CONTINUOUS RESERVOIR MODEL UPDATING USING AN ENSEMBLE KALMAN FILTER WITH A STREAMLINE-BASED COVARIANCE LOCALIZATION

A Thesis

by

ELKIN RAFAEL ARROYO NEGRETE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2006

Major Subject: Petroleum Engineering

CONTINUOUS RESERVOIR MODEL UPDATING USING AN ENSEMBLE KALMAN FILTER WITH A STREAMLINE-BASED COVARIANCE LOCALIZATION

A Thesis

by

ELKIN RAFAEL ARROYO NEGRETE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Akhil Dattagupta
Committee Members,	Bani K. Mallick
	J. Bryan Maggard
Head of Department,	Stephen A. Holditch

December 2006

Major Subject: Petroleum Engineering

ABSTRACT

Continuous Reservoir Model Updating Using an Ensemble Kalman Filter with a Streamline-Based Covariance Localization.

(December 2006)

Elkin Rafael Arroyo Negrete, B.S., Universidad Industrial de Santander, Colombia

Chair of Advisory Committee: Dr. Akhil Dattagupta

This work presents a new approach that combines the comprehensive capabilities of the ensemble Kalman filter (EnKF) and the flow path information from streamlines to eliminate and/or reduce some of the problems and limitations of the use of the EnKF for history matching reservoir models. The recent use of the EnKF for data assimilation and assessment of uncertainties in future forecasts in reservoir engineering seems to be promising. EnKF provides ways of incorporating any type of production data or time lapse seismic information in an efficient way. However, the use of the EnKF in history matching comes with its shares of challenges and concerns. The overshooting of parameters leading to loss of geologic realism, possible increase in the material balance errors of the updated phase(s), and limitations associated with non-Gaussian permeability distribution are some of the most critical problems of the EnKF. The use of larger ensemble size may mitigate some of these problems but are prohibitively expensive in practice.

We present a streamline-based conditioning technique that can be implemented with the EnKF to eliminate or reduce the magnitude of these problems, allowing for the use of a reduced ensemble size, thereby leading to significant savings in time during field scale implementation. Our approach involves no extra computational cost and is easy to implement. Additionally, the final history matched model tends to preserve most of the geological features of the initial geologic model.

A quick look at the procedure is provided that enables the implementation of this approach into the current EnKF implementations. Our procedure uses the streamline path information to condition the covariance matrix in the Kalman Update. We demonstrate the power and utility of our approach with synthetic examples and a field case. Our result

shows that using the conditioned technique presented in this thesis, the overshooting/undershooting problems disappears and the limitation to work with non-Gaussian distribution is reduced. Finally, an analysis of the scalability in a parallel implementation of our computer code is given.

DEDICATION

To the memory of my grandfathers and grandmothers, my very much loved Bertha Sanchez, Marcelina Alarcon, Joaquin Arroyo and Rafael Negrete. To my beloved mother, father and sisters (Doris Negrete, Jairo Arroyo, Anadela Arroyo, Jehydy Arroyo), to my always affectionate and loved wife (Alicia Cruz), and to my huge, loving and fun family.

ACKNOWLEDGMENTS

First of all, I like to sincerely thank my advisor and committee chair, Dr. Akhil Dattagupta, for his continuous encouragement, financial support, and especially for his academic and creative guidance. I was always greatly surprised by his amazing deep knowledge of the topics we studied in our group and the way he can always suggest new ideas that made my research work seem to be easily undertaken. I'd also like to thank Dr. Jonggeun Choe for his guidance during the early steps of this research.

My deepest gratitude to my friends in rooms 701 and 702: Deepak Devegowda, Dayo Oyerinde, Eduardo Jimenez, Ahmed Alhuthali, Ichiro Osako, Xialin Ma, Ajitabh Kumar, Jong UK Kim, Vegard Røine Stenerud, Rahul Mukerjee and Eileen Paclibon. It's been a privilege to work hand-in-hand with this amazing crew.

Last, but not least, I also want to mention my friends in Numerica Ltda (Alejo, Pipe, Felix, Jairo) and Ecopetrol-ICP Colombia, whom make me value the importance of science in my life; especially Dr Carlos Piedrahita for his constructive lectures in mathematics which opened a new and bright perspective of the right value of mathematics in engineering. Finally, Farid Majana, Carlos Montana, Eduardo A. Idrobo and Gabriel Alvarez for their priceless help and friendship.

Thank you very much.

TABLE OF CONTENTS

Page

ABSTRACT……..…….iii

DEDICATION……..……...v

ACKNOWLEDGMENTS …….……..……..vi

TABLE OF CONTENTS... ……..vii

LIST OF FIGURES ……..…….....……..……..ix

LIST OF TABLES ……..…….....…….....xvi

CHAPTER

I INTRODUCTION 1

1.1 Motivation and Literature Review 3
1.2 Objectives of Study 7

II INVERSE PROBLEM THEORY 9

2.1 The Forward Problem 9
2.1.1 Reservoir fluid flow: the forward equations in history matching 10
2.1.2 About the probabilistic nature of the solution of the flow equations..... 10
2.2 The Inverse Problem 12
2.3 Element of the Inverse Theory..... 14
2.3.1 Measurements or observed data..... 14
2.3.2 Prior information on model parameters 15
2.3.3 Likelihood function..... 16
2.3.4 The Bayes theorem and the posterior distribution 18
2.4 Solution of an Inverse Problem..... 19
2.4.1 Gradient based solution..... 20
2.4.2 Monte Carlo solution 20

III SOLUTION OF INVERSE PROBLEM USING KALMAN FILTERS 22

3.1 What Is a Filter and How It Is Connected to Inverse Problems..... 22
3.2 The Kalman Filter Equations 23
3.2.1 Minimum variance estimator - MVE..... 26
3.2.2 Maximum posterior estimate - MAP 29
3.3 The Ensemble Kalman Filter Applied to Reservoir Engineering 31

CHAPTER Page

3.3.1 The forecast step and time update step 32
3.4 Problems with the Assimilation of Different Types of Observation 34
3.4.1 Sequential processing..... 35

IV APPLICATION AND LIMITATIONS OF THE ENSEMBLE KALMAN FILTER TO HISTORY MATCHING 37

4.1 The Use of the Standard EnKF in History Matching..... 37
4.1.1 2D water flooding example using a log-normal permeability distribution..... 37
4.1.2 2D water flooding example using a bimodal permeability distribution 40
4.1.3 The Goldsmith field case study..... 44
4.2 Problems Posted by the Standard EnKF 50
4.3 Towards a Solution: Some Numerical Experiments with the EnKF 50
4.3.1 The normal score transform..... 51
4.3.2 Travel time inversion of water cut..... 55
4.3.3 Distance dependent conditioning of the covariance 62
4.4 The Use of the Streamline Assisted EnKF in History Matching..... 69
4.4.1 2D water flooding example using a bimodal permeability distribution 71
4.4.2 The PUNQ-S3 synthetic case study..... 73
4.4.3 The Goldsmith field case study..... 83
4.5 Parallel Implementation of the Algorithm and Scalability Analysis 86

V CONCLUDING REMARKS AND RECOMMENDATIONS 90

5.1 Conclusions..... 90
5.2 Recommendations..... 91

REFERENCES..... 93

APPENDIX..... 97

VITA 100

LIST OF FIGURES

FIGURE Page

1. Five random realizations of a 3D Gaussian random field. Random realizations belong to the first layer of an actual reservoir field. All this realization satisfy the initial data values taken from core, logs and seismic. The upper left figure shows the mean from one hundred realizations; notice how the mean is too smooth and lost the sharp permeability contrast. This clearly explains why the common practice of using the mean from the realization should be abandoned, even if it is accompanied by some analysis of error and resolution. 11

2. Water cut response from one hundred realization of the same reservoir field shown in figure above. Notice the answer to the question how much will be the water cut after 4 years is not longer a single value but a set of values that can be cast into a probability density function. 12

3. Prior Gaussian joint PDF. There is no clear relationship between N and m from the prior information. 15

4. Likelihood function. Notice there is not a single pair of values {N,m} that maximize the function. For this particular likelihood an infinite number of combination of the model parameter {N,m} can satisfy the observation. The prior will help us to bound the proper solution to the inverse problem.....18

5. The probabilities density $P_x(x)$ and $P_L(d|x)$ respectively represent the information on the prior model parameters (left) and the likelihood P.D.F (center) explaining how good a model x is explaining the observed data d_{obs}. Given the two states of information represented by $P_x(x)$ and $P_L(d|x)$ the posterior PDF (right) represent the combination of the two states of information.....21

6. Sensor which provides a measurement. 22

7. Signal reconstruction using an estimator/filter. 23

8. Permeability map for a 50x50x1 reservoir model; dots in the four corners represent the producers and the arrow in the center of the model represent the injector. 38

FIGURE Page

9. Initial water cut spread from 100 different realizations; the reference water cut is shown by the bold red line; (a) shows the water cut at well P1 (b) shows the water cut at well P2. 39

10. Water cut spread from 100 history matched realizations; the reference water cut is shown by the bold red line; (a) shows the water cut at well P1 (b) shows the water cut at well P2. 39

11. (a) Initial and (b) final log permeability distribution. True permeability is in the right upper corner of figure (a) and (b). The mean from the ensemble is next behin to the true permeability. Other permeabilities belong to some randomly selected members/realizations.....40

12. Nine spot waterflooding example; figure shows the porosity of the reference model and the position of the producer and injectors. 41

13. (a) Initial log perm map (b) Initial log perm histogram. True permeability is in the right upper corner of figure (a) and (b). The mean from the ensemble is next behin to the true permeability. Other permeabilities belong to some randomly selected members/realizations.....42

14. (a) Final log perm map (b) Final log perm histogram. True permeability is in the right upper corner of figure (a) and (b). The mean from the ensemble is next behin to the true permeability. Other permeabilities belong to some randomly selected members/realizations.....43

15. First row shows the initial water cut spread for 100 members at wells P1,P2,P4; second row shows the final spread at the same wells after history matching.....44

16. Goldsmith study area; well distribution producer with water cut information highlighted in yellow; Injectors in blue (from Datta-Gupta). 45

17. Gaussian random fields for the Goldsmith case (randomly selected members: 17 left, 58 center and 97 right.). Permeability maps generated using sequential Gaussian cosimulation conditioned to wells and seismic data. Below each map it shows the histogram of the permeability.....46

FIGURE Page

18. Updated permeability maps using the standard EnKF conditioned to water cut. Below each map it shows the histogram. Notice the standard EnKF is unable to preserve the prior geologic density function. 47

19. Initial (top row) and final, after standard EnKF (bottom row) water cut spread at wells P1, P7; 100 members in the ensemble; after assimilation the spread was greatly decreased around the observation shown in red. Unfortunately the permeability corresponding to these models have an unrealistic distribution; gray line around 4000 days shows the time up to which the information was assimilated. 48

20. (Left) Initial mean from the 100 members ensemble; (center) mean from the 100 members ensemble after standard EnKF update; (right) changes in the permeability. The updated mean (center) looks pretty sharply and perhaps unrealistic. The changes in the mean (right) shows that changes has indiscriminately occur different at spots all around the reservoir field. 49

21. Permeability map comparison (a) shows the natural logarithm of the permeability and its histogram (b) shows the normal score transformatio of the permeability and its histogram. 52

22. (a) Initial permeability map (b) permeability map after EnKF update with normal score transformation of the permeability. For this example the normal score seems to provide valid results. 53

23. Final water cut spread after assimilation of water cut at wells P1, P2, and P4. The spread correspond to a 100 models in the ensemble. Observation is shown by the bold red curve. 53

24. Updated permeability maps using the standard EnKF with normal score transformation of the permeability conditioned to water cut. Below each map it shows the histogram. Notice the normal score EnKF is able to preserve the prior geologic probability density. A closer look into the update permeability map reveals patches and high permeability contrast, which suggest the updated permeability may have some problems; although this results appears to be better than the standard EnKF I believe they should be used carefully. 55

25. An illustration of amplitude and peak travel time inversion. (a) Shows the classical amplitude inversion approach (b) shows the travel time inversion approach. (from Cheng & Datta-Gupta[23]). 56

FIGURE Page

26. Different case scenarios were one might need to compute the shift time for the observed water cut (in magenta). The vertical orange lines represent times we want to assimilate information. Horizontal arrows shows the shift time required for each curve at the different assimilation times.....57

27. (a) Updated permeability map using water cut (b) Updated permeability map using shift time. True permeability is in the right upper corner of figure (a) and (b). The mean from the ensemble is next behin to the true permeability. Other permeabilities belong to some randomly selected members/realizations. Althought there seems not to be a huge different beween the two approaches the final map for the shift time (b) do not present the overshooting problems preseted in (a). This suggest that use of the shift time may sligly decrease some overshooting problems.....59

28. First row shows the final water cut spread after assimilation of shift time; water cut shown at wells P1, P2, and P4. Second row shows the final water cut spread after assimilation of the water cut amplitude. Reference model show by the red bold line. Both cases the simulator was restarted from time zero.....60

29. Comparison of the RMS error; To compute the RMS we used the mean of the ensemble and compare its response with the true water cut for all wells. Curve labeled with WWCT shows the RMS error after assimilating water cut with out restarting the simulator; magenta and green shown the results for the shift time and water cut respectively both cases the simulator was restarted from zero.....61

30. (a) and (c) shows the cross covariance between shift time (first row measured at well P1, second row at well P4) and permeability at times 240 and 300 days respectively; (b) and (c) shows the cross covariance between water cut amplitude (first row measured at well P1, second row at well P4) and permeability at 240 and 300 days respectively; Notice how the similar the cross covariance are at early time 240 days. Later the cross covariance are not longer alike but have a similar trend. This may help to explain why both techniques results in a similar posterior permeability. 62

FIGURE Page

31. Example showing the cross covariance between the observation at a given well and the permeability at two different cells cells; upper scatter plot shows that according with the results from the ensemble the correlation between gridblock 23 and the oil production rate at well is almost linear; the lower scatter plot shows that for the grid block 132 the correlation is not well resolved by the covariance operator. 64

32. Gridlocks behind the saturation front crossed by streamlines arriving at well P8 at time 330 days for members 1 to 25. Reservoir model correspond to the nine spot examples with bimodal permeability distribution, shown the previous section.....67

33. Product of stacking regions selected by one hundred different realizations. The regions selected by the streamlines change during the assimilation period; Notice at early times 330-510 days the stacking regions cover more area; This is due to the fact that at early time when no too many points has been assimilated the permeability field of the realizations differ a lot between them; at late time all realization has similar characteristic therefore the stacked region are more uniform. 69

34. Streamline assisted ensemble Kalman filter flow chart. 70

35. a) Initial permeability map for randomly selected members in the ensemble, upper left corner shows the mean; b) posterior, after application of the standard EnKF, over/undershooting is evident in all the members; c) posterior, after application of the SL EnKF, problems with over/undershooting disappeared. 70

36. Comparison of the true PDF with respect to member 23; (Left) shows the Q-Q plot comparing the true vs member 23; notice that the Q-Q plot gives a straight line which mean that both PDFs can be considers the same; below the Q-Q plot it shows the histogram of the true permeability. (Center) shows the Q-Q plot comparing the initial PDF vs the posterior PDF after using the standard EnKF; the Q-Q plot does not show a straight line which means that the distribution are non longer equals; below the histogram of the posterior shows that the PDF changed from bimodal to a almost Gaussian distribution; (Right) shows the Q-Q plot comparing the initial PDF vs the posterior PDF after using the SL EnKF; the Q-Q plot shows an almost straight line which means that the posterior distribution was able to preserver most of the characteristic of the initial; below the histogram of the posterior shows that the PDF is still bimodal. 71

FIGURE Page

37. Comparison of the water cut spread for initial members (first row) the standard EnKF (second) row and the SL EnKF (third row). 72

38. Top surface map showing well locations (from Floris *et al*[20].). 74

39. Flow path information from streamlines. (a) Zones affected by the streamlines arriving at each producer at early time (b) Zones selected by streamlines arriving at each producer at later time. 76

40. (a) Initial spread for bottom hole pressure at well PRO-15 (b) gas to oil ratio at well PRO-1 and (c) water cut at producer PRO-11. Response from 100 members in the ensemble; red bold line shows the data for the reference model. 78

41. (a) Final bottom hole pressure spread at well PRO-15, after standard EnKF data assimilation, (b) gas to oil ratio at well PRO-1 and (c) water cut at producer PRO-11. Response from the updated 100 members in the ensemble; red bold line shows the data for the reference model. Gray line around 3000 days shows time up to which the information was assimilated; all models were re-run from time zero to avoid possible increase in the material balance error. 78

42. (a) Final bottom hole pressure spread at well PRO-15, after SL EnKF data assimilation, (b) gas to oil ratio at well PRO-1 and (c) water cut at producer PRO-11. Response from the updated 100 members in the ensemble; red bold line shows the data for the reference model. 78

43. RMS error for the dynamic variable. (a) Shows the global bottom hole pressure RMS error for the streamline assisted EnKF and for the standard EnKF, (b) shows the global RMS error of the gas oil ratio. The error in the dynamic variables is decreased to a similar extend for both techniques. 79

44. 60 members ensemble variance map; map computed using the updated permeabilities at assimilation time 2936 days; layer 2 and 4 were set to a transparent appearance to allow other layer being shown in detail; (a) 60 member's ensemble using the SL EnKF (b) 60 member's ensemble using the standard EnKF; the high strike in the variance map for the standard EnKF indicated that the posterior permeability in all the members are more apart from the true permeability as compared with the SL EnKF. 81

FIGURE Page

45. RMS error from the mean of the ensemble. For this synthetic case the reference permeability is known, thus the RMS error can be computed. Similar result can be inferred from the multidimensional variance plot map. (a) RMS error for the natural logarithm of the permeability after each assimilation step. (b) RMS error for the natural logarithm of the permeability for the standard EnkF and the streamline assisted EnKF for different number of members in the ensemble. 83

46. Regions selected by the streamlines arriving to producer P6 (black circle) at different times; (right) time 1680 days, (center) 2280 days (left) 3960 days. Region change with time because the schedule changes and the pressure field changes. This information was used to condition the covariance matrix. 84

47. Updated permeability map using the SL EnKF after water cut assimilation; below each map it shows the corresponding histograms. Notice the SL EnKF is able to preserve the prior density function; this is a required characteristic of any good history matching technique..... 85

48. Initial (top row) and final, after SL EnKF (bottom row) water cut spread at wells P1, P7 100 members in the ensemble; after assimilation the spread was greatly decreased around the observation shown in red; gray line around 4000 days shows the time up to which the information was assimilated..... 86

49. (Left) Initial mean from the 100 members ensemble; (center) mean from the 100 members ensemble after SL EnKF update.; (right) changes in the permeability it shows the changes have discriminately occur in few zones around the reservoir field. The changes are small, realistic and localized as compared with the one proposed by the standard EnKF (see Figure 20). 87

50. (a) total running time for the Goldsmith field case as a function of the number of processors. (b) Speed up factor for the same dataset; notice front the speed-up factor analysis the current implementation could in theory decrease the total running time by a factor of eight when using 16 CPUs. 88

51. Covariance error ratio for different values of correlation coefficient and size of the ensemble. 99

LIST OF TABLES

TABLE Page

1. Production and PVT data for a reservoir. It is possible to find m and N to fit this data (from Dake[37]) 17

2. Detail of six different numerical experiment using PUNQ-S3 75

3. Different types of correlation function defined using streamlines. First row used with pressure and second row used with water saturation..... 76

4. EnKF Running time for the Goldsmith field case. 87

CHAPTER I

INTRODUCTION

Proper characterization of the reservoir and the assessment of uncertainty is a crucial aspect of any optimal reservoir development plan and management strategy. To achieve this goal, it is necessary to reconcile geological models to the dynamic response of the reservoir through a process known as history matching. This procedure of history matching has been a topic of great interest[1-4] and an area of active research in the oil industry. Some significant developments have been made in the area of dynamic data integration. Most of these techniques involve computation of sensitivities which relate changes in production response at a well to a change in reservoir parameters. Techniques for automatic history matching that do not use sensitivity or gradient based approaches are stochastic algorithms such as Markov Chain Monte Carlo (MCMC) and simulated annealing. A recent development in the field of production data integration without the use of sensitivities or gradients is known as the ensemble Kalman Filter[5] (EnKF).

The EnKF is simulator independent. It can be linked to any existing reservoir simulator. Moreover, EnKF uses a sequential updating technique, that is, the data is assimilated as and it when it becomes available. The EnKF can assimilate the latest production data without re-running the simulator from the initial conditions. These characteristic makes it suitable for continuous model updating and clearly a more natural approach to reservoir characterization when compared to traditional history matching techniques. Furthermore, the increased application of down hole monitors, intelligent well systems and permanent sensors to continuously record pressure, well rates and temperature has provided a further boost to the sequential model updating approach via EnKF. These key features combined with the ease of implementation have generated significant interest[6-13] in its applicability to the area of reservoir characterization.

This thesis follows the style of *SPE Journal.*

In addition to all the persuasive properties discussed above, some important constrains are necessary in any history matching technique. A key requirement is that the final model should honor initial geological information and retain geologic realism. The EnKF works quite well when the initial parameters distribution is Gaussian. But it can have serious limitations with non-Gaussian distribution. The EnKF tends to transform multi-modal permeability distributions to a more normal or Gaussian distribution over a sequence of many updates[14]. This transformation leads to a loss of structure in the permeability field with the final model losing continuity of flow channels and barriers. The use of these 'history matched' suite of models for future forecasts or uncertainty analysis would certainly lead to erroneous interpretation and field sub-optimal development strategies. Failure of the EnKF to work with non-Gaussian distribution is related to the low order moment statistics used to characterize the model state.

In the past few years, we have seen the implementation of the EnKF in field-scale reservoir characterization, including some recent papers[13] that attempt to deal with some of the challenges pertaining to the use of the EnKF. When applied to synthetic cases, where a comparison can be made with the reference permeability and the final history matched permeability. Overshooting of permeability values have been reported[7-8]. Other common difficulty experienced when using the EnKF is known as the filter divergence. The effect of filter divergence is such that the distribution produced by the filter drifts away from the truth. Filter divergence normally occurs because the prior probability distribution becomes too narrow and the observations have progressively diminishing impact on the Kalman gain.

The most common approach to deal with filter divergence is to add some (white) noise to the prior ensemble of state vectors to “broader” its distribution and enhance the impact of the observations in the Kalman gain. Other authors[15] have proposed the use of a pair of ensemble where statistic from one ensemble is used to update the other. Treat the observation as random variables[16] by adding noise and the use of high values for the noise, can decrease the magnitude of the problem too. Some of EnKF problems and limitations can be controlled through the use of a much larger ensemble size. On previous implementations of the EnKF in reservoir history matching, an ensemble size of 60-100 models has been proposed. Our experience indicated that increasing the ensemble size to

more than 300 initial models reduces the severity of these problems. However, this approach is computationally demanding and in the case of field scale applications, may be prohibitively expensive to implement.

In this thesis, we describe an approach to address many of the currently reported challenges in the use of the EnKF applied to reservoir history matching. The unique feature of this proposed approach is that the final models, that constitute the ensemble, tend to retain the geological information that went into building them initially. Over a sequence of many EnKF updates, our approach tends to preserve the shape of the initial permeability distribution and consequently retains key geological features. The new technique also greatly decreases the severity of the overshooting problem reported in earlier implementations of the EnKF. Moreover, it allows the use of smaller ensemble size while providing improved results compared to the standard EnKF with similar ensemble size.

We have implemented and demonstrated the power of the approach through a set of synthetic example and further validated the technique through a field example. The results clearly indicate the technical feasibility of the approach for continuous model updating through the use of the Ensemble Kalman Filter for model calibration and reservoir characterization.

1.1 Motivation and Literature Review

The Kalman filter was first proposed in 1960 to solve optimal control problems. The traditional Kalman filter is more suitable for cases with a small number of variables and a linear relationship between model parameters and data observation. Application to nonlinear problems was at least partially solved by the development of the extended Kalman filter. Most forward problems in reservoir engineering are highly nonlinear and thus the solution of reservoir history matching using the traditional Kalman filter is not feasible. It was not until 1994 when Evensen[5] proposed a new methodology that made possible the use of Kalman filters for large nonlinear models. He noticed that to properly apply the Kalman equation it is necessary to have a stochastic description of the model state at any position in time. In fact, for normal distribution only the mean and the covariance are required. The equations that must be solved for such purpose are the

continuity equation for probability also known as the Kolmogorov equation. However, the solution of the Kolmogorov equation by direct numerical methods is impractical because of the size of most real problems. Then, Evensen using ideas from Epstein[17] (1969) proposed the use of an alternative solution of the Kolmogorov equation, using Monte Carlo methods. This technique is known as the ensemble Kalman filter (EnKF). Evensen[5] in his already classic paper discusses the stochastic dynamic nature of the atmospheric equations due to uncertainties in the initial conditions. He proposes to use the EnKF to incorporate new observations and improve the model forecast. Some additional clarification to the initial method as proposed by Evensen[5] (1994) was given by Burgers *et al*[16]. They make clear that it is necessary to treat the observations as a random variables. This implies that one should add noise with the correct statistic to the observation to create an ensemble of observations that is then used in updating the ensemble of model states. His recommendation helped to decrease problems related with the too low variance of the ensemble which ultimately can results in filter divergence problems. This problem will be discussed later.

From a different perspective and perhaps a more general point of view, other authors[18,19,20] in different areas of science have faced the problem of model parameter estimation also known as inverse problems. Inverse problems have been studied from a deterministic point[†] of view as well as from a probabilistic point of view. The first attempt to formalize the probabilistic approach to inverse problem was done by Backus[20] 1970.

Although, in principle they appear to be different problems, filtering a signal using Kalman filters and the probabilistic/Bayesian approach to inverse problems are the same. In his excellent book Tarantola[18] (2005) proved that the techniques, Kalman filter and the Bayesian approach to inverse linear problems, are actually equivalent. The same proof is later discussed in the petroleum engineering scene by Gu and Oliver[13] (2006). Further discussion of the similarities of both methods taking into account the ensemble nature of the EnKF is given by Anderson[21] (2003).

[†] The deterministic point of view of the inverse problem is also known as the frequentist approach. See J.A. Scales "Introduction to Geophysical Inverse Theory" for a discussion of this approach

Inverse problem solution using a Bayesian approach has a long history in petroleum engineering. On the other hand parameter estimation from the perspective of the Kalman filter is somewhat new in the petroleum engineering. Lorentzen *et al*[22] first used the EnKF in the petroleum industry to improve the forecast of the UBD (underbalanced drilling operations) or LHD (in low-head drilling) process by updating the relevant physical quantities such as gas holdup and Fanning friction factor. These quantities are updated using the measurements that become available during production such as injection rates, downhole pressure, choke pressure, and outlet rates.

