

A virtual laboratory for the teaching of Kriging Un laboratorio virtual para la enseñanza del Kriging

Fernando Giménez-Palomares,
Juan A. Monsoriu-Serra
UNIVERSITAT POLITÈCNICA DE VALÈNCIA
fgimenez@upv.es, jmonsori@fis.upv.es

José Q. Cuador-Gil
UNIVERSIDAD DE PINAR DEL RIO, CUBA
cuador@info.upr.edu.cu

Abstract

The Kriging is an interpolation method which in practical terms is to find the best estimate of the value of a variable. For this weighting performed, ie a weight attached to each measured value so as to minimize the variance of the estimation error. In this paper we present a virtual laboratory to study from a teaching perspective this technique in the case of dimension two.

El Kriging es un método de interpolación que en términos prácticos consiste en encontrar la mejor estimación posible del valor de una variable. Para esto efectúa una ponderación, es decir, atribuye un peso a cada valor medido de manera que minimice la varianza del error de estimación. En este trabajo presentamos un laboratorio virtual que permite estudiar desde un punto de vista docente esta técnica en el caso particular de dimensión dos.

Keywords: Kriging, semivariogram, geostatistics, virtual laboratory.

Palabras clave: Kriging, semivariograma, geoestadística, laboratorio virtual.

1 Introducción

EL Krigado puede interpretarse como un método que permite interpolar datos de variables distribuidas en el espacio en localizaciones donde no existen valores medidos a partir de los medidos en un conjunto limitado de localizaciones utilizando técnicas estadísticas. Existen muchas aplicaciones comerciales (Geo-EAS, GSLIB, GSTAT, VARIOWIN, VESPER, R+, SADA, WINGSLIB, GS+, S+, IDRISI, SURFER, ARCGIS GEOSTATISTICAL ANALIST, EASYKRIG,...) para el Krigado, pero todas ellas adolecen de un enfoque eminentemente profesional y poco adecuado para su uso como apoyo en la docencia. En el presente trabajo se muestran un laboratorio virtual que permite obtener y representar gráficamente información relacionada con el Krigado a partir de unas variables introducidas por el usuario, orientado preferentemente a su uso en la enseñanza. Como precedente podemos nombrar al artículo (Cuador et al, 2011) en donde también se presenta un laboratorio virtual sobre el Krigado pero enfocado al aspecto analítico del método.

Con la irrupción en el panorama educativo de las nuevas tecnologías y, en particular, con el uso del ordenador en las aulas, se han abierto nuevas posibilidades en la enseñanza de muchas asignaturas de ciencias e ingenierías (León, 1997) y (Bartolome, 2004). En este marco los laboratorios virtuales son una buena herramienta para la transmisión de conocimientos, fomentan la interactividad, permiten trabajar fácilmente con modelos que representan fenómenos matemáticos y físicos y pueden afianzar los conocimientos adquiridos por los alumnos (Vazquez, 2009). El laboratorio virtual que presentamos está constituido por una interfaz gráfica de usuario (GUI) de MATLAB[®] (TheMatworks, (n.d.)) y (Matpic, (n.d.)) diseñada con el objetivo de estudiar gráfica y analíticamente el Krigado.

2 Krigado

En 1971 el geo-matemático francés Georges Matheron y el ingeniero de minas sudafricano Daniel G. Krige comenzaron a desarrollar un método (conocido como Krigado o Kriging) para la interpolación de datos dispersos con el objeto de ser aplicado en la industria minera, en concreto en la estimación de mineral recobable, cosa que es muy importante ya que la variabilidad local puede afectar la ganancia al explotar una mina. Matheron se encargó de formalizar la mayor parte de los conceptos de la teoría que llamo Geoestadística. El propio Matheron la definió como la aplicación del formalismo de las variables aleatorias al reconocimiento y estimación de variables naturales. (Matheron, 1965) y (Matheron, 1970) constituyen las referencias clásicas sobre el método que desarrollaron.

Los ámbitos de aplicación de la Geoestadística son: minería, industria petrolífera, geología, meteorología, climatología, cartografía de suelos, edafología, hidrología, silvicultura, ecología, patología vegetal, epidemiología, entomología, ciencias ambientales, salud pública, etc.

