

Método UPD-L para estimar valores de una variable geominera medidos en un conjunto de puntos de R^n

Arístides Alejandro Legrá-Lobaina

Resumen

Se describió un algoritmo matemático para estimar en el punto P el valor de una variable dependiente U cuando se conocen valores puntuales de la misma medidos en m puntos P_i de R^n . La estimación se basó en una combinación lineal de los productos de cada U_i por una potencia de la distancia euclidiana de P a P_i . Se ilustra el uso de este estimador para la modelación de curvas, superficies e hiper-superficies geomineras.

Palabras clave: Modelación; estimador; interpolación; extrapolación; distancia Euclidiana.

UPD-L method to estimate values of geomining variables measured in a set of points of R^n

Abstract

Is described the mathematical algorithm to estimate in the point P then value of a dependent variable U when punctual values are measured in m points of R^n . The estimation is based on a lineal combination of the products of each U_i by a power of the euclidian distance from P to P_i . The use of this estimator is illustrated in the modeling of geo-mining curves, surfaces and hiper-surfaces.

Keywords: Modeling; estimator; interpolation; extrapolation; euclidean distance.

1. INTRODUCCIÓN

La modelación del comportamiento de cualquier propiedad de interés U en un dominio espacial consiste en construir un algoritmo matemático capaz de generar, mediante estimaciones (pronósticos), una red de puntos tales que faciliten el estudio del comportamiento de U en ese dominio y, tal vez, en un entorno adyacente.

En la actualidad son varios y disímiles los métodos que se proponen para realizar estas modelaciones. Todos ellos tienen o carecen de una o más de las siguientes propiedades:

1. Ser estimadores exactos (si se estima U en un punto donde se realizó una medición, el resultado de la estimación es el valor medido de U).
2. No presentar grandes oscilaciones en el comportamiento de las estimaciones en los puntos entre nodos.
3. Tener la capacidad de interpolar y extrapolar adecuadamente.
4. Ser capaces de calcular satisfactoriamente en espacios reales de n dimensiones.
5. Presentarse en forma de una expresión analítica.
6. Ser sencillo y algorítmicamente eficiente en su desarrollo y aplicación.
7. Disponer de expresiones para evaluar el error de estimación.

El objetivo del presente trabajo es describir un método (denominado UPD-L) que se basa en obtener una combinación lineal de los productos de la variable dependiente U por una potencia de la distancia euclidiana entre puntos y que cumple con varias de las propiedades mencionadas.

2. MÉTODO UPD-L

Sea un conjunto D de $m > 1$ puntos de un hiperespacio real de $n+1$ dimensiones:

$$D = \{(P_i; U_i) / i = 1, 2, 3, \dots, m; P_i = (x_{i1}; x_{i2}; \dots; x_{in};) \in R^n; U_i \in R\}$$

Donde, además: $P_i \neq P_j$ para $i \neq j$.

Paso 1:

Obtener $\text{Min} = \underset{i=1,2,\dots,m}{\text{Mínimo}}\{U_i\}$ y calcular el valor de: $\text{MinL} = \text{DespL} - \text{Min}$,

donde **DespL** es un número real positivo de manera que puedan resolverse los sistemas de ecuaciones lineales que se describen más

adelante. El valor seleccionado o calculado de Despl tendrá especial impacto en los resultados de las estimaciones.

Paso 2:

Obtener el conjunto de puntos:

$$D_T = \left\{ (P_i; V_i) / i = 1, 2, 3, \dots, m; P_i = (x_{i1}; \dots; x_{in};) \in R^n; V_i = U_i + \text{Min}L \right\}$$

Paso 3:

Resolver en L el sistema de ecuaciones lineales:

$$\begin{bmatrix} 0 & V_2 d_{12}^q & V_3 d_{13}^q & V_4 d_{14}^q & \dots & V_m d_{1m}^q \\ V_1 d_{21}^q & 0 & V_3 d_{23}^q & V_4 d_{24}^q & \dots & V_m d_{2m}^q \\ V_1 d_{31}^q & V_2 d_{32}^q & 0 & V_4 d_{34}^q & \dots & V_m d_{3m}^q \\ \dots & \dots & \dots & \dots & \dots & \dots \\ V_1 d_{m1}^q & V_2 d_{m2}^q & V_3 d_{m3}^q & V_4 d_{m4}^q & \dots & 0 \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ \dots \\ L_m \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ \dots \\ V_m \end{bmatrix}$$

Donde d_{ij} es la distancia euclidiana entre los puntos P_i y P_j de R^n , la cual está dada por:

$$d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{in} - x_{jn})^2 + \delta^2}$$

En este caso, q (mayor que 0 y diferente de 2) es la potencia a la que se elevará la distancia euclidiana y δ es un número real denominado factor de suavización.