In the area of reservoir characterization Nævdal[6] (2002) proposed the use of the ensemble Kalman filter to update near well reservoir model parameter as new measurements become available. Later Nævdal[7] (2003) again uses the EnKF and takes a further step to continually update the permeability field for a simplified real field reservoir simulation model. In the same paper he mentions the possibility of using parallel processing to update the reservoir models independently. After this paper other research groups around the world started giving increased attention to the use of EnKF in the area of history matching. Gu and Oliver[8] (2004) applied the EnKF to the well known PUNQ-S3 reservoir model showing that it is possible to use this methodology with very good results in three-phase reservoir model. Nevertheless, they found that overshooting of the permeability and porosity can occur and recommend further investigation of this issue. With the aim of comparing the randomized maximum likelihood RML and the EnKF Gao, Zafari and Reynolds[9] quantified the uncertainty of the EnKF for the PUNQ-S3 model and compared it against other history matching and uncertainty evaluation techniques.

Other interesting research in the area of history matching using the EnKF was published by Gu, Oliver[13] (2006). They discuss some of the limitations of the EnKF while working with non Gaussian members in the state variable. They analyze the effects of updating the water saturation using the EnKF in a one-dimensional problem. Since the saturation distribution is bimodal, it is not well modeled by the mean and variance. The updated water saturation presented overshooting and undershooting problems. They tested ideas to replace the water saturation for a proxy variable derived with technique like normal score transform, saturation front location and iterating the updated

permeability, to reduce the magnitude of over/undershooting problem in the water saturation. Other important contribution of this paper is that they clarify that the EnKF and the linear inversion problem using a Bayesian approach are equivalent. We will see later during the development of this thesis that using the normal score transform to solve problems related with non-Gaussian distributions does not always works.

Further evidence of the limitations of the EnKF to work with non-Gaussian distribution and its reliability characterizing the uncertainty in the case where the conditional probability distribution function is multi modal was given by Zafari, M., Reynolds, A.C[10] (2005).

In other areas of the reservoir engineering, Brouwer *et al*[23] (2004) uses the EnKF to successfully couple it with an optimal control algorithm to improve reservoir management. Another interesting paper on the use of the EnKF to incorporate time lapse seismic information is discussed by Skjervheim *et at*[11] (2005). He shows how the time lapse information provides an increased accuracy in the porosity field compared to the assimilation of only production data.

From the previous literature review it becomes clear that despite the increasing amount of publications promoting the use of the EnKF for history matching and uncertainty evaluation. The technique still has some problems and limitations that need to be addressed. This problems and limitations are not unique to the petroleum engineering world. The same challenges have been faced in other areas of engineering like optimal control theory (which was the place were the Kalman filter theory started) and oceanography and hydrology. Some proposed solutions to the limitations of the Kalman filter to work with non-Gaussian field are discussed by Anderson and Moore[24] (1979) they realized that any arbitrary distribution can be decomposed as the summation of individual Gaussian distribution. This property is known as Gaussian summation approximation. Unfortunately application of the Gaussian summation approximation along with an ensemble Kalman filter may be prohibitively expensive to implement. Attempts to use the EnKF with a non-Gaussian distribution in the area hydrologic data assimilation was studied by Reichele, R.H., McLaughlin, D.B., and Entekhabi, D.[25] (2002). References of papers discussing this limitation in the petroleum engineering were discussed above.

Other common difficulty experienced when using the EnKF is filter divergence. The effect of filter divergence is such that the distribution produced by the filter drifts away from the truth. Filter divergence normally occurs because the prior probability distribution becomes too narrow and the observation has progressively less impact on the Kalman gain. This phenomenon appears to be closely related with the overshooting/undershooting problems discussed above. Several books in optima control theory give different practical recommendation regarding this issue. For instance, see Anderson, B.D and Moore, J.B.[24] (1979), Kamen, E.W. and Su, J.K[26] and Brown and Hwang[27].

In this thesis we also documented some of the issues related to the parallel implementation of the EnKF. A summary of some numerical experiments will be discussed later. References of similar implementation are discuses below. A parallel implementation of the EnKF was first reported by Houtekamer and Mitchell[28] they parallelized the Kalman update step in a clever way; later Keppene[29] 2000 also document its experience with parallel application of the EnKF; in both papers the authors took advantages of the inherent parallelism in the EnKF by running each ensemble member on a different processor of a parallel computer. In this work we followed a similar approach to run each ensemble member in a different processor.

1.2 Objectives of Study

The main objective of this research is to study, analyze, apply and improve a data inversion technique known as the ensemble Kalman filter with specific emphasis in reservoir history matching and/or continuous model updating. Followings are the basic objectives:

- Provide a modified formulation to eliminate or reduce some of the problem of the traditional ensemble Kalman filter.
- Use shift time instead of water cut as a proxy variable in the state vector to check whether it may be beneficial to the EnKF.
- Test if the normal score transform is able to minimize the limitations in the EnKF, for non-Gaussian parameters distribution, while providing a reasonable updated these parameters.

- Use of additional information extracted from the streamline technology to condition the covariance matrix.
- Study the influence of some parameters like number of members in the ensemble and the type of covariance conditioning in the final result of the inversion process.
- Implementation of this procedure in a FORTRAN program for automatic history matching using Eclipse and/or Frontsim as forward simulators.
- Include basic parallel capabilities to the FORTRAN computer code using MPI, specifically running the forward simulator for each ensemble member in a different CPU.

CHAPTER II

INVERSE PROBLEM THEORY

2.1 The Forward Problem

In basic engineering courses we learnt that physical theories allow us to make predictions over a system. Given a complete description of a physical system, we can forecast or predict the outcome of some measurement of interest. The problem of predicting the results of a measurement is called the *forward problem*.

Perhaps the most common example, because we all dealt with it daily is the weather forecast. Engineers knowing the climate state at a point in time and space, are able to predict the future climate conditions. They use the governing laws that describe how pressure, temperature, moisture, air density, and wind will change with time at position

Forward problem of interest in petroleum engineering are many, As an example of some of then one can mention: a) production forecast in reservoir engineering (fluid flow in porous medium); b) well bore stability in drilling engineering (mechanics of solids/rocks and failure criteria); c) well stimulation using hydraulic fracture design and acid stimulation in production engineering (mechanics of solids/rocks; fluid mechanics; complex chemical reactions); The list continue to grow every day, a complete list is out of the scope of the present thesis.

Typically the forward problem is characterized by a set of partial differential equations PDE and a set of boundary conditions and/or initial conditions. The solution of this set of PDE gives a deterministic relationship between the model parameters **m** and the model response or data parameter **d**. This relationship can be expressed in a short notation as:

$$\mathbf{d} = g(\mathbf{m}) \quad \text{.......} \quad (2.1)$$

Notice that this short notation may represent a more complex set of equations or even the result from a very complicated computer program like a reservoir simulator. The (usually non linear) operator *g(·)* is called the forward operator.

2.1.1 Reservoir fluid flow: the forward equations in history matching

Following Aziz and Settari[30] the two phase formulation of fluid flow in a porous medium is given by

$$\nabla \bullet \left[\lambda_w (\nabla p_w - \gamma_w \nabla z)\right] = \frac{\partial}{\partial t}\left[\phi \frac{S_w}{B_w}\right] + q_w$$

$$\nabla \bullet \left[\lambda_n (\nabla p_n - \gamma_n \nabla z)\right] = \frac{\partial}{\partial t}\left[\phi \frac{S_n}{B_n}\right] + q_n \quad \text{..... (2.2)}$$

$$P_c = p_n - p_w = f(S_w)$$

$$S_w + S_n = 1$$

The solution of this set of equations for a given boundary and initial conditions can be easily accomplished nowadays by a reservoir simulator. During the development of this research work, we used the commercial reservoir simulators ECLIPSE and FRONTSIM.

2.1.2 About the probabilistic nature of the solution of the flow equations

We have seen that the forward problem describes the underlying principles as deterministic processes. The solution of our forward problem was given by **d**=g(**m**). Notice, it is assumed the model parameter **m** is known with certainty. This requirement is very strong, especially in reservoir engineering, because in actual cases it is hard to know **m** with certainty. To explain the consequences of such lack of certainty a brief explanation follows. What follows is greatly inspired by Epstein[17] (1969).

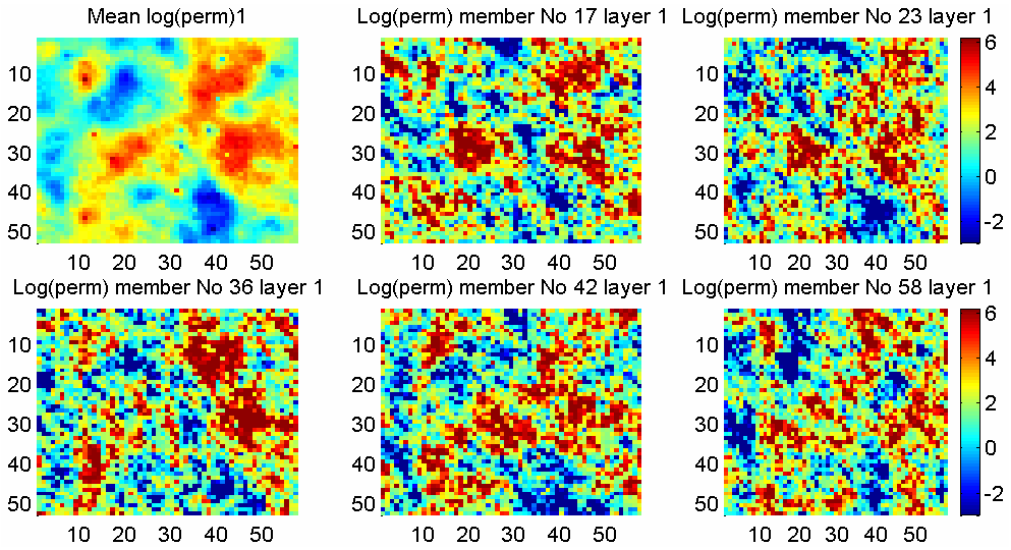


Figure 1. Five random realizations of a 3D Gaussian random field. Random realizations belong to the first layer of an actual reservoir field. All this realization satisfy the initial data values taken from core, logs and seismic. The upper left figure shows the mean from one hundred realizations; notice how the mean is too smooth and lost the sharp permeability contrast. This clearly explains why the common practice of using the mean from the realization should be abandoned, even if it is accompanied by some analysis of error and resolution.

It has long been recognized the impossibility of observing the reservoir either in sufficient detail or with sufficient accuracy to consider the model parameter and initial state[‡] as known with certainty. Thus, much effort has gone into the description of the model parameter using a statistical[§] point of view. One can use geostatistics to determine a single reservoir state which in terms of the available data is the "best" approximation of the unknown true state of the reservoir. It is also possible and probably a better practice to produce a set of equally probable reservoir states consistent with the all the static observation of model parameters using Gaussian random fields. Figure 1 shows a sequence of such probable reservoir states.

[‡] For instance, there is not enough data to properly initialize the model parameters like permeability, porosity, initial phase distribution, PTV data, fluid properties data etc.
[§] Stochastic techniques like sequential Gaussian simulation indicator methods, etc.

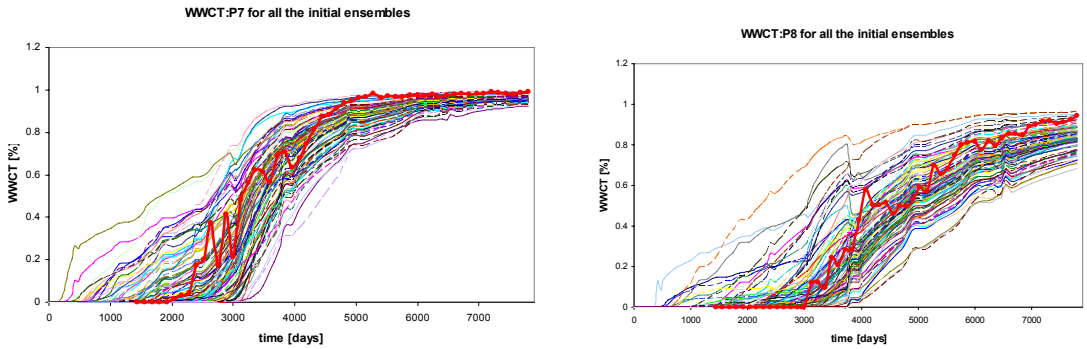


Figure 2. Water cut response from one hundred realization of the same reservoir field shown in figure above. Notice the answer to the question how much will be the water cut after 4 years is not longer a single value but a set of values that can be cast into a probability density function**.**

If we produce a forecast of the production data using the set of reservoir models (e.g see Figure 2) one can not say that the final outcome of any such system is right or wrong. Each reservoir model represents an individual member of an infinite ensemble of reservoir states which are consistent with the initial model parameter measurements. The different analyses will yield different forecast. The best we can do is to build a probability density function, to describe our model state, in a multidimensional space. This clearly shows that since the probabilistic nature of the initial state is unavoidable, then so also is the probabilistic nature of the prediction.

2.2 The Inverse Problem

We have seem that the forward operator $g(\cdot)$ allows us to predict some data measurement **d** if we know some model parameter **m**. Sometimes one is interested in finding the model parameter **m** given some observed measurement or data **d**. Since this problem is the inverse of the forward problem, it is called the *inverse problem*.

While the forward problem (if the physical laws used are deterministic) has a unique solution the inverse problem most of the times does not.

Perhaps the most common inverse problem in petroleum engineering is to infer permeability and skin factor from a well test. In this problem measurement of pressure $\mathbf{d}=\{p_1,p_2,\ldots,p_n\}^T$ are taken during a large enough shut in or drawdown period. It is assumed the pressure at the wellbore can be modeled using the well known radial

solution of the diffusivity equation, This is our forward model **d**=g(**m**), see equation below

$$p_i - p_{wf} = -70.6\frac{qB\mu}{kh}\left[\ln\left(-\frac{1688\phi\mu c_t r_w^2}{kt}\right) - 2s\right] \quad \text{..... (2.3)}$$

If enough information is provided one should be able to easily solve for the model parameters **m**=$\{k,s\}^T$ permeability and skin factor using the set of equations

$$k = \frac{1.151\alpha_p \Delta qB\mu}{h}\frac{1}{dP/d\ln(t)}$$

$$s = 1.151\left[\frac{p_{1hr} - p_{wf}}{-dP/d\ln(t)} - \log_{10}\left(\frac{k}{\phi\mu c_t r_w^2}\right) + 3.23\right] \quad \text{..... (2.4)}$$

Equations in 2.4 are the solution for a very simplified version of an inverse problem. It was solved without considering noise in the observation and without considering any prior information. A naïve interpretation of the inverse problem solution will reveal that the permeability is unique and deterministic. A more detailed analysis reveals that this depend of in the type of parameters we want to infer. For instance if the model parameters interest are now permeability, skin factor and porosity **m**=$\{k,s,\phi\}^T$ then there is not an unique solution for **m**. Further complication can arise if we also want to take into account the intrinsic random noise present in any measurement.

These types of problem are commonly known in the mathematical literature as ill-posed problem. To work around these complications one can then introduce additional elements to the inverse problem theory. For instance to reduce the number of plausible solutions of **m,** one may take advantage of any available prior information on the model parameters. To take care of the data uncertainty a non deterministic representation can be used. Because of the stochastic nature of the measurement or data, perhaps the most general (and simple) way to describe the data is using a probabilistic point of view.

Summarizing, the *inverse problem theory* refers as the family of techniques that allows us to determine a plausible set of values for the model parameter **m**, given an optional prior description of the model parameters $\mathbf{m}_{\text{prior}}$, some inexact observed data $\mathbf{d}_{\text{obs}}$ and an assumed theoretical relationship between the data and the model parameters $\mathbf{d}=g(\mathbf{m})$.

2.3 Elements of the Inverse Theory

We have briefly defined the elements of an inverse problem. It follows that the measurements or data, the prior information on model parameter and the information on the physical correlations between observable parameter and model parameter can all be described using probability density functions PDF. The probabilistic nature of each element will be briefly explained below.

2.3.1 Measurements or observed data

All physical measurements are subject to uncertainties. Therefore, when one reads a measurement from an instrument the result shall not be simply a deterministic value **d**, denoted as $\mathbf{d_{obs}}$ but a probability density function. If one assumes that the uncertainness due to the measuring instrument are independent of the input **d,** then the relationship between the "true (noise-free)" data **d**, and the observed data $\mathbf{d_{obs}}$ can be represented as,

$$\mathbf{d_{obs}} = \mathbf{d} + \boldsymbol{\varepsilon} \quad \text{..... (2.5)}$$

Where $\boldsymbol{\varepsilon}$ is an unknown error with a Gaussian distribution having a zero mean and a covariance matrix $\mathbf{C}_{\text{d}}$. The results of the measurement can be represented by a probability density $\rho_D(\mathbf{d})$

$$\rho_D(\mathbf{d}) = \frac{1}{(2\pi)^{n/2}\left|\mathbf{C_d}\right|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{d}-\mathbf{d_{obs}})^t \mathbf{C_d^{-1}} (\mathbf{d}-\mathbf{d_{obs}})\right] \quad \text{..... (2.6)}$$

2.3.2 Prior information on model parameters

The prior information helps to decrease or better say bound the number of plausible solutions in an inverse problem. By prior information we mean all the information that is obtained independent of the data measurements.

In reservoir engineering, the prior refers to the entire information gathered to build the static model, which may include well logs, core data, outcrops, seismic information, the expert criteria from a geologist, etc. Once the static model is built, it basically contains permeability and porosity distribution. It has long been recognized the impossibility of observing the reservoir either in sufficient detail or with sufficient accuracy to consider the parameters and initial state as known with certainty. Thus, much effort has gone into the description of the model parameters using a statistical point of view. It is possible to produce a set of probable reservoir states consistent with all the static observation of model parameters using Gaussian random fields. This probable set of reservoir models can be cast into a probability density known as the prior PDF. The prior can be represented by a probability density $\rho_M(\mathbf{m})$

$$\rho_M(\mathbf{m}) = \frac{\mathbf{1}}{(2\pi)^{M/2}|\mathbf{C_M}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{m}-\mathbf{m}_{prior})^t \mathbf{C_M^{-1}}(\mathbf{m}-\mathbf{m}_{prior})\right] \quad \ldots\ldots\ldots\ldots\ldots (2.7)$$

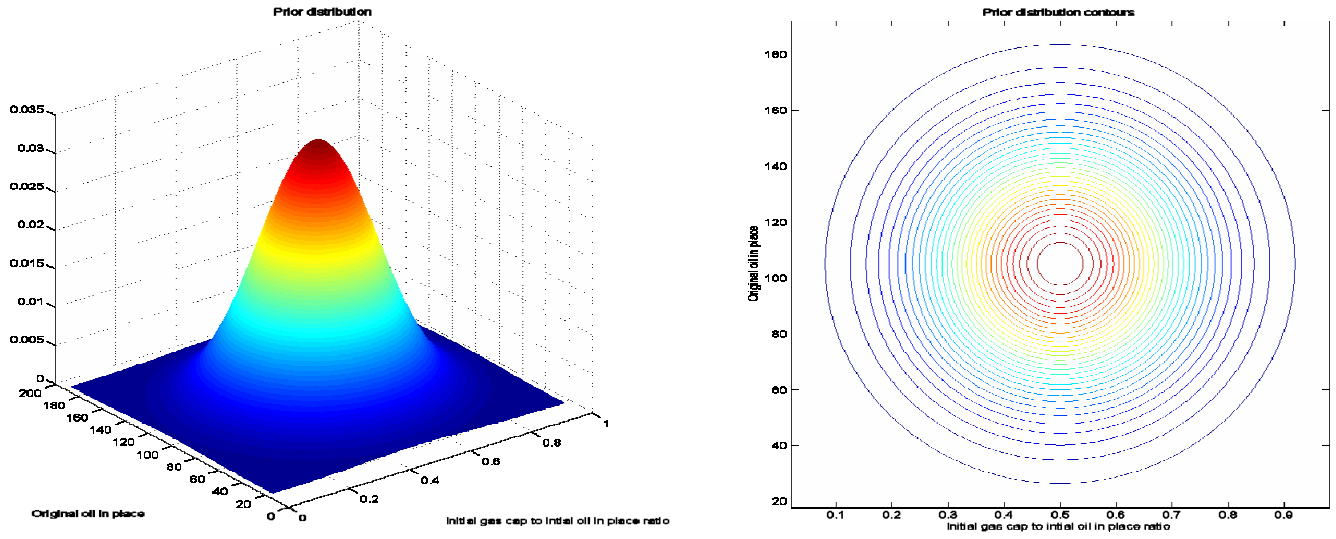


Figure 3. Prior Gaussian joint PDF. There is no clear relationship between N and m from the prior information.

Figure 3 shows the a joint probability distribution for two variables. This example corresponds to the prior information for original oil in place N and the initial gas cap to oil in place ratio m. The peak of the function shows the most probable prior value for N and m.

$$X = \begin{Bmatrix} N_{prior} \\ m_{prior} \end{Bmatrix} = \begin{Bmatrix} 105 \\ 0.5 \end{Bmatrix}$$ where N has units of MMstb.

$$C_x^{prior} = \begin{pmatrix} 900 & 0 \\ 0 & 0.0256 \end{pmatrix}$$

Notice in this particular example, the covariance matrix is diagonal, that is, the prior does not contain information about how the original oil in place is related with the initial gas cap to oil in place ratio. To illustrate the inverse theory, the prior information will be used along with the production data and the Schilthuis material balance equation[31] (MBE) to infer an updated estimates of N and m.

2.3.3 Likelihood function

The predicted values **d**=g(**m**) can not, in general, match the observed values for two reasons: measurement uncertainties and model uncertainties. The first one has to do with the inherent error associated with any measurement device. The second one has to do with the assumptions made during the analysis of the physical laws that describe certain problem.

If we were to have a forward model so good that the model uncertainties can be assumed to be zero, then we should be able to relate the observation to the true model parameters as follows

$$\mathbf{d}_{obs} = g(\mathbf{m}_{true}) + \boldsymbol{\varepsilon} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.8)$$

Equation above states that the observation $\mathbf{d_{obs}}$ are the response of the true model $\mathbf{m_{true}}$; provided that only measurement uncertainties are present. The problem is that we

do not know the true model. But, if we have an estimate of it say **m**, then g(**m**) is an estimate of the of the data and $\mathbf{d_{obs}}$ - g(**m**) is an estimate of **ε**. We assume that the error **ε**, follows a Gaussian ditribution with zero mean variance equal to var(**ε**)=$\mathbf{C}_D$.

$$p_{d|m}(\mathbf{d}\,|\,\mathbf{m}) = \rho(g(\mathbf{m})) = \frac{1}{(2\pi)^{N/2}|\mathbf{C_d}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{g(m)}-\mathbf{d_{obs}})^t \mathbf{C_d^{-1}}(\mathbf{g(m)}-\mathbf{d_{obs}})\right] \quad \ldots\ldots\ldots\ldots\ldots (2.9)$$

Equation 2.9 is known as the *likelihood* function; the likelihood function gives information of how good a model parameter **m** is explaining the observed data $\mathbf{d}_{obs}$.

Continuing with the example of the MBE; Table 1 shows the cumulative production data for a reservoir. It is possible to find a set of {N,m} values that fit the data in the table. Figure 4 shows the corresponding likelihood function for the MBE. In this particular case we have measurements of the mean reservoir pressure and cumulative production data (Table 1). The peak region in the function represents all plausible pairs of values {N,m} that best describe the given observation. Notice that there are several pairs of {N,m} that maximize the likelihood function. In other words, this means, for this type of problem, that the solution of the inverse problem using only the observation is not unique. To reduce the number of plausible solutions we can introduce the prior information.

Table 1. Production and PVT data for a reservoir. It is possible to find m and N to fit this data (from Dake[31]).

Pressure (psi)	Np (MMstb)	Rp (scf/stb)	Bo rb/stb	Rs scf/stb	Bg Rb/scf
3330			1.2511	510	0.00087
3150	3.295	1050	1.2353	477	0.00092
3000	5.903	1060	1.2222	450	0.00096
2850	8.852	1160	1.2122	425	0.00101
2700	11.503	1235	1.2022	401	0.00107
2550	14.513	1265	1.1922	375	0.00113
2400	17.73	1300	1.1822	352	0.00120

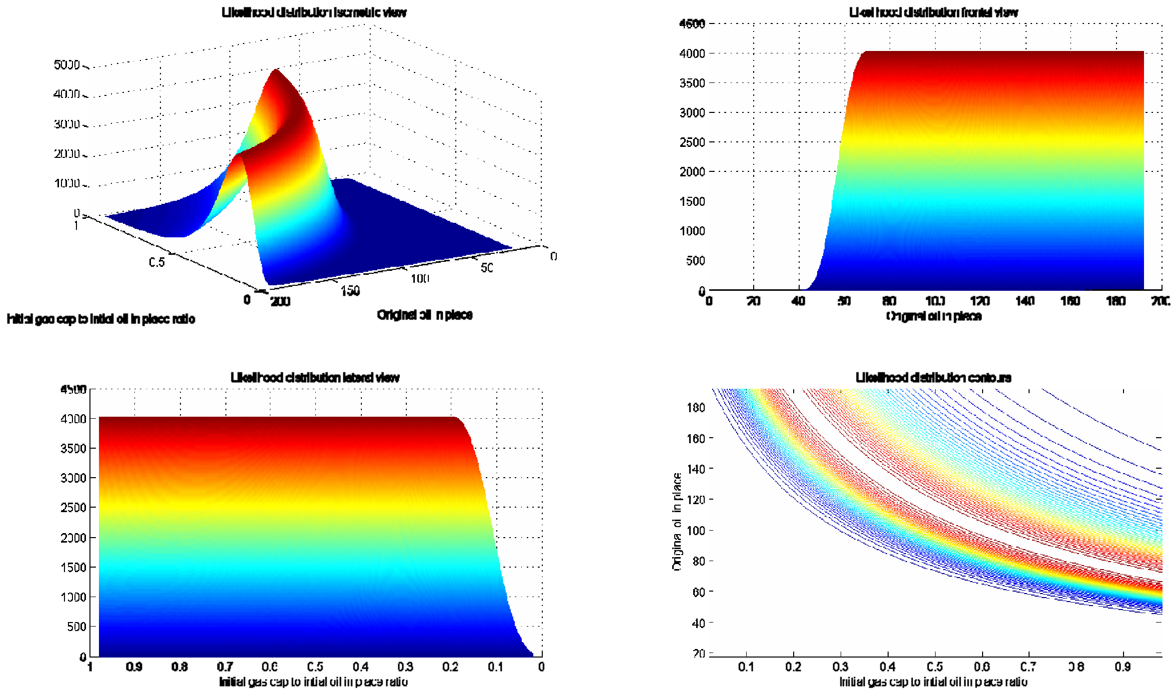


Figure 4. Likelihood function. Notice there is not a single pair of values {N,m} that maximize the function. For this particular likelihood an infinite number of combination of the model parameter {N,m} can satisfy the observation. The prior will help us to bound the proper solution to the inverse problem.