2.1 El Krigado y la variabilidad espacial

Sean $Z(x_i)$, $i = 1, 2, \dots, n$ un conjunto de valores medidos correspondiente a un muestreo espacial de puntos $\{x_i\}$, deseamos estimar un valor de Z en un punto x_0 . Se propone una combinación lineal de pesos asignados a cada localización con los valores $Z(x_i)$

$$Z^*(x_0) = \sum_{i=0}^n \lambda_i Z(x_i),$$

donde $Z^*(x_0)$ es el valor estimado de forma puntual en la localización x_0 . λ_i son los pesos o ponderadores, obtenidos de tal forma que proporcione un estimador:

1. insesgado $E[Z^*(x) - Z(x)] = 0$.
2. de varianza mínima $Var[Z^*(x) - Z(x)] = 0$ mínima.

Para lograr lo anterior el Krigado exige como primera etapa y fundamental un modelo que refleje las características de la variabilidad y correlación espacial de los datos medidos: el semivariograma. Este proceso implica el cálculo del semivariograma experimental y el ajuste a éste de modelos teóricos de semivariogramas.

El semivariograma experimental se define como la semi-varianza de los incrementos de la variable regionalizada, esto es:

$$\gamma(h) = \frac{1}{2Np(h)} \sum_{i=1}^{Np(h)} [Z(x_i) - Z(x_i + h)]^2,$$

donde $Np(h)$ es el número de pares a distancia h .

Los modelos teóricos de semivariogramas a ajustar son modelos conocidos y autorizados (Armstrong, 1998). El modelo de variabilidad finalmente obtenido es usado por el Krigado en el proceso de interpolación. El análisis de variabilidad y correlación espacial es un proceso que distingue a este interpolador, de ahí su popularidad en diferentes campos de las ingenierías, fundamentalmente en el campo de las geociencias.

2.2 Ecuaciones del Krigado Lineal

Para variables estacionarias, es decir aquellas que la media no depende de la posición. Se definen dos variantes de interpolación: el Krigado Simple para datos de esperanza nula o conocida y el Krigado Ordinario para datos con esperanza desconocida. A continuación se presenta el estimador y el sistema Krigado para estos casos:

Krigado Simple

$$\text{Estimador: } Z^*(x) = \sum_{i=1}^n \lambda_i Z(x_i) + m(1 - \lambda_i).$$

$$\text{Sistema: } \sum_{i=1}^n \lambda_i C(x_i, y_i) = C(x_j, x_0), \quad j = 1, 2, \dots, n.$$

Krigado Ordinario

$$\text{Estimador: } Z^*(x) = \sum_{i=1}^n \lambda_i Z(x_i).$$

$$\text{Sistema: } \sum_{i=1}^n \lambda_i C(x_i, y_i) - \mu = C(x_j, x_0), \quad j = 1, 2, \dots, n.$$

$$\sum_{i=1}^n \lambda_i = 1.$$

donde:

$Z^*(x)$ es el valor estimado en la posición x ,

λ_i son los pesos o ponderadores,
 m es la media de los datos medidos,
 x_i, x_j son las localizaciones de los datos,
 x_0 es la localización de la posición a estimar,
 $C(x_i, x_j)$ es la covarianza entre las localizaciones x_i y x_j ,
 $C(x_i, x_0)$ es la covarianza entre las localizaciones x_i y x_0 ,
 μ es el multiplicador de Lagrange, sin usos prácticos.

Los sistemas de Krigado anteriores tiene solución única, los mismos pueden ser escritos de forma matricial de la siguiente forma:

$$\begin{array}{ccc}
 \text{Krigado Simple} & & \text{Krigado Ordinario} \\
 \left(\begin{array}{ccccc} C_{11} & \dots & C_{1j} & \dots & C_{1N} \\ \vdots & & \vdots & & \vdots \\ C_{i1} & \dots & C_{ij} & \dots & C_{iN} \\ \vdots & & \vdots & & \vdots \\ C_{N1} & \dots & C_{Nj} & \dots & C_{NN} \end{array} \right) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_N \end{pmatrix} = \begin{pmatrix} C_{1,0} \\ \vdots \\ C_{i,0} \\ \vdots \\ C_{N,0} \end{pmatrix} & & \left(\begin{array}{cccccc} C_{11} & \dots & C_{1j} & \dots & C_{1N} & 1 \\ \vdots & & \vdots & & \vdots & \\ C_{i1} & \dots & C_{ij} & \dots & C_{iN} & 1 \\ \vdots & & \vdots & & \vdots & \\ C_{N1} & \dots & C_{Nj} & \dots & C_{NN} & 1 \\ 1 & \dots & 1 & \dots & 1 & 0 \end{array} \right) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_N \\ \mu \end{pmatrix} = \begin{pmatrix} C_{1,0} \\ \vdots \\ C_{i,0} \\ \vdots \\ C_{N,0} \\ 1 \end{pmatrix}
 \end{array}$$

