

Formulation of a correlated variables methodology for assessment of continuous gas resources with an application to the Woodford play, Arkoma Basin, eastern Oklahoma

Ricardo A. Olea⁽¹⁾, David W. Houseknecht⁽¹⁾, Christopher P. Garrity⁽¹⁾ and Troy A. Cook⁽²⁾

(1) U.S. Geological Survey, 12201 Sunrise Valley Dr., Mail Stop 956 Reston, VA 20192-0002, USA
olea@usgs.gov

(2) U.S. Geological Survey, Box 25046 Denver Federal Center Mail Stop 939 Denver, CO 80225-0459, USA

ABSTRACT

Shale gas is a form of continuous unconventional hydrocarbon accumulation whose resource estimation is unfeasible through the inference of pore volume. Under these circumstances, the usual approach is to base the assessment on well productivity through estimated ultimate recovery (EUR). Unconventional resource assessments that consider uncertainty are typically done by applying analytical procedures based on classical statistics theory that ignores geographical location, does not take into account spatial correlation, and assumes independence of EUR from other variables that may enter into the modeling. We formulate a new, more comprehensive approach based on sequential simulation to test methodologies known to be capable of more fully utilizing the data and overcoming unrealistic simplifications. Theoretical requirements demand modeling of EUR as areal density instead of well EUR. The new experimental methodology is illustrated by evaluating a gas play in the Woodford Shale in the Arkoma Basin of Oklahoma. Differently from previous assessments, we used net thickness and vitrinite reflectance as secondary variables correlated to cell EUR. In addition to the traditional probability distribution for undiscovered resources, the new methodology provides maps of EUR density and maps with probabilities to reach any given cell EUR, which are useful to visualize geographical variations in prospectivity.

Keywords: collocated cokriging, estimated ultimate recovery, gas shale, sequential stochastic simulation

Metodología para la evaluación de recursos de gas para el caso de yacimientos continuos usando múltiples variables correlacionadas, con un estudio de la arcilla Woodford, cuenca de Arkoma, Oklahoma oriental, EEUU

RESUMEN

La evaluación estocástica de la productividad de arcillas gasíferas se ha hecho hasta ahora usando estadística clásica, sin considerar la ubicación de pozos, la correlación espacial de los datos, ni la dependencia entre variables. Hemos usado simulación secuencial para explorar las supuestas bondades de un modelaje más integral, eliminando tales simplificaciones e incluyendo dependencia entre productividad, espesor neto y reflectancia de la vitrinita. Valiéndonos de la arcilla Woodford en Oklahoma, ilustramos prácticamente el modelaje y despliegue de resultados, los que incluyen histograma de recursos probables y mapas tanto de posibles producciones como de probabilidades de alcanzar ciertos niveles de productividad.

Palabras claves: arcillas gasíferas, cokrigado colocalizado, recuperación final por pozo, simulación estocástica secuencial

VERSIÓN ABREVIADA EN CASTELLANO

Introducción

Las arcillas gasíferas son uno de los tipos de roca para los que resulta difícil hacer una estimación de reservas sobre la base de un cálculo del volumen de poros. En estos casos es preferible hacer una evaluación usando como dato la productividad de los pozos en forma de estimaciones de la recuperación final por pozo (EUR) (Cook 2005). Respecto al tipo de modelaje, cuando se trata de tomar en cuenta la incertidumbre, la evaluación de arcillas gasíferas se ha hecho hasta ahora usando métodos propios de la estadística clásica (Crovelli 2000; Klett and Schmoker 2005), los cuales no toman en cuenta la ubicación geográfica de los pozos, la correlación espacial de los datos, ni dependencia entre variables.

Metodología

Dado que estas simplificaciones derivan en una utilización subóptima de la información disponible e ignoran importantes propiedades contenidas en los datos, hemos usado simulación secuencial (e.g. Gómez-Hernández and Cassiraga 1994) para explorar las supuestas bondades de un modelaje más integral. Considerando que cada pozo drena un volumen distinto, la correcta aplicación del nuevo método requiere trabajar con productividad por unidad de extensión de área en vez de la recuperación por pozo. Las áreas de drenaje se subdividen en celdas cuadradas de lado igual al diámetro de la zona estimulada, que es la que contribuye a la producción. La simulación secuencial es capaz de producir una serie de mapas del atributo de interés, llamados realizaciones, que se caracterizan por tener la misma probabilidad de ser el mapa verdadero pero desconocido. Es posible trabajar con una sola variable o con varias de ellas. En el caso multivariable, a la variable de mayor interés, EUR por celda en nuestro caso, se le denomina variable primaria. Si existe información y correlación con otras variables, la implementación usada en este caso (Remy et al. 2009) requiere tener el mapa de la variable secundaria, lo que obliga a generar realizaciones si no es el caso. Los pasos a seguir son:

1. Elegir un tamaño de celda no mayor que el diámetro de la zona fracturada alrededor de un pozo después de estimularlo.
2. Calcular el área de drenaje para cada pozo. Si es pozo es vertical, automáticamente el área de drenaje es igual a una celda. Si el pozo es horizontal, tomar un número de celdas consistentes con la orientación y el la longitud de la zona perforada.
3. Distribuir la producción por pozo asignando a cada celda que está formando parte del área de drenaje del pozo una recuperación igual a la recuperación estimada para el pozo (EUR) dividida por el número de celdas en su área de drenaje.
4. Si las celdas no tienen una distribución geográficamente homogénea, desaglomerar los datos de EUR por celda.
5. Calcular el modelo de semivariograma que mejor se ajusta al semivariograma experimental correspondiente al set de datos de EUR desaglomerados después de transformarlos para que sigan una distribución normal de media 0 y varianza 1 (puntajes normales).
6. Emplear los valores obtenidos para EUR por celda para calcular el coeficiente de correlación con la variable secundaria después de transformar ambas muestras a puntajes normales.
7. Usar simulación secuencial gaussiana para generar al menos 100 realizaciones de la variable secundaria correlacionada con la recuperación por celda.
8. Emplear los resultados de los cinco pasos anteriores para general realizaciones de EUR por celda usando cosimulación secuencial gaussiana, tantas como el número de realizaciones en el paso anterior, usando siempre una realización secundaria diferente. Si, como es comúnmente el caso, se requirió desagregar los datos, usar los datos del Paso 3 como la muestra primaria y los datos desagregados del Paso 4 como la distribución base para los efectos de la transformación automática de valores desde puntajes normales a valores en el espacio EUR.
9. Si el perímetro del área de estudio no es rectangular, eliminar los valores fuera de los límites de interés.
10. Desplegar los resultados como histogramas o distribuciones de probabilidad acumulada y mapas de probabilidad de tener recuperación mayor que ciertos valores críticos
11. Fin.

Estudio de caso

Hemos usado la arcilla Woodford de la cuenca de Arkoma en el estado de Oklahoma (e.g. Cardott 2008) (Figura 1) para hacer una demostración práctica de los distintos pasos del modelaje y despliegue de resultados. Esta arcilla es de edad Devónica–Carbonífera y llega a tener 100 metros de espesor neto. Para el estudio se dispuso de 962 pozos productores, todos con diámetros de drenaje de 400 m al ser medidos en planos perpendiculares al pozo. 85 por ciento de los pozos penetran la arcilla Woodford horizontalmente y se orientan aproximadamente en dirección norte-sur (Figura 2) para sacar ventajas de la orientación aproximadamente este-oeste de un sistema natural de fracturas (Miller and Young 2007). Del total de pozos, a 618 se les pudo estimar la recuperación total (EUR) hasta el momento de su agotamiento (Figura 3). La Figura 4 muestra un detalle del resultado del proceso de transformación de recuperación por pozo a recuperación por celda. Dada la alta concentración de pozos en el área más atractiva económicamente, fue necesario además desaglomerar los datos de recuperación por celda (Figura 5). La información de productividad fue complementada con valores de espesor neto e índice de reflectancia de la vitrinita, los que muestran una mejor distribución geográfica (Figura 6) y correlación con productividad en forma individual y conjunta en forma de espesor efectivo (Figura 7).

La Figuras 8 muestran los resultados de la simulación secuencial para el caso de la variable secundaria y la Figura 9 los de la cosimulación para la primaria. Estos mapas tienen en común la propiedad de reproducir los valores empíricos así como su histograma y semivariograma. Resumiendo los resultados en forma de histogramas, la simulación muestra que hay un 90% de probabilidad de que la cantidad de gas a producir sea entre 310 y $580 \cdot 10^9$ m³ (Figura 10), que en promedio supera por un tercio a la cantidad de gas recientemente estimada con el método tradicional (Tabla 1). Nuestro modelaje permite además entregar mapas de probabilidad para distintos niveles de recuperación por celda, los que indican, por ejemplo, que para un nivel de confianza de 90%, sólo un 9.6 por ciento de las celdas deberían ser capaces de producir al menos $2.8 \cdot 10^9$ m³ (Tabla 2 y Figura 11).

