Probabilistic Methodology for the Assessment of Original and Recoverable Coal Resources, Illustrated with an Application to a Coal Bed in the Fort Union Formation, Wyoming

Chapter 1 of
Section G, Coal Resource Assessment
Book 6, Modeling Techniques

Techniques and Methods 6–G1

Probabilistic Methodology for the Assessment of Original and Recoverable Coal Resources, Illustrated with an Application to a Coal Bed in the Fort Union Formation, Wyoming

By Ricardo A. Olea, Brian N. Shaffer, Jon E. Haacke, and James A. Luppens

Chapter 1 of
Section G, Coal Resource Assessment
Book 6, Modeling Techniques

Techniques and Methods 6–G1

U.S. Geological Survey, Reston, Virginia: 2021

Acknowledgments

Contents

Figures

Tables

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
Length		
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
Area		
acre	4,047	square meter (m^2)
Volume		
acre-foot (acre-ft)	1,233	cubic meter (m^3)
Mass		
ton, short (2,000 pounds)	0.9072	metric ton (t)
Density		
short ton/acre-foot	0.0007306	metric ton per cubic meter (t/m^3)

International System of Units to U.S. customary units

Multiply	By	To obtain
Length		
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
Area		
square meter (m^2)	0.0002471	acre
Volume		
cubic meter (m^3)	0.0008107	acre-foot (acre-ft)
Mass		
metric ton (t)	1.102	ton, short (2,000 pounds)
Density		
metric ton per cubic meter (t/m^3)	1,368.778	short ton/acre-ft

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$°C = (°F - 32) / 1.8.$$

Abbreviations and Symbols

abs. val.	absolute value
ASTM International	American Society for Testing and Materials
bst	billion short tons
CRIRSCO	Committee for Mineral Reserves International Reporting Standards
db	dry basis
eq.	equation
est. spec. gravity	estimated specific gravity
JORC	Joint Ore Reserves Committee
kst	thousand short tons
NAD	North American Datum
SEC	U.S. Securities and Exchange Commission
SME	Society for Mining, Metallurgy and Exploration
Ss	sandstone
USGS	U.S. Geological Survey
UTM	Universal Transverse Mercator
%	percent
±	$x \pm a$ means the interval from $x - a$ to $x + a$
>	greater than
≥	greater than or equal to
<	less than
≤	less than or equal to

Probabilistic Methodology for the Assessment of Original and Recoverable Coal Resources, Illustrated with an Application to a Coal Bed in the Fort Union Formation, Wyoming

By Ricardo A. Olea, Brian N. Shaffer, Jon E. Haacke, and James A. Luppens

Executive Summary

The U.S. Geological Survey (USGS) has been using its Circular 891[1] for evaluating uncertainty in coal resource assessments for more than 35 years. Calculated cell tonnages are assigned to four qualitative reliability classes depending exclusively on distance to the nearest drill hole. The main appeal of this methodology, simplicity, is also its main drawback. Reliability may depend so marginally on distance to the nearest drill hole that, over time, it has become evident that Circular 891 is inadequate for modeling reliability and is limited by other shortcomings. The present publication describes the use of geostatistics as an approach allowing a more satisfactory performance than that which is achieved following Circular 891. Geostatistics takes advantage of partly random and partly organized fluctuations in attributes such as coal thickness, coal density, and elevation of the top of a coal bed, borrowing concepts and tools that have been standard features in statistics and risk analysis for decades. Considering that readers interested in this study may not have the background to go directly into the details of the methodology, we start by explaining geostatistical concepts and modeling techniques. The remainder of the publication is devoted to formulating the assessment methodology, applying it to data from the Fillmore Ranch coal bed in the Fort Union Formation in Wyoming, and explaining the computer software applied for performing calculations and displays. The assessment methodology has been designed to report three different forms of resources: coal in place, coal mineable by surface mining methods, and coal mineable by underground mining methods. These three types of resources are reported graphically by displaying both the magnitude and the reliability of total coal resources and resources at the cell scale. In the case of the Fillmore Ranch coal bed example, there is a 90-percent probability that the resources in place are 9.687 ± 0.383 billion short tons (bst), while the coal available for underground mining is 2.279 ± 0.160 bst, and that available for surface mining is only 0.240 ± 0.025 bst because of the steep dip to the west away from the outcrop. These magnitudes are derived from numerical probability distributions not following any specific form.

Introduction

The search for optimality in the calculation of mining resources and for convenience in assigning a degree of reliability to such calculations is as old as mining itself (Matheron, 1963; Journel and Huijbregts, 1978). However, adoption of adequate methods, let alone agreement on standard methods, has been slow to come by despite international efforts (Joint Ore Reserves Committee, 2012; Committee for Mineral Reserves International Reporting Standards [CRIRSCO], 2013). More significant progress has resulted from regulating the qualification of competent persons for conducting the assessments than from focusing on the selection of assessment methods; most mining companies and regulatory agencies follow their own practices (Rendu, 2006).

In the Unites States, the Society for Mining, Metallurgy and Exploration (SME), as a member of CRIRSCO, supports the CRIRSCO recommendations, but public disclosures must be prepared in accordance with regulations enforced by the U.S. Securities and Exchange Commission (SEC), which recently announced its intention to more closely follow the CRIRSCO practices by 2021 (U.S. Securities and Exchange Commission [SEC], 2018). Yet, the U.S. Geological Survey (USGS) has been following in its research for more than three decades USGS Circular 891—referred to as Circular 891 hereinafter—for the assessment of coal resources (Wood and others, 1983). The SEC has presented objections to Circular 891 being the sole basis for making mining property disclosures (U.S. Securities and Exchange Commission [SEC], 2016).

---

[1]Wood, G.H., Jr., Kehn, T.M., Carter, M.D., and Culbertson, W.C., 1983, Coal resources classification system of the U.S. Geological Survey: U.S. Geological Survey Circular 891, 65 p.

Circular 891 represents an outdated coal resources classification approach, leaving the calculation of coal tonnage completely open and basing the reliability of an estimate of tonnage at a given location solely on the geographical distance from that location to the nearest drill hole, independent of the estimation methods. There are four reliability classes:

- **Measured**[2], for distances up to ¼ mile (mi), meaning highest degree of assurance in estimated values of rank, thickness, and extent;

- **Indicated**, for distances between ¼ and ¾ mi, meaning moderate degree of assurance in estimated values of rank, thickness, and extent;

- **Inferred**, for distances between ¾ and 3 mi, meaning low degree of assurance in estimated values of rank, thickness, and extent; and

- **Hypothetical**, for distances exceeding 3 mi, also meaning a low degree of assurance despite assuming continuity beyond inferred for rank, thickness, and extent of the coal bed.

Note that the classification applies to the properties of coal quality (rank) and coal quantity (thickness and extent) simultaneously, although their evaluations require two different types of assessments and completely different data. We regard the assessment of coal quality based on proximate analysis, ultimate analysis, and measurements of calorific values as a completely different modeling problem that has been studied and reported separately (Olea and Luppens, 2015; Olea and others, 2016; Karacan and Olea, 2018). This study addresses only the quantity part of coal assessment.

The main merit of the Circular 891 classification as related to the modeling of reliable coal assessments is its simplicity, but such simplicity is so extreme that the classification has always had serious drawbacks and has never accomplished its intended purpose, for the following reasons:

- Inconsistency—As a general practice, measures of reliability are specific to the procedures applied for making a prediction (Weisberg, 2014; Briggs, 2016). In Circular 891, the degree of reliability of the estimated tonnages is completely independent of the estimation method.

- Underuse—Uncertainty does depend on the distance to the nearest drill hole, but to a scant degree. The true dependence is not exclusively on such a distance and not directly on distance, but rather on functions of distance that describe the spatial continuity of the deposit. At the very least, the uncertainty depends on the distance to several of the closest data, and the dependency is not linear. These are the basis, for example, of the inverse distance estimation methods (Robinson and

Metternicht, 2006), which do not convey any measure of uncertainty. All the information about distances is implicit in the data, but Circular 891 does not use it.

- Accuracy—The terms "high," "moderate," and "low" to describe degree of reliability are vague, qualitative designations of uncertainty.

- Total resources—Reliability in the calculation of total tonnage in place is based on adding tonnages at the block level within each of the four classes. To the degree that those classes fail to determine the true reliability of thickness at the block level, the four accumulated values are also poor indicators of uncertainty in total resources and reserves.

- Geologic insensitivity—Distance to the nearest drill hole is independent of the geology of the coal bed. Consequently, two coal beds having significantly different geologic characteristics but penetrated by the same drill holes results in the same Circular 891 classification.

- Boundary insensitivity—Circular 891 classifications are insensitive to the complexity of the peripheral extent of a coal bed in the subsurface. Distances between drill holes penetrating or not penetrating the coal bed are not accounted for in the classification.

- Density insensitivity—Circular 891 classifications consider uncertainty in thickness and uncertainty in tonnage as being completely equivalent.

- Overlapping—In practice, the four reliability classes have not proved to be significantly different from one another in terms of the reliability they are supposed to predict. Figure 1 shows an example in which the distributions for the errors obtained when validating the calculations with additional drilling are about the same in all classes.

All these drawbacks prompted James Luppens, at the time with the USGS, to find a better method for evaluating uncertainty in coal resource assessments. Efforts and a literature search for ways to improve the Circular 891 methodology were unsuccessful; a more modern methodology for the accurate modeling of coal resources required a reformulation from the ground up. The modeling of geologic resources is not restricted to coal deposits. For example, the oil and gas industry has made significant improvements in reservoir modeling in the last decades using geostatistics (Pyrcz and Deutsch, 2014), a specialized branch of statistics considering spatial correlation. Geostatistics has also been extensively applied with success in ore mining (Journel and Huijbregts, 1978; David, 1988; Journel and Kyriakidis, 2004; Rossi and Deutsch, 2014), and by 2008 pioneering attempts had been made in coal mining (Hohn and McDowell, 2001; de Souza and others, 2004; Heriawan and Koike, 2008). Following these encouraging examples, the USGS in 2009 started formulating

---

[2]In the text, highlighting in bold indicates that a term is being defined. In the index, page numbers highlighted in bold indicate the location of a term being defined.

a geostatistical methodology to supersede the classification system of Circular 891. Extensive and successful testing indicated the new approach is more accurate than that of Circular 891 (Olea and others, 2011, 2014; Olea and Luppens, 2012, 2014). Recently, efforts by Olea and Luppens resulted in the approval by the American Society for Testing and Materials (ASTM International) of the same methodology described here as a standard practice for the evaluation of uncertainty in coal assessments (American Society for Testing and Materials [ASTM International], 2019).

This report is not one more test of the methodology; it is a user's manual for facilitating the understanding and application of the new methodology applying geostatistics specially tailored to coal resource assessments. Considering that not all users will have a background in statistics, the publication's first section, "Review of basic concepts," starts by defining basic terminology and explaining key concepts necessary for understanding the fundamentals of the approach, the reasons behind each step, and—above all—how to apply the methodology. The second section, "Probabilistic method for coal assessment," is a review of the tools used in the implementation of the assessment. The third section, "Practical application of the methodology," is an example using data from a coal bed in the Fort Union Formation in the Little Snake River coal field and Red Desert assessment area in south-central Wyoming, to show step by step the application of the tools and the form of the results that are typically obtained. Three other coal beds were also assessed but are not reported. The last section, "Workflow," summarizes the assessment procedure. The assessment of any coal bed utilizing this new methodology requires the use of several computer programs, comprising general-purpose programs, public domain software, and special-purpose programs coded for executing steps specific to this probabilistic coal assessment methodology. The special-purpose programs can be downloaded from Olea and Shaffer (2021).
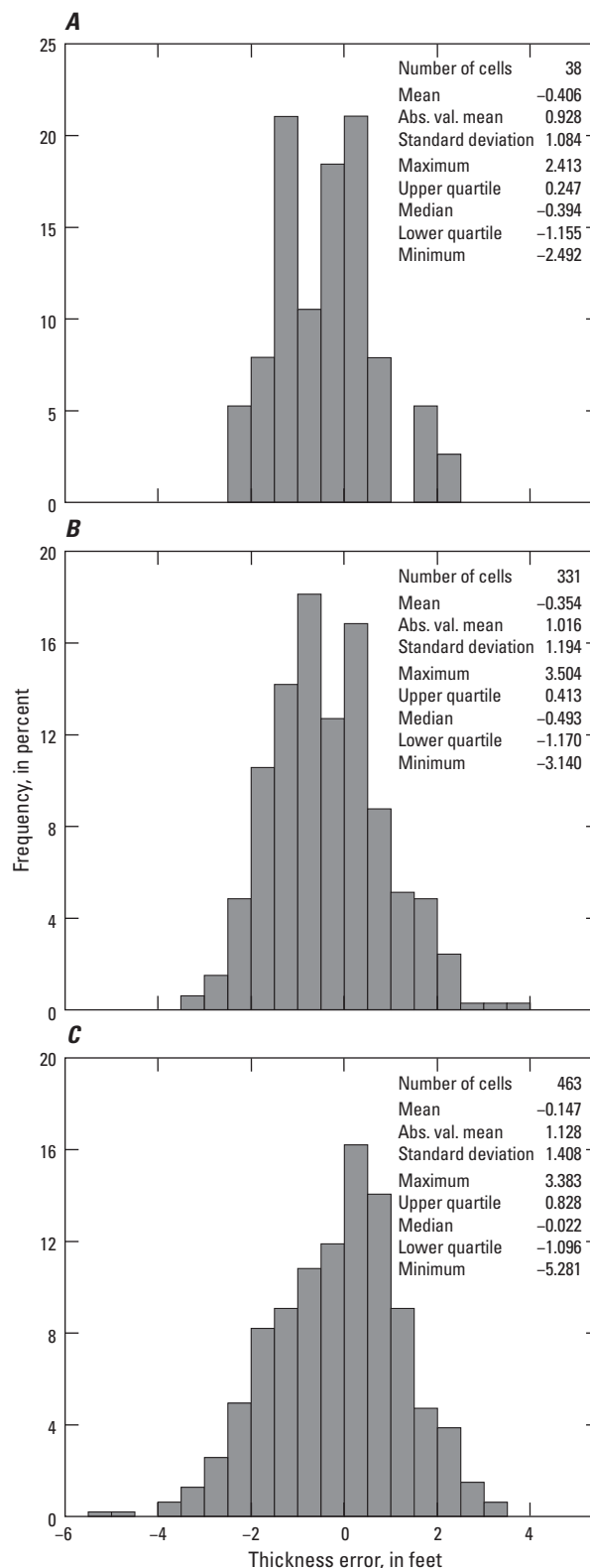
Figure 1. Histograms showing validation errors according to the Circular 891 classification for a lignite in Texas. *A*, Measured. *B*, Indicated. *C*, Inferred. Modified from Olea and others (2011, fig. 19). Abs. val., absolute value.

Review of Basic Concepts

A. Statistical Concepts Relevant to Modeling Uncertainty

Random Variables

Consider the following example: a meteorologist at Denver International Airport makes a temperature reading of 76.2 °F in early July. It is sunny, without a single cloud in the sky. The temperature was measured according to all standard practices. Hence, the temperature at the airport at that moment is 76.2 °F, with absolute confidence—no less, no more. This is an example of an exact value in statistical terms within instrumental precision, which is a single real number. The value of the temperature at the same place within 24 hours or a week later cannot be stated with the same confidence. There are no direct measurements for those values. The meteorologist can only analyze existing data and come up with predicted values for temperature and other information. A fundamental difference between the measured and the predicted conditions is uncertainty. Now, several values are possible: the temperature may be 69, or 73.6, or even 80 °F. Instead of a single number describing temperature in the future, there are several possibilities. Those possibilities are not all equally likely to occur. Given the season of the year, 110 °F is 12 degrees higher than the highest temperature ever recorded for that day, a value that is more unlikely to occur than values close to the average temperature for the month of July.

Statistics is in part a science devoted to making the best estimates or projections out of uncertain situations such as the one described above. Proper application of statistics requires an understanding of several concepts and formulations that will be covered in this section. Readers having a command of statistics may skip this section. The first concept is that of **probability**, which is a measure of the likelihood of the occurrence of an event. Such events can range from a simple situation, such as getting number 11 when rolling a pair of dice, to a complex situation, such as the election of a candidate for city mayor. Basic properties of probabilities, as defined in Hogg and others (2018), consist of the following:

a. The minimum value is zero, which is interpreted as an impossible occurrence.

b. The maximum value is 1, which is interpreted as the occurrence of the event with absolute certainty.

c. Probabilities of events that cannot take place simultaneously are additive.

It is also customary to measure probabilities in percent. A fundamental practice in statistics is to associate a probability with every possible outcome. Table 1 lists rounded temperature values that our example meteorologist would consider likely within one week. The set of outcomes (values) and their associated probabilities of occurrence are known as a **random variable**. This formulation has all the elements to numerically analyze uncertainty independent of the nature of the process being characterized or the form in which the probability values are obtained. For discrete events, such as sunny, cloudy, or rainy, individual probabilities ordinarily have a direct significance. For variables with a continuous numerical fluctuation in values, it is more meaningful to analyze intervals of values, called **confidence intervals**. In the example in table 1:

- There is 0.08 probability that the temperature will be less than or equal to 70 °F.

- The probability that the temperature will be between 70 °F and 80 °F is 0.91.

- There is a 0.38 probability that the temperature will be above 75 °F.

- The temperature will not go below 68 °F or above 83 °F.

Random variables and their properties have been widely used to secure profit at casinos, inform about weather predictions, evaluate investment portfolios, and assess gold mines, among many other applications. The use of random variables certainly can be extended to the assessment of coal deposits. Conceptually, the probabilistic assessment of coal resources does not require any additional knowledge of statistics beyond understanding the power and capabilities of tabulations like that in table 1. Most of what follows in this report is a description of the methods that can be applied to generate reliable probabilities, reliability that can be tested by collecting additional measurements. Let us continue with more definitions and explanations.

Table 1.   Hypothetical example of a random variable.

Temperature (°F)	Probability
68	0.01
69	0.02
70	0.05
71	0.06
72	0.08
73	0.12
74	0.15
75	0.13
76	0.11
77	0.07
78	0.05
79	0.05
80	0.04
81	0.03
82	0.02
83	0.01

Histograms of Random Variables

Probabilities commonly diminish away from a most likely value. Graphical displays usually are a good help when trying to have a collective insight into multiple numbers. Statistics makes extensive use of illustrations. Figure 2 is a histogram of the values in table 1. A **histogram** is a collection of bars with heights scaled proportionally to the probabilities (Salkind, 2007). The values in the example in table 1 ordinarily would have been generated running a weather prediction model, which incorporates physical information related to the problem of concern.
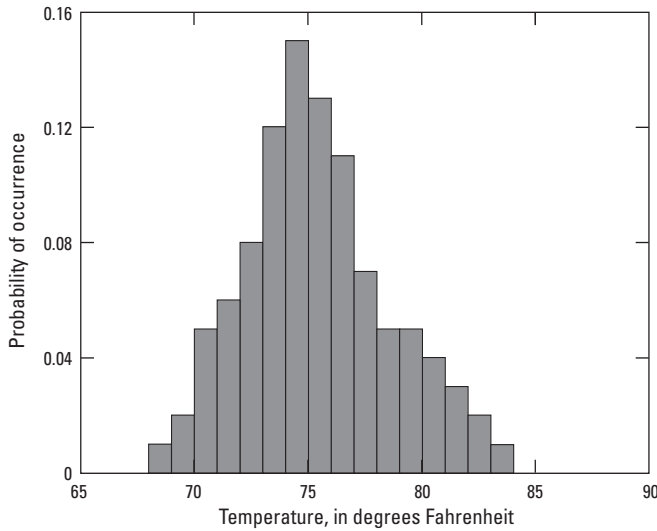
Figure 2.   Histogram for the values shown in table 1.

Normal Distributions

The collection of values making a random variable may originate from general analytical expressions, which are not necessarily linked to any natural phenomena. Such analytical expressions are called **probability distributions**. Discrete probability distributions, such as those providing the number of occurrence of "heads" that can be expected when tossing several balanced coins, directly provide the probability of occurrence associated with any event: 0.25 for two occurrences of "heads" when tossing two balanced coins. For continuous variables, such as atmospheric temperature or formation thickness, the collective probability for all those values at an interval is given by the area under the curve between the two values of interest. The most widely used continuous probability distribution is the **normal distribution** (fig. 3), also known as the **Gaussian distribution**, whose mathematical expression is (Everitt and Skrondal, 2010):

$$f(x) = \frac{1}{v \cdot \sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2} \cdot \left(\frac{x-u}{v} \right)^2 \right] \qquad (1)$$

where $\exp(w)$ is the natural exponential function 2.71828^w.

As can be observed in figure 3, the distribution is symmetric, with the central part of the peak around parameter u and the spread of the curve determined by v. A normal distribution with $u = 75$ and $v^2 = 10$ would follow closely the top of the bars in figure 2.
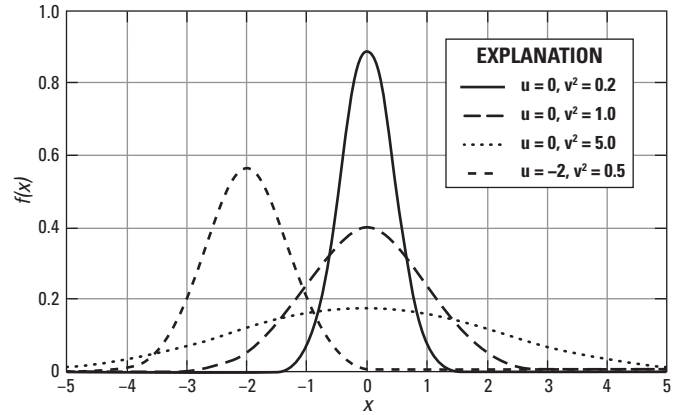
Figure 3.   Graph showing examples of normal distributions.

Histograms of Datasets

Histogram and probability distributions are not exclusively used for characterizing random variables modeling uncertainty. They can also be used for displaying datasets. Figure 4 shows the fluctuations in thickness values of an example coal bed. Explanation of the listed terms follows below. To avoid confusion, the label in the vertical axis is described as "frequency," instead of probability of occurrence, and is measured in percent.
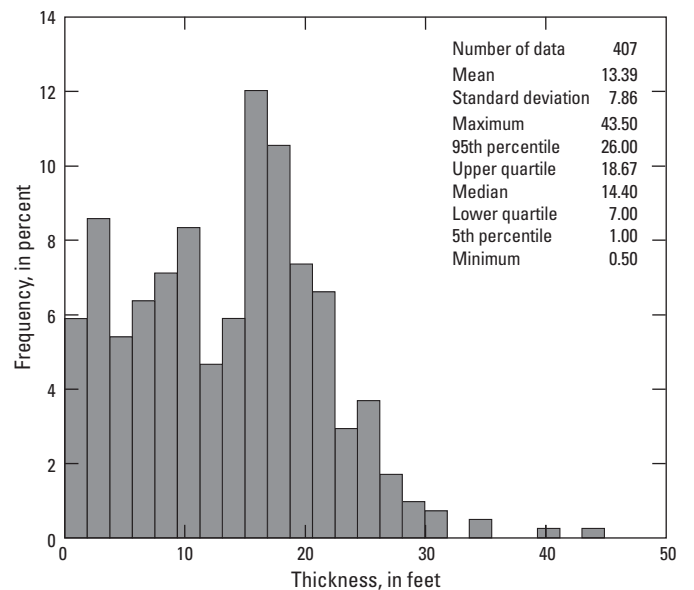
Figure 4.   Histogram showing fluctuation of relative abundance of coal bed thickness values.

Descriptors

Statistics, like any specialized body of knowledge, has developed a terminology for facilitating communication. In this report, an attempt is made to keep the terminology to a minimum as an aid in understanding for readers who may not have a background in statistics. The number of data, the first item in the list of values in figure 4, is also called **sample size**. All the other terms in figure 4 are intended to denote the centrality and dispersion of the data. Specifically:

The **mean** of a numerical dataset, m, is the arithmetic average of all values. It is calculated by

$$m = \frac{1}{n} \cdot \sum_{i=1}^{n} z_i \qquad (2)$$

where n is the sample size and z_i is each one of the values in the dataset. The mean conveys the number around which the values in the dataset fluctuate. The mean can take any value: positive, zero, or negative. It is parameter u in figure 3.

The **standard deviation** of a numerical dataset is the positive square root of the **variance**, s^2, which is given by

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (z_i - m)^2. \qquad (3)$$

The standard deviation is parameter v in figure 3. The variance and the standard deviation intend to inform how tightly the values cluster around the mean. When dealing with data, the higher the variance, the larger the dispersion. For random variables, the higher the variance, the larger the uncertainty.

There is a well-established practice in statistics to rank and then subdivide sets of values—either random variables or datasets—into groups of equal size. Special rules, not fully covered here, apply when the sample size is not a multiple of the number of groups. The general name for the boundaries—**dividers**—separating the groups is **quantiles**. The simplest situation is a subdivision into two groups, in which case the divider is called the **median**. The median of a collection of ranked numbers having an odd sample size, 201 for example, is the central value (101) separating them into two groups of the same size. When the sample size is even, say 200, the median is the average of the largest number in the group of lower values (100) and the smallest number in the group of higher values (101). In figure 4, the median is 14.4 feet.

Two other groupings of values are common, and both are mentioned in figure 4. For a set of values divided into four groups, each divider is called a **quartile**. The **lower quartile** is the divider between the groups containing the lowest and second-lowest values, and the **upper quartile** is the divider between the groups containing the highest and second-highest values. The central quartile coincides with the definition of the median, so the term "central quartile" is not used. Finally, there is the case of **percentiles** (Everitt and Skrondal, 2010), or dividers into 100 classes, which are more meaningful for continuous distributions or large collections of values. The 5th

percentile, for example, is the divider separating the 5 percent of lowest values from the 95 percent of highest values; in figure 4 the 5th percentile is 1.0. The 50th percentile is synonymous with the median, again the prevailing term.

Cumulative Frequency Distributions

The histogram is a helpful display, but it has two drawbacks that can be inconvenient in certain circumstances: (a) the appearance of the graph depends on the way the width of the bars is selected, and (b) unless the quantiles are provided, it is difficult to derive them from the display. The display of **cumulative probabilities (frequencies)** is an alternative or complement to displaying data in the form of histograms (Triola, 2014). Table 1 can be expanded to show the result of calculating the proportion of values less than or equal to any possible magnitude of the attribute (table 2). Figure 5 is a cumulative frequency distribution for the same values shown in figure 4.

Note in figure 5 that the scale in the ordinates (y-axis) is not linear. This is a consequence of the importance of the normal distribution. Here we follow the prevalent practice of using a scale in the ordinates in such a way that if the values follow a normal distribution, the bullets follow a straight line, which is an additional advantage of the cumulative frequency distribution over the histogram when it is important to decide whether the distribution follows a normal distribution or not. The necessary condition of symmetry required for a distribution to be normal is not difficult to assess visually in a histogram, but the exact peakness and spread of a normal distribution is harder to guess in a histogram than deciding if the points fall along a straight line, which certainly is not the case of figure 5. The red dashed line indicates how to find any percentile (in this example the 97th percentile) in case it is not listed somewhere.

Table 2.   Expansion of table 1 to show cumulative probabilities (frequencies).

Temperature (°F)	Probability	Cumulative probability
68	0.01	0.01
69	0.02	0.03
70	0.05	0.08
71	0.06	0.14
72	0.08	0.22
73	0.12	0.34
74	0.15	0.49
75	0.13	0.62
76	0.11	0.73
77	0.07	0.80
78	0.05	0.85
79	0.05	0.90
80	0.04	0.94
81	0.03	0.97
82	0.02	0.99
83	0.01	1.00
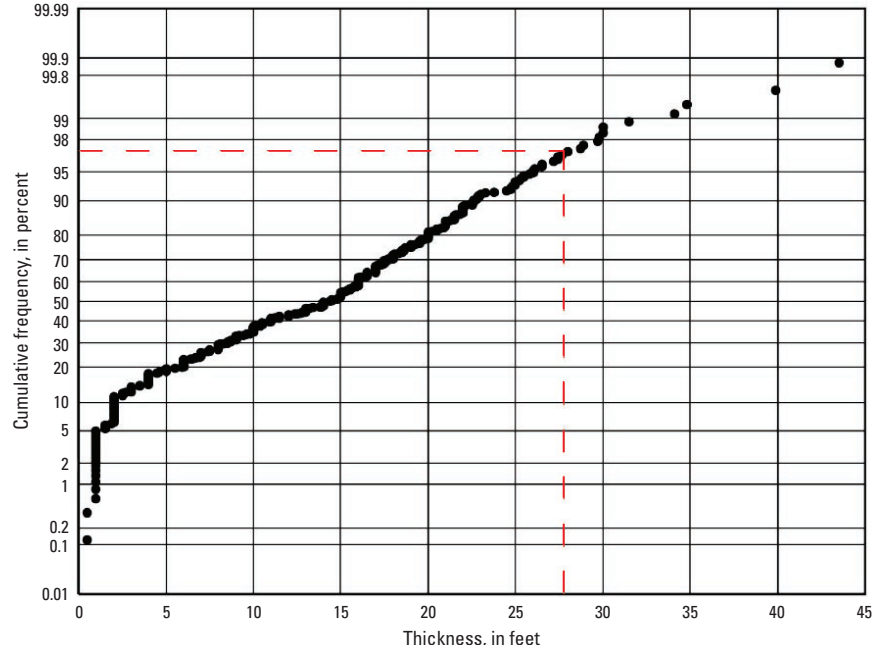
Figure 5. Cumulative frequency distribution for the same dataset displayed in figure 4. The red dashed line shows that the 97th percentile is 27.7 feet.

