Combining Decline-Curve Analysis and Geostatistics To Forecast Gas Production in the Marcellus Shale

Zhenke Xi and **Eugene Morgan,** Pennsylvania State University

Summary

To estimate the production potential at a new, prospective field site by means of simulation or material balance, one needs to collect various forms of costly field data and make assumptions about the nature of the formation at that site. Decline-curve analysis (DCA) would not be applicable in this scenario, because producing wells need to pre-exist in the target field. The objective of our work was to make first-order forecasts of production rates at prospective, undrilled sites using only production data from existing wells in the entire play. This is accomplished through the co-Kriging of decline-curve parameter values, where the parameter values are obtained at each existing well by fitting an appropriate decline model to the production history. Co-Kriging gives the best linear unbiased prediction of parameter values at undrilled locations, and also estimates uncertainty in those predictions. Thus, we obtained production forecasts at P10, P50, and P90, and we calculated the estimated ultimate recovery (EUR) at those same levels across the spatial domain of the play.

To demonstrate the proposed methodology, we used monthly gas-flow rates and well locations from the Marcellus shale-gas play in this research. Monitoring only horizontal and directional wells, the gas-production rates at each well were carefully filtered and screened. Also, we normalized the rates by perforation-interval length. We only kept production histories of 24 months or longer to ensure good decline-curve fits. Ultimately, we were left with 5,637 production records. Here, we chose Duong's decline model (Duong 2011) to represent the production decline in this shale-gas play, and fitting this decline curve was accomplished through ordinary least-squares (OLS) regression.

Interpolation was done by universal co-Kriging while considering the correlation between the four parameters in Duong's model, which also showed linear trends (the parameters showed dependency on the x and y spatial coordinates). Kriging gave us the optimal decline-curve coefficients at new locations (P50 curve), as well as the variance in these coefficient estimates (used to establish P10 and P90 curves). We were also able to map EUR for 25 years across the study area. Finally, the universal co-Kriging model was cross validated with a leave-one-out scheme, which showed significant, but not unreasonable, error in the decline-curve-coefficient prediction. The methods proposed were implemented and did not require various costly data, such as permeability and bottomhole pressure, thus giving operators a risk-based analysis of prospective sites. While we demonstrated the procedure on the Marcellus shale-gas play, it is applicable to any play with existing producing wells. We also made this analysis available to the public in a user-friendly web application (Xi and Morgan 2018).

Introduction

The unconventional shale-gas revolution has played an important role in supplying the increasing energy demand worldwide and has created benefits such as national energy security and jobs. However, the development of unconventional gas can be substantially risky because of the difficulty characterizing the reservoir and understanding how the reservoir behaves. Unlike conventional reservoirs, unconventional shale always has complicated transport mechanisms and ultralow permeability, while the induced hydraulic fractures act as additional factors in evaluating the well performance (Clarkson et al. 2012). As a result, effective methods that can forecast the unconventional gas production are critical for risk assessment and mitigation.

DCA is one of the basic tools used in production-rate forecasting and recovery estimation (Bhattacharya and Nikolaou 2013; Erdle et al. 2016). The advantage it has over other approaches (e.g., material balance or reservoir simulation) is that it requires less data and uses simple, empirical models. Arps' hyperbolic rate/time equation (Arps 1945), which is widely applied in conventional plays, can also capture the behavior of shale-gas rate decline with its hyperbolic decline constant b, which is sometimes greater than unity for unconventional resources. Cipolla et al. (2010) addressed issues in techniques of conventional production analysis when they are applied to unconventional reservoirs. Anderson and Liang (2011) developed a probabilistic approach on the basis of a "trilinear flow" model for unconventional-reservoir characterization and forecasting. Additional models that deal with complex fracture networks were developed by Diaz De Souza et al. (2012) and Moinfar et al. (2013). Furthermore, with the advancement of computing power and storage capacity of computer hardware, a variety of machine-learning or data-mining methods can make contributions to production prediction. For example, an expert system built on a multilayer artificial neural network (ANN) was developed by Ketineni et al. (2015) for the production-performance prediction of infill-drilling locations at an offshore field. Support vector machine, a sophisticated method for statistics, works well in pattern recognition, such as determining the location of sweet spots (Hauge and Hermansen 2017).

Although the previous methods might excel (given sufficient data and under appropriate assumptions), there are still limitations. For example, numerical history matching requires solvers with strong robustness and various types of field data (Oliver and Chen 2011). These conditions cannot always be satisfied because of the complexity of partial-differential equations or the scarcity of data in some newly developed plays. ANNs might be viable with less-refined data, but a large amount of data and different data types are still necessary. To find a proper ANN architecture, a significant amount of time is required for model training and cross validation. In addition, many such approaches only produce deterministic results without quantifying the uncertainty.

In this work, we designed a simple and elegant approach to forecast production without relying on field data, such as permeability and bottomhole pressure, with the ability to estimate the uncertainty of the predicted results. First, we fitted the production-decline curve for existing wells using Duong's model (Duong 2011). Then, a universal co-Kriging approach was adopted to interpolate the

This paper (SPE-197055-PA) was revised for publication from paper SPE-191793-18ERM-MS, first presented at the SPE/AAPG Eastern Regional Meeting, Pittsburgh, Pennsylvania, USA, 7–11 October 2018. Original manuscript received for review 25 September 2018. Revised manuscript received for review 15 February 2019. Paper peer approved 25 February 2019.

decline-curve parameters at undrilled sites. Finally, a heatmap of EUR was constructed with the interpolated parameters, enabling us to generate production forecasts up to an arbitrary number of months in the future for any site in the spatial domain of the play. Because Kriging models the decline-curve parameters as a Gaussian process over space, the variance of the predicted values was used to generate P10 and P90 bounds on forecasts and, thus, EUR. While we demonstrated the approach on the Marcellus shale-gas play, the methodology can be applied to any other play, conventional or unconventional, given that sufficient regional production data are available and an appropriate decline-curve model exists for the reservoir type.

Methodology

Data Cleaning. Processing raw data is a common step before extracting any interesting information from the data. To fit Duong's model (Duong 2011) to the flow rates that represent the true production behavior of the reservoir, we cleaned the data in the following ordered steps:

1. Drop data from vertical wells (keep only directional- and horizontal-well flow rates).
2. Order flow rate by increasing date.
3. Remove values that date from before the completion date.
4. Remove any zero values at the beginning (before first positive value) or remove first value if it is positive (might be a partial observation).
5. Remove data from before peak production, if there are any left.
6. Check for gaps, zero-values, or large changes (possibly created, for example, by shut-in or restimulation) in the record and only keep data until first instance.
7. Reject the well if the resulting record is less than 24 months long (to ensure sufficient sample size for robust regression).
8. Normalize the gas rate by perforation-interval length, lateral length, or similar.

Decline-Curve Analysis. Duong (2011) proposed a decline-curve model that is serviceable for unconventional reservoirs that exhibit long periods of linear flow caused by the pressure depletion within fracture networks. It models gas rate (q) as a function of time (t) and four parameters (a, m, q_i, q_∞) determined by means of two stages of sequential OLS regression. First, a and m were determined by OLS:

$$\ln\left(\frac{q}{G_p}\right) = m\ln(t) + a. \quad (1)$$

Then, q_i and q_∞ were determined by OLS regression:

$$q = q_i t_{am} + q_\infty, \quad (2)$$

where

$$t_{am} = t^{-m}\exp\left[\frac{a}{1-m}\left(t^{1-m} - 1\right)\right]. \quad (3)$$

In addition to catching the salient production traits of unconventional reservoirs, Duong's model is also advantageous in its numerical stability when fitting the curve. Because the four parameters were obtained by implementing OLS on Eqs. 1 and 2, they were guaranteed to be the solution for the global minimum of the squared error surface. This feature makes the fitting procedure faster than those using Arps' hyperbolic model, which involves nonlinear regression where the global minimum of squared error cannot be ensured even after many iterations.

Geostatistics Method. To calculate EUR at undrilled sites, the associated parameters must be predicted on the basis of known data points. Although a variety of mathematical interpolation techniques are available, such as spline and polynomial, the spatial correlation among observations was ignored. Therefore, methods from the Kriging family that would handle the underlying spatial structures were applied in our work.

Intuitively, the four parameters (a, m, q_i, q_∞) for each well should be mutually related because they collectively quantify the same decline curve. Meanwhile, a more reliable estimate can be obtained by using the spatial correlation between different variables (Trangmar et al. 1986). Thus, co-Kriging was used in our work to exploit the (potential) cross correlation between the four decline-curve parameters and their joint spatial correlation simultaneously. Co-Kriging is a well-established procedure detailed in many books (Journel and Huijbregts 1978; McBratney and Webster 1983; Vauclin et al. 1983). While co-Kriging takes advantage of multiple parameters, large-scale spatial trends in the data (external drift) must be accounted for before Kriging. This idea is imparted from universal Kriging, which allows for Kriging in the presence of nonstationarity (Matheron 1971). Therefore, we used the features of the two Kriging methods by first filtering out the external drift from the data and then co-Kriging the associated residuals, which led to an indirect but comprehensive hybrid method, "universal co-Kriging."

