



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS
MATEMÁTICAS Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA
APLICADA

GENERACIÓN DE UNA
METODOLOGÍA PARA LA
CARACTERIZACIÓN DE YACIMIENTOS
PETROLEROS MEDIANTE MÉTODOS
ESTADÍSTICOS Y NUMÉRICOS

T E S I S

QUE PARA OPTAR POR EL GRADO DE:
M A E S T R O E N C I E N C I A S

P R E S E N T A:
M I N U T T I M A R T Í N E Z C A R L O S

DIRECTORA DE TESIS:
DRA. SUSANA GÓMEZ GÓMEZ

INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN SISTEMAS

MÉXICO, D.F.

Abril, 2015



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Generación de una metodología para la caracterización de
yacimientos petroleros mediante métodos estadísticos y
numéricos

Minutti Martínez Carlos

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México

cminutti@gmail.com

Abril, 2015

Índice general

Introducción	v
1. Descripción del problema	1
1.1. Modelo PT-TPDP	1
1.2. Problema inverso	4
I Problema directo	9
2. Solución al modelo PT-TPDP	10
2.1. Daño y radio efectivo	12
3. Análisis de sensibilidad	14
3.1. Sensibilidad de los parámetros	14
3.2. Sensibilidad: parámetro-tiempo	18
3.3. Efecto general de los parámetros en la presión	24
II Problema inverso	30
4. Problema de optimización	31
4.1. Función objetivo	31
4.1.1. Espaciado logarítmico	32
4.1.2. Cambio de función objetivo	41
4.1.3. Presión de Laplace	43
4.2. Optimización secuencial	45
4.3. Transformación de parámetros	50
4.4. Análisis multivariado: Dependencia de los parámetros	56
4.5. Comparativo de resultados	59
5. Múltiples soluciones óptimas	67
5.1. Solución dual	67
5.2. Pseudo-mínimos	71

<i>ÍNDICE GENERAL</i>	I
6. Casos con datos reales	78
6.1. Pozo 1	78
6.2. Pozo 2	80
Conclusiones	81
Bibliografía	85

Índice de figuras

1.1. Curvas tipo: conjunto de datos y ajuste	6
1.2. Curvas tipo: comparación con la modificación de un parámetro (ω_v)	7
1.3. Curvas tipo: comparación con la modificación de un parámetro (κ)	8
3.1. Dispersión de τ_D y t_D vs presión	17
3.2. Desviación estándar de la presión, a lo largo del tiempo en Laplace	18
3.3. Dispersión de ω_f y ω_v vs presión	19
3.4. Dispersión de λ_{mf} , λ_{mv} y λ_{vf} vs presión	20
3.5. Dispersión de κ vs presión	21
3.6. Dispersión de s y C_D vs presión	21
3.7. Correlación de ω_f y ω_v con la presión, en diferentes tiempos	22
3.8. Correlación de λ_{mf} , λ_{mv} y λ_{vf} con la presión, en diferentes tiempos	22
3.9. Correlación de κ con la presión, en diferentes tiempos	23
3.10. Correlación de los parámetros de yacimiento con la presión, en diferentes tiempos	23
3.11. Correlación de los parámetros de pozo con la presión, en diferentes tiempos	24
3.12. Efecto promedio de ω_f y ω_v en la presión	25
3.13. Efecto promedio conjunto de ω_f y ω_v en la presión	26
3.14. Efecto promedio de λ_{mf} , λ_{mv} y λ_{vf} en la presión	27
3.15. Efecto promedio de κ en la presión	28
3.16. Efecto promedio de s y C_D en la presión	28
4.1. Boxplot - Error de estimación de ω 's y κ , sin y con espaciado logaritmico ($s = C_D = 0$)	35
4.2. Boxplot - Error de estimación de λ 's, sin y con espaciado logaritmico ($s = C_D = 0$)	36
4.3. Dispersión del valor ω_f a estimar vs estimación (espaciado logaritmico)	37
4.4. Dispersión del valor λ_{mf} a estimar vs estimación (espaciado logaritmico)	37
4.5. Dispersión del valor κ a estimar vs estimación (espaciado logaritmico)	38
4.6. Dispersión del valor λ_{mf} a estimar vs estimación (espaciado lineal)	38
4.7. Boxplot - Error de estimación de ω 's y κ , sin y con espaciado logaritmico	40
4.8. Boxplot - Error de estimación de λ 's, sin y con espaciado logaritmico	40
4.9. Dispersión de los valores s y C_D a estimar vs estimaciones (espaciado logaritmico)	41
4.10. Boxplot - Error de estimación de ω 's y κ , método no secuencial vs secuencial	51
4.11. Boxplot - Error de estimación de λ 's, método no secuencial vs secuencial	51
4.12. Forma general del cambio de la presión para ω 's e histograma para el muestreo correspondiente a su transformación	53

4.13. Forma general del cambio de la presión para κ e histograma para el muestreo correspondiente a su transformación	54
4.14. Forma general del cambio de la presión para λ 's e histograma para el muestreo correspondiente a su transformación	55
4.15. Correlación parcial (de Spearman) entre los errores de estimación y valores de los parámetros	58
4.16. Boxplot - Error de estimación de los parámetros de yacimiento, sin y con información de sensibilidad	61
4.17. Comparativo de la dispersión del valor a estimar de ω_f vs estimación, sin y con información de sensibilidad en la optimización	63
4.18. Comparativo de la dispersión del valor a estimar de λ_{mf} vs estimación, sin y con información de sensibilidad en la optimización	64
4.19. Comparativo de la dispersión del valor a estimar de κ vs estimación, sin y con información de sensibilidad en la optimización	65
5.1. Función objetivo con diferentes valores de κ	72
5.2. Función objetivo con diferentes valores de κ , para dos intervalos de tiempo	74
5.3. Función objetivo con diferentes valores de κ , para dos intervalos de tiempo	75
5.4. Función objetivo con diferentes valores de κ y ω_f	76
5.5. Función objetivo con diferentes valores de κ y ω_f aplicando restricción	77
6.1. Ajuste de doble y de triple porosidad para el pozo 1	79
6.2. Ajuste de doble y de triple porosidad para el pozo 2	80

Índice de cuadros

3.1. Correlación de Spearman entre los parámetros y la presión	16
4.1. Media y mediana del error de estimacion para espaciado lineal y logaritmico ($C_D = s = 0$)	33
4.2. Pruebas de hipótesis de igualdad de medias para espaciado lineal y logaritmico ($C_D = s = 0$)	35
4.3. Media y mediana del error de estimacion para espaciado lineal y logaritmico	39
4.4. Pruebas de hipótesis de igualdad de medias para espaciado lineal y logaritmico	39
4.5. Media y mediana del error de estimacion para las distintas funciones objetivo	43
4.6. Pruebas de hipótesis de igualdad de medias para el uso de las distintas funciones objetivo	44
4.7. Media y mediana del error de estimacion para $\bar{p}_{wD}(\tau_D)$ vs $\tau_D \bar{p}_{wD}(\tau_D)$	45
4.8. Pruebas de hipótesis de igualdad de medias para $\bar{p}_{wD}(\tau_D)$ vs $\tau_D \bar{p}_{wD}(\tau_D)$	46
4.9. Media y mediana del error de estimacion para optimización con Algoritmo Genetico vs Optimización por Intervalos	48
4.10. Pruebas de hipótesis de igualdad de medias para optimización con Algoritmo Genetico vs Optimización por Intervalos	48
4.11. Media y mediana del error de estimación para optimización con método no secuencial vs secuencial	49
4.12. Pruebas de hipótesis de igualdad de medias para optimización con método no secuencial vs secuencial	50
4.13. Media y mediana del error de estimación para optimización con un muestreo uniforme sobre los parámetros vs muestreo con $\Delta p^*(\cdot)$	56
4.14. Pruebas de hipótesis de igualdad de medias para optimización con muestreo uniforme sobre los parámetros vs muestreo con $\Delta p^*(\cdot)$	57
4.15. Media y mediana del error de estimación para optimización sin y con información de sensibilidad	60
4.16. Pruebas de hipótesis de igualdad de medias para optimización sin y con información de sensibilidad	61
4.17. Intervalos de confianza para la media del error de estimación, sin y con información de sensibilidad	62
6.1. Parámetros estimados para el Pozo 1, con los modelos de doble y de triple porosidad	79
6.2. Parámetros estimados para el Pozo 2, con los modelos de doble y de triple porosidad	81

Introducción

La caracterización de los yacimientos de petróleo a través de pruebas de presión en un pozo es un problema importante dentro de la industria de los hidrocarburos. En estos ensayos, las tasas de flujo y las presiones se miden (comúnmente en el fondo del pozo), y los resultados se utilizan para caracterizar los depósitos con la ayuda de modelos matemáticos *ad-hoc*.

El comportamiento de presión en un yacimiento pueden ser modelado mediante la combinación de la ecuación de difusión con el principio de la conservación de la masa de fluido, la ley de Darcy, que describe el flujo de un fluido a través de un medio poroso, y una ecuación de estado (véase, por ejemplo, Chen *et al.* [7], capítulo 2). La mayoría de estos modelos consisten de una ecuación diferencial parcial o de un sistema de ecuaciones diferenciales parciales.

Estos modelos de interpretación son la base para la determinación de las propiedades de la formación mediante la resolución de problemas inversos. De esta forma, si estas propiedades son conocidas con suficiente exactitud, la productividad de un yacimiento puede ser estimada.

Un yacimientos naturalmente fracturados (YNF) se caracteriza frecuentemente por la presencia de dos tipos diferentes de medios porosos, una red de fracturas interconectadas rodeadas por una matriz comparativamente de baja permeabilidad, que consiste en roca. Las fracturas tienen una capacidad de flujo alta, pero una baja capacidad de almacenamiento, y la matriz tiene un alto almacenamiento, pero una capacidad de flujo baja; esto implica la existencia de muy diferentes porosidades y permeabilidades en cada uno de los componentes (fracturas y matriz) de la formación.

En las últimas décadas los YNF han recibido mucha atención: la recuperación y manejo del petróleo de YNF es un problema importante. Desde 1960, cuando un artículo seminal de Barenblatt, Zheltov y Kochina [3] apareció, ha habido una cantidad considerable de actividad científica para diagnosticar y caracterizar los YNF.

La mayoría de los modelos para la caracterización de YNF se basan en la teoría de Warren y Root [20], quienes introdujeron el concepto de doble porosidad. A partir de este trabajo, y teniendo en cuenta los distintos tipos de matriz-fractura, y los regímenes de flujo o geometrías de yacimiento, se han desarrollado diversos modelos. Sin embargo, varios autores han hecho hincapié en la limitación del enfoque de doble porosidad para el tratamiento de yacimientos más complejos (véase, por ejemplo Abdassah y Ershaghi [1], Dehghani *et al.* [9], Gurbinar *et al.* [11], y Wu *et al.* [22]).

Es así que en 1986, Abdassah y Ershaghi [1] introdujeron el concepto de triple porosidad, considerando un modelo en el que se produce el flujo entre la porosidad en estado no estacionario entre el sistema de fracturas y dos tipos de bloques de matriz. Un breve resumen de las teorías existentes para el tratamiento de los YNF, discutiendo sus respectivas ventajas e inconvenientes, se puede encontrar en un artículo de Wu *et al.* [22].

En cuanto a considerar medios de porosidad adicionales, Camacho *et al.* [6] proponen un mo-

delo de triple porosidad para el análisis de presiones transitorias en yacimientos naturalmente fracturados vugulares (YNFV). Los vúgulos (cavidades) pueden tener un efecto sustancial sobre la permeabilidad si el depósito contiene un sistema interconectado de los mismos. Al igual que Kazemi [15], Camacho *et al.* [6] consideran un YNFV con un pozo circular (en el centro de un yacimiento), en donde tres componentes independientes pero que interactúan (matriz, fracturas y vúgulos) coexisten. Ellos mostraron que los vúgulos pueden afectar el comportamiento de las pruebas de pozo. Por lo tanto, los vúgulos se deben tener en cuenta con el fin de caracterizar completamente un YNFV.

El estudio de los procesos de flujo en yacimientos naturalmente fracturados vugulares ha recibido recientemente mucha atención debido a que un número de estos depósitos se han encontrado en todo el mundo y especialmente en México, con producciones significativas de petróleo y gas.

Camacho *et al.* [6] son los únicos que consideran flujo primario en ambas redes de fracturas y vúgulos, lo que genera un modelo de triple porosidad - doble permeabilidad (TPDP). Existe evidencia de que en algunas rocas vugulares las cavidades se tocan formando una red de cavidades conectadas directamente, creando vías de flujo libres de la resistencia de la matriz (ver por ejemplo Zhang *et al.* [23] y [24], Nair [16] y Vik *et al.* [19]). El modelo TPDP de Camacho *et al.* [6] depende de dos relaciones de almacenamiento, uno para las fracturas y otro para las cavidades, y tres parámetros de flujo que tienen en cuenta la transferencia de fluidos entre la matriz, fracturas y vúgulos. También depende de la razón de permeabilidad entre las fracturas y vúgulos.

El modelo de TPDP consiste en un sistema de tres ecuaciones diferenciales parciales que se pueden resolver analíticamente para obtener las presiones en los tres medios considerados (fracturas, matriz, vúgulos) y también la presión del pozo. Las soluciones analíticas se obtienen en el espacio de Laplace con la ayuda de la transformada de Laplace. El algoritmo de Stehfest [18] se aplica a continuación para regresar las soluciones a tiempo real. El problema de regresión no lineal para el modelo de Camacho *et al.* [6] equivale a la búsqueda de los parámetros del modelo que realizan el mejor ajuste de la curva de presión medida en el pozo que se estudia. Para comprobar la bondad del ajuste calculamos también la curva de la log-derivada de la presión (i.e. la derivada de la presión con respecto al logaritmo del tiempo), la cual ha mostrado ser altamente conveniente para exponer el efecto de los parámetros de yacimiento a través de la presión.

Debido a que cualquier método de minimización en la solución del problema inverso de identificación de parámetros requerirá la solución del modelo muchas veces, el uso del algoritmo de Stehfest para regresar numéricamente cada solución en el espacio de Laplace a tiempo real puede hacer que la optimización sea inaceptablemente lenta, y debido al mal-condicionamiento de este procedimiento inverso, se pueden producir curvas de presión que muestran un comportamiento oscilatorio. Para evitar este inconveniente, algunos autores proponen llevar a cabo el proceso de identificación de parámetros en el espacio de Laplace (ver [21], [5] y [17]). Es así que si se realiza la identificación de parámetros en el espacio de Laplace, es conveniente utilizar el método propuesto por Bourgeois y Horne [5] para la transformación de los datos de presión al espacio de Laplace.

El problema inverso que corresponde al modelo de TPDP que aquí se presenta es altamente complejo, lo cual es debido a la precisión requerida en el ajuste tanto a la curva de presión como a la curva de la log-derivada para determinar correctamente los múltiples parámetros que están involucrados en un yacimiento, ya que los parámetros tienen diferencias significativas entre ellos en cuanto al efecto en magnitud de la presión, así como en los tiempos en que cada uno se manifiesta con mayor influencia, lo cual es debido al propio modelo y los diferentes aspectos numéricos involucrados.

Estudiando a detalle estas interacciones entre tiempo-presión-parámetro es posible realizar una

caracterización más precisa, veloz y automatizada del yacimiento. Es así que en este trabajo se pretende determinar las diferentes características en el proceso de optimización que permitan identificar adecuadamente y en un tiempo razonable los parámetros que caracterizan a un yacimiento, esto mediante estudios en la función objetivo, análisis de sensibilidad, escalamiento de parámetros, y eliminación de soluciones múltiples producto de infinitos mínimos inherentes al modelo y a la precisión numérica, todo esto desarrollado tanto en el espacio de Laplace como en tiempo real.

En Gómez *et al.* [13] y [12] se aborda el problema de identificación de parámetros utilizando el modelo de Camacho *et al.* [6], haciendo énfasis en la existencia de múltiples soluciones (es decir, diferentes conjuntos de parámetros que reproducen los datos medidos con suficiente precisión) y la necesidad de un adecuado algoritmo de optimización global para poder identificar las distintas soluciones.

Mediante la metodología desarrollada en este trabajo es posible solventar el problema de múltiples soluciones al conocer explícitamente la forma de estas, y al identificar con suficiente precisión los distintos parámetros.

Capítulo 1

Descripción del problema

El petróleo se encuentra almacenado dentro de un medio poroso (roca), que presenta tres medios de flujo: fracturas, vóculos (cavidades) y matriz (roca porosa propia de la roca almacén). El flujo a través de estos medios (y por lo tanto la presión en el pozo) depende de varios parámetros, por ejemplo las proporciones presentes de aceite en cada una de los medios, o la facilidad que tiene el petróleo de moverse en cada una de ellos (permeabilidad).

Cuando se perfora un pozo en un yacimiento petrolero, es posible realizar una prueba de presión del pozo, la cual consiste en cerrar la producción del mismo y registrar el aumento de presión.

Al considerar aspectos como la geometría del pozo, la estructura del medio poroso y las características de flujo del petróleo, se puede generar un modelo matemático que a partir de los parámetros que describen el yacimiento explica como se ve afectada la presión en el pozo a lo largo del tiempo.

Es por medio de estos parámetros que describen al yacimiento que es posible estimar la presión a tiempos mayores, y por medio de esta presión, determinar la productividad del pozo. Aunque estos parámetros no se conocen a priori, es posible estimarlos al comparar los distintos valores de presión que genera el modelo matemático contra los datos medidos durante la prueba de presión en el pozo; por lo tanto es necesario un modelo que sea capaz de describir adecuadamente los cambios de presión en el pozo y un método para estimar los parámetros del modelo que corresponden a los datos observados.

1.1. Modelo de Penetración Total: Triple Porosidad - Doble Permeabilidad

Como se ha mencionado anteriormente, varios autores comentan la incapacidad de los modelos de doble porosidad para describir adecuadamente yacimientos de mayor complejidad (como lo son los YNF). Es así que en Camacho *et al.* [6] se introduce un modelo sin dimensiones para la simulación de alta porosidad secundaria, principalmente porosidad vugular y de fracturas, en yacimientos naturalmente fracturados. Se propone un modelo de triple porosidad para presión no estacional donde se considera la transferencia de fluido entre la matriz, los vóculos y las fracturas, siendo la presión directamente proporcional a la diferencia en el volumen medio de la presión macroscópica. El modelo considera la interacción entre estos 3 medios, e incluye la posibilidad de tener flujo

primario a través del sistema de vóculos y/o de fracturas.

Durante el desarrollo del modelo se supone lo siguiente:

- Las propiedades de la roca son constantes en cada medio.
- El yacimiento es de espesor uniforme con fronteras inferior y superior impermeables.
- El fluido es ligeramente compresible con viscosidad constante y el flujo de fluido es de una sola fase e isotérmico.
- El flujo del fluido en el pozo es radial, desde las fracturas y vóculos.

Las variables adimensionales del modelo están dadas por:

$$p_{Dj} = \frac{2\pi hk_T}{q\mu B}(p_0 - p_j), \quad (1.1a)$$

$$t_D = \frac{k_T}{(\Phi_f c_f + \Phi_m c_m + \Phi_v c_v)\mu r_w^2} t, \quad (1.1b)$$

$$r_D = \frac{1}{r_w} r \quad (1.1c)$$

donde p_j y p_{Dj} es la presión en el medio j , con y sin dimensiones respectivamente, donde $j = f$ para fracturas, $j = v$ para vóculos o $j = m$ para matriz, con p_0 como el valor inicial de presión, considerado igual en todos los medios. De forma similar t_D y r_D representan el tiempo y radio del pozo, adimensionales.

Denotamos:

$$k_T = k_f + k_v$$

como la permeabilidad total, donde k_f y k_v son las permeabilidades de fracturas y vóculos, respectivamente. La definición de los parámetros del modelo está dada por:

$$\kappa = \frac{k_f}{k_T}, \quad (1.2a)$$

$$\lambda_{mf} = \sigma_{mf} r_w^2 \frac{k_m}{k_T}, \quad (1.2b)$$

$$\lambda_{mv} = \sigma_{mv} r_w^2 \frac{k_m}{k_T}, \quad (1.2c)$$

$$\lambda_{vf} = \sigma_{vf} r_w^2 \frac{k_{vf}}{k_T}, \quad (1.2d)$$

$$\omega_f = \frac{\Phi_f c_f}{\Phi_f c_f + \Phi_m c_m + \Phi_v c_v}, \quad (1.2e)$$

$$\omega_v = \frac{\Phi_v c_v}{\Phi_f c_f + \Phi_m c_m + \Phi_v c_v} \quad (1.2f)$$

donde k_{vf} se define como:

$$k_{vf} = \begin{cases} k_v & \text{si } p_v > p_f \\ k_f & \text{de otra forma} \end{cases}$$

El parámetro λ_{ij} es el factor de interacción entre el medio i y j ; κ es el cociente de la permeabilidades y w_j se le conoce como razón o capacidad de almacenamiento (ambos, κ y w_j expresados como proporción). Nótese que en las definiciones de λ_{mf} y λ_{mv} , hemos usado k_m debido a que en ausencia de fuerzas capilares, se espera que bajo las condiciones de producción primaria, el fluido va desde la matriz hacia fracturas y vóculos.

Además, al considerar a ω_m como la razón de almacenamiento en matriz, y debido a que $w_f + w_v + w_m = 1$, si solo trabajamos con ω_f y ω_v , entonces es evidente que debe cumplirse la restricción:

$$\omega_f + \omega_v \leq 1$$

Los parámetros Φ , c y σ representan la porosidad, compresibilidad y el factor de flujo interporoso, y aunque no son conocidos, no haremos mayor énfasis en ellos, ya que no son parámetros con los que se trabaje directamente. Los parámetros h , q , μ y B son el espesor de la formación, tasa de flujo del petróleo, viscosidad del petróleo y un factor de volumen, los cuales son conocidos pero de igual forma no son relevantes para el trabajo aquí presentado.

Usando las variables adimensionales, la ecuación diferencial para la presión $p_{Df}(r_D, t_D)$ en las fracturas es la siguiente:

$$\kappa \frac{1}{r_D} \frac{\partial}{\partial r_D} \left(r_D \frac{\partial p_{Df}}{\partial r_D} \right) + \lambda_{mf}(p_{Dm} - p_{Df}) + \lambda_{vf}(p_{Dv} - p_{Df}) = w_f \frac{\partial p_{Df}}{\partial t_D} \quad (1.3)$$

Para la presión $p_{Dm}(r_D, t_D)$ en el bloque de la matriz, asumiendo que la permeabilidad de la matriz es despreciable en comparación con la de facturas y vóculos, la ecuación es:

$$-\lambda_{mv}(p_{Dm} - p_{Dv}) - \lambda_{mf}(p_{Dm} - p_{Df}) = (1 - \omega_f - \omega_v) \frac{\partial p_{Dm}}{\partial t_D} \quad (1.4)$$

y la presión $p_{Dv}(r_D, t_D)$ para vóculos es determinada por la ecuación

$$(1 - \kappa) \frac{1}{r_D} \frac{\partial}{\partial r_D} \left(r_D \frac{\partial p_{Dv}}{\partial r_D} \right) + \lambda_{mv}(p_{Dm} - p_{Dv}) - \lambda_{vf}(p_{Dv} - p_{Df}) = \omega_v \frac{\partial p_{Dv}}{\partial t_D} \quad (1.5)$$

Para obtener la solución del modelo de flujo del yacimiento, el sistema de EDP (1.3)-(1.5) requiere de las condiciones iniciales y de frontera apropiadas. Para establecer estas condiciones deben introducirse dos fenómenos: daño (también conocido como skin) y almacenamiento.

En la literatura de ingeniería petrolera el efecto del daño (denotado por s), es usado para caracterizar una condición cercana a la región del pozo. El daño no posee dimensiones físicas y puede ser negativo, positivo o cero. Siempre que exista un aumento o disminución en la caída de presión entre la caída predicha con respecto a la ley de Darcy y la caída medida por una prueba buildup¹ o drawdown², se asume que la diferencia es causada por el daño.

Además, cuando la producción de un pozo se cierra, el flujo en el pozo continúa después de cerrar. Este régimen de flujo se refiere como flujo posterior o almacenamiento del pozo, y puede afectar el análisis de los datos de presión. El efecto de almacenamiento del pozo se controla típicamente por la compresibilidad del fluido en el pozo y se denota por la constante $C \geq 0$.

¹Una prueba de pozo buildup consiste en el histórico del aumento de presión a lo largo del tiempo una vez que se ha cerrado la producción.

²Una prueba de pozo drawdown consiste en el histórico de la caída de presión a lo largo del tiempo mientras el pozo se encuentra en producción.

Teniendo en cuenta la presencia del daño mecánico (skin) tanto para las fracturas y cavidades (vúgulos), y el efecto adimensional del almacenamiento del pozo, es necesario introducir nuevos parámetros al modelo, s_f , s_v y C_D . Los parámetros s_f , s_v están relacionados con el daño mecánico de fracturas y vúgulos respectivamente, mientras que C_D es el coeficiente adimensional de almacenamiento del pozo, definido por

$$C_D = \frac{1}{2\pi h(\Phi_f c_f + \Phi_m c_m + \Phi_v c_v) r_w^2} C \quad (1.6)$$

Tomando en cuenta lo anterior, deben satisfacerse las siguientes condiciones iniciales y de frontera

$$p_{Dj}(r_D, t_D)|_{r_D \rightarrow \infty} = 0, \quad j = f, v, m \quad (1.7a)$$

$$\left[\kappa r_D \frac{\partial p_{Df}}{\partial r_D} + (1 - \kappa) r_D \frac{\partial p_{Dv}}{\partial r_D} \right]_{r_D=1} = -1, \quad (1.7b)$$

$$\left[p_{Df} - s_f r_D \frac{\partial p_{Df}}{\partial r_D} \right]_{r_D=1} = \left[p_{Dv} - s_v r_D \frac{\partial p_{Dv}}{\partial r_D} \right]_{r_D=1} = p_{wD}(t_D), \quad (1.7c)$$

$$p_{Dj}(r_D, t_D)|_{t_D=0} = 0, \quad j = f, v, m \quad (1.7d)$$

donde $p_{wD}(t_D)$ es la presión adimensional de flujo del pozo.

1.2. El problema de identificación de parámetros del modelo

En estadística la regresión no lineal es un análisis en el que se pretende ajustar una función no lineal en los parámetros a un conjunto de datos, donde el ajuste generalmente se realiza mediante un método de optimización en el que la función objetivo es una medida del error cometido entre el ajuste y los datos. Ya que una prueba de pozo genera valores de presión, los cuales se considera corresponden a un modelo e incluyen un error en la medición, el problema de identificación de parámetros (problema inverso) de este modelo es un problema de regresión no lineal.

El objetivo del análisis de una prueba de pozo es determinar estimados razonables de los parámetros del yacimiento dado un modelo que describe el comportamiento de la presión durante la prueba de pozo en base a estos parámetros, en nuestro caso el modelo de TPDP de Camacho *et al.* [6] que se describió en 1.1, cuyo vector de parámetros es:

$$\mathbf{x} = (\mathbf{x}_r, \mathbf{x}_w) = (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa, s, C_D) \in \mathbb{R}^8$$

donde $\mathbf{x}_r = (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa, s)$ son los parámetros del yacimiento y $\mathbf{x}_w = (s, C_D)$ los parámetros de pozo³. Los parámetros de yacimiento y de pozo son los que se deben de identificar en el problema inverso.

Para identificar los parámetros en el problema inverso es necesario contar con los datos de presión sin dimensiones, ya que la presión generada por el modelo es adimensional. Por lo tanto se considera el problema de pasar de datos con dimensiones a sin dimensiones, donde es necesario

³Diferenciamos los parámetros de yacimiento (\mathbf{x}_r) que están relacionados con la estructura física del yacimiento y los parámetros de pozo (\mathbf{x}_w) que están asociados únicamente a efectos generados por el pozo, por lo que no representan ningún efecto propio del yacimiento.

determinar los parámetros α_t y α_p , que corresponden a la transformación del tiempo y la presión de dimensiones a sin dimensiones, i.e. $p_{wD}(t_D) = \alpha_p \cdot p_w(\alpha_t \cdot t)$.

Dado que existe información que no se conoce para pasar los valores de presión y tiempo de dimensiones a sin dimensiones, es necesario estimar los parámetros α_t y α_p en el caso de problemas reales. Sin embargo, estos parámetros de escala no afectan a la determinación de los parámetros del yacimiento, y no son necesarios para trabajar con datos sintéticos ya que el modelo genera valores sin dimensiones, por lo que en este trabajo solo se utilizarán los parámetros en \mathbf{x} .

De esta forma el problema de regresión no lineal aquí abordado, consiste en estimar los parámetros de yacimiento a través del modelo de Camacho *et al.* [6] (lo que corresponde al problema inverso del modelo).

Típicamente el problema inverso involucra optimización, siendo también este nuestro caso; para ello es necesario contrastar los datos de presión generados por el modelo ($\text{MOD}(\mathbf{x})$) y los obtenidos por la prueba de pozo (data) a través de una función objetivo $F(\mathbf{x})$, la cual mide una distancia entre los datos del modelo y los observados, limitando los valores que puede tomar $\mathbf{x} \in [L, U]$.

$$\begin{aligned} \text{mín } F(\mathbf{x}) &= \|\text{MOD}(\mathbf{x}) - \text{data}\| \\ \text{s.a. } &L \leq \mathbf{x} \leq U \end{aligned}$$

Generalmente se utiliza como función objetivo la función de mínimos cuadrados, pero es importante elegir adecuadamente esta función, situación que se estudiará en el Capítulo 4.

Debido a que el modelo correspondiente a las EDP (1.3)-(1.5) devuelve una solución en el espacio de Laplace, existe la posibilidad de realizar el proceso de optimización en este espacio, al llevar los datos de presión (medidos en tiempo real) a Laplace con el algoritmo de Bourgeois y Horne [5], o llevar la solución del modelo de Laplace a tiempo real a través del algoritmo de Stehfest [18].

El primer método involucra un error en los datos de presión, debido a la interpolación y extrapolación que debe asumir el algoritmo [5], y el segundo método, aunque involucra un error menor, aumenta significativamente el tiempo de cómputo, debido a que la solución en tiempo real requiere resolver el modelo al menos 8 veces en Laplace, para cada punto en tiempo real.

Al trabajar con datos sintéticos conviene utilizar la optimización en Laplace, ya que los datos de presión generados también están en Laplace, por lo que no hay error por el paso de dimensiones. Sin embargo la metodología, análisis y pruebas aquí desarrollados no están influenciados por el espacio de trabajo que se elige.

Sin importar si el ajuste se realiza en Laplace o en tiempo real, las curvas que utilizaremos para comparar el ajuste final que se obtuvo por medio de la optimización, son conocidas como curvas tipo.

Técnica de ajuste por curvas tipo

El ajuste por curvas tipo son un método de interpretación de pruebas de pozo de decremento de presión (drawndown) e incrementode presión (buildup), que permiten observar gráficamente algunas de las características relacionadas a los parámetros de yacimiento, por lo que el ajuste a estas curvas también es utilizado para determinar si existe una adecuada identificación de los parámetros en el problema inverso.

Las curvas tipo ajustan básicamente los datos de presión medidos durante la prueba de pozo a lo largo del tiempo (o más correctamente de la diferencia de presión con respecto a la presión

inicial, i.e. $p(t) = p_j - p_0$, con p_j la presión en tiempo j) y de la log-derivada de la misma, es decir:

$$\frac{\partial}{\partial \log(t)} p(t) \quad (1.8)$$

ambas expresadas en variables con o sin dimensiones, y coordenadas logarítmicas.

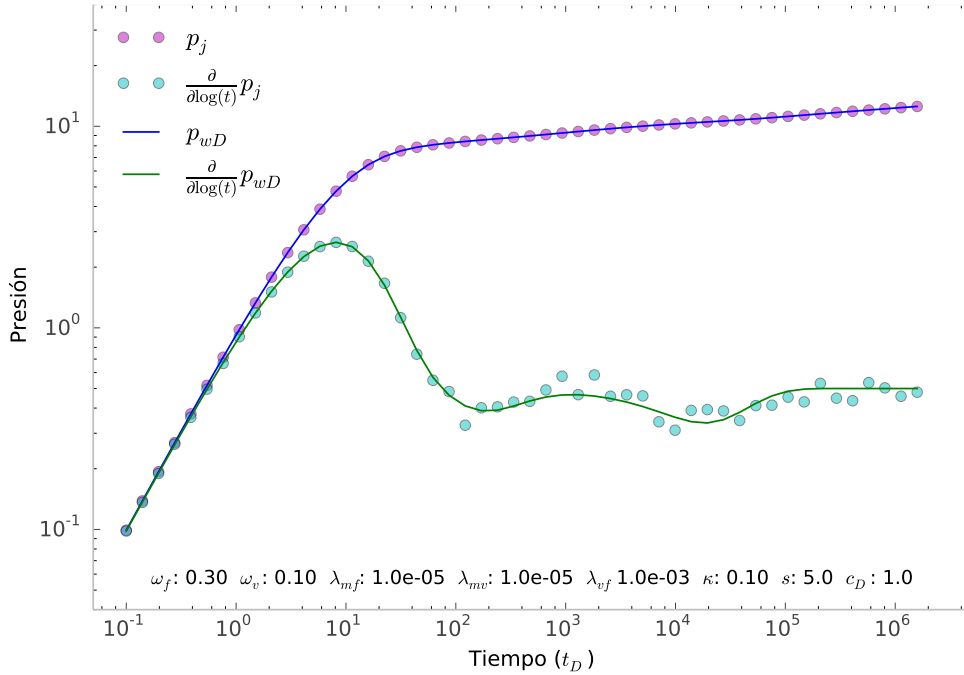


Figura 1.1: Curvas tipo: conjunto de datos y ajuste

En la Figura 1.1 se muestra un ejemplo de ajuste de datos por curva tipo para datos sintéticos, los cuales son generados a partir del modelo, a los que se les agregó ruido Gaussiano.

