SPi	Journal Code:				Article ID				Dispatch: 10.01.18	CE: Alice Sucaldito
	A	S	M	B	2	3	0	6	No. of Pages: 19	ME:

Received: 21 April 2017 | Revised: 28 November 2017 | Accepted: 14 December 2017

DOI: 10.1002/asmb.2306

RESEARCH ARTICLE

WILEY

The Hubbert diffusion process: Estimation via simulated annealing and variable neighborhood search procedures—application to forecasting peak oil production

Istoni da Luz Sant'Ana[1] | Patricia Román Román[2] | Francisco Torres-Ruiz[2]

[1]Department of Mathematics, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

[2]Department of Statistics and Operations Research, University of Granada, Granada, Spain

Correspondence
Francisco Torres-Ruiz, Department of Statistics and Operations Research, Faculty of Sciences, University of Granada, 18071 Granada, Spain.
Email: fdeasis@ugr.es

Funding information
Ministerio de Economía y Competitividad, Spain, Grant/Award Number: MTM2014-58061-P; CNPQ, Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil

Accurately charting the progress of oil production is a problem of great current interest. Oil production is widely known to be cyclical: in any given system, after it reaches its peak, a decline will begin. With this in mind, Marion King Hubbert developed his peak theory in 1956 based on the bell-shaped curve that bears his name. In the present work, we consider a stochastic model based on the theory of diffusion processes and associated with the Hubbert curve. The problem of the maximum likelihood estimation of the parameters for this process is also considered. Since a complex system of equations appears, with a solution that cannot be guaranteed by classical numerical procedures, we suggest the use of metaheuristic optimization algorithms such as simulated annealing and variable neighborhood search. Some strategies are suggested for bounding the space of solutions, and a description is provided for the application of the algorithms selected. In the case of the variable neighborhood search algorithm, a hybrid method is proposed in which it is combined with simulated annealing. In order to validate the theory developed here, we also carry out some studies based on simulated data and consider 2 real crude oil production scenarios from Norway and Kazakhstan.

KEYWORDS
diffusion processes, Hubbert curve, oil production model, simulated annealing, variable neighborhood search

1 | INTRODUCTION

For several decades now, the forecasting of oil production has been a problem of great interest given its fundamental role in the world's economy. In fact, the growth rate of the world economy is directly linked to oil consumption.

Given that petroleum is a nonrenewable and finite resource, it is imperative that we are able to forecast future oil production and to predict the precise time at which production will peak. According to the concept of peak oil, in any given country (and in the world as a whole), oil production rates will eventually reach, or may have already reached, a maximum. After that point, production rates will start to decline.

Given the undeniable fact that oil fuels the world's economy, reaching production peak has unavoidable implications for economic growth. Some possible consequences are a slowdown of economic growth, the need to resort to more efficient energy usage, and the development of alternative energy sources among others. Many studies have dealt with the consequences of peak oil, and some examples of these are the archives of the Association for the Study of Peak Oil[1] or

the so-called Hirsch Report, which provides a good overview on the issue of peak oil, its implications, and its possible mitigation.[2] Curtis[3] provided an analysis of policy responses to climate change and peak oil, whereas Lutz et al[4] analyzed the economic effects of peak oil.

The pioneer of historical evaluations on crude oil depletion was the geologist Marion King Hubbert, who, in 1956, correctly estimated that oil production in the USA would peak around 1970. In recent years, many authors have been stimulated by the economic and political implications of energy problems and have turned their attention to Hubbert's theory in an attempt to apply his analysis to other countries and thus forecast the evolution of world oil production. Some experts conclude that production has already peaked, eg, the works of Bakhtiari[5] and Deffeyes,[6] whereas others argue that the peak will occur soon; concretely, in the work of Towler,[7] it was concluded that the peak is not likely to occur before 2018, and that this deadline will be further extended by rising oil prices and technology developments.

The debate concerning oil depletion has become broader with an abundance of analyses and predictions about the date of peak oil. Many of these models refer to Hubbert's approach and try to extend and update it. We must note that, although the peak theory pioneered by Hubbert was implemented in the oil production context, the related depletion analysis can also be applied to other nonrenewable resources such as natural gas,[8] phosphorus,[9] and lithium.[10]

This paper considers a stochastic process related to the Hubbert curve to deal with oil production, and it is structured as follows. Section 2 is devoted to a brief review of some of the models proposed so far for modeling oil production and estimating peak and peak time. Section 3 describes how the Hubbert process is obtained from a general expression and how the reparametrization of the logistic curve is carried out. Section 4 deals with how the parameters of the model are estimated using maximum likelihood and with the subsequent estimation of peak and peak time. The complexity of the likelihood system of equations leads to a direct estimation by maximizing the likelihood function. Section 5 deals with this problem through the application of simulated annealing (SA) and variable neighborhood search (VNS) algorithms. First, a brief summary of the algorithms is provided, and then their adaptation to the problem at hand is presented. Some strategies are suggested for bounding the space of solutions, and a description is provided for the application of the algorithms selected. In the case of the VNS algorithm, a hybrid method is proposed in which it is combined with SA. In order to validate the methodology described, Section 6 describes a simulation study. Finally, in Section 7, we present 2 applications to real data by considering crude oil production data from Norway and Kazakhstan. These examples show the possibilities that the process affords for the modelization of oil production and help answer the question of when peak production can be expected.

2 | BRIEF SUMMARY OF OIL PRODUCTION MODELS

Marion King Hubbert was the first researcher who developed a theory for the study of oil production. In 1956, he applied his theory to crude oil production in the US lower 48 states and correctly predicted that its peak would be reached around 1970.[11] At first, he did not provide a functional form for his prediction but instead fit past production to a bell-shaped curve in which the area under the curve was equal to his estimates of the amount of total oil available. Later, in 1959, he specified a functional form for the curve.[12] His starting point was the logistic curve, stating that cumulative production would follow a logistic curve, and thus that yearly production would follow the first derivative of the logistic curve (which was named, since that moment, the Hubbert's curve).

Probably, the leading proponent of Hubbert's theories is Kenneth Deffeyes who has published several books on the subject, the latest in 2010.[13] However, some aspects of these theories have also led other researchers to extend and/or modify his original model.

One such aspect is related to the fact that Hubbert provides a forecast with only one peak in oil production, which seems valid for countries with a large number of oil fields and basins. In the work of Laherrère,[14] examples were presented showing that oil production in several countries (France, UK, and others) cannot be represented by a single Hubbert cycle, and a model characterized by several cycles was introduced. This approach has been labeled *multiple-Hubbert modeling* and has been extended in the work of Maggio and Cacciola[15] and applied to world oil production[16] and to oil production in Brazil.[17]

In addition, some authors have focused on the symmetry of the curve and on modifying this aspect. For example, the works of Hallock et al[18] and Hallock et al[19] used a modified version of the bell-shaped curve, which peaks at 60% of ultimate production instead of the typical 50%. This method implies an asymmetric shape of production and a steeper rate of decline than increase. Symmetric and asymmetric linear and exponential models were considered in the work of Brandt,[20] where they are compared with the Hubbert curve.

Another widely discussed aspect is that Hubbert's method assumes that oil production is only time-dependent and therefore does not take into account the effect of possible technological and/or economic factors. This has led to the modification of the model via the introduction of economic variables and econometric models. For example, the works of Reynolds[21,22] included prices and costs in Hubbert's model. In the work of Kaufmann,[23] the effect of geological, economic, and political factors on oil production in the US lower 48 states between 1947 and 1985 was analyzed with a method that combines curve fitting and econometric models. Brecha[24] provided a hybrid approach to the peak-oil question with 2 models in which the use of logistic curves for cumulative production is supplemented with data on projected extraction costs and historical rates of capacity increase. In addition, an econometric model based on a system of simultaneous equations was developed in the work of Kemp and Kasim.[25] Other economic approaches to this subject have considered models in which production increases with demand, advancing technology, reserve additions, and site development.[26] In the work of Guseo et al,[27] a generalized Bass model was introduced to treat global oil growth as a natural diffusion process linked to exogenous variables like price, technology, and strategic interventions.

As regards to probabilistic modeling of oil production, the literature contains few references to the subject. In the work of Michel,[28] a probabilistic model to predict oil production was suggested based on field size (modeled by a Pareto distribution), which takes into account the process of launching production (modeled by a gamma distribution).

3 | THE HUBBERT DIFFUSION PROCESS

3.1 | The Hubbert curve

It is well known that the Hubbert curve is obtained from the derivative of the logistic function, for which we consider the general expression

$$l(t) = \frac{k}{\eta + \alpha^t}, \qquad t \in \mathbb{R}; \eta, k > 0,\ 0 < \alpha < 1. \tag{1}$$

By deriving (1) with respect to t and imposing that $l'(t_0) = x_0$ (here t_0 represents the initial observation time), the following expression for the Hubbert curve is obtained:

$$x(t) = l'(t) = x_0 \left(\frac{\eta + \alpha^{t_0}}{\eta + \alpha^t} \right)^2 \alpha^{t - t_0},\ t \in \mathbb{R}; \eta > 0,\ 0 < \alpha < 1. \tag{2}$$

In the context of oil production, $x(t)$ usually represents the number of barrels produced per day. Parameter η has no units, whereas, for α, the units are $\exp(\text{days}^{-1})$.

The maximum of (2) is, probably, the main feature of the Hubbert curve when it comes to modeling oil production. In fact, the maximum is known as the peak of production and is achieved at time instant

$$t_{\max} = \ln \eta / \ln \alpha, \tag{3}$$

usually known as peak time. Furthermore, $t_{\max} > t_0$, ie, the peak occurs after the initial observation time if and only if $0 < \eta < \alpha^{t_0} < 1$. In addition, its value is

$$x(t_{\max}) = x_0 \frac{(\eta + \alpha^{t_0})^2}{4\eta\alpha^{t_0}}. \tag{4}$$

Another important feature of the curve is given by its inflection points. It can be seen that the curve exhibits 2 inflection points symmetric around $t_{\max}$. Concretely,

$$t_{\text{inf},1} = t_{\max} + \frac{\ln\left(2 + \sqrt{3}\right)}{\ln \alpha},$$
$$t_{\text{inf},2} = t_{\max} + \frac{\ln\left(2 - \sqrt{3}\right)}{\ln \alpha}.$$

These points verify $t_{\text{inf},1} < t_{\max} < t_{\text{inf},2}$. Furthermore, $t_{\text{inf},1} > t_0$ if and only if $\eta < \alpha^{t_0}(2 - \sqrt{3}) < 2 - \sqrt{3}$.