To construct maps of decline-curve parameters, and the subsequent EUR, over the spatial domain of the play, it is best to perform this universal co-Kriging on blocks rather than specific points. The size of the block should be chosen to balance utility of the resulting map and computation time, while the size should also be at a useful scale to support operational decisions and give sufficient spatial resolution. However, too many blocks results in large memory and computational-time requirements for the Kriging process, and might produce an unnecessary level of detail. Here, we recommend scaling the blocks to be roughly twice the average lateral length, such that if a well pad were placed in the center of a block, the lateral(s) would still be contained within the block, unless chosen to be exceedingly long. In "block Kriging" the mean and variance of the predicted parameters are determined for each block. Because we performed co-Kriging, the covariances were also determined. Thus, we determined the P10, P50, and P90 EUR by numerically integrating Duong's decline curve with its parameters sampled from a multivariate normal distribution. To better explain this integrated approach, a case study on the Marcellus Shale is presented in the following section.

Case Study

Data Source. Monthly production data were obtained for all the Marcellus wells (11,904 wells out of 15,988 have production data, and their positions are shown in **Fig. 1**), along with various header data for each well (e.g., geographical coordinates, completion date), from DrillingInfo (2018) and the West Virginia Geologic and Economic Survey (2018). The data have been cleaned by the procedure

previously introduced. Furthermore, we used flow rates only from horizontal and directional wells (records from vertical wells were dropped) because they play a major role in massive unconventional shale-gas extraction. Besides, normalizing the flow rates with the perforation-interval length can lead to an equitable comparison between wells. Data from 5,637 wells met all the conditions and, thus, were later used for curve fitting. The data size after each processing stage is summarized in **Table 1.**

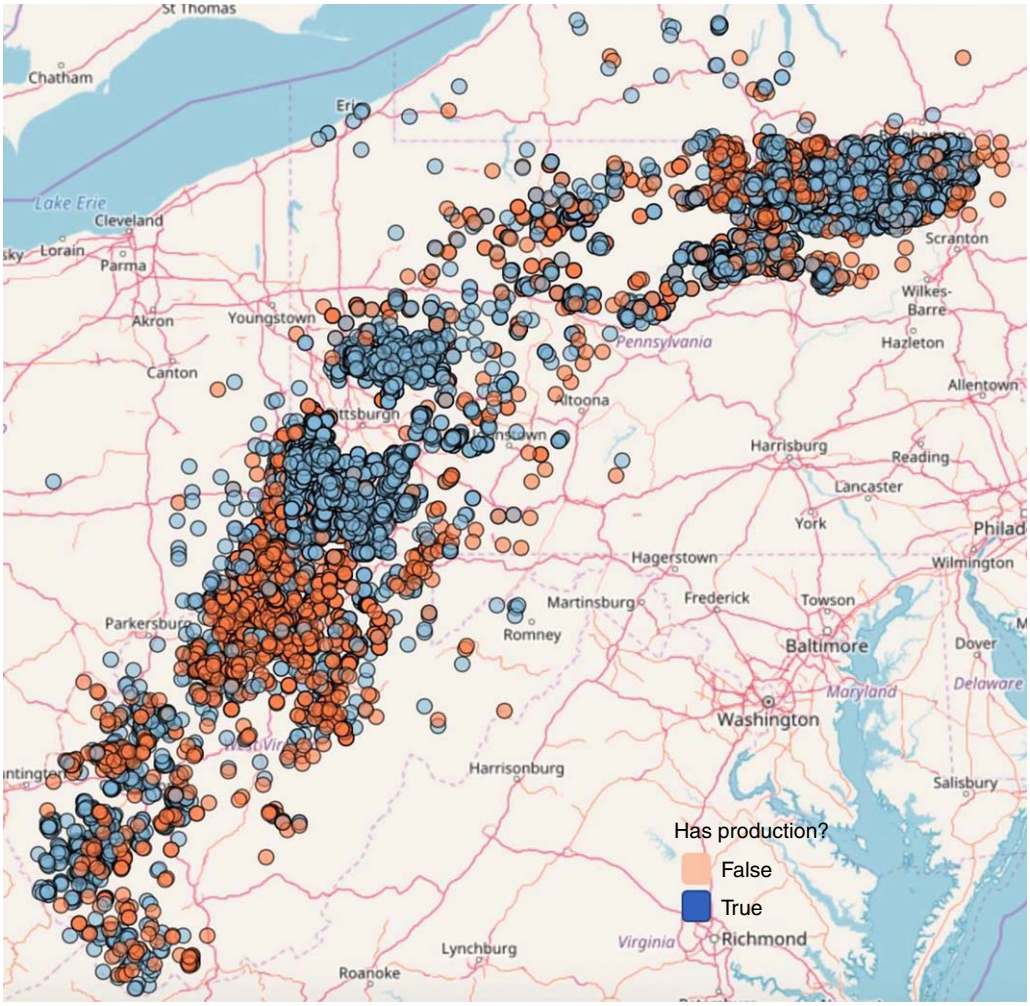

Fig. 1—Marcellus well locations in the study area [data source: DrillingInfo (2018), West Virginia Geological and Economic Survey (2018)]. Wells denoted with blue points have production data and are used in the analysis in this paper (red points do not have available data).

Criteria	Number of Wells
Total	15,990
Production data available	11,904
Passed filtering conditions	8,527
Horizontal and directional wells	5,637

Table 1—Number of wells available for each filtering stage.

Parameter Visualization. The distributions of parameters from fitting Duong's model are shown in **Fig. 2.** We found that parameters a and m followed normal distributions, while q_i and q_∞ appeared to be log-normal (the log-transform of these parameters followed normal distributions). As a side observation, if one included the vertical wells in the analysis, q_i and q_∞ would become bimodal, and the rest of the analysis would proceed by clustering the wells on the basis of these two variables and performing universal co-Kriging on each mode with an additional indicator variable included to predict which mode each block should resemble.

In **Fig. 3,** the correlation coefficient of the parameter pairs is shown in the upper triangle and the lower triangle shows the scatter-plot matrix for each parameter pair. We found good correlations between a and m (Fig. 3d), and separately between $\log(q_i)$ and $\log(q_\infty)$ (Fig. 3l), because they, respectively, displayed a linear trend and because the associated correlation-coefficient values (0.873 in Fig. 3a and 0.77 in Fig. 3i) were high (note that all the log transforms used in this analysis were natural-log transforms). This was an indicator that the value of one parameter can depend on other parameters, to some extent. As a result, co-Kriging should be used in further interpolation because predictions will borrow strength from these correlations in addition to the spatial correlations.

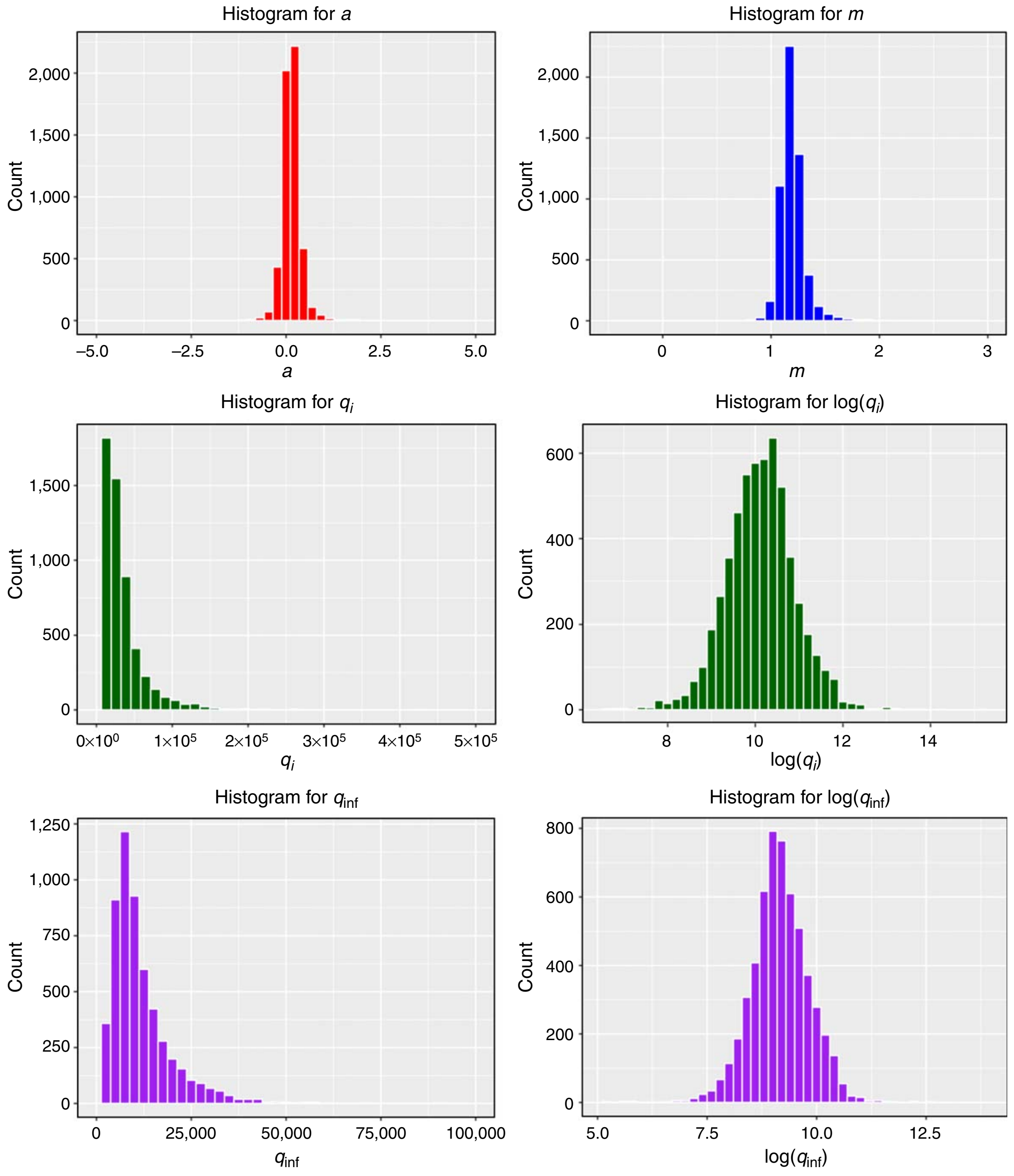

Fig. 2—Histograms of the values of Duong's decline-curve model parameters as fitted to the available well-production data.