Paso 4:

Escribir la fórmula para estimar el valor de $U_d(P)$ para cualquier punto $P = (x_1, x_2, \dots, x_n)$:

$$U_d(P) = L_1 V_1 d_{PP_1}^q + L_2 V_2 d_{PP_2}^q + L_3 V_3 d_{PP_3}^q + \dots + L_m V_m d_{PP_m}^q - \text{Min}L$$

Paso 5:

Obtener (evaluando en la fórmula del paso 5) el conjunto de puntos:

$$C = \left\{ (P_i; W_i) / i = 1, 2, 3, \dots, m; P_i = (x_{i1}; \dots; x_{in};) \in R^n; W_i = U_d(P_i) - U_i \right\}$$

Paso 6:

Obtener $\text{Min} = \underset{i=1,2,\dots,m}{\text{Mínimo}} \{W_i\}$ y calcular el valor de: $\text{Min}C = \text{Despl} - \text{Min}$.

Paso 7:

Obtener el conjunto de puntos:

$$C_T = \{ (P_i; O_i) / i = 1, 2, 3, \dots, m; P_i = (x_{i1}, \dots, x_{in};) \in R^n; O_i = W_i + MinC \}$$

Paso 8:

Resolver en K el sistema de ecuaciones lineales:

$$\begin{bmatrix} 0 & O_2 d_{12}^q & O_3 d_{13}^q & O_4 d_{14}^q & \dots & O_m d_{1m}^q \\ O_1 d_{21}^q & 0 & O_3 d_{23}^q & O_4 d_{24}^q & \dots & O_m d_{2m}^q \\ O_1 d_{31}^q & O_2 d_{32}^q & 0 & O_4 d_{34}^q & \dots & O_m d_{3m}^q \\ \dots & \dots & \dots & \dots & \dots & \dots \\ O_1 d_{m1}^q & O_2 d_{m2}^q & O_3 d_{m3}^q & O_4 d_{m4}^q & \dots & 0 \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \\ K_3 \\ \dots \\ K_m \end{bmatrix} = \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ \dots \\ O_m \end{bmatrix}$$

Donde d_{ij} es la distancia euclidiana entre los puntos P_i y P_j de R^n , la cual está dada por:

$$d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{in} - x_{jn})^2}$$

Paso 9:

Escribir la fórmula para estimar el valor de $U_c(P)$ para $P=(x_1, x_2, \dots, x_n)$:

$$U_c(P) = K_1 O_1 d_{PP_1}^q + K_2 O_2 d_{PP_2}^q + K_3 O_3 d_{PP_3}^q + \dots + K_m O_m d_{PP_m}^q - MinC$$

Paso 10:

Se establece que: $U(P) = U_d(P) - U_c(P)$

El valor $U_c(P)$ constituye una **corrección** del efecto de distorsión del valor estimado $U_d(P)$. La escala de $U_c(P)$ depende de $DespL$ y puede ser en el futuro uno de los elementos para evaluar el error de la estimación.

2.1. Aplicaciones sencillas

En todos los casos se ha tomado $DespL=1$.

Caso 1: Modelar una curva. Sean los datos ($m=10, n=1$) descritos en la Tabla 1.

Tabla 1. Valores $(X_i, Y_i), i=1, 2, \dots, 10$

X	Y	X	Y	X	Y
1	-3	5	6	8	5
2	4	6	8	9	9
3	5	7	7	10	20
4	4				

Al aplicar el método ($q=1$ y $\delta=0$) se obtiene que $MinL=4$, $MinC=1$ y las expresiones:

