Bayesian Deconvolution of Oil Well Test Data Using Gaussian Processes

J. Andrés Christen, Bruno Sansó, Mario Santana and Jorge X. Velasco-Hernández

CIMAT, UCSC, CIMAT and IMP

Abstract

We use Bayesian methods to infer an unobserved function that is convolved with a known kernel, using noisy realizations of an observable process. Our method is based on the assumption that the function of interest is a Gaussian process. Thus, the resulting convolution is also a Gaussian process. This fact is used to obtain inferences about the unobserved process, effectively providing a deconvolution method. We apply the methodology to the problem of estimating the parameters of an oil reservoir from well-test pressure data. From a system of linear ordinary differential equations, we write the equation that governs the dynamics of pressure in the well as the convolution of an unknown process with a known kernel. The unknown process describes the structure of the well. This is modeled with a Gaussian process whose mean function is obtained as a linear combination of specific bases. A purposedly designed directional Monte carlo method is used to sample from the posterior distribution of the parameters. Applications to data from Mexican oil wells show very accurate results.

KEYWORDS: Oil well test data; Deconvolution; Bayesian Inference; Inverse Problems; Gaussian Processes; Simulation.

1 Introduction

Pressure and flow rate measurements from oil well testing experiments are used to evaluate the productivity of the well and estimate the properties of the reservoir formation. As described in Kuchuk et al. (2010) when a well is subject to a change in production rate, it creates pressure diffusion in permeable formations. This pressure diffuses from the wellbore into the formation and, by monitoring the transient pressure changes, one can obtain valuable information on properties and characteristics of the reservoir. The resulting data consist of time series of noisy observations. As in many geophysical and petroleum engineering problems, we are in the presence of a problem where direct measurement of the quantities of interest is either impossible or too costly. Thus we resort to a mathematical model that provides pressure as a function of the geological characteristics of the well. We refer to this model as the "forward" model. The problem of interest in this paper is to infer the geological structure using measurements of pressure. Thus we are interested in the "inverse" problem. As seen, for example, in Andrecut (2009); Duru and Horne (2011), the forward model can be stated as the convolution of an unobserved field, say $V(t)$, with a known kernel, say, $k(t)$, where $k(t)$ corresponds to the controlled flow rate of the well. Thus, the pressure, $p(t)$ can be essentially written as

$$p(t) = \int_0^t k(t-z)V(z)dz \tag{1}$$

where k is a known control kernel.

Pressure tests commonly use flow rate as pulses. Therefore, if the well is in a stable state, then, at some time t_e, the well is shut-in. The resulting pressure build up is measured as a function of time. The goal is to use such measurements to infer $V(z)$. This corresponds to important characteristics of the formation that include average reservoir pressure, formation permeability wellbore damage or stimulation, skin effects, productivity indexes, drainage areas, among others (Kuchuk et al., 2010; Gringarten, 2008). More specifically, in this paper we consider the problem of inferring $V(t)$ using a collection of noisy observations $Y(t_1), \ldots, Y(t_n)$ from $p(t)$. Here, k is a step function representing the flow, ie. q_0 before t_e and 0 thereafter.

Deconvolution is a technique for well-test interpretation. It allows the reconstruction of a constant rate drawdown response for all production and shut-in periods combined (Gringarten, 2008; Kuchuk et al., 2010). Well-test analysis and interpretation techniques are problematic since there is a variety of alternative methods and interpretation criteria are heterogeneous, generating concerns about the reliability of the results (Gringarten, 2008). Well-test analysis involves the following three steps (von Schroeter et al., 2001): i) estimating the reservoir properties from data; ii) matching the shape of the response function against a library of curve type, and iii) fitting the parameters of this model to the data. Deconvolution addresses the first step. In the case of a single flow period with constant production rate the response function can be obtained as the derivative of the pressure drop with respect to the logarithm of time. If deconvolution techniques are not applied, the standard method for obtaining such derivative is by numerically differentiating the pressure curve. This presents the problem of error amplification, since the signal, the pressure curve itself, is measured with error (von Schroeter et al., 2001).

Deconvolution is a reliable alternative to numerical differentiation, but it is still affected by measurement error. A number of approaches have been proposed to address the ill-posed problem of deconvolving pressure data. In the petroleum engineering literature the methods by von Schroeter et al. (2001); Levitan (2005); Levitan et al. (2006); Ilk et al. (2007); Onur et al. (2008) have been applied with varied success. The main concern of all of these methods is the problem of estimation robustness given the noisy nature of data.

The mathematical model proposed in Equation (1) is valid under the following hypotheses: i) The wellbore/reservoir system is at equilibrium prior to the test and pressure is constant and equal to p_0 throughout the reservoir; ii) the oil is a single phase flow and it is slightly compressible; iii) wellbore storage and skin effects are constant throughout the whole production history (Onur et al., 2008). Under such conditions, many researchers have addressed the problem of estimation robustness of the reservoir properties, or more generally the problem of deconvolving (1) under noisy measurements (Fasana and Piombo, 1997; Hansen, 2002; Cheng et al., 2005; Ramm and Smirnova, 2005; Pimonov et al., 2009; Andrecut, 2009; Pimonov et al., 2010; Kuchuk et al., 2010).

In this paper we take a Bayesian approach to estimating the reservoir properties. We assume a Gaussian process prior for $V(z)$. The mean of such process is derived from a system of linear ordinary differential equations that governs the dynamics of the pressure in the reservoir. This is presented in the next section. In Section 3 we describe in detail our inferential approach, which is based on the properties of Gaussian processes. In Section 4 we present the results for wells in a Mexican reservoir. In the final Section we discuss our findings and possible extensions of our method.

2 The forward model

A clear understanding of the justification for the foward model in equation (1) is fundamental for effectively inferring V. We start with a system of ordinary differential equations (ODE) for a set of state variables $\boldsymbol{X}$. These variables describe the characteristics of the well. We assume that they are linearly related to pressure. Thus

$$\frac{d\boldsymbol{X}(t)}{dt} = \boldsymbol{A}\boldsymbol{X}(t) + \boldsymbol{b}k(t) \quad \text{and} \quad p(t) = \boldsymbol{c}'\boldsymbol{X}(t) + \alpha k(t).$$

In these equations only the controlled flow k is known. $\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}$ and α are unknown. p is observed, but subject to measurement error. It is well known (Bay, 1999, chapter 6) that the solution to this system of ODE has the form

$$p(t) = \boldsymbol{c}' e^{\boldsymbol{A}(t-t_0)}\boldsymbol{X}(t_0) + \alpha k(t) - \boldsymbol{c}' \int_0^{t-t_0} e^{\boldsymbol{A}(z)}\boldsymbol{b}k(t-z)dz, \tag{2}$$

for $t \in [t_0, T]$. The system is conceived as stable and, accordingly, all the eigenvalues of $\boldsymbol{A}$ are assumed to have nonpositve real parts (Bay, 1999, p. 278).