Finally, we consider the area under the curve, known as ultimate recoverable resources (URR), in oil production and often used as a tool for estimating the parameters of the curve. Its expression is given by

$$URR = -x_0 \frac{(\eta + \alpha^{t_0})^2}{\eta \alpha^{t_0} \ln \alpha}. \tag{5}$$

3.2 | Obtaining the diffusion process

In order to model Hubbert-type behaviors from a stochastic point of view, our contribution is to consider a diffusion process whose mean function is (2). This expression verifies the ordinary differential equation

$$x'(t) = r(t)x(t), \quad x(t_0) = x_0, \tag{6}$$

where

$$r(t) = -\ln \alpha \frac{\alpha^t - \eta}{\eta + \alpha^t},$$

which can be viewed as a generalization of the Malthusian growth model with a time-dependent fertility rate $r(t)$. A stochastic version of this model is given by the nonhomogeneous lognormal diffusion process (or lognormal diffusion process with exogenous factors). This is a diffusion process $\{X(t); t \geq t_0\}$, taking values on $\mathbb{R}^+$ and with infinitesimal moments

$$A_1(x,t) = h(t)x$$
$$A_2(x) = \sigma^2 x^2,$$

being the solution of the stochastic differential equation

$$dX(t) = h(t)X(t)dt + \sigma X(t)dW(t)$$
$$X(t_0) = X_0,$$

where $W(t)$ is a standard Wiener process independent on X_0, $t \geq t_0$, and $h(t)$ is a continuous and bounded function.

An explanation of the main features of the process can be found in the work of Gutiérrez et al,[29] where the authors made a detailed theoretical analysis of the process. Concerning other potential fields of application, in the work of Gutiérrez et al,[30] an inferential analysis was performed to assess the usefulness of the process in economics.

Among the characteristics of the process, we will focus on the mean function and its conditioned version on a value y in a previous time instant s whose expressions are

$$m_X(t) = E[X(t)] = E[X_0] \exp\left(\int_{t_0}^{t} h(u)du\right)$$
$$m_X(t|y,s) = E[X(t)|X(s) = y] = y \exp\left(\int_{s}^{t} h(u)du\right). \tag{7}$$

These functions verify $m'_X(t) = m_X(t)h(t)$ and $m'_X(t|y,s) = m_X(t|y,s)h(t)$, that is, the same ordinary differential equation (6) verified by the Hubbert curve (2). This question leads us to define the Hubbert diffusion process as the particular case of the nonhomogeneous lognormal diffusion process considering $h(t) = r(t)^*$.

In this way, all the features of the Hubbert process can be obtained from those established in the work of Gutiérrez et al[29] for the lognormal diffusion process with exogenous factors. In particular, the transition probability density function is, for $s < t$,

$$f(x,t|y,s) = \frac{1}{\sigma x \sqrt{2\pi(t-s)}} \exp\left\{\frac{-\left[\ln \frac{x}{y} - 2\ln \frac{\eta+\alpha^s}{\eta+\alpha^t} - \left(\ln \alpha - \frac{\sigma^2}{2}\right)(t-s)\right]^2}{2\sigma^2(t-s)}\right\}, \tag{8}$$

*Note that $r(t)$ is a decreasing continuous function verifying $-\ln \alpha < r(t) < \ln \alpha$, $\forall t \in \mathbb{R}$.

which corresponds to that of a lognormal variable, ie,

$$X(t)|X(s) = y \sim \Lambda_1 \left[\ln y + 2 \ln \frac{\eta + \alpha^s}{\eta + \alpha^t} + \left(\ln \alpha - \frac{\sigma^2}{2} \right) (t - s), \sigma^2 (t - s) \right].$$

In order to calculate the finite-dimensional distributions of this process, the distribution of X_0 must be fixed. By considering a degenerate distribution at t_0, ie, $P[X_0 = x_0] = 1$, or a lognormal distribution $\Lambda_1(\mu_0, \sigma_0^2)$, all finite-dimensional distributions are lognormal[†]. Concretely, $\forall n \in \mathbb{N}$ and $t_1 < t_2 < \ldots < t_n$, and denoting by Λ_n the n-dimensional lognormal distribution, we have

$$(X(t_1), \ldots, X(t_n))^T \sim \Lambda_n(\mu, \Sigma),$$

where the components of vector $\mu = (\mu_1, \ldots, \mu_n)^T$ and matrix $\Sigma = (\sigma_{ij})$, ie, $i, j = 1, \ldots, n$, are

$$\mu_i = \mu_0 + 2 \ln \frac{\eta + \alpha^{t_0}}{\eta + \alpha^{t_i}} + \left(\ln \alpha - \frac{\sigma^2}{2} \right) (t_i - t_0), \quad i = 1, \ldots, n$$

and

$$\sigma_{ij} = \sigma_0^2 + \sigma^2 \left(\text{Min}(t_i, t_j) - t_0 \right), \quad i, j = 1, \ldots, n,$$

respectively. Finally, from (7), the mean functions of the Hubbert process result in the forms

$$m_X(t) = E[X_0] \left(\frac{\eta + \alpha^{t_0}}{\eta + \alpha^t} \right)^2 \alpha^{(t - t_0)}, \; t \geq t_0 \tag{9}$$

and

$$m_X(t|y, s) = y \left(\frac{\eta + \alpha^s}{\eta + \alpha^t} \right)^2 \alpha^{(t-s)}, \; t \geq s, \tag{10}$$

respectively, which are Hubbert curves of the type here introduced and can be used for predictions purposes within the context of this model.

4 | INFERENCE ON THE PROCESS

In the context of oil production, the prediction of peak and peak time is a problem of great interest. Furthermore, obtaining future production values can be very useful in real situations. This last question can be addressed using the mean functions (9) and (10). Note that these functions and the peak time and peak given by (3) and (4), respectively, are functions expressed in terms of the parameters of the process. The parameters of the process must therefore be previously estimated if we intend to make estimations in real-life situations.

Let us then examine in this section the ML estimation of the parameters of the model from which we can obtain, by virtue of Zehna's theorem[‡], the corresponding for the aforementioned parametric functions.

4.1 | Likelihood function

Let us consider a discrete sampling of the process based on d sample paths at times t_{ij}, $(i = 1, \ldots, d, j = 1, \ldots, n_i)$ with $t_{i1} = t_1$, $i = 1, \ldots, d$. Denote by $\mathbf{x} = \{x_{ij}\}_{i=1,\ldots,d; j=1,\ldots,n_i}$ the observed values of the $X(t_{ij})$ variables of the process.

The likelihood function depends on the choice of the initial distribution. In the following, we will consider the general case when the initial distribution is lognormal, ie, $X(t_1) \sim \Lambda_1(\mu_1, \sigma_1^2)$. The transition probability density function (8) can be rewritten as

$$f(x_{ij}, t_{ij}|x_{ij-1}, t_{i,j-1}) = \frac{1}{\sigma x_{ij} \sqrt{2\pi (t_{ij} - t_{i,j-1})}} \times \exp \left(- \frac{\left[\ln \left(\frac{x_{ij}}{x_{i,j-1}} \right) - 2T_{ij}^{\eta,\alpha} - \left(\ln \alpha - \frac{\sigma^2}{2} \right) (t_{ij} - t_{i,j-1}) \right]^2}{2\sigma^2 (t_{ij} - t_{i,j-1})} \right),$$

[†]Note that the former case is a particular case of the second with $\mu_0 = \log x_0$ and $\sigma_0^2 = 0$.

[‡]Zehna's theorem states that if $\widehat{\theta}$ is a maximum likelihood (ML) estimator for θ, then $g(\widehat{\theta})$ is a ML estimator for $g(\theta)$ (see the work of Rohatgi and Saleh[31]).

where

$$T_{ij}^{\eta,\alpha} = \ln \frac{\eta + \alpha^{t_{i,j-1}}}{\eta + \alpha^{t_{ij}}},$$

from which, and denoting $N = \sum_{i=1}^{d} n_i$, the log-likelihood function of the sample is

$$\begin{aligned} L_{\mathbf{x}}\left(\mu_1, \sigma_1^2, \eta, \alpha, \sigma^2\right) = &-\frac{N}{2}\ln(2\pi) - \frac{d}{2}\ln\sigma_1^2 - \frac{N-d}{2}\ln\sigma^2 - \sum_{i=1}^{d}\ln x_{i1} \\ &- \frac{1}{2\sigma_1^2}\sum_{i=1}^{d}[\ln x_{i1} - \mu_1]^2 - \sum_{i=1}^{d}\sum_{j=2}^{n_i}\ln x_{ij} - \frac{1}{2}\sum_{i=1}^{d}\sum_{j=2}^{n_i}\ln(t_{ij} - t_{i,j-1}) \\ &- \frac{1}{2\sigma^2}\left[Z_1 + 4\left(Y_1^{\eta,\alpha} - Y_2^{\eta,\alpha}\right) + \left(\ln\alpha - \frac{\sigma^2}{2}\right)\left[\left(\ln\alpha - \frac{\sigma^2}{2}\right)Z_2 - 2[Z_3 - 2R^{\eta,\alpha}]\right]\right], \end{aligned} \tag{11}$$

where

$$Y_1^{\eta,\alpha} = \sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\left(T_{ij}^{\eta,\alpha}\right)^2}{t_{ij} - t_{i,j-1}}, \quad Y_2^{\eta,\alpha} = \sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right)T_{ij}^{\eta,\alpha}}{t_{ij} - t_{i,j-1}}, \quad R^{\eta,\alpha} = \sum_{i=1}^{d}\ln\frac{\eta + \alpha^{t_{i1}}}{\eta + \alpha^{t_{in_i}}}$$

$$Z_1 = \sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\ln^2\left(\frac{x_{ij}}{x_{i,j-1}}\right)}{t_{ij} - t_{i,j-1}}, \quad Z_2 = \sum_{i=1}^{d}(t_{in_i} - t_{i1}), \quad Z_3 = \sum_{i=1}^{d}\ln\left(\frac{x_{in_i}}{x_{i1}}\right).$$

4.2 | Obtaining the ML estimates

From (11), the maximum likelihood (ML) estimates of μ_1 and σ_1^2 are

$$\widehat{\mu}_1 = \frac{1}{d}\sum_{i=1}^{d}\ln x_{i1} \quad \text{and} \quad \widehat{\sigma}_1^2 = \frac{1}{d}\sum_{i=1}^{d}(\ln x_{i1} - \widehat{\mu}_1)^2.$$

However, estimating η, α, and σ^2 poses some difficulties. In fact, the resulting system of equations is exceedingly complex and does not have an explicit solution, and numerical procedures must be employed. Nevertheless, it is impossible to carry out a general study of the system of equations to check the conditions of convergence of the chosen numerical method since it is dependent on sample data.