External Drift. To ensure the stationarity of our data before geostatistical modeling, we fitted a background trend (external drift) for universal Kriging. Because we have r multiple response variables (in this case, the four decline-curve parameters) and p multiple explanatory variables [in this case, spatial coordinates x and y, which are the Universal Transverse Mercator (UTM) Eastings and Northings (both in km), respectively, plus an interaction term xy], we fitted a multivariate multiple linear regression model for the background trend:

$$Y = \beta Z + \varepsilon, \quad (4)$$

where Y is an $n \times r$ matrix (n being the number of observations, or wells, in this application) containing each response variable as a column, Z is an $n \times (p+1)$ matrix containing the explanatory variables plus a series of ones (for fitting an intercept) as columns, β is a $(p+1) \times r$ matrix containing the regression coefficients, and ε is an $n \times r$ matrix of error terms. The regression coefficient estimates, their p-values, and the adjusted R^2 values are reported in **Table 2.**

Table 2 shows that while the adjusted R^2 values were fairly low for all four models, the coefficients were generally significant (low p-values), although also generally low in magnitude. Nevertheless, these models had better adjusted goodness of fit (higher R^2) than simpler models with fewer terms, including models of just an intercept (or mean, which would be used for simple Kriging). This criterion for a higher adjusted R^2 was the reason for including the interaction term xy.

Fitting Linear Model of Coregionalization. Predictions can be strengthened by using the strong correlation between a and m and between q_i and q_∞. A linear model of coregionalization models not only the spatial autocorrelation of individual variables, but also the between-variable spatial dependence. Here, we fitted a Matern variogram model (Diggle and Ribeiro 2007) with nugget to all direct and cross variograms, as shown in **Fig. 4.**

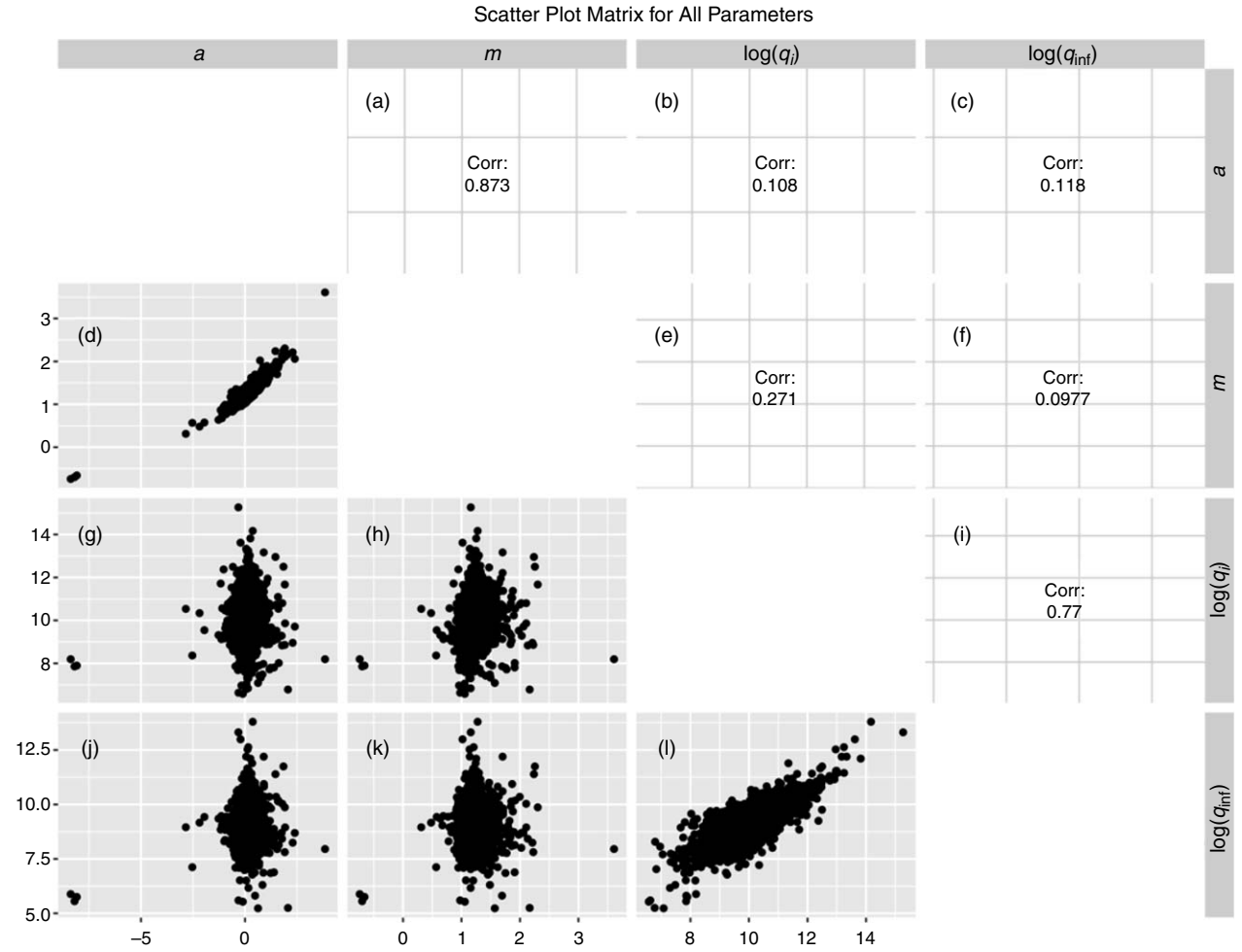

Fig. 3—Scatter plot for fitted decline-curve parameters (lower triangular part) and associated correlation-coefficient values (upper triangular part). Each scatter-plot panel corresponds to its correlation coefficient in a mirror position [e.g., panel (d) is associated with the value in panel (a)].

Response	Intercept	*x* Coefficient	*y* Coefficient	*xy* Coefficient	Adjusted R^2
a	-4.4×10^{0} (8.9×10^{-3})	7.0×10^{-3} (2.0×10^{-2})	-1.4×10^{-6} (2.6×10^{-2})	9.4×10^{-4} (1.1×10^{-2})	0.04
m	8.7×10^{-2} (9.0×10^{-1})	3.5×10^{-3} (4.9×10^{-3})	-6.9×10^{-7} (9.7×10^{-3})	1.9×10^{-4} (2.0×10^{-1})	0.04
$\ln(q_i)$	3.9×10^{1} (3.5×10^{-25})	3.8×10^{-3} (5.7×10^{-1})	4.4×10^{-7} (7.6×10^{-1})	-7.2×10^{-3} (5.3×10^{-19})	0.19
$\ln(q_\infty)$	3.9×10^{1} (4.0×10^{-33})	-1.6×10^{-2} (4.8×10^{-3})	4.5×10^{-6} (3.5×10^{-4})	-7.3×10^{-3} (1.5×10^{-24})	0.16

Table 2—Coefficient estimates and their *p*-values (in parentheses, lower values mean more statistical significance), along with the adjusted R^2 values, for the multivariate multiple linear regression.

Results and Discussion

Block Co-Kriging With External Drift. Figs. 5 and 6 show the mean and variance, respectively, of each decline-curve parameter. They were estimated as 5 × 5 km blocks, spanning the domain of the Marcellus. The noticeably low *a* and *m* predictions in the northern part of the Marcellus were controlled by two wells: API 37-105-21728 and 37-105-21722, which showed radically sharp decline after 3 months. One can also see that larger q_i and q_∞ values generally exist around southwest Pennsylvania and in the northeast, around Bradford County, Pennsylvania (Fig. 5). Fig. 6 shows small variances for all the parameters in proximity to the existing well locations, and increasingly large variances toward the fringe of the Marcellus, with the largest values being in the far southwest and northern extents.