For suitably designed well test experiments, we have that $p(t)$ is a stable process before a given time t_e. Thus $\boldsymbol{c}' e^{\boldsymbol{A}(t-t_0)} \boldsymbol{X}(t_0)$ is approximately constant, and equal to $p(t_e)$, for $t < t_e$. Thus, Equation (2) takes the form

$$p(t) = p(t_e) + \alpha k(t) - \int_0^{t-t_0} k(t-z) V(z) dz,$$

where $p(t_e)$ is the unknown initial pressure and $V(t) = \boldsymbol{c}' e^{\boldsymbol{A}t} \boldsymbol{b}$. Since $V(t)$ must be a real number, at most one eigenvalue of $\boldsymbol{A}$ may be zero and the rest should be negative (Bay, 1999). We can then write

$$V(t) = \sum_{j=1}^{q-1} \beta_j \exp(-\lambda_j t) \tag{3}$$

for some constants β_j and $\lambda_j > 0$, and q the (unknown) dimension of $\boldsymbol{A}$. Andrecut (2009) suggests a form for the (negative) eigenvalues $\lambda_j = \alpha (j/q)^\gamma$ where typically $\alpha = 10^\gamma T^{-1}$, $\gamma = 3$ and $q = 10$ to 20 eigenvalues are considered. We take a similar approach defining

$$\lambda_j = \frac{j^\gamma}{T - t_1}.$$

Avoiding the eigenvalues to depend on the dimension q makes the fitted models nested, and coeficient estimation remains coherent across various dimensions q. Keeping α and γ fixed maintains the linear Gaussian estimation problem.

Following the discussion in the introduction, we take $k(t) = q_0 1(t < t_e)$, where q_0 is the constant production rate before shut-in. This is an ideal scenario in which the flow is stopped instantaneously at time t_e, and $k(t)$ is known without error.

Under this assumption the forward model is given as

$$p(t) = p(t_e) + \alpha k(t) - \int_{\max(0,t-t_e)}^{t-t_0} V(z)dz. \tag{4}$$

3 Inference

The common approach to analyzing oil well pressure data is to notice that the derivative of $p(t)$ is

$$p'(t) = \begin{cases} -q_0 V(t-t_0) & \text{for } \; t_0 \le t \le t_e \\ -q_0(V(t-t_0) - V(t-t_e)) & \text{for } \; t_e \le t \le T. \end{cases} \tag{5}$$

This derivative is numerically calculated from pressure measurements and is then used to estimate $V(t)$ which in turn is used to plot $tV(t)$ in log-log scale (von Schroeter et al., 2001; Bourdet, 2002; Schlumberger, 2002). This latter graph is the basic tool to interpreting the data. Note, however, that for t greater that t_e, which is the relevant part of the experiment, $V(t)$ is not precisely $p'(t)$. However, assuming (as in the previous Section) that before the experiment the process is stable we may think that $V(t-t_0) \approx 0$ for $t_e \le t \le T$ and therefore

$$V(z) = \frac{p'(z+t_e)}{q_0} \quad \text{for} \;\; 0 \le z \le T - t_e.$$

$zV(z)$ is plotted in log-log scale and since for $t \ge t_e$ the pressure increases, then $zV(z)$ (and $V(z)$) should always be positive. Note that the initial time, t_0, is assumed to be distant enough to have a stable process at t_e, and its actual value is now unimportant, as far as the inference is concerned.

Numerical derivation of noisy data can be very unstable and necessary involves an assumption on the level of smoothness of the implicit function been differentiated.

Here we take a more consistent modeling approach, in which the existence, and smoothness properties, of the derivative process is guarantee through the use of a twice differentiable correlation function in a Gaussian Process model for the pressure data. We explain our approach in the following.

Using the expansion in (3) we can integrate the forward model in (4). For the j-th term in the expansion we have

$$-\int_0^{t-t0} k(t-z)\exp(-\lambda_j z)dz = \frac{q_0}{\lambda_j}\{\exp(-\lambda_j(t-t_0)) - \exp(-\lambda_j \max(0, t-t_e))\}.$$

For convenience we normalize the regressors to have $f_j(t_e) = 0$. Again, the stability assumptions at t_e imply that $\exp(-\lambda_j(t-t_0)) \approx 0$ for $t_e - \epsilon \leq t < t_e$. Accordingly we define the regressors as

$$f_j(t) = \frac{q_0}{\lambda_j}(1 - \exp(-\lambda_j \max(0, t-t_e)),$$

where $j = 1, 2, \ldots, q-1$ and $t \in [t_1, T]$ $(t_1 < t_e)$ and $f_0(t) = 1$. These bases are not orthonormal but are clearly linearly independent ($\sum_{j=0}^{q-1} a_j f_j(t) = 0$ would require $\sum_{j=1}^{q-1} a_j \lambda_j^n = 0$ for all $n = 0, 1, \ldots$ which is impossible for different λ_j's). What remains to be done is to infer the β_j's from data in (3). This will be done in the next Sections.

3.1 The Gaussian process approach

Gaussian process regressions are a powerful tool to making inference about functions and their derivaties, as illustrated in Rasmussen (2003). To estimate $V(t)$, we start with the observation equation. That is, given $p(t)$

$$y_i = p(t_i) + e_i, \quad e_i \sim N(0, \sigma_1^2) \text{ and } cov(e_i, e_j) = 0.$$

We then use a Gaussian process to describe $p(t)$, that is

$$p(t) \sim N(\boldsymbol{f}(t)\boldsymbol{\beta}, \sigma_0^2) \ \text{ and } \ cov(p(t_i), p(t_j)) = \sigma_0^2 \rho_2(t_i, t_j), \tag{6}$$

where $\boldsymbol{f}(t) = (1, f_1(t), f_2(t), \ldots, f_{q-1}(t))$. In other words, pressure time series is a Gaussian Process with covariance function $\sigma^2 \rho_2(t_i, t_j)$ (ρ_2 is a correlation function to be defined below, ρ_2 will depend on a small number of parameters). We further assume that $p(t)$ and e_i are independent for all t. That is, $y_i = p(t_i) + e_i$ with $cov(y_i, y_j) = \sigma_0^2 \rho_2(t_i, t_j)$ for $t_i \neq t_j$. We define $\sigma^2 = Var(y_i) = \sigma_0^2 + \sigma_1^2 = cov(y_i, y_i)$. Our model can be expressed in a standard form as

$$\boldsymbol{Y} = \boldsymbol{F}\boldsymbol{\beta} + \boldsymbol{\epsilon} \ \text{ where } \ \boldsymbol{\epsilon} \sim MN_n(\boldsymbol{0}, \sigma^2 \boldsymbol{K}). \tag{7}$$

Thus, $\boldsymbol{Y} = (y_i)$ is a Gaussian process with mean $\boldsymbol{F}\boldsymbol{\beta}$ and variance-covariance matrix $\sigma^2 \boldsymbol{K}$; $\boldsymbol{F}$ is the design matrix obtained by stacking the n regressor vectors $\boldsymbol{f}(t_i)$. Let $\boldsymbol{K} = (\rho_2^0(t_i, t_j)), i, j = 1, 2, \ldots, n$ be a $n \times n$ correlation matrix with

$$\rho_2^0(t_i, t_j) = \begin{cases} 1 & \text{for } \ t_i = t_j \\ \eta \rho_2(t_i, t_j) & \text{otherwise,} \end{cases} \tag{8}$$

where $\eta = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2} = \frac{\sigma_0^2}{\sigma^2}$. This is a covariance function with a discutinuity arising from the observational error, where $\frac{\sigma_1^2}{\sigma_0^2 + \sigma_1^2}$ is the "sill" and η is the limit at zero time distance. While the correlation for the underlaying pressure Gaussian process ρ_2 is conceived continuos at zero, ρ_2^0 is not. Equation (7) can be used as a likelihood to infer the parameters that define the mean and covariance functions of $p(t)$, and, consequently, the predictive distribution of $p(t)$ and $V(t)$. This is, in esense, the approach suggested in Rasmussen (2003) for inference on the derivative of a curve.