One alternative would be using stochastic optimization procedures like SA and VNS. These algorithms are designed to solve problems of the type $\min_{\theta\in\Theta} g(\theta)$, g being the target function to be optimized, and are often more appropriate than classical numerical methods since they impose fewer restrictions on the space of solutions Θ and on the analytical properties of g.

In our case, once μ_1 and σ_1^2 have been estimated, the problem becomes maximizing function $L_{\mathbf{x}}(\widehat{\mu}_1, \widehat{\sigma}_1^2, \eta, \alpha, \sigma^2)$. Since the algorithms aforementioned are usually formulated for minimization problems, from (11), the target function we will consider is

$$g_{\mathbf{x}}(\eta, \alpha, \sigma^2) = \frac{N-d}{2}\ln\sigma^2 + \frac{1}{2\sigma^2}\left[Z_1 + 4\left(Y_1^{\eta,\alpha} - Y_2^{\eta,\alpha}\right) + \left(\ln\alpha - \frac{\sigma^2}{2}\right)\left[\left(\ln\alpha - \frac{\sigma^2}{2}\right)Z_2 - 2[Z_3 - 2R^{\eta,\alpha}]\right]\right]. \tag{12}$$

4.3 | Fitting and forecasting. Peak time and peak

Fitting the observed data and predicting the future behavior of the process can be done from the ML estimates of mean functions (9) and (10). Note that, in the case of calculating predictions at future time instants, it is more reasonable to use the conditioned version of the mean function since it employs more updated information than that provided by the initial distribution of the process.

In addition, and given that the mean functions of the process are Hubbert functions, forecasts for peak time can be performed by means of point estimation, substituting the estimates of the parameters in the time instant at which their maxima are achieved. We note that this time instant is the same in both cases and coincides with (3). Therefore, the ML

estimation of peak time is

$$\widehat{t}_{\max} = \frac{\ln \widehat{\eta}}{\ln \widehat{\alpha}}.$$

The estimation of the peak is obtained by substituting that of peak time in (9) or (10). Note that we have 2 possibilities depending on the mean function chosen, resulting in

$$\widehat{E[X(t_1)]}\frac{(\widehat{\eta} + \widehat{\alpha}^{t_1})^2}{4\widehat{\eta}\widehat{\alpha}^{t_1}}, \tag{13}$$

which coincides with (4) for a degenerate initial distribution, and

$$y\frac{(\widehat{\eta} + \widehat{\alpha}^{s})^2}{4\widehat{\eta}\widehat{\alpha}^{s}}, \tag{14}$$

when the conditioned version is used.

At this point, an important remark must be made about the values of time instants. In fact, when dealing with high values for the time instants of the process, it is possible that the estimated value of η is close to zero (note that $\eta = \alpha^{t_{\max}}$). In these situations, common in real cases, the accuracy in the estimates can cause detrimental effects on the final results. One way to tackle this problem is to consider a new diffusion process $\{Y(t); t \geq t_0 - k\}$ obtained from $\{X(t); t \geq t_0\}$ by considering a shift of length k in time, ie, $Y(t) = X(t + k)$. It is not difficult to see that the new process is also a Hubbert diffusion process whose parameters α and σ remain invariant, whereas η becomes $\eta' = \alpha^{-k}\eta$. Furthermore, the expressions of the mean function and the peak are the same whether we use η or η', whereas, for peak time, it is enough to undo the changes made in time. In practice, we recommend considering $k = t_0$ so that the original data can be interpreted as observations of the new process with initial time instant $t_0 = 0$.

It is important to note that the value of the peak is independent of the use of η or η'. In addition, if we note $m_Y(t|y, s)$ the conditioned mean function of process $Y(t)$, then $m_Y(t|y, s) = m_X(t + x|y, s + k)$.

5 | APPLICATION OF SA AND VNS FOR ESTIMATING THE PARAMETERS OF THE HUBBERT DIFFUSION PROCESS

5.1 | A brief summary of SA and VNS

SA is a local metaheuristic algorithm introduced in the work of Kirkpatrick et al[32] and inspired by the metallurgical process of annealing as studied in statistical mechanics. The algorithm performs an iterative exploration of solution space Θ searching for improvements on the value of the target function, say g, and intends to avoid an attraction towards local minima. Concretely, let θ be the solution for a given iteration, θ' a new value selected in a neighborhood of θ in the next iteration, and $\Delta = g(\theta') - g(\theta)$. If $\Delta \leq 0$, then θ' is selected as the new solution. Otherwise, it could be accepted with probability $p = \exp(-\Delta/T)$, where T is a scale factor called *temperature*. This selection procedure, usually referred to as Metropolis algorithm, was described in the work of Metropolis et al.[33]

When applying SA, temperature T is gradually decreased in such a way that, at the beginning, the cooling process allows to select solutions, which worsen the target function. As T decreases, such solutions are not longer accepted. Theoretical studies have shown that, during an infinitely slow cooling, the process converges (as T approaches zero) to a global minimum with probability one.

Therefore, the application of SA requires the following.

1. Initializing the parameters of the algorithm: initial solution (θ_0), initial and final temperature (T_0, T_F), chain length for each application of the Metropolis algorithm (L), cooling procedure, and a stopping condition.
2. Applying the Metropolis procedure L times.
3. Checking the stopping rule and checking if the final temperature has been achieved. If both conditions are not met, then decrease T and return to the previous step.

As regards to the VNS algorithm introduced in the work of Mladenović and Hansen,[34] its goal is to explore several neighborhoods in Θ when a local optimum is found for g through a local search method. The algorithm is applied in 2 phases: first, a structure of neighborhoods, say N_k, ie, $k = 1, \ldots, k_{\max}$, is determined in the solution space and an initial solution θ_0 is chosen. For $k = 1, \ldots, k_{\max}$, the second phase uses a local search method to determine a new θ^* solution

in $N_k(\theta_0)$. If θ^* improves the target function, then $\theta_0 = \theta^*$, and the search is recommenced from $N_1(\theta_0)$. If not, the search continues with $N_{k+1}(\theta_0)$. Note that the final solution is a local minimum with respect to the $k_{\max}$ neighborhood structures, and therefore finding a global minimum is much more probable than when using a single structure. Also note that the procedure here described changes neighborhood every time there is an improvement in g. Another variation exists depending on the different ways that the neighborhood structure may change when a local optimum has been reached and on several local search methods (for other possibilities, see the work of BoussaïD et al[35]).

Both algorithms have experienced remarkable popularity in the last years, having been applied to several fields of research[36,37] and Zoraghi[38] used them in the context of ML estimation for distributions, whereas in the works of Román-Román et al[39] and Román-Román and Torres-Ruiz,[40] they helped estimate the parameters of Gompertz-type and Richards diffusion processes, respectively.

5.2 | Bounding the solution space

The parametric space Θ linked to the target function (12), on which the selected algorithms must operate, is continuous and unbounded. Concretely,

$$\Theta = \{(\eta, \alpha, \sigma) : \eta > 0, 0 < \alpha < 1, \sigma > 0\}.$$

The drawback is that the solution space might not be explored with enough depth. This requires us to find arguments for bounding said space.

Regarding parameter σ, it is known that when it has high values it leads to sample paths with great variability around the mean of the process. Thus, excessive variability in available paths would make a Hubbert-type modeling inadvisable. Some simulations performed for several values of σ have led us to consider that $0 < \sigma < 0.1$ so that we may have paths compatible with a Hubbert-type growth. On the other hand, whereas α is bounded, there does not seem to be an upper bound for η. However, as noted in Section 3, the Hubbert curve has an inflection point before $t_{\max}$, which is visualized ($t_{\inf} > t_0$) if and only if $\eta < \alpha^{t_0}(2 - \sqrt{3}) < 2 - \sqrt{3}$, so an upper bound for η is found.

Additionally, when an estimation of the URR is available, some refinements can be made for bounding α.

- From (5), the following second-degree equation is obtained

$$x_0\eta^2 + \alpha^{t_0}(2x_0 + URR\ln\alpha)\eta + x_0\alpha^{2t_0} = 0.$$

  In order to have a solution, the discriminant of the equation must verify

$$\alpha^{2t_0} URR\ln\alpha(URR\ln\alpha + 4x_0) \geq 0,$$

  from which we deduce $\alpha \leq \alpha_1$, where $\alpha_1 = \exp(-4x_0/URR)$.
- Now, let T_F be the final observation time (it can be before or after $t_{\inf}$ and/or $t_{\max}$) and consider c defined by

$$c = \int_{t_0}^{T_F} x(t)dt = \eta URR\frac{\alpha^{t_0} - \alpha^{T_F}}{(\eta + \alpha^{t_0})(\eta + \alpha^{T_F})},$$

  from which we obtain a new second-degree equation, concretely,

$$M\eta^2 + \alpha_0^t\left[(M-1) + (M+1)\alpha^h\right]\eta + M\alpha^{2t_0+h} = 0,$$

  where $h = T_F - t_0$ and $M = c/URR$.

  Following a similar argument to the one established for the previous bound, the following expression is now verified, ie, $(M-1)^2 - 2(M^2+1)\alpha^h + (M+1)^2\alpha^{2h} \geq 0$, resulting $\alpha \leq \alpha_2$ being

$$\alpha_2 = \left(\frac{M-1}{M+1}\right)^{2/h}.$$

- From the previous bounds, we consider $0 < \alpha < \alpha^* = \text{Min}(\alpha_1, \alpha_2)$.