2.3.4 The Bayes theorem and the posterior distribution

The goal of the inverse problem is to combine the data/observation, the likelihood function and the prior information to find a *posterior* estimate of the model parameters. The posterior should honor, to a certain extent, the initial information. This can be archived using the well known Bayes theorem.

$$p_{m|d}(\mathbf{m}\,|\,\mathbf{d}) = \frac{p_{d|m}(\mathbf{d}\,|\,\mathbf{m})\rho_M(\mathbf{m})}{p_d(\mathbf{d})} \quad \text{...} \quad (2.10)$$

For a small size problem it is possible to compute a discrete representation of the posterior using equation 2.10 or 2.11 (see Figure 5 right). Substitution of equations 2.7 and 2.9 into equation 2.10 gives us an expression of the posterior estimate.

$$p_{m|d}(\mathbf{m}|\mathbf{d}) = \frac{1}{cte}\exp\left[-\frac{1}{2}(\mathbf{g(m)}-\mathbf{d_{obs}})^t \mathbf{C_d^{-1}}(\mathbf{g(m)}-\mathbf{d_{obs}})\right]+\left[-\frac{1}{2}(\mathbf{m}-\mathbf{m}_{prior})^t \mathbf{C_M^{-1}}(\mathbf{m}-\mathbf{m}_{prior})\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots (2.11)$$

Where *cte* represents a normalization constant. As an example let us consider again the inverse problem using the MBE. From Figure 5 it is easy to see that information has been gained on the model parameters, because of the production data and the theoretical information on the physics of the problem. In the prior information there was no correlation between N and m. After incorporating the production data and the physics of the problem, we can easily see that when N increases m decreases. This relationship is in agreement with the definition of $m=GB_{gi}/NB_{oi}$. Here *G* represent the initial gas in place Bg and Bo the gas and oil formation volume factors and N the initial oil in place (for more details see Dake[31]).

After a detailed examination of the posterior in Figure 5, it can be inferred that the values that maximize the posterior distribution are given by

$$\begin{Bmatrix} N_{posterior} \\ m_{posterior} \end{Bmatrix} = \begin{Bmatrix} 109.2 \\ 0.517 \end{Bmatrix}$$

The value that maximizes the posterior distribution is commonly known as the *maximum a- posteriori estimate* or MAP estimate.

2.4 Solution of an Inverse Problem

It is clear from the previous discussion that the solution of a very general inverse problem is given by equation 2.11. As expected the solution is not a single value but a complete PDF. The problem is that evaluating 2.11 can be computationally too demanding. Thus, because most of the times one can not utilize the complete PDF, one might be more interested in finding a measurement of central tendency like the MAP and a measurement of dispersion like the covariance. Many times, especially in history matching, the posterior PDF is not Gaussian but it has several modes. Thus finding the mean and the covariance may not be good enough.

To find a solution for the inverse problem one can use basically two techniques: a gradient base technique or a more rigorous one as a Monte Carlo technique.

2.4.1 Gradient based solution

Because most of the times a complete evaluation of equation 2.11 is too expensive; one may be interested in finding the maximum posterior estimate MAP and a measurement of the variance around this maximum. Our goal if to find a peak in the surface given by equation 2.11 and thus, the MAP estimate can be obtained by finding the model parameter **m** that maximizes 2.11.

For the sake of simplicity here we will only consider the case where the relationship between **d** and **m** is linear.

$$\mathbf{d} = \mathbf{Gm} \quad \text{..... (2.12)}$$

Where **G** is known as the sensitivity matrix. After minimizing the exponential part of equation 2.11 with respect to **m** and thus maximizing equation 2.11 the solution for the linear case is given by

$$\tilde{\mathbf{m}} = \mathbf{m}_{prior} + \mathbf{C}_M \mathbf{G}^t (\mathbf{G}\mathbf{C}_M \mathbf{G}^t + \mathbf{C}_d)^{-1} (\mathbf{d}_{obs} - \mathbf{G}\mathbf{m}_{prior}) \quad \text{..... (2.13)}$$

2.4.2 Monte Carlo solution

It is clear from equation 2.11 that the most general solution of an inverse problem provides a probability distribution over the model parameter space. It is only when the posterior probability distribution is very simple (for instance, is has only one maximum) that an analytical technique can be used to characterize it.

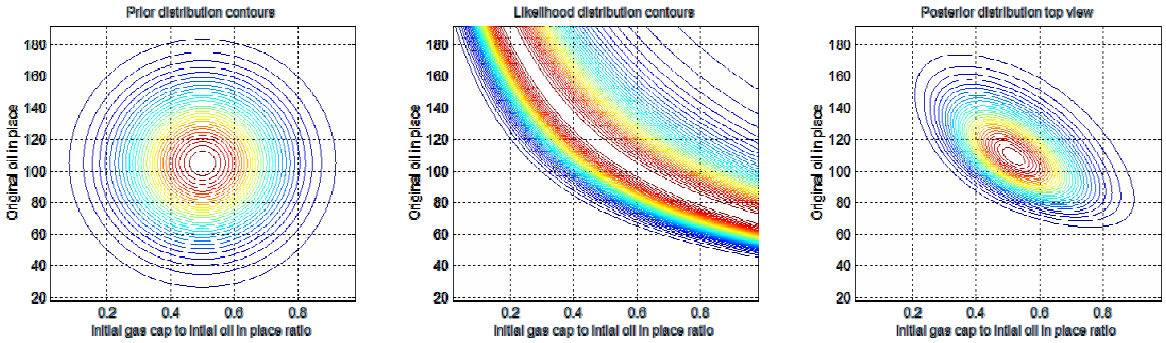


Figure 5. The probabilities density $P_x(x)$ and $P_L(d|x)$ respectively represent the information on the prior model parameters (left) and the likelihood P.D.F (center) explaining how good a model x is explaining the observed data d_{obs}. Given the two states of information represented by $P_x(x)$ and $P_L(d|x)$ the posterior PDF (right) represent the combination of the two states of information.

For a more general and complete solution, one needs to perform an extensive exploration of the model parameter space. Monte Carlo solution refers to an approach that uses random sequences as input. The basic idea is to try to get an estimate of the complete posterior PDF by sampling from it. Thus a continuous PDF is approximated by a discrete description of it. Some examples of Monte Carlo solution techniques are:

1. Simulated annealing
2. Markov Chain Monte Carlo
3. Ensemble Kalman filter

A detailed description of each one of these techniques is out of the scope of the present thesis. Only the Ensemble Kalman filter will be discussed in detail.

CHAPTER III

SOLUTION OF INVERSE PROBLEM USING KALMAN FILTERS

Reader may be wondering how a filter is related with the inverse problem theory and how a filter is related with history matching. This relation will be clarified in this chapter by starting with some basic definition to will clarify the relation between the inverse theory and filters.

3.1 What Is a Filter and How Is It Connected to Inverse Problems

Consider a signal y(t) which is a real value function of the continuous time variable t. Suppose that there is another signal d(t) that is generated from y(t). That is in mathematical terms

$$d(t) = g(y(t)) + \varepsilon(t) \quad \ldots\ldots (3.1)$$

where ε(t) is a noise or disturbance term. These types of relationship given by eqaution 3.1 arise in many applications. For example, in a communication system, one can think of the signal y(t) as the electromagnetic waves arriving to an AM radio and the other signal d(t) as the pressure waves that make the sound we hear when listening to the radio. In other applications, d(t) may be a measurement of the signal y(t) obtained from a sensor as illustrated in figure 6. Notice the similarity between equation 3.1 and our forward model equation 2.1.

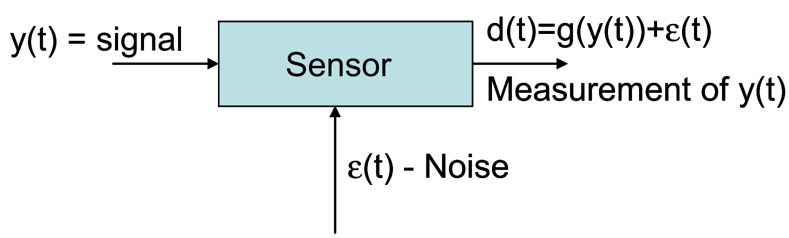


Figure 6. Sensor which provides a measurement.

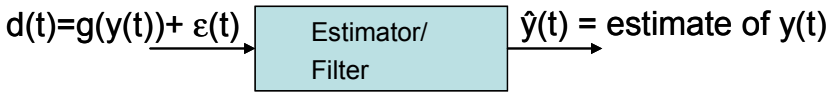


Figure 7. Signal reconstruction using an estimator/filter.

The determination of y(t) from d(t) is a type of filtering problem. In particular a device which produces an estimate $\tilde{y}(t)$ of y(t) is called an estimator or *filter*. Such a device is illustrated in Figure 7

Notice that the idea of reconstruction of the signal y(t) from a measurement d(t) has a resemblance to the inverse problem discussed in the previous chapter. Thus, from this point of view a filtering problem is basically an inverse problem. Our purpose is then to use the mathematical formulation proposed in filter theory known as a Kalman filter to estimate our permeability based on measurements from the field. We will also see later in this chapter that the Kalman filter equations are actually the same equations discussed earlier under a Bayesian framework.

3.2 The Kalman Filter Equations

We start by assuming our forward model can be cast in the form

$$\mathbf{y}_{k+1} = g(\mathbf{y}_k) + \mathbf{v}_k \quad \text{..... (3.2)}$$

Where $\mathbf{y}_k$ represents the model at time k and is the known as the *state vector* and $\mathbf{v}_k$ is a Gaussian noise with zero mean and covariance matrix $\mathbf{Q}_k$.

Here it may be worth to notice the small differences in the variables used in the equation 3.2 compared with equation 2.1. Notice in previous chapter we use **m** to design the model parameter while here we use **y**. The reason for such change is that in the Kalman filter theory the state vector is augmented to include the model parameters, as well as the data/measurement. Thus, a detailed description of the state vector variable **y** is as follows. The joint model-observation state vector can include three types of parameters. For the reservoir case, these types are the static variables, $\mathbf{m}_k^s$ (e.g permeabilities, porosity), dynamic variables varying with time $\mathbf{m}_k^d$ (e.g pressure, phase saturations) and the production data, usually measured at the wells, $\mathbf{d}_k$ (e.g bottom-hole

pressure, well production rates, water cuts). Also notice that the effect of the forward operator *g(·)* over $\mathbf{m}_k^s$ is null; that is, noting happens to $\mathbf{m}_k^s$ when moving from time *k* to time *k+1*.

$$\mathbf{y}_k = \begin{Bmatrix} \mathbf{m}_\mathbf{k}^\mathbf{s} \\ \mathbf{m}_\mathbf{k}^\mathbf{d} \\ \mathbf{d}_\mathbf{k} \end{Bmatrix} \quad \text{..... (3.3)}$$

The state vector at time *k* is defined by equation (3.3) where superscript index *p* denotes prior, *s* stands for static, and *d* stands for dynamic.

The observation of the process is assumed to occur at discrete points in time *k* in accordance with the linear relationship

$$\mathbf{d}_k = \mathbf{H}\mathbf{y}_k + \boldsymbol{\varepsilon}_k \quad \text{..... (3.4)}$$

Where again $\mathbf{d}_k$ represents the data $\mathbf{y}_k$ represents the state vector and $\boldsymbol{\varepsilon}_k$ represents the noise in the observation at time *k*. This noise is assumed to be Gaussian with zero mean and a covariance matrix $\mathbf{C_d}$. The basic function of multiplying by matrix **H** is to select the rows in $\mathbf{y}_k$ corresponding to the calculated production data $\mathbf{d}_k$. Matrix **H** is given by equation 3.5 where I is simply the identity matrix.

$$\mathbf{H} = [0 \quad \mathrm{I}] \quad \text{..... (3.5)}$$

One can assume that we have an initial estimate of the process at some point in time t_k, and that this estimate is based on all our knowledge about the process prior to t_k. This prior estimate will be denoted as $\mathbf{y}_{prior}$. It is also common to assume that one can know the covariance matrix associated with the prior, $\mathbf{C_y}$. If $\mathbf{y}_k$ represents the true state vector, one can define the estimated error and its associated error covariance matrix to be

$$\mathbf{e_k} = \mathbf{y_k} - \mathbf{y_{k,prior}} \quad \text{..... (3.6)}$$

$$\mathbf{C_y} = E[\mathbf{e_k e_k^t}] = E[(\mathbf{y_k} - \mathbf{y_{k,prior}})(\mathbf{y_k} - \mathbf{y_{k,prior}})^t] \quad \text{.................................. (3.7)}$$

It is also worth to notice that since the state vector $\mathbf{y}_k$ contains static and dynamic variables, its covariance matrix should contain equivalent information. A closer look inside the covariance matrix of the state vector reveals

$$\mathbf{C_y} = \begin{pmatrix} \mathbf{C}_{M^s} & \mathbf{C}_{M^s,M^d} & \mathbf{C}_{M^s,d} \\ \mathbf{C}^t_{M^d,d} & \mathbf{C}_{M^d} & \mathbf{C}_{M^d,d} \\ \mathbf{C}^t_{M^s,d} & \mathbf{C}^t_{M^d,d} & \mathbf{C}^{cal}_{d} \end{pmatrix} \quad \text{.. (3.8)}$$

where $\mathbf{C}_{M^s}$ is the covariance matrix of the static variables (in our case permeability, porosity), for the sake of simplicity let's consider only one type of static variable, the size of this matrix being MxM where M is the number of grid blocks, matrix $\mathbf{C}_{M^d}$ is the covariance matrix of the dynamic variables (i.e phase saturation and pressure) again for simplicity let's consider only one dynamic variable, the size of the matrix being MxM where M again is the number of grid block, $\mathbf{C}_d$ is the covariance matrix of the calculated data, the size of the matrix being NxN where N is the number of measurements at a given time (bottom hole pressure, oil production rate, water cut, etc). The non–diagonal elements in the matrix like $\mathbf{C}_{M^s,d}$ and $\mathbf{C}_{M^d,d}$ are the cross-covariance matrix between the static variable and the data (i.e permeability and water cut, permeability and bottom hole pressure)

We now seek to use the measurement $\mathbf{d}_{k,obs}$ to improve the prior estimate. We choose a linear blending of the noisy measurements and the prior estimate in accordance with the equation

$$\tilde{\mathbf{y}}_k = \mathbf{y}_{k,prior} + \mathbf{K}_k(\mathbf{d}_{k,obs} - \mathbf{H}\mathbf{y}_{k,prior}) \quad \text{... (3.9)}$$

Where $\tilde{\mathbf{y}}_k$ is the posterior or updated estimate of the state vector and $\mathbf{K}_k$ is a blending factor (yet to be determined) known as the *Kalman gain*. Two different approaches can be used to determine the Kalman gain and both approaches are equivalent. With the aim of proving a proof that both are equivalent we first will show the Kalman filter using the classical approach (commonly used in optimal control literature) and then the Bayesian approach (this is the same approach used in the previous chapter)

3.2.1 Minimum variance estimator - MVE

The problem now is to find a particular $\mathbf{K}_k$ that yields an updated estimate $\tilde{\mathbf{y}}_k$ that is optimal in some sense. With this goal, we then form an expression for the error covariance matrix associated with the updated (posterior) estimate

$$\tilde{\mathbf{C}}_{\mathbf{y}} = E[\tilde{\mathbf{e}}_{\mathbf{k}}\tilde{\mathbf{e}}_{\mathbf{k}}^{\mathbf{t}}] = E[(\mathbf{y}_{\mathbf{k}} - \tilde{\mathbf{y}}_{\mathbf{k}})(\mathbf{y}_{\mathbf{k}} - \tilde{\mathbf{y}}_{\mathbf{k}})^t] \quad \text{..... (3.10)}$$

If we substitute the updated $\tilde{\mathbf{y}}_k$ by its equivalent expression (equation 3.9) into equation 3.10, the result is

$$\tilde{\mathbf{C}}_{\mathbf{y}} = E[(\mathbf{y}_{\mathbf{k}} - \mathbf{y}_{k,prior} - \mathbf{K}_k(\mathbf{d}_{k,obs} - \mathbf{H}\mathbf{y}_{k,prior}))(\mathbf{y}_{\mathbf{k}} - \mathbf{y}_{k,prior} - \mathbf{K}_k(\mathbf{d}_{k,obs} - \mathbf{H}\mathbf{y}_{k,prior}))^t] \quad \text{... (3.11)}$$

Recall from equation 3.6 that ($\mathbf{y}_k$-$\mathbf{y}_{k,prior}$) is the prior estimation error $\mathbf{e}_k$, that is uncorrelated with the measurement error $\boldsymbol{\varepsilon}_k$. Also the observation need to be treat as random variables which means that they can be expressed as $d_{k,obs}$=$\mathbf{H}\mathbf{y}_k$+$\boldsymbol{\varepsilon}_k$. Substiting $\mathbf{e}_k$ and $d_{k,obs}$ into 3.11 results

$$\tilde{\mathbf{C}}_{\mathbf{y}} = E[(\mathbf{e}_k - \mathbf{K}_k(\mathbf{H}\mathbf{e}_k - \boldsymbol{\varepsilon}_k))(\mathbf{e}_k - \mathbf{K}_k(\mathbf{H}\mathbf{e}_k - \boldsymbol{\varepsilon}_k))^t] \quad \text{..... (3.12)}$$

To continue with our proof, here we need to use a couple of very simple matrix rules

$$(AB)^t = B^t A^t$$
$$\frac{d[trace(AB)]}{dA} = B^t \qquad \text{..... (3.13)}$$
$$\frac{d[trace(ACA^t)]}{dA} = 2AC^t$$

Using rule one from equation 3.13 and noting taht $\boldsymbol{\varepsilon}_k$ is uncorrelated with $\mathbf{e}_k$, also the covariance matrices $\mathbf{C}_d$ and $\mathbf{C}_y$ are defined as $\mathbf{C}_d=E[\boldsymbol{\varepsilon}_k\boldsymbol{\varepsilon}_k^t]$ and $\mathbf{C}_y=E[\mathbf{e}_k\mathbf{e}_k^t]$ equation 3.12 can be rewritten

$$\widetilde{\mathbf{C}}_y = \mathbf{C}_y - \mathbf{K}_k\mathbf{H}\mathbf{C}_y - \mathbf{C}_y\mathbf{H}_k^t\mathbf{K}_k^t + \mathbf{K}_k(\mathbf{H}\mathbf{C}_y\mathbf{H}^t + \mathbf{C}_d)\mathbf{K}_k^t \qquad \text{..... (3.14)}$$

Because we want to find an optimal $\mathbf{K}_k$, we can use straightforward calculus and minimize the individual terms along the major diagonal of $\widetilde{\mathbf{C}}_y$, as these terms represent the estimation error variance for the elements of the posterior state vector.

$$\frac{d[trace\,\widetilde{\mathbf{C}}_y]}{d\mathbf{K}_k} = -2\left(\mathbf{H}\mathbf{C}_y\right)^t + 2\mathbf{K}_k(\mathbf{H}\mathbf{C}_y\mathbf{H}^t + \mathbf{C}_d) \qquad \text{..... (3.15)}$$

Now we set the derivative in equation 3.15 equal to zero and solve for the optimal Kalman gain $\mathbf{K}_k$.

$$\mathbf{K}_k = \mathbf{C}_y\mathbf{H}^t(\mathbf{H}\mathbf{C}_y\mathbf{H}^t + \mathbf{C}_d) \qquad \text{..... (3.16)}$$

This particular $\mathbf{K}_k$ minimizes the mean square estimation error and is called the *Kalman gain*. The posterior estimated $\tilde{\mathbf{y}}$ obtained by equation 3.17, given below, is known as the *minimum variance estimator* or MVE. We have now means of assimilating the measurements to updated the state vector. The posterior estimate and the posterior covariance may now be computed using 3.17

$$\tilde{\mathbf{y}} = \mathbf{y}_{prior} + \mathbf{C}_y\mathbf{H}^t(\mathbf{H}\mathbf{C}_y\mathbf{H}^t + \mathbf{C}_d)^{-1}(\mathbf{d}_{obs} - \mathbf{H}\mathbf{y}_{prior}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.17)$$

$$\tilde{\mathbf{C}}_y = \mathbf{C}_y - \mathbf{C}_y\mathbf{H}_k^t(\mathbf{H}\mathbf{C}_y\mathbf{H}^t + \mathbf{C}_d)^{-1}\mathbf{H}_k^t\mathbf{C}_y$$

Some comments on equation 3.17 are in order. Notice that it resembles in great detail the equation 2.13. Which suggest that the MVE is equivalent to the MAP estimate. In fact we will prove that both equations are equivalent. It is also interesting to notice that this derivation has a very close resemblance with the minimum variance estimator found in the Kriging equation. This suggests that the Kalman equation can be deduced from the Kriging equation as well.

If we want to compare equation 3.17 with equation 2.13 it is necessary to rewrite equation 3.17 such that each one of the disguised elements in equation 3.13 is explicitly shown. Let us start by analyzing the terms $\mathbf{C}_y\mathbf{H}^t$ and $\mathbf{H}\mathbf{C}_y\mathbf{H}^t$. Recall that the state vector matrix $\mathbf{C}_y$ is composed of a set of sub-matrices as shown in equation 3.8. The complete $\mathbf{C}_y$ matrix is a very large matrix. Fortunately this matrix always appears multiplied times matrix measurement matrix $\mathbf{H}$. Thus, in practice we never need to compute the whole matrix $\mathbf{C}_y$. Equation 3.18 and 3.19 show in detail the sub-matrices in $\mathbf{C}_y$ that we require to compute. Notice that in practice we never need to compute the sub matrices $\mathbf{C}_{M^s}$, $\mathbf{C}_{M^s,M^d}$

$$\mathbf{C}_y\mathbf{H}^t = \begin{pmatrix} \mathbf{C}_{M^s} & \mathbf{C}_{M^s,M^d} & \mathbf{C}_{M^s,d} \\ \mathbf{C}^t_{M^d,d} & \mathbf{C}_{M^d} & \mathbf{C}_{M^d,d} \\ \mathbf{C}^t_{M^s,d} & \mathbf{C}^t_{M^d,d} & \mathbf{C}^{cal}_{d} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{M^s,d} \\ \mathbf{C}_{M^d,d} \\ \mathbf{C}^{cal}_{d} \end{pmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.18)$$

$$\mathbf{H}\mathbf{C}_y\mathbf{H}^t = \left(\mathbf{C}^{cal}_{d}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.19)$$

If we substitute the terms $\mathbf{C}_y\mathbf{H}^t$ and $\mathbf{H}\mathbf{C}_y\mathbf{H}^t$ into equation 3.17 and express $\mathbf{y}_k$ in a vector form we can rewrite the equation as

$$\tilde{\mathbf{y}}_k = \begin{Bmatrix} \tilde{\mathbf{m}}_{\mathbf{k}}^{\mathbf{s}} \\ \tilde{\mathbf{m}}_{\mathbf{k}}^{\mathbf{d}} \\ \tilde{\mathbf{d}}_{\mathbf{k}} \end{Bmatrix} = \begin{Bmatrix} \mathbf{m}_{\mathbf{k}}^{\mathbf{s}} \\ \mathbf{m}_{\mathbf{k}}^{\mathbf{d}} \\ \mathbf{d}_{\mathbf{k}} \end{Bmatrix} + \begin{pmatrix} \mathbf{C}_{M^s,d} \\ \mathbf{C}_{M^d,d} \\ \mathbf{C}_{d}^{cal} \end{pmatrix} \left(\mathbf{C}_{d}^{cal} + \mathbf{C}_d\right)^{-1} \left(\mathbf{d}_{k,obs} - \mathbf{H}g(\mathbf{y}_{k,prior})\right) \quad \text{..... (3.20)}$$

Our goal is to understand how equation 3.20 is related to equation 2.13. We move our attention to the static variables $\mathbf{m}_k^s$ only. Then we get an expression closer to the posterior estimate using the Bayesian approach discussed in the previous chapter. For the sake of clarity equation 2.13 is shown here (see equation 3.22)

$$\tilde{\mathbf{m}} = \mathbf{m}_{prior} + \mathbf{C}_{M^s,d}(\mathbf{C}_{d}^{cal} + \mathbf{C}_d)^{-1}(\mathbf{d}_{obs} - \mathbf{H}g(\mathbf{y}_{k,prior})) \quad \text{..... (3.21)}$$

$$\tilde{\mathbf{m}} = \mathbf{m}_{prior} + \mathbf{C}_M\mathbf{G}^t(\mathbf{G}\mathbf{C}_M\mathbf{G}^t + \mathbf{C}_d)^{-1}(\mathbf{d}_{obs} - \mathbf{G}\mathbf{m}_{prior}) \quad \text{..... (3.22)}$$

While equation 3.21 shows the MVE for the static variable, equation 3.22 shows the MAP estimate from the Bayesian framework. Notice the similarities between both these equations. By comparing equation 3.21 to equation 3.22, we notice that the cross-covariance matrix between the static variable and the data (i.e permeability and water cut, permeability and bottom hole pressure) $\mathbf{C}_{M^s,d}$ is actually approximately equivalent to the multiplication of the covariance matrix of the prior times the sensitivity matrix $\mathbf{C}_M\mathbf{G}$. Specifically we get the following,

$$\begin{aligned} \mathbf{C}_M\mathbf{G}^t &\approx \mathbf{C}_{M^s,d} \\ \mathbf{G}\mathbf{C}_M\mathbf{G}^t &\approx \mathbf{C}_d^{cal} \end{aligned} \quad \text{..... (3.23)}$$

3.2.2 Maximum posteriori estimate - MAP

In the previous proof our model parameter was substituted by the concept of state of information or state vector. The state vector describes the joint state of the static model parameter $\mathbf{m}^s$, the dynamic model parameters $\mathbf{m}^d$ and the calculated observation $\mathbf{d}_{cal}$. In the previous chapter we have seen that each one of the elements in the state vector can be treated using a Gaussian probability distribution, then this concept can be extended and

we can describe the complete vector using a probability distribution too. If we replace the complete state vector **y** in place of the model parameter **m** in equation 2.11, we get equation 3.24.

$$p_{m|d}(\mathbf{y}\,|\,\mathbf{d}) = \frac{1}{cte}\exp^{\left[-\frac{1}{2}(\mathbf{Hy_k}-\mathbf{d_{obs}})^t\,\mathbf{C_d^{-1}}\,(\mathbf{Hy_k}-\mathbf{d_{obs}})\right]+\left[-\frac{1}{2}(\mathbf{y}-\mathbf{y}_{prior})^t\,\mathbf{C}_y^{-1}(\mathbf{y}-\mathbf{y}_{prior})\right]} \quad \ldots\ldots\ldots\ldots (3.24)$$

Where *cte* represents a normalization constant. The exponential part of equation 3.24 can be isolated in a function S($\mathbf{y}_k$)

$$2S(\mathbf{y}_k) = (\mathbf{Hy_k}-\mathbf{d_{obs}})^t\,\mathbf{C_d^{-1}}\,(\mathbf{Hy_k}-\mathbf{d_{obs}}) + (\mathbf{y}-\mathbf{y}_{prior})^t\,\mathbf{C}_y^{-1}(\mathbf{y}-\mathbf{y}_{prior}) \quad \ldots\ldots\ldots (3.25)$$

Our goal is to find the MAP by maximizing equation 3.24. This is equivalent to minimizing the objective function S(**y**) in equation 3.25. Finding the minimum is equivalent to find a value for **y** that makes the derivative of the objective function equal to zero $\partial S/\partial \mathbf{y}_k = 0$. Equations 3.26 to 3.28 show some of the steps in the minimization of 3.24

$$\frac{\partial S}{\partial \mathbf{y}_k} = \mathbf{H}^t\mathbf{C}_d^{-1}(\mathbf{Hy}_k - \mathbf{d}_{obs}) + \mathbf{C}_y^{-1}(\mathbf{y}_k - \mathbf{y}_{prior}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.26)$$

$$(\mathbf{H}^t\mathbf{C}_d^{-1}\mathbf{H} + \mathbf{C}_y^{-1})y_k = \mathbf{H}\mathbf{C}_d^{-1}\mathbf{d}_{obs} + \mathbf{C}_y^{-1}\mathbf{y}_{prior} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.27)$$

$$\tilde{\mathbf{y}} = \mathbf{y}_k = (\mathbf{H}^t\mathbf{C}_d^{-1}\mathbf{H} + \mathbf{C}_y^{-1})^{-1}\mathbf{H}^t\mathbf{C}_d^{-1}\mathbf{d}_{obs} + (\mathbf{H}^t\mathbf{C}_d^{-1}\mathbf{H} + \mathbf{C}_y^{-1})^{-1}\mathbf{C}_y^{-1}\mathbf{y}_{prior} \quad \ldots\ldots\ldots (3.28)$$

If one uses the following two identities[**] (equation 3.29) one can rewrite equation 3.28 into a more familiar form as shown in equation 3.30.