A known feature of Gaussian processes is that, under enough regularity conditions, they are mean square differentiable. More precisely, $p'(t)$ would also be a Gaussian process (and therefore $V(t)$) if the the covariance function of $p(t)$ is differenciable and $\frac{\partial^2 cov(t_i,t_j)}{\partial t_i \partial t_j}\Big|_{t_i=t}\Big|_{t_j=t}$ is finite for all $t \in [t_0, T]$. In which case $p'(t)$ is a Gaussian process with $cov(p'(t_i), p'(t_j)) = \frac{\partial^2 cov(t_i,t_j)}{\partial t_i \partial t_j}$; and $cov(p(t_i), p'(t_j)) = \frac{\partial cov(t_i,t_j)}{\partial t_j}$ (see Christakos, 1992, p. 57-75). We notice in passing that, for $p(t)$ to be mean square differentiable, its covariance function can not have a "nugget", as this would introduce a discontinuity.

Holsclaw et al. (2011) argue that a better estimate of the derivative of a function, and the associate uncertainty bands, is acheived by imposing a Gaussian prior on the derivative itself, rather than the original function. That is, assume defining $V(t)$ (ie. $p'(t)$) as a Gaussian process itself and work upwards in the opposite direction. Thus, let

$$V(t) \sim N(\nabla \boldsymbol{f}(t)\boldsymbol{\beta}, \sigma^2/a) \quad \text{and} \quad cov(V(t_i), V(t_j)) = \frac{\sigma^2}{a}\rho_0(t_i, t_j),$$

ρ_0 is a correlation function (we require an arbitrary matching constant a, as will be clear below). Then, as stated in general terms in Gihman and Skorohod (1979) p. 251, for a suitable kernel $q_c(t, z)$, $p(t) = p(t_e) - \int_{-\infty}^{\infty} q_c(t, z)V(z)dz$ is a Gaussian process with mean function $\boldsymbol{f}(t)'\boldsymbol{\beta}$ and covariance function

$$cov(p(t_i), p(t_j)) = \frac{\sigma^2}{a}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q_c(t_i, z_1)q_c(t_j, z_2)\rho_0(z_1, z_2)dz_1dz_2,$$

$(t_i, t_j \in [t_e, T])$ if and only if this last expression is finite for all $t \in [t_e, T]$. From (4) we see that in our case $q_c(t, z) = q_0 I_{(t-t_e, t-t_0)}(z)$; since this creates a bounded support, the latter requirement is fulfilled. Also, the covariance function of $p(t)$ results in

$$cov(p(t_i), p(t_j)) = \frac{\sigma^2 q_0^2}{a}\int_{t_j-t_e}^{t_j-t_0}\int_{t_i-t_e}^{t_i-t_0} \rho_0(z_1, z_2)dz_1.dz_2,$$

This can be used to obtain a likelihood for the mean and covariance parameters similar the one obtained from Equation (7).

It is interesting to notice that, under the stationarity assumption, the correlation function for $V(t)$ should have the form $\rho_0(t_i, t_j) = \gamma_0(|t_i - t_j|)$ (where γ_0 is an isotropic correlation function) and the integrated process will, in general, not be stationary. Nevertheless, in our case we have that $cov(p(t_i), p(t_j)) = \frac{\sigma^2 q_0^2}{a} h(|t_i - t_j|)$ with

$$h(d) = -2\gamma_2(d) + \gamma_2(d + (t_e - t_0)) + \gamma_2(-d + (t_e - t_0)) - 2\gamma_1(0)(-d + (t_e - t_0)),$$

where $\gamma_1(d)$ and $\gamma_2(d)$ are first and second antiderivatives of $\gamma_0(d)$ (ie. $\gamma_2'(d) = \gamma_1(d)$ and $\gamma_1'(d) = \gamma_0(d)$). Therefore $p(t)$ has an isotropic covariance for $t \in [t_e, T]$ and it *is also stationary*.

To obtain a useful parametrization of $h(d)$ we recall that t_0 is a distant time chosen so that the process is stable at $t_e - \epsilon$ and therefore we are assuming that $T - t_e < t_e - t_0$ which leads to $-d + (t_e - t_0) > 0$. We define $(t_e - t_0) = \alpha R$ where $\gamma_0(R) \approx 0.05$ (the correlation range) and $\alpha = \frac{T - t_e}{\alpha_1 R}$ and $\alpha_1 \in (0, 1]$; α_1 is a unit less bounded parameter. Therefore the correlation function is parametrized in terms of R and α_1 and is given by $h(d) = -2\gamma_2(d) + \gamma_2(d + \alpha R) + \gamma_2(-d + \alpha R) - 2\gamma_1(0)(-d + \alpha R)$; $0 < d < \alpha R$. Finally, from (6) we require $\sigma_0^2 \rho_2(d) = \frac{\sigma^2 q_0^2}{a} h(d)$, that is $\sigma_0^2 = \frac{\sigma^2 q_0^2}{a} h(0)$ or $a = \frac{\sigma^2 q_0^2}{\sigma_0^2} h(0)$ *ie*

$$\rho_2(d) = \frac{h(d)}{h(0)}.$$

This is the resulting correlation function for the pressure process $p(t)$, which is parametrized in terms of the range for γ_0, R, and the α_1 parameter above. We also need to define the limit at zero $\eta = \frac{\sigma_0}{\sigma_0 + \sigma_1}$ (representing the expected noise-to-signal ratio) in (8) to complete our definition for the correlation structure for the processes y_i, $p(t)$ and $V(t)$. Assuming that these parameters are fixed, our inference focuses on $\tau = \sigma^{-2}$ and $\boldsymbol{\beta}$.

To make inference about $V(z)$ using only data on $p(t)$, we need the correlation between $p(t_i)$ and $V(t_j)$. As mentioned above we have $cov(p(t_i), p'(t_j)) = -\sigma_0^2 \frac{h'(|t_i - t_j|)}{h(0)}$. From (5) we see that $cov(p(t_i), p'(t_j)) = q_0 cov(p(t_i), V(t_j - t_e)) - q_0 cov(p(t_i), V(t_j - t_0))$, but $V(t - t_0) \approx 0$, thus the second term will be far less significant than the first. Therefore we assume

$$cov(p(z_i + t_e), V(z_j)) = \sigma_0^2 \, \frac{-h'(|z_i - z_j|)}{q_0 h(0)},$$

for $z_i, z_j \in [0, T - t_e]$, and indeed $cov(V(z_i), V(z_j)) = \sigma_0^2 \, \frac{\gamma_0(|z_i - z_j|)}{q_0^2 h(0)}$.