T1 Table 1 contains, for several values of α and η, the bounds provided by the proposed method. We can see how, for $\eta = 0.01$, the bound provided by α_2 is preferable, regardless of the value of α. This situation changes as η grows since the range of values of α for which the bound provided by α_1 is preferable increases. In such a case, as η grows, the amplitude of the intervals provided by α_1 decreases.

TABLE 1 Bounds for α for several values of this parameter and η

	$t_0 = 0,\ t_f = 50$							
	$\eta = 0.01$		$\eta = 0.025$		$\eta = 0.05$		$\eta = 0.075$	
α	α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2
0.05	0.8891	0.8088	0.7519	0.8388	0.5807	0.8619	0.4594	0.8756
0.10	0.9136	0.8088	0.8031	0.8388	0.6585	0.8619	0.5500	0.8756
0.15	0.9283	0.8088	0.8347	0.8388	0.7088	0.8619	0.6111	0.8756
0.20	0.9388	0.8088	0.8579	0.8388	0.7467	0.8619	0.6584	0.8756
0.25	0.9470	0.8088	0.8763	0.8388	0.7776	0.8619	0.6977	0.8756
0.30	0.9538	0.8088	0.8917	0.8388	0.8037	0.8619	0.7315	0.8756
0.35	0.9596	0.8088	0.9049	0.8388	0.8265	0.8619	0.7614	0.8756
0.40	0.9647	0.8088	0.9164	0.8388	0.8468	0.8619	0.7883	0.8756
0.45	0.9691	0.8088	0.9268	0.8388	0.8651	0.8619	0.8127	0.8756
0.50	0.9731	0.8088	0.9361	0.8388	0.8818	0.8619	0.8353	0.8756
0.55	0.9768	0.8088	0.9446	0.8388	0.8972	0.8619	0.8562	0.8756
0.60	0.9801	0.8088	0.9525	0.8388	0.9114	0.8619	0.8758	0.8756
0.65	0.9832	0.8088	0.9598	0.8388	0.9248	0.8619	0.8941	0.8756
0.70	0.9861	0.8088	0.9666	0.8388	0.9373	0.8619	0.9115	0.8756
0.75	0.9887	0.8090	0.9729	0.8388	0.9491	0.8619	0.9280	0.8756
0.80	0.9912	0.8132	0.9789	0.8395	0.9603	0.8621	0.9437	0.8757
0.85	0.9936	0.8546	0.9846	0.8522	0.9709	0.8660	0.9586	0.8776
0.90	0.9958	0.9398	0.9900	0.9148	0.9810	0.9019	0.9730	0.9000
0.95	0.9979	0.9915	0.9951	0.9821	0.9907	0.9715	0.9867	0.9644
	$\eta = 0.1$		$\eta = 0.15$		$\eta = 0.2$		$\eta = 0.25$	
α	α_1	α_2	α_1	α_2	α_1	α_2	α_1	α_2
0.05	0.3714	0.8853	0.2568	0.8989	0.1893	0.9085	0.1470	0.9158
0.10	0.4671	0.8853	0.3518	0.8989	0.2782	0.9085	0.2290	0.9158
0.15	0.5341	0.8853	0.4228	0.8989	0.3485	0.9085	0.2969	0.9158
0.20	0.5874	0.8853	0.4818	0.8989	0.4089	0.9085	0.3569	0.9158
0.25	0.6323	0.8853	0.5331	0.8989	0.4629	0.9085	0.4117	0.9158
0.30	0.6716	0.8853	0.5791	0.8989	0.5122	0.9085	0.4627	0.9158
0.35	0.7067	0.8853	0.6210	0.8989	0.5580	0.9085	0.5107	0.9158
0.40	0.7386	0.8853	0.6598	0.8989	0.6010	0.9085	0.5563	0.9158
0.45	0.7679	0.8853	0.6960	0.8989	0.6417	0.9085	0.5998	0.9158
0.50	0.7952	0.8853	0.7301	0.8989	0.6803	0.9085	0.6417	0.9158
0.55	0.8206	0.8853	0.7624	0.8989	0.7173	0.9085	0.6820	0.9158
0.60	0.8446	0.8853	0.7931	0.8989	0.7529	0.9085	0.7211	0.9158
0.65	0.8672	0.8853	0.8224	0.8989	0.7871	0.9085	0.7590	0.9158
0.70	0.8887	0.8853	0.8505	0.8989	0.8202	0.9085	0.7959	0.9158
0.75	0.9092	0.8853	0.8776	0.8989	0.8522	0.9085	0.8318	0.9158
0.80	0.9288	0.8854	0.9037	0.8989	0.8834	0.9085	0.8669	0.9158
0.85	0.9476	0.8865	0.9289	0.8995	0.9136	0.9088	0.9012	0.9161
0.90	0.9657	0.9016	0.9533	0.9078	0.9431	0.9141	0.9347	0.9198
0.95	0.9831	0.9595	0.9769	0.9538	0.9719	0.9511	0.9677	0.9501

5.3 | Choosing the main options for applying SA and VNS

Once the solution space has been bounded, we specify the choice of the initial parameters of the algorithms and the stopping conditions to apply them to the estimation of parameters in the Hubbert diffusion process.

For SA, we consider the following.

1. The initial solution is chosen randomly in the bounded subspace $\Theta' = (0, 2-\sqrt{3}) \times (0, \alpha^*) \times (0, 0.1)$. Note that we have considered the shorter space based on the previous bounding. This depends on the availability of data about URR. If these data are not available, then we must replace Θ' with $\Theta^* = (0, 2-\sqrt{3}) \times (0, 1) \times (0, 0.1)$.

2. For the application of SA, the initial temperature must be high so that, at the beginning, there is a high probability, ie, p_0, of accepting values that increase the value of the target function. In our case and following Busetti,[41] we have considered $p_0 = 0.9$ so that $T_0 = -\Delta g^+ / \log(p_0)$. Δg^+ denotes the average increase in the target function when values that produce an increase are accepted after considering N values in the solution space. In our case, we have considered $N = 100$.
3. For the cooling process, we have considered a geometric scheme in which the current temperature is multiplied by a constant γ $(0 < \gamma < 1)$, ie, $T_i = \gamma T_{i-1}$, $i \geq 1$. The usual values for γ are between 0.8 and 0.99. For this case, we have set the constant at $\gamma = 0.95$.
4. The selected length of the chain for the application of the Metropolis algorithm is $L = 50$. Therefore, in each step, a chain of 50 solutions will be generated before checking the stopping rule and modifying the temperature if necessary.
5. The selected stopping rule is twofold. First, it checks whether the latest 50 values generated are equal, in which case, the algorithm is stopped. Otherwise, it continues until the temperature reaches a value close to zero (0.1 in this case).

As aforementioned, for the application of the VNS algorithm, we must take into account the neighborhood structure and the local search method. We have made the following choices.

- **Neighborhood structure.** After selecting a value for $k_{\max}$ (in the applications developed in this paper, we have selected $k_{\max} = 5$), we proceed as follows.

  Let Θ' be the solution subspace aforementioned. Given an initial solution $\theta_0 = (\alpha_0, \eta_0, \sigma_0)$, we consider quantities

$$h_{11} = \frac{\eta_0}{k_{\max}}, \quad h_{12} = \frac{2-\sqrt{3}-\eta_0}{k_{\max}}, \quad h_{21} = \frac{\alpha_0}{k_{\max}} \quad h_{22} = \frac{\alpha^* - \alpha_0}{k_{\max}}, \; h_{31} = \frac{\sigma_0}{k_{\max}}, \quad h_{32} = \frac{0.1-\sigma_0}{k_{max}}$$

  from which the neighborhood structure is given by

$$N_k(\theta) = [\eta_0 - k h_{11}, \eta_0 + k h_{12}] \times [\alpha_0 - k h_{21}, \alpha_0 + k h_{22}] \times [\sigma_0 - k h_{31}, \sigma_0 + k h_{32}]$$

  for $k = 1, \ldots, k_{\max}$.

TABLE 2 Absolute relative errors ($\times 10^{-3}$) between the real and the estimated log-likelihood function after applying simulated annealing (SA) and variable neighborhood search (VNS)–SA from simulated data