EUR. The EUR at P10, P50, and P90 levels was estimated by drawing a random sample size of 1,000 ft from a multivariate normal distribution for *a, m,* $\ln(q_i)$, and $\ln(q_\infty)$. The mean values and covariance matrix were constructed from the universal co-Kriging-predicted values in each block. The samples of the log-transformed variables were back transformed by means of

$$q_i = \exp(\ln(q_i) + 0.5\{\mathrm{var}[\ln(q_i)]\}), \quad \ldots\ldots (5)$$

$$\text{and } q_\infty = \exp(\ln(q_\infty) + 0.5\{\mathrm{var}[\ln(q_\infty)]\}), \quad \ldots\ldots (6)$$

to correct for bias. Each sample was then used to calculate a 25-year EUR value (in Mscf):

$$EUR = \int_1^{25\times 12} q(t; a, m, q_i, q_\infty)\mathrm{d}t, \quad \ldots\ldots (7)$$

where $q(t; a, m, q_i, q_\infty)$ represents Duong's (2011) model and t is measured in months. The 10th, 50th, and 90th percentiles of the resulting EUR vector produced the P90, P50, and P10, respectively. Because we normalized the flow rates earlier, using perforation-interval length in units of 1,000 ft, the unit of our predicted EUR was Bcf per 1,000 ft. These are mapped in **Fig. 7,** where we observed that a larger EUR generally coincides with larger q_i and q_∞ (compare to Fig. 5).

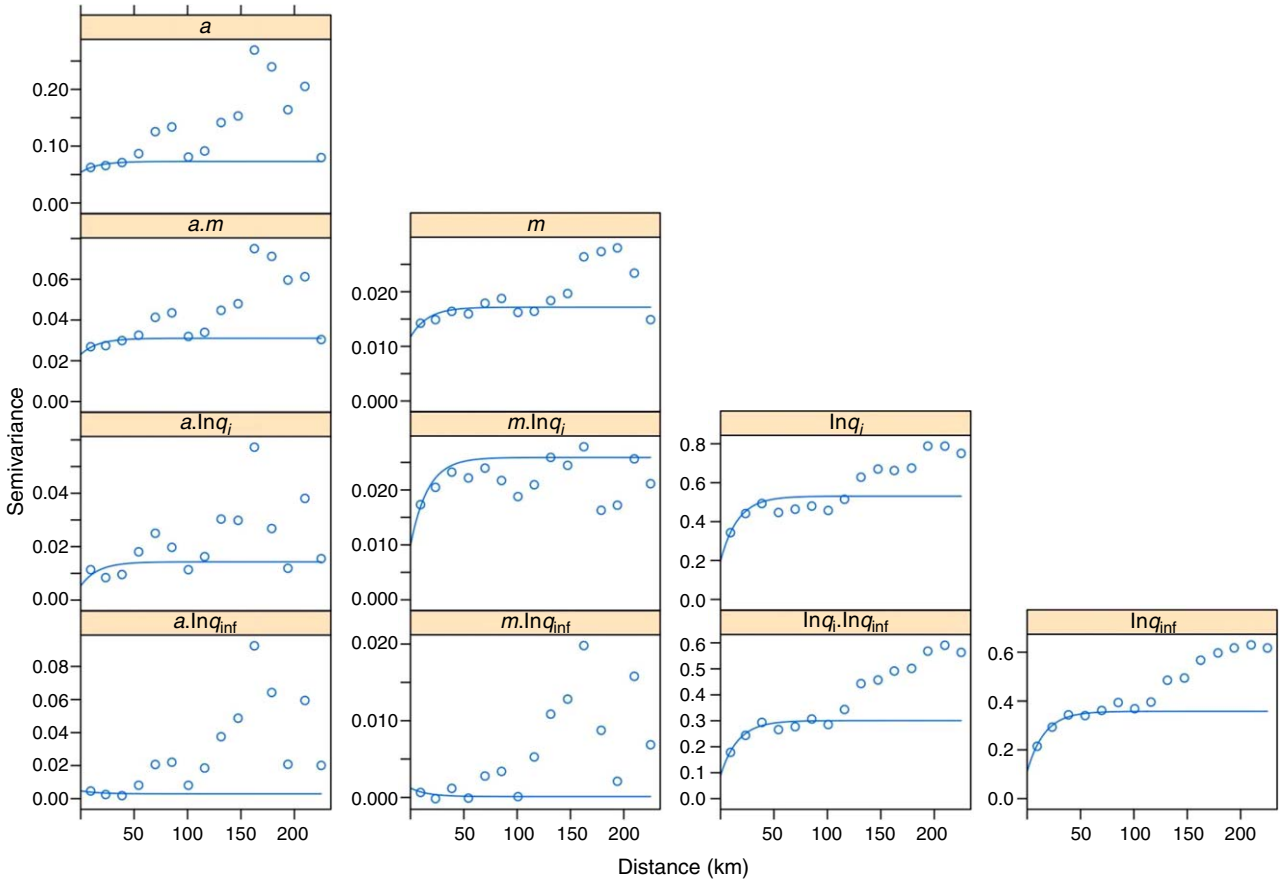

Fig. 4—Fitted direct variograms (on the diagonal) and cross variograms (off-diagonal). Points represent the binned sample semivariances, and the curves are the fitted Matern models (Diggle and Ribeiro 2007). The titles at the top of each panel indicate the decline-curve parameter (or combination of parameters, in the off-diagonal cross variograms) used for the associated semivariogram.

Male et al. (2016) predicted the EUR for wells with 25-year production in the Marcellus shale play and reported a median value of 3.9 Bcf. The median of our predicted EUR (P50) was 1.39 in Bcf/1,000 ft, and the median of the perforation-interval length in our data set was 3,842 ft. Therefore, we estimated our median EUR in Bcf by calculating their product, 5.34 Bcf, which was fairly close to 3.9 Bcf.

Validation. Leave-one-out cross validation offers a rational way to test the quality of the universal co-Kriging previously mentioned. In this validation scheme, a single well (with known decline-curve parameters) was removed from the population of wells, and the universal co-Kriging model was refit to the remaining data. Prediction using this new model was made at the single wellsite, and the predicted decline-curve parameters were compared against their known values at the site. In addition, the predictions of the decline parameters from the background trend (Eq. 4) were also recorded for comparison. We repeated this process for 100 randomly selected wells. This sample size of 100 was chosen as a balance between increasing statistical robustness of the analysis of errors and decreasing computational time (this process is computationally costly, with these 100 samples taking approximately 32 processor hours). Histograms of the resulting percent error of each parameter (difference between the true value and the predicted value divided by the true value) are presented in **Fig. 8.**

