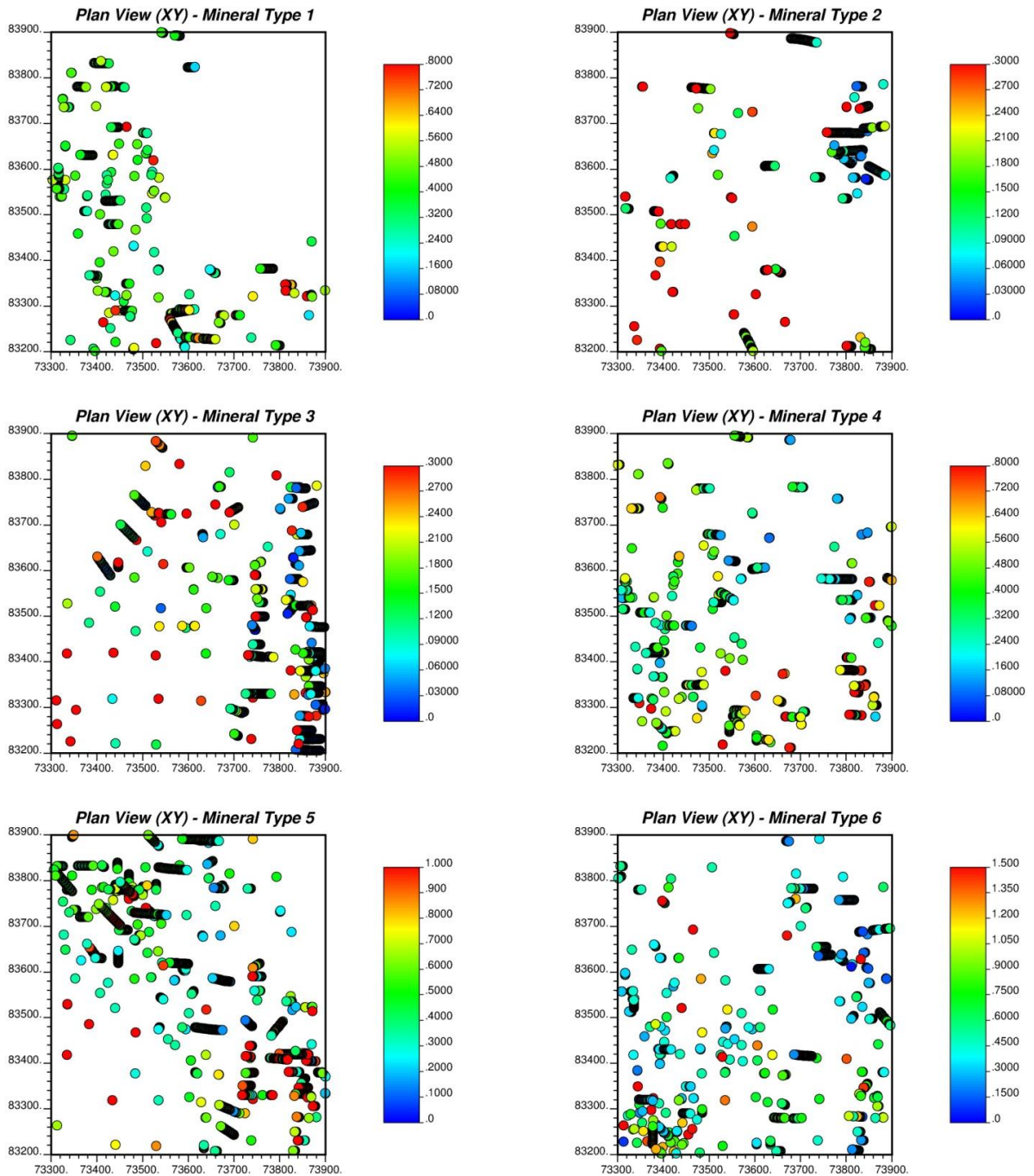
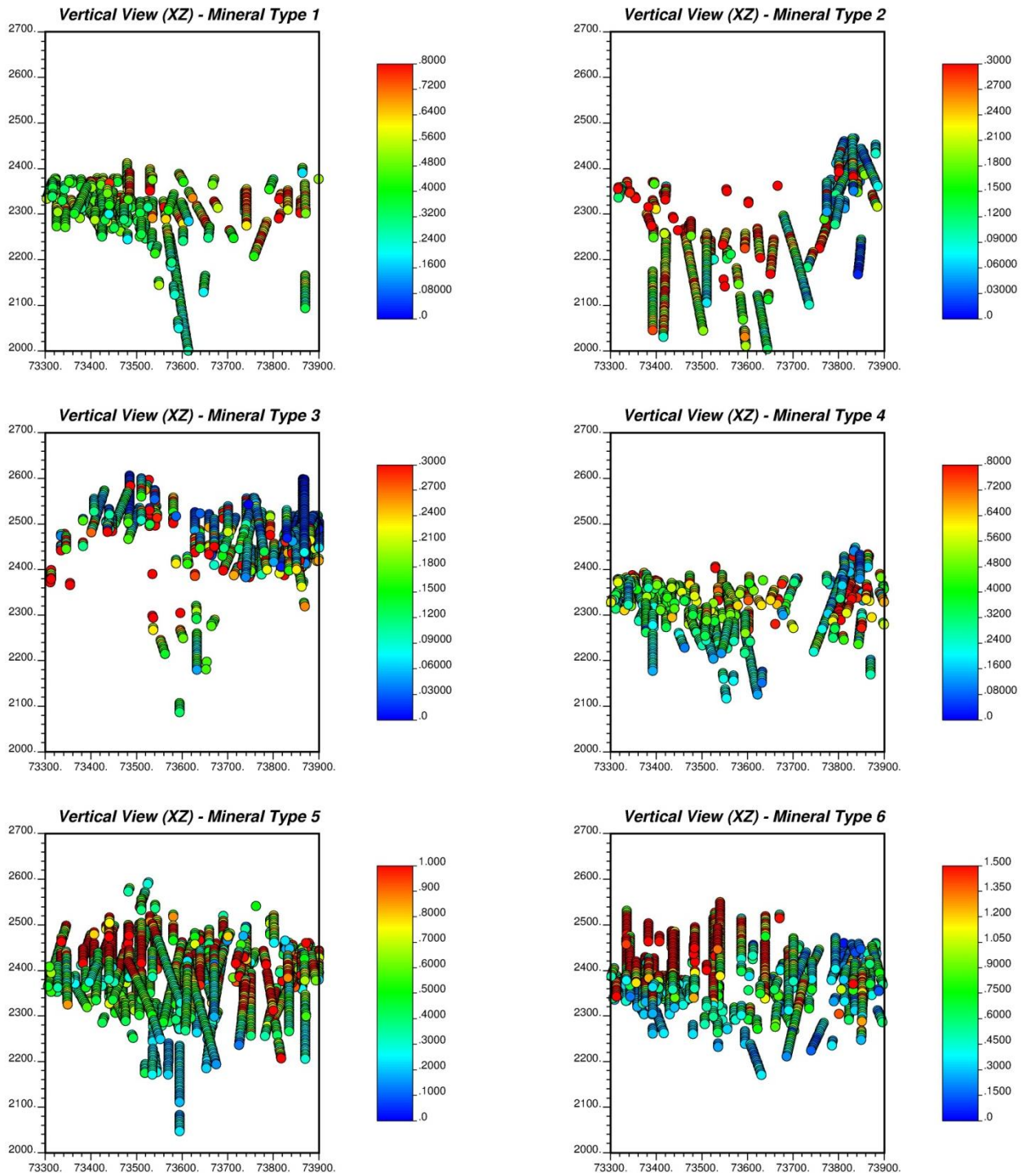
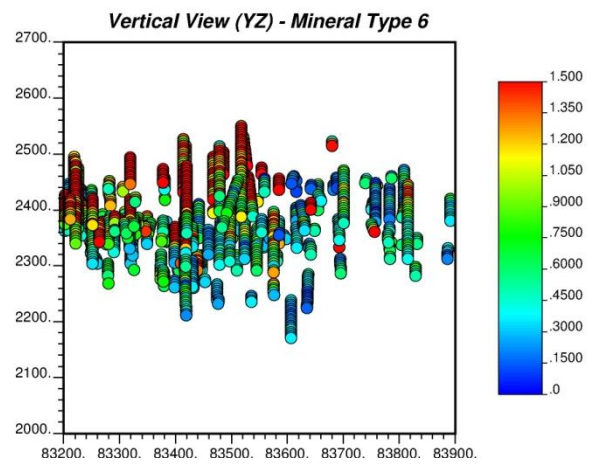
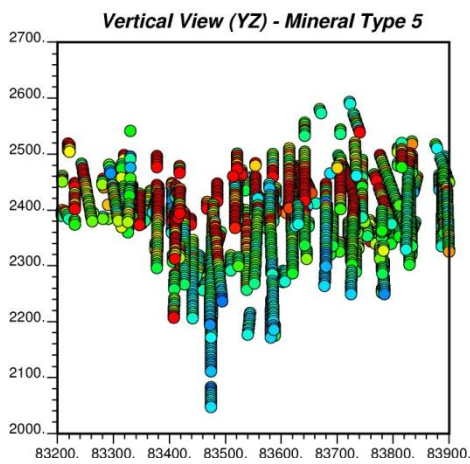
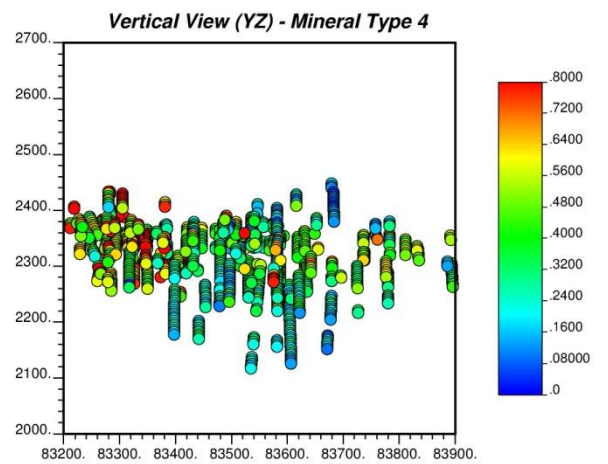
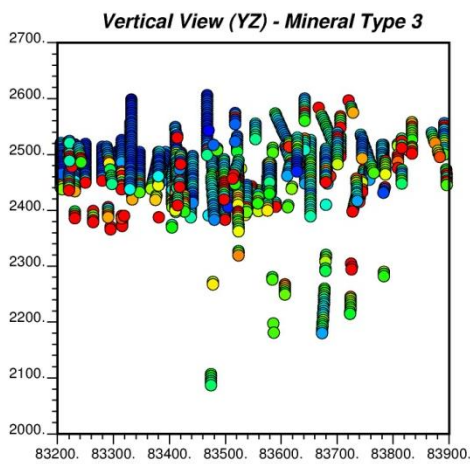
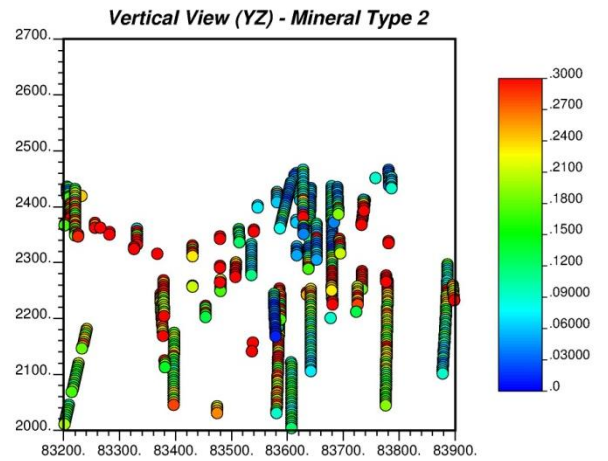
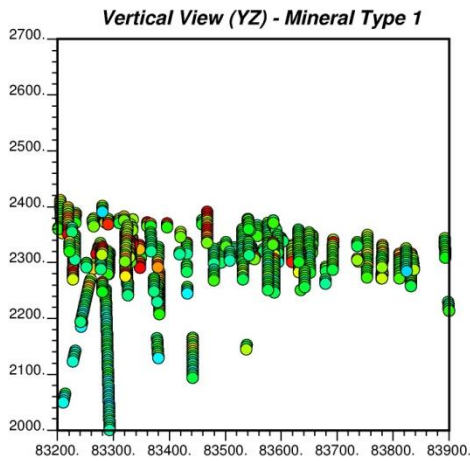


Figura 15: Proyección XY (Planta).





**Figura 16: Proyección XZ (Perfil EW)**



**Figura 17: Proyección YZ (Perfil NS).**

## D.2 Análisis de contacto

Se presenta el análisis de contacto de cada unidad de estimación con respecto al resto de las unidades. Para esto, se calcula el correlograma omnidireccional a distancias entre 0 y 100 metros. En caso de existir un contacto blando, el decaimiento de la curva descrita por el correlograma correspondiente debiese ser gradual. Esto no es posible de apreciar claramente en ninguno de los gráficos de análisis de contacto por lo que no se tiene evidencia de contactos blandos. Por este motivo no se consideran datos de unidades de estimación contiguas para el proceso de simulación.

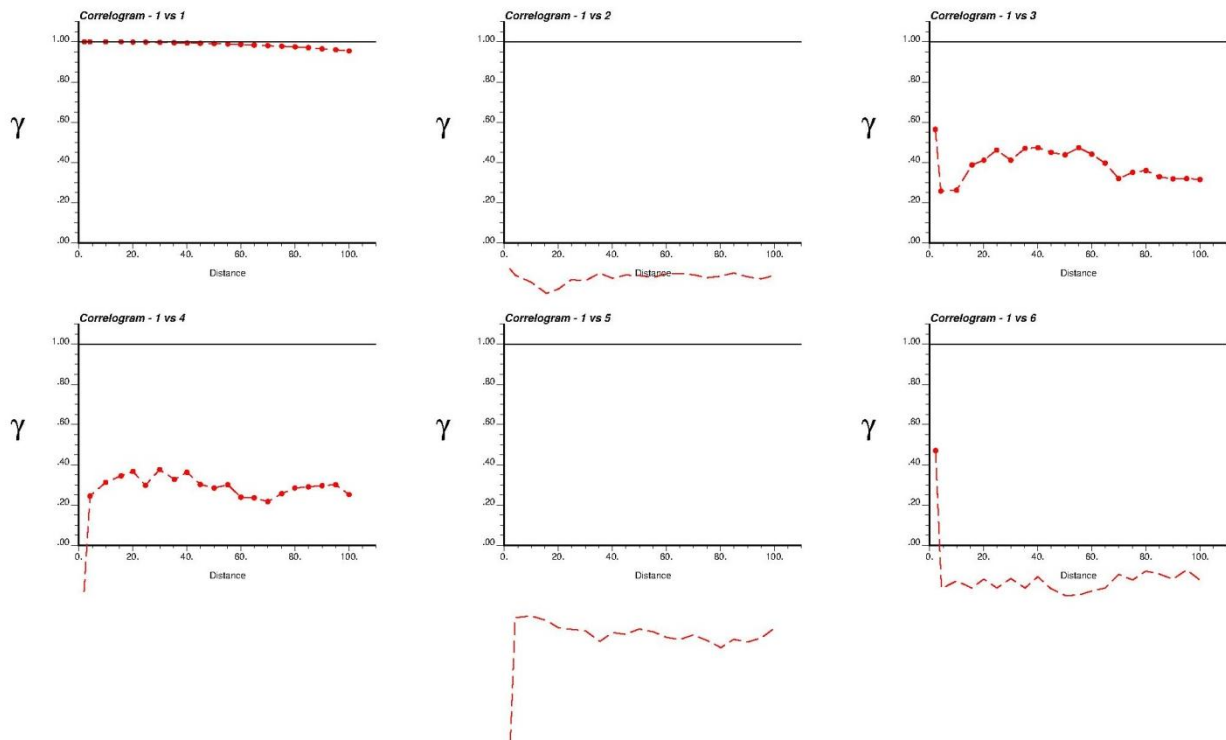
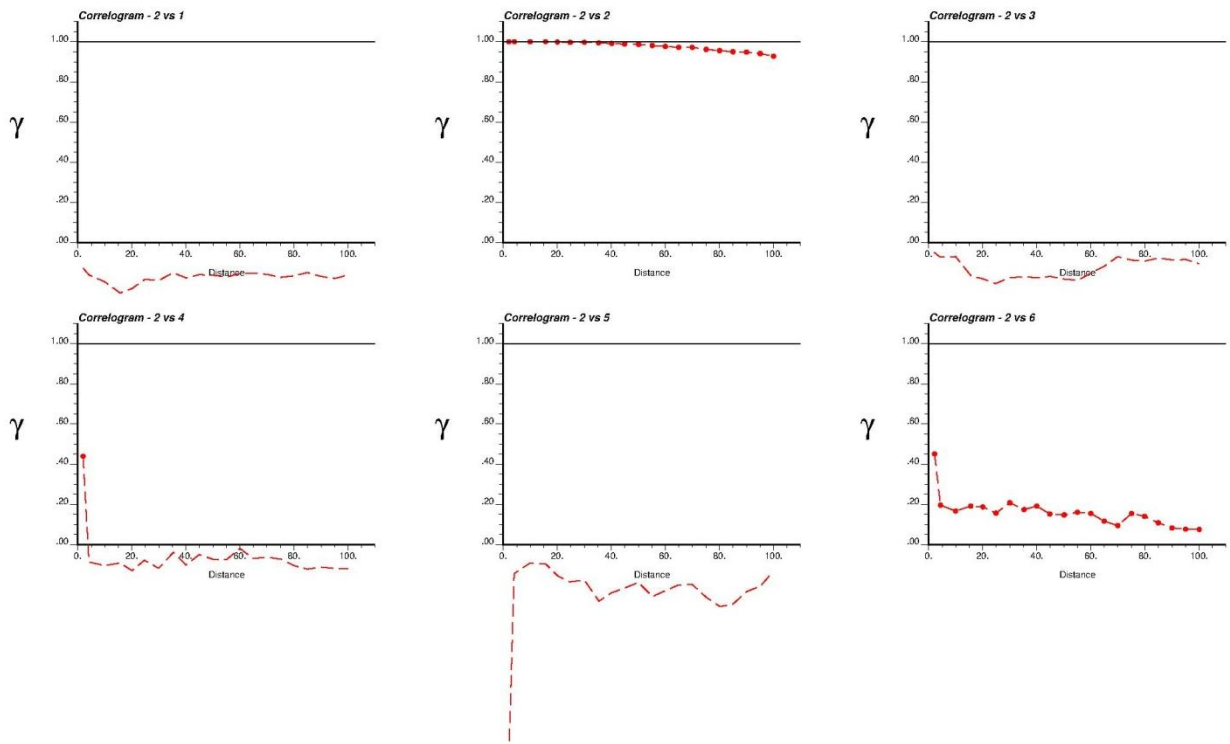
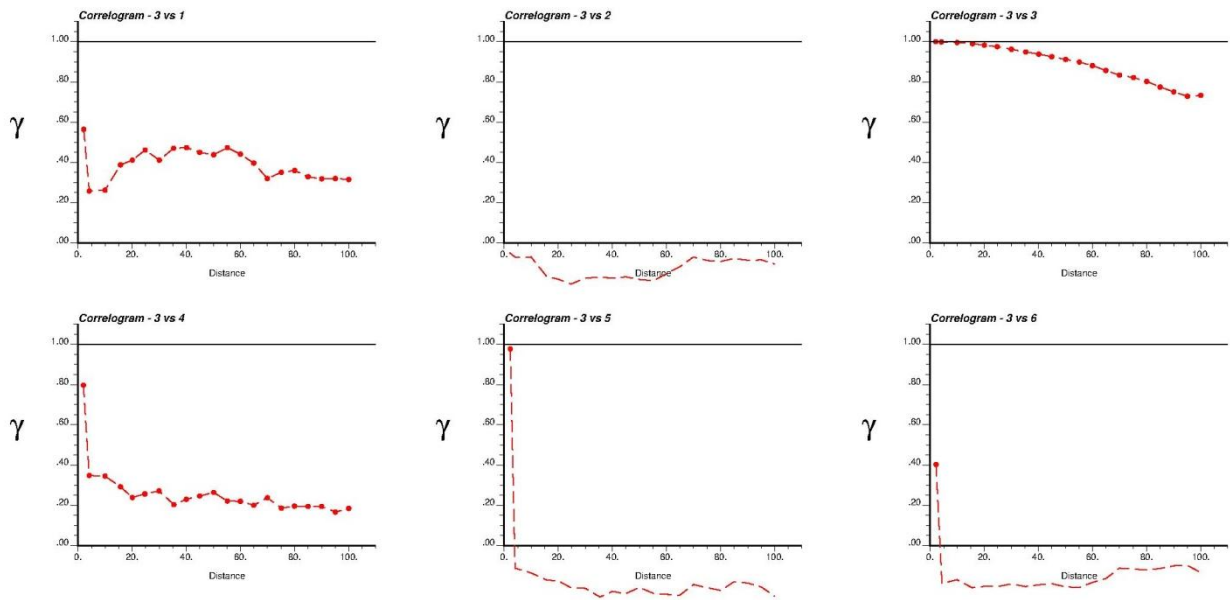


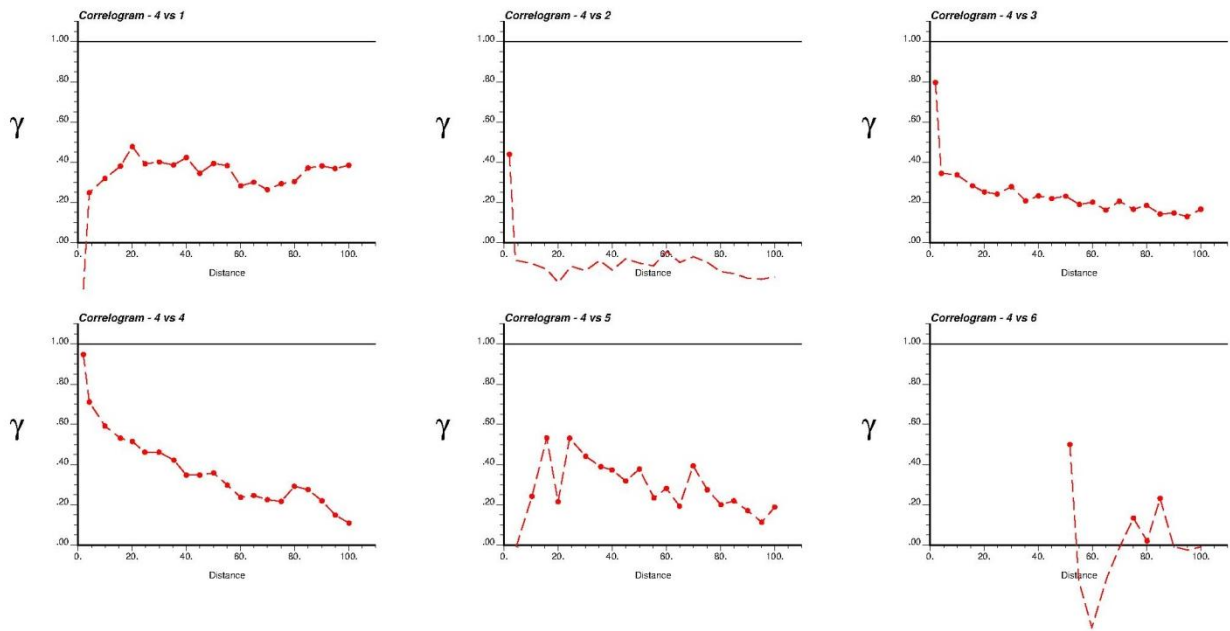
Gráfico 122: Correlogramas – UE1.



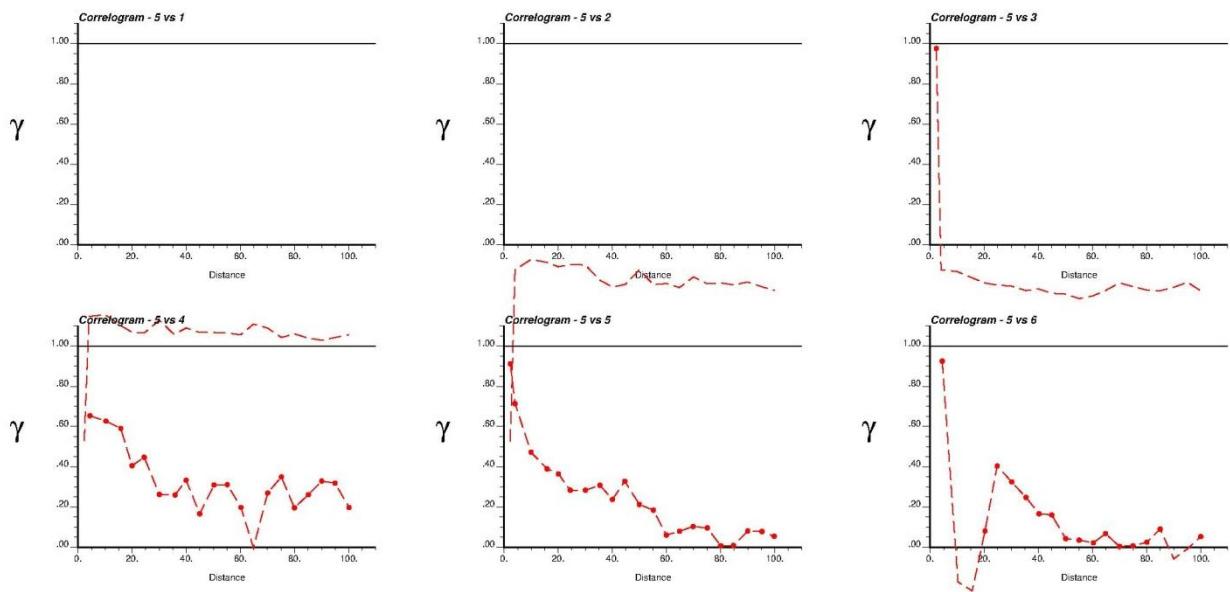
**Gráfico 123: Correlogramas – UE2.**



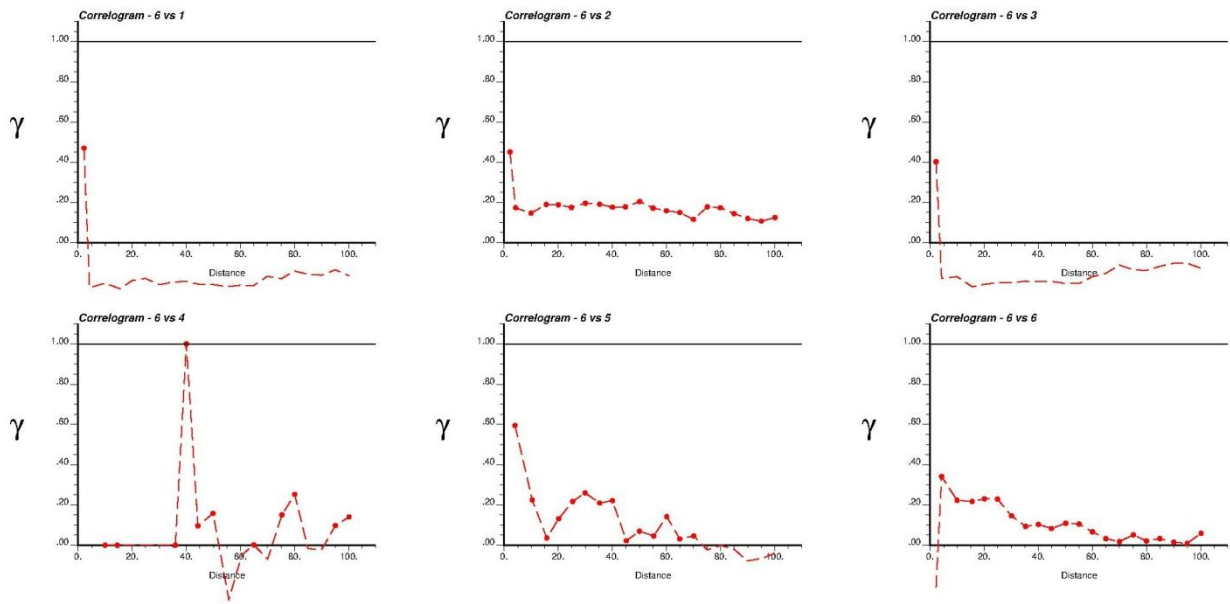
**Gráfico 124: Correlogramas – UE3.**



**Gráfico 125: Correlogramas – UE4.**



**Gráfico 126: Correlogramas – UE5.**

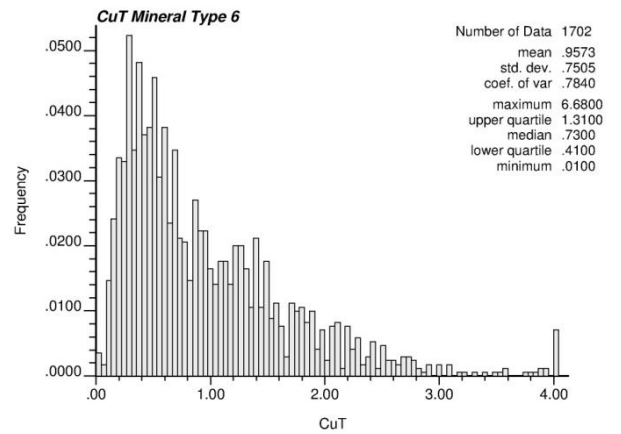
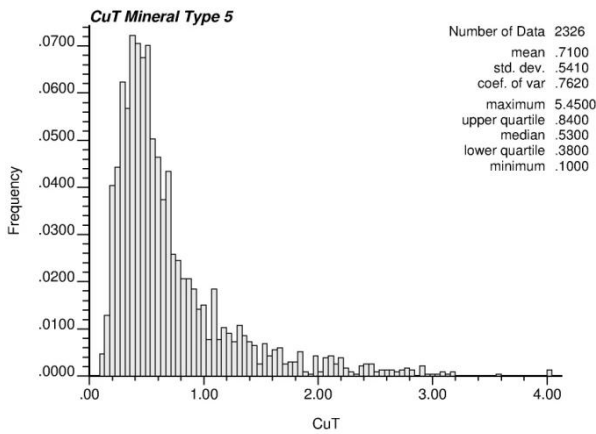
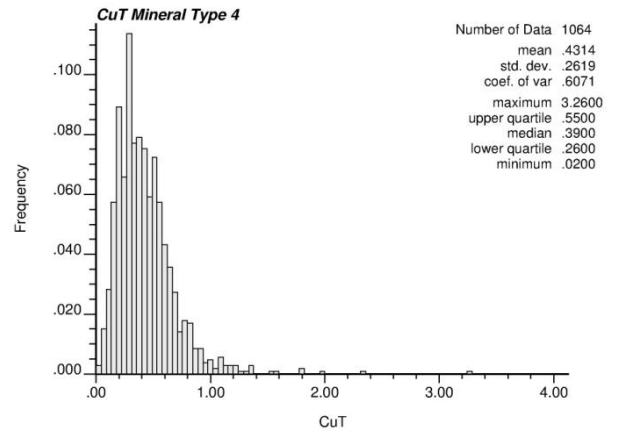
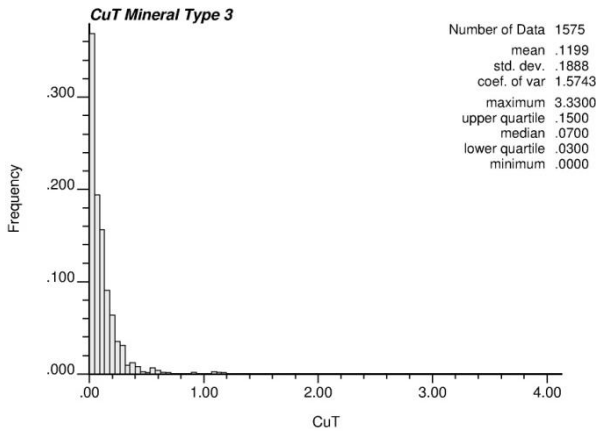
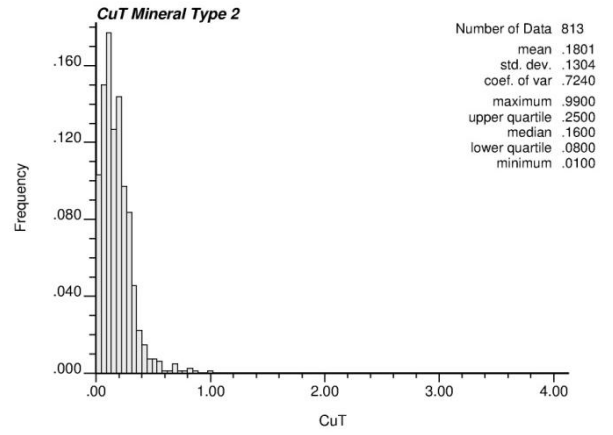
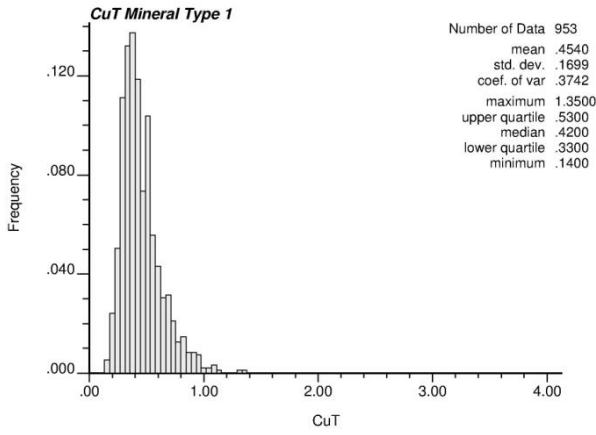


**Gráfico 127: Correlogramas – UE6.**

## **ANEXO E: Caso de estudio II – Histogramas**

El Gráfico 128 muestra los histogramas para cada una de las unidades de estimación considerando los mismos límites en cada uno: desde 0% hasta 4% de ley de cobre total. Es posible apreciar una clara distribución lognormal de la ley en las unidades 1, 3, 4 y 6.

Se puede apreciar además de los histogramas, la estrechez de los valores en torno a la media de las unidades 1, 2 y 3, tal cual se indicaba por medio de la varianza de cada unidad de estimación. Es claro también que la unidad de estimación 6 es la que posee un mayor rango de valores, mientras que la unidad 1 es la que posee el menor rango. Por último, no se aprecia a partir de los histogramas presencia de valores aberrantes que escapen drásticamente al rango de valores de cada una de las distribuciones por unidad de estimación.

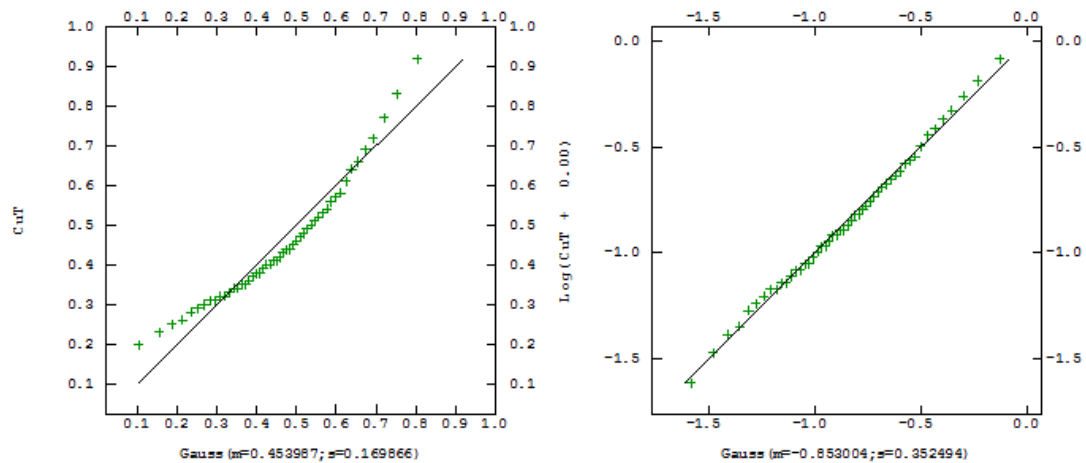


**Gráfico 128: Histogramas.**

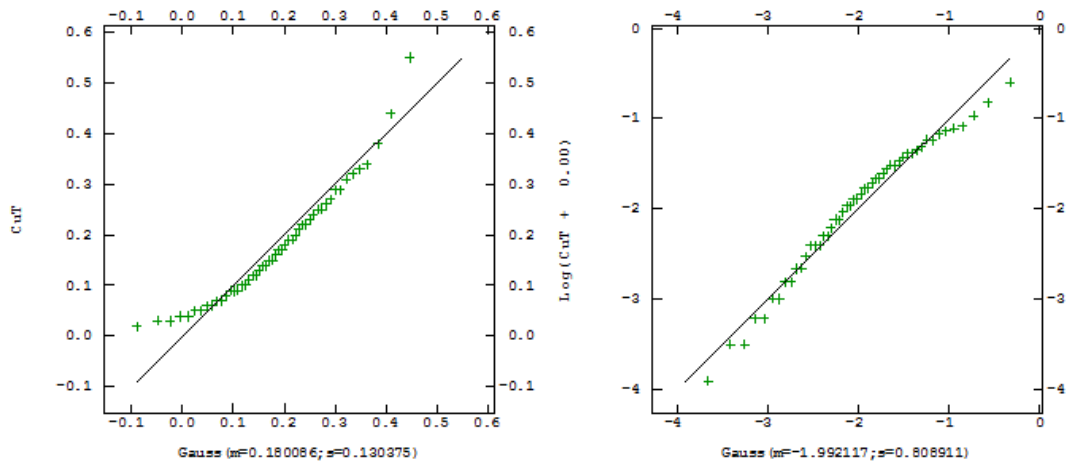


## ANEXO F: Caso de estudio II – Gráficos cuantil vs cuantil

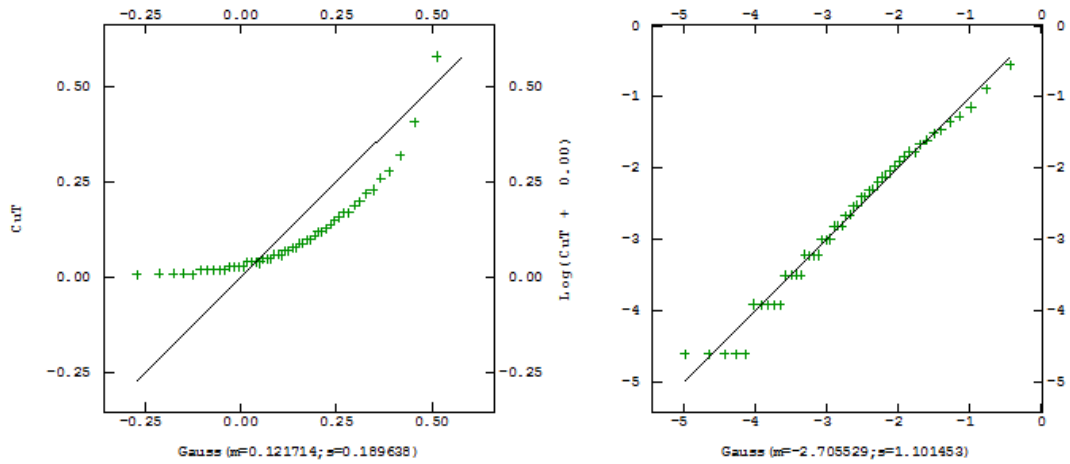
Mediante los gráficos cuantil versus cuantil se chequea la distribución de cada población. Dado que se tiene a priori una distribución lognormal dada por los histogramas, se chequea el gráfico cuantil vs cuantil para cada unidad de estimación en comparación con su distribución teórica normal y lognormal. Los siguientes gráficos muestran un buen ajuste a una distribución lognormal para cada una de las unidades de estimación (UE). Tal como se explica en la sección de histogramas, la unidad 2 es la que posee mayor diferencia con respecto a su distribución lognormal teórica, sin embargo se considerarán todas las unidades de estimación como lognormal. Es importante destacar que para la UE-3, se filtraron 23 datos con valor 0 para poder calcular su gráfico cuantil vs cuantil lognormal.



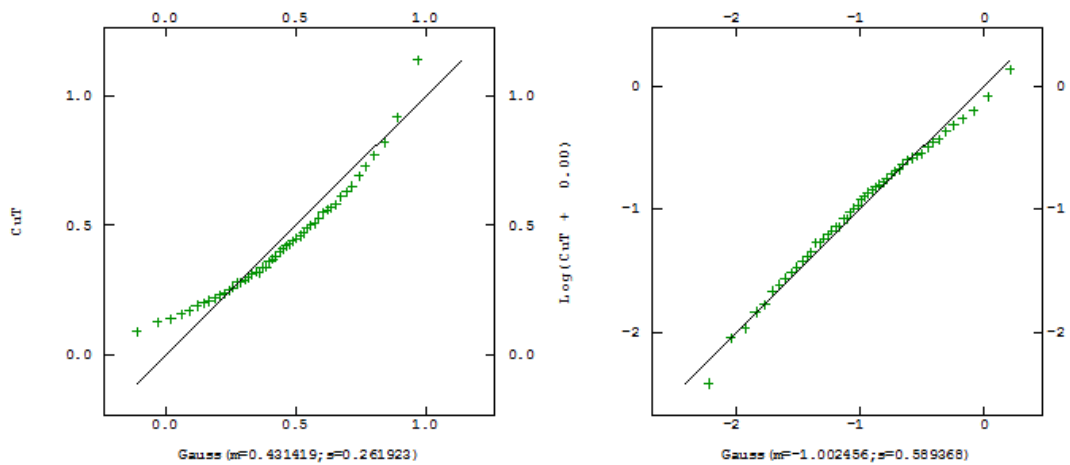
**Gráfico 129: QQ-Plot – UE 1. (Izq: Gaussiano - Der: Lognormal).**



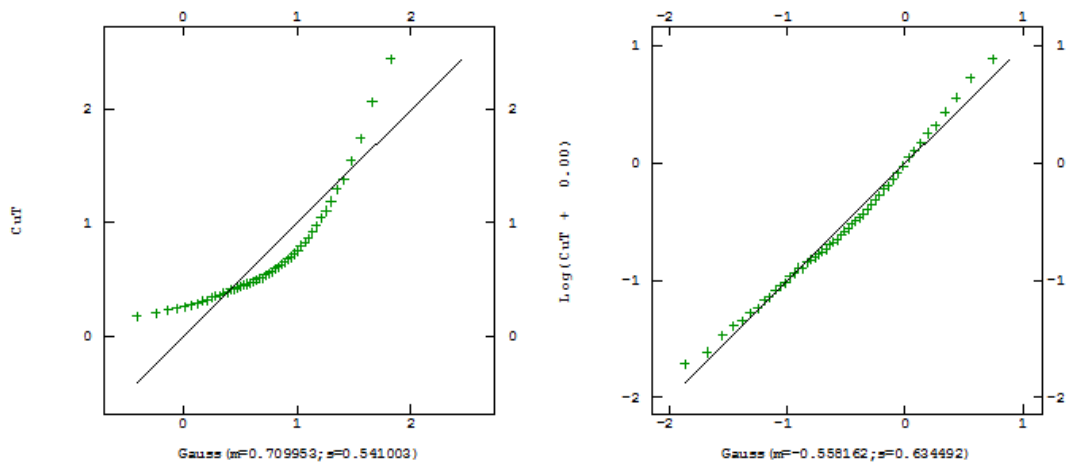
**Gráfico 130: QQ-Plot – UE 2. (Izq: Gaussiano - Der: Lognormal).**



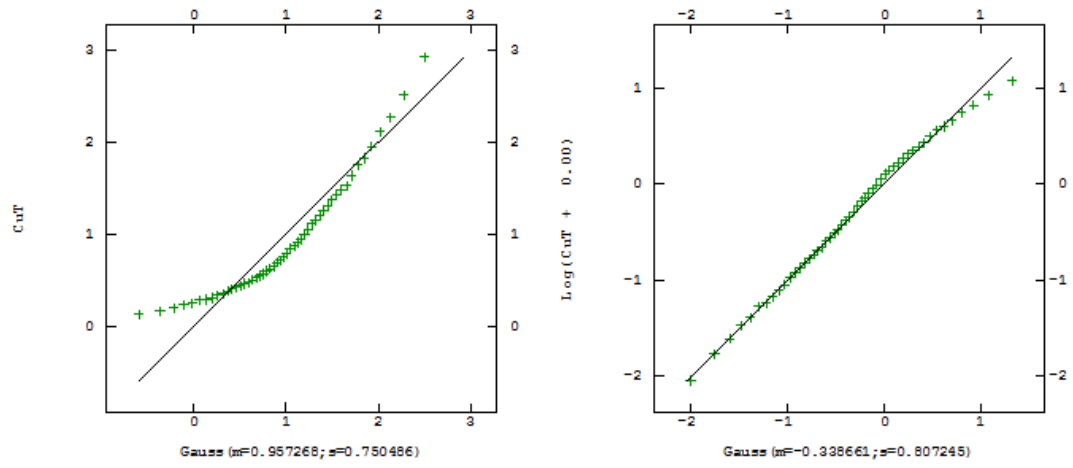
**Gráfico 131: QQ-Plot – UE 3. (Izq: Gaussiano - Der: Lognormal).**



**Gráfico 132: QQ-Plot – UE 4. (Izq: Gaussiano - Der: Lognormal).**



**Gráfico 133: QQ-Plot – UE 5. (Izq: Gaussiano - Der: Lognormal).**

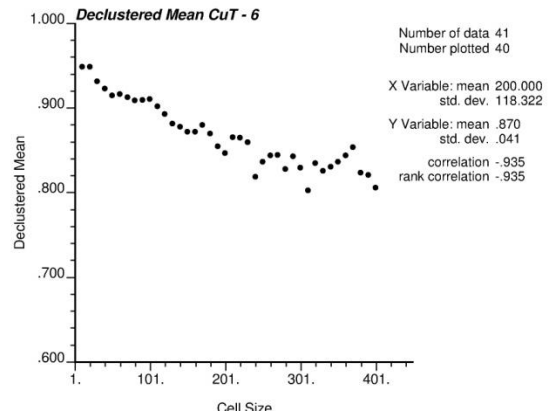
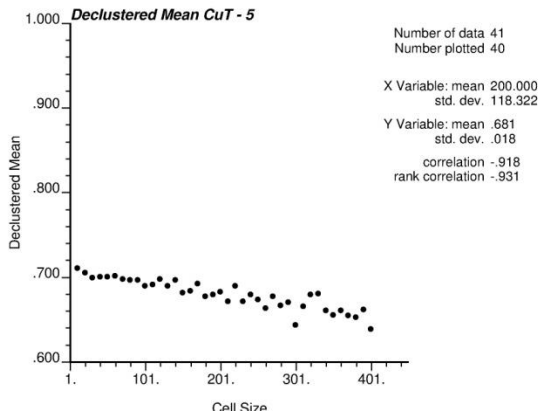
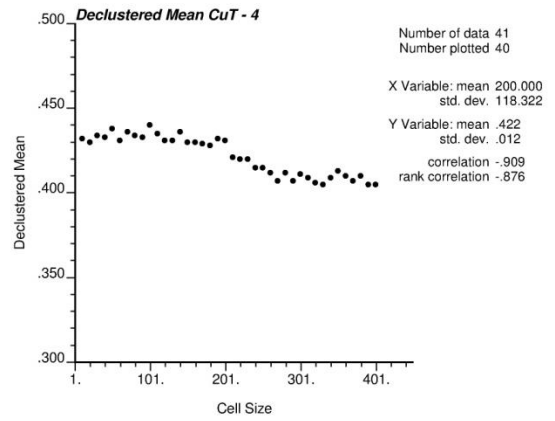
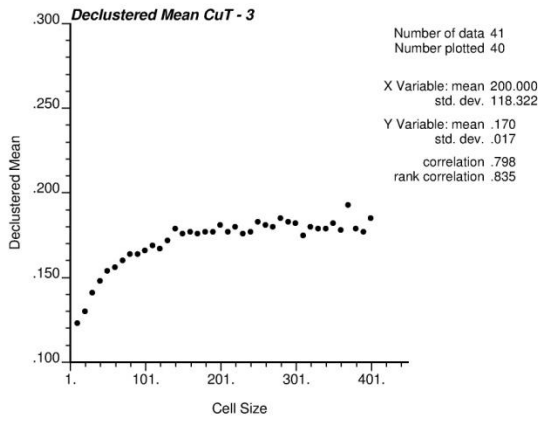
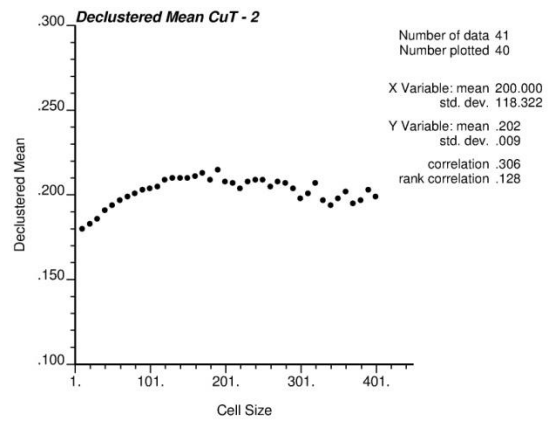
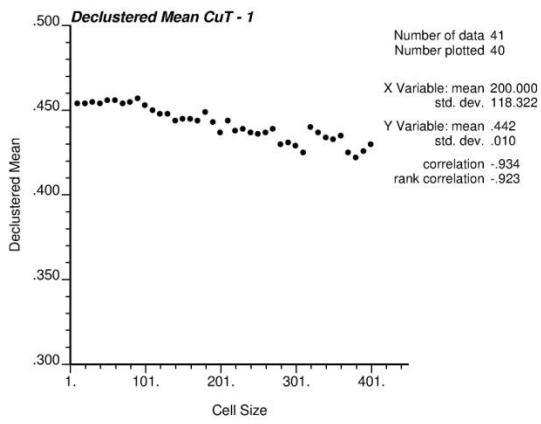


**Gráfico 134: QQ-Plot – UE 6. (Izq: Gaussiano - Der: Lognormal).**

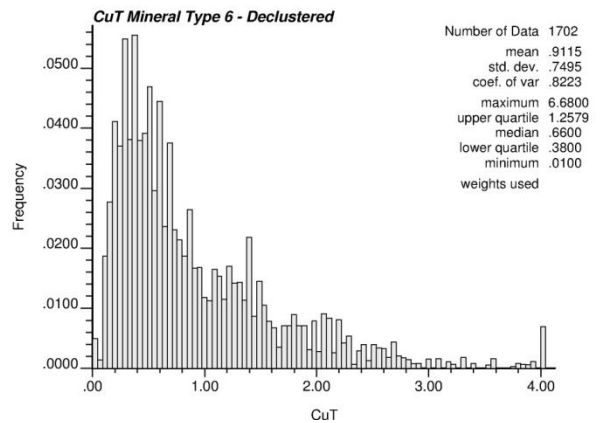
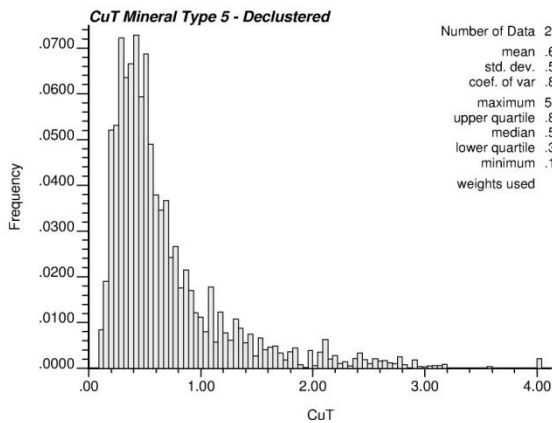
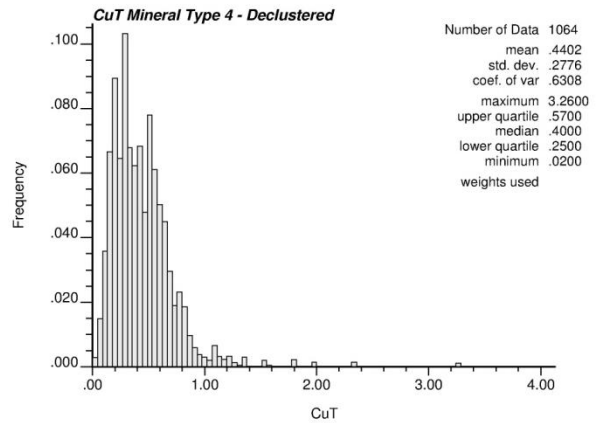
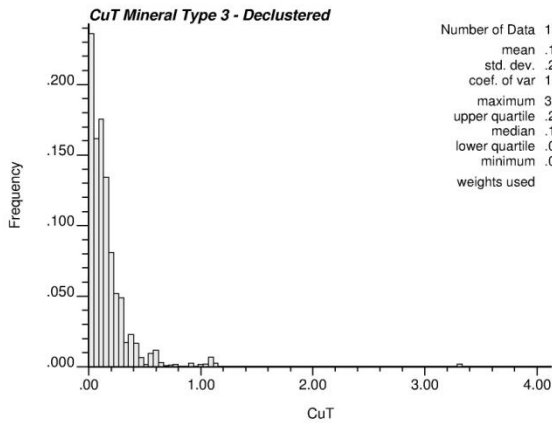
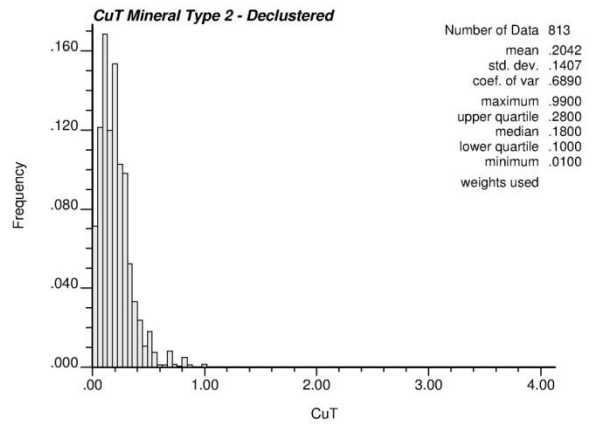
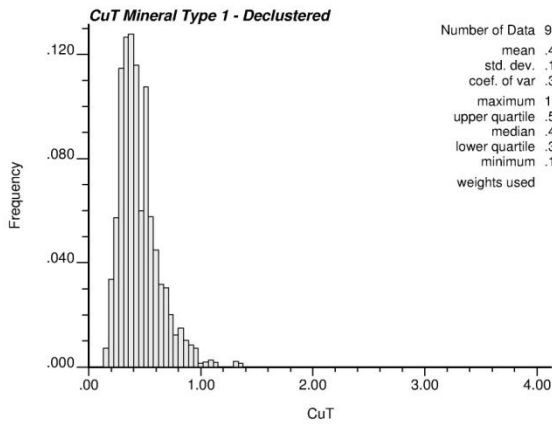
## ANEXO G: Caso de estudio II – Desagrupamiento y estadísticas desagrupadas

El proceso de desagrupamiento permite asignar un ponderador (peso) a cada una de las muestras de tal manera que se corrigen los efectos de un muestreo irregular. De esta manera se asignan ponderadores de bajo valor a aquellas muestras en zonas de muestreo preferente y ponderadores de alto valor a aquellas muestras que abarquen una mayor zona del espacio muestreado. Se utiliza un desagrupamiento por el método de celdas, el cual subdivide el dominio en celdas de igual tamaño, todas las cuales tienen un mismo ponderador. El valor ponderador asignado a cada muestra es inversamente proporcional al número de muestras en la respectiva celda. Para realizar este proceso por tanto se debe elegir primeramente un tamaño de celda, para lo cual se puede graficar la media desagrupada en función del tamaño de celda (Gráfico 135).

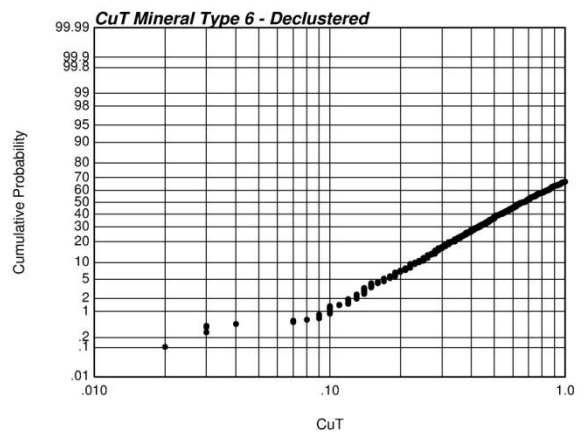
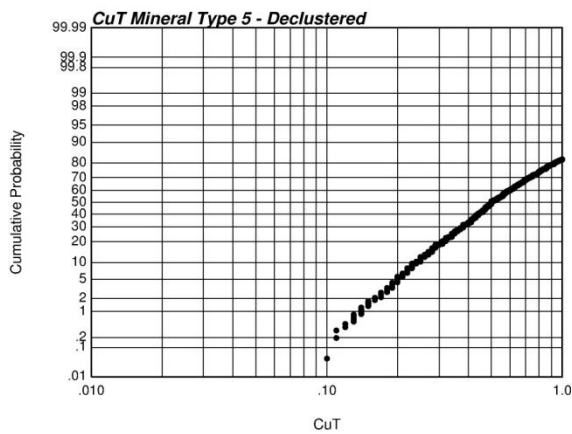
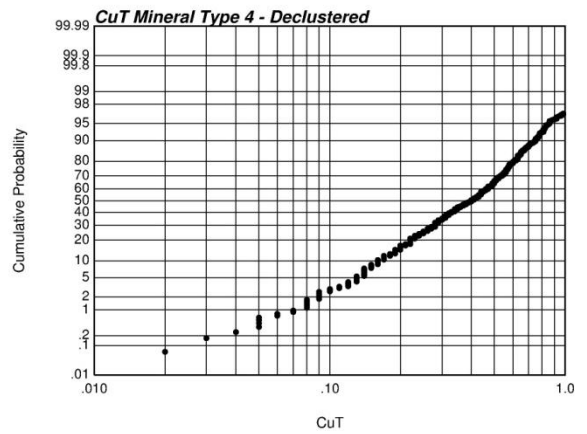
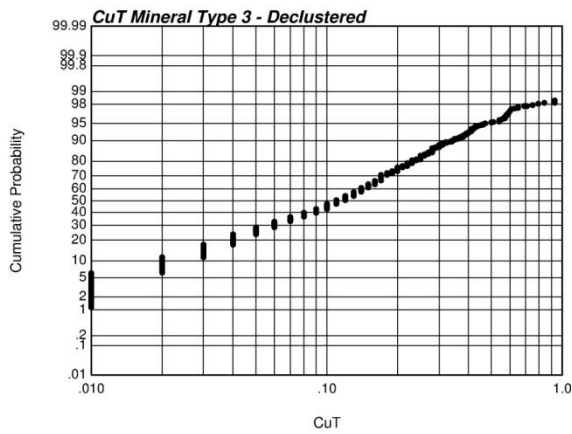
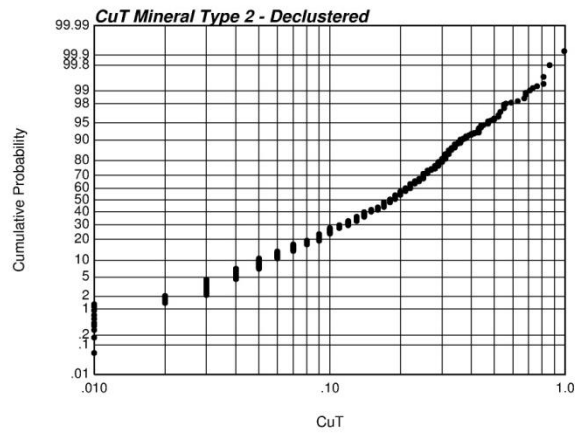
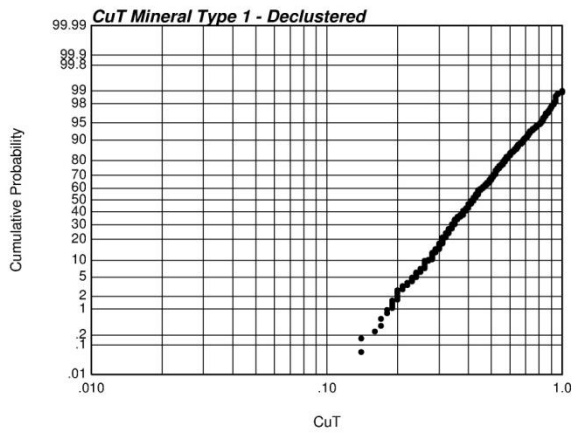
Se presenta la ley media desagrupada utilizando una anisotropía (1, 1, 0.25) en las coordenadas (X, Y, Z) respectivamente. La coordenada X varía desde 10 hasta 400 metros con un paso incremental de 10 metros, es decir, la celda mínima es de  $10 \times 10 \times 2.5 \text{ m}^3$  y la celda final de  $400 \times 400 \times 100 \text{ m}^3$ . Es posible apreciar dos tipos de tendencias en el Gráfico 135: una disminución de la ley media desagrupada a medida que aumenta el tamaño de la celda (UE1, UE4, UE5 y UE6) y, un aumento de dicha ley a medida que aumenta el tamaño de celda (UE 2 y UE3). El primer caso se produce cuando los ponderadores tienden a privilegiar las zonas con baja ley, es decir, se tiene un muestreo más denso en zonas de alta ley (lo que es común en la práctica y por ende apreciable en 4 de 6 UE en este caso). El segundo escenario corresponde al caso contrario, es decir, cuando el muestreo es preferencial en zonas de baja ley. Si bien existen diversos criterios para determinar el tamaño de celda, los cuales buscan compensar el muestreo preferencial, en este caso se utilizará un criterio relativamente neutral: aquella celda que ajuste a una pseudo malla regular de  $100 \times 100$  en planta posible de ver en ciertas zonas de las UE.