Adicionalmente se presentan las curvas de ajuste, que corresponden a los mismos datos, pero sin ruido, por lo que este podría considerarse un caso ideal en el problema de regresión no lineal. Por supuesto la derivada se ve considerablemente más afectada por el ruido que la presión, sin embargo es de gran importancia comparar contra la derivada, ya que es ahí donde es más fácil contrastar ambas curvas, así como también las características del yacimiento.

Puede apreciarse en las curvas que existe una zona en la derivada donde el ajuste no parece ser tan bueno, sin embargo, ya que se tratan de los mismos datos con y sin ruido, esto es puramente resultado de la aleatoriedad en el ruido, que debe considerarse como una posibilidad al momento de contrastar ajustes.

En la Figura 1.2 se presenta un ejemplo de curva tipo en el cual se contrastan dos curvas para un juego de parámetros, en el cual solo se ha modificado ligeramente a ω_v ; sin embargo se observa

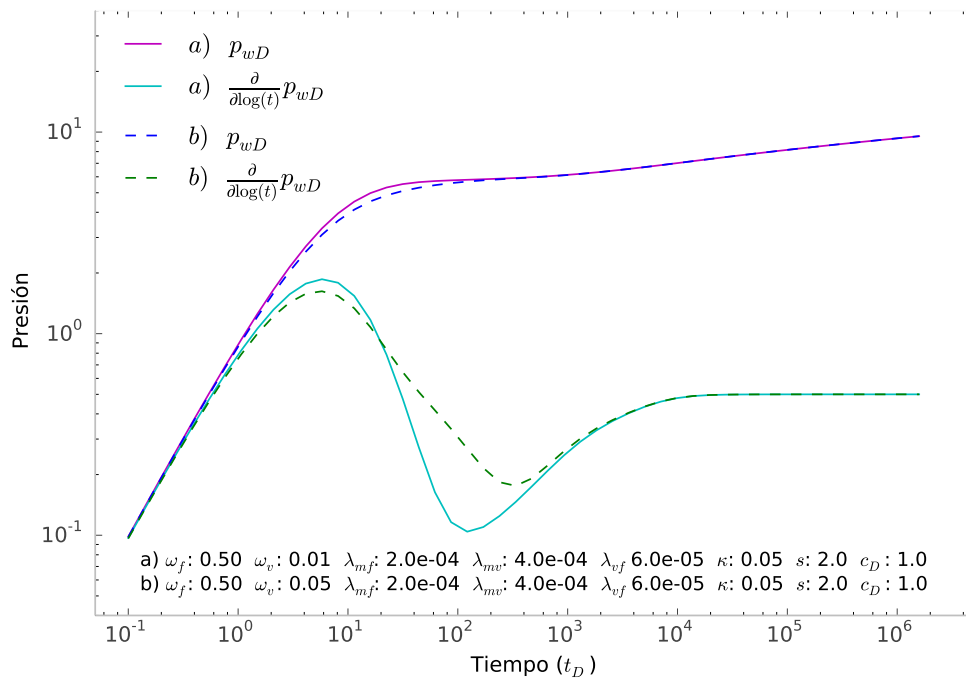


Figura 1.2: Curvas tipo: comparación con la modificación de un parámetro (ω_v)

que este cambio afecta de forma importante a la curva, y si se considera que $0 \leq \omega_j \leq 1$, la estimación del valor $\omega_v = 0.05$ en lugar de $\omega_v = 0.01$ no corresponde a un error de importancia.

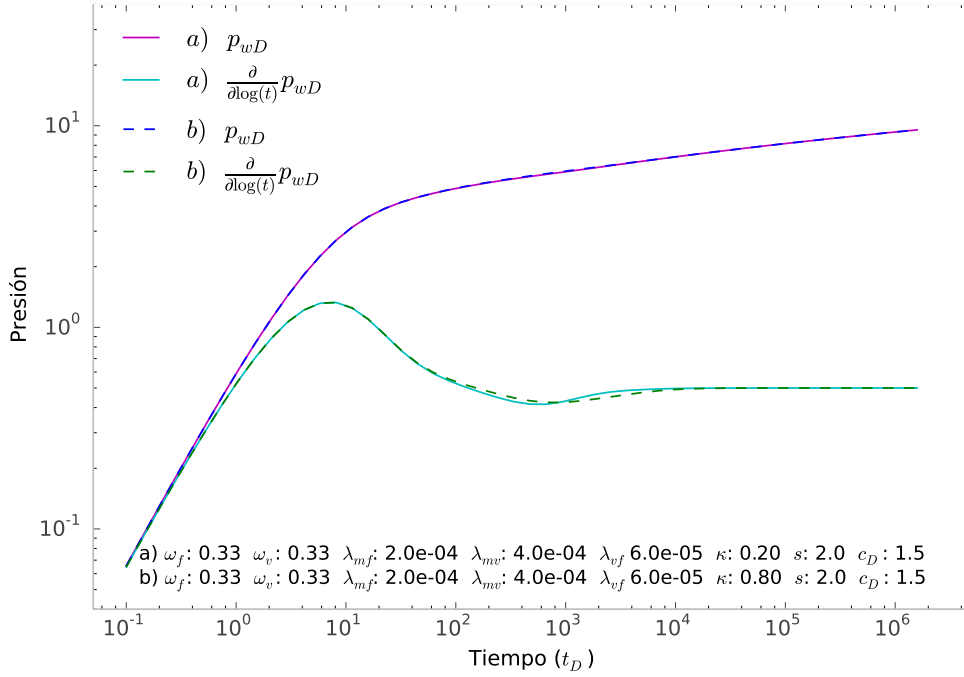


Figura 1.3: Curvas tipo: comparación con la modificación de un parámetro (κ)

En una situación opuesta observamos la Figura 1.3, donde el parámetro κ se ha modificado de manera importante, de $\kappa = 0.2$ a $\kappa = 0.8$ ($0 \leq \kappa \leq 1$), y a pesar de ello la diferencia entre ambas curvas casi no se percibe. Esto podría generar conclusiones opuestas sobre la permeabilidad κ , ya que en el primer caso se considera que esta se encuentra principalmente en vórgulos y en el segundo en fracturas.

Los ejemplos anteriores son el resultado de una combinación de diferentes situaciones, donde el efecto de cada parámetro en cuanto al cambio en magnitud de la presión es distinto entre parámetros, así como un efecto en el cual cada parámetro influye de manera no necesariamente lineal, y donde también entra en juego la compleja interacción entre los mismos. Adicionalmente, cada parámetro tiene un rol distinto dependiendo el intervalo de tiempo que se este estudiando.

Los efectos antes mencionados se estudian en detalle en el Capítulo 3. Sin embargo es claro que en ocasiones el ajuste de la curva a los datos requiere un alto grado de precisión para una adecuada estimación, por lo que es necesario el uso de diferentes técnicas que tomen en cuenta todo lo anterior durante el proceso de optimización.

Parte I

Problema directo

Capítulo 2

Solución al modelo Penetración Total: Triple Porosidad - Doble Permeabilidad

La solución del sistema de EDP (1.3)-(1.5) para las condiciones iniciales y de frontera (1.7a)-(1.7d) puede ser obtenida analíticamente en el espacio de Laplace a través de la transformada de Laplace.

Denotemos con

$$\mathbf{x} = (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa, s_f, s_v, C_D) \in \mathfrak{R}^9,$$

al vector que describe los parámetros del modelo, y $\bar{p}_{wD}(\mathbf{x}, \tau_D)|_{C_D=0}$ a la solución sin dimensiones en el espacio de Laplace, para el valor τ_D (que por conveniencia, le llamaremos tiempo en Laplace) en el modelo (1.3)-(1.5) con las condiciones iniciales y de frontera (1.7a)-(1.7d), sin considerar el almacenamiento en pozo.

Definimos:

$$\Lambda_f(\mathbf{x}, \tau_D) = \lambda_{mf} + \lambda_{vf} + \omega_f \tau_D \quad (2.1a)$$

$$\Lambda_m(\mathbf{x}, \tau_D) = (1 - \omega_f - \omega_v) \tau_D + \lambda_{mf} + \lambda_{mv} \quad (2.1b)$$

$$\Lambda_v(\mathbf{x}, \tau_D) = \lambda_{mv} + \lambda_{vf} + \omega_v \tau_D \quad (2.1c)$$

Utilizando $\Lambda_f(\mathbf{x}, \tau_D)$, $\Lambda_m(\mathbf{x}, \tau_D)$ y $\Lambda_v(\mathbf{x}, \tau_D)$ definimos:

$$b_{11}(\mathbf{x}, \tau_D) = \frac{\lambda_{mf}^2}{\Lambda_m(\mathbf{x}, \tau_D)} - \Lambda_f(\mathbf{x}, \tau_D), \quad (2.2a)$$

$$b_{12}(\mathbf{x}, \tau_D) = \frac{\lambda_{mf} \lambda_{mv}}{\Lambda_m(\mathbf{x}, \tau_D)} + \lambda_{vf}, \quad (2.2b)$$

$$b_{22}(\mathbf{x}, \tau_D) = \frac{\lambda_{mv}^2}{\Lambda_m(\mathbf{x}, \tau_D)} - \Lambda_v(\mathbf{x}, \tau_D) \quad (2.2c)$$

Para simplificar la notación, en adelante omitimos la dependencia (\mathbf{x}, τ_D) . Entonces:

$$h_0 = (\kappa b_{22} - (1 - \kappa) b_{11})^2 + 4\kappa (1 - \kappa) b_{12}^2 \quad (2.3)$$

Con b_{11} , b_{12} , b_{22} y h_0 definimos

$$g_1 = -\frac{(1-\kappa)b_{11} + \kappa b_{22} + \sqrt{h_0}}{2\kappa(1-\kappa)} \quad (2.4a)$$

$$g_2 = -\frac{(1-\kappa)b_{11} + \kappa b_{22} - \sqrt{h_0}}{2\kappa(1-\kappa)} \quad (2.4b)$$

$$g_3 = \frac{\kappa b_{22} - (1-\kappa)b_{11} + \sqrt{h_0}}{2\kappa(1-\kappa)b_{12}}, \quad (2.5a)$$

$$g_4 = \frac{\kappa b_{22} - (1-\kappa)b_{11} - \sqrt{h_0}}{2\kappa(1-\kappa)b_{12}} \quad (2.5b)$$

$$K_1 = Y_0(\sqrt{g_1}) + s_f \sqrt{g_1} Y_1(\sqrt{g_1}), \quad (2.6a)$$

$$K_2 = Y_0(\sqrt{g_2}) + s_f \sqrt{g_2} Y_1(\sqrt{g_2}) \quad (2.6b)$$

$$K_{11} = (\kappa + (1-\kappa)\kappa g_3) \sqrt{g_1} Y_1(\sqrt{g_1}) \quad (2.7a)$$

$$K_{12} = (\kappa + (1-\kappa)\kappa g_4) \sqrt{g_2} Y_1(\sqrt{g_2}) \quad (2.7b)$$

$$K_{21} = (1-\kappa g_3) Y_0(\sqrt{g_1}) + (s_f - s_v \kappa g_3) \sqrt{g_1} Y_1(\sqrt{g_1}) \quad (2.7c)$$

$$K_{22} = (1-\kappa g_4) Y_0(\sqrt{g_2}) + (s_f - s_v \kappa g_4) \sqrt{g_2} Y_1(\sqrt{g_2}) \quad (2.7d)$$

donde Y_j es la función de Bessel de segundo tipo y orden j .

Utilizando lo anterior, la solución en el espacio de Laplace de la presión sin dimensiones y sin almacenamiento en pozo, está dada por:

$$\bar{p}_{wD}(\mathbf{x}, \tau_D)|_{C_D=0} = \frac{K_{22}(\mathbf{x}, \tau_D)K_1(\mathbf{x}, \tau_D) - K_{21}(\mathbf{x}, \tau_D)K_2(\mathbf{x}, \tau_D)}{\tau_D[K_{11}(\mathbf{x}, \tau_D)K_{22}(\mathbf{x}, \tau_D) - K_{12}(\mathbf{x}, \tau_D)K_{21}(\mathbf{x}, \tau_D)]} \quad (2.8)$$

Al considerar el efecto de almacenamiento en pozo, la presión sin dimensiones, en el espacio de Laplace tiene la forma:

$$\bar{p}_{wD}(\mathbf{x}, \tau_D) = \frac{\bar{p}_{wD}(\mathbf{x}, \tau_D)|_{C_D=0}}{\tau^2 C_D \bar{p}_{wD}(\mathbf{x}, \tau_D)|_{C_D=0} + 1} \quad (2.9)$$

Es así que la solución sin dimensiones, en tiempo real sin dimensiones $p_{wD}(\mathbf{x}, t_D)|_{C_D=0}$ y $p_{wD}(\mathbf{x}, t_D)$ están dadas por:

$$p_{wD}(\mathbf{x}, t_D)|_{C_D=0} = \mathcal{L}^{-1} \{ \bar{p}_{wD}(\mathbf{x}, \tau_D)|_{C_D=0} \}_{t_D} \quad (2.10)$$

$$p_{wD}(\mathbf{x}, t_D) = \mathcal{L}^{-1} \{ \bar{p}_{wD}(\mathbf{x}, \tau_D) \}_{t_D} \quad (2.11)$$

donde \mathcal{L}^{-1} es la transformada inversa de Laplace.

2.1. Daño y radio efectivo

Es necesario remarcar que la solución $\bar{p}_{wD}(\mathbf{x}, \tau_D)$ en (2.9) es solo válida para $s_f \geq 0$ y $s_v \geq 0$, es decir daño en fracturas y vóculos mayor a cero, ya que de otra forma pueden producirse valores de presión negativos, lo cual no obedece a un sentido físico.

Para considerar la situación donde $s_f < 0$ o $s_v < 0$, es necesario aplicar el concepto de radio efectivo propuesto por Gringarten *et al.* [10] y Agarwal *et al.* [2], que considera que la influencia del daño es equivalente a considerar un pozo sin daño ($s = 0$) pero con un radio efectivo (diferente al radio del pozo) $r'_w = r_w e^{-s}$. Para este enfoque hay que asumir en el modelo que $s = s_f = s_v$.

Dado que en las definiciones (1.1b), (1.2b)-(1.2d) y (1.6) aparece r_w , entonces es necesario definir un escalamiento para estas variables:

$$t'_D = t_D e^{2s}, \quad \lambda'_{ij} = \lambda_{ij} e^{-2s}, \quad C'_D = C_D e^{2s} \quad (2.12)$$

Dado lo anterior, en adelante se considera que los parámetros a identificar serían:

$$\mathbf{x} = (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa, s, C_D) \in \mathfrak{R}^8 \quad (2.13)$$

Al calcular la presión con estos nuevos parámetros, obtenemos que:

$$p_{wD}(\mathbf{x}, t_D) = p_{wD}(\mathbf{x}, t'_D) |_{C_D=C'_D, \lambda_{ij}=\lambda'_{ij}, s=0} \quad (2.14)$$

Transformando a Laplace esta expresión y definiendo $\tau'_D = \tau_D e^{-2s}$, se tiene:

$$\bar{p}_{wD}(\mathbf{x}, \tau_D) = e^{-2s} \cdot \bar{p}_{wD}(\mathbf{x}, \tau'_D) |_{C_D=C'_D, \lambda_{ij}=\lambda'_{ij}, s=0} \quad (2.15)$$

Mediante las dos expresiones anteriores podemos determinar la presión para el espacio de Laplace así como en tiempo real sin dimensiones, para el caso $s < 0$. Aunque el concepto de radio efectivo se aplica también para $s \geq 0$, es recomendable solo utilizarlo cuando el daño es negativo, ya que los parámetros re-escalados pueden producir soluciones numéricas menos estables.

El supuesto $s = s_f = s_v$, se debe a que si se utilizan dos daños diferentes, se asume que la transformación del radio efectivo sería de la forma $r'_w = r_w e^{-(\alpha s_f + (1-\alpha)s_v)}$ ¹, donde la problemática consiste en determinar el valor apropiado de α . En Gómez *et al.* [14] se hace una propuesta para α , tal que $s = \kappa s_f + (1 - \kappa)s_v$.

Considerando lo anterior, se deduce que si se cumple $r'_w = r_w e^{-(\alpha s_f + (1-\alpha)s_v)}$, dados los daños (s_f^*, s_v^*) , existe un radio efectivo $r_w e^{-s^*}$ tal que si:

$$\begin{aligned} \mathbf{x}_1^* &= (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa, s_f^*, s_v^*, C_D) \\ \mathbf{x}' &= (\omega_v, \omega_f, \lambda'_{mf}, \lambda'_{mv}, \lambda'_{vf}, \kappa, 0, 0, C'_D) \\ \mathbf{x}_2^* &= (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa, s^*, s^*, C_D) \end{aligned}$$

entonces se cumple:

$$p_{wD}(\mathbf{x}_1^*, t_D) = p_{wD}(\mathbf{x}', t'_D) = p_{wD}(\mathbf{x}_2^*, t_D), \quad \forall t_D$$

¹Para verificar este supuesto, en este trabajo se realizó un experimento generando 100,000 juegos de parámetros aleatorios, a continuación se optimiza el valor de s para determinar el radio efectivo, tal que $r'_w = r_w e^{-s}$ es el radio efectivo para el conjunto de parámetros de prueba, y en el 100% de los casos se obtuvo que $\min(s_f, s_v) \leq s \leq \max(s_f, s_v)$, por lo que la evidencia numérica apoya este supuesto.

ya que el juego de parámetros \mathbf{x}_2^* necesariamente tiene un radio efectivo $r_w e^{-s^*}$, que corresponde al radio efectivo para \mathbf{x}_1^* .

En conclusión, ya que para cualesquiera s_f y s_v , existe un radio efectivo s , que genera los mismos valores de presión para el mismo conjunto de parámetros de yacimiento, entonces suponer $s = s_f = s_v$ genera un modelo equivalente, donde la estimación de los parámetros de yacimiento no se ve afectada por este supuesto, ya que el valor que se encuentre de s , será el de radio efectivo. Por lo tanto, de aquí en adelante utilizaremos un solo valor para el daño (s).

Capítulo 3

Análisis de sensibilidad

Al estudiar la sensibilidad de los distintos parámetros con respecto a la respuesta que genera el modelo (presión en el pozo), es posible generar métodos de optimización ad-hoc que tomen en cuenta esta información, y así estimar estos parámetros de forma más precisa y en menor tiempo de cómputo.

3.1. Sensibilidad de los parámetros

El modelo presentado en 1.1, correspondiente al sistema de EDP (1.3)-(1.5) devuelve la presión $\bar{p}_{wD}(\mathbf{x}, \tau_D)$, correspondiente a la solución sin dimensiones en el espacio de Laplace, y por medio de inversión numérica (algoritmo de Stehfest) es posible obtener la presión $p_{wD}(\mathbf{x}, t_D)$, correspondiente la solución sin dimensiones, en tiempo real sin dimensiones.

En los experimentos aquí presentados se estudiará el efecto de los parámetros

$$\mathbf{x} = (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa, s, C_D) \in \mathbb{R}^8$$

con respecto a la presión en Laplace $\bar{p}_{wD}(\mathbf{x}, \tau_D)$ y la presión en tiempo real $p_{wD}(\mathbf{x}, t_D)$ mediante la generación de $n = 1,000,000$ conjuntos de parámetros aleatorios, para tiempos también distribuidos aleatoriamente, considerando en adelante (para este y posteriores experimentos) los siguientes rangos para cada parámetro, los cuales están basados en la realidad del fenómeno que se modela:

$$\begin{aligned} 0 &\leq \omega_i &\leq 1, & \quad i = f, v \\ 1e-9 &\leq \lambda_i &\leq 1e-3, & \quad i = mf, mv, vf \\ 0 &\leq \kappa &\leq 1 \\ -4 &\leq s &\leq 50 \\ 0 &\leq C_D &\leq 1500 \\ 1e-8 &\leq \tau_D &\leq 1e+2 \\ 1e-2 &\leq t_D &\leq 1e+8 \end{aligned}$$

Para demostrar que la muestra de tamaño $n = 1,000,000$ es suficientemente densa para modelar todas las características de los parámetros de yacimiento¹ $\mathbf{x}_r = (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa) \in \mathfrak{R}^6$, si re-definimos cada parámetro \mathbf{x}_r , con $\mathbf{x}_{ri}^* = \frac{\mathbf{x}_{ri} - \min(\mathbf{x}_{ri})}{\max(\mathbf{x}_{ri}) - \min(\mathbf{x}_{ri})}$, entonces \mathbf{x}_r sigue una distribución uniforme, tal que $0 \leq \mathbf{x}_{ri}^* \leq 1$, y podemos determinar la probabilidad de que un nuevo valor cualquiera \mathbf{x}'_r , este alejado de \mathbf{x}_r^* , por un error mayor que p , dada la muestra actual.

Ya que una muestra n en una distribución uniforme en \mathfrak{R}^6 puede considerarse como un problema equivalente a una muestra $n' = 10$ de una distribución uniforme en \mathfrak{R} , entonces la probabilidad de que el error en los parámetros de yacimiento sea mayor al 15 % es:

$$P(|\mathbf{x}_r^* - \mathbf{x}'_r| > 0.15) = (1 - 0.15)^{10} = 0.1969$$

con un valor esperado:

$$E(|\mathbf{x}_r^* - \mathbf{x}'_r|) = \frac{1}{10 + 2} = 0.0833$$

En conclusión, para cualquier valor que no se incluya en la muestra, el error promedio en los parámetros de yacimiento con la muestra actual de $n = 1,000,000$, es de a lo más 8.4 %, y no más del 20 % de nuevos conjuntos de parámetros tendrán un error superior al 15 %, por lo que podemos considerar a esta muestra lo suficientemente densa.

Con esto en mente, realizamos el experimento de generar los n conjuntos de parámetros aleatorios² obteniendo el valor de presión para cada realización, y determinando la correlación de Spearman³ de los parámetros con los distintos valores de presión.

Al realizar múltiples veces el experimento se observa que la precisión de la estimación se encuentra en los 2 primeros decimales, por lo que esta es la precisión con la que se presentan los resultados en el Cuadro 3.1.

Se incluye en los resultados la correlación de las log-derivadas, tanto en tiempo real ($\partial p_{wD}(t_D)$), como en Laplace ($\partial \tau_D \bar{p}_{wD}(\tau_D)$), ya que como se ha comentado, la curva de la log-derivada es utilizada para poder observar el efecto de algunos de los parámetros, por lo que es de interés la correlación que muestran con los mismos.

También se agrega el valor $\tau_D \bar{p}_{wD}(\tau_D)$, que corresponde a la presión obtenida en Laplace, multiplicada por el valor de Laplace τ_D ; esto con el fin de compararlo contra la presión en tiempo real $p_{wD}(t_D)$. Observamos en el Cuadro 3.1 que los valores de presión generados por $p_{wD}(t_D)$ están correlacionados en mayor medida con $\tau_D \bar{p}_{wD}(\tau_D)$, de lo que lo están con $\bar{p}_{wD}(\tau_D)$, e igualmente las correlaciones con los parámetros son más similares entre $\tau_D \bar{p}_{wD}(\tau_D)$ y $p_{wD}(t_D)$, que entre $\bar{p}_{wD}(\tau_D)$ y $p_{wD}(t_D)$.

De lo anterior podemos deducir que es más adecuado trabajar el problema inverso en el espacio de Laplace utilizando $\tau_D \bar{p}_{wD}(\tau_D)$ que si usamos $\bar{p}_{wD}(\tau_D)$; esto no solo por la similitud a la presión en tiempo real, sino porque la sensibilidad a los parámetros es mayor. Bourgeois [4] también

¹Recordemos que los parámetros de yacimiento son nuestro principal interés, y que los de pozo son efectos secundarios, propios del pozo.

²Todos los parámetros se muestrean a partir de una distribución uniforme a lo largo de su rango, a excepción de los λ 's en los cuales la distribución uniforme es en el rango de su logaritmo en base 10, es decir, el principal valor de interés es el exponente.

³La correlación de Spearman al igual que la de Pearson, mide el grado de relación ente dos variables expresándolo mediante un coeficiente $\rho \in [-1, 1]$. pero a diferencia de la correlación de Pearson, la de Spearman tiene la capacidad de medir una relación no lineal, por lo que es más apropiada para esta situación, donde los parámetros no necesariamente mantienen una relación lineal con la presión.

	$\bar{p}_{wD}(\tau_D)$	$\tau_D \bar{p}_{wD}(\tau_D)$	$p_{wD}(t_D)$	$\partial \tau_D \bar{p}_{wD}(\tau_D)$	$\partial p_{wD}(t_D)$
ω_f	0.00	0.00	0.00	-0.00	0.00
ω_v	0.00	0.00	0.00	0.00	0.01
λ_{mf}	0.00	0.00	0.00	0.00	0.02
λ_{mv}	0.00	0.00	0.00	-0.01	0.02
λ_{vf}	0.00	0.00	0.00	-0.01	0.03
κ	0.00	0.00	0.00	-0.02	0.02
s	0.03	0.18	0.19	-0.13	0.07
C_D	-0.04	-0.11	-0.11	0.05	-0.05
τ_D	-1.00	-0.94	-0.93	0.62	-0.58
t_D	1.00	0.94	0.93	-0.62	0.58
$\bar{p}_{wD}(\tau_D)$	1.00	0.96	0.95	-0.63	0.60
$\tau_D \bar{p}_{wD}(\tau_D)$	0.96	1.00	1.00	-0.73	0.66
$p_{wD}(t_D)$	0.95	1.00	1.00	-0.75	0.68
$\partial \tau_D \bar{p}_{wD}(\tau_D)$	-0.63	-0.73	-0.75	1.00	-0.97
$\partial p_{wD}(t_D)$	0.60	0.66	0.68	-0.97	1.00

Cuadro 3.1: Correlación de Spearman entre los parámetros y la presión

recomienda utilizar $\tau_D \bar{p}_{wD}(\tau_D)$ debido a la definición de la transformada de Laplace:

$$L\{p_w(t)\} = \bar{p}_w(\tau) = \int_{t=0}^{\infty} e^{-\tau t} p_w(t) dt$$

por lo que la dimension de $\bar{p}_w(\tau)$ es *tiempo · presión*, siendo más conveniente trabajar con la expresión $\tau \bar{p}_w(\tau)$, cuya dimension es de *presión* (en este caso *psi*), al igual que $p_w(t)$. A esta expresión Bourgeois [4] le llama **presión de Laplace**.

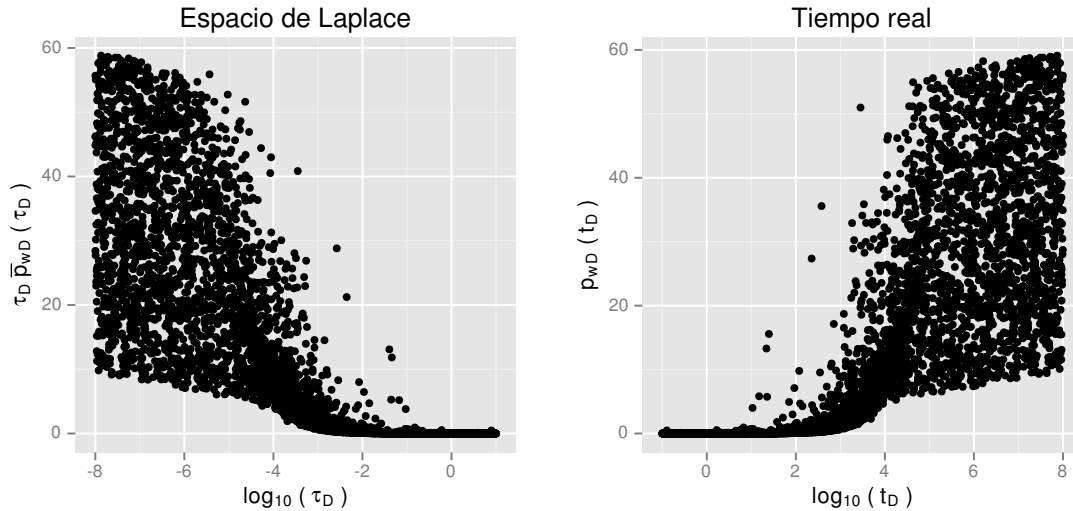
Otra ventaja al utilizar la *presión de Laplace* tiene que ver con disminuir la probabilidad de errores numéricos. En el experimento generado anteriormente tenemos que el rango de estas expresiones es:

$$\begin{aligned} \bar{p}_{wD}(\tau_D) &= [6.67e - 06, 6e + 9] \\ \tau_D \bar{p}_{wD}(\tau_D) &= [6.67e - 05, 6e + 1] \end{aligned}$$

por lo que la *presión de Laplace* posee un rango mucho más estable numéricamente al momento de operar con otras expresiones, es así que en adelante, cuando se trabaje en el espacio de Laplace, se utilizará exclusivamente la *presión de Laplace*.

Otro resultado que puede observarse de las correlaciones obtenidas, tiene que ver con el efecto que tienen los parámetros en la presión. Se aprecia que la influencia de los parámetros de pozo $\mathbf{x}_w = (s, C_D)$ es mucho mayor que la de los parámetros de yacimiento. Además la presión se encuentra principalmente dominada por los valores de tiempo τ_D y t_D .

En la Figura 3.1 observamos la dispersión tanto del parámetro de tiempo en el espacio de Laplace como en tiempo real, con respecto a la presión. Puede notarse que en efecto la *presión de Laplace* está altamente influenciada por los valores de tiempo y además, existe una mayor dispersión

Figura 3.1: Dispersión de τ_D y t_D vs presión

para valores de τ_D pequeños que para valores grandes. Esta mayor dispersión en tiempos cortos es clave para el proceso de optimización en el problema inverso, ya que implica que al existir mayor variabilidad en tiempos cortos de τ_D , que en tiempos largos, también existe “más información” de los parámetros en estos tiempos, ya que al ser más amplio el rango en el que se mueve la presión, es más sencillo detectar la variación que genera un parámetro en tiempos cortos, que en tiempos largos.

El caso de t_D vs $p_{wD}(t_D)$ es análogo, ya que $\tau_D \approx 1/t_D$; con la salvedad de que en este caso los tiempos largos son los más informativos para los parámetros. Es así que en adelante, cuando intervienen las variables de tiempo, solo se presentaran los análisis para τ_D o t_D , según se considere más conveniente; ya que los efectos que se presentan en cada espacio de tiempo son similares, considerando la anti-simetría entre ambos casos.

En la Figura 3.2 se muestra el resultado de calcular la desviación estándar para los valores de presión obtenidos con los parámetros aleatorios, para diferentes valores de tiempo. Se observa que en efecto la variabilidad de la presión es mayor a tiempos cortos de τ_D , es así que una buena práctica para el estudio del efecto de los parámetros en la presión, es obtener una muestra con espaciado logarítmico a lo largo de τ_D , obteniendo mayor densidad de puntos en tiempos cortos en Laplace que en tiempos largos. En los futuros experimentos de este trabajo también se tomara en cuenta este espaciado logarítmico, al igual que para el proceso de optimización en el problema inverso.

En las Figuras 3.3 - 3.6 se observa el efecto de los distintos parámetros en relación a los valores de presión. Puede apreciarse que en general no aparece ninguna tendencia, a diferencia de lo que sucede con la presión y el tiempo. El daño (s) es la única excepción (Figura 3.6), donde se nota como aumentan los valores de presión conforme aumenta el daño, lo cual es consistente con el

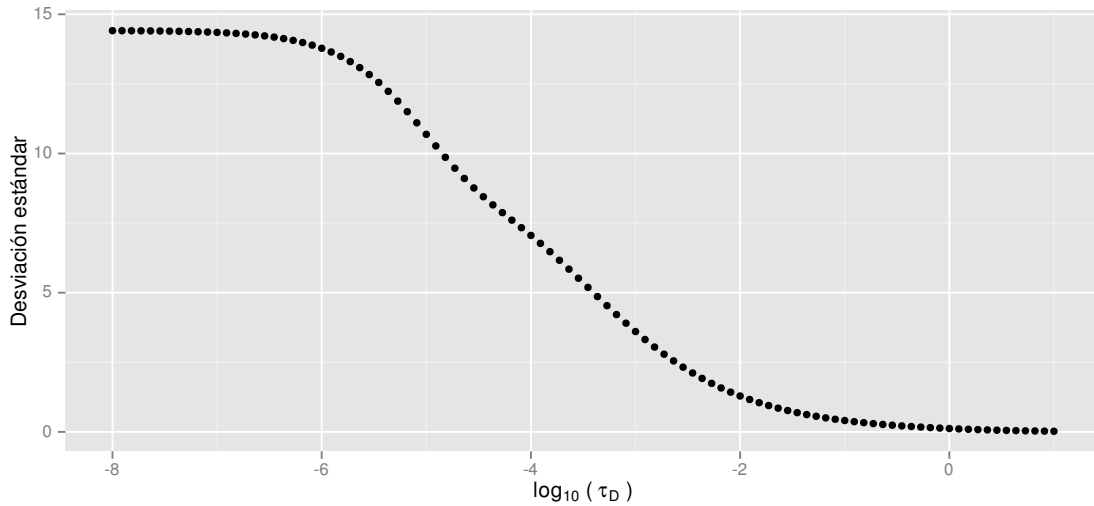


Figura 3.2: Desviación estándar de la presión, a lo largo del tiempo en Laplace

problema físico, ya que el daño tiende a mejorar o entorpecer el flujo en el pozo, generando un desplazamiento general en los valores de presión obtenidos.

Al considerar los análisis anteriores, se observa que el efecto de cada uno de los parámetros puede verse fuertemente afectado por el impacto que tiene el valor de tiempo en la presión, provocando que solo se aprecie el efecto de los dos parámetros de mayor influencia. Es así que para un adecuado estudio del efecto de los parámetros de yacimiento, es necesario separar el efecto del tiempo en cada parámetro, ya que cada uno puede tener una influencia distinta a lo largo del tiempo.