Conclusiones

Cuando se tienen datos de recuperación por pozo y medidas de otros atributos correlacionados con EUR, la simulación secuencial permite un modelaje estocástico integrado haciendo uso simultáneo de toda la información disponible. Este tipo de modelaje es además más satisfactorio teóricamente que métodos más simples actualmente en uso por considerar la ubicación de los pozos, la geometría de los volúmenes drenados, correlación espacial y correlación entre variables. Las discrepancias entre los resultados de nuestra metodología y la estimación oficial indican que las diferencias en el modelaje son significativas, requiriéndose estudios adicionales para clarificar las diferencias. En el intertanto, los valores oficiales deben considerarse como los únicos válidos legalmente. Nuestro método y los usados anteriormente tienen en común el despliegue final de resultados en forma de un histograma resumiendo el rango de valores probables para los recursos totales y sus respectivos intervalos de confianza. Resultados adicionales propios sólo de la simulación son mapas de EUR por celda y mapas de probabilidad de alcanzar ciertas metas de productividad por celda, los que pueden usarse para analizar la prospectividad de distintos sectores de un área de estudio.

Introduction

Oil and gas fields traditionally have been “conventional” accumulations that have resulted from the concentration of hydrocarbons after buoyant migration from source rocks to porous reservoirs. As hydrocarbon resources become scarcer, industry is exploiting source rocks, which commonly are fine-grained sedimentary rocks rich in organic material and with poor flow properties. Currently, shale gas is one of the most attractive types of these “unconventional” resources. Exploitation is possible due to a combination of factors, the most relevant being higher prices, development of horizontal drilling, and technical capability for massive hydraulic rock fracturing.

The increasing relevance of unconventional resources has been followed with interest by government and industry, resulting in the need of stochastically assessing the importance of the accumulations. The U.S. Geological Survey (USGS) pioneered the attempts with a methodology formulated by consultant John D. Grace (Donald Gautier, personal communication, 2010; NOGA Assessment Team, 1995) that was implemented and documented by James Schmoker (1995, 1999). The method is based on well productivity through estimated ultimate recovery (EUR) data obtained by analysis of production decline curves (Cook 2005). This USGS official approach has undergone several rounds of improvements but remains essentially an analytic stochastic method of resource assessment based on mathematical equations derived from classical probability theory (Crovelli 2000; Klett and Schmoker 2005). It requires as input the EUR probability distribution and other parameters such as total and untested area. Both the selection of the specific distributions and the value of their parameters can be highly subjective (Salazar *et al.* 2010). Important assumptions implicit in the methodology are lack of spatial correlation and stochastic independence of all variables involved in the modeling. The assessment is independent from the geographical location of each individual well.

Common and challenging situations in any assessment are those where drilling is concentrated in one or more areas, leaving extensive portions of the assessment unit with minimal production information or none at all. It is not unusual, however, to have data for other geological attributes more regularly scattered throughout the assessment unit. In addition, if at least one of those attributes is correlated to productivity, there are geostatistical methods that have been proven to reduce uncertainty in such situations (e.g. Goovaerts 1997). In this contribution, we explore the potential of geostatistics for modeling shale-gas

resources, taking into account the individual well locations plus the possibility of considering correlation between EUR values and geological variables, such as thermal maturity and formation thickness. In addition, the new methodology allows the simultaneous handling of production from vertical and horizontal wells. This new approach is part of a continuous effort by the USGS to enhance its mathematical models through internal research and feedback from the scientific community. We include practical results using real data in an effort to go beyond testing the theoretical merits of the formulations. The numbers, however, should be used only to help judge the weaknesses and merits of the proposed methodology; they are not in any way intended to supersede the magnitude of the resources released in the official USGS assessment (Houseknecht *et al.* 2010).

Methodology

Among the methods already available for the modeling of spatially correlated attributes and their associated uncertainty, we have selected sequential simulation (e.g. Gómez-Hernández and Cassiraga 1994). In such a numerical approach, values are drawn from univariate uncertainty distributions, often at a regular grid of locations. The sites need to be visited at random, skipping data locations, where the measured values are retained. Thus, the resulting grids honor the data. For sites lacking experimental data, the simulation generates values that become data in the inference of the remaining grid cells. The final result is called a realization. Realizations are not unique. One can generate as many realizations as desired. They are called realizations because each one has an equal chance of being the exact version of the unique but unknown reality that would be obtained by drilling, exploiting, and measuring every single grid cell. Properly generated realizations have the property of following the same probability distributions and low order moments as an unbiased sample, such as mean, median, and semivariogram. For the practical application of sequential simulation, we have used the implementations in the SGeMS software (Remy *et al.* 2009).

Different implementations of sequential simulation satisfy different needs. For the modeling of correlated variables, we selected sequential Gaussian cosimulation (e.g. Verly 1993), so called because all univariate distributions of cell uncertainty are normal. When the attribute to be modeled is not normal, a transformation to normal scores following a standard normal distribution is necessary and the modeling is done at the normal score space. SGeMS offers three alterna-

tives. We selected the Markov-model-1 type (MM1) (Shmaryan and Journel 1999) because it does not require modeling of the cross-semivariogram and still produces realistic results. We have only a few wells with simultaneous measurements for our primary and secondary variables, a situation which does not allow modeling of the cross-semivariogram. In the MM1 approach, the only required knowledge is the correlation coefficient between variables plus the semivariogram for the primary attribute—productivity in our case. Finally, the software requires an exhaustive knowledge for the correlated secondary variable. To satisfy such a requirement, we used sequential Gaussian simulation (e.g. Deutsch 2002) to first generate as many realizations of the secondary variable as realizations to be produced later of the primary variable.

EUR values indicate the expectation of productivity by individual wells, which are a mixture of vertical and predominantly horizontal wells. The length of the horizontal completions (laterals) varies between a few hundred feet to nearly a mile (100–1500 m). For any given well, the longer the lateral, the higher is the productivity. Thus, in an attempt to infer the productivity of an entire assessment unit, it is incorrect to combine the productivity of wells with different lateral lengths. What is relevant is the inherent production potential of the gas shale per some constant unit area. In this study, we decided to homogenize the well productivity data by breaking them into productivity by cell—the preferred term among assessors—which is equivalent to the concept of support used by geostatisticians. A cell is a projection onto the land surface of a certain portion of rock inferred to be drained in the subsurface by a well. Assuming that production comes primarily from the rock cylinder fractured during the stimulation of a well, it is possible to relate the production to square cells of constant area. If the producing unit is thinner than the stimulation diameter, both for horizontal and vertical wells, the height of the drained volume is equal to the producing unit thickness. For laterals, if one subdivides the stimulated rock cylinders into contiguous segments of length equal to the stimulation diameter, then those segments project into a map as square cells. The number of cells per vertical well is automatically equal to one.

In both classical statistics and geostatistics, the use of unbiased samples is a necessary condition for obtaining representative results. The EUR sample that results from the Woodford drilling is highly biased toward spots most economically attractive (sweet spots). Hence, the second problem with the well EUR data and their related cell EUR values is that they are biased samples because of preferential drilling at the most prospective areas and no drilling at the less

favorable spots. The unbiased, normalized cell EUR sample required in cosimulation was obtained by re-sampling of the cell EUR values at a sparse density (e.g. Olea 2007) plus extrapolation at the lower tail of the cumulative distribution following an approach analogous to that used in the inference of distributions of geochemical data with values below detection limit (Olea 2008).

The following procedure summarizes the main steps in the proposed stochastic assessment approach.

Procedure

Within the boundaries of the assessment unit:

1. Select a square assessment cell with side not larger than the diameter of the fractured zone around a well.
2. Calculate the drainage area for every well. If the well is vertical, the drainage is equal to one cell. Otherwise, use the orientation and length of the lateral to model the drainage area by the closest number of contiguous cells centered along the lateral.
3. For every cell associated with a well, assign a cell ultimate production equal to the well EUR divided by the number of cells in the drainage area.
4. If the cells are not uniformly scattered over the assessment area, prepare a declustered dataset of cell EURs.
5. Model the semivariogram for the normal scores of the declustered cell EURs.
6. Use the cells with measurements for both EUR and the correlated secondary variable to obtain the correlation coefficient between the normal scores of these two variables.
7. Apply sequential Gaussian simulation to generate at least 100 realizations for the secondary variable(s) correlated to cell EUR.
8. Use the results from the previous five steps to run a sequential Gaussian cosimulation. Generate as many realizations as there are realizations in Step 7 using different secondary variable realizations to generate each of the cell EUR realizations. In the likely case that declustering is necessary, use the cell EUR dataset in Step 3 as the primary dataset and the declustered dataset in Step 4 for the reference distribution.
9. If assessment area is not rectangular, trim the cell EUR realizations to discard results outside the boundaries.
10. Use the results in Step 10 to summarize the results as histograms, cumulative probability distributions, or probability maps.
11. End.

The woodford shale

The Upper Devonian-Lower Mississippian Woodford Shale has been the focus in recent years of significant gas exploitation in the Oklahoma part of the Arkoma Basin, south-central United States (Figure 1). Eventually, over 1 trillion standard cubic feet of gas (TSCF) ($28.3 \times 10^9 \text{ m}^3$) should be produced from the approximately 1,000 wells drilled during 2005–2009.