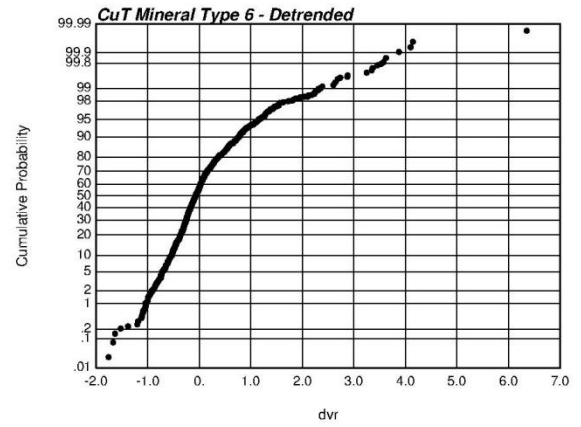
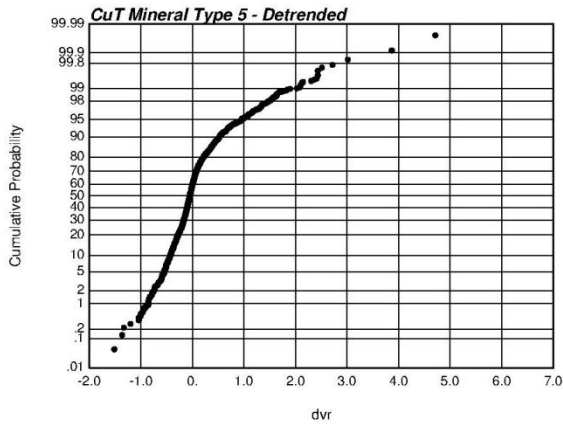
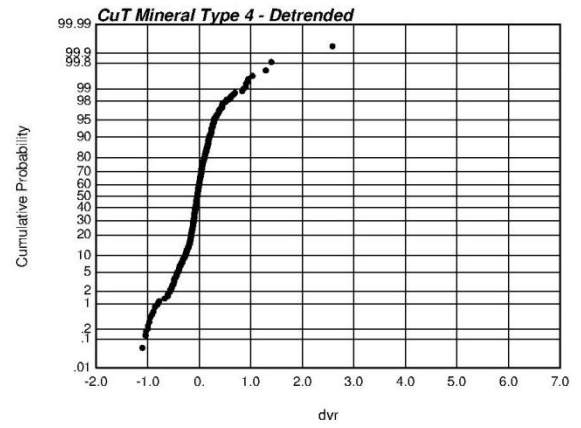
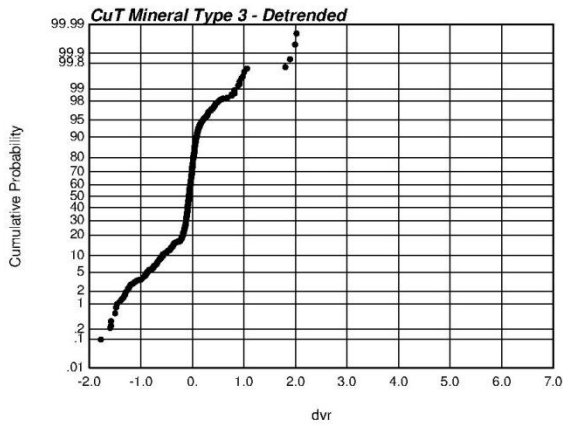
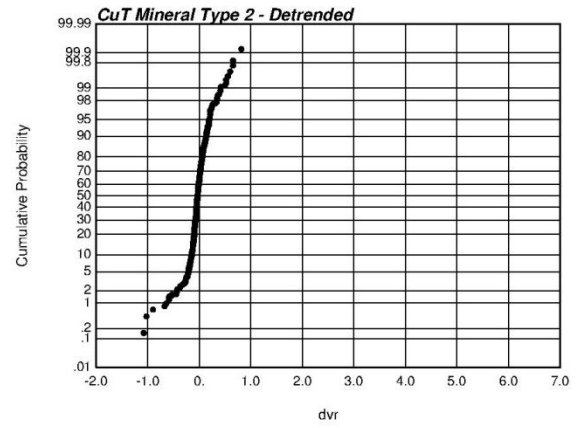
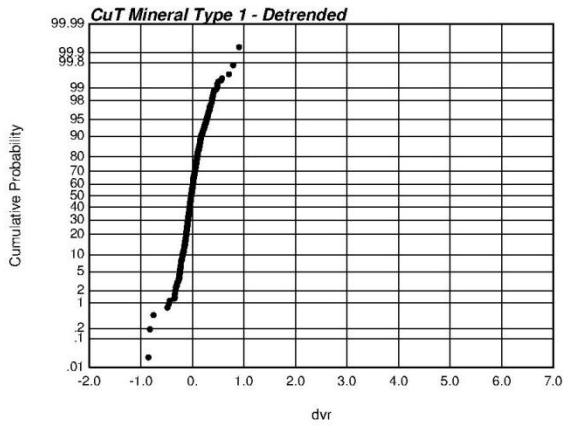
**Gráfico 135: Media desagrupada por tamaño de celda.**



**Gráfico 136: Histogramas desagrupados.**

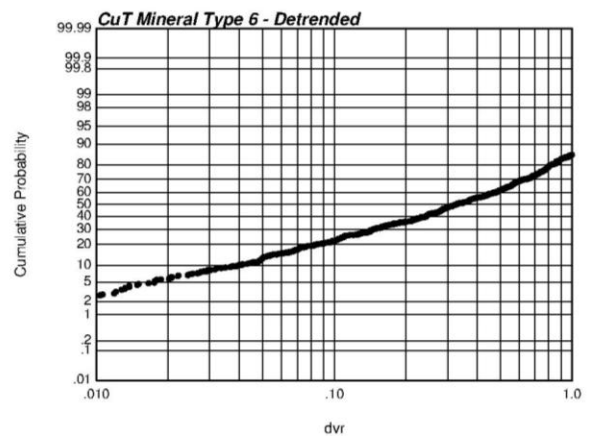
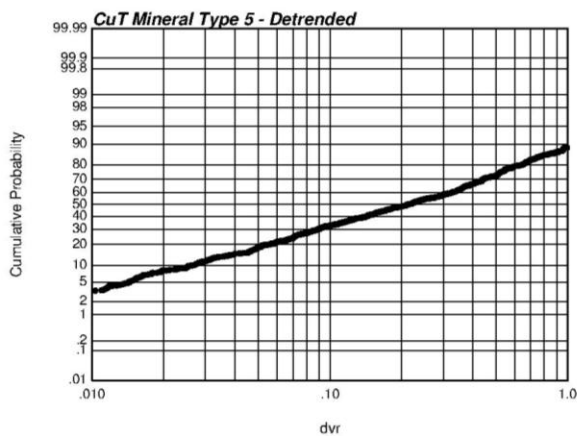
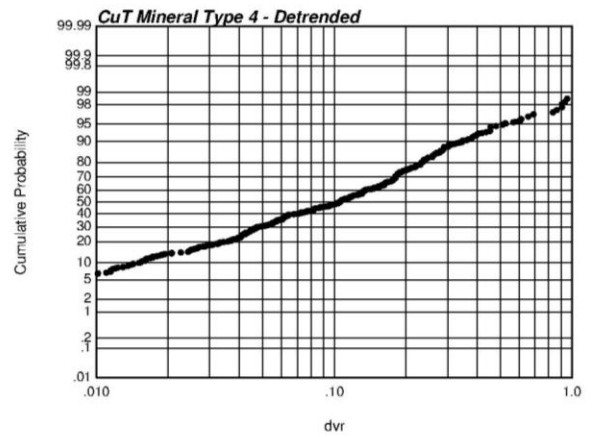
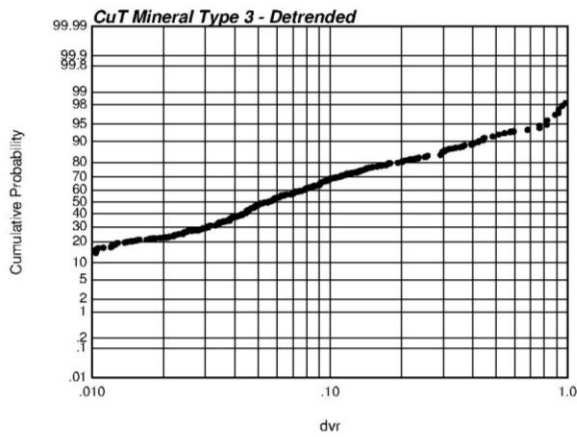
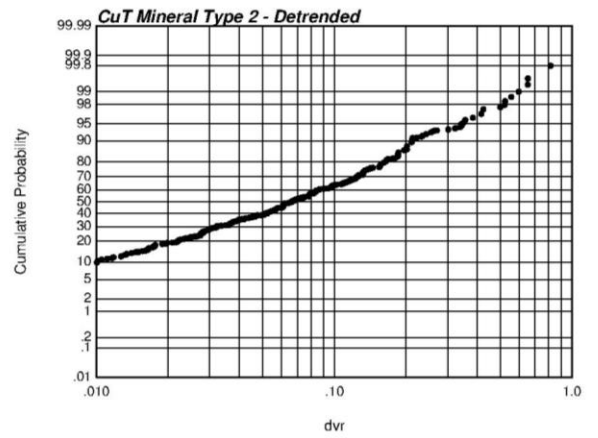
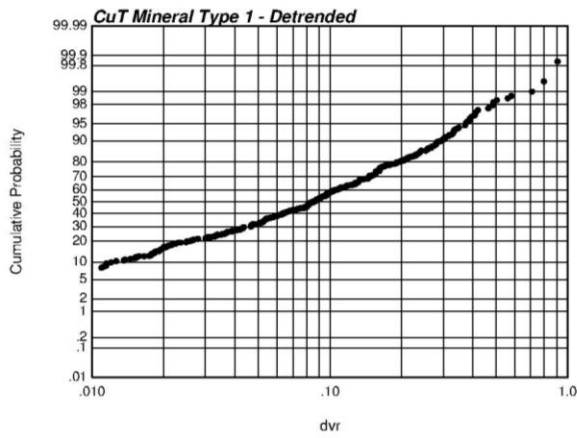


**Gráfico 137: Gráficos de probabilidad desagrupado – Lognormal.**



**Gráfico 138: Gráficos de probabilidad desagrupado de residuos – Normal.**





**Gráfico 139: Gráficos de probabilidad desagrupado de residuos – Lognormal.**

## ANEXO H: Caso de estudio II – Validación hipótesis multigaussiana

Se presentan las estadísticas de los valores normalizados de la ley desagrupada de cobre para cada una de las unidades de estimación para los compósitos originales como para los residuos.

### *CuT by Tmin - Arithmetic - Declustered Nscore*

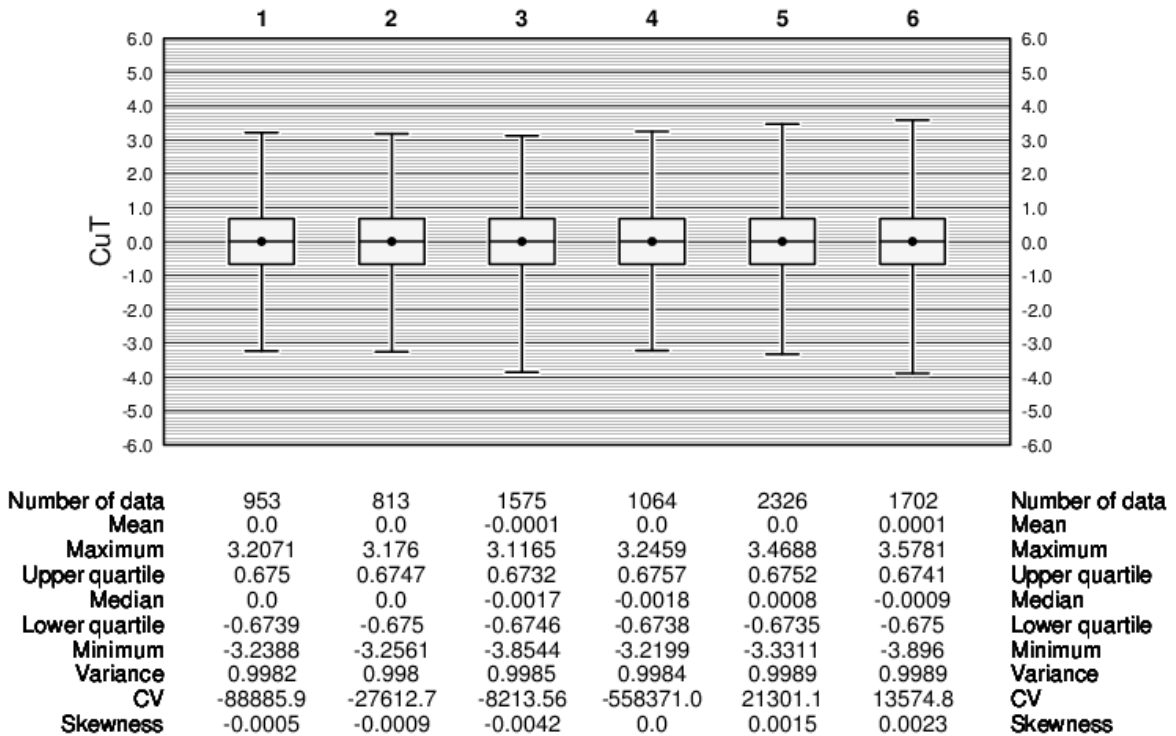
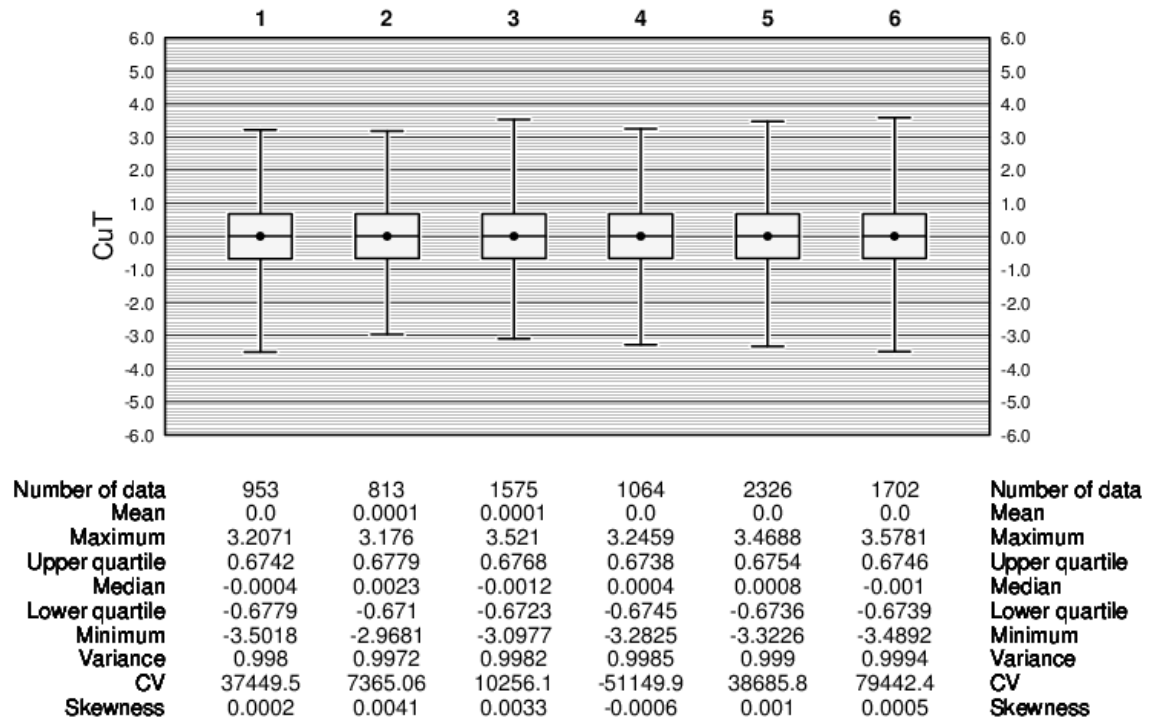


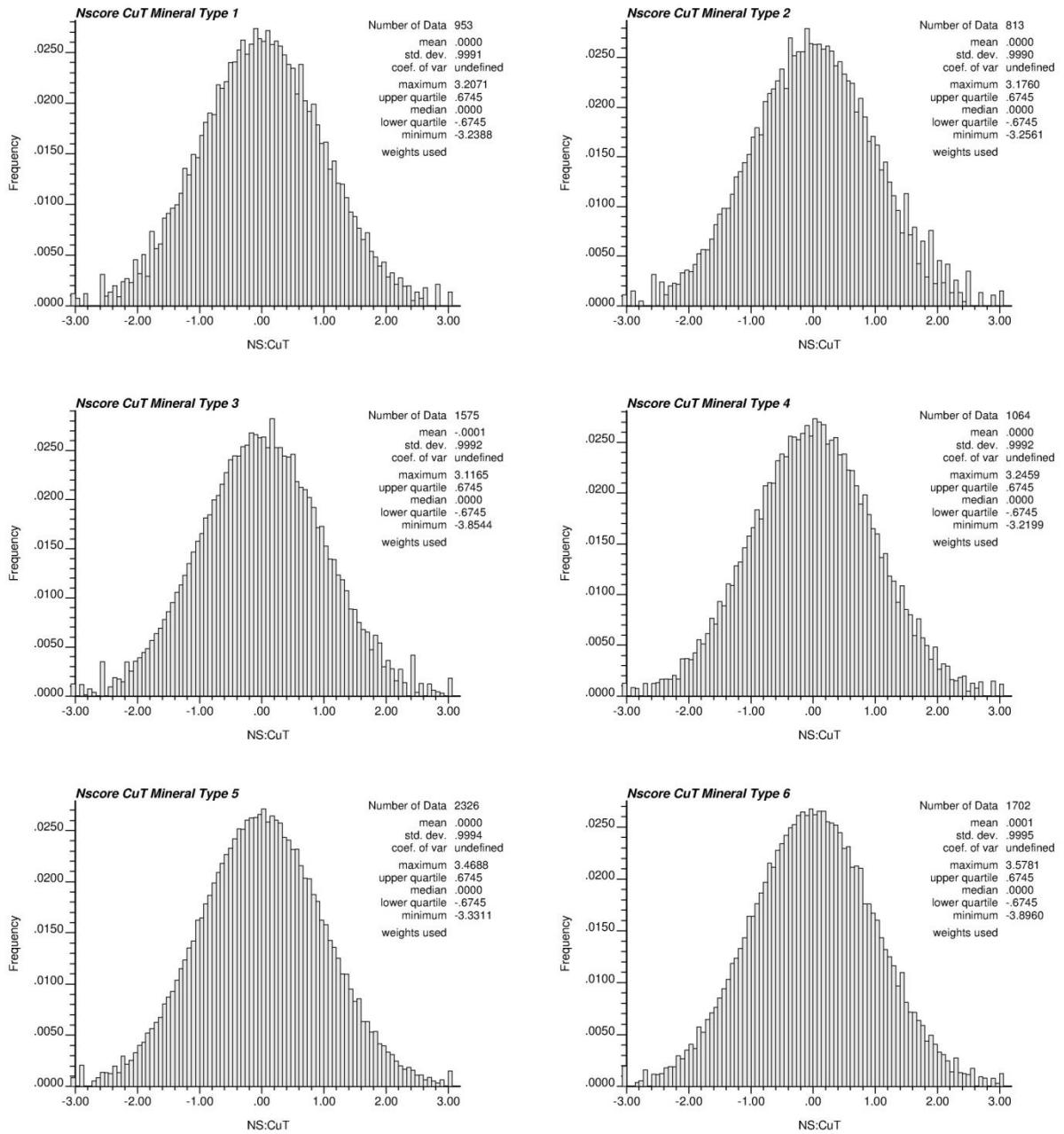
Gráfico 140: Box plot de CuT valores normalizados por unidad de estimación.

**CuT by Tmin - Arithmetic - Declustered Nscore - Detrended**

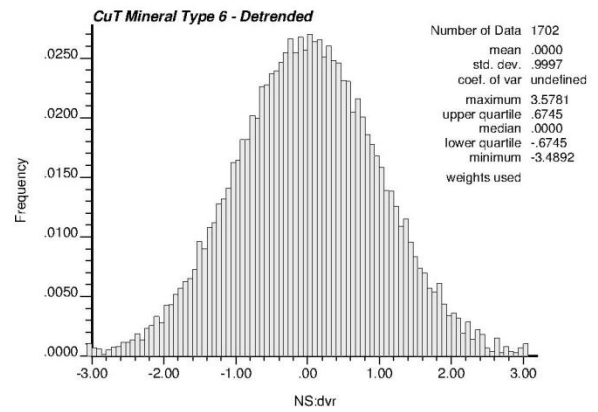
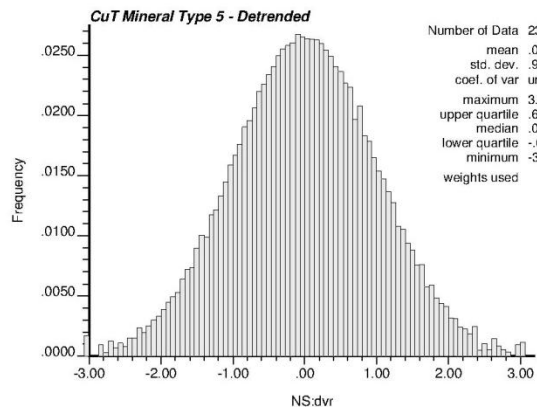
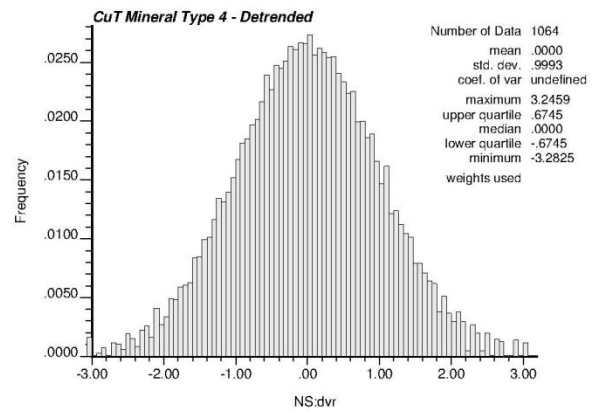
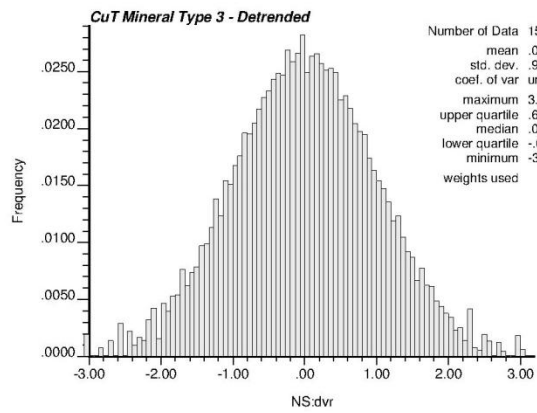
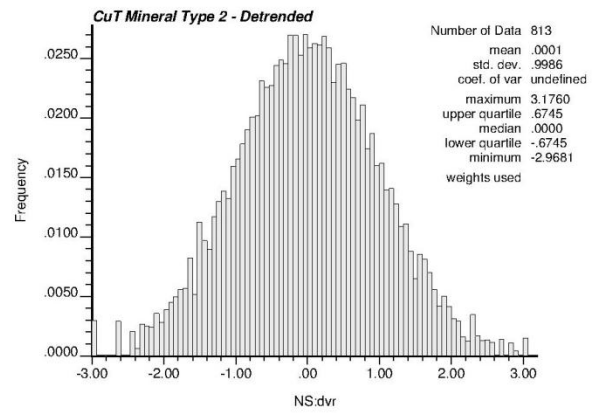
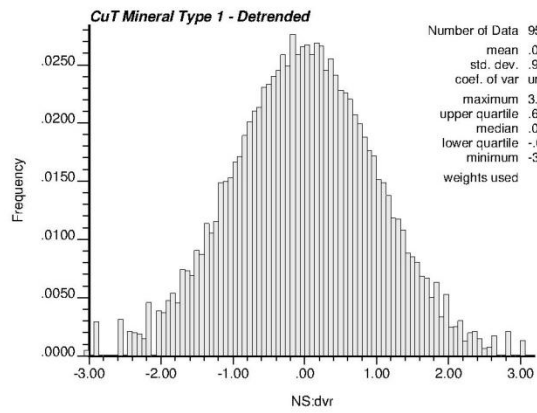


**Gráfico 141: Box plot de CuT valores normalizados de residuos por unidad de estimación.**

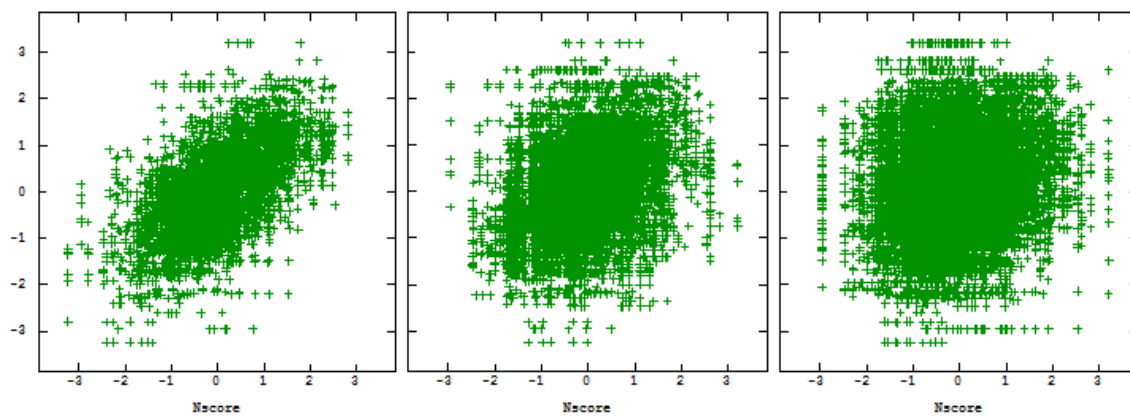
Se presenta en esta sección las estadísticas de orden 1 (histogramas) para los compósitos originales y los residuos, y de orden 2 (nubes de correlación diferida) sólo para los compósitos originales. Esto permite validar la hipótesis de gaussianidad de las distribuciones univariable (para compósitos originales y residuos) y bivariable (sólo para compósitos originales), siendo por extensión válida la hipótesis multigaussiana.



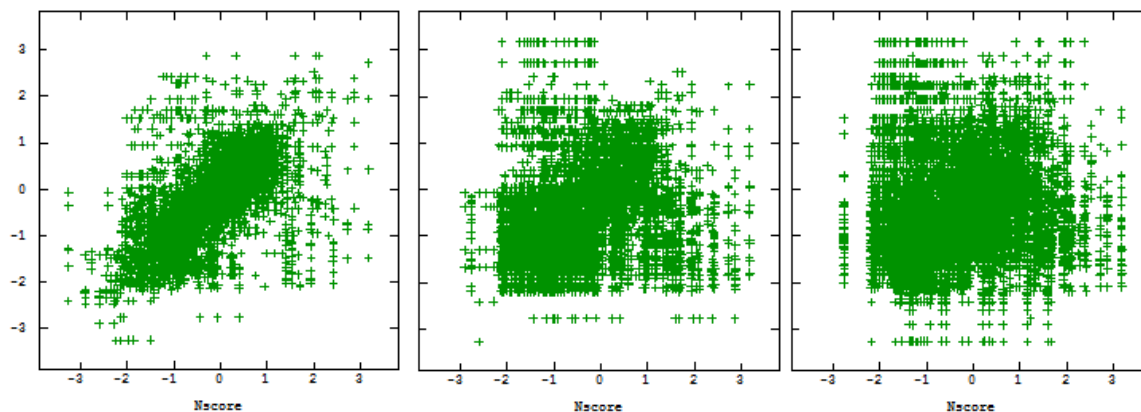
**Gráfico 142: Histogramas valores Gaussianos – Compósitos originales.**



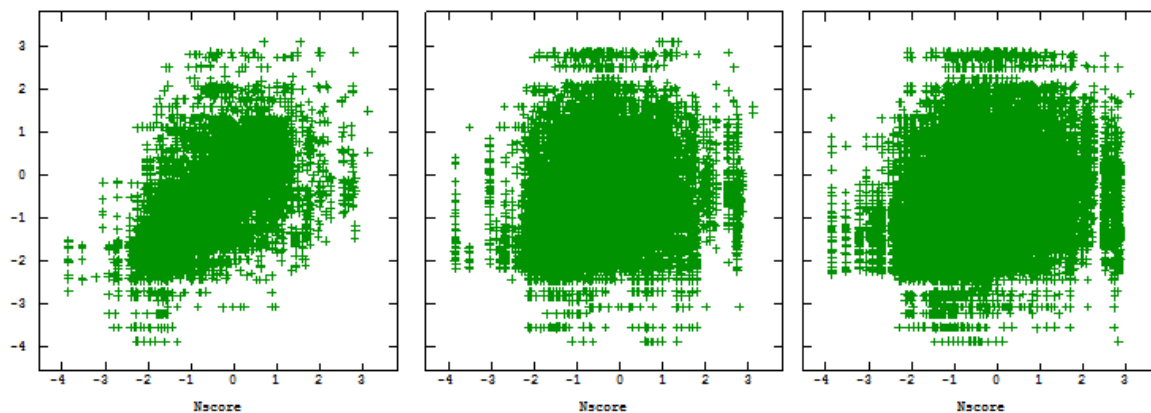
**Gráfico 143: Histogramas valores Gaussianos - Resíduos.**



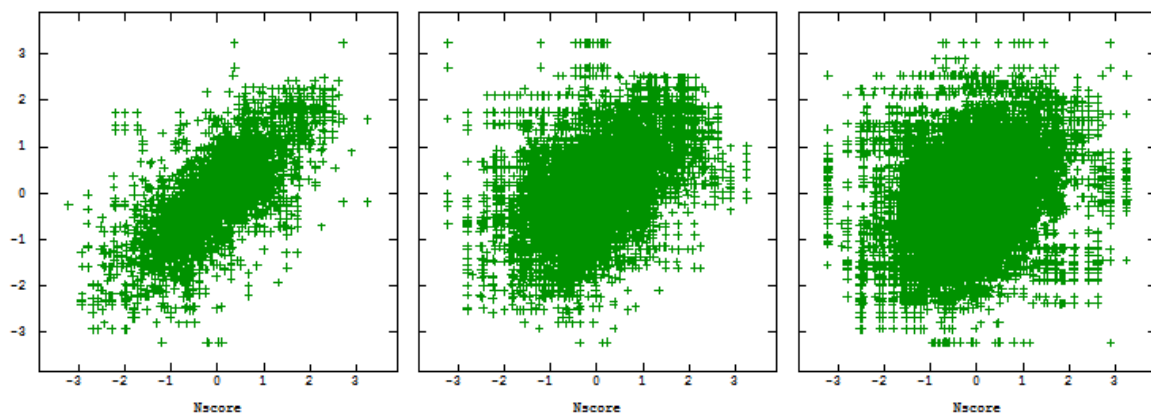
**Gráfico 144: Nube de correlación diferida–UE1 (Izq:25-50m - Cen:50-75m - Der:100-125m).**



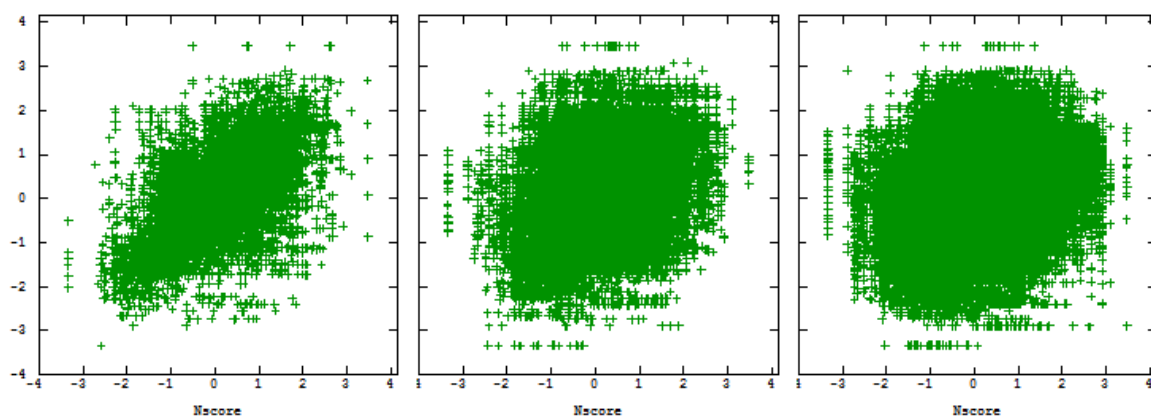
**Gráfico 145: Nube de correlación diferida–UE2 (Izq:25-50m - Cen:50-75m - Der:100-125m).**



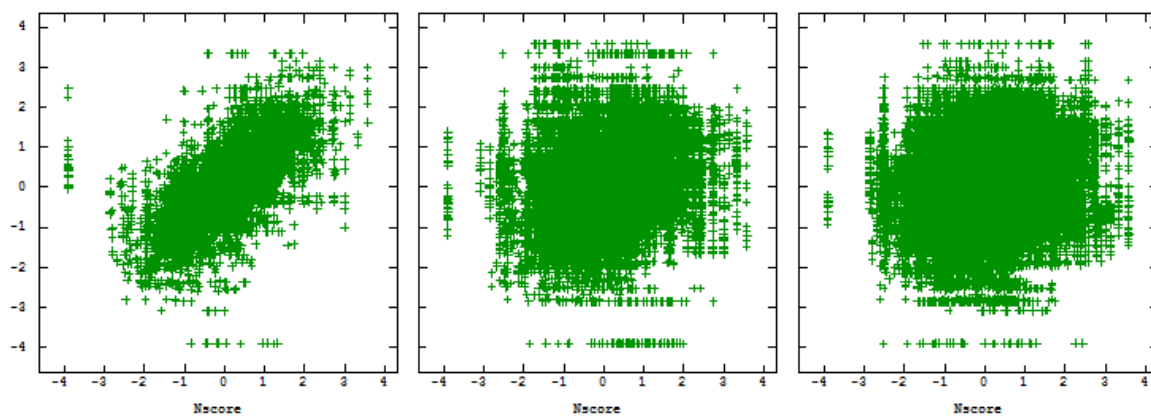
**Gráfico 146: Nube de correlación diferida–UE3 (Izq:25-50m - Cen:50-75m - Der:100-125m).**



**Gráfico 147:Nube de correlación diferida–UE4 (Izq:25-50m - Cen:50-75m - Der:100-125m).**



**Gráfico 148:Nube de correlación diferida–UE5 (Izq:25-50m - Cen:50-75m - Der:100-125m).**



**Gráfico 149:Nube de correlación diferida–UE6 (Izq:25-50m - Cen:50-75m - Der:100-125m).**

## ANEXO I: Caso de estudio II – Mapas variográficos

Mediante el mapa variográfico se visualiza el variograma experimental para todas las direcciones en un plano mediante una escala de colores. De esta manera se logra identificar más fácilmente posibles direcciones de anisotropía. Para cada unidad de estimación se utiliza la misma búsqueda en planta mediante 6 pasos de 50 metros tanto en X como en Y, de esta manera se estudia un alcance de 300x300 en XY. Ahora bien, dado que en la vertical las UE tienen distintas potencias, se utiliza una búsqueda distinta dependiendo del caso, manteniéndose constante el largo del paso (10 metros) pero variando el número de pasos según la Tabla 44.

UE	N° de pasos	Alcance
1	20	200
2	20	200
3	25	250
4	15	150
5	25	250
6	20	200

Tabla 44: N° de pasos en Z por UE.

### I.1 Compósitos originales

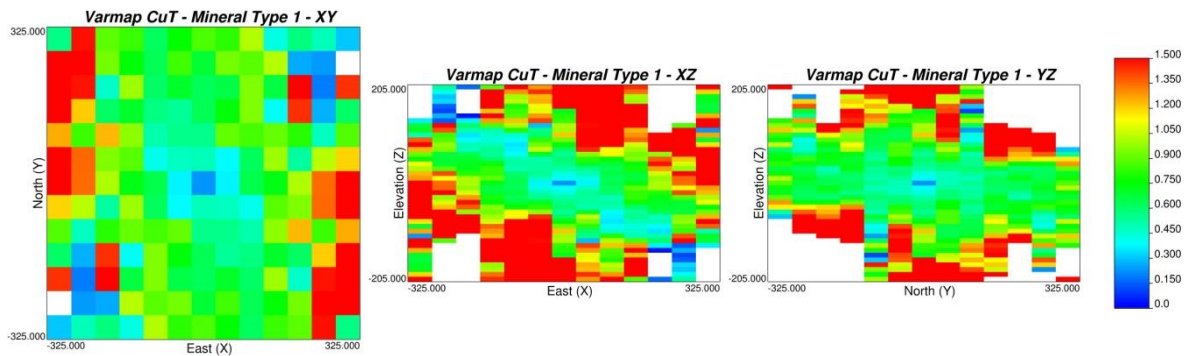
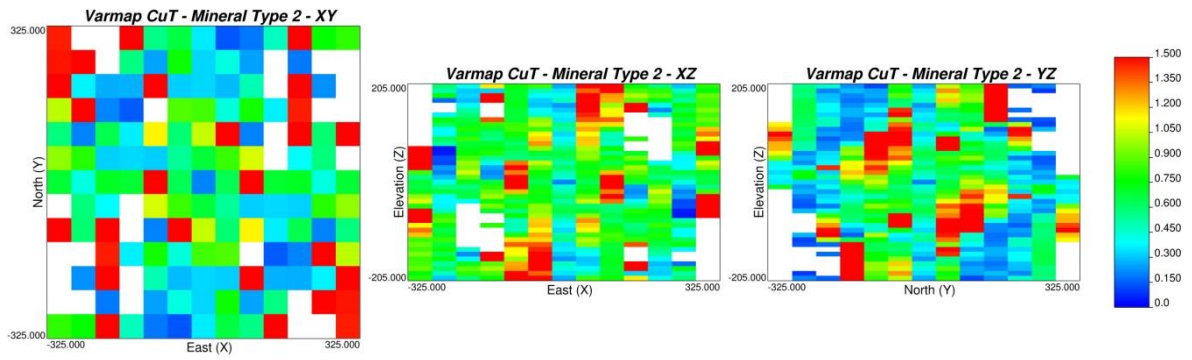
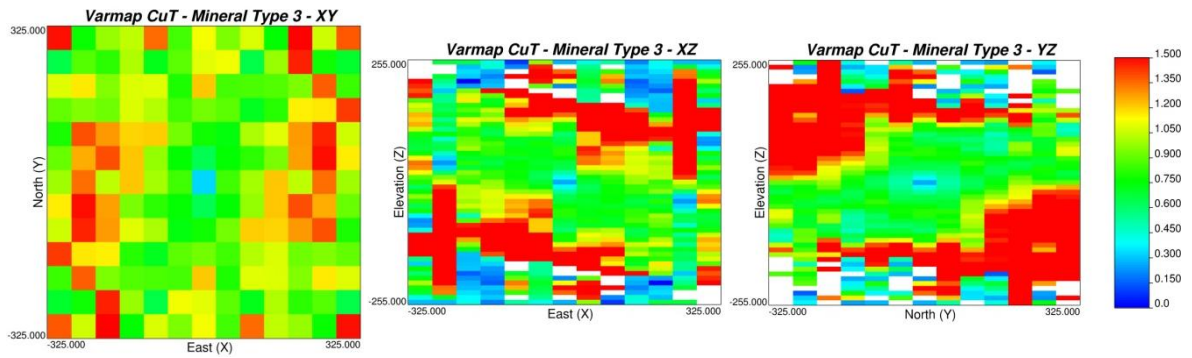


Figura 18: Mapa variográfico – UE1 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.

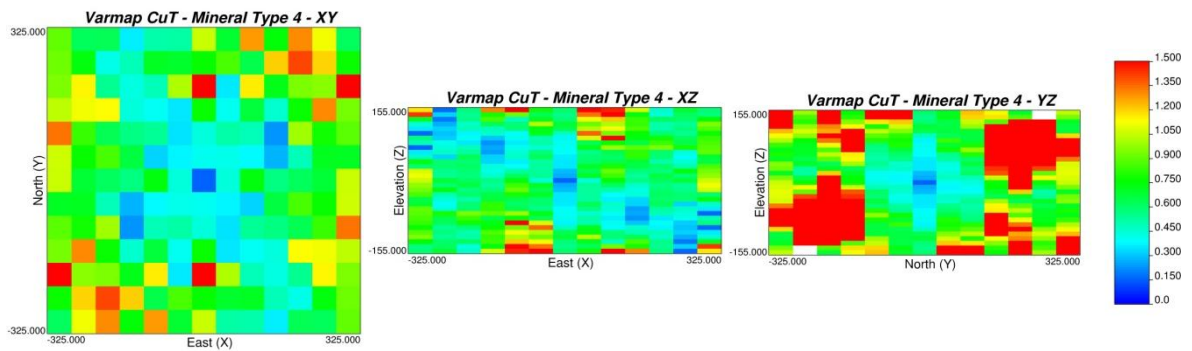




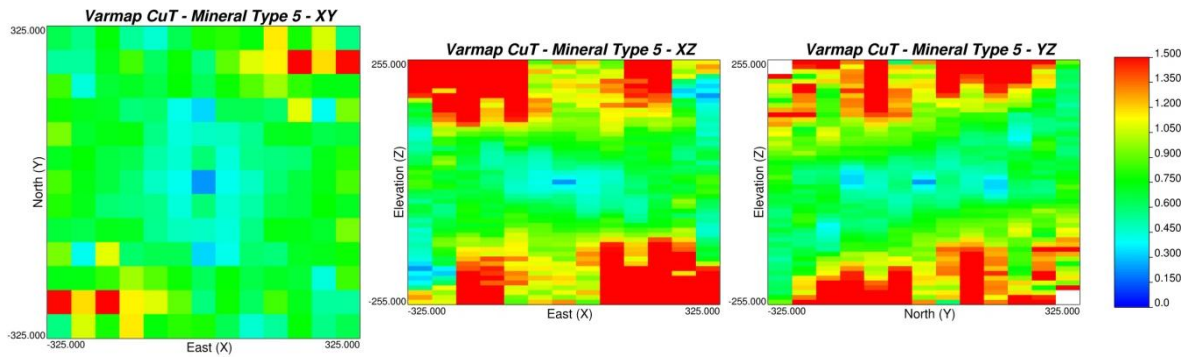
**Figura 19: Mapa variográfico – UE2 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.**



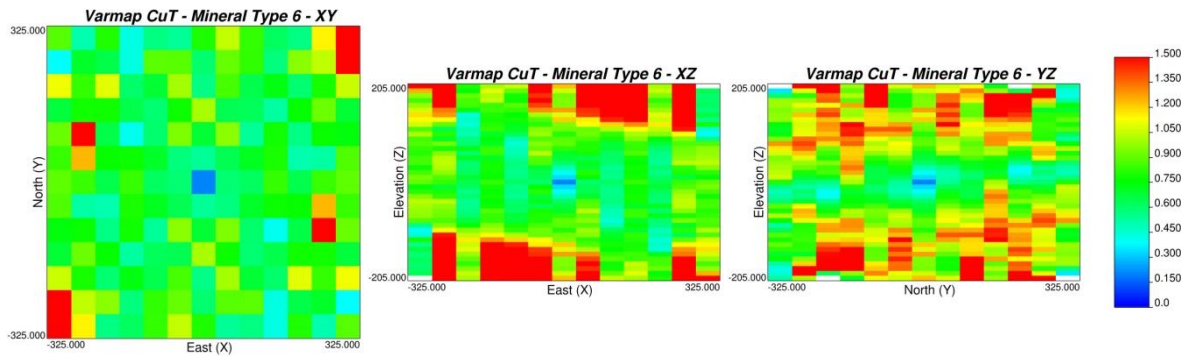
**Figura 20: Mapa variográfico – UE3 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.**



**Figura 21: Mapa variográfico – UE4 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.**



**Figura 22: Mapa variográfico – UE5 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.**



**Figura 23: Mapa variográfico – UE6 (Izq: XY - Cen: XZ - Der: YZ) – Compósitos.**

Los mapas variográficos para la UE1 sugieren anisotropía geométrica entre el plano horizontal rotado en un dip de  $20^\circ$  y su vertical. En el caso de la UE2 y UE6 los mapas variográficos no sugieren anisotropía a simple vista por lo que se realiza un estudio en mayor detalle mediante el cálculo de sus variogramas experimentales. Para la UE3 y UE5, los mapas sugieren una anisotropía geométrica entre el plano horizontal y la vertical, además en la UE5 existe la posibilidad de tener una anisotropía geométrica en planta con una elipse rotada en  $340^\circ$ . En la UE4, el mapa XY podría indicar cierta anisotropía geométrica en planta, lo cual será revisado en mayor detalle en el Anexo J.

## I.2 Residuos

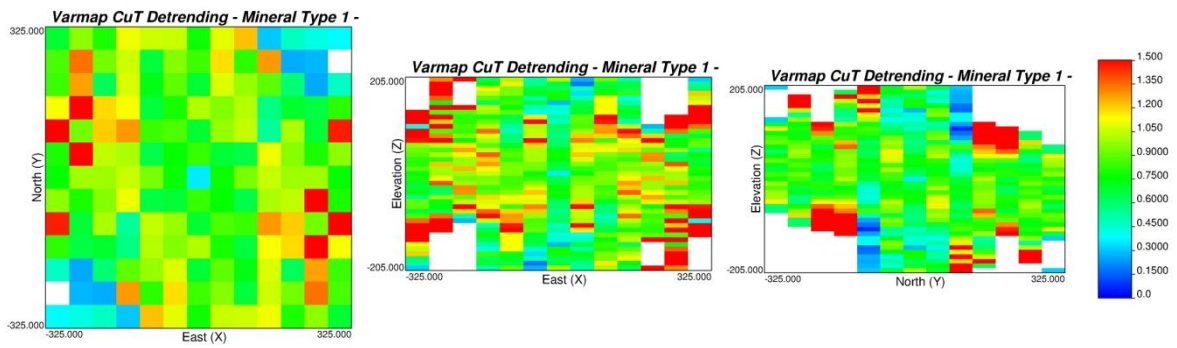


Figura 24: Mapa variográfico – UE1 (Izq: XY - Cen: XZ - Der: YZ) – Residuos.

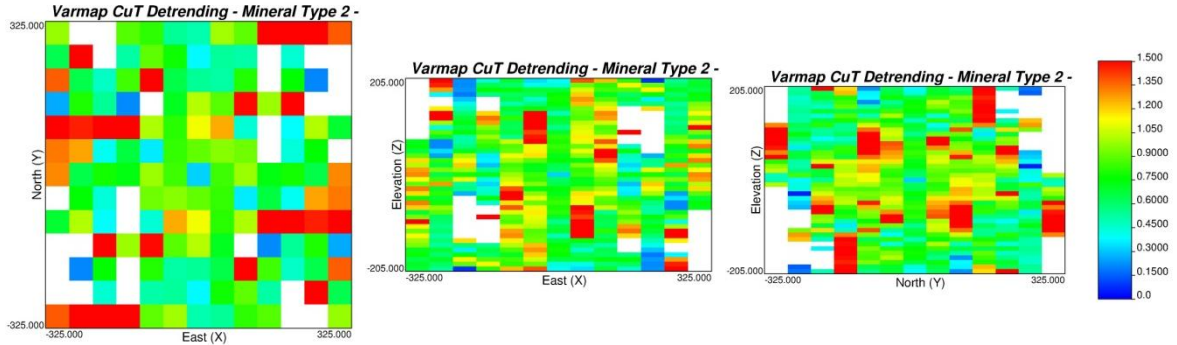


Figura 25: Mapa variográfico – UE2 (Izq: XY - Cen: XZ - Der: YZ) – Residuos.

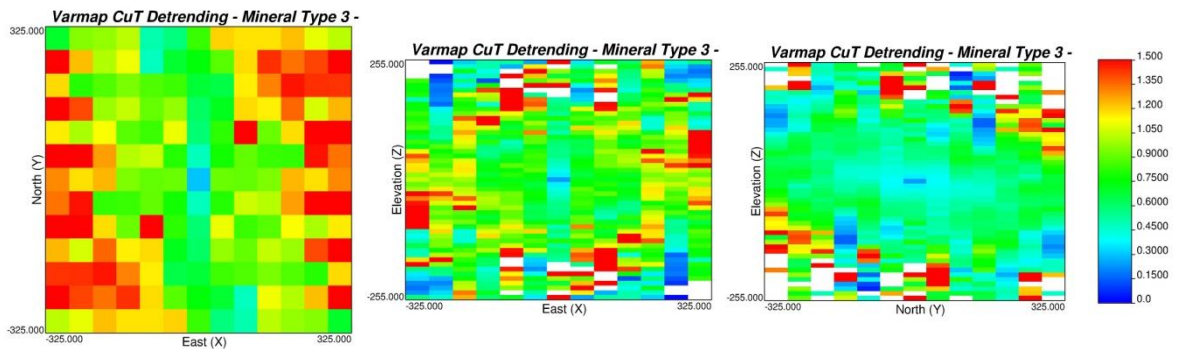
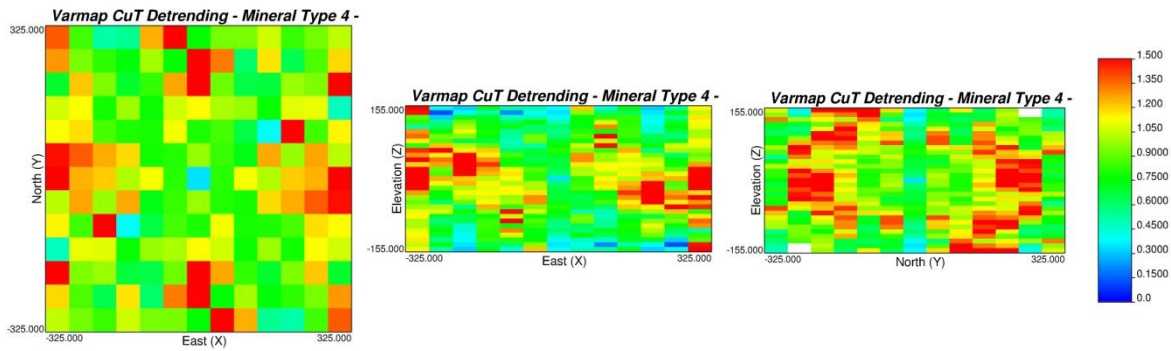
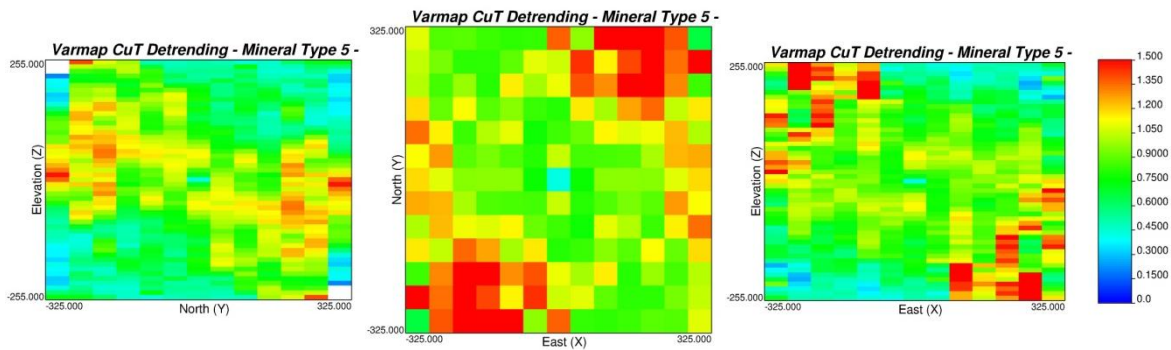


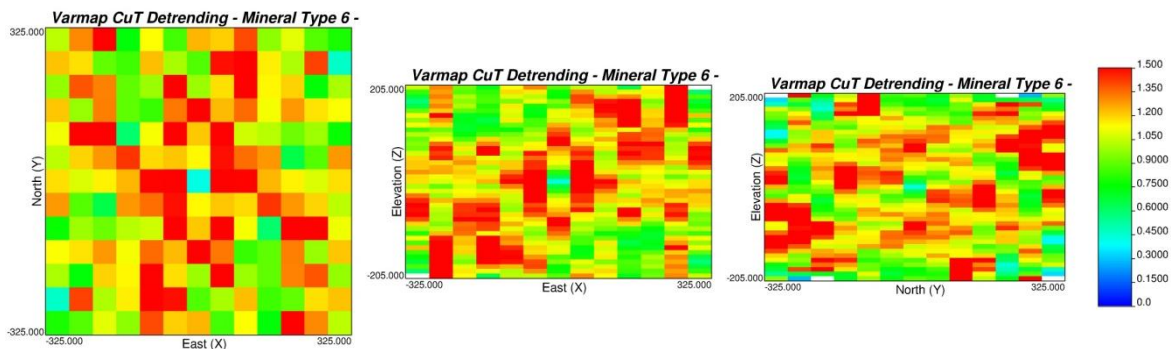
Figura 26: Mapa variográfico – UE3 (Izq: XY - Cen: XZ - Der: YZ) – Residuos.



**Figura 27: Mapa variográfico – UE4 (Izq: XY - Cen: XZ - Der: YZ) – Residuos.**



**Figura 28: Mapa variográfico – UE5 (Izq: XY - Cen: XZ - Der: YZ) – Residuos.**



**Figura 29: Mapa variográfico – UE6 (Izq: XY - Cen: XZ - Der: YZ) – Residuos.**

Los mapas variográficos para la UE1 y UE5 sugieren anisotropía geométrica entre el plano horizontal rotado en un dip de  $20^\circ$  y su vertical. En el caso de la UE2, UE4 y UE6 los mapas variográficos no sugieren anisotropía a simple vista. Para la UE3 los mapas sugieren una anisotropía geométrica en planta con una elipse rotada en  $340^\circ$ .

## ANEXO J: Caso de estudio II – Variogramas experimentales

Una vez determinados los mapas variográficos en XY, XZ e YZ para cada UE y definido un primer indicio de los tipos de anisotropía posibles de enfrentar en cada caso, se calculan los variogramas experimentales en cada UE considerando en cada caso 3 tipos distintos: variograma omnidireccional que considera una única dirección (a la izquierda en cada gráfico), variograma omnihorizontal que considera una dirección horizontal y una vertical (al centro en cada gráfico) y variograma direccional que considera 4 direcciones en planta (a la derecha en cada gráfico).

Dirección	Azimut [°]	Tolerancia Azimut [°]	Banda Horizontal [m]	Dip [°]	Tolerancia Dip [°]	Banda Vertical [m]	Color
Omnidireccional	0	90	∞	0	90	∞	Rojo
Omnihorizontal - XY	0	90	∞	0	22.5	10	Rojo
Omnihorizontal - Z	0	22.5	10	90	22.5	10	Negro
Direccional - N0E	0	22.5	25	0	22.5	10	Rojo
Direccional - N45E	45	22.5	25	0	22.5	10	Negro
Direccional - N90E	90	22.5	25	0	22.5	10	Verde
Direccional - N135E	135	22.5	25	0	22.5	10	Azul

**Tabla 45: Parámetros variogramas experimentales.**

Los parámetros para el cálculo de los variogramas experimentales se presentan en la Tabla 45. Para los comósitos originales, en la UE1, tanto el variograma omnihorizontal como el direccional están rotados con un dip de 20°, y en el caso de la UE5, el variograma direccional está rotado con un azimut de 340°. En el caso de los residuos, en la UE1 y UE5, tanto el variograma omnihorizontal como el direccional están rotados con un dip de 20°, y en el caso de la UE3, el variograma direccional está rotado con un azimut de 340°.

El número de pasos utilizados en cada variograma así como la distancia del paso se presentan en la Tabla 46 y Tabla 47, donde se muestra además el rango de estudio del correspondiente variograma el cual es equivalente al número de pasos por el largo del paso. Se utiliza una tolerancia del paso equivalente a la mitad de éste en cada caso.

UE	Omnidireccional			Omnihorizontal - XY			Omnihorizontal - Z			Direccional		
	N° Pasos	Paso [m]	Rango [m]	N° Pasos	Paso [m]	Rango [m]	N° Pasos	Paso [m]	Rango [m]	N° Pasos	Paso [m]	Rango [m]
1	16	25	400	16	25	400	6	25	150	6	30	180
2	16	25	400	16	25	400	30	5	150	5	60	300
3	16	25	400	16	25	400	30	5	150	8	40	320
4	16	25	400	16	25	400	30	5	150	8	40	320
5	16	25	400	16	25	400	30	5	150	8	40	320
6	16	25	400	16	25	400	30	5	150	8	40	320

Tabla 46: Número y distancia del paso – Compósitos originales.

UE	Omnidireccional			Omnihorizontal - XY			Omnihorizontal - Z			Direccional		
	N° Pasos	Paso [m]	Rango [m]	N° Pasos	Paso [m]	Rango [m]	N° Pasos	Paso [m]	Rango [m]	N° Pasos	Paso [m]	Rango [m]
1	16	25	400	16	25	400	10	15	150	6	30	180
2	16	25	400	16	25	400	8	20	160	5	60	300
3	16	25	400	16	25	400	8	20	160	8	40	320
4	16	25	400	16	25	400	8	20	160	8	40	320
5	16	25	400	16	25	400	10	15	150	8	40	320
6	16	25	400	16	25	400	8	20	160	8	40	320

Tabla 47: Número y distancia del paso - Residuos.

### J.1 Compósitos originales

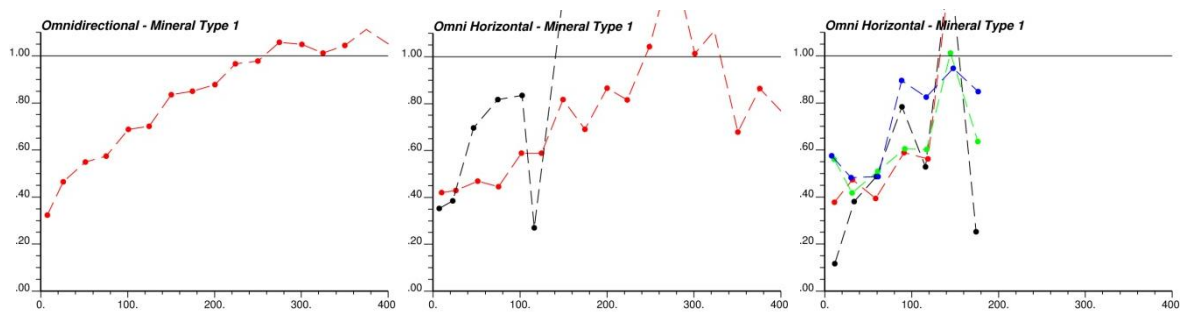


Gráfico 150: Variogramas experimentales – UE1 – Compósitos.

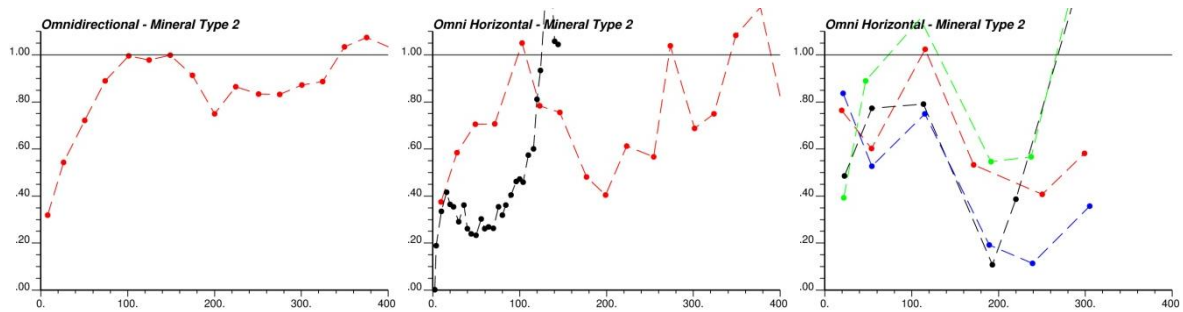
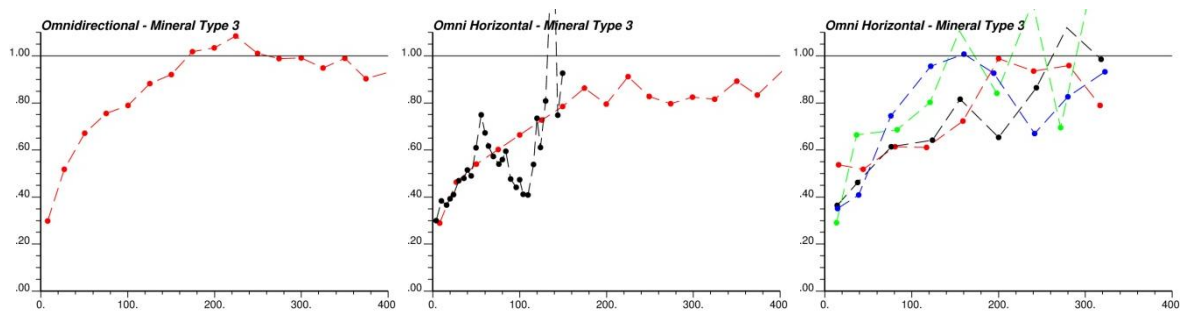
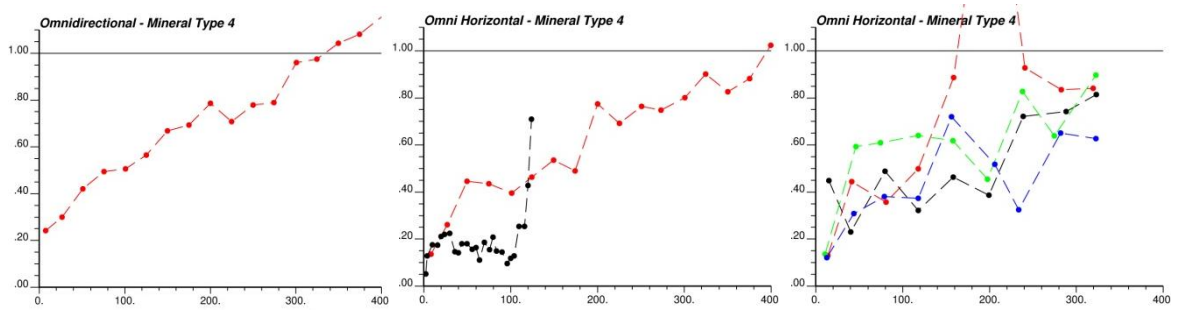


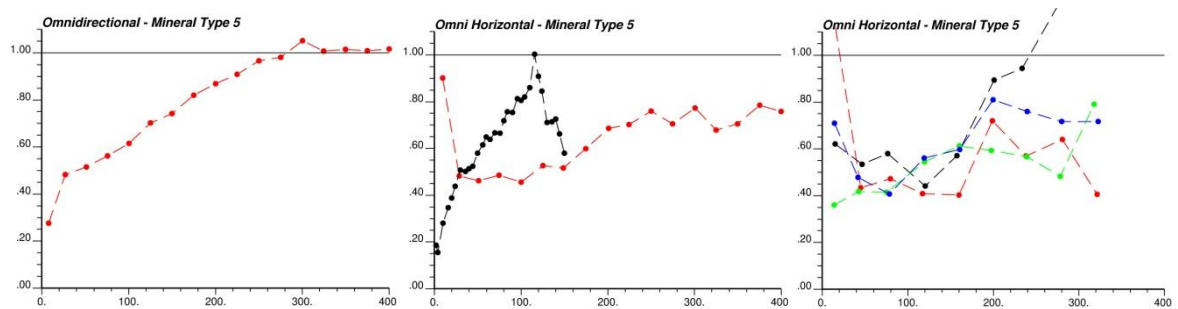
Gráfico 151: Variogramas experimentales – UE2 – Compósitos.