3.2. Sensibilidad: parámetro-tiempo

Para poder estudiar el efecto de cada parámetro a lo largo del tiempo, se repite el experimento de la sección anterior, con la diferencia de que los tiempos ahora se distribuyen con espaciado logarítmico a lo largo del intervalo de τ_D , en lugar de un valor aleatorio, y el análisis de la correlación de cada parámetro con la presión se realiza para cada valor de tiempo, en lugar de utilizar todo el intervalo.

De esta forma, el efecto que tienen los parámetros sobre la presión puede apreciarse mucho más claramente. Ya que solo nos interesa la magnitud del efecto, se utiliza el valor absoluto de la correlación ($|\rho| = [0, 1]$).

En la Figura 3.7 se muestra el efecto de los parámetros ω_f y ω_v (para tiempos reales), los cuales son muy similares entre sí, siendo las diferencias observadas debidas a la aleatoriedad⁴.

Observamos que el efecto más importante de estos parámetros ocurre en tiempos reales cortos

⁴Al repetir múltiples veces el experimento, las diferencias entre ambas curvas cambian ligeramente, pero la forma general y los rangos de valores se mantienen.

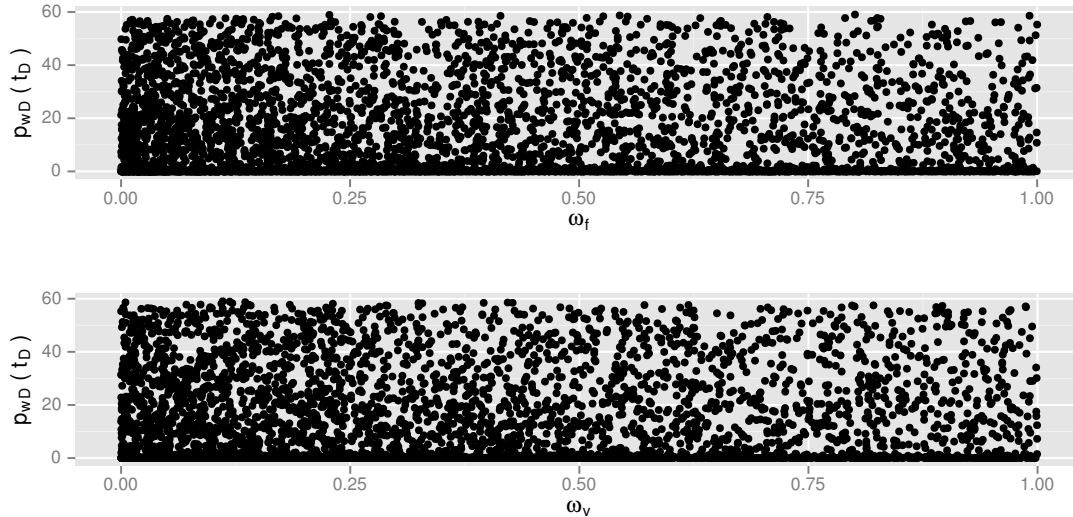


Figura 3.3: Dispersión de ω_f y ω_v vs presión

e intermedios (por lo tanto tiempos largos e intermedios en Laplace), por lo que será más fácil detectar estos parámetros al utilizar los valores de tiempo mencionados.

En la Figura 3.8 se presenta el efecto de los parámetros λ 's. λ_{mf} y λ_{mv} tienen un efecto muy similar entre sí, que se presenta en los tiempos intermedios y largos reales, de igual forma sucede con λ_{vf} , sin embargo su efecto tiene una mayor influencia en la presión.

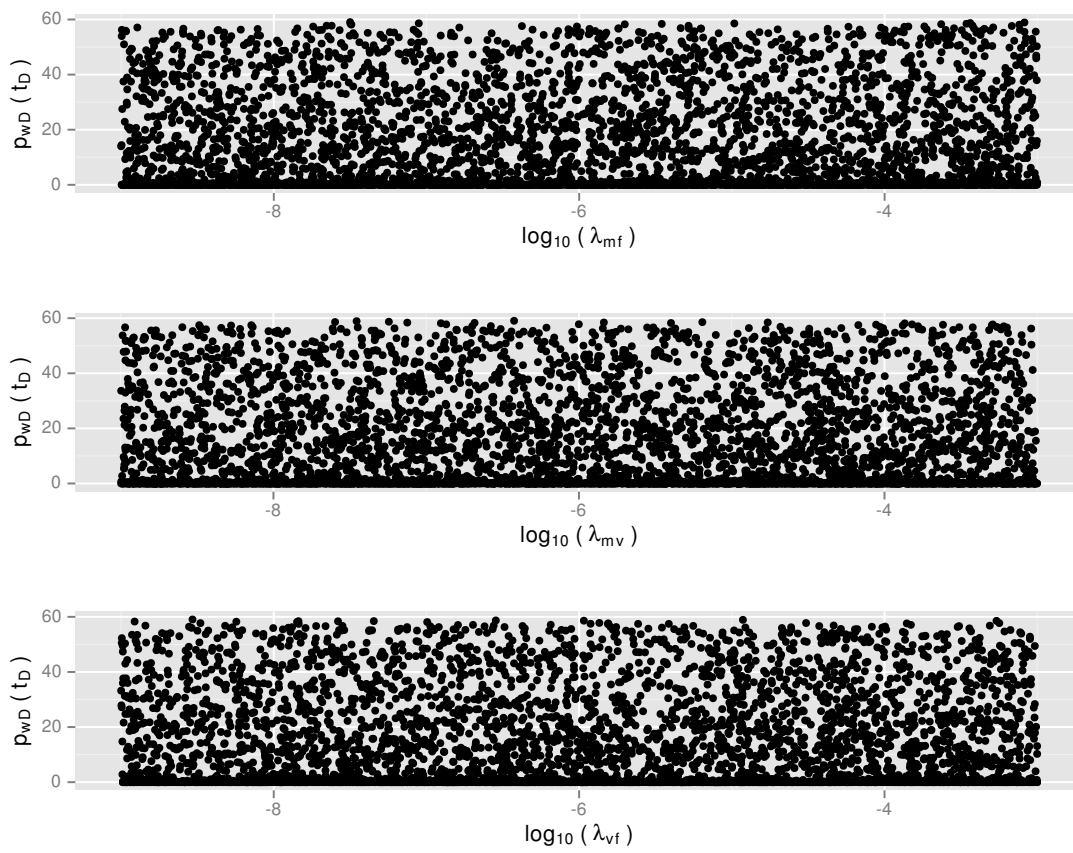
El efecto de κ se muestra en la Figura 3.9, cuya influencia más importante ocurre a tiempos cortos e intermedios, de forma similar a los ω 's,

En la Figura 3.10 se presentan todos los parámetros de yacimiento, para comparar la magnitud de su influencia, siendo los λ 's los de mayor influencia en tiempos largos reales y posteriormente los ω 's en tiempos cortos reales, seguidos de κ .

Para estudiar el caso de la Figura 3.11 donde se muestran los parámetros de pozo (s , C_D), hay que tener en cuenta que existe una diferencia entre la correlación que tiene un parámetro con la presión, y la magnitud de ese cambio. En un punto determinado de tiempo, un parámetro puede ser muy influyente en los cambios de presión, y aun así no generar cambios de gran magnitud en la misma, dificultando su detección.

Es así que en la Figura 3.11 se ejemplifica un fenómeno en el cual un parámetro que tiene un efecto de gran magnitud, suprime el efecto de otro(s) parámetro(s). El parámetro de daño (s) desplaza la presión a lo largo de todo el rango de tiempo, sin embargo, al tomar en cuenta el almacenamiento (C_D), este influye fuertemente en los valores de la presión, dando la impresión de que el efecto del daño aparece en tiempos largos reales. Sin embargo lo que en realidad sucede es que el efecto de almacenamiento es más fuerte en magnitud que el del daño, evitando que se observe el mismo.

El estudio del efecto que posee el almacenamiento en los valores de presión es de gran impor-

Figura 3.4: Dispersión de λ_{mf} , λ_{mv} y λ_{vf} vs presión

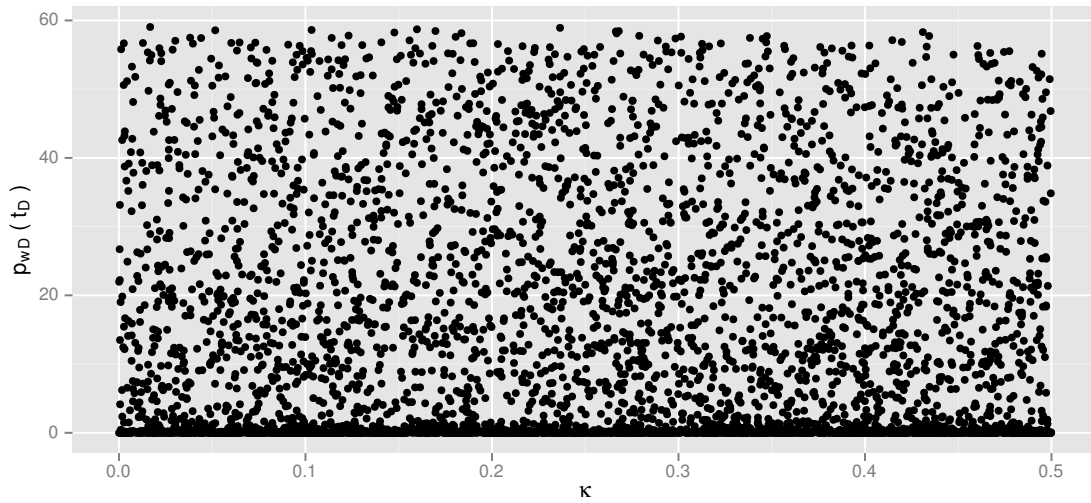


Figura 3.5: Dispersión de κ vs presión

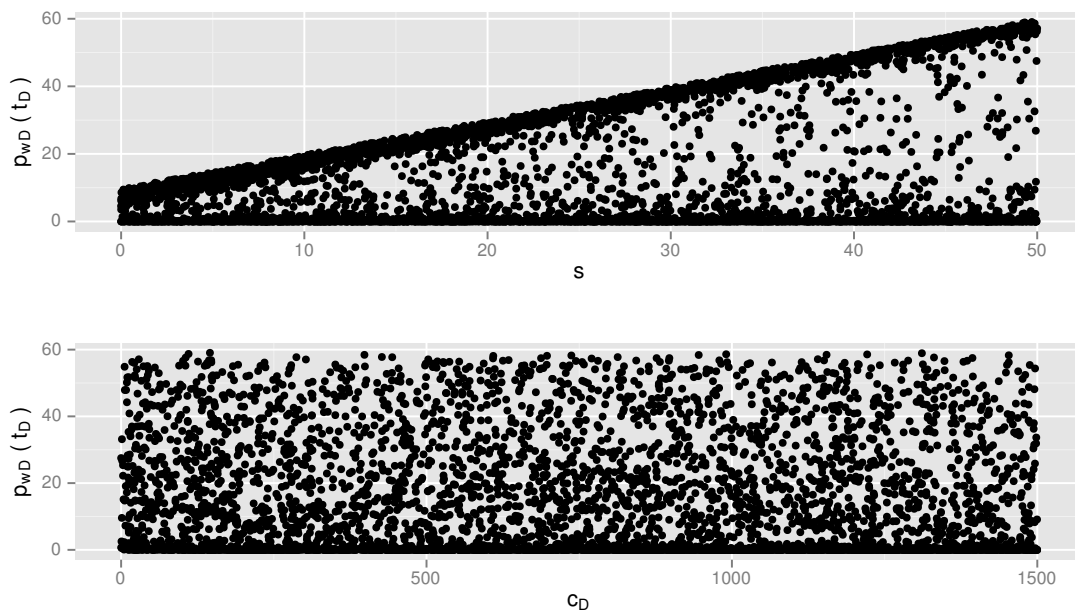


Figura 3.6: Dispersión de s y C_D vs presión

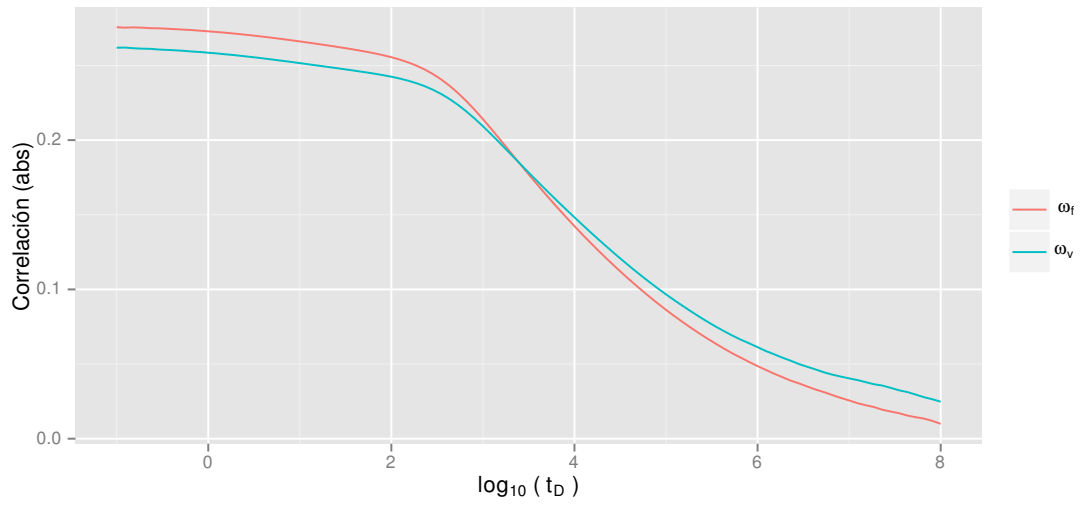


Figura 3.7: Correlacion de ω_f y ω_v con la presion, en diferentes tiempos

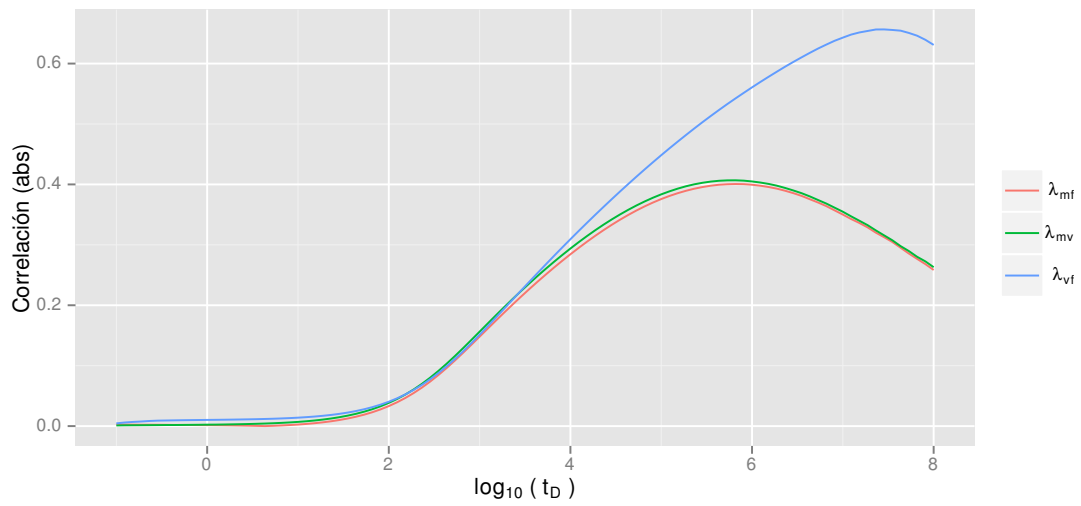


Figura 3.8: Correlacion de λ_{mf} , λ_{mv} y λ_{vf} con la presion, en diferentes tiempos

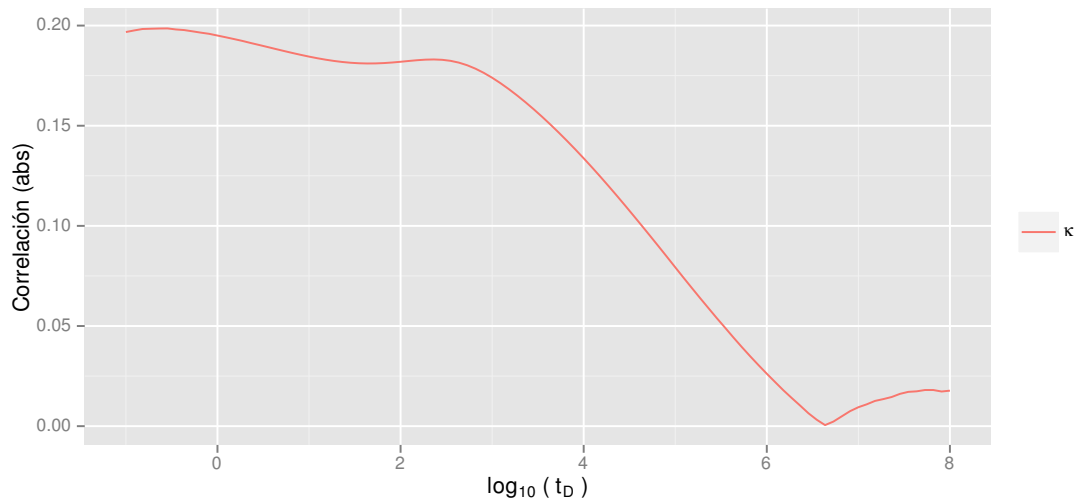


Figura 3.9: Correlacion de κ con la presion, en diferentes tiempos

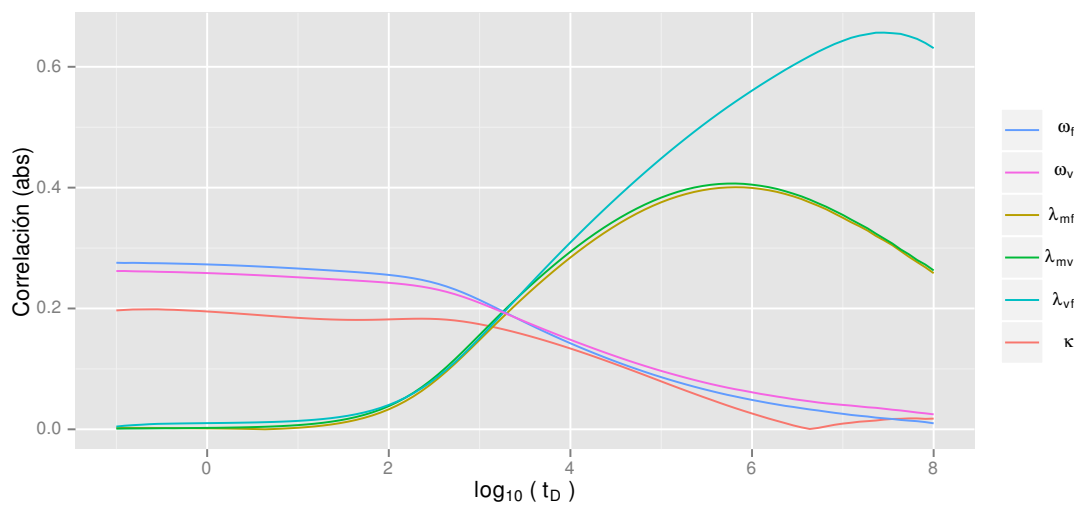


Figura 3.10: Correlacion de los parámetros de yacimiento con la presion, en diferentes tiempos

tancia para la correcta determinación de los parámetros en el problema inverso, ya que este puede suprimir el efecto de los parámetros de yacimiento; esto muestra la conveniencia de estudiar los parámetros de yacimiento y de pozo por separado, ya que si se tomaran en conjunto, daría la impresión de que ningún parámetro tiene efecto a tiempos cortos reales, a excepción del almacenamiento.

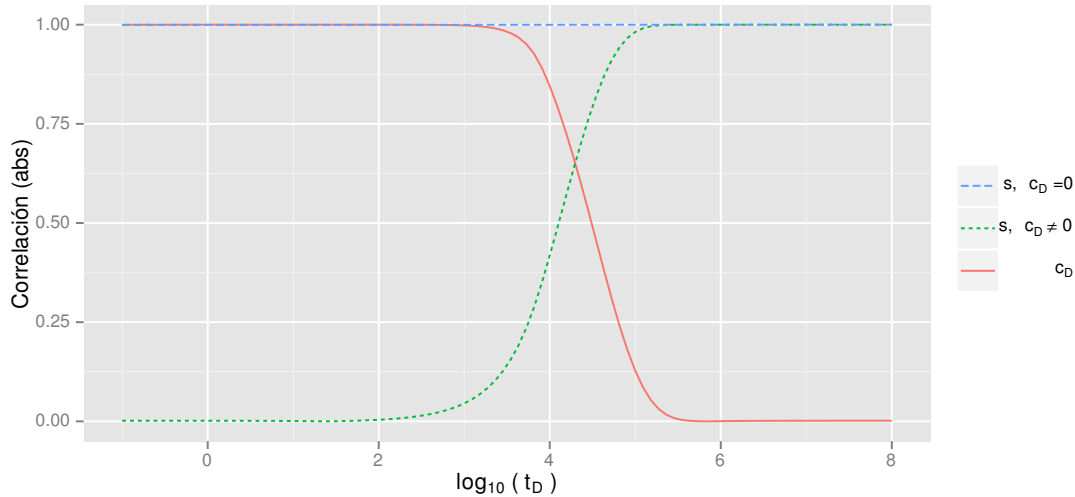


Figura 3.11: Correlación de los parámetros de pozo con la presión, en diferentes tiempos

3.3. Efecto general de los parámetros en la presión

En las secciones anteriores se estudió la influencia de cada parámetro en la presión, y se determinó cómo cada uno influye de forma distinta en diferentes puntos de tiempo.

Ahora se estudia la magnitud de estos cambios en la presión, ya que un parámetro puede ser el que más afecta a la presión en un punto determinado de tiempo, pero si ese cambio que genera en la presión es muy pequeño, conlleva a un problema al momento de estimar el valor real del parámetro.

El experimento nuevamente consiste en generar valores aleatorios de los parámetros y de los tiempos (los tiempos con espaciado logarítmico), y obtener el valor de presión correspondiente. Posteriormente, para valores específico de un parámetro se mide el valor de presión promedio, por lo que para cada valor del parámetro se tiene su efecto medio en la presión cuanto el resto de parámetros (y tiempos) varían en su rango. Esto nos permite conocer la influencia promedio en la presión que tiene un parámetro en sus diferentes valores.

En la Figura 3.12 se muestra el resultado del experimento para los parámetros ω_f y ω_v . Se observa que el efecto en la presión es mayor cuando el parámetro se encuentra en sus valores extremos, y mucho menor en valores intermedios, también observamos que la velocidad de cambio

de la presión es mayor al acercarnos a valores extremos. Por esto mismo se deduce que es más probable tener errores mayores en la estimación de estos parámetros cuando su valor es intermedio, que cuando se encuentra en los extremos, ya que alrededor de valores intermedios se producen cambios muy similares en la presión, por lo que en esta zona es más difícil diferenciar entre distintos valores del parámetro.

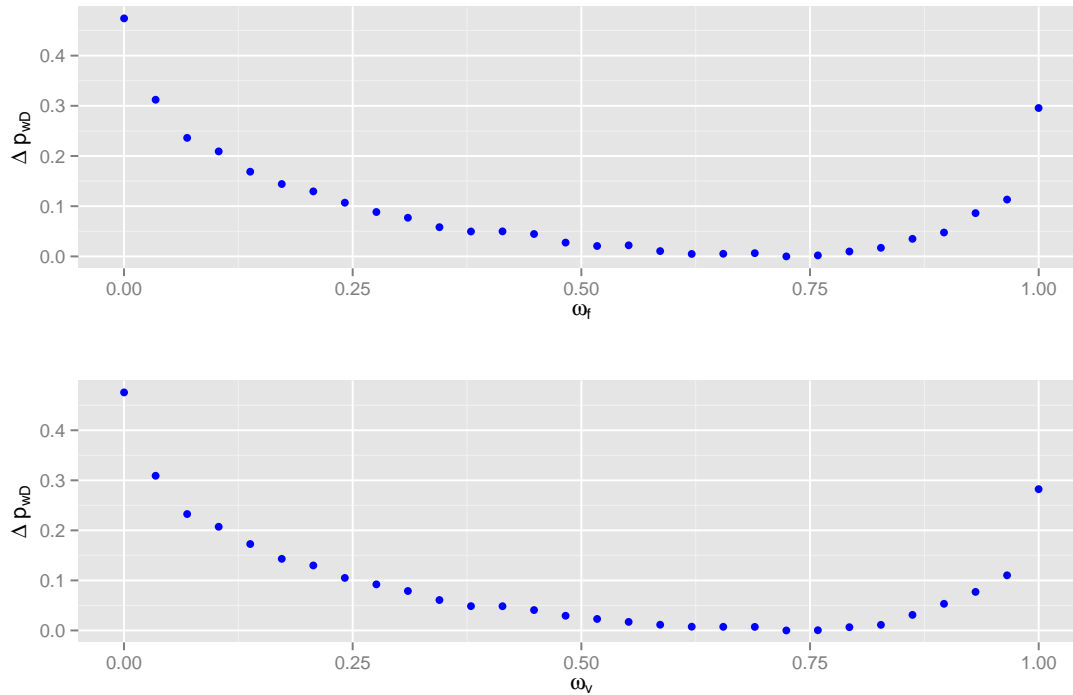


Figura 3.12: Efecto promedio de ω_f y ω_v en la presión

El efecto mínimo del parámetro ocurre en $\omega \approx 0.72$. Aunque se podría esperar que este efecto ocurra más probablemente en un valor como 0.5, la restricción $\omega_f + \omega_v \leq 1$ tiene un efecto en la generación de los valores aleatorios, donde por ejemplo si $\omega_f = 0.5$, entonces $\omega_v \leq 0.5$, por lo que esto influye en el valor que “en promedio” afecta menos al cambio de presión.

Debido a esta dependencia entre ambos parámetros, es necesario analizar su efecto conjunto. En la Figura 3.13 se presenta el resultado de variar todos los parámetros aleatoriamente y medir su efecto medio en el cambio de presión para un conjunto (ω_f, ω_v) .

Puede verse que el mayor cambio en la presión ocurre cuando ambos valores (ω_f, ω_v) son pequeños, y el menor cambio ocurre para $(\omega_f, \omega_v) = (0.5, 0.5)$. Es también alrededor de este punto donde los cambios en la presión son más similares entre sí, por lo que para valores cercanos a $(0.5, 0.5)$ será más difícil la estimación de estos parámetros con buena precisión, ya que diferentes pares (ω_f, ω_v) generan valores de presión similares.

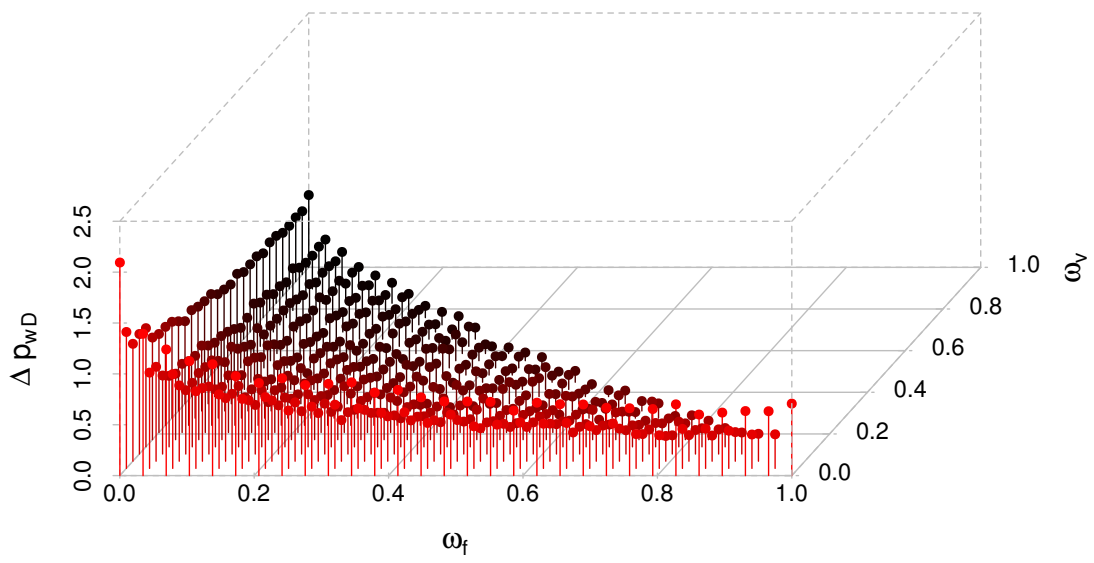


Figura 3.13: Efecto promedio conjunto de ω_f y ω_v en la presión

En la Figura 3.14 se presenta el resultado del experimento para λ_{mf} , λ_{mv} y λ_{vf} . El efecto de mayor relevancia en estos parámetros se encuentra en valores pequeños del mismo, y disminuye conforme se acerca más a su límite superior. Sin embargo debemos esperar tener mayores errores de estimación para valores pequeños del parámetro, ya que es en esa zona donde casi no se produce cambio en la presión al variar el parámetro, lo que hace difícil discriminar entre los diferentes valores.

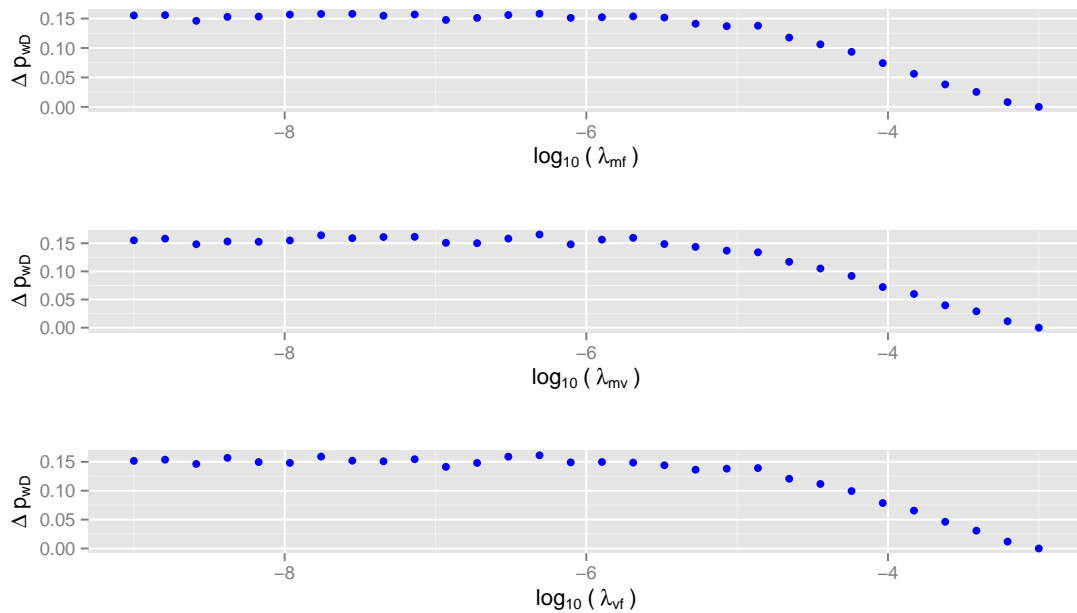


Figura 3.14: Efecto promedio de λ_{mf} , λ_{mv} y λ_{vf} en la presión

El caso de κ es muy similar al de ω_f y ω_v , siendo su mayor efecto en valores extremos y el menor en el valor central ($\kappa = 0.5$).

En cuanto a los parámetros de pozo, se observa que s aumenta su efecto conforme aumenta de valor, al contrario de lo que hace C_D .

La primera conclusión que podemos obtener de los experimentos se refiere al efecto en magnitud de cada parámetro, siendo el de mayor influencia C_D , seguido de s . Al comparar el efecto que tienen en el cambio de presión, se observa que el efecto de los parámetros de pozo está muy por arriba a los de yacimiento; esto debe tomarse en cuenta al momento de abordar el problema inverso, ya que una mala estimación de estos parámetros hará imposible la correcta estimación de los parámetros de yacimiento, que son los de principal interés.

En cuanto a los parámetros de yacimiento, los de mayor influencia en magnitud son los ω 's, seguidos por κ , y los λ 's son los parámetros de menor influencia, por lo que la determinación de estos parámetros presenta la mayor dificultad.