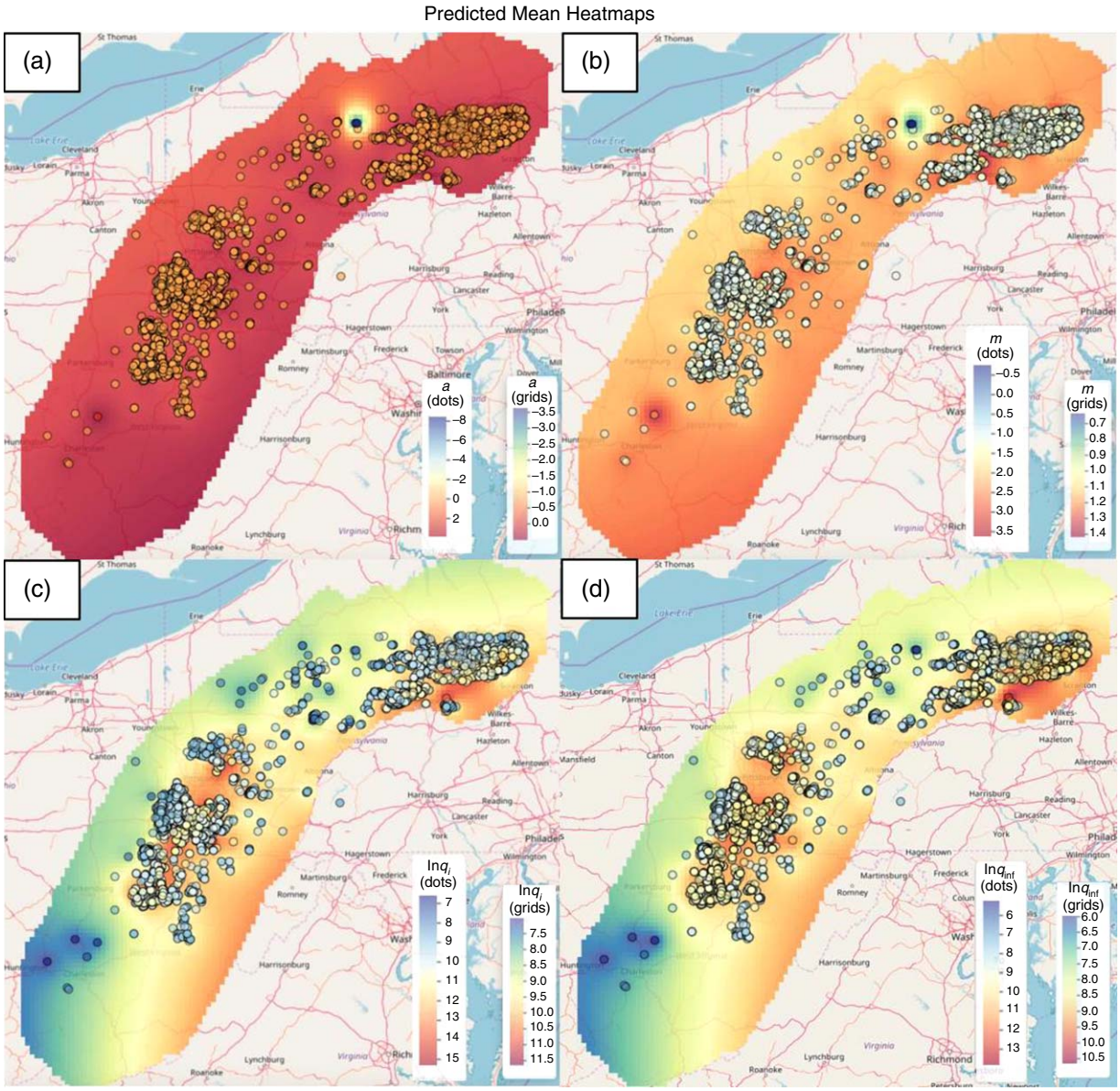

Fig. 5—Predicted mean for each type of parameter for Duong's (2011) decline-curve model: Panels (a), (b), (c), and (d) show the mean values for *a, m*, log(q_i), and log(q_{inf}), repectively (all the log transforms used in this analysis are natural-log transforms). Legends are provided in the lower right-hand corner of each panel; the left side represents dots or existing wellsites, and the right side represents grids or undeveloped sites. Well-location data were provided by West Virginia Geological and Economic Survey (2018) and DrillingInfo (2018).

Difference-of-mean hypothesis tests indicate whether the Kriged predictions are statistically different from the true values. Here, we ran one-sample *t*-tests on the percent error of each parameter to see if the mean percent error was significantly different from zero. The null hypothesis is that the mean percent error equals zero, and the alternative hypothesis is that it is not equal to zero. The results (estimate of mean percent error and associated *p*-value of the hypothesis test) are reported in **Table 3.** Because the hypothesis tests were conducted at the 95% confidence level, any *p*-values less than 0.05 indicate a statistically significant difference from zero mean (one can reject the null hypothesis in favor of the alternative). Table 3 shows that most of the *p*-values for Kriging were well above this 0.05 threshold, with ln(q_i) being the only exception, whereas all the *p*-values for the background trend were below this threshold. In other words, on average, the universal co-Kriging gives predictions that are not significantly different from the true values (except in the case of q_i), and has a much greater predictive accuracy than the background-linear-trend model alone, which gives predictions that differ strongly from the true values.

Besides these hypothesis tests, which relate to the central tendency and standard deviation of each histogram in Fig. 8, it is also worth discussing the extremes of the histograms. For most cases, the percent error rarely exceeded ±20%. However, for parameter *a*, we saw some outlier cases (four of them, for Kriging) where the percent error can exceed ±1,000% and increase as high as –3,800%. These outliers can be explained by the very small magnitude of the true *a* values for these four cases because they are the closest to zero. The true *a* value is in the denominator of the percent error calculation.

Two of the wells sampled for this leave-one-out cross validation were examined in more detail in **Fig. 9**—the well with the least overall percent error ("Best") and the well with the greatest overall percent error ("Worst"). The overall percent error was taken as the mean absolute percent error across the decline-curve parameters, except for *a* because of its outlier values, as previously discussed. Fig. 9 gives a sense of the range of discrepancy between the Kriged decline curves, the initial OLS decline-curve fits, and the cumulative production from the data. In these examples, at the end of the production data (53 months for "Best" and 63 months for "Worst"),

the percent error between the Kriged decline-curve prediction and the data was –5.72% for the "Best" case and 182% for the "Worst" case. Similarly, comparing the 25-year EUR from the Kriged decline curve and the OLS-fitted decline curve calculated –6.88% for the "Best" case and 223% for the "Worst" case. Extending this error analysis to all 100 wells in the leave-one-out cross validation produced mean percent errors of 14.0% when compared with the cumulative production from the data, and 14.1% when compared with the 25-year EURs, indicating that the risk of excessively large errors witnessed in the worst-case scenario was low.

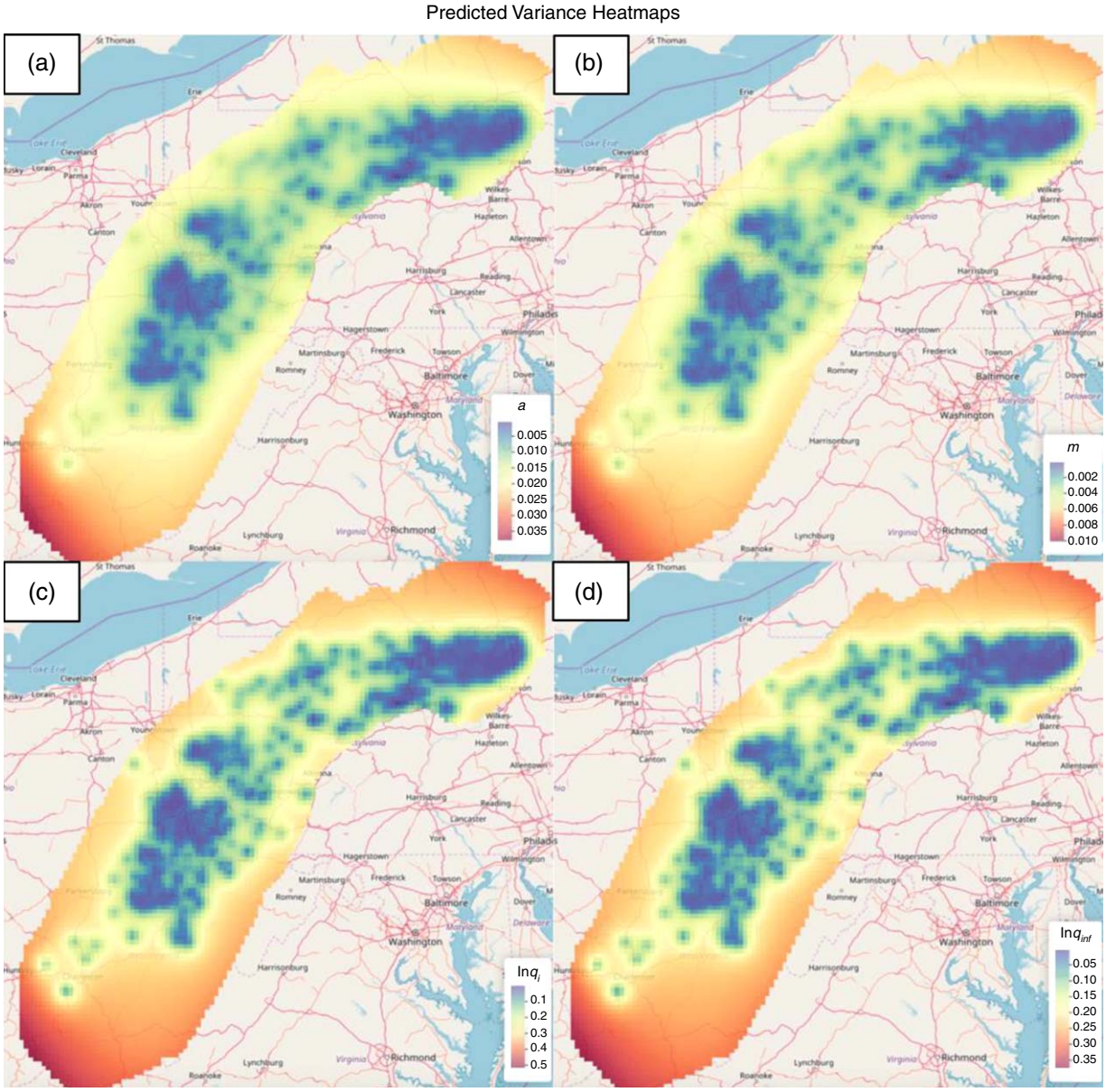

Fig. 6—Predicted variance for each type of parameter for Duong's (2011) decline-curve model: Panels (a), (b), (c), and (d) show the mean values for *a, m*, log(q_i), and log(q_{inf}), repectively (all the log transforms used in this analysis are natural-log transforms).

Geological Interpretation. Several authors (Wrightstone 2009; Erenpreiss et al. 2012; Wang and Carr 2013; EIA 2017) conducted geological mapping in the Marcellus Shale regarding its depth, thickness, and thermal maturity **(Figs. 10 and 11),** which helped interpret the P50 EUR heatmap constructed in our study. From a geological perspective on unconventional plays, high-hydrocarbon-generation potential and a thick source rock attribute to a high gas production (Zou et al. 2010; Mallick and Achalpurkar 2014). Also, thermal maturity was a key factor of the hydrocarbon-generation potential, and was correlated to formation depth because of the geothermal and pressure gradients (the Marcellus Formation top contour vs. thermal maturity zones in the left side of Fig. 10). In the overmature zone, and those within the gas-generation window (gas prone), an in-situ natural-gas-accumulation region formed (the gray and pink zone shown in the left side of Fig. 10) while a considerable amount of natural-gas reserves was unlikely to be found in the immature zone (the light-yellow area shown in the left side of Fig. 10). Hence, high gas production from the wells located in the middle of the play resulted from the source rock that tended to generate more gas than oil. The wells in the northeast produce a significant amount of dry gas from overmatured shale. Similarly, abundant yields of dry gas were found along the southwest boundary of the play despite the few number of wells that have been drilled there so far. In addition, the thicker the source rock, the more hydrocarbon-productive kerogens it might bear. Thus, the shale thickness was another critical factor to consider when assessing the production potential. In

Fig. 11, the P50 EUR heatmap (right side) revealed that the high-yield areas previously mentioned (middle, northeast, southwest) always show up in the bright-green and pink zones in the isopach map (Fig. 11, left side) where the associated Marcellus Shale gross thickness is considerable.