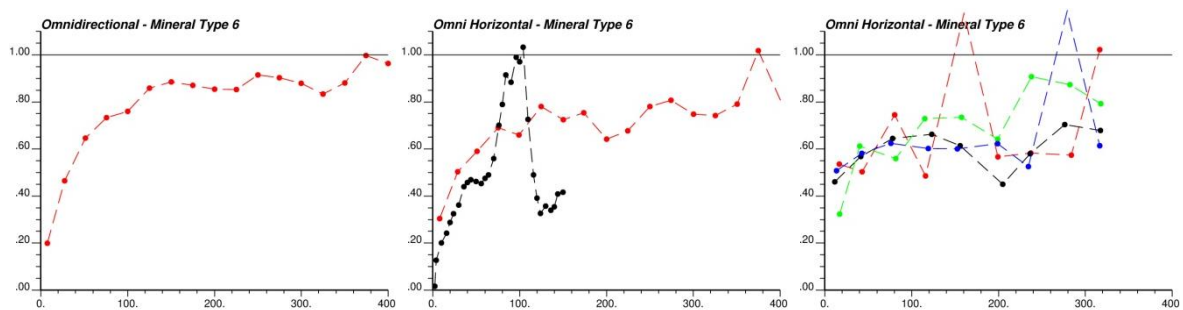
**Gráfico 152: Variogramas experimentales – UE3 – Compósitos.**



**Gráfico 153: Variogramas experimentales – UE4 – Compósitos.**



**Gráfico 154: Variogramas experimentales – UE5 – Compósitos.**



**Gráfico 155: Variogramas experimentales – UE6 – Compósitos.**

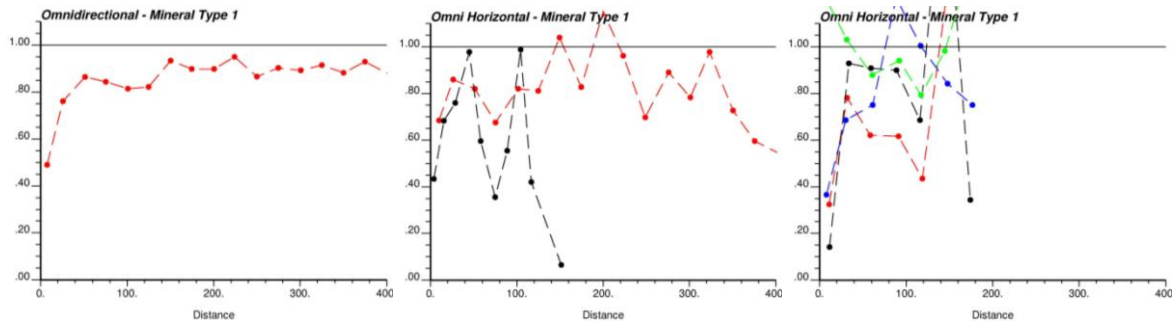
Se tiene el siguiente resultado para cada una de las UE en base a los variogramas experimentales.

- **UE1:** El desacople del variograma en su parte horizontal y vertical muestra una anisotropía geométrica donde el vertical tiene menor alcance que el horizontal. Un estudio direccional no sugiere anisotropía en planta por lo tanto se opta por el modelo omnihorizontal con dip de 20°.
- **UE2:** Si bien el variograma omnihorizontal presenta ciertas diferencias entre su parte vertical y horizontal, al realizar un ajuste sobre los primeros puntos y no modelar la caída de la parte vertical, se tiene un comportamiento similar entre ambas partes. De esta manera se opta por el modelo omnidireccional en este caso.
- **UE3:** El paso del variograma omnidireccional a omnihorizontal muestra anisotropía geométrica donde la componente vertical tiene menor alcance que la horizontal. Más aún, un análisis direccional muestra una anisotropía geométrica en planta donde la dirección N45E corresponde al alcance horizontal máximo y la dirección N135E al alcance horizontal mínimo. Se opta por tanto por el modelo direccional.
- **UE4:** En esta UE se tiene el mismo caso que para la UE2, por lo tanto se opta por el modelo omnidireccional. Es importante notar que es en este variograma donde se aprecia de mejor manera un comportamiento que denota presencia de deriva, es decir, un gráfico que aumenta linealmente por sobre la varianza de los datos.
- **UE5:** El variograma omnihorizontal muestra una clara anisotropía geométrica entre la componente horizontal y vertical, donde la primera tiene mayor alcance que la segunda. Dado que el análisis direccional no muestra anisotropías (al contrario de lo que se sugería inicialmente), se opta por el modelo omnihorizontal.
- **UE6:** En este caso, tanto en el variograma omnihorizontal como en el direccional no se logran apreciar anisotropías dado que todas las direcciones tienen el mismo comportamiento, por lo cual se opta por el modelo omnidireccional.

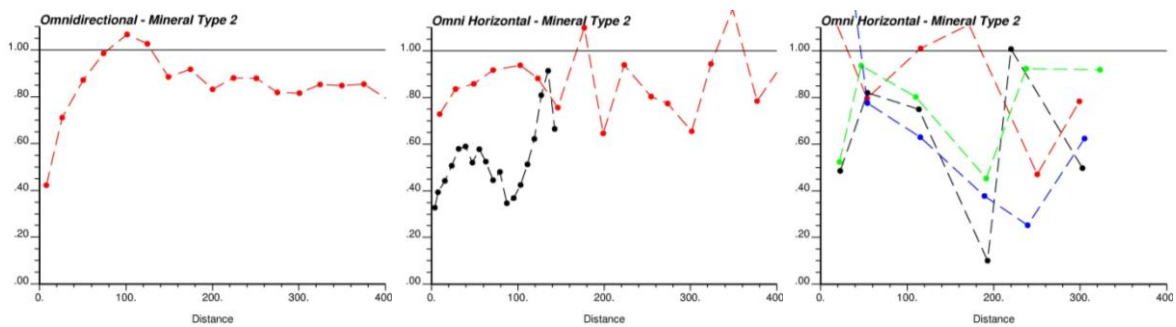
Una vez definido el tipo de variograma para cada UE se procede con la siguiente etapa del estudio consistente en determinar los modelos variográficos correspondientes.



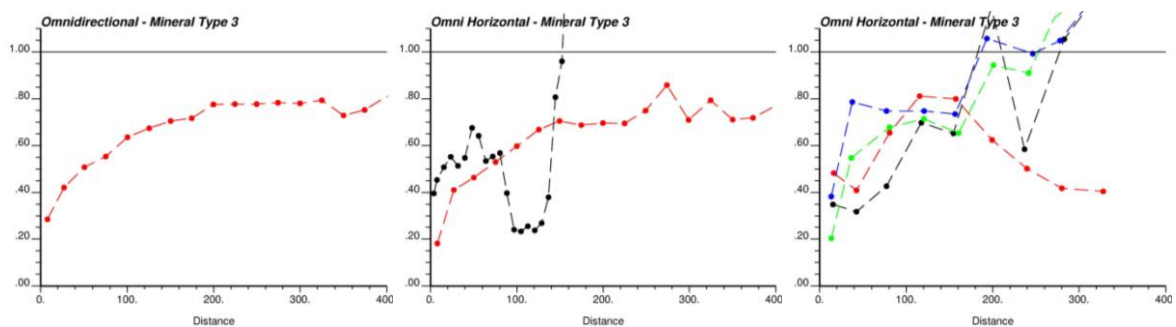
## J.1 Residuos



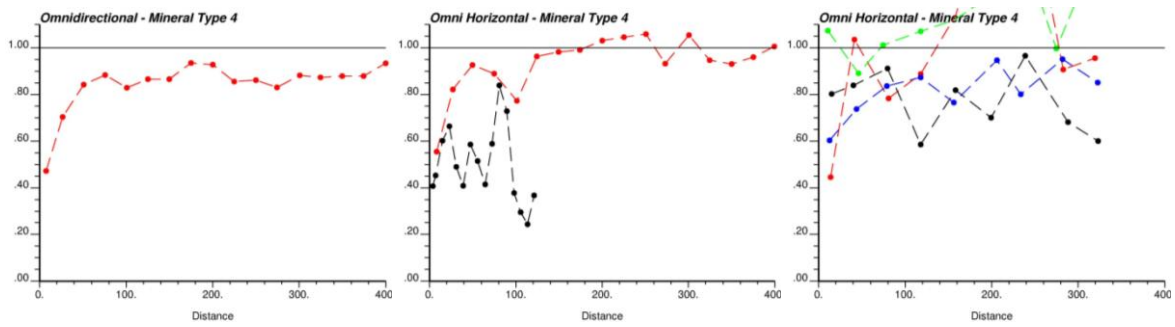
**Gráfico 156: Variogramas experimentales – UE1 – Residuos.**



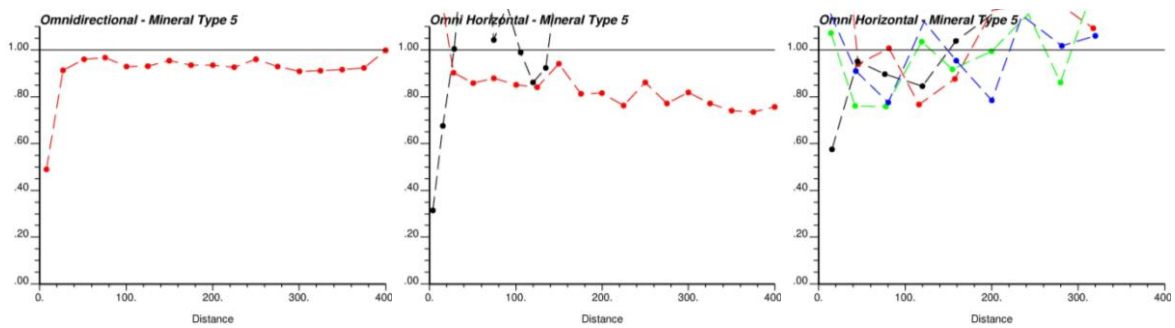
**Gráfico 157: Variogramas experimentales – UE2 – Residuos.**



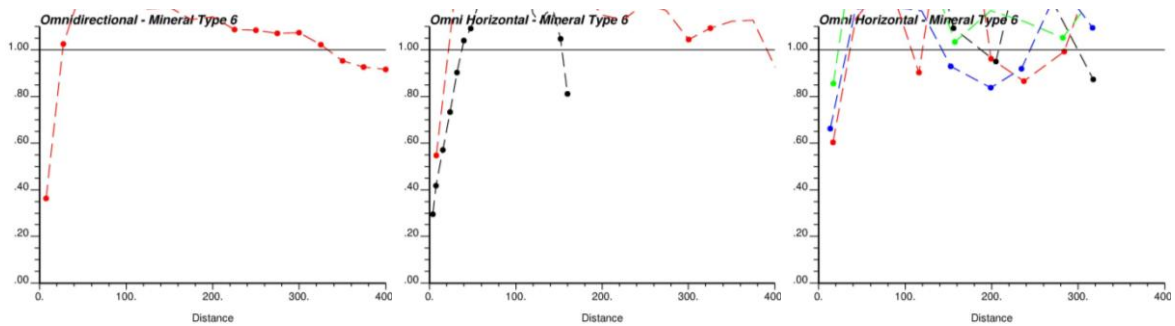
**Gráfico 158: Variogramas experimentales – UE3 – Residuos.**



**Gráfico 159: Variogramas experimentales – UE4 – Residuos.**



**Gráfico 160: Variogramas experimentales – UE5 – Residuos.**



**Gráfico 161: Variogramas experimentales – UE6 – Residuos.**

Se tiene el siguiente resultado para cada una de las UE en base a los variogramas experimentales.

- **UE1:** Si bien para el caso de los compósitos originales se opta por un modelo omnihorizontal, el variograma omnihorizontal de los residuos no muestra dicha anisotropía geométrica. Más aún, el variograma direccional no sugiere diferencias entre direcciones por lo cual se opta por un modelo omnidireccional.
- **UE2:** Para este caso, la situación es a la inversa que en la UE1. En los compósitos originales se opta por un modelo omnidireccional, sin embargo, en este caso, se aprecia una anisotropía geométrica entre el plano horizontal y la vertical por lo que se opta por un modelo omnihorizontal.
- **UE3:** Si bien en el caso de los compósitos originales se opta por un modelo direccional, en este caso se utiliza un modelo omnihorizontal dado que el variograma experimental direccional no evidencia una clara anisotropía en el plano horizontal, mas si se aprecia anisotropía geométrica entre el plano horizontal y la vertical.
- **UE4:** Al igual que en el caso de los compósitos originales, aun cuando hay diferencias entre la parte horizontal y vertical del variograma, si se realiza el ajuste sobre los primeros puntos (los cuales son los más importantes en el modelo) y no se considera la caída de la parte vertical en el variograma, se tiene un comportamiento similar entre el plano horizontal y la vertical por lo cual se opta por un modelo omnidireccional. Es interesante que para este caso si se alcanza una meseta, lo que muestra que el método de residuos en efecto filtra el efecto de la deriva de los compósitos originales en el variograma.
- **UE5:** Si bien se opta por un modelo omnihorizontal en los compósitos originales, en este caso no se aprecia una anisotropía geométrica entre el plano horizontal y la vertical por lo cual se opta por un modelo omnidireccional.
- **UE6:** Al igual que en el caso de los compósitos originales, tanto en el variograma omnihorizontal como en el direccional no se logran apreciar anisotropías dado que todas las direcciones tienen el mismo comportamiento, por lo cual se opta por el modelo omnidireccional.

Una vez definido el tipo de variograma para cada UE se procede con la siguiente etapa del estudio consistente en determinar los modelos variográficos correspondientes.

## **ANEXO K: Caso de estudio II – Histograma por UE**

Se presentan los histogramas acumulados con una escala de 0% a 4% de ley de cobre total. Cada gráfico está presentado por UE donde se muestran los resultados por SK y RDR más sus sensibilizaciones en el siguiente orden: SK, Cmean-0.1, Cmean-0.5, Cmean-1, Cmean-2, Nrej-10, Nrej-100, Nrej-1000, Nrej-10000, Tol-1, Tol-2, Tol-3, Rtype-0, Rtype-1, Referror-0.01, Referror-0.1, Referror-1, Referror-10, lo cual suma un total de 18 histogramas en cada gráfico por UE. Se debe hacer notar que esta disposición permite visualizar el comportamiento de los resultados a medida que se cambia un parámetro y que dada la representación de todos los escenarios de cada sensibilización existen histogramas idénticos dado que corresponden al valor base del parámetro respectivo. Tales histogramas idénticos son los siguientes: Cmean-1, Nrej-1000, Tol-2, Rtype-1 y Referror-0.1. Cada gráfico muestra los 100 histogramas acumulados de cada realización (líneas moradas) y el histograma acumulado de los datos originales desagrupados (línea negra). Los resultados correspondientes al Kriging de residuos se presentan en un mismo gráfico considerando las seis UE.

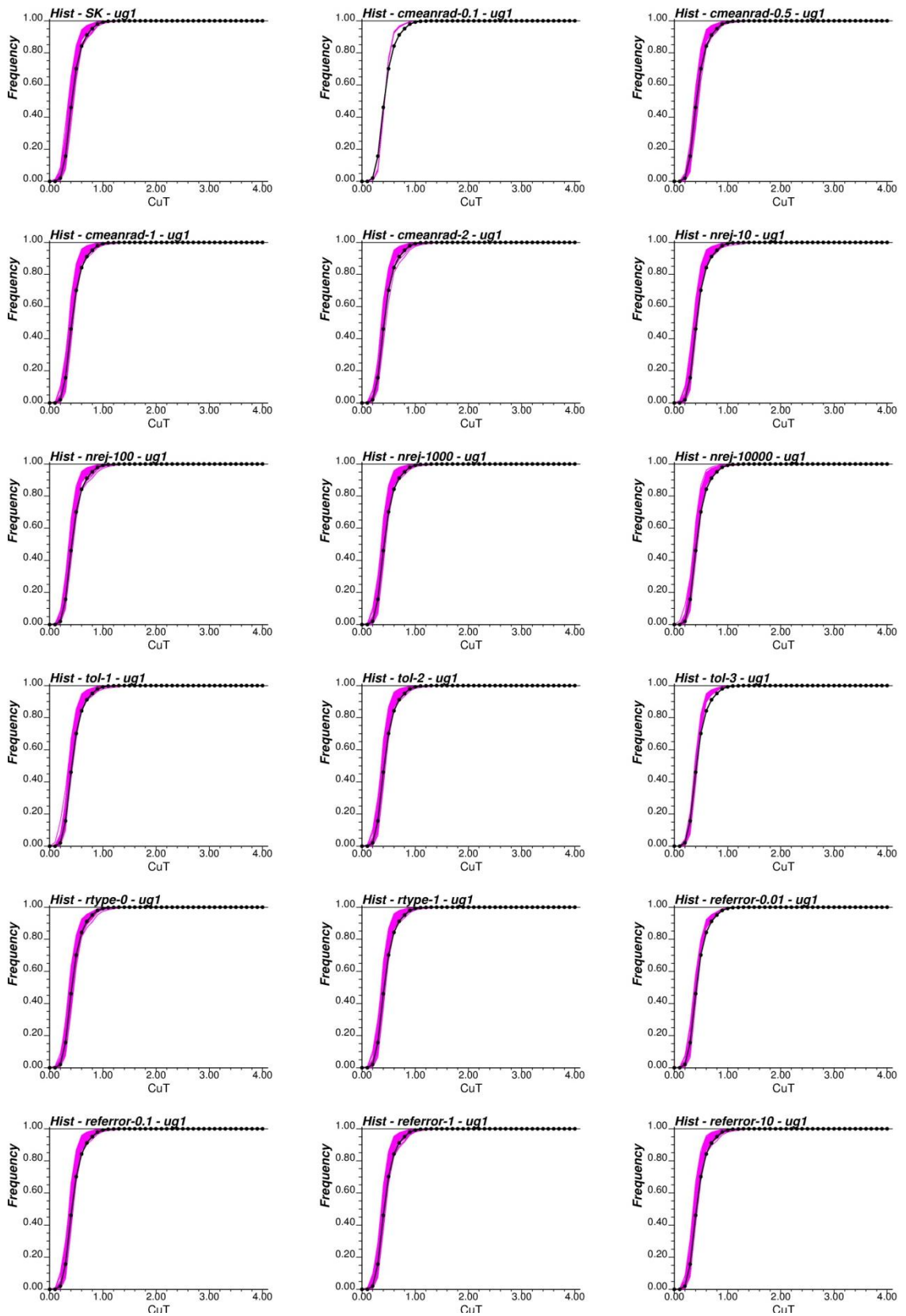
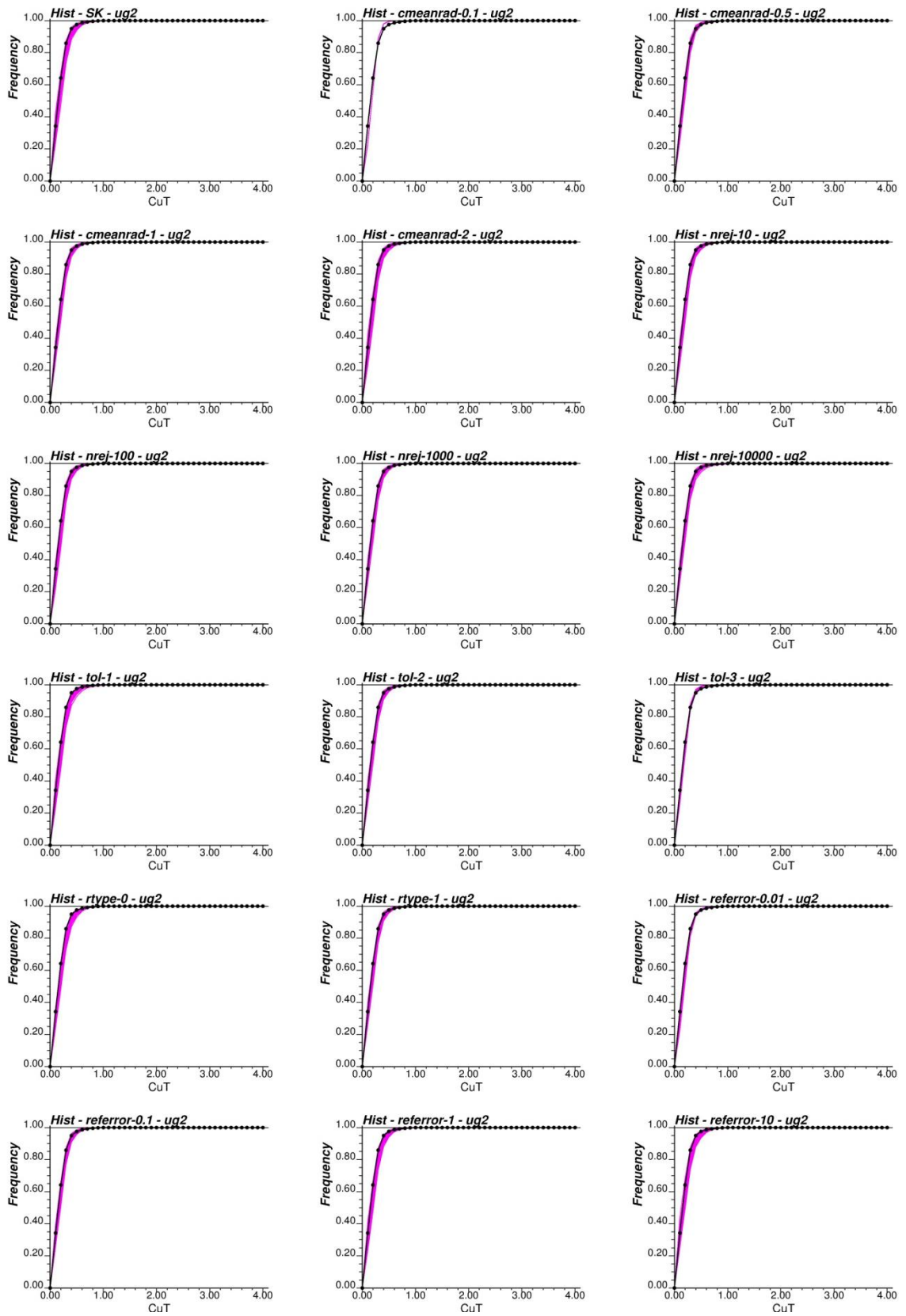
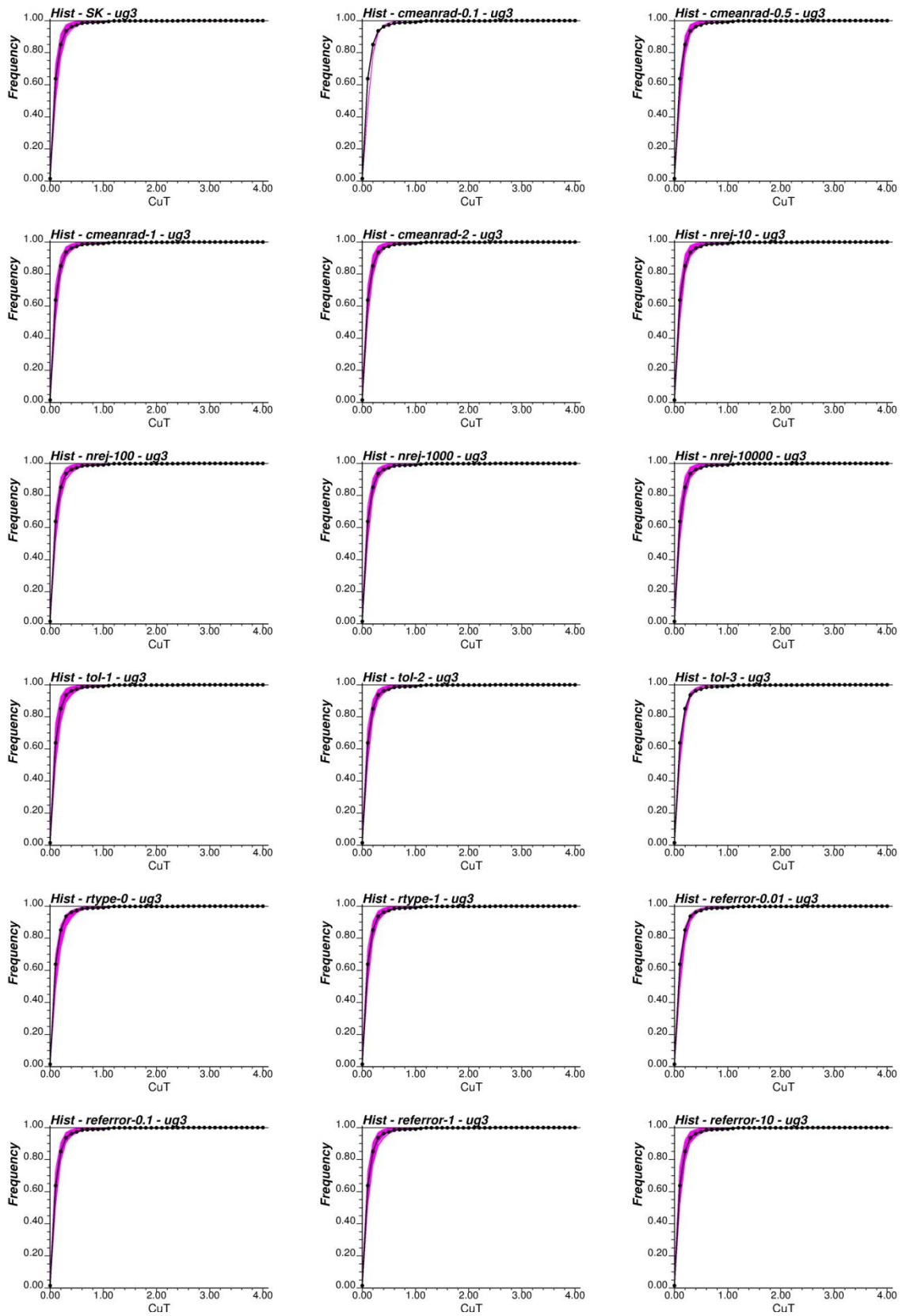


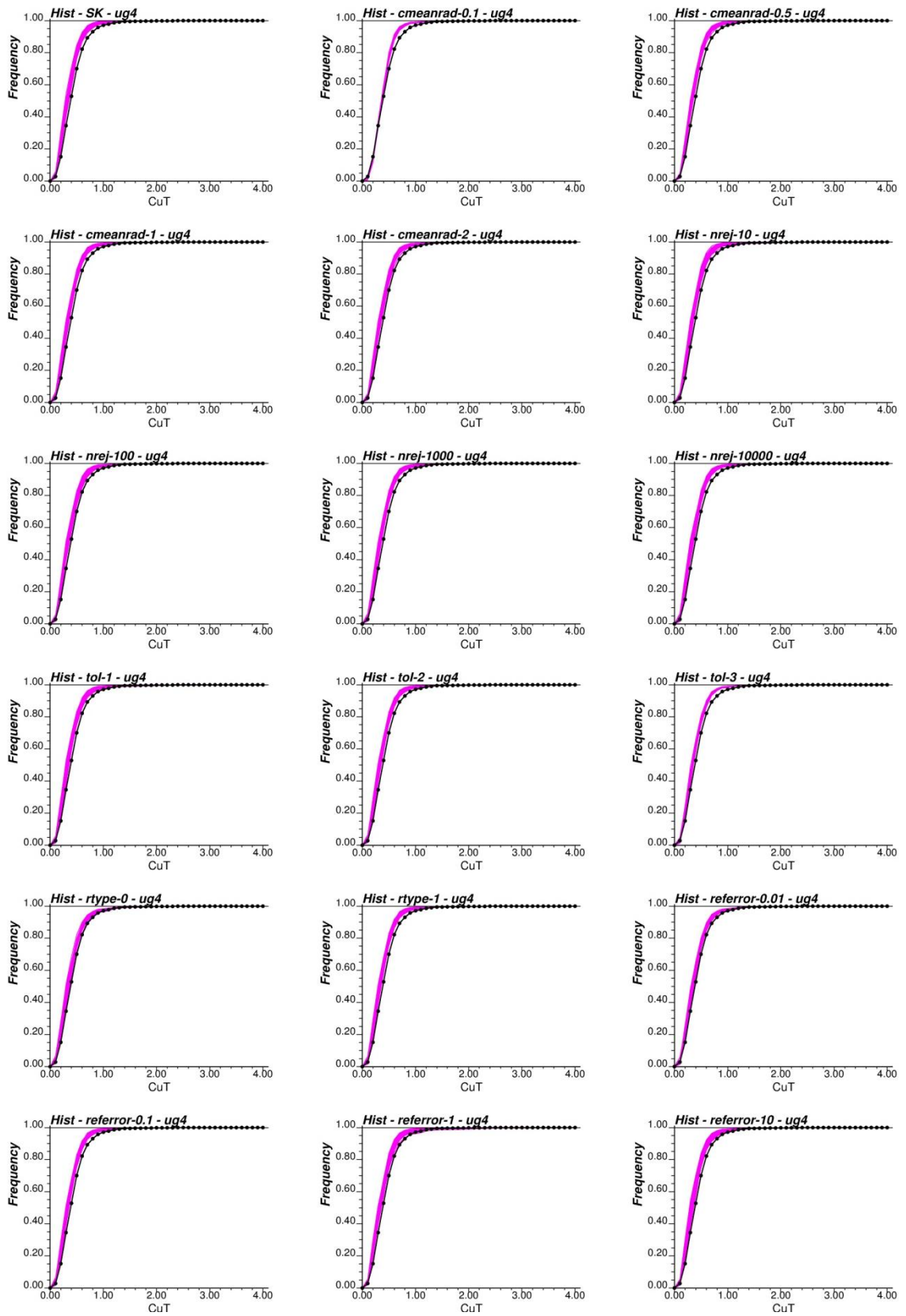
Gráfico 162: Histograma acumulado – UE 1.



**Gráfico 163: Histograma acumulado – UE 2.**



**Gráfico 164: Histograma acumulado – UE 3.**



**Gráfico 165: Histograma acumulado – UE 4.**



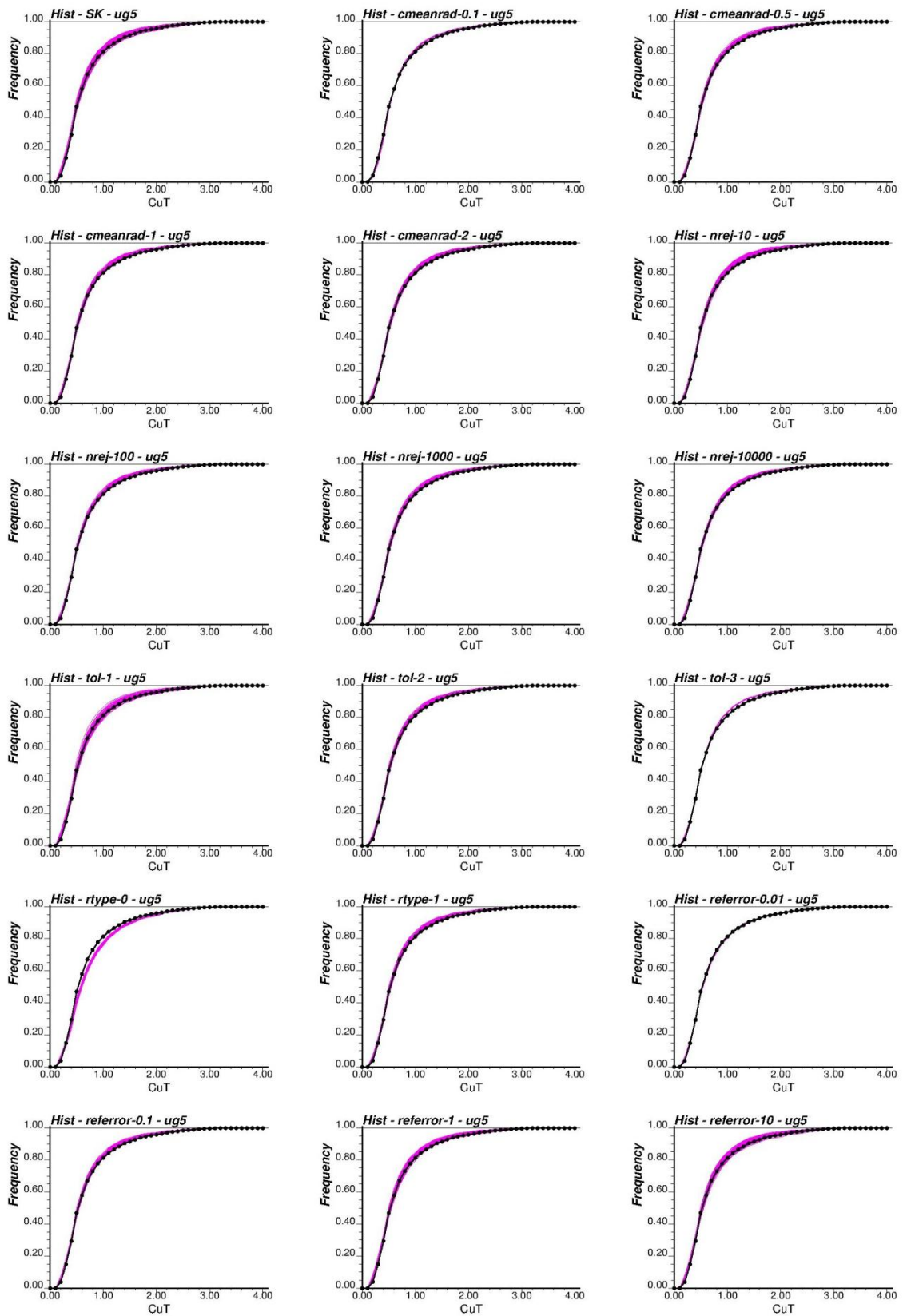


Gráfico 166: Histograma acumulado – UE 5.

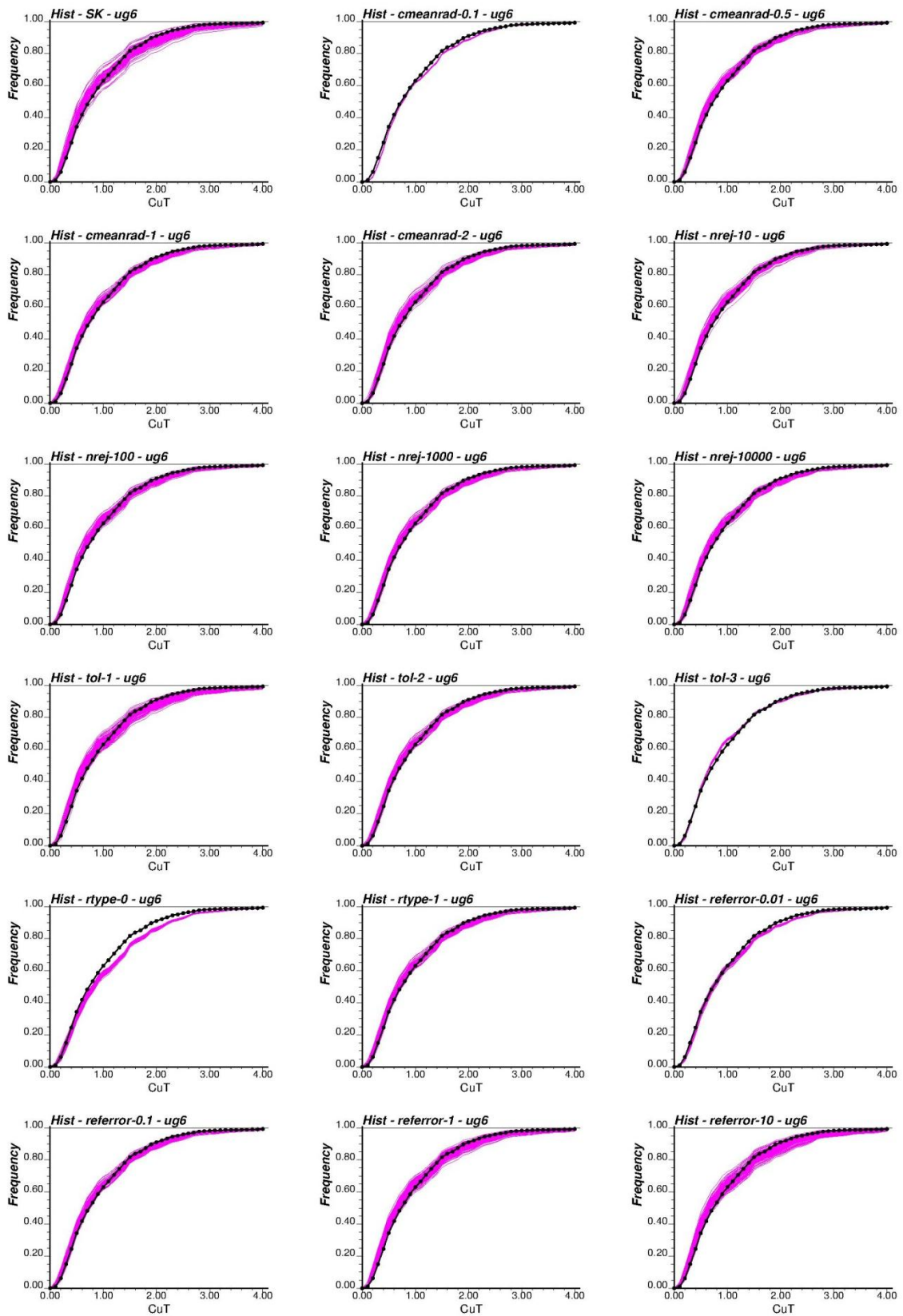
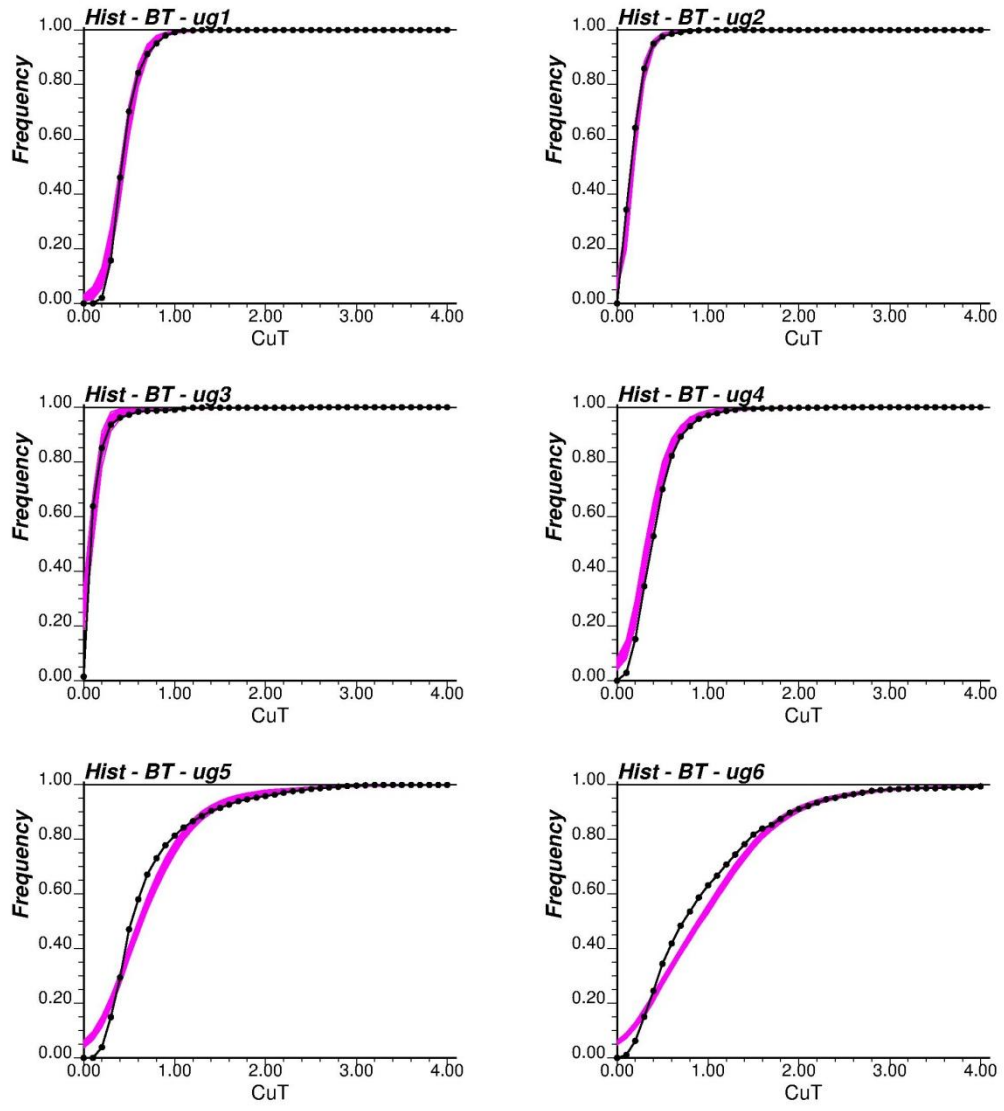


Gráfico 167: Histograma acumulado – UE 6.

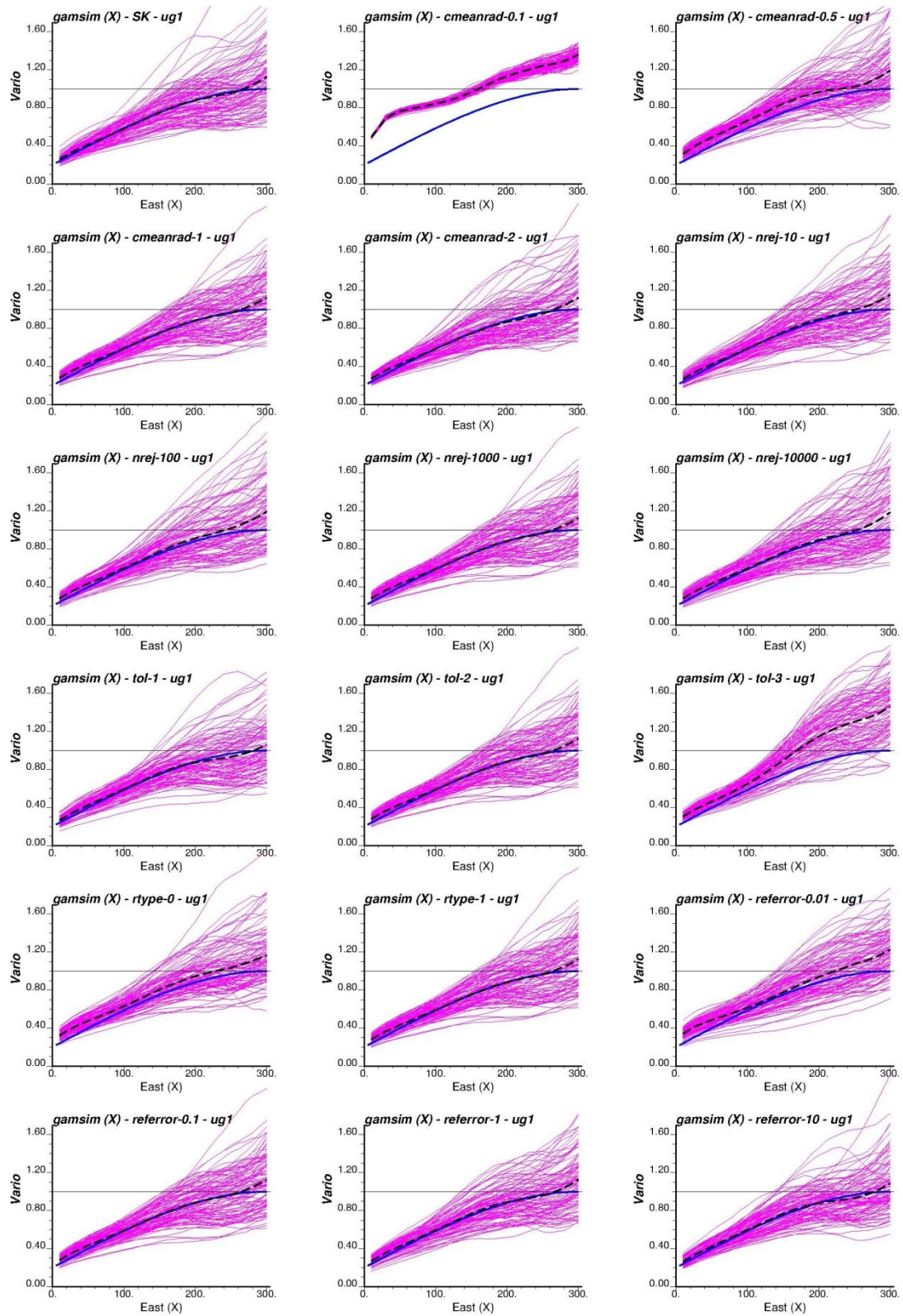


**Gráfico 168: Histograma acumulado por UE – Kriging de residuos.**

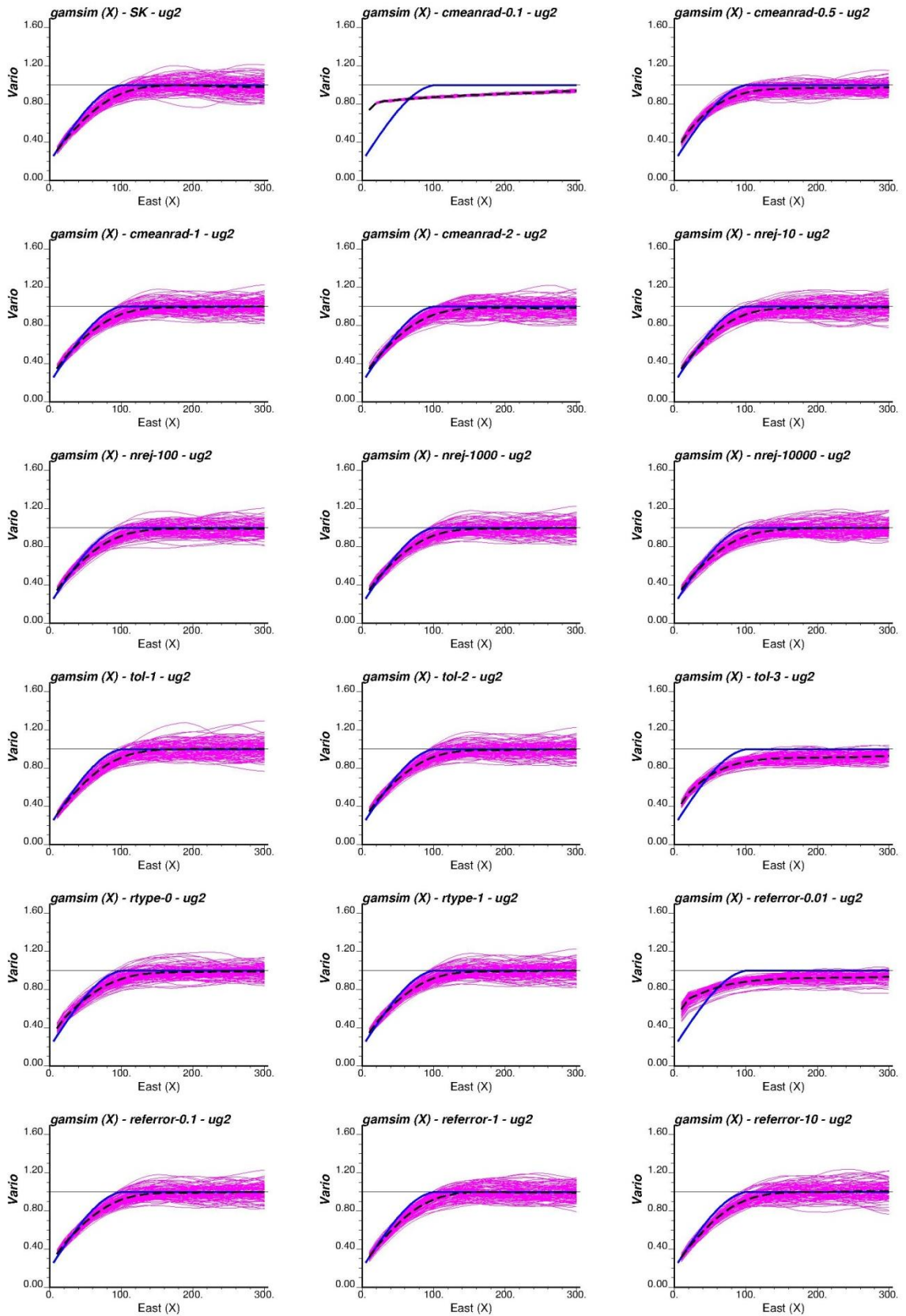
## **ANEXO L: Caso de estudio II – Variograma por UE**

Se presentan los variogramas experimentales de los resultados de las simulaciones en el mismo formato en que se presentan los histogramas, es decir, considerando 18 variogramas en cada gráfico por UE y ocupando el mismo orden para su presentación, junto con un gráfico que presenta los resultados del Kriging de residuos por UE. Cada variograma muestra, además de los 100 variogramas experimentales de cada simulación (líneas moradas), el variograma modelado correspondiente (línea azul) y el promedio de los 100 variogramas experimentales (línea negra segmentada). Para el cálculo de cada variograma experimental se utilizan 30 pasos de 10 metros, realizando dicho cálculo en las direcciones X, Y y Z.

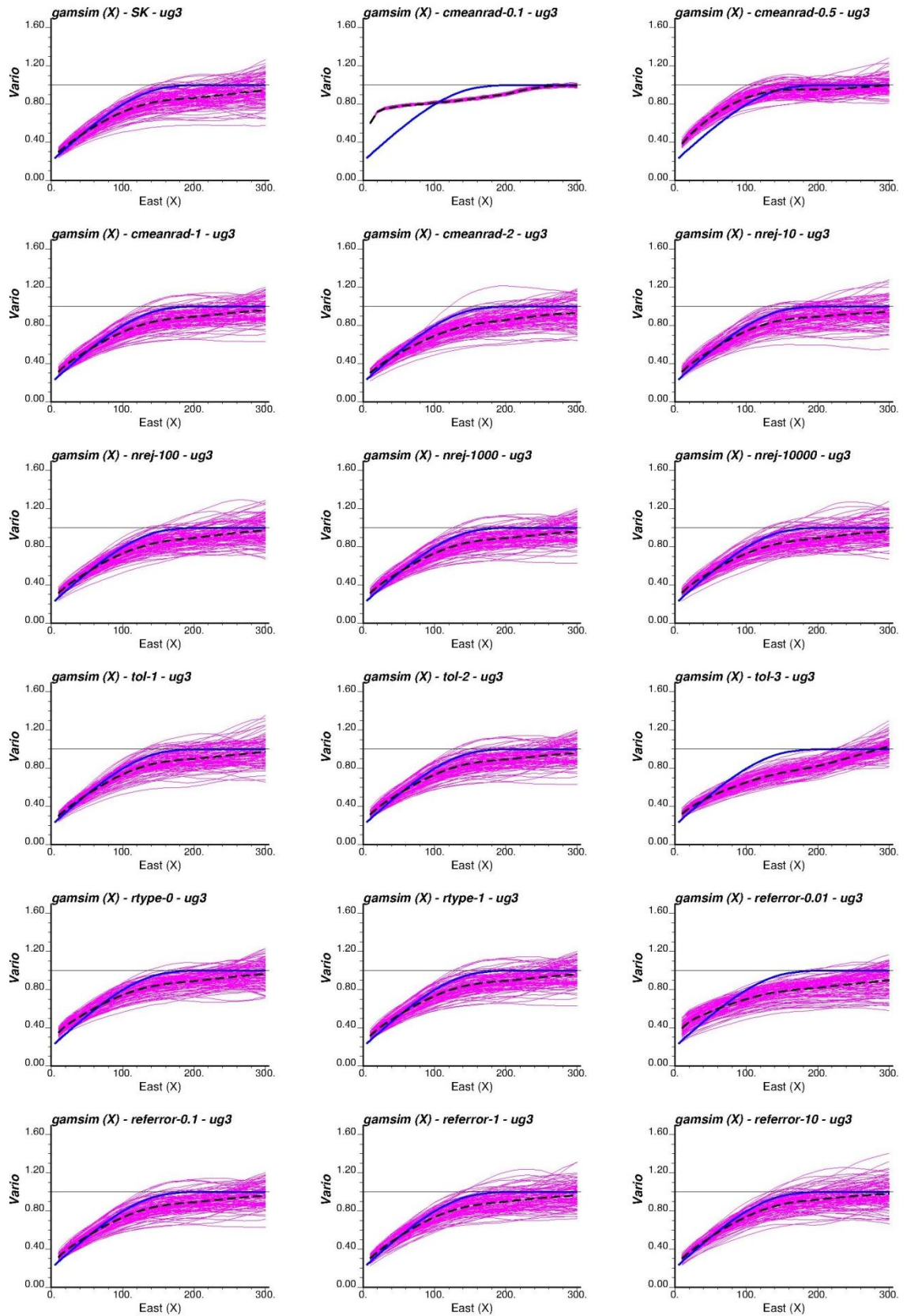
*Dirección X*



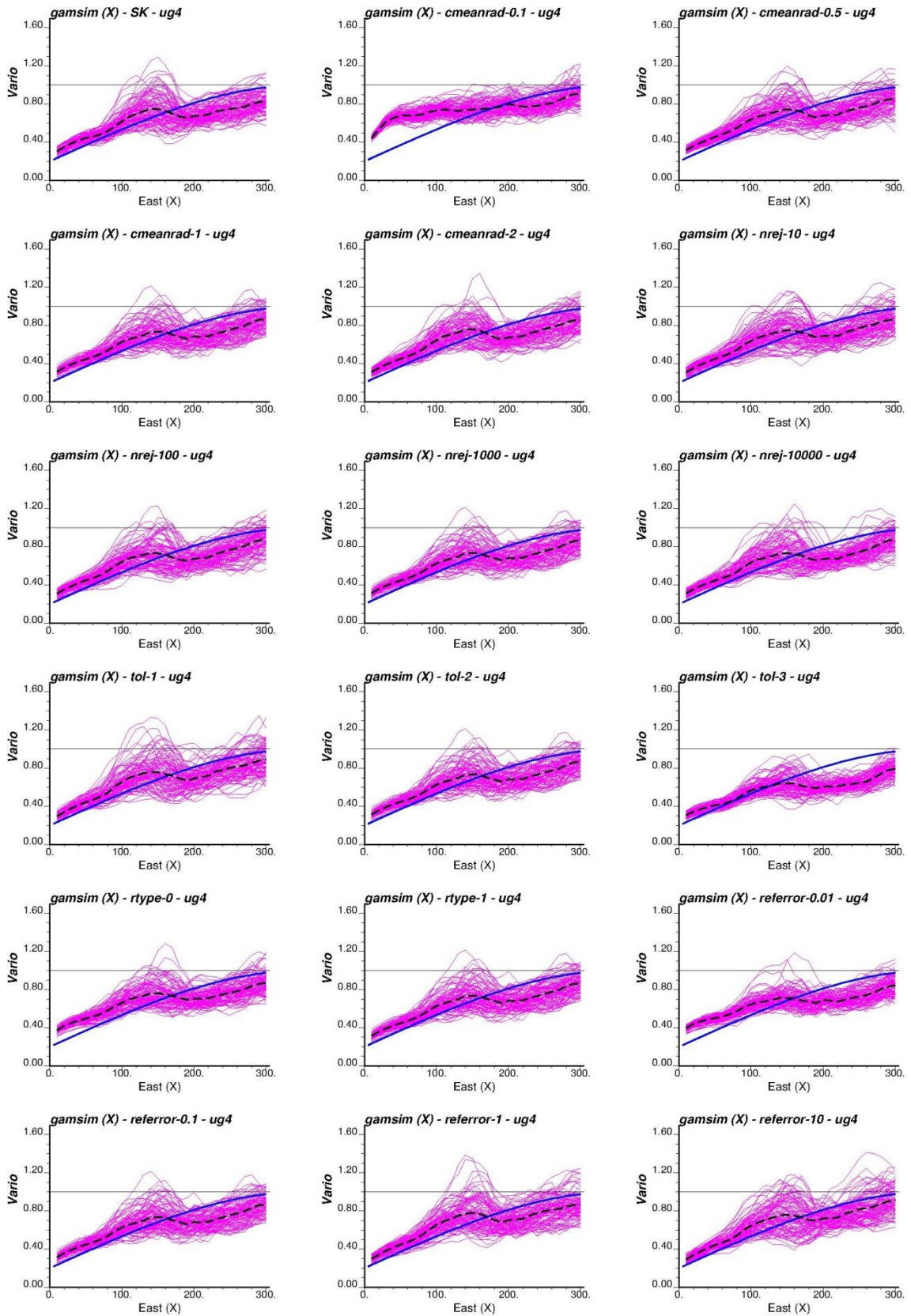
**Gráfico 169: Variograma experimental – Dirección X – UE 1.**