Otra conclusión de relevancia está dada por la curva que genera cada parámetro en su efecto en

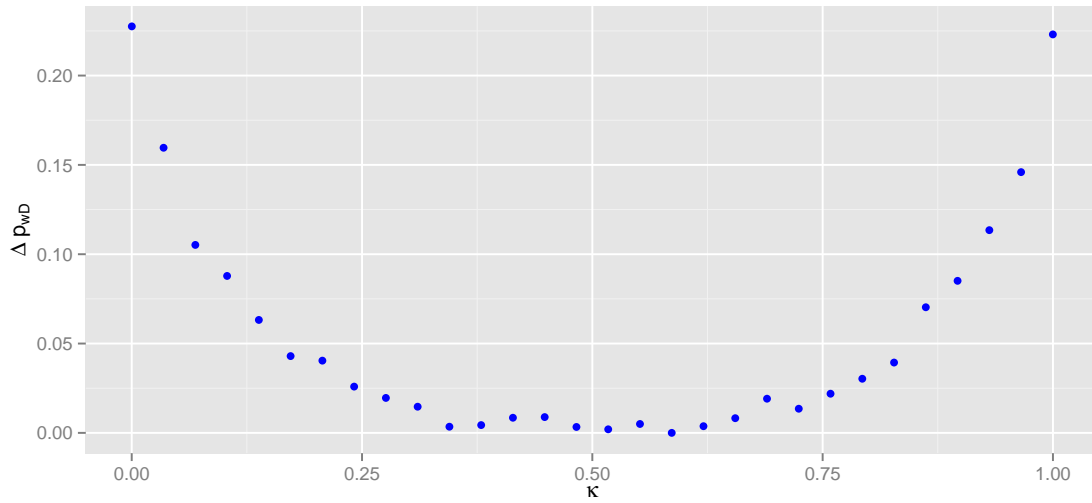


Figura 3.15: Efecto promedio de κ en la presión

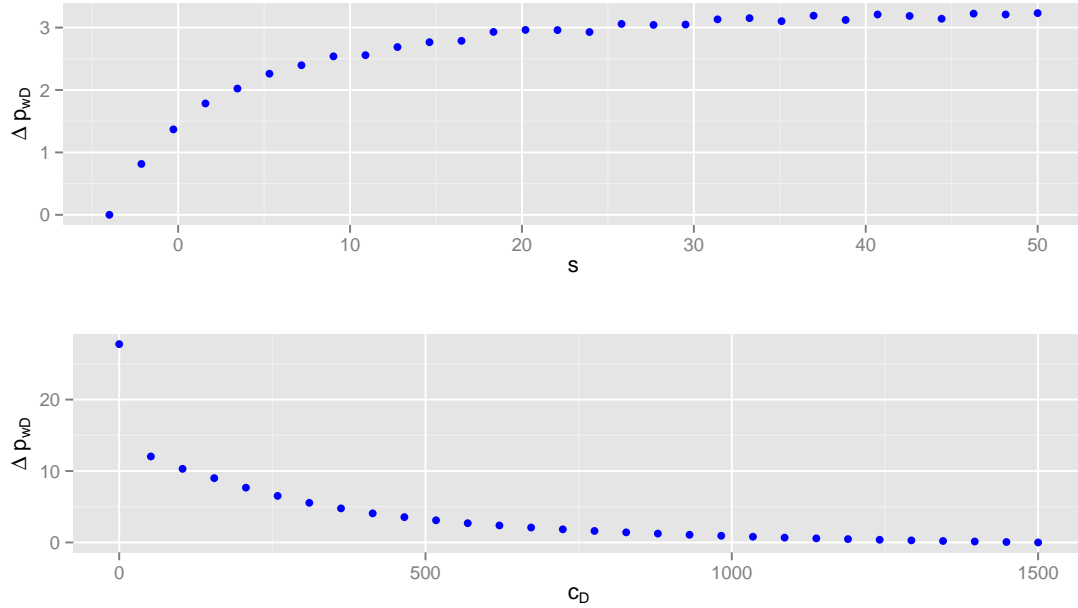


Figura 3.16: Efecto promedio de s y C_D en la presión

la presión. Esta información nos dice, por ejemplo, que las estimaciones de κ en valores cercanos a 0.5 pueden tener un mayor error, debido al débil efecto que tiene un cambio del parámetro alrededor de este punto. Por lo que valores cercanos a los extremos serán más fácil de estimar.

También nos dice que un error de 0.05 en un extremo del parámetro es de mucha mayor importancia que el mismo error alrededor de 0.5, por lo que debe tomarse en cuenta la posibilidad de una transformación de los parámetros, haciendo más lineal su efecto, y con esto tendríamos medidas de error más apropiadas. Además un efecto lineal de los parámetros es de gran importancia en la optimización, ya que un optimizador trabajará de forma más eficiente si el paso que da en cada valor del parámetro, tiene el mismo efecto en la función objetivo sin importar el valor de este.

Considerando los análisis realizados hasta ahora, se puede vislumbrar que la información obtenida por medio de los análisis de sensibilidad puede ser utilizada para realizar una optimización *ad-hoc*; tema del Capítulo 4.

Parte II

Problema inverso

Capítulo 4

Problema de optimización

En el problema inverso es necesario el recurso de la optimización para poder estimar los parámetros del modelo que generan la correspondiente información que se obtuvo durante el proceso de toma de datos; esto se realiza mediante el uso de una función objetivo, la cual determina una medida de discrepancia (o de distancia) entre dos conjuntos de datos de presión, en nuestro caso, los datos del modelo de TPDP y los de la prueba de pozo. El propósito directo de la optimización es determinar el conjunto de parámetros que minimizan el valor de la función objetivo.

La optimización es un campo de estudio muy grande y es por ello que existen muchos aspectos y métodos que pueden considerarse durante un proceso de optimización. Mediante los análisis de sensibilidad realizados en el Capítulo 3, se pretende determinar algunas de las características que más relevancia tienen durante el proceso de optimización y que están asociadas a los distintos métodos que se pueden usar para el problema inverso que nos ocupa.

Sin embargo, ya que se está desarrollando una metodología, se puede utilizar el mismo procedimiento en otros modelos, e incluso en distintos problemas inversos.

4.1. Función objetivo

Como se ha mencionado, en el problema inverso de identificación de parámetros, la función objetivo es una medida de discrepancia entre la solución que nos proporciona un modelo y los datos medidos. Sin embargo la función objetivo no solo involucra la expresión que mide la diferencia entre los dos conjuntos de datos, sino también la muestra de datos que se elige utilizar, la cual, por conveniencia también puede ser expresada con alguna transformación de los mismos datos.

En relación a lo anterior, tanto la presión generada por el modelo como la medida corresponden a valores específicos de tiempo, de los cuales mediante sensibilidad hemos determinado que existen valores de tiempo que proporcionan mayor información que otros, en el sentido de que la presión en esos tiempos es más sensible a cambios en los parámetros del modelo, y de ello concluimos que la forma más adecuada de trabajar estos datos es mediante espaciado logarítmico, a lo largo del tiempo.

También se determinó la similitud entre las expresiones $\tau_D \bar{p}_{wD}(\tau_D)$ y $p_{wD}(t_D)$, refiriéndonos a ellas como *presión de Laplace* y *presión en tiempo real*.

Utilizando los dos anteriores resultados, probaremos el efecto de distintos tipos de función

objetivo en la optimización, así como de la muestra y transformación de los datos de presión (medidos y generados). Para esto se realiza un experimento, en donde generamos datos sintéticos de presión para 5,000 combinaciones de parámetros aleatorios, que representan el valor exacto de los parámetros para cada experimento, por lo que les llamaremos conjuntos de parámetros óptimos, para los cuales posteriormente estimamos su valor mediante el uso de optimización con un Algoritmo Genético, y dado que se conocen los parámetros que generaron cada uno de los casos sintéticos, podemos medir los errores de estimación que se cometen al aplicar los distintos cambios propuestos, y es así que determinamos la confiabilidad de los resultados obtenidos mediante el uso de pruebas de hipótesis.

El rango para la generación de los valores de los parámetros y su optimización nuevamente es:

$$\begin{aligned} 0 &\leq \omega_i \leq 1, & i = f, v \\ 1e-9 &\leq \lambda_i \leq 1e-3, & i = mf, mv, vf \\ 0 &\leq \kappa \leq 1 \\ -4 &\leq s \leq 50 \\ 0 &\leq C_D \leq 1500 \\ 1e-8 &\leq \tau_D \leq 1e+2 \end{aligned}$$

donde consideramos distribuciones uniformes tanto para la generación de los valores de los parámetros como para su optimización, a excepción de los parámetros λ 's, donde la distribución uniforme es en relación a la variable $\log_{10}(\lambda_i)$.

En estos experimentos la optimización es realizada exclusivamente en el espacio de Laplace, debido a que el tiempo de cómputo es mucho menor¹. Además cada optimización se detiene únicamente cuando se ha alcanzado un número de iteraciones (generaciones) fijo, que es el mismo para todas las optimizaciones realizadas.

Ya que también se probarán expresiones diferentes para la función objetivo, es necesario hacer explícito, que al menos que se exprese lo contrario, la función objetivo utilizada durante la optimización es:

$$F(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - y_i}{y_i} \right)^2 \quad (4.1)$$

donde \mathbf{y} son los datos medidos durante la prueba de pozo (en este caso, datos sintéticos), en el espacio de Laplace y \mathbf{x} los generados por el modelo. La razón del uso de esta función se aclarará cuando se estudien los diferentes tipos de función objetivo en la sección 4.1.2.

4.1.1. Espaciado logarítmico

Realizando el experimento antes mencionado², con 5,000 optimizaciones de distintos juegos de parámetros, se compara el efecto de utilizar el espaciado logarítmico en los tiempos empleados en

¹Ya que a solución del modelo está dada en Laplace, para efectuar una optimización en tiempo real, es necesario realizar una inversión numérica mediante el Algoritmo de Stehfest, lo que conlleva calcular al menos otros 8 valores de presión en Laplace para cada valor de presión generado.

²Refiriéndonos al proceso de generar valores aleatorios de los parámetros para posteriormente estimarlos a través de optimización, utilizando para ello los datos de presión generados.

la función objetivo, como los sugieren los análisis de sensibilidad realizados.

Para diferenciar el efecto que tienen los parámetros de pozo (s, C_D) sobre las diferentes pruebas (al igual que en la optimización), se realiza un experimento donde se asumen como $s = C_D = 0$, y otro donde se agregan a la estimación, variándolos aleatoriamente a lo largo de su rango.

Para medir el error de estimación de cada parámetro utilizamos la expresión:

$$err(\hat{\theta}) = \frac{|\theta - \hat{\theta}|}{R(\theta)}$$

donde $\hat{\theta}$ corresponde a la estimación del valor exacto del parámetro θ y $R(\theta)$ el valor de su rango, por ejemplo, $R(C_D) = 1500 - 0 = 1500$.

Para el caso de λ 's, se mide el error en relación a $\log_{10}(\lambda)$, por lo que $R(\log_{10}(\lambda)) = 6$.

Es así que en el Cuadro 4.1 se muestra la media y mediana del error, obtenidas para cada parámetro, considerado espaciado lineal y logarítmico en el tiempo.

Adicionalmente se incluye el valor μ , el cual hace referencia a la media del error de estimación de los parámetros de yacimiento, para cada juego de parámetros, es decir para cada optimización i , se calcula:

$$\mu_i = \frac{1}{6} \left(err(\hat{\omega}_f)_i + err(\hat{\omega}_v)_i + err(\hat{\lambda}_{mf})_i + err(\hat{\lambda}_{mv})_i + err(\hat{\lambda}_{vf})_i + err(\hat{\kappa})_i \right)$$

por lo que es una medida de utilidad, ya que nos proporciona el error general o conjunto, cometido en cada optimización. Es así que μ es la media, de las medias del error conjunto.

	ω_f	ω_v	λ_{mf}	λ_{mv}	λ_{vf}	κ	μ
Media del error							
Lineal	0.1235	0.1253	0.2981	0.2919	0.2731	0.1777	0.2149
Logarítmico	0.0777	0.0792	0.1717	0.1775	0.1447	0.1086	0.1266
Mediana del error							
Lineal	0.0672	0.0703	0.2353	0.2302	0.2019	0.1257	0.2064
Logarítmico	0.0312	0.0303	0.0881	0.0922	0.0641	0.0517	0.1024

Cuadro 4.1: Media y mediana del error de estimacion para espaciado lineal y logaritmico ($C_D = s = 0$)

La media nos devuelve el error esperado de estimación para cada parámetro por separado, y μ en conjunto. La mediana es una medida que describe el valor por debajo del cual el 50 % de las estimaciones se encuentra.

Es así que, el error de estimación esperado para ω_f es del 12 % para espaciado lineal y 8 % para espaciado logarítmico. El error conjunto esperado de estimación es de 21 % para espaciado lineal y 13 % para espaciado logarítmico. A su vez el 50 % de las estimaciones tienen un error conjunto de estimación menor al 20 % para el espaciado lineal y menor al 10 % para espaciado logarítmico.

Como referencia, ya que el error de estimación es tal que $err(\hat{\theta}) \in [0, 1]$, y por construcción $\theta/R(\theta) \sim U(0, 1)$; si la estimación del parámetro θ fuera aleatoria, es decir, $\hat{\theta} \sim U(0, 1)$, entonces el valor esperado del error sería:

$$E(err(\hat{\theta})) = E \left(\left| \hat{\theta} - \frac{\theta}{R(\theta)} \right| \right) = \frac{1}{3}$$

Por lo tanto, entre más cercano es el error medio de estimación a 0.33, más equivale a una estimación aleatoria. Puede verse que para el espaciado lineal, el valor de las λ 's es muy cercano a una estimación aleatoria, pero es significativamente mejor con el espaciado logarítmico.

Lo anterior mantiene sentido con los análisis de sensibilidad del Capítulo 3, ya que un espaciado logarítmico en Laplace, genera mayor densidad de puntos en tiempos largos reales, los cuales (como se mostró en los análisis de sensibilidad), son donde principalmente se presenta el efecto de mayor sensibilidad en la presión para los λ 's.

Al considerar el error medio de estimación conjunto (μ), se observa que este se redujo en 41 %, lo cual es una mejora importante.

Para probar de forma confiable el efecto de usar el espaciado logarítmico, se realiza una prueba de hipótesis de igualdad de medias³ para cada parámetro, y para el error conjunto (μ).

En el Cuadro 4.2 se muestran los valores p -value⁴ para estas pruebas de hipótesis, siendo $1-p$ -value el máximo nivel de confianza para el cual se rechaza la hipótesis de igualdad de medias, por ejemplo, si el p -value es menor a 0.05, decimos que se rechaza la hipótesis de que ambas medias sean iguales, con un 95 % de confianza.

Adicionalmente se incluyen los intervalos de confianza al 95 % para la diferencia de medias ($\mu_{lin} - \mu_{log}$), por lo que si el cero se encuentra fuera de este intervalo, equivale a rechazar la hipótesis de igualdad de medias al 95 % de confianza⁵. Los valores límites del intervalo (que tan lejanos se encuentran del cero) también nos dicen para que variables es más fuerte la diferencia en sus medias.

Dicho lo anterior, de los resultados de las pruebas de hipótesis determinamos que existe diferencia estadística entre los valores medios de los errores de estimación a más del 99 % de confianza cuando contrastamos el uso de espaciado logarítmico en lugar de espaciado lineal, obteniendo una mejora en las estimaciones de todos los parámetros con el espaciado logarítmico.

De los intervalos de confianza podemos determinar que los parámetros λ 's fueron los más beneficiados por este cambio, seguidos de κ (ambos con efecto principalmente en tiempos intermedios y largos, reales -donde se aumentó la densidad de información-) y por último ω 's (con efecto en tiempos cortos e intermedios, reales).

En la Figuras 4.1 y 4.2 se presenta de forma gráfica (mediante diagramas de cajas o box plot⁶) la reducción del error para cada uno de los diferentes parámetros al usar el espaciado logarítmico.

³Al obtener la media de una muestra (media muestral), se intenta estimar con ella a la media poblacional (es decir, la media que se obtendría si se pudiera estudiar todos los posibles casos). Sin embargo, dado que se está utilizando una muestra para generar los casos sintéticos, existe un error en la estimación de la media poblacional, error que está asociado a la muestra utilizada; es por ello que es necesario una prueba que considere esta aleatoriedad en la medición, y así poder diferenciar las soluciones obtenidas con cierta "probabilidad" u intervalo de confianza. A esta prueba se le conoce como prueba de igualdad de medias.

⁴El p -value expresa la probabilidad de obtener un valor tan extremo como el que se ha obtenido, suponiendo que ambas medias son iguales, por lo que $1-p$ -value es el nivel de confianza con la que concluimos que ambas medias son diferentes.

⁵Tanto las pruebas de hipótesis como los intervalos de confianza presentados, están basadas en la distribución t de Student, que se encuentra definida en el rango de $(-\infty, \infty)$, y aunque la diferencia de medias del error es tal que $-1 \leq \mu_1 - \mu_2 \leq 1$, ya que queremos probar cuando existe suficiente evidencia para considerar a ambas medias diferentes, si al utilizar una distribución de colas pesadas como la t de Student el valor cero no se encuentra en el intervalo, mucho menos sucederá con una distribución definida en $[-1, 1]$, por lo que si concluimos con esta distribución que existe diferencia estadística entre ambas medias, entonces podemos tener la confianza de que también así será con una distribución definida únicamente en $[-1, 1]$.

⁶En el diagrama de cajas se representan los cuartiles Q_1 , Q_2 , Q_3 , o dicho de otra forma, los valores donde se encuentran el 25, 50 y 75 por ciento de los datos.

	<i>p-value</i>	Intervalo de confianza (95%)	
		Límite Inferior	Límite Superior
ω_f	$< 2.2e-16$	0.0403	0.0513
ω_v	$< 2.2e-16$	0.0405	0.0517
λ_{mf}	$< 2.2e-16$	1.0429	1.1714
λ_{mv}	$< 2.2e-16$	0.9692	1.0964
λ_{vf}	$< 2.2e-16$	1.1040	1.2319
κ	$< 2.2e-16$	0.0628	0.0752
μ	$< 2.2e-16$	0.0846	0.0921

Cuadro 4.2: Pruebas de hipótesis de igualdad de medias para espaciado lineal y logarítmico ($C_D = s = 0$)

También puede observarse que la dispersión del error se reduce, por lo que tendremos menor probabilidad de cometer errores grandes en la estimación.

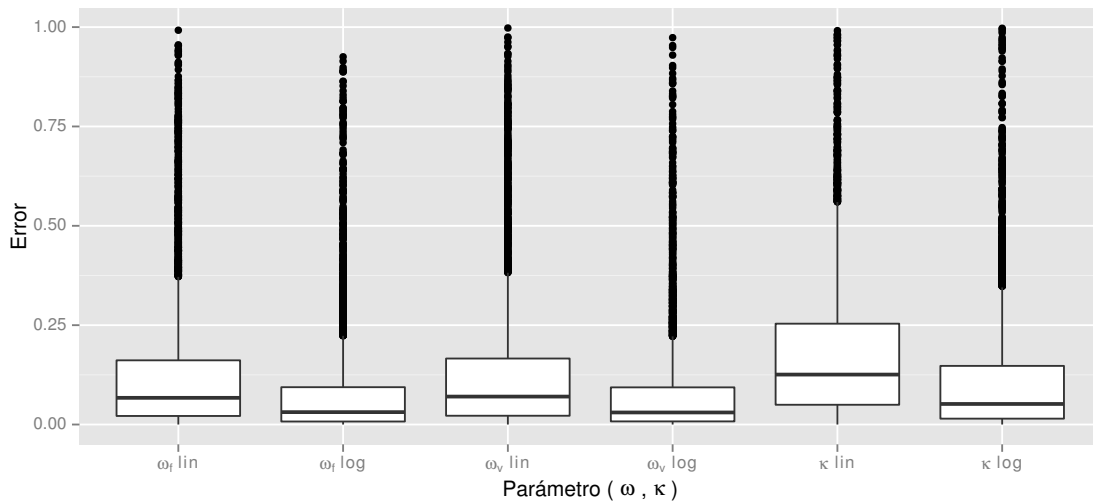


Figura 4.1: Boxplot - Error de estimacion de ω 's y κ , sin y con espaciado logarítmico ($s = C_D = 0$)

Otro resultado de interés se muestra en las Figuras 4.3 - 4.5, donde se presenta la dispersión de los parámetros a estimar *vs* la estimación realizada. Los ejemplos mostrados corresponden a los parámetros ω_f , λ_{mf} y κ , con la optimización realizada mediante el espaciado logarítmico. Se muestra en color más oscuro las zonas donde se presentan más casos, y con color claro donde ocurrieron menos casos, por lo que el situación ideal correspondería a una recta de la forma $y = x$.

Se aprecia que los resultados son consistentes con los análisis de sensibilidad del Capítulo 3, ya que ω_f y κ presentan mejores estimaciones en sus valores extremos y λ_{mf} en los valores más grandes, que como comentamos anteriormente, corresponde a las zonas donde el parámetro tiene un velocidad de cambio mayor en su efecto en la presión.

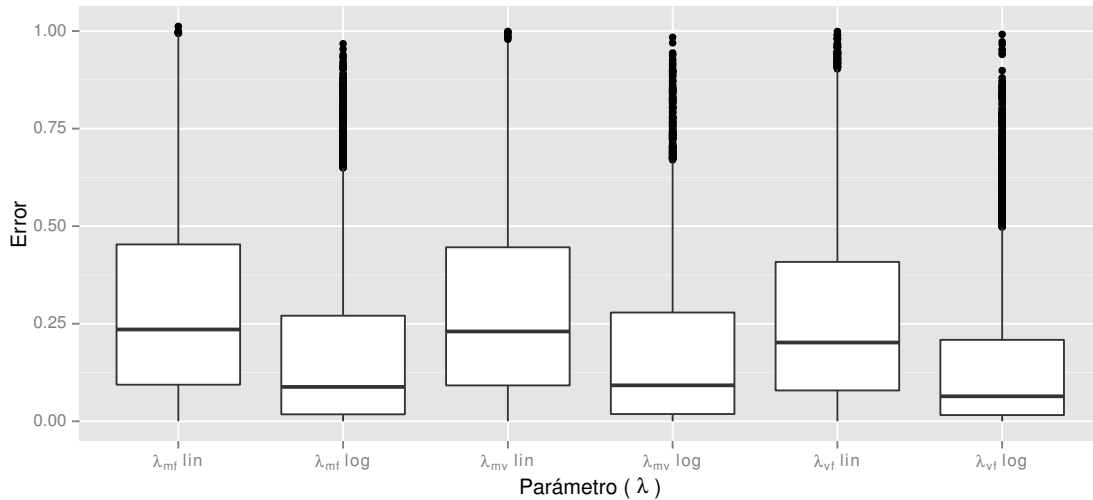


Figura 4.2: Boxplot - Error de estimacion de λ 's, sin y con espaciado logaritmico ($s = C_D = 0$)

Adicionalmente, en la Figura 4.5 se muestra la dispersión de las estimaciones de κ , es de notar que si $\kappa > 0.5$ (flujo principalmente en fracturas), entonces generalmente $\hat{\kappa} > 0.5$, es decir, incluso si existe error de estimación, este tiende a mantener el flujo principal en el mismo sistema de cavidades, por lo que es menos probable que la estimación esté por debajo de 0.5. El caso equivalente sucede para $\kappa < 0.5$.

Es así que se observa que el espaciado logarítmico produce estimaciones consistentes con lo que predicen los análisis de sensibilidad. Sin embargo, en la Figura 4.6 se presenta el efecto del parámetro λ_{mf} al utilizar espaciado lineal en la optimización, y se aprecia claramente que tiende a sobre-estimar el valor del parámetro, situación que se produce debido a la poca información que tiene el espaciado lineal en los tiempos largos reales, donde se produce el efecto de los λ 's.

Es de esta forma que el espaciado logarítmico no solo produce mejores estimaciones, sino que utiliza datos de tiempos que son de mayor utilidad al momento de estimar los parámetros menos sensibles (λ 's).

Inclusión de los parámetros de pozo

En el Cuadro 4.3 se presentan los errores de estimación al agregar los parámetros de pozo (s , C_D) en la estimación.

Se puede ver que el error de estimación de los parámetros de yacimiento, crece considerablemente. Esto se debe a que el almacenamiento obscurece gran parte de la respuesta del yacimiento, afectando la zona de respuesta de los otros parámetros involucrados.

El error promedio conjunto (μ) era de 21% para el espaciado lineal cuando no se incluían los parámetros de pozo y ahora es de 27%. Para el espaciado logarítmico cambia de 13% a 20%.

Debido a que el efecto de los parámetros de pozo es muy grande, estos se pueden identificar

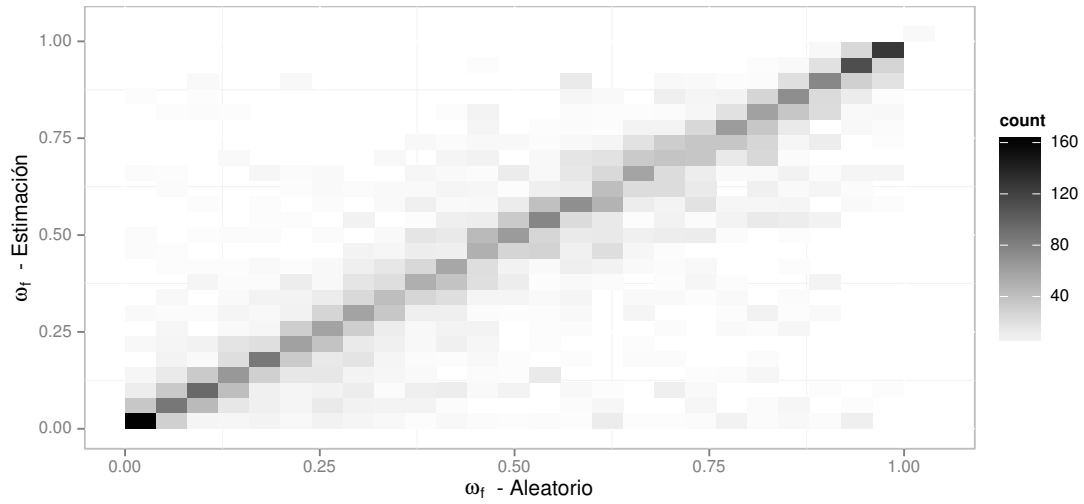


Figura 4.3: Dispersión del valor ω_f a estimar *vs* estimación (espaciado logarítmico)

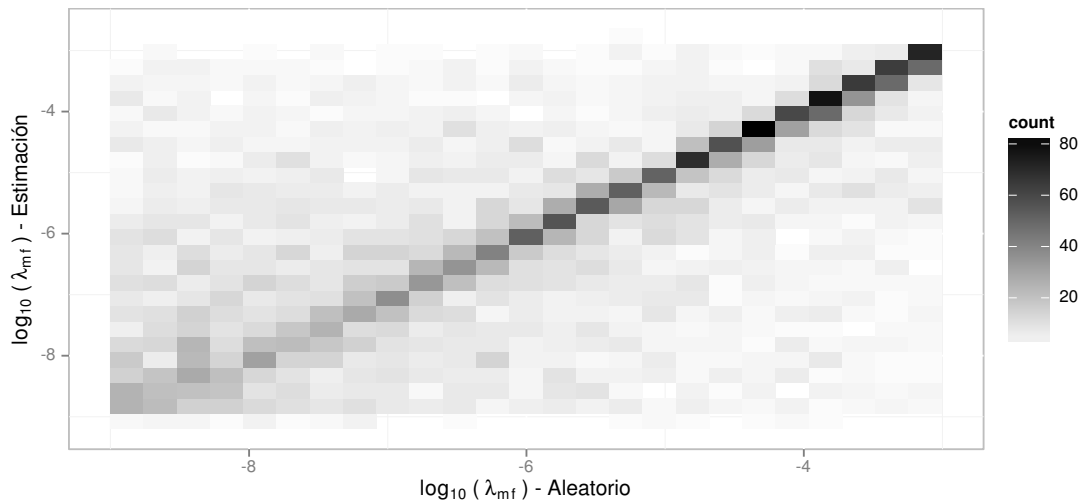


Figura 4.4: Dispersión del valor λ_{mf} a estimar *vs* estimación (espaciado logarítmico)

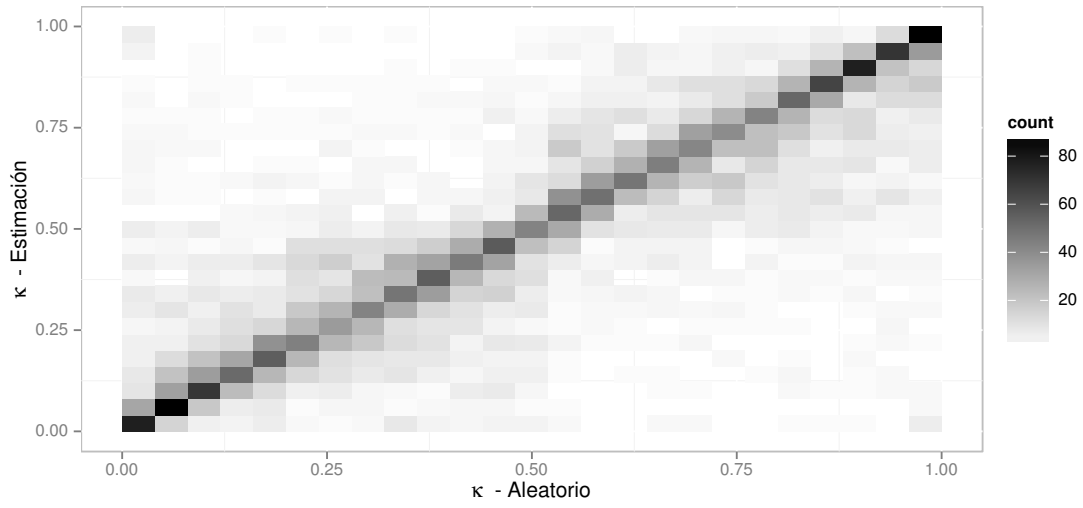


Figura 4.5: Dispersión del valor κ a estimar *vs* estimación (espaciado logarítmico)

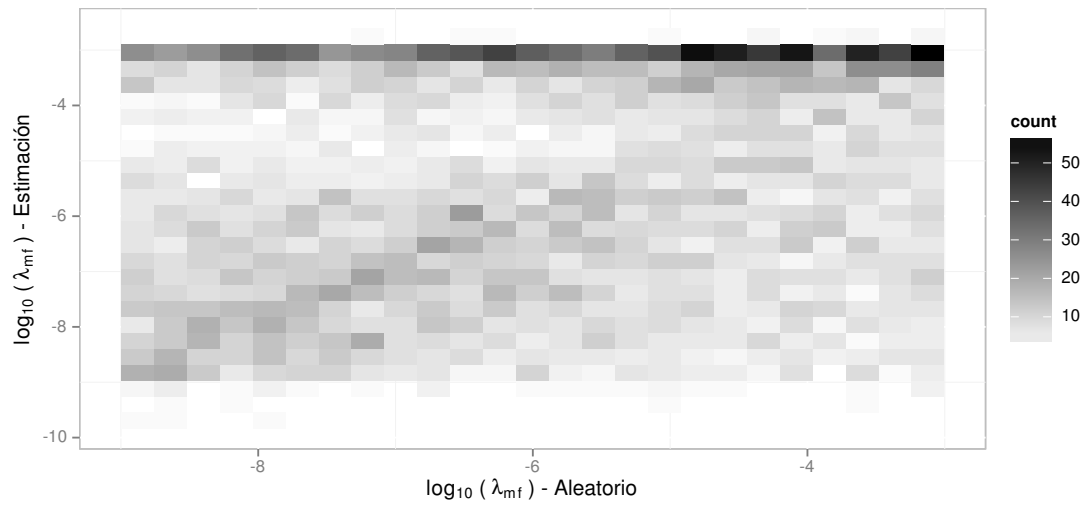


Figura 4.6: Dispersión del valor λ_{mf} a estimar *vs* estimación (espaciado lineal)

	ω_f	ω_v	λ_{mf}	λ_{mv}	λ_{vf}	κ	s	C_D	μ
Media del error									
Lineal	0.1950	0.2130	0.3377	0.3445	0.3357	0.2160	$2.10e-04$	$4.42e-09$	0.2737
Logarítmico	0.1789	0.1791	0.2364	0.2350	0.2008	0.2014	$7.76e-05$	$1.66e-05$	0.2052
Mediana del error									
Lineal	0.1354	0.1499	0.2995	0.3055	0.2936	0.1669	$8.74e-06$	$1.51e-10$	0.2688
Logarítmico	0.1149	0.1139	0.1666	0.1655	0.1274	0.1484	$1.34e-06$	$2.83e-07$	0.1984

Cuadro 4.3: Media y mediana del error de estimacion para espaciado lineal y logaritmico

fácilmente, obteniendo un error muy pequeño en su estimación. Sin embargo, debido a esta misma razón, a pesar de tener un error pequeño, este error puede afectar suficientemente a otros parámetros, en especial a los menos sensibles, como lo son los λ 's, que tienen un error cercano al 33% en el caso de espaciado lineal, equivalente al error de una estimación aleatoria.

A pesar del aumento en el error de estimación, el espaciado logarítmico sigue generando una importante mejora, con una reducción del error conjunto del 26%.

	<i>p-value</i>	Intervalo de confianza (95%)	
		Límite Inferior	Límite Superior
ω_f	$2.386e-05$	0.0086	0.0236
ω_v	$< 2.2e-16$	0.0261	0.0416
λ_{mf}	$< 2.2e-16$	0.5791	0.6804
λ_{mv}	$< 2.2e-16$	0.5749	0.6769
λ_{vf}	$< 2.2e-16$	0.7447	0.8459
κ	$7.533e-05$	0.0074	0.0219
s	$3.273e-08$	$8.58e-05$	$1.80e-04$
C_D	0.2761	$-4.64e-05$	$1.32e-05$
μ	$< 2.2e-16$	0.0647	0.0721

Cuadro 4.4: Pruebas de hipótesis de igualdad de medias para espaciado lineal y logaritmico

En el Cuadro 4.4 se presenta nuevamente los resultados para las pruebas de hipótesis de diferencia de medias para cada parámetro. Se reporta que para los parámetros de yacimiento y el parámetro s existe diferencia estadística a más del 99% de confianza, al usar el espaciado logarítmico; además, son nuevamente los λ 's los parámetros más beneficiados por este cambio, seguidos de ω 's y κ .

Para el parámetro C_D no se reporta diferencia estadística por el uso del espaciado logarítmico, lo cual es de esperar debido a la alta precisión de estimación que ya se tiene en este parámetro, y a que su efecto es en tiempos cortos, reales (recordando que el espaciado logarítmico genera mayor densidad de puntos en tiempos largos, reales).