	$\sigma = 0.05$															
	$\eta = 0.01$		$\eta = 0.025$		$\eta = 0.05$		$\eta = 0.075$		$\eta = 0.1$		$\eta = 0.15$		$\eta = 0.20$		$\eta = 0.25$	
α	SA	VNS	SA	VNS	SA	VNS	SA	VNS	SA	VNS	SA	VNS	SA	VNS	SA	VNS
0.05	2.150	0.470	1.440	0.110	0.750	0.001	0.310	0.001	0.110	0.010	0.010	0.001	0.040	0.001	0.040	0.001
0.15	2.750	0.050	1.620	0.070	1.090	0.058	0.050	0.010	0.328	0.012	0.292	0.018	0.245	0.011	0.123	0.049
0.25	2.450	0.040	0.840	0.010	0.915	0.013	0.440	0.020	0.574	0.017	0.005	0.005	0.211	0.001	0.066	0.065
0.35	1.500	0.010	1.020	0.040	0.822	0.027	0.310	0.001	0.277	0.012	0.139	0.009	0.170	0.005	0.127	0.043
0.45	6.640	0.270	1.990	0.001	0.313	0.007	0.280	0.001	0.321	0.004	0.346	0.005	0.007	0.064	0.077	0.067
0.55	9.290	0.100	1.970	0.030	0.228	0.037	0.560	0.001	0.420	0.057	0.067	0.002	0.259	0.019	0.099	0.100
0.65	7.489	2.060	7.620	0.530	1.348	0.039	0.120	0.040	1.028	0.117	0.502	0.095	0.153	0.085	0.074	0.068
0.75	2.470	0.170	5.240	0.050	0.404	0.064	2.270	0.040	3.651	0.034	2.144	0.738	0.027	0.029	0.518	0.426
0.85	6.330	0.700	3.530	0.350	1.022	0.078	0.090	0.010	0.536	0.053	0.938	0.135	0.392	0.295	0.904	0.342
0.95	0.590	0.150	0.300	0.010	0.458	0.146	0.030	0.020	0.343	0.105	0.259	0.236	0.018	0.017	0.028	0.013
	$\sigma = 0.07$															
	$\eta = 0.01$		$\eta = 0.025$		$\eta = 0.05$		$\eta = 0.075$		$\eta = 0.1$		$\eta = 0.15$		$\eta = 0.20$		$\eta = 0.25$	
α	SA	VNS	SA	VNS	SA	VNS	SA	VNS	SA	VNS	SA	VNS	SA	VNS	SA	VNS
0.05	0.180	0.010	0.250	0.010	0.310	0.001	0.120	0.001	0.090	0.001	0.009	0.001	0.009	0.001	0.020	0.001
0.15	1.000	0.010	0.052	0.030	0.267	0.009	0.110	0.001	0.009	0.006	0.069	0.013	0.012	0.004	0.096	0.005
0.25	0.440	0.010	0.240	0.001	0.033	0.018	0.070	0.001	0.011	0.016	0.043	0.005	0.034	0.006	0.082	0.010
0.35	0.260	0.010	0.640	0.010	0.139	0.026	0.260	0.001	0.132	0.010	0.003	0.017	0.022	0.022	0.016	0.009
0.45	2.300	0.010	0.310	0.001	0.322	0.005	0.070	0.010	0.336	0.031	0.338	0.028	0.031	0.007	0.041	0.051
0.55	6.600	0.030	0.970	0.040	0.598	0.022	0.040	0.030	0.046	0.040	0.039	0.022	0.013	0.044	0.007	0.002
0.65	3.350	0.010	0.590	0.160	0.517	0.125	0.090	0.070	0.091	0.084	0.162	0.133	0.063	0.022	0.376	0.108
0.75	0.500	0.010	0.750	0.080	0.181	0.136	0.330	0.015	1.325	1.256	1.609	0.069	0.866	0.317	1.584	0.658
0.85	2.000	0.440	1.390	0.600	0.211	0.060	0.050	0.020	0.019	0.016	0.429	0.151	0.186	0.184	0.715	0.570
0.95	0.330	0.080	0.090	0.030	0.154	0.064	0.051	0.021	0.074	0.039	0.512	0.183	0.483	0.356	0.654	0.595

- **Local search.** The local search method we have selected is SA. This choice allows us to perform a hybrid procedure, which has proven to be useful in several applications (see, for instance the works of Abbasi et al[37] or Román-Román and Torres-Ruiz[40]).
- **Initial solution.** In order to apply the algorithm, we have considered, as an initial solution, the one found by applying the SA algorithm.

6 | SIMULATION EXAMPLE

In order to validate the procedures suggested in previous sections, we have performed a simulation study with the following pattern: 50 sample paths were simulated, each one obtained through the existing relation between the Hubbert process and the Wiener process. These sample paths are linked by the expression

$$X(t) = X_{t_0}\left(\frac{\eta+\alpha^{t_0}}{\eta+\alpha^{t}}\right)^2 \alpha^{t-t_0}\exp\left(\sigma W(t-t_0) - \frac{\sigma^2}{2}(t-t_0)\right).$$

We have considered 501 data at time instants $t_i = (i-1)\cdot 0.1$, ie, $i = 1, \ldots, 501$. We have proposed a set of values for parameters $\eta = (0.01, 0.025, 0.05, 0.1, 0.15, 0.2, 0.25)$ (note that $\eta < 2-\sqrt{3}$), $\alpha = 0.05 + (i-1)\cdot 0.1$, ie, $i = 1, \ldots, 10$, and $\sigma = (0.05, 0.07)$, chosen arbitrarily. However, we wanted these values to be able to mirror real cases (as the ones described in the next section). The initial distribution is degenerate at value $x_0 = 100$. After the simulation, we chose a sample from the first value, using a step equal to 1. Hence, a sample of 51 data was obtained for each sample path. Furthermore, the whole procedure was replicated 50 times.

Table 2 shows, for each combination of the parameters, the absolute relative errors ($\times 10^{-3}$) between the real and the estimated log-likelihood function after applying the SA and the hybrid VNS-SA algorithm. Note that both methods provide good estimates of the parameters in terms of the relative error in the likelihood function. Nevertheless, we must remark

T2

TABLE 3 Estimated values of the parameters after applying the variable neighborhood search–simulated annealing algorithm from simulated data

							$\sigma = 0.05$					
	$\eta = 0.01$			$\eta = 0.025$			$\eta = 0.05$			$\eta = 0.075$		
α	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$
0.05	0.0101	0.0501	0.0599	0.0248	0.0500	0.0547	0.0500	0.0501	0.0511	0.0746	0.0501	0.0509
0.15	0.0100	0.1503	0.0526	0.0249	0.1502	0.0531	0.0504	0.1501	0.0510	0.0747	0.1501	0.0502
0.25	0.0099	0.2501	0.0516	0.0250	0.2505	0.0513	0.0500	0.2499	0.0506	0.0745	0.2499	0.0505
0.35	0.0101	0.3501	0.0509	0.0247	0.3500	0.0514	0.0505	0.3498	0.0512	0.0751	0.3501	0.0498
0.45	0.0098	0.4503	0.0524	0.0249	0.4502	0.0507	0.0500	0.4500	0.0508	0.0745	0.4502	0.0497
0.55	0.0099	0.5498	0.0509	0.0249	0.5499	0.0506	0.0500	0.5499	0.0506	0.0748	0.5502	0.0501
0.65	0.0101	0.6501	0.0503	0.0249	0.6497	0.0510	0.0502	0.6501	0.0503	0.0744	0.6499	0.0496
0.75	0.0101	0.7501	0.0508	0.0248	0.7500	0.0508	0.0502	0.7500	0.0501	0.0742	0.7498	0.0498
0.85	0.0092	0.8467	0.0506	0.0238	0.8480	0.0507	0.0511	0.8499	0.0502	0.0739	0.8497	0.0496
0.95	0.0142	0.9476	0.0503	0.0253	0.9496	0.0505	0.0521	0.9498	0.0503	0.0729	0.9491	0.0498
		$\eta = 0.1$			$\eta = 0.15$			$\eta = 0.20$			$\eta = 0.25$	
α	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$
0.05	0.1001	0.0409	0.0505	0.1499	0.0499	0.0505	0.0201	0.0502	0.0504	0.2482	0.0499	0.0505
0.15	0.1004	0.1499	0.0506	0.1506	0.1500	0.0506	0.2005	0.1500	0.0502	0.2476	0.1498	0.0510
0.25	0.1004	0.2499	0.0508	0.1505	0.2499	0.0505	0.2005	0.2498	0.0505	0.2470	0.2498	0.0503
0.35	0.1003	0.3498	0.0507	0.1509	0.3498	0.0504	0.2005	0.3498	0.0505	0.2491	0.3498	0.0506
0.45	0.1004	0.4499	0.0508	0.1505	0.4499	0.0502	0.1992	0.4497	0.0503	0.2496	0.4495	0.0502
0.55	0.0995	0.5498	0.0503	0.1511	0.5499	0.0504	0.2009	0.5500	0.0503	0.2476	0.5497	0.0507
0.65	0.1004	0.6498	0.0503	0.1505	0.6497	0.0502	0.2003	0.6497	0.0503	0.2474	0.6499	0.0504
0.75	0.1011	0.7501	0.0505	0.1497	0.7496	0.0504	0.1988	0.7497	0.0505	0.2470	0.7496	0.0504
0.85	0.1012	0.8500	0.0505	0.1511	0.8498	0.0503	0.1991	0.8495	0.0504	0.2440	0.8493	0.0502
0.95	0.1027	0.9500	0.0502	0.1501	0.9488	0.0503	0.1939	0.9476	0.0503	0.2285	0.9469	0.0504

that using VNS with an initial solution given by SA improves the estimation of the parameters since a noticeable decrease in the relative error is observed. Finally, Tables 3 and 4 show the estimates of the parameters provided by VNS.

7 | REAL APPLICATION

In this section, we consider real data taken from the *US Energy Information Administration*[42] concerning crude oil production (including lease condensate) from Norway and Kazakhstan. The aim is to fit a Hubbert diffusion process to both cases by applying the methodology developed in the previous sections and then to obtain stochastic models showing the behavior of oil production in both countries. Thanks to these models, we can:

- obtain the estimate mean function of the production,
- make forecasts about oil production in the future, and
- estimate production peak time and peak.

In the Norwegian case (Figure 1), the collected data range is 1980-2014, whereas for Kazakhstan (Figure 2), it is 1992-2014. In these graphs, we observe that, in the case of Norway, production peak has already occurred, specifically in 2001 (with a production of 3226 thousand barrels per day (Mbbl/d). After that year, its annual production started an uninterrupted decline with some exceptions such as 2014 when production increased by 2.5% compared with 2013. This kind of behavior suggests that using the peak theory developed by Hubbert might be a good fit. A similar behavior may be expected for oil production in Kazakhstan if no exceptional factors influence its annual production, although its peak has not yet been reached. We will now analyze both cases, starting with Norway.

The data observed for Norway allows us to consider 2 different situations.

- **Scenario 1:** The peak is visualized. In this case, we have taken the whole observed sample path. This situation allows us to evaluate the capacity of the model for fitting the data.

TABLE 4 Estimated values of the parameters after applying the variable neighborhood search–simulated annealing algorithm from simulated data (cont.)