In our example we use the Matèrn correlation function with smoothness parameter equal to 1.5. This function can be expressed as $M(d) = (d + 1)e^{-d}, \quad d \geq 0$. It is a convenient correlation function, computationally simple to evaluate, that corresponds to a mean square differentiable processes. Following our previous discussion, we will use the parametrization

$$\gamma_0(d) = M\left(c \, \frac{d}{R}\right). \tag{9}$$

where $c = 4.75$ is such that $\gamma_0(R) \approx 0.05$; that is, this is a standardized correlation with "range" R. From (9), we have $\gamma_1(d) = \frac{R}{c} M_1 \left(c \, \frac{d}{R}\right)$, $\gamma_1(0) = -2\frac{R}{c}$ and $\gamma_2(d) = \frac{R^2}{c^2} M_2 \left(c \, \frac{d}{R}\right)$, $\gamma_2(0) = 3\frac{R^2}{c^2}$. Figure 1 illustrates the shapes of the correlation function that results from the integrated approach.

3.2 Prior, posterior and predictive distributions

Let $\boldsymbol{\mu}, \boldsymbol{A}^{-1}, a$ and b be NormalGamma parameters of the prior for $\boldsymbol{\beta}$ and $\tau = \sigma^{-2}$. Using the Cholesky decomposition one may efficiently calculate $\hat{\boldsymbol{\beta}} = (\boldsymbol{F}'\boldsymbol{K}^{-1}\boldsymbol{F} + \boldsymbol{A})^{-1}(\boldsymbol{F}'\boldsymbol{K}^{-1}\boldsymbol{Y} + \boldsymbol{A}\boldsymbol{\mu})$, $q \times 1$ which is known as the generalized LE estimator. The marginal posterior distribution of $\boldsymbol{\beta}|\boldsymbol{Y}$ is a multivariate non-central student-t distribution with $2a + n$ degrees of freedom, mean $\hat{\boldsymbol{\beta}}$, and scale matrix $\boldsymbol{B} =$

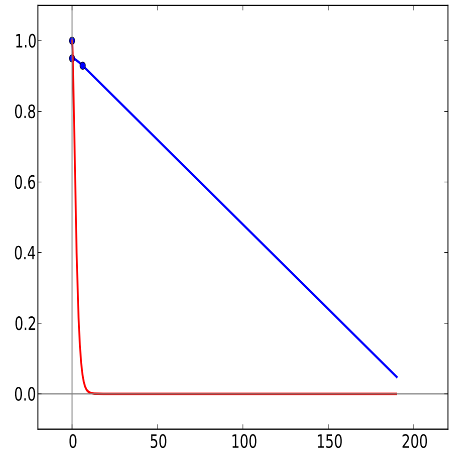


Figure 1: $\rho_2^0(d)$ correlation function for the measured pressure process $p(t)$, with $\eta = 0.95$ (sill $= 0.05$) and range $R = (T - t_e)/16$ and $\alpha_1 = 0.5$ (blue); $T - t_e = 100$. Corresponding Matèrn correlation $\gamma_0(d)$ (red), proportional to the correlation of the impulse response process $V(z)$.

$\left(\boldsymbol{F}'\boldsymbol{K}^{-1}\boldsymbol{F} + \boldsymbol{A}\right)(2a+n)\, b_p^{-1}$, where $b_p = b + \frac{1}{2}\left[\boldsymbol{y}'\boldsymbol{K}^{-1}\boldsymbol{y} + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu} - \hat{\boldsymbol{\beta}}' B\hat{\boldsymbol{\beta}}\right]$. If we wish to calculate the predictive distribution of the random variable Z, that is, the conditional distribution $Z|\boldsymbol{Y}$, assuming jointly $\boldsymbol{Y}, Z|\boldsymbol{\beta}, \sigma^2$ is also a multivariate normal, let $\kappa(t_i) = \sigma^{-2} cov(y_i, Z|\boldsymbol{\beta}, \sigma^2)$, $v^2 = \sigma^{-2} var(Z|\boldsymbol{\beta}, \sigma^2)$ and let $\boldsymbol{m}$ be a $q \times 1$ vector such that $E(Z|\boldsymbol{\beta}, \sigma^2) = \boldsymbol{m}'\boldsymbol{\beta}$. Then the predictive distribution of Z is a noncentral student-t distribution with $2a + n$ degrees of freedom, mean

$$E(Z|\boldsymbol{Y}) = m_z = \boldsymbol{\kappa}'\boldsymbol{K}^{-1}\boldsymbol{Y} + ut_z^{-1}\hat{\boldsymbol{\beta}}'\boldsymbol{T}_z^{-1}\boldsymbol{c},$$

and variance

$$var(Z|\boldsymbol{Y}) = \left((2a+n-2)t_z\right)^{-1} c_0,$$

where $\boldsymbol{c} = \boldsymbol{m} - \boldsymbol{F}'\boldsymbol{K}^{-1}\boldsymbol{\kappa}$, $\boldsymbol{\kappa} = (\kappa(t_i)), i = 1, 2, \ldots, n$, $u = \left(v^2 - \boldsymbol{\kappa}'\boldsymbol{K}^{-1}\boldsymbol{\kappa}\right)^{-1}$, $\boldsymbol{T}_z = u\boldsymbol{c}\boldsymbol{c}' + B$, $t_z = u\left(1 - u\boldsymbol{c}'\boldsymbol{T}_z^{-1}\boldsymbol{c}\right)$, and $c_0 = 2b_p + \left[\left(\boldsymbol{\kappa}'\boldsymbol{K}^{-1}\boldsymbol{Y}\right)^2 - m_z^2\right] t_z + \hat{\boldsymbol{\beta}}'\left(\boldsymbol{B} - \boldsymbol{B}\boldsymbol{T}_z^{-1}\boldsymbol{B}\right)\hat{\boldsymbol{\beta}} + 2u\boldsymbol{\kappa}'\boldsymbol{K}^{-1}\boldsymbol{Y}\boldsymbol{c}'\boldsymbol{T}_z^{-1}\boldsymbol{B}\hat{\boldsymbol{\beta}}$.

For the predictive distribution of $p(t)$ we have $\kappa(t_i) = \sigma^{-2}\rho_2(|t - t_i|)$, $v^2 = \eta$ and $\boldsymbol{m} = \boldsymbol{f}(t)'$. For the predictive distribution of $V(t - t_e)$ we have $\kappa(t_i) = \frac{-h'(|t-t_i|)}{q_0 h(0)}\eta$, $v^2 = \frac{\eta}{q_0^2 h(0)}$ and $\boldsymbol{m} = \nabla \boldsymbol{f}(t)'$, for $t \in [t_1, T]$.

3.3 Model Selection

Selecting among the possible number of bases q may be easily tackled using the Bayesian Model Selection ideas (Hoeting et al., 1999) that basically include in the parameter space the now unknown parameter q, ie. the number of bases in our model. These in fact represent an increasing number of eigenvalues for the matrix $\boldsymbol{A}$ and therefore it is pointless to experiment with other but the full sequence of bases up to bases $q - 1$ (as opposed to usual variable selection in linear models, Hoeting et al., 1999). We also would like to include in the model selection the correlation parameters R and α_1. Given a discrete grid of the total number of base functions q, R and α_1 the marginal likelihood of each model configuration is

$$l(\boldsymbol{Y}|q = h, R = r, \alpha_1 = a) = \int\int f(\boldsymbol{Y}|\boldsymbol{F}_h, \boldsymbol{K}_{r,a}, \boldsymbol{\beta}, \tau) f(\boldsymbol{\beta}, \tau) d\boldsymbol{\beta} d\tau$$

where $f(\boldsymbol{y}|\boldsymbol{F}_h, \boldsymbol{K}_{r,a}, \boldsymbol{\beta}, \tau)$ is our Multivarite Normal model with h bases and correlation parameters r, a, that is $\boldsymbol{Y} \sim MN_h(\boldsymbol{F}_h\boldsymbol{\beta}, \tau^{-1}\boldsymbol{K}_{r,a})$ and $f(\boldsymbol{\beta}, \tau)$ is the Normal-Gamma prior for $\boldsymbol{\beta}, \tau$. The former expression is proportional to

$$\frac{(2\pi)^{h/2}}{|\boldsymbol{F}_h'\boldsymbol{K}_{r,a}^{-1}\boldsymbol{F}_h + \boldsymbol{A}|^{1/2}}(b + 0.5S^2_{h,r,a})^{-(a+n/2)}.$$