A Woodford Shale gas assessment unit defined recently by the USGS extends from a geologic hinge at the southern margin of the Cherokee Platform on the northwest to the northern boundary of the Ouachita Thrust Belt on the south (Figure 1). Although the Woodford Shale or a coeval formation is present in both the Cherokee Platform and the Ouachita Thrust Belt, geological parameters that are known to influence gas resource potential are significantly different than those in the Arkoma Basin. On the Cherokee Platform, the Woodford Shale is thinner and less thermally mature than in the Arkoma Basin. In the Ouachita Thrust Belt, the coeval Arkansas Novaculite comprises interbeds of black shale and chert and is of higher thermal maturity. To the northeast and southwest, the Woodford Shale is truncated by erosion on the flanks of the Ozark Uplift and Arbuckle Uplift, respectively. Thus, the northern, western, and southern boundaries of the assessment unit are clearly defined by geology. On the east, an arbitrary boundary was set at the

Oklahoma-Arkansas state border for purposes of this research, even though a coeval shale formation (Chattanooga) is present in the Arkansas part of the basin.

The Woodford Shale in the Arkoma Basin ranges from less than 50 to more than 300 ft thick (15–100 m), and is characterized by high gamma-ray (>150 API units) well-log response throughout most of its thickness. The high gamma-ray (HGR) log response is a direct reflection of its high content of organic matter, with total organic carbon content ranging up to 14 percent. Thermal maturity increases from less than 1 percent vitrinite reflectance along the western and northwestern margins of the basin to more than 4 percent along the southern margin of the basin adjacent to the Oklahoma-Arkansas state border (Cardott, 2008). Prolific gas production from the Woodford Shale has been established in areas where vitrinite reflectance ranges between about 1.0 and 2.5 percent (Cardott, 2008). Areas characterized by vitrinite reflectance of less than 1 percent have little gas production and actually produce oil and condensate locally, probably indicating that gas generation has not occurred in those low thermal maturity areas. Relatively little gas production has been established in areas where vitrinite reflectance values exceed 3.0 percent, although relatively few wells have tested those areas. For purposes of this research, we assume that the most favorable range of thermal maturity is 1 to 3 percent vitrinite reflectance and that gas resource potential decreases in areas of less than 1 and more than 3 percent vitrinite reflectance.

Current depth of burial of the Woodford Shale in the Oklahoma part of the Arkoma Basin ranges from less than 1,000 ft (300 m) along the northern basin margin to more than 15,000 ft (4600 m) at the southern basin margin (Figure 1). The most prolific gas production from the Woodford Shale has been established in areas where the reservoir is 5,000 to 8,000 ft (1520–2440 m) deep. The orientation of the main system of natural fractures is approximately east-west (Miller and Young 2007), so most horizontal wells have laterals oriented closely to north-south to intersect a maximum number of natural fractures (Figure 2).

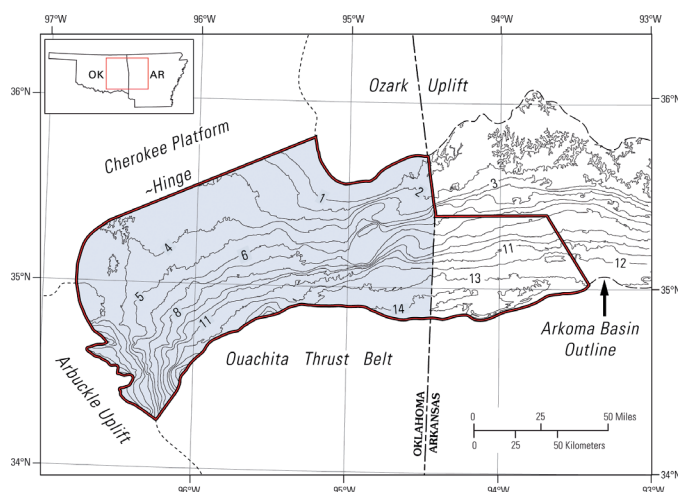


Figure 1. Location map of portion of interest in Arkoma Basin and adjacent geologic provinces. Boundary of Woodford Shale assessment unit from Houseknecht et al. (2010) is outlined in brown and area considered in this study is colored light blue. Generalized contours are thickness of Woodford Shale overburden, in thousands of feet.

Figura 1. Mapa mostrando parte de la Cuenca de Arkoma y zonas circundantes. La línea café es el límite de la arcilla Woodford de acuerdo a Houseknecht et al. (2010) y el área de estudio es aquella en celeste. Las curvas de nivel, en miles de pies (305 m), corresponden al espesor de la sobrecarga superpuesta a la arcilla Woodford.

Results

By the end of 2009, there were 983 wells drilled into the Woodford play to produce gas. Of those, only 21 (2 percent) were dry and abandoned wells. Because some of these wells failed for mechanical or economic reasons, for the purpose of modeling, we ignored the possibility of drilling non-producing wells, a standard practice in continuous assessments (Has-

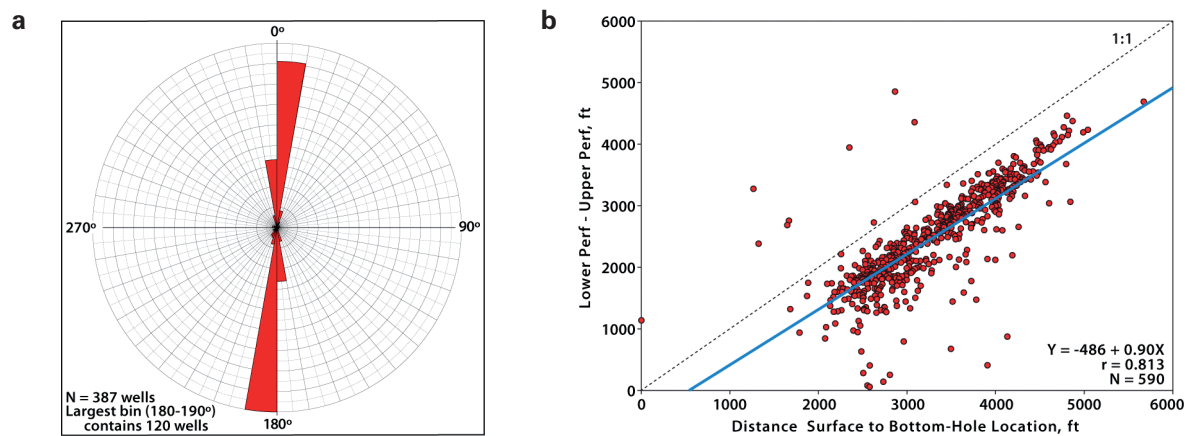


Figure 2. Geometry of horizontal wells: (a) rose diagram showing azimuth of laterals for Woodford Shale wells prepared extracting information from IHS Energy, Inc. commercial database; (b) scatterplot of horizontal projection of distance from surface to bottom-hole location versus lateral length (lower perforation minus upper perforation) for wells in the neighboring Fayetteville Shale play to the east. Blue line is regression of lateral length on horizontal distance downloaded from IHS Energy, Inc. commercial database and perforation lengths obtained from online data available from the Arkansas Geological Survey.

Figura 2. Características geométricas de los pozos horizontales: (a) diagrama mostrando el azimut de los pozos horizontales de acuerdo a información obtenida de la base de datos de la compañía IHS Energy, Inc.; (b) diagrama de dispersión de la proyección horizontal de la distancia desde la boca del pozo hasta su extremo inferior versus la longitud del intervalo perforado en el caso de pozos produciendo desde la arcilla Fayetteville al este de muestra zona de estudio. La recta azul es la línea de regresión. La distancia horizontal fue obtenida de IHS Energy, Inc. y la longitud del intervalo perforado desde un sitio en la Internet del Servicio Geológico del Estado de Arkansas. Un pie (ft) es aproximadamente igual a 0.305 m.

kett and Brown 2005). Among the 962 producers, 85 percent are horizontal wells. Preparation of well EURs requires production data during a minimum time before a decline curve analysis can be done (Cook 2005). Figure 3 displays the fraction of producers for which it was possible to calculate EUR values. These are our primary variable data.