							$\sigma = 0.07$					
		$\eta = 0.01$			$\eta = 0.025$			$\eta = 0.05$			$\eta = 0.075$	
α	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$
0.05	0.0101	0.0502	0.0711	0.0253	0.0499	0.0696	0.0501	0.0498	0.0711	0.0754	0.0501	0.0701
0.15	0.0101	0.1499	0.0716	0.0253	0.1499	0.0696	0.0499	0.1500	0.0699	0.0751	0.1501	0.0695
0.25	0.0101	0.2502	0.0698	0.0251	0.2498	0.0691	0.0497	0.2497	0.0701	0.0751	0.2501	0.0696
0.35	0.0099	0.3501	0.0696	0.0251	0.3501	0.0695	0.0503	0.3500	0.0693	0.0751	0.3502	0.0701
0.45	0.0101	0.4502	0.0699	0.0252	0.4502	0.0699	0.0499	0.4501	0.0703	0.0749	0.4501	0.0698
0.55	0.0101	0.5506	0.0703	0.0252	0.5501	0.0704	0.0498	0.5498	0.0702	0.0745	0.5499	0.0697
0.65	0.0103	0.6507	0.0691	0.0251	0.6502	0.0696	0.0502	0.6503	0.0698	0.0738	0.6502	0.0698
0.75	0.0103	0.7505	0.0699	0.0254	0.7506	0.0693	0.0506	0.7505	0.0702	0.0753	0.7505	0.0698
0.85	0.0093	0.8474	0.0702	0.0233	0.8471	0.0701	0.0496	0.8499	0.0699	0.0749	0.8498	0.0701
0.95	0.0135	0.9493	0.0698	0.0239	0.9505	0.0695	0.0480	0.9500	0.0697	0.0739	0.9494	0.0697
		$\eta = 0.1$			$\eta = 0.15$			$\eta = 0.20$			$\eta = 0.25$	
α	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$
0.05	0.1003	0.0502	0.0704	0.1507	0.0499	0.0694	0.2009	0.0501	0.0697	0.2501	0.0499	0.0701
0.15	0.0998	0.1500	0.0699	0.1500	0.1500	0.0692	0.1988	0.1500	0.0701	0.2468	0.1499	0.0695
0.25	0.1000	0.2501	0.0699	0.1504	0.2500	0.0698	0.2001	0.2498	0.0700	0.2479	0.2501	0.0699
0.35	0.1001	0.3501	0.0699	0.1499	0.3501	0.0697	0.2004	0.3502	0.0698	0.2466	0.3501	0.0702
0.45	0.1005	0.4503	0.0702	0.1509	0.4502	0.0698	0.2006	0.4500	0.0699	0.2449	0.4503	0.0701
0.55	0.0999	0.5502	0.0701	0.1509	0.5502	0.0703	0.1993	0.5503	0.0700	0.2479	0.5499	0.0700
0.65	0.0999	0.6501	0.0699	0.1510	0.6505	0.0697	0.2001	0.6501	0.0700	0.2474	0.6502	0.0698
0.75	0.1007	0.7505	0.0700	0.1512	0.7503	0.0699	0.2002	0.7503	0.0699	0.2458	0.7501	0.0700
0.85	0.1001	0.8503	0.0697	0.1505	0.8506	0.0699	0.1999	0.8504	0.0698	0.2437	0.8500	0.0699
0.95	0.0980	0.9499	0.0700	0.1490	0.9489	0.0701	0.1950	0.9473	0.0701	0.2488	0.9470	0.0701

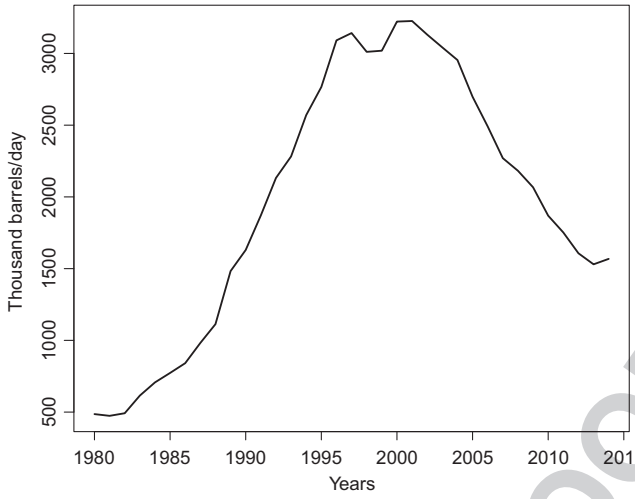

FIGURE 1 Observed crude oil production data for Norway

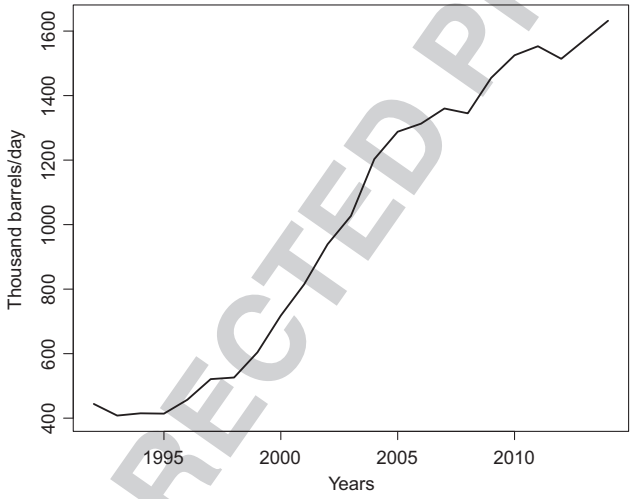

FIGURE 2 Observed crude oil production data for Kazakhstan

- **Scenario 2:** The peak is not visualized. In this case, we consider the truncated data of the previous one at a specific time T_F before the visualization of the peak (in this case, we have chosen $T_F = 1999$). This scenario allows us to consider the predictive capabilities of the model to estimate peak time and to predict the future behavior of oil production.

As we mentioned in Section 5, in order to find a bound for parameter α (for all other parameters, we have considered the fixed bounds mentioned in said section), we must have an estimation of the URR value. To this end, we have considered the sum of the estimated proved reserves, also collected by the US Energy Information Administration,[42] and the accumulated crude oil production until 2014. Table 5 contains the results obtained after applying the hybrid VNS-SA algorithm. Note that, for parameter η, its estimated value in the table corresponds to that of the shifted process by considering $k = t_0 = 1980$. The bounds obtained for α were 0.8724 for scenario 1 and 0.8846 for scenario 2. T5

TABLE 5 Estimated values and standard errors for Norway, considering both scenarios

				Observed values		Maximum Likelihood Estimation		
Scenario	$\widehat{\eta}$	$\widehat{\alpha}$	$\widehat{\sigma}$	**Peak time**	**Peak**	**Peak time**	**Peak a)**	**Peak b)**
1	.0407	.8638	.0634	2001	3226	2001.858	3226.144	----
	(.00216)	(.00177)	(.00016)			(.0003)	(.40167)	
2	.0393	.8607	.0731	2001	3226	2001.579	3338.133	3133.323
	(.00355)	(.00510)	(.00039)			(.00418)	(1.01908)	(.33548)

Table 5 also contains the observed values of peak time and peak, together with their respective estimates. Under scenario 1, peak estimation (*Peak a)* in the table) was performed using expression (13). This case allows us to validate the methodology proposed since peak was already reached in 2001. Indeed, by considering the truncated data series prior to 2001, the method predicts a peak time that is very close to the value observed. When considering scenario 2, we also included the estimation made by using (14). In this latter case, value x_s was the production in 1999 ($x_s = 3019$ Mbbl/d), and the estimated value was noted by *(Peak b)* in the table). This value allows us to perform more accurate predictions regarding production peak. Note that no exact distributions are available for the ML estimators of the parameters of the model. For this reason, and with a view to provide an error value for our estimations, we considered a joint asymptotic distribution of estimators. With this as a starting point, estimation errors could be obtained for the parameters and for any parametric function, such as peak and peak time, by applying the delta method (see Appendix for more details). Table 5 shows these error values below each estimation.

Once the model for crude oil production has been estimated, we may infer what the production will be for the next years. In this case, we have predicted production values until 2040. To this end, we have considered the conditional version of the mean given by (10). Figures 3 and 4 represent the observed values and the predicted ones under both scenarios. In the first case, the mean function is conditioned on the value recorded at the last observation time instant ($s = 2014$, $x_s = 1568$ Mbbl/d), whereas for the second case, we have considered ($s = 1999$, $x_s = 3019$ Mbbl/d). In Figure 4, we have also represented the observed values until 2014 to visualize the overall behavior of production from the initial

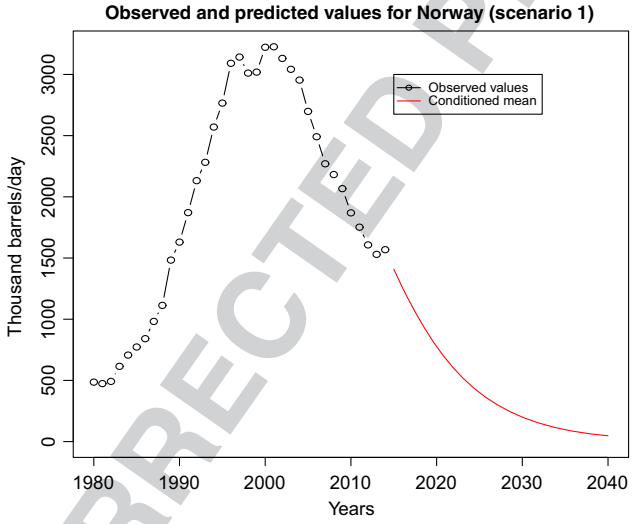

FIGURE 3 Observed and predicted values for Norway (scenario 1) [Colour figure can be viewed at wileyonlinelibrary.com]

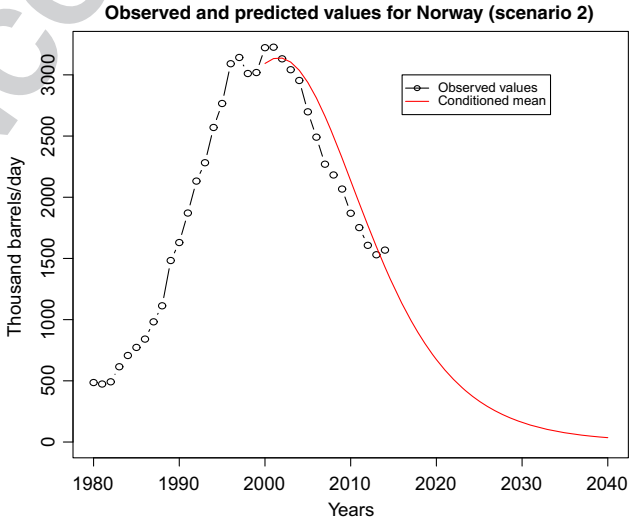

FIGURE 4 Observed and predicted values for Norway (scenario 2) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 6 Predicted values and confidence intervals of oil production for Norway from 2015 to 2040 when considering scenario 1

