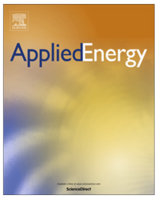
Applied Energy 195 (2017) 344–355

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy



Spatial variability of tight oil well productivity and the impact of technology

J.B. Montgomery [a,b,*], F.M. O'Sullivan [a]

[a] MIT Energy Initiative, Massachusetts Institute of Technology, United States
[b] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, United States

HIGHLIGHTS

- Forecasts of tight oil production assume rates of technological improvement.
- A spatial error model and regression-kriging are proposed for estimating this.
- These are more accurate than existing methods, which overestimate technology's role.
- Productivity has been equally impacted by changes in technology and location.

ARTICLE INFO

Article history:
Received 22 July 2016
Received in revised form 1 March 2017
Accepted 9 March 2017
Available online 21 March 2017

Keywords:
Hydraulic fracturing
Unconventional oil and gas
Bakken and Three Forks Shale
Spatial autocorrelation
Linear regression
Kriging

ABSTRACT

New well productivity levels have increased steadily across the major shale gas and tight oil basins of North America since large-scale development began a decade ago. These gains have come about through a combination of improved well and hydraulic fracturing design, and a greater concentration of drilling activity in higher quality acreage, the so called "sweets spots." Accurate assessment of the future potential of shale and tight resources depends on properly disentangling the influence of technology from that of well location and the associated geology, but this remains a challenge. This paper describes how regression analysis of the impact of design choices on well productivity can yield highly erroneous estimates if spatial dependence is not controlled for at a sufficiently high resolution. Two regression approaches, the spatial error model and regression-kriging, are advanced as appropriate methods and compared to simpler but widely used regression models with limited spatial fidelity. A case study in which these methods are applied to a large contemporary well dataset from the Williston Basin in North Dakota reveals that only about half of the improvement in well productivity is associated with technology changes, but the simpler regression models substantially overestimate the impact of technology by attributing location-driven improvement to design changes. Because of the widespread reliance on these less spatially resolved regression models, including by the U.S. Energy Information Administration to project shale gas and tight oil resource potential, the overestimate of technology's role in well productivity has important implications for future resource availability and economics, and the development choices of individual operators.

1. Introduction

Oil and gas produced from shale and tight rock formations is playing an increasingly important role in global and domestic energy markets. Due to increased production of oil from North Dakota, Texas, and other states, the United States is now considered by some to be the world's "swing producer," supplanting OPEC in this traditional oil market balancing role [1]. In North Dakota, which includes the most active part of the Williston tight oil basin, crude oil production grew from 98 thousand barrels a

---

Abbreviations: CV, cross-validation; FE, fixed effects; IQR, interquartile range; STA, surface trend analysis; EIA, Energy Information Administration (U.S.); kft, thousand feet; MASE, mean absolute scaled error; Mbbl, thousand barrels; Mbbl/d, thousand barrels a day; MMgal, million gallons; MMlb, million pounds; NS, nonspatial; RK, regression-kriging; SEM, spatial error model.

[*] Corresponding author at: MIT Energy Initiative, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States.

E-mail addresses: jbm@mit.edu (J.B. Montgomery), frankie@mit.edu (F.M. O'Sullivan).

http://dx.doi.org/10.1016/j.apenergy.2017.03.038
0306-2619/© 2017 Elsevier Ltd. All rights reserved.

day (Mbbl/d) in 2005 to 1174 Mbbl/d a decade later [2]. Additionally, the U.S. power sector has drastically increased its reliance on domestically produced natural gas, especially from shale [3]. Although these formations have long been known to contain abundant oil and gas, the "tightness," or low permeability, of the rock led many to view production from them as not economically viable [4]. However, commercial rates of production turned out to be possible using long horizontally-drilled wells combined with hydraulic fracturing—in which fluid and sand is pumped into wells to break apart rock and create pathways for fluid flow—and this has led to the rapid expansion of shale gas and tight oil production in the past decade [5–7].

In recent years there has been a sustained downturn in oil and gas prices, leading to substantial uncertainty about future levels of production from shale gas and tight oil formations [8]. The future outlook for these resources now depends largely on the capacity of industry to improve the economics of extraction through higher productivity. Thus far there have been signs of this happening, with production per drilling rig increasing as the number of active drilling rigs has fallen precipitously, as shown for the Williston basin in Fig. 1(a) [9]. Although some of this trend can be attributed to more efficient drilling, much of it is driven by a rise in average new well productivity (Fig. 1(b)) [7,10,11].