We calculate $l(\boldsymbol{Y}|q = h, R = r, \alpha_1 = a)$ for a series of examples and for $q = 4, 5, \ldots, 25$, $R = (T - t_e)2^{-k}$ and $\alpha_1 = 2^{-k}$ for $k = 1, 2, 3, 4$; see Figure 2. We plot $\log l(\boldsymbol{Y}|q = h, R = r, \alpha_1 = a) + C$ for an arbitray constant C (we choose C to make the minimum log likelihood at -600). After less than 23 bases $\boldsymbol{F}_h'\boldsymbol{K}_{r,a}^{-1}\boldsymbol{F}_h + \boldsymbol{A}$ becomes singular, meaning that the bases become numerically linearly dependent

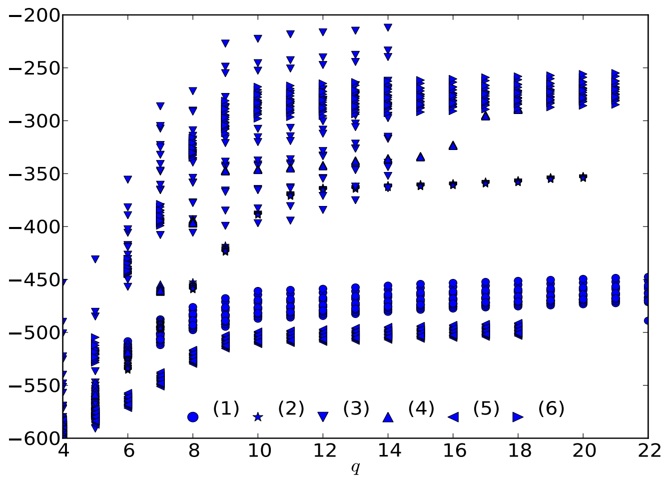


Figure 2: $\log P(q = h, R = r, \alpha_1 = a|\boldsymbol{Y}) = C + \log l(\boldsymbol{Y}|q = h, R = r, \alpha_1 = a)$ (all shifted to a minimum of -600) for $q = 4, 5, \ldots, 25$, $R = (T - t_e)2^{-k}$ and $\alpha_1 = 2^{-k}$ for $k = 1, 2, 3, 4$ and a constant prior for q. The maximum posterior probabilities for all R and α_1 combinations are depicted at each q and for each data set from wells (1) to (6) (owned by Pemex, in Tabasco, Mexico, see Section 4.1 for more details on these data sets).

(since the λ_j's become similar). In all cases there is a sharp increase in log likelihood until about $q = 10$ where they level out. We see that in most cases $q = 10-12$ should be enough (this has been also noticed by more heuristic means in Andrecut, 2009). With respect to the correlation parameters we see that in all cases $R = (T - t_e)/16$ and $\alpha_1 = 2$ have the highest likelihood. We use this guidelines for the examples presented in Sections 4.1 and 4.2.

3.4 Positivity Constraint of $V(z)$: MCMC

As discussed in Section 3, $V(z)$ must be positive for every $z > 0$ and strictly speaking we have to consider only the values of $\boldsymbol{\beta}$ that satisfy the constraint $V(z) > 0$. In many real pressure test data examples, the error is sufficiently low, resulting in a posterior distribution that is basically fully supported where $V(z) > 0$. That is, without explicitly restricting the support of $\boldsymbol{\beta}$, $V(z) > 0$ and we may proceed as in Section 3.2 with a fully analytic method. In examples where such unconstrained analytic calculation results in a posterior predictive distribution of $V(z)$ with a noticeable portion below zero, we need to explicitly restrict $V(z) > 0$, which results in a Truncated Multivariate Normal distribution for $\boldsymbol{\beta}|\tau$, with a complex support, and we to resort to a MCMC method for its analysis.

The support of this distribution is difficult to establish analytically. Thus, we resort to numerically checking if $V(z_i) > 0$, for a preset number of times z_i's, for any $\boldsymbol{\beta}$ being considered within our sampling scheme. In Section 4.2 we will illustrate the need to impose the constraints on $\boldsymbol{\beta}$ in order to capture the complexities of various systems and avoid inferences that lead to negative $V(z)$. We use a novel Optimal Direction Gibbs (ODG)sampler to simulate from the truncated Normal distribution. This algorithm is a generalized Gibbs sampler that simulates from the full conditional distribution of a set of directions to minimize the mutual information (correlation) between steps of the MCMC (Christen et al., 2012).

To simulate from the predictive distribtion of $V(z)$, defined by

$$f(V(z)|\boldsymbol{Y}) = \int\int f(V|z, \boldsymbol{Y}, \boldsymbol{\beta}, \tau) f(\boldsymbol{\beta}, \tau|\boldsymbol{Y}) d\boldsymbol{\beta} d\tau,$$

we simulate $(\boldsymbol{\beta}^{(i)}, \tau^{(i)})$ from the (truncated) posterior distribution of $f(\boldsymbol{\beta}, \tau|\boldsymbol{Y})$ using the ODG, and in turn, for every z_i, we simulate $V(z_i)$ from $f(V|z_i, \boldsymbol{Y}, \boldsymbol{\beta}^{(i)}, \tau^{(i)})$, which is an univariate truncated normal distribution. Note that simulating from an univariate truncated normal is straightforward using an efficient $\Phi(x)$ inversion method (Wichura, 1988) which rivals other *ad hoc* methods in terms of computational efficiency (the former is in fact the current default in R for simulating from the standard normal).

4 Results

In this section we present our results using pressure data from wells in a naturally fractured reservoir of Mexico. As mentioned in Section 1, after estimation of the logarithmic derivative curve from data, one has to match the shape of this response function to a type curve. These curve types characterize different flow regimes that can be roughly classified as "onset", "intermidiate phase" and "late" periods. Kuchuk et al. (2010) list several flow regimes for the different time periods. For example, at the pressure experiment onset the flow regime may be radial, linear, spherical, elliptical; at a more intermidiate phase it can be channel linear, radial composite, dual porosity, and later they may become steady-state, pseudo-steady state, among others. We refer the reader to Kuchuk et al. (2010, p. 48) for an ample discussion and more details on the characteristics of these flow regimes.

4.1 A well with fractal pressure behaviour

Well A (owned by Pemex, in Tabasco, Mexico) is located in the Jujo-Tecominoacán oil field where its main production arises from the Upper Jurassic Kimeridgian formation. This field is a naturally-fractured reservoir with an average depth of about 5000 m. The oil produced there is light in an area of c. $70 Km^2$. Other similar pressure profile data from the same oil field (with the same alleged fractal characteristics) were used in the model selection discussion in Section 3.3; see Figure 2.