[**] See Tarantola[18] 2005 page 250 for a detailed proof of the identities.

$$(\mathbf{H}^t\mathbf{C}_d^{-1}\mathbf{H}+\mathbf{C}_y^{-1})^{-1}\mathbf{H}^t\mathbf{C}_d^{-1}=\mathbf{C}_y\mathbf{H}^t(\mathbf{C}_d+\mathbf{H}\mathbf{C}_y\mathbf{H}^t)^{-1}$$
$$(\mathbf{H}^t\mathbf{C}_d^{-1}\mathbf{H}+\mathbf{C}_y^{-1})^{-1}=\mathbf{C}_y-\mathbf{C}_y\mathbf{H}^t(\mathbf{C}_d+\mathbf{H}\mathbf{C}_y\mathbf{H}^t)^{-1}\mathbf{H}\mathbf{C}_y \quad \text{………………….. (3.29)}$$

$$\tilde{\mathbf{y}}=\mathbf{y}_{prior}+\mathbf{C}_y\mathbf{H}^t(\mathbf{H}\mathbf{C}_y\mathbf{H}^t+\mathbf{C}_d)^{-1}(\mathbf{d}_{obs}-\mathbf{H}\mathbf{y}_{prior}) \quad \text{………………………. (3.30)}$$

Notice that our MAP estimate (equation 3.30) gives the same expression that we found using the MVE (equation 3.17)

Notice that almost all elements in equation 3.30 are known. $\mathbf{y}_{\text{prior}}$ is known from geostatistical modeling; $\mathbf{C}_{\text{d}}$ is known from the measurement device specifications; $\mathbf{d}_{\text{obs}}$ is known from production data; $\mathbf{Hy}_{\text{prior}}$ by running the simulator given a prior. The only unknown element in equation 3.30 in the state vector covariance matrix $\mathbf{C}_{\text{y}}$. The next section deals with a Monte Carlo technique that allows us to compute $\mathbf{C}_{\text{y.}}$

3.3 The Ensemble Kalman Filter Applied to Reservoir Engineering

So far we have seen that in general, it is always possible to find the posterior estimate $\tilde{\mathbf{y}}$ if we known the state vector covariance matrix $\mathbf{C}_{\text{y.}}$. Unfortunately it is typically unknown and there is not a simple way to compute $\mathbf{C}_{\text{y.}}$. This limitation restricted the use of Kalman filter to be used primary for linear problems. It was not until 1994 that Evensen proposed the use of a Monte Carlo technique to describe the PDF and then estimate $\mathbf{C}_{\text{y}}$ at any time.

To work around the problem of knowing the state vector covariance matrix at any point in time, the ensemble Kalman filter (EnKF) uses an ensemble of state vectors instead of a single state vector. The statistics (mean and covariance) are then computed from the ensemble. The ensemble of state vectors can be represented by the equation 3.31.

$$\boldsymbol{\Psi}_k^p=\left\{y_{k,1}^p \quad y_{k,2}^p \quad \dots \quad y_{k,N_e}^p\right\} \quad \text{……………………………………………... (3.31)}$$

Each state vector represents an individual member of an infinite ensemble of possible states that are consistent with initial measurement from core, well logs and

seismic. Each member can be generated using any of the well known geostatistical techniques like sequential Gaussian simulation or indicator simulation. Now since each model produces a different forecast and we do not know a-priori which one will be close to the true, if any, the best we can do is to construct a probability density from the ensemble.

3.3.1 The forecast step and time update step

Kalman filters have two main steps: a forecast/time step and an update/analysis step. In this particular research the forecast step is carried out by a commercial reservoir streamline simulator. This action can be represented as in our equation notation as

$$\begin{Bmatrix} \mathbf{m}_k^d \\ \mathbf{d}_k \end{Bmatrix} = g(\mathbf{m}^s, \mathbf{m}_{k-1}^d) \quad \text{(3.32)}$$

Where the forward operator $g(\cdot)$ represent a numerical solution of the porous media fluid flow equations moving forward from time step k-1 to the time step k. In the previous section we proof that the optimal Kalman update equation in matrix from is given by equation 3.33

$$\mathbf{\Psi}_k^u = \mathbf{\Psi}_k^p + \mathbf{K}\left(\mathbf{D}_k - \mathbf{H}\mathbf{\Psi}_k^p\right) \quad \text{(3.33)}$$

The superscript u denotes updated and p denotes prior. Here the matrix **K** is known as the Kalman gain, matrix **D** represents an ensemble of perturbed observations as defined by equations 3.36 and 3.37 below. The basic function of multiplying by matrix **H** is to select rows in $\mathbf{\Psi}_k^p$ corresponding to the calculated production data $\mathbf{d}_k$. Matrix **H** is given by equation 3.34 where **I** is simply the identity matrix.

$$\mathbf{H} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \quad \text{(3.34)}$$

The Kalman gain matrix is given by equation 3.35 as follows

$$\mathbf{K} = \mathbf{C}_{\Psi}^{p}\mathbf{H}^{T}\left(\mathbf{H}\mathbf{C}_{\Psi}^{p}\mathbf{H}^{T} + \mathbf{C}_{D}\right)^{-1} \quad \text{(3.35)}$$

Where $\mathbf{C}_{\Psi}^{p}$ represents an estimate of the state vector covariance matrix $\mathbf{C}_{y,}$ and $\mathbf{C}_{D}$ represents the observation covariance matrix. Typically we assume the errors in the observation are not correlated and therefore, $\mathbf{C}_{D}$ is a diagonal matrix. The ensemble of perturbed observations can be represented as follows

$$\mathbf{D}_{k} = \left\{\mathbf{d}_{k,1} \quad \mathbf{d}_{k,2} \quad \ldots \quad \mathbf{d}_{k,N_e}\right\} \quad \text{(3.36)}$$

$$\mathbf{d}_{k,i} = \mathbf{d}_{k} + \boldsymbol{\varepsilon}_{i} \quad \text{(3.37)}$$

Where $\mathbf{d}_k$ represent a vector of any type of production data measured at time k, and $\boldsymbol{\varepsilon}_i$ represent the noise in the observation for each member i. The noise is assumed to be normally distributed with zero mean and a covariance given by $\mathbf{C}_D$

Because the true state vector is unknown, we approximate it with the mean of the ensemble using equation (3.38) below. The covariance matrix C_{Ψ}^{p} can be approximated at any point in time using equation (3.39). Because in the Kalman gain equation this covariance always appears multiplied by the matrix **H,** in practice there is no need to compute the whole covariance matrix but only a small portion of it.

$$\bar{y}^{p} = \frac{1}{N_e}\sum_{i=1}^{N_e} y_{i}^{p} \quad \text{(3.38)}$$

$$C_{\Psi}^{p} = \frac{1}{N_e - 1}\sum_{i,j=1}^{N_e}(y_{i}^{p} - \bar{y}^{p})(y_{j}^{p} - \bar{y}^{p})^{T} \quad \text{(3.39)}$$

Furthermore in the new approach proposed here, we will constrain the values in the cross-covariance matrix C_{Ψ}^{p} using streamlines trajectories. To account for the conditioning using streamline the covariance matrix is redefined as

$$\mathbf{HC}_{\Psi}^{p} = \rho \circ \left(\frac{1}{N_e - 1} \sum_{i,j=1}^{N_e} (\mathbf{y}_i^p - \overline{\mathbf{y}}^p)(\mathbf{H}\mathbf{y}_j^p - \mathbf{H}\overline{\mathbf{y}}^p)^T \right) \quad \text{..... (3.40)}$$

Where ρ is a correlation function representing the flow path information extracted from the streamlines. The details will be given in the next chapter. The operation ρ° in equation 3.40 denotes the Schur product operator. Recall that the Hadamard or Schur product of two matrices ***A*** and ***B*** with the same size, denoted by ***(A∘B),*** is defined as ***(A∘B)*** = $(a_{ij}b_{ij})$ for each element of the matrices. More detailed description of the correlation function ρ and the benefits of using it will be given in next sections.

3.4 Problems with the Assimilation of Different Types of Observation

The traditional way of calculating the Kalman gain would involve the computation of the eigenvalue decomposition directly from the NxN (were N is the number of available observations at a given time) matrix ($\mathbf{C}_D + \mathbf{H}\mathbf{C}_y\mathbf{H}^t$); see Evensen[14] (2004) for a more detailed description of these operations. The pseudo inversion of this matrix has some scalability issues. It is important to note that when different measurement types (i.e bottom hole pressure, and water cut) are assimilated simultaneously, the calculated and observed data need to be scaled to ensure similar variability. This is required to ensure that the eigenvalues corresponding to each of the measurement types have similar magnitude. Unfortunately, scaling the calculated data is not always possible because we do not know the true variance of the data. The standard approach for resolving this is to assimilate different measurement types, which normally have uncorrelated errors, sequentially one dataset at a time. This ensures that the results are not affected by poor scaling, which in the worst case may result in the truncation of all eigenvalues corresponding to measurements of one kind. Details of how to compute and implement a sequential assimilation algorithm for a Kalman filter is discussed by Anderson, B.D., Moore, J.B[24].

What follows is a brief description of the algorithm used to assimilate different dataset types.

3.4.1 Sequential processing

Sequential processing is the name given to the procedure in which the observation vector is processed one component type at the time (i.e. for our particular case one component type could be water cut or bottom hole pressure)

It is possible to partition the observation vector $\mathbf{d}_{k,obs}$ into sub vectors

$$\mathbf{d}_{k,obs} = \left\{\mathbf{d}^1_{k,obs}, \mathbf{d}^2_{k,obs}, \cdots, \mathbf{d}^I_{k,obs}\right\} \quad \text{..... (3.41)}$$

$$\mathbf{C}^i_D = E[(\boldsymbol{\varepsilon}^i_k)(\boldsymbol{\varepsilon}^j_k)^t] = \mathbf{C}_D \delta_{ij} \quad \text{for i,j=1,2,…,I ..... (3.42)}$$

Where the super index *i* in the **d** vector correspond to the type of observation; (i.e water cut or bottom hole pressure, etc). The implication of equation 3.42 is that and the necessary assumption of the sequential processing is that $E[(\boldsymbol{\varepsilon}^i_k)(\boldsymbol{\varepsilon}^j_k)^t] = 0$ for i≠j or in simple words there should not be cross-correlation or dependence between each measurement type (i.e water cut, bottom hole pressure).

Since the observations were partitioned, then it is also required to partition the measurement matrix $\mathbf{H}=\{\mathbf{H}^1, \mathbf{H}^2, \ldots, \mathbf{H}^I\}$. This allows us to re-write the measurement equation 3.4 and it's covariance matrix as

$$\mathbf{d}^i_k = \mathbf{H}^i \mathbf{y}_k + \boldsymbol{\varepsilon}^i_k \quad \text{..... (3.43)}$$

We can now define more precisely the notion of sequential processing of the observation vector. Instead of processing $\mathbf{d}_{k,obs}$ as a single data vector, the components $\mathbf{d}^1_{k,obs}, \mathbf{d}^2_{k,obs}, \cdots, \mathbf{d}^I_{k,obs}$ are processed one at a time, or sequentially. Instead of calculating the posterior $\tilde{\mathbf{y}}_k = E[\mathbf{y}_k \mid \mathbf{y}_{k-1}, \mathbf{d}_k]$ in terms of the prior and the complete observations, first an intermediate quantity $\tilde{\mathbf{y}}^1_k = E[\mathbf{y}_k \mid \mathbf{y}_{k-1}, \mathbf{d}^1_k]$ is calculated in terms of the prior and the first set of observation, then $\tilde{\mathbf{y}}^2_k = E[\mathbf{y}_k \mid \mathbf{y}_{k-1}, \mathbf{d}^1_k, \mathbf{d}^2_k]$ and so on until finally

$$\tilde{\mathbf{y}}_k = E[\mathbf{y}_k \mid \mathbf{y}_{k-1}, \mathbf{d}^1_k, \mathbf{d}^2_k, \cdots, \mathbf{d}^I_k] \quad \text{..... (3.44)}$$

Each intermediate estimate of the posterior $\tilde{\mathbf{y}}_k^i$ are achieved by direct application of the Kalman update equations,

$$\mathbf{K}^i = \mathbf{C}_y^{i-1}(\mathbf{H}^i)^t\left((\mathbf{H}^i)(\mathbf{C}_y^{i-1})(\mathbf{H}^i)^t + \mathbf{C}_D^i\right)^{-1} \quad \text{(3.45)}$$

$$\mathbf{y}_k^i = \mathbf{y}_k^{i-1} + \mathbf{K}^i\left(\mathbf{d}_k^i - \mathbf{H}^i\mathbf{y}_k^{i-1}\right) \quad \text{(3.46)}$$

Here the measurement update equation is initialized by $\mathbf{y}_k^o = \tilde{\mathbf{y}}_{k/k-1}$ and $\mathbf{C}_y^0 = \mathbf{C}_{y,k/k-1}$. After assimilating measurement type 1 (say water cut), the updated set of permeabilities is used to re-run the simulator, for all the members in the ensemble, from *k-1* to *k*. With the new measurement of type 2 (say bottom hole pressure) it is possible now to compute $\mathbf{C}_y^2$ and then again run the simulator from *k-1* to *k* and so on until all measurements types are assimilated. A clear disadvantage of sequential updating is that it will require to re run the simulator as many times as measurement types.

The scaling problem and/or sequential EnKF when using different types of observation has been studied in detail in the weather forecasting scene[5,28]. However, to our knowledge, this issued has not been mentioned in any EnKF History Matching paper to date.

CHAPTER IV

APPLICATION AND LIMITATIONS OF THE ENSEMBLE KALMAN FILTER TO HISTORY MATCHING

This chapter will provided examples of the EnKF used in history matching. Examples where the EnKF provides a satisfactory as well as unsatisfactory estimate of the true permeability will be discussed. The chapter also presents a brief discussion of the different approaches adopted while trying to reduce/solve the problem and limitations of the EnKF during the present work. This thesis briefly documents three different approaches. Special attention is paid to the approach where the covariance is conditioned using streamlines. Its advantageous characteristic and the detailed description of its implementation will be discussed in sufficient depth.

To better reference what we mean during the discussion, we will refer as the standard EnKF to the approach proposed by Evensen[5] 1994 and later introduce into reservoir history matching by Nævdal *et al*[6]. We will refer as the streamline-assisted ensemble Kalman filter or simply SL EnKF to the conditioning approach proposed here.

4.1 The Use of the Standard EnKF in History Matching

This section describes the results when using the standard EnKF for two synthetic examples and one field case. The first synthetic example has a Gaussian permeability distribution the second synthetic example and the field cases have a multimodal permeability distribution

4.1.1 2D water flooding example using a log-normal permeability distribution

This example shows the application of the EnKF for the data assimilation/history matching of data water cut. The reservoir model consists of a 5-spot example divided into 50x50x1 finite difference grid blocks. The four produces are set to a constant rate production. Figure 8 shows the reference permeability map of the model along with the well distribution.

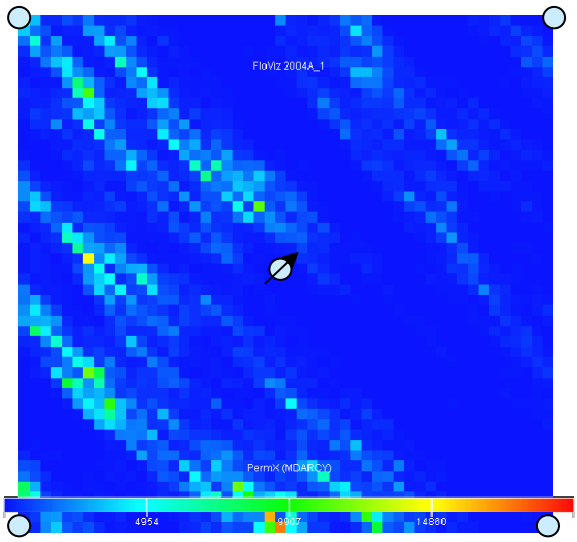
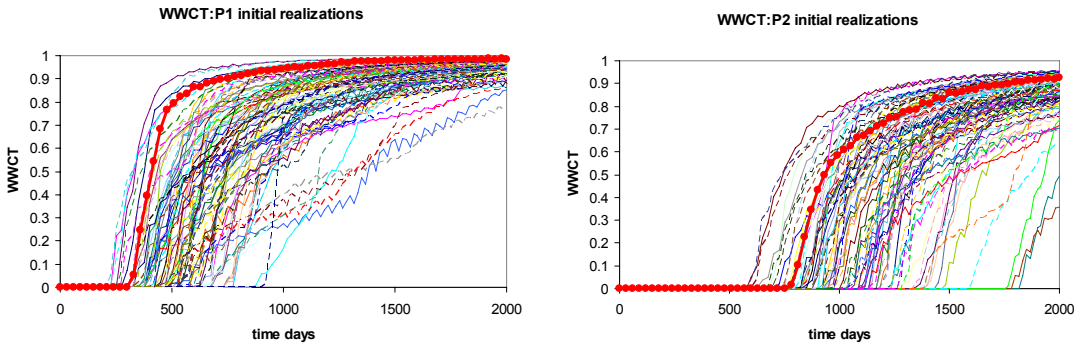


Figure 8. Permeability map for a 50x50x1 reservoir model; dots in the four corners represent the producers and the arrow in the center of the model represent the injector.

To generate the reference permeability model and the permeability of the initial members in the ensemble we used the sequential Gaussian simulation utility of GSLIB[32]. One hundred (100) different permeability realizations were generated. Each realization uses the same variogram and was conditioned to the permeability at wells position. One of the realizations was taken to be the true or the reference permeability.

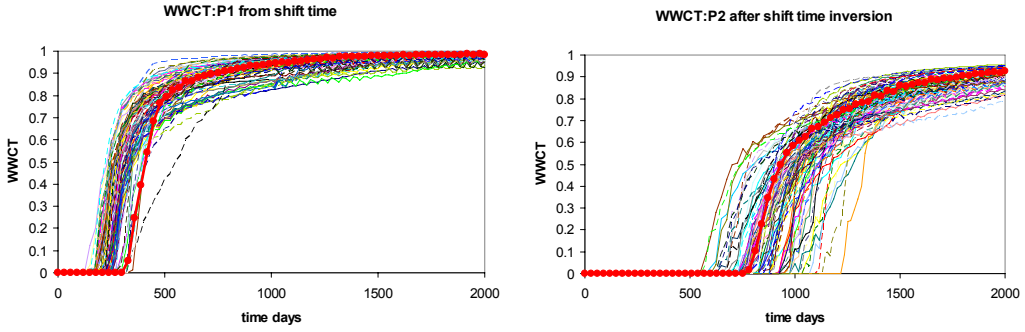
Using a commercial streamline simulator (FRONTSIM), we run our model for 2000 days. The production data from the reference permeability was used to history match all the other models. The production data used in this case was water cut. The static variable of interest was the horizontal permeability $k_h=k_x=k_y$

The final spread of the water cut is greatly decreased after history matching. Figure 9 shows the initial response from the 100 different realizations. Figure 10 shows the response from the 100 updated realizations. Notice the final answer is not longer a deterministic response but a statistical one. If one is asked about the value of the water cut at 700 days at well P2, the response could be drawn from the mean of all the ensembles but it also has an uncertainty analysis on it. One may respond the water cut will be 20% ± 10%. This is a convenient and neat characteristic of the EnKF. It automatically allows to provide an uncertainty analysis in the forecast.



(a) (b)

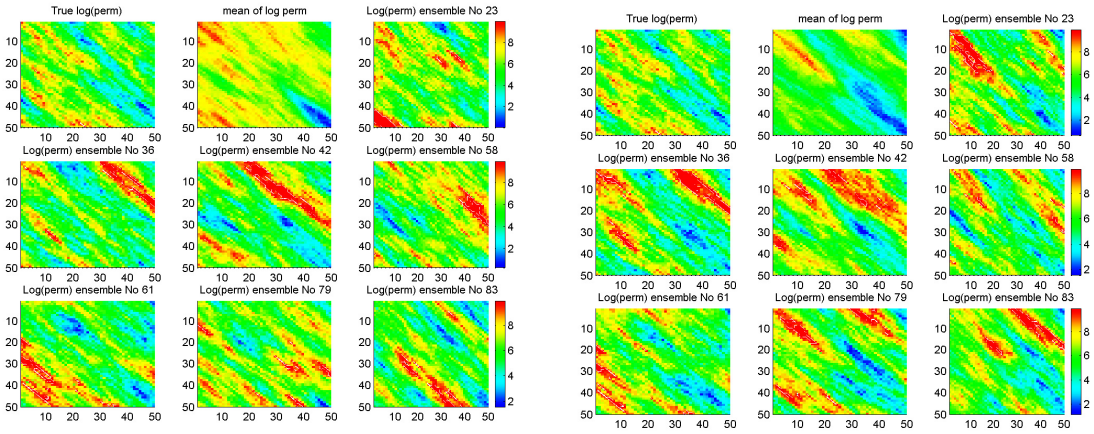
Figure 9. Initial water cut spread from 100 different realizations; the reference water cut is shown by the bold red line; (a) shows the water cut at well P1 (b) shows the water cut at well P2.



(a) (b)

Figure 10. Water cut spread from 100 history matched realizations; the reference water cut is shown by the bold red line; (a) shows the water cut at well P1 (b) shows the water cut at well P2.

Because the goal of the history matching is not only to match the data but also to provide a realistic representation of the permeability, it is always a good idea to check the final permeability distribution. Figure 11 shows the initial (before EnKF) and final permeability distribution (after EnKF). Notice how the mean and other members in the ensemble are able to recover the high and low permeability streaks along the main diagonal of the model.



(a) (b)

Figure 11. (a) Initial and (b) final log permeability distribution. True permeability is in the right upper corner of figure (a) and (b). The mean from the ensemble is next behin to the true permeability. Other permeabilities belong to some randomly selected members/realizations.

This example shows that the standard EnKF works fine when the permeability distribution is Gaussian. This example shows the application of the EnKF with water cut only but any type of measurement can be integrated using the same ideas. Current papers in the topic of using the EnKF for history matching purpose have shown similar cases, all of them with satisfactory result. Although a review of the literature of the application of the EnKF gives the impression that the EnKF methodology is perfectly suited to history matching, the EnKF still has some limitations that require to be addressed, before it can be applied to realistic reservoir models, as discussed next.

4.1.2 2D water flooding example using a bimodal permeability distribution

Real reservoir permeability field can be very complex. It is pretty common to find reservoirs with different facies, channels, barriers, fractures etc. These complex geological characteristics are not easy to model using Gaussian random fields. Geostatisticians have found ways of characterizing these complex models (i.e indicator simulator[33]). Nowadays a typical permeability and porosity distribution for a reservoir model is not Gaussian but multimodal. Since the EnKF theory relies on the first and second moments (mean and variance) of the PDF only, one might be interested to test whether the EnKF is able to properly history match models with a non-Gaussian

distribution of the permeability and/or porosity. For such purpose we set up the numerical experiment presented in this section.

The model consists of a two dimensional square with an area of 1681 ft^2. The model is divided into 41x41 grid blocks. The model is a two-phase incompressible case. Our reservoir has eight producers and one injector as shown in Figure 12. The boundary conditions are of Newman type, fixed production rate all the producers.

The porosity and permeability fields were extracted from a small segment of a real reservoir field. The real field case has different facies, each one with a distribution of permeabilities with a different probability density function. Information for the extracted segment contains three different facies. To generate the initial models we run the sequential Gaussian simulation algorithm to generate permeability within each facies. The permeability within each facies has a different mean and a variance and therefore the final permeability is multimodal. Information to build the variograms was taken from the initial extracted segment at the well position and other few random points. From the 101 realizations one was randomly chosen to be the true case. This true model was then run for 2000 days and the information about water cuts and bottom-hole pressures were saved to be later used as the observation that we want to match using the EnKF.