There are two important factors to recognize behind increases in well productivity. First, oil and gas operating companies have been scaling up well designs in an effort to increase well production through greater reservoir access. There has been a shift toward longer lateral lengths—which tends to mean a greater number of hydraulic fracturing "stages" at which fractures are initiated from—and larger volumes of both the water-based fracturing fluid pumped to create fractures and the sand, or proppant, carried by this fluid in order to keep the fractures "propped" open after water has flowed back [7,12–15]. In addition to this, operating companies have been "high-grading," or focusing their drilling efforts on the locations in a field with the most favorable geology and highest expected production [14,16–19]. Technology-driven improvements to productivity may be transferable to future wells in all parts of a field but high-grading amounts to simply exploiting the lowest-cost resource first. To understand changes in resource economics and realistically forecast future production it is therefore critical to be able to distinguish accurately between the influence of location and technology choices on well productivity [20,21].

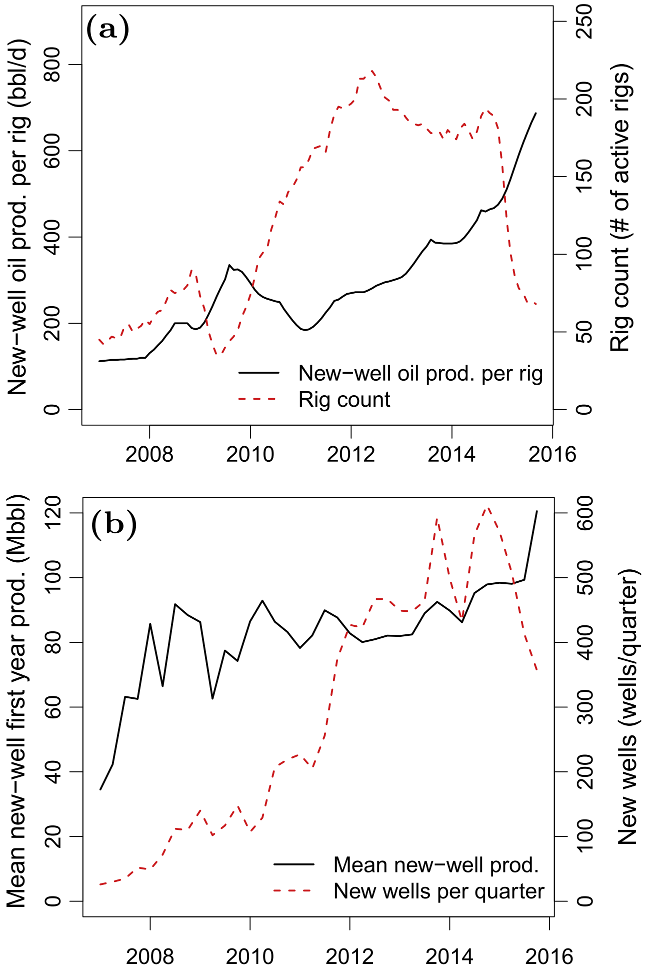


Fig. 1. Two perspectives on productivity in Williston tight oil basin. (a) Productivity of drilling rigs, measured as production per active rig (Source: U. S. Energy Information Administration [9]). (b) Productivity of new wells, measured as mean first year production of new wells in each fiscal quarter.

Multivariate statistical analysis remains an important approach to understanding the role that technology choices have played on well productivity. There are large datasets of production and engineering data available, due to the large number of wells that have already been drilled in these formations [22–24]. Additionally, there are limitations to physics-based modeling approaches due to frequently inadequate well-level geological data and the challenges of simulating fracture propagation and complex flow behavior in low permeability rock [25,26]. As a result, multivariate regression modeling has been widely adopted to infer the impact of technology on tight oil and shale gas well productivity [3,23,27–35].

An important modeling challenge associated with this is how to control for location, since reservoir quality, and hence well productivity, is spatially dependent. Some authors have chosen to simply ignore this feature and use *nonspatial* models, but this makes it unclear how reliable their results are [27–30]. At the other end of the spectrum, location or functions of location can be included as independent variables in a regression model, using *surface trend analysis* [34]. Another approach to control for location lies in between these, and assumes geological homogeneity within a small sample of wells [31–33], or within *fixed effects* regions [3,23,34,35]. For example, the U.S. Energy Information Administration (EIA) assumes county-level fixed effects, in which the difference in each well's productivity from the mean in its county is attributed to the influence of technology.

Implicit in all of these approaches is an assumption that spatial variability can be neglected below some arbitrary scale and this assumption will not overly influence results. However, important properties in shale and tight reservoirs have been found to vary considerably over even relatively short distances [26]. Furthermore, the tendency of operating companies to high-grade drilling activity alongside the scaling up of technology parameters creates a risk of conflating these impacts and potentially under- or overestimating the amount of technological improvement actually made. No study has specifically considered the potential of different controls for location to influence inference results and it is difficult to compare estimates between studies since different datasets and assumptions have been used. Studies that have adopted some controls for location have generally concluded that differences in well location play an important role on well productivity, but a lack of robust controls for location has made it difficult to quantify this relationship in the past [33–37].

In other domains with spatially dependent data, such as ecology, soil science, and urban energy consumption, regression-kriging and spatial error models have been used to explicitly incorporate spatial autocorrelation, or the spatial clustering of similar observations, into estimates [38–41]. These approaches have not