We assumed that hydraulic fracturing is effective within a radius of 200 m from the wellbore, based on

industry submissions to state regulatory agencies. Consequently, we have taken as assessment cell a square 400 m on a side (Step 1 of the Procedure). According to the information summarized in Figure 2a, wells are almost exactly oriented in a north-south direction. Locations of a well at the surface and of its bottom hole are public domain information for the Woodford play, but the parameter of interest is not: the length of the completion (lateral). Figure 2b shows

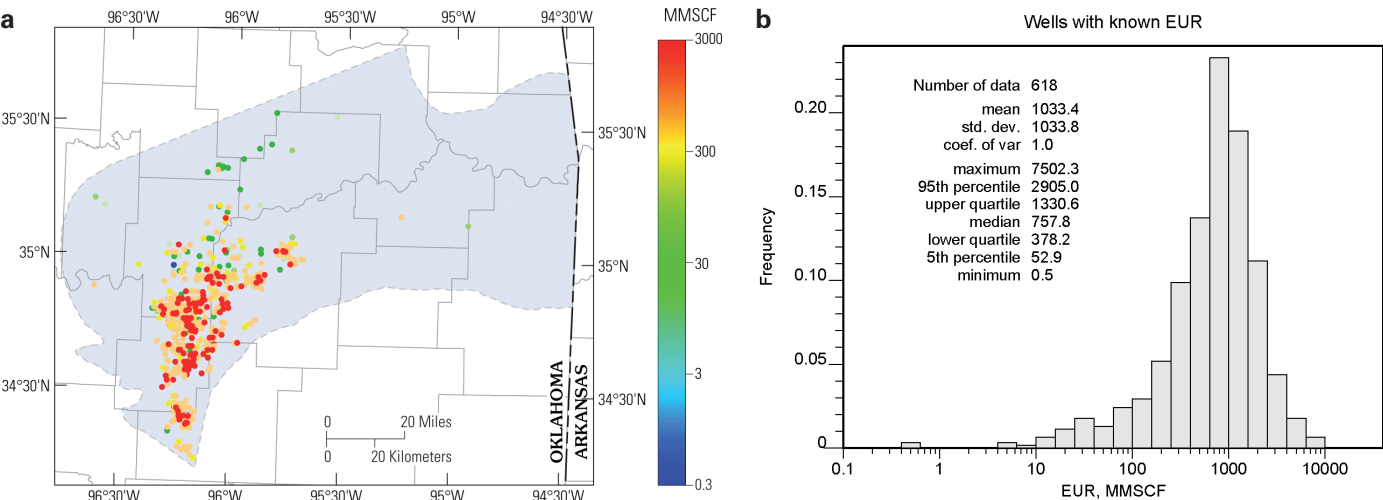


Figure 3. Well EUR data generated using the method outlined in Cook (2005) and taken from (Houseknecht et al. 2010): (a) posting evidencing a sweet spot in the southwest covering about 20 percent of the study area; (b) histogram.

Figura 3. Recuperación final por pozo (EUR) estimada de acuerdo a Cook (2005) y que corresponde a los mismos valores usados en la evaluación oficial hecha por Houseknecht et al. (2010). Un MMSCF es aproximadamente igual a $28.3 \cdot 10^3 \text{ m}^3$: (a) ubicación de los pozos mostrando una concentración de pozos de alta recuperación en el suroeste en un área que cubre aproximadamente 20 por ciento del total; (b) histograma de los mismos valores.

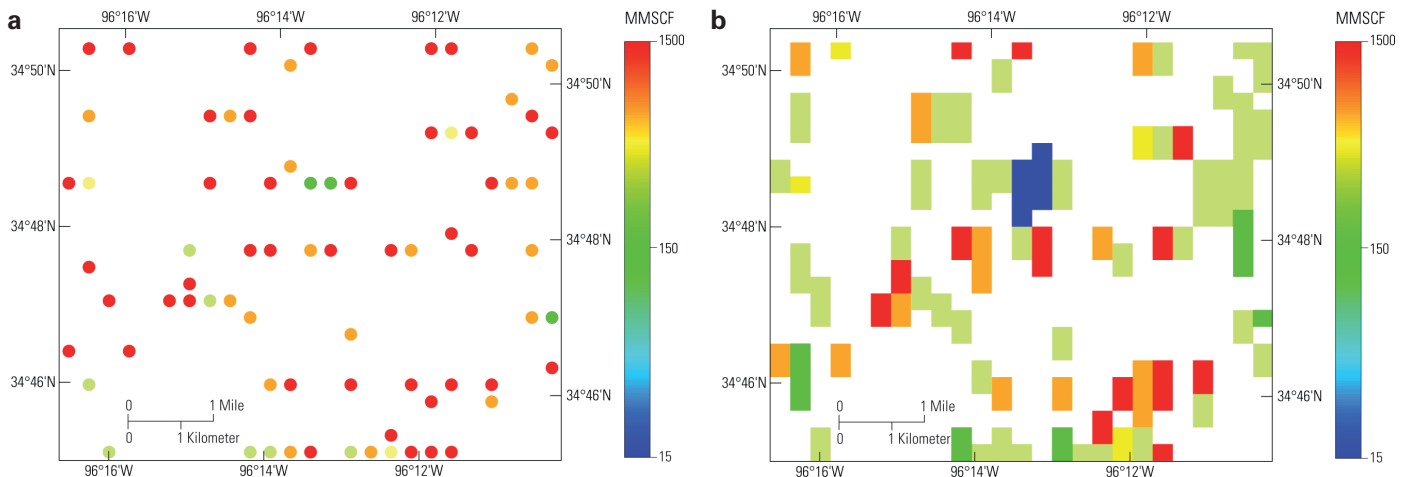


Figure 4. Detail near the center of the sweet spot: (a) posting of center of lateral length showing well EUR; (b) well drainage area showing cell EUR.

Figura 4. Detalle cerca del centro del área más atractiva y densamente perforada. Un MMSCF es aproximadamente igual a $28.3 \cdot 10^3 \text{ m}^3$: (a) ubicación del centro del intervalo en producción usando el color para indicar la recuperación final esperada (EUR) por pozo; (b) mapa de drenaje y recuperación de los mismos pozos.

a correlation from the neighboring Fayetteville play to the east that we used to infer the lateral length. Following Step 2 of the modeling Procedure, we decomposed the horizontal producing interval into 400 m square cells aligned along a north-south direction. The actual area drained by each well is so small at the scale of the assessment unit that Figure 4 shows only a close-up in the middle of the well concentration in Figure 3 to illustrate the effect of the conversion from well EURs to cell EURs at Step 3.

At the play level, wells are distributed in a way that departs radically from the regular, stratified, or random patterns required to have an unbiased spatial sample. Consequently, Step 4 cannot be ignored. An unbiased sample is required to model the semivariogram, to transform the data to follow a standard normal distribution, and to convert the results back to volumes of gas in the cosimulation. Figure 5a depicts the cumulative distribution both for all cells and for the declustered sample. Figure 5b is the omnidirectional

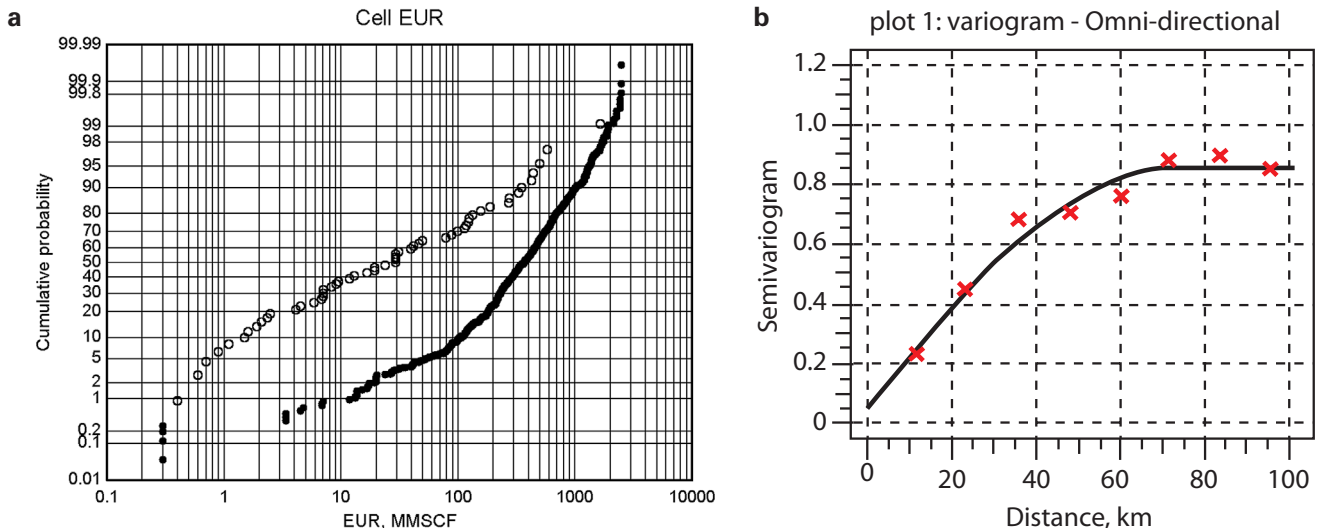


Figure 5. Declustered cell EURs: (a) cumulative frequency for all cells (bullets) and for the declustered sample (open circles); (b) semivariogram for the normal scores of declustered cell EURs, where the crosses denote experimental values and the solid line the fitted model that is spherical with a nugget of 0.05, sill of 0.8, and range of 57.1 km.

Figura 5. Producción final por celda luego de un proceso de desaglomeración: (a) distribución de probabilidad acumulada para el caso de todas las celdas (puntos) y para el caso de la muestra desaglomerada (círculos). Un MMSCF es aproximadamente igual a $28.3 \cdot 10^3 \text{ m}^3$; (b) semivariograma de la transformación normal de los valores de productividad final por celda desagregada en que las cruces indican valores experimentales y la curva es el modelo que mejor se ajusta a estos valores y que tiene un efecto de pepita de 0.05, meseta de 0.8 y alcance de 57.1 km.