$$\begin{aligned}
 Y_d = & (4,888888888888874)(-3+MinL)(d_{1P}) \\
 & (-0,374999999999978)(4+MinL)(d_{2P}) \\
 & (-0,111111111111114)(5+MinL)(d_{3P}) \\
 & (0,1875)(4+MinL)(d_{4P}) \\
 & (0)(6+MinL)(d_{5P}) \\
 & (-0,125)(8+MinL)(d_{6P}) \\
 & (-0,0454545454545471)(7+MinL)(d_{7P}) \\
 & (0,333333333333334)(5+MinL)(d_{8P}) \\
 & (0,269230769230763)(9+MinL)(d_{9P}) \\
 & (-0,171296296296294)(20+MinL)(d_{10P}) - MinL \\
 Y_c = & (0,111111111110927)(2,53130849614536E-13+MinC)(d_{1P}) \\
 & (0)(-6,3948846218409E-14+MinC)(d_{2P}) \\
 & (0)(-3,19744231092045E-14+MinC)(d_{3P}) \\
 & (0)(-4,44089209850063E-14+MinC)(d_{4P}) \\
 & (0)(-5,6843418860808E-14+MinC)(d_{5P}) \\
 & (0)(-6,75015598972095E-14+MinC)(d_{6P}) \\
 & (0)(-9,2370555648813E-14+MinC)(d_{7P}) \\
 & (0)(-1,47437617670221E-13+MinC)(d_{8P}) \\
 & (0)(-7,46069872548105E-14+MinC)(d_{9P}) \\
 & (0,111111111111)(-1,70530256582424E-13+MinC)(d_{10P}) - MinC
 \end{aligned}$$

Al aplicar las expresiones a una red de puntos entre $x=1$ y $x=10$ (interpolación) se obtiene el gráfico que se muestra en la Figura 1.

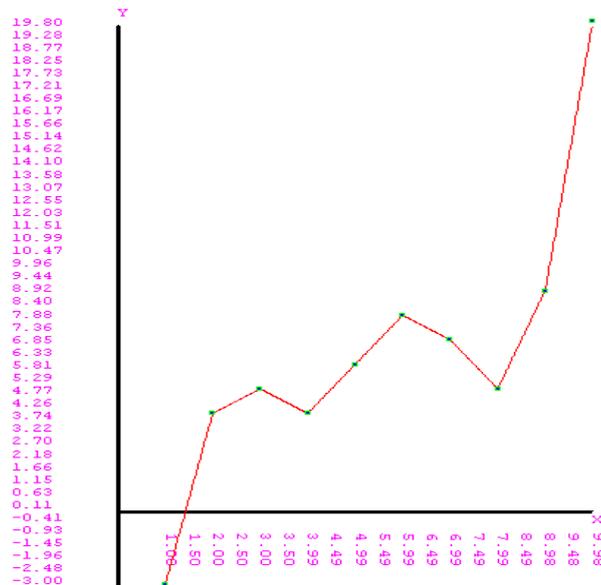


Figura 1. Curva estimada mediante UPD-L.

Nótese que cuando hay tres o más puntos colineales, en cualquiera de los modelos, se anulan los coeficientes de los puntos interiores del conjunto de puntos colineales.

Caso 2: Modelar una superficie. Los datos ($m=4$, $n=2$) en la Tabla 2.

Tabla 2. Valores (X_{1i}, Y_{1i}, U_i) , $i=1, \dots, 4$

X_1	X_2	U
0	0	3
1	0	2
0	1	4
1	2	1

Para $q=1$ y $\delta=0$ se obtiene que $\text{MinL}=0$, $\text{MinC}=1$ y las expresiones:

$$Z_d = (0,22427223071011)(3+\text{MinL})(d_{1P}) \\ (-0,343179100908368)(4+\text{MinL})(d_{2P}) \\ (0,359212573827628)(2+\text{MinL})(d_{3P}) \\ (1,63424873159007)(1+\text{MinL})(d_{4P}) - \text{MinL}$$

$$Z_c = (0,249758004607877)(-8,88178419700125E-16+\text{MinC})(d_{1P}) \\ (0,0462319545286316)(-8,88178419700125E-16+\text{MinC})(d_{2P}) \\ (0,188071133331337)(8,88178419700125E-16+\text{MinC})(d_{3P}) \\ (0,342430069141358)(0+\text{MinC})(d_{4P}) - \text{MinC}$$

En la Figura 2 se muestra el gráfico obtenido al aplicar el modelo en una red rectangular de puntos, contenidos en la región plana, delimitada por una frontera que contiene a la frontera natural de los puntos.