Year	Mean	Lower Limit	Upper Limit
2015	1409.457	1409.443	1409.471
2016	1260.912	1260.887	1260.938
2017	1123.215	1123.181	1123.250
2018	996.756	996.715	996.797
2019	881.552	881.506	881.598
2020	777.338	777.289	777.387
2021	683.641	683.591	683.691
2022	599.847	599.796	599.897
2023	525.254	525.204	525.304
2024	459.119	459.070	459.168
2025	400.687	400.641	400.734
2026	349.217	349.173	349.262
2027	303.999	303.957	304.041
2028	264.363	264.324	264.402
2029	229.689	229.653	229.725
2030	199.408	199.374	199.441
2031	173.001	172.970	173.032
2032	150.004	149.975	150.032
2033	129.997	129.971	130.023
2034	112.609	112.586	112.633
2035	97.510	97.489	97.531
2036	84.407	84.388	84.427
2037	73.044	73.027	73.062
2038	63.196	63.180	63.211
2039	54.663	54.649	54.677
2040	47.274	47.261	47.286

TABLE 7 Estimated values and standard errors for Kazakhstan

			Maximum Likelihood Estimation	
$\hat{\eta}$	$\hat{\alpha}$	$\hat{\sigma}$	Peak time	Peak
.0563	.9173	.0646	2025.413	2058.396
(.00910)	(.00556)	(.00026)	(.03383)	(2.27192)

time of observation. Note that, after reaching its peak around 2001, Norway begins to exhibit a decline in production, and our forecasts indicate the exhaustion of its resources close to 2040, provided that a new oil cycle does not begin (something that could happen, eg, after the discovery of new resources). In addition, by using the joint asymptotic distribution of the ML estimators of the parameters and applying the delta method, asymptotic confidence intervals can be obtained for our predictions (note that the conditioned mean function is also a parametric function). Table 6 displays said intervals for scenario 1 (the one based on the real situation today). T6

As for Kazakhstan, it is clear that only scenario 2 can be considered since production peak has not yet occurred. Table 7 contains the results[§]. In this case, production peak was estimated by using (14), considering $s = 2014$ and the production in that year (1632 Mbbl/d). In a similar way to Norway, Figure 5 shows the observed values and the forecasts made until 2040, whereas Table 8 contains the confidence intervals for our predictions. The results show that oil production will continue to grow until a peak time, which will occur around 2025, with a predicted value of 2058.396 Mbbl/d. After that time instant, a clear decrease in production is observed, although predicted values do not yet allow to deduce the time when resources will be exhausted. T7 F5 T8

[§]In this case, the bound of parameter α is 0.9603.

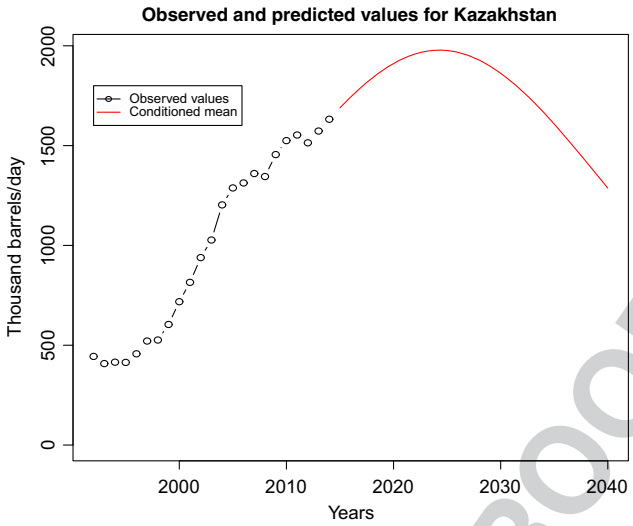

FIGURE 5 Observed and predicted values for Kazakhstan [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 8 Predicted values and confidence intervals of oil production for Kazakhstan from 2015 to 2040

Year	Mean	Lower Limit	Upper Limit
2015	1694.600	1694.375	1694.826
2016	1754.204	1753.716	1754.692
2017	1810.137	1809.350	1810.925
2018	1861.734	1860.610	1862.858
2019	1908.348	1906.852	1909.844
2020	1949.373	1947.471	1951.275
2021	1984.255	1981.917	1986.593
2022	2012.508	2009.708	2015.308
2023	2033.729	2030.446	2037.012
2024	2047.610	2043.830	2051.391
2025	2053.948	2049.662	2059.233
2026	2052.647	2047.855	2057.438
2027	2043.727	2038.436	2049.017
2028	2027.320	2021.545	2033.096
2029	2003.667	1997.429	2009.906
2030	1973.110	1966.437	1979.782
2031	1936.080	1929.008	1943.152
2032	1893.088	1885.657	1900.519
2033	1844.708	1836.963	1852.454
2034	1791.563	1783.551	1799.574
2035	1734.305	1726.078	1742.533
2036	1673.607	1665.215	1681.998
2037	1610.139	1601.635	1618.642
2038	1544.561	1535.996	1553.126
2039	1477.511	1468.933	1486.088
2040	1409.589	1401.045	1418.132

8 | CONCLUSIONS

A diffusion process associated with the Hubbert curve has been proposed to study crude oil production and to predict production peak and peak time. The inferential study is carried out on the basis of discrete sampling via the maximum-likelihood method. Since a complex system of equations appears, which cannot be solved via classical numerical procedures, we suggest using metaheuristic optimization algorithms such as simulated annealing and variable neighborhood search to directly optimize the likelihood function. One of the fundamental problems for the application of these methods is the space of solutions since, in this case, it is continuous and unbounded, which could lead to unnecessary calculation and long algorithm running times. To this end, a strategy is suggested for bounding the space of solutions, which uses the information on the characteristics of the model provided by the sample data. Simulations were performed to test the validity of the bounding method for the space of solutions, showing that it is indeed very useful. The suggested bounding procedure was used in the application of the SA and VNS algorithms to estimate the parameters of the process. Both yielded good results, and the use of a hybrid VNS-SA algorithm led to substantial improvement (in terms of absolute relative errors between the estimated an the real likelihood function) when compared with the results of SA alone.

Finally, 2 applications to real oil production are made. Concretely, we considered oil production data from Norway and Kazakhstan. The former case allowed us to validate the procedure and methodology proposed since peak already took place in 2001. As a matter of fact, by considering the truncated data series prior to 2001, the method predicts a peak time that is very close to the value observed in real life. Since Kazakhstani oil production has not yet reached its peak, we forecast that its growth trend will continue until a peak time, which will occur with high probability around 2025. For the 2 countries, we have considered forecasts of oil production until 2040 based on the values observed in 2014. For Norway, forecasts indicate the exhaustion of its resources at a time that is close to 2040. For Kazakhstan, a clear decrease in production is observed, although predicted values do not yet allow to deduce the time when resources will be exhausted.

ACKNOWLEDGEMENTS

This work was supported in part by the Ministerio de Economía y Competitividad, Spain under grant MTM2014-58061-P and by CNPQ, Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil.

REFERENCES

1. Association for the Study of Peak Oil. http://peak-oil.org/. Accessed April 1, 2017.
2. Hirsch RL, Bezdek R, Wendling R. Peaking of World Oil Production: Impacts, Mitigation and Risk Management. http://www.netl.doe.gov/publications/others/pdf/Oil_Peaking_NETL.pdf. Accessed April 1, 2017.
3. Curtis F. Peak globalization: climate change, oil depletion and global trade. *Ecol Econ*. 2009;69:427-434. https://doi.org/10.1016/j.ecolecon.2009.08.020
4. Lutz C, Lehr U, Wiebe KS. Economic effects of peak oil. *Energy Policy*. 2012;48:829-834. https://doi.org/10.1016/j.enpol.2012.05.017
5. Bakhtiari AM. World oil production capacity model suggests output peak by 2006-07. *Oil Gas J*. http://www2.energybulletin.net/node/147. Accessed April 1, 2017.
6. Deffeyes KS. *Beyond Oil: The View from Hubbert's Peak*. New York, NY: Hill and Wang Publishers; 2006.
7. Towler B. *The Future of Energy*. 1st ed. New York, NY: Academic Press; 2014.
8. Soldo B. Forecasting natural gas consumption. *Appl Energy*. 2012;92:26-37. https://doi.org/10.1016/j.apenergy.2011.11.003
9. Déry P, Anderson B. Peak phosphorus. *Energy Bulletin*. 2007;13. http://www.resilience.org/stories/2007-08-13/peak-phosphorus
10. Vikström H, Davidsson S, Höök M. Lithium availability and future production outlook. *Appl Energy*. 2013;110:252-266. https://doi.org/10.1016/j.apenergy.2013.04.005
11. Hubbert MK. *Nuclear Energy and the Fossil Fuels*. Houston, TX: Shell Development Company, Exploration and Production Research Division; 1956.
12. Hubbert MK. *Techniques of Prediction with Application to the Petroleum Industry*. Houston, TX: Shell Development Company, Exploration and Production Research Division; 1959.
13. Deffeyes KS. *When Oil Peaked*. New York, NY: Hill and Wang Publishers; 2010.
14. Laherrère JH. The Hubbert curve: its strengths and weaknesses. http://dieoff.org/page191.htm. Accessed April 1, 2017.
15. Maggio G, Cacciola G. A variant of the Hubbert curve for world oil production forecasts. *Energy Policy*. 2009;37(11):4761-4770. https://doi.org/10.1016/j.enpol.2009.06.053