B. Formulations Relevant to Modeling Spatial Random Variables

Reliable use of random variables requires adequate generation of possible values and associated probabilities. The methods used in this report are part of what is called geostatistics. **Geostatistics** is a branch of statistics specialized in the modeling of attributes having some degree of continuity in spaces up to the third dimension. Such attributes sometimes are called **regionalized variables** (Chilès and Delfiner, 2012). Some of the geostatistical methods are special adaptations or direct borrowings from older branches of knowledge such as time series analysis, signal processing, and pattern recognition. Geostatistics is a collection of modeling tools that has evolved from the work by the French engineer Georges Matheron (1963, 1965), who was primarily interested in having accurate estimates and assessments of uncertainty in ore mining. Geostatistics is widely applied in other disciplines, even outside the earth sciences and engineering. An application to coal resource assessments is a minor expansion by comparison.

Correlation is a term denoting interdependence (Everitt and Skrondal, 2010). Two main differences between geostatistics and classical statistics that are relevant to this report are: (a) in geostatistics, it is necessary to keep track of the location where the measurements have been taken, and (b) a set of measurements, coal bed thicknesses for example, needs to be analyzed in terms of spatial distance—geographical distance in the two dimensions of our interest—for investigating the degree of spatial correlation.

Semivariograms

The degree of spatial correlation is measured in this report in terms of the **semivariogram**, which is the mean of square differences known only exactly for synthetic datasets generated using a certain semivariogram, but is never known for a partial sampling of some natural process. For partial samplings, such as coal-bed thickness data, the semivariogram needs to be estimated from the same data to be used in the estimations of the regionalized variable. The most widely used estimator of the semivariogram, $\gamma^*(\mathbf{h})$, is (Chilès and Delfiner, 2012)

$$\gamma^*(\mathbf{h}) = \frac{1}{2 \cdot N(\mathbf{h})} \cdot \sum_{i=1}^{N(\mathbf{h})} \left[z(\mathbf{s}_i) - z(\mathbf{s}_i + \mathbf{h}) \right]^2 \qquad (4)$$

where $z(\mathbf{s}_i)$ is a measurement at location $\mathbf{s}_i$ and $N(\mathbf{h})$ is the number of pairs of values a distance $\mathbf{h}$ apart, which is a distance considering both magnitude and direction, say N. 30° E., in two dimensions. Two necessary conditions for the application of equation 4 are (a) the dataset must be free of preferential sampling (for example, higher density sampling where the coal bed is thicker), and (b) the mean value must fluctuate around a constant value, that is, there can be no systematic trend, a requirement that may be selectively satisfied only along certain orientations (Olea, 2006). Being a quadratic estimator, the estimated semivariogram cannot take negative values. A **quadratic estimator** is an analytical expression involving products of up to two variables, which may be a product of a variable by itself—a square. Figure 6 shows a typical collection of estimated semivariogram values, which

increase to reach a plateau, sl, that is known as the **sill**. The distance, **a**, at which such a plateau is reached is the **range**. Although $\gamma(0) = 0$, often the semivariogram seems to start from a positive value called the **nugget effect**, ne, in figure 6.
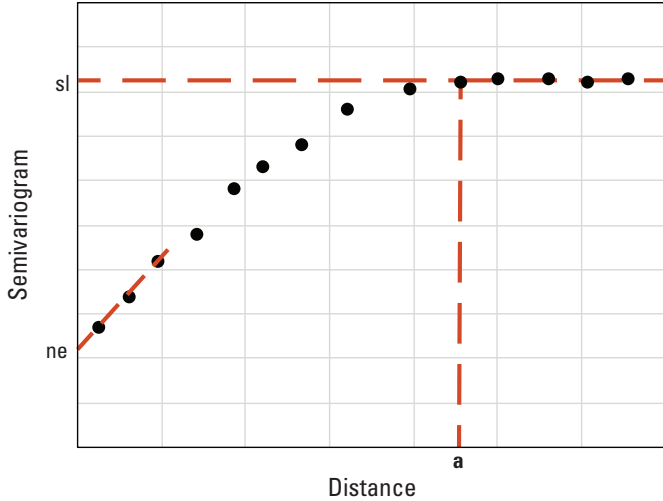
Figure 6. Diagrammatic example of a semivariogram. Dots denote estimated values applying equation 4; semivariogram value ne is the nugget effect, semivariogram value sl is the sill, and distance value **a** is the range.

A semivariogram is said to be **omnidirectional** or isotropic in the special case when the parameters of the semivariogram are the same along any direction. The isotropy and lack of trend of any semivariogram must be investigated by applying equation 4 along at least four different directions, such as N.-S., N. 45° E., E.-W., and N. 135° E. (Olea, 2006).

The semivariogram components are related to important characteristics of regionalized variables. When the maximum distance between locations in the dataset is significantly larger than the range, the sill approaches the variance of the dataset. The range coincides with the concept of zone of influence; data closer than the range are spatially correlated. It only makes sense to apply geostatistics when at least a portion of the distances from the estimation location to the data is less than the range. A corollary is that a properly designed sampling of a regionalized variable should have a significant proportion of neighbors in closer proximity than a distance equal to the range. Finally, the nugget effect is the result of a combination of two factors: (a) measurement errors, and (b) short distance fluctuations at a distance smaller than the shortest distance among the estimated semivariogram values because of absence of closer data pairs. In the presence of a dataset where possible measurement errors are considered nonexistent to negligible, the nugget effect is purely the result of sudden fluctuations at distances below the minimum feasible sampling distance. Thus, the lower the nugget effect—with the minimum being zero—the smoother is a map

of the corresponding variable (Srivastava, 2013). Conversely, a horizontal, pure-nugget-effect semivariogram denotes lack of spatial correlation, a situation in which geostatistics is not applicable.

Kriging

Spatial correlation is revealed by the fact that, on average, there are other small values close to a small value, and the same is true for large values. This form of association is partly random and partly systematic. The random part of the variability does not allow accurate modeling of any large, spatially correlated dataset in terms of classical regression equations. It has been accepted both in theory and in practice for at least 30 years that geostatistics is the approach that consistently produces the most accurate regionalized variables models, plus a reliability measure as a bonus (Caers, 2011). The two main families of methods are estimation and simulation.

Geostatistical estimation methods are generically called kriging. **Kriging** is a generalization of minimum mean square error estimation that considers spatial correlation. Unlike classical regression, kriging does not provide an analytical expression—regression equation—that can be evaluated multiple times at different locations. Kriging is the solution of a system of equations, with the solution to each system of equations providing just two numbers associated with a specific estimation location $\mathbf{s}_0$: a minimum mean square error estimation $z^*(\mathbf{s}_0)$, and a **kriging variance** $\sigma^2(\mathbf{s}_0)$. There are different forms of kriging depending on assumptions about the nature and knowledge about the regionalized variable process. **Ordinary kriging**, which is the most widely used form of kriging, assumes a constant but unknown mean. In such a case,

$$z^*\left(\mathbf{s}_0\right) = \sum_{i=1}^{k} \lambda_i \cdot z\left(\mathbf{s}_i\right) \tag{5}$$

and

$$\sigma^2\left(\mathbf{s}_0\right) = \left(\sum_{i=1}^{k} \lambda_i \cdot \gamma\left(\mathbf{s}_0,\mathbf{s}_i\right)\right) - \mu \quad, \tag{6}$$

where

k	is the size of a subset of the dataset consisting of the observations closest to $\mathbf{s}_0$,
$\gamma(\mathbf{s}_0,\mathbf{s}_i)$	is the semivariogram for the distance between location $\mathbf{s}_0$ and $\mathbf{s}_i$,
λ_i	is a weight determined by solving a system of linear equations comprising semivariogram terms, such as equation 7, and
μ	is a Lagrange multiplier, which also is determined by solving the same system of equations.

For **ordinary kriging**, the system of equations is (Olea, 2012)

$$
\left.
\begin{array}{c}
\sum_{i=1}^{k} \lambda_i \cdot \gamma\left(\mathbf{s}_1, \mathbf{s}_i\right) - \mu = \gamma\left(\mathbf{s}_0, \mathbf{s}_1\right) \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\sum_{i=1}^{k} \lambda_i \cdot \gamma\left(\mathbf{s}_k, \mathbf{s}_i\right) - \mu = \gamma\left(\mathbf{s}_0, \mathbf{s}_k\right) \\
\sum_{i=1}^{k} \lambda_i = 1
\end{array}
\right\}.
\tag{7}
$$

This system results from minimizing the mean square error. This minimum mean square error is the kriging variance. Among all properties of kriging, those of relevance to this report include the following:

- Minimum error—By construction, of all methods to estimate regionalized variables, kriging provides the smallest errors in minimum mean square error sense.

- Unbiasedness—On average, when performing multiple estimations at different locations with the same data, the average error tends to be zero when extra data are available to make validation possible.

- Exact interpolation—When the estimation location $\mathbf{s}_0$ coincides with that of some measurement $\mathbf{s}_i$, then $z^*(\mathbf{s}_0) = z(\mathbf{s}_i)$ and $\sigma^2(\mathbf{s}_0) = 0$, which properly implies that there is no uncertainty.

- Intolerance to duplicated locations—When two measurements $z(\mathbf{s}_i)$ and $z(\mathbf{s}_j)$ are taken at the same location, then $\mathbf{s}_i = \mathbf{s}_j$ and two lines in the system of equations are the same. The system is said to be singular and has no solution or multiple solutions, both of which are an inconvenience to be avoided.

- Declustering power—Although not directly evident from the form of the system of equations, kriging is capable of automatically taking care of the collective influence of a group of measurements close by. Suppose a location $\mathbf{s}_i$ receives a weight of 0.19. If four other measurements are added in close proximity, now all five measurements will receive weights closely adding to 0.19, that is, about 0.04 each. This property is unique to kriging.

- Screen effect—This is another property easier to visualize by running examples than to demonstrate analytically. Suppose a measurement at $\mathbf{s}_i$ receives a weight of 0.22. If one measurement is added at a location between $\mathbf{s}_i$ and the estimation location $\mathbf{s}_0$, the weight at $\mathbf{s}_i$ will drop dramatically, for example to 0.03. This property is behind the justification not to use all available data in equations 5–7 (Leuangthong and others, 2008).

Despite these numerous positive properties and some additional advantages not mentioned here, kriging suffers from a few disadvantages, all relevant when modeling uncertainty:

- Remember that the objective is to measure the uncertainty of any magnitude calculated in the assessment of a coal bed. The estimate and the variance produced by kriging at $\mathbf{s}_0$ are exactly the mean and the variance of the random variable for the attribute of interest at $\mathbf{s}_0$, say, thickness or tonnage. Yet, the mean and the variance in general do not completely determine the form of a probability distribution unless the distribution is determined by these two parameters. This is the special case of the normal distribution in equation 1 and figure 3. Hence, at the price of making a distributional assumption, this drawback can be overcome.

- The mathematical consequence of producing the smallest possible error is that kriging tends to avoid predicting extreme low and extreme high values when it should. This results in **smoothing**, which is the property of making the estimated variable appear to have gentler fluctuations than the actual attribute. This property can be easily verified by calculating the semivariogram for both the data and the estimated values. The latter always has a lower sill. Thus, although kriging maps are optimal in a minimum square error sense, for drawing reliable conclusions kriging offers a distorted view of reality because the optimality of the estimation for each individual location does not translate into an optimality over the entire area of interest as a whole (Leuangthong and others, 2008). In contrast to the incomplete determination of the probability distribution, smoothing has not found a direct solution.

- Finally, there is the problem that the kriging variance is independent from the individual measurements $z(\mathbf{s}_i)$ used in the estimation (eq. 6). This is another undesirable property, describing a situation in disagreement with reality, which dictates that uncertainty is a result of both the local sampling density and the fluctuation in the local data.

Efforts to improve kriging directly have been unsuccessful. Improvements have come from techniques collectively called simulation, although some of the simulation techniques still use kriging. Both kriging and simulation are used in the methodology explained in the following section.

Simulation

Stochastic simulation, like kriging, is not one formulation, but a family of approaches. Let us consider a collection of random variables, one per location over the space of interest, which is two-dimensional (fig. 7). **Stochastic simulation** is the use of these random variables for modeling purposes. In the case of two-dimensional regionalized variables, the objective is the generation of maps that, as a whole,

are as accurate as possible without losing the perspective of honoring the data. The maps are obtained by drawing one value at a time from every random variable. The resulting outcome is called a **realization.** Similar to kriging, stochastic simulation is also a numerical technique providing now one number per location at a time. Multiple approaches have been formulated in the last 45 years for drawing the values.

Most readers may have gone through the experiment of distributing the same values of a geologic attribute posted on a piece of paper and asking several individuals to manually produce a contour map. As in figure 8, invariably the result is to have as many different maps as participants, with the maps all being mathematically correct. This result is possible because of the freedom for interpolating values throughout the dataset. Statistically, this freedom is the uncertainty. In a way, stochastic simulation is a tool for automatically generating maps—realizations—that are compatible with all that is known about an attribute of interest. Stochastic simulation uses multiple realizations covering the realm of possibilities as the approach for modeling uncertainty. As for how many are sufficient, it depends on the sampling density and the complexity of the attribute to be modeled. A rule of thumb is to consider that 100 to 200 realizations are sufficient (Deutsch, 2018). The best way to investigate the adequacy of the number of realizations used is to conduct a sensitivity analysis and to check for stability in the results of interest.

Several possibilities are available for satisfying the requirement of having realizations that are realistic as a whole. Here, the choice has been to rely on the most widely applied approach among simulation methods using the semivariogram and kriging as the driving engine: sequential simulation (Gómez-Hernández and Cassiraga, 1994). The building block of sequential simulation is a simple theorem stating that the joint probability distribution $\text{Prob}[A \cap B]$ of two events, A and B, can be decomposed as the product of two simpler distributions:

$$\text{Prob}[A \cap B] = \text{Prob}[A] \cdot \text{Prob}[B|A], \qquad (8)$$

where $\text{Prob}[B|A]$ denotes the probability of B when A has occurred. Recursively, the expression can be expanded to any number of events, values of coal thickness at certain locations in our case.

We employ two forms of sequential simulation in the assessment methodology: sequential Gaussian simulation and sequential indicator simulation (Pyrcz and Deutsch, 2014). **Sequential Gaussian simulation** requires that all random variables follow a normal distribution, certainly not the case for most of those in figure 7. For consistency with this assumption, the original data must be transformed to values following a normal distribution (Leuangthong and others, 2008), and for convenience, the mean is brought to zero and the variance to one. These transformed values receive the name of **normal scores**, corresponding to the dotted curve in figure 3. Then a random visitation schedule is decided for all locations making what will be a realization. The first location is kriged and the result $(z^*(\mathbf{s}_0), \sigma^2(\mathbf{s}_0))$ is used to define a normal distribution. Then two steps follow that separate conventional kriging from sequential Gaussian simulation: (a) the value going into the realization is not $z^*(\mathbf{s}_0)$, but a value $z_g(\mathbf{s}_0)$ drawn at random from the normal distribution $(z^*(\mathbf{s}_0), \sigma^2(\mathbf{s}_0))$, which is the $\text{Prob}[A]$ part of equation 8; and (b) the value $z_g(\mathbf{s}_0)$ is added to the dataset; this is the $|A$ part at the end of equation 8, so as to be able to have $\text{Prob}[B|A]$ in the simulation of the following location. The random drawing of values according to the results of kriging and the addition of the simulated values to the original data continue until simulating the value for the last location. All results are in terms of normal scores. Consequently, the last step is a transformation of the simulated values back to the original scale, say, thickness, which are not necessarily normal after the back transformation.

Sequential indicator simulation is quite similar to sequential Gaussian simulation, except that instead of working with normal scores the method works with categorical values and makes no distributional assumptions.

The ability of sequential simulation to generate different results with the same data is a consequence of the random elements in the procedure—first, the random drawing of values based on the results from kriging, and second, the change of the visitation schedule, thus adding different
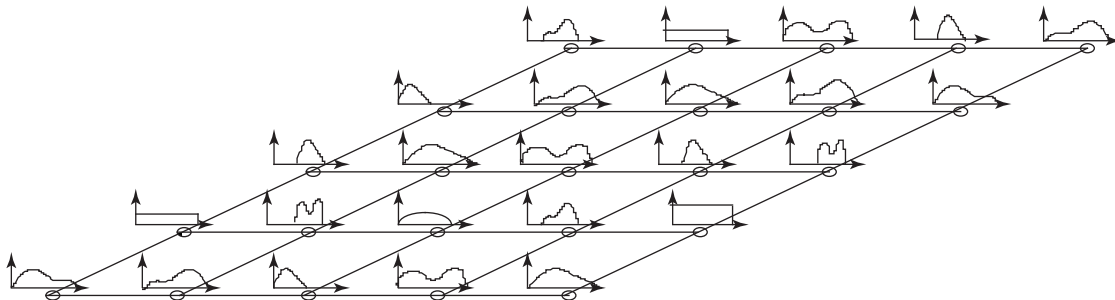
Figure 7.   Diagrammatic view of a collection of random variables over a two-dimensional space. From Christakos and others (2005, fig. 21). Used with permission of Springer Nature.

simulated values to the original data in changing sequences. From realization to realization, even if a few simulated values may be the same at some locations, in general simulated values $z_g(\mathbf{s}_i)$ will be different from realization to realization. However, in a properly implemented simulation, the values at the location of a measurement are all equal to the measurement.

For examples, the reader can inspect the "Practical application of the methodology" section in this report. Some positive properties of sequential simulation are as follows:

- All data are honored, so all realizations coincide in value at the sampling locations.

- The realizations do not present the smoothing of kriging, which is revealed by the same estimated semivariogram for both the data and the realizations.

- Instead of having one map per attribute, now there are several maps, all with the same likelihood of being the correct answer. This result is the key for the probabilistic assessment of uncertainty using simulation. When fixing the attention at one location, there may

be as many possible values as realizations generated (fig. 9). The simulated values constitute the outcomes of the random variable for a probabilistic assessment. Figure 9 is no longer one of those conceptual distributions in figure 7; figure 9 is equivalent to figure 2, used to illustrate the concept of the random variable. Validation of these distributions with additional observations have confirmed excellent capacity for predicting errors (for example, Olea, 2018).

- Measures of dispersion in the nodal distributions, for example, the standard deviation, depend not only on the semivariogram and the sampling pattern, but also on the data—complexity that has proved to be key in the improvements behind simulation.

- By adding the resources per each realization, it is possible to have a random variable assessing the uncertainty for total resources, total coal in place in the case of our interest. This is something not possible to obtain even by using kriging.
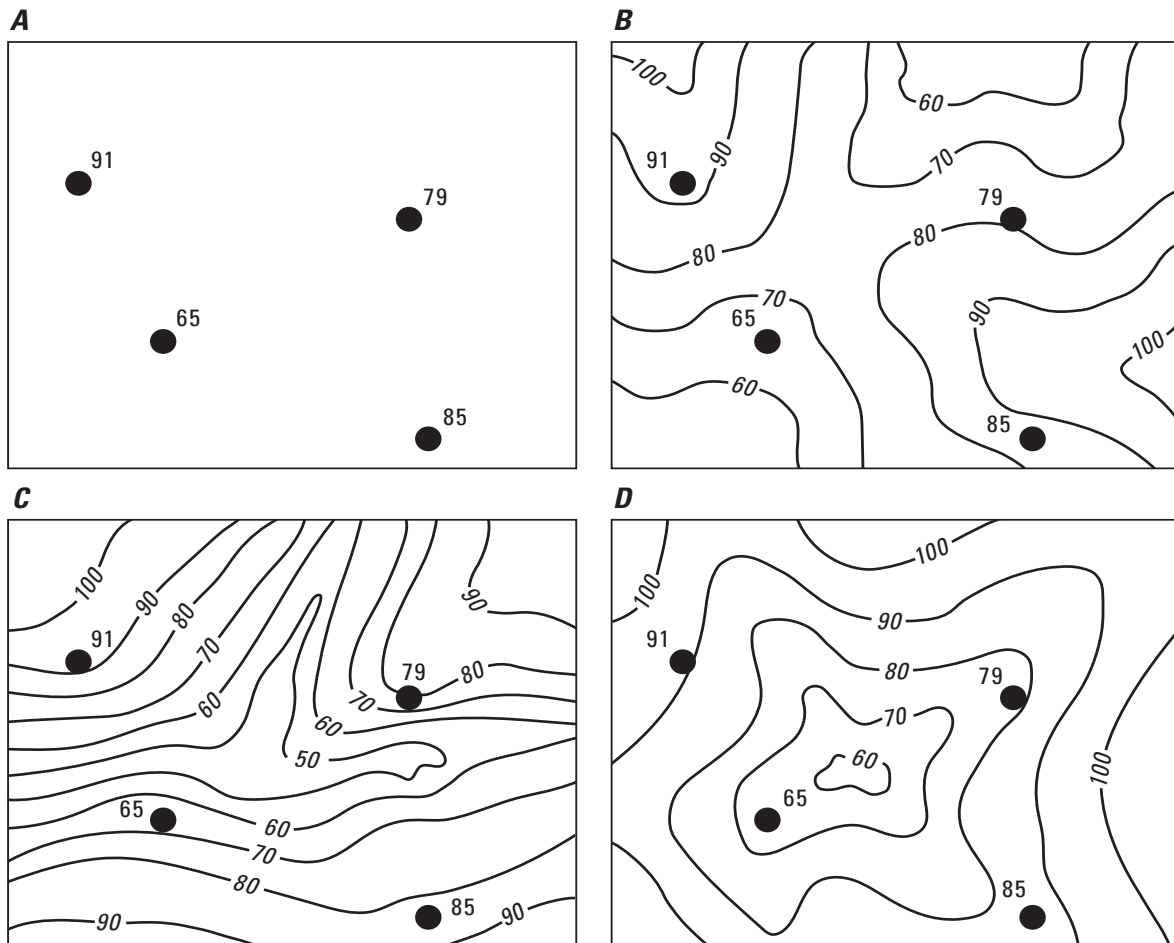
Figure 8. Example of different contour maps (*B*, *C*, *D*) derived from the same data in map *A*. From Olea and others (2011, fig. 2).

However, there are three main disadvantages to simulation:

1. It is not possible to simultaneously minimize the prediction error and have a realistic global rendition of the attribute of interest; the simulation mean square errors are larger than the kriging mean square errors on average (Olea, 2009). This is an unavoidable fact.

2. The modeling is more demanding than kriging, about as many times more demanding in terms of processing time and space allocation than kriging as the number of realizations. The solution for this inconvenience is to use large and fast computers, which were not available when kriging was introduced.

3. The other inconvenience to some users was already mentioned: existence of multiple maps, all of the same nature, instead of a maximum of two maps using kriging, with one map showing estimated values and the other map showing kriging variance. This problem, however, can be solved by preparing summary maps. Thus, for example, by collecting mean values at every cell—12.18 in the case of the location displayed in figure 9—it is possible to have what is called an **E-type map** (Srivastava, 2013), which is none other than the same kriging map of analytical estimates. By mapping all standard deviation values—7.821 in figure 9—it is possible to have a map that is equivalent, yet not equal, to the kriging standard deviation map.

In summary, despite some imperfections, geostatistics has evolved enough to offer all the elements needed to overcome the drawbacks of modeling uncertainty using Circular 891 that were mentioned in the Introduction.

Probabilistic Method for Coal Assessment

The methodology can be subdivided into three parts:

A. Modeling the coal bed thickness.

B. Conversion from thickness to tonnage.

C. Probabilistic assessment of the resources.

The methodology relies on the application of kriging and sequential simulation described in the previous section, with the ultimate objective of yielding results that are the basis for the probabilistic assessment of the attributes of interest.

A. Modeling the Coal Bed Thickness

The goal is straightforward: it is necessary to generate coal thickness realizations. In practice, the task is more complex because of practical and theoretical considerations summarized in figure 10.

Collect, check and prepare data

↓

Model the semivariograms

↓

Model the coal bed boundary

↓

Model the coal bed thickness
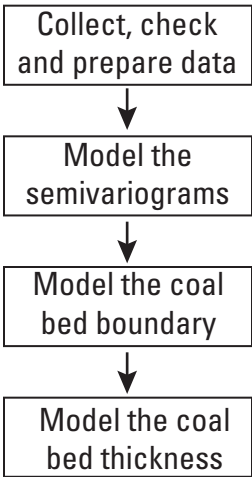
Figure 10. Diagram showing basic steps in the modeling of coal bed thickness. For details on the data, see the "Workflow" section (p. 48).

The Data

The critical information is coal bed thickness data at multiple locations. It is required that these data include not only the actual thickness of the coal bed, but also locations where the coal bed is absent, if the coal bed does not extend throughout the entire study area. If the locations are provided

Number of realizations	100
Mean	12.180
Standard deviation	7.821
Maximum	37.800
95th percentile	32.993
Upper quartile	14.096
Median	9.597
Lower quartile	7.582
5th percentile	4.725
Minimum	2.362

Figure 9. Histogram showing an example of an actual random variable for cell tonnage at a coal bed.

as non-Cartesian coordinates (for example, latitude and longitude) it is necessary to convert those locations to a Cartesian coordinate system.

If the data include information indicating that the coal bed is absent, the thickness values will be considered in two different ways. Those values larger than zero will be used to model the coal bed thickness. All available thickness data will be used to model the geographical extent of the coal bed. For that purpose, it is necessary to prepare a binary variable that is 0 when the location is outside the coal bed and 1 otherwise, namely, each time the value of thickness is larger than 0.

Modeling the Semivariogram

It is necessary to conduct two separate modeling processes, one for thickness and another for the binary variable, being careful to satisfy the two conditions for the proper application of equation 4. Detection of trend and anisotropy requires modeling the semivariogram along multiple directions, which is an important consideration in the semivariogram modeling.

The requirement of the absence of preferential sampling, say, more drilling where the coal bed is thicker, may demand a preprocessing of the data only for the purpose of modeling the semivariogram. There are several approaches for compensating the preferential sampling, but this publication follows a method by Olea (2017) that does not use weights. The first step is always to verify if the original dataset shows preferential drilling. Such a step can be decided by observing a posting of the dataset, or observing a scatterplot of attribute value versus distance to the nearest drill hole, or both. If the compensation is required, the method discards a subset of redundant points. The compensated dataset is always smaller than the original dataset with preferential sampling.

The next step is the estimation of semivariogram values, $\gamma^*(\mathbf{h})$. The increments in $\mathbf{h}$ are usually multiples of a value of the same order of magnitude as the average distance to the closest neighbor in the dataset. For example, if that average distance is 1.5 kilometers (km), a reasonable possibility is to estimate the semivariogram $\gamma^*(\mathbf{h})$ for distances of 1, 2, 3, …, km until reaching values surpassing the range, say 15 km if the range is 10 km. There is specialized software for conducting this modeling. The details are in the "Workflow" section.

The final step is the fitting of an analytical expression to the collection of $\gamma^*(\mathbf{h})$ values. For an example, see the lower figure on page 21. This step is necessary for the proper functioning of the kriging system of equations (eq. 7). Interpolation of semivariogram values or use of an arbitrary analytical model may result in multiple solutions or negative kriging variances. Mathematicians have found several expressions that both follow the never-decreasing form of the semivariograms for distances below the range and assure unique kriging solutions with nonnegative variances. The most commonly chosen expressions are shown in figure 11 and defined in equations 9–11.