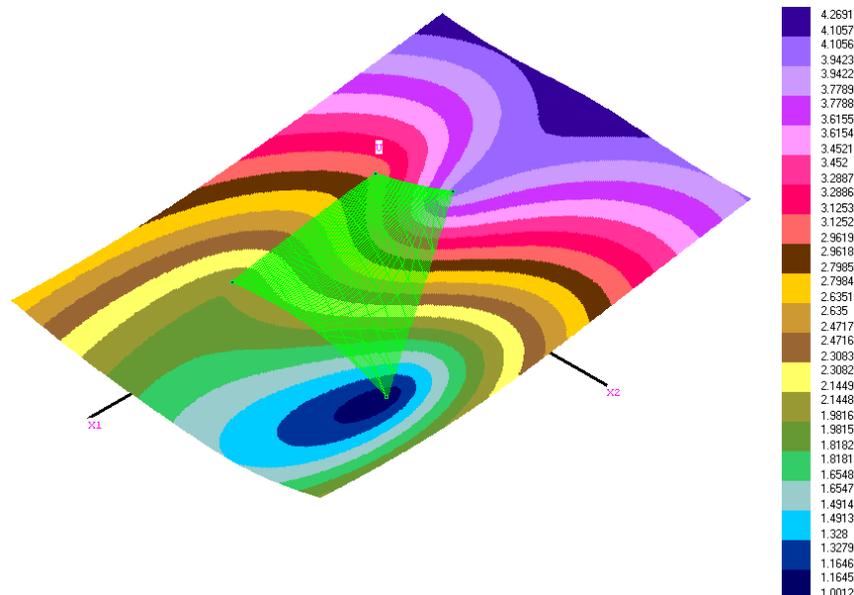


Figura 2. Superficie estimada mediante UPD-L.

Nótese que se ha estimado coherentemente en puntos interiores a la frontera (interpolación) y los valores están entre 1 y 4. También se ha estimado a U en puntos fuera de la frontera (extrapolación), pero en este caso los valores estimados crecen en la medida en que los puntos $(X_1; X_2)$ se alejan de la frontera natural determinada por los datos.

Mediante un gráfico de isofranjas se muestran en la Figura 3 los valores estimados de la corrección U_c .

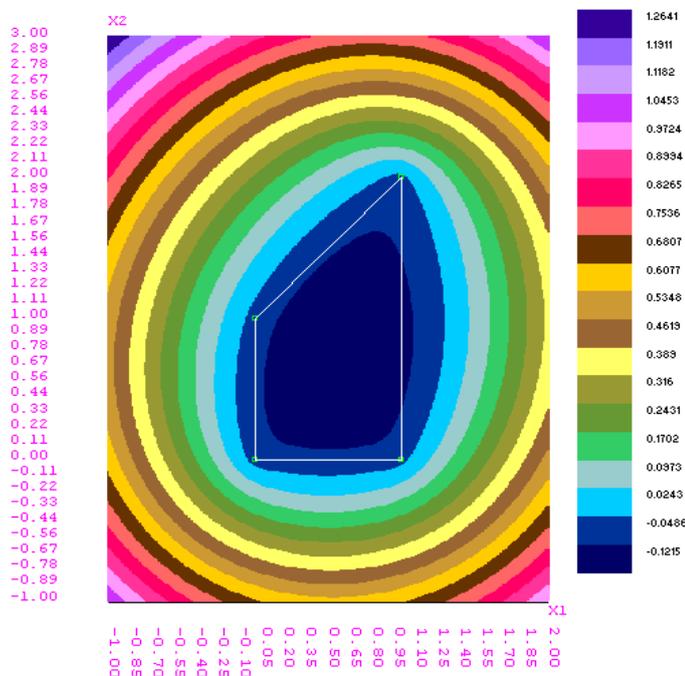


Figura 3. Isofranjas de valores de corrección.

Con el fin de completar el propósito de este trabajo a continuación se muestran (Figura 4) los gráficos de isofranjas obtenidos por el método UPD-L ($q=1$ y $\delta=0$) y por otros cinco métodos de estimación bien conocidos: Splines Bicúbico y Bilineal (Xia *et al.* 2013), Inverso del Cuadrado de la Distancia (Babak y Deutsch 2008), Kriging Puntual (Geovariances 2010) y finalmente Interpolación Lineal con Triangulación (Legrá-Lobaina, Atanes-Beatón & Guilarte-Fuentes 2014), a partir de los datos de concentración de Ni (%) en un bloque de un yacimiento laterítico y usando configuraciones genéricas. Nótese la aparente similitud de los resultados de los métodos UPD-L, Splines Bicúbico, Bilineal y Kriging.