Flow regimes depend on several factors among which the internal reservoir geology plays a prominent role. Reservoir heterogeneity has a large impact on fluid flow and therefore on pressure profiles in shut in experiments. In particular, for naturally fractured reservoirs (a special case of a heterogeneous reservoir), fluid flows across complex rock fracture structures (as opposed to flow in more standard homogeneous media) and may be characterized as possessing a "fractal" structure

(Flamenco-Lopez and Camacho-Velazquez, 2003). Well pressure tests conducted in naturally fractured reservoirs or, more generally, in reservoirs with pronounced heterogeneity, result in estimated pressure profiles (specifically in the log-log plot of $zV(z)$) that can not be properly characterized from the library of type curves, based on the standard flow regimes briefly listed in the previous Section.

Several approaches have been developed to account for the impact of a fractal geometry in a reservoir, in order to study the characteristics of the pressure curve and pressure derivative (Chang and Yortsos, 1990). However results are not conclusive since most of these models postulate power law relations in porosity and permeability increasing the number of parameters to estimate. We postulate that deconvolution of the form developed here may provide an interesting alternative to help in the identification of the flow regime in the case of wells with fractal pressure behaviour.

In Figure 3 we show the results of our estimation. In contrast in Figure 4 we present the estimation using numerical differentiation as implemented in a popular commercial software for well tests analysis. A key feature of a fractal transient response is that the log-log plot of the pressure derivative versus time is linear (Acuna et al., 1992). That is, initially the log-log plot of the $zV(z)$ function seems linear, typically between $\log_1 0$ times of -1 to 0 (first cycle) and 0 to 1 (second cycle). This transient linear response can be clearly seen in our estimation in Figure 3(d).

Through independent tests (JXVH, personal communication) we have found evidence of a flow regime in the pressure curve that produces a logarithmic derivative with a linear transient phase in an interval of at least one logartithmic cycle. Our deconvolution of the signal is consistent with this finding, as oppose to the traditional method estimates shown in Figure 4; the linear transient phase is totally blurred in the latter. The improvement is remarkable and the practical implications are interesting.

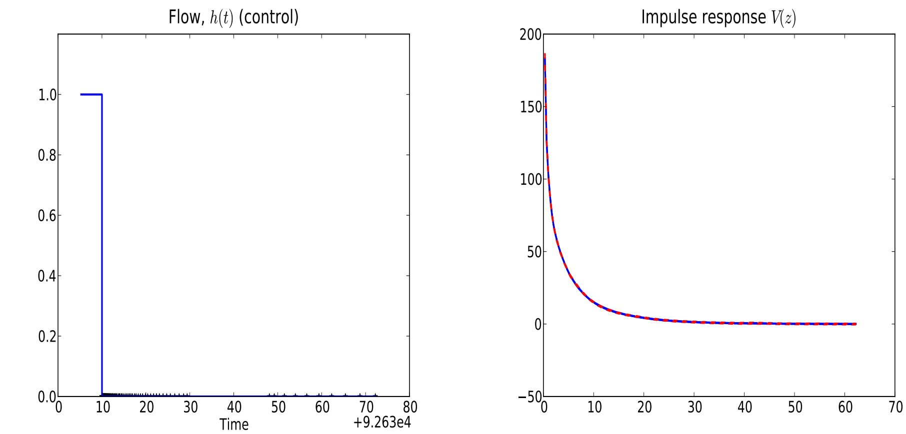
We now have a robust curve estimation against which we can validate flow regimes that would take into account a fractal structure in the reservoir. Moreover, we have probability renges that quantify our unciertanty regarding the estimation of $zV(z)$, namely the predictive distribution of $V(z)$ and $p(t)$ depicted in the red dashed lines in Figure 3. This provides a formal indication to where the flow regime is safely estimated and the proper time window to establish the regime flow, if at all possible.

Deconvolution for this well indicates the existence of a pressure behavior that can be described as fractal (Acuna et al., 1992). The technique developed here is now being used in the actual analysis of this type of behavior in mature oil fileds in Mexico. A forthcoming paper will present the general conclusions regarding the application of our methodology in a comprehensive suit of examples.

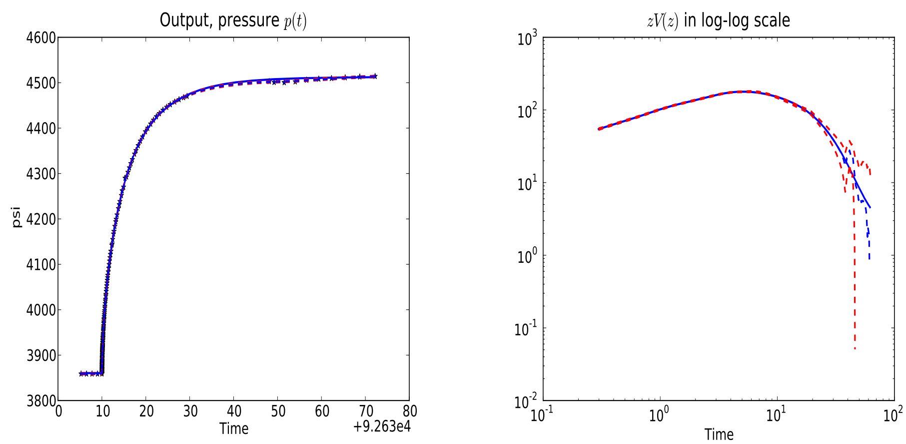
4.2 MCMC for positivity constraint

The pressure profile data set for well B (owned by Pemex, Mexico) has an extreme jump and the unrestricted solution has a negative section for V. For $R = (T - t_e)/16$ and $\alpha_1 = 0.5$ we consider the likelihood of each model $J = 3, 4, \ldots, 15$; the log likelihoods are plotted in Figure 5(a). There is a sharp increase at $J = 5$ bases after which the log likelihood levels out. The analysis becomes quite numerically unstable for $J \geq 5$ and calculations can only be carried out reliabily for $J < 9$ bases. We therefore choose to experiment the fit with positivity constraits for $J = 5$. The unconstrained fit with $J = 5$ bases for the pressure data and the corresponding log-log plot of $zV(z)$ are presented in Figures 5(b) and (c).

The optimal direction Gibbs MCMC is run with this data and 5,000 simulations are drawn from the corresponding truncated normal and predictive distributions, as explained in Section 3.4. The results are shown in Figure 6.



(a) (b)



(c) (d)

Figure 3: Section of test data for well A. Output flow (control) (a), generalized least squares estimates (solid blue), predictive mean (dashed blue) and predictive $\pm 2\sigma$ bands (dashed red) for $V(z)$ (b), $p(t)$, along with the measured pressure data, (c) and $zV(z)$ in log-log scale (d).