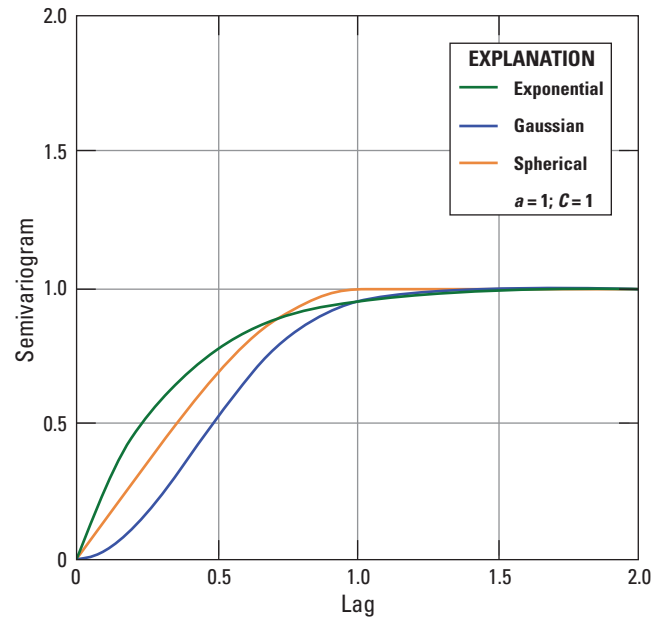The **spherical semivariogram** model is

$$\gamma(\mathbf{h}) = \begin{cases} C \cdot \left[\dfrac{3}{2} \cdot \dfrac{\mathbf{h}}{\mathbf{a}} - \dfrac{1}{2} \cdot \left(\dfrac{\mathbf{h}}{\mathbf{a}} \right)^3 \right], & \text{if } \| \mathbf{h} \| \leq \| \mathbf{a} \|, \\ C, & \text{if } \| \mathbf{h} \| > \| \mathbf{a} \| \end{cases} \qquad (9)$$

the **exponential semivariogram** model is

$$\gamma(\mathbf{h}) = C \cdot \left[1 - \exp\left(-3 \cdot \dfrac{\mathbf{h}}{\mathbf{a}} \right) \right], \qquad (10)$$

and the **Gaussian semivariogram** model is

$$\gamma(\mathbf{h}) = C \cdot \left[1 - \exp\left(-3 \cdot \left(\dfrac{\mathbf{h}}{\mathbf{a}} \right)^2 \right) \right]. \qquad (11)$$

Figure 11.   Graph showing the three most common analytical semivariogram models used in solving kriging systems of equations (eq. 7) (from Olea, 2009).

Coal Bed Boundary

In practical terms, the area of study is always a rectangle with sides parallel to the axis of the Cartesian coordinates, ordinarily east-west and north-south. Although coal beds are locally continuous, it is necessary to subdivide the study area into squares, or pixels, which should be small enough so that their edges are not noticeable in displays at normal resolution. These small squares are called **cells**, their centers are the **nodes**, and the collection of nodes makes a **grid**. The cell size should not be decided purely on the basis of aesthetic considerations, however. In addition, it is important to choose the cell size of any study area to simultaneously minimize the need for computer resources and maximize the number of

the geologic features of interest for modeling, to the extent reasonable within the resolution of the drilling (Hengl, 2006, 2009; Pyrcz and Deutsch, 2014).

All coal beds have a limited extent. Consequently, unless an assessment comprises a laterally continuous portion of the bed, there is always the problem of establishing the boundaries of a coal bed, which may include inner gaps. The boundary is ordinarily known without uncertainty only along outcrops. At other locations, the determination of the bed boundaries may add as much or more uncertainty to an assessment as the modeling of the thickness. Here, the modeling of the bed boundaries is an integral part of the method, but it is done prior to and separately from the modeling of thickness and tonnage.

The modeling of the bed boundaries is done in two stages. First, the region within the study area where it is likely that the coal bed is present is determined. Then, multiple scenarios are generated inside the region using the binary variable. Kriging is acceptable as the right tool for modeling the region. The estimated values allow determining where it is reasonable to expect the presence of the coal bed because this binary-variable estimated value is equal to the probability that the location is inside the coal bed (Leuangthong and others, 2008). The extent of the region is further limited by using the kriging variance to discard areas where the results are too uncertain to decide whether the coal bed is present or absent. Sequential indicator simulation is applied within the unique region for generating multiple realizations for the areal extent of the coal bed. For an example, see pages 22–24.

Modeling the Coal Bed Thickness

The coal bed thickness is also modeled in two stages, but for different reasons than those that pertain to the boundary modeling. First, sequential Gaussian simulation is used to generate thickness realizations within each of the coal bed boundary realizations. The random components in any simulation software are controlled by a number supplied by the modeler called the **random seed**.

The second step takes care of the purely mathematical nature of geostatistical formulations. Differing from a weather forecasting model, stochastic simulation is a general mathematical procedure. In our case, the only items that are considered are the coal thickness measurements, their locations, and the style of the data fluctuation, as specified by the semivariogram. Physical, chemical, and geologic properties of the attribute being mapped do not enter the calculations because it has not been found how to simultaneously consider the statistics and equations controlling natural processes, if any. Such a generality may result in violating some natural laws—in our case, not having a coal bed thickness gradually approaching zero when it is known that the bed pinches out. With the impossibility of incorporating geologic knowledge into the geostatistical methods, the alternative is to refine the maps post-simulation. Here, a procedure has been implemented for allowing the modelers to selectively refine

areas where they postulate that the coal bed gradually thins out and disappears. The modeling is controlled by a maximum distance from any cell to the perimeter of the coal bed. If such a distance is less than this critical distance, the thickness is recalculated as a linear proportion (cell to perimeter distance)/(maximum distance) of the thickness at the closest cell among those not being modified (fig. 12).
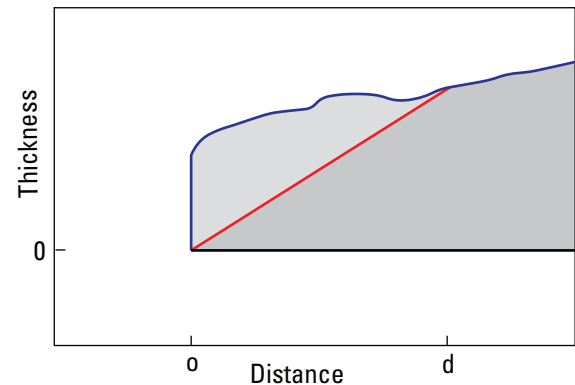
Figure 12. Diagrammatic cross section of the coal bed model. The red line shows a linear pinchout within distance od for correcting the original simulation. Top of coal bed is shown in blue.

B. Conversion from Thickness to Tonnage

The form of the conversion depends on the availability of coal density measurements (fig. 13). If there are no measurements, it is necessary to apply some average value observed at a similar type of coal in some other place, such as 1,770 tons/acre-foot for subbituminous lignite (Wood and others, 1983). The tonnage realizations turn out to be a scaling of the thickness realizations after multiplying them by the conversion factor.
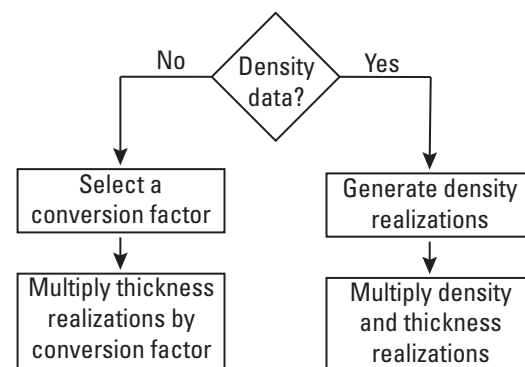

Figure 13. Flowchart summarizing the conversion from thickness to tonnage realizations.

A more realistic modeling is possible when density measurements are available. All locations must be in the same coordinate system used for the thickness values. If there is preferential sampling, a correction must be performed to be able to obtain realistic semivariograms. Sequential Gaussian simulation must be applied to obtain as many realizations as performed for thickness using the same study area boundary, nodes, and coal bed boundary realizations. Taking into account the area associated with every cell, the tonnage maps are obtained by multiplying cell by cell the values of the thickness realizations by those in the density realizations. The pairing must be at random; however, because there is no correlation between the two types of realizations if the precaution has been taken of selecting different random seeds for each set of realizations, then a pairing in the order in which they were generated is fine: the first thickness realization with the first density realizations, the second with the second, and so on until pairing of the last realization of each attribute.

C. Probabilistic Assessment of the Resources

This methodology considers three types of resources (fig. 14). Conceptually the objectives are clear, but more work is necessary to address some technical issues explained below.

Calculations for all three types of assessments are obtained by differently processing the same in-place tonnage realizations, assuring consistency. This methodology provides probabilistic estimates at two different scales for each type of assessment: (a) total resources and (b) resources cell by cell.
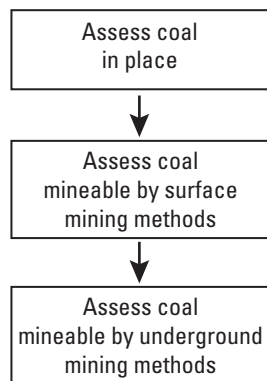
Figure 14. Diagram showing different types of assessments that are possible to obtain based on the tonnage realizations.

Coal in Place

By the definition of the concept of a realization, each tonnage realization is one equally likely characterization of the coal bed. Addition of the tonnage cell by cell provides the total tonnage in place regardless of any economic or technical

consideration. The collection of modeled tonnage values obtained from all the realizations is a numerical characterization of a random variable providing the magnitude and uncertainty of the total resources, a random variable that conceptually is no different than the example in table 1 of the section on basic concepts. The random variable for total tonnage is unique; one histogram or cumulative distribution allows us to see all the information there is. For an example, see the upper figure on page 34. The results should be used to check the stability of the modeling to changes in the number of realizations.

Generation of random variables characterizing cell resources only requires inspecting the same tonnage realizations in a different way. This time, instead of analyzing the realizations individually, it is necessary to observe all realizations simultaneously by fixing the attention on one location at a time. Evidently, at each location, there will be up to as many values as the total number of realizations. If some of the cells are outside the coal bed, the number of values will be less than the number of realizations. Again, this collection of values characterizes, at each location, a random variable modeling the magnitude and uncertainty in the cell tonnage (fig. 9). Considering that the realizations ordinarily comprise several thousand cells, there is need to summarize the results, preferably in graphical form, but in a more condensed form than one histogram per location. For the magnitude of the coal in place location by location, the most efficient way is to display a map of the mean values at each random variable, an E-type map. In term of reliability, there are several options using measures of dispersion, such as the standard deviation or some percentiles. In this methodology we use the difference between the 5th and the 95th percentiles, here called **5–95 spread**. In a scale of 0 to 1, this spread defines a confidence interval having a 0.9 probability of containing the true value. See the map on page 34 and the upper figure on page 35 for an example displaying both a map and a cumulative distribution of the spread. A cumulative tonnage axis has been added for maximizing the functionality of the cumulative distribution. The normal scaling in figure 5 has been abandoned in favor of having a linear scale in the cumulative tonnage.

Coal Mineable by Surface Mining Methods

The display of final results is done in the same form as in the case of coal in place. What is different are the tonnage realizations used for preparing the results, which include some discriminating criteria. This discrimination requires some additional modeling for the purpose of characterizing four attributes: ground surface elevation, thickness of the oxidized (weathered) zone, elevation of the top of the coal bed, and thickness of the **cover**, or the overburden between the ground surface and the top of the coal bed.

Today, digital ground-surface elevation maps are readily available. To utilize digital elevation data, they must be adjusted to follow a grid having the same boundary and nodes as the tonnage realizations. The surface is unique, denoting that the surface is known with certainty.

The modeling of the top of the coal bed is analogous to the modeling of thickness and density when having multiple measurements: preparation of the data, modeling of the semivariogram, and generation of as many realizations as coal bed outlines. Thickness of the cover is obtained directly as the difference between the elevation of the ground surface and the elevation of the top of the coal bed.

The modeling of the oxidized (weathered) zone can be done in two different ways, depending on data available. If in fact there are no measurements for thickness of the oxidized zone, the practice is to borrow an average value from some other coal bed that is considered analogous to the coal bed of interest and has some known measurements. This constant value is certainly an oversimplification of reality with an optimistic degree of reliability. If there are multiple measurements of oxidation thickness, it is better to proceed in a way analogous to the generation of coal-bed thickness realizations. In this case, multiple surfaces are generated that, combined together, offer an adequate influence of fluctuations in oxidation thickness on the uncertainty in mineable resources.

Upon completing this additional modeling, it is possible to go ahead with the assessment of the mineable resources. Tonnage realizations for modeling the surface mining resources have the following minimum restrictions relative to the tonnage realizations used for assessing coal in place:

- The coal bed thickness must be above a minimum value.

- The thickness of the cover must be greater than the thickness of the oxidation zone.

- The thickness of the cover must be less than or equal to a maximum value.

The cells passing these three tests compose the surface mineable resources. Restrictions can be extended with additional considerations, such as resources already extracted or environmental, societal, or technical restrictions (Shaffer, 2017; Shaffer and others, 2019).

Coal Mineable by Underground Mining Methods

This final stage is more straightforward because all the preparatory work can be borrowed from the surface mining modeling. The restrictions on the in-place tonnage are now:

- The coal bed must exceed a cut-off thickness.

- If the coal bed is thicker than a maximum value, only that maximum may be entered into the assessment. This simulates a maximum mining height based on equipment capabilities.

- The cover must be over the maximum value considered in surface mining.

- The cover must not exceed another maximum value— typically the maximum depth of cover where coal can be safely removed without excessive ground-control issues.

Practical Application of the Methodology

Part 1. Geology of the Fillmore Ranch Coal Bed

The methodology is illustrated by applying it to the Fillmore Ranch coal bed, which is in south-central Wyoming (fig. 15) and lies within a stratigraphic interval that is up to 13,000 ft thick and is composed of coal-bearing Upper Cretaceous, Paleogene, and lower Neogene sedimentary rocks (Scott and others, 2019). The Fillmore Ranch coal bed is in the upper portion of the China Butte Member of the Paleogene Fort Union Formation (fig. 16). The Fort Union Formation within the assessment area was deposited in a fluvial paleoenvironment that featured fluvial channel, floodplain, lacustrine, and paludal environments (Hettinger and others, 2008). The Fillmore Ranch coal bed is laterally continuous throughout most of the assessment area (also termed study area or map area), with pods of thicker coal present in the southern and east-central parts. The coal bed crops out along the eastern part of the assessment area and dips to the west from the outcrop as much as 20°. The structural dip flattens to less than 5° along the western edge of the assessment area (Scott and others, 2019).

## Part 2. Preparation of a Thickness Model for the Coal Bed

A. Data

Available thickness data for the Fillmore Ranch coal bed consist of 3,031 drill holes plus 159 digitized points marking the line of outcrop (Haacke, 2018). The drill holes can be subdivided into (1) 902 holes penetrating the coal bed (fig. 17); (2) 222 holes where the coal bed is interpreted to be absent due to nondeposition or to syndepositional erosion based on projected bases of coal elevations; and (3) 1,907 holes penetrating deeper coal beds beyond the outcrop, providing additional information about the absence of the Fillmore Ranch coal bed. Figure 18 displays maps of both thickness and a binary variable denoting only presence or absence of the coal bed, with 6 holes lying south of the map area and 52 beyond the eastern margin of the map area in figure 18A. The notation 0^+ denotes a small thickness value larger than zero. All the drill holes are in south-central Wyoming.

In the data, the locations are in Cartesian coordinates in meters, and thickness is given in feet. The inconsistency was kept to preserve the original units in the processing. On displaying the maps, however, the coordinate system was replaced with the township and range grid of the Public Land Survey System (U.S. Geological Survey, 2018), which is a standard format used on maps in USGS publications.

The maps in figure 18 represent an area measuring 82 by 120 km and extending slightly into Colorado. The average distance to the nearest drill hole in figure 18B is 1,063 meters (m) and the minimum distance to the nearest drill hole is 30 m. A grid having the selected spacing of 100 m requires 820 columns and 1,200 rows, or 984,000 nodes, to cover the same area in figure 18 with the pixel maps that will be required for reporting the results of the assessment. This spacing is within the highest resolution compatible with the drilling density.
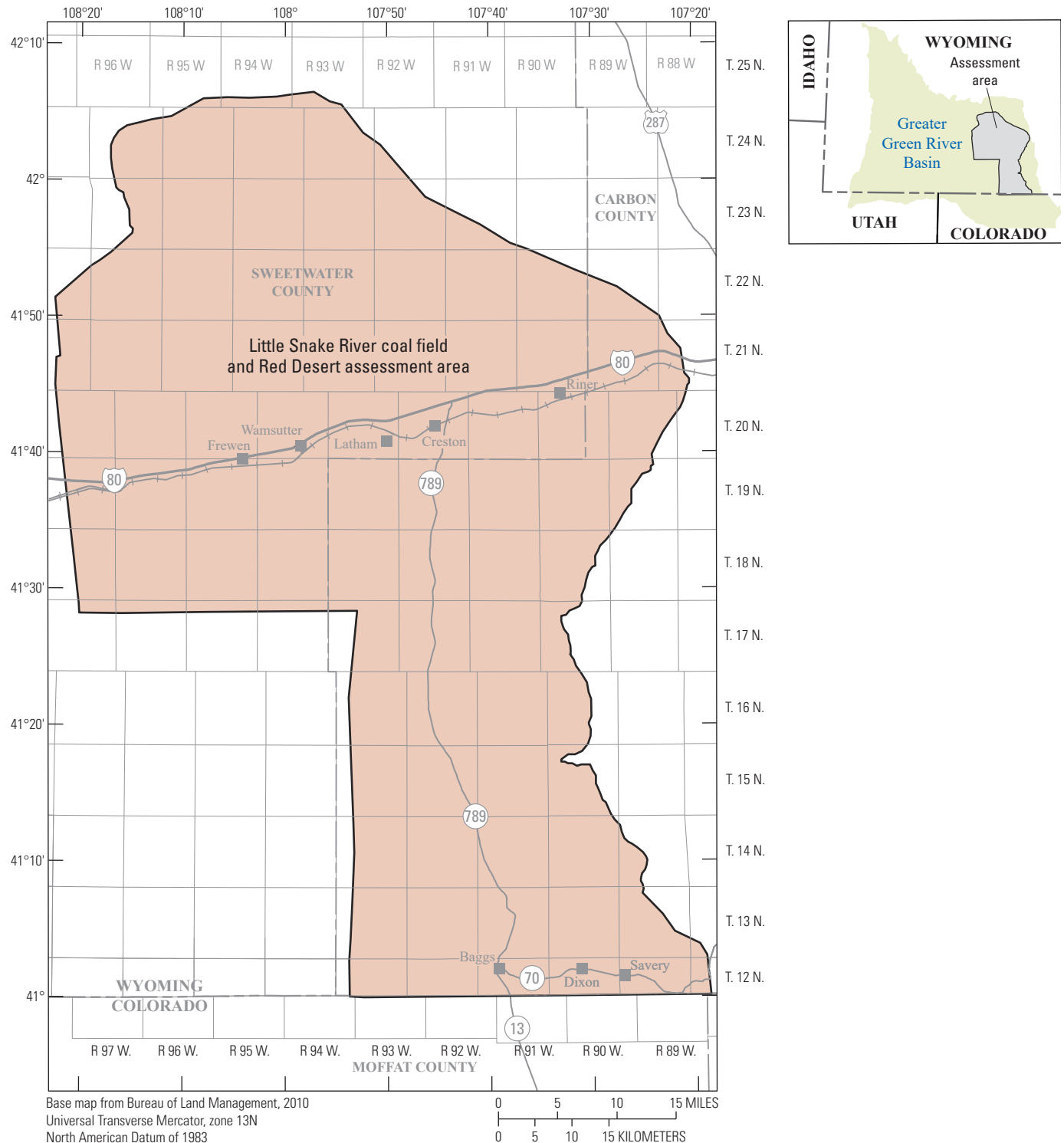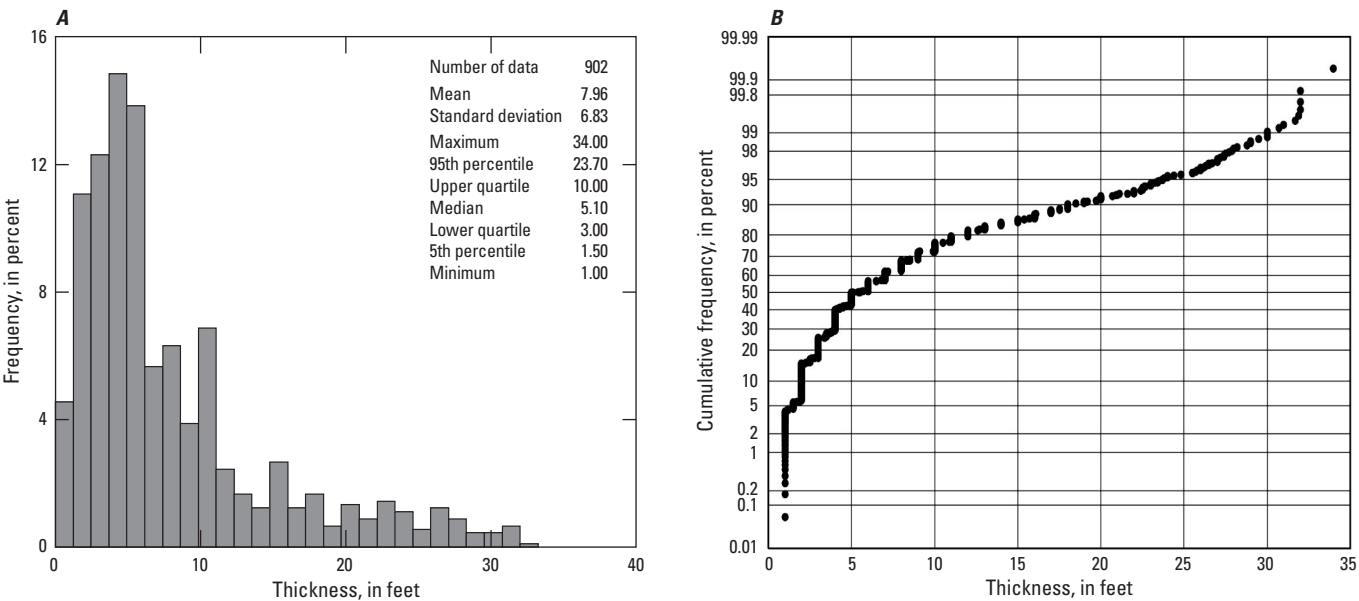

Figure 15.   Map showing location of the assessment area of this study, which contains the Fillmore Ranch coal bed within the Little Snake River coal field and Red Desert assessment area, south-central Wyoming. Modified from Scott and others (2019).

EXPLANATION				
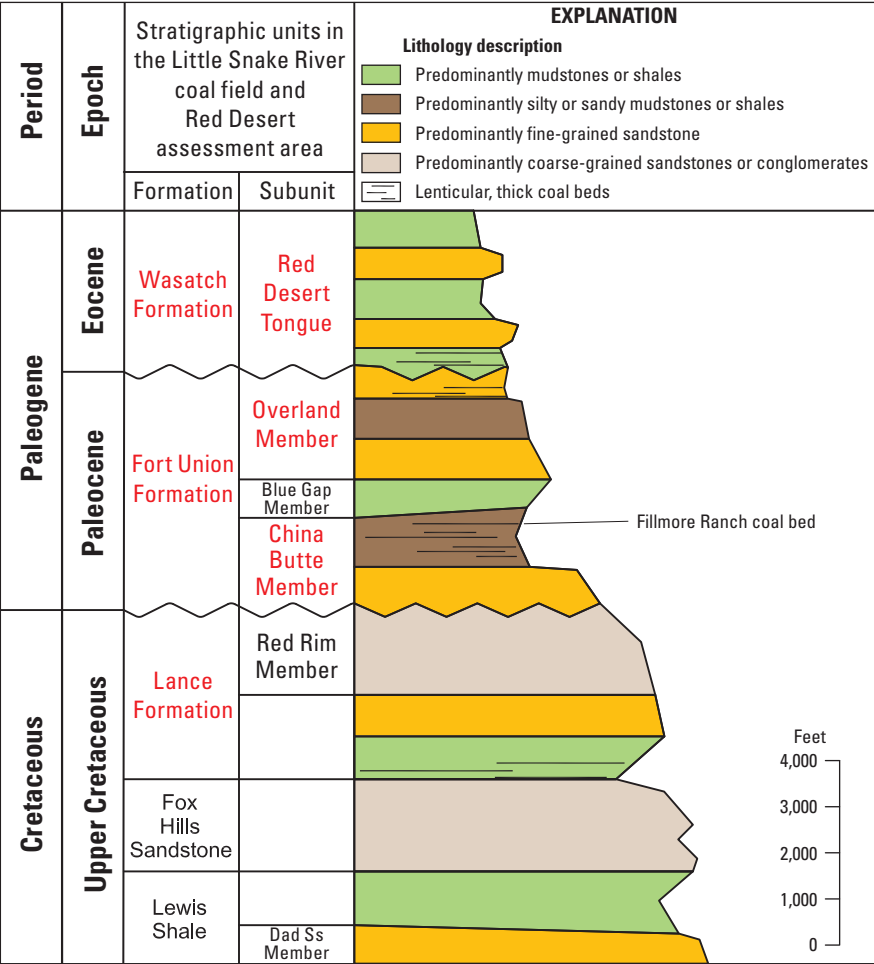
Figure 16.   Generalized geologic column showing the stratigraphic position of the Fillmore Ranch coal bed within the study area. Modified from Pipiringos (1961) and Hettinger and others (2008). Ss, sandstone.

Number of data	902
Mean	7.96
Standard deviation	6.83
Maximum	34.00
95th percentile	23.70
Upper quartile	10.00
Median	5.10
Lower quartile	3.00
5th percentile	1.50
Minimum	1.00

Figure 17.   Thickness data for the Fillmore Ranch coal bed, taken from Haacke (2018). *A,* Histogram. *B,* Graph of cumulative distribution.
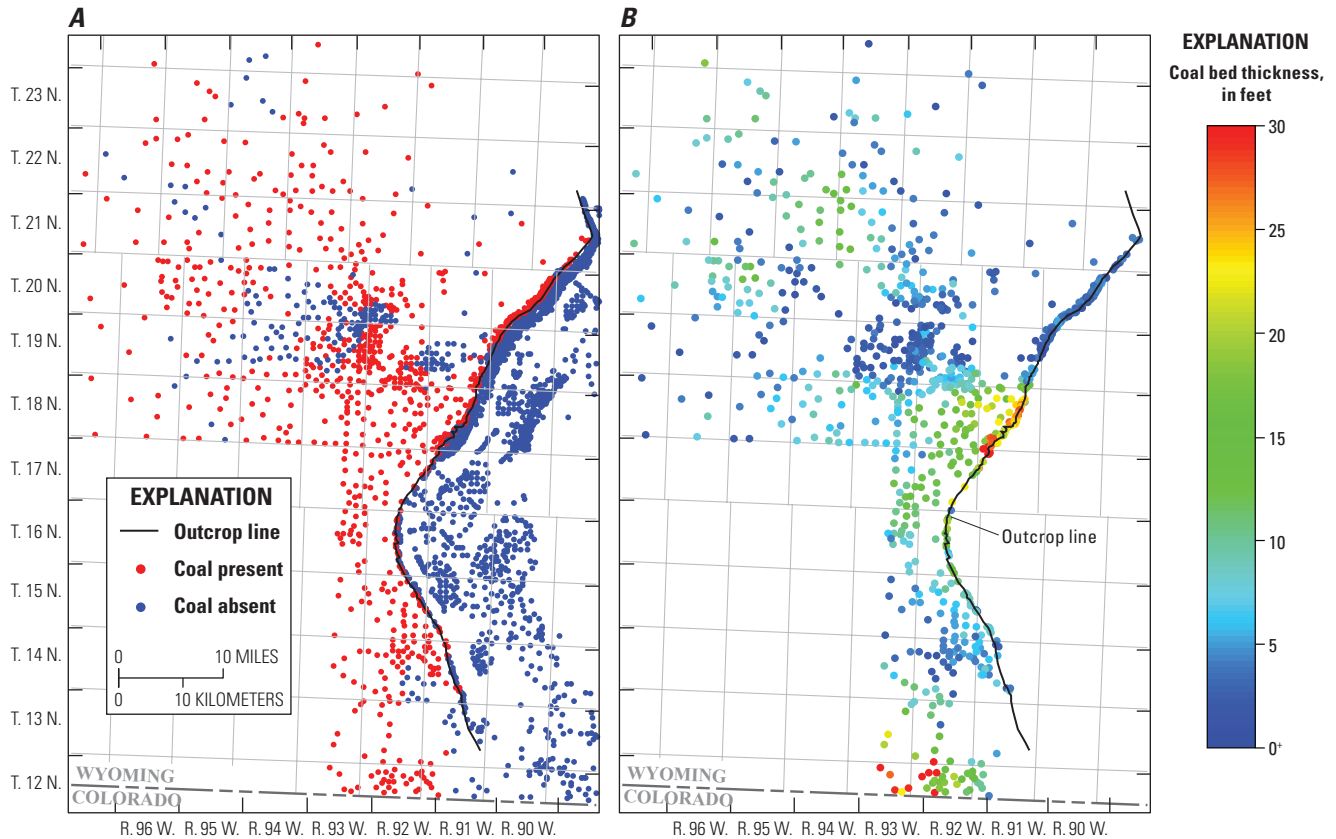
Figure 18. Maps showing display of outcrop and variables for the Fillmore Ranch coal bed. *A*, Posting of indicators of the presence or absence of the coal bed. *B*, Posting of coal-bed thickness values, where 0+ denotes a value close to but larger than zero.