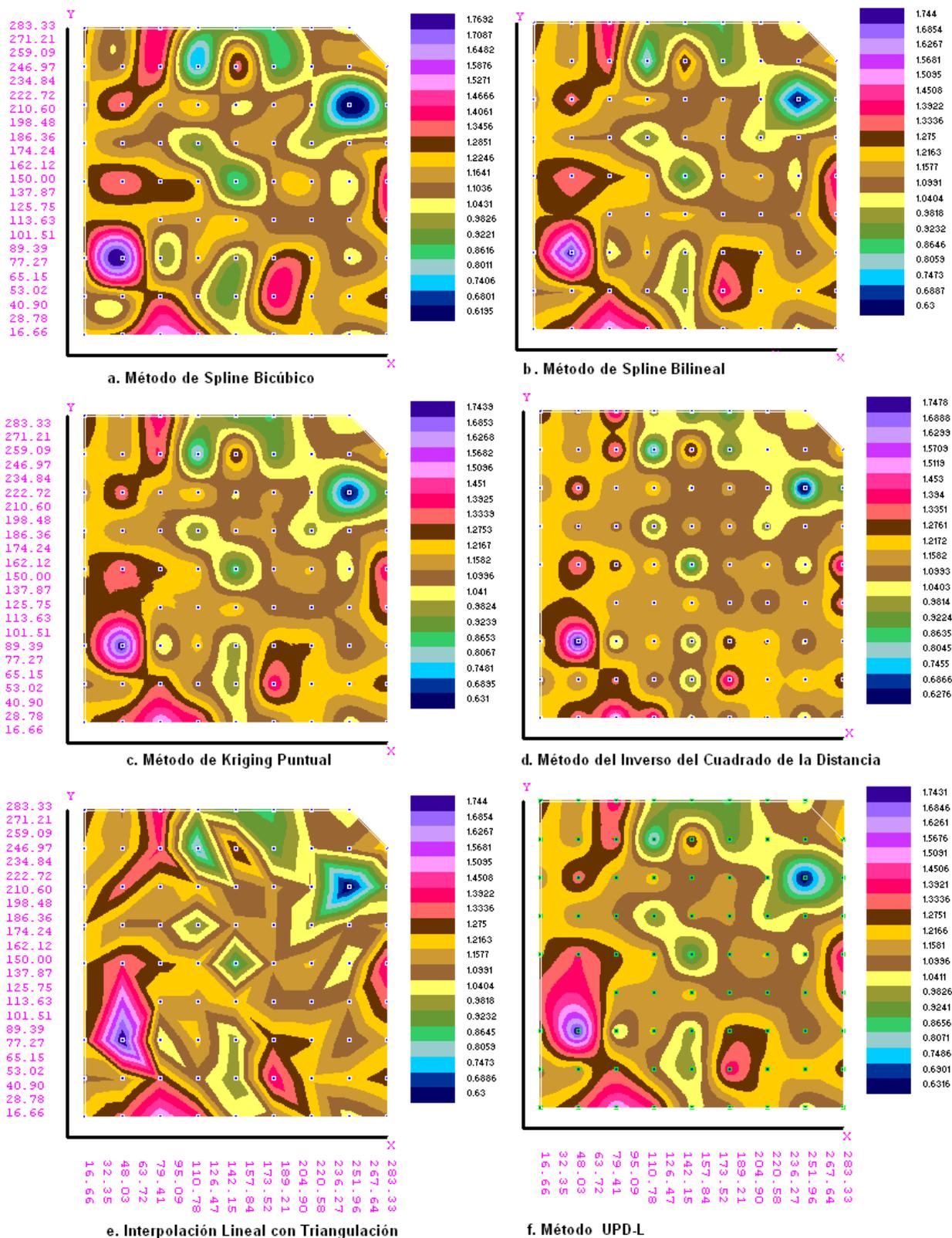


Figura 4. Isofranjas para comparar los modelos de la concentración del Ni en un bloque de un yacimiento laterítico, obtenidos por seis métodos.

Caso 3: Modelar una hiper-superficie; los datos (n=3) en la Tabla 3.