En las Figuras 4.7 y 4.8 se presentan los diagramas de caja para los errores de estimación de los parámetros de yacimiento. Nuevamente se aprecia que no solo se reduce el error de estimación con el espaciado logarítmico, sino también la dispersión, disminuyendo la probabilidad de errores

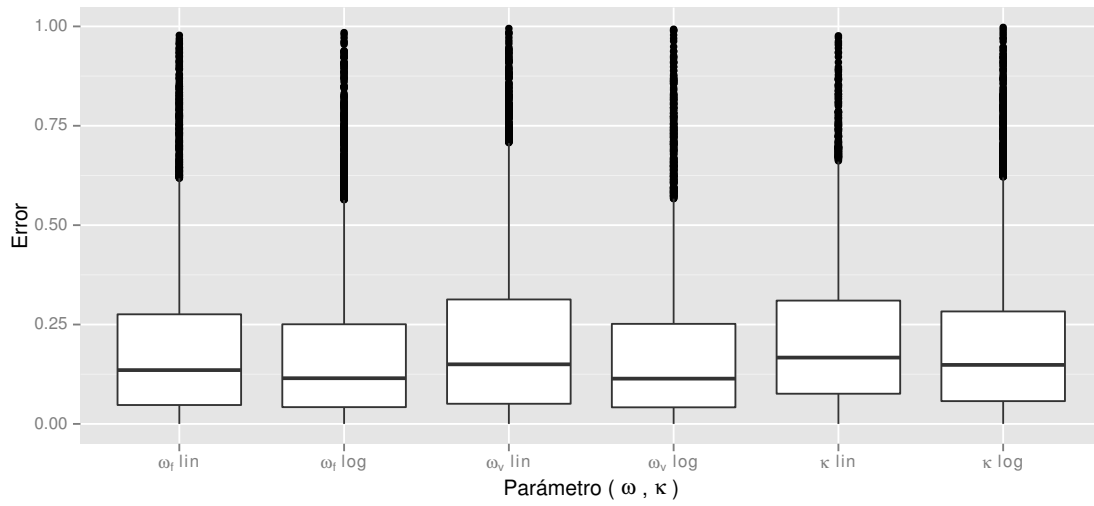


Figura 4.7: Boxplot - Error de estimacion de ω 's y κ , sin y con espaciado logaritmico

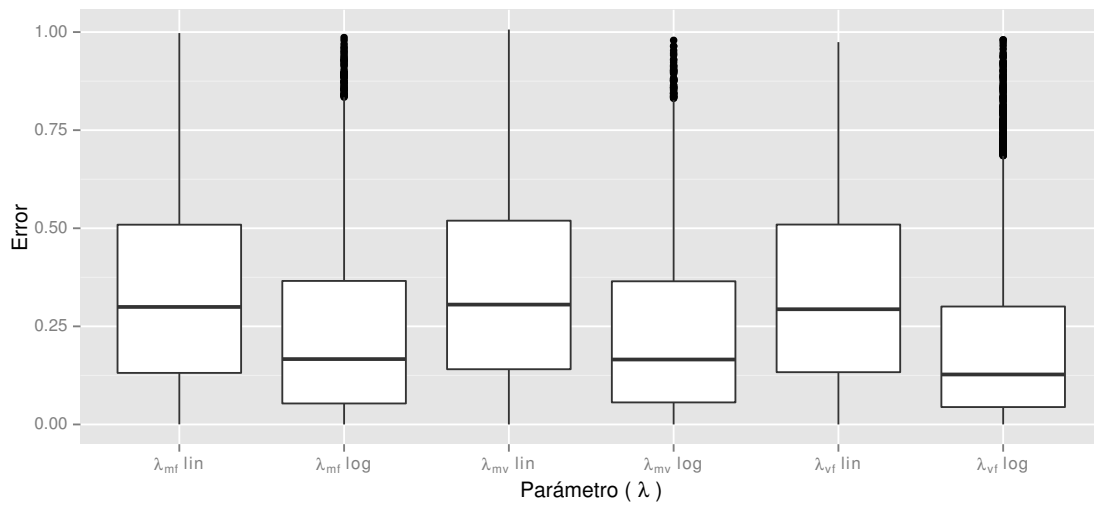


Figura 4.8: Boxplot - Error de estimacion de λ 's, sin y con espaciado logaritmico

grandes.

En cuanto a la dispersión de los valores de cada parámetro *vs* estimación, las forma general de las gráficas para los parámetros de yacimiento son muy similares a las de las Figuras 4.3 - 4.5, presentadas anteriormente, por lo que solo se agrega la Figura 4.9, donde se incluyen a los parámetros de pozo s y C_D , para los cuales observamos que su estimación es muy precisa, ya que la gráfica es de la forma $y = x$.

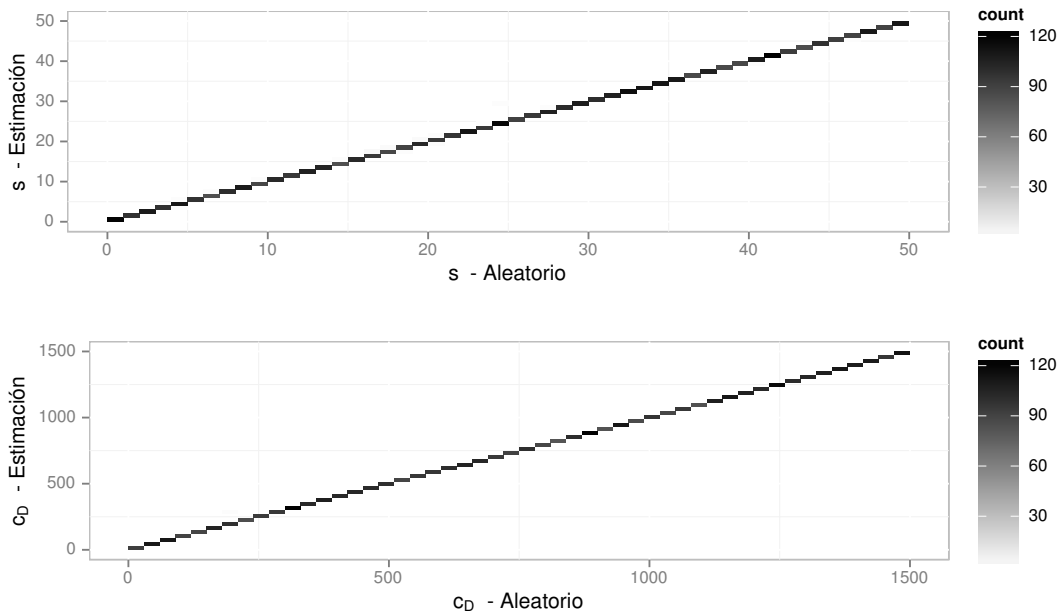


Figura 4.9: Dispersión de los valores s y C_D a estimar *vs* estimaciones (espaciado logarítmico)

4.1.2. Cambio de función objetivo

Aunque en la solución de problemas inversos de identificación de parámetros típicamente se usa la función de mínimos cuadrados (MC) como función objetivo, en la sección anterior se utilizó la función de la ecuación (4.1), ya que esta permite medir únicamente el efecto del espaciado logarítmico, sin que interfiera el efecto que genera una función como la de mínimos cuadrados, donde los valores de mayor magnitud tienen una influencia más marcada; y debido a que el espaciado logarítmico toma en cuenta una mayor densidad de puntos en tiempos cortos de Laplace, los cuales corresponden a valores de presión mayores, este efecto influye en un mayor ajuste de la función en tiempos cortos de Laplace, al usar espaciado logarítmico, que cuando no se utiliza.

Para mostrar el efecto de la función objetivo de MC y la necesidad de un cambio en su expresión, se realiza un experimento en el cual se generan 400,000 conjuntos de parámetros con su respectiva presión, en diferentes valores de tiempo τ_D , y se varían los parámetros con un error no mayor del

5%. A la presión calculada con esta variación la llamamos presión con error.

Ya que la presión original y la de error son generadas con una pequeña variación en los parámetros, una buena función objetivo debería tener un error (ajuste) “similar” en todos los tiempos, lo que implica que la correlación de la función objetivo y los tiempos debe ser cero.

Para realizar esta prueba se utilizan las siguientes funciones objetivo:

$$fsq(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (4.2)$$

$$fwsq(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - y_i}{y_i} \right)^2 \quad (4.3)$$

$$fprop(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{y_i} \right| \quad (4.4)$$

$$fexp(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n e^{\left| \frac{x_i - y_i}{y_i} \right|} \quad (4.5)$$

$$flog(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n (\log(x_i) - \log(y_i))^2 \quad (4.6)$$

donde \mathbf{y} son los datos (sintéticos) medidos durante la prueba de pozo, en el espacio de Laplace y \mathbf{x} los generados por el modelo.

Las correlaciones de Spearman obtenidas por el experimento son:

	τ_D	fsq	$fwsq$	$fprop$	$fexp$	$flog$
Correlación	1.0000	-0.9197	-0.0059	-0.0059	-0.0059	-0.0092

Claramente se aprecia que la función de mínimos cuadrados (fsq) es la única que se ve afectada por el valor del tiempo en que es utilizada, por lo que bajo este criterio, cualquiera de las otras funciones propuestas es válida.

Tras realizar varias veces el experimento se observa que la precisión se encuentra en los dos primeros decimales, sin embargo se muestran cuatro, para poder apreciar que las funciones objetivo que involucran internamente una proporción, generan el mismo efecto.

Debido a lo anterior, en los siguientes experimentos solo utilizamos las funciones fsq , $fwsq$ y $flog$. Para la función basada en proporción se selecciona la función de Mínimos Cuadrados Ponderados ($fwsq$), debido a que es la que se puede relacionar mejor con Mínimos Cuadrados (fsq) y a diferencia de las otras dos funciones, la derivada está definida en su mínimo, algo necesario para optimizadores que utilizan derivadas.

Para probar el efecto en la estimación de los parámetros, se realizan los experimentos equivalentes de la sección anterior, nuevamente con 5,000 optimizaciones de conjuntos de parámetros óptimos aleatorios, utilizando el espaciado logarítmico, que se ha demostrado es mejor. Sin embargo, ya que en la sección anterior se determinó el impacto que tienen los parámetros de pozo en la estimación de los parámetros de yacimiento, ahora se realiza únicamente el experimento donde ya se incluyen los parámetros de pozo.

En el Cuadro 4.5 se presentan las medias y medianas obtenidas para los distintos errores de estimación.

Al utilizar $flog$ o $fwsq$ en lugar de fsq , se obtiene una reducción del 10% en el error de estimación conjunto (μ), y además, en el Cuadro 4.6 se presentan los p -value para las pruebas de igualdad de medias, para los errores de estimación en los diferentes parámetros, contrastando las posibles combinaciones de función objetivo.

	ω_f	ω_v	λ_{mf}	λ_{mv}	λ_{vf}	κ	s	C_D	μ
Media del error									
fsq	0.1978	0.2069	0.2544	0.2530	0.2213	0.2273	$6.97e-05$	$2.18e-04$	0.2268
$fwsq$	0.1789	0.1791	0.2364	0.2350	0.2008	0.2014	$7.76e-05$	$1.66e-05$	0.2052
$flog$	0.1797	0.1843	0.2336	0.2394	0.2061	0.2017	$5.87e-05$	$1.36e-06$	0.2075
Mediana del error									
fsq	0.1414	0.1508	0.1928	0.1873	0.1561	0.1711	$1.04e-06$	$4.52e-05$	0.2221
$fwsq$	0.1149	0.1139	0.1666	0.1655	0.1274	0.1484	$1.34e-06$	$2.83e-07$	0.1984
$flog$	0.1185	0.1177	0.1599	0.1634	0.1360	0.1463	$7.76e-05$	$1.66e-05$	0.2017

Cuadro 4.5: Media y mediana del error de estimacion para las distintas funciones objetivo

En conclusión, al contrastar la función objetivo de mínimos cuadrados (fsq) con la de mínimos cuadrados ponderados ($fwsq$) o con la logarítmica ($flog$), se observa que existe diferencia estadística, con más del 99% de confiabilidad en los diferentes parámetros, en donde el uso de las nuevas funciones propuestas ($fwsq$ y $flog$) implican una mejora en la estimación de los parámetros, a excepción del daño (s), donde no hay suficiente evidencia de que el cambio de función objetivo influya en su estimación. Sin embargo, este es un parámetro en el que ya se tiene bastante precisión.

Adicionalmente, se observa en los intervalos de confianza que los parámetros λ 's son los más beneficiados por el cambios de función objetivo, siendo más fácilmente apreciable en la mediana, que en la media.

En cuanto al contraste de la función objetivo de mínimos cuadrados ponderados *vs* la función logarítmica, no hay evidencia estadística que indique una diferencia en las estimaciones con el uso de una u otra, por lo que son igualmente válidas.

4.1.3. Presión de Laplace

En los análisis de sensibilidad realizados en el Capítulo 3 ha sido de utilidad usar la *Presión de Laplace* en lugar de únicamente el valor que devuelve la solución del modelo de TPDP, es decir $\tau_D \bar{p}_{wD}(\tau_D)$ en lugar de $\bar{p}_{wD}(\tau_D)$, ya que la *presión de Laplace* mantiene mayor similitud con el valor de presión en tiempo real (como ya se ha mostrado en los análisis de sensibilidad).

Sin embargo, aunque el uso de la *presión de Laplace* dentro de la función objetivo tiene la conveniencia de una mayor estabilidad numérica, no tiene efecto en la estimación de los parámetros cuando se utiliza una función objetivo que involucra una proporción, ya que se cancela el efecto de τ_{D_i} , como se observa en el caso de la función objetivo de Mínimos Cuadrados Ponderados (ecuación

	Intervalo de confianza (95 %)		
	<i>p-value</i>	Límite Inferior	Límite Superior
	fsq vs fwsq		
ω_f	3.49e-07	0.0117	0.0263
ω_v	2.62e-13	0.0204	0.0352
λ_{mf}	7.51e-09	0.1073	0.2174
λ_{mv}	1.80e-09	0.1115	0.2191
λ_{vf}	1.48e-10	0.1244	0.2339
κ	2.19e-11	0.0183	0.0335
s	0.7283	-5.27e-05	3.68e-05
C_D	4.20e-07	0.0001	0.0003
μ	< 2.2e-16	0.0178	0.0253
	fsq vs flog		
ω_f	1.02e-06	0.0108	0.0254
ω_v	4.02e-09	0.0151	0.0301
λ_{mf}	3.16e-12	0.1422	0.2532
λ_{mv}	4.57e-09	0.1086	0.2175
λ_{vf}	1.94e-07	0.0916	0.2021
κ	3.76e-11	0.0180	0.0332
s	0.1855	-5.28e-06	2.72e-05
C_D	4.05e-09	0.0001	0.0003
μ	< 2.2e-16	0.0156	0.0231
	fwsq vs flog		
ω_f	0.8142	-0.0064	0.0081
ω_v	0.1724	-0.0023	0.0126
λ_{mf}	0.2308	-0.0930	0.0224
λ_{mv}	0.9362	-0.0541	0.0587
λ_{vf}	0.2602	-0.0240	0.0887
κ	0.9392	-0.0071	0.0077
s	0.3976	-6.28e-05	2.49e-05
C_D	0.3173	-4.50e-05	1.46e-05
μ	0.2551	-0.0016	0.0061

Cuadro 4.6: Pruebas de hipótesis de igualdad de medias para el uso de las distintas funciones objetivo

4.1), que hemos trabajado:

$$\begin{aligned} fwsq(\mathbf{x}, \mathbf{y}) &= \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - y_i}{y_i} \right)^2 = \frac{1}{n} \sum_{i=1}^n \left(\frac{\tau_{Di} \bar{p}_{wD}(\tau_{Di}) - \tau_{Di} \bar{p}_{wD}^*(\tau_{Di})}{\tau_{Di} \bar{p}_{wD}^*(\tau_{Di})} \right)^2 \\ &= \frac{1}{n} \sum_{i=1}^n \left(\frac{\bar{p}_{wD}(\tau_{Di}) - \bar{p}_{wD}^*(\tau_{Di})}{\bar{p}_{wD}^*(\tau_{Di})} \right)^2 \end{aligned}$$

donde $\bar{p}_{wD}^*(\tau_{Di})$ hace referencia a los datos medidos de la prueba de pozo en el espacio de Laplace, al tiempo τ_{Di} .

Es así que el uso de la *presión de Laplace* solo tiene influencia en la estimación cuando se utiliza una función objetivo como la de Mínimos Cuadrados, por lo que es esta la función que utilizaremos para medir efecto del uso de la *presión de Laplace*.

En el Cuadro 4.7 se muestra el resultado del experimento de 5,000 optimizaciones, con valores aleatorios para los parámetros de yacimiento y de pozo. En la media del error conjunto (μ) se observa que el uso de la *presión de Laplace* genera una reducción del error conjunto en 4.5%.

	ω_f	ω_v	λ_{mf}	λ_{mv}	λ_{vf}	κ	s	C_D	μ
	Media del error								
$\bar{p}_{wD}(\tau_D)$	0.2170	0.2203	0.2631	0.2541	0.2338	0.2367	$1.57e-04$	$5.23e-03$	0.2375
$\tau_D \bar{p}_{wD}(\tau_D)$	0.1978	0.2069	0.2544	0.2530	0.2213	0.2273	$6.97e-05$	$2.18e-04$	0.2268
	Mediana del error								
$\bar{p}_{wD}(\tau_D)$	0.1628	0.1651	0.2088	0.1958	0.1783	0.1860	$5.75e-08$	$7.71e-04$	0.2337
$\tau_D \bar{p}_{wD}(\tau_D)$	0.1414	0.1508	0.1928	0.1873	0.1561	0.1711	$1.04e-06$	$4.52e-05$	0.2221

Cuadro 4.7: Media y mediana del error de estimacion para $\bar{p}_{wD}(\tau_D)$ vs $\tau_D \bar{p}_{wD}(\tau_D)$

En cuanto a magnitud, la *presión de Laplace* produce la mejora de menor relevancia entre las distintas propuestas que se han estudiado hasta el momento, sin embargo, en el Cuadro 4.8 se observa que este cambio es estadísticamente significativo a más del 98% de confianza en la mayoría de los parámetros (a excepción de λ_{mv} y s), por lo que es aconsejable su uso dentro de la optimización.

4.2. Optimización secuencial

Como resultado de los análisis de sensibilidad expuestos en el Capítulo 3, se determinó que cada parámetro tiene diferente efecto en cuanto al cambio que genera en la presión, y la magnitud de ese cambio. También se establece de forma general los tiempos para los que cada parámetro tiene mayor influencia, es decir, son estos tiempos donde es más fácil determinarlo.

Por ejemplo, el parámetro C_D es el que mayor influencia tiene en la presión (al grado de que mientras este no sea estimado con suficiente precisión, es casi imposible determinar cualquiera de los parámetro de yacimiento) y se observa que su efecto es a tiempos cortos reales.

Considerando esta información para cada parámetro, es posible realizar una optimización secuencial, la cual pretende estimar primero los parámetros de mayor influencia (como lo es el

	<i>p-value</i>	Intervalo de confianza (95 %)	
		Límite Inferior	Límite Superior
ω_f	$4.09e-07$	0.0117	0.0265
ω_v	0.0005	0.0058	0.0210
λ_{mf}	$1.86e-05$	0.0601	0.1615
λ_{mv}	0.0592	-0.0019	0.0991
λ_{vf}	$1.49e-05$	0.0633	0.1678
κ	0.0178	0.0016	0.0172
s	0.1254	$-2.43e-05$	$1.99e-04$
C_D	$< 2.2e-16$	0.0045	0.0055
μ	$2.836e-09$	0.0071	0.0142

Cuadro 4.8: Pruebas de hipótesis de igualdad de medias para $\bar{p}_{wD}(\tau_D)$ vs $\tau_D \bar{p}_{wD}(\tau_D)$

almacenamiento C_D , el cual puede estimarse con alta precisión aun si no se han estimado adecuadamente los parámetros de yacimiento), dando prioridad para ello, a los tiempos donde tiene mayor influencia el parámetro en cuestión.

Con este enfoque es posible determinar de una forma más precisa los distintos parámetros, debido a que la búsqueda se puede hacer de forma más eficiente, ya que al dar prioridad a los parámetros más sensibles, estos generan menos ruido al momento de determinar los menos sensibles.

Considerando los resultados de los análisis de sensibilidad, el orden de optimización es:

$$C_D \rightarrow s \rightarrow (\omega_f, \omega_v) \rightarrow \kappa \rightarrow \lambda_{vf} \rightarrow (\lambda_{mf}, \lambda_{mv})$$

donde los parámetros ω_f y ω_v deben determinarse de forma conjunta debido a la restricción $\omega_f + \omega_v \leq 1$ (además su efecto en la presión es de magnitud similar), y por último los parámetros λ_{mf} y λ_{mv} que son los que tienen el menor efecto en la presión, y de magnitud similar.

Para los parámetros C_D , ω 's y κ se utilizan principalmente tiempos cortos reales, y para λ 's, tiempos largos. El parámetro s tiene efecto a lo largo de todo el rango de tiempo, pero debido a que el parámetro C_D interfiere con su efecto (Figura 3.11), es mejor utilizar tiempos largos⁷.

Existen varias formas de implementar una optimización de naturaleza secuencial. Sin embargo, debido a la dependencia (lineal y no lineal) que puede existir entre los parámetros, es importante que el proceso secuencial sea iterativo hasta lograr que no se generen cambios significativos en todos los parámetros.

En este trabajo la optimización de carácter secuencial se lleva a cabo usando un proceso similar al Algoritmo Genético, al cual llamaremos Optimización por Intervalos. Se utiliza este método ya que es más fácilmente (y eficientemente) transformable a una optimización de tipo secuencial. Sin embargo, mantiene similitudes con el Algoritmo Genético aquí desarrollado, ya que está basada en métodos aleatorios.

La implementación de la Optimización por Intervalos consta de tres características:

- **Diversidad:** Hace referencia a la exploración simultánea de múltiples puntos en el espacio de búsqueda, con la intención de cubrir con suficiente densidad el espacio. Al igual que como

⁷Ya que el parámetro C_D afecta también a los parámetros de yacimiento que involucran tiempos cortos reales (ω 's y κ), es conveniente usar una combinación de tiempos cortos y largos para estos parámetros.

se implementó en el Algoritmo Genético, la variabilidad de aproximaciones a la solución se realiza mediante múltiples optimizaciones que se llevan a cabo al mismo tiempo, con diferentes puntos iniciales (aleatorios), que a su vez son determinados por poblaciones iniciales aleatorias. La diversidad de soluciones tiene como finalidad encontrar el mínimo global (con buen ajuste a los datos), esperando que al explorar múltiples soluciones de forma separada, la probabilidad de que todas las aproximaciones converjan a puntos que no son mínimos globales sea mínima.

- **Aleatoriedad:** La aleatoriedad en el algoritmo está dada principalmente por el mecanismo de búsqueda de nuevos (mejores) estimados del parámetro, donde en cada iteración se prueba una nueva aproximación a la solución, seleccionada aleatoriamente en el rango del parámetro, eligiendo finalmente aquella con menor función objetivo.
- **Convergencia:** En este algoritmo la convergencia consiste en disminuir proporcionalmente el rango de cada parámetro alrededor de la solución actual, conforme se aumenta el número de iteraciones. Es decir, si LL y UL son el límite inferior y superior para el parámetro θ ; en la iteración i de un total de n iteraciones, el rango de búsqueda se encuentra dado por $[\hat{\theta} - (\hat{\theta} - LL) \cdot (1 - i/n), \hat{\theta} + (UL - \hat{\theta}) \cdot (1 - i/n)]$, donde $\hat{\theta}$ es el último estimador de θ con mejor función objetivo. La búsqueda de una nueva estimación se puede realizar de forma conjunta para todos los parámetros o de manera individual.

Es así que el método de optimización anteriormente descrito puede implementarse fácilmente de manera secuencial o no secuencial, simplemente dando un orden en el que se realiza la búsqueda de cada parámetro, o haciendo una búsqueda simultánea en todos los parámetros.

Para validar la comparación entre la Optimización por Intervalos, con el Algoritmo Genético, se realiza un experimento de optimización (incluyendo las mejoras ya estudiadas en las secciones anteriores) con 5,000 conjuntos de parámetros óptimos aleatorios con el método por Intervalos de manera no secuencial, y se contrasta el resultado con el del Algoritmo Genético, considerando el mismo tiempo de optimización en ambos métodos.

En el Cuadro 4.9 se muestra la media y la mediana para el error de estimación en ambos métodos de optimización, y se observa que existe poca diferencia entre la media del error de cada parámetro, y lo mismo ocurre con la mediana.

En el Cuadro 4.10 se presenta el resultado de las pruebas de hipótesis de igualdad de medias, para el contraste de ambos métodos, y se observa que para los parámetros de yacimiento únicamente κ posee una diferencia estadísticamente significativa, pero al considerar los valores de las medias, la reducción del error para κ es de solo 4.4%, y de 3.7% en el error conjunto⁸ (μ).

Debido a esto, se considera que el cambio en el algoritmo de optimización no presenta mejora relevante en la disminución del error de estimación, siendo así equivalentes las formas no secuenciales de ambos métodos de optimización, por lo que se considera válido realizar la comparación entre el Algoritmo Genético y la Optimización por Intervalos secuencial, para medir el efecto de esta característica secuencial.

Es así que en el Cuadro 4.11 se muestran los valores de la media y mediana para los errores de estimación de 5,000 conjuntos de parámetros óptimos aleatorios sintéticos, contrastando el

⁸Una situación donde hay diferencia estadística en la prueba de igualdad de medias poblacionales, a pesar de que no existe mucha diferencia entre los valores de las medias (muestrales) estudiadas, se debe a que la estimación de la media poblacional es muy precisa (tiene un error pequeño), generalmente por la poca variabilidad en su estimación y/o tamaños de muestra grande, por lo que también nos indica que es una buena estimación de la media poblacional.

	ω_f	ω_v	λ_{mf}	λ_{mv}	λ_{vf}	κ	s	C_D	μ
	Media del error								
Genético	0.1789	0.1791	0.2364	0.2350	0.2008	0.2014	$7.76e-05$	$1.66e-05$	0.2052
Intervalos	0.1814	0.1822	0.2202	0.2212	0.1876	0.1926	$9.97e-05$	$9.64e-05$	0.1975
	Mediana del error								
Genético	0.1149	0.1139	0.1666	0.1655	0.1274	0.1484	$1.34e-06$	$2.83e-07$	0.1984
Intervalos	0.1306	0.1299	0.1626	0.1617	0.1234	0.1427	$1.90e-06$	$4.53e-06$	0.1899

Cuadro 4.9: Media y mediana del error de estimacion para optimización con Algoritmo Genetico vs Optimización por Intervalos

	<i>p-value</i>	Intervalo de confianza (95 %)	
		Límite Inferior	Límite Superior
ω_f	0.4802	-0.0095	0.0045
ω_v	0.3945	-0.0101	0.0040
λ_{mf}	0.8647	-0.0600	0.0504
λ_{mv}	0.6178	-0.0410	0.0690
λ_{vf}	0.2568	-0.0232	0.0870
κ	0.0148	0.0017	0.0160
s	0.7643	-0.0002	0.0001
C_D	0.2545	$-2.17e-04$	$5.75e-05$
μ	$3.25e-05$	0.0041	0.0114

Cuadro 4.10: Pruebas de hipótesis de igualdad de medias para optimización con Algoritmo Genetico vs Optimización por Intervalos

efecto de una estimación mediante la optimización a través de un Algoritmo Genérico (todos los parámetros se optimizan al mismo tiempo) *vs* Optimización por Intervalos secuencial (se modifica un parámetro a la vez y en el orden propuesto anteriormente), en ambos métodos ya se considera el uso de espaciado logarítmico y de una función objetivo de MCP (es decir, se utilizan las mejoras que se han estudiado hasta el momento), y se ha ajustado el número de iteraciones del algoritmo secuencial para que en ambos casos se tenga el mismo tiempo de optimización.

	ω_f	ω_v	λ_{mf}	λ_{mv}	λ_{vf}	κ	s	C_D	μ
	Media del error								
No Secuencial	0.1789	0.1791	0.2364	0.2350	0.2008	0.2014	$7.76e-05$	$1.66e-05$	0.2052
Secuencial	0.1266	0.1304	0.1789	0.1783	0.1371	0.1400	$2.67e-05$	$2.64e-05$	0.1485
	Mediana del error								
No Secuencial	0.1149	0.1139	0.1666	0.1655	0.1274	0.1484	$1.34e-06$	$2.83e-07$	0.1984
Secuencial	0.0608	0.0611	0.0996	0.1015	0.0739	0.0670	$2.55e-07$	$2.53e-08$	0.1339

Cuadro 4.11: Media y mediana del error de estimación para optimización con método no secuencial *vs* secuencial

Aunque de forma general se observa una reducción del 30% en el error conjunto (μ), en este caso vemos que al contrario de resultados anteriores, existe una diferencia importante entre los valores de la mediana de cada experimento, situación que no se presentaba anteriormente.

En cuanto a la diferencia en la media, debe tenerse en cuenta que aun si de forma general existe un bajo error de estimación, y algunas de las estimaciones tienen un fuerte error, esto elevará considerablemente el valor de la media, ya que la media es una medida que se ve altamente afectada por valores extremos, por lo que puede observarse que en este caso es más apropiado utilizar la mediana para medir el efecto generado por la optimización secuencial.

Tomando en cuenta lo anterior, al utilizar la mediana para medir el efecto en los parámetros de yacimiento al utilizar una optimización de carácter secuencial, observamos que la disminución en los errores de estimación va desde el 39% (λ_{mv}) al 55% (κ), lo que representa una mejora importante con respecto a un método no secuencial. Además, a pesar de que la optimización no es exhaustiva, se obtienen errores de estimación aceptables.

En el Cuadro 4.12 se muestran los resultados para la prueba de hipótesis de igualdad de medias, donde concluimos que en todos los parámetros de yacimiento existe diferencia estadística entre ambos métodos y esto se garantiza a más del 99% de confianza, a favor del algoritmo secuencial. Además, en los intervalos de confianza se observa que el efecto de mejora es más notorio en los λ 's.

Por último, en las Figuras 4.10 y 4.11 se muestra mediante un diagrama de cajas el efecto en el error de estimación de cada uno de los métodos de optimización estudiados, y se observa que tanto la mediana, como la dispersión del error se reduce con la optimización secuencial, por lo que disminuye la probabilidad de cometer errores grandes en la estimación.

Para ejemplificar numéricamente el efecto de reducir la probabilidad de errores grandes, estimamos la probabilidad de que el error conjunto (μ) sea mayor a $1/3$ (el valor que corresponde a una estimación aleatoria). Obtenemos:

	<i>p-value</i>	Intervalo de confianza (95 %)	
		Límite Inferior	Límite Superior
ω_f	$< 2.2e-16$	0.0453	0.0592
ω_v	$< 2.2e-16$	0.0417	0.0557
λ_{mf}	$< 2.2e-16$	0.4624	0.5880
λ_{mv}	$< 2.2e-16$	0.4772	0.6019
λ_{vf}	$< 2.2e-16$	0.4662	0.5839
κ	$< 2.2e-16$	0.0542	0.0687
s	0.0708	$-4.32e-06$	$1.06e-04$
C_D	0.7419	$-6.83e-05$	$4.87e-05$
μ	$< 2.2e-16$	0.0529	0.0605

Cuadro 4.12: Pruebas de hipótesis de igualdad de medias para optimización con método no secuencial *vs* secuencial

$$P(\mu_{nosec} \geq \frac{1}{3}) \approx \frac{1}{5000} \sum_{i=1}^{5000} I(\mu_{nosec} \geq \frac{1}{3}) = 0.1030$$

$$P(\mu_{sec} \geq \frac{1}{3}) \approx \frac{1}{5000} \sum_{i=1}^{5000} I(\mu_{sec} \geq \frac{1}{3}) = 0.0468$$

Por lo tanto la probabilidad de cometer un error de estimación mayor a $1/3$ con la optimización de tipo secuencial (μ_{nosec}) es de 0.0468, que es menor a la mitad de la probabilidad generada por el algoritmo no secuencial.

4.3. Transformación de parámetros

En los análisis de sensibilidad realizados en el Capítulo 3 se observó que para algunos rangos de los parámetros no se produce un cambio significativo en la presión, por lo que se predecían mayores errores de estimación para esos valores (ya que al no existir suficiente diferencia en los valores de presión al variar el parámetro, es difícil detectar un cambio en el mismo), situación que se confirmó en el capítulo actual con los resultados presentados en las Figuras 4.3-4.5.

Considerando la situación anterior, puede esperarse que una transformación en los parámetros ayude a resolver este problema. A continuación probaremos si efectivamente una transformación en los parámetros resulta efectiva.

Sin embargo, en un algoritmo de optimización con base aleatoria, el realizar una transformación de los parámetros es equivalente a cambiar el mecanismo de muestreo de los parámetros. En el caso aquí estudiado, se debe aumentar la densidad de muestra para valores de los parámetros que generan poca diferencia entre sí, en los cambios de presión.