2.3 Semivariogramas y modelos La siguiente figura muestra la gráfica típica de un semivariograma

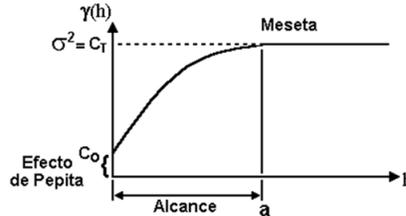


Figura 1: Semivariograma.

Se definen las siguientes variables:

- Alcance: distancia a la cual el variograma se estabiliza
- Meseta: valor constante que toma el variograma en distancias mayores al rango
- Efecto pepita: valor no-nulo de semivariancia al que en ocasiones se llega al extrapolar la curva del semivariograma para la distancia cero.

No es posible usar directamente el semivariograma $\gamma(h)$ experimental en el proceso de estimación geoestadística. Es necesario ajustar a modelos teóricos conocidos (y autorizados) al $\gamma(h)$ experimental que deben de ser definidos positivos. Los modelos más usados son los siguientes:

- Esférico: $\gamma(h) = \begin{cases} C(\frac{3h}{2a} - \frac{h^3}{2a^3}), & h \leq a \\ C, & h > a \end{cases}$
- Exponencial: $\gamma(h) = C(1 - e^{-h/a}), \quad h \geq 0.$

Se cumple que para $h = a'$, $\gamma(h) = 0.95C$.

- Gaussiano: $\gamma(h) = C(1 - e^{-h^2/a^2})$, $h \geq 0$.
- Función potencia: $\gamma(h) = h^\alpha$, $0 < \alpha \leq 2$.
- Lineal: $\gamma(h) = (C/a)h$, donde $C = \sigma^2$.

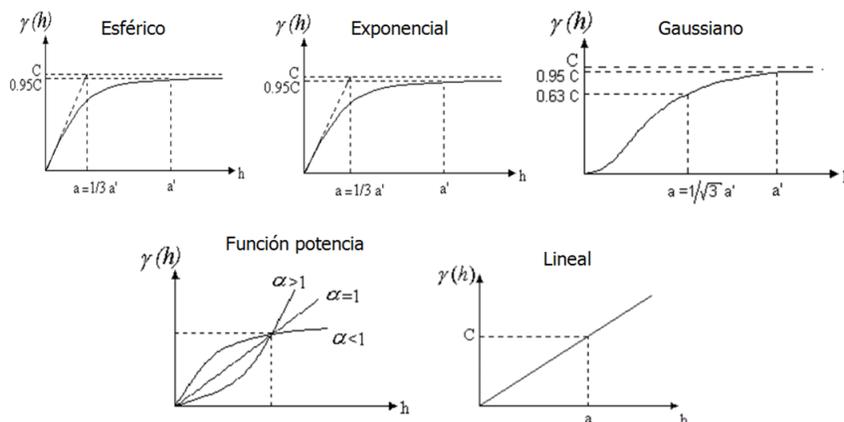


Figura 2: Modelos de semivariogramas teóricos.

2.4 Factores a considerar en la asignación de los ponderadores

- Cercanía a la posición que está siendo estimada. Se puede utilizar todos los datos disponibles (vecindad única) o sólo una parte de ellos (vecindad móvil)
- Redundancia entre los valores de datos
- Continuidad o variabilidad espacial
- Isotropía y anisotropía: si existe o no una dirección preferencial. Si se sospecha que puede ocurrir el fenómeno de anisotropía es interesante realizar semivariogramas considerando por separado varias direcciones del espacio (semivariogramas direccionales).