Tabla 3. Valores $(X_{1i}, X_{2i}, X_{3i}, U_i)$, $i=1, \dots, 8=m$

X1	X2	X3	U
0	0	0	3
0	0	1	4
1	0	0	5
1	0	1	6
0	1	0	5
0	1	1	4
1	1	0	3
1	1	1	2

Al aplicar el método ($q=1,5$ y $\delta=0,25$) se obtiene que $\text{MinL}=-1$, $\text{MinC}=1,39991266980416$ y el modelo:

$$U_d = (1,71494814422119)(3+\text{MinL})(d_{1P}) \\ (0,406721246494629)(4+\text{MinL})(d_{2P}) \\ (-0,247392202368689)(5+\text{MinL})(d_{3P}) \\ (-0,639860271686667)(6+\text{MinL})(d_{4P}) \\ (-0,716201335960565)(5+\text{MinL})(d_{5P}) \\ (-0,218357598294565)(4+\text{MinL})(d_{6P}) \\ (0,777329877037474)(3+\text{MinL})(d_{7P}) \\ (3,76439230303347)(2+\text{MinL})(d_{8P}) - \text{MinL}$$

$$U_c = (-0,483340244579029)(0,428737036055315+\text{MinC})(d_{1P}) \\ (-0,127053592041432)(0,152520467435469+\text{MinC})(d_{2P}) \\ (0,383458101964458)(-0,123696101184342+\text{MinC})(d_{3P}) \\ (1,17599337268755)(-0,399912669804165+\text{MinC})(d_{4P}) \\ (1,1152433403972)(-0,358100667980288+\text{MinC})(d_{5P}) \\ (0,360581035986942)(-0,0818840993604533+\text{MinC})(d_{6P}) \\ (-0,132577906006455)(0,194332469259346+\text{MinC})(d_{7P}) \\ (-0,4800843920506)(0,470549037879198+\text{MinC})(d_{8P}) - \text{MinC}$$

En la Figura 5 se muestran gráficos de tres y ocho isofranjas obtenidas al realizar cortes horizontales al modelo en una red de puntos contenidos en una región 3D, en forma de ortoedro, que contiene los datos.