**Gráfico 170: Variograma experimental – Dirección X – UE 2.**

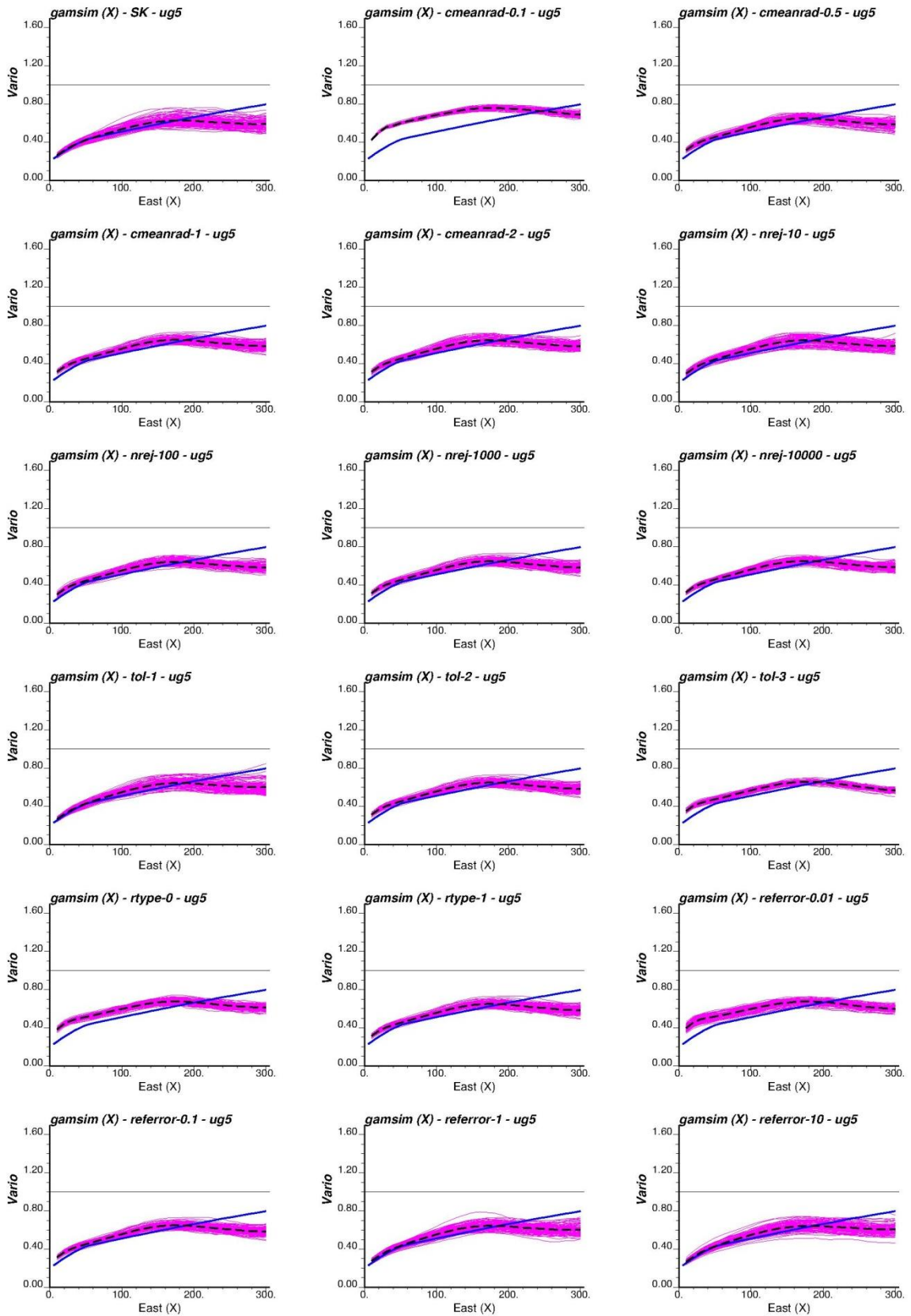


**Gráfico 171: Variograma experimental – Dirección X – UE 3.**

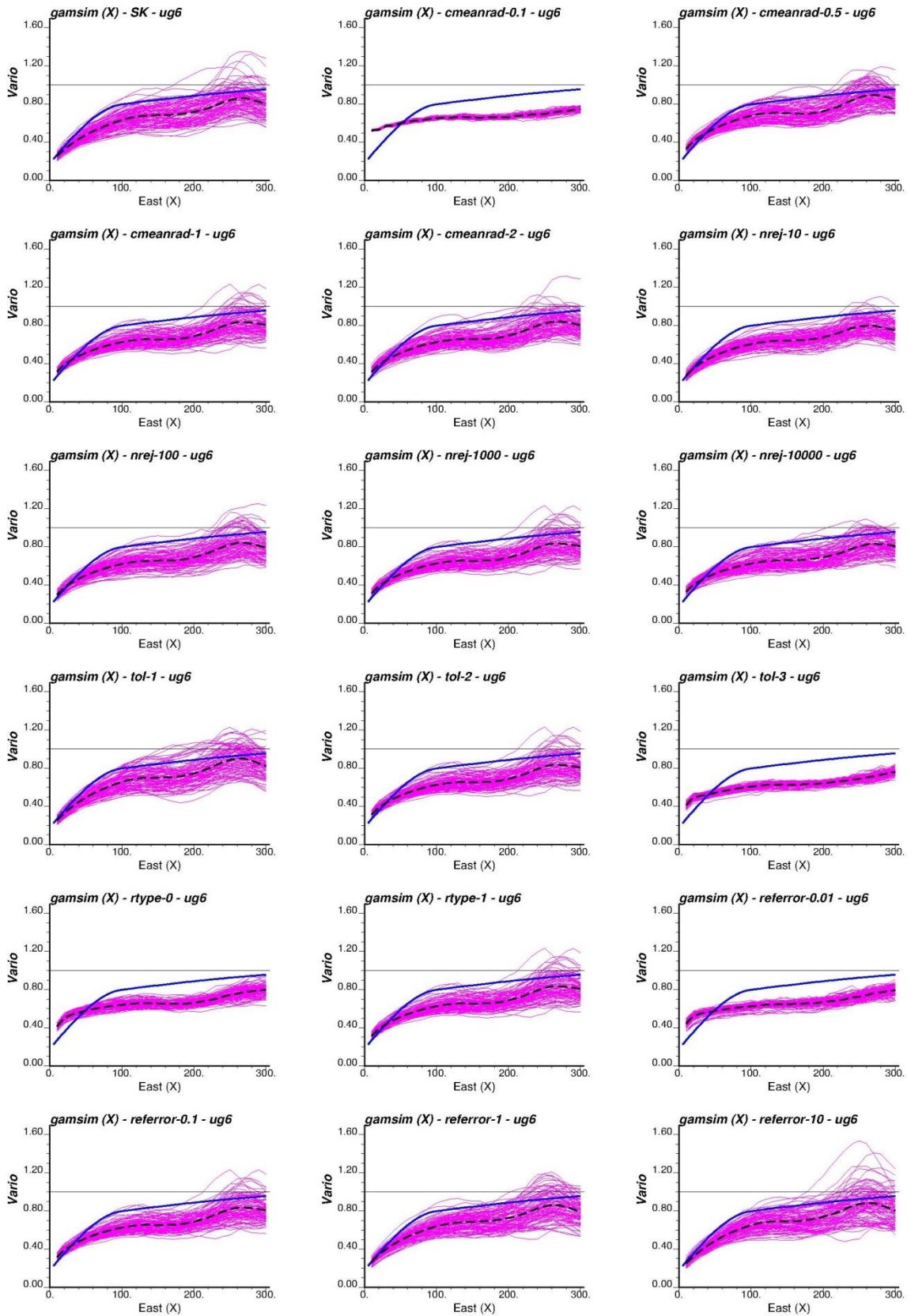


**Gráfico 172: Variograma experimental – Dirección X – UE 4.**

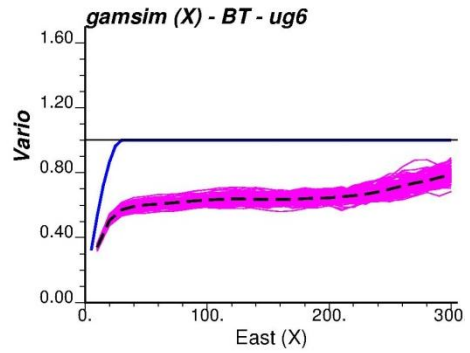
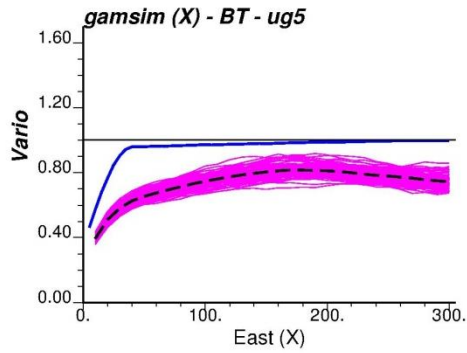
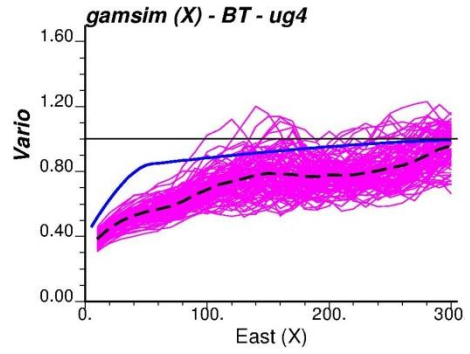
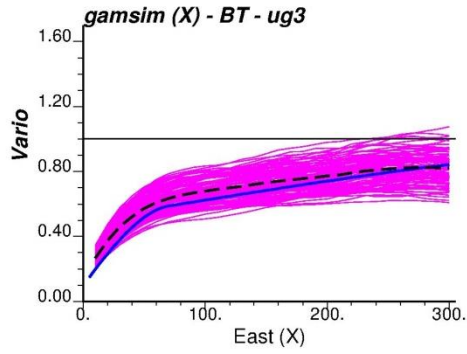
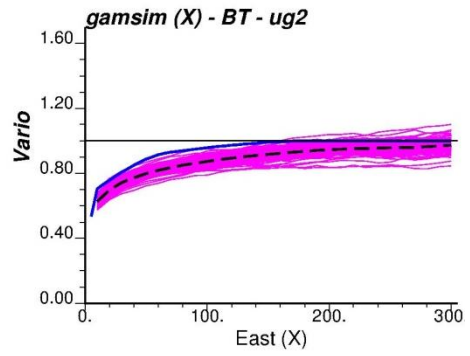
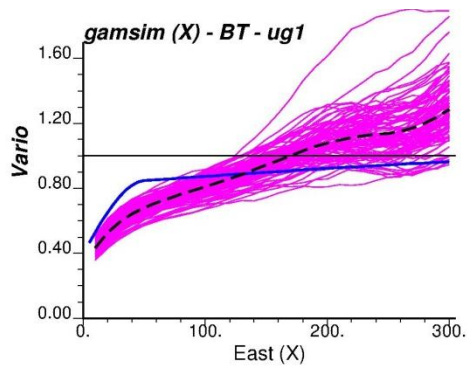




**Gráfico 173: Variograma experimental – Dirección X – UE 5.**

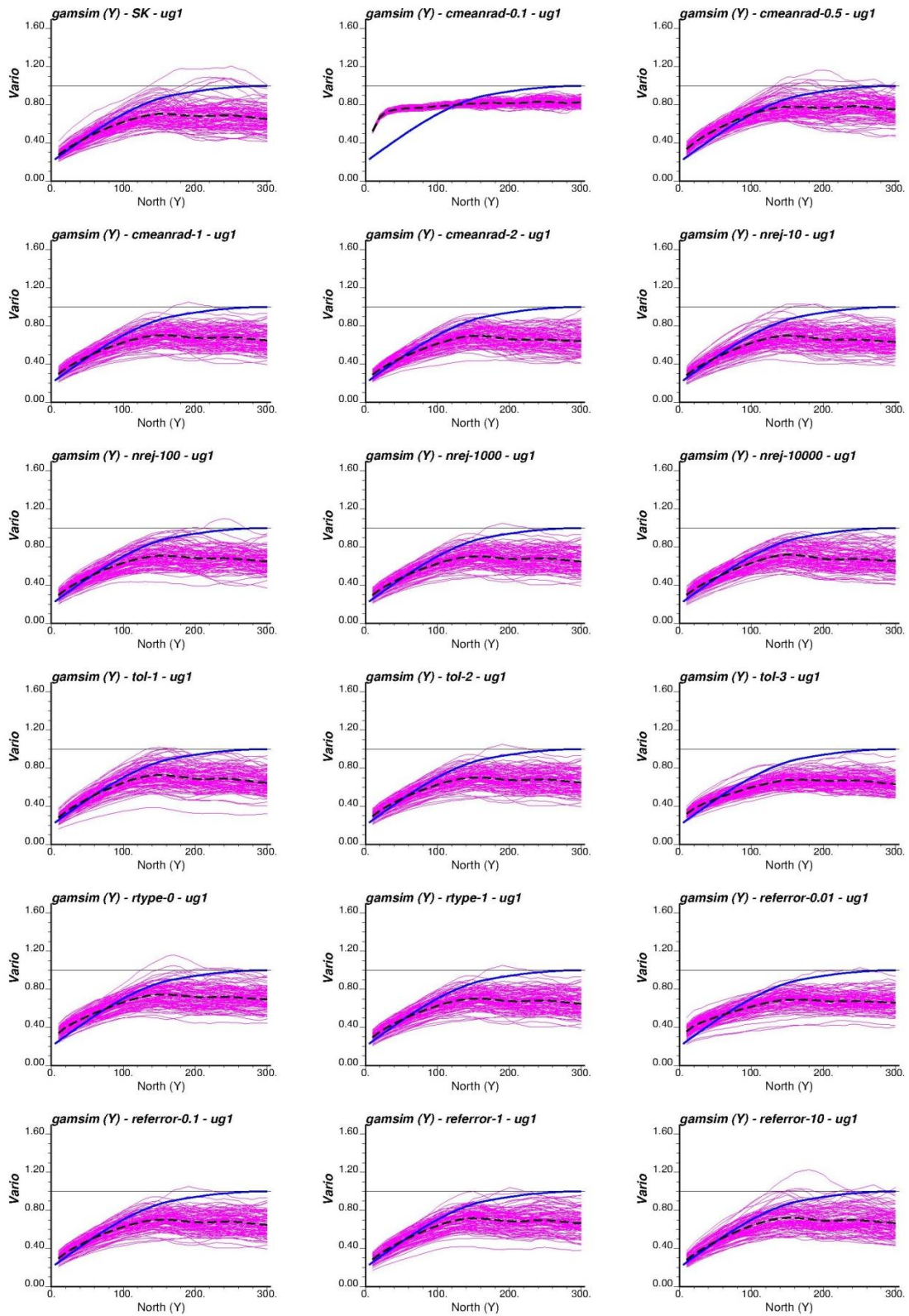


**Gráfico 174: Variograma experimental – Dirección X – UE 6.**

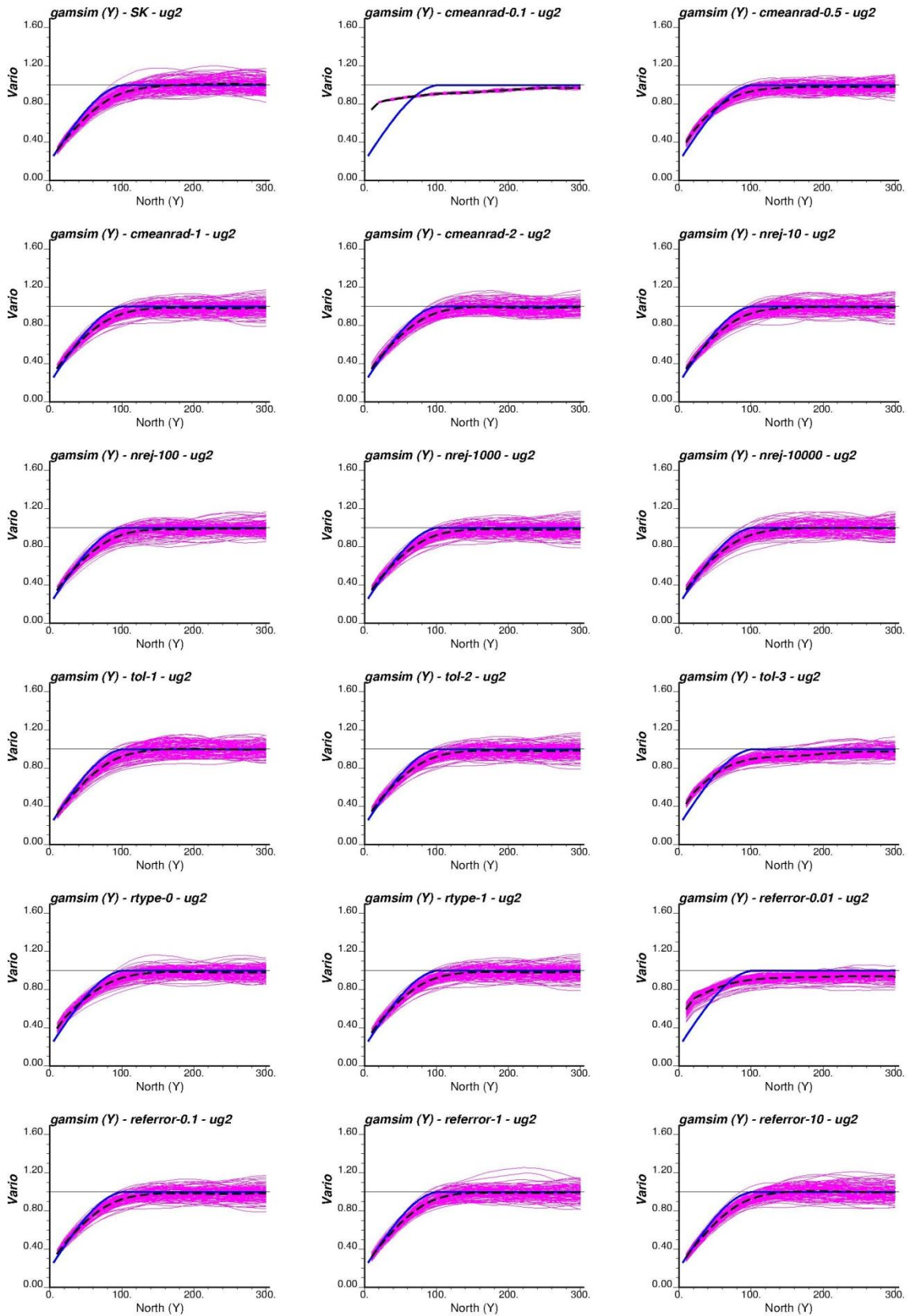


**Gráfico 175: Variograma experimental por UE – Dirección X – Kriging de residuos.**

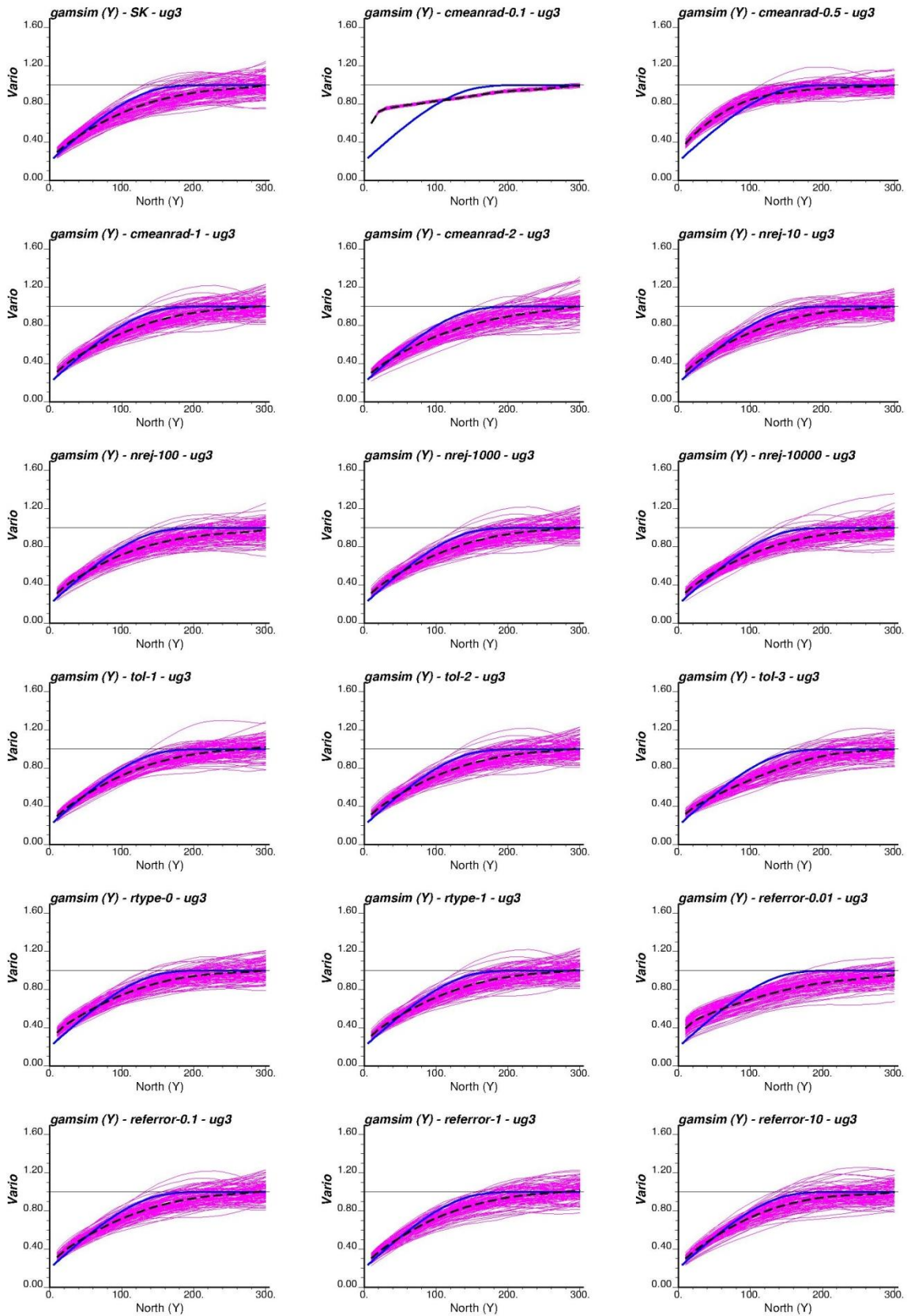
*Dirección Y*



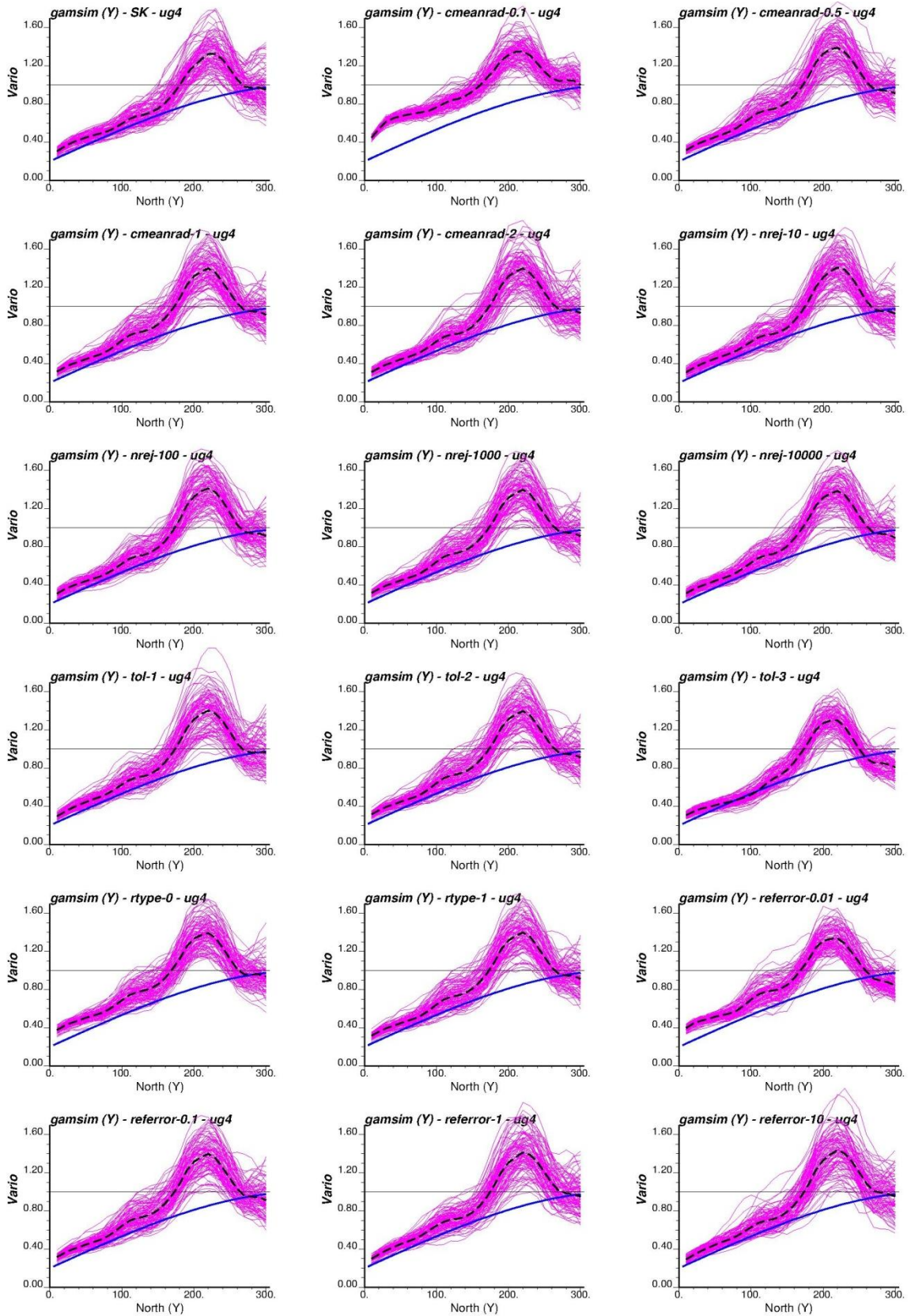
**Gráfico 176: Variograma experimental – Dirección Y – UE 1.**



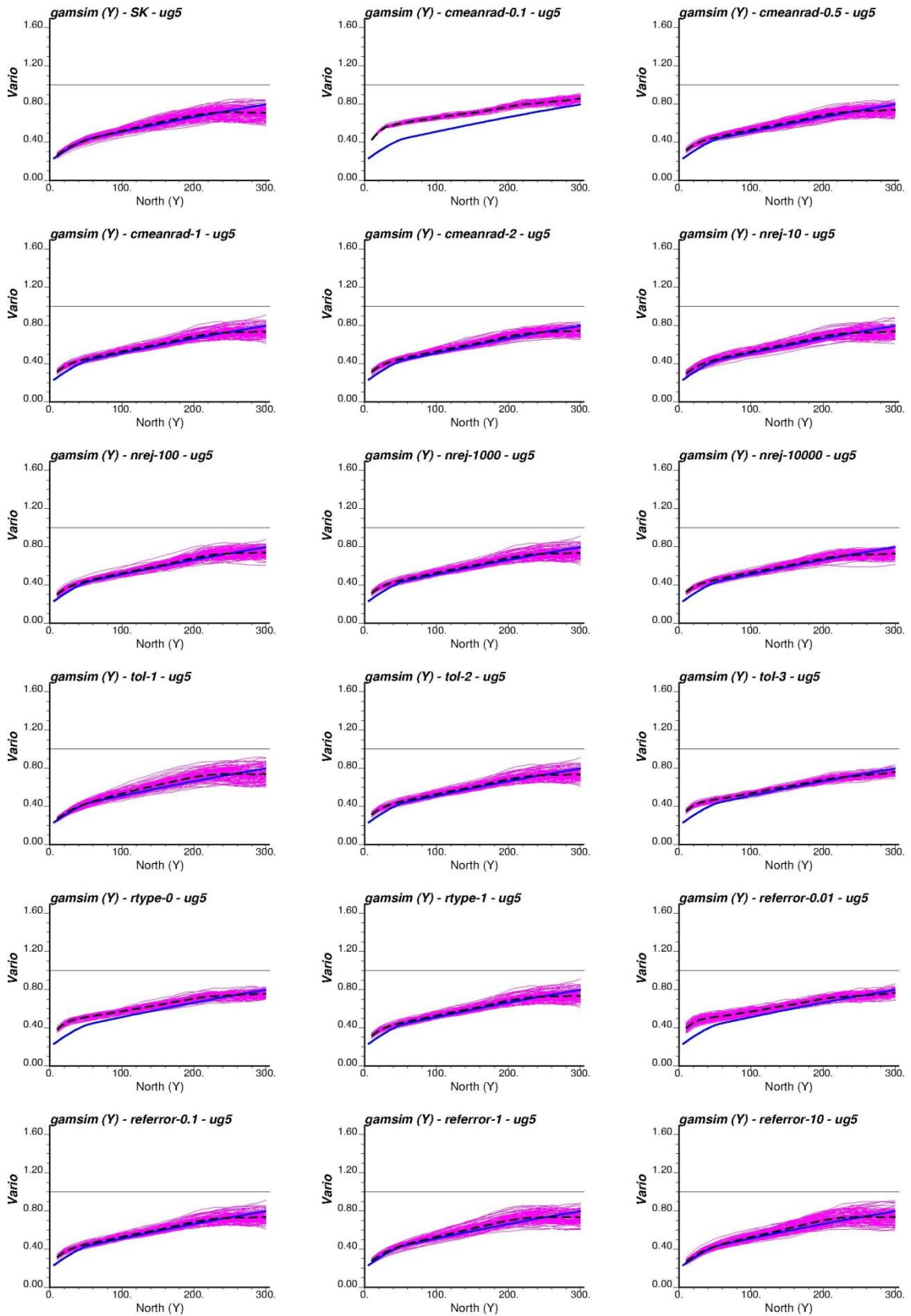
**Gráfico 177: Variograma experimental – Dirección Y – UE 2.**



**Gráfico 178: Variograma experimental – Dirección Y – UE 3.**

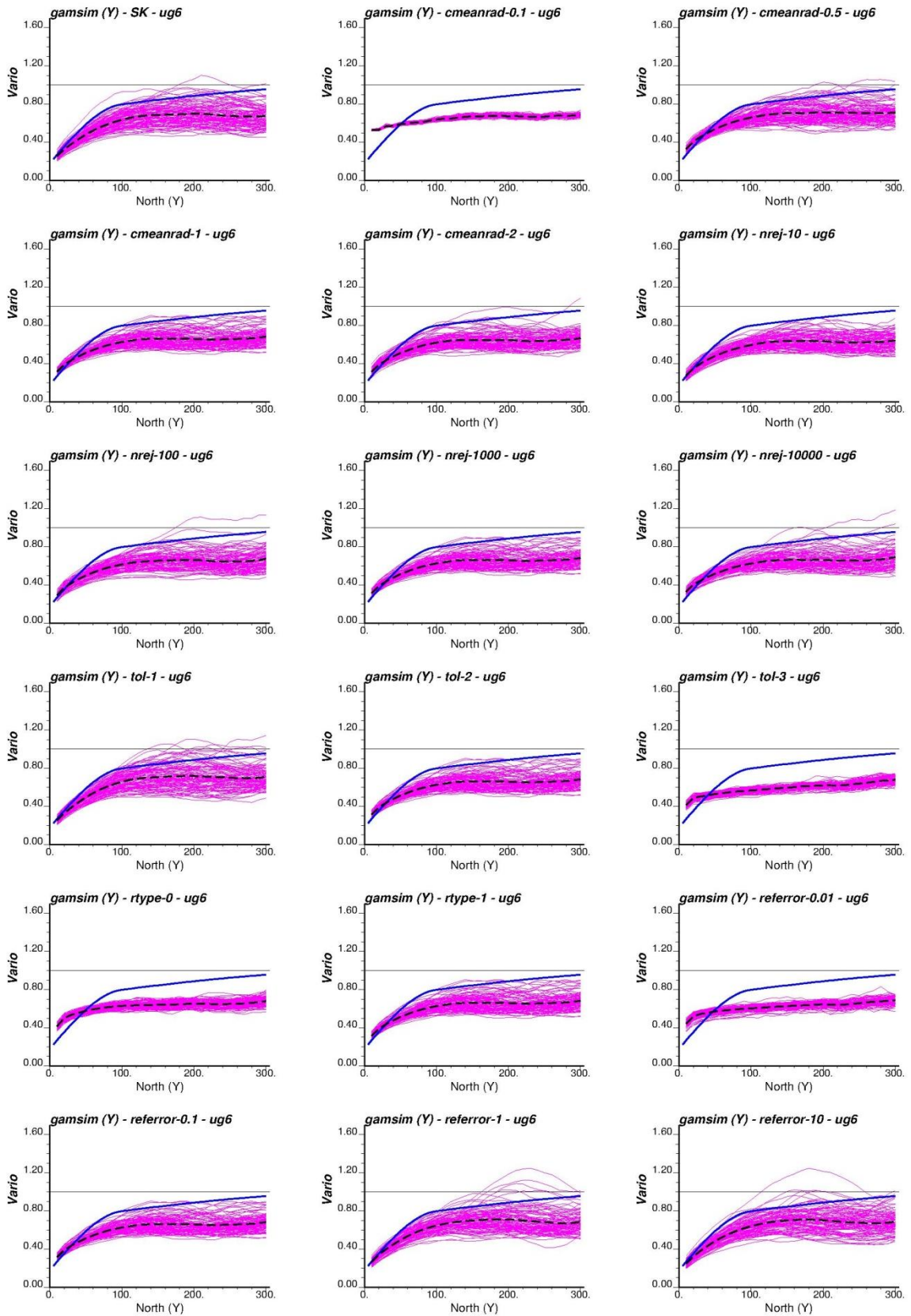


**Gráfico 179: Variograma experimental – Dirección Y – UE 4.**

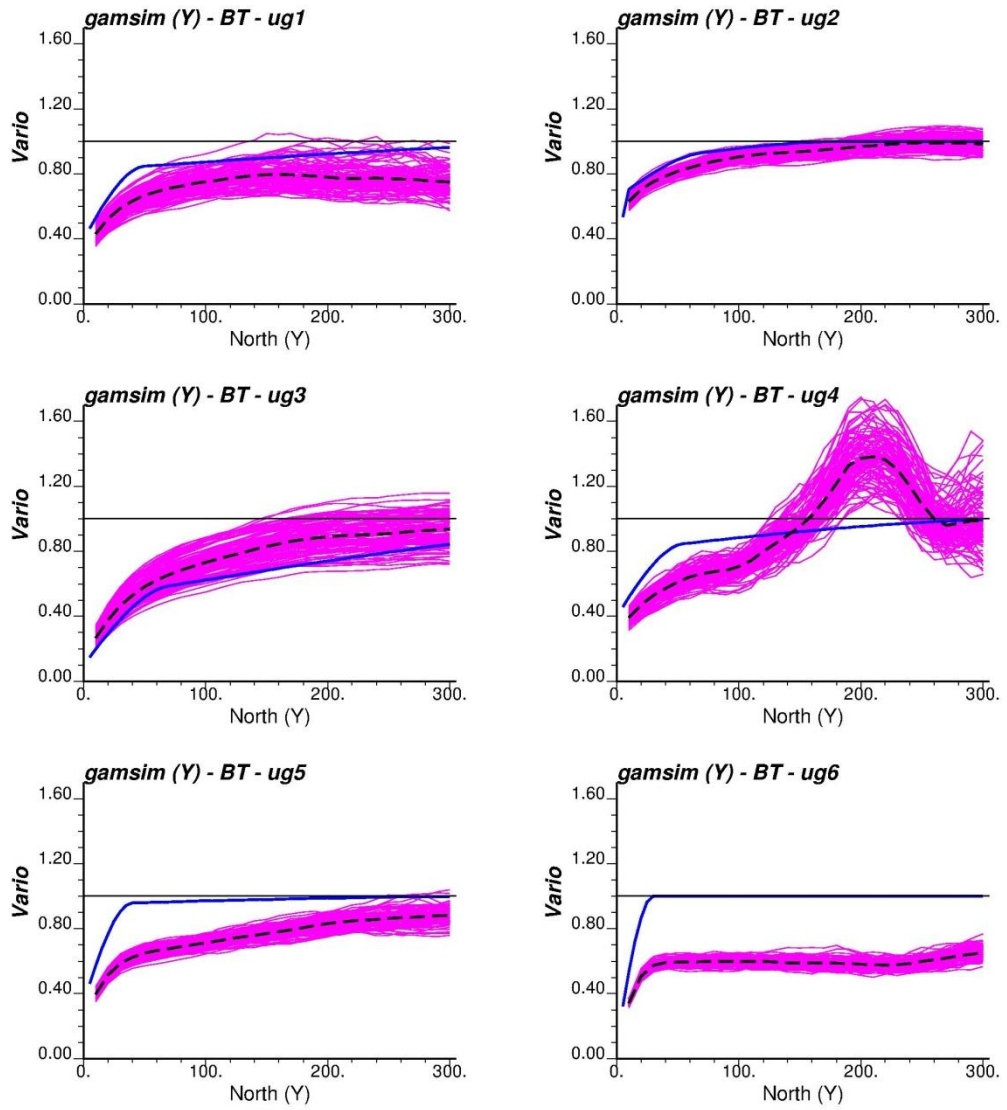


**Gráfico 180: Variograma experimental – Dirección Y – UE 5.**





**Gráfico 181: Variograma experimental – Dirección Y – UE 6.**



**Gráfico 182: Variograma experimental por UE – Dirección Y – Kriging de residuos.**

Dirección Z

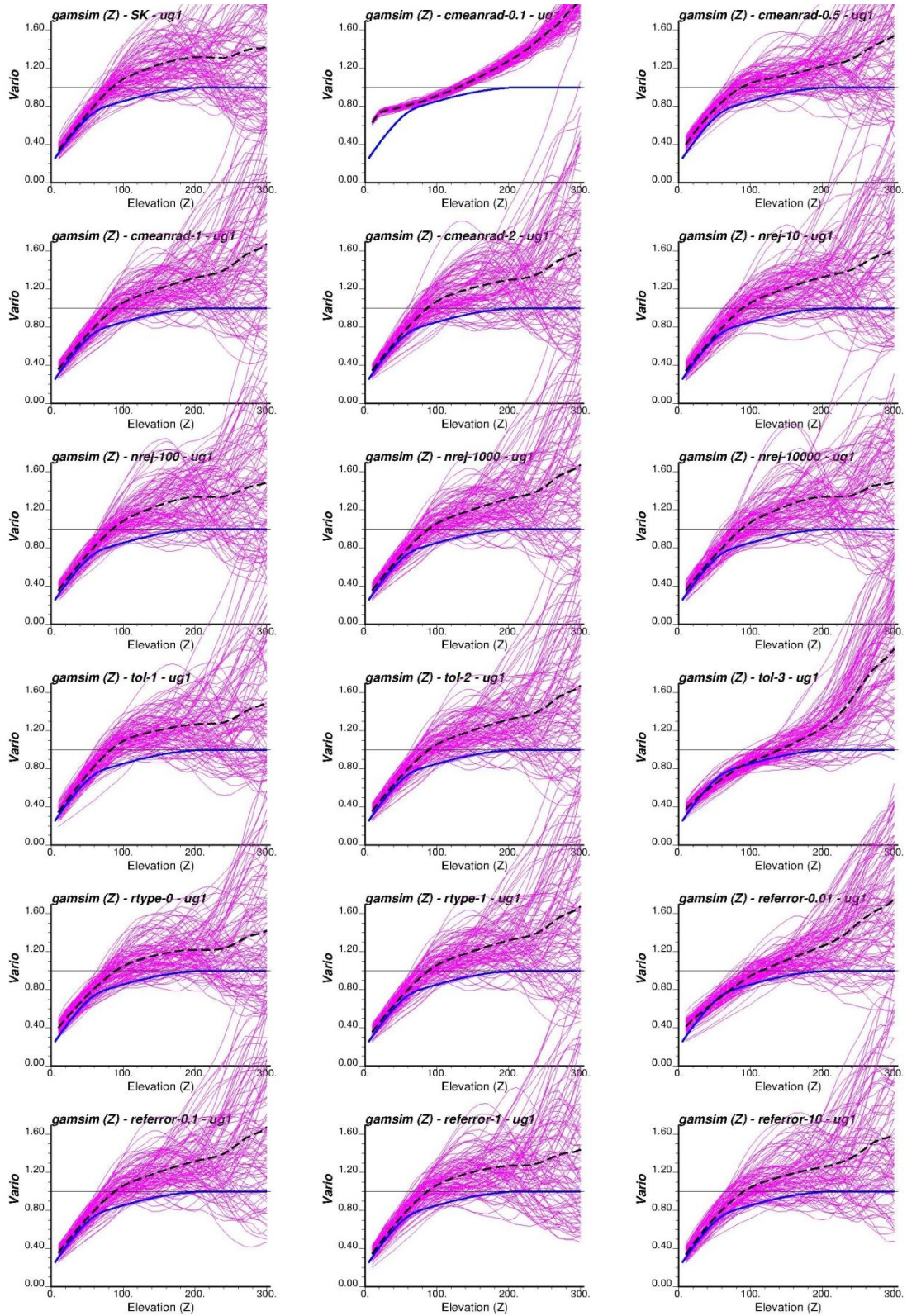
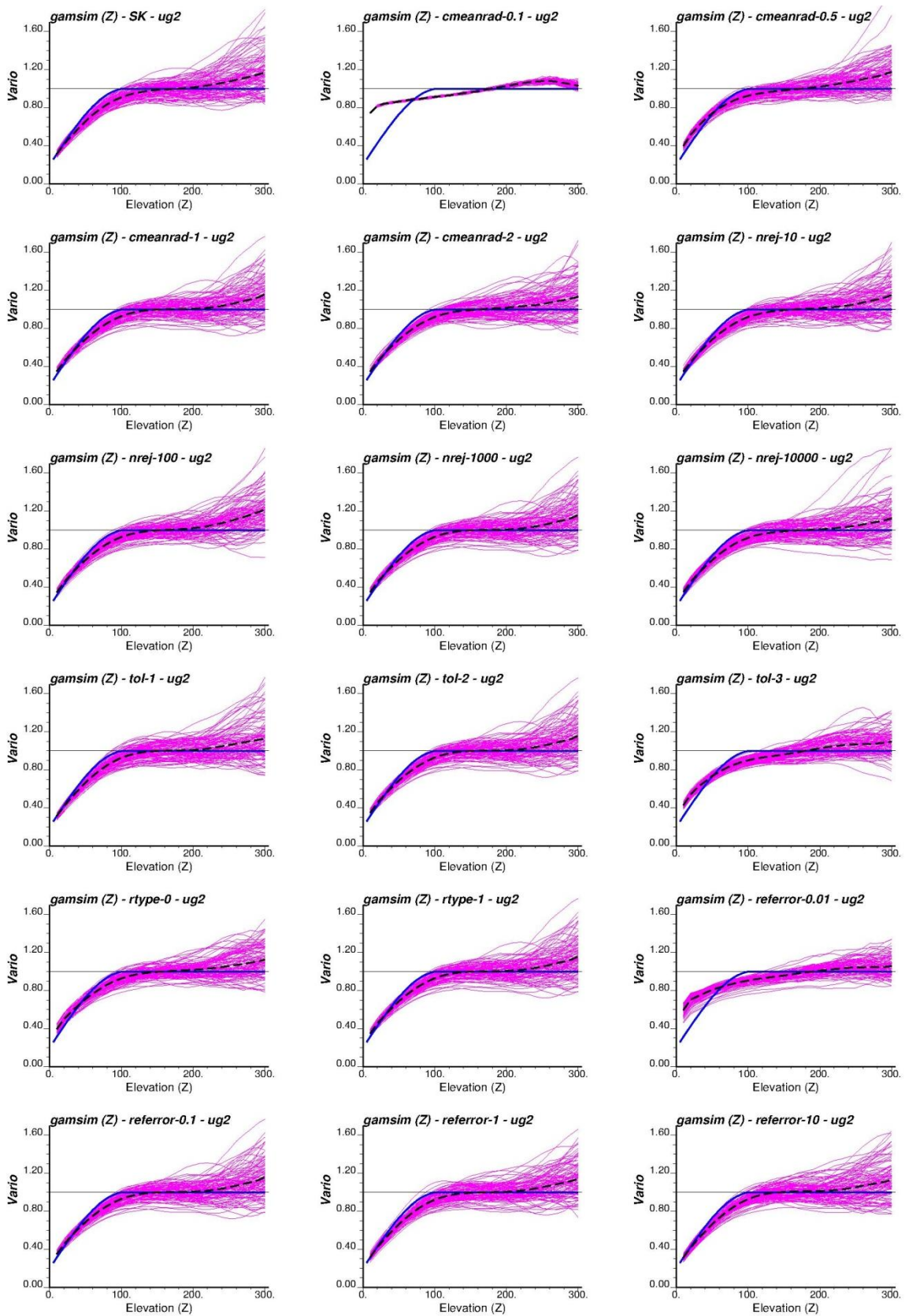
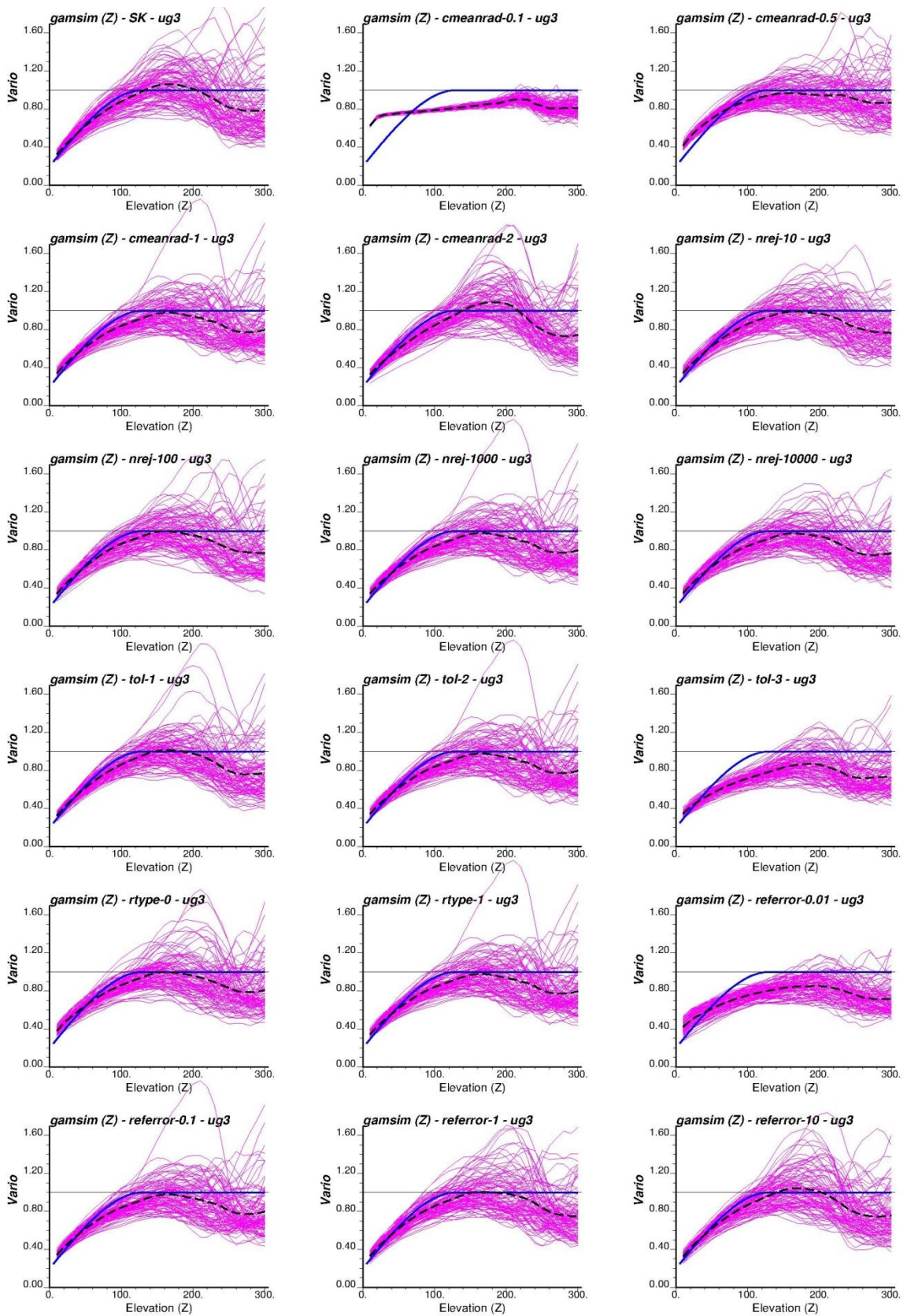


Gráfico 183: Variograma experimental – Dirección Z – UE 1.



**Gráfico 184: Variograma experimental – Dirección Z – UE 2.**



**Gráfico 185: Variograma experimental – Dirección Z – UE 3.**

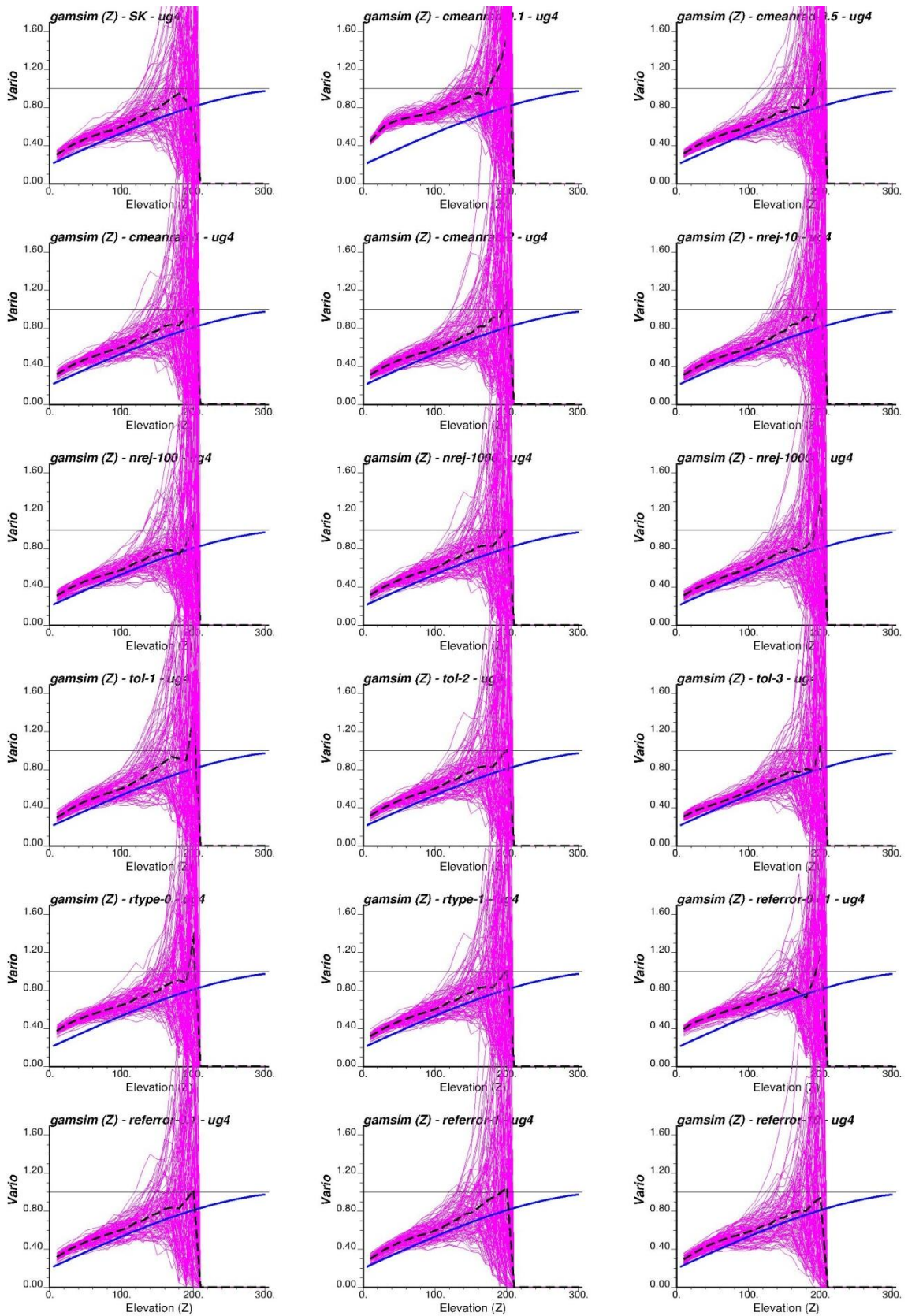


Gráfico 186: Variograma experimental – Dirección Z – UE 4.

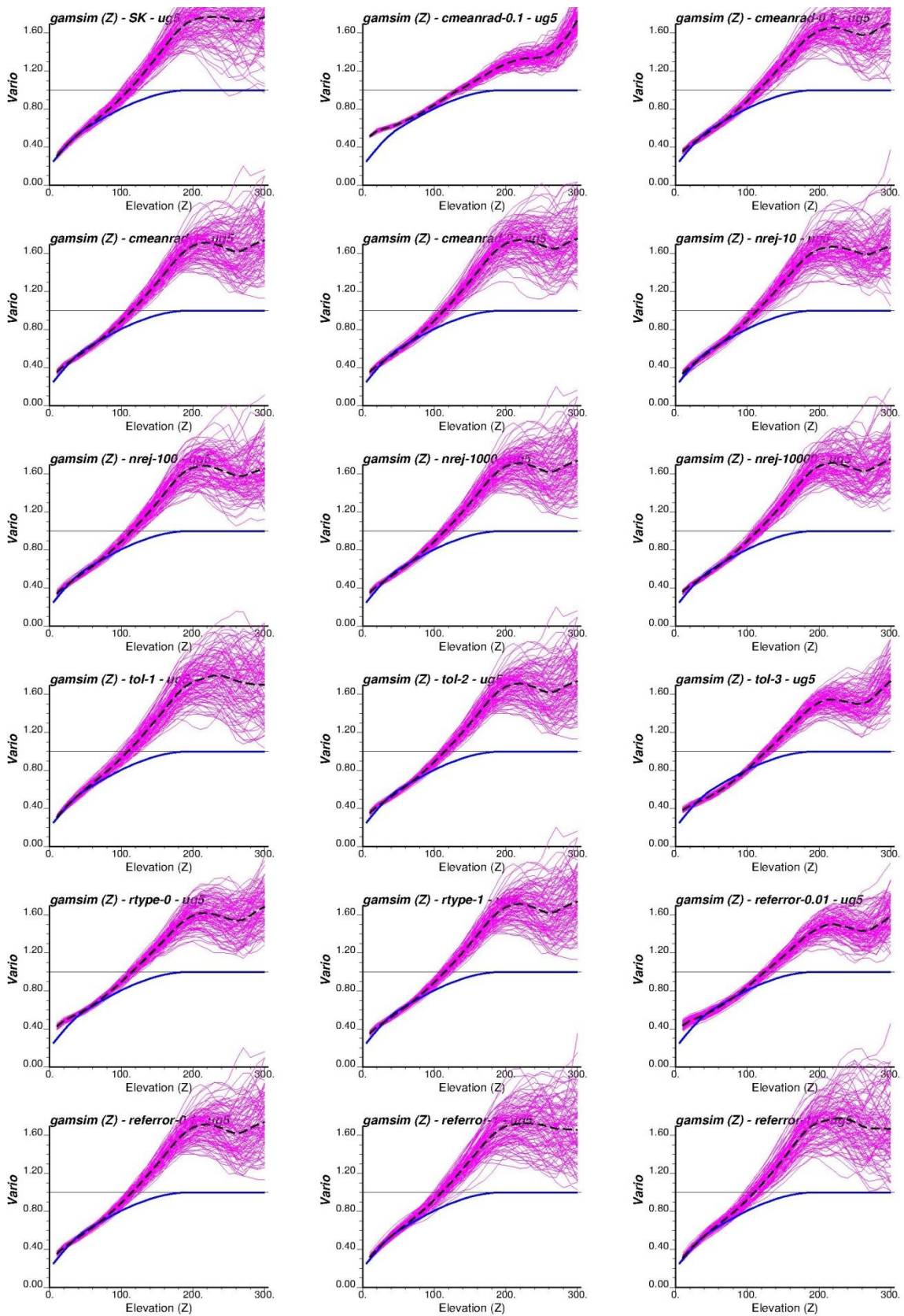
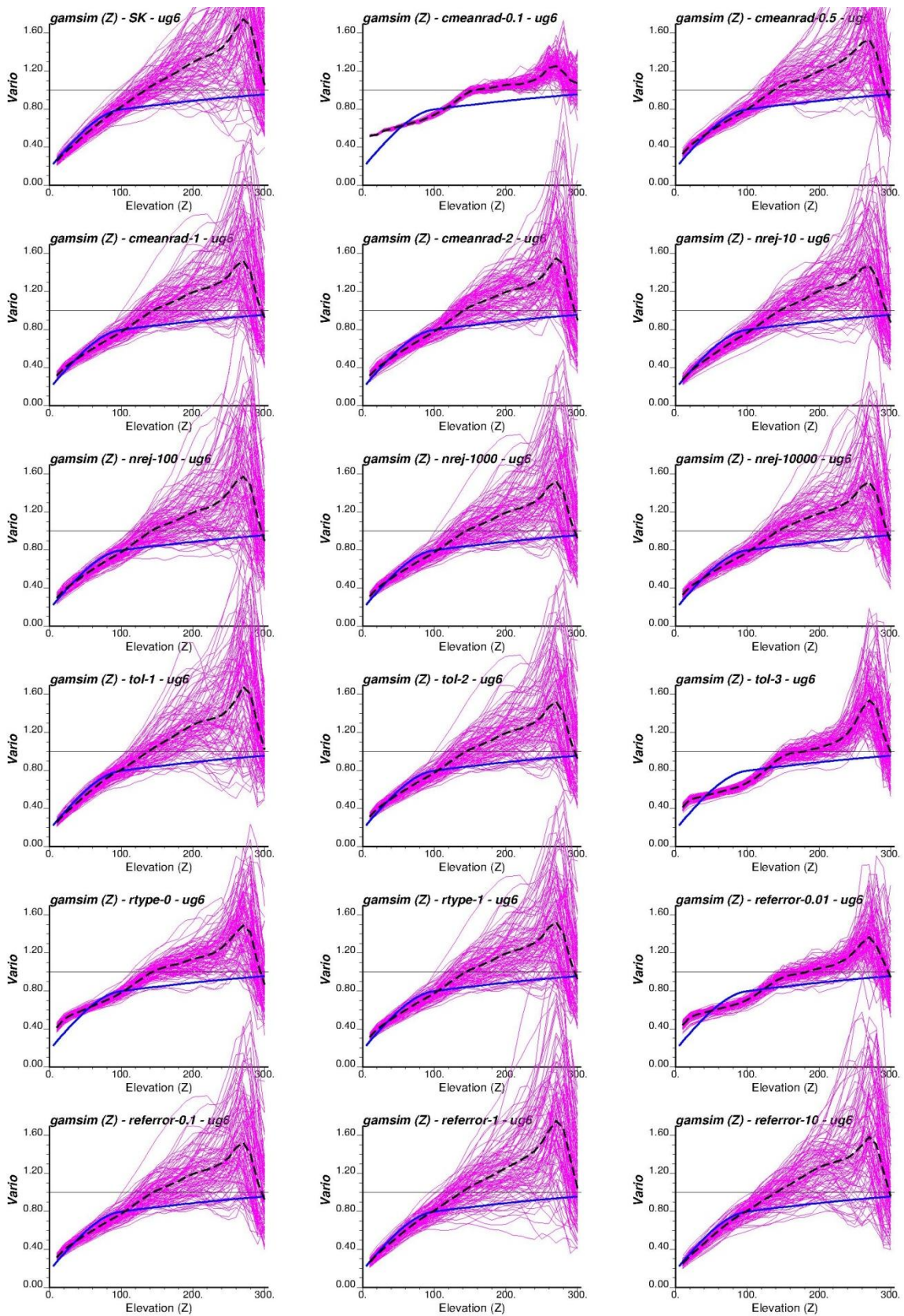
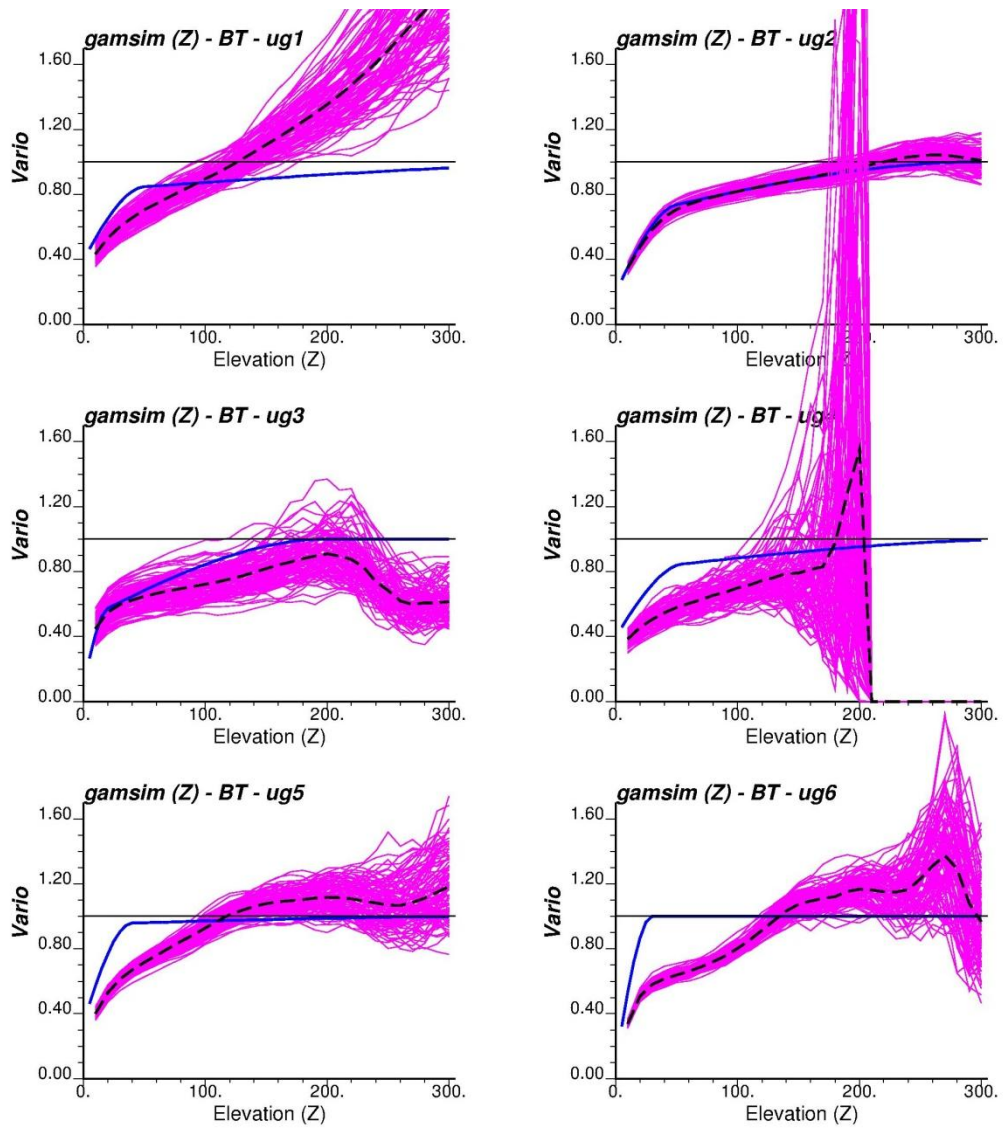


Gráfico 187: Variograma experimental – Dirección Z – UE 5.