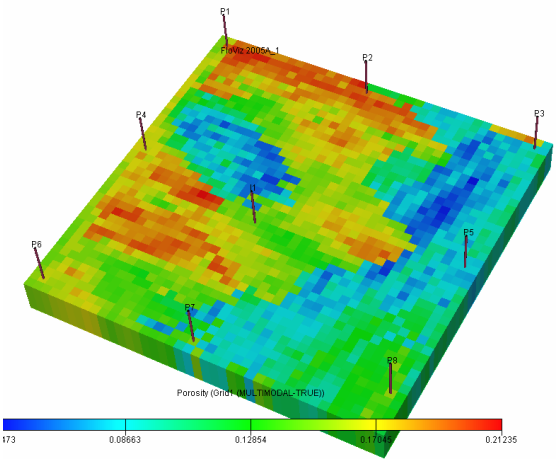
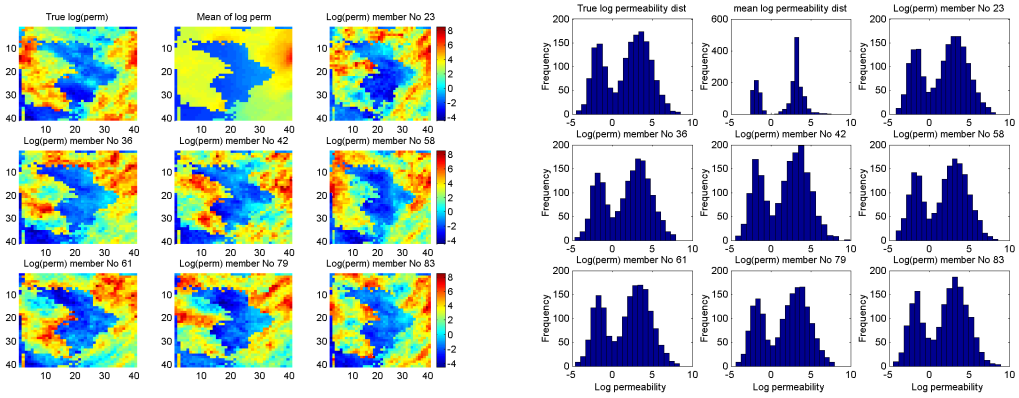


Figure 12. Nine spot waterflooding example; figure shows the porosity of the reference model and the position of the producer and injectors.

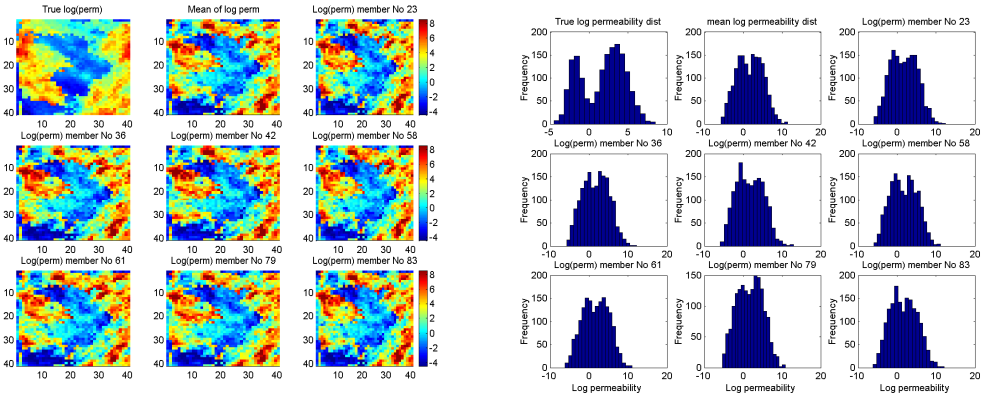


(a) (b)

Figure 13. (a) Initial log perm map (b) Initial log perm histogram. True permeability is in the right upper corner of figure (a) and (b). The mean from the ensemble is next behin to the true permeability. Other permeabilities belong to some randomly selected members/realizations.

The logarithm of the permeability field for some random selected members is shown in Figure 13. Notice that the permeability distribution is not log-normally distributed. Instead the permeability distribution is bimodal. The first mode in the probability density function (PDF) correspond to those permeabilities related to facies 1, the second mode in the PDF correspond to those gridlocks with permeabilities belonging to facies 3. Permeabilities related to facie 2 are some where in the middle of the PDF.

This initial distribution of the permeability was chosen to show some of the problems found with the use of the EnKF when applied to state vector variables with a non-normal distribution. Figure 14 shows the final permeability distribution after the application of the standard EnKF. Notice how the initial bimodal distribution is not respected by the EnKF and is transformed to a Gaussian one. This tendency of transforming the permeability distribution can make the use of the standard EnKF unsuitable for actual reservoir cases with complex geology and heterogeneous distribution (channels and barriers can not be properly preserved by the filter).



(a) (b)

Figure 14. (a) Final log perm map (b) Final log perm histogram. True permeability is in the right upper corner of figure (a) and (b). The mean from the ensemble is next behin to the true permeability. Other permeabilities belong to some randomly selected members/realizations.

Interestingly although the final permeability of all the members became Gaussian and with some over/undershooting problems (see Figure 14), the final members are able to close match the production data. Figure 15 shows that in general the EnKF is able to greatly reduce the spread of the initial members around the observations. This clearly confirms that in history matching problems, the plausible set of models that match the data is infinite, but not all the models matching the data should be considered valid. The question of how to check if a model should be considered a valid solution is still an open question and no unique answer can be given. In actual history matching studies where comparison of the posterior probability with a reference model is impossible, the most straightforward and practical way we have to test the validity of the final models is to compare then with the initial/prior models to check if they are able to preserve the geological realism that went into the prior models.

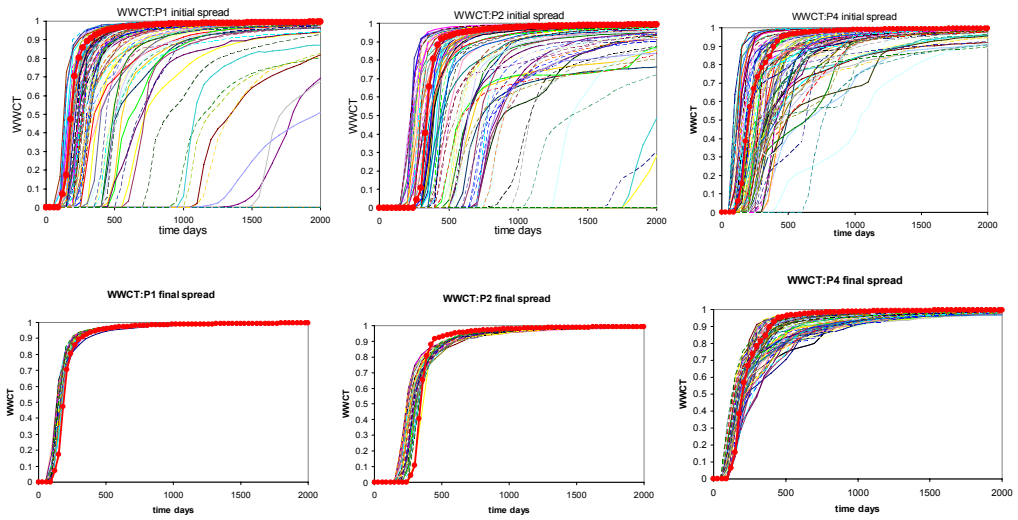


Figure 15. First row shows the initial water cut spread for 100 members at wells P1,P2,P4; second row shows the final spread at the same wells after history matching.

Summarizing the result from this experiment, the standard EnKF was able to match the data but present serious limitations with the posterior PDF. The posterior is unable to preserve the important geological characteristics that went into the model. Results suggest that the standard EnKF has limitations when applied to reservoir models with a non-Gaussian distribution of the permeability.

4.1.3 The Goldsmith field case study

We will discuss the application of the standard EnKF to a field case. The field case is from the Goldsmith San Andres Unit GSAU, a dolomite formation in west Texas. We matched 20 years of water flood production history. The pilot area (see **Figure 16**) consists of nine inverted five-spot patters covering approximately 320 acres with an average thickness of 100 ft.

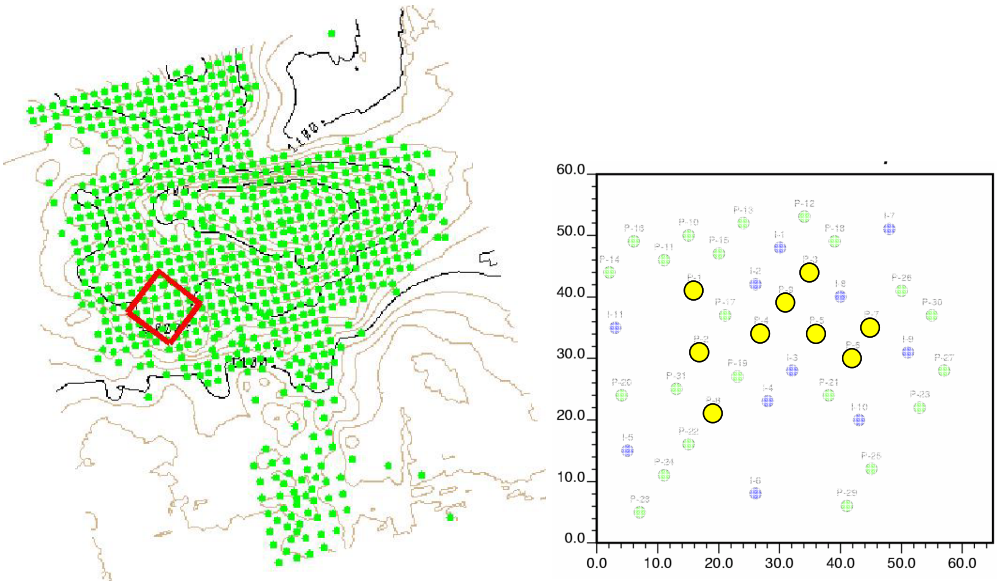


Figure 16. Goldsmith study area; well distribution producer with water cut information highlighted in yellow; Injectors in blue (from Datta-Gupta).

The area has more than 50 years of production history before the initiation of the CO_2 project in 1996. Due to practical difficulties describing the correct boundary conditions for the pilot area, wells around the pilot area were included in this study. The study area includes 11 injectors and 31 producers. Production history information from only 9 producers is used. Because only these have significant water cut response. The detailed production rate and the well schedule, including infill drilling, well conversions, and well shut-in can be found elsewhere[34] The study area was discretized into 58x53x10 grid blocks. The initial 100 realizations of porosity and permeability were obtained using sequential Gaussian co-simulation conditioned to well and seismic data.

Figure 17 shows the logarithm permeability distribution map and the histograms for members 17, 58 and 97. It can be inferred from the figure that the initial distribution is not Gaussian, maybe due to the presence of channels and/or barrier. Although here we only show members 17,58 and 97, the same type of distribution is common in all the initial ensemble members. Actual reservoirs fields can rarely be described using a simple unimodal Gaussian distribution.

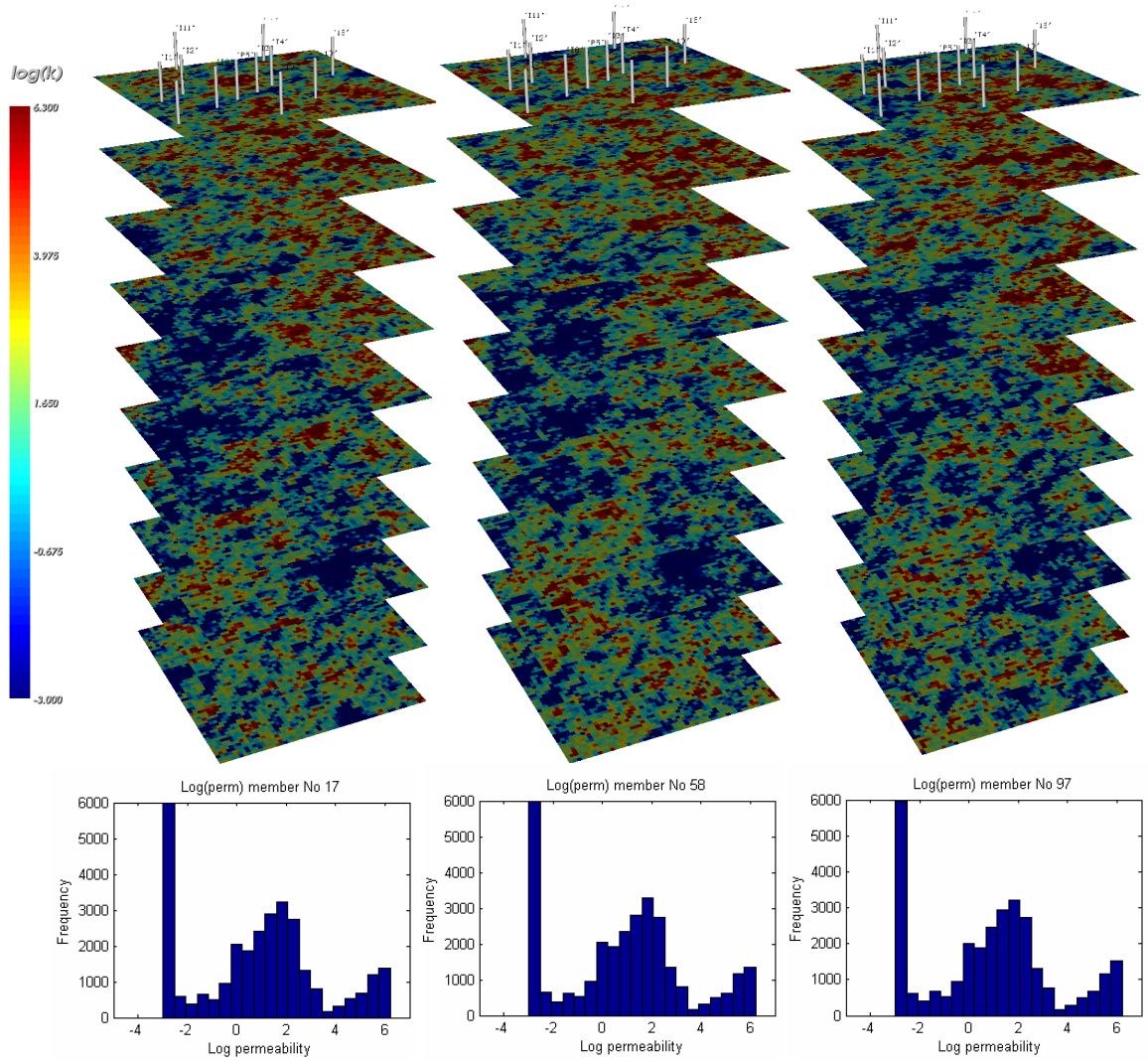


Figure 17. Gaussian random fields for the Goldsmith case (randomly selected members: 17 left, 58 center and 97 right.). Permeability maps generated using sequential Gaussian cosimulation conditioned to wells and seismic data. Below each map it shows the histogram of the permeability.

The problems posted by the EnKF while working with non-normally distributed permeability field becomes more evident in this field case. Figure 18 shows some of the posterior members after assimilation of water cut measurements. Notice the permeability distribution after a sequence of assimilation steps became totally normally distributed. The EnKF was unable to preserve the initial density function of the permeability, specifically high permeability channels, and barriers. The effects of using the standard EnKF with this type of reservoirs models can be very misleading. For instance initially the members have permeabilities ranging from around 0.005 md to 500 md. After

assimilation, using a standard EnKF, due to under/over shooting of the permeability field these values went to a range of $6x10^{-6}$ md and $9x10^{6}$ md which are clearly very large and unrealistic change.

On the other hand an analysis of the ability of the updated members to match the observation reveals that the spread in the water cut of the initial members was greatly decreased around the observation although the permeability became totally unrealistic. Figure 19 shows the initial members spread for water cut (top) and the water cut for the updated members (bottom).

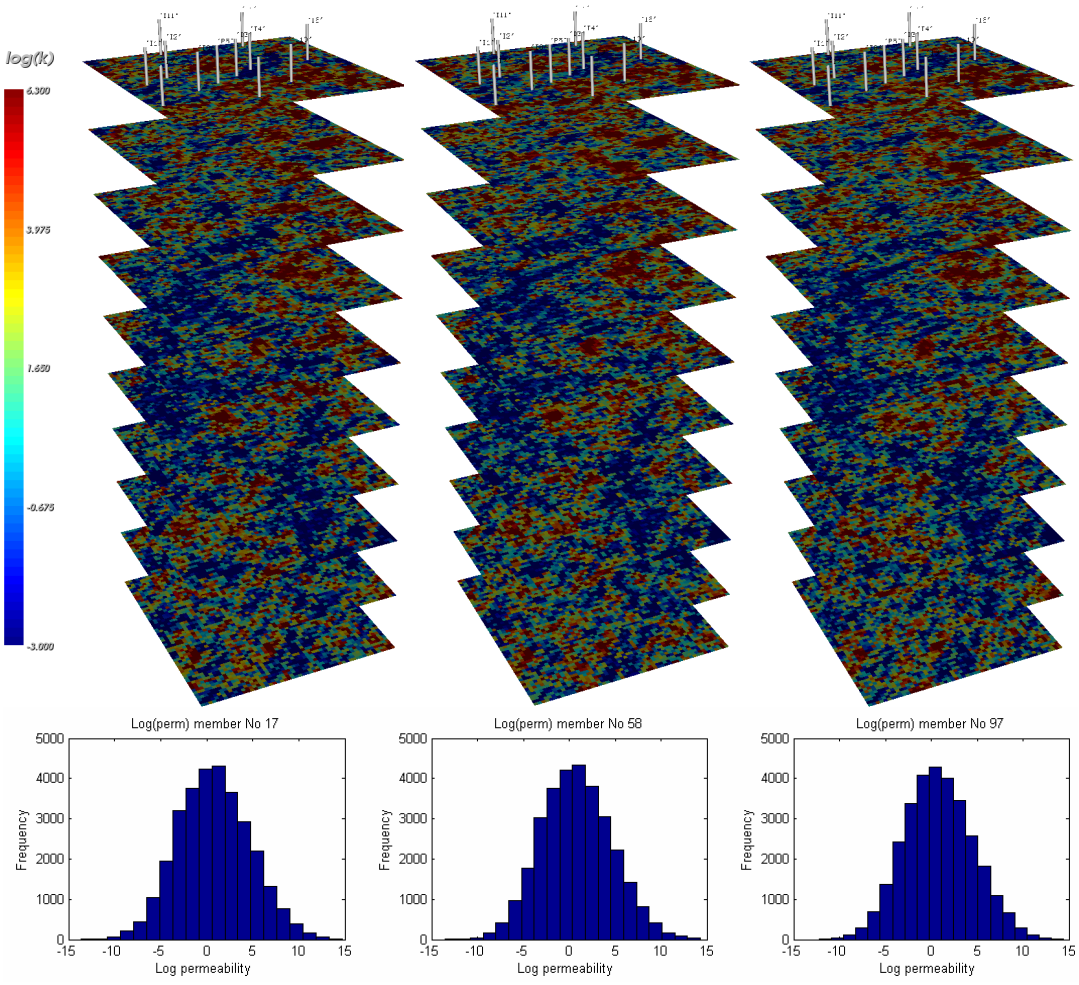


Figure 18. Updated permeability maps using the standard EnKF conditioned to water cut. Below each map it shows the histogram. Notice the standard EnKF is unable to preserve the prior geologic density function.

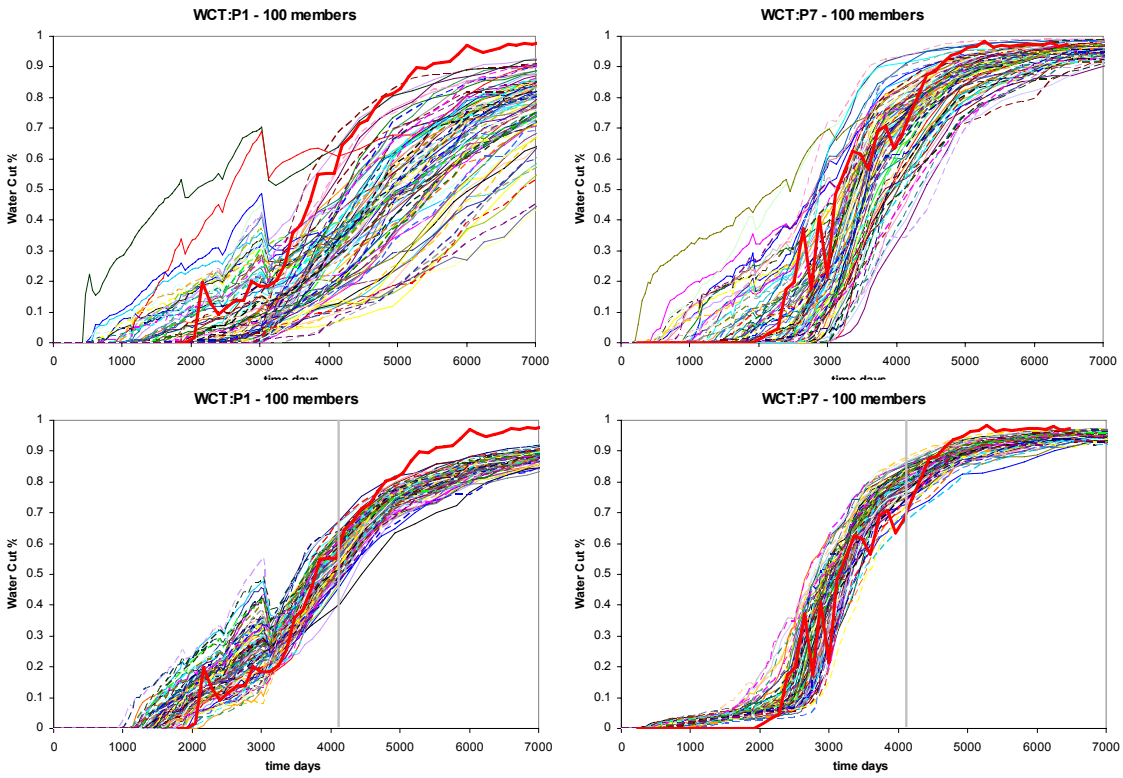


Figure 19. Initial (top row) and final, after standard EnKF (bottom row) water cut spread at wells P1, P7; 100 members in the ensemble; after assimilation the spread was greatly decreased around the observation shown in red. Unfortunately the permeability corresponding to these models have an unrealistic distribution; gray line around 4000 days shows the time up to which the information was assimilated.

Due to the very high dimensionality of the problem and the small number of constraints, it is not surprising that in history matching it will be always possible to find a set of model parameters that satisfy the data but have a totally unrealistic description of the model parameters. This behavior demonstrates that judging the success of the EnKF based on the ability of the final members to reproduce the history is inadequate. A detailed study of the final permeability field is in general recommended. Unfortunately, for actual cases where the true permeability is unknown this is not a trivial task.

In the present field case, the wells that contain water cut information are located towards the center of the model; injectors are located around the producers (see Figure 16). This suggest that the flow of water arriving to each wells should most probably travel across the volume surrounded by the injectors and the wells. A small amount of water should be traveling close to the boundaries of the model were stagnation points can

occurs. This naturally leads us to think that major changes proposed by the standard EnKF should be preferentially found in the area subscribed by the injectors and wells with observations. Unfortunately, this is not the case. In this particular example the standard EnKF has produced changes in all parts of the model irrespective of the wells and injector position. Figure 20 (right) shows the changes to the initial mean after updating of the permeability. Notice how the standard EnKF produce changes even in zones were clearly it should be zones with stagnation points. The changes at the stagnation points do not seems to be supported by a realistic relation between the water cut and the permeability at those areas. We will see later that the SL EnKF presents a major advantage in this issue; changes proposed by the SL EnKF seem to be better guided towards the area where the water is flowing.

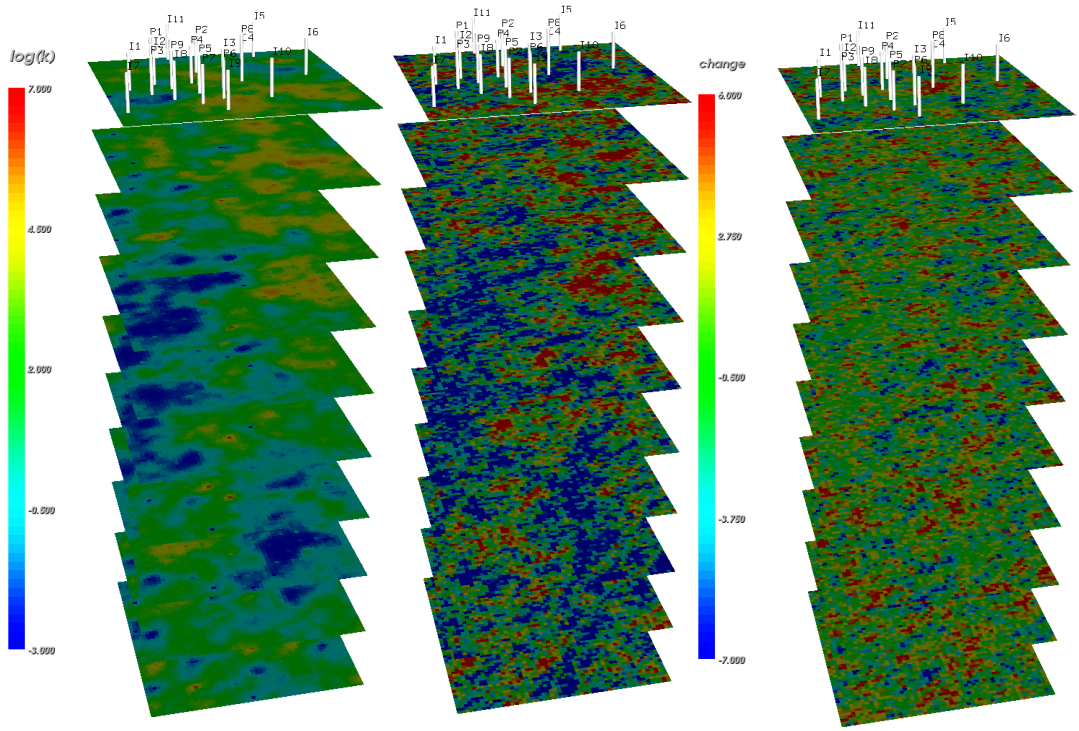


Figure 20. (Left) Initial mean from the 100 members ensemble; (center) mean from the 100 members ensemble after standard EnKF update; (right) changes in the permeability. The updated mean (center) looks pretty sharply and perhaps unrealistic. The changes in the mean (right) shows that changes has indiscriminately occur different at spots all around the reservoir field.

4.2 Problems Posed by the Standard EnKF

From the previous example it became clear that the standard EnKF have some problems and limitations that need to be solved before it can be applied to real reservoir cases. During the introduction and in the literature review, a brief discussion on the characteristics of the problems were presented. In a nutshell the problems considered in this thesis are:

- Over/undershooting problem. The distribution produced by the filter generates some unrealistic, very high or very low, values of the static variable. We believe that the over/undershooting is close related with the EnKF problem known as the filter divergence.
- Limitations to work with non-Gaussian distributions. Recall that one of the assumptions used while deriving the Kalman filter equations was the restriction on the prior and the likelihood distribution be describe by a Gaussian probability distribution. Since any Gaussian PDF can be completely described by its mean and covariance the EnKF takes advantage of its characteristic to describe the multidimensional state vector by its mean and its covariance only. The direct consequence of such assumption is that if the prior distribution is not Gaussian, it will not be respected by the filter and the posterior will become Gaussian, particularly for reasonable ensembles size.