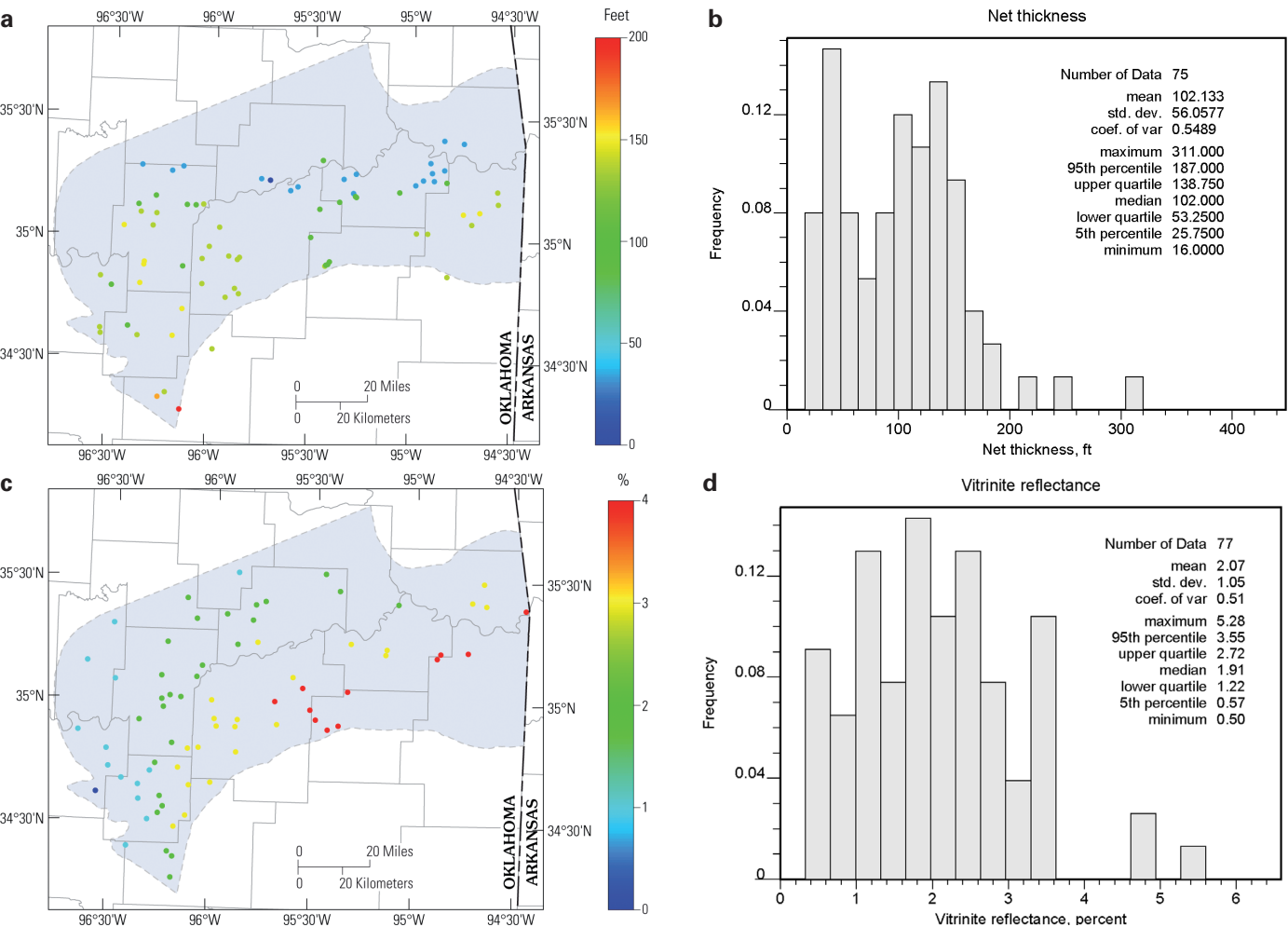


Figure 6. Secondary data: (a) posting of net HGR thickness; (b) histogram of net HGR thickness; (c) posting of vitrinite reflectance (modified from Cardott (2008)); (d) histogram of vitrinite reflectance.

Figura 6. Información secundaria: (a) mapa de valores de espesor neto. Un pie (ft) es aproximadamente igual a 0.305 m; (b) histograma de los mismos valores; (c) mapa de valores de reflectancia de la vitrinita tomados de Cardott (2008); (d) histograma de los mismos valores.

semivariogram for cell EUR after the normal score transformation required for the cosimulation software and mentioned in Step 5.

The play has information for four attributes other than EUR: depth to the top of the Woodford, total Woodford thickness, net HGR thickness, and vitrinite reflectance. A preliminary analysis eliminated total thickness and depth from further consideration for being redundant with the other attributes and having a weaker correlation to the primary variable. Figure 6 shows that both variables are sparsely sampled, but with a more regular geographical distribution than the primary variable, an important consideration when it comes to using them in cosimulation.

The relationship between vitrinite reflectance and gas generation potential is not linear, which precludes directly using the variable in cosimulation. Instead, we defined the weighting function in Figure 7a to prepare a new secondary variable, combining the capa-

bility of net HGR thickness and vitrinite reflectance to predict the amount of gas expected at any given cell, a synthetic variable that we called effective thickness. As seen in Figure 7b, the correlation coefficient in normal score space between effective thickness and cell EUR is 0.73 (Step 6). Obviously, to a net HGR thickness approaching zero corresponds a nearly null productivity, which in normal score space would translate into large negative values for both attributes, say, (-2,-2). This conceptual point helps to confirm that the isolated outlier (-0.37,0.05) is not an error, but rather a consequence of the biased sampling. A sensitivity analysis shows that addition of this point to the regression increases the correlation coefficient to 0.93 and, conversely, removing the outlier from Figure 7b, the correlation coefficient for the remaining 8 points drops to 0.36. In the modeling, we discounted the two extreme possibilities; the correlation coefficient was set to fluctuate between 0.5 and 0.73. In an official as-

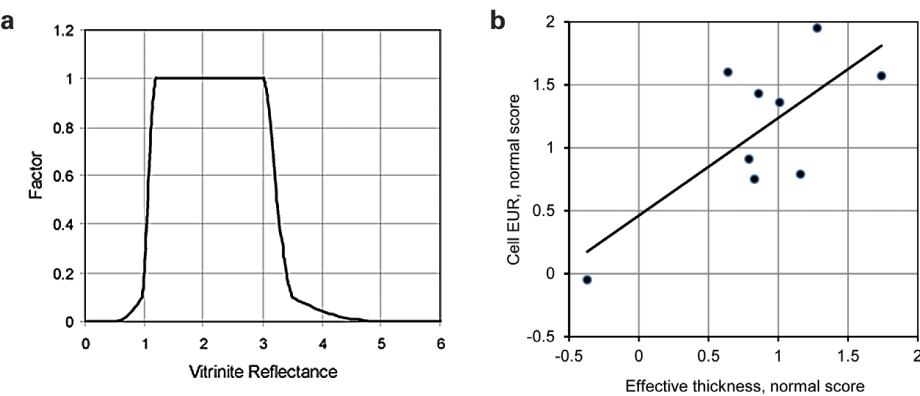


Figure 7. Effective thickness: (a) Weighting function used to convert net thickness to effective thickness; (b) correlation between the primary and the secondary variable, whose correlation coefficient is 0.73.
Figura 7. Espesor efectivo: (a) transformación usada para convertir de espesores netos a espesores efectivos; (b) diagrama de dispersión entre las variables primaria y secundaria y regresión lineal con un coeficiente de correlación de 0.73.

assessment it would be desirable to base the correlation coefficient on more than 9 points. Here, our primary objective is to illustrate the possibilities offered by multivariate assessments if indeed these variables are correlated. Our feeling is, however, that a correlation of about 0.6 between effective thickness and EUR is consistent with the implicit assumptions made in the official assessment.

Figure 8 displays two extreme scenarios among the 100 realizations of effective thickness required in Step 7. Each map contains 126,047 cells. The modeling was completed by entering all previous results into the co-simulation program and employing them to generate 100 possible realizations of cell EUR fluctuation across the assessment unit. Figure 9 shows two extreme and two intermediate situations. These maps offer a spatial rendering of resource potential that the standard USGS methodology cannot provide. They are of great value when required to expand geologic assessments into economic evaluations.