B. Spatial Continuity

Proper estimation of the semivariograms requires data without preferential sampling, which is not the case for any of the variables in figure 18. For regionalized variables, preferential sampling results in clusters of sampling locations at those places where the variable is within the interval of preferential interest. Such clusters can be observed directly in a posting of the data (fig. 18) or in a scatterplot of the variable versus the distance to the nearest observation (fig. 19*A*). Given the skewness in the thickness distribution, the percentile rank of thickness was used instead of thickness directly. From the inspection of figures 18*B* and 19*A*, it is safe to postulate that there are clusters up to 600 m in diameter. A **Q-Q plot** is a scatterplot of the quantiles (percentiles) of a set of values against those of a second set. According to the Q-Q plot in figure 20, most of the discarding takes place among the upper third of highest values between 8 and 30 ft when using a stratification with square cells having sides of 700 m. A 700-m declustering cell size is the minimum cell size that produced a compensated sample without clusters (figs. 19*B* and 21B).

The indicator data have the problem that the number of drill holes at locations where the coal bed is absent is more than twice the number of drill holes where the coal bed is present, whereas the extent of the areas covered by those two types of data is approximately the same (fig. 18*A*). This proportion of indicator values was corrected using declustering square cells having 1,500 m on the side. The declustered indicator data are posted in figure 21*A*.

Figure 22 and table 3 have the two semivariograms of ultimate interest. In the case of thickness, the modeling has been done for the normal scores of thickness because that is the type of semivariogram that will be required in the modeling. No significant trend or variability in results was observed for different orientations. In such a situation, it is acceptable to combine estimated values regardless of orientation, in which case the resulting semivariogram is said to be omnidirectional.
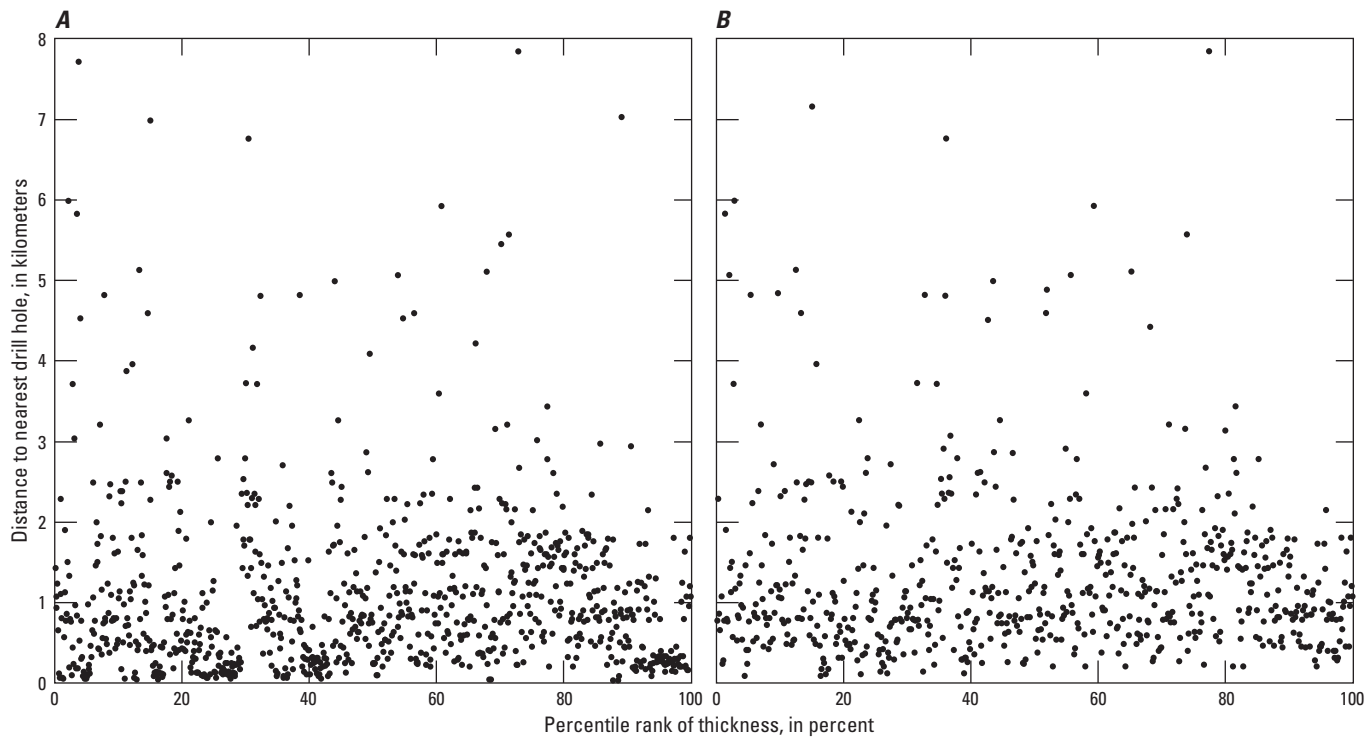
Figure 19.   Scatterplots of thickness versus distance to the nearest drill hole. *A*, Original data. *B*, Result after compensation for preferential sampling.
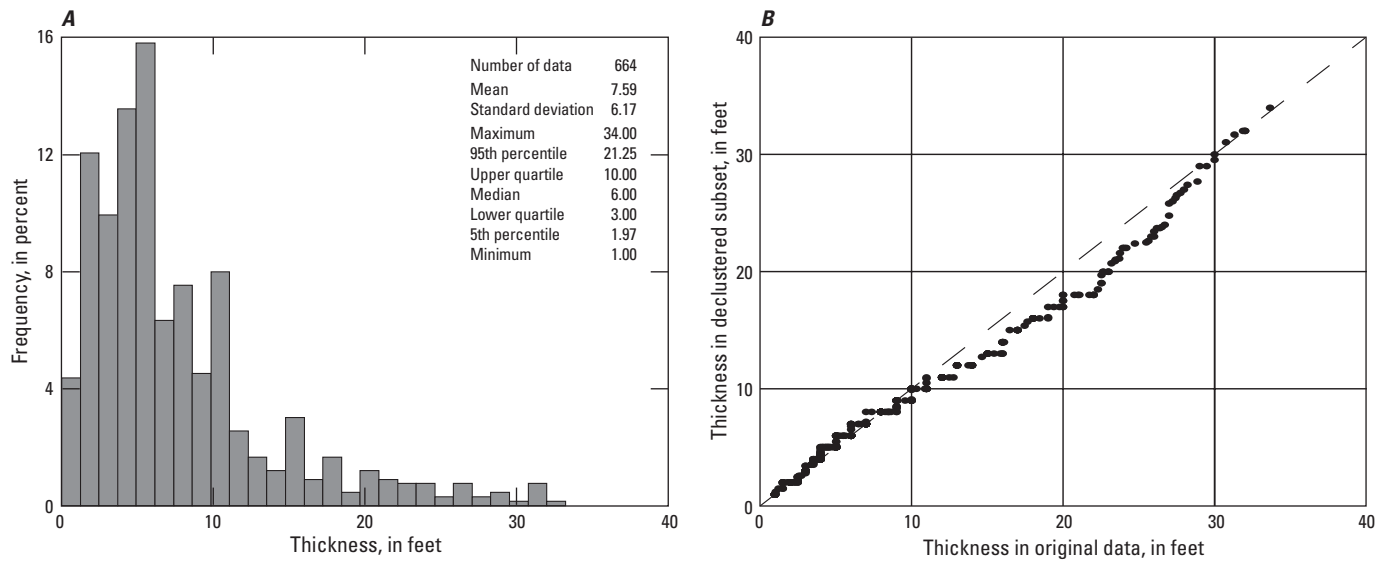
Figure 20.   Graphs showing thickness data after compensation for preferential sampling. *A*, Histogram. *B*, Q-Q plot.
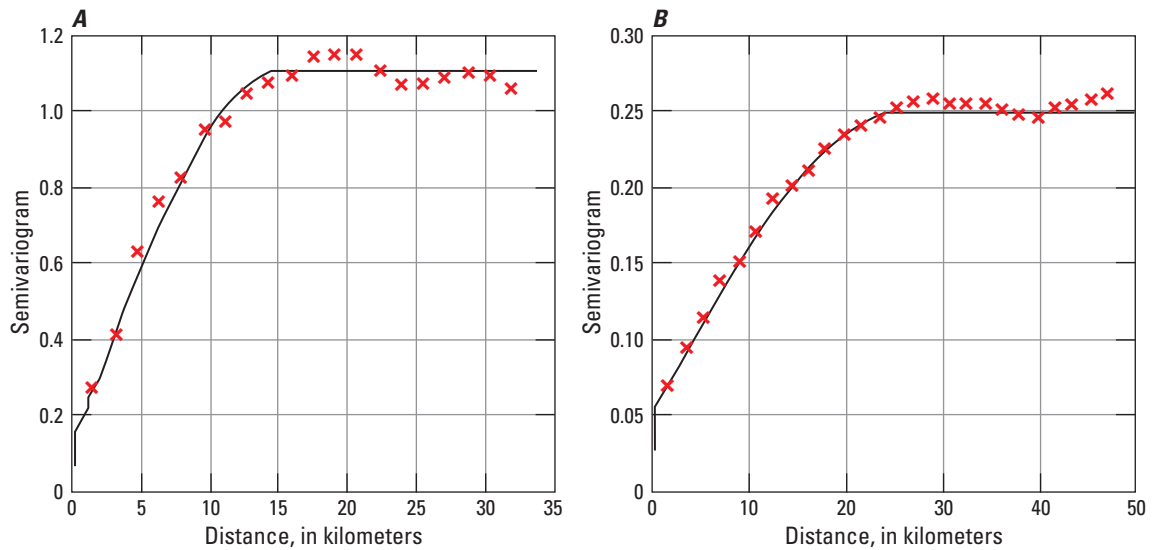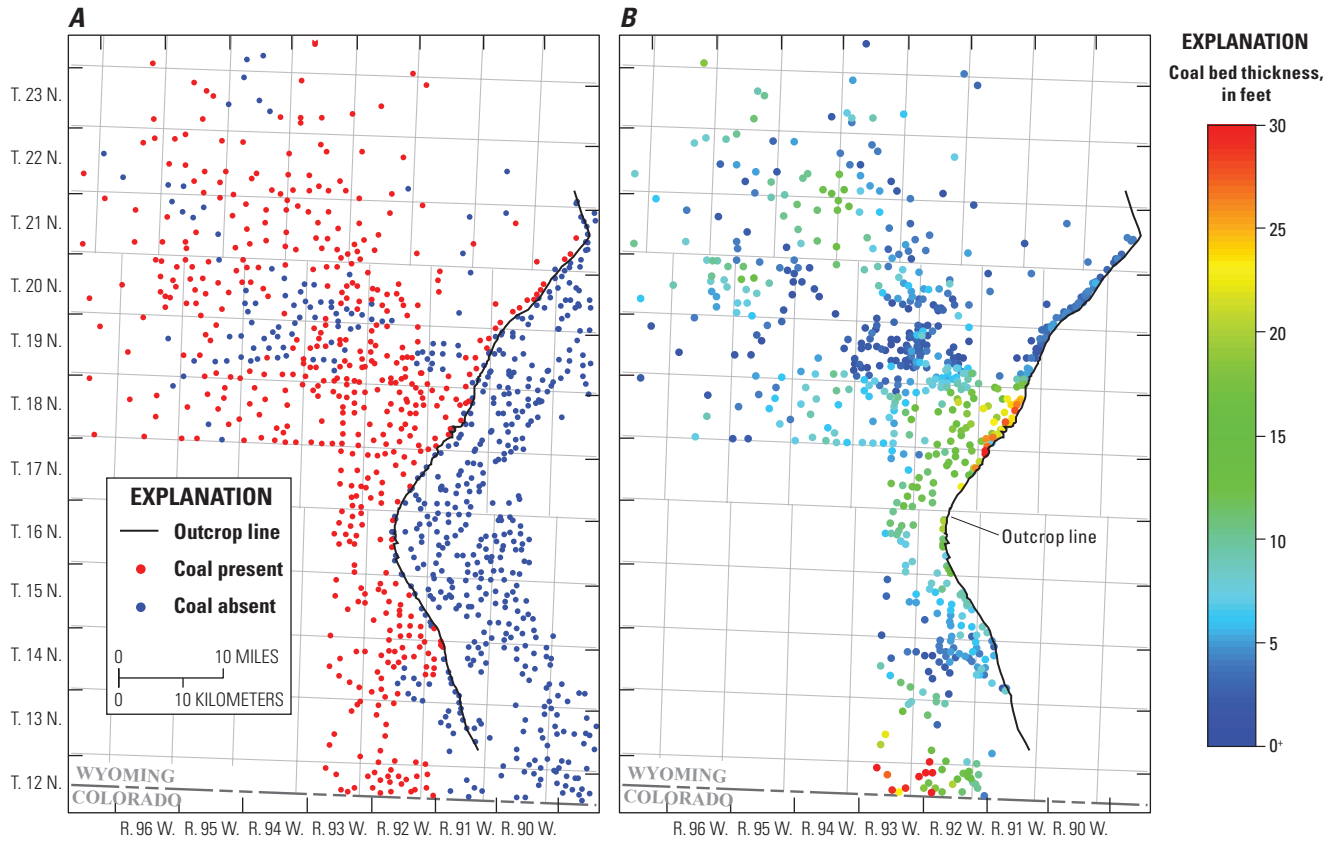
Figure 21.   Maps showing posting of the data after declustering, for the Fillmore Ranch coal bed. *A*, Indicator. *B*, Thickness, where 0+ denotes a value close to but larger than zero.

Figure 22.   Omnidirectional semivariograms based on declustered data. *A*, Normal scores of thickness. *B*, Indicator of presence or absence of the coal bed. In both semivariograms, the red "x" symbols denote estimated values applying equation 4.

Table 3.   Parameters of the semivariogram models shown in figure 22.

	Type	Nugget	Sill	Range (kilometers)
Normal scores	Spherical	0.12	1.11	15.3
Indicator	Spherical	0.05	0.25	26.0

C. Lateral Extension of the Modeling

Presence-or-absence values are used to decide where the modeling of the coal bed is unfeasible. Modeling of the coal bed should be avoided when (a) it is clear that a location is outside the deposit because there is a value at the location indicating that the coal bed is missing, or the location is surrounded by missing values, or (b) the mathematical modeling is not possible because there are no data near the location or the data are insufficient in number. Ordinary kriging is used to numerically assess both possibilities (fig. 23), with the estimate providing information to evaluate case (a) and the standard deviation for case (b). The final decision is always partly heuristic. Based on geologic knowledge about the Fillmore Ranch coal bed boundary, it was decided to discard from the modeling those locations having an estimated indicator less than or equal to 0.48, or having a standard error above 0.28. Figure 24 shows the result of imposing such a restriction on the maps in figure 23. Note that these cutoffs are far from the extreme values of the respected variables.
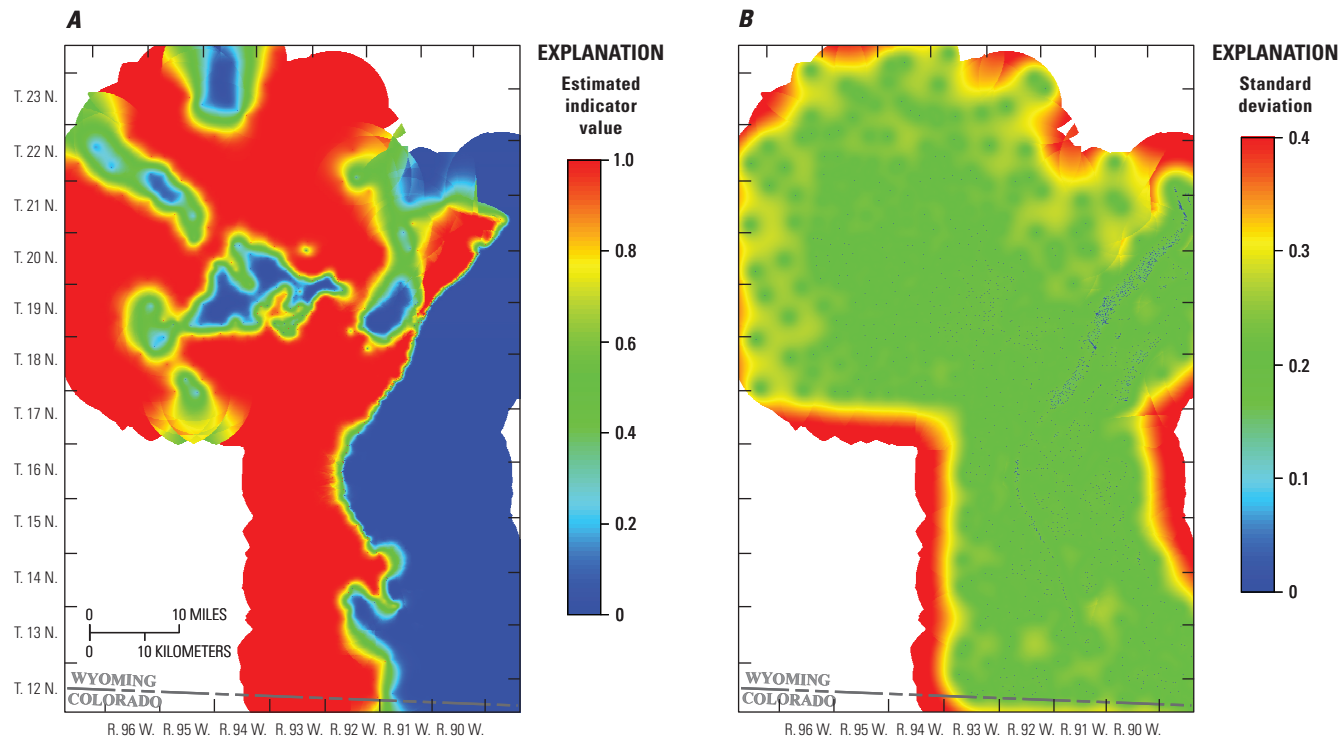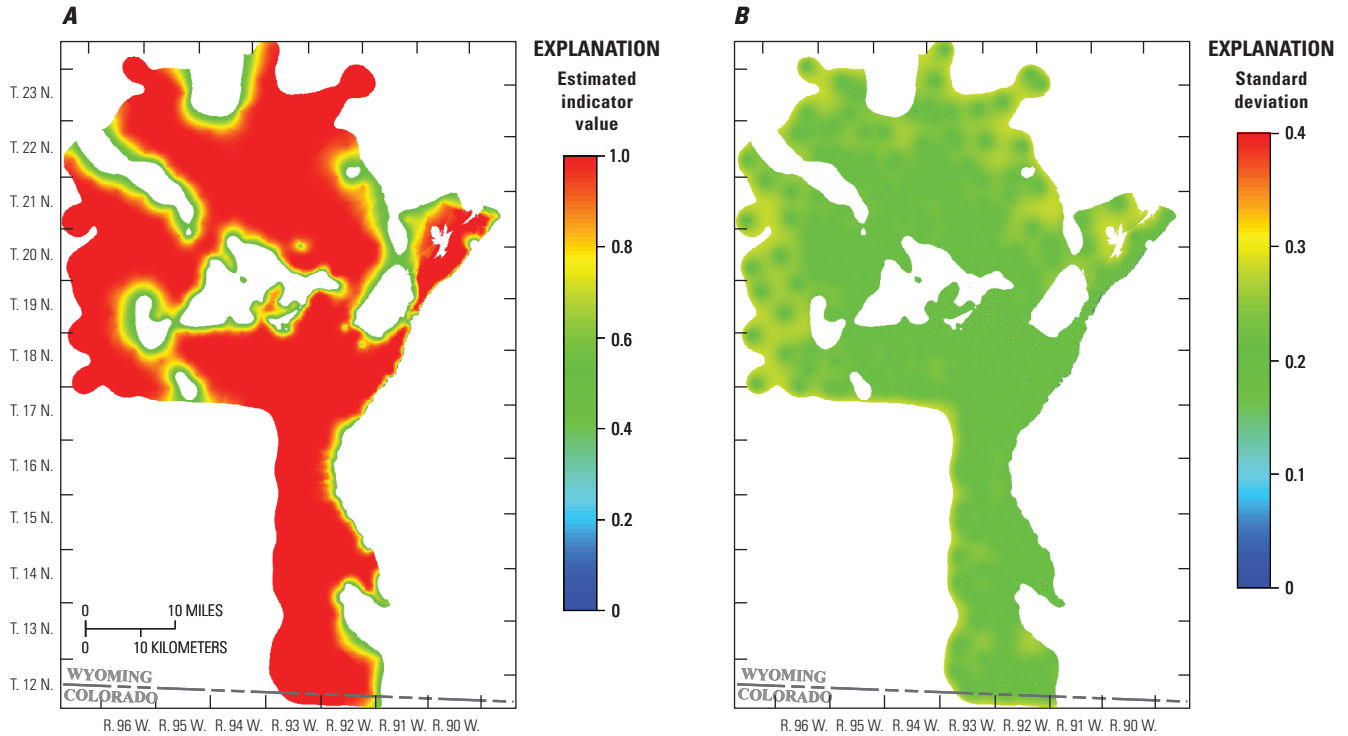

Figure 23.   Indicator maps showing the presence or absence of the Fillmore Ranch coal bed by ordinary kriging, with blank areas indicating lack of data for making any predictions. *A*, Estimated indicator value. *B*, Standard deviation.

Figure 24. Indicator maps showing the region where modeling of the Fillmore Ranch coal bed should not take place, which is the same blank area in both maps. *A*, Estimated indicator value. *B*, Standard deviation.

D. Modeling the Coal Bed Boundary

The previous section deals with determining the locations where the modeling is sufficiently reliable, which is not necessarily synonymous with the coal bed being present, although it should be close because of the way that the modeling region was prepared. The objective here is to decide which parts of the modeling region are indeed inside the coal bed. The uncertainty is resolved through multiple realizations prepared by applying sequential indicator simulation. We decided to start with the minimum number of recommended realizations, 100 (Deutsch, 2018), with the idea to increase the number if the stability of the results recommends generating a larger number of realizations. Figure 25 displays the first 4 realizations out of the total of 100 prepared, realizations that in figure 26 determine the coal bed boundary of thickness realizations prepared by sequential Gaussian simulation.
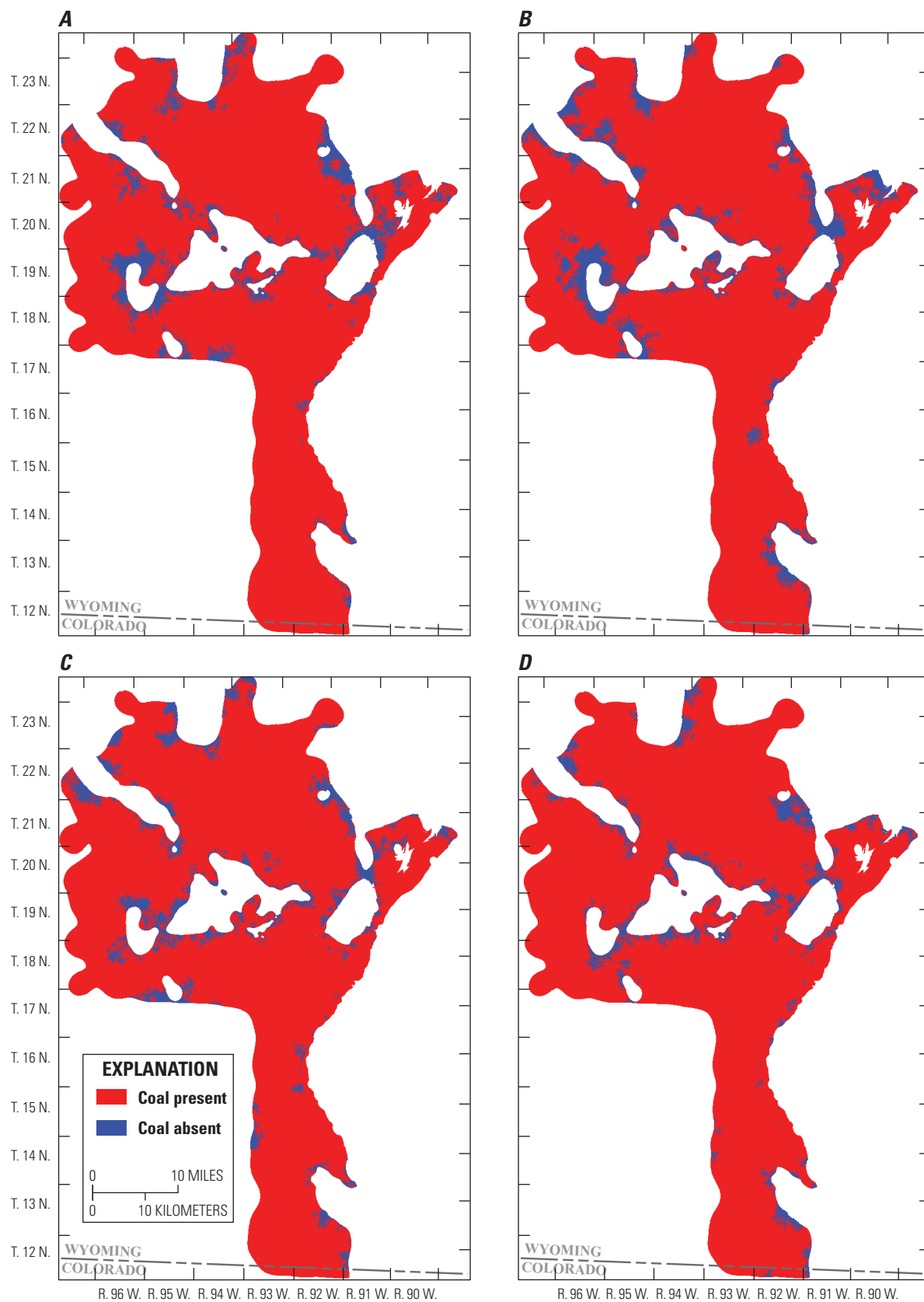
Figure 25.   Maps showing the first four realizations for the coal bed boundary out of 100 realizations.
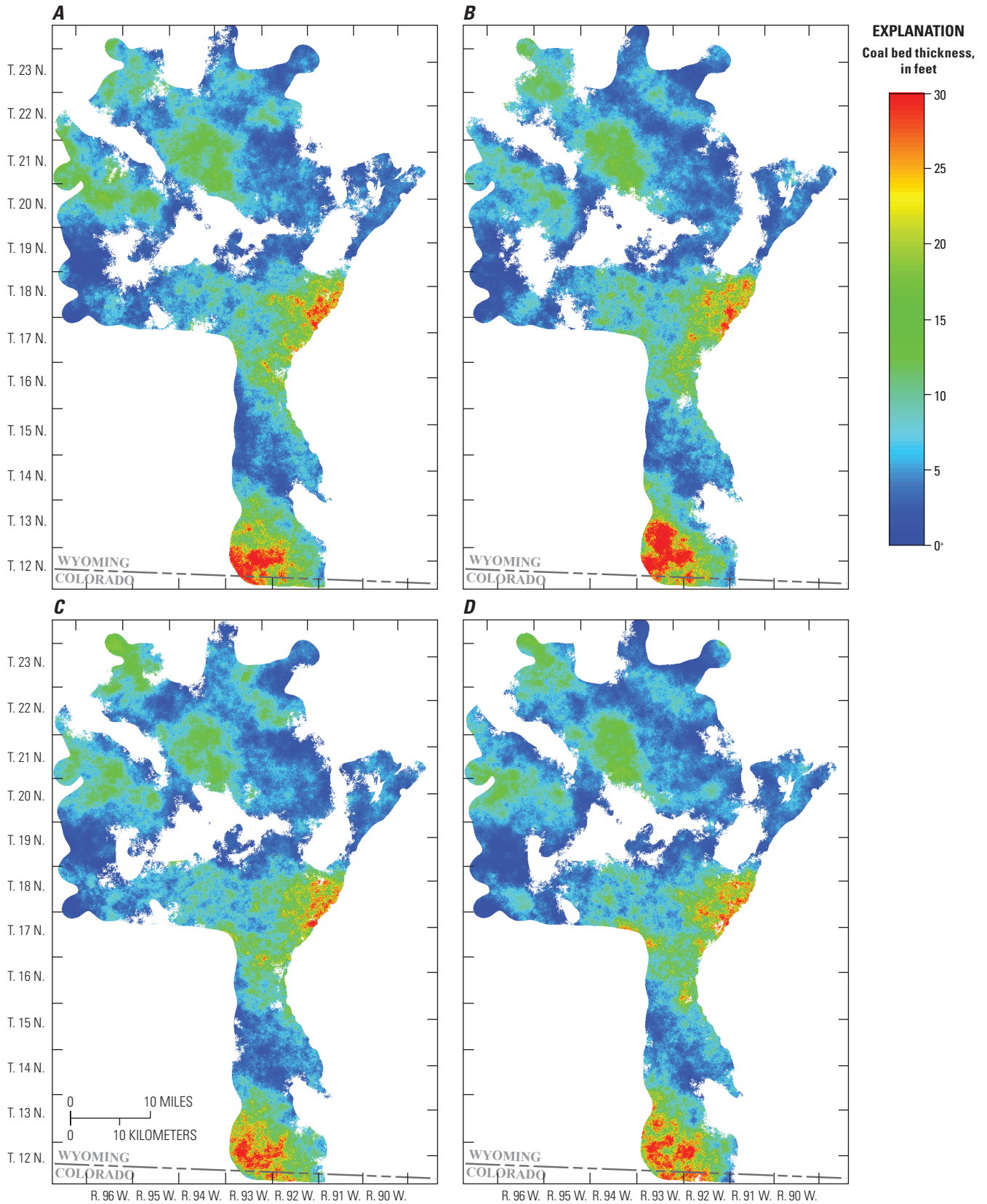
Figure 26. Maps showing coal bed thickness for the same realizations shown in figure 25. 0+ denotes a value close to but larger than zero.