Para más detalles sobre el método del Krigado se pueden consultar las referencias (Isaaks y Srivastava, 1989), (Olea, 1999) y (Yarus y Chambers, 2006).

3 El laboratorio virtual Krig

Hemos desarrollado una interfaz gráfica de usuario (GUI) de MATLAB[®] para estudiar el Krigado 2-dimensional. El aspecto gráfico de la aplicación puede verse en la Figura 3.

En la parte lateral izquierda se encuentran los campos donde el usuario introduce los parámetros de entrada. En la parte lateral derecha se encuentran tres ventanas gráficas en donde se generan varias gráficas que proporcionan información sobre el Krigado. En la parte lateral izquierda por abajo aparece también un recuadro con los resultados numéricos.

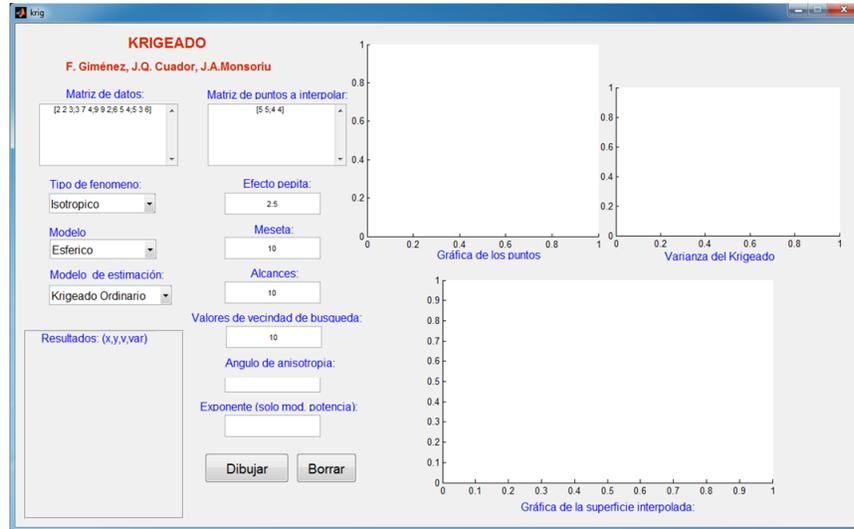


Figura 3: El laboratorio virtual Krig.

Los parámetros de entrada de la aplicación son:

- Matriz de datos: las filas tienen la forma $(x_i, y_i, Z(x_i))$
- Matriz de puntos a interpolar: las filas tienen la forma (x_i, y_i)
- Tipo de fenómeno: isotrópico o anisotrópico
- Modelo: esférico, exponencial, gaussiano, función potencia o lineal
- Modelo de estimación: Krigado ordinario o simple
- Efecto pepita
- Meseta
- Alcances
- Valores de vecindad de búsqueda (número total)
- Ángulo de anisotropía (sólo si la hay)
- Exponente: se refiere al valor de si se trata del modelo de la función potencia (sólo si se trata de este modelo)

Como salida se muestran:

- Ventana con la gráfica correspondiente a los datos conocidos (en verde) y los interpolados (en rojo). Encima de cada bolita aparece el valor de Z en dicho punto.
- Ventana con la gráfica en pseudocolores de la superficie interpolada $(x, y) \rightarrow Z(x, y)$ y los puntos correspondientes a los datos iniciales
- Ventana con la gráfica en pseudocolores de la varianza del Krigado $(x, y) \rightarrow var(Z(x, y))$ (lo

cual proporciona una estimación del error cometido)

- Tabla de valores con los resultados de la interpolación.

3.1 Ejemplo

La siguiente tabla muestra el porcentaje en tanto por ciento en contenido en fosforo de una finca correspondiente a un muestreo que se ha realizado:

Tabla 1:

Coordenadas (hm)			Coordenadas (hm)		
		% Fosforo			% Fosforo
1.6294	1.8116	0.3759	1.3115	0.0714	0.2631
0.2540	1.8268	0.3928	1.6983	1.8680	0.3349
1.2647	0.1951	0.2821	1.3575	1.5155	0.3567
0.5570	1.0938	0.2106	1.4863	0.7845	0.3049
1.9150	1.9298	0.2120	1.3110	0.3424	0.2776
0.3152	1.9412	0.3849	1.4121	0.0637	0.2335
1.9143	0.9708	0.3709	0.5538	0.0923	0.3850
1.6006	0.2838	0.2196	0.1943	1.6469	0.3980
0.8435	1.8315	0.2063	1.3897	0.6342	0.2890
1.5844	1.9190	0.3780	1.9004	0.0689	0.2314