4.3 Towards a Solution: Some Numerical Experiments with the EnKF

Three methods attempting to solve the overshooting problem and/or the limitations with non-Gaussian distribution are considered in this thesis. What follows is a brief description of some of the ideas attempted during a quest of the methodology that may reduce/solve the problems and limitations previously discussed. The techniques presented here are

1. Normal score transformation
2. Using travel time instead of amplitude for matching dynamic data.

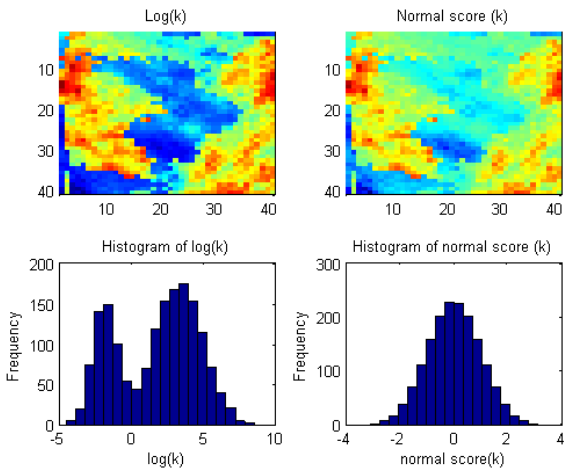
3. Distance dependent conditioning of the covariance

4.3.1 The normal score transform

One naïve and obvious way to avoid the limitations of the standard EnKF to work with non Gaussian distribution is to use the normal score transform. The normal score transform has been used in the past in the EnKF context. Gu and Oliver[13] used it in trying to solve the overshooting problems related to the water saturation in a one dimensional water flood problem. The idea is to replace the model parameter variable, which presents a non-Gaussian distribution, by a proxy variable with Gaussian distribution. The approach will require a non parametric transformation know as the *normal score transform*[32]. The non parametric transformation is constructed from the CDF of gridblock permeabilities. The new permeability variable after transformation will be normally distributed. Figure 21 shows the permeability maps for one member of the bimodal example before and after applying the normal score transform. Notice that before the normal score transformation the permeability has a bimodal distribution and after normal score transformation the permeability has a Gaussian distribution. This non parametric transformation and its inverse can be expressed as

$$\begin{aligned} \mathbf{m}^* &= ns(\mathbf{m}) \\ \mathbf{m} &= bns(\mathbf{m}^*) \end{aligned} \quad \text{.....(4.1)}$$

Where the m* is the normal scored variable and m is any model parameter (i.e permeability). Details of the transformation function $ns(\cdot)$ and its inverse function $bns(\cdot)$ can be found in any geostatistics book[13]



(a) (b)

Figure 21. Permeability map comparison (a) shows the natural logarithm of the permeability and its histogram (b) shows the normal score transformatio of the permeability and its histogram.

Consider an example where we want to update permeability by conditioning to water cut. The state vector for this particular example can be written as

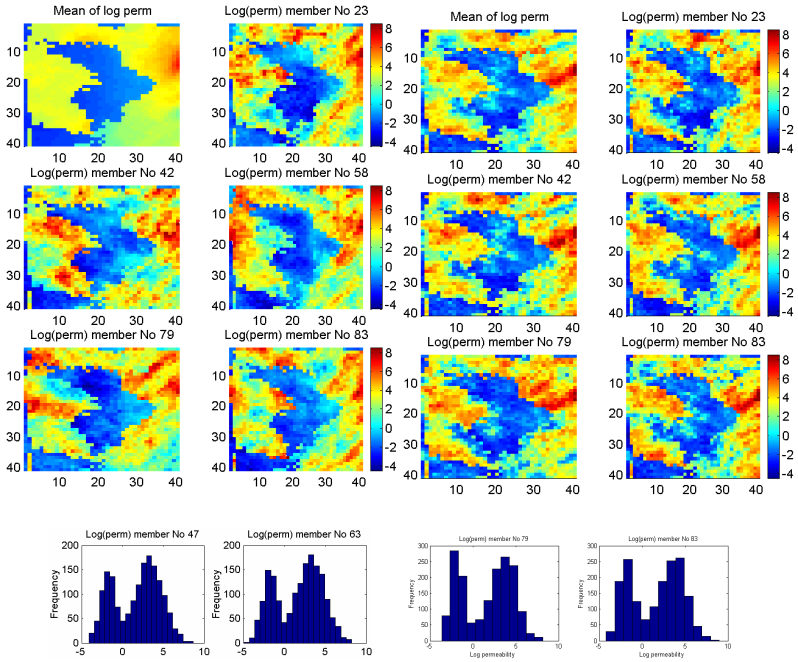
$$y = \left[\log(k_1),\cdots,\log(k_M), p_1,\cdots,p_M, S_{w_1},\cdots,S_{w_M}, WCT_{P1},\cdots WCT_{PN}\right]^t \quad \text{.....(4.2)}$$

Using the new transformed variable (see equation 4.1) the state vector can be redefined as follows

$$y = \left[k_1^*,\cdots,k_M^*, p_1,\cdots,p_M, S_{w_1},\cdots,S_{w_M}, WCT_{P1},\cdots WCT_{PN}\right]^t \quad \text{..... (4.3)}$$

A typical procedure to couple the normal score transformation with the EnKF will require that after we transform the permeability field, we use Kalman update equation (see equation 3.33) and assimilate any observation. After the assimilation we can use the back-transformation function *bns*(·) and get the permeability field of interest. Then, do the time update up to next available observation and repeat the same algorithm again.

We applied the ideas presented here to the bimodal case discussed in the previous section. We will show result where the application of the normal score transform gave acceptable result and another example where the normal score fails to provide a good



(a) (b) (c)

Figure 22. (a) Initial permeability map (b) permeability map after EnKF update with normal score transformation of the permeability. For this example the normal score seems to provide valid results.

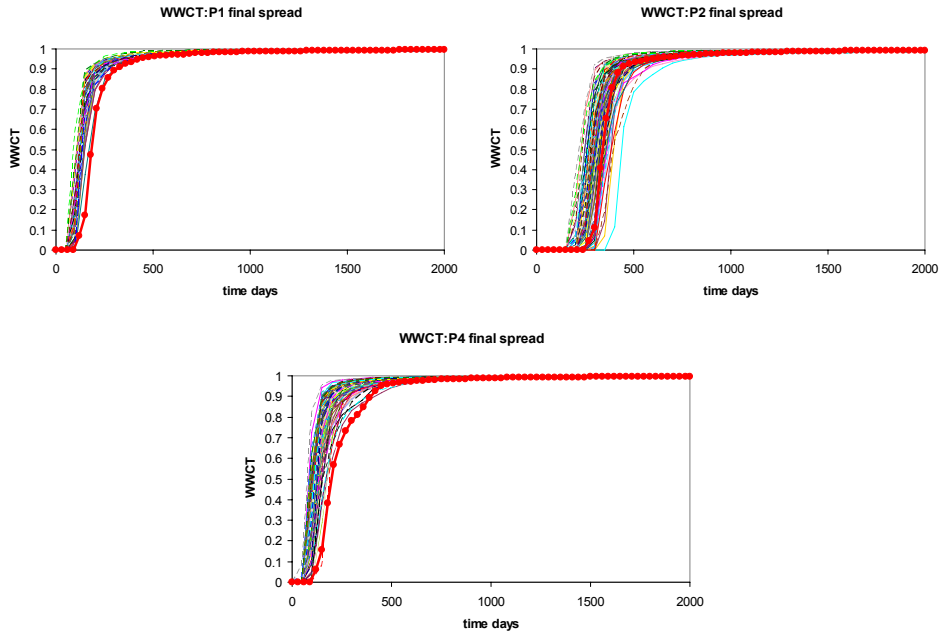


Figure 23. Final water cut spread after assimilation of water cut at wells P1, P2, and P4. The spread correspond to a 100 models in the ensemble. Observation is shown by the bold red curve.

representation of the final permeability. Figure 22 shows the final result for the permeability and its histogram. At a first glance, one might be tempted to think that the normal score approach is working fine, and there is not a easy way to say the contrary in this particular example.

Figure 23 shows the final spread of the water cut after using the EnKF with normal score transformation of the permeability. Notice that the spread around the observation appears to be biased for wells P1 and P4 measurements. The spread around the observation in P2 does not present any apparent problem.

Although the previous example suggests that the normal score transform can provide acceptable results, other experiments carried out with the normal score transformation seem to have some serious limitations. The result presented below correspond to the Goldsmith field case previously discussed in section 4.1.3. We re-run the Goldsmith case with the normal score transformation of the permeability. Figure 24 shows the final permeability map after using EnKF with normal score transformation of the permeability. Unrealistic patches and high sharp changes in the updated permeability suggest that these results should be used with caution.

The result presented here shows a case where there is no clear evidence that the normal score fails and also an example where it seems to fail. From the examples discussed here it is hard to give a definitive conclusion. To properly understand the effect of using the normal score transform a more detailed investigation should be carried out in a future research.

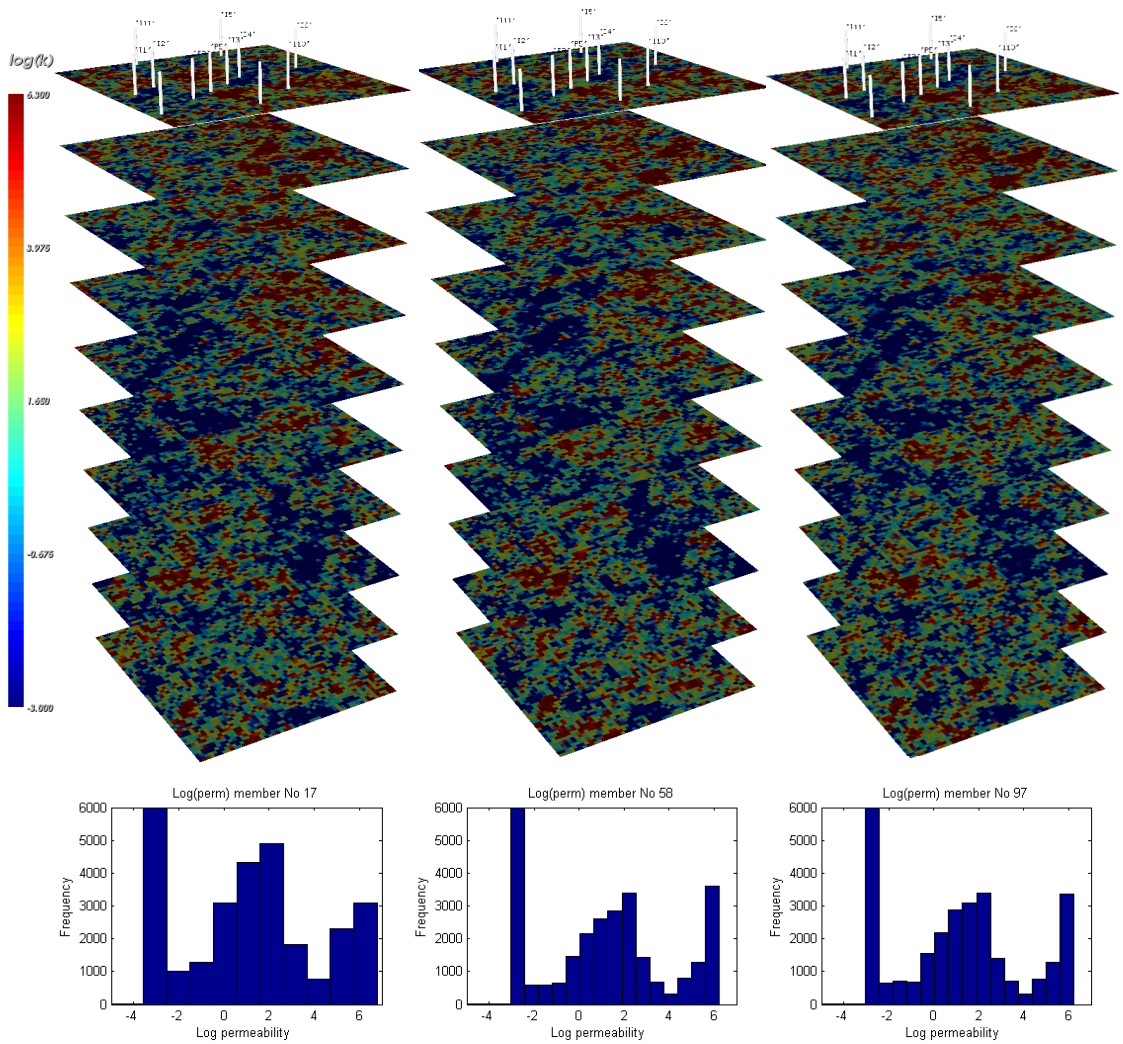


Figure 24. Updated permeability maps using the standard EnKF with normal score transformation of the permeability conditioned to water cut. Below each map it shows the histogram. Notice the normal score EnKF is able to preserve the prior geologic probability density. A closer look into the update permeability map reveals patches and high permeability contrast, which suggest the updated permeability may have some problems; although this results appears to be better than the standard EnKF I believe they should be used carefully.

4.3.2 Travel time inversion of water cut

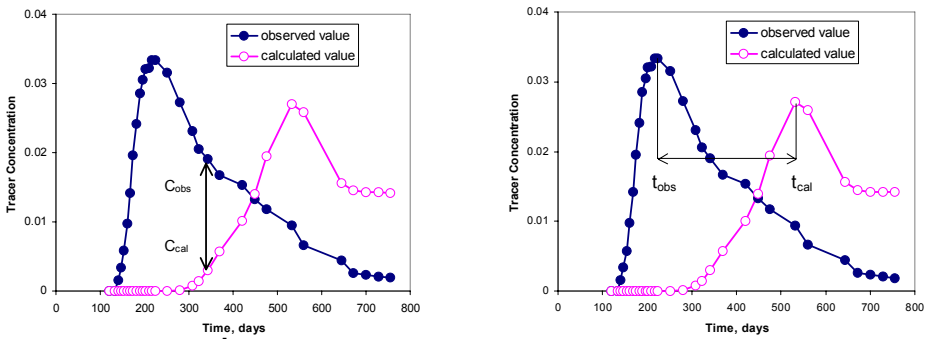
It is well known that the function between the water cut and the permeability is highly non linear. This non-linear relationship may be not be well modeled by the covariance and the mean. It was thought that if we replace the observation water cut by the shift time in the water cut curve, the relationship will be more linear and therefore, the

overshooting problems may disappear. Before we show any result we will try to explain some basic concepts related with the shift time.

What we have done in the previous examples when assimilating water cut measurement can be referred as *amplitude inversion*, because we want to minimize the difference between the observed water cut and the calculated water cut. Another way to approach the same problem is known as *travel-time inversion*. Here we attempts to match the observed data at some reference time, for example the breakthrough time or the peak arrival time. Figure 25 shows a comparison between amplitude inversion and travel time inversion. Here we minimize the horizontal distance between to peaks in Figure 25 (b) which will be called the shift time. It can be shown[35] that the amplitude inversion is highly non-linear compared to travel time inversion which has quasi-linear properties.

Using the shift times instead of the water cut, the state vector (equation 4.2) can be redefined as follows

$$y = \left[\log(k_1),\cdots,\log(k_M), p_1,\cdots,p_M, S_{w_1},\cdots,S_{w_M}, ST_{P1},\cdots ST_{PN}\right]^t \quad \text{..... (4.4)}$$



(a) (b)

Figure 25. An illustration of amplitude and peak travel time inversion. (a) Shows the classical amplitude inversion approach (b) shows the travel time inversion approach. (from Cheng & Datta-Gupta[23]).

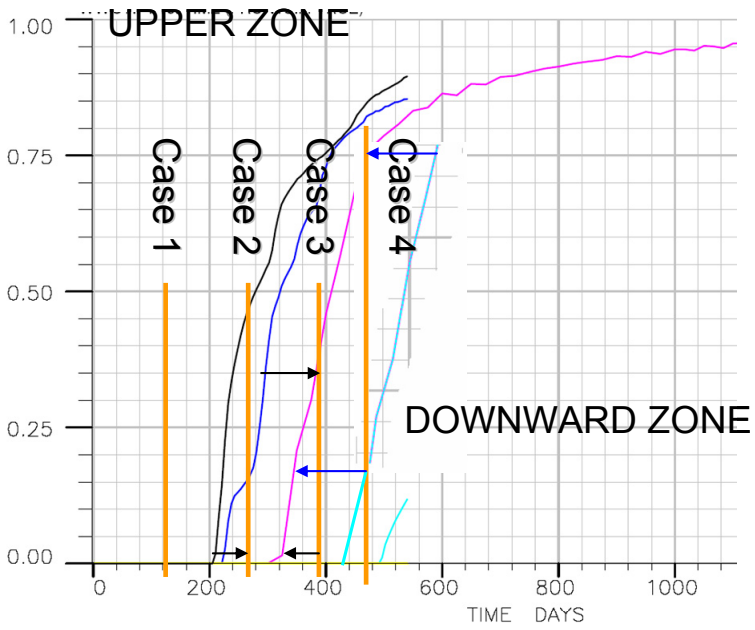


Figure 26. Different case scenarios were one might need to compute the shift time for the observed water cut (in magenta). The vertical orange lines represent times we want to assimilate information. Horizontal arrows shows the shift time required for each curve at the different assimilation times.

Where *ST* refers to the shift time computed from the water cut. Some comment in the calculation of the shift time using the EnKF : 1) since shift time match only one point in the curve (i.e break trough time or a peak), traditionally other authors have used a more general idea known as the generalized travel time. 2) In the standard EnKF we re-start the simulation from a given point in time, let say t_1, up to some other point in time, let say t_2. This means that we will be able to predict the water cut response in that interval t_2-t_1. To compute the shift time of the calculated water cut to the observed water cut, it will require the calculated water cut curve from t_0 to t_{end}; This means we will need to re run each model from t_0 to $t_{end.}$ Application of travel-time inversion appears then, in principle, more expensive than amplitude inversion when using the EnKF where we do not need to restart the simulator from the beginning of time[††]. This difficulty makes the use of the EnKF with travel time inversion more CPU intensive compared to the traditional EnKF

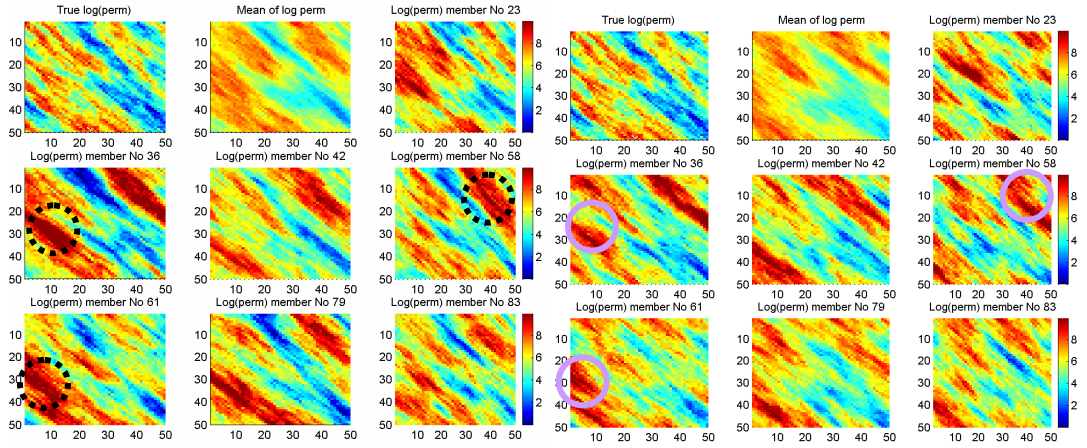
[††] Restarting the simulator from the beginning of time was also investigated by Wen and Chen[12] They restart the simulator to avoid possible material balance error in the saturations due to the update of the dynamic variables.

with amplitude inversion. To better explain the ideas discussed here, let's consider an example with the water cut curve as shown in Figure 26. Notice the observation we want to match is shown in magenta whereas the vertical orange lines represent different times at which we want to assimilate the data. Other water cut curves in black, blue and cyan are the response from some randomly selected members. Now let's consider the case where we want to assimilate information at each time:

1. Case 1: Notice the observation in magenta at time shown by case 1 is zero. Also all the members water cut response is zero. In this case then we can say nothing about the shift time value. Therefore the shift time for these points will be zero in all the members and the assimilation will not have any effect in the permeability
2. Case 2: At this point the observed water cut is zero but the calculated water cuts from some members are no. Response in black and blue are different from zero. The best we can do at this point is to compute the shift time for curve in back and blue as the time from the breakthrough of these two curves up to the current position in time.
3. Case 3: At this point the observation has a value of water cut of around 30%. Notice the calculated water cut, blue curve, has also a value for the water cut of about 75%. Now the idea is to shift the blue curve in time such that the magenta curve and the blue curve super impose one into the other; to do that we can compute the shift time as the value shown by the arrow (related with case 3) in the figure. It is worth to note a difficulty here. Since we need to know the blue curve from time zero up to time given by case 3, thus in general shift time calculations require to re-run the simulator from time zero every time we want to assimilate some measurement.
4. Case 4: At this position in time the observation has a water cut of about 75%, the shift time for back and blue curves can be computed following the same approach explained in case 3. The calculation of shift time is found with the curve in cyan. Notice that at this position in time we only

known the cyan water cut from time zero to time 440 days. The problem is that to compute the real shift time we require to know the cyan water cut at a later time (here up to time 600 days). Since we have run our simulator up to 400 days only, it will be impossible to compute the real shift time. To compute an approximation of the shift time we can extrapolate the cyan curve. Even though this extrapolation gives us a crude value of the real shift time, it seems it is good enough to provide a valuable application of the EnKF.

Our current EnKF implementation has an algorithm to consider the four cases just discussed. We apply these concepts to the reservoir case shown in section 4.4.1 A comparison of the result from the application of the EnKF using shift time is shown in Figure 27. Notice that the small overshooting problems shown in (a) , highlighted by the small black circle, are reduced a little bit by the application of the shift time (see b).



(a) (b)

Figure 27. (a) Updated permeability map using water cut (b) Updated permeability map using shift time. True permeability is in the right upper corner of figure (a) and (b). The mean from the ensemble is next behin to the true permeability. Other permeabilities belong to some randomly selected members/realizations. Althought there seems not to be a huge different beween the two approaches the final map for the shift time (b) do not present the overshooting problems preseted in (a). This suggest that use of the shift time may sligly decrease some overshooting problems.

Figure 28 shows the updated water cut spread at wells P1,P2, and P4. First row shows the result after assimilation of shift time. Second row shows the result after assimilation of water cut. Both figures present a very similar behavior and it is hard to draw a conclusion from the plots. Perhaps a better way to measure the advantages of using shift time in the final match is by measuring the RMS error using equation 4.5

$$RMS_{global} = \sqrt{\frac{1}{W}\sum_{j=1}^{M}\frac{1}{N}\sum_{i=1}^{N}\left(d_{obs,i} - d_{cal,i}\right)_j^2} \quad \text{..... (4.5)}$$

Where d_{obs} is a vector of the dynamic variable of interest d_{cal} is a vector of the response from the simulator using the posterior mean from the ensemble as the input model parameters, i is a time index and N is the number of time data points in the observation, W is the number of wells with observation of type water cut.

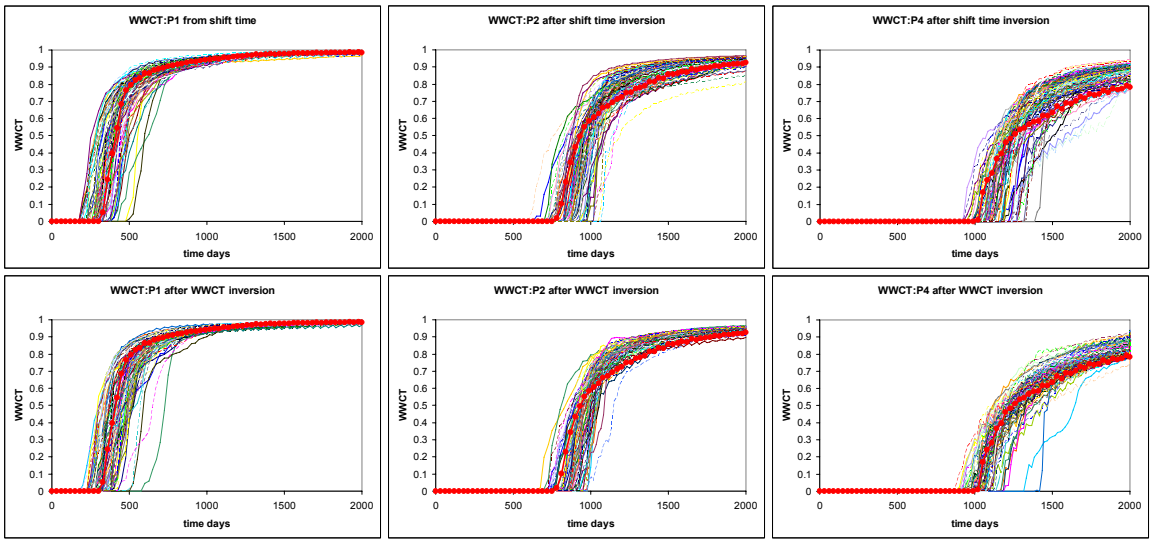


Figure 28. First row shows the final water cut spread after assimilation of shift time; water cut shown at wells P1, P2, and P4. Second row shows the final water cut spread after assimilation of the water cut amplitude. Reference model show by the red bold line. Both cases the simulator was restarted from time zero.