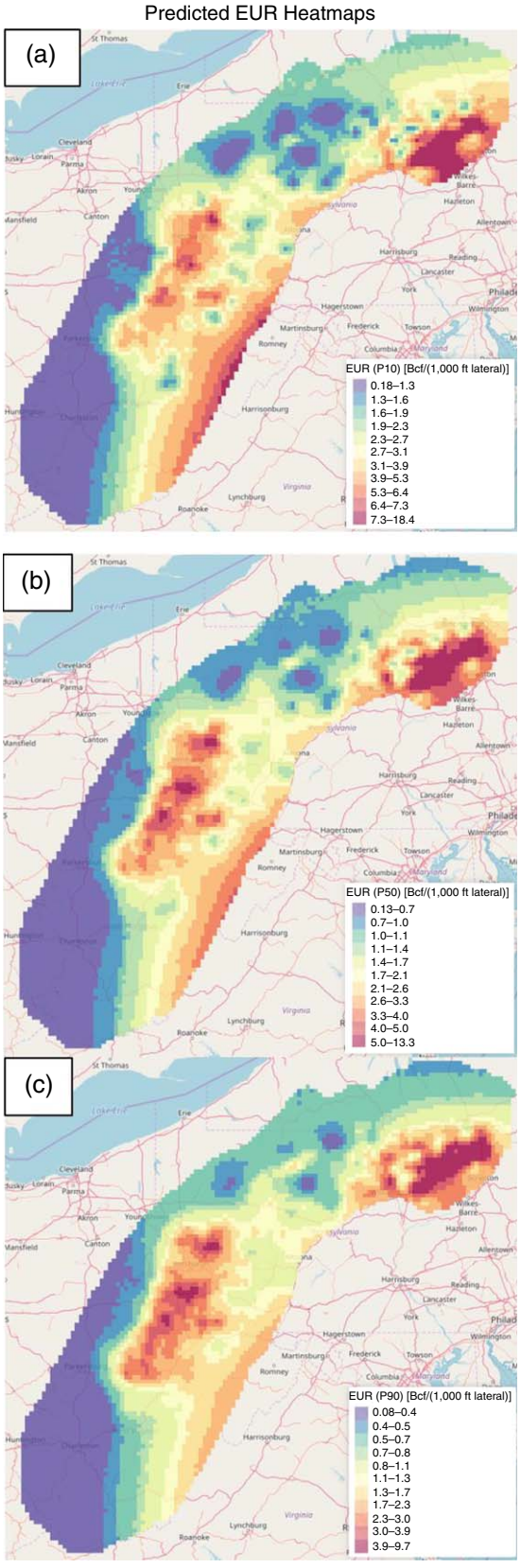

Fig. 7—Predicted EUR heatmaps across the Marcellus Shale. Panels (a), (b), and (c) stand for P10, P50, and P90 EUR, respectively. The units are Bcf per 1,000 ft of interval length.

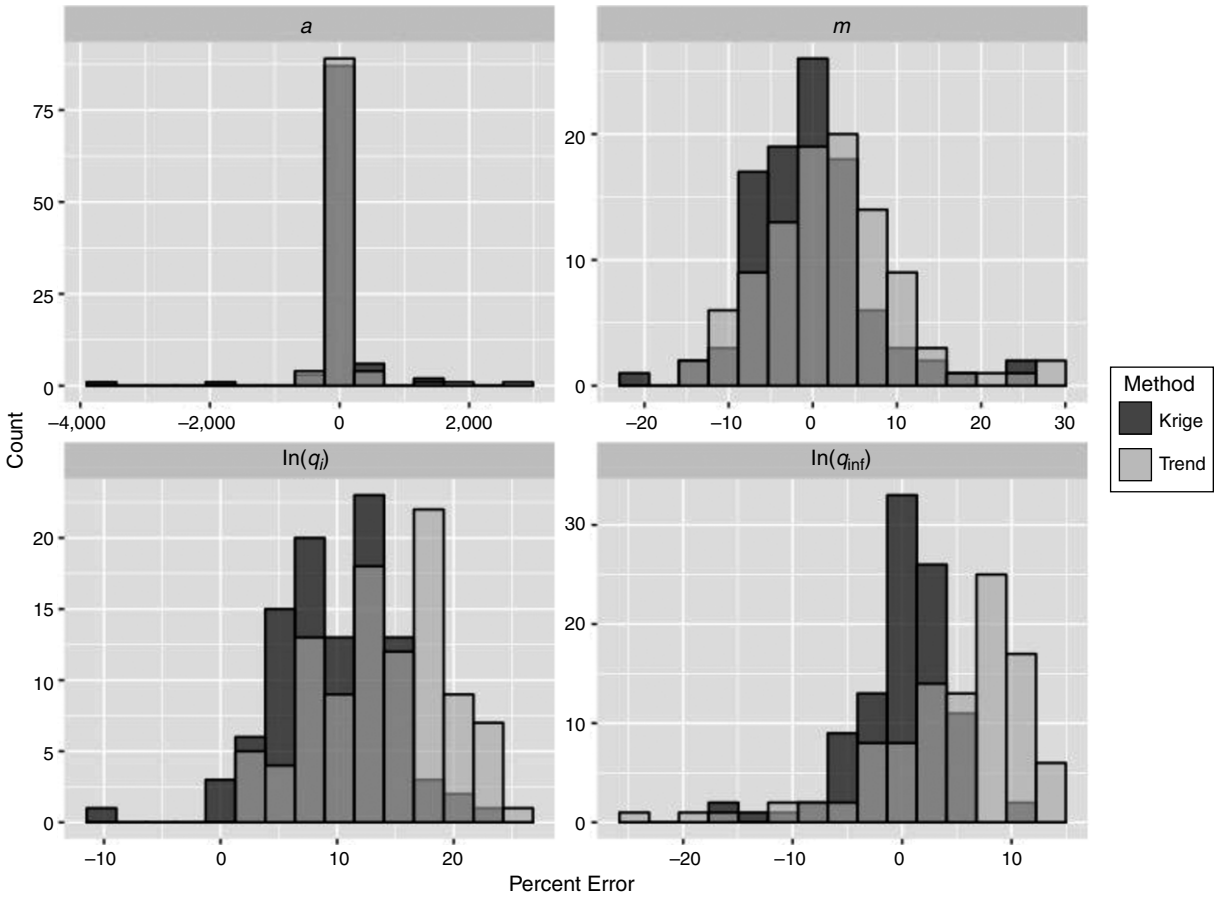

Fig. 8—Histograms of percent error for each decline-curve parameter and each method—universal co-Kriging (dark gray) and background trend (light gray). Note that the light grey bars (for Trend) are transparent so as not to hide the dark grey bars, thus producing the medium grey color. The decline-curve parameter is indicated at the top of each frame. Note that the *x*- and *y*-axis scales vary from frame to frame.

	Kriged		Trend	
Parameter	Mean	*p*-value	Mean	*p*-value
a	–3.62	0.94	107.37	0.003
m	–0.44	0.53	2.32	0.006
$\ln(q_i)$	9.79	7.56×10^{-36}	13.97	5.63×10^{-45}
$\ln(q_\infty)$	–0.10	0.83	4.44	6.35×10^{-9}

Table 3—One-sample *t*-test results for each decline-curve parameter and for the universal co-Kriging predictions and the background trend predictions.

As a result, the high-EUR zones in our results were located in the northeast, middle, and southeast part in our mapping area, sharing similar locations with high thermal maturity and great thickness (see comparisons in Figs. 10 and 11). In other words, our production forecasts and reserves estimates were in agreement with the associated geological evidence. Besides qualitatively validating the heat-map with geological variables, it was beneficial to incorporate them as additional inputs in co-Kriging to strengthen the predictions of the decline-curve parameters for undrilled wells, taking advantage of the spatial correlations among geological variables. Specifically, we might add the shale thickness, vitrinite reflectance ($\%R_o$), or other factors at operating wellsites as one of the dependent variables in co-Kriging interpolation, and use maps of such properties away from the wells to increase the accuracy of, and reduce uncertainty in, our decline-curve parameter predictions.

Website Application. In order to increase the transparency of this work and make the findings available in a comprehensive format, we created a website application (Xi and Morgan 2018). This utility allows one to click anywhere in the Marcellus Shale to view the predicted decline curves (P10, P50, P90) along with specific EUR values for the selected block. Furthermore, one can view the fitted decline-curve parameter values (as well as goodness of fit) at each well in geographical format.