Se desea calcular aproximadamente el contenido en fosforo en los puntos de coordenadas (0.3, 0.4), (1, 0.8) y (1.5, 1.5). Se supone que ha habido un estudio previo del semivariograma correspondiente y se ha determinado los siguientes valores: efecto pepita 0.0001, meseta 0.005, alcances 0.6, valores de vecindad de búsqueda 2 y que es un fenómeno isotrópico.

La siguiente figura muestra el resultado obtenido con el laboratorio virtual Krig usando el Krigado ordinario y modelo esférico.

El usuario fácilmente puede volver a ejecutar el programa con otros datos iniciales y/u otro método de Krigado y observar cual es la influencia que tienen estos cambios en la aproximación obtenida. Por ejemplo para el Krigado Simple y modelo gaussiano se obtiene la figura:

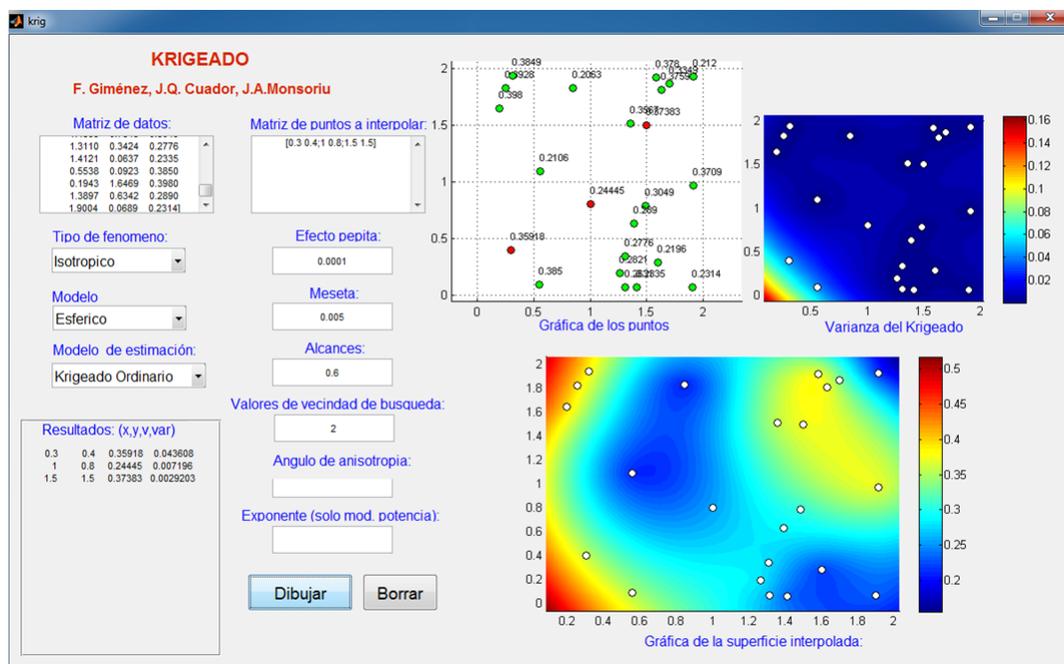


Figura 4: Resultados del ejemplo para Krigado ordinario y modelo esférico.