E. First Step in the Modeling of the Coal Bed Thickness

One hundred thickness maps were generated within each of the boundary maps. For conciseness, only the first four are displayed.

F. Second Step in the Modeling of the Coal Bed Thickness

Figure 27 is a partial rendition of the post-processed thickness realizations after using a critical distance between 4 and 5 km for pinching out the realizations in figure 26 north of T. 18 N. between R. 90 W. and R. 96 W., where it is postulated that the Fillmore Ranch coal bed thins out gradually. To the west of the map area there are no drill holes for properly characterizing the coal bed. To the east, as shown in figure 18, the coal bed is truncated at an outcrop.

Part 3. Tonnage in Place

G. Coal Density

Conversion of thickness maps into tonnage maps requires considering the volume of prisms, with the height given by the thickness and the base given by the cell area, which is 100 m by 100 m in our case. Density is also necessary to complete the conversion. Despite its importance, density is not always measured, in which case it is necessary to apply some conversion factor, a practice that is discouraged. The accuracy of such a conversion using a constant is misleading because when density measurements are taken, it is observed that the values fluctuate. There are 16 values for this assessment obtained from a correlation with ash content (table 4). Although the sample size is small (figs. 28 and 29), modeling them is better than using an approximate constant. These values are also used because they help illustrate the methodology more generally. The semivariogram necessary for the modeling of density is shown in figure 30. The results using sequential Gaussian simulation can be observed in figure 31.
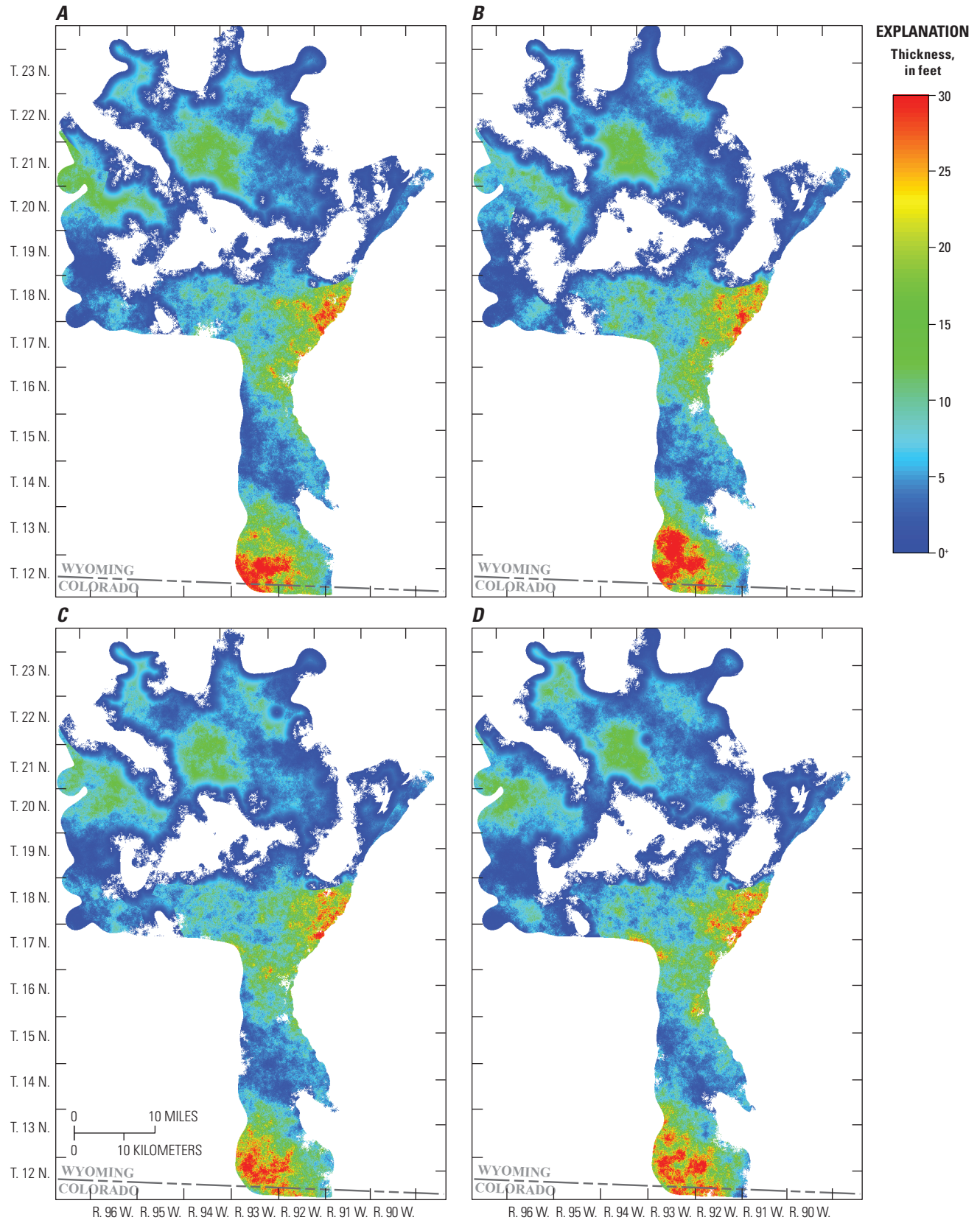
Figure 27. Maps from figure 26 after applying the pinchout correction. 0+ denotes a value close to but larger than zero.

Table 4.   Estimated values of coal specific gravity used for the calculation of coal densities.

[Coordinates are based on the Universal Transverse Mercator (UTM) grid system, zone 13 north, North American Datum (NAD) 1983. m, meters; db, dry basis; %, in percent; Est. spec. gravity, estimated specific gravity]

| Hole name	Coordinates		Calculations	
	Easting, m	Northing, m	Ash-raw (db), %	Est. spec. gravity
17-091-7-1AS	275515.00	4594275.00	7.42	1.324
RR92 (EW20)	282006.18	4606109.23	8.38	1.334
RR105 (EW30)	282099.03	4605572.08	11.27	1.363
CB 15 (AM614)	281891.64	4603889.27	8.17	1.332
CB 18 (2AS)	281365.25	4603213.31	8.16	1.332
CB 22 (AM612)	281270.77	4602921.15	7.43	1.324
CB 40 (AM602)	280733.01	4601321.29	9.89	1.349
CB 41 (3AS)	280294.47	4600950.26	10.03	1.350
CB 42 (AM603)	280616.51	4600867.11	6.29	1.313
CB 55 (AM651)	279396.94	4599761.95	7.18	1.322
CB 57 (3AS)	279455.50	4599696.53	6.61	1.316
CB 58 (AM655)	278944.13	4599290.49	6.27	1.313
CB 72 (2AS)	277834.36	4598391.76	7.12	1.321
CB 96 (3AS)	276212.83	4595383.03	6.22	1.312
CB 97 (AM660)	276355.21	4595457.87	7.06	1.321
17-092-13-1AS	274375.00	4592800.00	7.26	1.323
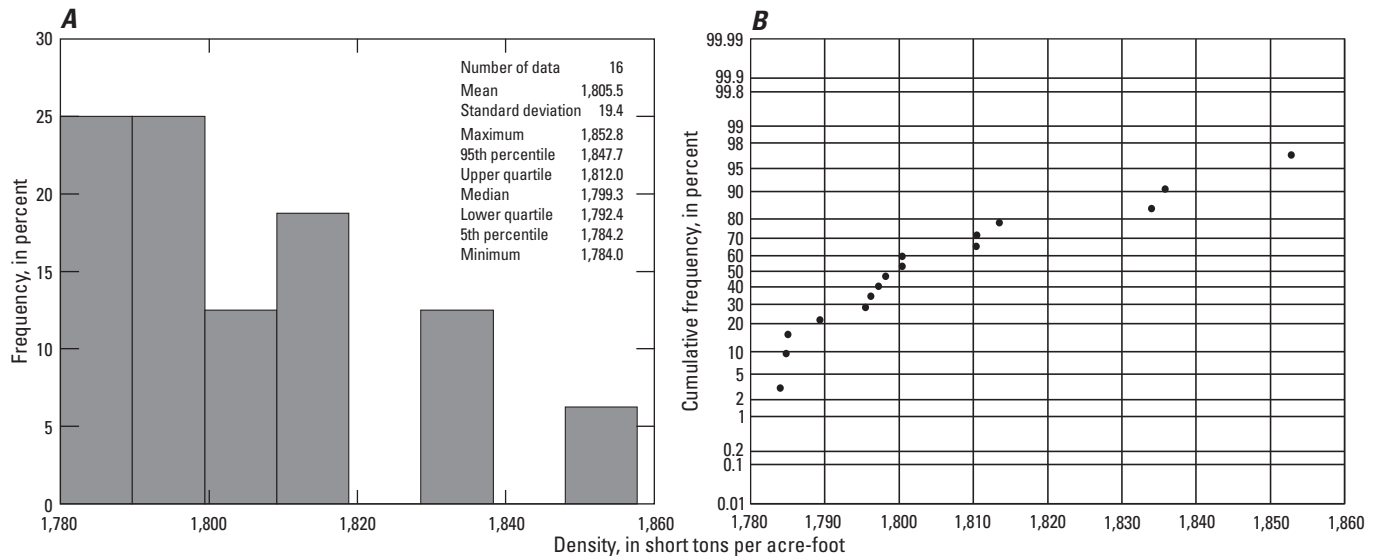
Figure 28.   Coal density data, which are listed in table 4 . *A*, Histogram. *B*, Cumulative distribution.
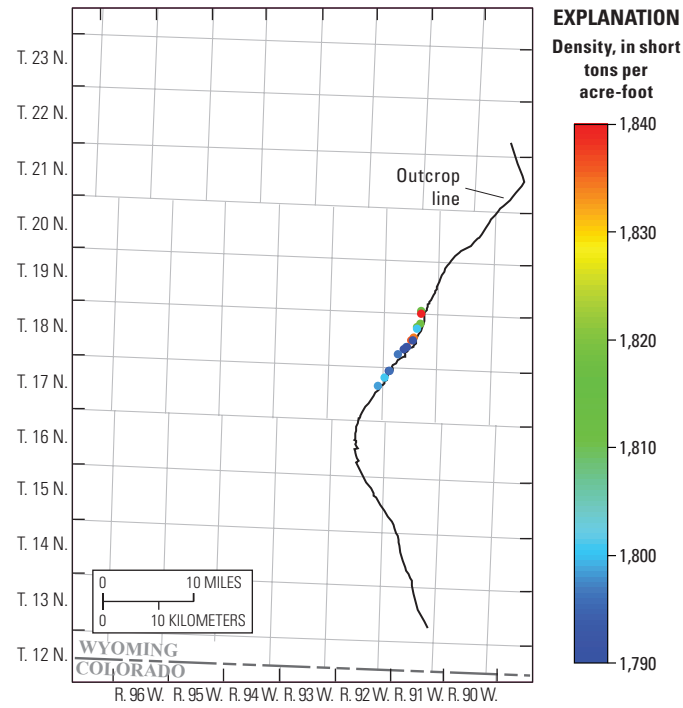
Figure 29.   Map showing location of coal density measurements.
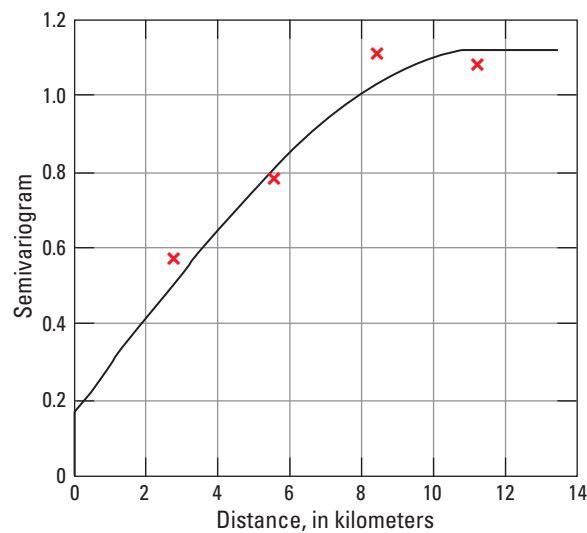
Figure 30.   Semivariogram for the normal scores of coal density. The red "x" symbols denote estimated values, and the line is a spherical model with a nugget of 0.17, a sill of 1.11, and a range of 11.3 kilometers.
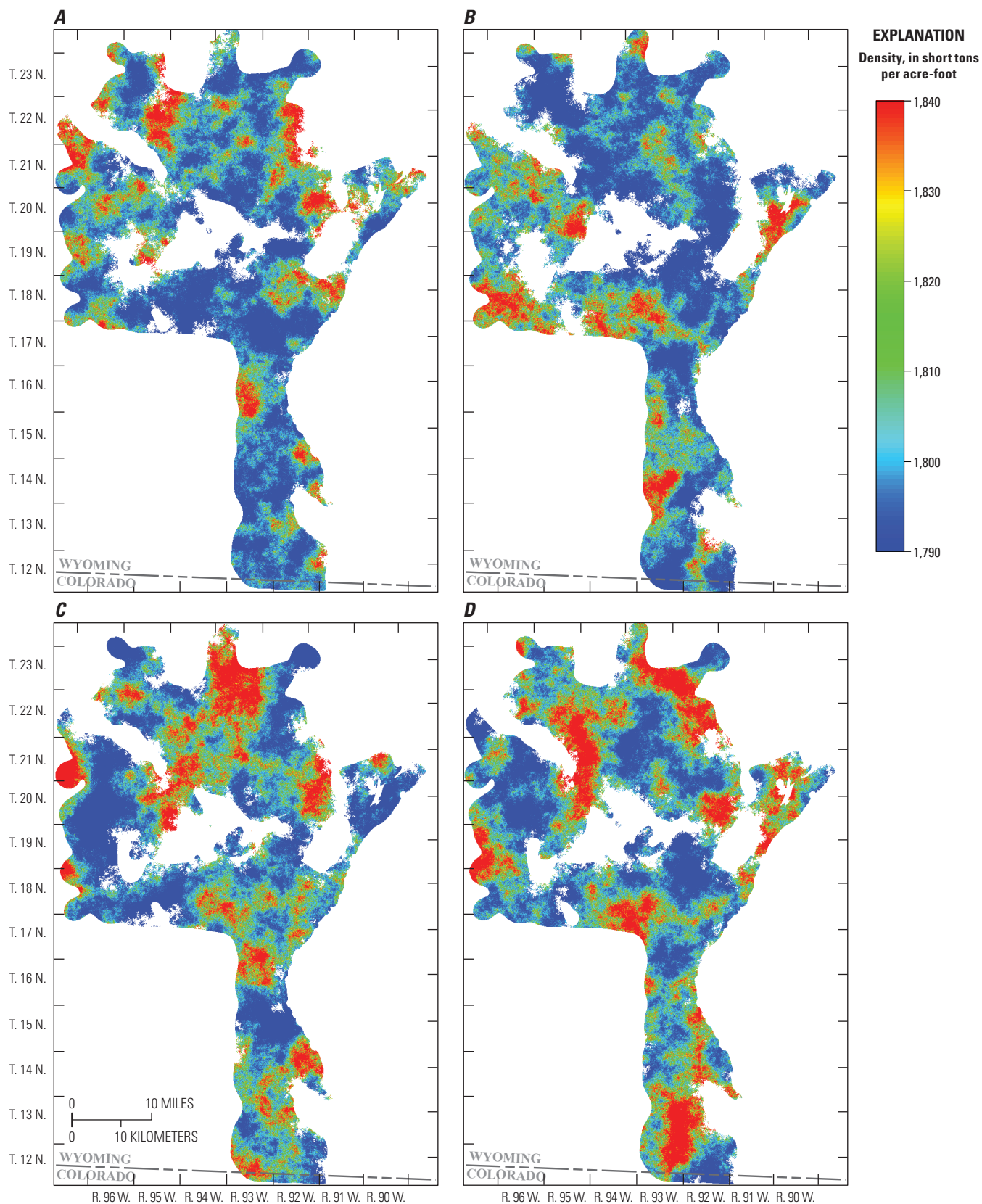
Figure 31. Maps displaying the first four realizations of coal density out of a total of 100 realizations.

H. Tonnage Maps

The maps in figure 32*A–C* are a partial display of the tonnage maps that result from the node-by-node multiplication of thickness and density realizations. It is assumed that the two types of realizations are not correlated, which justifies pairing the realizations at random. Applying the computer program *EtypeSGeMSsim* to all tonnage realizations, it is possible to obtain several statistics node by node. The map in figure 32*D* is an E-type map of tonnage.

I. Magnitude and Uncertainty of Total Original Resources

It is always important to assess the total tonnage of coal in place and its uncertainty. Once the tonnage realizations are obtained, providing answers is quite straightforward. Each tonnage realization is an equally likely map of the actual tonnage and is compatible with all that is known about the coal bed. The sum of the tonnage, cell by cell, is then one possible value for the total tonnage in place within the area of study. There are 100 values of total tonnage, one for each tonnage realization. The collection of these 100 values (fig. 33) can be interpreted as a numerical approximation of the random variable modeling the magnitude of the original resources. From the properties of random variables and confidence intervals, by considering the 5th and the 95th percentiles, one possible textual form of summarizing the results is to report that there is 90-percent probability that the tonnage in place is 9.687 ± 0.383 billion short tons (bst), or equivalently, 9.687 bst ± 4 percent. The value 9.687 is the average of the 5th and the 95th percentiles, and 0.383 is half the difference between the same values.

Using a deterministic methodology implemented applying the computer program *PC/Cores*, Scott and others (2019) reported original coal resources of 9.667 bst for the same coal bed, which lands in the central part of our probability distribution. This general agreement between the two types of assessments, also observed in other coal assessments, implies that the adoption of the methodology in this report should not bring significant differences in total tonnages in place relative to coal deposits assessed in the past. Quantification of the uncertainty is the main difference and advantage of the methodology in this report.

Figure 34 shows the results of a sensitivity analysis to verify the adequacy of having selected 100 realizations in the modeling to obtain the responses of interest. Both the coal tonnage in place and its standard deviation stabilize at about 50 realizations. For a higher number of realizations, the fluctuations in tonnage are below 0.1 percent and are no more than 1 percent for the standard deviation. Consequently, the use of 100 realizations is a safe choice in this case.

J. Magnitude and Uncertainty in the Cell Tonnage

Understanding the fluctuation of tonnage from cell to cell is as critical in mining as knowledge about the total resources. The same realizations used in the assessment in the previous section can be used for analyzing the resources at the cell level, bringing internal consistency between the evaluations both at the global and the local scales. Some summaries are in order when wanting to avoid examining all random variables cell by cell. As for the cell tonnage magnitude, the most adequate summary can be taken from figure 32, either the realization with the median total tonnage (fig. 32*B*) or the map resulting from averaging all realizations (fig. 32*D*).

The reliability of each value of cell tonnage can be taken from any descriptor characterizing the dispersion of the values at each node. In this methodology, we have selected the spread between the 5th and the 95th percentiles, which, by definition, provides the magnitude of the interval containing the true tonnage with 90-percent probability. These values were already obtained when running the program *EtypeSGeMSsim* to prepare the map in figure 32*D*. Figure 35 is a map of the 5–95 spread for the Fillmore Ranch coal bed in the study area and figure 36*A* is a histogram of the same values.

For certain purposes, the results may be easier to manipulate and interpret if summarized as a cumulative distribution (fig. 36*B*), supplemented with the cumulative mean cell tonnage (fig. 32*D*). For example, if the cells are divided in three groups, each having the same number of cells, figure 36*B* indicates that the 5–95 spread for the group having the least uncertainty is always below 18.0 thousand short tons (kst) and contains 0.915 bst of the resources in place. The group having the intermediate uncertainty has a 5–95 spread between 18.0 and 33.5 kst and contributes 2.51 bst (3.425 – 0.915). The group having the highest uncertainty has the most tonnage; the 5–95 spread fluctuates between 33.5 and 118.8 kst and the tonnage in this group is 6.287 bst (9.712 – 3.425). Note that, by construction, there is no overlapping in the uncertainty values associated with the tonnage classes.

K. Comparison to Circular 891

Traditionally, the prevailing form of displaying reliability cell by cell has been the mapping of the Circular 891 classes. Figure 37 shows such a map for our study area of the Fillmore Ranch coal bed. The map also discloses the tonnage associated with each class. Circular 891 is method-free in terms of estimating cell tonnages. The tonnages reported here are the ones resulting from the geostatistical methodology summarized by the map in figure 32*D*.

Figure 36*B* can also be used for discretizing the map in figure 35 after ignoring the color lines applicable to a previous example. While the map in figure 35 is only one, different numbers of classes and dividers can result in

multiple categorical maps. Selection of classes and dividers, if any, should be made on the basis of the characteristics of each deposit. It is not the intent of this report to perpetuate the four classes in Circular 891, nor the idea to replace the fixed dividers of ¼, ¾, and 3 miles by a new set of 5–95 spreads. Figure 38 was generated selecting the same tonnages per class used in figure 37 for facilitating comparisons. In this case, per figure 36B, the dividers in terms of 5–95 spread are 18.917 kst for 1.031 bst and 46.788 kst for 5.354 bst (1.031 + 4.323). Figures 38 and 39A dispel any notion that 5–95 spread reliability is equivalent to the distance to the closest drill hole.

Geostatistical uncertainty, in the presence of systematic drilling as is the case here, is primarily sensitive to the complexity of the geology. Correlation of the E-type value of attributes with reliability of the modeling is so common that it has been given a name: proportional effect (Manchuk and others, 2009; Emery, 2012). For the Fillmore Ranch coal bed, this correlation is 0.89 (fig. 39B).