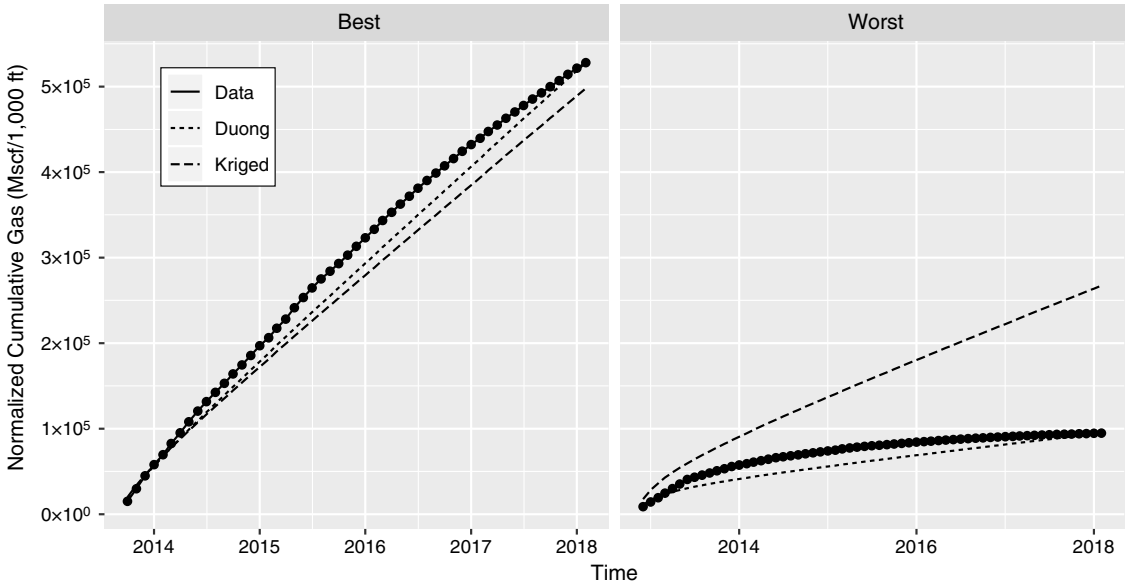

Fig. 9—Two examples of Kriging-predicted decline curves (dashed lines) plotted alongside their associated well-production data (solid lines with points) and OLS-fitted Duong's model (dotted lines). The "Best" case represented the well from the leave-one-out cross validation with the least amount of overall error between the Kriged decline-curve parameters and the initial OLS parameter estimates for Duong's model. The "Worst" case represented the greatest overall error. In both cases, the initial OLS parameter estimates were withheld from the geostatistical model fitting, so the model predictions were completely uninformed by the data at these wells.

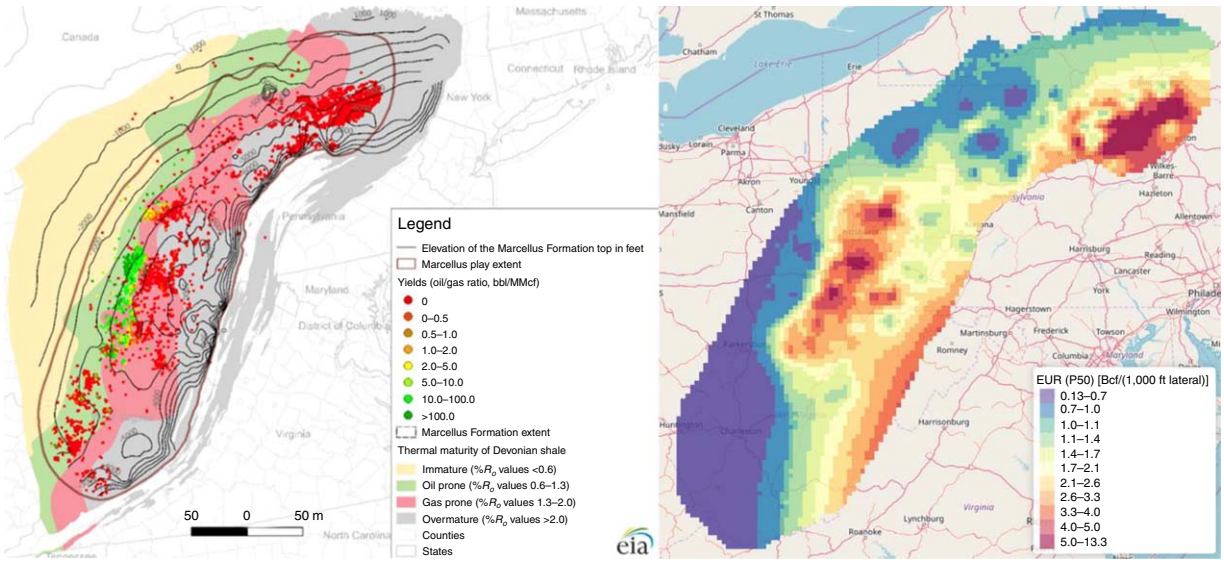

Fig. 10—Thermal maturity and initial yields in the Marcellus Shale are shown on the left side: The light-yellow area represents the immature zone, the green zone denotes oil prone, the pink zone denotes gas prone, and the gray zone denotes the overmature zone. On the right side, the heatmap of the predicted P50 EUR indicates that high-production areas (orange red) tend to be located within the gas-prone zone and the overmature zone, where gas is the major product from the source rock (e.g., the northeastern high-EUR area exists mainly within the overmature zone). The left side of the figure was sourced from the US Energy Information Administration (2017) on the basis of DrillingInfo (2018) and the US Geological Survey (2018).

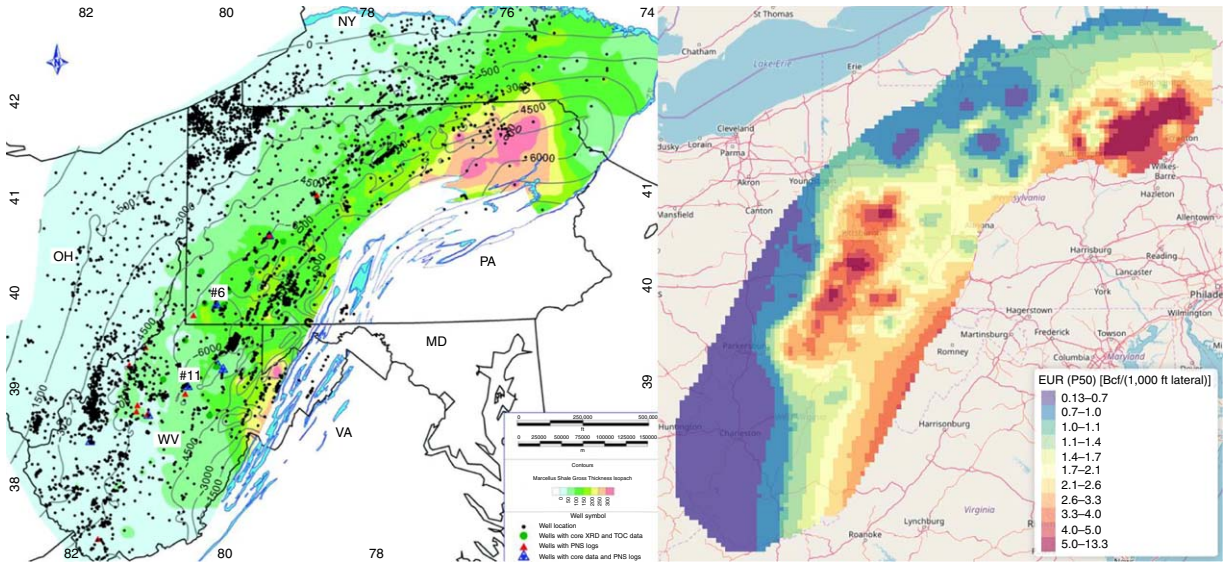

Fig. 11—Color-filled Marcellus isopach map (Wang and Carr 2013) is shown on the left side, where the formation thickness increases as the color varies from light green to bright green to pink. On the right side, the heatmap of the predicted P50 EUR indicates that high-production areas (orange red) show up in locations with considerable shale thickness (see the isopach map). These areas include the northeast, middle, and southeast boundary of the entire shale play.

Conclusions

The approach we proposed in this paper is a combination of DCA and geostatistical methods. First, the parameters that describe the natural-gas-production variation were obtained by decline-curve fitting. According to the statistics of these parameters, we chose a proper geostatistical method (universal co-Kriging) to interpolate the parameters for undrilled sites. Finally, we estimated the EUR and associated decline curve with predicted parameters while quantifying the uncertainty at the same time. Without requesting costly types of data, our approach implemented and provided operators with a good reference for early-stage explorations.

The median value of our predicted EUR (P50) was close to that reported by other authors (Male et al. 2016). The heatmaps generated in our case study showed that the promising sites with high EUR are mainly located in the middle and northeast of the study area, which was consistent with other geological evidence in the Marcellus. Although some areas along the southeast boundary of the EUR heatmap also appeared to be sweet spots, the reliability of their EUR might suffer from the relatively high variance of the associated decline-curve parameters. Furthermore, the validation not only indicated the superiority of the universal co-Kriging over the background-trend prediction, but also showed reasonable percent errors in predicted parameters despite the "magnitude" issue for parameter a.