As a whole, the 100 scenarios are rich in information about uncertainty surrounding EUR in the Woodford play. The tables and the last two figures offer different ways of summarizing such information as part of the assessment proper. As mentioned earlier, 346 producing wells without EUR data were not used in the modeling. The gas in those wells should not count as undiscovered resources. Figure 10b is the sum of the 0.64 TSCF ($18.1 \cdot 10^9 \text{ m}^3$) of the 618 wells in Figure 1 plus the gas estimated by the 100 realizations for those cells penetrated by the 346 wells without EUR data. Table 1 compares the results obtained here with the new methodology under scrutiny and those in the official assessment using the same well EUR data but processing the information with the latest version of Schmoker's approach (Houseknecht *et al.* 2010). It should be noted that the official methodology models the undiscovered resources in two stages. First, the input probability distributions are used to analytically calculate mean and variance for the undiscovered re-

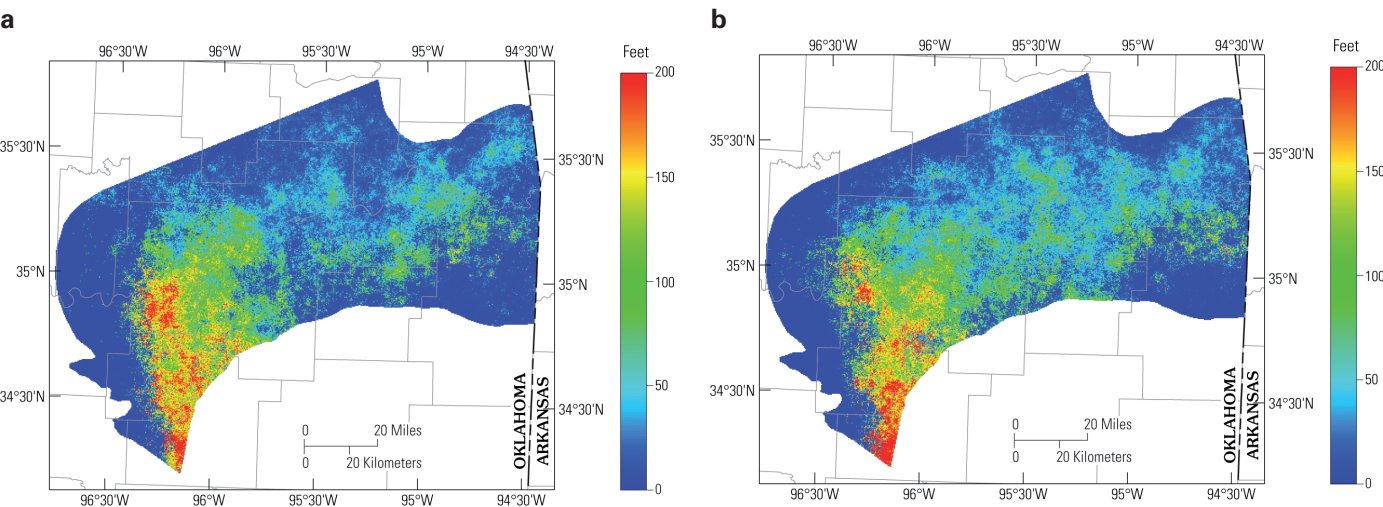


Figure 8. Possible maps for effective thickness: (a) scenario resulting in the realization in Figure 9a; (b) scenario resulting in the realization in Figure 9c.
Figura 8. Posibles mapas de espesor efectivo. Un pie (ft) es aproximadamente igual a 0.305 m: (a) mapa usado en la preparación del mapa de producción final por celda de la Figura 9a; (b) mapa usado en la preparación del mapa de producción final por celda de la Figura 9c.

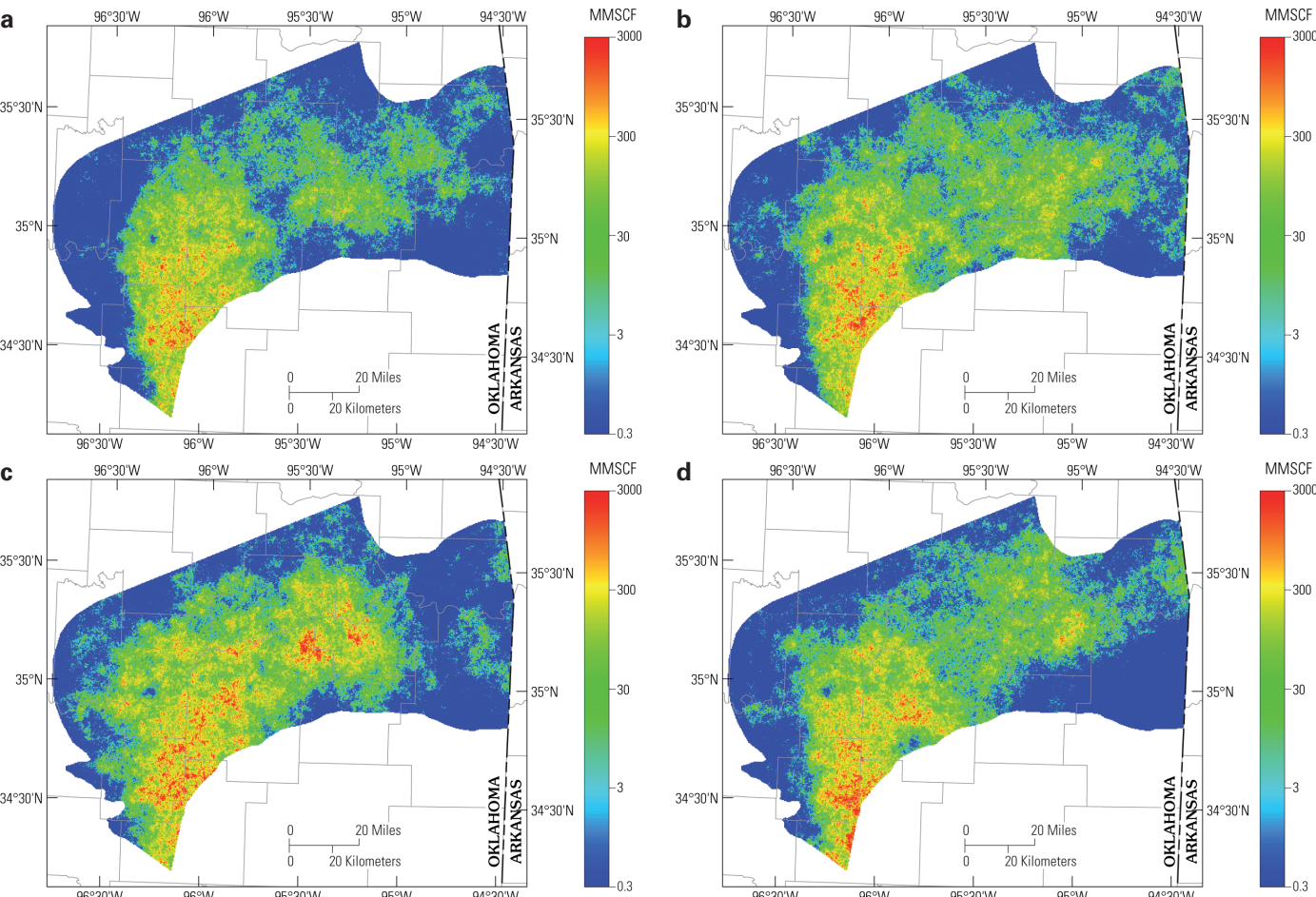


Figure 9. Selected scenarios of cell EUR: (a) case in which there is only 0.05 probability in a scale 0–1 that the undiscovered resources will be lower; (b) case in which there is equal probability that the undiscovered resources will be lower or higher; (c) case with 0.95 probability that the undiscovered resources will be lower; (d) map resulting from averaging all 100 possible scenarios.
Figura 9. Selección de mapas de producción final por celda. Un MMSCF es aproximadamente igual a $28.3 \cdot 10^3 \text{ m}^3$: (a) caso en que hay sólo un 5 por ciento de probabilidad de que los verdaderos recursos por descubrir sean menores que los mostrados por el mapa; (b) caso en que es igualmente probable que los recursos reales sean menores o mayores que los indicados por el mapa; (c) caso en que hay un 95 por ciento de probabilidad de que los verdaderos recursos sean menores que los del mapa; (d) mapa que resulta de sacar un promedio celda por celda de la colección de 100 mapas igualmente probables de ser el mapa correcto.

sources, which then are used to define a lognormal distribution with those two parameters. Unlike the distribution in our Figure 10c, such a lognormal distribution is not the natural result from the modeling. Rather, it is an assumption that is mathematically inconsistent with the fact that all of the input distributions were triangular except for the EUR distributions.

The percentiles in the first row of Table 1 are the values read from the assumed lognormal distribution. The practical and theoretical differences between the official methodology and the one proposed here do not seem to be irrelevant. In terms of the $P_{95} - P_5$ spread, according to Table 1, ours is 13% narrower than the interval in the official assessment. Salazar

	F_5 , TSCF	F_{50} , TSCF	F_{95} , TSCF	Mean, TSCF
USGS 2010	6.1	10.2	17.0	10.7
This study	11.0	13.8	20.5	14.3

Table 1. Comparison of the 2010 official U.S. Geological Survey (USGS) assessment of undiscovered resources (Houseknecht et al. 2010) and EUR values from Figure 10c. F_5 stands for the 5th percentile and F_{95} for the 95th percentile of the cumulative distribution.
Tabla 1. Comparación entre la evaluación oficial del Servicio Geológico de EEUU (USGS) de recursos por descubrir (Houseknecht et al. 2010) y los mismos recursos de acuerdo a nuestros resultados resumidos en la Figura 10c. F_5 indica el quinto percentil y F_{95} el percentil 95. Un TSCF es aproximadamente igual a $28.3 \cdot 10^9 \text{ m}^3$.