El caso de ω 's se presenta en la Figura 4.12, en la primer gráfica (línea sólida) la forma general de la función de cambio en la presión con respecto al valor de ω 's, que llamamos $\Delta p(\omega)$, siendo esta una generalización de los análisis de sensibilidad que se resumen en las Figuras 3.12 y 3.13,

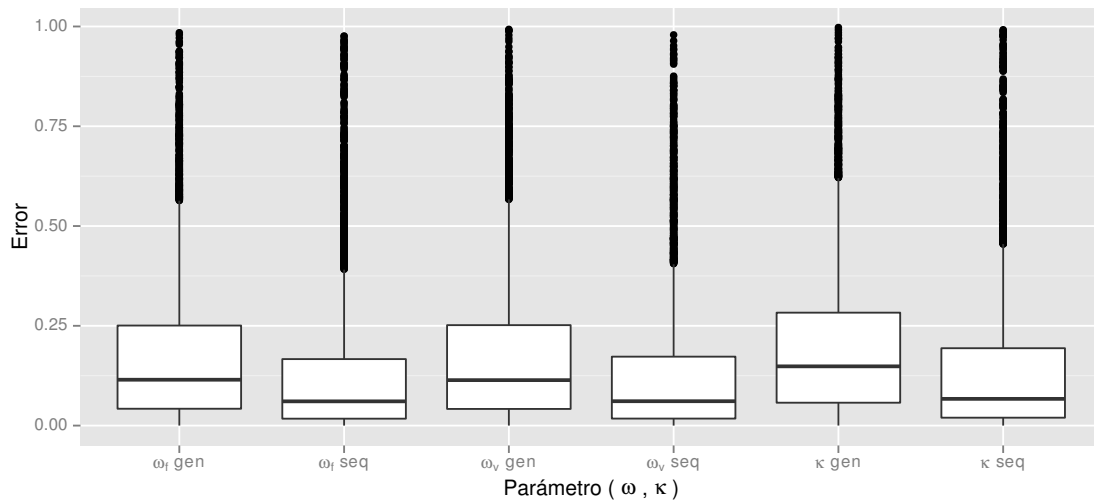


Figura 4.10: Boxplot - Error de estimacion de ω 's y κ , método no secuencial *vs* secuencial

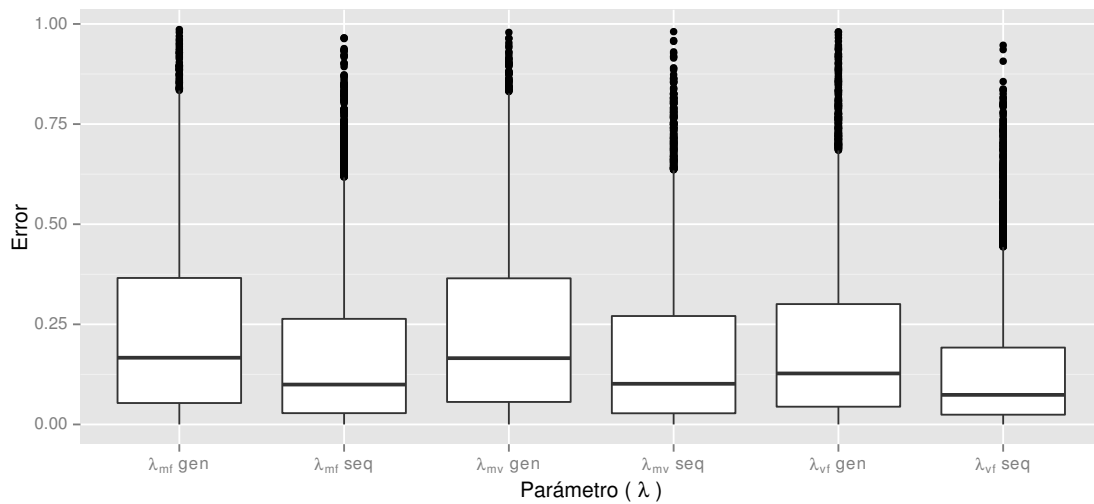


Figura 4.11: Boxplot - Error de estimacion de λ 's, método no secuencial *vs* secuencial

donde observamos que los valores centrales de los ω 's son los que tienen mayor similitud entre sí, y valores extremos los que poseen mayor diferenciación, es decir, el cambio en la curva es más fuerte en valores extremos que en valores centrales.

Para describir este efecto, se propone como forma general para el efecto de ω 's en Δp la función:

$$\Delta p(\omega) = \left(\omega - \frac{1}{\sqrt{2}} \right)^2$$

Ya que en este caso la muestra de los parámetros ω 's que queremos generar tiene una relación inversamente proporcional con el efecto del mismo en Δp , entonces podemos utilizar una modificación de esta misma curva para generar la función de densidad de probabilidad (f.d.p.) que será utilizada para el muestreo del parámetro:

$$\Delta p^*(\omega) = \begin{cases} \omega^2 & \text{si } 0 \leq \omega \leq \frac{1}{\sqrt{2}} \\ \left(\omega - \frac{2}{\sqrt{2}} \right)^2 & \text{si } \frac{1}{\sqrt{2}} < \omega \leq 1 \end{cases}$$

Esta función se presenta en la primer gráfica de la Figura 3.12 (línea punteada). Es así que:

$$P(\omega_1 \leq \omega \leq \omega_2) \propto \int_{\omega_1}^{\omega_2} \Delta p^*(\omega) d\omega$$

donde $\omega_1, \omega_2 \in [0, 1]$, con $\omega_1 \leq \omega_2$.

Utilizando $\Delta p^*(\omega)$ como el kernel de la f.d.p., generamos un muestreo por el método de Aceptación-Rechazo⁹, dando como resultado el histograma de la Figura 4.12 que corresponde a una muestra de 50,000 valores ω , generados aleatoriamente a partir de la función de distribución que produce $\Delta p^*(\omega)$.

En la Figura 4.13 se presenta el caso análogo para κ , donde la función $\Delta p(\kappa)$ representa el efecto general del parámetro con respecto al cambio de presión (Figura 3.15 en los análisis de sensibilidad), teniendo de esta forma:

$$\Delta p(\kappa) = \left(\kappa - \frac{1}{2} \right)^2$$

$$\Delta p^*(\kappa) = \begin{cases} \kappa^2 & \text{si } 0 \leq \kappa \leq \frac{1}{2} \\ (\kappa - 1)^2 & \text{si } \frac{1}{2} < \kappa \leq 1 \end{cases}$$

donde $\Delta p^*(\kappa)$ es el kernel de la f.d.p. a partir de la cual generamos los nuevos valores de κ . El la Figura 4.13 se muestra al histograma para esta función de densidad, utilizando 50,000 muestras de κ .

⁹Si tenemos una función de densidad de probabilidad $f(\theta)$ para la variable aleatoria θ , para generar muestras con la distribución $f(\theta)$ por el método de Aceptación-Rechazo se obtienen valores aleatorios y^* tal que y^* sigue una distribución uniforme de la forma $y^* \sim U(0, f_{\text{máx}})$, donde $f(\theta) \leq f_{\text{máx}}, \forall \theta \in R(\theta)$. A su vez se generan valores θ^* tal que $\theta^* \sim U(R(\theta))$. Si $y^* \leq f(\theta^*)$, entonces se acepta a θ^* como una realización de θ para la distribución $f(\theta)$, de lo contrario se rechaza.

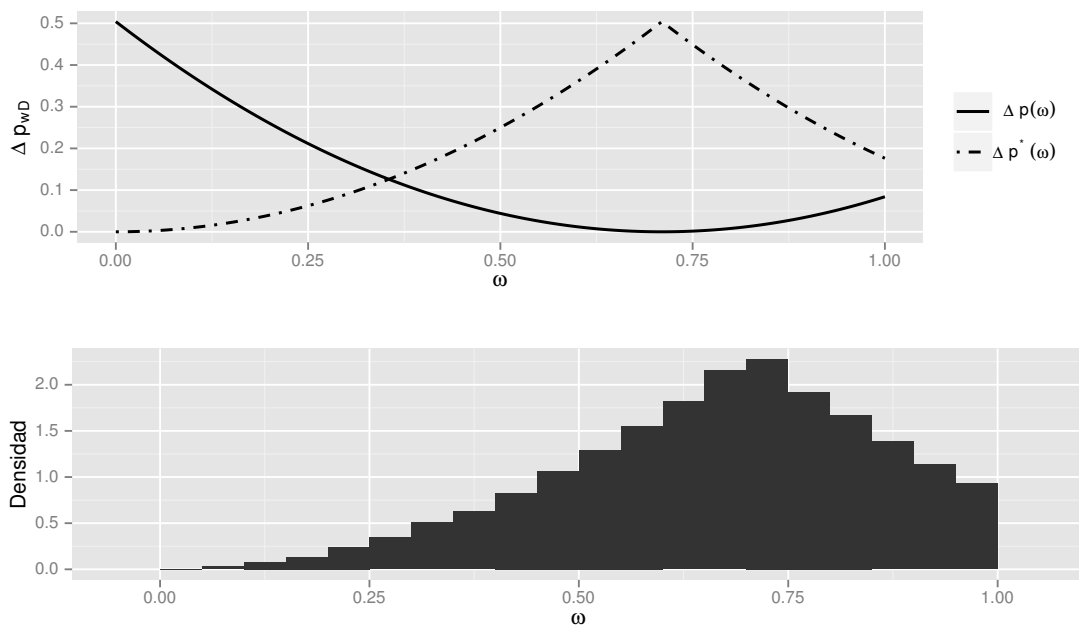


Figura 4.12: Forma general del cambio de la presión para ω 's e histograma para el muestreo correspondiente a su transformación

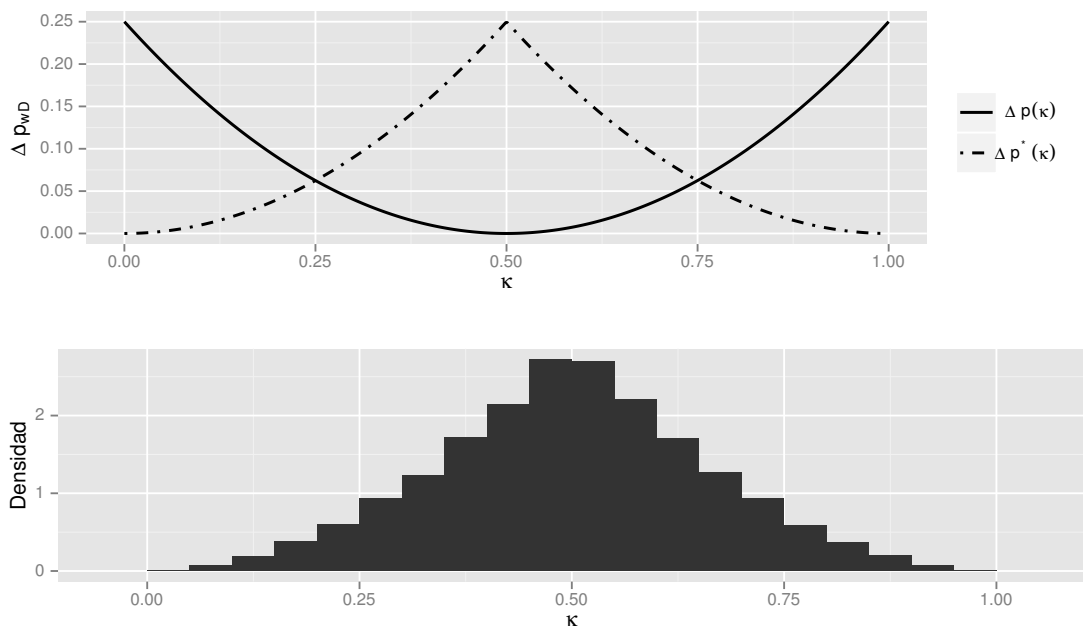


Figura 4.13: Forma general del cambio de la presión para κ e histograma para el muestreo correspondiente a su transformación

En cuanto a los λ 's, en los análisis de sensibilidad (Figura 3.14) se determinó que es alrededor de valores pequeños del parámetro donde se produce un menor cambio en Δp , por lo que son estos valores los que requieren de mayor densidad de muestreo, pero a su vez son también los valores que poseen mayor magnitud en Δp , por lo que es esta misma función la que determina el kernel de la función de densidad, tal que $\Delta p(\lambda) = \Delta p^*(\lambda)$, con:

$$\Delta p(\lambda) = -\frac{(\log_{10}(\lambda) + 9)^3}{1440} + 0.15$$

El efecto de los λ 's en el cambio de presión ($\Delta p(\lambda)$) y el histograma que corresponde a su función de densidad $\Delta p^*(\lambda)$ con 50.000 muestras, se presentan en la Figura 4.14.

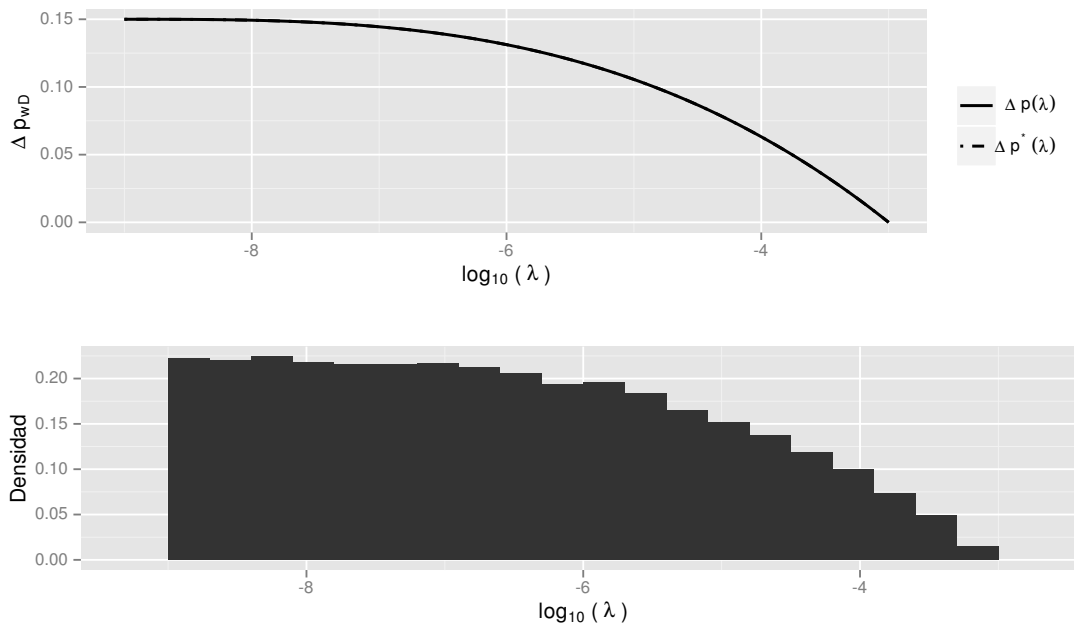


Figura 4.14: Forma general del cambio de la presión para λ 's e histograma para el muestreo correspondiente a su transformación

Utilizando las funciones de distribución que se han generado ($\Delta p^*(\cdot)$) para muestrear los parámetros de yacimiento con mayor densidad en los valores que presentan mayor problema de identificabilidad, se realiza el experimento de generar 5,000 conjuntos de parámetros óptimos aleatorios, los cuales se contrastan con los últimos cambios que se incluyen en la sección 4.2 (es decir la optimización de tipo secuencial), que involucran un muestreo uniforme sobre todos los parámetros.

La media y mediana de estos experimentos se muestran en el Cuadro 4.13, donde se observa que los valores entre ambos experimentos son muy similares, por lo que a primera vista no parece mostrar ventajas el realizar estas transformaciones.

En el Cuadro 4.14 se muestra las pruebas de hipótesis de igualdad de medias para los distintos parámetros. En el p -value (> 0.05) observamos que no existe evidencia estadística para concluir en ninguno de los parámetros que la transformación propuesta de los mismos influye de alguna forma en su identificación.

También debe considerarse que el algoritmo de optimización utilizado (Optimización por intervalos) se basa en el rango de los parámetros, por lo que es equivalente a que todos los parámetros se encuentren en la misma escala, ya que cada paso en la optimización se realiza en proporción al rango del parámetro. Por lo que ya incluye un escalamiento inicial.

Es así que a pesar de que se tiene mayor densidad de muestreo en aquellos valores de los parámetros que presentan mayor dificultad de identificación, no se generó mejora en su estimación, lo cual es debido a que la variación de los parámetros en estos rangos genera un cambio tan pequeño en la presión, que es difícil diferenciarlos.

Sin embargo en un problema real, donde los datos de presión contienen ruido, puede que no sea posible detectar estos pequeños cambios en la presión, es por ello que es más conveniente reportar un intervalo de confianza para cada uno de los parámetros, en el que consideremos con cierta confianza o “probabilidad”, se encuentra el verdadero valor del mismo¹⁰.

	ω_f	ω_v	λ_{mf}	λ_{mv}	λ_{vf}	κ	s	C_D	μ
	Media del error								
Muestreo uniforme	0.1266	0.1304	0.1789	0.1783	0.1371	0.1400	$2.67e-05$	$2.64e-05$	0.1485
Muestreo con $\Delta p^*(\cdot)$	0.1275	0.1299	0.1795	0.1820	0.1405	0.1395	$7.76e-05$	$1.66e-05$	0.1498
	Mediana del error								
Muestreo uniforme	0.0608	0.0611	0.0996	0.1015	0.0739	0.0670	$7.62e-06$	$6.28e-07$	0.1339
Muestreo con $\Delta p^*(\cdot)$	0.0556	0.0609	0.0992	0.1086	0.0806	0.0667	$2.61e-07$	$2.44e-08$	0.1355

Cuadro 4.13: Media y mediana del error de estimación para optimización con un muestreo uniforme sobre los parámetros *vs* muestreo con $\Delta p^*(\cdot)$

4.4. Análisis multivariado: Dependencia de los parámetros

Con la intención de estudiar la dependencias que existen entre los diferentes parámetros y sus errores de estimación, se obtiene la correlación parcial¹¹ de Spearman entre los errores de estimación de cada parámetro y entre el valor exacto de los diferentes parámetros, utilizando 5,000 conjuntos aleatorios de los parámetros, con su respectiva estimación y error de estimación.

En la Figura 4.15 se representa numérica y gráficamente las correlaciones antes mencionadas, antecediendo con la letra E a los errores de estimación de los parámetros para diferenciarlos del valor del parámetro. En la parte superior derecha de la gráfica se representa mediante un círculo la correlación entre cada una de las variables, siendo mayor el tamaño del mismo cuanto más fuerte

¹⁰Siempre teniendo en cuenta que un error de por ejemplo 10% en el rango de un parámetro que es poco sensible a los cambios de presión, es menos importante que el mismo error en los valores más sensibles del parámetro.

¹¹La correlación parcial mide el grado de relación entre dos variables, una vez que se ha eliminado el efecto, o la relación indirecta causada por otras variables.

	<i>p-value</i>	Intervalo de confianza (95 %)	
		Límite Inferior	Límite Superior
ω_f	0.7984	-0.0076	0.0059
ω_v	0.8679	-0.0061	0.0072
λ_{mf}	0.9912	-0.0684	0.0676
λ_{mv}	0.1454	-0.1185	0.0175
λ_{vf}	0.4518	-0.0865	0.0385
κ	0.9055	-0.0066	0.0075
s	0.2836	$-1.58e-05$	$5.41e-05$
C_D	0.3157	$-2.46e-05$	$7.61e-05$
μ	0.5078	-0.0051	0.0025

Cuadro 4.14: Pruebas de hipótesis de igualdad de medias para optimización con muestreo uniforme sobre los parámetros *vs* muestreo con $\Delta p^*(\cdot)$

es la correlación, y a la vez, si su color tiende al rojo, la correlación es negativa, y positiva si tiende al azul.

En la parte inferior izquierda de la gráfica se muestra numéricamente el valor de la correlación (ρ), donde $\rho \in [-1, 1]$, siendo mayor la correlación cuando $|\rho|$ es más cercano a 1.

Dicho lo anterior, se observa que la correlación más fuerte se da entre los valores de ω_f y ω_v , lo cual no es de particular interés, ya que esto es resultado de la restricción $\omega_f + \omega_v \leq 1$. Por esta misma razón tampoco es de relevancia la correlación que existe entre el error de estimación de ω_f ($E-wf$) y ω_v ($E-wv$).

En cuanto al valor de los parámetros, puede verse que no existe ninguna otra correlación significativa con el resto de los parámetros, lo cual es de esperar ya en los experimentos el valor de los parámetros se elige aleatoriamente.

Las siguientes correlaciones significativas son las que relacionan a la función objetivo (*obj*) con los errores de estimación de los parámetros s y C_D . Cuanto mayor es el error de estimación de s , mayor es el valor de la función objetivo, y lo mismo sucede con C_D . Estas correlaciones son fuertes debido a que s y C_D son los parámetros que mayor efecto tienen en la presión, y aunque C_D es el parámetro de mayor efecto, la correlación entre s y *obj* es mayor debido a que s tiene efecto a lo largo de todo el rango de tiempo, mientras que C_D solo tiene efecto en los tiempos cortos reales.

Es de interés que el error de C_D y s tiene tanta influencia con el valor de la función objetivo a pesar de que el máximo error de estimación obtenido para ambos parámetros es de $2.54e-05$ y $1.21e-03$, respectivamente.

En cuanto a los parámetros C_D y s también se observa que el error de estimación de C_D está correlacionado negativamente con s y positivamente con C_D , lo que quiere decir que existe mayor dificultad de estimar adecuadamente C_D cuando el valor de este es mayor y/o cuando el valor de s es menor. Esto hace sentido con los análisis de sensibilidad de C_D (Figura 3.16), donde se observaba que entre mayor era el valor del parámetro, menor era el cambio en la presión al rededor del mismo, lo cual dificulta la estimación.

Por otra parte, el error de estimación de κ se observa que influye en la estimación de los parámetros ω 's. Entre mayor es el error de estimación de κ , mayor tiende a ser el error de estimación de los ω 's. Además, el valor de κ está relacionado negativamente con el error ω_f y positivamente

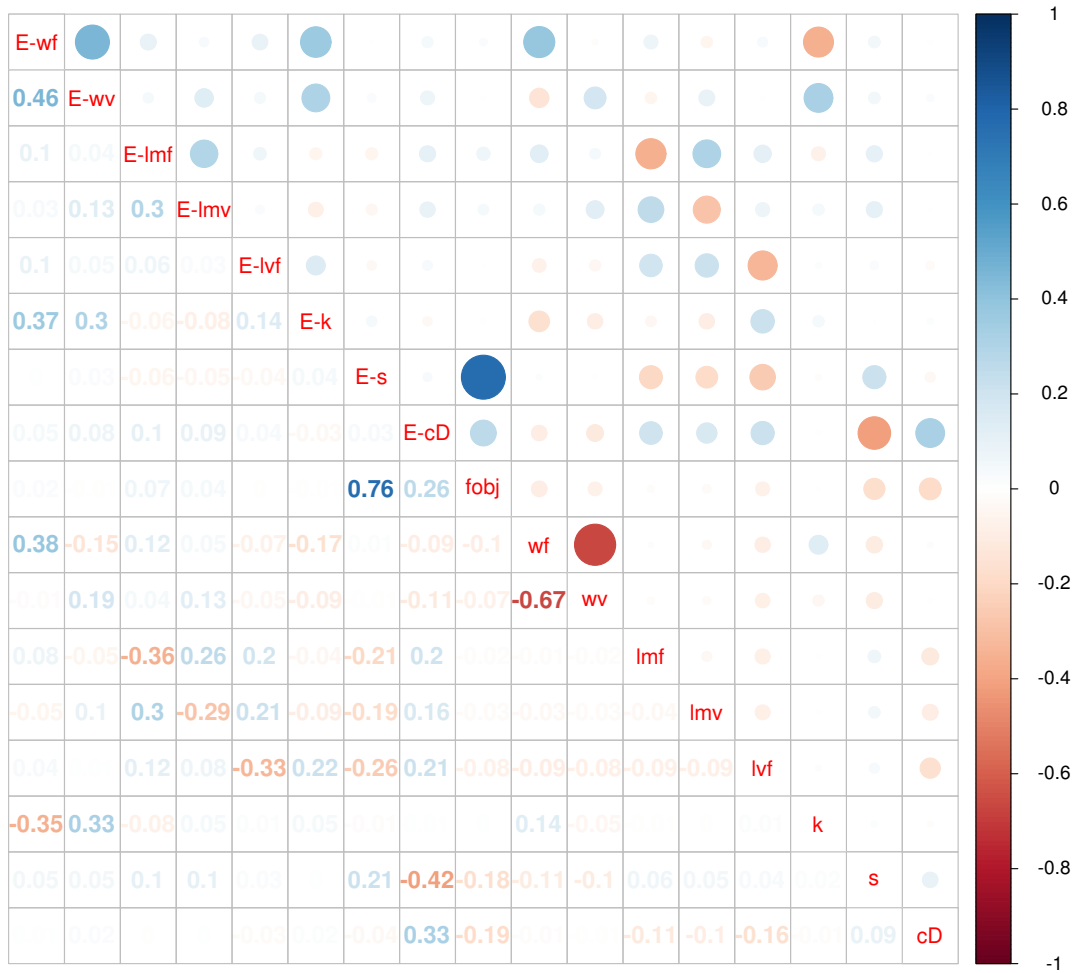


Figura 4.15: Correlación parcial (de Spearman) entre los errores de estimación y valores de los parámetros

con ω_v , por lo que entre más cercano es el valor de κ a 1, tiende a ser mayor el error de estimación de ω_v , y entre más cercano es κ a 0, tiende a crecer el error de estimación en ω_f .

Considerando la definición de los parámetros, cuando $\kappa \rightarrow 1$, el flujo o permeabilidad ocurre principalmente en fracturas, por lo que ω_v al ser un parámetro relacionado con los vóculos, no tiene mucha influencia en cambios en la presión, por lo que se dificulta su estimación. De manera similar ocurre cuando $\kappa \rightarrow 0$ para estimar ω_f .

Por ultimo, los errores de estimación de λ_{mf} y λ_{mv} están altamente relacionados entre si, pero no lo están con λ_{vf} . Adicionalmente se observa que para todo λ , su error de estimación crece conforme el valor del parámetro es menor, o dicho de otra forma, es más difícil estimar el parámetro de manera precisa conforme su valor es menor, lo cual también hace sentido con los análisis de sensibilidad antes realizados (Figura 3.14), donde se observaba que alrededor de valores pequeños del parámetro, se produce el mismo efecto en la presión, por lo que es difícil diferenciar entre estos valores del parámetro.

A su vez, también se observa que existe una relación entre el error de estimación de λ_{mf} ($E-lmf$) y el valor de λ_{mv} (lmv). Lo mismo ocurre con el error de estimación de λ_{mv} y el valor de λ_{mf} .

De los resultados anteriores podemos concluir que:

- Ya que la función objetivo se ve altamente afectada por los parámetros C_D y s , una vez que se considere se tiene suficiente precisión en estos parámetros, deberían mantenerse fijos para que la variación que producen en la función objetivo no dificulte la estimación de otros parámetros.
- El error de estimación de C_D tiende a crecer conforme el valor del parámetro es mayor.
- Los errores de estimación de ω 's están relacionados fuertemente con el valor de κ y su error de estimación, por lo que es necesaria una buena estimación de este parámetro para determinar adecuadamente ω 's. Además, si el valor encontrado de κ es cercano a 1, entonces se puede esperar un mayor error de estimación en ω_v que en ω_f , y lo contrario si el valor de κ es cercano a 0.
- Existe correlación entre los errores de estimación de λ_{mf} y λ_{mv} , pero no con λ_{vf} , por lo que se puede esperar mayor precisión en λ_{vf} .
- Si el valor estimado de λ es pequeño, se puede esperar un mayor error de estimación que si es grande.

4.5. Comparativo de resultados

En los experimentos realizados en este capítulo se fijo un tiempo de optimización que es igual en todos los experimentos. Sin embargo, en los resultados que se presentan en adelante se aumenta este tiempo de optimización para comparar la precisión a la que logra llegar cada método.

A continuación se estudia el efecto que tiene la incorporación de todas las características propuestas hasta el momento (a excepción de la transformación de los parámetros, que no mostró mejora), comparando contra lo que consideraríamos una primera aproximación al problema, o dicho de otra forma, una situación de desconocimiento del comportamiento del problema, donde elegimos las

opciones más intuitivas o comunes, es decir: utilizando un espaciado lineal en los tiempos, una función objetivo de mínimos cuadrados, y $\bar{p}_{wD}(\tau_D)$ en lugar de la **presión de Laplace** ($\tau_D \bar{p}_{wD}(\tau_D)$) dentro de la función objetivo, y utilizando como optimizador un Algoritmo Genético.

Diferenciaremos ambas situaciones con el nombre de optimización con y sin información de sensibilidad.

Para alcanzar el punto de convergencia en ambos experimentos, se utiliza 5 veces el tiempo de optimización original, para 5,000 casos sintéticos, con parámetros aleatorios.

El resultado de la media y mediana de estos experimentos se muestra en el Cuadro 4.15, donde podemos apreciar que en la media de los parámetros de yacimiento tenemos reducciones del error que van desde 48% (λ_{mf} y λ_{mv}) hasta 63% (ω 's) y 56% en el error conjunto (μ).

En la mediana se presentan reducciones del 70% (λ_{mf} y λ_{mv}) al 83% (ω 's y κ), y 62% para el error conjunto (μ), es decir, en la mediana observamos que en todos los parámetros hay menos de la tercera parte del error original.

	ω_f	ω_v	λ_{mf}	λ_{mv}	λ_{vf}	κ	s	C_D	μ
	Media del error								
Sin sensibilidad	0.2530	0.2650	0.3249	0.3213	0.3248	0.2679	$3.23e-04$	$2.65e-02$	0.2928
Con sensibilidad	0.0974	0.0987	0.1695	0.1678	0.1266	0.1107	$3.07e-06$	$1.42e-07$	0.1284
	Mediana del error								
Sin sensibilidad	0.1991	0.2096	0.2887	0.2826	0.2862	0.2152	$4.32e-05$	$1.80e-02$	0.2864
Con sensibilidad	0.0364	0.0362	0.0862	0.0859	0.0664	0.0358	$7.19e-08$	$1.17e-08$	0.1075

Cuadro 4.15: Media y mediana del error de estimación para optimización sin y con información de sensibilidad

En los valores de la media del error de estimación observamos que cuando no se utiliza la información derivada de los análisis de sensibilidad, los únicos parámetros que son determinados con precisión, son los más sensibles, C_D y s . Sin embargo, debido al gran efecto que tienen en la presión estos parámetros, la precisión obtenida no es suficiente para que permita identificar adecuadamente el resto de los parámetros. Esta situación es más evidente con los λ 's, donde el valor de 32% de error de estimación, prácticamente coincide con el valor de una estimación aleatoria, un valor de 1/3, por lo que puede decirse que en este caso no existe estimación real para estos parámetros cuando no se utiliza la información de sensibilidad¹².

En la Figura 4.16 se presentan los diagramas de cajas que comparan el error de estimación de los parámetros de yacimiento, sin y con información de sensibilidad durante la optimización.

En el Cuadro 4.16 se muestra las pruebas de hipótesis de igualdad de medias, y como era de esperarse, todos los parámetros tienen diferencia estadísticamente significativa a más del 99% de confianza, por lo que consideramos que la mejora que generan las características agregadas al proceso de identificación de los parámetros es evidente, y no resultado de la aleatoriedad.

En el Cuadro 4.17 se presentan los intervalos de confianza para la media del error de estimación, para cada uno de los parámetros, con un 95% de confianza; contrastando el uso o no de la información de sensibilidad en el proceso de optimización.

¹²El determinar con suficiente precisión los parámetros más sensibles, para así lograr posteriormente determinar aquellos menos sensibles, es uno de los objetivos que se alcanza con la optimización de tipo secuencial.

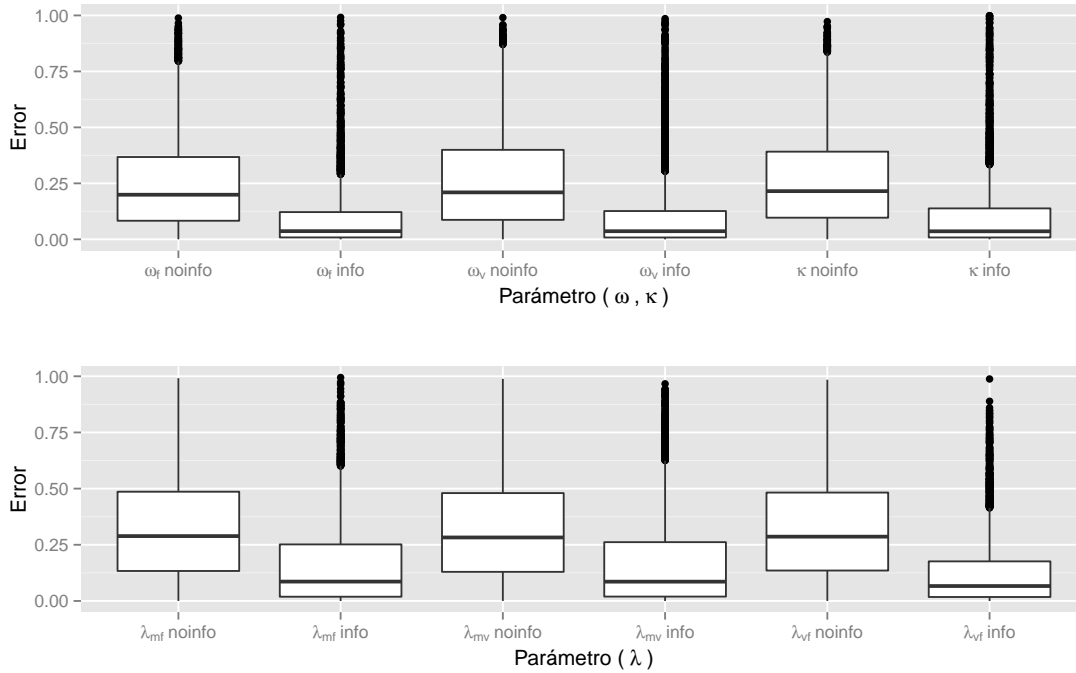


Figura 4.16: Boxplot - Error de estimación de los parámetros de yacimiento, sin y con información de sensibilidad

	<i>p-value</i>	Intervalo de confianza (95 %)	
		Límite Inferior	Límite Superior
ω_f	$< 2.2e-16$	0.1485	0.1628
ω_v	$< 2.2e-16$	0.1590	0.1735
λ_{mf}	$< 2.2e-16$	1.2982	1.4221
λ_{mv}	$< 2.2e-16$	1.2707	1.3926
λ_{vf}	$< 2.2e-16$	1.4984	1.6211
κ	$< 2.2e-16$	0.1497	0.1648
s	$< 2.2e-16$	$2.97e-04$	$3.43e-04$
C_D	$< 2.2e-16$	$2.65e-02$	$1.42e-07$
μ	$< 2.2e-16$	0.1607	0.1680

Cuadro 4.16: Pruebas de hipótesis de igualdad de medias para optimización sin y con información de sensibilidad

	Intervalo de confianza (95 %)		
	Media	Límite Inferior	Límite Superior
	Sin información de sensibilidad		
ω_f	0.2530	0.2472	0.2588
ω_v	0.2650	0.2590	0.2709
λ_{mf}	0.3249	0.3186	0.3311
λ_{mv}	0.3213	0.3151	0.3276
λ_{vf}	0.3248	0.3185	0.3310
κ	0.2679	0.2621	0.2738
s	$3.23e-04$	$3.00e-04$	$3.46e-04$
C_D	$2.65e-02$	$2.57e-02$	$2.73e-02$
μ	0.2928	0.2903	0.2953
	Con información de sensibilidad		
ω_f	0.0974	0.0933	0.1014
ω_v	0.0987	0.0946	0.1028
λ_{mf}	0.1695	0.1640	0.1751
λ_{mv}	0.1678	0.1624	0.1732
λ_{vf}	0.1266	0.1223	0.1308
κ	0.1107	0.1060	0.1153
s	$3.07e-06$	$2.28e-06$	$3.87e-06$
C_D	$1.42e-07$	$1.17e-07$	$1.66e-07$
μ	0.1284	0.1258	0.1311

Cuadro 4.17: Intervalos de confianza para la media del error de estimación, sin y con información de sensibilidad

Se observa, por ejemplo, que el verdadero valor de la media del error de estimación para el error conjunto (μ) se encuentra con un 95 % de confiabilidad entre el 29.03 % y el 29.53 % para el caso donde no se utiliza la información de sensibilidad, y entre el 12.58 % y el 13.11 % cuando sí se utiliza esta información.