**Gráfico 188: Variograma experimental – Dirección Z – UE 6.**



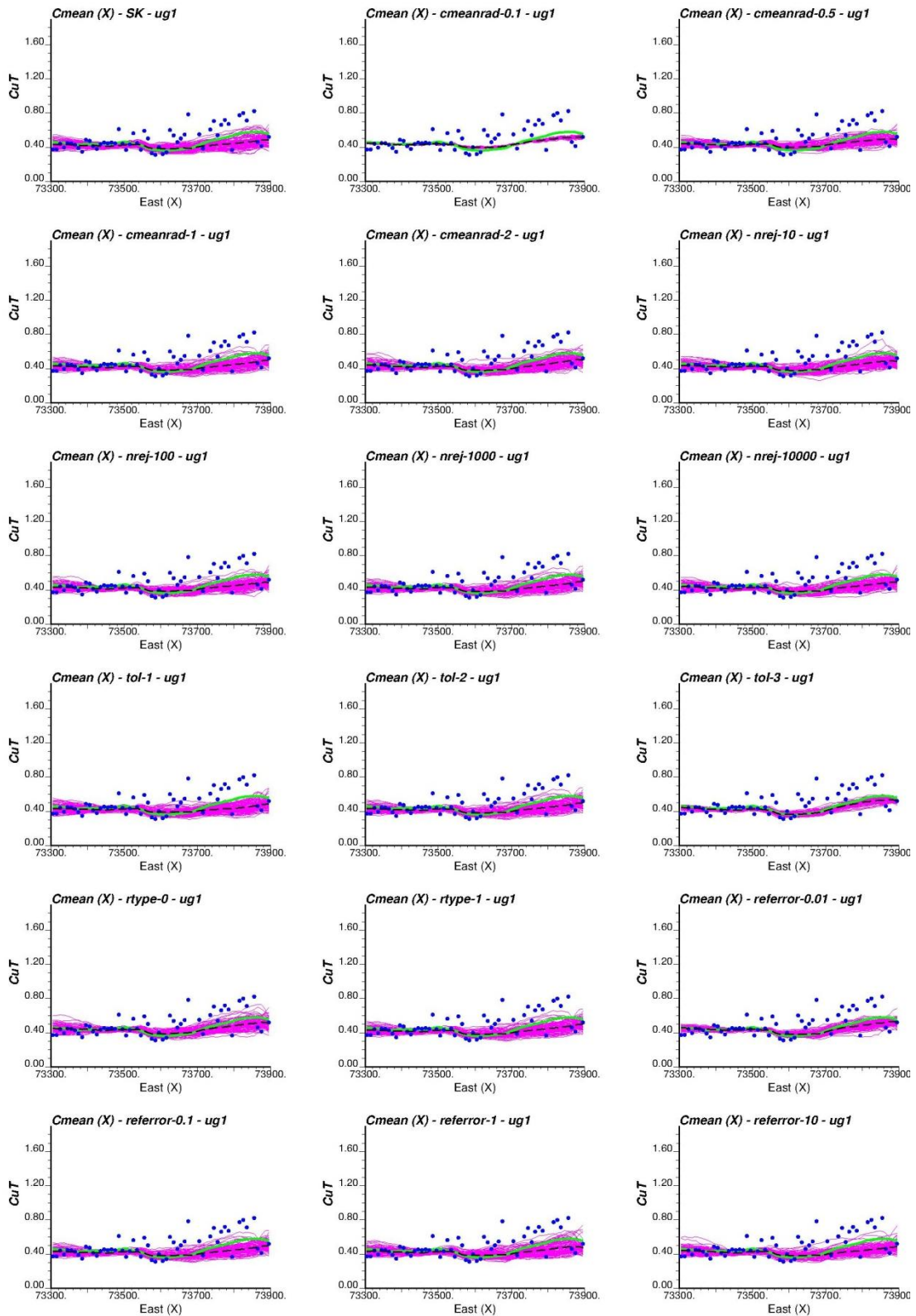


**Gráfico 189: Variograma experimental por UE – Dirección Z – Kriging de residuos.**

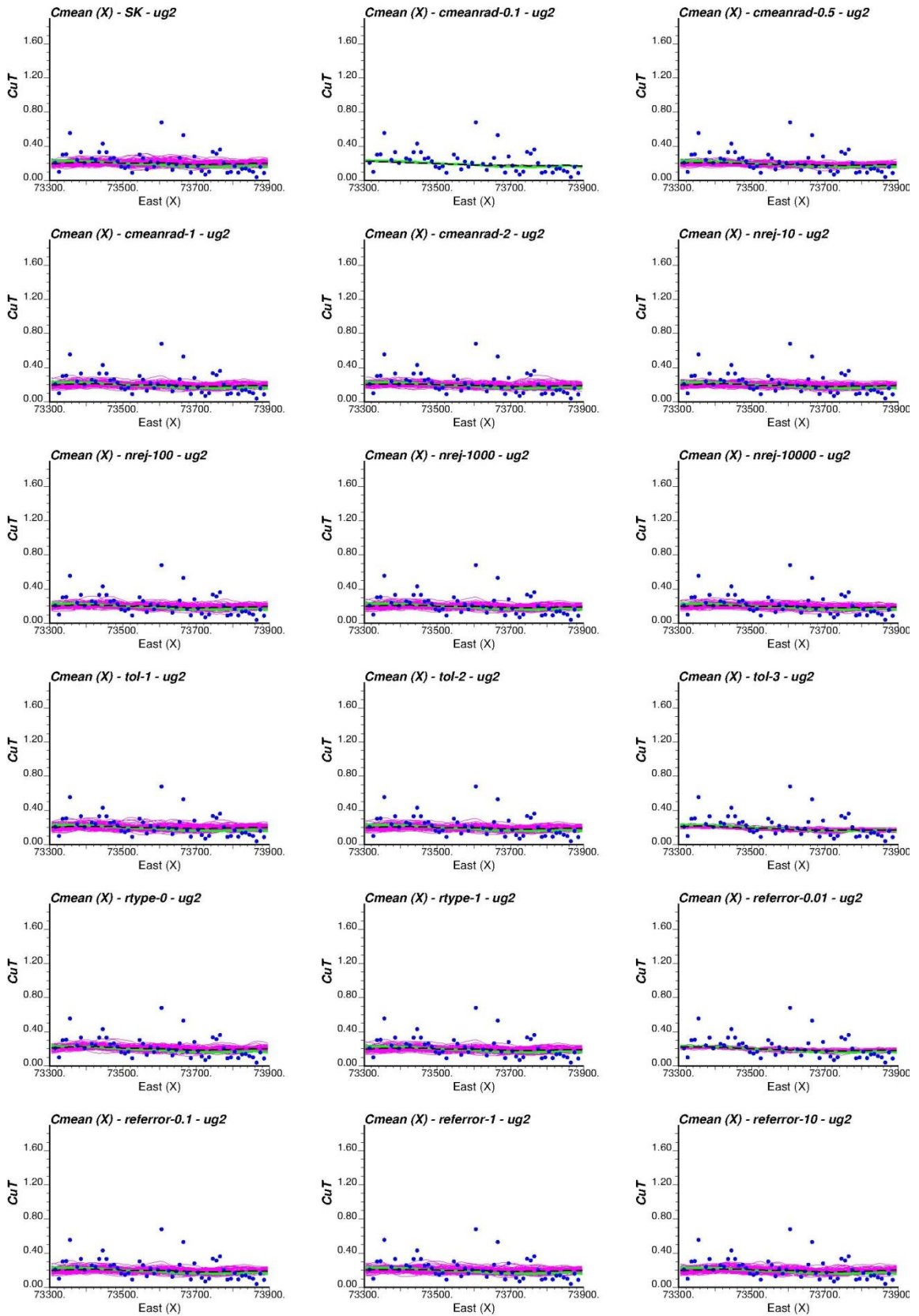
## **ANEXO M: Caso de estudio II – Media local por UE**

Se presenta la media local de los resultados de las simulaciones en el mismo formato en que se presentan los histogramas, es decir, considerando 18 medias locales en cada gráfico por UE y ocupando el mismo orden para su presentación junto con un gráfico que presenta los resultados del Kriging de residuos por UE. Cada media local muestra, además de las 100 medias locales de cada simulación (líneas moradas), los compósitos (puntos azules), la media de las simulaciones (línea negra segmentada) y el modelo de deriva (línea verde). El cálculo se realiza a la resolución de la grilla para cada una de las direcciones X, Y y Z.

*Dirección X*



**Gráfico 190: Media local – Dirección X – UE 1.**



**Gráfico 191: Media local – Dirección X – UE 2.**

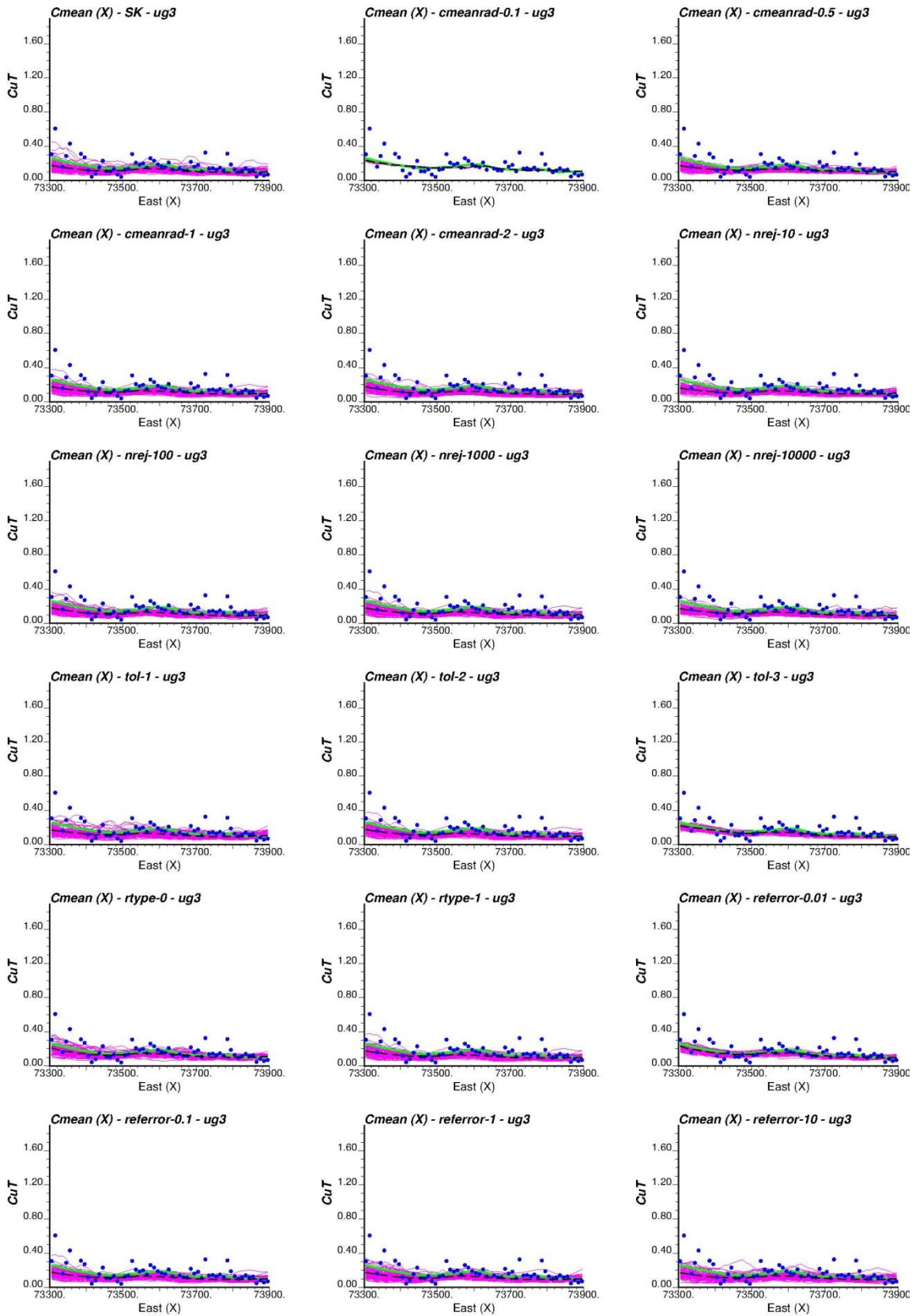


Gráfico 192: Media local – Dirección X – UE 3.

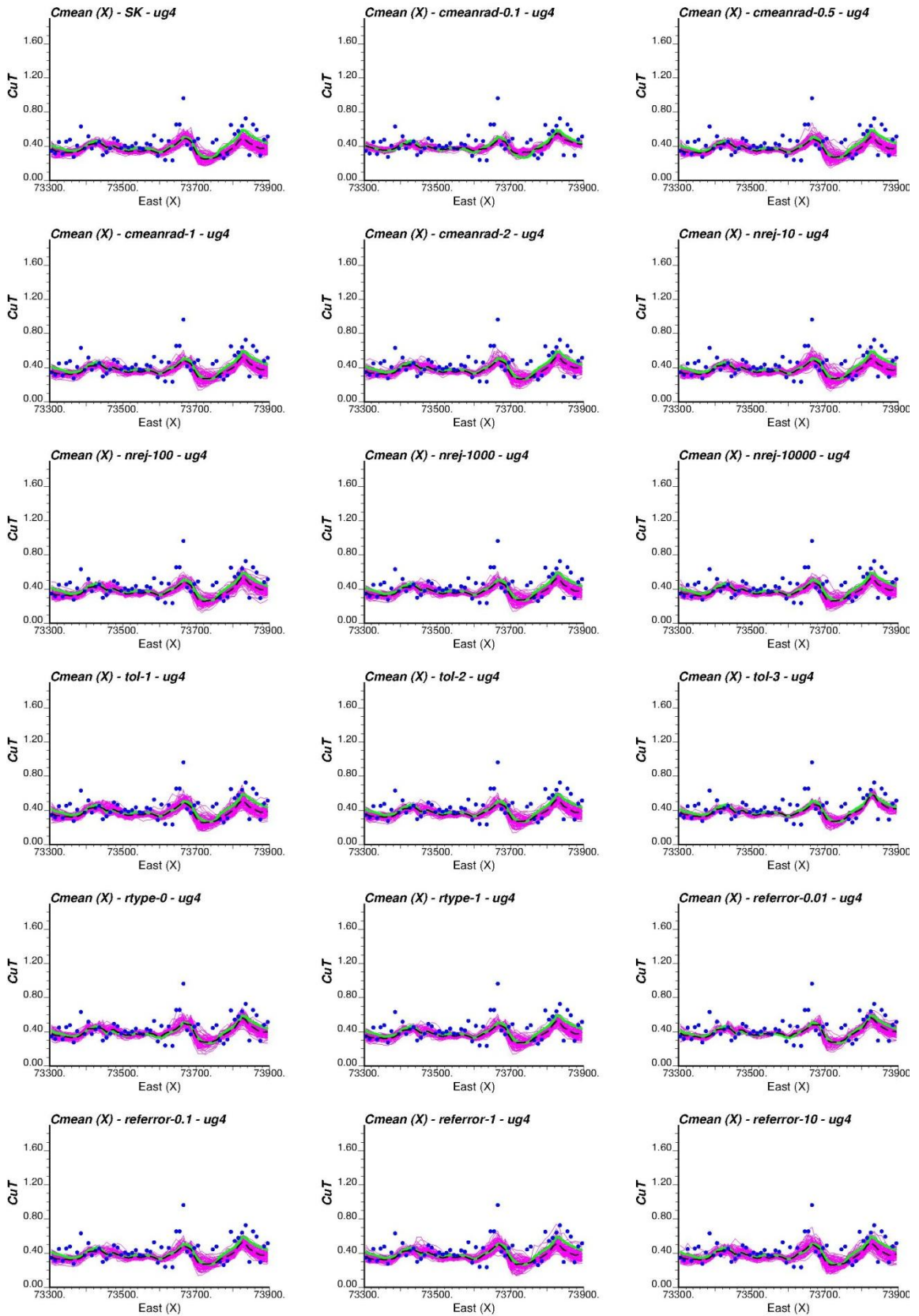
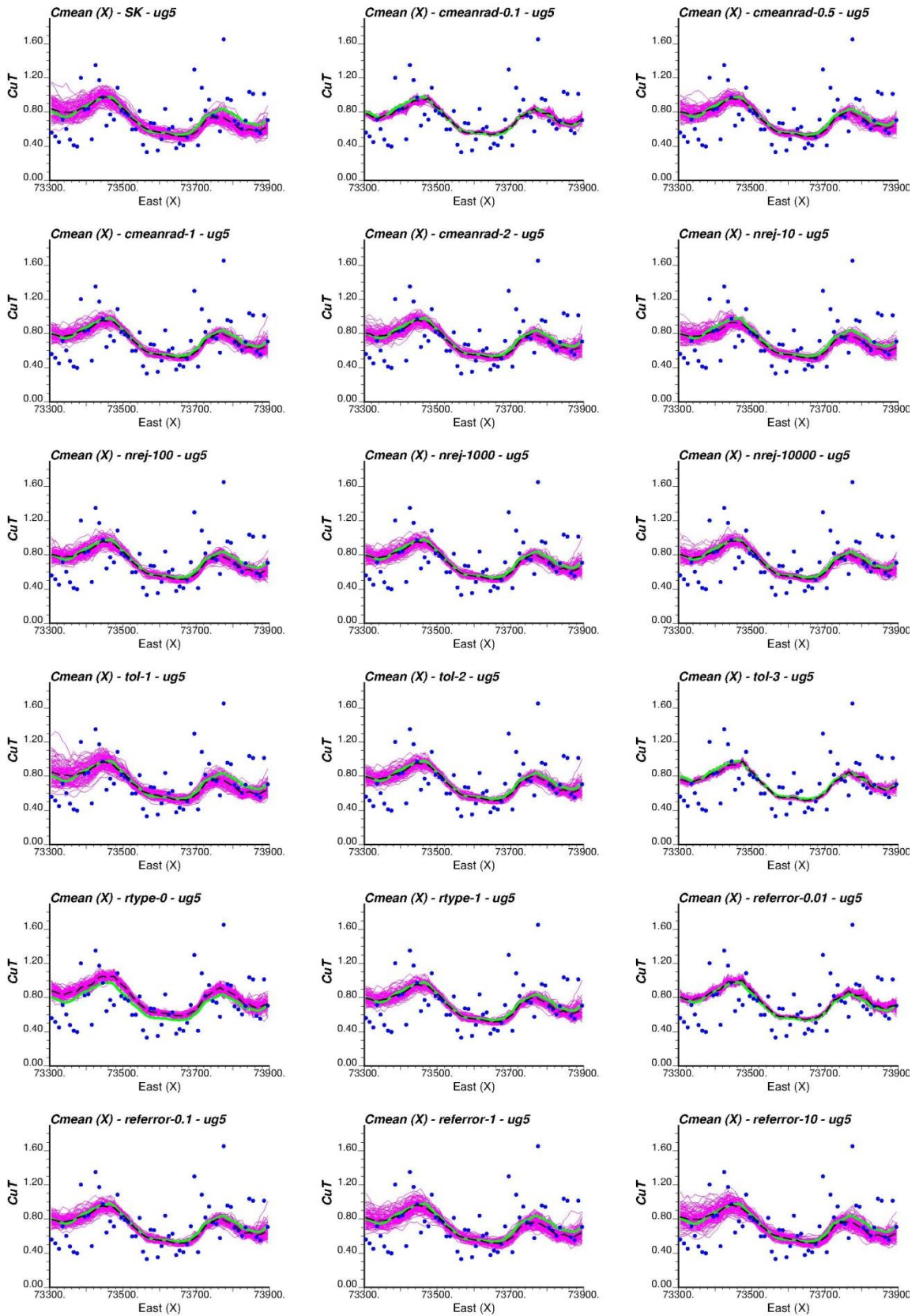


Gráfico 193: Media local – Dirección X – UE 4.



**Gráfico 194: Media local – Dirección X – UE 5.**

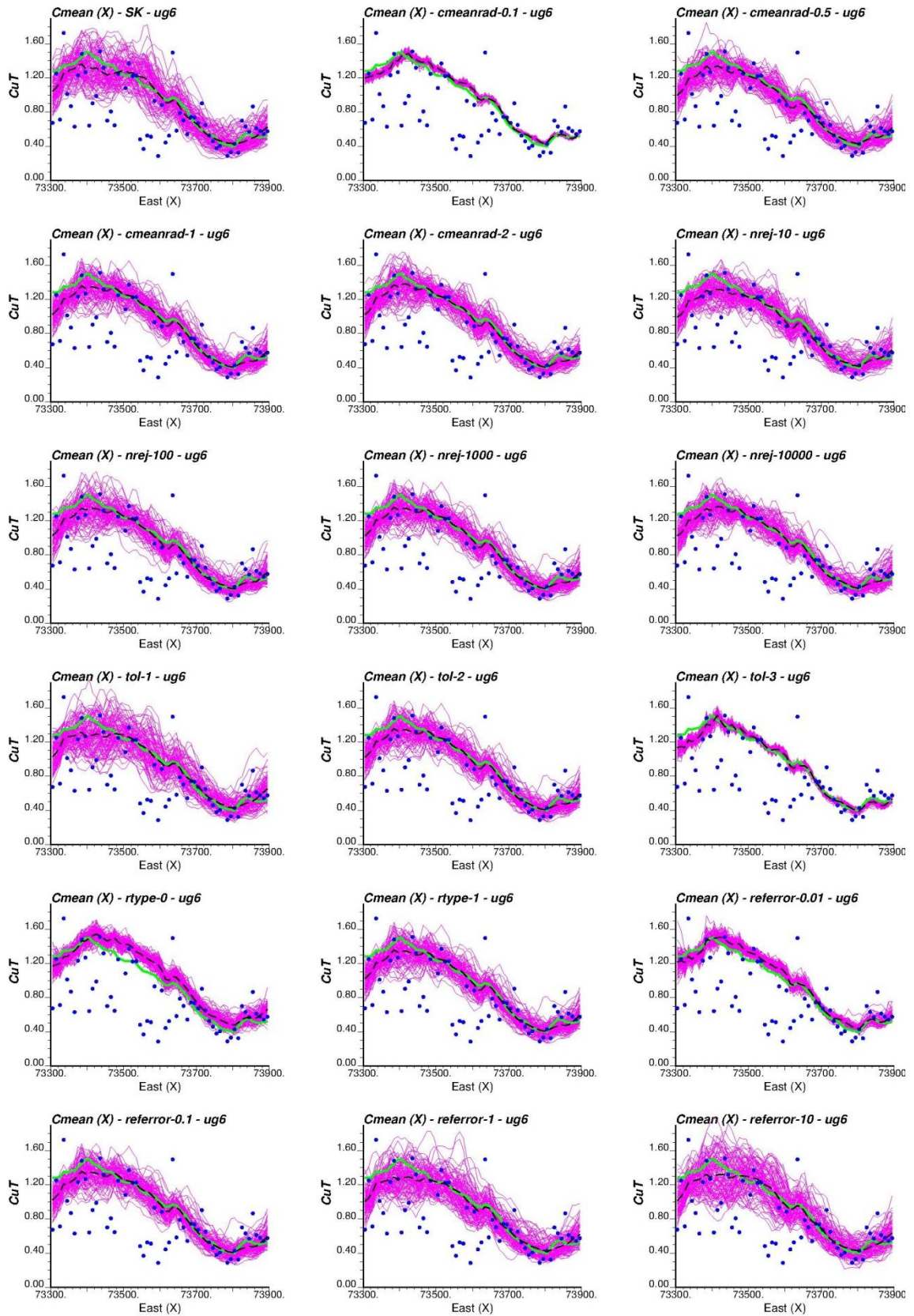
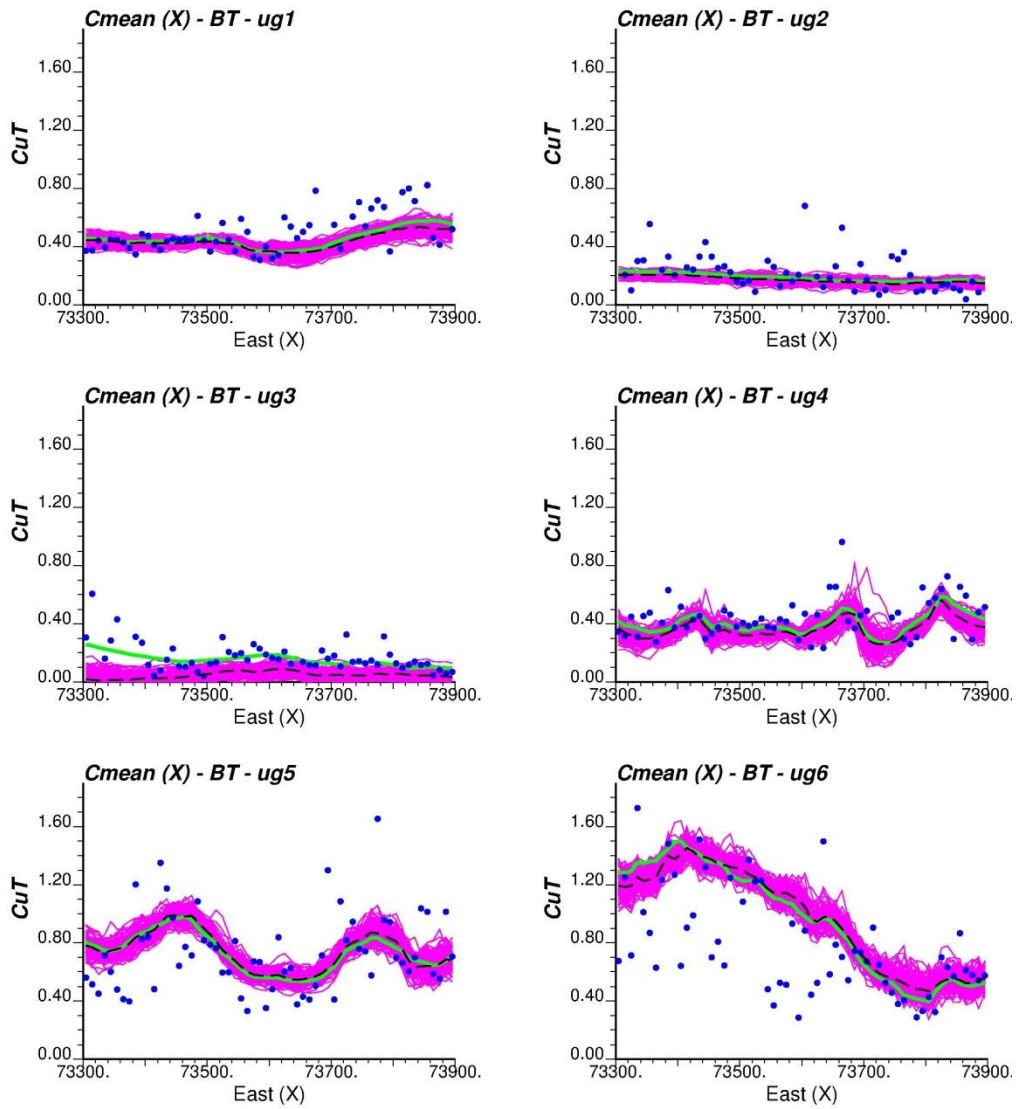


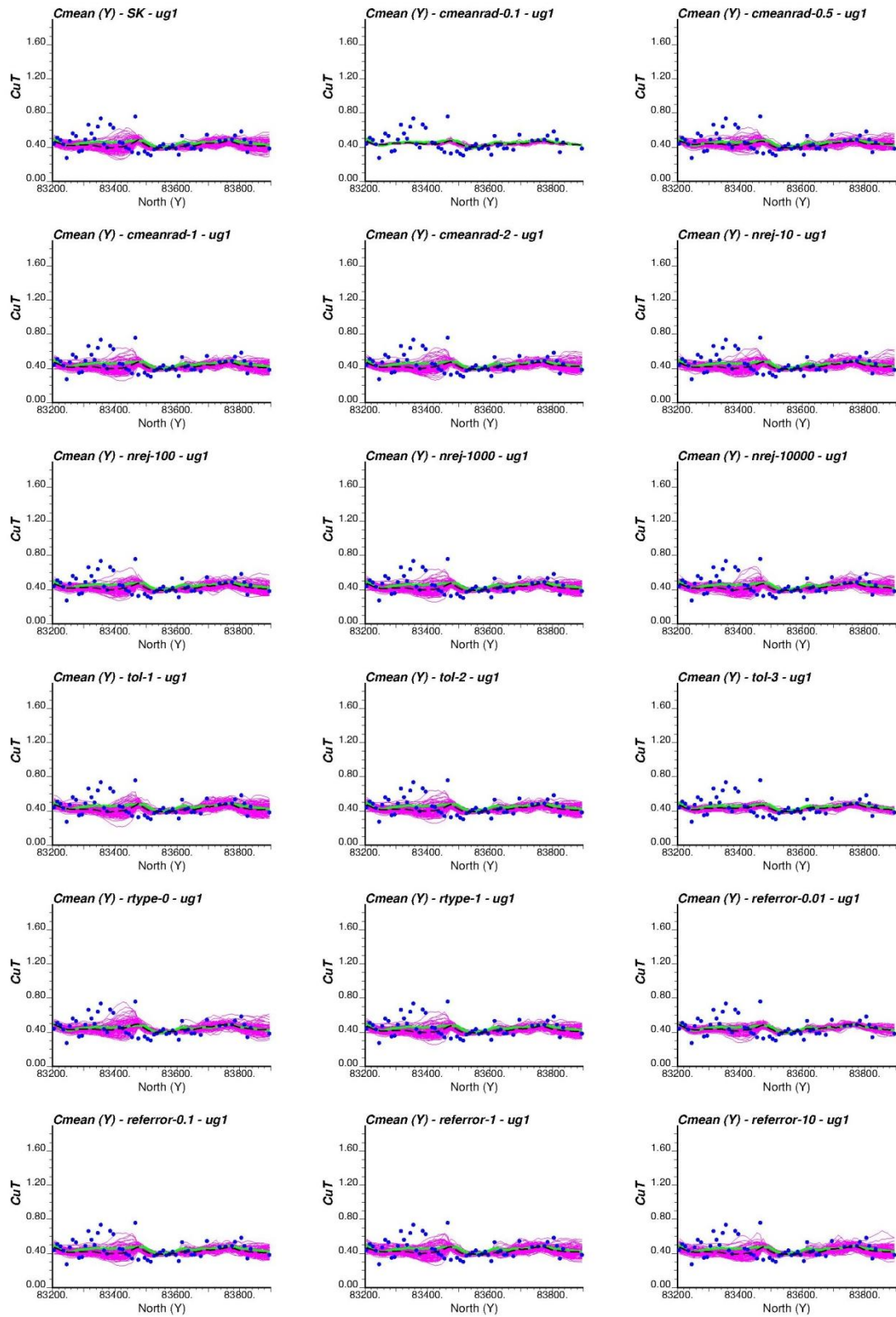
Gráfico 195: Media local – Dirección X – UE 6.



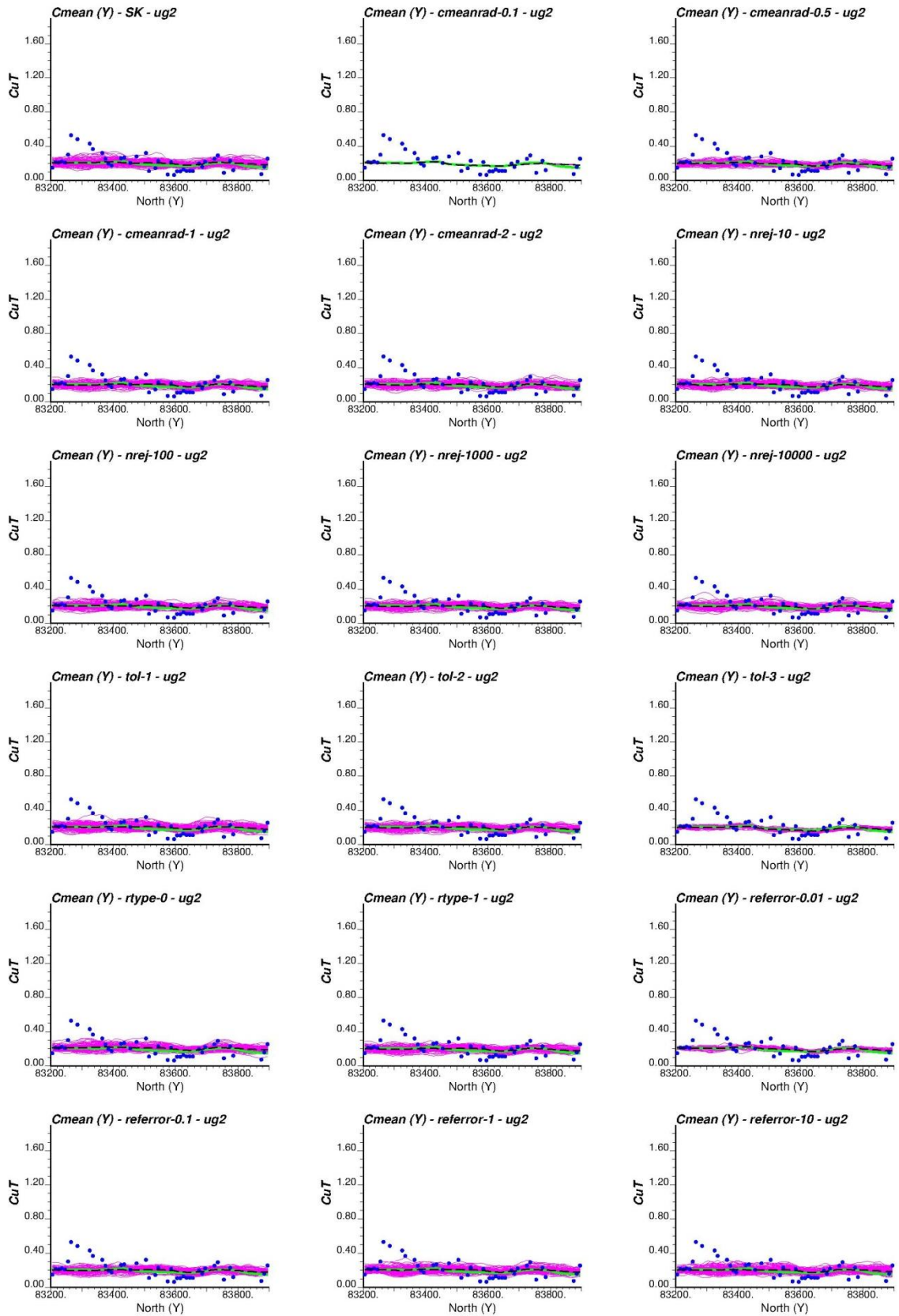


**Gráfico 196: Media local por UE – Dirección X – Kriging de residuos.**

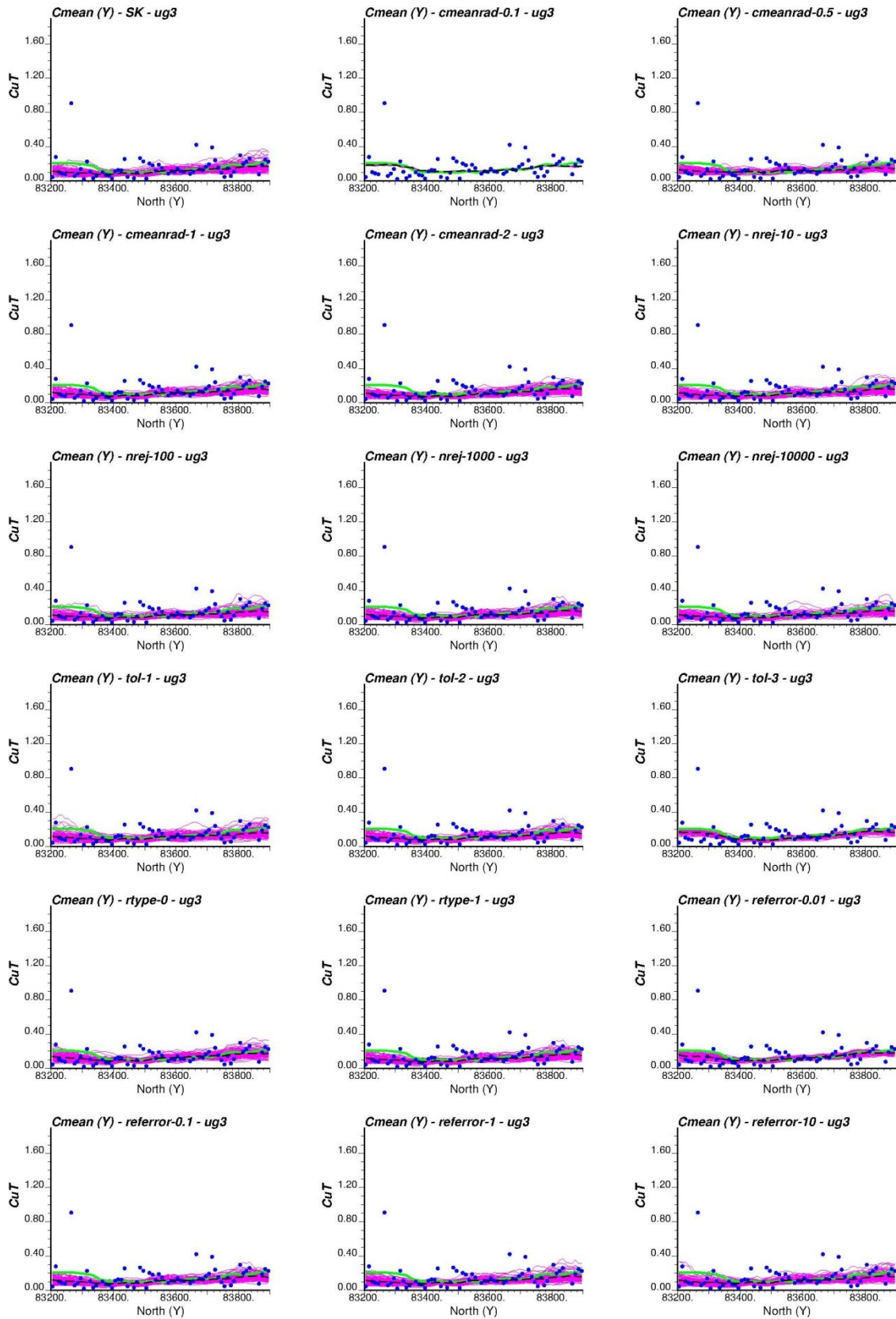
*Dirección Y*



**Gráfico 197: Media local – Dirección Y – UE 1.**



**Gráfico 198: Media local – Dirección Y – UE 2.**



**Gráfico 199: Media local – Dirección Y – UE 3.**

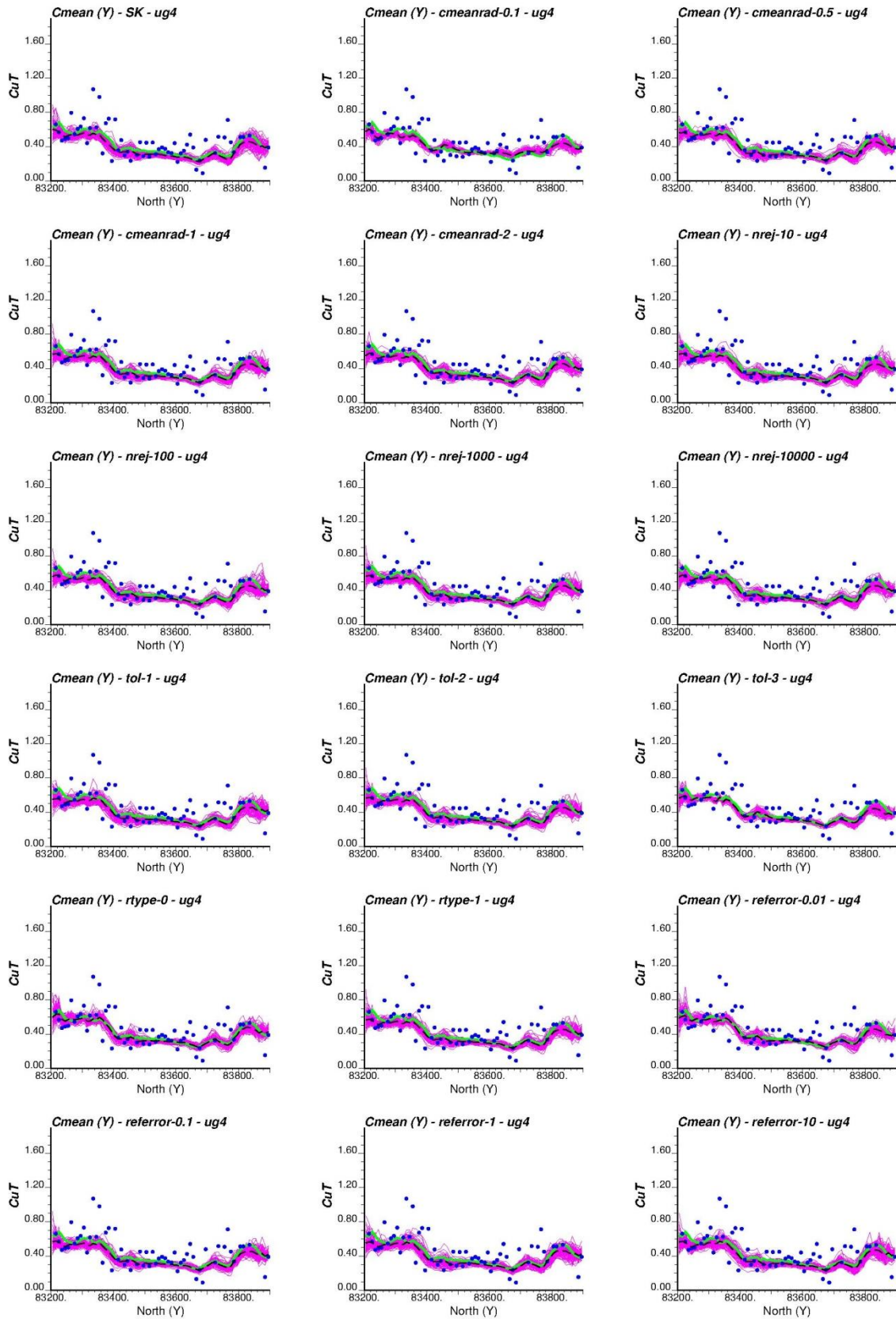


Gráfico 200: Media local – Dirección Y – UE 4.

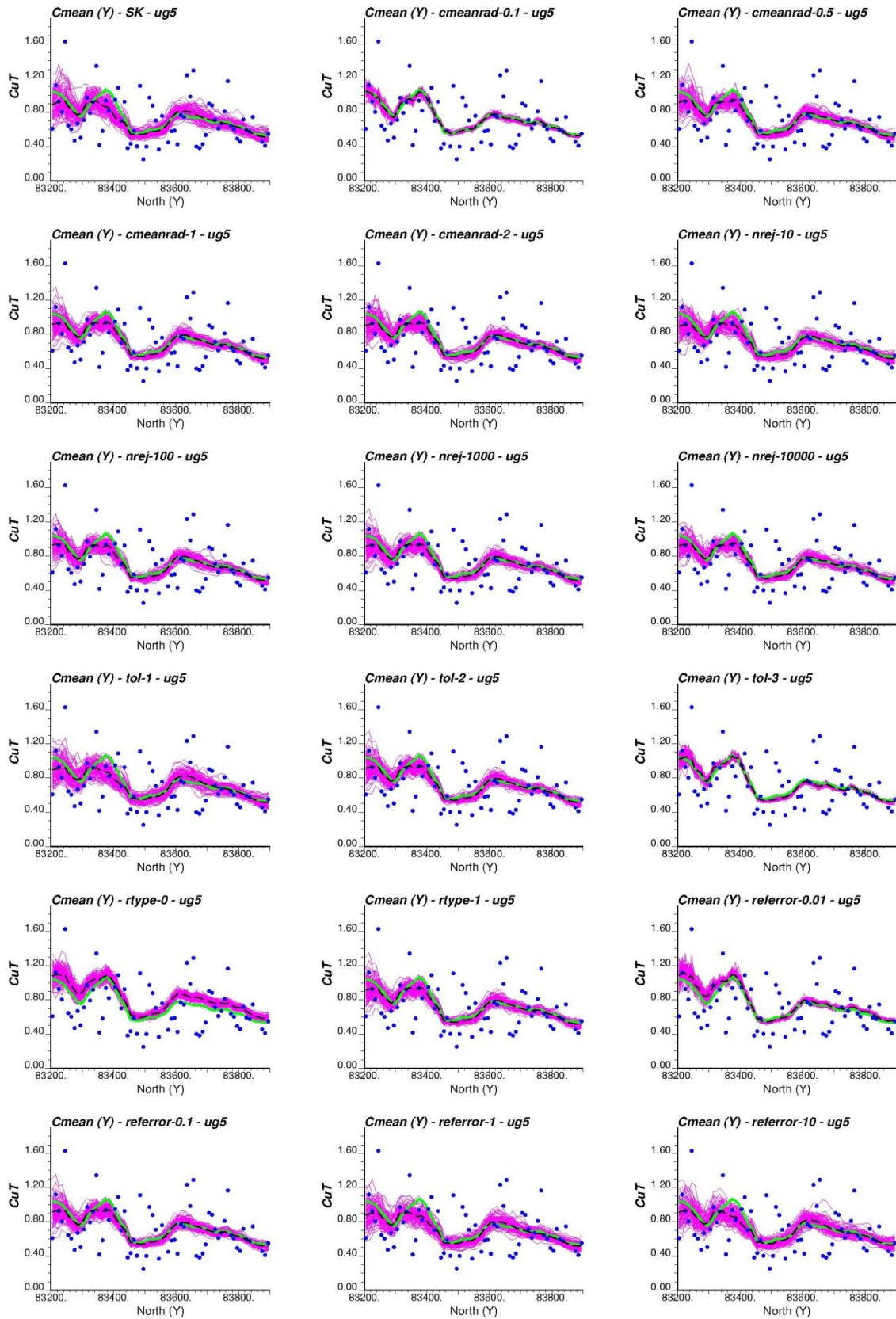


Gráfico 201: Media local – Dirección Y – UE 5.

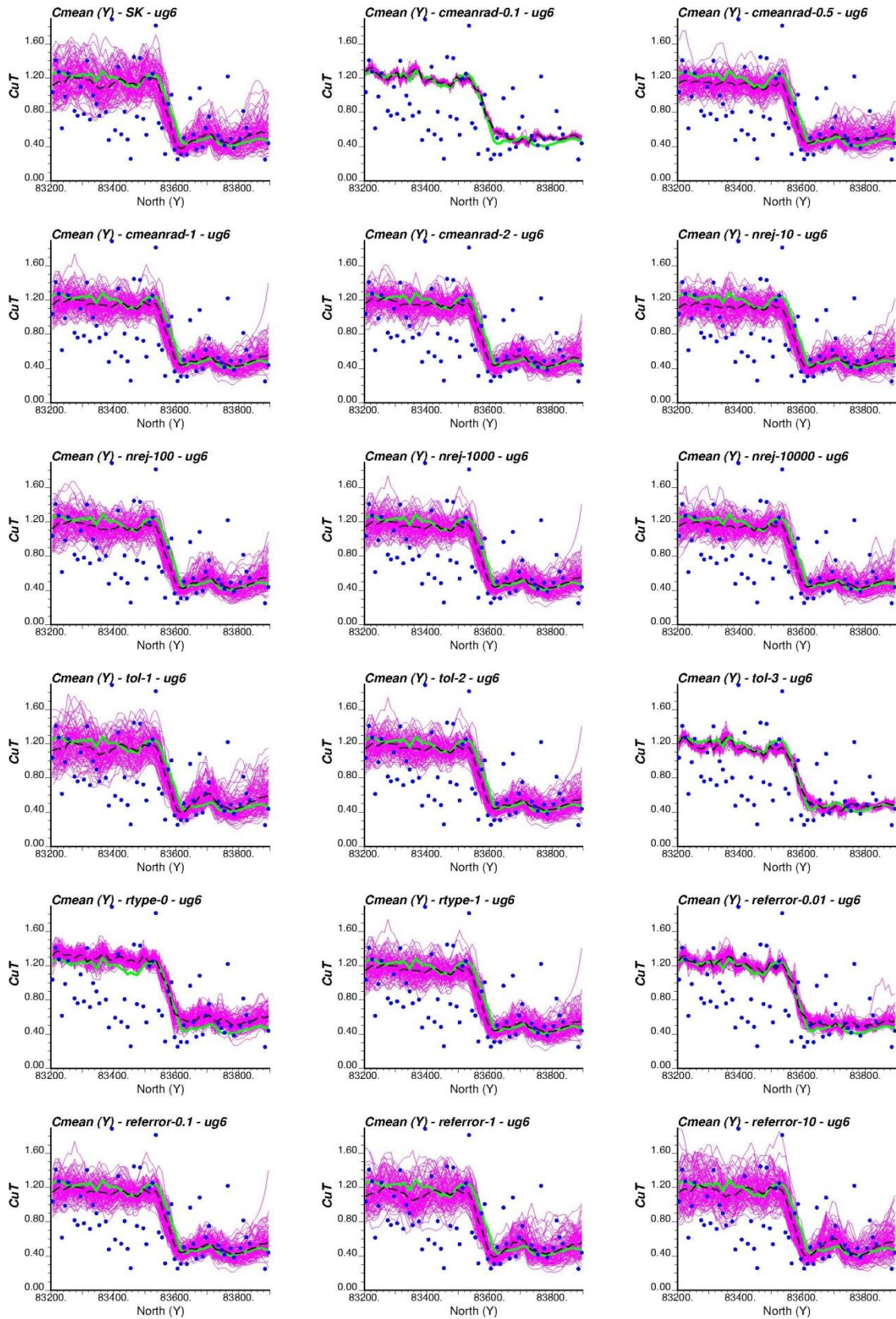
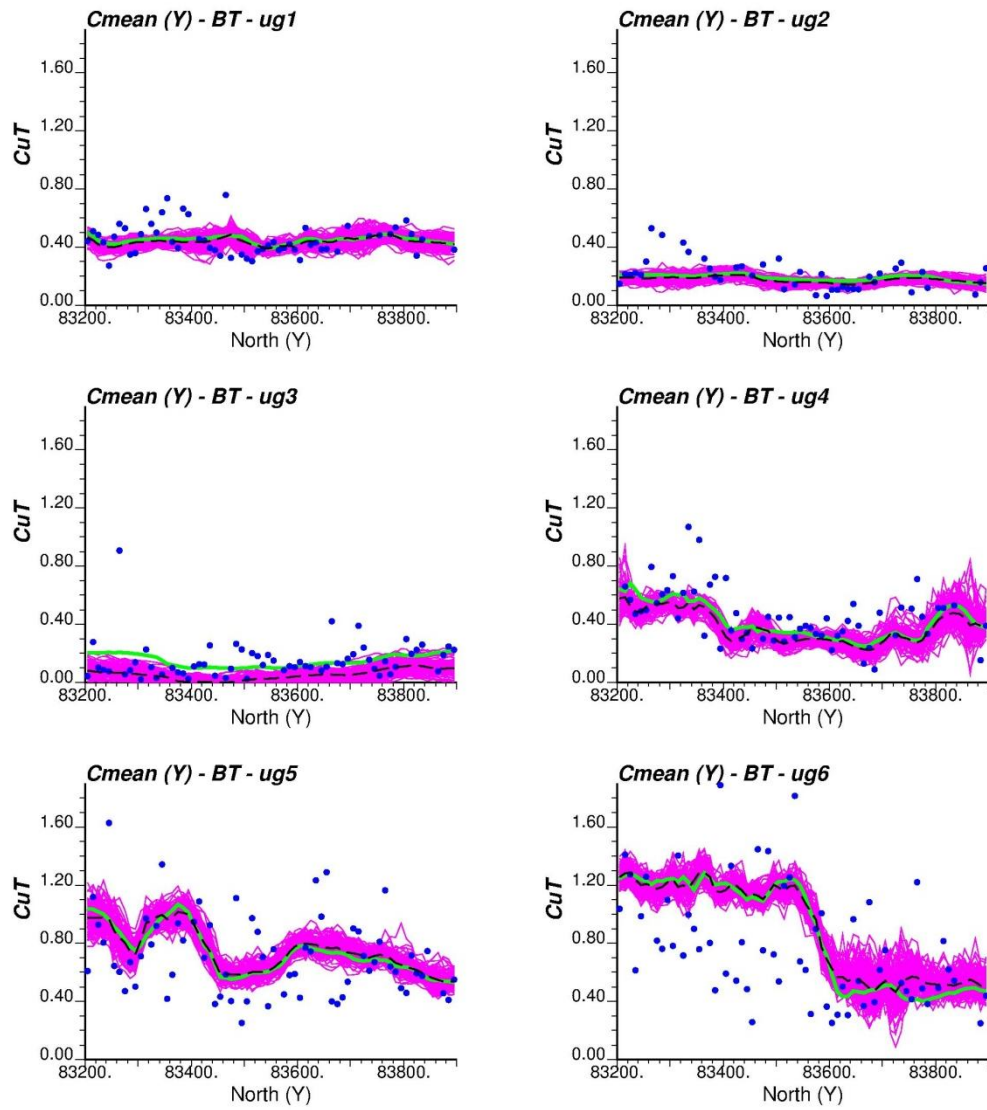


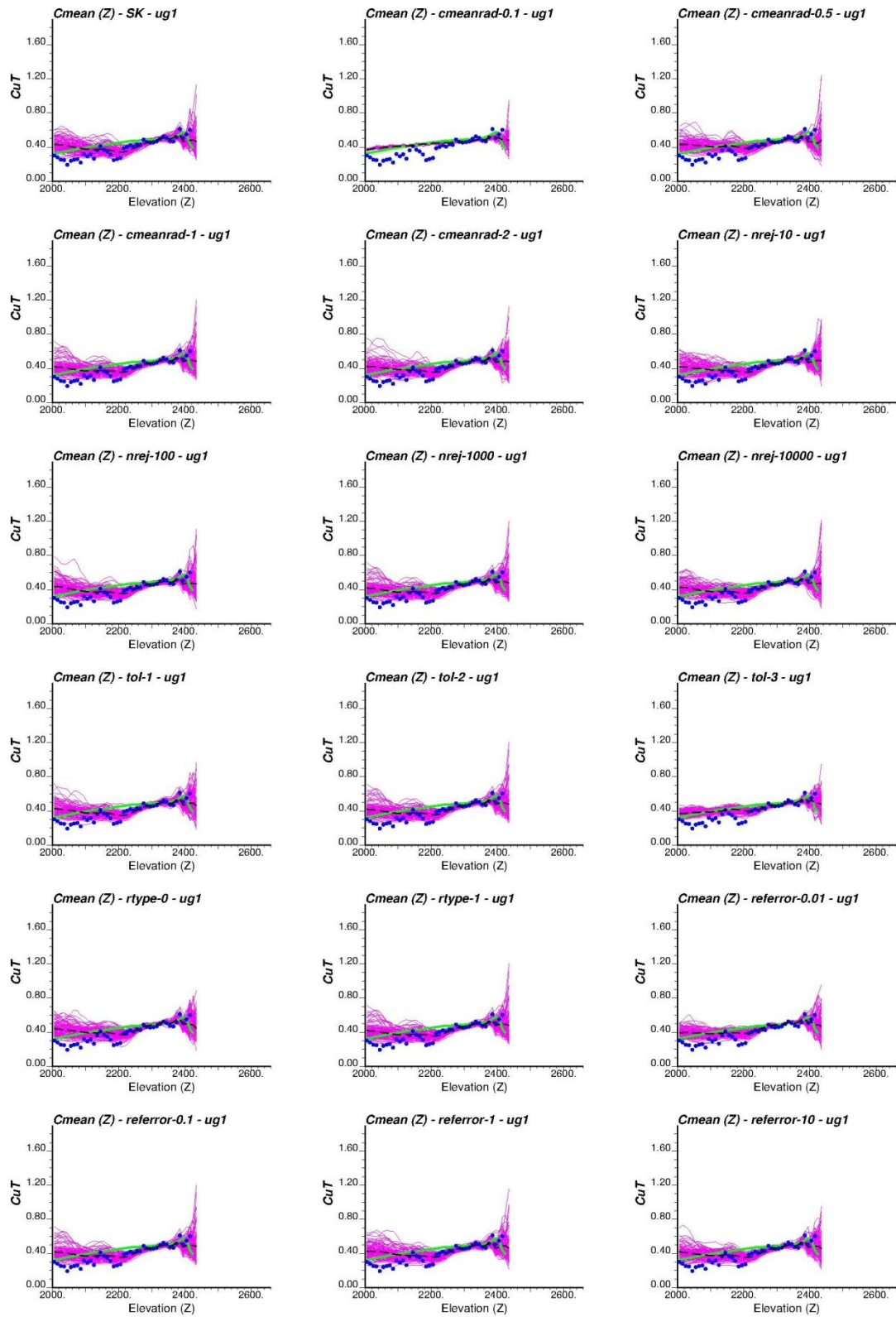
Gráfico 202: Media local – Dirección Y – UE 6.