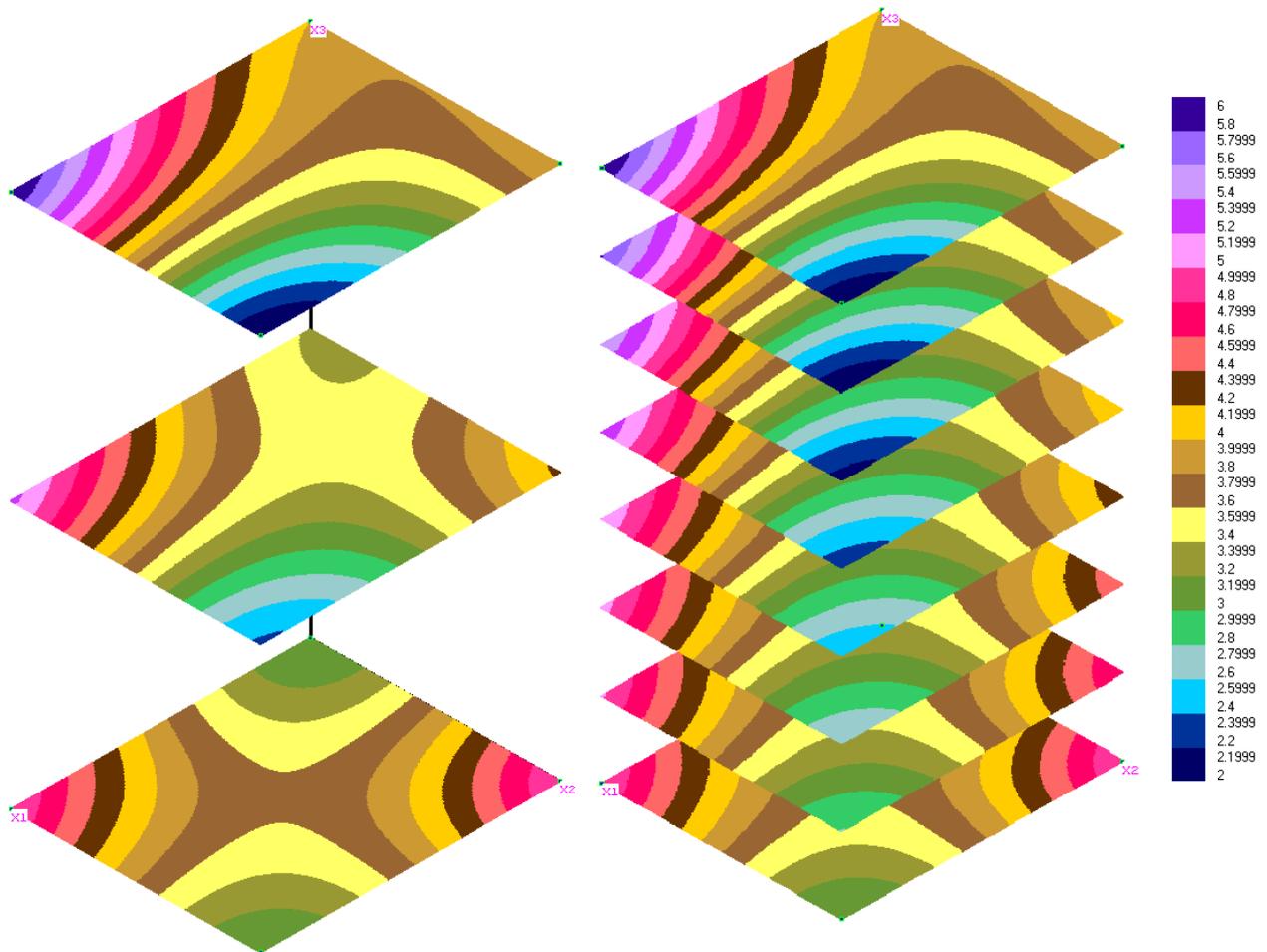


Figura 5. Hiper-superficie estimada con el método UPD-L

3. CONCLUSIONES

El método UPD-L es un interpolador exacto cuando $\delta \neq 0$ y para las configuraciones usadas no presenta oscilaciones pronunciadas en el comportamiento de las estimaciones en los puntos entre nodos. Tiene la capacidad de interpolar y extrapolar en espacios reales de n dimensiones que contengan a los datos. Se presenta en forma de dos expresiones analíticas que resultan fáciles de calcular, su aplicación es algorítmicamente eficiente.

Todavía no se ha definido una expresión para evaluar el error de estimación con la capacidad de cuantificarlo de manera precisa tal que, comparando los valores de los errores en diferentes puntos, se pueda determinar cuáles son las estimaciones con valor práctico.

Aún es necesario profundizar en ciertas propiedades matemáticas de este método, tales como su comportamiento ante ciertas situaciones matemáticas específicas. Por ejemplo, este método es capaz de generar un modelo con solo dos puntos de coordenadas diferentes; sin embargo,

es ineficaz cuando dos de los puntos dados como datos tienen iguales coordenadas. Un aspecto de especial interés será determinar criterios prácticos para determinar bajo qué condiciones las extrapolaciones son pronósticos válidos.

También será necesario estudiar generalizaciones (por ejemplo, vectoriales) y variantes del método que permitan modelar fenómenos naturales variados con características anisotrópicas especiales. Puede suceder que ciertas superficies generadas por L tengan alguna relación especial con los modelos físicos de una superficie elástica.

4. REFERENCIAS

- XIA, P.; TAHARA, T.; KAKUE, T.; AWATSUJI, Y.; NISHIO, K.; URA, S.; KUBOTA, T. & MATOBA, O. 2013: Performance comparison of bilinear interpolation, bicubic interpolation, and B-spline interpolation in parallel phase-shifting digital holography. *Optical Review* 20(2): 193-197.
- BABAK, O. & DEUTSCH, C. 2008: Statistical approach to inverse distance interpolation. *Stochastic Environmental Research and Risk Assessment* 23(5): 543-553.
- GEOVARIANCES. 2010: ISATIS Software. Beginner's Guide. Versión 9.0. Geovariances & Ecole des Minas de Paris, France.
- LEGRÁ-LOBAINA, A. A.; ATANES-BEATÓN, D. M. & GUILARTE-FUENTES, C. 2014: Contribución al método de interpolación lineal con triangulación de Delaunay. *Minería y Geología* 30(2): 58-72.

Arístides Alejandro Legrá Lobaina. alegra@ismm.edu.cu

Doctor en Ciencias Técnicas. Profesor Auxiliar.

Departamento de Matemáticas.

Instituto Superior Minero Metalúrgico, Holguín, Cuba