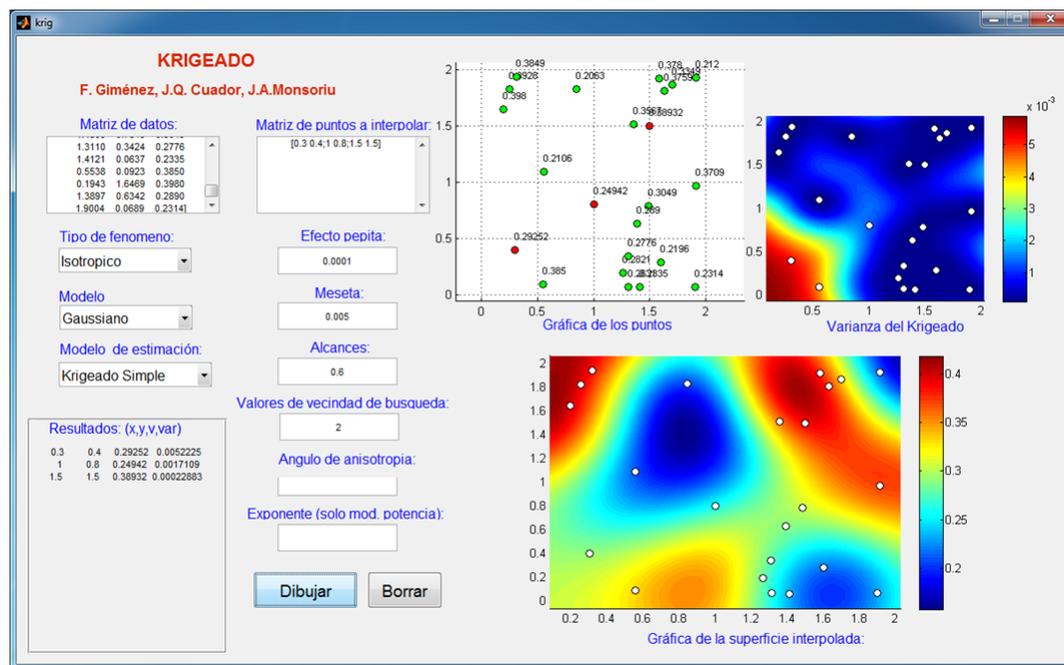


Figura 5: Resultados del ejemplo para Krigado simple y modelo gaussiano.

4 Conclusiones

El laboratorio virtual Krig ofrece a los estudiantes la posibilidad de estudiar de forma interactiva el método del Krigado. Éstos pueden experimentar de manera sencilla la influencia que tienen en la estimación de los datos del muestro de las distintas variables que intervienen en el proceso. Además es posible su utilización para el análisis científico de problemas reales y la posterior discusión de los resultados obtenidos. Creemos que puede ser una buena herramienta docente

para que los alumnos de asignaturas de cálculo numérico en ingenierías puedan afianzar sus conocimientos sobre el tema.

Agradecimientos

Los autores agradecen al Instituto de Ciencias de la Educación de la Universitat Politècnica de València por su ayuda al Equipo de Innovación y Calidad Educativa MOMA.

Referencias

-  [Bartolomé A. R. \(2004\).](#)
Nuevas Tecnologías en el Aula. Guía de Supervivencia.
ICE-Universitat de Barcelona, Ed. Graó de IRIF, S.L.
-  [Cuador J. Q., Monsoriu J. A., Giménez F., Castro J. C. \(2011\).](#)
Promoting mathematical skills using the instructive program Krigeado.
2nd Promotion and Innovation with New Technologies in Engineering
Education (FINTDI 2011), Teruel, IEEE Xplore Digital Library.
-  [Isaaks E. H., Srivastava R. M. \(1989\).](#)
An introduction to Applied Geostatistics.
Ed. Oxford University Press.
-  [León J. A., Vizcarro C. \(1997\).](#)
Nuevas tecnologías para el aprendizaje.
Ediciones Pirámide.
-  [Matheron G. \(1965\).](#)
Les variables régionalisées et leur estimation.
Paris, Ed. Masson.
-  [Matheron G. \(1970\).](#)
La Théorie des Variables Régionalisées et ses Application.
Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau, Fascicule 5,
Ecole de Mines de Paris.
-  [Matpic, \(n.d.\).](#)
<http://www.matpic.com/>
-  [Olea R. A. \(1999\).](#)
Geostatistics for Engineers and Earth Scientists, Kluwer Academic Publishers.
-  [Vázquez C. \(2009\).](#)
Laboratorios Virtuales, Innovación y experiencias educativas.
Revista digital CSIF-enseñanza Andalucía.
-  [The Mathworks, \(2008\).](#)
INC. Matlab R2008a User's Guide.
The Mathworks, INC. Natick, MA. USA.
-  [Yarus J. M., Chambers R. L. \(2006\).](#)
Practical Geostatistics-An Armchair Overview for Petroleum Reservoir Engineers.
Distinguished Author Series, Society of Petroleum Engineers.