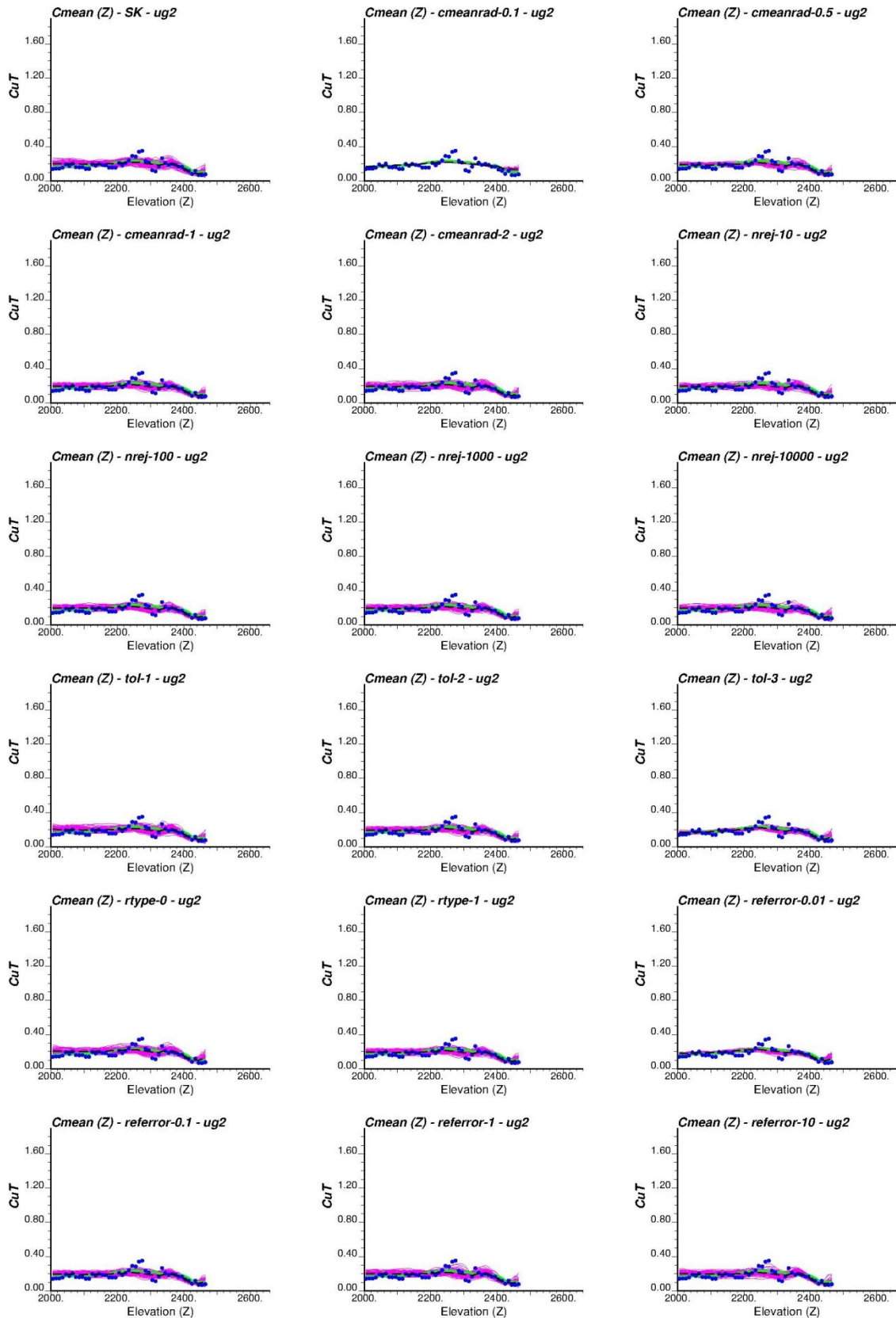
**Gráfico 203: Media local por UE – Dirección Y – Kriging de residuos.**



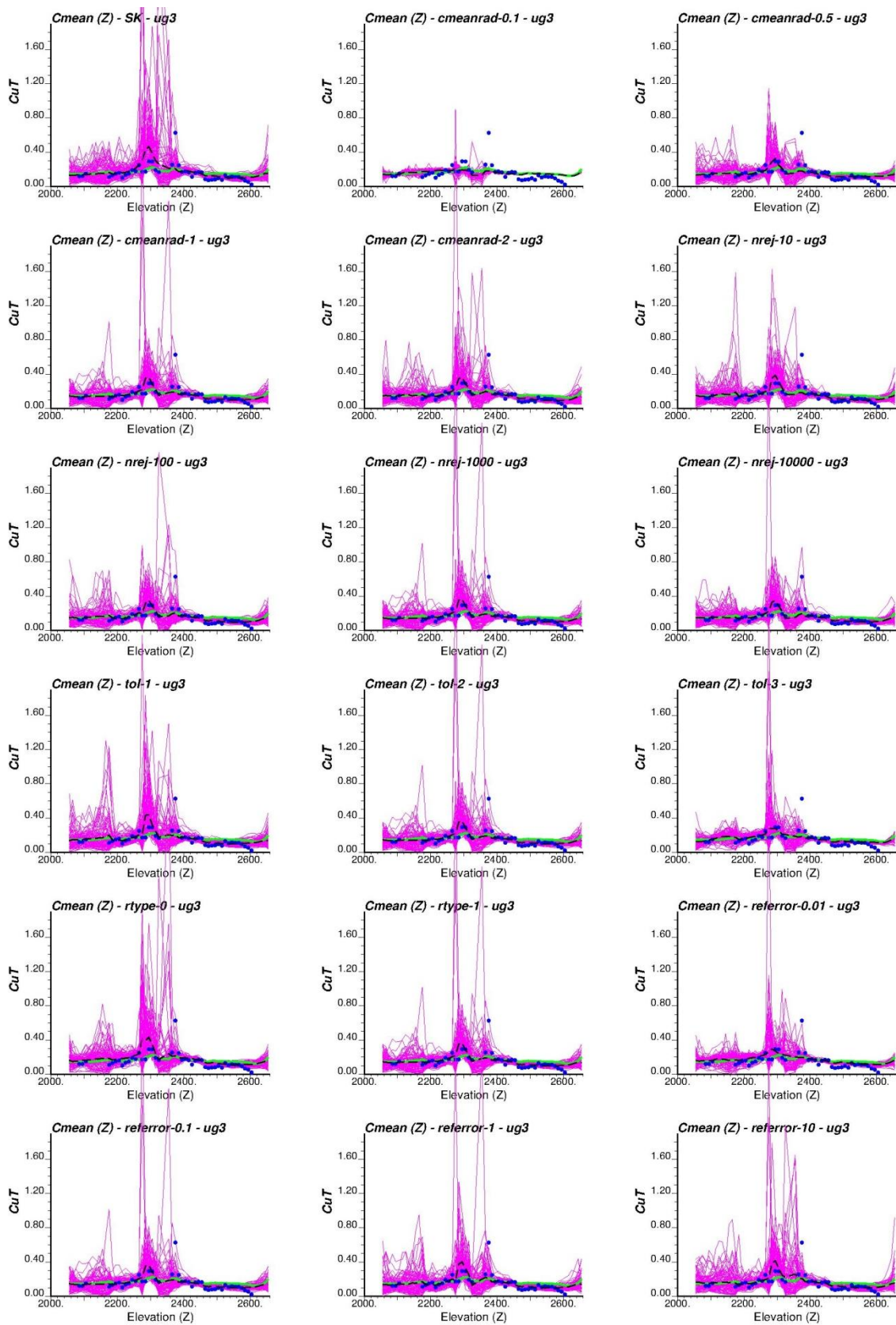
*Dirección Z*



**Gráfico 204: Media local – Dirección Z – UE 1.**



**Gráfico 205: Media local – Dirección Z – UE 2.**



**Gráfico 206: Media local – Dirección Z – UE 3.**

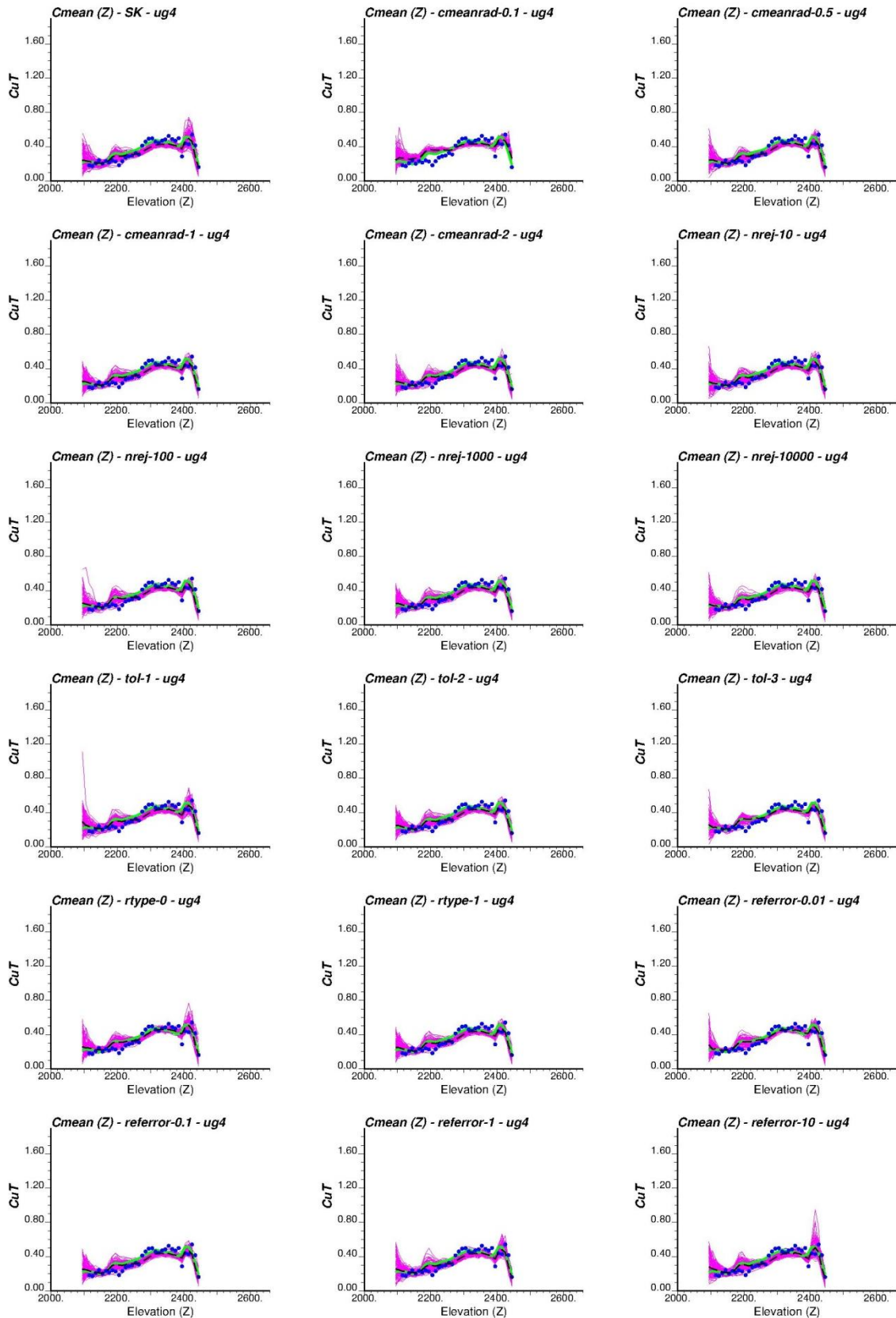
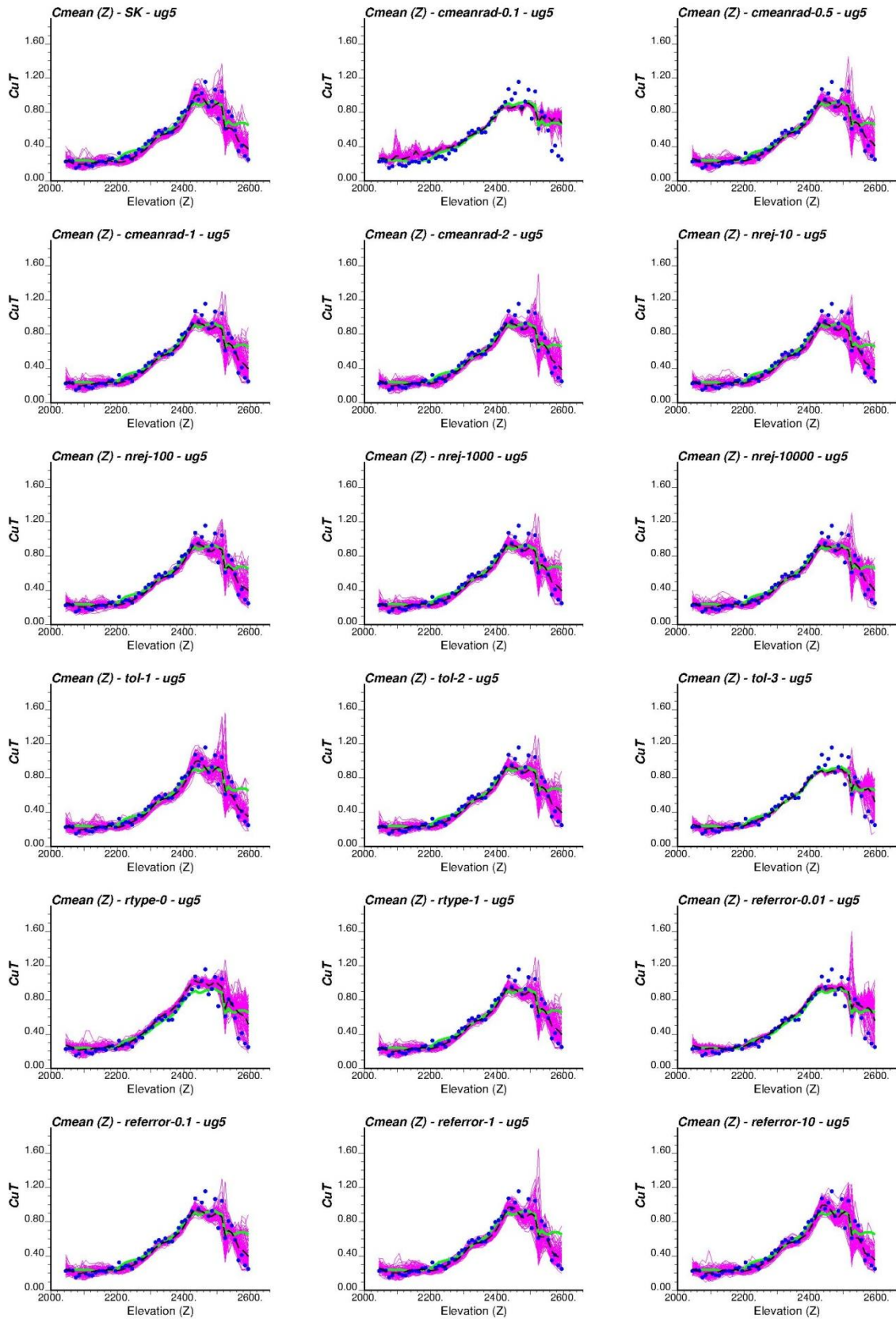


Gráfico 207: Media local – Dirección Z – UE 4.



**Gráfico 208: Media local – Dirección Z – UE 5.**

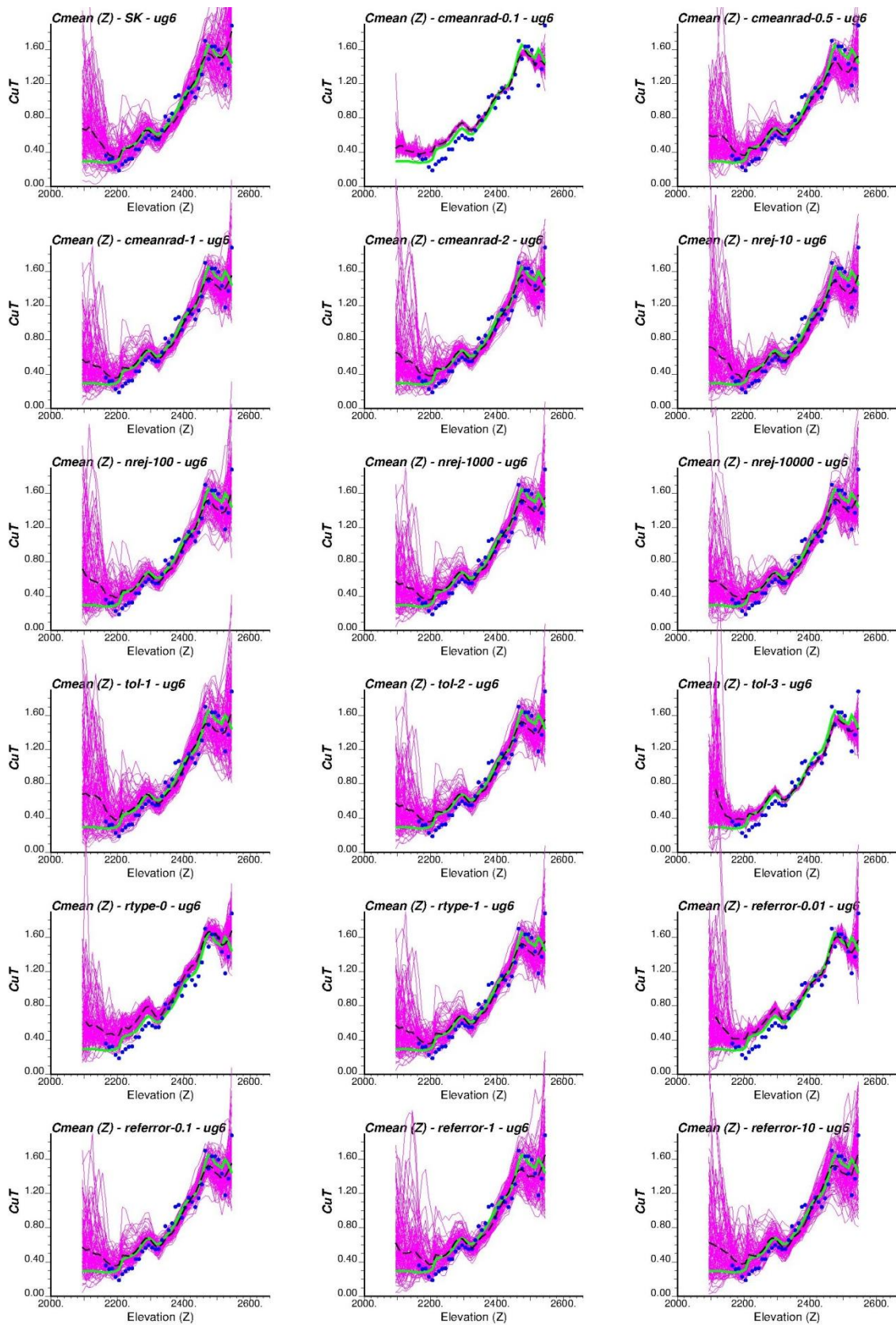
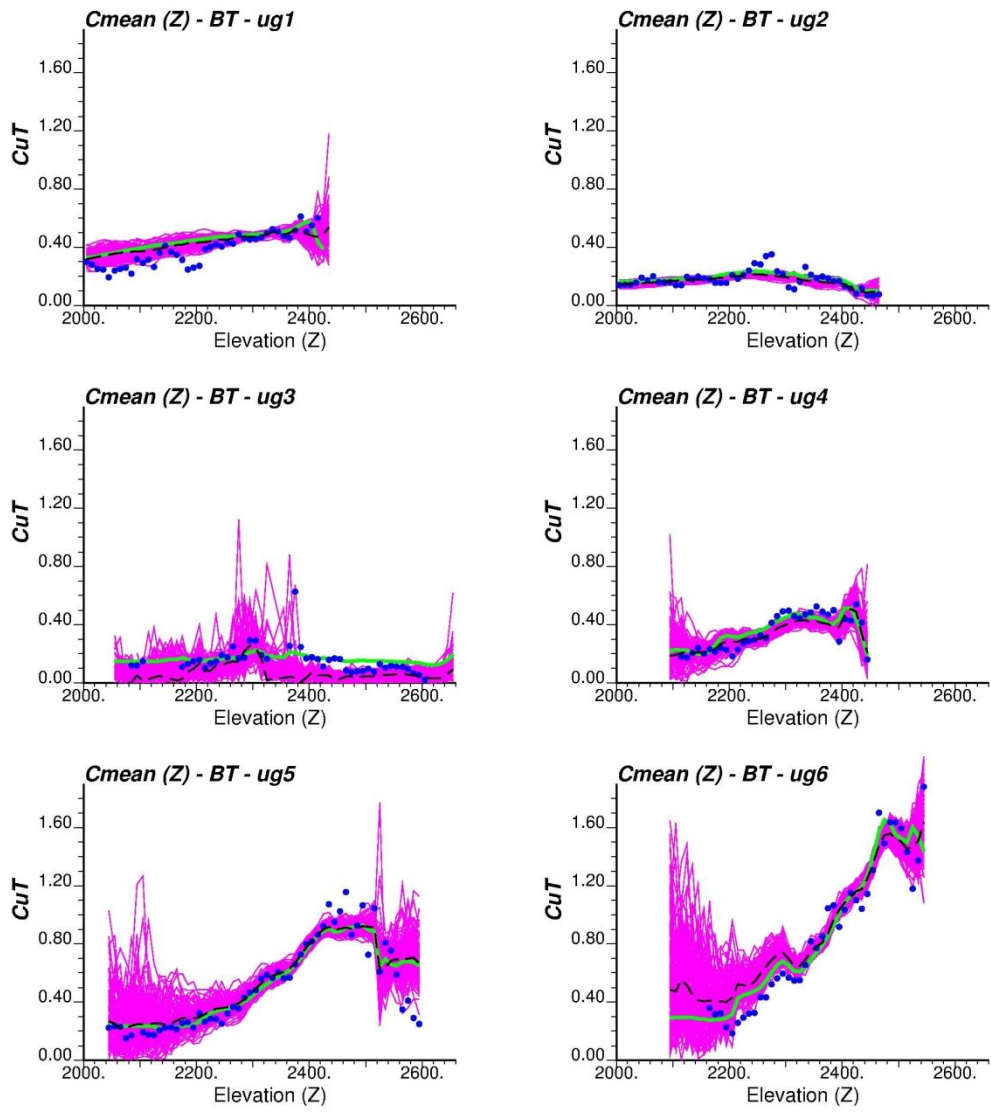


Gráfico 209: Media local – Dirección Z – UE 6.



**Gráfico 210: Media local por UE – Dirección Z – Kriging de residuos.**

## ANEXO N: Caso de estudio II – Coeficiente de determinación ( $R^2$ ) por UE

Coeficiente de determinación ( $R^2$ [%])							
N°	Sensibilización	UE1	UE2	UE3	UE4	UE5	UE6
1	SK	35.224	12.317	13.007	27.156	43.244	26.253
2	Cmean-0.1	48.783	25.950	10.892	34.608	18.963	14.947
3	Cmean-0.5	40.785	21.315	15.788	39.033	41.670	21.275
4	Cmean-1.0	42.121	16.387	15.625	41.019	40.450	24.701
5	Cmean-2.0	37.889	13.580	14.348	44.200	39.684	23.496
6	Nrej-10	42.433	15.170	15.515	39.194	44.112	25.650
7	Nrej-100	37.309	12.913	12.388	41.465	41.058	26.116
8	Nrej-1000	42.121	16.387	15.625	41.019	40.450	24.701
9	Nrej-10000	42.243	16.031	14.213	43.616	42.092	23.089
10	Tol-1	37.258	11.766	12.967	41.326	45.879	25.095
11	Tol-2	42.121	16.387	15.625	41.019	40.450	24.701
12	Tol-3	52.794	26.187	14.113	44.644	20.274	14.177
13	Rtype-0	40.637	13.605	10.172	42.381	34.921	21.933
14	Rtype-1	42.121	16.387	15.625	41.019	40.450	24.701
15	Refferror-0.01	47.818	24.684	13.191	44.316	27.418	16.875
16	Refferror-0.1	42.121	16.387	15.625	41.019	40.450	24.701
17	Refferror-1	38.904	11.254	14.647	40.090	44.362	25.183
18	Refferror-10	39.890	11.807	13.701	38.228	43.829	24.531
19	BT	46.417	22.374	2.067	33.719	25.825	18.472

Tabla 48: Coeficiente de determinación ( $R^2$ ) por UE.

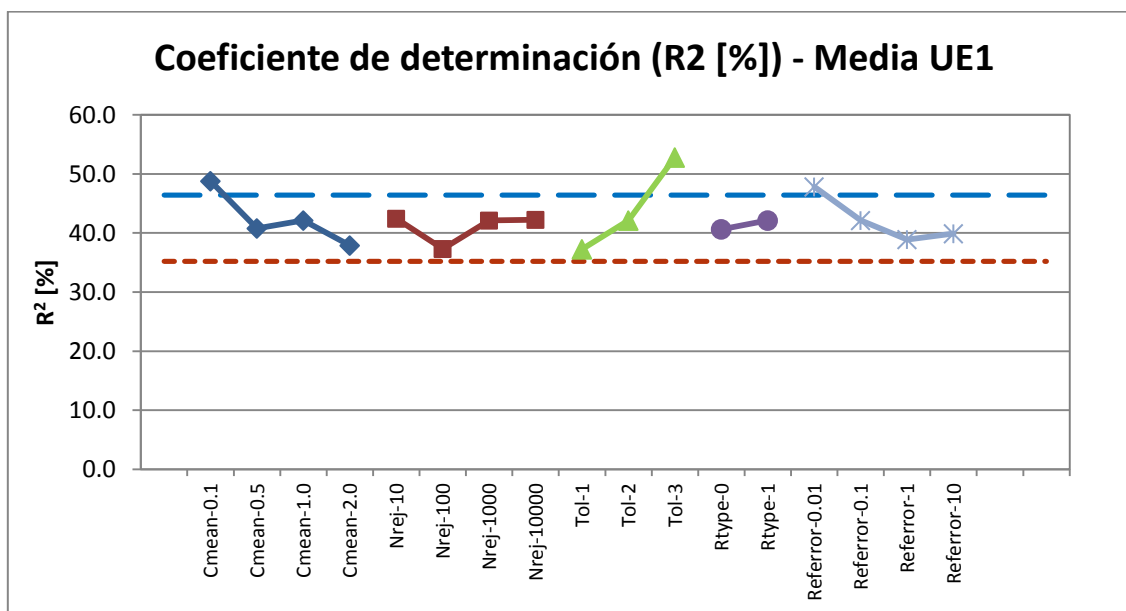
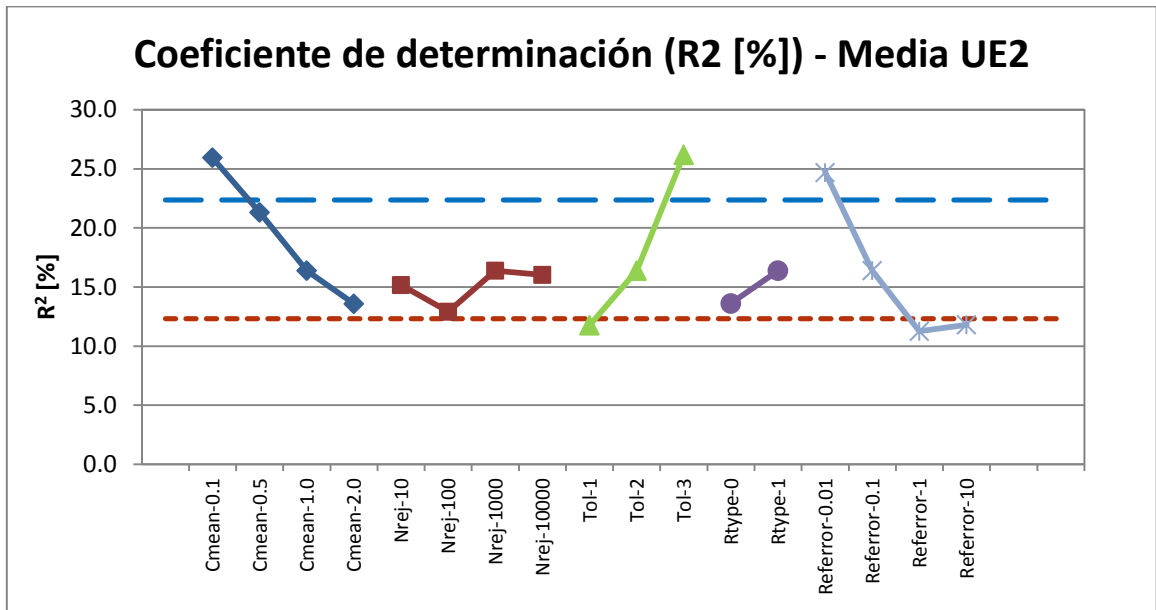
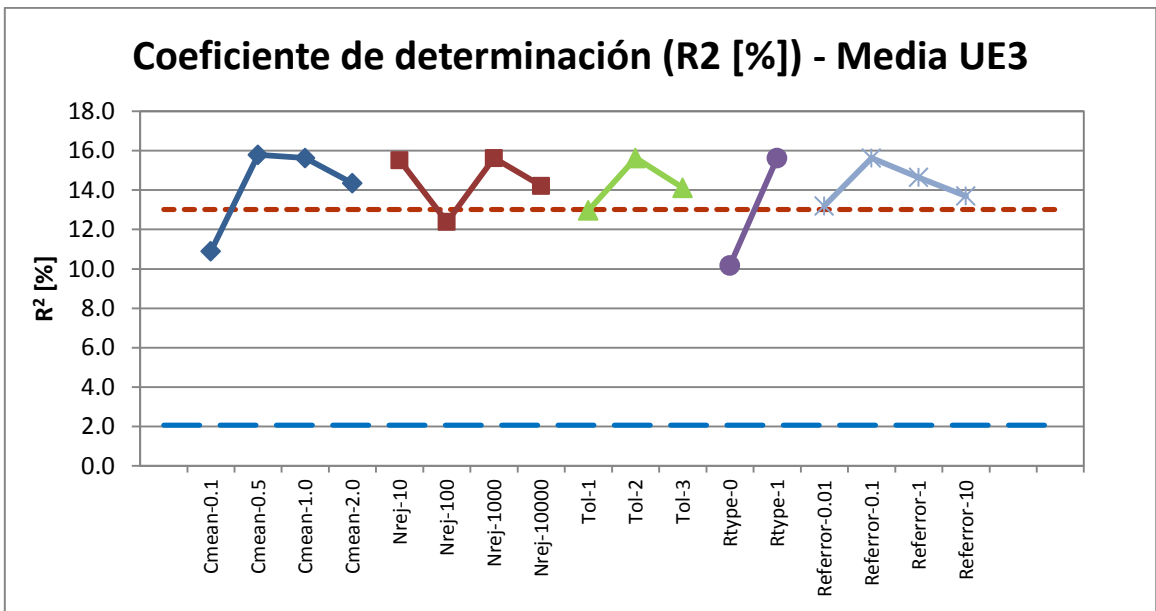


Gráfico 211: Coeficiente de determinación ( $R^2$ ) – Media UE1.

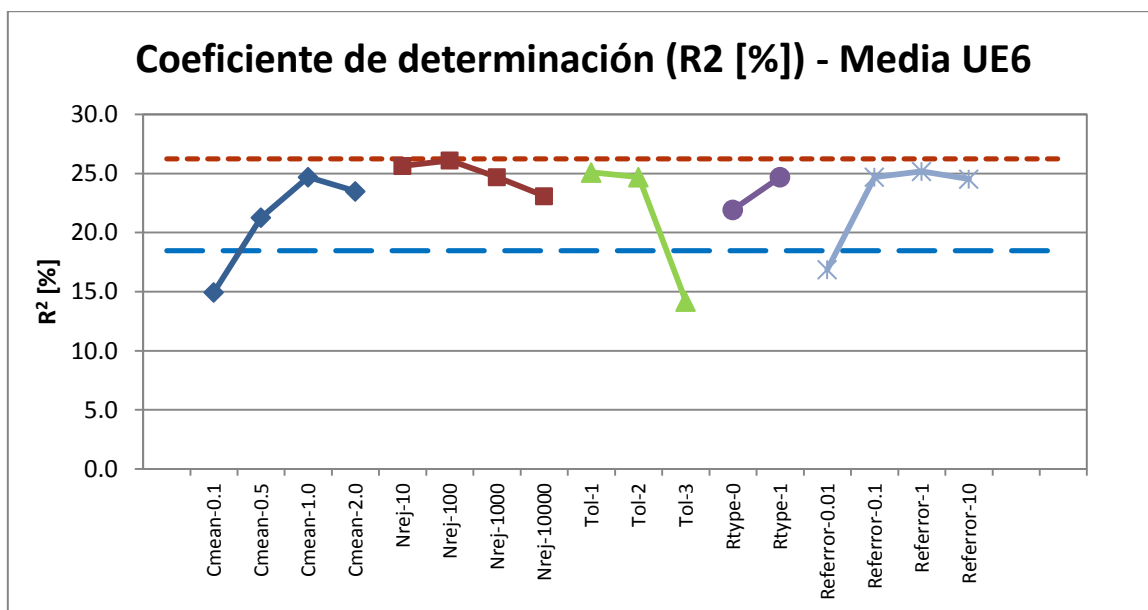




**Gráfico 212: Coeficiente de determinación (R<sup>2</sup>) – Media UE2.**



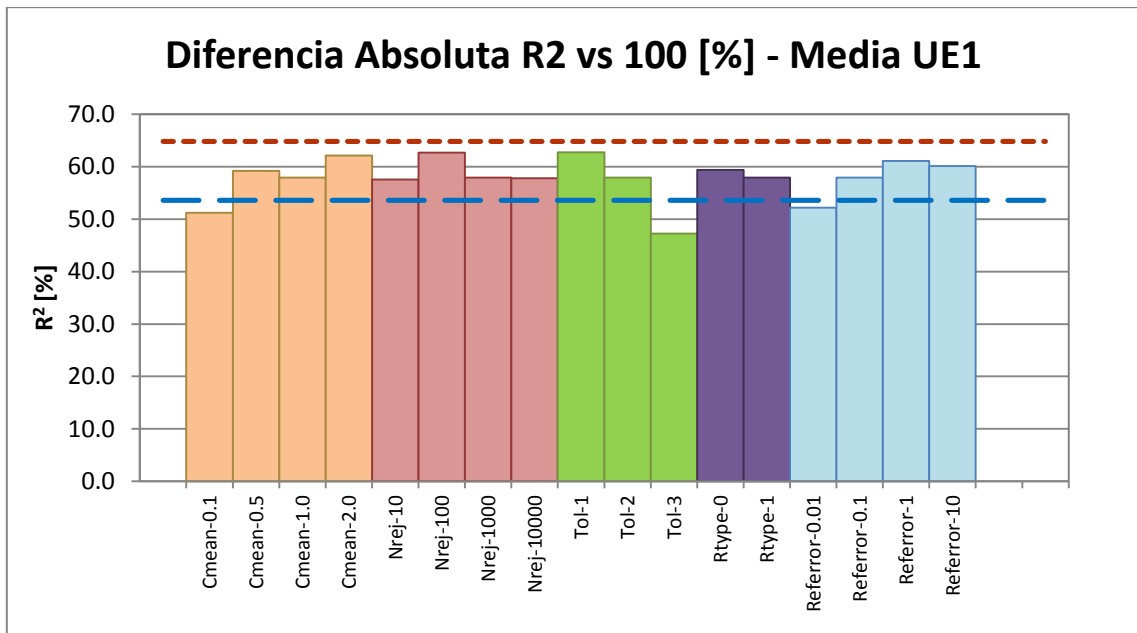
**Gráfico 213: Coeficiente de determinación (R<sup>2</sup>) – Media UE3.**



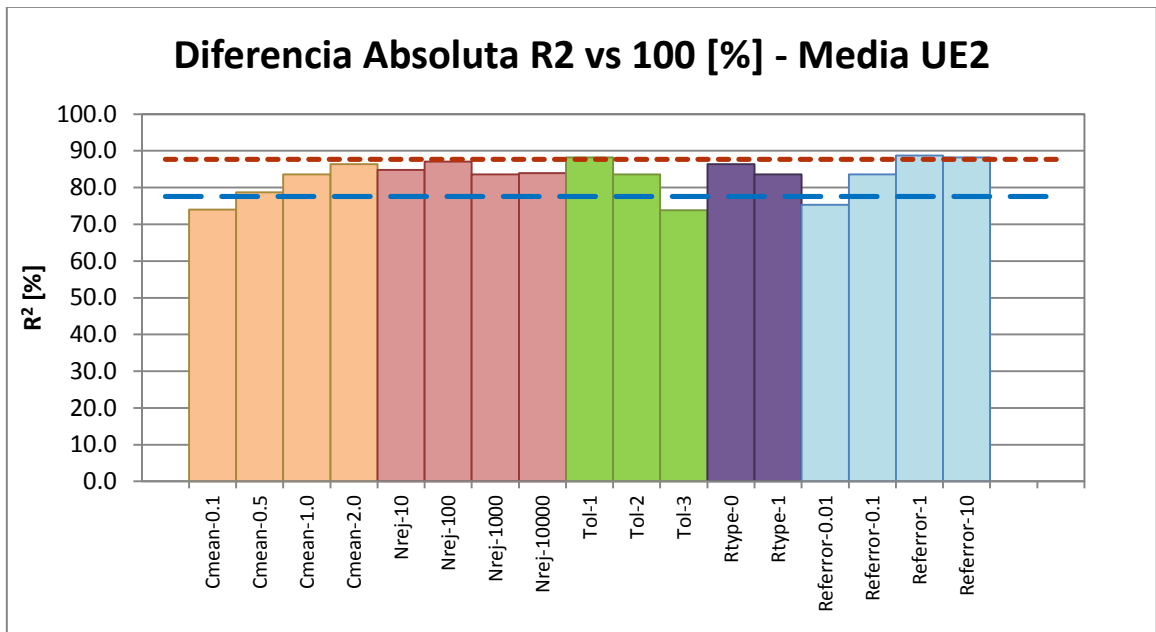
**Gráfico 214: Coeficiente de determinación (R<sup>2</sup>) – Media UE6.**

Diferencia Absoluta R2 vs 100 [%]							
Nº	Sensibilización	UE1	UE2	UE3	UE4	UE5	UE6
1	SK	64.776	87.683	86.993	72.844	56.756	73.747
2	Cmean-0.1	51.217	74.050	89.108	65.392	81.037	85.053
3	Cmean-0.5	59.215	78.685	84.212	60.967	58.330	78.725
4	Cmean-1.0	57.879	83.613	84.375	58.981	59.550	75.299
5	Cmean-2.0	62.111	86.420	85.652	55.800	60.316	76.504
6	Nrej-10	57.567	84.830	84.485	60.806	55.888	74.350
7	Nrej-100	62.691	87.087	87.612	58.535	58.942	73.884
8	Nrej-1000	57.879	83.613	84.375	58.981	59.550	75.299
9	Nrej-10000	57.757	83.969	85.787	56.384	57.908	76.911
10	Tol-1	62.742	88.234	87.033	58.674	54.121	74.905
11	Tol-2	57.879	83.613	84.375	58.981	59.550	75.299
12	Tol-3	47.206	73.813	85.887	55.356	79.726	85.823
13	Rtype-0	59.363	86.395	89.828	57.619	65.079	78.067
14	Rtype-1	57.879	83.613	84.375	58.981	59.550	75.299
15	Referror-0.01	52.182	75.316	86.809	55.684	72.582	83.125
16	Referror-0.1	57.879	83.613	84.375	58.981	59.550	75.299
17	Referror-1	61.096	88.746	85.353	59.910	55.638	74.817
18	Referror-10	60.110	88.193	86.299	61.772	56.171	75.469
19	BT	53.583	77.626	97.933	66.281	74.175	81.528

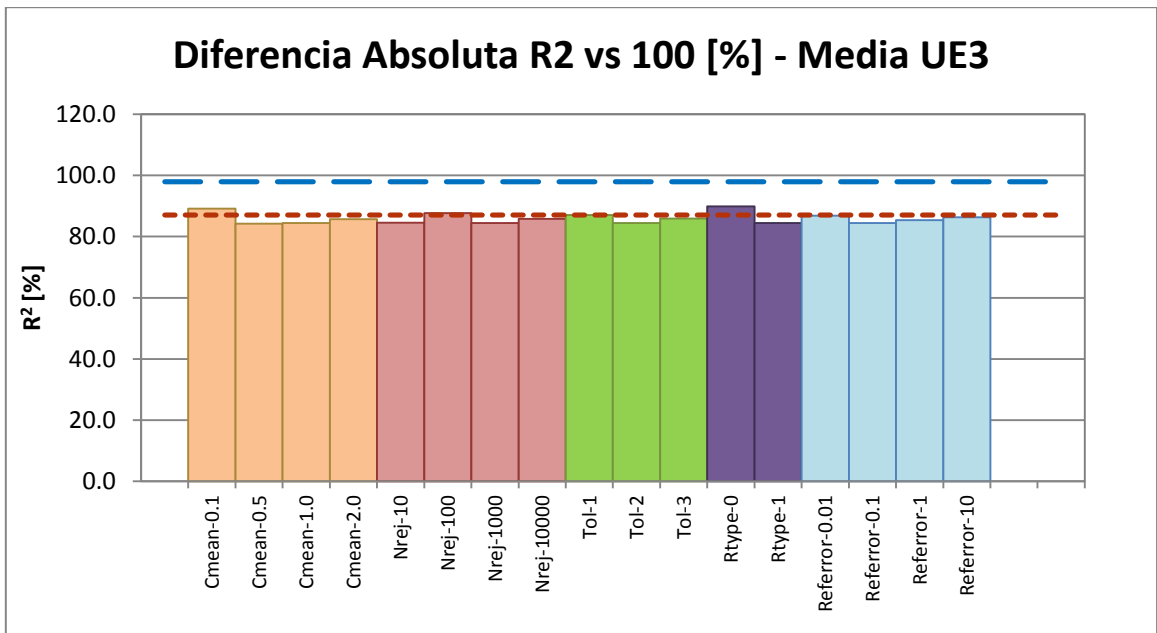
**Tabla 49: Diferencia absoluta R2 vs 100 [%] por UE.**



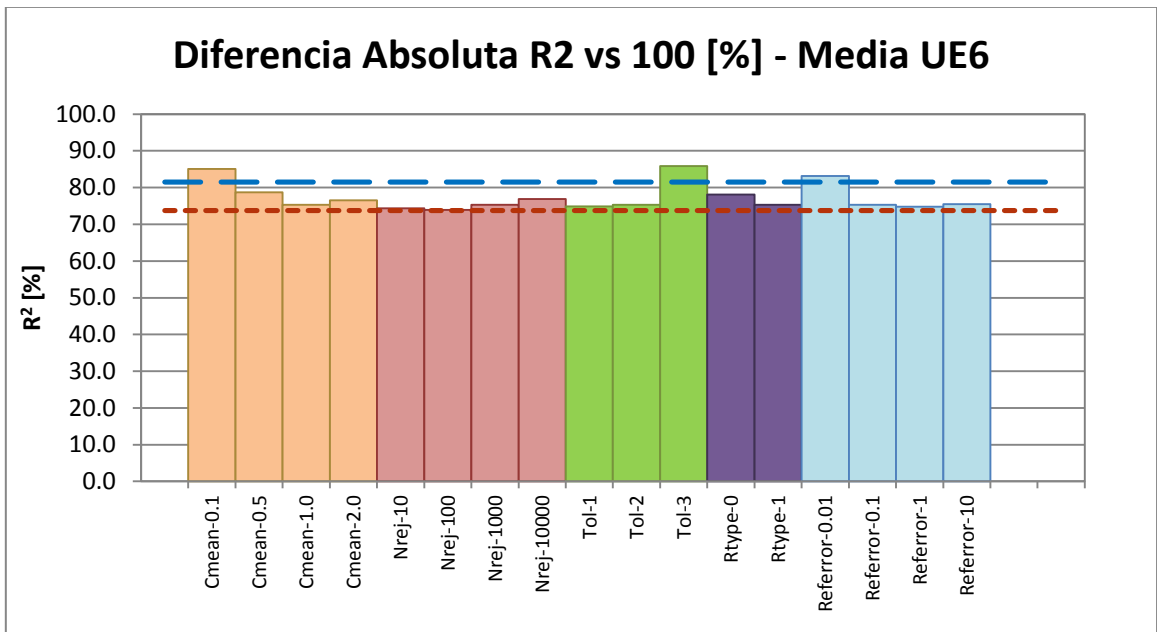
**Gráfico 215: Diferencia Absoluta R2 vs 100 [%] – Media UE1.**



**Gráfico 216: Diferencia Absoluta R2 vs 100 [%] – Media UE2.**



**Gráfico 217: Diferencia Absoluta R2 vs 100 [%] – Media UE3.**

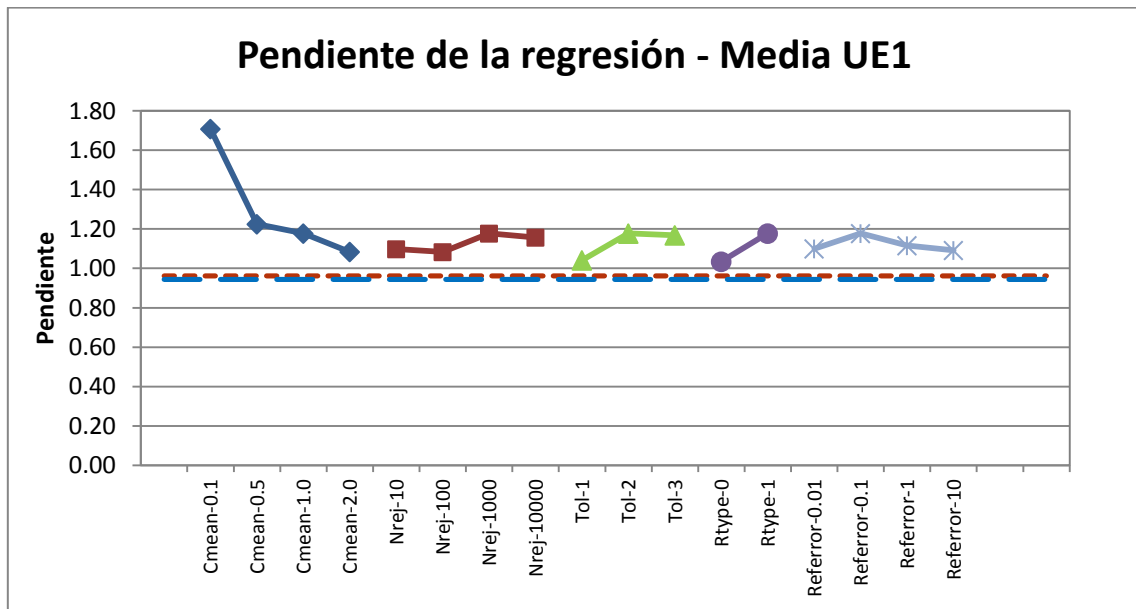


**Gráfico 218: Diferencia Absoluta R2 vs 100 [%] – Media UE6.**

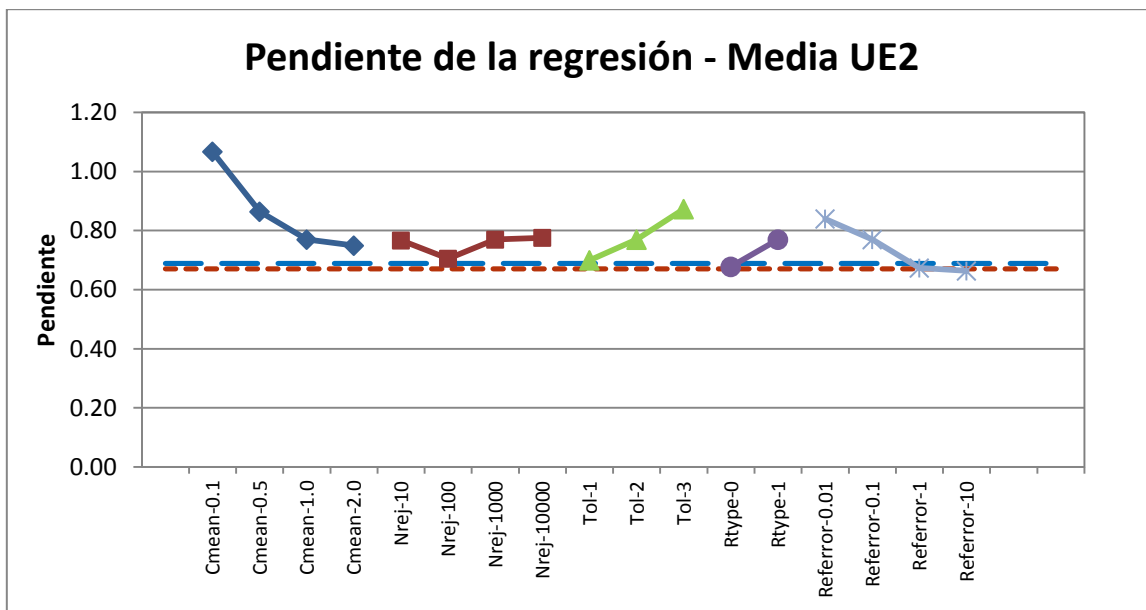
**ANEXO O: Caso de estudio II – Pendiente de la regresión por UE**

Pendiente de la regresión							
N°	Sensibilización	UE1	UE2	UE3	UE4	UE5	UE6
1	SK	0.962	0.670	0.835	0.950	1.187	0.920
2	Cmean-0.1	1.708	1.067	0.953	1.202	0.799	0.686
3	Cmean-0.5	1.224	0.864	1.175	1.012	1.195	0.914
4	Cmean-1.0	1.177	0.770	1.082	1.022	1.201	0.926
5	Cmean-2.0	1.084	0.749	1.012	1.096	1.186	0.891
6	Nrej-10	1.098	0.767	1.093	1.052	1.265	0.962
7	Nrej-100	1.083	0.705	0.948	1.060	1.200	0.950
8	Nrej-1000	1.177	0.770	1.082	1.022	1.201	0.926
9	Nrej-10000	1.157	0.776	1.032	1.066	1.195	0.873
10	Tol-1	1.041	0.700	0.919	1.060	1.229	0.957
11	Tol-2	1.177	0.770	1.082	1.022	1.201	0.926
12	Tol-3	1.168	0.873	0.933	1.016	0.719	0.640
13	Rtype-0	1.034	0.678	0.675	0.895	0.997	0.802
14	Rtype-1	1.177	0.770	1.082	1.022	1.201	0.926
15	Referror-0.01	1.100	0.840	0.779	0.966	0.861	0.726
16	Referror-0.1	1.177	0.770	1.082	1.022	1.201	0.926
17	Referror-1	1.116	0.674	0.946	1.019	1.265	0.946
18	Referror-10	1.091	0.664	0.923	0.913	1.236	0.903
19	BT	0.944	0.689	0.251	0.763	0.902	0.797

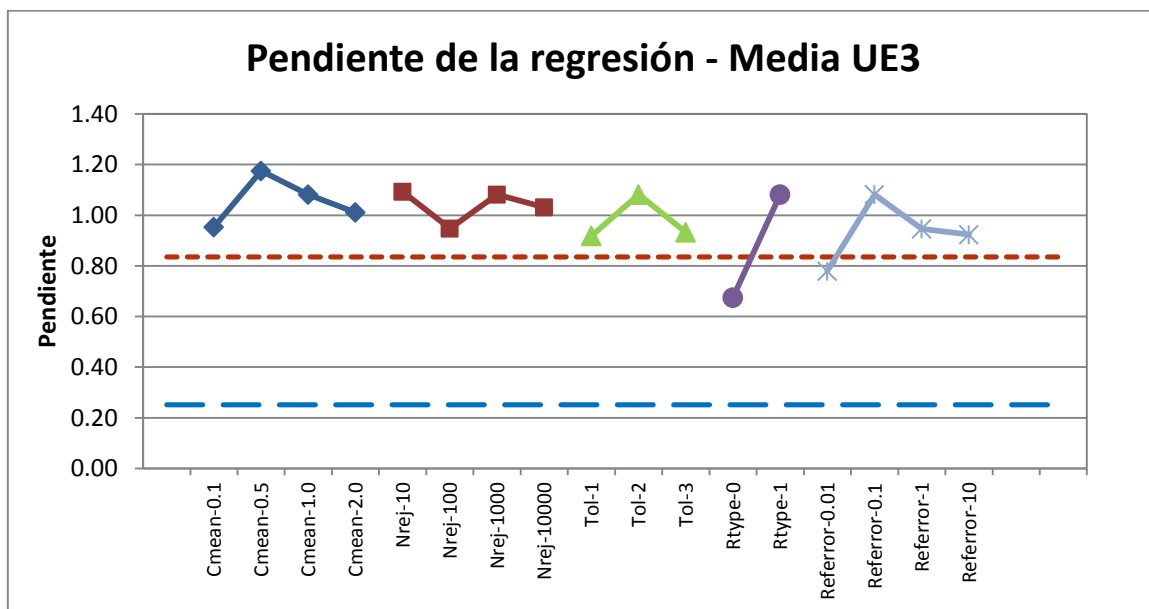
**Tabla 50: Pendiente de la regresión por UE.**



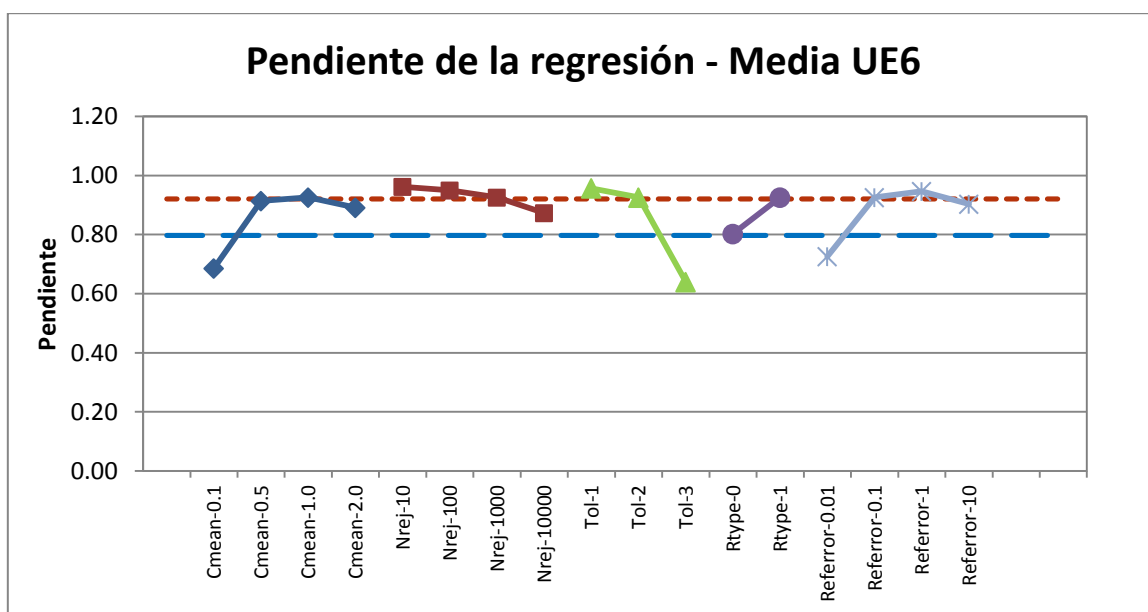
**Gráfico 219: Pendiente de la regresión – Media UE1.**



**Gráfico 220: Pendiente de la regresión – Media UE2.**



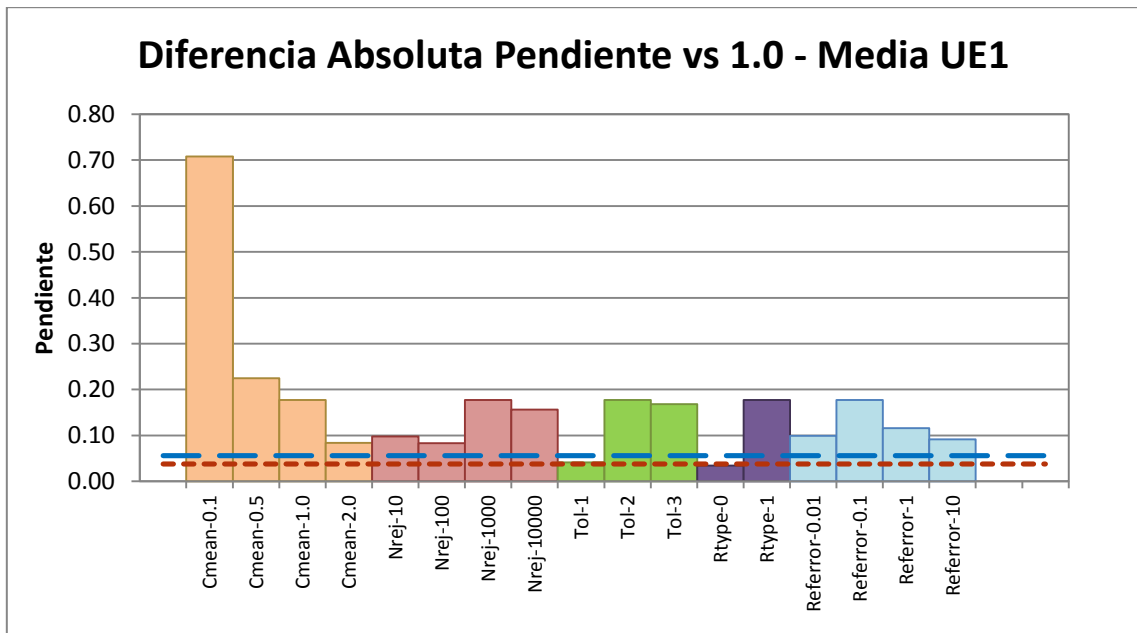
**Gráfico 221: Pendiente de la regresión – Media UE3.**



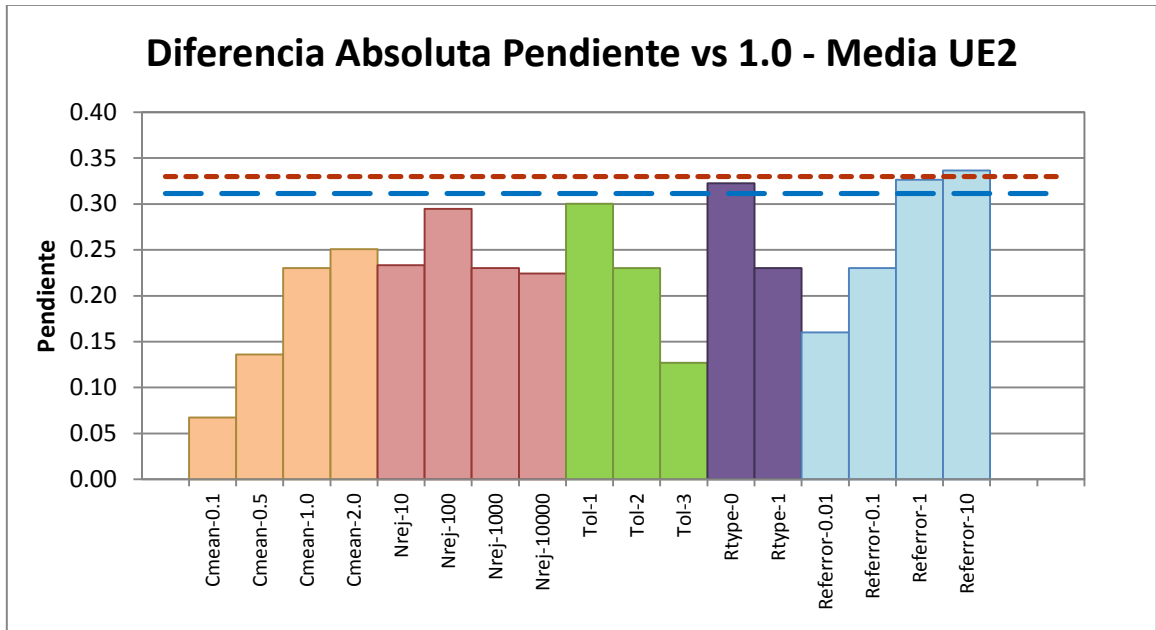
**Gráfico 222: Pendiente de la regresión – Media UE6.**

Diferencia Absoluta Pendiente vs 1.0							
Nº	Sensibilización	UE1	UE2	UE3	UE4	UE5	UE6
1	SK	0.038	0.330	0.165	0.050	0.187	0.080
2	Cmean-0.1	0.708	0.067	0.047	0.202	0.201	0.314
3	Cmean-0.5	0.224	0.136	0.175	0.012	0.195	0.086
4	Cmean-1.0	0.177	0.230	0.082	0.022	0.201	0.074
5	Cmean-2.0	0.084	0.251	0.012	0.096	0.186	0.109
6	Nrej-10	0.098	0.233	0.093	0.052	0.265	0.038
7	Nrej-100	0.083	0.295	0.052	0.060	0.200	0.050
8	Nrej-1000	0.177	0.230	0.082	0.022	0.201	0.074
9	Nrej-10000	0.157	0.224	0.032	0.066	0.195	0.127
10	Tol-1	0.041	0.300	0.081	0.060	0.229	0.043
11	Tol-2	0.177	0.230	0.082	0.022	0.201	0.074
12	Tol-3	0.168	0.127	0.067	0.016	0.281	0.360
13	Rtype-0	0.034	0.322	0.325	0.105	0.003	0.198
14	Rtype-1	0.177	0.230	0.082	0.022	0.201	0.074
15	Referror-0.01	0.100	0.160	0.221	0.034	0.139	0.274
16	Referror-0.1	0.177	0.230	0.082	0.022	0.201	0.074
17	Referror-1	0.116	0.326	0.054	0.019	0.265	0.054
18	Referror-10	0.091	0.336	0.077	0.087	0.236	0.097
19	BT	0.056	0.311	0.749	0.237	0.098	0.203

**Tabla 51: Diferencia absoluta pendiente vs 1.0 por UE.**

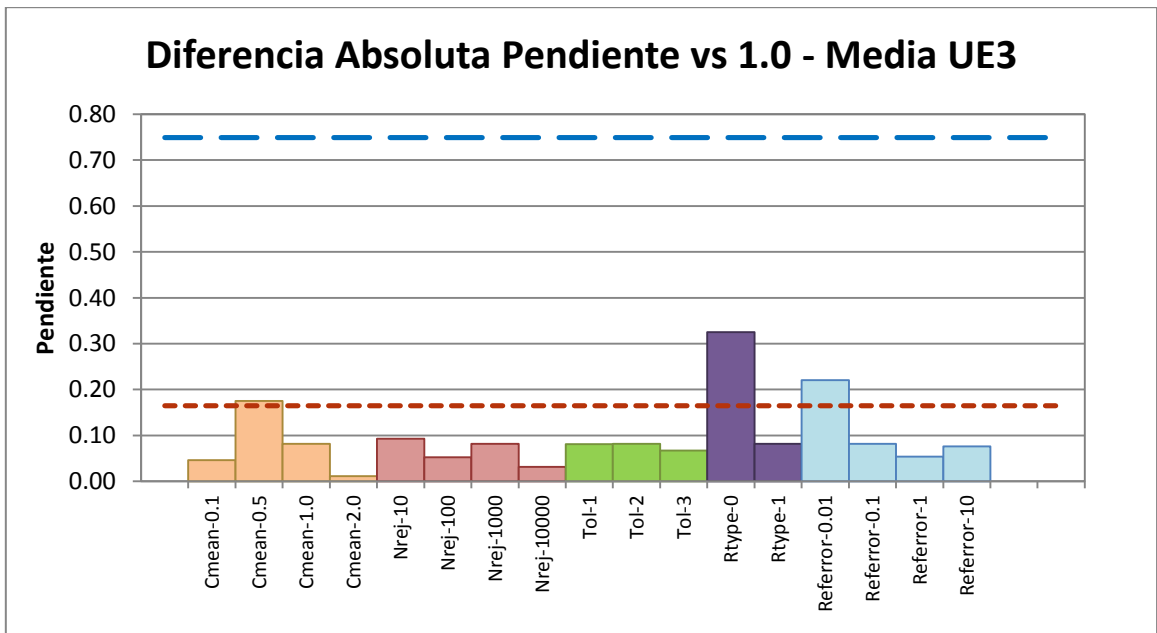


**Gráfico 223: Diferencia absoluta pendiente vs 1.0 – Media UE1.**

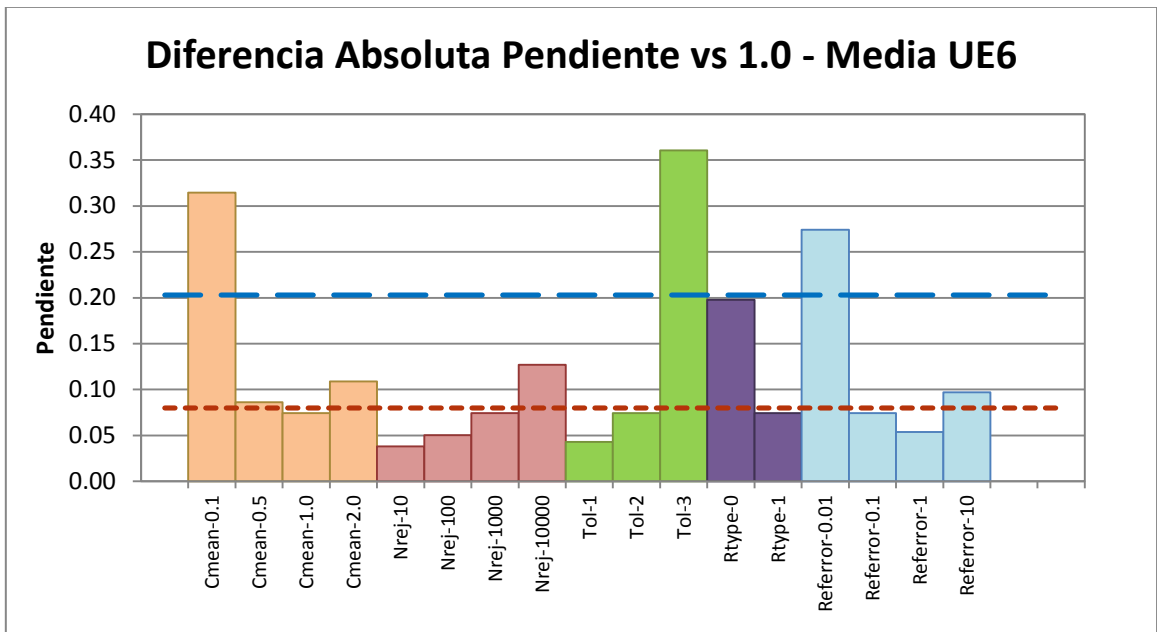


**Gráfico 224: Diferencia absoluta pendiente vs 1.0 – Media UE2.**





**Gráfico 225: Diferencia absoluta pendiente vs 1.0 – Media UE3.**

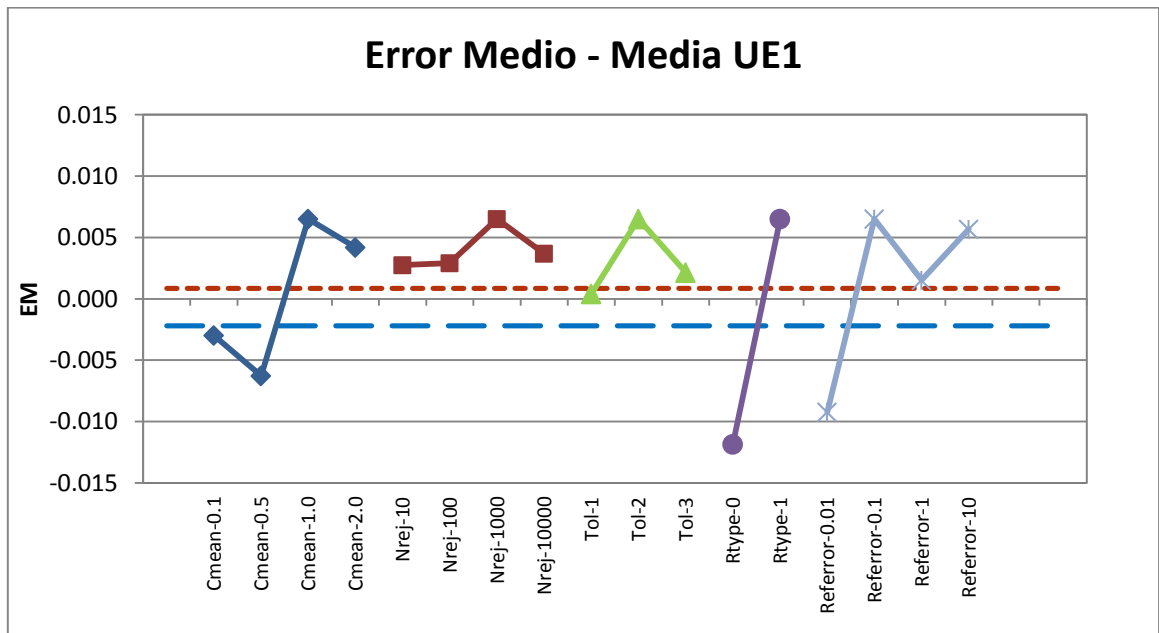


**Gráfico 226: Diferencia absoluta pendiente vs 1.0 – Media UE6.**

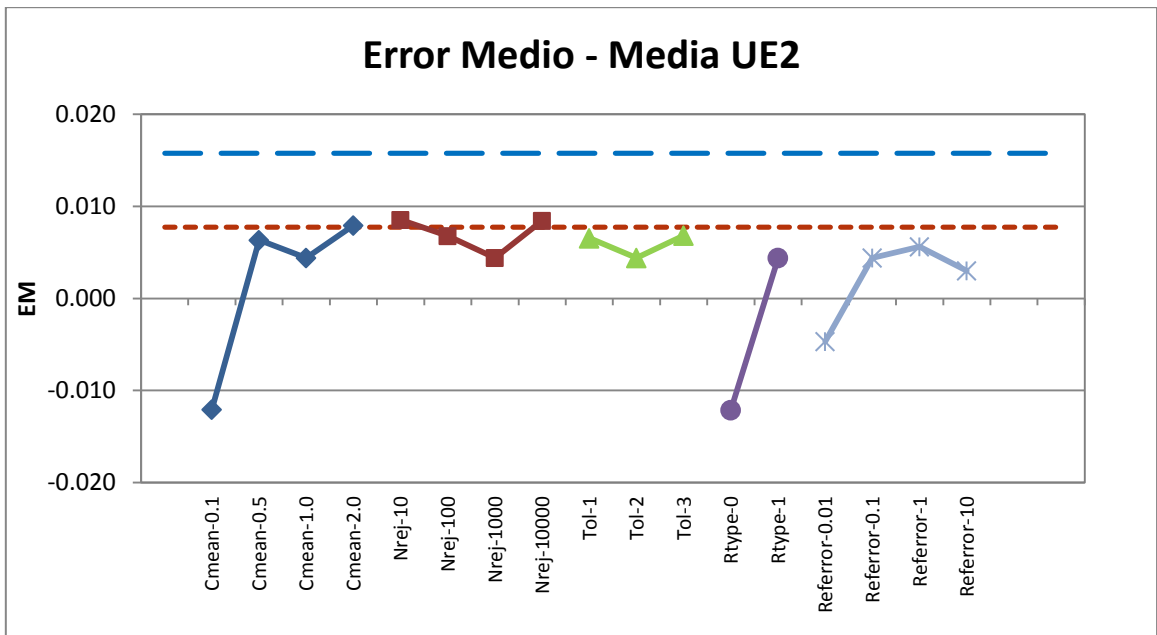
**ANEXO P: Caso de estudio II – Error medio por UE**