Above all, differently from the distance classes, when validation is possible, modeling of uncertainty by geostatistics has provided results in excellent agreement with measured values (Olea, 2018).
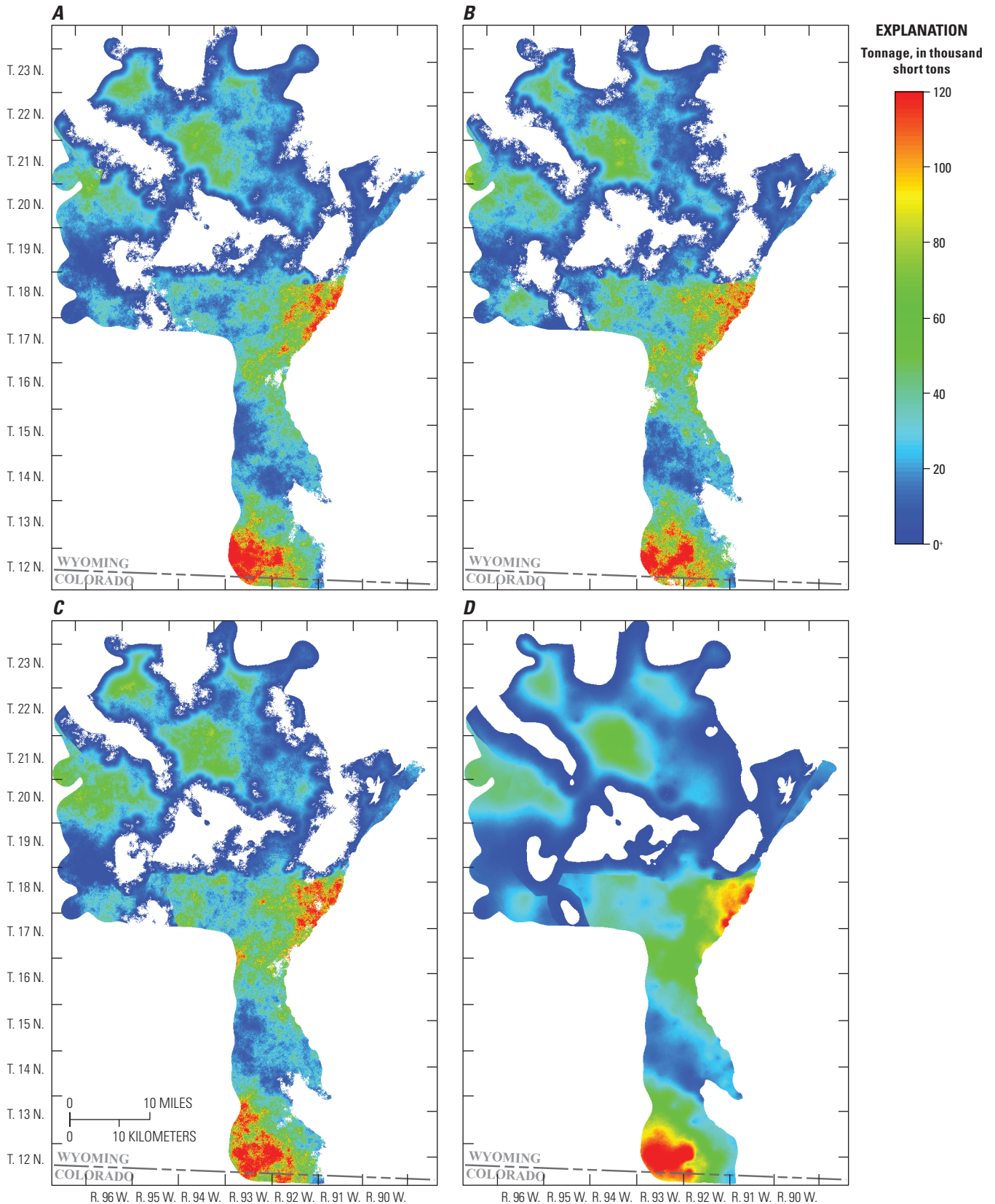
Figure 32. Maps showing likely tonnage for the Fillmore Ranch coal bed, selected to illustrate the range in variability (*A–C*) and an average map (*D*). *A*, Realization 42 with the fifth lowest tonnage (9.304 billion short tons [bst]). *B*, Realization 45 with the median tonnage (9.724 bst). *C*, Realization 33 with the fifth highest tonnage (9.910 bst). *D*, E-type map showing averaging of all 100 realizations. 0+ denotes a value close to but larger than zero.
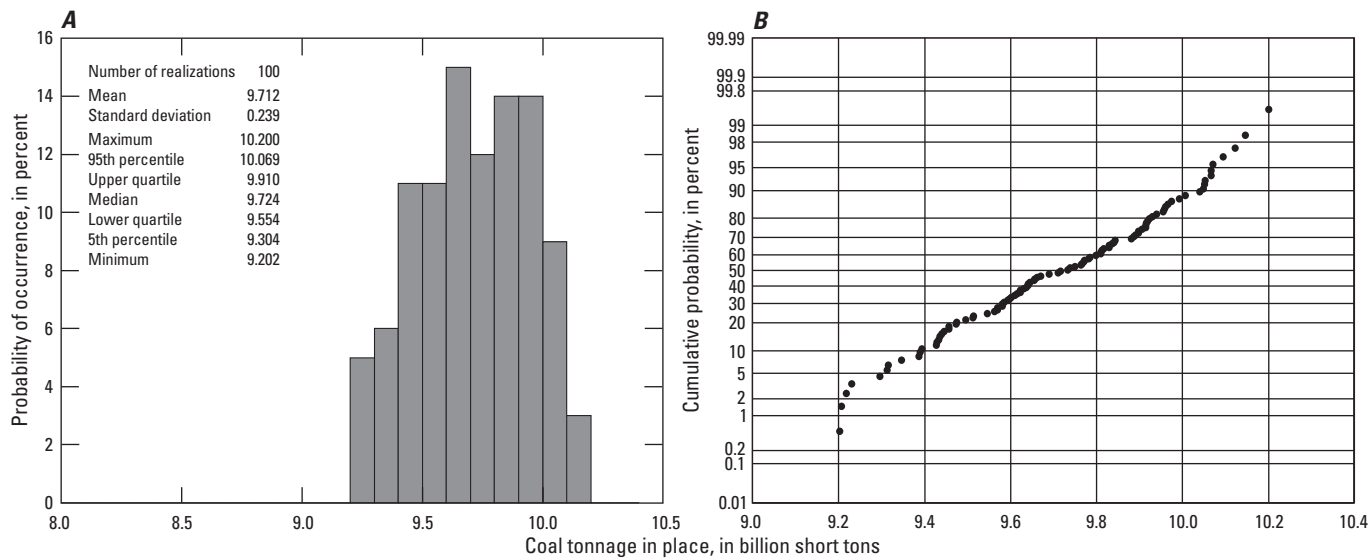
Figure 33.   Probabilistic assessment of total coal tonnage in place. *A*, Histogram. *B*, Cumulative distribution.
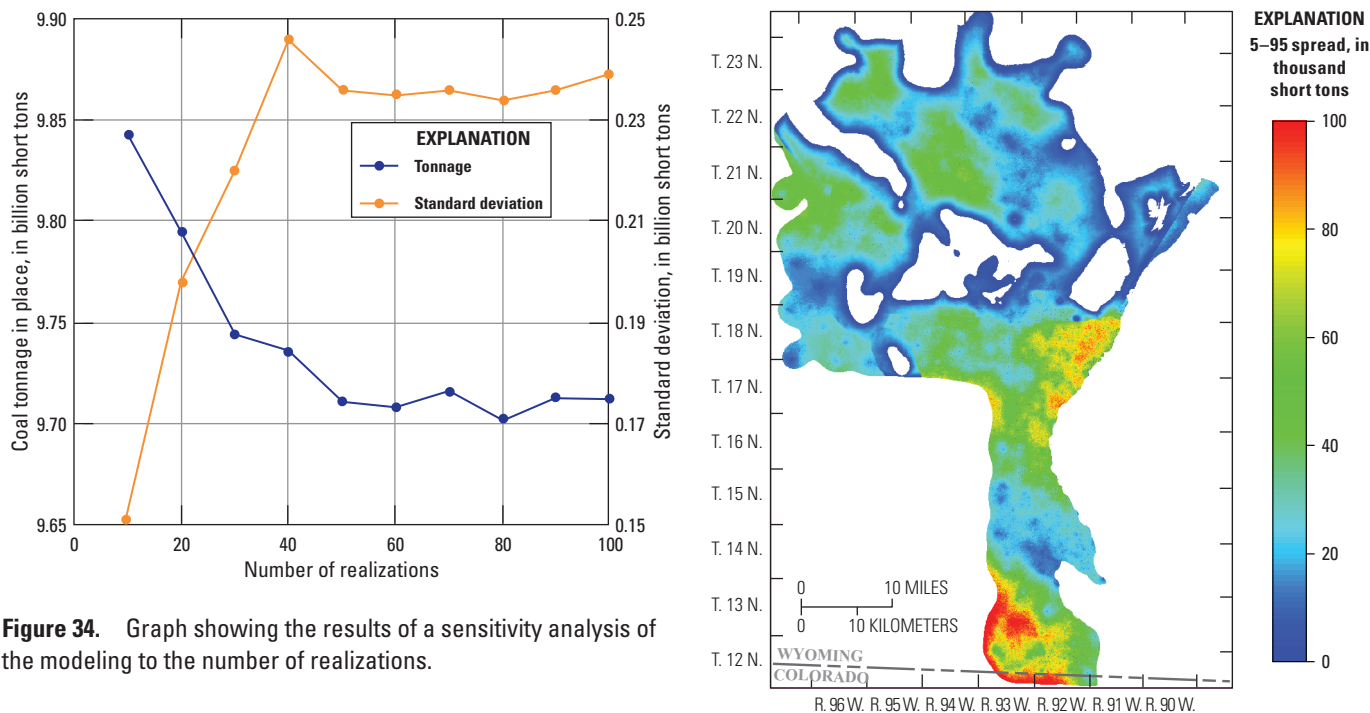
Figure 34.   Graph showing the results of a sensitivity analysis of the modeling to the number of realizations.


Figure 35.   Map of the 5–95 spread in cell tonnage for the Fillmore Ranch coal bed.
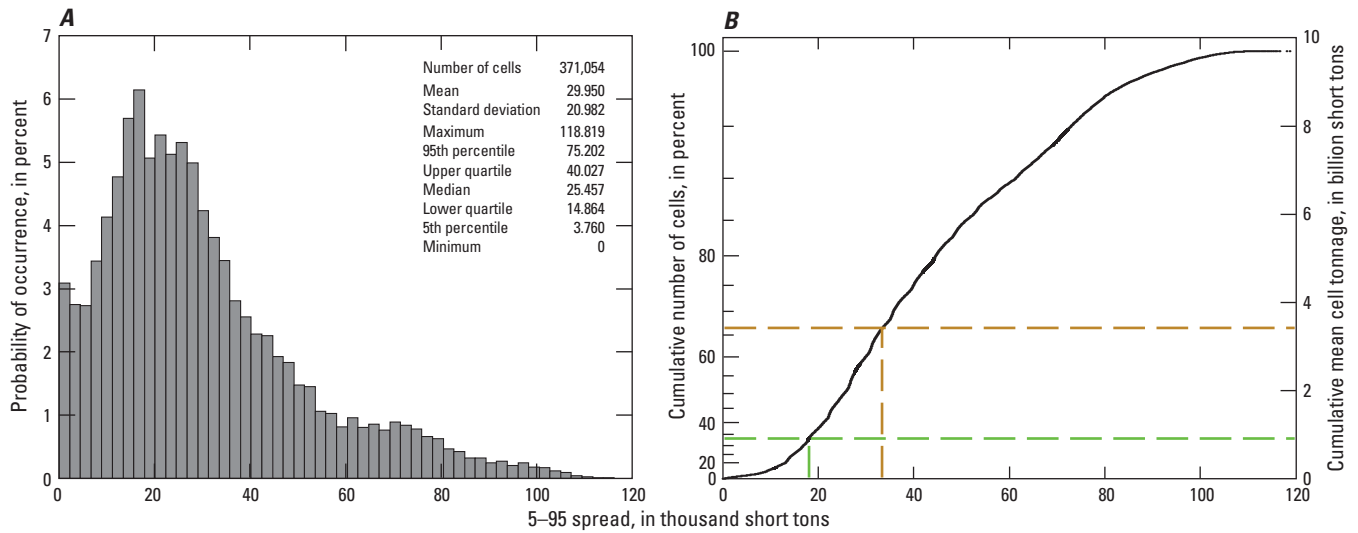
Figure 36. Graphs showing uncertainty in cell tonnage. *A*, Histogram of the values in figure 35. *B*, Cumulative distribution of the same values in terms of cumulative cell frequency (left axis) and cumulative cell tonnage taken from figure 32*D* (right axis), with the green dashed line showing the third of the cells having the lowest tonnage uncertainty and the orange dashed line showing the third of the cells having the highest tonnage uncertainty.
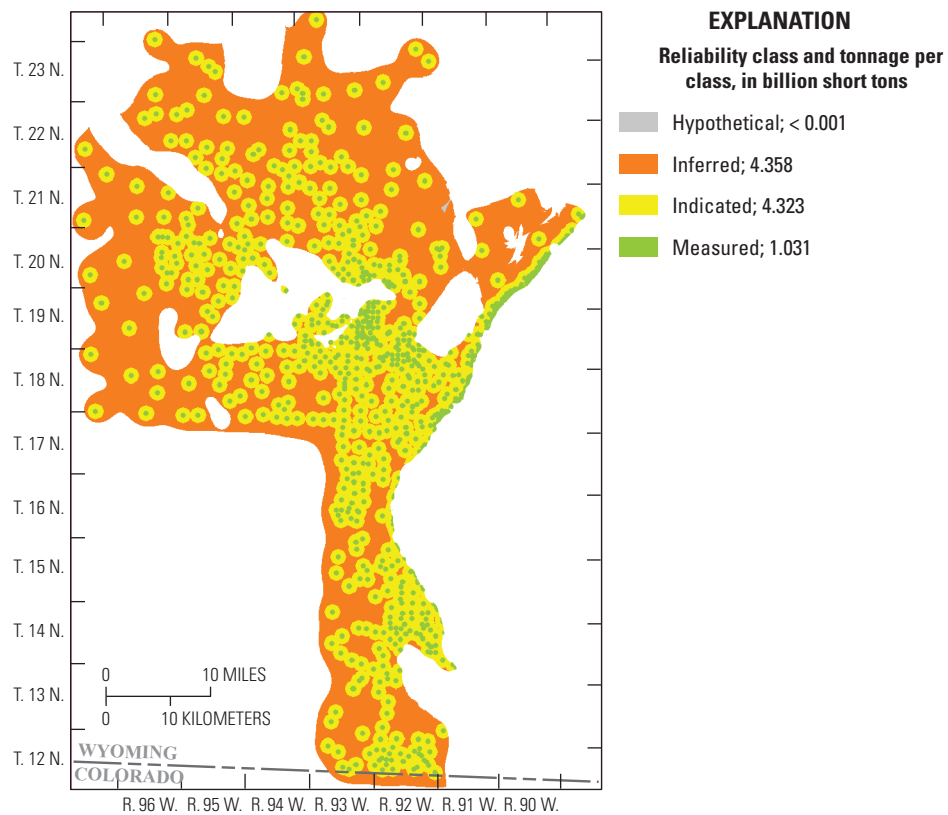
Figure 37. Map showing Circular 891 reliability classes and associated cell tonnage according to the map in figure 32*D*.
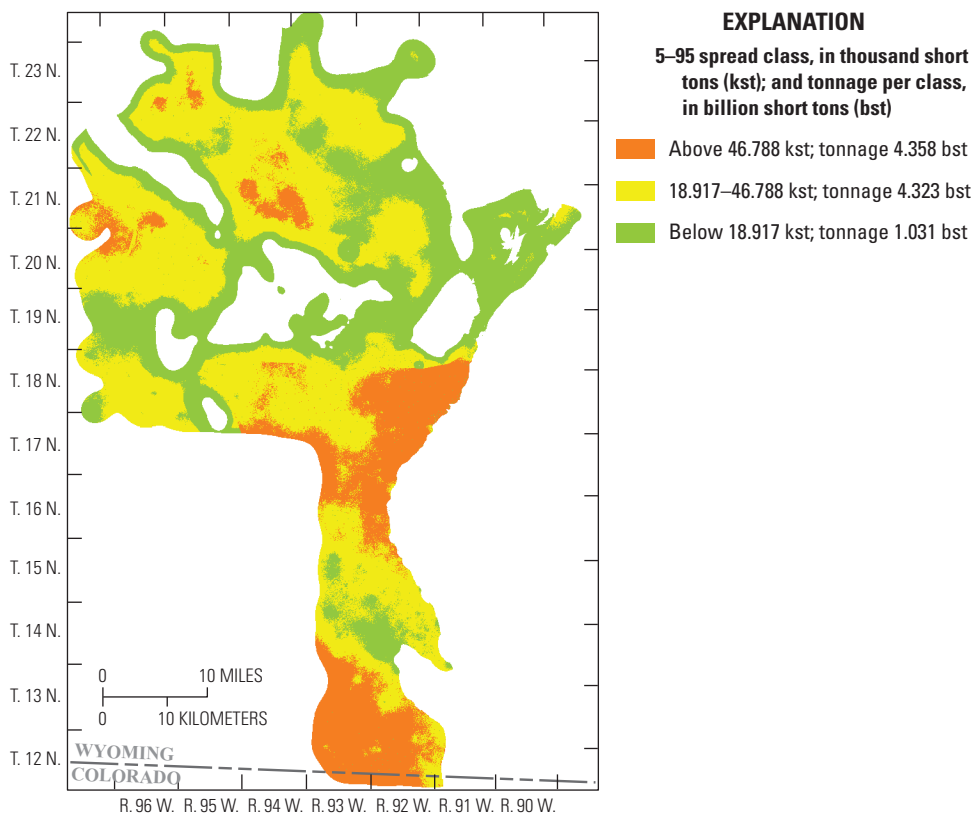
Figure 38. Categorical version of the map in figure 35 showing tonnage per class, the same as that shown in figure 37. The larger the 5–95 spread, the higher the uncertainty.
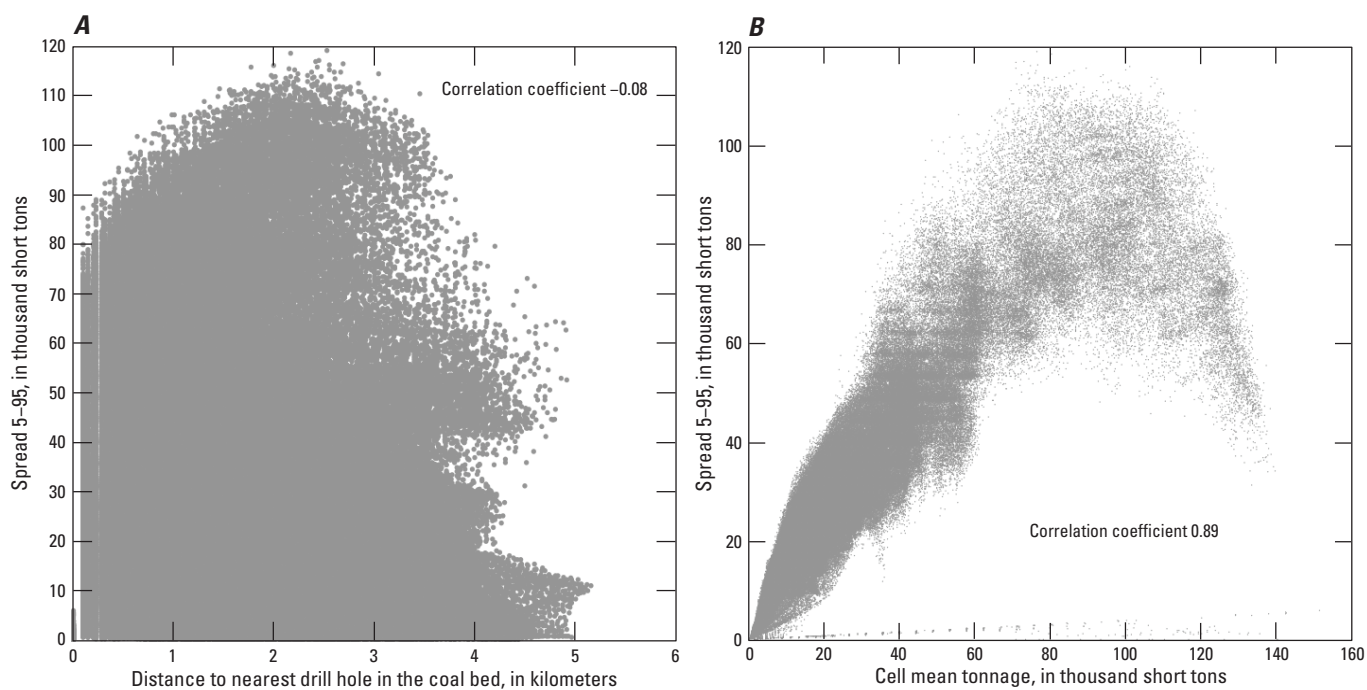

Figure 39. Scatterplots of 5–95 spread. *A,* Versus the distance to the nearest drill hole according to figure 37. *B,* Versus the cell mean tonnage shown in figure 32*D.*

Part 4. Mineable Tonnage

Two conventional methods of coal extraction are considered here for the mining of the Fillmore Ranch coal bed: surface mining and underground mining. The mining method to apply depends on the coal bed thicknesses and the thickness of the cover. Cover is the overburden—the material between the ground surface and the top of the coal bed. Therefore, modeling the cover requires modeling these two surfaces.

L. Ground Surface Elevation

Ground surface elevation can be taken from digital elevation maps, available today at a finer resolution than the cell size of 100 m used in this study. Therefore, no estimation is necessary. For all practical purposes, the map in figure 40 is error free.
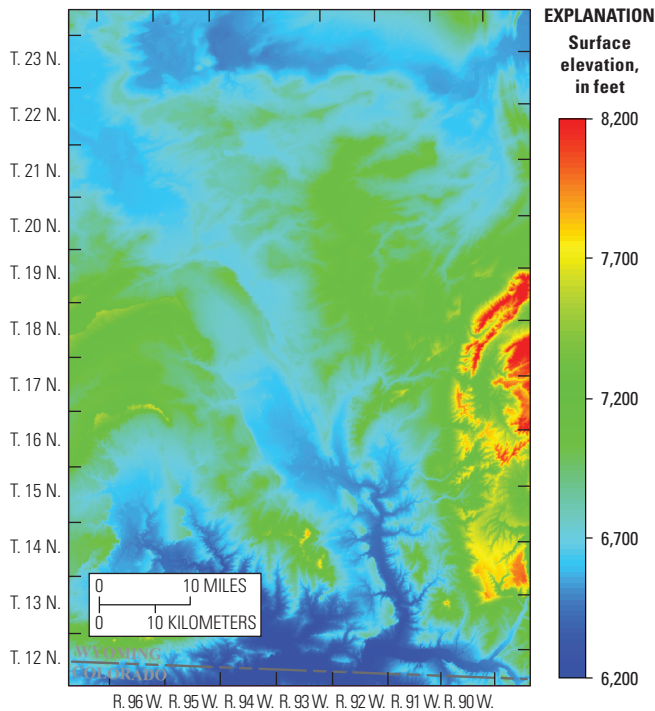
**Figure 40.**   Ground surface elevation map of the study area.

M. Elevation of the Top of the Coal Bed

The top elevation is known only at the same locations of the drill holes providing thickness values (fig. 18*B*). Figure 41 is a posting of the elevation values and figure 42 displays their frequency distribution.
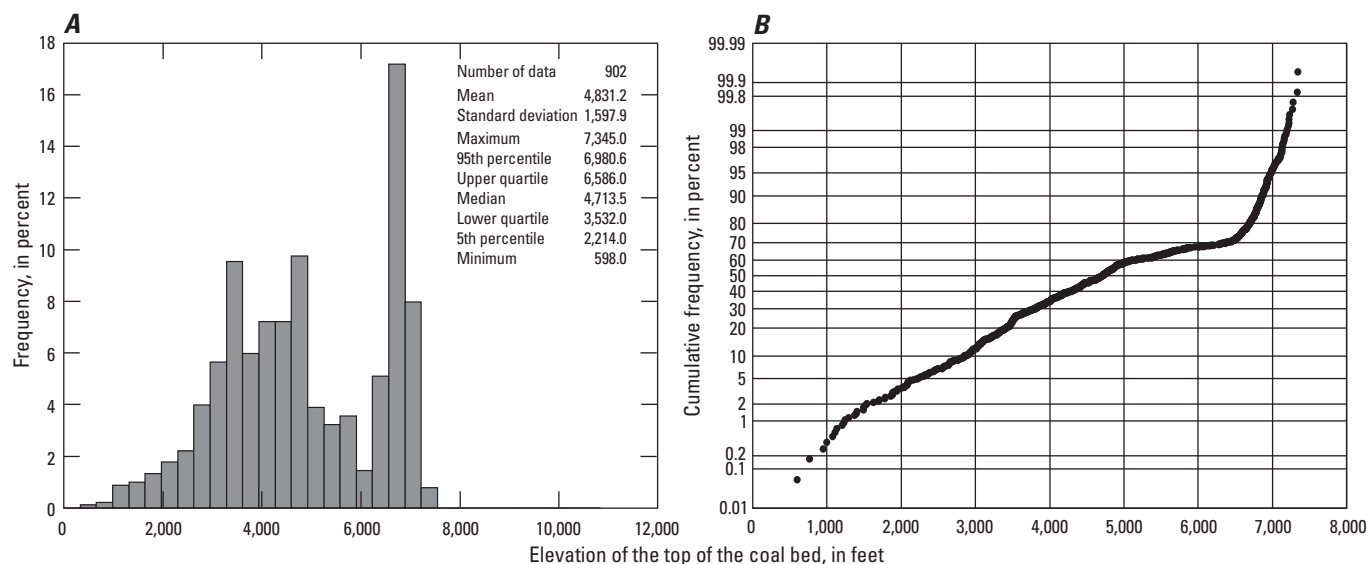
The elevation data for the top of the coal bed have the same preferential sampling problems as the thickness data. Figure 43 displays the semivariogram for the normal scores of the elevation of the coal bed top after applying the same

declustering procedure applied to thickness. In this case, in addition to the preferential sampling, there is a clear trend to have lower elevations away from the outcrop.
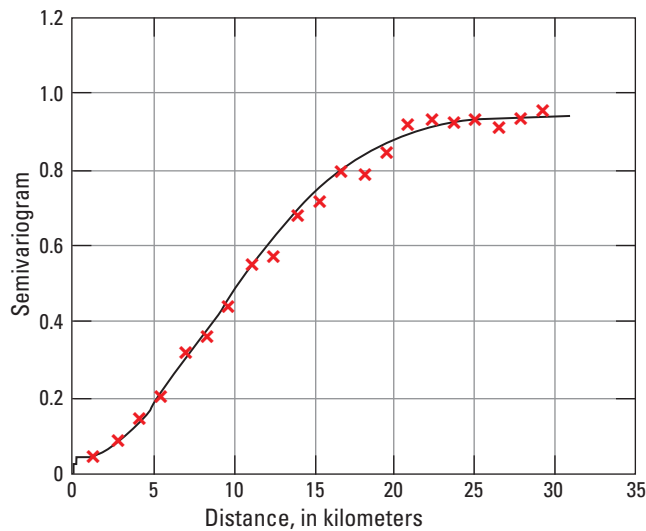
Figure 44 displays the first four realizations for the elevation of the top of the Fillmore Ranch coal bed out of a total of 100 realizations.
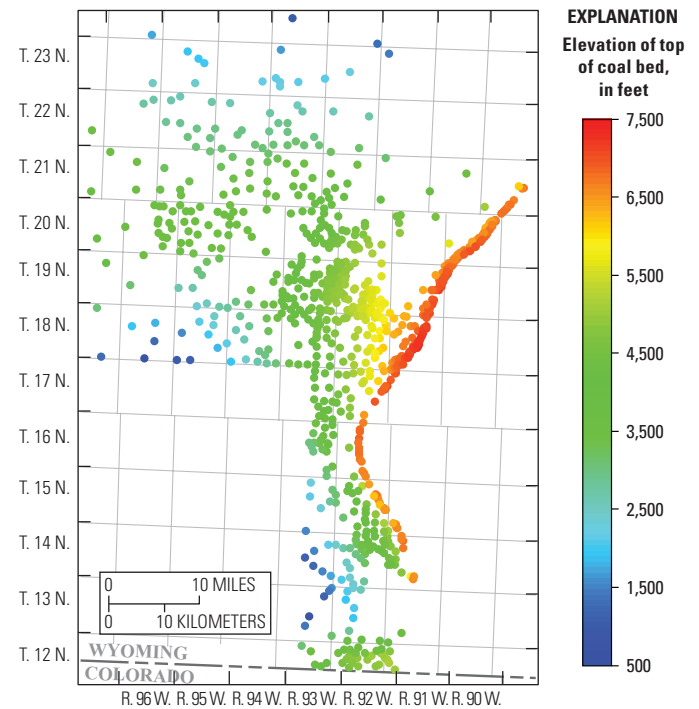

Figure 41.   Map showing posting of data for the elevation of the top of the Fillmore Ranch coal bed.

A

Number of data	902
Mean	4,831.2
Standard deviation	1,597.9
Maximum	7,345.0
95th percentile	6,980.6
Upper quartile	6,586.0
Median	4,713.5
Lower quartile	3,532.0
5th percentile	2,214.0
Minimum	598.0

B

Elevation of the top of the coal bed, in feet

Figure 42.   Graphs showing elevation of the top of the Fillmore Ranch coal bed. *A*, Histogram. *B*, Cumulative distribution.

Figure 43.   Semivariogram of the declustered normal scores for the elevation of the top of the Fillmore Ranch coal bed in the N. 58° E. direction without trend. The red "x" symbols denote calculated values, and the continuous curve is a fitted Gaussian model with a nugget of 0.04, a sill of 0.9, and a range of 21.2 kilometers.
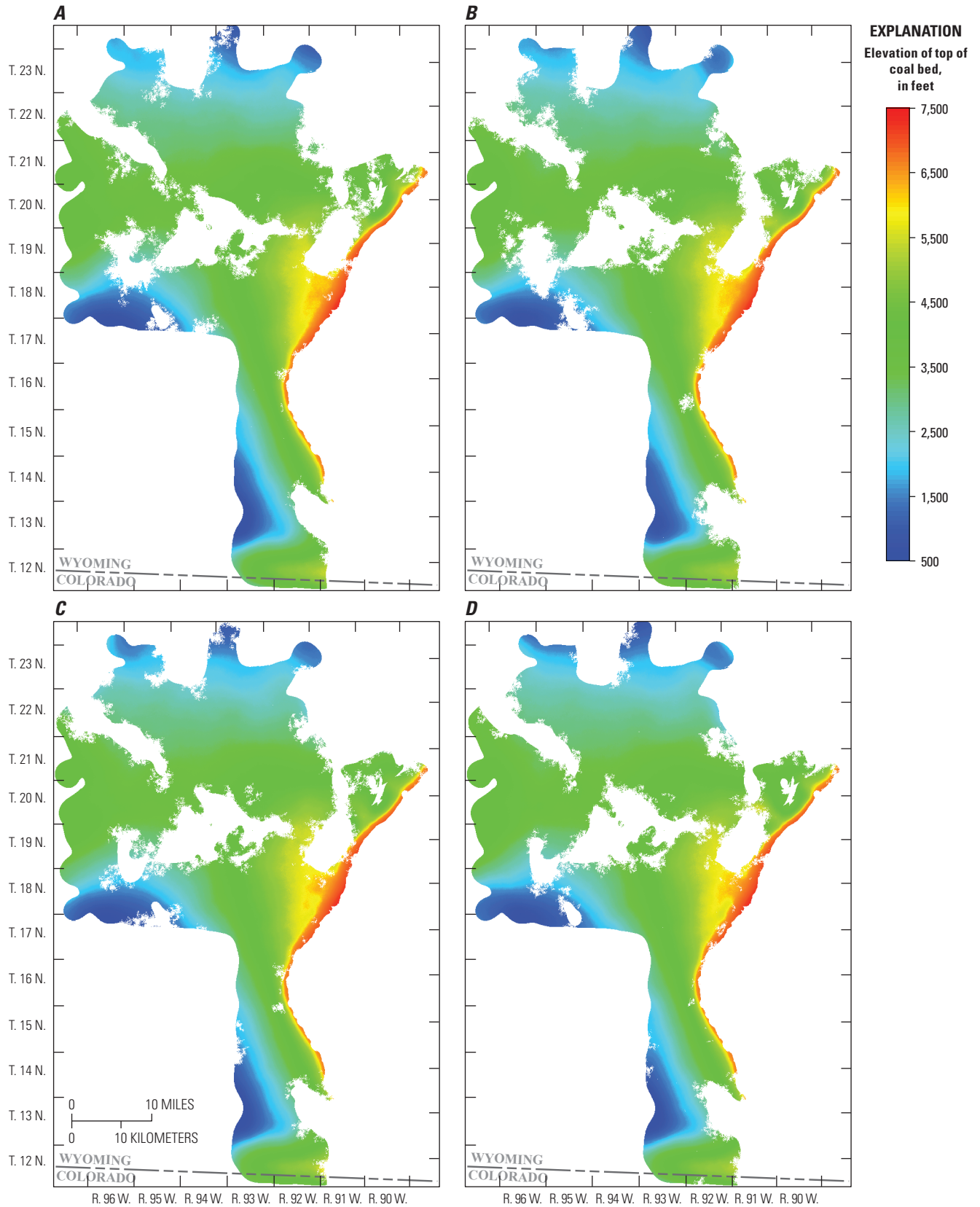
Figure 44. Maps displaying the first four realizations for the elevation of the top of the Fillmore Ranch coal bed.