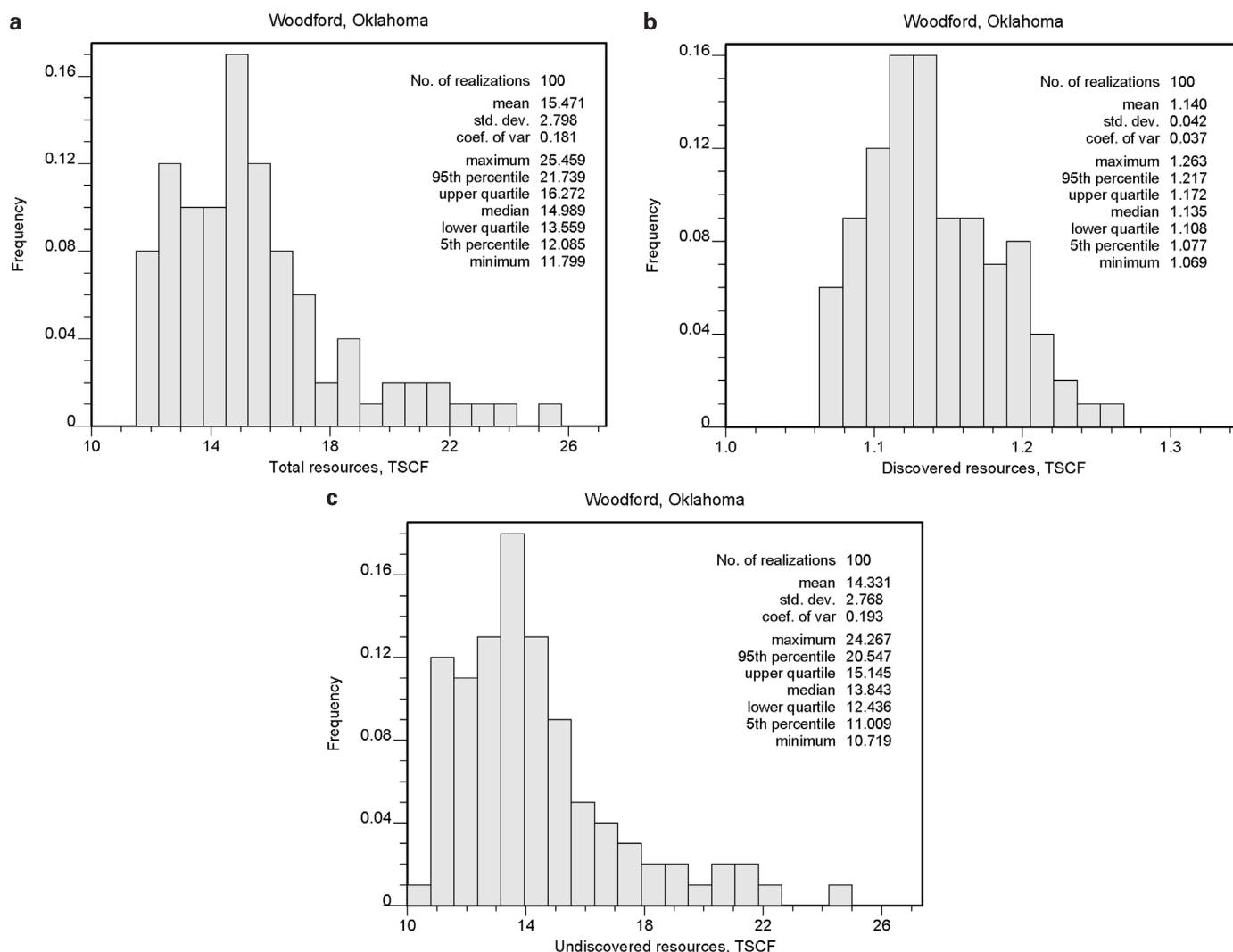


Figure 10. Histogram of play resources: (a) total resources; (b) discovered resources; (c) undiscovered resources.

Figura 10. Histogramas del área estudiada. Un TSCF es aproximadamente igual a $28.3 \cdot 10^9 \text{ m}^3$: (a) recursos totales; (b) recursos ya perforados; (c) recursos por descubrir.

et al. (2010) have suggested that spreads in the official unconventional USGS assessments are already too narrow primarily as a result of using low variance triangular distributions in the modeling. This perspective will remain conjectural until it is mathematically shown that the probability of the true undiscovered resources to fall outside the reported intervals exceeds 10%. Thus, the fact that our $P_{95} - P_5$ spread is even narrower is not necessarily wrong. To start with, we have not used any triangular distributions. In addition, considering that the spread in the distribution of any assessed commodity is a measure of uncertainty, the correct implication is that our modeling is more precise than the results that can be obtained with the official methodology or some variant of it like that of Salazar and others. This should be the logical consequence of using not only information about estimated

ultimate recovery but also measurements of net HGR thickness and vitrinite reflectance, considering interdependences among geologic variables, and extracting information about spatial correlation.

Our results do follow the expectations of Salazar *et al.* (2010) in the sense that distributions assessing the uncertainty about undiscovered resources in USGS assessments should be more skewed to be consistent with the fact that mineral resources tend to be highly and positively skewed. Our distribution is more skewed than the one in the official assessment, as evidenced by the fact that in the official assessment $(P_{95} - P_{50}) / (P_{50} - P_5)$ is 1.7 while our ratio is 2.4, a 41 percent increase.

Additional advantages of the experimental technique are illustrated by Table 2 and Figure 11. The ability to estimate spatial productivity of the play and to

Cell EUR MMSCF	Total resources area percent	Undiscovered resources area percent
20	19.5	17.7
100	11.2	9.6
200	4.6	3.4
500	0.3	0.1

Table 2. Drilling effort at different productivity levels. All values are percentages of the assessment unit area with at least a 90 percent probability to drain a cell with a productivity at least as high as the value in the first column.

Tabla 2. La columna central indica, para el caso de recursos totales, el porcentaje de celdas que con al menos 90% de probabilidad tendrán la productividad indicada en la columna de la izquierda. La columna de la derecha indica lo mismo para el caso de nuevos recursos. Un TSCF es aproximadamente igual a $28.3 \cdot 10^9 \text{ m}^3$.

summarize the proportion of the total and untested parts of the assessment unit provides information that would be of great value to both public and private sector entities facing a range of decisions regard-

ing resource development, land management, and energy policy. Results such as these that incorporate spatial analysis into resource estimation are not possible using current methodology.