Error Medio							
N°	Sensibilización	UE1	UE2	UE3	UE4	UE5	UE6
1	SK	0.001	0.008	0.005	0.052	0.067	0.199
2	Cmean-0.1	-0.003	-0.012	0.004	0.040	0.025	0.066
3	Cmean-0.5	-0.006	0.006	0.015	0.032	0.050	0.186
4	Cmean-1.0	0.007	0.004	0.017	0.031	0.049	0.165
5	Cmean-2.0	0.004	0.008	0.017	0.028	0.051	0.168
6	Nrej-10	0.003	0.009	0.017	0.039	0.058	0.187
7	Nrej-100	0.003	0.007	0.015	0.033	0.057	0.182
8	Nrej-1000	0.007	0.004	0.017	0.031	0.049	0.165
9	Nrej-10000	0.004	0.008	0.015	0.037	0.047	0.168
10	Tol-1	0.000	0.007	0.013	0.032	0.053	0.197
11	Tol-2	0.007	0.004	0.017	0.031	0.049	0.165
12	Tol-3	0.002	0.007	0.014	0.026	0.013	0.094
13	Rtype-0	-0.012	-0.012	-0.004	0.003	-0.046	0.019
14	Rtype-1	0.007	0.004	0.017	0.031	0.049	0.165
15	Referror-0.01	-0.009	-0.005	0.001	0.010	-0.027	0.023
16	Referror-0.1	0.007	0.004	0.017	0.031	0.049	0.165
17	Referror-1	0.002	0.006	0.011	0.034	0.063	0.195
18	Referror-10	0.006	0.003	0.011	0.026	0.063	0.187
19	BT	-0.002	0.016	0.072	0.034	0.016	0.064

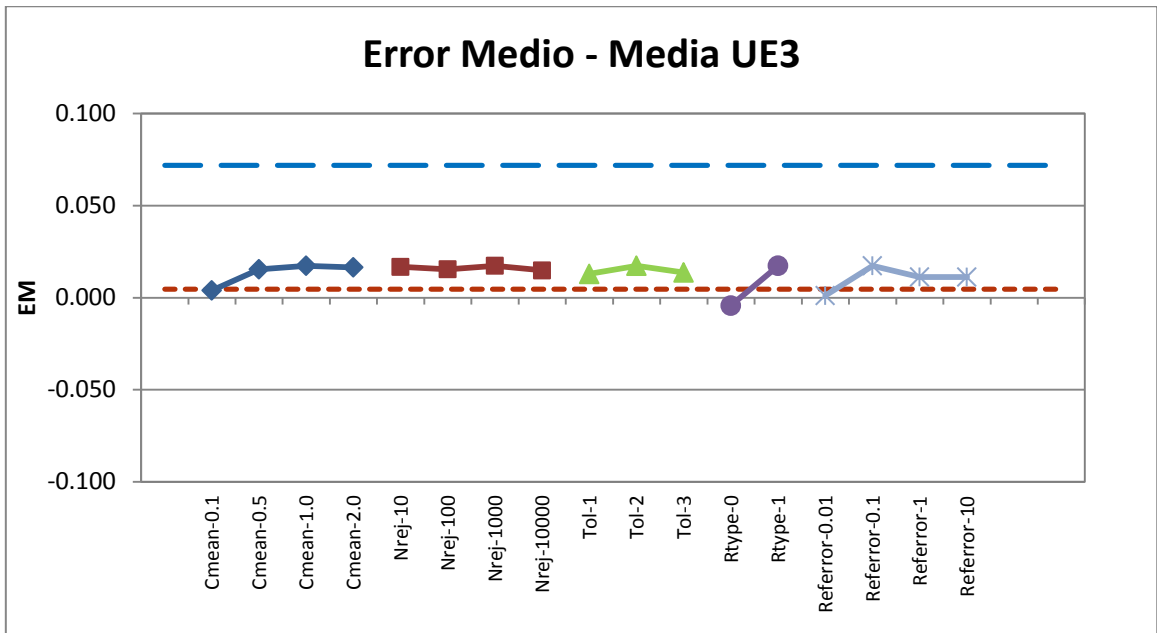
**Tabla 52: Error medio por UE.**



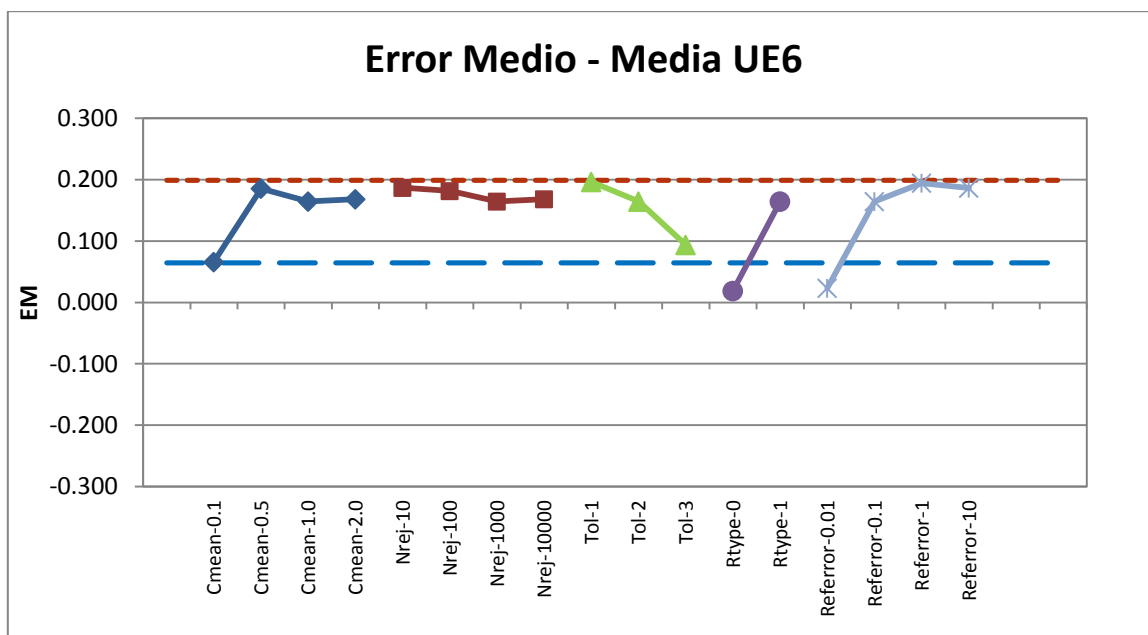
**Gráfico 227: Error medio – Media UE1.**



**Gráfico 228: Error medio – Media UE2.**



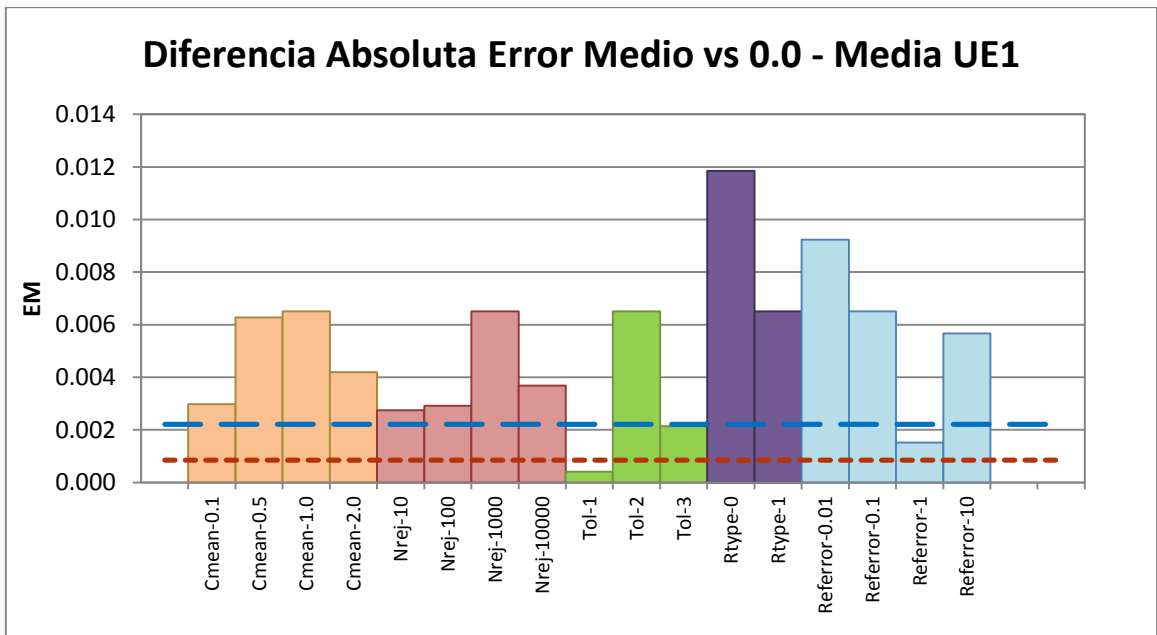
**Gráfico 229: Error medio – Media UE3.**



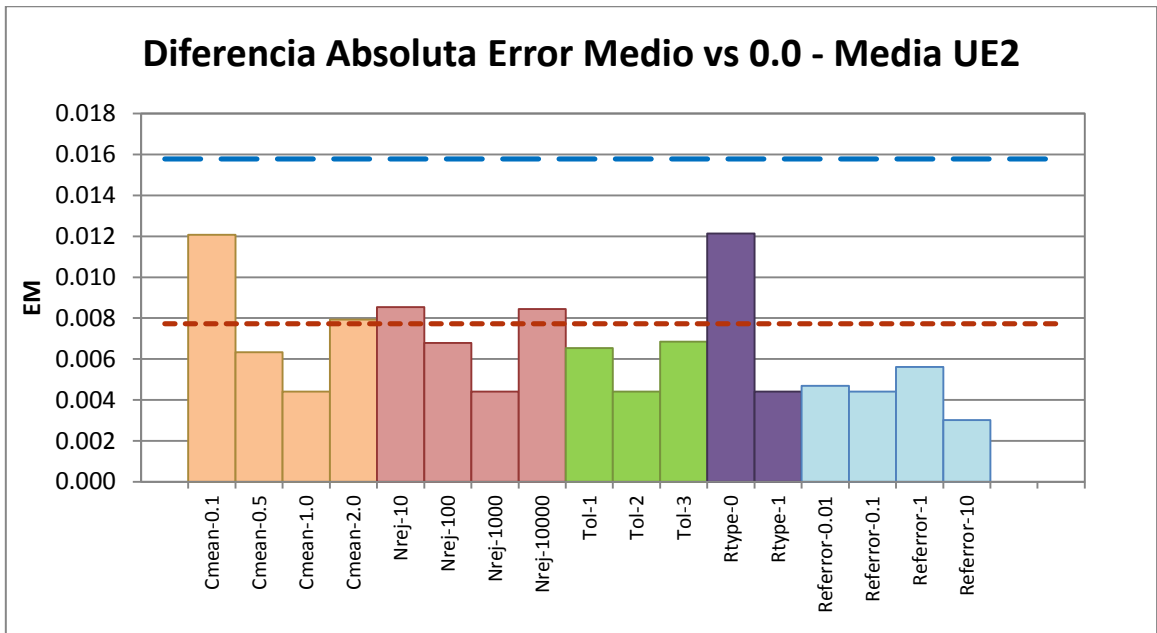
**Gráfico 230: Error medio – Media UE6.**

Diferencia Absoluta Error Medio vs 0.0							
N°	Sensibilización	UE1	UE2	UE3	UE4	UE5	UE6
1	SK	0.001	0.008	0.005	0.052	0.067	0.199
2	Cmean-0.1	0.003	0.012	0.004	0.040	0.025	0.066
3	Cmean-0.5	0.006	0.006	0.015	0.032	0.050	0.186
4	Cmean-1.0	0.007	0.004	0.017	0.031	0.049	0.165
5	Cmean-2.0	0.004	0.008	0.017	0.028	0.051	0.168
6	Nrej-10	0.003	0.009	0.017	0.039	0.058	0.187
7	Nrej-100	0.003	0.007	0.015	0.033	0.057	0.182
8	Nrej-1000	0.007	0.004	0.017	0.031	0.049	0.165
9	Nrej-10000	0.004	0.008	0.015	0.037	0.047	0.168
10	Tol-1	0.000	0.007	0.013	0.032	0.053	0.197
11	Tol-2	0.007	0.004	0.017	0.031	0.049	0.165
12	Tol-3	0.002	0.007	0.014	0.026	0.013	0.094
13	Rtype-0	0.012	0.012	0.004	0.003	0.046	0.019
14	Rtype-1	0.007	0.004	0.017	0.031	0.049	0.165
15	Referror-0.01	0.009	0.005	0.001	0.010	0.027	0.023
16	Referror-0.1	0.007	0.004	0.017	0.031	0.049	0.165
17	Referror-1	0.002	0.006	0.011	0.034	0.063	0.195
18	Referror-10	0.006	0.003	0.011	0.026	0.063	0.187
19	BT	0.002	0.016	0.072	0.034	0.016	0.064

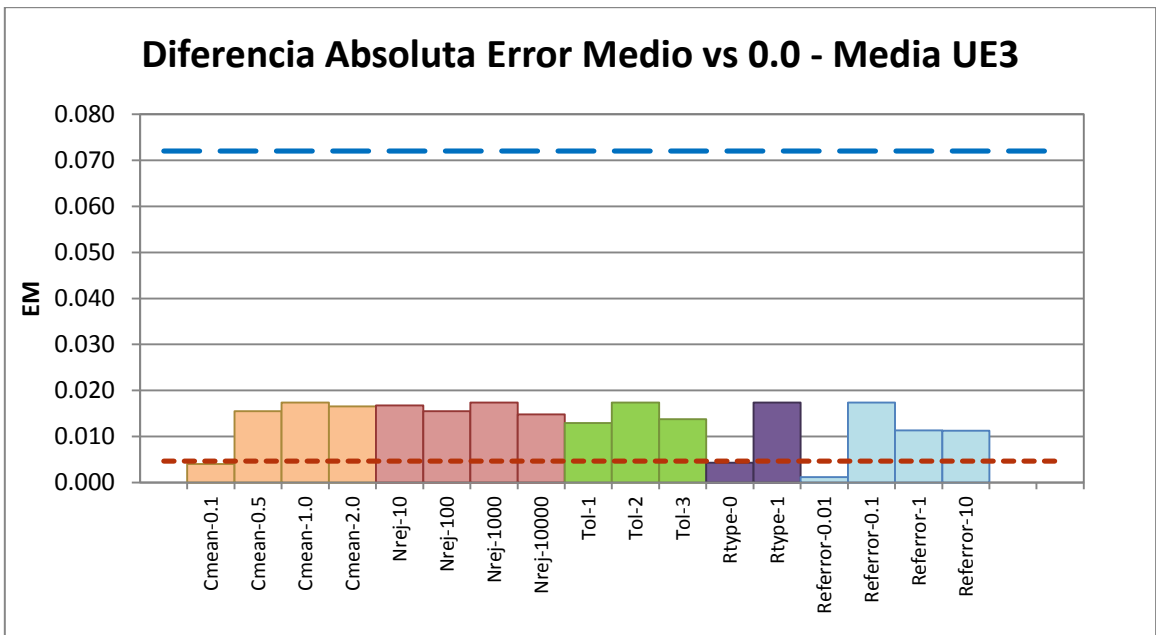
**Tabla 53: Diferencia absoluta error medio vs 0.0 por UE.**



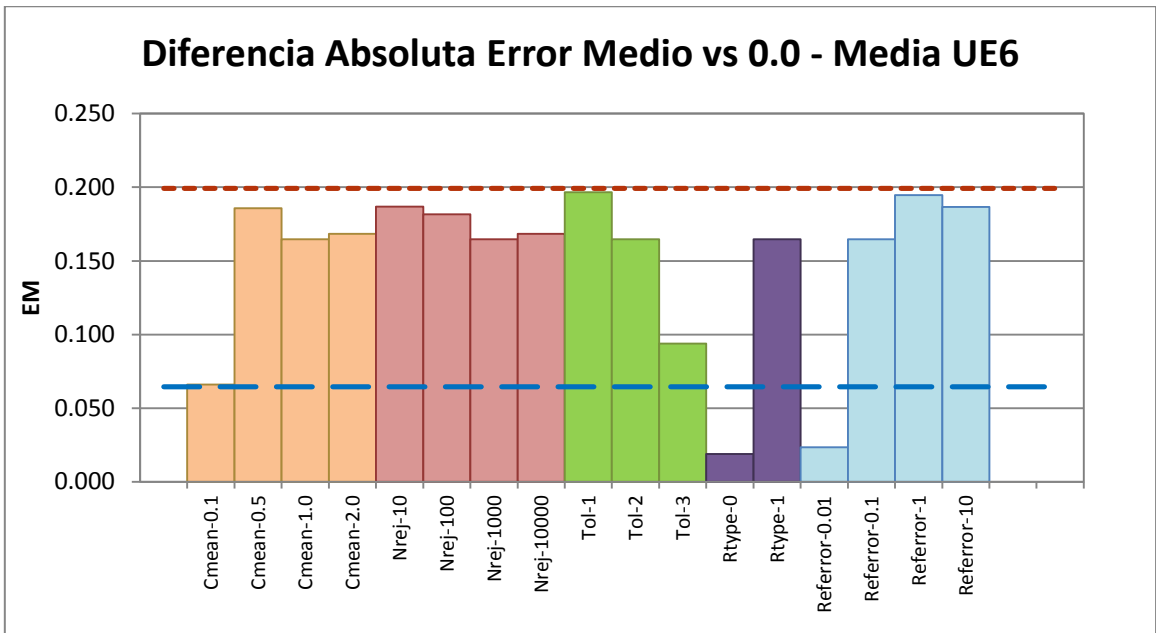
**Gráfico 231: Diferencia absoluta error medio vs 0.0 – Media UE1.**



**Gráfico 232: Diferencia absoluta error medio vs 0.0 – Media UE2.**



**Gráfico 233: Diferencia absoluta error medio vs 0.0 – Media UE3.**



**Gráfico 234: Diferencia absoluta error medio vs 0.0 – Media UE6.**

# ANEXO Q: Caso de estudio II – Accuracy plot por UE

## Q.1 Accuracy plot

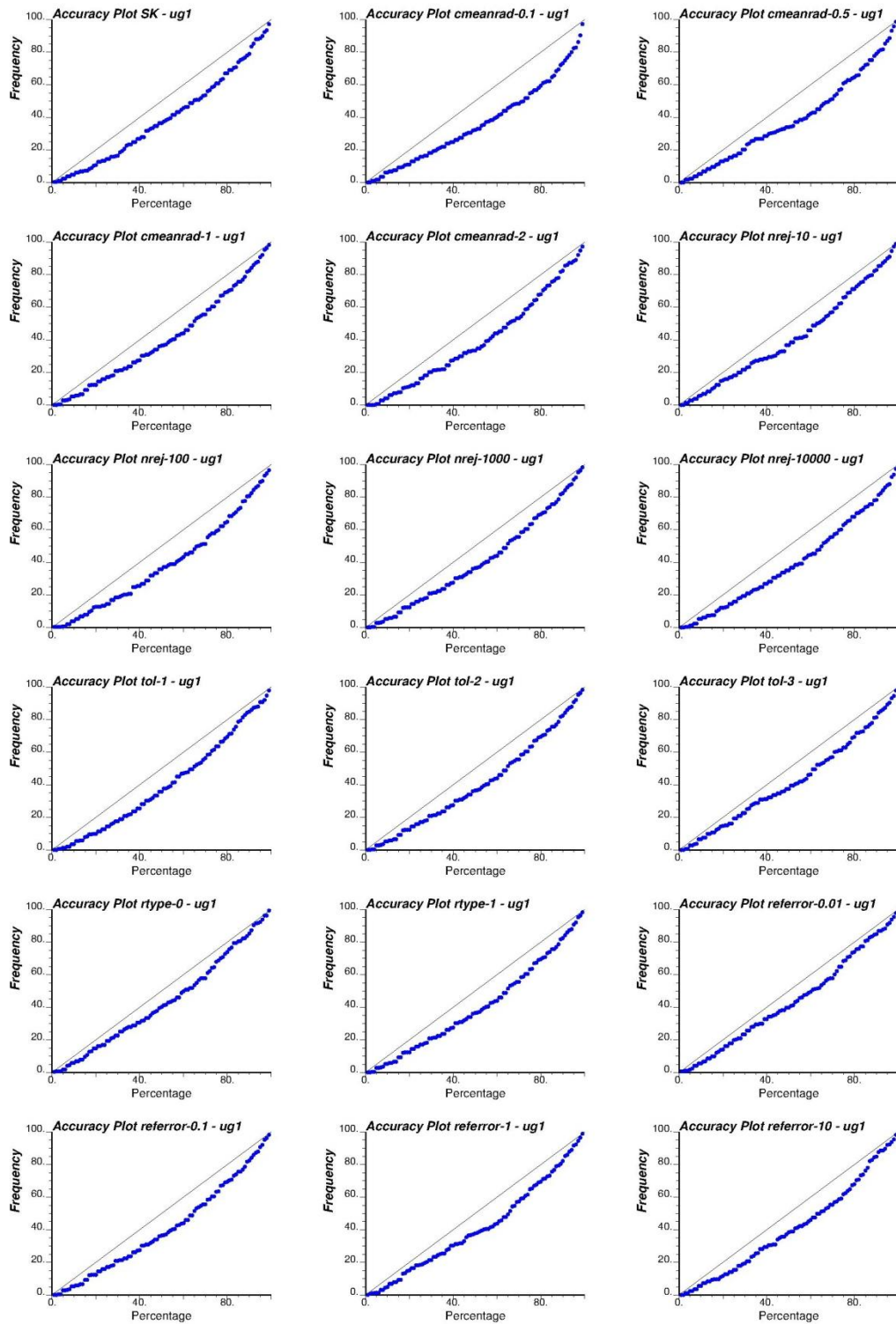
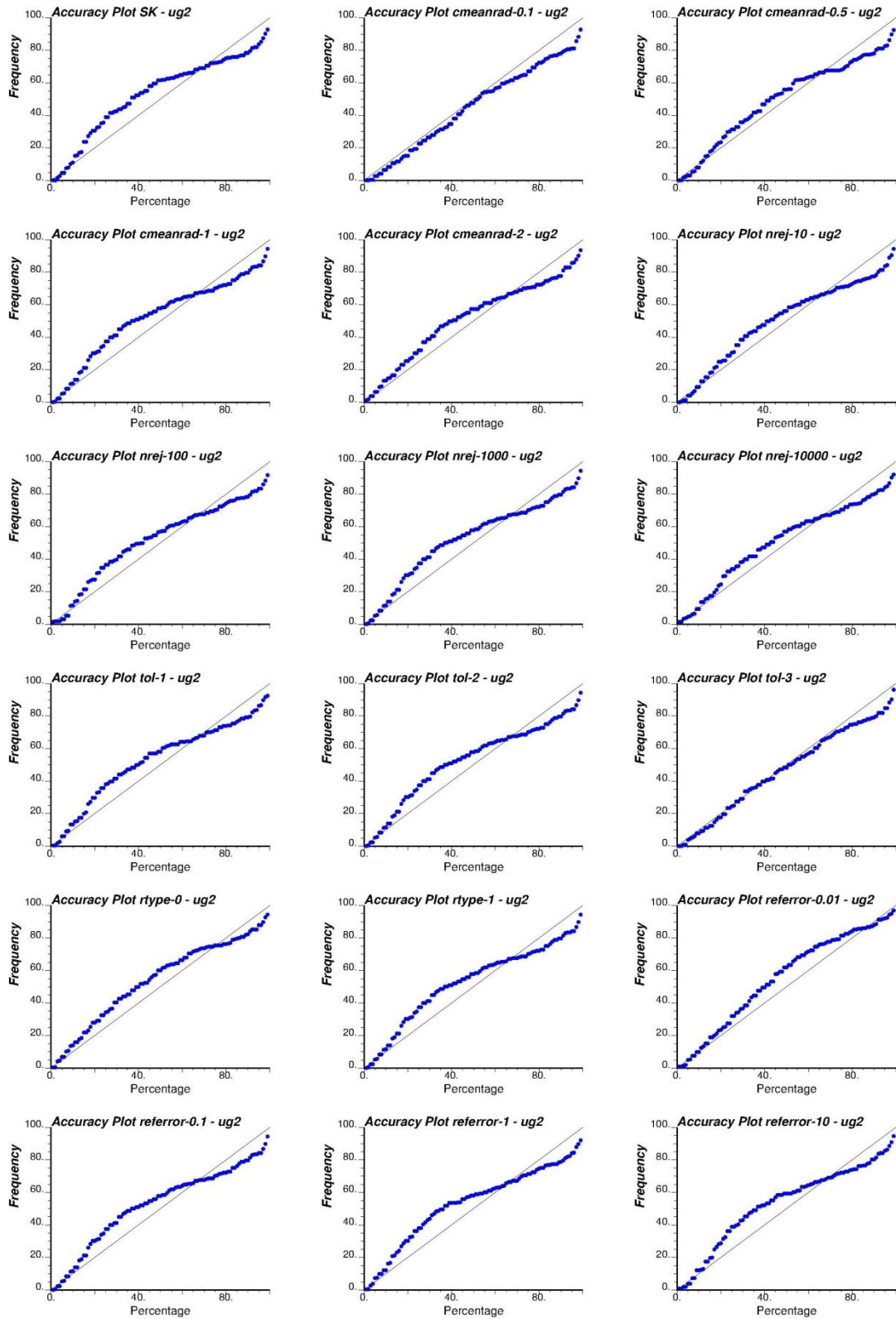
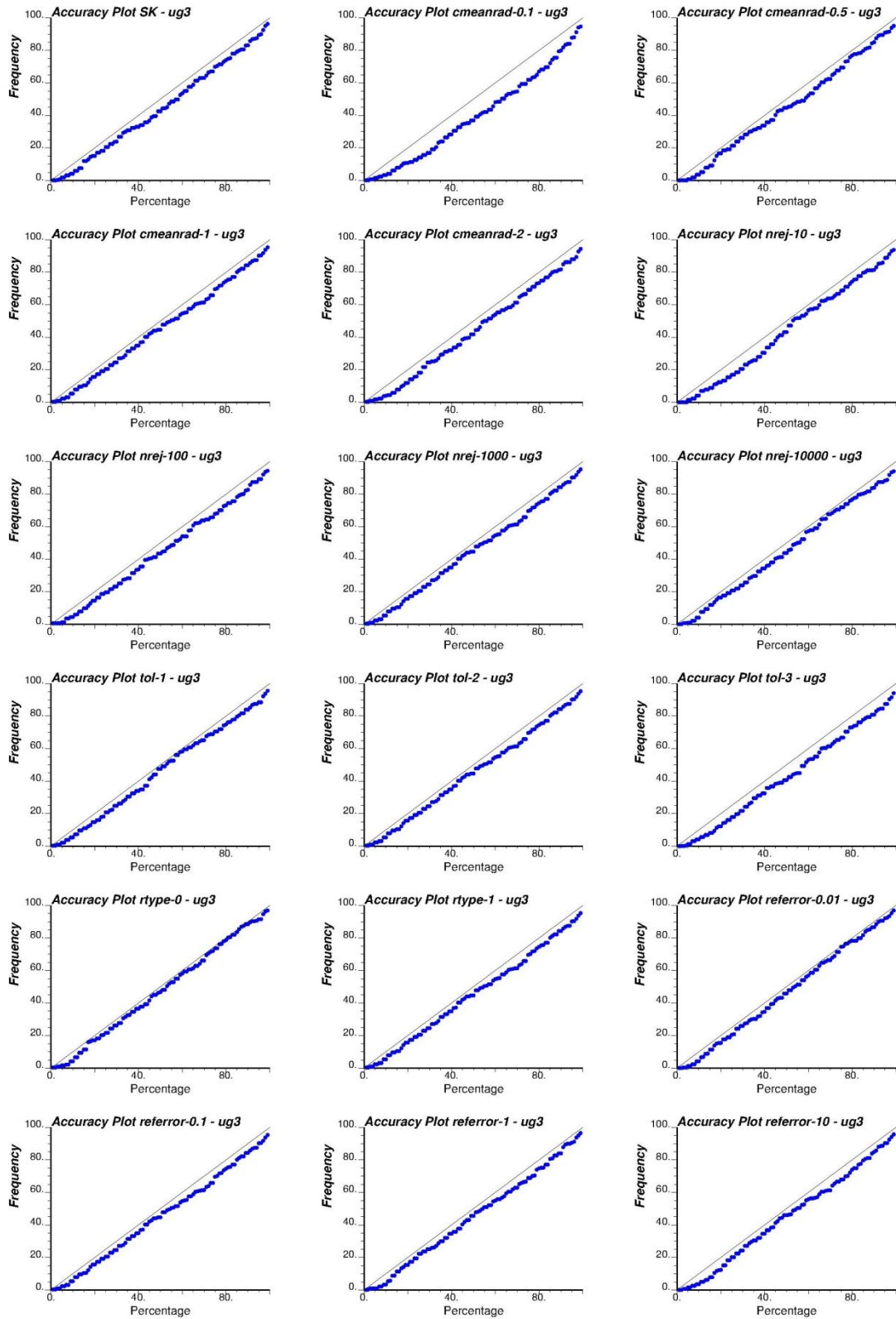


Gráfico 235: Accuracy plot – UE1.

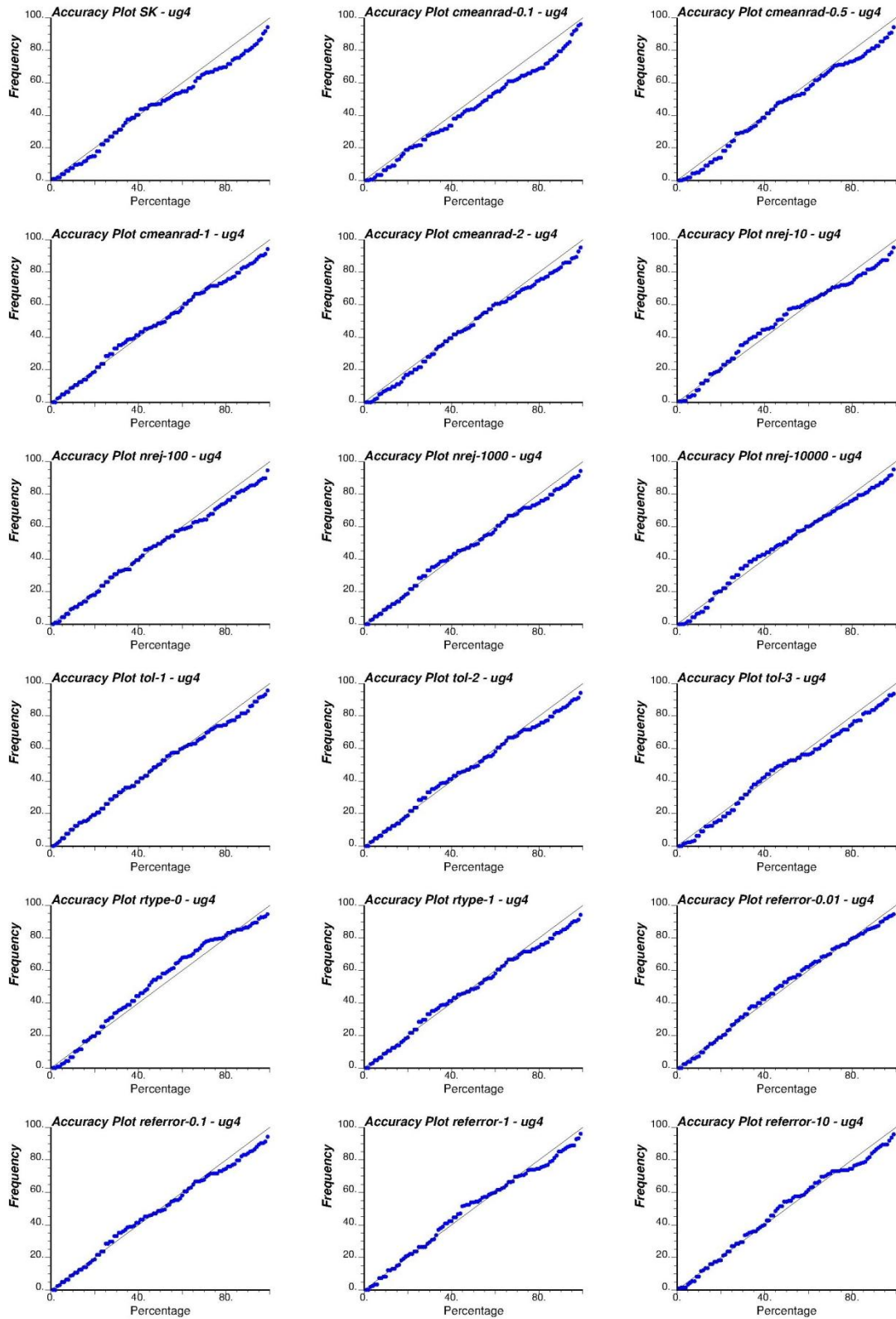


**Gráfico 236: Accuracy plot – UE2.**

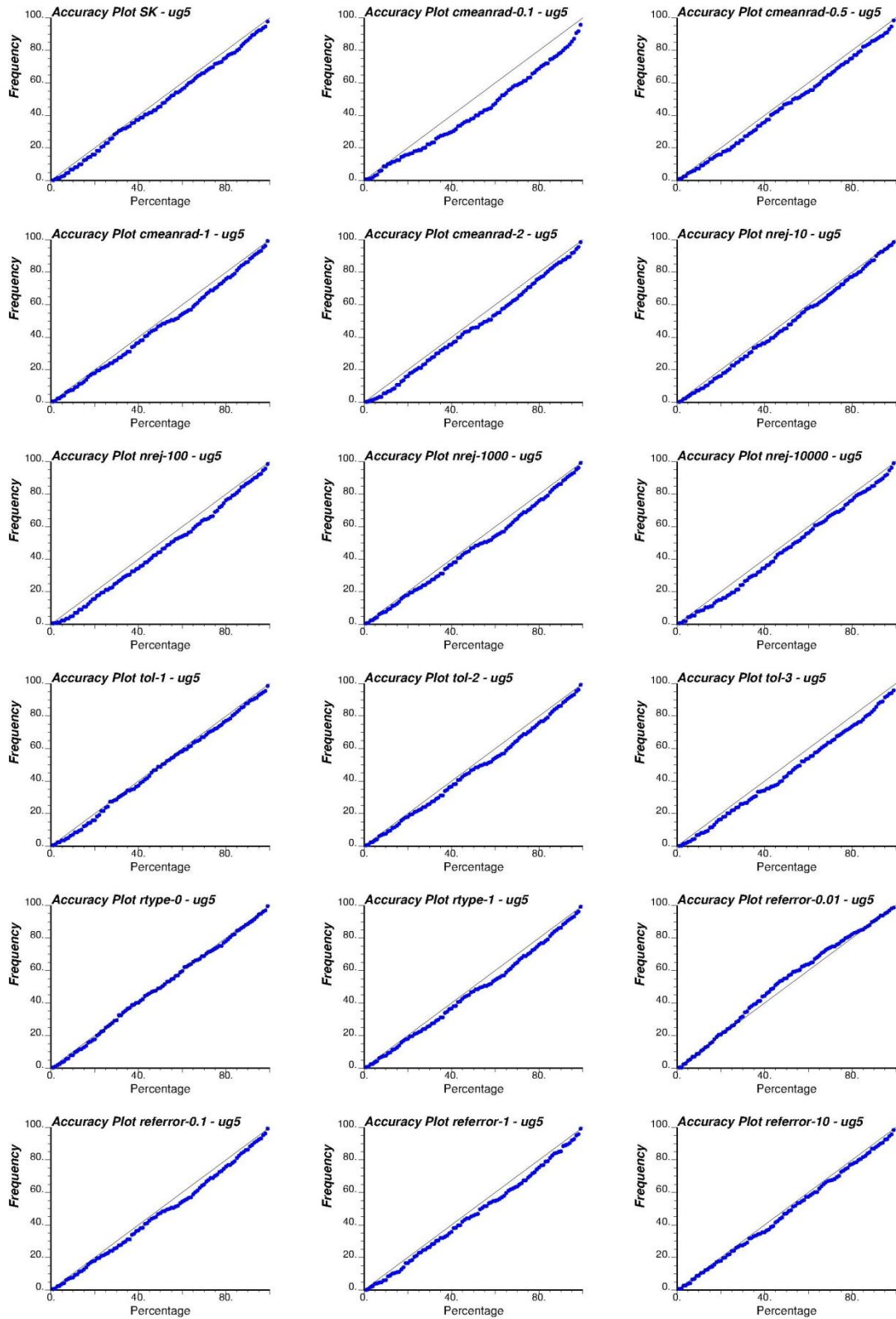




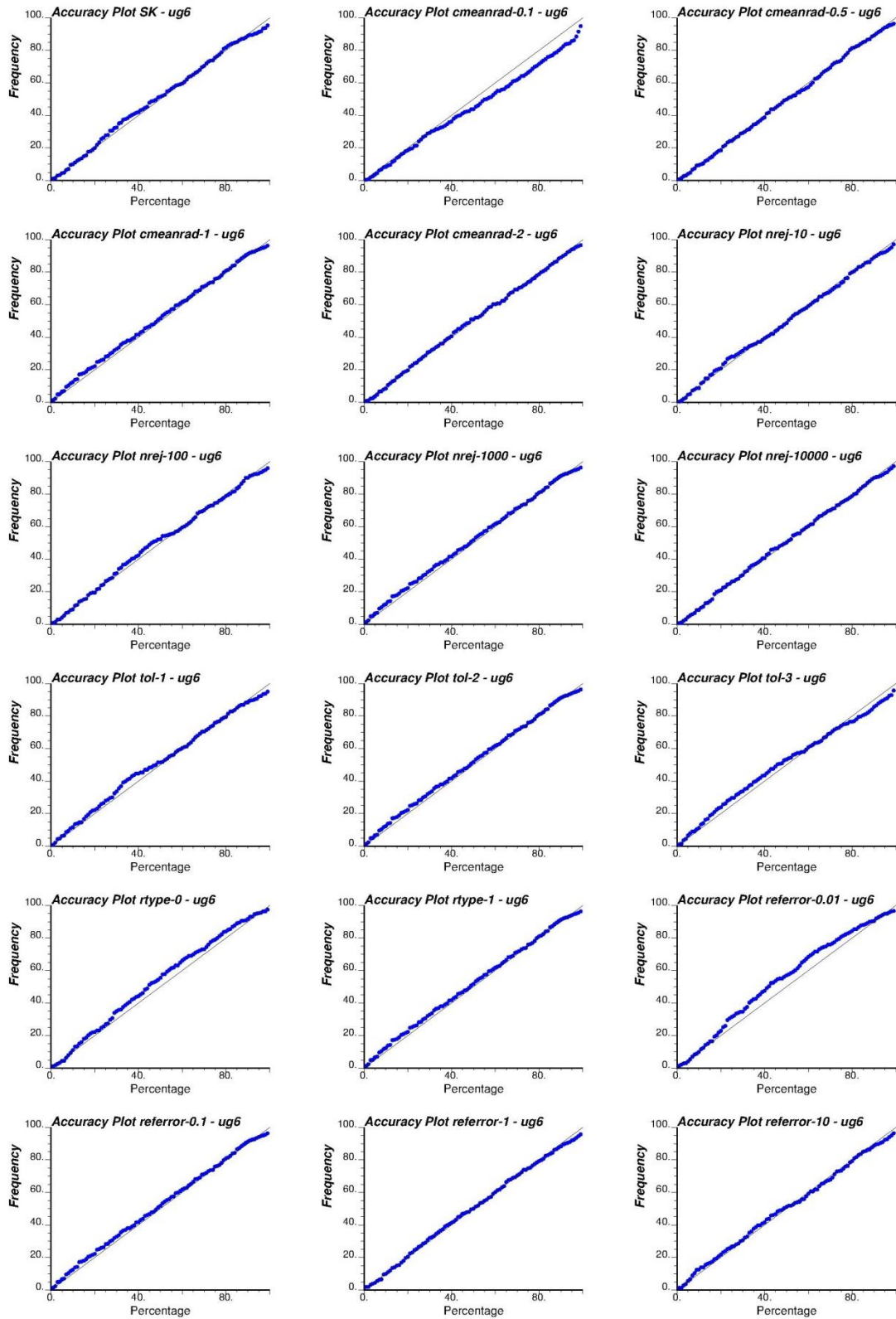
**Gráfico 237: Accuracy plot – UE3.**



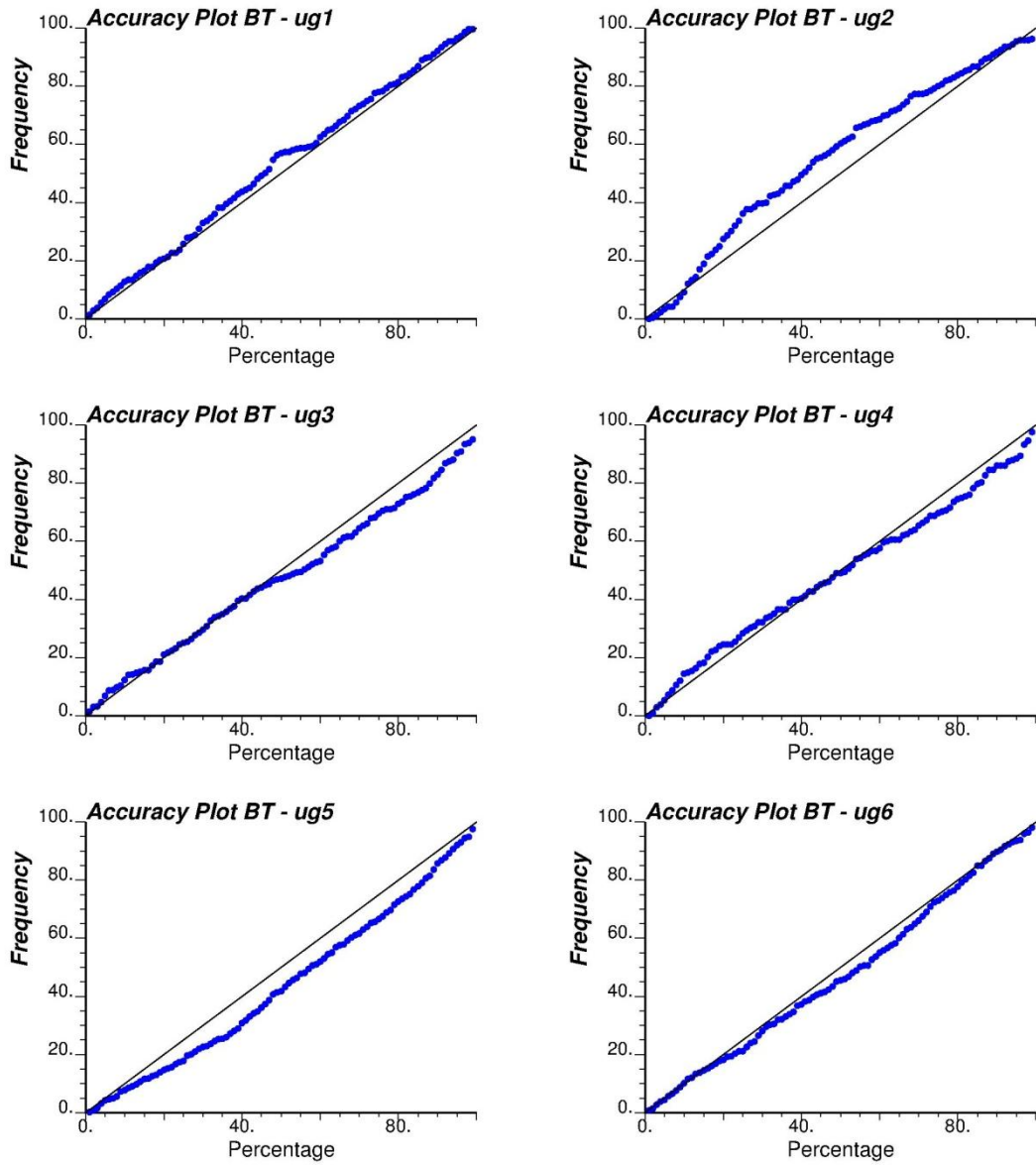
**Gráfico 238: Accuracy plot – UE4.**



**Gráfico 239: Accuracy plot – UE5.**



**Gráfico 240: Accuracy plot – UE6.**



**Gráfico 241: Accuracy plot – Kriging de resíduos.**

## Q.2 Pendiente de la regresión – Accuracy plot

Pendiente de la regresión - Accuracy Plot							
N°	Sensibilización	UE1	UE2	UE3	UE4	UE5	UE6
1	SK	0.948	0.814	0.981	0.902	0.984	0.977
2	Cmean-0.1	0.845	0.910	0.948	0.918	0.901	0.900
3	Cmean-0.5	0.901	0.858	0.990	0.951	0.978	1.002
4	Cmean-1.0	0.960	0.812	0.973	0.931	0.977	0.970
5	Cmean-2.0	0.956	0.826	0.986	0.963	0.995	0.986
6	Nrej-10	0.961	0.849	0.999	0.922	1.006	0.971
7	Nrej-100	0.943	0.820	0.977	0.930	0.998	0.976
8	Nrej-1000	0.960	0.812	0.973	0.931	0.977	0.970
9	Nrej-10000	0.943	0.846	0.996	0.942	0.999	0.988
10	Tol-1	0.993	0.817	0.990	0.938	0.999	0.951
11	Tol-2	0.960	0.812	0.973	0.931	0.977	0.970
12	Tol-3	0.934	0.918	0.972	0.954	0.968	0.911
13	Rtype-0	0.991	0.869	1.023	1.003	1.001	1.014
14	Rtype-1	0.960	0.812	0.973	0.931	0.977	0.970
15	Referror-0.01	0.977	0.990	1.023	0.983	1.015	1.006
16	Referror-0.1	0.960	0.812	0.973	0.931	0.977	0.970
17	Referror-1	0.951	0.790	0.990	0.941	0.983	0.974
18	Referror-10	0.969	0.829	0.992	0.953	0.981	0.955
19	BT	1.006	0.987	0.896	0.893	0.976	0.990

Tabla 54: Pendiente de la regresión por UE – Accuracy plot.

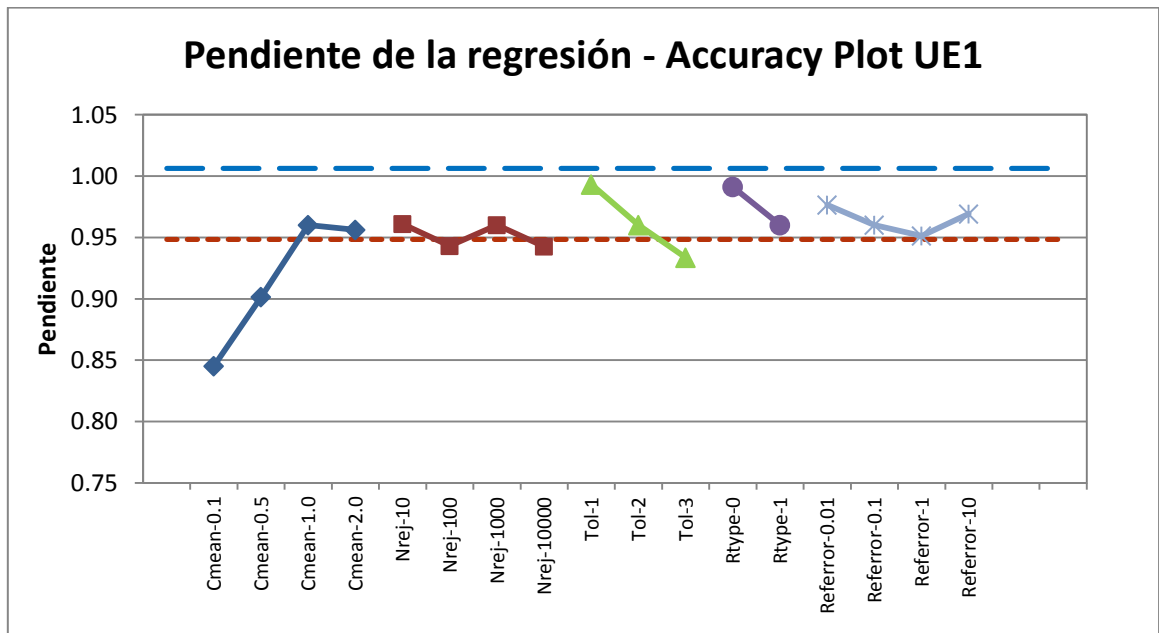
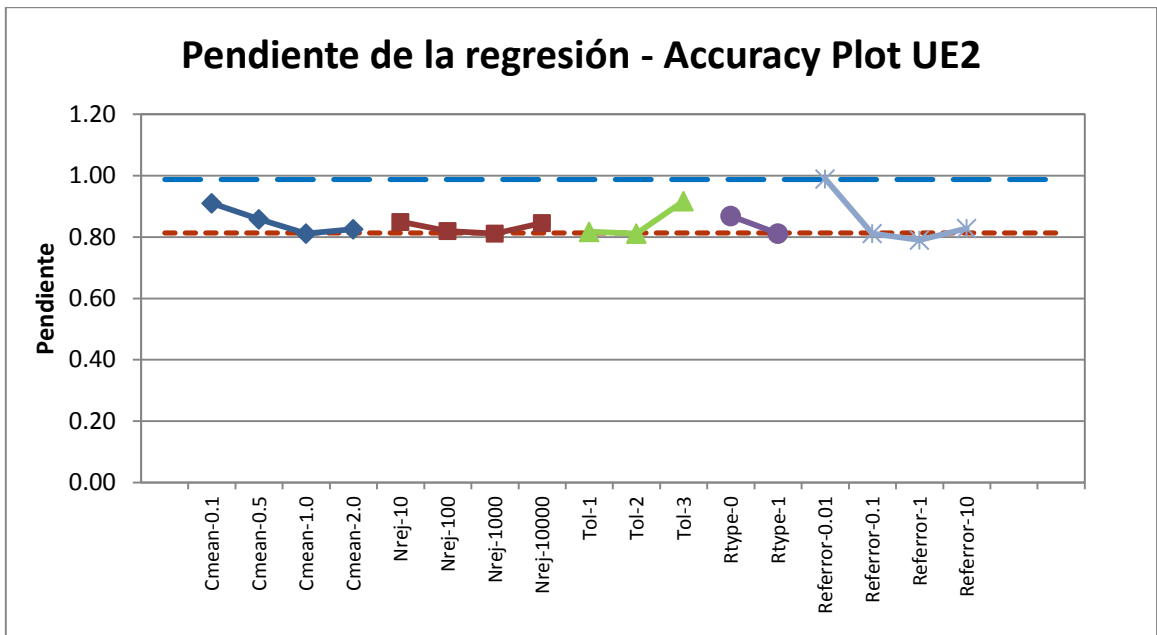
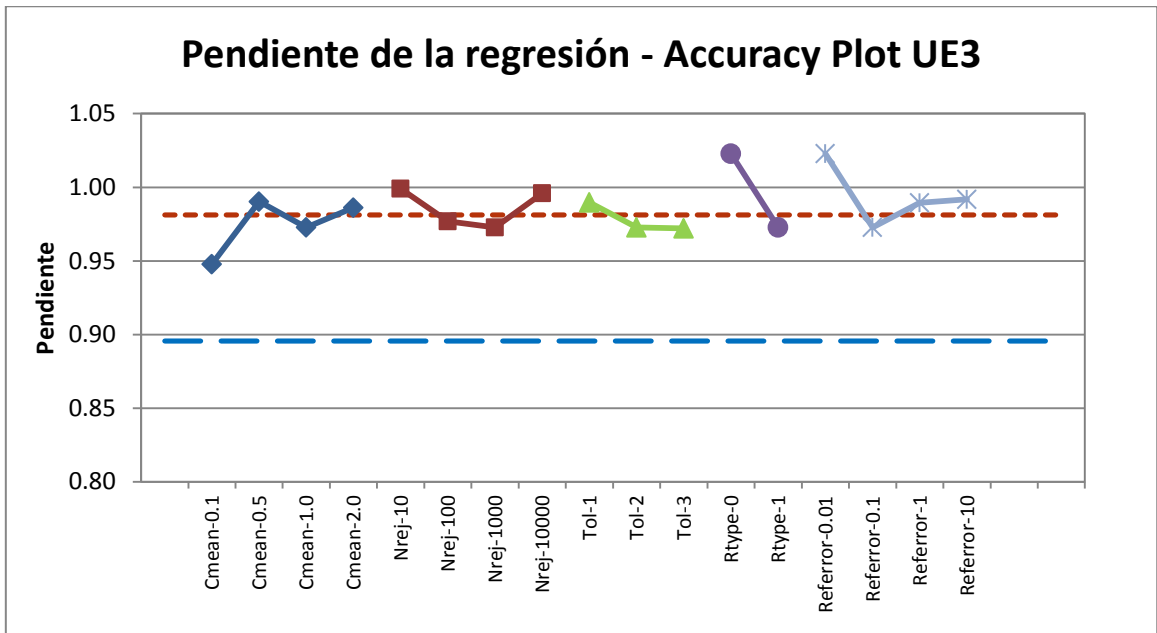


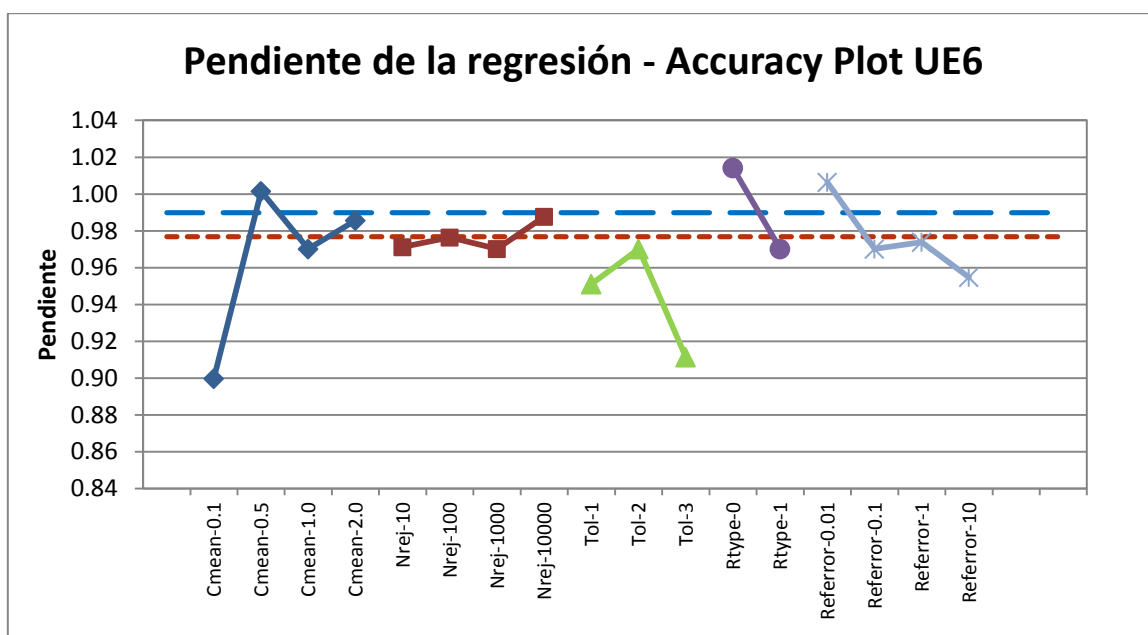
Gráfico 242: Pendiente de la regresión – Accuracy plot UE1.