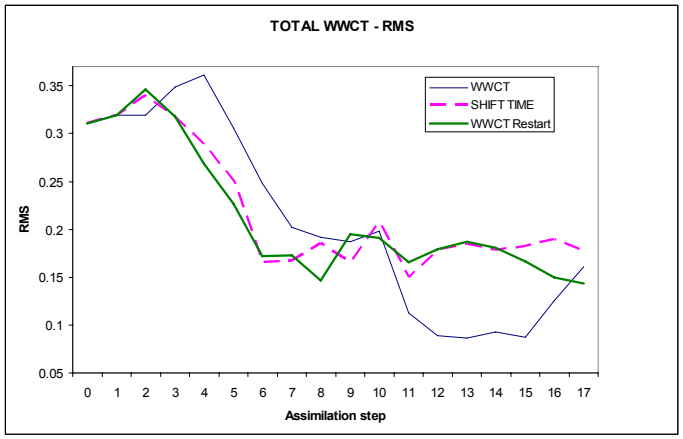
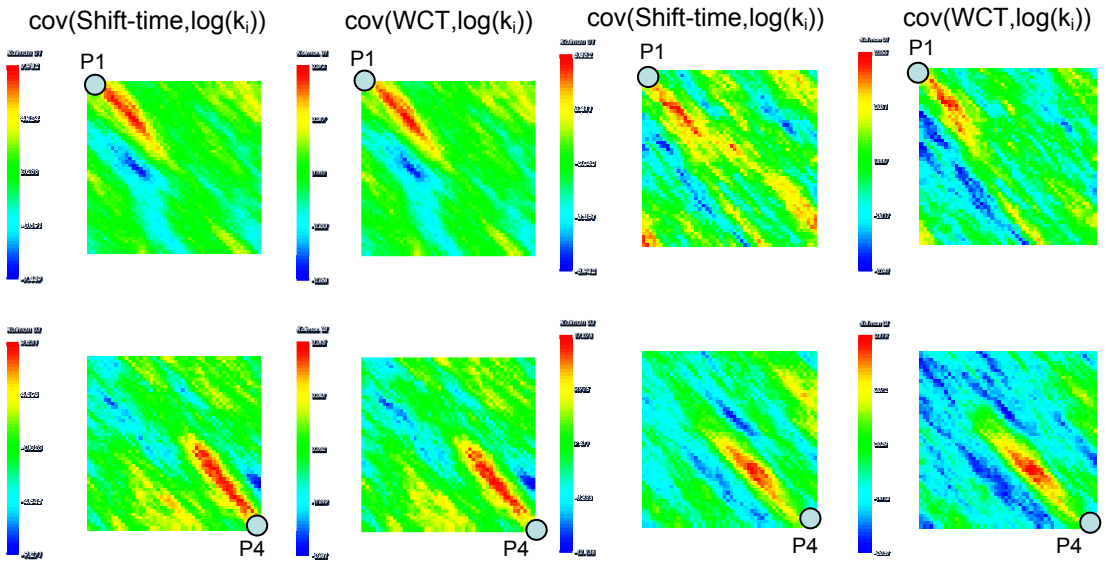


Figure 29. Comparison of the RMS error; To compute the RMS we used the mean of the ensemble and compare its response with the true water cut for all wells. Curve labeled with WWCT shows the RMS error after assimilating water cut with out restarting the simulator; magenta and green shown the results for the shift time and water cut respectively both cases the simulator was restarted from zero.

Figure 29 shows a comparison of the result for amplitude and travel time inversion. Using equation 4.5, both techniques seem to decrease the RMS error to similar levels. From the point of view of matching the data thus we can say nothing about which one does better. But from the point of view of a more reliable update of permeability, perhaps the travel time inversion produces a more realistic description of the permeability.

A comment on the linearity of the relationship between the shift time and permeability as compared to the relationship between the water cut amplitude and permeability. Since it appears the linearity should be more evident in the shift time we decided to plot the cross covariance function between all the grid blocks and the shift time at wells P1 and P4. Figure 30 shows a comparison of the cross-covariance map at two different times during the inversion process. Surprisingly the cross-covariance cov(log(k_i),shift-time) at early time has the same structure when compared to the cross-covariance cov(log(k_i),water cut). Later in time they start to veer one from another; but still the same general trend is present in both. Does this plot suggest that the relationship is not as linear as expected? or does our experiment did not have enough number of members in the ensemble to clarify the expected linearity? This question should be investigated in a future research.



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

Figure 30. (a) and (c) shows the cross covariance between shift time (first row measured at well P1, second row at well P4) and permeability at times 240 and 300 days respectively; (b) and (c) shows the cross covariance between water cut amplitude (first row measured at well P1, second row at well P4) and permeability at 240 and 300 days respectively; Notice how the similar the cross covariance are at early time 240 days. Later the cross covariance are not longer alike but have a similar trend. This may help to explain why both techniques results in a similar posterior permeability.

Summarizing, at least for this example the use of the shift time seems to have a slightly better performance in terms of the reduction of possible overshooting problems. The advantages regarding the matching of the production data (RMS error see Figure 29) are not evident.

4.3.3 Distance dependent conditioning of the covariance

To better understand the effect of the conditioning of the covariance matrix $\mathbf{C}_y$, let us start by analyzing some of the sub matrices in $\mathbf{C}_y$. Recall from equation 3.8 and 3.21 that the sub matrix that ultimately drives the changes in the static variables (i.e permeability, porosity) is known as $C_{M^s,d}$. This matrix contains the cross-covariance relationship between the static variable (model parameter) and the dynamic variable (data, observation). For completeness purpose, equations 3.8 and 3.21 are rewritten here

$$\mathbf{C_y} = \begin{pmatrix} \mathbf{C}_{M^s} & \mathbf{C}_{M^s,M^d} & \mathbf{C}_{M^s,d} \\ \mathbf{C}^t_{M^d,d} & \mathbf{C}_{M^d} & \mathbf{C}_{M^d,d} \\ \mathbf{C}^t_{M^s,d} & \mathbf{C}^t_{M^d,d} & \mathbf{C}^{cal}_d \end{pmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{ (4.6)}$$

$$\tilde{\mathbf{m}} = \mathbf{m}_{prior} + \mathbf{C}_{M^s,d}(\mathbf{C}^{cal}_d + \mathbf{C}_d)^{-1}(\mathbf{d}_{obs} - \mathbf{H}g(\mathbf{y}_{k,prior})) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{ (4.7)}$$

Let's consider a hypothetical case where we may be interested in updating the permeability field using bottom hole pressure (BHP), water cut (WCT) and oil production rate (OPR). A detailed description of $C_{M^s,d}$ for this particular example is shown in equation 4.8. It reveals each one of the components of this matrix.

$$C_{M^s,d} = \begin{pmatrix} \mathrm{cov}(\log(k_1), BHP) & \mathrm{cov}(\log(k_1), WCT) & \mathrm{cov}(\log(k_1), OPR) \\ \vdots & \vdots & \vdots \\ \mathrm{cov}(\log(k_M), BHP) & \mathrm{cov}(\log(k_M), WCT) & \mathrm{cov}(\log(k_M), OPR) \end{pmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{ (4.8)}$$

Where the index M in the permeability refers to the index for each grid block. It is well known that a cross-covariance between the variables log(k) and BHP, cov(log(k),BHP) will measure the correlation between these two variables. This cross covariance needs to be calculated at every grid block. Figure 31 shows a cartoon of the approach used in the EnKF to compute the cross covariance matrix. It can be explained as follows: Think about each point, in the x-axis of the scatter plot, as set of plausible value for the log of the permeability at grid block 23. Now since in the EnKF we have several realization of the permeability field then the permeability at grid block 23 has different values as shown in the figure. Notice that each point in the y-axis of the scatter plot represent a value for the oil production rate at the given well. Since each realization produces a different forecast of the oil production rate, the production at the well will be different for each realization. The upper scatter plot shows a cross plot of the different values of the permeability vs different values of the oil production rate. It is also important to notice from the figure that one of the grid blocks is near by the well while the other grid block in the example is far way from the well. It is clear from basic

statistical theory that one needs to sample the population with a large number of experiments to be able to properly characterize its PDF. This suggest that unless we estimate the cross-covariance of permeability and oil production rate from a very large ensemble, one may be estimating the true value of the cross covariance without enough information.

It is also important to analyze the effects of the distance between the well with the observation and the grid block of interest. Since the relationship between the oil production rate WOPR and the permeability for the far apart grid block (grid cell 132) is highly non-linear, then the computed cross-covariance can be easily overestimated, unless we use large number of samples. Using a small ensemble size may then overestimates its value and ultimately generates overshooting problems. Appendix I contain a mathematical proof of ideas here discussed. The proof was inspired in the work of Hamill, T.M., Whitaker, J.S.[36]

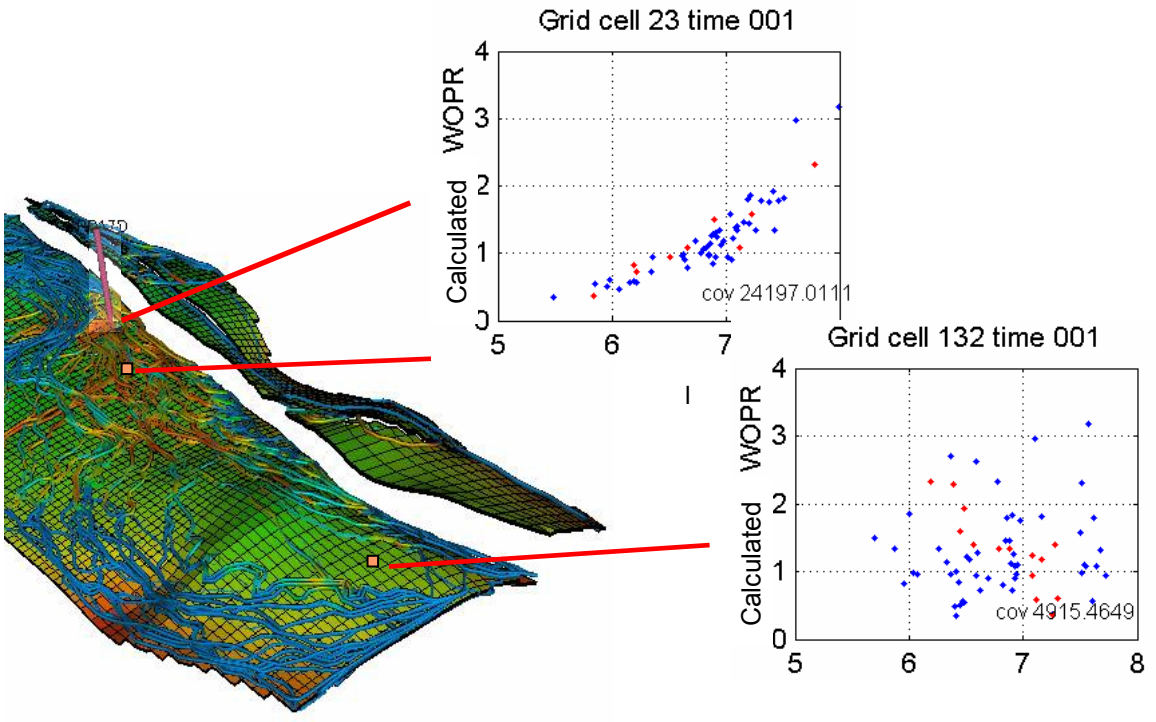


Figure 31. Example showing the cross covariance between the observation at a given well and the permeability at two different cells cells; upper scatter plot shows that according with the results from the ensemble the correlation between gridblock 23 and the oil production rate at well is almost linear; the lower scatter plot shows that for the grid block 132 the correlation is not well resolved by the covariance operator.

Figure 31 suggests that when using a small ensemble size it may be beneficial to condition the covariance matrix such that those cross covariance values related with grid blocks far away from the observation be damped. The remaining question is how to decide if a grid block is far away enough such that the cross covariance should be small or close to zero. What follows is a description of our proposed approach to decide which grid block of the cross covariance should be damped

A general rule for any history matching technique is to change the parameters where the uncertainties in the effect of the change is better known and/or changes in the parameters will have the largest influence on the solution. Thus, it is vital to identify zones in the reservoir where changes will have the greatest impact and/or where the relationship between the production data and the parameters of interest is better known. We can then perform the required changes in these zones preferentially. Prior knowledge of the governing physical phenomena can be used to infer these zones. For instance in primary depletion, the bottom-hole pressure is mainly affected by reservoir parameters with in the zone defined by the radius of investigation. In waterflooding scenarios wells water cut are primary affected by reservoir properties with in the swept zones. When sufficient spread is provided the covariance matrix can point out zones where the dependency with the observation is stronger. In the absence of well conditioned covariance; (i.e filter divergence) other more physical-phenomena can provide ways to identify these zones. Information from streamlines can be used for this purpose. The use of streamlines to decide zones with dominant changes have been proved to be useful in the past[1]. Streamlines assisted history matching was first proposed by Emanuel and Milliken[2]. They used streamlines to identify the grid blocks that affect each well in relation to a production response. In other areas of science where EnKF has been applied, like in weather forecasting, the medium under consideration is more homogeneous, conditioning of the covariance have been used[15,28,36]. Houtekamer and Mitchell[15] proposed to use a cut off radius, beyond which the effect of a given observation is small over their model parameters. The corresponding zones of the covariance outside the cut off radius are damped or not used in the computation of the cross-covariance matrix. By

doing so they avoided the estimation of small correlation associated with remote grid blocks.

The technique is supported by the fact that unless a high number of members are provided, the correlation of some remote grid blocks with some measurement may be noisy or over/underestimated. To filter remote observations, they used a Schur product of the covariance times a correlation function. A similar idea is used here, but instead of using a correlation function to decide which grid blocks are strongly correlated to an observation, we use streamlines. As an advantage, our approach is better suited for reservoir-like problems compared with the cut off radius approach since breakthrough time and water cut curves are often related with not only those grid blocks defined by a cut off radius but with grid block that may be far away from the observation and/or beyond the cut off radius, specially for highly heterogeneous reservoirs. (Typically those grid blocks correspond to those streamline that may follow a channel like path. See Figure 32)

Zones corresponding to each observation are selected by streamlines from several simulation models (see Figure 32). Our final set of zones is determined by stacking zones from each members of the ensemble to produce a single set. Additional constrains can be further applied based on water front movement and breakthrough time and/or minimum number of members in the ensemble affecting a given region. Since the EnKF is a sequential updating technique, each time we want to assimilate observations, streamlines are re-generated. Therefore, different flow path mapping may happen for different arrangement of the injector/producer through time (see Figure 33).

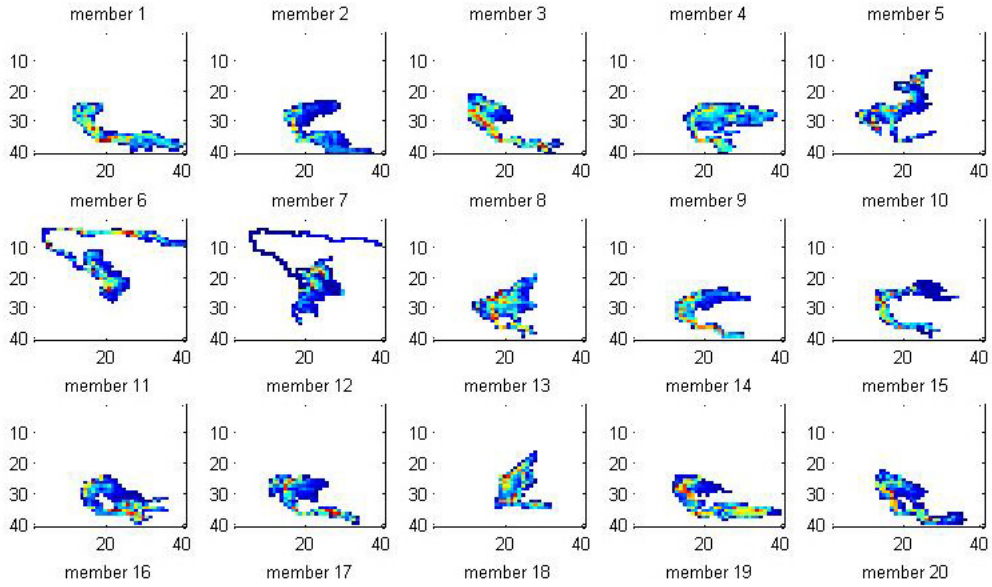


Figure 32. Gridlocks behind the saturation front crossed by streamlines arriving at well P8 at time 330 days for members 1 to 25. Reservoir model correspond to the nine spot examples with bimodal permeability distribution, shown the previous section.

An outline of the procedure for our proposed approach is given in the flow chart in Figure 34. Descriptions of the major steps are as follows,

1. **Ensemble Forecast Step**. We start with an ensemble of reservoir models conditioned to static data. For each member of the ensemble, we simulate the production response up to next available observation time, using either a streamline or a finite-difference simulator. For finite-difference models, we also trace the streamlines for each member of the ensemble. The streamline tracing is performed using a generalization of the Pollock algorithm using the total phase fluxes from the simulator.[37,38]
2. **Computation and Conditioning of the Cross-covariance Matrix.** For each member, we utilize the streamlines to associate grid blocks or regions that contribute to each producer at given time. Next, we stack the selected grid blocks/regions from all ensemble members to define a common region of influence for all members. Using the production responses at current time step from each member, we compute the cross-covariance matrix including only the grid blocks within the common region of influence.

3. **Computation of the Kalman Gain.** Using the covariance of the observed data, the model response and the cross-covariance from the previous step, compute the Kalman gain (Equation 3.16).
4. **The Kalman Update.** Update the reservoir model using the Kalman update equation (Equation. 4.9, discussed later). Repeat all steps again until all production data are assimilated.

As can be inferred from the flow chart the implementation of the approach explained here can be easily incorporated in any existing EnKF implementation, since it only requires an additional step as compared to a standard EnKF implementation.

To account for the conditioning using streamline the covariance matrix is redefined as

$$\mathbf{C}_{\Psi}^{p}\mathbf{H}^{t} = \rho \circ \left(\frac{1}{N_e - 1} \sum_{i,j=1}^{N_e} (\mathbf{y}_i^p - \overline{\mathbf{y}}^p)(\mathbf{H}\mathbf{y}_j^p - \mathbf{H}\overline{\mathbf{y}}^p)^T \right) \quad \text{..... (4.9)}$$

Where ρ is a correlation function representing the flow path information extracted from the streamlines (see Figure 33). More specific information of how to build this matrix will be given in the next section. The operation $\rho\circ$ in equation 3.40 denotes the Schur product operator. Recall that the Hadamard or Schur product of two matrices ***A*** and ***B*** with the same size, denoted by ***(A∘B),*** is defined as ***(A∘B)*** = $(a_{ij}b_{ij})$ for each element of the matrices.

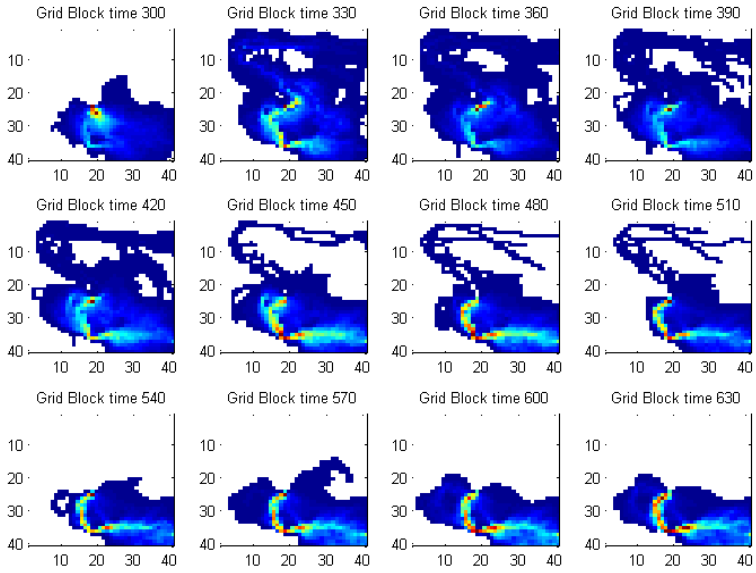


Figure 33. Product of stacking regions selected by one hundred different realizations. The regions selected by the streamlines change during the assimilation period; Notice at early times 330-510 days the stacking regions cover more area; This is due to the fact that at early time when no too many points has been assimilated the permeability field of the realizations differ a lot between them; at late time all realization has similar characteristic therefore the stacked region are more uniform.

4.4 The Use of the Streamline Assisted EnKF in History Matching

In this thesis the SL EnKF approach is tested using a set of two reservoirs where having serious convergence problems when using the conventional EnKF history matching techniques. We will present the result using the standard EnKF and our new approach to highlight major advantages of the new methodology. This section discusses the performance of the covariance conditioning technique here known as the SL EnKF. To check its performance, we start by re-running some of the examples with problems shown in the previous standard EnKF section.

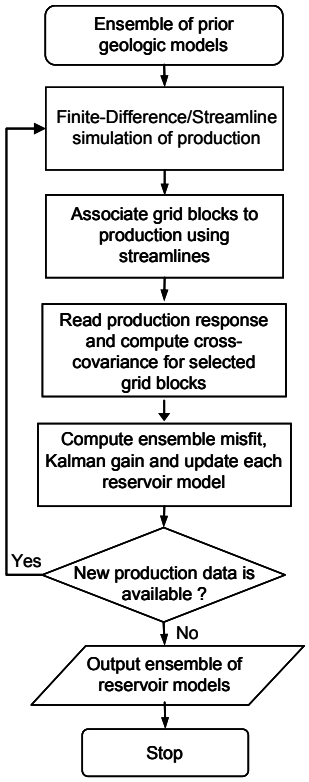


Figure 34. Streamline assisted ensemble Kalman filter flow chart.

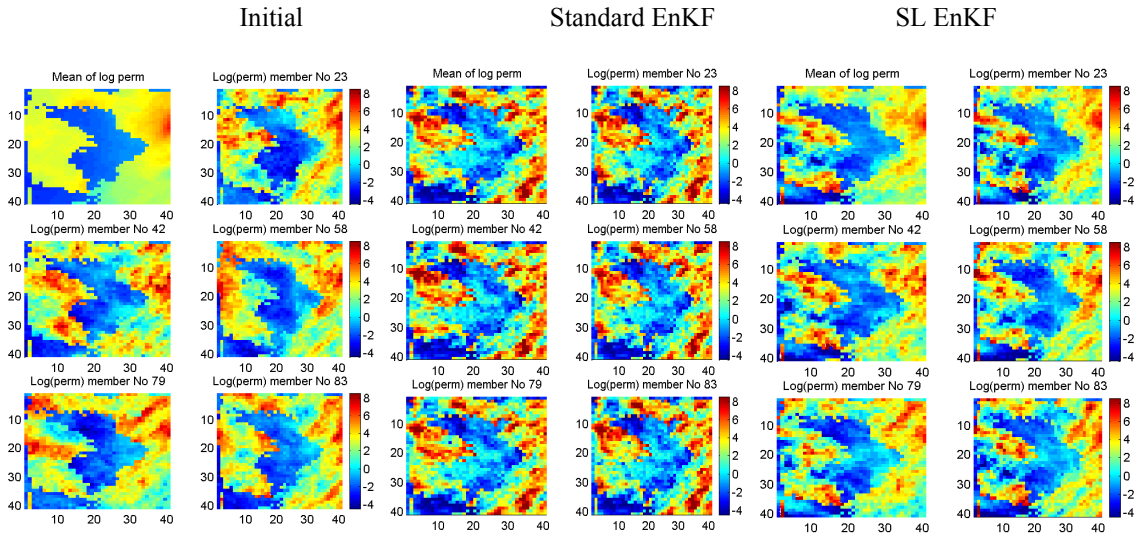


Figure 35. a) Initial permeability map for randomly selected members in the ensemble, upper left corner shows the mean; b) posterior, after application of the standard EnKF, over/undershooting is evident in all the members; c) posterior, after application of the SL EnKF, problems with over/undershooting disappeared.

4.4.1 2D water flooding example using a bimodal permeability distribution

See section 4.1.2, "2D Water flooding example using a bimodal permeability distribution", for a description of the reservoir model used in this example. The idea of rerunning the model using the SL EnKF is to check whether the SL EnKF will reduce/solve the previous overshooting and Gaussian distribution related limitations problems reported by the standard EnKF.

Figure 35 shows a comparison of the result between the standard and the SL EnKF. It is clear that while the standard EnKF produces over/undershooting of the permeability, the SL EnKF does not. The over/undershooting seems to disappear when the SL EnKF is used. Figure 36 shows a comparison of the initial and posterior PDF using quantile to quantile (Q-Q) plot. It is obvious from the figure that the SL EnKF has a better performance

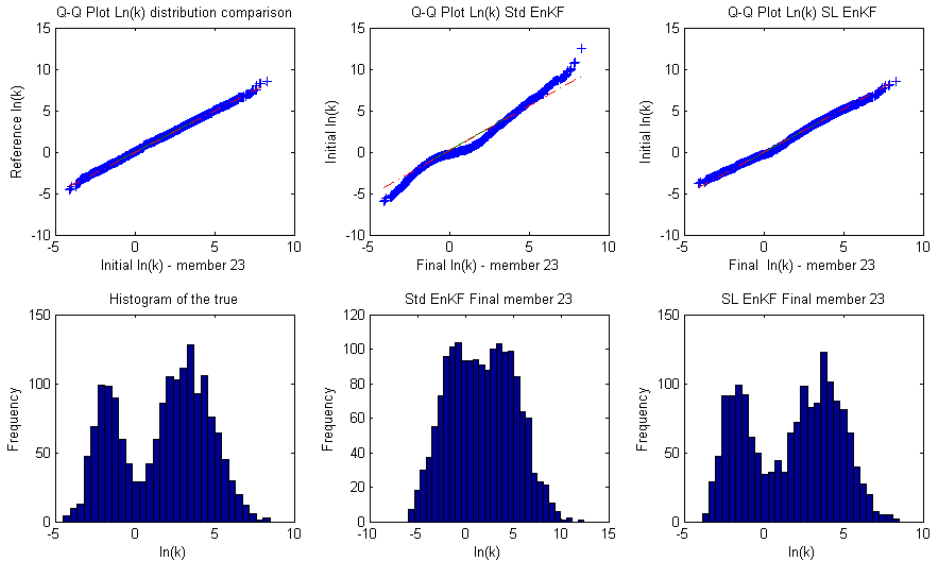


Figure 36. Comparison of the true PDF with respect to member 23; (Left) shows the Q-Q plot comparing the true vs member 23; notice that the Q-Q plot gives a straight line which mean that both PDFs can be considers the same; below the Q-Q plot it shows the histogram of the true permeability. (Center) shows the Q-Q plot comparing the initial PDF vs the posterior PDF after using the standard EnKF; the Q-Q plot does not show a straight line which means that the distribution are non longer equals; below the histogram of the posterior shows that the PDF changed from bimodal to a almost Gaussian distribution; (Right) shows the Q-Q plot comparing the initial PDF vs the posterior PDF after using the SL EnKF; the Q-Q plot shows an almost straight line which means that the posterior distribution was able to preserver most of the characteristic of the initial; below the histogram of the posterior shows that the PDF is still bimodal.

Figure 37 first row shows the initial and final spread distribution for the water cut at wells P1,P2 and P4. Second row shows the final spread after using the standard EnKF. The third row shows the final spread after using the SL EnKF. A naïve analysis of the spread second and third row will assume that the standard EnKF is better. But one needs to recall that during any history matching process user needs to trade how good does he fit the data while preserving the prior. It is clear that since the spread of the ensemble is better preserved by the SL EnKF, then the filter divergence problem and consequently the over/undershooting problem disappear.