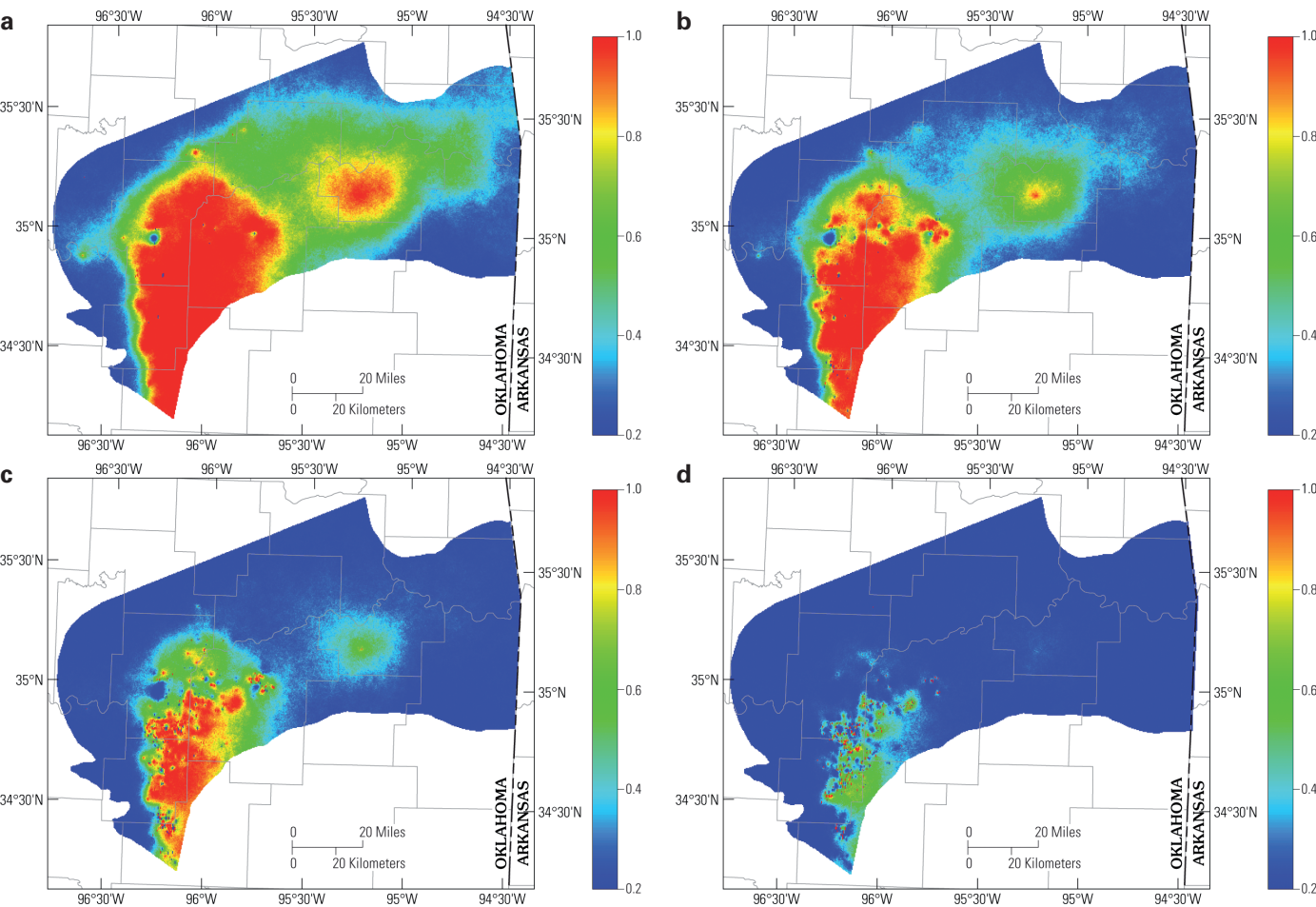


Figure 11. Maps showing the probability that the cell EURs will be at least: (a) 20 MMSCF; (b) 100 MMSCF; (c) 200 MMSCF; (d) 500 MMSCF.

Figura 11. Mapas con la probabilidad de que la recuperación final por celda sea al menos: (a) $0.57 \cdot 10^6 \text{ m}^3$; (b) $2.8 \cdot 10^6 \text{ m}^3$; (c) $5.6 \cdot 10^6 \text{ m}^3$; (d) $14 \cdot 10^6 \text{ m}^3$.

Conclusions

Advances in spatial statistics allow today the application of sequential simulation for the modeling of uncertainty in unconventional accumulations considering geographical location, spatial correlation, and quantitative information of attributes correlated with estimated ultimate recovery. Instead of directly using well EURs, productivity density in the form of cell EUR allows additional decoupling of the influence on productivity of purely geologic factors from the length of the completed interval in horizontal wells.

The final result of the modeling is a series of EUR density maps characterizing multiple scenarios, all following the same histogram and style of spatial variation revealed by the data. In our case, it is necessary to compensate for preferential drilling at a sweet spot. After such corrective step, the declustered data should follow the same distribution and semivariogram as the true but unknown map. The information contained in the modeled maps can be summarized in two different ways: (a) conventional global summary of uncertainty in the form of histograms; (b) probability maps and productivity tables that are an alternative display of the spatial information rendered by the scenarios individually. They offer synoptic ways to visualize the prospectivity of different areas across a play.

Further studies will be required to evaluate the significance of the discrepancy in the results in Table 1. Higher estimates of undiscovered resources for the Woodford play using cosimulation should not be interpreted as an indication that systematic application of cosimulation to other areas will result in larger estimates than those obtained modeling essentially the same data but ignoring attributes, their correlations, and geographic location of the wells.

We hope the results presented here will encourage the collection and analysis of data for a spectrum of geologic attributes that may influence productivity in unconventional petroleum resource plays. While multivariate assessment of unconventional resources following the procedure outlined here or other similar approaches holds the potential for spatially focused estimates of undiscovered resources, a better understanding of the geologic influences on productivity is essential.

Acknowledgments

During the preparation of the manuscript, we benefited from the reviews of Jorge Cifuentes (ENAP-SIPETROL, Chile), Michael Pyrcz (Chevron Energy Technology Company R&D), and a third anonymous reviewer.

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the United State Government.

References

- Cardott, B.J. 2008. Overview of Woodford gas-shale play in Oklahoma. Oklahoma Gas Shales Conference. Electronically available at http://www.ogs.ou.edu/pdf/Cardott_Overview.pdf. Accessed in July 2011.
- Cook, T. 2005. Calculation of Estimated Ultimate Recovery (EUR) for Wells in Continuous-Type Oil and Gas Accumulations of the Uinta-Piceance Province. In *Petroleum Systems and Geologic Assessment of Oil and Gas in the Uinta-Piceance Province, Utah and Colorado*, Chapter 23, U.S. Geological Survey Digital Data Series DDS-69-D, 5 p. Electronically available at http://pubs.usgs.gov/dds/dds-069/dds-069-d/REPORTS/69_D_CH_23.pdf. Accessed in July 2011.
- Crovelli, R.A. 2000. *Analytic resource assessment method for continuous (unconventional) oil and gas accumulation—the ACCESS method*. U.S. Geological Survey Open-File Report 00-044, 34 p.
- Deutsch, C.V. 2002. *Geostatistical reservoir modeling*. Oxford University Press, New York, 384 p.
- Gómez-Hernández, J.J. and Cassiraga, E.F. 1994. Theory and practice of sequential simulation. In Armstrong, M. and Dowd, P.A., eds., *Geostatistical simulation*, Kluwer Academic Publishers, Dordrecht, 111–124.
- Goovaerts, P. 1997. *Geostatistics for natural resources evaluation*. Oxford University Press, New York, 483 p.
- Haskett, W.J. and Brown, P. J. 2005. Evaluation of unconventional resource plays. Society of Petroleum Engineers Paper 96879, 11 p.
- Houseknecht, D.W., Coleman, J.L., Milici, R.C., Garrity, C.P., Rouse, W.A., Fulk, B.R., Paxton, S.T., Abbot, M.M., Mars, J.C., Cook, T.A., Schenk, C.J., Charpentier, R.R., Klett, T.R., Pallastro, R.M., and Ellis, G.S. 2010. *Assessment of undiscovered natural gas resources of the Arkoma Basin province and geologically related areas*. U. S. Geological Survey Fact Sheet 2010-3043, 4p. Electronically available at <http://pubs.usgs.gov/fs/2010/3043/pdf/FS10-3043.pdf>. Accessed in July 2011.
- Klett, T.R. and Schmoker, J.W. 2005. U.S. Geological Survey input-data form and operational procedure for assessment of continuous petroleum accumulations, 2002. In *Petroleum Systems and Geological Assessment of Oil and Gas in the Southwestern Wyoming Province, Wyoming, Colorado, and Utah*, Chapter 18, U.S. Geological Survey Digital Data Series DDS-69-D, 8 p.
- Miller, R. and Young, R. 2007. Characterization of the Woodford Shale in outcrop and subsurface in Pontotoc and Coal Counties, Oklahoma. American Association of Petroleum Geologists, Search and Discovery Article #50052, 3 posters.
- NOGA Assessment Team. 1995. *1995 National Assessment of United States Oil and Gas Resources*. U.S. Geological Survey Circular 1118, 20 p.

- Olea, R.A. 2007. Declustering of clustered preferential sampling for global histogram and semivariogram inference. *Mathematical Geology*, 39 (5), 453–467.
- Olea, R.A. 2008. Inference of distributional parameters for compositional samples containing nondetects. Proceedings of CoDaWork'08, Girona, Spain, <http://dugi-doc.udg.edu/handle/10256/708>. Accessed in July 2011.
- Remy, N., Boucher, A., and Wu, J. 2009. *Applied geostatistics with SGeMS: a user's guide*. Cambridge University Press, New York, 284 p. and one CD-ROM.
- Salazar, J., McVay, D.A., and Lee, W.J. 2010. Development of an improved methodology to assess potential unconventional resources. *Natural Resources Research*, DOI: 10.1007/s11053-010-9126-9. Accessed in September 2010.
- Schmoker, J.W. 1995. Method for assessing continuous type (unconventional) hydrocarbon accumulations. In Gautier, D.L., Dolton, G.L., Takahashi, K.I., Varnes, K.L., eds. *1995 Natural Assessment of United States Oil and Gas Resources—Results, Methodology, Supporting Data*. U.S. Geological Survey Digital Data Series DDS-30, one CD-ROM.
- Schmoker, J.W. 1999. *U.S. Geological Survey assessment model for continuous (unconventional) oil and gas accumulations—The "FORSPAN" model*. U.S. Geological Survey Bulletin 2168, 9 p., only electronically available at <http://pubs.usgs.gov/bul/b2168/>. Accessed in July 2011.
- Shmalyan, L.E. and Journel, A.G. 1999. Two Markov models and their applications. *Mathematical Geology* 31 (8), 965–988.
- Verly, G. 1993. Sequential Gaussian cosimulation: A simulation method integrating several types of information. In Soares, A., editor, *Geostatistics Tróia '92*. Kluwer Academic Publishers, Dordrecht, vol. 1, p. 543–554.

Recibido: diciembre 2010

Revisado: mayo 2011

Aceptado: julio 2011

Publicado: octubre 2011