16. Nashawi IS, Malallah A, Al-Bisharah M. Forecasting world crude oil production using multicyclic hubbert model. *Energy Fuels*. 2010;24(3):1788-1800. https://doi.org/https://doi.org/10.1021/ef901240p
17. Saraiva TA, Szklo A, Pereira-Lucena AF, Chavez-Rodriguez MF. Forecasting Brazil's crude oil production using a multi-Hubbert model variant. *Fuel*. 2014;115:24-31. https://doi.org/10.1016/j.fuel.2013.07.006
18. Hallock JL, Tharakan PJ, Hall CAS, Jefferson M, Wu W. Forecasting the limits to the availability and diversity of global conventional oil supply. *Energy*. 2004;29:1673-1696. https://doi.org/10.1016/j.energy.2004.04.043
19. Hallock JL, Wu W, Hall CAS, Jefferson M. Forecasting the limits to the availability and diversity of global conventional oil supply: validation. *Energy*. 2014;64:130-153. https://doi.org/10.1016/j.energy.2013.10.075
20. Brandt AR. Testing Hubbert. *Energy Policy*. 2007;35:3074-3088. https://doi.org/10.1016/j.enpol.2006.11.004
21. Reynolds DB. The mineral economy: how prices and costs can falsely signal decreasing scarcity. *Ecol Econ*. 1999;31:155-166. https://doi.org/10.1016/S0921-8009(99)00098-1
22. Reynolds DB. Using non–time–series to determine supply elasticity: how far do prices change the Hubbert curve? *OPEC Rev*. 2002;26(2):147-167. https://doi.org/10.1111/1468-0076.00111
23. Kaufmann RK. Oil production in the lower 48 states: reconciling curve fitting and econometric models. *Resour Energy*. 1991;13(1):111-127. https://doi.org/10.1016/0165-0572(91)90022-U
24. Brecha RJ. Logistic curves, extraction costs and effective peak oil. *Energy Policy*. 2012;51:586-597. https://doi.org/10.1016/j.enpol.2012.09.016
25. Kemp A, Kasim S. An econometric model of oil and gas exploration development and production in the UK continental shelf: a systems approach. *Energy J*. 2003;24(2):113-142. 10.5547/ISSN0195-6574-EJ-Vol24-No2-5
26. Holland SP. Modeling peak oil. *Energy J*. 2008;29(2):61-79. https://doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No2-4
27. Guseo R, Dalla Valle A, Guidolin M. World oil depletion models: price effects compared with strategic or technological interventions. *Technol Forecast Soc Chang*. 2007;74(4):452-469. https://doi.org/10.1016/j.techfore.2006.01.004
28. Michel B. Oil production: a probabilistic model of the Hubbert curve. *Appl Stoch Model Bus Ind*. 2011;27(4):434-449. https://doi.org/10.1002/asmb.851
29. Gutiérrez R, Rico N, Román P, Torres F. Approximate and generalized confidence bands for some parametric functions of the lognormal diffusion process with exogenous factors. *Sci Math Jpn*. 2006;64:313-329.
30. Gutiérrez R, Román P, Torres F. Inference and first-passage-times for the lognormal diffusion process with exogenous factors: application to modelling in economics. *Appl Stoch Model Bus Ind*. 1999;15(4):325-332. https://doi.org/10.1002/(SICI)1526-4025(199910/12)15:4<325::AID-ASMB397>3.0.CO;2-F
31. Rohatgi VK, Saleh AK. *An Introduction to Probability and Statistics*. 3rd ed. New York, NY: John Wiley & Sons; 2015.
32. Kirkpatrick S, Gelatt D, Vecchi MP. Optimization by simulated annealing. *Science*. 1983;220:671-680. https://doi.org/10.1126/science.220.4598.671
33. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller AH. Equation of state calculations by fast computing machines. *J Chem Phys*. 1953;21:1087-1092. https://doi.org/10.1063/1.1699114
34. Mladenović N, Hansen P. Variable neighborhood search. *Comput Oper Res*. 1997;24(11):1097-1100. https://doi.org/10.1016/S0305-0548(97)00031-2
35. Boussaïd I, Lepagnot J, Siarry P. A survey on optimization metaheuristics. *Inf Sci*. 2013;237:82-117. https://doi.org/10.1016/j.ins.2013.02.041
36. Vera JF, Díaz-García JA. A global simulated annealing heuristic for the three-parameter lognormal maximum likelihood estimation. *Comput Stat Data Anal*. 2008;52:5055-5065. https://doi.org/10.1016/j.csda.2008.04.033
37. Abbasi B, Niaki S, Khalife MA, Faize Y. A hybrid variable neighborhood search and simulated annealing algorithm to estimate the three parameters of the Weibull distribution. *Expert Syst with Appl*. 2011;38:700-708. https://doi.org/10.1016/j.eswa.2010.07.022
38. Zoraghi N, Abbasi B, Niaki STA, Abdi M. Estimating the four parameters of the Burr III distribution using a hybrid method of variable neighborhood search and iterated local search algorithms. *Appl Math Comput*. 2012;218:9664-9675. https://doi.org/10.1016/j.amc.2012.03.003
39. Román-Román P, Romero D, Rubio MA, Torres-Ruiz F. Estimating the parameters of a Gompertz-type diffusion process by means of simulated annealing. *Appl Math Comput*. 2012;218:5121-5131. https://doi.org/10.1016/j.amc.2011.10.077
40. Román-Román P, Torres-Ruiz F. A stochastic model related to the Richards-type growth curve. Estimation by means of simulated annealing and variable neighborhood search. *Appl Math Comput*. 2015;266:579-598. https://doi.org/10.1016/j.amc.2015.05.096
41. Busetti F. http://aiinfinance.com/saweb.pdf. 2017.
42. US Energy Information Administration. http://www.eia.gov/beta/international/data/browser. Accessed April 1, 2017.

How to cite this article: da Luz Sant'Ana I, Román Román P, Torres-Ruiz F. The Hubbert diffusion process: Estimation via simulated annealing and variable neighborhood search procedures—application to forecasting peak oil production. *Appl Stochastic Models Bus Ind*. 2018;1–19. https://doi.org/10.1002/asmb.2306

APPENDIX

This appendix summarizes the calculation of the estimation errors for the parameters of the model and for several parametric functions: peak time, peak, and predictions.

In general, on the basis of the properties of the ML estimators and if θ is a multidimensional parameter, it is known that its ML estimator $\widehat{\theta}$ is asymptotically distributed as a normal distribution with mean θ and covariance matrix $I(\theta)^{-1}/N$, where $I(\theta)$ is Fisher's information matrix associated with a sample of size N.

Accordingly, and by applying the delta method, if $g(\theta)$ is a function of the parameter, then

$$\sqrt{N}\left(g(\widehat{\theta}) - g(\theta)\right) \xrightarrow{D} N\left[0; \nabla g(\theta)^T I(\theta)^{-1} \nabla g(\theta)\right],$$

where $\nabla g(\theta)$ represents the derivative vector of $g(\theta)$ with respect to θ.

In the problem at hand, we consider the likelihood (11) resulting from ignoring the initial data (therefore, sample size is $N - d$) and parametric vector $\theta = (\eta, \alpha, \sigma)^T$. Thus, Fisher's information matrix is given by

$$I(\theta) = \frac{1}{\sigma^2}\begin{pmatrix} 4M_1^{\eta,\alpha} & 4M_3^{\eta,\alpha} + 2\frac{X_1^{\eta,\alpha}}{\alpha} & -X_1^{\eta,\alpha} \\ 4M_3^{\eta,\alpha} + 2\frac{X_1^{\eta,\alpha}}{\alpha} & 4M_2^{\eta,\alpha} + \frac{Z_2}{\alpha^2} + 4\frac{X_2^{\eta,\alpha}}{\alpha} & -X_2^{\eta,\alpha} - \frac{Z_2}{2\alpha} \\ -X_1^{\eta,\alpha} & -X_2^{\eta,\alpha} - \frac{Z_2}{2\alpha} & \frac{N-d}{2\sigma^2} + \frac{Z_2}{4} \end{pmatrix},$$

where

$$M_1^{\eta,\alpha} = \sum_{i=1}^{d}\sum_{j=2}^{n_i} \frac{\left(W_{ij}^{\eta,\alpha}/S_{ij}^{\eta,\alpha}\right)^2}{t_{ij} - t_{i,j-1}},\quad M_2^{\eta,\alpha} = \sum_{i=1}^{d}\sum_{j=2}^{n_i} \frac{\left(V_{ij}^{\eta,\alpha}/S_{ij}^{\eta,\alpha}\right)^2}{t_{ij} - t_{i,j-1}},\quad M_3^{\eta,\alpha} = \sum_{i=1}^{d}\sum_{j=2}^{n_i} \frac{V_{ij}^{\eta,\alpha} W_{ij}^{\eta,\alpha}}{t_{ij} - t_{i,j-1}}$$

$$Z_2 = \sum_{i=1}^{d}(t_{in_i} - t_{i1}),\quad X_1^{\eta,\alpha} = \sum_{i=1}^{d}\frac{W_i^{\alpha}}{S_i^{\eta,\alpha}},\quad X_2^{\eta,\alpha} = \sum_{i=1}^{d}\frac{V_i^{\eta,\alpha}}{S_i^{\eta,\alpha}},$$

with

$$\begin{aligned} S_{ij}^{\eta,\alpha} &= (\eta + \alpha^{t_{ij-1}})(\eta + \alpha^{t_{ij}}),\quad S_i^{\eta,\alpha} = (\eta + \alpha^{t_{i1}})(\eta + \alpha^{t_{in_i}}), \\ W_{ij}^{\alpha} &= \alpha^{t_{ij}} - \alpha^{t_{i,j-1}},\quad W_i^{\alpha} = \alpha^{t_{in_i}} - \alpha^{t_{i1}}, \\ V_{ij}^{\eta,\alpha} &= t_{i,j-1}\alpha^{t_{i,j-1}-1}(\eta + \alpha^{t_{ij}}) - t_{ij}\alpha^{t_{ij}-1}(\eta + \alpha^{t_{i,j-1}}), \\ V_i^{\eta,\alpha} &= t_{i1}\alpha^{t_{i1}-1}(\eta + \alpha^{t_{in_i}}) - t_{in_i}\alpha^{t_{in_i}-1}(\eta + \alpha^{t_{i1}}). \end{aligned}$$

The elements in the diagonal of matrix $I(\theta)^{-1}/(N - d)$ provide variances for the estimations of parameters, whereas the delta method provides those of parametric functions (3), (13), and (14), which determine peak time and peak (unconditioned and conditioned). For the purpose of our predictions, function (10) was used.