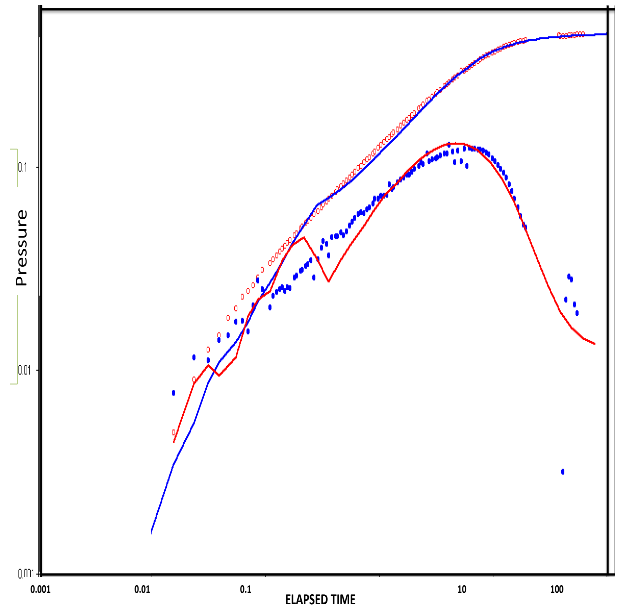


Figure 4: Pressure data for well A, as in Figure 3 (red dots), in log-log scale. Expert estimation of $p(t)$ (blue curve) and $zV(z)$ (red curve), in log-log scale. Blue dots represent a numerical "derivative" obtained by interpolating the pressure data (red dots), this "derivative data" is then fitted to obtain $zV(z)$ using simple heuristics, as implemented in a well known commercial software.

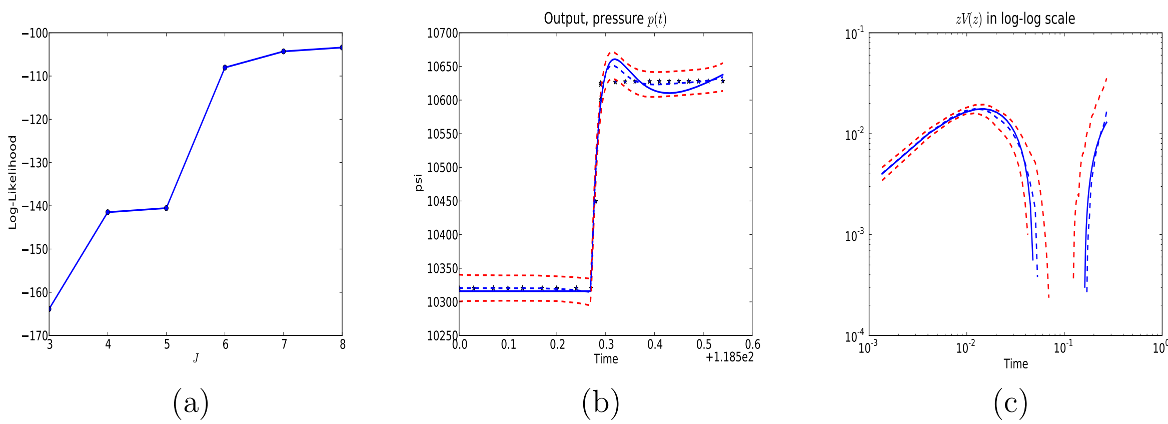
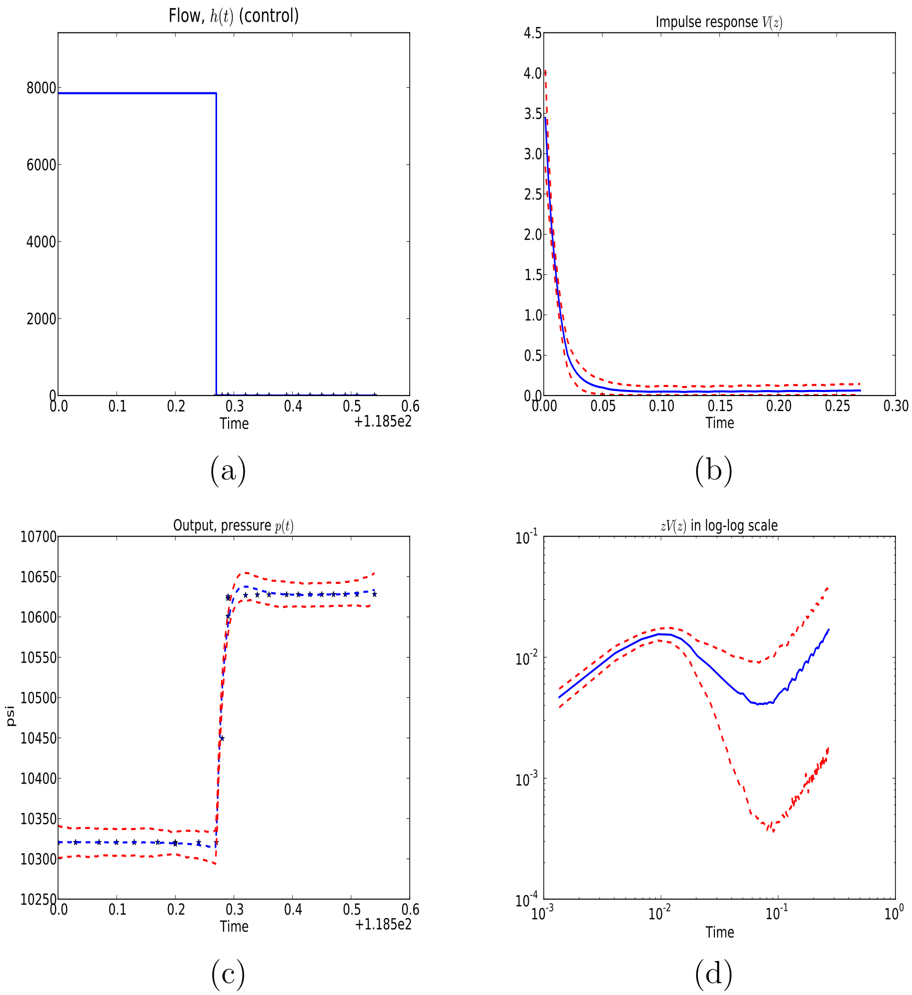


Figure 5: Pressure profile data set for well B, (a) Log probabilities of each model with $R = (T - t_e)/16$ and $\alpha_1 = 0.5$ and $q = 3, 4, \ldots, 8$ bases, see Section 3.3. (b) Pressure fit for the unconstrained case and (c) the corresponding log-log plot of $zV(z)$. Note that the decay in the generalized least squares fit (blue line) at the top of the pressure jump corresponds to a negative (derivative) $V(z)$ and therefore the undefined section in the log-log plot of $zV(z)$. The prior (and posterior) distribution for $\boldsymbol{\beta}$ should be constrain to ensure that $V(z)$ remains positive, leading to a truncated normal distribution (see Figure 6).



(a) (b) (c) (d)

Figure 6: Pressure profile data for well B analyzed with a positivity constrain for $V(z)$ using MCMC. Output flow (control) (a), predictive mean (solid blue) and predictive 5% and 95% bands (dashed red) for $V(z)$ (b), $p(t)$, along with the measured pressure data, (c) and $zV(z)$ in log-log scale (d); this last plot is not defined in the whole time range since solutions are restricted to maintain $V(z)$ positive. Since $V(z)$ is positive $p(t)$ is increasing for every fixed $\boldsymbol{\beta}$, however the predictive mean of $p(t)$ has an additional term accounting for the correlation in the model, which results in a slight perturbation in the predictive distribution of p around the pressure jump in (c).