N. Cover

Figure 45 shows the difference in elevation between the ground surface and the top of the coal bed, which is the thickness of the cover. Figure 46 displays the same information grouped into classes for the purpose of assessing the mineable resources. No measurements of oxidation depth were taken in the field for the Fillmore Ranch coal bed. A few geophysical logs were inspected to postulate an oxidation depth of 50 ft. The minimum extension of the oxidized zone can be used as a justification for not having generated realizations for the oxidized zone in this study.

O. Surface Mining

It is assumed here that a cell is worth recovering by surface mining methods when: (a) the coal bed thickness is greater than 3 ft; (b) the thickness of the cover is over 50 ft, which prevents the coal from being oxidized; and (c) the thickness of the cover does not exceed 300 ft. The selected results displayed in figure 47 were obtained by applying these criteria to the same 100 tonnage realizations partly displayed in figures 32*A*–*C*.

Figures 48–50 summarize the information that can be extracted from the 100 realizations partly displayed in figure 47. The resources are now a minor portion of the tonnage in place: there is a 90-percent probability that the true tonnage available for surface mining is 0.240 ± 0.025 bst, or equivalently, 0.240 bst ± 10 percent. Proportionally, there has been a significant decrease in reliability relative to the ± 4-percent dispersion for tonnage in place, which agrees with (a) the higher cell tonnage uncertainty, denoted by the higher average 5−95 spread of the retained cells (figs. 36*A* and 50*A*); and (b) the additional uncertainty caused by the modeling of the cover. The tonnage values reported here are as negligible as the tonnage of 0.213 bst reported by Scott and others (2019).

P. Underground Mining

Underground mining methods are the preferred choice when the coal bed thickness is greater than 8 ft and the cover thickness is between 300 and 3,000 ft. In addition, it is assumed that the extraction will be done by longwall mining methods, which currently can mine a maximum thickness of 15 ft. The methodology is flexible enough to allow changes in these parameters. Figure 51 has the results in terms of the cells available for underground mining.

The realizations in figure 27 show that the coal bed is thicker than 15 ft primarily in the areas possible to exploit by underground mining. The 15-ft maximum mining height, however, limits cell tonnage to about 67 kst, which significantly reduces the recoverable resources.

Figures 52–54 summarize the uncertainty about the resources available for extraction by underground mining methods. The resources are now intermediate in magnitude between the tonnage in place and the tonnage available for surface mining: 2.279 ± 0.160 bst, or equivalently, 2.279 bst ± 7 percent, with 90-percent probability. The underground mineable tonnage of 2.387 bst obtained by Scott and others (2019) lands in the central part of the distribution in figure 52*A*. However, Scott and others (2019) discarded 0.265 bst because of land-use restrictions.

The constraint of a maximum 15-ft mining height has the effect of reducing the 5–95 spread values in such a way that several cells with the highest tonnage spread (fig. 35) now have the lowest uncertainty (fig. 53), which is a perfectly logical inversion, not a mathematical artifact. Let us consider a cell that is a candidate for underground mining and that has a 5th percentile of 16 ft and a 95th percentile of 33 ft. For this cell, the unrestricted-thickness 5–95 spread would be 17 ft, or approximately a 75-kst tonnage spread, which is a high value. Considering the maximum mining height constraint for longwall mining methods, both percentiles are now 15 ft, hence their difference is zero, which is in perfect agreement with the fact that, with 90-percent probability, the 15-ft thickness, equivalent to about 67 kst, is now guaranteed. The effect over all cells is to have the lowest mean spread, lower even than those shown in figures 36*A* and 50*A*.
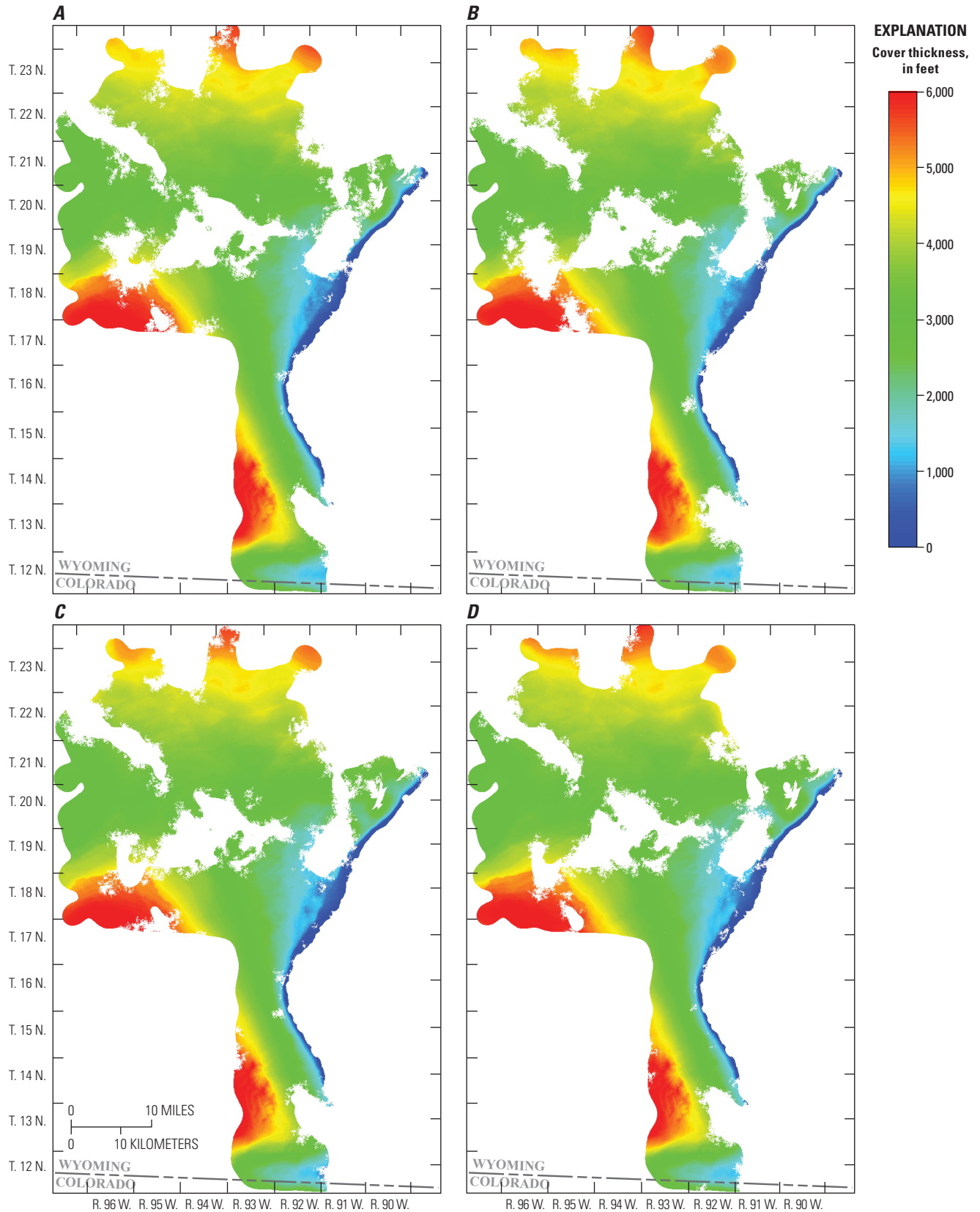
Figure 45.   Maps displaying the first four realizations for the cover thickness above the Fillmore Ranch coal bed.
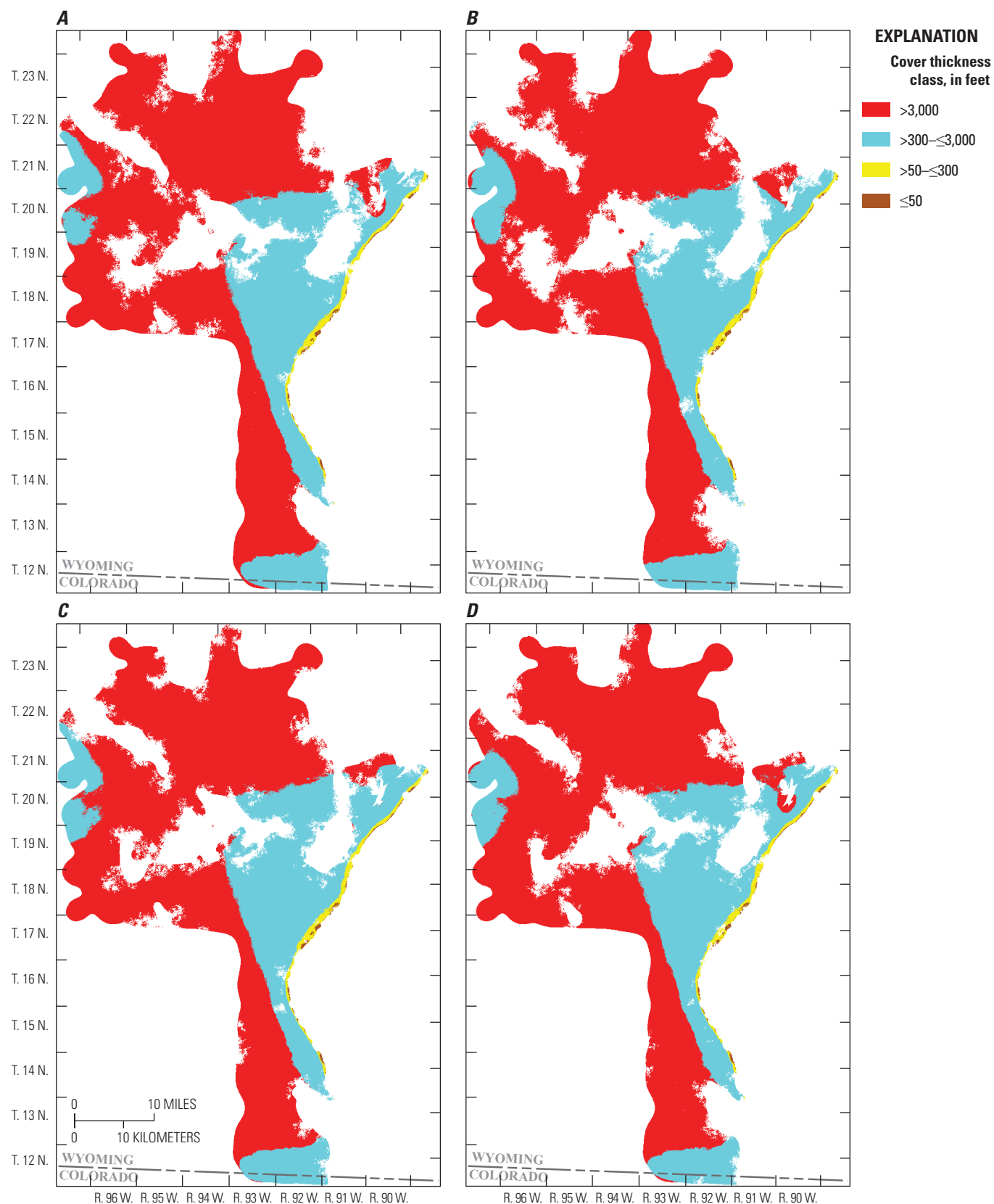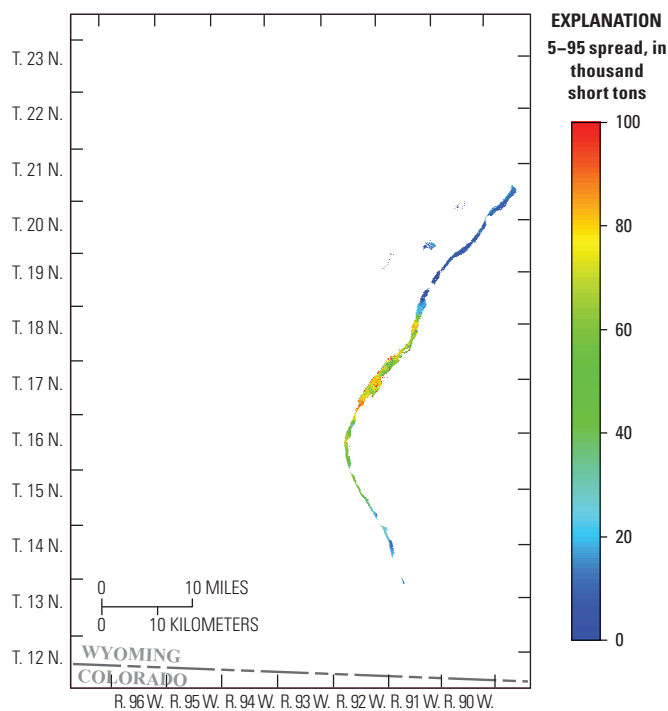
Figure 46.   Maps displaying the first four realizations of thickness classes for the cover above the Fillmore Ranch coal bed.
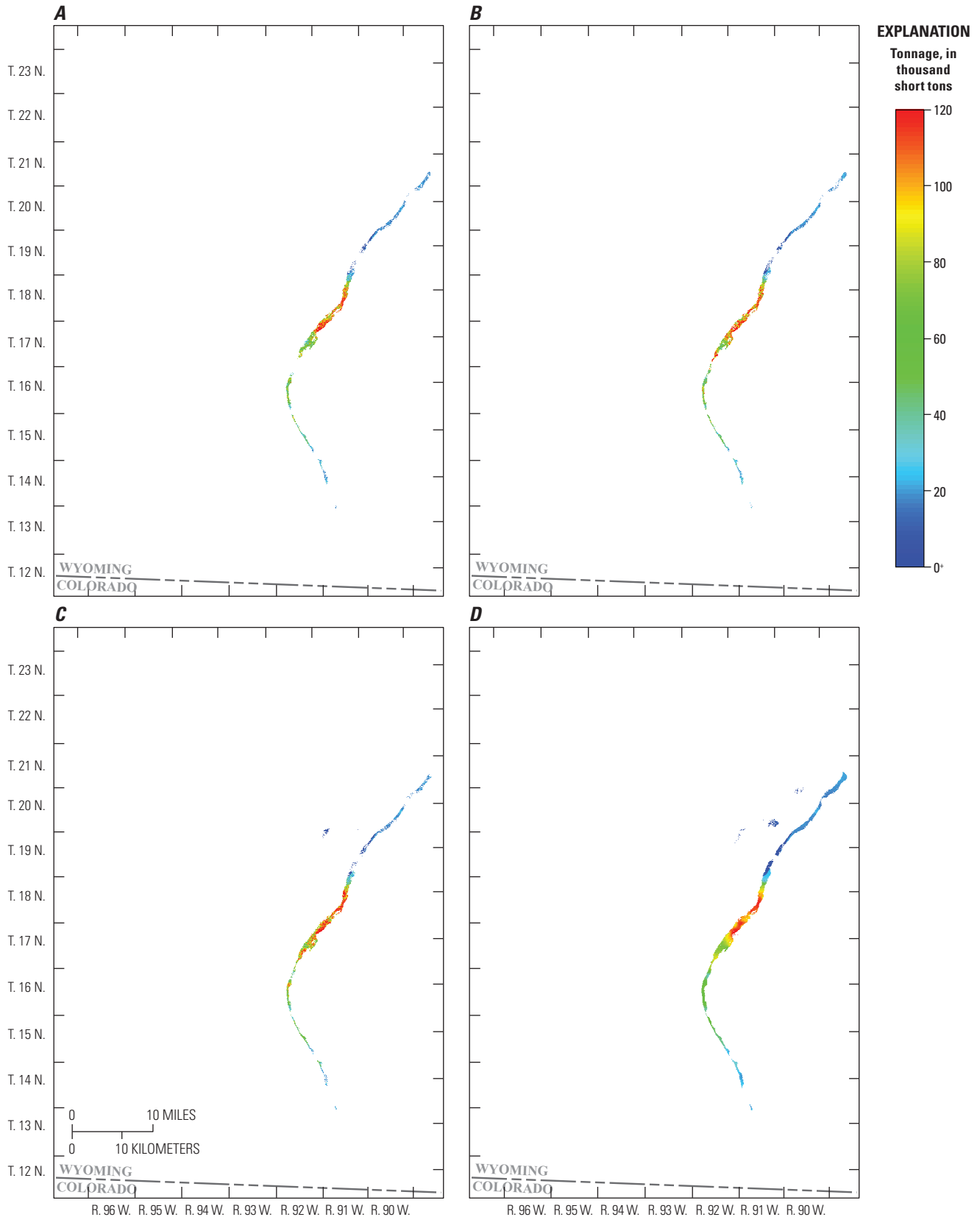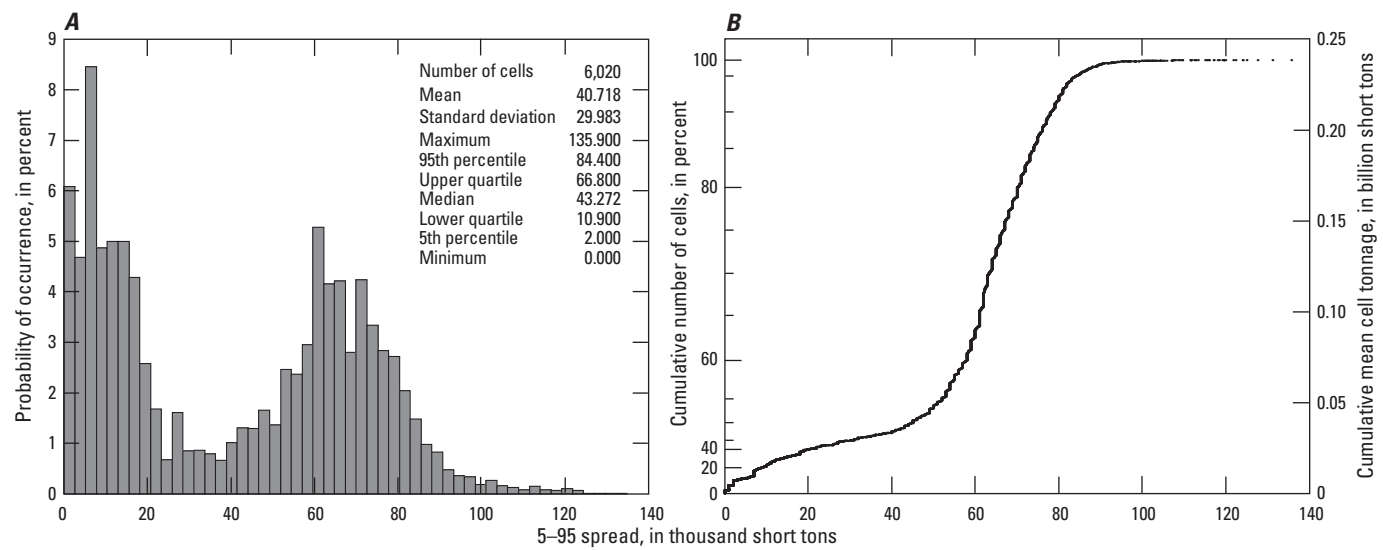
Figure 47.   Maps showing coal resources recoverable by surface mining, displayed as three realizations according to the magnitude of the tonnage (*A–C*) and as an average of all realizations (*D*). *A*, Realization 89 with the fifth lowest tonnage (0.223 billion short tons [bst]). *B*, Realization 93 with the median tonnage (0.239 bst). *C*, Realization 57 with the fifth highest tonnage (0.256 bst). *D*, E-type realization. 0+ denotes a value close to but larger than zero.
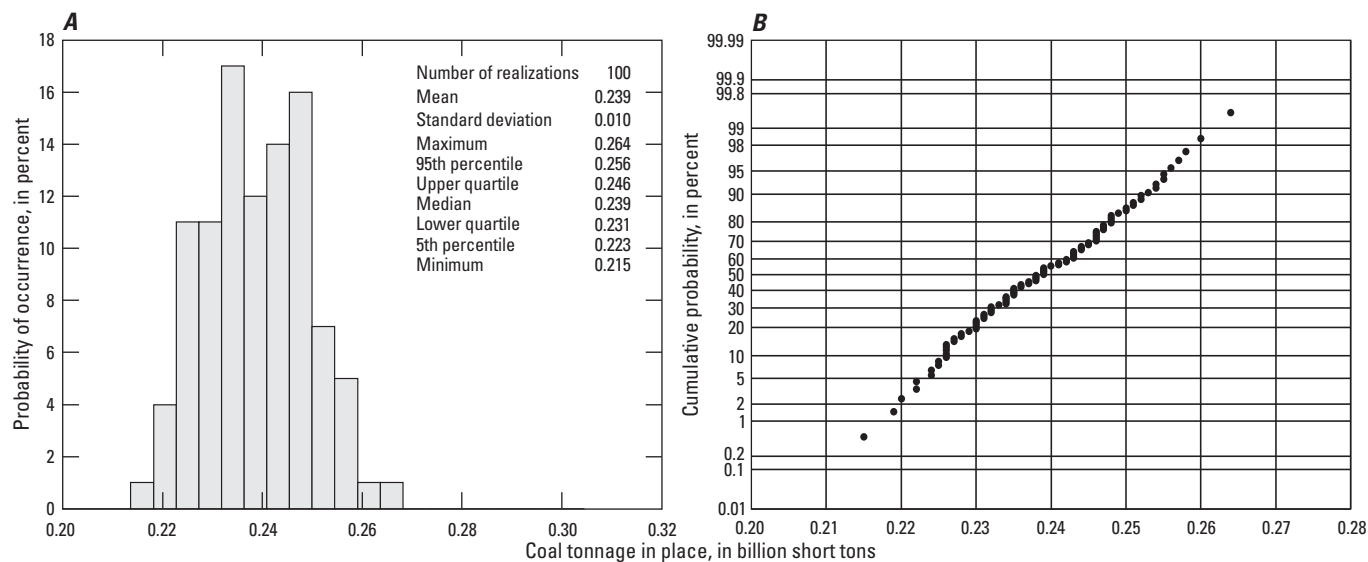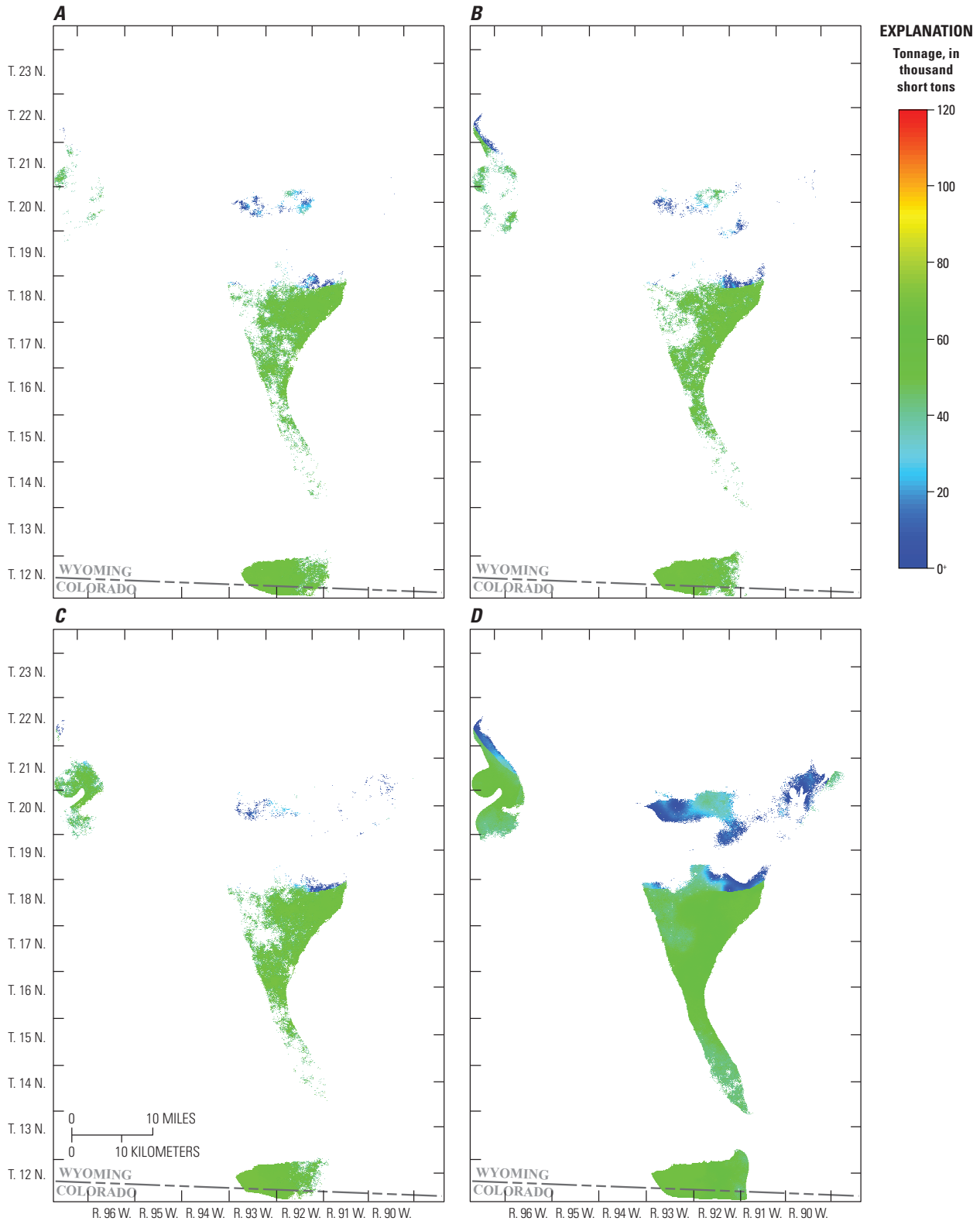
Figure 48.   Graphs showing probabilistic assessment of the coal resources available for recovery by surface mining methods. *A*, Histogram. *B*, Cumulative distribution.
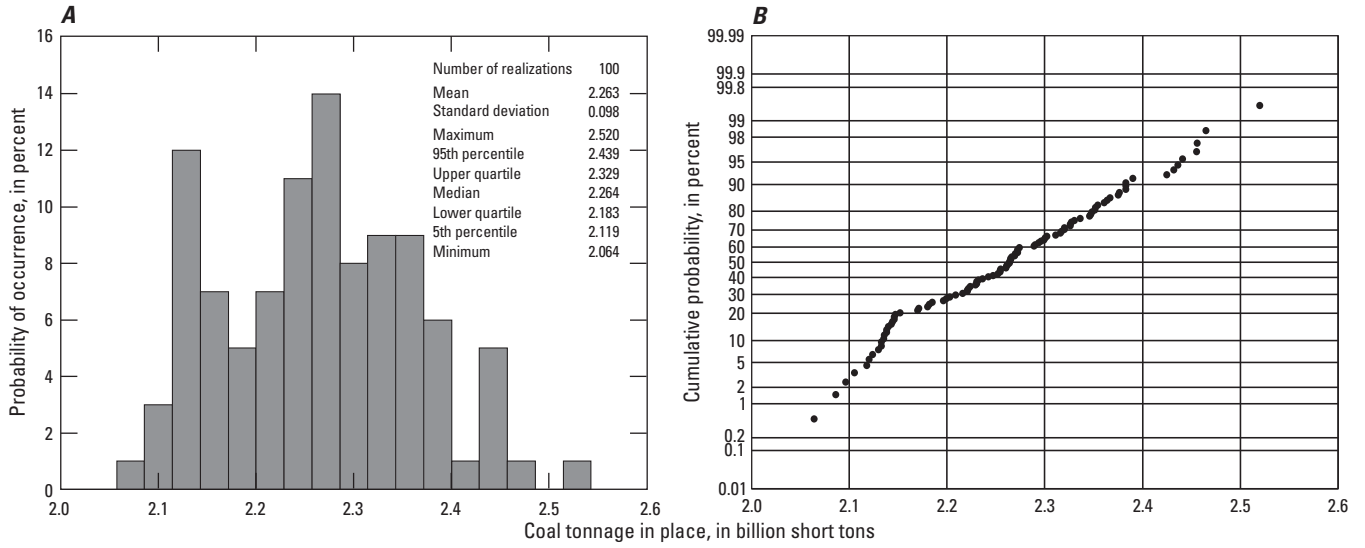

Figure 49.   Map of the 5–95 spread for the cells fitting the criteria for surface mining recovery.