En las Figuras 4.17 - 4.19 se muestra el comparativo de la dispersión del valor a estimar *vs* el valor estimado, para los caso con y sin información de sensibilidad durante el proceso de optimización, para los parámetros ω_g , λ_{mf} y κ . Se observa que, como se ha mencionado anteriormente, la estimación de λ es aleatoria, ya que la gráfica no muestra ningún patrón en las estimaciones (recordando que una línea diagonal es el caso ideal).

En estas gráficas también se observa que cuando se utiliza la información de sensibilidad, los rangos de los parámetros donde hay menor precisión, son los mismos rangos que se predecían en los análisis de sensibilidad: valores centrales para ω 's y κ , y valores pequeños para λ ; situación derivada del propio efecto que tienen los parámetros sobre la presión.

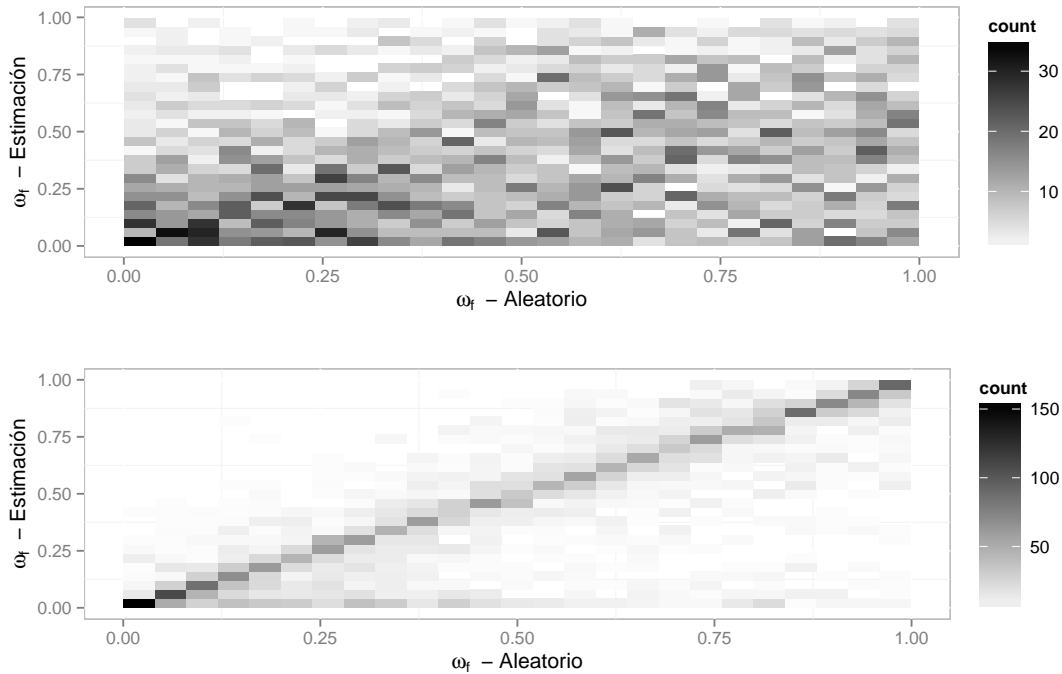


Figura 4.17: Comparativo de la dispersión del valor a estimar de ω_f *vs* estimación, sin y con información de sensibilidad en la optimización

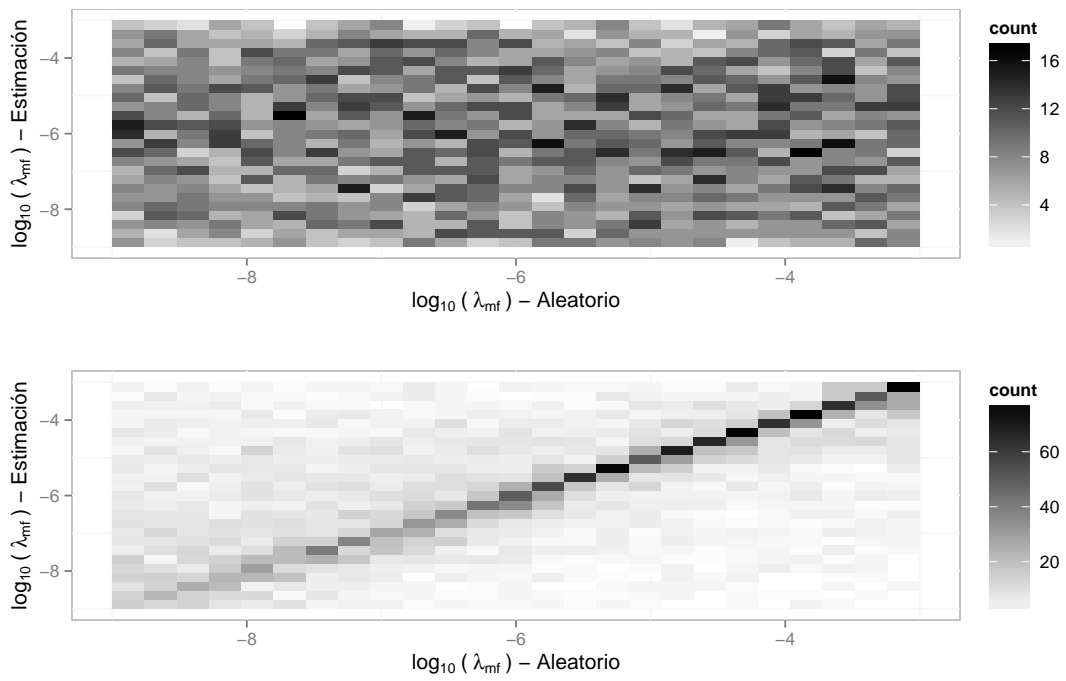


Figura 4.18: Comparativo de la dispersión del valor a estimar de λ_{mf} vs estimación, sin y con información de sensibilidad en la optimización

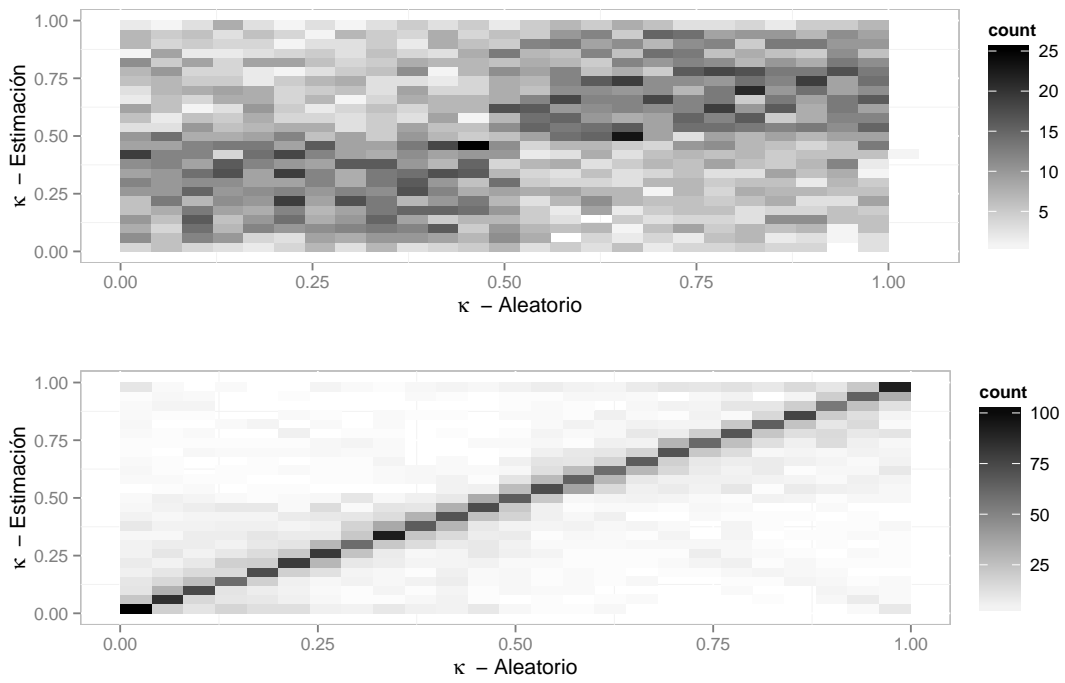


Figura 4.19: Comparativo de la dispersión del valor a estimar de κ vs estimación, sin y con información de sensibilidad en la optimización

En cuanto a la probabilidad de cometer errores de estimación grandes, tenemos que:

$$P(\mu_{noinfo} \geq \frac{1}{3}) \approx \frac{1}{5000} \sum_{i=1}^{5000} I(\mu_{noinfo} \geq \frac{1}{3}) = 0.3150$$

$$P(\mu_{info} \geq \frac{1}{3}) \approx \frac{1}{5000} \sum_{i=1}^{5000} I(\mu_{info} \geq \frac{1}{3}) = 0.0342$$

por lo que la inclusión de la información de sensibilidad en el proceso de optimización nos permite reducir 9 veces la probabilidad de cometer errores grandes en la estimación, obteniendo 3.5% de probabilidad de tener errores mayores a 1/3.

De los resultados anteriores podemos ver que ha sido de gran utilidad analizar el comportamiento de la solución del modelo y crear métodos *ad-hoc* que tomen en cuenta esta información durante el proceso de identificación de los parámetros, llegando a errores de menos de la tercera parte que cuando no se realiza este proceso, siendo vital para problemas inversos que tienen una gran complejidad.

También puede apreciarse que esta metodología para abordar un problema inverso no depende del método de optimización utilizado, ya que aunque se use un método más avanzado (que por ejemplo incluya el gradiente de la función objetivo para elegir la dirección de búsqueda de la solución), las características aquí estudiadas también pueden ser incorporadas en estos otros métodos, mejorando sus resultados.

Capítulo 5

Múltiples soluciones óptimas

Teniendo en cuenta que el problema inverso aquí abordado involucra un problema de optimización, puede llegarse a soluciones óptimas (conjuntos de parámetros), muy diferentes del conjunto de parámetros que generó los datos que se busca reproducir (en nuestro caso, datos de presión).

Algunas de estas soluciones óptimas son debidas a que en la solución analítica existe más de un conjunto de parámetros que genera exactamente los mismos valores de presión, lo que se traduce en una función objetivo con múltiples mínimos globales. Otras soluciones o más correctamente pseudo-soluciones óptimas, son causadas por la precisión numérica involucrada en el proceso de optimización, donde un conjunto de parámetros que genera valores suficientemente similares a los datos, es considerado erróneamente como solución.

En cuanto a la primer situación, es necesario conocer bajo qué condiciones es que se presentan y cuáles son estas otras soluciones, para que sean consideradas. En relación a las soluciones debidas a la precisión numérica, en este capítulo se propone una metodología (aplicable a cualquier método de optimización), que pretende disminuir su probabilidad de ocurrencia.

5.1. Solución dual

En el presente trabajo se ha utilizado el modelo TPDP de la sección 1.1, en el cual hemos determinado que existe una solución que genera los mismos valores de presión al intercambiar los parámetros que corresponden a fracturas y vórgulos¹, por lo que a esta solución la llamamos **dual**. Es decir, ambos conjuntos de parámetros producen los mismos datos.

Proposición 1. *Si denotamos:*

$$\begin{aligned}\mathbf{x} &= (\omega_v, \omega_f, \lambda_{mf}, \lambda_{mv}, \lambda_{vf}, \kappa, s, C_D) \in \mathbb{R}^8, \\ \mathbf{x}^* &= (\omega_v^*, \omega_f^*, \lambda_{mf}^*, \lambda_{mv}^*, \lambda_{vf}^*, \kappa^*, s, C_D) \in \mathbb{R}^8\end{aligned}$$

¹Conceptualmente, en el modelo se puede intercambiar las fracturas por los vórgulos (y viceversa), sin producir diferencia en su planteamiento.

donde

$$\begin{aligned}\omega_f^* &= \omega_v, \\ \omega_v^* &= \omega_f, \\ \lambda_{mf}^* &= \lambda_{mv}, \\ \lambda_{mv}^* &= \lambda_{mf}, \\ \kappa^* &= 1 - \kappa\end{aligned}$$

entonces:

$$p_{wD}(\mathbf{x}^*, t_D) = p_{wD}(\mathbf{x}, t_D)$$

Demostración. Utilizando la solución del modelo expuesta en el Capítulo 2. Definimos:

$$\Lambda_f(\mathbf{x}, \tau_D) = \lambda_{mf} + \lambda_{vf} + \omega_f \tau_D \quad (5.1a)$$

$$\Lambda_m(\mathbf{x}, \tau_D) = (1 - \omega_f - \omega_v) \tau_D + \lambda_{mf} + \lambda_{mv} \quad (5.1b)$$

$$\Lambda_v(\mathbf{x}, \tau_D) = \lambda_{mv} + \lambda_{vf} + \omega_v \tau_D \quad (5.1c)$$

Por lo que en el dual se tiene:

$$\Lambda_f(\mathbf{x}^*, \tau_D) = \lambda_{mf}^* + \lambda_{vf} + \omega_f^* \tau_D = \lambda_{mv} + \lambda_{vf} + \omega_v \tau_D = \Lambda_v(\mathbf{x}, \tau_D) \quad (5.2a)$$

$$\begin{aligned}\Lambda_m(\mathbf{x}^*, \tau_D) &= (1 - \omega_f^* - \omega_v^*) \tau_D + \lambda_{mf}^* + \lambda_{mv}^* \\ &= (1 - \omega_v - \omega_f) \tau_D + \lambda_{mv} + \lambda_{mf} = \Lambda_m(\mathbf{x}, \tau_D)\end{aligned} \quad (5.2b)$$

$$\Lambda_v(\mathbf{x}^*, \tau_D) = \lambda_{mv}^* + \lambda_{vf} + \omega_v^* \tau_D = \lambda_{mf} + \lambda_{vf} + \omega_f \tau_D = \Lambda_f(\mathbf{x}, \tau_D) \quad (5.2c)$$

Utilizando $\Lambda_f(\mathbf{x}, \tau_D)$, $\Lambda_m(\mathbf{x}, \tau_D)$ y $\Lambda_v(\mathbf{x}, \tau_D)$ definimos:

$$b_{11}(\mathbf{x}, \tau_D) = \frac{\lambda_{mf}^2}{\Lambda_m(\mathbf{x}, \tau_D)} - \Lambda_f(\mathbf{x}, \tau_D), \quad (5.3a)$$

$$b_{12}(\mathbf{x}, \tau_D) = \frac{\lambda_{mf} \lambda_{mv}}{\Lambda_m(\mathbf{x}, \tau_D)} + \lambda_{vf}, \quad (5.3b)$$

$$b_{22}(\mathbf{x}, \tau_D) = \frac{\lambda_{mv}^2}{\Lambda_m(\mathbf{x}, \tau_D)} - \Lambda_v(\mathbf{x}, \tau_D) \quad (5.3c)$$

Equivalentemente:

$$\begin{aligned} b_{11}(\mathbf{x}^*, \tau_D) &= \frac{\lambda_{mf}^{*2}}{\Lambda_m(\mathbf{x}^*, \tau_D)} - \Lambda_f(\mathbf{x}^*, \tau_D) \\ &= \frac{\lambda_{mv}^2}{\Lambda_m(\mathbf{x}, \tau_D)} - \Lambda_v(\mathbf{x}, \tau_D) = b_{22}(\mathbf{x}, \tau_D) \end{aligned} \quad (5.4a)$$

$$\begin{aligned} b_{12}(\mathbf{x}^*, \tau_D) &= \frac{\lambda_{mf}^* \lambda_{mv}^*}{\Lambda_m(\mathbf{x}^*, \tau_D)} + \lambda_{vf} \\ &= \frac{\lambda_{mv} \lambda_{mf}}{\Lambda_m(\mathbf{x}, \tau_D)} + \lambda_{vf} = b_{12}(\mathbf{x}, \tau_D) \end{aligned} \quad (5.4b)$$

$$\begin{aligned} b_{22}(\mathbf{x}^*, \tau_D) &= \frac{\lambda_{mv}^{*2}}{\Lambda_m(\mathbf{x}^*, \tau_D)} - \Lambda_v(\mathbf{x}^*, \tau_D) \\ &= \frac{\lambda_{mf}^2}{\Lambda_m(\mathbf{x}, \tau_D)} - \Lambda_f(\mathbf{x}, \tau_D) = b_{11}(\mathbf{x}, \tau_D) \end{aligned} \quad (5.4c)$$

Para simplificar la notación, en adelante omitimos la dependencia (\mathbf{x}, τ_D) , utilizando \star , para referirnos al dual. Entonces, definimos:

$$h_0 = (\kappa b_{22} - (1 - \kappa) b_{11})^2 + 4\kappa(1 - \kappa) b_{12}^2 \quad (5.5)$$

Siendo el dual:

$$\begin{aligned} h_0^* &= (\kappa^* b_{22}^* - (1 - \kappa^*) b_{11}^*)^2 + 4\kappa^*(1 - \kappa^*) b_{12}^{*2} \\ &= ((1 - \kappa) b_{11} - \kappa b_{22})^2 + 4(1 - \kappa)\kappa b_{12}^2 = h_0 \end{aligned} \quad (5.6)$$

Con b_{11} , b_{12} , b_{22} y h_0 definimos

$$g_1 = -\frac{(1 - \kappa) b_{11} + \kappa b_{22} + \sqrt{h_0}}{2\kappa(1 - \kappa)}, \quad (5.7a)$$

$$g_2 = -\frac{(1 - \kappa) b_{11} + \kappa b_{22} - \sqrt{h_0}}{2\kappa(1 - \kappa)} \quad (5.7b)$$

$$g_3 = \frac{\kappa b_{22} - (1 - \kappa) b_{11} + \sqrt{h_0}}{2\kappa(1 - \kappa) b_{12}}, \quad (5.8a)$$

$$g_4 = \frac{\kappa b_{22} - (1 - \kappa) b_{11} - \sqrt{h_0}}{2\kappa(1 - \kappa) b_{12}} \quad (5.8b)$$

donde observamos que los duales equivalen a:

$$g_1^* = g_1, \quad g_2^* = g_2, \quad g_3^* = -g_4, \quad g_4^* = -g_3 \quad (5.9a)$$

Adicionalmente definimos:

$$K_1 = Y_0(\sqrt{g_1}) + s\sqrt{g_1} Y_1(\sqrt{g_1}) = K_1^* \quad (5.10a)$$

$$K_2 = Y_0(\sqrt{g_2}) + s\sqrt{g_2} Y_1(\sqrt{g_2}) = K_2^* \quad (5.10b)$$

$$K_{11} = (\kappa + (1 - \kappa) \kappa g_3) \sqrt{g_1} Y_{11} \quad (5.11a)$$

$$K_{12} = (\kappa + (1 - \kappa) \kappa g_4) \sqrt{g_2} Y_{12} \quad (5.11b)$$

$$K_{21} = (1 - \kappa g_3) Y_{01} + (s - s \kappa g_3) \sqrt{g_1} Y_{11} \quad (5.11c)$$

$$K_{22} = (1 - \kappa g_4) Y_{02} + (s - s \kappa g_4) \sqrt{g_2} Y_{12} \quad (5.11d)$$

Con los duales:

$$K_{11}^* = ((1 - \kappa) - (1 - \kappa) \kappa g_4) \sqrt{g_1} Y_{11} \quad (5.12a)$$

$$K_{12}^* = ((1 - \kappa) - (1 - \kappa) \kappa g_3) \sqrt{g_2} Y_{12} \quad (5.12b)$$

$$K_{21}^* = (1 + (1 - \kappa) g_4) Y_{01} + (s + s(1 - \kappa) g_4) \sqrt{g_1} Y_{11} \quad (5.12c)$$

$$K_{22}^* = (1 + (1 - \kappa) g_3) Y_{02} + (s + s(1 - \kappa) g_3) \sqrt{g_2} Y_{12} \quad (5.12d)$$

Para obtener el numerador y denominador de la solución, calculamos:

$$K_{22}K_1 - K_{21}K_2 = \kappa(g_3 - g_4)(s\sqrt{g_1}Y_{11} + Y_{01})(s\sqrt{g_2}Y_{12} + Y_{02}), \quad (5.13a)$$

$$K_{11}K_{22} = \kappa(\kappa g_3 - g_3 - 1)(\kappa g_4 - 1)(\sqrt{g_1} Y_{11}(Y_{02} + s\sqrt{g_2}Y_{12})), \quad (5.13b)$$

$$K_{12}K_{21} = \kappa(\kappa g_4 - g_4 - 1)(\kappa g_3 - 1)(\sqrt{g_2} Y_{12}(Y_{01} + s\sqrt{g_1}Y_{11})) \quad (5.13c)$$

$$K_{22}^*K_1^* - K_{21}^*K_2^* = (1 - \kappa)(g_3 - g_4)(s\sqrt{g_1}Y_{11} + Y_{01})(s\sqrt{g_2}Y_{12} + Y_{02}), \quad (5.14a)$$

$$K_{11}^*K_{22}^* = (1 - \kappa)(\kappa g_3 - g_3 - 1)(\kappa g_4 - 1)(\sqrt{g_1} Y_{11}(Y_{02} + s\sqrt{g_2}Y_{12})), \quad (5.14b)$$

$$K_{12}^*K_{21}^* = (1 - \kappa)(\kappa g_4 - g_4 - 1)(\kappa g_3 - 1)(\sqrt{g_2} Y_{12}(Y_{01} + s\sqrt{g_1}Y_{11})) \quad (5.14c)$$

por lo que:

$$K_{22}K_1 - K_{21}K_2 = \frac{\kappa}{1 - \kappa}(K_{22}^*K_1^* - K_{21}^*K_2^*), \quad (5.15a)$$

$$K_{11}K_{22} = \frac{\kappa}{1 - \kappa}(K_{11}^*K_{22}^*), \quad (5.15b)$$

$$K_{12}K_{21} = \frac{\kappa}{1 - \kappa}(K_{12}^*K_{21}^*) \quad (5.15c)$$

Es así que la solución del modelo es:

$$\bar{p}_{wD}(\mathbf{x}, \tau_D)|_{C_D=0} = \frac{K_{22}(\mathbf{x}, \tau_D)K_1(\mathbf{x}, \tau_D) - K_{21}(\mathbf{x}, \tau_D)K_2(\mathbf{x}, \tau_D)}{\tau_D[K_{11}(\mathbf{x}, \tau_D)K_{22}(\mathbf{x}, \tau_D) - K_{12}(\mathbf{x}, \tau_D)K_{21}(\mathbf{x}, \tau_D)]} \quad (5.16)$$

$$= \frac{\kappa(1 - \kappa)}{\tau_D \kappa(1 - \kappa)} \frac{K_{22}(\mathbf{x}, \tau_D)K_1(\mathbf{x}, \tau_D) - K_{21}(\mathbf{x}, \tau_D)K_2(\mathbf{x}, \tau_D)}{K_{11}(\mathbf{x}, \tau_D)K_{22}(\mathbf{x}, \tau_D) - K_{12}(\mathbf{x}, \tau_D)K_{21}(\mathbf{x}, \tau_D)} \quad (5.17)$$

$$= \bar{p}_{wD}(\mathbf{x}^*, \tau_D)|_{C_D=0} \quad (5.18)$$

Por lo tanto:

$$p_{wD}(\mathbf{x}^*, t_D) = p_{wD}(\mathbf{x}, t_D)$$

donde Y_{ij} es la función de Bessel de segundo tipo y orden i , para el valor $\sqrt{g_j}$. \square

Por lo que se demuestra que para todo conjunto de parámetros \mathbf{x} siempre existe otra solución \mathbf{x}^* , que genera los mismos valores de presión en el modelo de TPDP, donde a esta otra solución, la llamamos dual.

Desde un principio asumimos $s = s_f = s_v$, debido a que se ha demostrado en la sección 2.1 que dada la teoría de radio efectivo, este supuesto no influye en la estimación de los parámetros de yacimiento, que son los parámetros de nuestro interés.

Como resultado de la existencia de la solución dual, se puede concluir que durante el proceso de identificación de parámetros del modelo de TPDP (el problema inverso), se puede restringir $0 \leq \kappa \leq 0.5$, ya que siempre existirá otra solución idéntica en $0.5 \leq \kappa \leq 1$, la cual puede ser descrita a partir de la primer solución obtenida, por lo que no es necesario explorar todo el espacio de solución del parámetro κ .

En cuanto a los análisis realizados en el Capítulo 4 se tomó en cuenta la solución dual para medir el error de estimación, es decir, si se observó que la optimización había llegado a la solución dual, se determinaba la solución original para así medir el error. Aunque este problema se puede eliminar al limitar el rango de κ , se utilizó el rango completo del parámetro para asegurar que se estudia adecuadamente el efecto de las interacciones entre los parámetros.

5.2. Pseudo-mínimos

En el problema inverso se presenta un proceso de minimización de la función objetivo, en el cual un conjunto de parámetros \mathbf{x}^* es solución si:

$$f_{obj}(\mathbf{x}^*) \leq f_{obj}(\mathbf{x}), \forall \mathbf{x}$$

donde f_{obj} es una función objetivo que mide la discrepancia entre los datos del modelo y los datos medidos, mediante una norma, por lo que $f_{obj}(\mathbf{x}) \geq 0$ para todo \mathbf{x} .

En una situación ideal \mathbf{x}^* es solución si $f_{obj} = 0$. Sin embargo, en la práctica, debido a la precisión numérica de la máquina, al ruido en los datos, a que se resuelve el problema de forma aproximada, e incluso a la forma de la función objetivo, es necesario definir un valor f_{\min} , tal que se considera solución a \mathbf{x}^* si:

$$f_{obj}(\mathbf{x}^*) \leq f_{\min}$$

o alternativamente se puede establecer una condición de paro en la que después de un número de iteraciones n , no se presenta una reducción significativa en el valor de la función objetivo.

Ambos métodos de decisión involucran establecer una condición donde interviene una tolerancia ϵ para *inferir* que se ha llegado a una aproximación a la solución óptima y así considerar a un conjunto de parámetros como solución del problema inverso. No obstante, bajo estas condiciones es posible considerar erróneamente como solución a un conjunto de parámetros que reproduzca con suficiente similitud los datos.

Es así que llamaremos pseudo-mínimos a los falsos mínimos que se presentan a causa de la precisión numérica antes mencionada.

Para ejemplificar estos pseudo-mínimos simulamos una situación donde conocemos el valor de los parámetros de pozo y de yacimiento, a excepción de κ , el cual es el único parámetro a estimar.

Para medir la discrepancia entre la solución del modelo y los datos medidos, considerando a \mathbf{p}_d

para los datos de presión medidos y \mathbf{p}_m para la presión del modelo, utilizamos la función objetivo:

$$f(\mathbf{p}_d, \mathbf{p}_m) = \frac{1}{n} \sum_{i=1}^n \left| \frac{\mathbf{p}_m - \mathbf{p}_d}{\mathbf{p}_d} \right|$$

aunque puede usarse cualquier otra de las funciones objetivo definidas anteriormente en (4.2)-(4.6), obteniendo resultados similares.

Si solo pretendemos estimar κ , entonces podemos considerar que κ^* es una buena estimación del valor real de κ si $f(\mathbf{p}_d, \mathbf{p}_m(\kappa^*)) < \epsilon$.

Utilizamos los siguientes valores en los parámetros para generar los datos sintéticos:

$$\begin{aligned} \omega_f &= 0.1028, \quad \omega_v = 0.1232, \quad \lambda_{mf} = 1.73e-06, \quad \lambda_{mv} = 7.32e-08, \quad \lambda_{mf} = 1.16e-04, \\ \kappa &= 0.6538, \quad s = 0, \quad C_D = 0 \end{aligned}$$

Si consideramos desconocido a κ , al determinar la función objetivo con diferentes valores de κ , obtenemos la Figura 5.1, donde parecen existir dos mínimos, uno en el valor $\kappa = 0.65$ (el verdadero valor), y otro en $\kappa = 0.29$. Sin embargo este otro mínimo no pertenece a un mínimo analítico, sino es resultado de aspectos numéricos, es decir, un pseudo-mínimo.

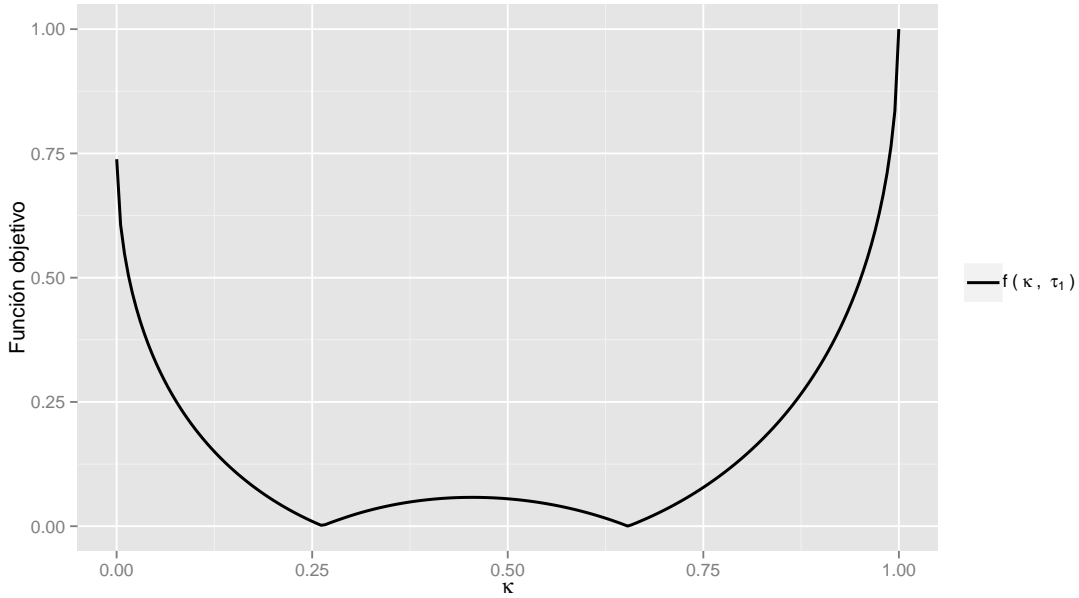


Figura 5.1: Función objetivo con diferentes valores de κ

Para poder abordar un método que nos permita evitar los pseudo-mínimos, es necesario comprender la causa por la que se producen.

Si el valor de la función objetivo es muy cercano a cero, implica que en los tiempo τ_{Di} , en los cuales se está comparando los valores de presión $\bar{p}_{wD}(\tau_{Di})$ vs la presión medida (o generada en

este caso), existe una gran similitud del valor de presión en ambos valores de tiempo utilizados. Sin embargo, si se trata de un pseudo-mínimo, entonces deben existir valores de τ_{D_i} para los cuales los valores de presión difirieren en mayor medida, con respecto a los datos.

Aunque sería ideal poder considerar principalmente los valores de τ_{D_i} donde los valores de presión difieren mayormente, no nos es posible conocer de antemano esos valores. Sin embargo sabemos por los análisis de sensibilidad que existen tiempos en los cuales ciertos parámetros tienen mayor o menor efecto.

Es así que si contrastamos funciones objetivo con rangos de tiempo que están relacionados al efecto de distintos parámetros, es posible determinar diferencias en las funciones objetivo que nos permitan evitar los pseudo-mínimos.

Por ejemplo, en la gráfica de la Figura 5.1 se utilizó en la función objetivo 10 tiempos espaciados logarítmicamente en el intervalo:

$$\tau_1 = [1e-8, 1e+2]$$

Si adicionalmente graficamos la función objetivo con 10 tiempos espaciados logarítmicamente en el intervalo:

$$\tau_2 = [1e-6, 1e-4]$$

entonces obtenemos la gráfica de la Figura 5.2, donde se puede apreciar que la nueva función objetivo logra diferenciar el pseudo-mínimo del mínimo analítico.

Nótese que $\tau_2 \subset \tau_1$, por lo que el nuevo intervalo de tiempo τ_2 (seleccionado con la información de los análisis de sensibilidad) no considera tiempos fuera del intervalo original, pero en el rango τ_2 todos los parámetros tienen efecto en la presión, y por lo tanto no incluye tiempos donde solo algunos de los parámetros son los de principal influencia.

Para poder comparar adecuadamente las funciones objetivo con los distintos tiempos se escalan las funciones objetivo, tal que $\max\{f\} = 1$.

Sin embargo, no siempre este nuevo rango tiene que diferenciar el mínimo real. Al realizar un nuevo experimento con el conjunto de parámetros:

$$\begin{aligned} \omega_f &= 0.5530, \omega_v = 0.3632, \lambda_{mf} = 6.38e-04, \lambda_{mv} = 8.28e-05, \lambda_{mf} = 9.86e-09, \\ \kappa &= 0.3695, s = 0, C_D = 0 \end{aligned}$$

y los mismos rangos τ_1 y τ_2 definidos anteriormente, observamos en la Figura 5.3 que tanto la función objetivo con τ_1 como la de τ_2 muestran dos mínimos cada una, y sin embargo solo coinciden en uno de ellos, por lo que necesariamente este debe ser el mínimo real, y los otros son pseudo-mínimos, ya que el mínimo real debe ser el mismo sin importar los valores de tiempo que se utilicen.