5 Discussion

The analysis of pressure profiles in shut in oil well experiments is crucial in understanding the geology of oil fields which eventually leads to establishing probable reserves, well production policies and total prospective output. A formal analysis of the pressure profile data and the impulse response function $V(z)$ is therefore much needed. The novel approach taken here, namely, modeling the pressure function with a Gaussian process, enables to formally infer the $V(z)$ function, establishing clearly the assumptions needed to be assumed on $V(z)$. For the precise structure of the problem, if $V(z)$ is assumed stationary, $p(z)$ results also stationary. Moreover, our formal approach permits quantifying our uncertainty in estimating $p(t)$ and $V(z)$ (by calculating the corresponding predictive distributions) and potentially this could lead to 1) a more informative analysis of the impulse response function $V(z)$ in a time window with low levels of uncertainty and 2) better prospective analysis of the well production regime and potential.

The functional form of the correlation function and its parameter values are an important part of any Gaussian process data analysis approach. We used the Matèrn correlation function with smoothness parameter fixed to 1.5. This guarantees the desired smoothness. Moreover, it has a simple analytic form and so do its first and second derivatives and antiderivatives. The correlation parameters, namely R and α_1, were established by a model comparison approach. We believe this is a reasonable and effective strategy, that avoids the use of MCMC schemes that are difficult to tune.

Still a more integrated approach is needed, in the sense of a formal definition of the flow profiles and a systematic (model based) assignment of and estimated impulse response function to a specific flow profile. This in turn could be embedded in a general oil production inference method for the well. However, a quantitative understanding of the mapping between the inner geology and the flow profiles should first be established in terms of, for example, more precise functional forms for the impulse response functions (the log-log plots of $zV(z)$ are only characterized in broad qualitative terms) before a model based approach is attempted. Our approach here presented therefore represents the first steps towards a mathematical (probabilistic) interpretation of pressure profile data, in an important field where a rather heuristic and qualitative strategy is currently in practice.

6 Acknowledgements

JXVH acknowledges partial financial support from SENER-CONACyT grant 143935 to the Mexican Petroleum Institute and Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo, grant P711RT0278. JAC acknowledges partial financial support from CONACyT (Ciencia Básica) grant 128477, JAC and BS acknowledges partial financial support from a UC-MEXUS joint grant.

References

Acuna, A., Ershaghi, I., Yortsos, Y. C., Program, P. E., and Angeles, L. (1992). Fractal analysis of pressure transients i n the geysers geothermal field. *Proceedings of the 17th Workshop on Geothermal Reservoir Engineering*, SGP-TR-141:87–93.

Andrecut, M. (2009). Pressure rate deconvolution methods for well test analysis. *Modern Physics Letters B*, 23(8):1027–1051.

Bay, J. (1999). *Fundamentals of linear state space systems*. McGraw-Hill.

Bourdet, D. (2002). Well test analysis : the use of advanced interpretation models . In Cubitt, J., editor, *Handbook of Petroleum Exploration and Production*, Handbook of Petroleum Exploration and Production, chapter 3, page 416. Elsevier, Amsterdam.

Chang, J. and Yortsos, Y. C. (1990). Pressure-transient analysis of fractal reservoirs. *SPE Form Eval., 5*, page 631.

Cheng, Y., McVay, D., and Lee, W. (2005). A practical nonlinear regression technique for horizontal well test interpretation. *Petroleum Science and Technology*, 23:241–369.

Christakos, G. (1992). *Random Field Models in Earth Sciences.* Academic Press, Inc.

Christen, J., Fox, C., Pérez-Ruiz, D., and Santana-Cibrian, M. (2012). On optimal direction gibbs sampler. *TBA*.

Duru, O. O. and Horne, R. N. (2011). Simultaneous interpretation of pressure, temparature, and flow-rate data using Bayesian inversion methods. *SPE Reservoir Evaluation and Engineering*, pages 225–238.

Fasana, A. and Piombo, B. (1997). Identification of linear mechanical systems by deconvolution techniques. *Mechanical systems and signal processing*, 11(3):351–373.

Flamenco-Lopez, F. and Camacho-Velazquez, R. (2003). Determination of fractal parameters of fracture networks using pressure-transient data. *SPE Reservoir Evaluation and Engineering*, SPE82607:39–47.

Gihman, I. I. and Skorohod, A. V. (1979). *The theory of stochastic processes*, volume 1. New York: Springer.

Gringarten, A. C. (2008). From straight lines to deconvolution: the evolution of the stete of the art in well test analysis. *SPE102079*, pages 1–10.

Hansen, P. (2002). Deconvolution and regularization with toeplitz matrices. *Numerical Algorithms*, 29:323–378.

Hoeting, J., Madigan, D., Raftery, A., and Volinsky, C. (1999). Bayesian model averaging: A tutorial. *STATISTICAL SCIENCE*, 14(4):382–401.

Holsclaw, T., Sansó, B., Lee, H., Higdon, D., Heitmann, K., Alam, U., and Habib, S. (2011). Gaussian process modeling of derivative curves. Technical Report UCSC-SOE-11-02, Applied Mathematics and Statistics, University ofCalifornia Santa Cruz.

Ilk, D., Valkó, P. P., and Blasingame, T. A. (2007). A Deconvolution Method Based on Cumulative Production for Continuously Measured Flowrate and Pressure Data. *SPE111269*, pages 1–10.

Kuchuk, F., Onur, M., and Hollander, F. (2010). *Pressure transient formation and well testing: convolution, deconvolution and nonlinear estimation*, volume 57 of *Developments in Petroleum Science*. Elsevier.

Levitan, M. M. (2005). Practical Application of Pressure / Rate Deconvolution to Analysis of Real Well Tests. *SPE84390*, (April):1—-9.

Levitan, M. M., Crawford, G. E., and Hardwick, A. (2006). Practical Considerations for Pressure-Rate Deconvolution of Well-Test Data. *SPE90680*, (March):35–47.

Onur, M., Çnar, M., Ilk, D., Valko, P. P., Blasingame, T. A., and Hegeman, P. (2008). An Investigation of Recent Deconvolution Methods for Well-Test Data Analysis. *SPE 102575*, (June):226–247.

Pimonov, E., Ayan, C., Onur, M., and Kuchuk, F. (2010). A New Pressure / Rate-Deconvolution Algorithm To Analyze Wireline- Formation-Tester and Well-Test Data. *SPE123982*, (August):603–613.

Pimonov, E., Onur, M., and Kuchuk, F. (2009). A new robust algorithm for solution of pressure/rate deconvolution problem. *Journal of Inverse Ill-Posed Problems*, 17:611–627.

Ramm, A. and Smirnova, A. (2005). On deconvolution problems: numerical aspects. *Journal of Computational and Applied Mathematics*, 176:445–460.

Rasmussen, C. (2003). Gaussian processes to speed up hybrid monte carlo for expensive bayesian integrals. In Bernardo, J. M., m. J. Bayarri, Berger, J. O., Dawid, P., and West, D. H. A. F. M. S. M., editors, *Bayesian Statistics 7*. Oxford University Press.

Schlumberger (2002). *Well Test Interpretation*. Schlumberger testing Services. Schlumberger.

von Schroeter, T. V., Hollaender, F., Gringarten, A. C., and Álzaga Ruiz, H. (2001). Deconvolution of Well Test Data as a Nonlinear Total Least Squares Problem. *SPE71574*, pages 1—-12.

Wichura, M. J. (1988). Algorithm as 241: The percentage points of the normal distribution. *Applied Statistics*, 37:477–484.