Number of cells	6,020
Mean	40.718
Standard deviation	29.983
Maximum	135.900
95th percentile	84.400
Upper quartile	66.800
Median	43.272
Lower quartile	10.900
5th percentile	2.000
Minimum	0.000

Figure 50.   Graphs showing summary of cell tonnage reliability. *A*, Histogram of the values shown in figure 49. *B*, Cumulative distribution of the same values (left vertical axis) combined with cumulative tonnage taken from the values shown in figure 47*D* (right vertical axis).

Figure 51.   Maps showing a selective display of 100 realizations of underground mining resources by magnitude of the tonnage (*A–C*) and as an average of all realizations (*D*). *A,* Realization 40 with the fifth lowest tonnage (2.119 billion short tons [bst]). *B,* Realization 27 with the median tonnage (2.264 bst). *C,* Realization 7 with the fifth highest tonnage (2.439 bst). *D,* E-type realization. 0+ denotes a value close to but larger than zero.

A

Number of realizations	100
Mean	2.263
Standard deviation	0.098
Maximum	2.520
95th percentile	2.439
Upper quartile	2.329
Median	2.264
Lower quartile	2.183
5th percentile	2.119
Minimum	2.064

B

Coal tonnage in place, in billion short tons

Figure 52.   Graphs showing probabilistic assessment of the coal resources available for recovery by underground mining methods. *A*, Histogram. *B*, Cumulative distribution.
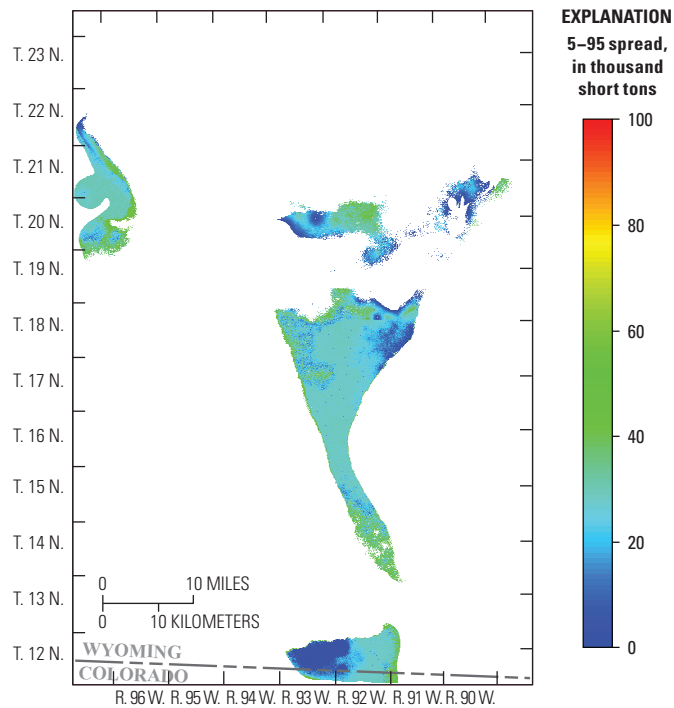
Figure 53.   Map of the 5–95 spread according to 100 realizations selectively displayed in figure 51.
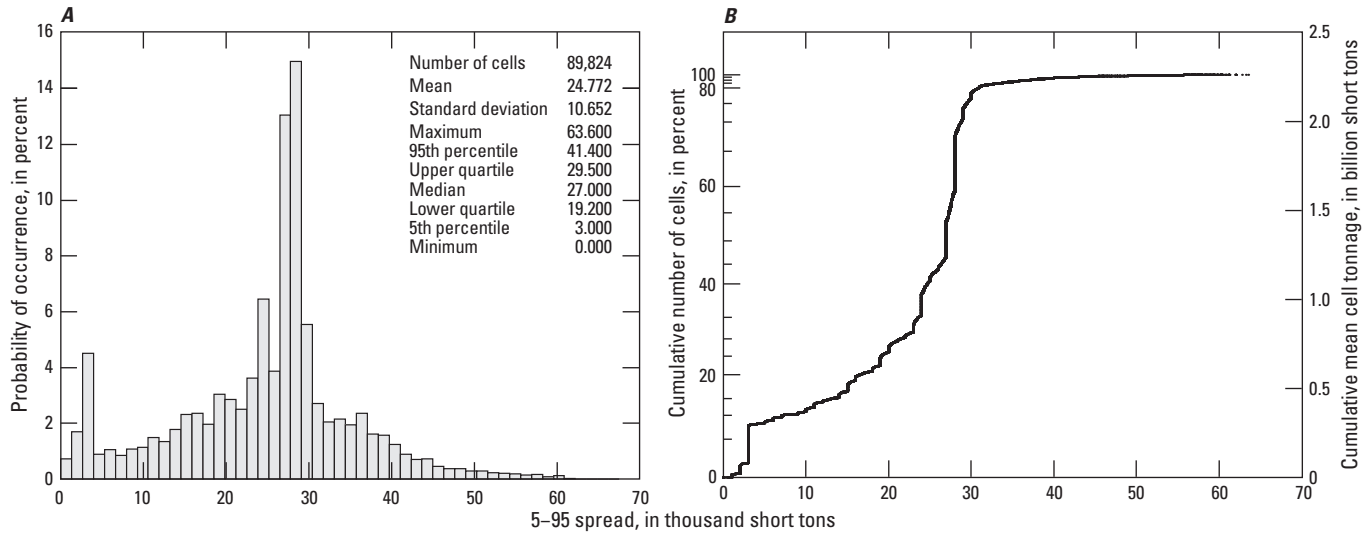
Figure 54.   Graphs showing uncertainty in cell tonnage for the underground mining resources. *A*, Histogram of the values shown in figure 53. *B*, Cumulative distribution of the same values (left vertical axis) plus cumulative tonnage taken from the values shown in figure 51*D* (right vertical axis).

Workflow

A. Data

Minimum data for a collection of drill holes inside and outside a coal bed must include the following:

1. Easting.

2. Northing.

3. Coal bed thickness.

4. Elevation of the top of the coal bed, if the coal bed is present.

   In addition, it is necessary to have either:

5a. Coal bed density measurements

   or

5b. A constant value for the average coal bed density.

   In addition, it is necessary to have either:

6a. Measurements of the oxidation zone thickness

   or

6b. A constant value for the average oxidation zone thickness.

   Finally, it is required to have:

7. A ground-surface elevation grid having nodes coinciding with those to be used in the modeling.

The Fillmore Ranch coal bed data are taken from Haacke (2018), with the exception of the density data that were prepared on the basis of information listed in table 4.

B. Processing

All utility programs mentioned below can be downloaded from a site specified in Olea and Shaffer (2021). Graphical display of the results obtained with the software GSLIB (Deutsch and Journel, 1997) were modified by changing the Encapsulated PostScript code of the illustrations to satisfy the standards for USGS publications. For the same reasons, displays of semivariograms produced with the software SGeMS (Remy and others, 2009) were altered graphically using Adobe Illustrator.

The steps in the modeling are as follows:

1. If the coal bed is not laterally continuous within the study area, convert the set of thickness values to a binary variable denoting presence or absence of the coal bed by using a code of 1 to denote presence and 0 to denote absence, an operation that can be done in Excel. In the following steps, it will be understood that all modeling related to the binary variable will be skipped if the coal bed is laterally continuous within the study area.

2. Display the histogram and the cumulative distribution of thickness where the coal bed is present, applying programs *histplt* and *probplt* in GSLIB (fig. 17).

3. Post both the thickness and the binary values, applying program *locmap* in GSLIB (fig. 18).

4. Apply utility program *nearest* to the drill holes inside the coal bed. Use GSLIB program *scatplt* to display the results as a scatterplot of thickness rank versus distance to the nearest drill hole (fig. 19). Use this scatterplot and posting of thickness values inside the coal bed for deciding if there is clustering. Go to Step 6 if no preferential sampling is detected.

5. Execute utility program *decluster* for obtaining datasets free of preferential sampling both for thickness and the binary variable, which will be used solely for semivariogram modeling. Display the results with GSLIB programs *histplt, qpplt* and *locmap* (figs. 20, 21).

6. Apply option *Data Analysis→Variogram* in public domain software SGeMS to model the semivariograms for the binary variable and the normal scores of thickness (fig. 22). Use as data the declustered dataset in case of having detected preferential sampling.

7. Define the perimeter of a rectangular study area and subdivide it into square cells.

8. Select option *Algorithm→Estimation→kriging* in software SGeMS for running an ordinary kriging mapping of the binary variable. Apply utility *kvar2ksd* for converting kriging variance to kriging standard deviation. Display the results using module *pixelplt* in GSLIB (fig. 23).

9. Running utility program *trim*, use the previous results to define the area where modeling is possible because there are enough data indicating high likelihood that the cells are inside the coal bed. Display the results with *pixelplt* (fig. 24).

10. Applying option *Algorithm→Simulation→sisim* in SGeMS, generate 100 realizations of the binary variable.

11. Applying utility program *g2gtrim*, limit the extension of the realizations to the region defined in Step 9 and display some of the results with *pixelplt* (fig. 25).

12. Select option *sgsim* in SGeMS's *Algorithm→Simulation* to generate 100 realizations of thickness, using a different random seed than that used in Step 10. In what follows, it will be understood that every set of realizations must be generated using a different random seed.

13. Run utility program *blankgridsis* to limit the extension of the thickness simulations to those regions determined in the previous step, partly displaying the modified realizations with *pixelplt* (fig. 26).

14. If necessary, where required, apply utility program *pinchout* for wedging out the thickness realizations toward the boundaries of the coal bed, mapping the results with *pixelplt* (fig. 27).

15. If there is only an average value for coal density, run utility program *multgridconst* to multiply all 100 thickness realizations by that constant value to convert the thickness realizations into tonnage realizations. Go to Step 19.

16. When a dataset of density values is available, repeat Steps 2–6, this time using the density values (figs. 28–30).

17. Select option *Algorithm→Simulation→sgsim* within SGeMS to produce 100 realizations of coal density and display some of the grids with *pixelplt* after trimming using the same realizations in Step 11 (fig. 31).

18. Multiply cell-by-cell one thickness realization by one density realization using utility program *thick2ton*.

19. Apply utility program *resourcesCsis* for calculating total tonnage per tonnage realization.

20. Rank the total tonnage per realization with Excel.

21. Use these ranks for selective display of the tonnage realizations with *pixelplt* (figs. 32*A–C*).

22. Display the results in Step 19 with *histplt* in the form of a histogram and with *probplt* as a cumulative distribution (fig. 33).

23. Run utility *EtypeSGeMSsis* for generating a grid with some cell-by-cell statistics. Show the resulting mean values as an E-type map with *pixelplt* (fig. 32*D*).

24. Use the results in Step 19 to prepare a sensitivity analysis on the number of realizations to check that the results of the modeling are within a stable interval (fig. 34). If the results of the modeling are not within a stable interval, go back to Step 10 to generate additional realizations.

25. Use the results in Step 23 to map the 5–95 spread with *pixelplt* (fig. 35).

26. Display the results in Step 23 for 5−95 spread as a histogram with *histplt* and as a cumulative distribution combining programs *reliabcurve* and *scatplt* (fig. 36).

27. Display the ground-surface elevation map with *pixelplt* (fig. 40).

28. Repeat Steps 16 and 17, using instead the data for the top of the coal bed (figs. 41–43).

29. Applying *Algorithm→Simulation→sgssim* in SGeMS for generating 100 realizations of the top of the coal bed, process the realizations with *blankgridsis* to match the coal bed areal extension for the tonnage realizations, and display some of the results with *pixelplt* (fig. 44).

30. Execute utility program *dif2surf* to have realizations with the difference between the ground surface elevation and each of the realizations for the elevation of the top of the coal bed. Display a few of the resulting cover realizations with *pixelplt* (fig. 45).

31. If there is only a constant value for the oxidation depth, go to Step 34.

32. Using the data for the thickness of the oxidation zone, repeat Steps 16 and 17.

33. Select option *Algorithm→Simulation→sgssim* in SGeMS, generate 100 realizations for the thickness of the oxidation zone, and display a few with *pixelplt*.

34. Process the cover realizations with utility program *coverclass*. Partially display the resulting realizations with *pixelplt* (fig. 46).

35. Process the thickness, tonnage, and cover realizations plus oxidation thickness or oxidation realizations with utility program *mining* to obtain realizations for the tonnage available for both surface and underground mining methods.

36. Repeat Steps 19–23 and 25–26, this time using the realizations of tonnage available for surface mining instead of the realizations for in-place tonnage (figs. 47–50).

37. Complete the assessment, repeating the previous step for the tonnage available for underground mining instead of that for surface mining (figs. 51–54).

Conclusions

Application of geostatistical techniques that have been used to advantage to improve modeling in the oil industry and in ore mining has resulted in a state-of-the-art methodology for the assessment of coal resources. Advantages of the new methodology described in this publication include the following:

- Consistency—The same set of tonnage realizations are the basis for assessing both the magnitude of the resources and their uncertainty in the modeling. The current practice is to use different methods, with distance to the nearest drill hole being the prevailing approach to evaluate reliability.

- Full usage of data—Thickness data implicitly have plenty of information that is not used when considering only distance to the nearest drill hole, disregarding the distance to all other drill holes and the style of fluctuations. Geostatistics makes use of all that information without requiring additional data.

- Cell tonnage accuracy—While a nearest distance criterion provides a qualitative index of reliability, use of the concept of random variable allows a quantitative assessment of uncertainty cell by cell. For example, now it is possible to report the length of confidence intervals containing the true value of tonnage with 90-percent probability.

- Total tonnage uncertainty—A different random variable extracted from the same realizations allows evaluation of uncertainty specific to total tonnage, which is quantitative and different for resources in place, resources mineable by surface methods, and resources mineable by underground mining methods. The magnitude and uncertainty of the resources can now be reported, for example, as 9.687 ± 0.383 billion short tons with 90-percent probability.

- Sensitivity to geology—The semivariogram allows the capturing of geological differences among different coal beds. Hence, even if these beds are penetrated by the same drill holes, the reliability will be different.

- Sensitivity to boundary conditions—Drill holes inside and outside the boundaries of the coal bed are used for evaluating both the magnitude and the uncertainty of the resources.

- Sensitivity to density—When coal density values are available in sufficient number to conduct mapping, it is possible to have fluctuations in tonnage reliability different from the fluctuations in reliability associated with thickness.

- Confidence intervals—Proper application of geostatistics has proven to provide confidence intervals containing the true cell tonnage with probabilities validated by additional drilling, confirming the realistic power of the modeling.

Preparation of an assessment following the methodology presented in this report results in multiple plots and maps, most of which may be considered intermediate results. There are two types of final results for either resources in place or mineable resources: total tonnage and cell tonnage. There are two aspects for either kind of tonnage: the magnitude of the resource and the reliability of its modeling. The magnitude of the resource can be viewed in the form of a map, such as those in figures 32*B* or 32*D*, or as a distribution of possible values, such as figure 33. By providing the entire distribution, it is possible for the reader to select at will percentiles of interest, if any. Uncertainty is also characterized by maps and probability distributions, such as those in figures 35 and 36.

While the deterministic tonnages in place, tonnages available for surface mining, and tonnages available for underground mining that the USGS has been assessing are in general agreement with the probabilistic distributions resulting from the methodology presented here, there is a

drastic departure in terms of the 5−95 spread used here and the distance categories used by the USGS to quantify uncertainty in tonnage at the block scale.

The methodology detailed in this report fully satisfies the requirements and needs of the USGS. It was not possible to cover here the multiple variants, simplifications, or additions that other agencies or companies may be interested in. We leave it up to the reader to use this report as a starting point for preparing alternative approaches. Also, modelers willing to take this report as a starting point could use it for the assessment of other mineral commodities.

References Cited

American Society for Testing and Materials, 2019, Standard practice for statistical modeling of uncertainty in assessment of in-place coal resources: ASTM International, Designation D8215-19a, 12 p., accessed June 2020 at https://www.astm.org/Standards/D8215.htm.

Briggs, W., 2016, Uncertainty—The soul of modeling, probability & statistics: Switzerland, Springer, 258 p.

Caers, J., 2011, Modeling uncertainty in the earth sciences: Chichester, United Kingdom, Wiley-Blackwell, 227 p.

Chilès, J.-P., and Delfiner, P., 2012, Geostatistics—Modeling spatial uncertainty (2d ed.): Hoboken, N.J., John Wiley & Sons, 734 p.

Christakos, G., Olea, R.A., Serre, M., Yu, H.L., and Wang, L.L., 2005, Interdisciplinary public health reasoning and epidemic modelling—The case of Black Death: Berlin, Springer, 319 p.

Committee for Mineral Reserves International Reporting Standards [CRIRSCO], 2013, International reporting template for the public reporting of exploration results, mineral resources and mineral reserves: CRIRSCO, 41 p., accessed June 2020 at http://www.crirsco.com/templates/crirsco_international_reporting_template_2013.pdf.

David, M., 1988, Handbook of applied advanced geostatistical ore reserve estimation: Amsterdam, Elsevier Science Publishers, Developments in Geomathematics 6 [series no. 6], 216 p.

Deutsch, C.V., 2018, All realizations all the time, in Sagar, B.S.D., Cheng, Q., and Agterberg, F., eds., Handbook of mathematical geosciences—Fifty years of IAMG: Cham, Switzerland, Springer, p. 131–142.

Deutsch, C.V., and Journel, A.G., 1997, GSLIB—Geostatistical software library and user's guide (2d ed.): New York, Oxford University Press, 369 p. and 1 CD.

Emery, X., 2012, Co-simulating total and soluble copper grades in an oxide ore deposit: Mathematical Geosciences, v. 44, no. 1, p. 27–46.

Everitt, B.S., and Skrondal, A., 2010, The Cambridge dictionary of statistics (4th ed.): Cambridge, United Kingdom, Cambridge University Press, 480 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 simulations, Proceedings of the Geostatistical Simulation Workshop, Fontainebleau, France, 27−28 May 1993: Dordrecht, The Netherlands, Kluwer Academic Publishers, p. 111−124.

Haacke, J.E., 2018, Coal drill hole database for the Little Snake River coal field and Red Desert area, Wyoming: U.S. Geological Survey data release, accessed June 2020 at https://doi.org/10.5066/F7RN36RG.

Hengl, T., 2006, Finding the right pixel size: Computers & Geosciences, v. 32, no. 9, p. 1283–1298.

Hengl, T., 2009, A practical guide to geostatistical mapping: Luxembourg, Office for Official Publications of the European Communities, 270 p., accessed June 2020 at https://www.onlineprogrammingbooks.com/free-engineering-ebook-a-practical-guide-to-geostatistical-mapping/.

Heriawan, M.N., and Koike, K., 2008, Uncertainty assessment of coal tonnage by spatial modeling of seam distribution and coal quality: International Journal of Coal Geology, v. 76, no. 3, p. 217–226.

Hettinger, R.D., Honey, J.G., Ellis, M.S., Barclay, C.S.V., and East, J.A., 2008, Geologic map of Upper Cretaceous and Tertiary strata and coal stratigraphy of the Paleocene Fort Union Formation, Rawlins-Little Snake River area, south-central Wyoming: U.S. Geological Survey Scientific Investigations Map 3053, 3 sheets. [Also available at https://doi.org/10.3133/sim3053.]

Hogg, R.V., McKean, J.W., and Craig, A.T., 2018, Introduction to mathematical statistics (8th ed.): Boston, Pearson, 768 p.

Hohn, M.E., and McDowell, R.R., 2001, Uncertainty in coal property valuation in West Virginia—A case study: Mathematical Geology, v. 33, no. 2, p. 191–216.

Joint Ore Reserves Committee [JORC], 2012, The JORC Code, 2012 Edition—Australasian code for reporting of exploration results, mineral resources and ore reserves: Australasian Institute of Mining and Metallurgy, Australian Institute of Geoscientists, and Minerals Council of Australia, accessed June 2020 at jorc.org/docs/jorc_code2012.pdf.

Journel, A.G., and Huijbregts, C.J., 1978, Mining geostatistics: London, Academic Press, 600 p.

Journel, A.G., and Kyriakidis, P.C., 2004, Evaluations of mineral reserves—A simulation approach: New York, Oxford University Press, 216 p.

Karacan, C.Ö., and Olea, R.A., 2018, Mapping of compositional properties of coal using isometric logratio transformation and sequential Gaussian simulation—A comparative study for spatial ultimate analyses data: Journal of Geochemical Exploration, v. 186, p. 36–49.

Leuangthong, O., Khan, K.D., and Deutsch, C.V., 2008, Solved problems in geostatistics: Hoboken, N.J., John Wiley & Sons, 207 p.

Manchuk, J.G., Leuangthong, O., and Deutsch, C.V., 2009, The proportional effect: Mathematical Geosciences, v. 41, no. 7, p. 799–816.

Matheron, G., 1963, Principles of geostatistics: Economic Geology, v. 58, no. 8, p. 1,246–1,266.

Matheron, G., 1965, Les variables régionalisées et leur estimation; une application de la théorie des fonctions aléatories aux Sciences de la Nature: Paris, Masson et Cie, 305 p.

Olea, R.A., 2006, A six-step practical approach to semivariogram modeling: Stochastic Environmental Research and Risk Assessment, v. 20, no. 5, p. 307–318.

Olea, R.A., 2009, A practical primer on geostatistics (ver. 1.4, December 2018): U.S. Geological Survey Open-File Report 2009–1103, 346 p., accessed June 2020 at https://doi.org/10.3133/ofr20091103.

Olea, R.A., 2012, Geostatistics for engineers and earth scientists: Norwell, Mass., Kluwer Academic Publishers, 303 p.

Olea, R.A., 2017, Resampling of spatially correlated data with preferential sampling for the estimation of frequency distributions and semivariograms: Stochastic Environmental Research and Risk Analysis, v. 31, no. 2, p. 481–491.

Olea, R.A., 2018, Advances in sensitivity analysis of uncertainty to changes in sampling density when modeling spatially correlated attributes, *in* Sagar, B.S.D., Cheng, Q., and Agterberg, F., eds., Handbook of mathematical geosciences—Fifty years of IAMG: Cham, Switzerland, Springer, p. 375–393.

Olea, R.A., and Luppens, J.A., 2012, Sequential simulation approach to modeling of multi-seam coal deposits, with an application to the assessment of a Louisiana lignite: Natural Resources Research, v. 21, no. 4, p. 443–459.

Olea, R.A., and Luppens, J.A., 2014, Modeling uncertainty in coal resource assessments, with an application to a central area of the Gillette coal field, Wyoming: U.S. Geological Survey Scientific Investigations Report 2014–5196, 46 p., accessed July 2020 at https://doi.org/10.3133/sir20145196.

Olea, R.A., and Luppens, J.A., 2015, Mapping of coal quality using stochastic simulation and isometric logratio transformation with an application to a Texas lignite: International Journal of Coal Geology, v. 152, pt. B, p. 80−93.

Olea, R.A., and Shaffer, B.N., 2021, Computer programs for the assessment of coal resources: U.S. Geological Survey software release, https://doi.org/10.5066/P971RL9L.

Olea, R.A., Luppens, J.A., Egozcue, J.J., and Pawlowsky-Glahn, V., 2016, Calorific value and compositional ultimate analysis with a case study of a Texas lignite: International Journal of Coal Geology, v. 162, p. 27–33.

Olea, R.A., Luppens, J.A., and Tewalt, S.J., 2011, Methodology for quantifying uncertainty in coal assessments with an application to a Texas lignite deposit: International Journal of Coal Geology, v. 85, no. 1, p. 78–90.

Olea, R.A., Luppens, J.A., and Tewalt, S.J., 2014, Moving away from distance classifications as measures of resource uncertainty, *in* Pardo-Igúzquiza, Eulogio, Guardiola-Albert, Carolina, Heredia, Javier, Moreno-Merino, Luis, Durán, J.J., and Vargas-Guzmán, J.A., eds., Mathematics of Planet Earth, Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences: Heidelberg, Germany, Springer, p. 585−588.

Pipiringos, G.N., 1961, Uranium-bearing coal in the central part of the Great Divide Basin, chap. A *of* Uranium-bearing coal in the Great Divide Basin, Sweetwater County, Wyoming: U.S. Geological Survey Bulletin 1099–A, p. A-1–A-104, 5 pls.
[Also available at https://doi.org/10.3133/b1099A.]

Pyrcz, M.J., and Deutsch, C.V., 2014, Geostatistical reservoir modeling (2d ed.): New York, Oxford University Press, 433 p.

Remy, N., Boucher, A., and Wu, J., 2009, Applied geostatistics with SGeMS—A user's guide: New York, Cambridge University Press, 264 p. and 1 CD.

Rendu, J.-M., 2006, Reporting mineral resources and mineral reserves in the United States of America—Technical and regulatory issues, *in* Proceedings, Sixth International Mining Geology Conference: Australasian Institute of Mining and Metallurgy, p. 11−20.

Robinson, T.P., and Metternicht, G., 2006, Testing the performance of spatial interpolation techniques for mapping soil properties: Computers and Electronics in Agriculture, v. 50, no. 2, p. 97–108.

Rossi, M.E., and Deutsch, C.V., 2014, Mineral resource estimation: Dordrecht, The Netherlands, Springer, 332 p.

Salkind, N.J., 2007, Statistics for people who (think they) hate statistics: Thousand Oaks, Calif., Sage Publishers, 402 p.

Scott, D.C., Shaffer, B.N., Haacke, J.E., Pierce, P.E., and Kinney, S.A., 2019, Coal geology and assessment of coal resources and reserves in the Little Snake River coal field and Red Desert assessment area, Greater Green River Basin, Wyoming: U.S. Geological Survey Professional Paper 1836, 169 p., accessed July 2020 at https://doi.org/10.3133/pp1836.

Shaffer, B.N., 2017, Assessing U.S. coal resources and reserves: U.S. Geological Survey Fact Sheet 2017–3067, 6 p. [Also available at https://doi.org/10.3133/fs20173067.]

Shaffer, B.N., Pierce, P.E., Kinney, S.A., Olea, R., and Luppens, J.A., 2019, Assessment of coal resources and reserves in the Little Snake River coal field and Red Desert assessment area, Greater Green River Basin, Wyoming: U.S. Geological Survey Fact Sheet 2019–3053, 6 p. [Also available at https://doi.org/10.3133/fs20193053.]

Souza, L.E., de, Costa, J.F.C.L., and Koppe, J.C., 2004, Uncertainty estimate in resources assessment—A geostatistical contribution: Natural Resources Research, v. 13, no. 1, p. 1–15.

Srivastava, R.M., 2013, Geostatistics—A tool for data analysis, spatial prediction and risk management in the coal industry: International Journal of Coal Geology, v. 112, p. 2–13.

Triola, M.F., 2014, Elementary statistics (12th ed.): Boston, Pearson, 822 p. and 1 CD.

U.S. Geological Survey [USGS], 2018, Do US Topos and The National Map have a layer that shows the Public Land Survey System (PLSS)?: USGS website, accessed June 2020 at https://www.usgs.gov/faqs/do-us-topos-and-national-map-have-a-layer-shows-public-land-survey-system-plss.

U.S. Securities and Exchange Commission [SEC], 2016, Modernization of property disclosures for mining registrants: SEC Releases 33-10098 and 34-78086, File S7-10-16, 296 p., accessed June 2020 at https://www.sec.gov/rules/proposed/2016/33-10098.pdf.

U.S. Securities and Exchange Commission [SEC], 2018, SEC adopts rules to modernize property disclosures required for mining registrants: SEC Press Release 2018-248, October 31, 2018, accessed June 2020 at https://www.sec.gov/news/press-release/2018-248.

Weisberg, H.I., 2014, Willful ignorance—The mismeasure of uncertainty: Hoboken, N.J., John Wiley & Sons, 434 p.

Wood, G.H., Jr., Kehn, T.M., Carter, M.D., and Culbertson, W.C., 1983, Coal resources classification system of the U.S. Geological Survey: U.S. Geological Survey Circular 891, 65 p. [Also available at https://doi.org/10.3133/cir891.]

Index

A page number in bold denotes a reference to a definition.

Olea and others—**Probabilistic Methodology for the Assessment of Original and Recoverable Coal Resources**—TM 6–G1