De esta forma podemos determinar que un mínimo real de la función objetivo será el mismo para cualquier intervalo de tiempo que se utilice, por lo que una posible restricción que se puede imponer para llegar al mínimo real es que, \mathbf{x}^* es mínimo si:

$$f(\mathbf{x}^*, \tau_1) \leq f(\mathbf{x}, \tau_1) \quad \text{y} \quad f(\mathbf{x}^*, \tau_2) \leq f(\mathbf{x}, \tau_2), \quad \forall \mathbf{x} \in R(\mathbf{x}) \quad (5.19)$$

donde $f(\mathbf{x}, \tau_i)$, $i = 1, 2$ es el valor de la función objetivo en \mathbf{x} para el rango de tiempo τ_i , tal que $\tau_1 \neq \tau_2$.

Esto implica que en un algoritmo de optimización, si se tiene un punto \mathbf{x}_1 , se considera a otro punto \mathbf{x}_2 como mejor estimador solo si la función objetivo es menor en ambos rangos de tiempo.

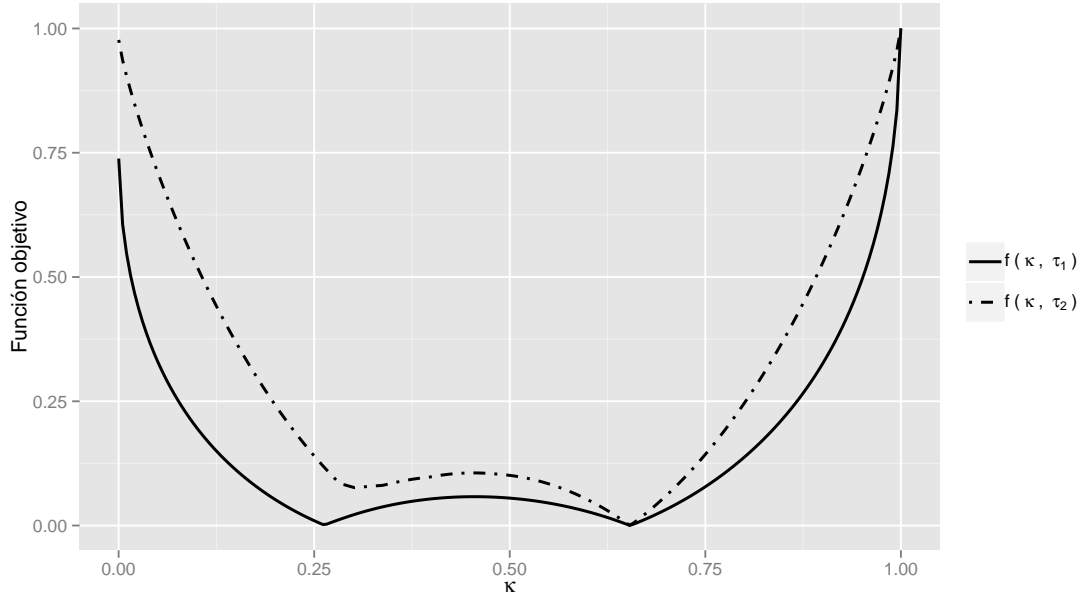


Figura 5.2: Función objetivo con diferentes valores de κ , para dos intervalos de tiempo

Según el análisis anterior queda claro que el problema no es tomar mayor densidad de puntos de tiempo (En el ejemplo anterior si se utilizan 500 valores de tiempos en lugar de 10, no se genera diferencia observable en la gráfica de la Figura 5.3), sino considerar tiempos en los que influyen de forma diferente, en distintos parámetros.

Considerando la restricción (5.19), ahora proponemos como ejemplo una situación donde se estiman κ y ω_f , utilizando como datos sintéticos los generados con los parámetros:

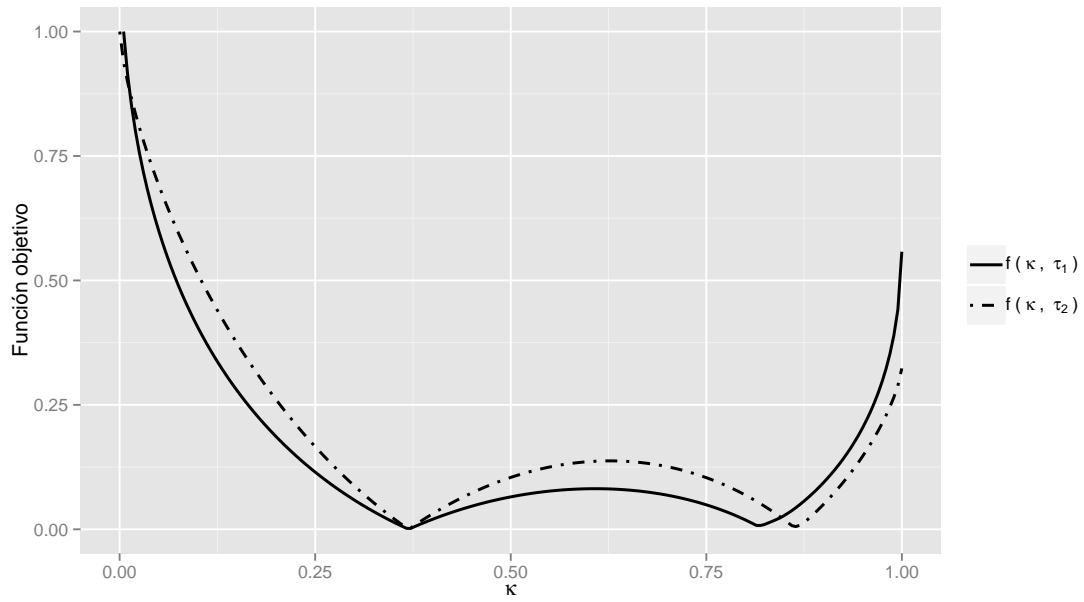
$$\begin{aligned} \omega_f &= 0.4981, \quad \omega_v = 0.0128, \quad \lambda_{mf} = 6.33e-04, \quad \lambda_{mv} = 7.82e-04, \quad \lambda_{mf} = 4.69e-04, \\ \kappa &= 0.6907, \quad s = 0, \quad C_D = 0 \end{aligned}$$

En la Figura 5.4 se muestra la función objetivo al variar κ y ω_f , utilizando el rango de tiempo τ_1 , y marcando una línea vertical en el punto $(\omega_f = 0.4981, \kappa = 0.6907)$, que corresponde al verdadero mínimo de la función objetivo.

Sin embargo se puede apreciar que existe un valle en la función objetivo en el cual se podría llegar a diferentes soluciones si no se tiene suficiente precisión en el valor del mínimo f_{\min} .

Al aplicar la restricción que hemos establecido en (5.19) obtenemos una zona factible de búsqueda, es decir, el área que cumple con la restricción, esta área se muestra en la Figura 5.5, donde se observa que se elimina el valle que podía llevar a soluciones diferentes de la verdadera.

Es así que un proceso de minimización que considera la restricción (5.19) nos llevaría a la solución real.

Figura 5.3: Función objetivo con diferentes valores de κ , para dos intervalos de tiempo

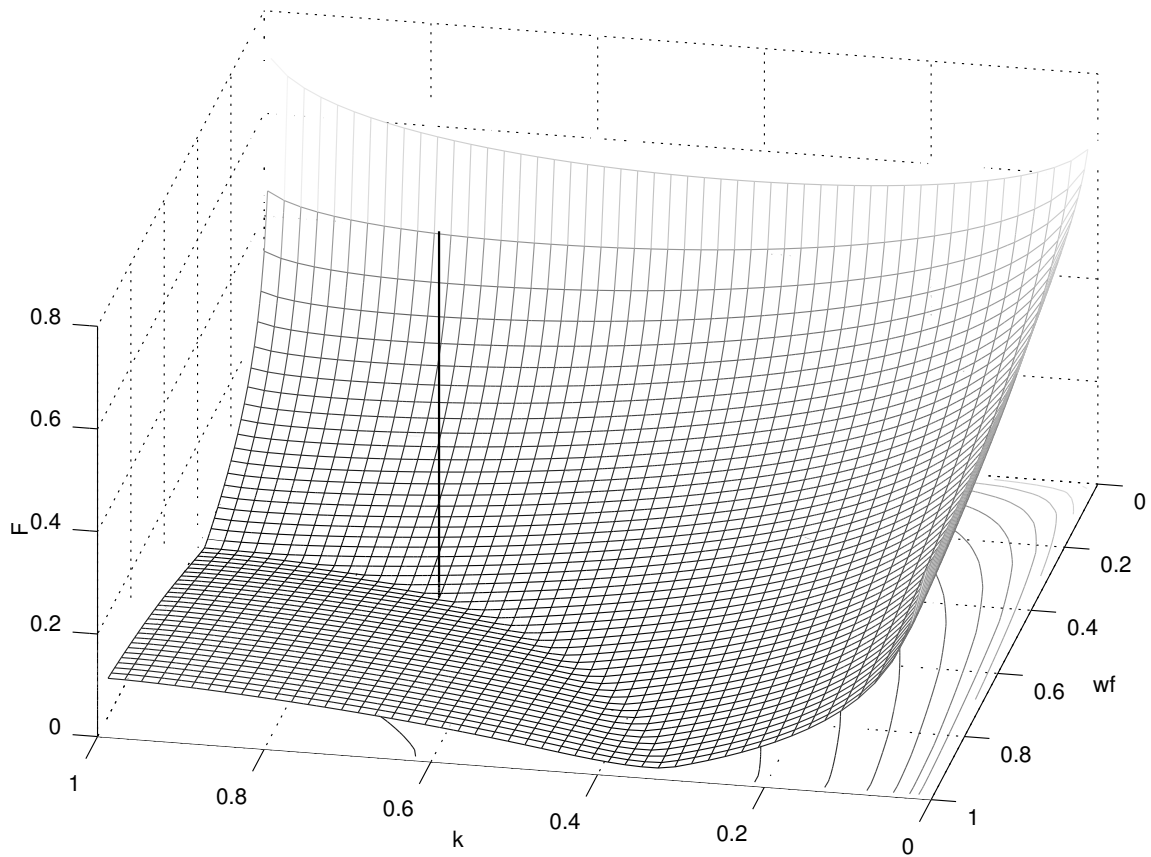


Figura 5.4: Función objetivo con diferentes valores de κ y ω_f

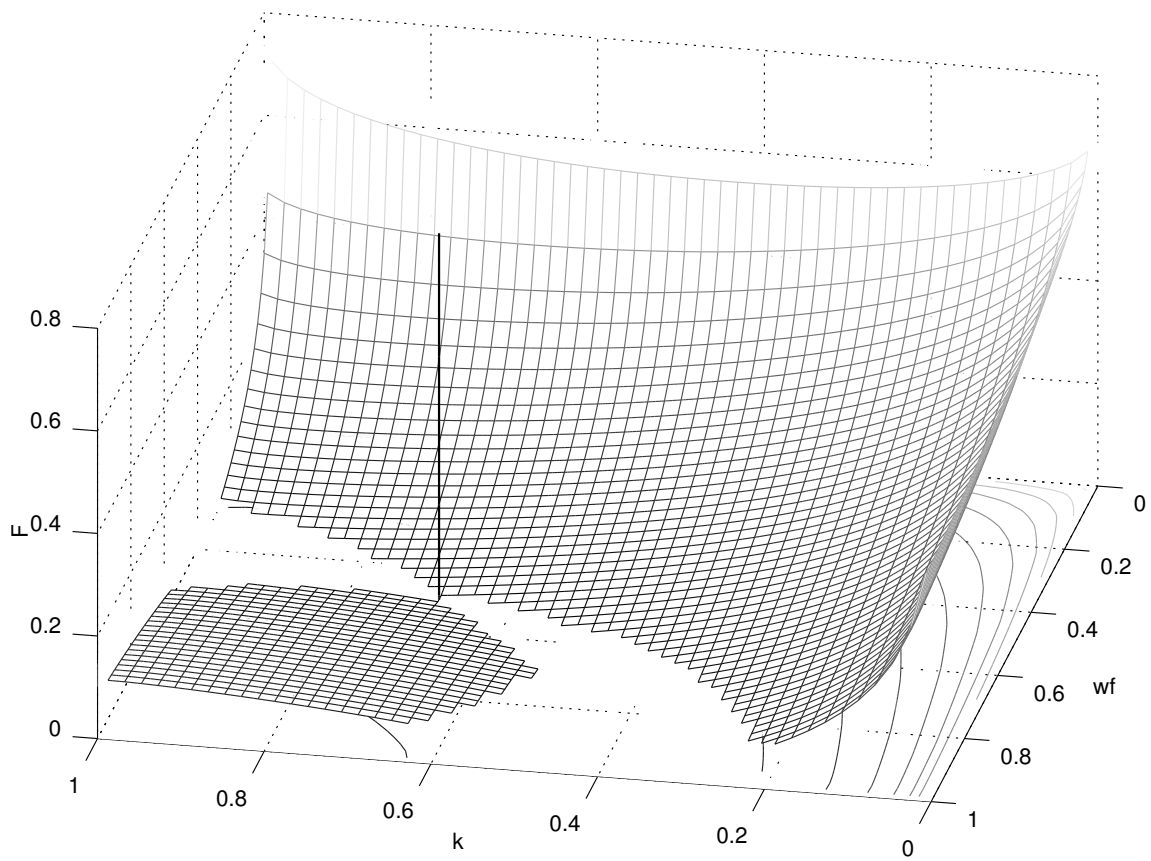


Figura 5.5: Función objetivo con diferentes valores de κ y ω_f aplicando restricción

Capítulo 6

Casos con datos reales

Con datos reales de pruebas de pozo reales, que provienen de yacimientos mexicanos, se realizan ajustes que toman en cuenta toda la metodología de análisis desarrollada a lo largo de este trabajo.

Se han seleccionado dos pozos, los cuales ejemplifican dos situaciones de interés cuando se comparan los ajustes que se obtienen con el modelo de Triple Porosidad [6] (ecuaciones (1.3)-(1.5)), con respecto a un modelo de Doble Porosidad, utilizado por los simuladores comerciales, y es conocido como modelo de Warren y Root [20], para el cual Da Prat et al. [8] incluyen el efecto de daño.

Los ajustes presentados tanto para doble como para triple porosidad han sido realizados con el mismo método de optimización y tomando en cuenta los análisis de sensibilidad en ambos casos.

6.1. Pozo 1

En la Figura 6.1 se muestran los datos de presión de una prueba de pozo y el ajuste obtenido con los modelos de doble porosidad (W&R) y triple porosidad. Se observa que en los primeros tiempos el ajuste de doble porosidad no es muy bueno, sin embargo el ajuste de triple porosidad es muy cercano a los datos a lo largo de todo el rango de tiempo.

En el Cuadro 6.1 se presentan los parámetros obtenidos con el ajuste de los dos diferentes modelos.

En este ejemplo los parámetros obtenidos con ambos modelos son similares, debido a que en el modelo de triple porosidad se observa un valor de $\kappa = 0.99$, que es similar al valor que asume el modelo de doble porosidad ($\kappa = 1$), y el valor de $\omega_f \approx 0$ indica que prácticamente no existen fracturas.

Es así que en este ejemplo se muestra que el modelo de triple porosidad contiene como casos especiales a modelos de doble o de una porosidad, sin embargo el modelo de triple porosidad proporciona mayor información, ya que identifica un mayor número de parámetros.

Además, en el valor de la función objetivo (f_{obj}) se observa que el modelo de triple porosidad ha logrado un mejor ajuste a los datos.

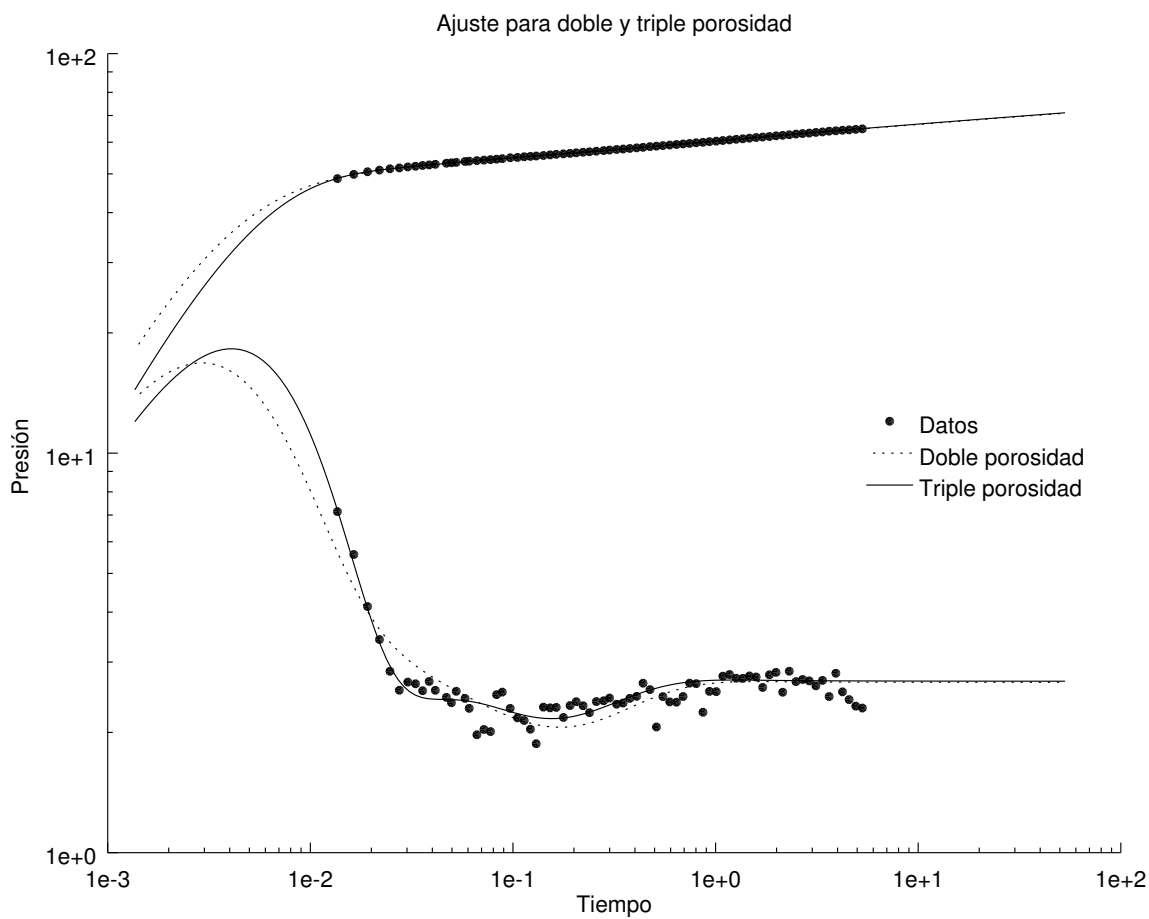


Figura 6.1: Ajuste de doble y de triple porosidad para el pozo 1

Doble porosidad	Triple porosidad (PT)
$\omega = 0.4872$	$\omega_v = 0.4925$
	$\omega_f = 0.0019$
$\lambda = 5.95e-07$	$\lambda_{mv} = 8.94e-07$
	$\lambda_{mf} = 1.34e-07$
	$\lambda_{vf} = 1.00e-05$
	$\kappa = 0.9913$
$s = 3.3009$	$s = 3.4447$
$C_D = 1207.2$	$C_D = 1137.0$
$f_{obj} = 3.44e-06$	$f_{obj} = 4.50e-07$

Cuadro 6.1: Parámetros estimados para el Pozo 1, con los modelos de doble y de triple porosidad

6.2. Pozo 2

En la Figura 6.2 se presentan los ajustes de los modelos de doble y de triple porosidad para las mediciones de la prueba de pozo, al que llamaremos Pozo 2.

Se aprecia que el ajuste del modelo de doble porosidad no logra reproducir el segundo valle que comienza a aparecer al final de los datos. Estos dos valles son característicos del efecto de la triple porosidad.

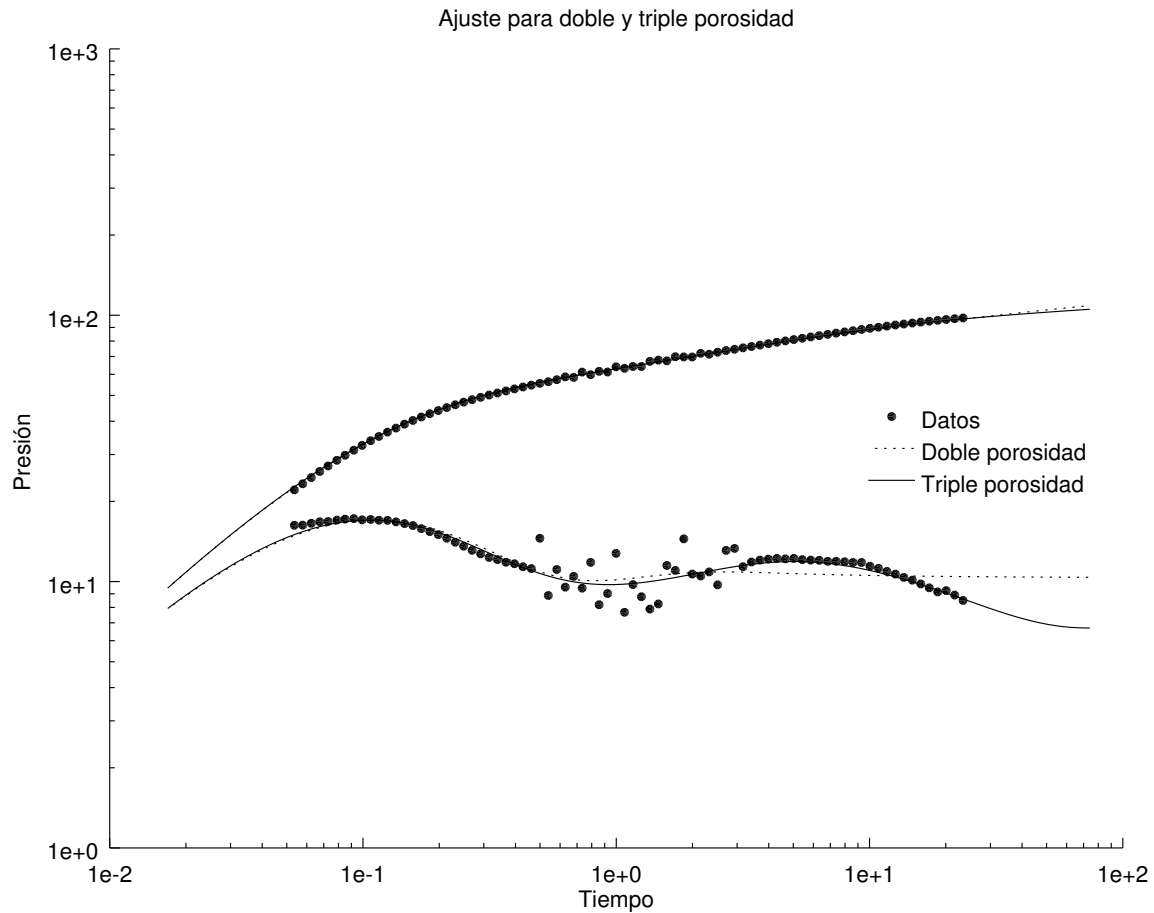


Figura 6.2: Ajuste de doble y de triple porosidad para el pozo 2

En el Cuadro 6.2 se presentan los parámetros obtenidos para los modelos de doble y de triple porosidad.

En este ejemplo el modelo de doble porosidad no es adecuado para reproducir los datos de la prueba de pozo, por lo que no es de extrañar que los valores obtenidos de ω , sean tan diferentes en ambos modelos.

Doble porosidad	Triple porosidad (PT)
$\omega = 0.4994$	$\omega_v = 0.0403$
	$\omega_f = 0.0021$
$\lambda = 1.88e-05$	$\lambda_{mv} = 4.98e-09$
	$\lambda_{mf} = 1.33e-05$
	$\lambda_{vf} = 4.88e-08$
	$\kappa = 0.1708$
$s = -2.7338$	$s = -4.0000$
$C_D = 1500.0$	$C_D = 1500.0$
$f_{obj} = 2.00e-04$	$f_{obj} = 6.98e-05$

Cuadro 6.2: Parámetros estimados para el Pozo 2, con los modelos de doble y de triple porosidad

Es el modelo de triple porosidad el que ha logrado obtener un ajuste de los datos adecuado, por lo que en este ejemplo se muestra la necesidad de un modelo de triple porosidad.

Conclusiones

En este trabajo se realizaron diferentes experimentos para comprender los efectos que tiene cada uno de los parámetros de un modelo matemático, que hay que identificar a través de un problema inverso. Para poder llevar a cabo esta identificación es necesario contar con datos medidos, en este caso la presión de una prueba de pozo.

Al estudiar el efecto de los parámetros nos fue posible diseñar métodos para estimarlos con mayor precisión.

Adicionalmente la metodología generada no solo es aplicable al modelo estudiado, sino que puede ser utilizada en diferentes modelos que simulan distintos fenómenos y con diferentes algoritmos de optimización heurísticos y deterministas.

De lo anterior obtenemos las siguientes conclusiones:

- Un análisis de sensibilidad de la presión a los parámetros del modelo estudiado nos permitió conocer la influencia que tiene cada parámetro, así como su efecto en relación a otras variables, en este caso el tiempo.
- El análisis de sensibilidad también nos permite predecir cuáles rangos de los valores de los parámetros podrían presentar mayor error de estimación, esto nos puede permitir también evaluar la eficiencia del optimizador, ya que si presenta errores de estimación, debe ser principalmente en los rango esperados.
- Los análisis de sensibilidad nos indicaron que es adecuado en este problema utilizar un espaciado logarítmico, lo que generó una reducción en el error medio de estimación de los parámetros de yacimiento desde 37 % a 47 %, cuando no se incluyen los parámetros de pozo ($s = C_D = 0$), y de 7 % a 40 % cuando se incluyen.
- Igualmente, de los análisis de sensibilidad observamos los rangos de valores de presión que genera el modelo estudiado, por lo que se determinó que una función objetivo de mínimos cuadrados ponderados es más apropiada que una de mínimos cuadrados tradicional, ya que está da mayor importancia a valores más grandes, los cuales están relacionados con los parámetros más sensibles (parámetros de pozo). Al cambiar la función objetivo se obtienen reducciones en el error medio de estimación de los parámetros de yacimiento, desde 7 % a 13 %, y para la mediana del error desde 12 % hasta 24 %.
- Dadas las correlaciones obtenidas, se determinó que en el espacio de Laplace era más adecuado el uso de la *presión de Laplace* ($\tau_D \bar{p}_{wD}(\tau_D)$) para el proceso de optimización, ya que esta mantiene mayor similitud con la presión en tiempo real $p_{wD}(t_D)$.

- Al utilizar la información de sensibilidad se desarrolló una optimización de tipo secuencial, la cual determina en primer lugar los parámetros de mayor influencia para así lograr la precisión necesaria en los de menor influencia. Todo esto utilizando para cada parámetro los rangos de tiempo en los que tiene mayor efecto, logrando reducciones en el error medio de estimación desde 24 % a 32 %, y de 39 % a 55 % para la mediana del error.
- Utilizando en conjunto las propuestas anteriores se encontró que el error medio de estimación se reduce desde 48 % (λ_{mf} y λ_{mv}) hasta 63 % (ω 's). En la mediana se presentan reducciones del 70 % (λ_{mf} y λ_{mv}) al 83 % (ω 's y κ). Lo anterior en contraste a una situación donde no se aplica la metodología propuesta.
- Con la metodología propuesta y un algoritmo de optimización heurístico, finalmente se obtiene un error medio de estimación para los parámetros de yacimiento de 10 % (ω 's) al 17 % (λ_{mf} y λ_{mv}), y de 4 % (ω 's) al 9 % (λ_{mf} y λ_{mv}) para la mediana del error.
- Se estudió la posibilidad de una transformación en los parámetros (diseñada para cada parámetro, basándose en los análisis de sensibilidad) para mejorar la precisión en las estimaciones, pero no se encontraron mejoras debidas a este cambio, por lo que se descartó esta idea. Se considera que la falta de mejora se debe a la poca sensibilidad de algunos de los parámetros y a que el algoritmo de optimización por intervalos se basa en el rango de los parámetros, por lo que es equivalente a tener todos los parámetros en la misma escala.
- La información del efecto de cada parámetro en diferentes tiempos nos permitió proponer una restricción adicional en el problema de optimización, para lograr evitar pseudo-mínimos (mínimos falsos debidos a problemas de precisión), para así llegar a los mínimos verdaderos.
- Con la metodología desarrollada fue posible hacer ajustes en datos reales, de pozos en yacimientos mexicanos, obteniendo buen ajuste con el modelo estudiado (modelo de triple porosidad) y se observó que el modelo de doble porosidad es un caso especial del de triple porosidad (por lo que este modelo sirve para casos de una, dos y tres porosidades).

Bibliografía

- [1] D. Abdassah and I. Ershaghi. Triple-porosity systems for representing naturally fractured reservoirs. *SPE Formation and Evaluation* 2, pages 113 – 127, 1986.
- [2] Ram G. Agarwal, Rafi Al-Hussainy, and H.J. Ramey Jr. An investigation of wellbore storage and skin effect in unsteady liquid flow: I. analytical treatment. *Society of Petroleum Engineers Journal*, 10:279 – 290, 1970.
- [3] G. I. Barenblatt, I. P. Zheltov, and I. N. Kochina. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks. *Journal of Applied Mathematics and Mechanics*, pages 1286 – 1303, 1960.
- [4] M. Bourgeois, Stanford University. Department of Petroleum Engineering, and E. Acquitaine. *Well Test Interpretation Using Laplace Space Type Curves*. Stanford University, 1992.
- [5] M.J. Bourgeois and R.N. Horne. Well-test-model recognition with laplace space. *SPE Formation Evaluation*, 8(1):17 – 25, 1993.
- [6] R. Camacho-Velázquez, M. Vásquez-Cruz, R. Castrejon-Aivar, and V. Arana-Ortiz. Pressure transient and decline curve behavior in naturally fractured-vuggy carbonate reservoirs. *SPE Reservoir Evaluation & Engineering* 8, pages 95 – 111, 2005.
- [7] Z. Chen, G. Huan, and Y. Ma. *Computational methods for multiphase flows in porous media*. Society for Industrial and Applied Mathematics, 2006.
- [8] G. Da Prat, H. Cinco-Ley, and H.J. Jr. Ramey. Decline curve analysis using type curves for two-porosity systems. *SPEJ*, pages 354 – 362, June 1981.
- [9] K. Dehghani, P. Harris, K. Edwards, and W. Dees. Modeling a vuggy carbonate reservoir. *McElroy Field, West Texas, AAPG Bulletin* 83, pages 19 – 42, 1999.
- [10] A.C. Gringarten, D.P. Bourdet, P.A. Landel, and V.J. Kniazeff. A comparison between different skin and wellbore storage type-curves for early-time transient analysis. In *SPE 8205*, Las Vegas, September 1979.
- [11] O. Gurpınar, J. Kalbus, and D. List. Numerical modeling of a large, naturally fractured oil complex. In *SPE International Petroleum Conference and Exhibition in Mexico*, Villahermosa, Mexico, 1-3 February 2000.

- [12] S. Gómez, R. Camacho, M. Vásquez, G. Ramos, N. del Castillo, and A. Mesejo. Well test characterization of naturally fractured vuggy reservoirs, with a global optimization method. In *OTC-24762-MS. Offshore Technology Conference-Asia*, 25-28 March 2014.
- [13] S. Gómez, G. Ramos, R. Camacho A. Mesejo, M. Vásquez, and N. del Castillo. Well test analysis of naturally fractured vuggy reservoirs with an analytical triple porosity-double permeability model and a global optimization method. *Oil & Gas Science and Technology*, 69(4):653 – 671, 2014.
- [14] S. Gómez, G. Ramos, A. Mesejo, R. Camacho, M. Vásquez, and N. Castillo. Study of the characterization of naturally fractured vuggy reservoirs, with totally penetrated wells using global optimization. synthetic and real data. Technical Report 156, IIMAS-UNAM, October 2013.
- [15] H. Kazemi. Pressure transient analysis of naturally fractured reservoirs with uniform fracture distribution. *SPE Journal* 9, pages 451 – 462, 1969.
- [16] N.G. Nair. *Measurement and modeling of multiscale flow and transport through large-vug Cretaceous carbonates*. Proquest, Umi Dissertation Publishing, 2008.
- [17] M. Onur. Numerical laplace transformation of sampled data for well-test analysis. *SPE Reservoir Evaluation & Engineering*, 4:268 – 277, 1998.
- [18] H. Stehfest. Algorithm 386: Numerical inversion of laplace transforms. *Communications of the ACM*, 13(1):47 – 49, 1970.
- [19] B. Vik, K. E. Sylta, and A. Skauge. Connectivity in vuggy carbonates, new experimental methods and applications. *Transport in Porous Media*, 93:561 – 575, 2012.
- [20] J. Warren and P. Root. The behavior of naturally fractured reservoirs. *SPE Journal* 3, pages 245 – 255, 1963.
- [21] D.J. Wilkinson. Pressure transient parameter estimation in the laplace domain. In *67th Annual Technical Conference and Exhibition of the SPE*, pages 601 – 611, Washington, DC, October 1992.
- [22] Y.-H. Wu, H. Liu, and G. Bodvarsson. A triple-continuum approach for modeling flow and transport processes in fractured rock. *Journal of Contaminant Hydrology* 73, pages 145 – 179, 2004.
- [23] L. Zhang, S.L. Bryant, J.W. Jennings, T.J. Arbogast, and R. Paruchuri. Multiscale flow and transport in highly heterogeneous carbonates. In *SPE Annual Technical Conference and Exhibition*, pages 26 – 29, Houston, Texas, September 2004.
- [24] L. Zhang, N. Nair, J.W. Jennings, and S.L. Bryant. Models and methods for determining transport properties of touching-vug carbonates. In *SPE Annual Technical Conference and Exhibition*, Dallas, Texas, October 2005.