Last, some extensions and modifications might be applied in our future work, such as adopting a modified Duong's (2011) model (Joshi and Lee 2013) to avoid potentially unrealistic forecasts or using a pattern-recognition method (unsupervised learning) to further narrow the range of promising areas on the basis of our current analysis.

Nomenclature

a = intercept constant defined in Duong's model
b = hyperbolic decline constant
EUR = estimated ultimate recovery, Mscf
G_p = cumulative gas production, Mscf
m = slope defined in Duong's model
n = number of observations, or wells in this research
p = number of different explanatory variables
q = gas rate, Mscf/month
q_∞ = gas rate at infinite time, Mscf/month
q_i = initial gas rate, Mscf/month
r = number of different response variables
t = time in accordance with gas rate, month
t_{am} = one parameter in Duong's model
x = Eastings in UTM, km
y = Northings in UTM, km
Y = matrix containing response variables from all observations
Z = matrix containing explanatory variables plus a series of ones
β = matrix containing regression coefficients
ε = matrix containing error terms

Acknowledgments

This work was made possible by the Deike Research Grant, College of Earth and Mineral Sciences, Pennsylvania State University. The authors thank DrillingInfo for the donated academic license.

References

Anderson, D. M. and Liang, P. 2011. Quantifying Uncertainty in Rate Transient Analysis for Unconventional Gas Reservoirs. Presented at the North American Unconventional Gas Conference and Exhibition, The Woodlands, Texas, 14–16 June. SPE-145088-MS. https://doi.org/10.2118/145088-MS.

Arps, J. J. 1945. Analysis of Decline Curves. *Trans Metall Soc AIME* **160:** 228–247. https://doi.org/10.2118/945228-G.

Bhattacharya, S. and Nikolaou, M. 2013. Analysis of Production History for Unconventional Gas Reservoirs With Statistical Methods. *SPE J.* **18** (5): 878–896. SPE-147658-PA. https://doi.org/10.2118/147658-PA.

Cipolla, C. L., Lolon, E. P., Erdle, J. C. et al. 2010. Reservoir Modeling in Shale-Gas Reservoirs. *SPE Res Eval & Eng* **13** (4): 638–653. SPE-125530-PA. https://doi.org/10.2118/125530-PA.

Clarkson, C. R., Jensen, J. L., and Chipperfield, S. 2012. Unconventional Gas Reservoir Evaluation: What Do We Have To Consider? *J Nat Gas Sci Eng* **8:** 9–33. https://doi.org/10.1016/j.jngse.2012.01.001.

Diaz De Souza, O. C., Sharp, A., Martinez, R. C. et al. 2012. Integrated Unconventional Shale Gas Reservoir Modeling: A Worked Example From the Haynesville Shale, De Soto Parish, North Louisiana. Presented at the SPE Americas Unconventional Resources Conference, Pittsburgh, Pennsylvania, 5–7 June. SPE-154692-MS. https://doi.org/10.2118/154692-MS.

Diggle, P. and Ribeiro, P. J. 2007. *Model-Based Geostatistics*. New York: Springer-Verlag. https://doi.org/10.1007/978-0-387-48536-2.

DrillingInfo. 2018. DrillingInfo. https://info.drillinginfo.com (accessed 2 May 2018).

Duong, A. N. 2011. Rate-Decline Analysis for Fracture-Dominated Shale Reservoirs. *SPE Res Eval & Eng* **14** (3): 377–387. SPE-137748-PA. https://doi.org/10.2118/137748-PA.

EIA (Energy Information Administration). 2017. *Marcellus Shale Play: Geology Review*, Washington, DC: US Department of Energy, https://www.eia.gov/maps/pdf/MarcellusPlayUpdate_Jan2017.pdf (accessed 15 August 2018).

Erenpreiss, M., Wickstrom, L., Riley, R. A. et al. 2012. Mapping the Regional Organic Thickness of the "Marcellus Shale" Hamilton Group. Presented at the Geological Society of America 46th Annual Meeting, Dayton, Ohio, 23–24 April.

Erdle, J., Hale, B., Houze, O. et al. 2016. *Monograph 4: Estimating Ultimate Recovery of Developed Wells in Low-Permeability Reservoirs*. Houston: Society of Petroleum Evaluation Engineers.

Hauge, V. L. and Hermansen, G. H. 2017. Machine Learning Methods for Sweet Spot Detection: A Case Study. In *Geostatistics Valencia 2016*, ed. J. J. Gómez-Hernández, J. Rodrigo-Ilarri, M. E. Rodrigo-Clavero, E. Cassiraga, and J. A. Vargas-Guzmán, Chap. 19, 573–588. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-46819-8_38.

Joshi, K. and Lee, J. 2013. Comparison of Various Deterministic Forecasting Techniques in Shale Gas Reservoirs. Presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, 4–6 February. SPE-163870-MS. https://doi.org/10.2118/163870-MS.

Journel, A. G. and Huijbregts, C. J. 1978. *Mining Geostatistics*. London: Academic Press.

Ketineni, S. P., Ertekin, T., Anbarci, K. et al. 2015. Structuring an Integrative Approach for Field Development Planning Using Artificial Intelligence and Its Application to an Offshore Oilfield. Presented at the SPE Annual Technical Conference and Exhibition, Houston, 28–30 September. SPE-174871-MS. https://doi.org/10.2118/174871-MS.

Male, F., Marder, M. P., Browning, J. et al. 2016. Marcellus Wells' Ultimate Production Accurately Predicted From Initial Production. Presented at the SPE Low Perm Symposium, Denver, 5–6 May. SPE-180234-MS. https://doi.org/10.2118/180234-MS.

Mallick, M. and Achalpurkar, M. P. 2014. Factors Controlling Shale Gas Production: Geological Perspective. Presented at the Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, 10–13 November. SPE-171823-MS. https://doi.org/10.2118/171823-MS.

Matheron, G. 1971. *The Theory of Regionalised Variables and Its Applications*. Paris: Les Cahiers du Centre de Morphologie Mathématique.

McBratney, A. B. and Webster, R. 1983. Optimal Interpolation and Isarithmic Mapping of Soil Properties: V. Co-Regionalization and Multiple Sampling Strategy. *Eur J Soil Sci* **34** (1): 137–162. https://doi.org/10.1111/j.1365-2389.1983.tb00820.x.

Moinfar, A., Varavei, A., Sepehrnoori, K. et al. 2013. Development of a Coupled Dual Continuum and Discrete Fracture Model for the Simulation of Unconventional Reservoirs. Presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, 18–20 February. SPE-163647-MS. https://doi.org/10.2118/163647-MS.

Oliver, D. S. and Chen, Y. 2011. Recent Progress on Reservoir History Matching: A Review. *Computat Geosci* **15** (1): 185–221. https://doi.org/10.1007/s10596-010-9194-2.

Trangmar, B. B., Yost, R. S., and Uehara, G. 1986. Application of Geostatistics to Spatial Studies of Soil Properties. *Adv Agron* **38:** 45–94. https://doi.org/10.1016/S0065-2113(08)60673-2.

Vauclin, M., Vieira, S. R., Vachaud, G. et al. 1983. The Use of Cokriging With Limited Field Soil Observations 1. *Soil Sci Soc Am J* **47** (2): 175–184. https://doi.org/10.2136/sssaj1983.03615995004700020001x.

Wang, G. and Carr, T. R. 2013. Organic-Rich Marcellus Shale Lithofacies Modeling and Distribution Pattern Analysis in the Appalachian Basin. *AAPG Bull* **97** (12): 2173–2205. https://doi.org/10.1306/05141312135.

West Virginia Geologic and Economic Survey (WVGES). 2018. WVGES Marcellus Wells. http://www.wvgs.wvnet.edu/www/datastat/Marcellus/Downloads/WVGES%20Marcellus%20Wells.xlsx (accessed 7 May 2018).

Wrightstone, G. R. 2009. Marcellus Shale—Geologic Controls on Production. *Search Discov* **10206:** 1–10.

Xi, Z. and Morgan, E. 2018. Virtual Asset 1.0: Marcellus Shale. https://shinysrv.ems.psu.edu/eum19/Virtual_Asset_1_0/ (accessed 13 May 2019).

Zou, C., Dong, D., Wang, S. et al. 2010. Geological Characteristics and Resource Potential of Shale Gas in China. *Pet Explor Dev* **37** (6): 641–653. https://doi.org/10.1016/S1876-3804(11)60001-3.

Zhenke Xi is an MS degree candidate in petroleum and natural gas engineering at Pennsylvania State University. He holds a BE degree in resources exploration engineering at Jilin University, China. Zhenke's research interests include data analytics and software development for unconventional gas production and reservoir simulation. He is a member of SPE.

Eugene Morgan is an assistant professor of petroleum engineering at Pennsylvania State University. His current research interests include big data analytics for the oil and gas industry, as well as stochastic inversion techniques for seismic-survey data. Morgan holds PhD and MS degrees in civil and environmental engineering from Tufts University. He is a member of SPE.