**Gráfico 243: Pendiente de la regresión – Accuracy plot UE2.**



**Gráfico 244: Pendiente de la regresión – Accuracy plot UE3.**

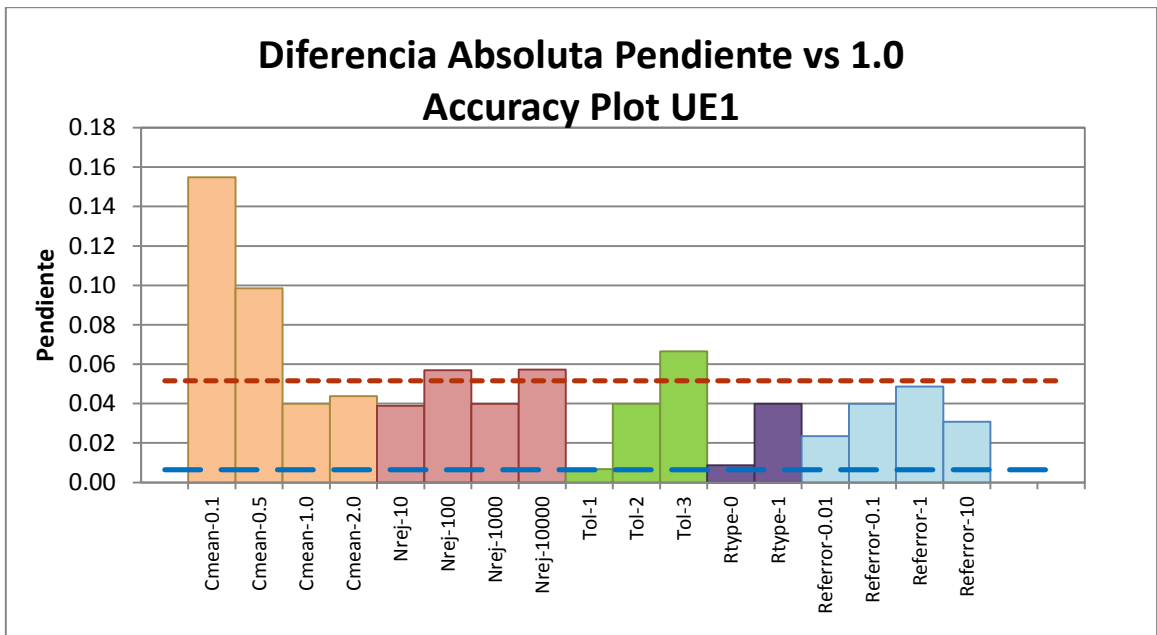


**Gráfico 245: Pendiente de la regresión – Accuracy plot UE6.**

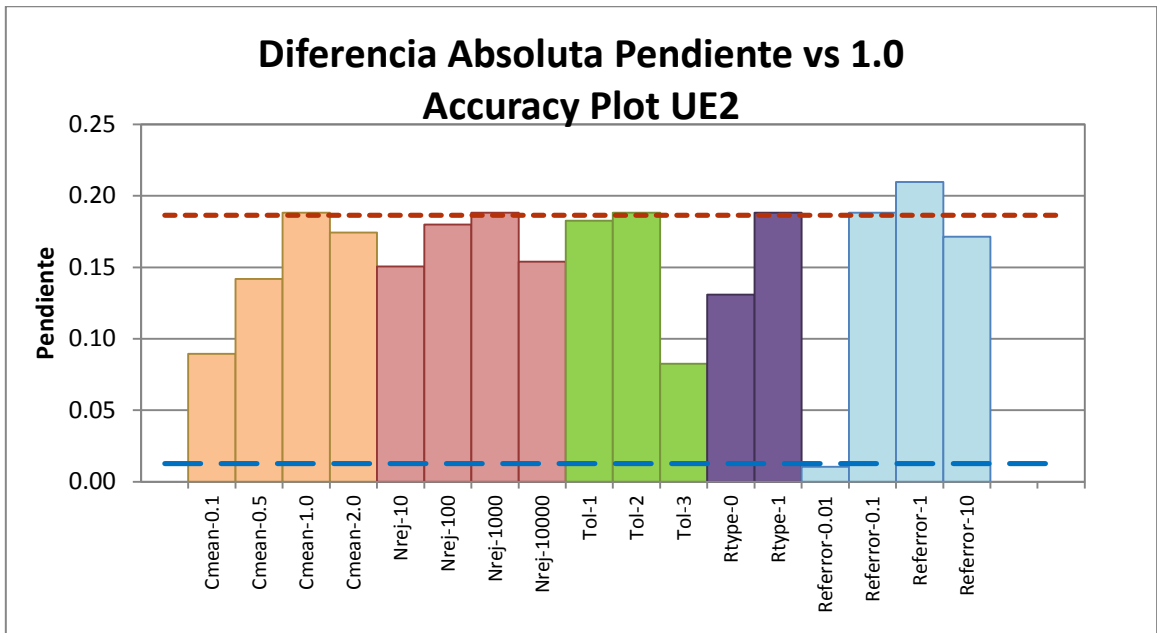
Diferencia Absoluta Pendiente vs 1.0 - Accuracy Plot							
N°	Sensibilización	UE1	UE2	UE3	UE4	UE5	UE6
1	SK	0.052	0.186	0.019	0.098	0.016	0.023
2	Cmean-0.1	0.155	0.090	0.052	0.082	0.099	0.100
3	Cmean-0.5	0.099	0.142	0.010	0.049	0.022	0.002
4	Cmean-1.0	0.040	0.188	0.027	0.069	0.023	0.030
5	Cmean-2.0	0.044	0.174	0.014	0.037	0.005	0.014
6	Nrej-10	0.039	0.151	0.001	0.078	0.006	0.029
7	Nrej-100	0.057	0.180	0.023	0.070	0.002	0.024
8	Nrej-1000	0.040	0.188	0.027	0.069	0.023	0.030
9	Nrej-10000	0.057	0.154	0.004	0.058	0.001	0.012
10	Tol-1	0.007	0.183	0.010	0.062	0.001	0.049
11	Tol-2	0.040	0.188	0.027	0.069	0.023	0.030
12	Tol-3	0.066	0.082	0.028	0.046	0.032	0.089
13	Rtype-0	0.009	0.131	0.023	0.003	0.001	0.014
14	Rtype-1	0.040	0.188	0.027	0.069	0.023	0.030
15	Referror-0.01	0.023	0.010	0.023	0.017	0.015	0.006
16	Referror-0.1	0.040	0.188	0.027	0.069	0.023	0.030
17	Referror-1	0.049	0.210	0.010	0.059	0.017	0.026
18	Referror-10	0.031	0.171	0.008	0.047	0.019	0.045
19	BT	0.006	0.013	0.104	0.107	0.024	0.010

**Tabla 55: Diferencia absoluta pendiente vs 1.0 por UE – Accuracy plot.**

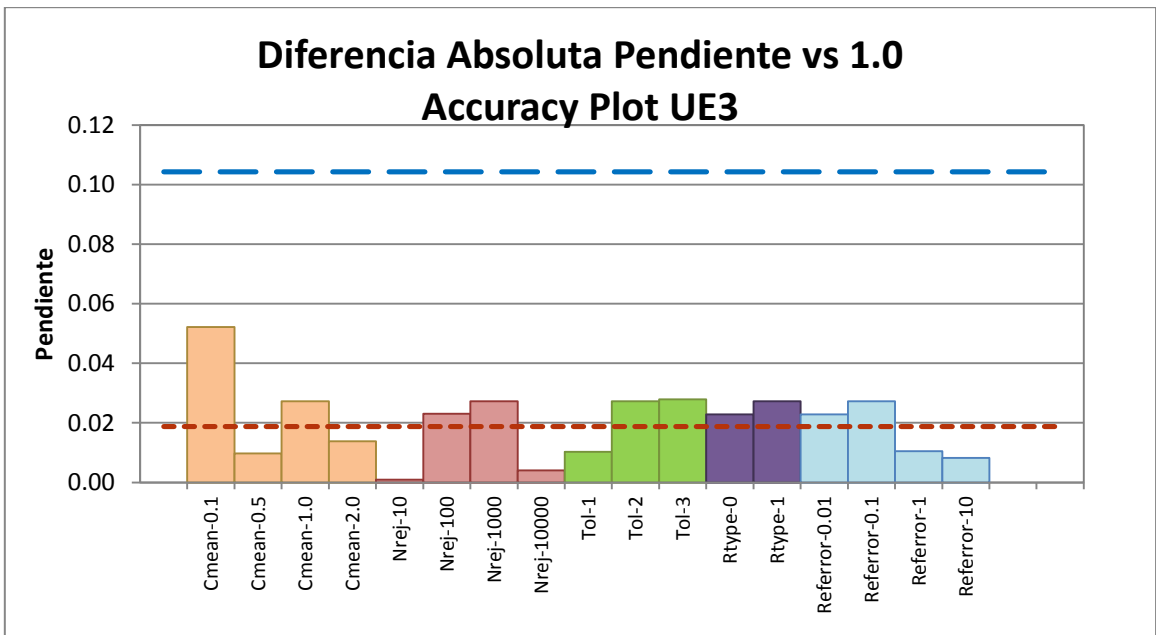




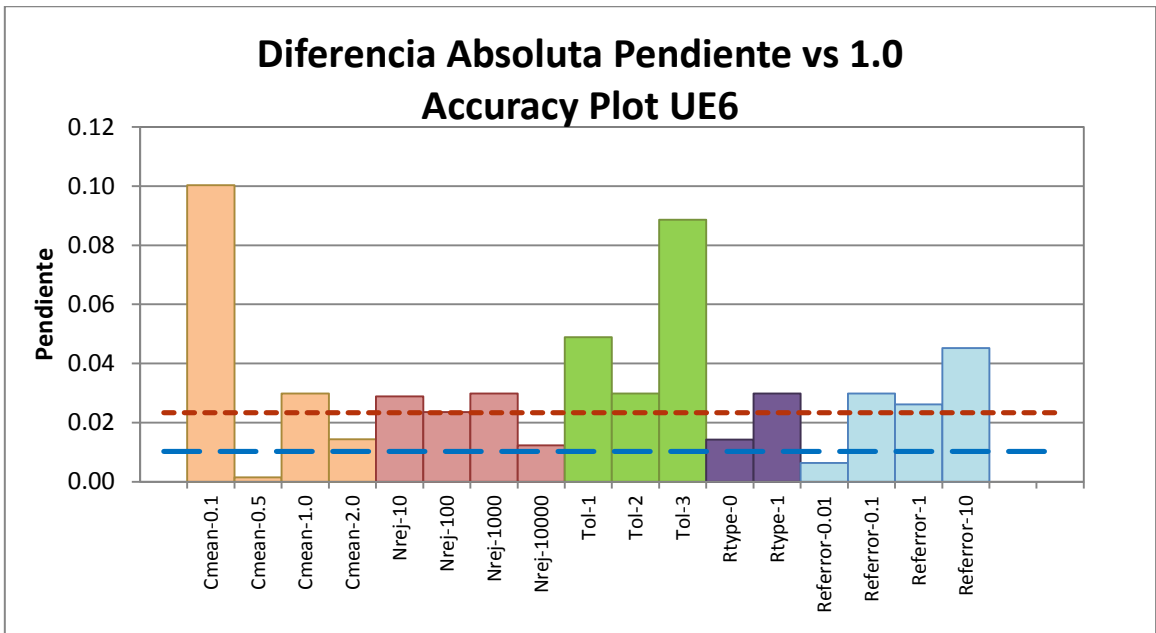
**Gráfico 246: Diferencia absoluta pendiente vs 1 – Accuracy plot UE1.**



**Gráfico 247: Diferencia absoluta pendiente vs 1 – Accuracy plot UE2.**



**Gráfico 248: Diferencia absoluta pendiente vs 1 – Accuracy plot UE3.**



**Gráfico 249: Diferencia absoluta pendiente vs 1 – Accuracy plot UE6.**

**ANEXO R: Caso de estudio II – Tiempo de simulación por UE**

Tiempo de simulación [segundos] - UE1					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	3.490	4.320	<b>3.793</b>	0.014
3	Cmean-0.5	5.440	9.000	<b>7.148</b>	0.493
4	Cmean-1.0	5.990	10.650	<b>8.062</b>	1.023
5	Cmean-2.0	6.020	11.580	<b>8.251</b>	1.219
6	Nrej-10	4.350	5.760	<b>4.783</b>	0.069
7	Nrej-100	4.620	6.510	<b>5.299</b>	0.297
8	Nrej-1000	5.050	8.550	<b>6.819</b>	0.440
9	Nrej-10000	13.680	38.350	<b>23.494</b>	33.024
10	Tol-1	3.560	4.930	<b>4.003</b>	0.143
11	Tol-2	5.650	9.410	<b>7.648</b>	0.535
12	Tol-3	5.730	9.950	<b>7.368</b>	1.040
13	Rtype-0	12.010	14.430	<b>13.275</b>	0.221
14	Rtype-1	4.980	8.250	<b>6.695</b>	0.409
15	Referror-0.01	20.260	27.000	<b>23.135</b>	1.739
16	Referror-0.1	5.520	10.780	<b>7.575</b>	0.689
17	Referror-1	3.710	4.660	<b>4.028</b>	0.040
18	Referror-10	3.730	4.660	<b>3.961</b>	0.033

**Tabla 56: Tiempo de simulación – UE1.**

Tiempo de simulación [segundos] - UE2					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	6.380	7.550	<b>6.754</b>	0.062
3	Cmean-0.5	11.560	14.550	<b>13.000</b>	0.416
4	Cmean-1.0	13.230	20.790	<b>16.709</b>	3.550
5	Cmean-2.0	12.760	20.940	<b>16.728</b>	3.459
6	Nrej-10	8.070	9.860	<b>8.720</b>	0.130
7	Nrej-100	7.600	10.200	<b>8.429</b>	0.299
8	Nrej-1000	11.780	17.710	<b>13.992</b>	1.098
9	Nrej-10000	34.930	79.860	<b>47.113</b>	53.705
10	Tol-1	6.080	7.100	<b>6.542</b>	0.018
11	Tol-2	12.980	17.710	<b>15.081</b>	0.878
12	Tol-3	10.550	15.260	<b>12.447</b>	0.905
13	Rtype-0	29.610	36.110	<b>32.717</b>	1.718
14	Rtype-1	11.470	15.600	<b>13.259</b>	0.638
15	Referror-0.01	50.440	72.560	<b>58.303</b>	15.564
16	Referror-0.1	12.850	17.350	<b>14.821</b>	0.982
17	Referror-1	6.260	7.530	<b>6.633</b>	0.042
18	Referror-10	6.380	7.550	<b>6.754</b>	0.062

**Tabla 57: Tiempo de simulación – UE2.**

Tiempo de simulación [segundos] - UE3					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	3.780	4.760	<b>4.014</b>	0.020
3	Cmean-0.5	9.130	12.650	<b>11.081</b>	0.477
4	Cmean-1.0	9.580	13.870	<b>11.544</b>	0.674
5	Cmean-2.0	9.860	14.090	<b>11.723</b>	0.733
6	Nrej-10	6.520	7.910	<b>6.960</b>	0.071
7	Nrej-100	6.290	7.740	<b>6.754</b>	0.067
8	Nrej-1000	9.700	14.840	<b>11.655</b>	0.909
9	Nrej-10000	21.930	50.010	<b>36.444</b>	35.730
10	Tol-1	5.150	5.850	<b>5.375</b>	0.011
11	Tol-2	10.060	13.820	<b>11.704</b>	0.632
12	Tol-3	9.530	14.650	<b>11.696</b>	1.284
13	Rtype-0	22.680	26.990	<b>24.415</b>	1.040
14	Rtype-1	8.700	12.000	<b>10.212</b>	0.535
15	Referror-0.01	39.420	53.820	<b>44.717</b>	6.282
16	Referror-0.1	10.140	14.490	<b>11.755</b>	0.770
17	Referror-1	5.350	6.850	<b>5.970</b>	0.100
18	Referror-10	5.410	6.790	<b>5.909</b>	0.077

**Tabla 58: Tiempo de simulación – UE3.**

Tiempo de simulación [segundos] - UE4					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	2.930	3.870	<b>3.172</b>	0.019
3	Cmean-0.5	3.370	4.600	<b>3.791</b>	0.039
4	Cmean-1.0	3.340	4.430	<b>3.628</b>	0.040
5	Cmean-2.0	3.200	4.230	<b>3.602</b>	0.052
6	Nrej-10	2.760	4.230	<b>2.960</b>	0.040
7	Nrej-100	2.620	3.290	<b>2.729</b>	0.008
8	Nrej-1000	3.280	4.460	<b>3.693</b>	0.044
9	Nrej-10000	5.910	11.080	<b>8.321</b>	1.048
10	Tol-1	2.620	3.210	<b>2.762</b>	0.006
11	Tol-2	3.530	4.240	<b>3.839</b>	0.023
12	Tol-3	3.620	5.320	<b>4.086</b>	0.086
13	Rtype-0	4.700	6.100	<b>5.267</b>	0.128
14	Rtype-1	3.180	3.900	<b>3.442</b>	0.019
15	Referror-0.01	7.660	8.740	<b>8.056</b>	0.040
16	Referror-0.1	3.430	4.650	<b>3.915</b>	0.063
17	Referror-1	2.640	3.810	<b>2.934</b>	0.047
18	Referror-10	2.640	3.810	<b>2.934</b>	0.047

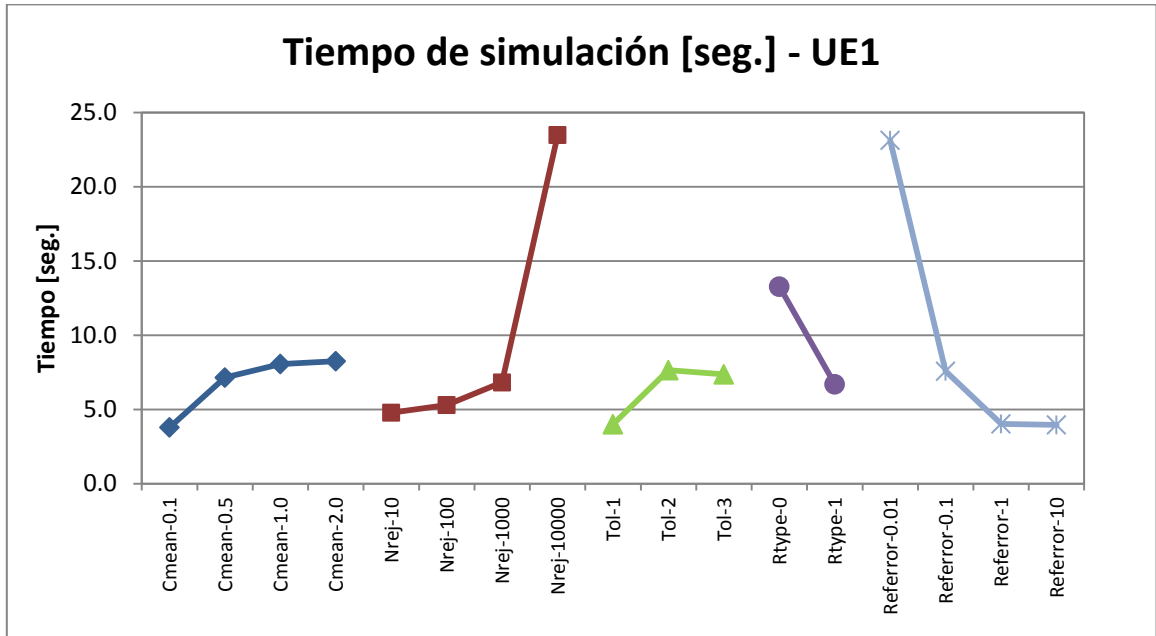
**Tabla 59: Tiempo de simulación – UE4.**

Tiempo de simulación [segundos] – UE5					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	8.350	9.140	<b>8.696</b>	0.027
3	Cmean-0.5	11.530	13.290	<b>12.395</b>	0.143
4	Cmean-1.0	10.950	12.790	<b>11.692</b>	0.121
5	Cmean-2.0	10.860	13.260	<b>12.026</b>	0.246
6	Nrej-10	5.550	6.970	<b>5.854</b>	0.070
7	Nrej-100	5.380	6.130	<b>5.592</b>	0.017
8	Nrej-1000	11.200	14.350	<b>12.411</b>	0.383
9	Nrej-10000	50.230	62.740	<b>55.720</b>	6.642
10	Tol-1	3.960	4.600	<b>4.179</b>	0.011
11	Tol-2	11.760	13.420	<b>12.575</b>	0.102
12	Tol-3	14.820	19.110	<b>16.053</b>	0.513
13	Rtype-0	14.540	16.210	<b>15.085</b>	0.158
14	Rtype-1	10.980	12.950	<b>11.895</b>	0.224
15	Referror-0.01	35.380	37.810	<b>36.662</b>	0.279
16	Referror-0.1	11.030	13.980	<b>12.764</b>	0.401
17	Referror-1	4.100	5.410	<b>4.510</b>	0.089
18	Referror-10	3.960	5.180	<b>4.291</b>	0.043

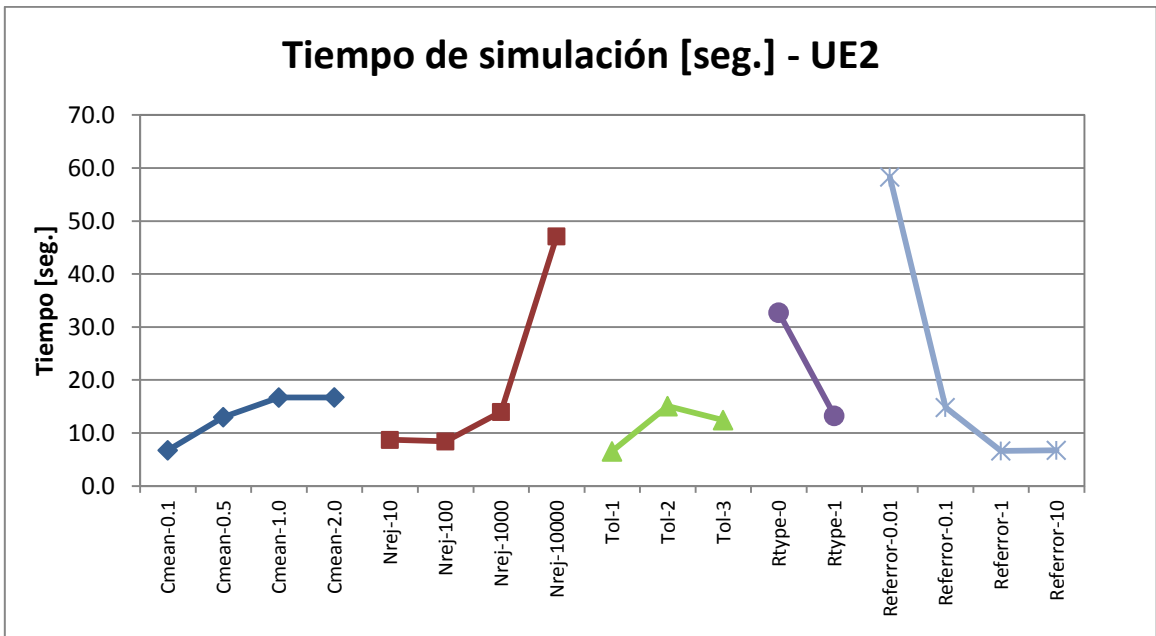
**Tabla 60: Tiempo de simulación – UE5.**

Tiempo de simulación [segundos] – UE6					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	5.570	6.470	<b>5.942</b>	0.019
3	Cmean-0.5	11.170	14.160	<b>12.500</b>	0.292
4	Cmean-1.0	10.060	12.010	<b>11.118</b>	0.158
5	Cmean-2.0	10.610	13.450	<b>11.655</b>	0.239
6	Nrej-10	4.820	6.470	<b>5.201</b>	0.101
7	Nrej-100	4.740	5.430	<b>4.943</b>	0.013
8	Nrej-1000	10.920	14.200	<b>12.765</b>	0.473
9	Nrej-10000	50.950	74.550	<b>60.573</b>	28.842
10	Tol-1	3.320	3.900	<b>3.480</b>	0.009
11	Tol-2	10.260	13.700	<b>11.936</b>	0.893
12	Tol-3	15.160	19.190	<b>16.446</b>	0.793
13	Rtype-0	10.780	13.120	<b>11.706</b>	0.277
14	Rtype-1	10.860	13.280	<b>11.945</b>	0.228
15	Referror-0.01	23.790	28.830	<b>25.828</b>	1.618
16	Referror-0.1	11.390	13.730	<b>12.671</b>	0.279
17	Referror-1	3.670	5.130	<b>3.976</b>	0.039
18	Referror-10	3.450	4.340	<b>3.575</b>	0.010

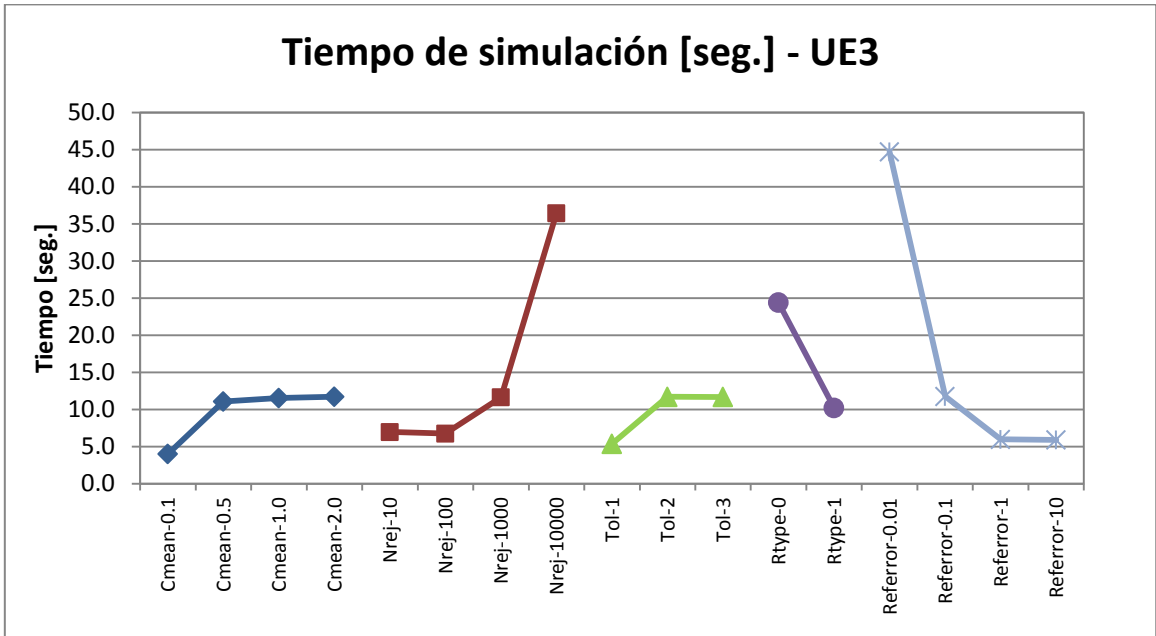
**Tabla 61: Tiempo de simulación – UE6.**



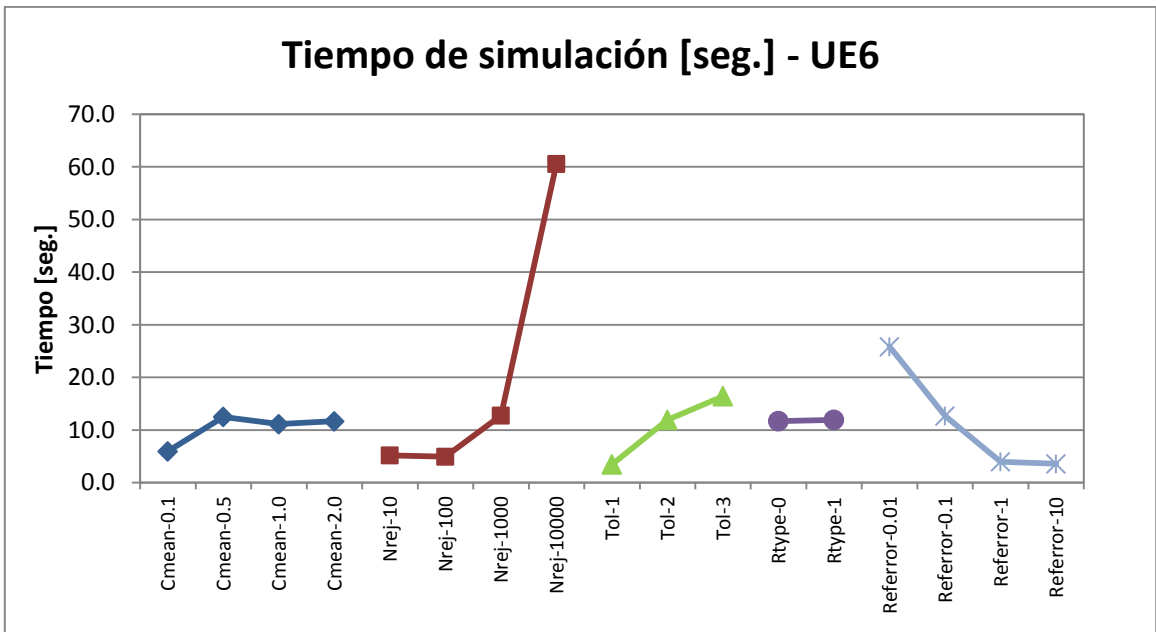
**Gráfico 250: Tiempo de simulación medio – UE1.**



**Gráfico 251: Tiempo de simulación medio – UE2.**



**Gráfico 252: Tiempo de simulación medio – UE3.**



**Gráfico 253: Tiempo de simulación medio – UE6.**

**ANEXO S: Caso de estudio II – Nodos rechazados por UE**

Nodos rechazados [%] - UE1					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	2.095	3.330	<b>2.753</b>	0.060
3	Cmean-0.5	3.976	15.184	<b>8.903</b>	5.131
4	Cmean-1.0	3.783	15.454	<b>9.875</b>	5.550
5	Cmean-2.0	3.933	18.271	<b>10.036</b>	6.612
6	Nrej-10	28.987	42.893	<b>35.404</b>	11.316
7	Nrej-100	12.650	30.458	<b>19.148</b>	10.487
8	Nrej-1000	3.783	15.454	<b>9.875</b>	5.550
9	Nrej-10000	2.105	10.275	<b>5.200</b>	3.621
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	3.783	15.454	<b>9.875</b>	5.550
12	Tol-3	3.706	13.493	<b>7.377</b>	5.485
13	Rtype-0	43.776	54.026	<b>49.294</b>	4.719
14	Rtype-1	3.783	15.454	<b>9.875</b>	5.550
15	Referror-0.01	50.320	64.512	<b>57.087</b>	9.307
16	Referror-0.1	3.783	15.454	<b>9.875</b>	5.550
17	Referror-1	0.000	0.000	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

**Tabla 62: Nodos rechazados – UE 1.**

Nodos rechazados [%] - UE2					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	0.000	0.000	<b>0.000</b>	0.000
3	Cmean-0.5	4.307	7.436	<b>5.703</b>	0.415
4	Cmean-1.0	5.420	10.656	<b>7.616</b>	1.052
5	Cmean-2.0	4.907	11.262	<b>7.644</b>	1.288
6	Nrej-10	28.560	36.415	<b>32.267</b>	2.726
7	Nrej-100	13.274	19.497	<b>16.047</b>	1.822
8	Nrej-1000	5.420	10.656	<b>7.616</b>	1.052
9	Nrej-10000	1.923	6.468	<b>3.486</b>	0.635
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	5.420	10.656	<b>7.616</b>	1.052
12	Tol-3	3.163	6.520	<b>4.290</b>	0.419
13	Rtype-0	45.418	52.274	<b>48.721</b>	1.372
14	Rtype-1	5.420	10.656	<b>7.616</b>	1.052
15	Referror-0.01	44.967	54.496	<b>49.152</b>	3.814
16	Referror-0.1	5.420	10.656	<b>7.616</b>	1.052
17	Referror-1	0.000	0.001	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

**Tabla 63: Nodos rechazados – UE 2.**



Nodos rechazados [%] - UE3					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	0.912	1.241	<b>1.099</b>	0.004
3	Cmean-0.5	4.286	10.075	<b>7.416</b>	1.110
4	Cmean-1.0	5.577	11.368	<b>8.171</b>	1.575
5	Cmean-2.0	5.912	11.448	<b>8.402</b>	1.401
6	Nrej-10	28.799	38.331	<b>32.856</b>	3.882
7	Nrej-100	12.259	20.267	<b>16.487</b>	3.198
8	Nrej-1000	5.577	11.368	<b>8.171</b>	1.575
9	Nrej-10000	1.958	6.508	<b>4.183</b>	0.958
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	5.577	11.368	<b>8.171</b>	1.575
12	Tol-3	3.947	9.322	<b>6.198</b>	1.315
13	Rtype-0	44.715	52.706	<b>47.972</b>	2.655
14	Rtype-1	5.577	11.368	<b>8.171</b>	1.575
15	Referror-0.01	50.794	64.232	<b>55.798</b>	6.005
16	Referror-0.1	5.577	11.368	<b>8.171</b>	1.575
17	Referror-1	0.000	0.089	<b>0.005</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

**Tabla 64: Nodos rechazados – UE 3.**

Nodos rechazados [%] - UE4					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	4.911	8.950	<b>6.333</b>	0.616
3	Cmean-0.5	5.571	12.453	<b>9.065</b>	1.856
4	Cmean-1.0	5.916	12.788	<b>8.705</b>	1.678
5	Cmean-2.0	5.054	11.707	<b>8.376</b>	2.145
6	Nrej-10	26.065	36.680	<b>31.465</b>	4.549
7	Nrej-100	12.080	20.934	<b>15.667</b>	2.745
8	Nrej-1000	5.916	12.788	<b>8.705</b>	1.678
9	Nrej-10000	2.517	7.667	<b>4.799</b>	1.089
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	5.916	12.788	<b>8.705</b>	1.678
12	Tol-3	5.609	14.952	<b>9.383</b>	3.045
13	Rtype-0	39.466	46.980	<b>43.096</b>	2.021
14	Rtype-1	5.916	12.788	<b>8.705</b>	1.678
15	Referror-0.01	46.147	57.749	<b>50.373</b>	4.030
16	Referror-0.1	5.916	12.788	<b>8.705</b>	1.678
17	Referror-1	0.000	0.000	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

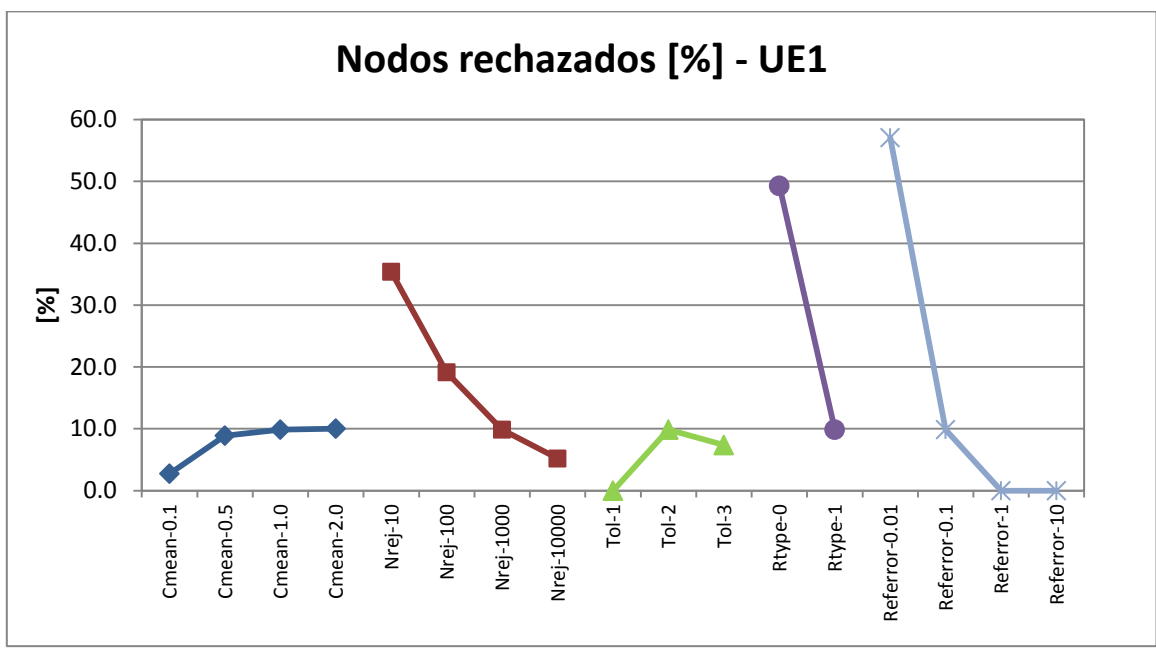
**Tabla 65: Nodos rechazados – UE 4.**

Nodos rechazados [%] – UE5					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	11.633	13.320	<b>12.350</b>	0.123
3	Cmean-0.5	16.588	20.808	<b>18.702</b>	0.754
4	Cmean-1.0	16.977	20.622	<b>18.896</b>	0.571
5	Cmean-2.0	16.945	20.894	<b>19.003</b>	0.746
6	Nrej-10	43.530	49.585	<b>45.961</b>	0.937
7	Nrej-100	26.697	31.831	<b>29.331</b>	1.144
8	Nrej-1000	16.977	20.622	<b>18.896</b>	0.571
9	Nrej-10000	10.900	14.484	<b>12.472</b>	0.515
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	16.977	20.622	<b>18.896</b>	0.571
12	Tol-3	20.627	23.657	<b>22.127</b>	0.422
13	Rtype-0	47.878	51.588	<b>49.658</b>	0.636
14	Rtype-1	16.977	20.622	<b>18.896</b>	0.571
15	Referror-0.01	65.582	68.948	<b>67.043</b>	0.411
16	Referror-0.1	16.977	20.622	<b>18.896</b>	0.571
17	Referror-1	0.002	0.490	<b>0.088</b>	0.006
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

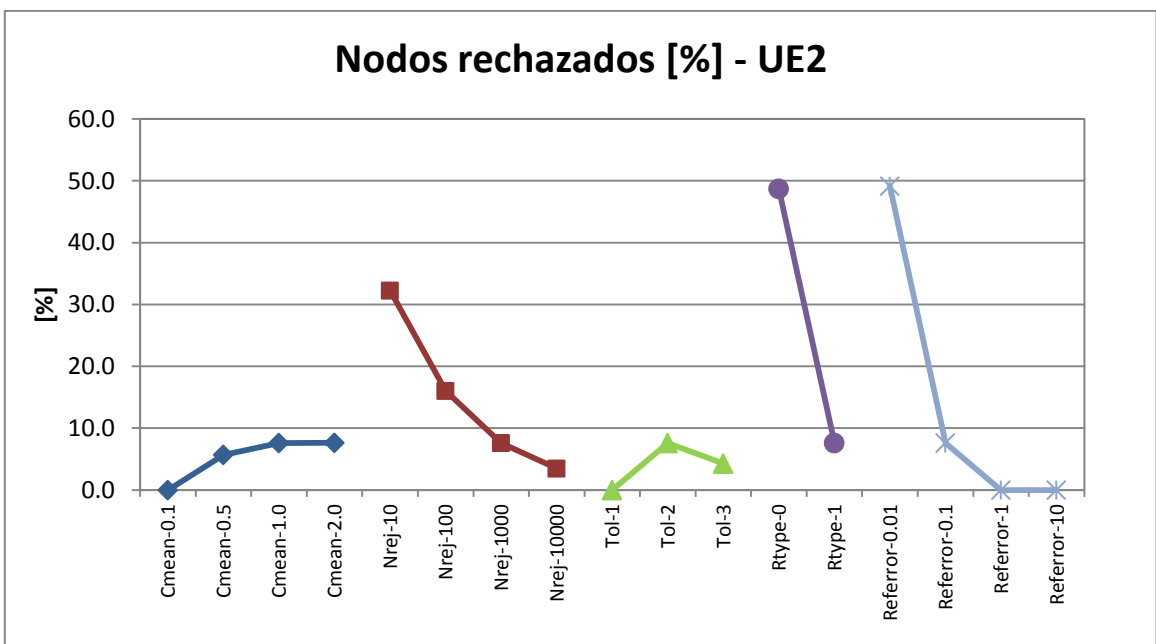
**Tabla 66: Nodos rechazados – UE5.**

Nodos rechazados [%] – UE6					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	10.937	12.602	<b>11.780</b>	0.124
3	Cmean-0.5	28.411	40.895	<b>34.262</b>	4.715
4	Cmean-1.0	29.382	38.946	<b>34.265</b>	3.368
5	Cmean-2.0	29.846	39.613	<b>34.117</b>	3.572
6	Nrej-10	58.265	66.084	<b>62.062</b>	2.868
7	Nrej-100	41.524	51.713	<b>46.310</b>	4.426
8	Nrej-1000	29.382	38.946	<b>34.265</b>	3.368
9	Nrej-10000	21.813	30.725	<b>25.817</b>	3.279
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	29.382	38.946	<b>34.265</b>	3.368
12	Tol-3	33.916	38.382	<b>35.911</b>	1.131
13	Rtype-0	53.585	62.009	<b>57.250</b>	2.094
14	Rtype-1	29.382	38.946	<b>34.265</b>	3.368
15	Referror-0.01	72.935	77.926	<b>75.143</b>	0.918
16	Referror-0.1	29.382	38.946	<b>34.265</b>	3.368
17	Referror-1	0.269	2.037	<b>0.804</b>	0.128
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

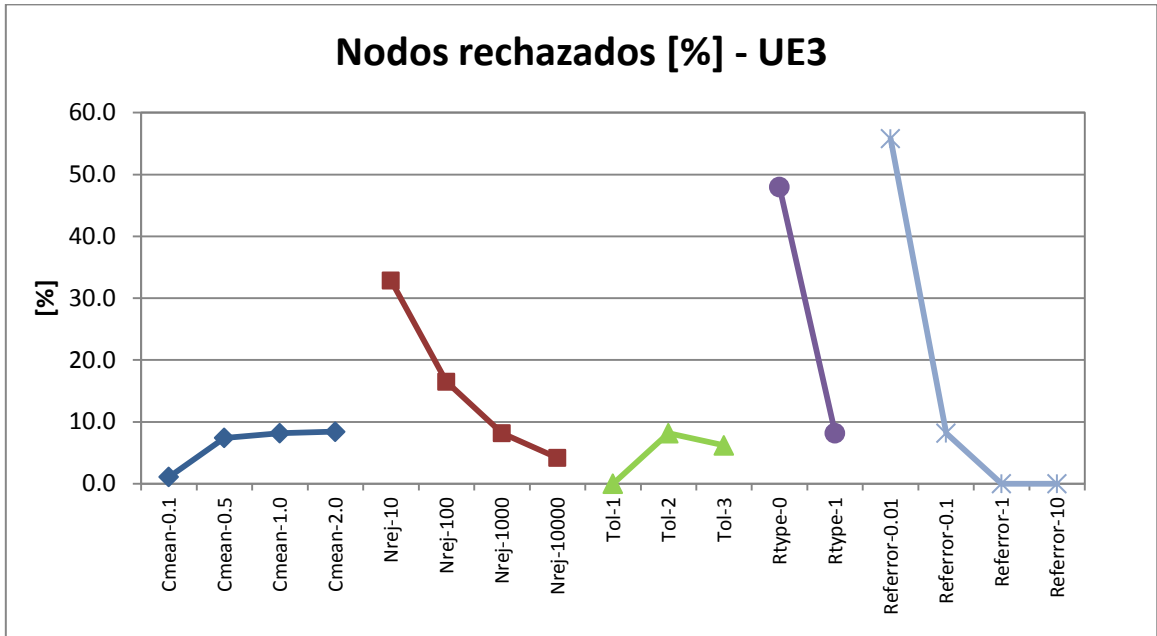
**Tabla 67: Nodos rechazados – UE6.**



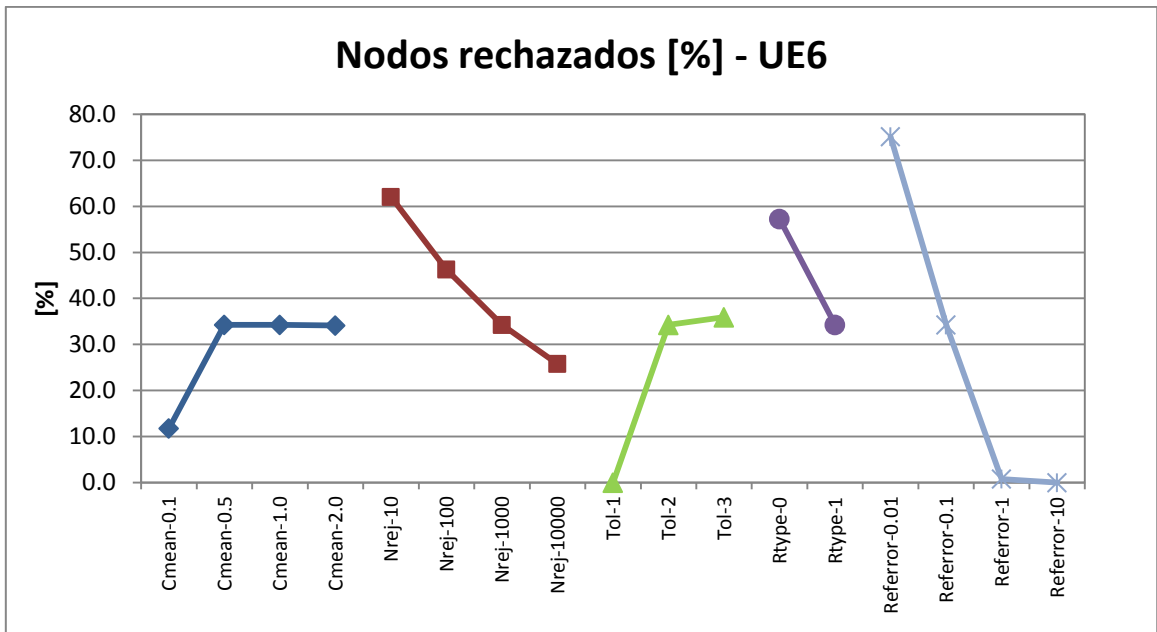
**Gráfico 254: Nodos rechazados medio – UE 1.**



**Gráfico 255: Nodos rechazados medio – UE 2.**



**Gráfico 256: Nodos rechazados medio – UE 3.**



**Gráfico 257: Nodos rechazados medio – UE6.**

## ANEXO T: Caso de estudio II – Nodos no simulados por UE

Nodos no simulados [%] - UE1					
Nº	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	0.000	0.000	<b>0.000</b>	0.000
3	Cmean-0.5	0.000	0.000	<b>0.000</b>	0.000
4	Cmean-1.0	0.000	0.000	<b>0.000</b>	0.000
5	Cmean-2.0	0.000	0.000	<b>0.000</b>	0.000
6	Nrej-10	0.000	0.044	<b>0.005</b>	0.000
7	Nrej-100	0.000	0.000	<b>0.000</b>	0.000
8	Nrej-1000	0.000	0.000	<b>0.000</b>	0.000
9	Nrej-10000	0.000	0.000	<b>0.000</b>	0.000
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	0.000	0.000	<b>0.000</b>	0.000
12	Tol-3	0.106	3.264	<b>0.728</b>	0.425
13	Rtype-0	0.000	0.022	<b>0.000</b>	0.000
14	Rtype-1	0.000	0.000	<b>0.000</b>	0.000
15	Referror-0.01	0.149	6.009	<b>1.963</b>	1.436
16	Referror-0.1	0.000	0.000	<b>0.000</b>	0.000
17	Referror-1	0.000	0.000	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

**Tabla 68: Nodos no simulados – UE 1.**

Nodos no simulados [%] - UE2					
Nº	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	0.000	0.000	<b>0.000</b>	0.000
3	Cmean-0.5	0.000	0.000	<b>0.000</b>	0.000
4	Cmean-1.0	0.000	0.000	<b>0.000</b>	0.000
5	Cmean-2.0	0.000	0.000	<b>0.000</b>	0.000
6	Nrej-10	0.027	0.429	<b>0.161</b>	0.007
7	Nrej-100	0.000	0.000	<b>0.000</b>	0.000
8	Nrej-1000	0.000	0.000	<b>0.000</b>	0.000
9	Nrej-10000	0.000	0.000	<b>0.000</b>	0.000
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	0.000	0.000	<b>0.000</b>	0.000
12	Tol-3	0.048	0.645	<b>0.212</b>	0.009
13	Rtype-0	0.002	0.426	<b>0.069</b>	0.005
14	Rtype-1	0.000	0.000	<b>0.000</b>	0.000
15	Referror-0.01	3.736	10.935	<b>6.675</b>	1.951
16	Referror-0.1	0.000	0.000	<b>0.000</b>	0.000
17	Referror-1	0.000	0.000	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

**Tabla 69: Nodos no simulados – UE 2.**

Nodos no simulados [%] - UE3					
Nº	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	0.000	0.000	<b>0.000</b>	0.000
3	Cmean-0.5	0.000	0.000	<b>0.000</b>	0.000
4	Cmean-1.0	0.000	0.000	<b>0.000</b>	0.000
5	Cmean-2.0	0.000	0.000	<b>0.000</b>	0.000
6	Nrej-10	0.010	0.296	<b>0.060</b>	0.002
7	Nrej-100	0.000	0.005	<b>0.000</b>	0.000
8	Nrej-1000	0.000	0.000	<b>0.000</b>	0.000
9	Nrej-10000	0.000	0.000	<b>0.000</b>	0.000
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	0.000	0.000	<b>0.000</b>	0.000
12	Tol-3	0.180	1.925	<b>0.678</b>	0.116
13	Rtype-0	0.000	0.388	<b>0.036</b>	0.004
14	Rtype-1	0.000	0.000	<b>0.000</b>	0.000
15	Referror-0.01	2.895	13.458	<b>6.505</b>	3.596
16	Referror-0.1	0.000	0.000	<b>0.000</b>	0.000
17	Referror-1	0.000	0.000	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

**Tabla 70: Nodos no simulados – UE 3.**

Nodos no simulados [%] - UE4					
Nº	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	0.000	0.000	<b>0.000</b>	0.000
3	Cmean-0.5	0.000	0.000	<b>0.000</b>	0.000
4	Cmean-1.0	0.000	0.000	<b>0.000</b>	0.000
5	Cmean-2.0	0.000	0.000	<b>0.000</b>	0.000
6	Nrej-10	0.000	0.000	<b>0.000</b>	0.000
7	Nrej-100	0.000	0.000	<b>0.000</b>	0.000
8	Nrej-1000	0.000	0.000	<b>0.000</b>	0.000
9	Nrej-10000	0.000	0.000	<b>0.000</b>	0.000
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	0.000	0.000	<b>0.000</b>	0.000
12	Tol-3	0.507	5.877	<b>2.126</b>	0.855
13	Rtype-0	0.000	0.038	<b>0.001</b>	0.000
14	Rtype-1	0.000	0.000	<b>0.000</b>	0.000
15	Referror-0.01	0.000	0.067	<b>0.005</b>	0.000
16	Referror-0.1	0.000	0.000	<b>0.000</b>	0.000
17	Referror-1	0.000	0.000	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

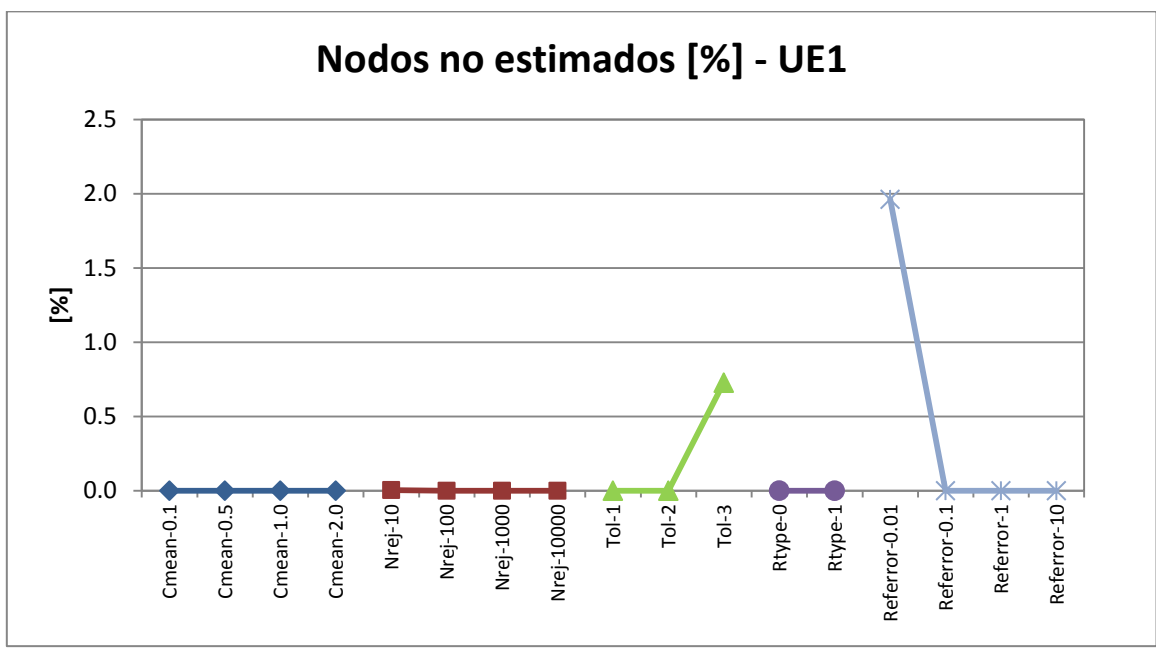
**Tabla 71: Nodos no simulados – UE 4.**

Nodos no simulados [%] – UE5					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	0.000	0.000	<b>0.000</b>	0.000
3	Cmean-0.5	0.000	0.000	<b>0.000</b>	0.000
4	Cmean-1.0	0.000	0.000	<b>0.000</b>	0.000
5	Cmean-2.0	0.000	0.000	<b>0.000</b>	0.000
6	Nrej-10	0.507	1.597	<b>0.878</b>	0.051
7	Nrej-100	0.000	0.027	<b>0.003</b>	0.000
8	Nrej-1000	0.000	0.000	<b>0.000</b>	0.000
9	Nrej-10000	0.000	0.000	<b>0.000</b>	0.000
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	0.000	0.000	<b>0.000</b>	0.000
12	Tol-3	4.701	6.714	<b>5.518</b>	0.130
13	Rtype-0	0.162	0.909	<b>0.360</b>	0.016
14	Rtype-1	0.000	0.000	<b>0.000</b>	0.000
15	Referror-0.01	11.488	15.287	<b>13.464</b>	0.521
16	Referror-0.1	0.000	0.000	<b>0.000</b>	0.000
17	Referror-1	0.000	0.000	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

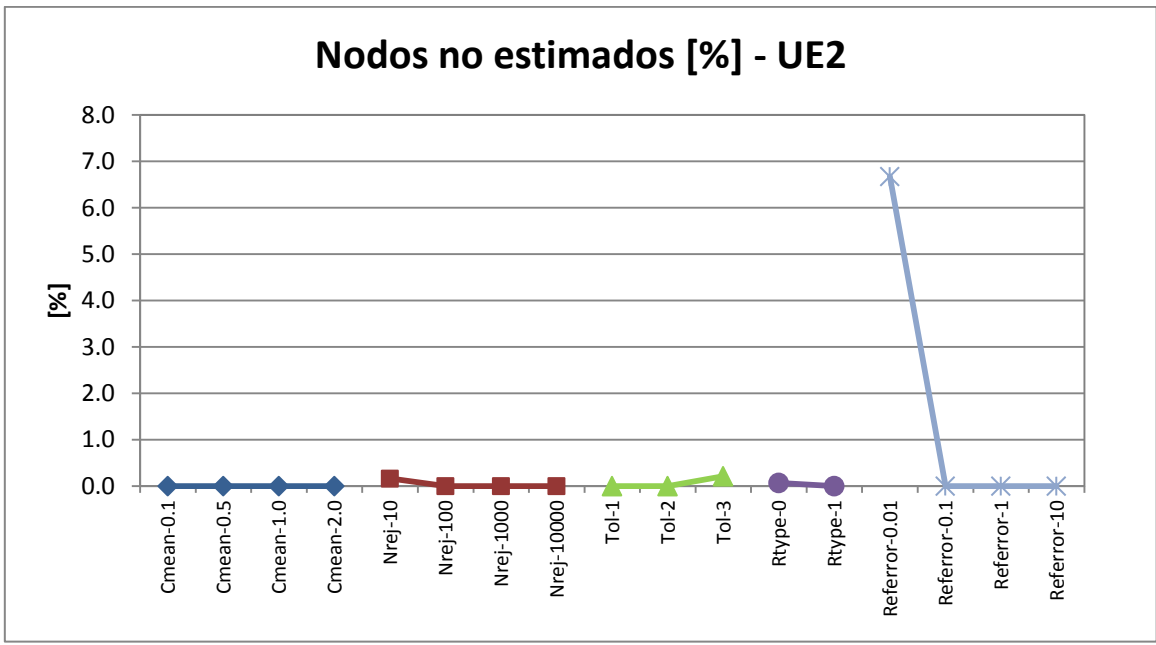
**Tabla 72: Nodos no simulados – UE5.**

Nodos no simulados [%] – UE6					
N°	Sensibilización	Mínimo	Máximo	Media	Varianza
2	Cmean-0.1	0.000	0.000	<b>0.000</b>	0.000
3	Cmean-0.5	0.000	0.015	<b>0.000</b>	0.000
4	Cmean-1.0	0.000	0.004	<b>0.000</b>	0.000
5	Cmean-2.0	0.000	0.004	<b>0.000</b>	0.000
6	Nrej-10	1.638	4.442	<b>2.880</b>	0.413
7	Nrej-100	0.004	0.292	<b>0.064</b>	0.003
8	Nrej-1000	0.000	0.004	<b>0.000</b>	0.000
9	Nrej-10000	0.000	0.000	<b>0.000</b>	0.000
10	Tol-1	0.000	0.000	<b>0.000</b>	0.000
11	Tol-2	0.000	0.004	<b>0.000</b>	0.000
12	Tol-3	9.671	13.811	<b>11.362</b>	0.509
13	Rtype-0	0.487	2.141	<b>1.162</b>	0.120
14	Rtype-1	0.000	0.004	<b>0.000</b>	0.000
15	Referror-0.01	14.332	21.836	<b>17.133</b>	1.978
16	Referror-0.1	0.000	0.004	<b>0.000</b>	0.000
17	Referror-1	0.000	0.000	<b>0.000</b>	0.000
18	Referror-10	0.000	0.000	<b>0.000</b>	0.000

**Tabla 73: Nodos no simulados – UE6.**

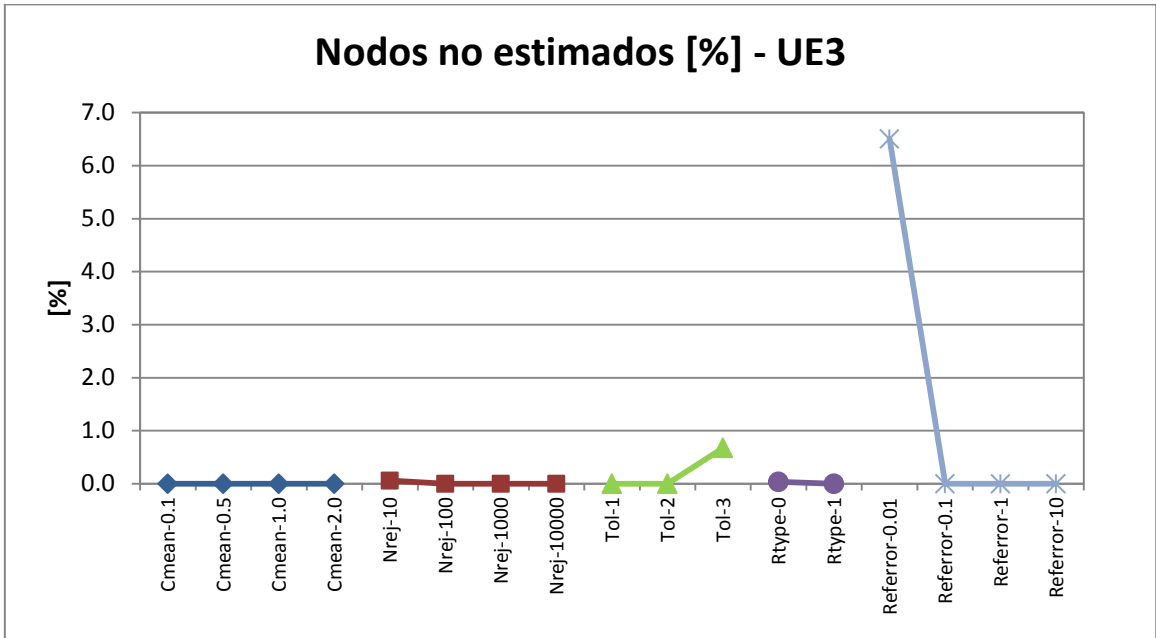


**Gráfico 258: Nodos no simulados medio – UE 1.**

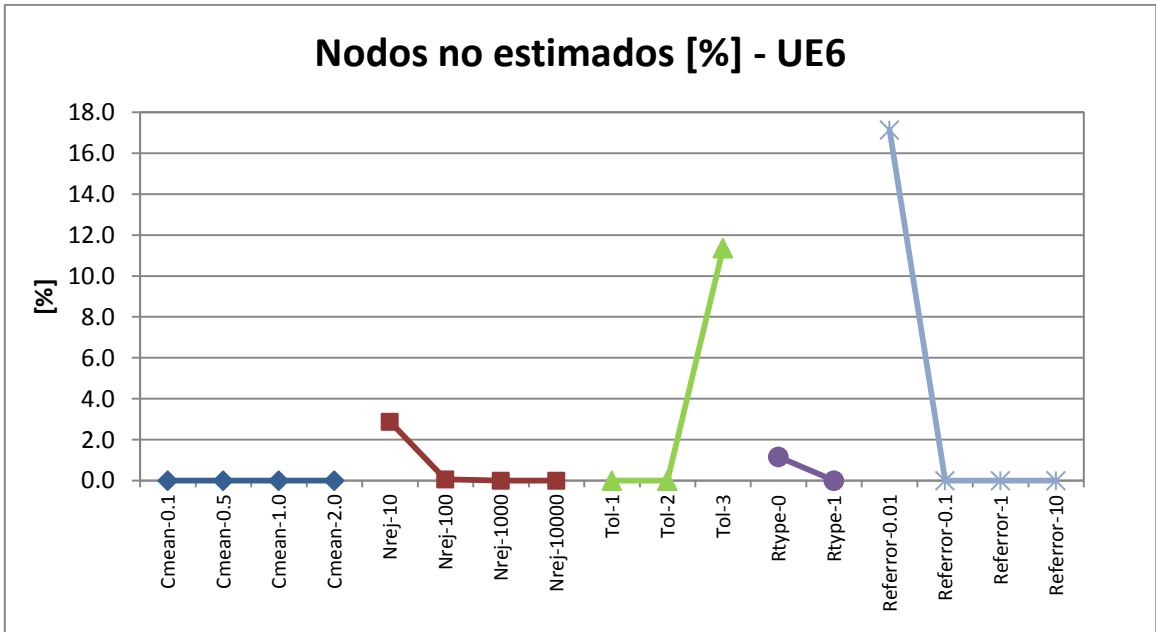


**Gráfico 259: Nodos no simulados medio – UE 2.**





**Gráfico 260: Nodos no simulados medio – UE 3.**



**Gráfico 261: Nodos no simulados medio – UE6.**