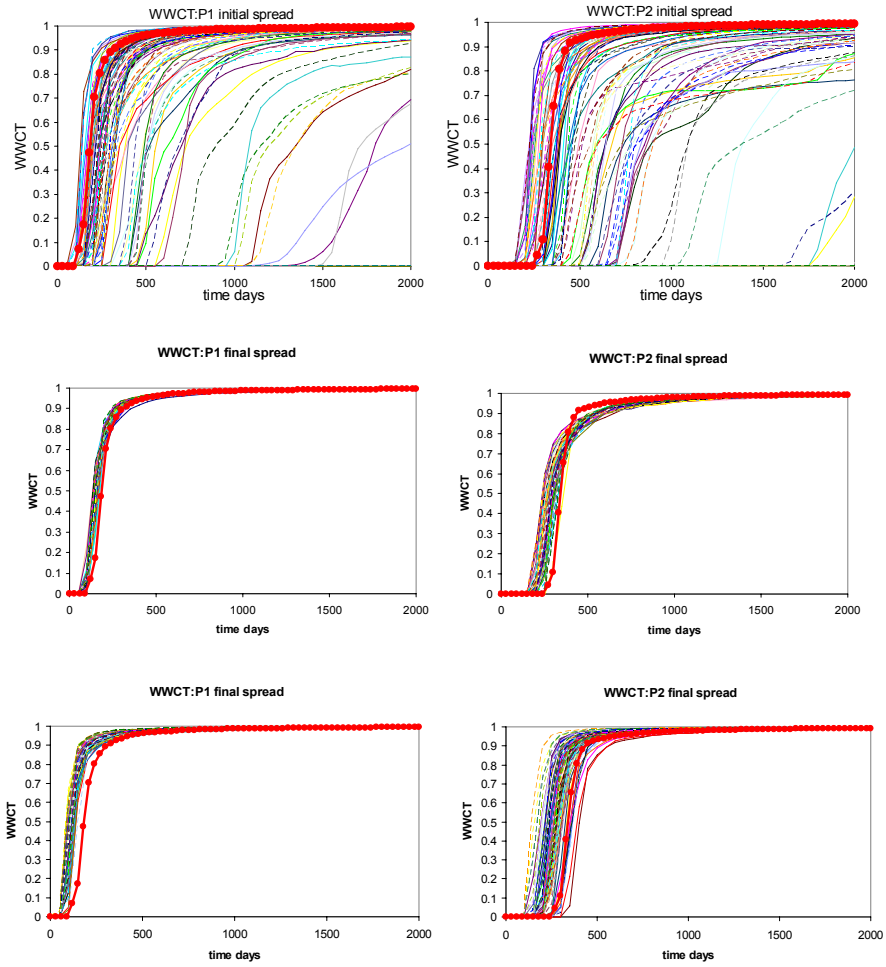


Figure 37. Comparison of the water cut spread for initial members (first row) the standard EnKF (second) row and the SL EnKF (third row).

It is not completely clear why the SL EnKF reduces the magnitude of the problem related with the Gaussian assumptions. Certanly part of the answer is tied to the fact that the changes in the SL EnKF are much smaller in magnitude compared to the standard EnKF.

Summarizing the result from this experiment, the SL EnKF was able to preserve the non-Gaussian distribution and also it was able to decrease/solve the overshooting problems while still matching the data.

4.4.2 The PUNQ-S3 synthetic case study

The previous example presented a qualitative description of the advantages of using the SL EnKF. The next example tries to quantify the advantages of using the SL EnKF. The SL EnKF will be applied to the popular PUNQ-S3 model. The PUNQ-S3 reservoir model was developed in the European Union by a group of companies and universities. Detailed description of the model can be found elsewhere[39,40]. Here we will give a brief description of the model. The PUNQ-S3 is a small size reservoir model. It consists of 19x28x5 gridlocks from which 1761 are active. The gridblock sizes are Δx=180 ft, Δy=180 ft, Δz ranges from 1.3 to 8.8 ft. The reservoir has a small gas cap in the center of a dome shape structure; it has a fault to the east and south and a strong aquifer zone to the west and north (see Figure 38). The field initially contains 6 production wells located around the Gas Oil contact. Due to the strong aquifer, no injection wells are present. All six producing wells were produced as follows: an extended well testing during the first year, then a shut-in period lasting the following 3 years, and finally a 4-year production period. The well testing period consists of 4 time windows, each of which is 3-month long with a constant flow rate. The oil production rate is fixed at 150 sm^3/day within the 4-year production period. All wells have a 2-week shut-in each year to collect shut-in pressure. The reservoir properties and an ECLIPSE input data file can be downloaded from reservoir project website. In our history matching examples we used the streamline simulator FRONTSIM. Because of the differences in the simulators we did not used the 16.5 years of production data given in the PUNQ-S3 website. Instead, to avoid any bias in observed data, we regenerated our reference production data using FRONTSIM. The initial ensemble members were generated using

GSLIB[13]. Values for normalized porosity at well locations and anisotropy information used to generate the true model are given in the PUNQ-S3 website. The horizontal and vertical permeabilities were calculated using a deterministic relationship, estimated from the well data and proposed elsewere[40]

$$\log(k_h) = 9.02\phi + 0.77$$

$$k_v = 0.31k_h + 3.12 \quad \text{..... (4.10)}$$

We history matched the model using the standard EnKF and the SL EnKF; data was assimilated for 8 years period (2936 days). Using the updated models we rerun all the models from time zero and forecasted the next 8.5 years. Rerunning the models from time zero was done to account for possible material balance error during the updating of the EnKF phases.

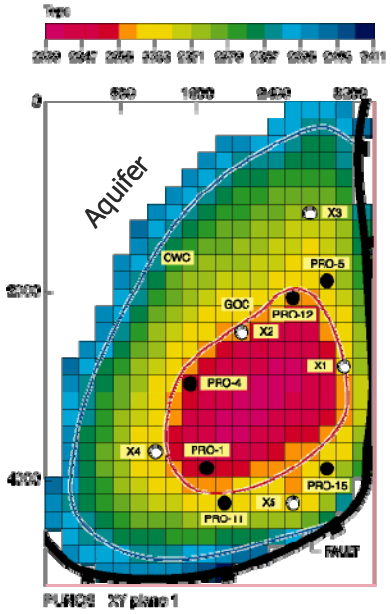


Figure 38. Top surface map showing well locations (from Floris *et al*[20].).

This section contains a comparison between the standard EnKF and the streamline assisted, SL EnKF. The document compares the RMS error of the dynamic variables like BHP and GOR, for both techniques. It also compares the variance map of the natural

logarithm of the permeability for both techniques. A summary of the cases we run is shown in **Table 2**.

Table 2. Detail of six different numerical experiment using PUNQ-S3.

Name	Conditioned covar		No members			Dynamic variable			Static variable	
	YES	NO	100	60	30	WCT	BHP	GOR	PERMX	PORO
Std EnKF 100		•	•			•	•	•	•	
Std EnKF 60		•		•		•	•	•	•	
Std EnKF 30		•			•	•	•	•	•	
SL EnKF 100	•		•			•	•	•	•	
SL EnKF 60	•			•		•	•	•	•	
SL EnKF 30	•				•	•	•	•	•	

4.4.2.1 Generation of the streamline based correlation function

In this section we will explain the details of how to build the correlation function ρ to condition the covariance. To generate the correlation function ρ, streamlines are traced from the producer. Since the streamlines are traced based on the pressure field, the regions covered by the correlation function ρ generated from the streamlines can be interpreted as the zones where the radius of investigation of the pressure drawdown has reached an end. At early assimilation times, streamlines cover regions around the well only as shown in Figure 39.(a). Later in time, streamlines cover the whole reservoir model as shown in Figure 39.(b). Streamlines arriving to each producer are affecting a different region. For water cut like measurement it is possible to further condition the flow path information by identifying those streamlines that have brokenthrough. Identification of the saturation front along streamlines can be approximated using the relationship proposed by Datta-Gupta and King (1985)[41]. Recall the saturation front can be estimated using the streamline time of flight as

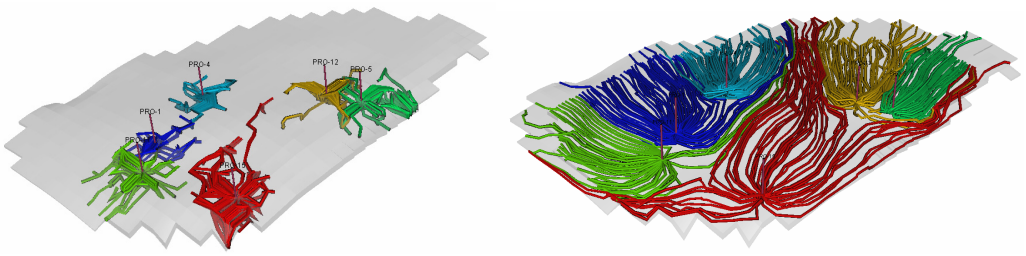
$$\frac{df_w}{dS_w} \leq \frac{\tau}{t} \quad \text{.....} \quad (4.11)$$

Where df_w/dS_w is the derivative of the fractional flow curve with respect to water saturation, τ is the time of flight along the streamlines; and t is the simulation time.

It is possible to define several types of the correlation function based in the streamline technology. In this work only two types of correlation functions were investigated. Table 3 contains a description of the criteria used to build the function.

Table 3. Different types of correlation function defined using streamlines. First row used with pressure and second row used with water saturation.

$\rho_{ij} = \begin{cases} 1 \\ 0 \end{cases}$	If grid block *i* is crossed by any streamline arriving at well *j* If grid block *i* is not crossed by streamline arriving at well *j*
$\rho_{ij} = \begin{cases} 1 \\ 0 \end{cases}$	If grid block *i* is crossed by streamlines arriving at well *j* and the grid block is behind the saturation front If grid block *i* is not crossed by streamline arriving at well *j* or the grid block is up stream of the water saturation front



(a) (b)

Figure 39. Flow path information from streamlines. (a) Zones affected by the streamlines arriving at each producer at early time (b) Zones selected by streamlines arriving at each producer at later time.

These conditions are better explained graphically as follows: A picture of the regions selected by the streamline arriving at producers at time 1 day is shown in Figure 39 (a). Notice that each ensemble member may generate a very different streamline path. After stacking all the different regions from all members, a single region is defined for each producer at each time step. Final regions belonging to different producers may overlap because they were generated by stacking regions from different members. We can think of the correlation function ρ_{ij} as a matrix with the column *j* (representing well *j*) filled with ones at the grid positions *i* selected in Figure 39 (b). Other grid blocks in the same column the correlation function is set equal to zero. A similar procedure is repeated for all others producer *j* until matrix ρ_{ij} is completed. Following the same procedure we can build the correlation function at each assimilation time.

4.4.2.2 Comparison of results from the Std EnKF and the SL EnKF

Our goal is to qualify the effects of using the SL EnKF as compared to the standard EnKF. For such purposes we run the same history matching process using different sizes on the ensemble. We use 30, 60, and 100 members in the ensemble. Each case was run using standard EnKF and the SL EnKF. Finally we compared the RMS error in the production data and RMS error in the permeability to check whether the impact of conditioning the covariance matrix is favorable.

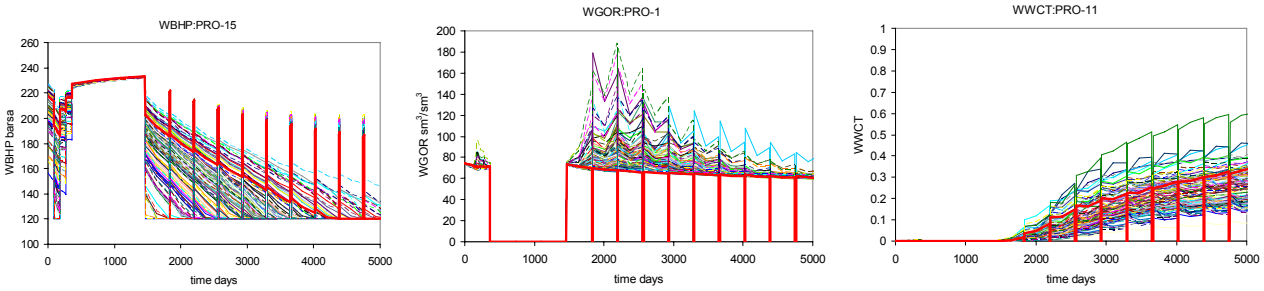
4.4.2.2.1 Comparison of dynamic variables

We start by analyzing the spread of bottom hole pressure (BHP), gas oil ratio (GOR) and water cut (WCT). Figure 40 to Figure 42 shows the initial and final bottom-hole pressure, gas to oil ratio and water cut spread for some selected wells. The spread is computed by running all the posterior models from time zero to 16.5 years.

Another interesting plot is generated when taking the mean of the ensemble to generate the response and compute the RMS error. Since after each assimilation we can generate a new mean, it is possible to check the RMS error after more and more information is incorporated/history matched. Figure 43 shows the RMS for the same data at different assimilation times. Notice both standard and streamline assisted EnKF reach similar RMS error. Computation of the RMS error from the mean of the ensemble after each assimilation step was carried out. The RMS error can be computed using equation 4.7 and 4.8

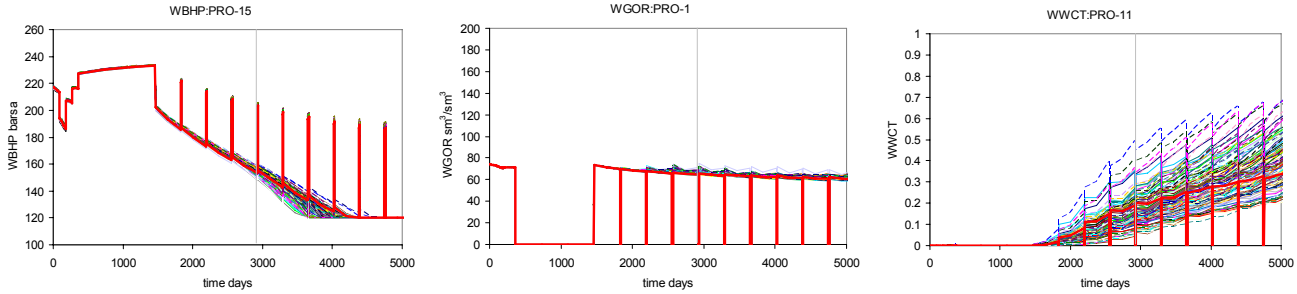
$$RMS_{well} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(d_{obs,i} - d_{cal,i}\right)^2} \quad \text{.....(4.7)}$$

$$RMS_{global} = \sqrt{\frac{1}{W}\sum_{j=1}^{W}\frac{1}{N}\sum_{i=1}^{N}\left(d_{obs,i} - d_{cal,i}\right)_j^2} \quad \text{.....(4.8)}$$



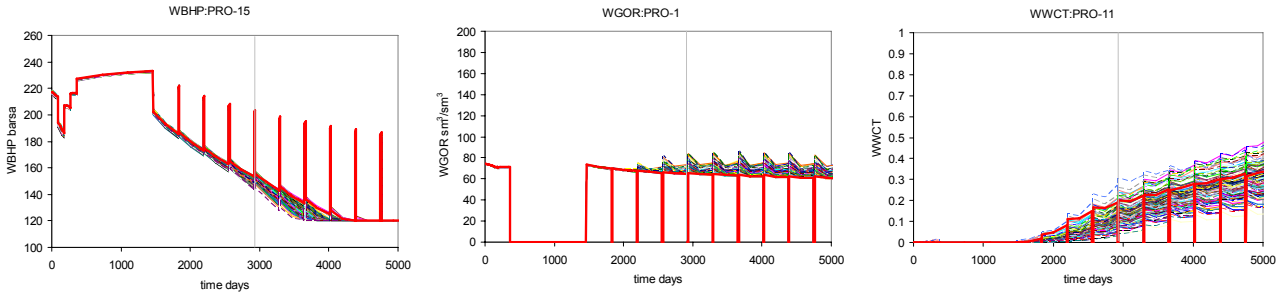
(a) (b) (c)

Figure 40. (a) Initial spread for bottom hole pressure at well PRO-15 (b) gas to oil ratio at well PRO-1 and (c) water cut at producer PRO-11. Response from 100 members in the ensemble; red bold line shows the data for the reference model.



(a) (b) (c)

Figure 41. (a) Final bottom hole pressure spread at well PRO-15, after standard EnKF data assimilation, (b) gas to oil ratio at well PRO-1 and (c) water cut at producer PRO-11. Response from the updated 100 members in the ensemble; red bold line shows the data for the reference model. Gray line around 3000 days shows time up to which the information was assimilated; all models were re-run from time zero to avoid possible increase in the material balance error.



(a) (b) (c)

Figure 42. (a) Final bottom hole pressure spread at well PRO-15, after SL EnKF data assimilation, (b) gas to oil ratio at well PRO-1 and (c) water cut at producer PRO-11. Response from the updated 100 members in the ensemble; red bold line shows the data for the reference model.

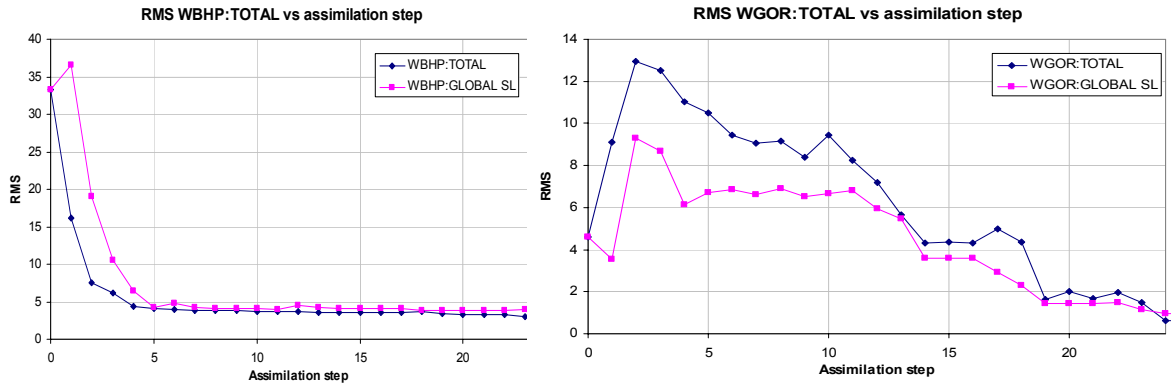


Figure 43. RMS error for the dynamic variable. (a) Shows the global bottom hole pressure RMS error for the streamline assisted EnKF and for the standard EnKF, (b) shows the global RMS error of the gas oil ratio. The error in the dynamic variables is decreased to a similar extend for both techniques.

Here d_{obs} is a vector of the dynamic variable of interest, d_{cal} is a vector of the response from the simulator using the posterior mean from the ensemble as the input parameters, i is a time index and N is the number of time data points in the observation, W is the number of wells with observation of a given type. It is possible to compute the RMS error per well (see equation 4.7) or a global RMS error (see equation 4.8).

Major error for the GOR belongs to well PRO-1. This well is located at the top of the formation near the gas oil contact. And thus the GOR in this well is very sensitive to changes in pressure. The error in the bottom hole pressure decreases with a few assimilation steps while the error in the GOR requires more assimilation step. This suggests that the relationship between the BHP and the log of perm is more linear than the relationship between the GOR and the log of perm.

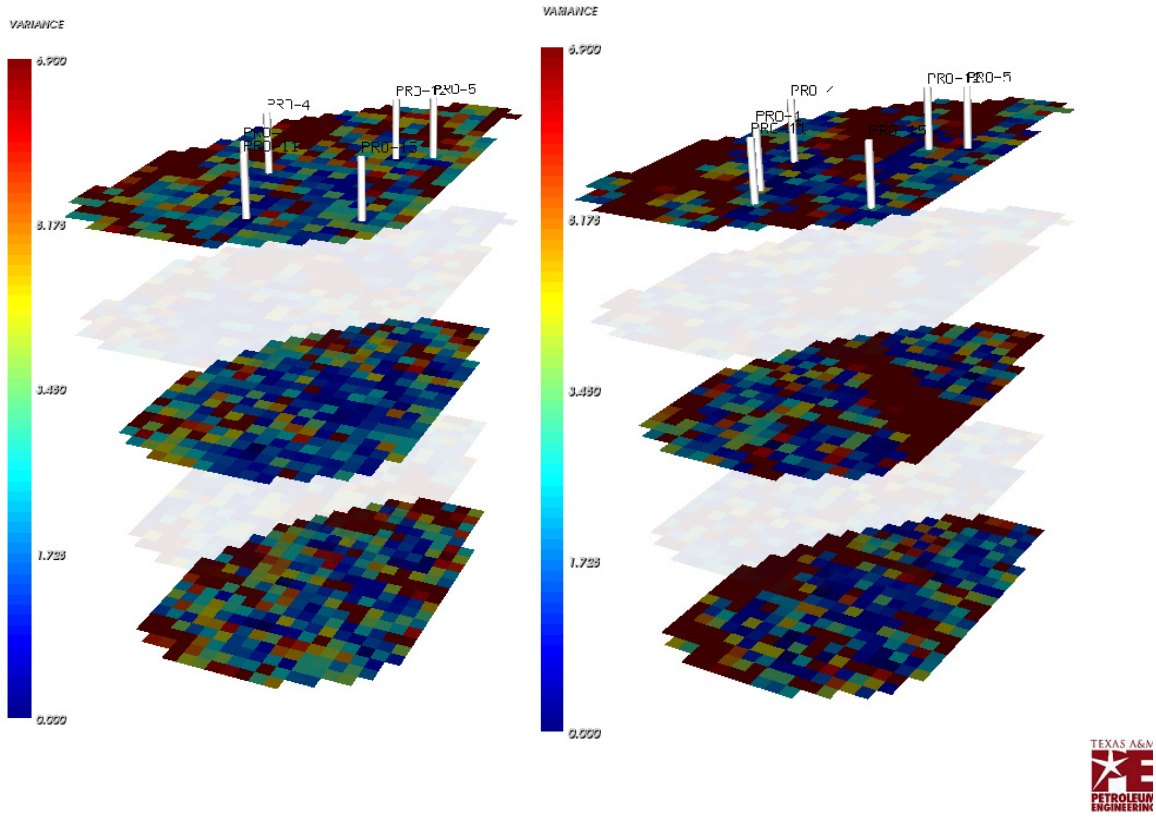
Both methodologies decrease the difference between the observed and calculated as we increase the number of assimilation steps. The RMS error falls down very quickly at the beginning and then it tends to present an asymptotic behavior. The standard EnKF and the SL EnKF go to the same level of RMS error. This means that we can say nothing about which one is more reliable based on the RMS error of the dynamic variable. It will be shown that the advantages of the SL EnKF are more evident when we analyze the error in the static variable

4.4.2.2.2 Comparison of static variables

The purpose here is to check whether the mean from the updated members are close to the real permeability. After computing a grid block variances from the ensemble those grid blocks with high values of the permeability indicates that the updated permeabilities are far away from the true; those grid blocks with a lower value of the variance indicate that the updated permeabilities are in good agreement with the true. To compute the variance we use the reference permeability as our true and the members from the ensemble as one sample from the random variable. Thus the variance at each grid block can be easily computed as

$$\mathrm{var}(\ln(k)) = \frac{1}{M-1}\sum_{i=1}^{M}\left(\ln(k_i) - \ln(k_{ref})\right)^2 \quad \text{.....(4.9)}$$

Where *ln(k)* is the natural logarithm of the permeability for the ensemble member, *i* and k_{ref} is the reference permeability. This operation is computed at every grid block. Figure 44 shows a comparison of the variance map. The maps were computed using the updated permeability members at time 2936 days. As expected the variance is small around the well position. This is especially evident at layer 4 and 3 where the wells are completed. A closer look of the variance maps, from both techniques, reveals that the in the overall there are more zones with small variance when using the SL EnKF that when using the standard EnKF. Although here we show the maps from the 60 members ensemble only, similar results were obtained with ensembles size of 30 and 100 members. Previous results are confirmed when computing the RMS error of the permeability field using equation 4.10.



(a) (b)

Figure 44. 60 members ensemble variance map; map computed using the updated permeabilities at assimilation time 2936 days; layer 2 and 4 were set to a transparent appearance to allow other layer being shown in detail; (a) 60 member's ensemble using the SL EnKF (b) 60 member's ensemble using the standard EnKF; the high strike in the variance map for the standard EnKF indicated that the posterior permeability in all the members are more apart from the true permeability as compared with the SL EnKF.

Other possible way to check for errors in the estimated permeability is to compare the mean permeability from the ensemble. Using the mean from the ensemble it is possible to compute an RMS error for the static variables as follows

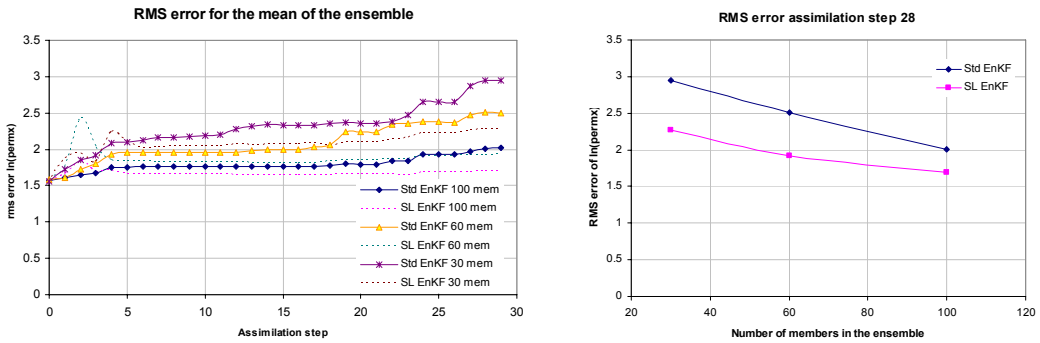
$$RMS_{\ln(k)} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\ln(k_{true,i}) - \ln(k_{mean,i})\right)^2} \quad \text{.....(4.10)}$$

Where M is the number of grid blocks k_{true} is the reference of true permeability, k_{mean} is the mean permeability from the ensemble.

The RMS error was computes for all the experiment shown in Table 2. Figure 45 (a) shows the result for this calculation. Notice that the RMS error for the natural logarithm of the permeability is less for the streamline assisted EnKF as compared to the standard EnKF, especially at later assimilation steps. This tendency is the same for all the cases, with different ensemble size, tested here. Figure 45 (b) shows a comparison of the RMS for different number of members in the ensemble. Again the SL EnKF presents better result. The difference on the RMS error becomes more evident as the number of members in the ensemble decreases. It seems based on the tendency of the curves that given a very high number of members the streamline assisted EnKF and the standard EnKF will have similar results. This behavior is expected and could be explained as follows; it is clear that if we have a large number of members with a good spread distribution the calculation of the cross-covariance at remote points will become small or zero by themselves. Thus, the effect of using the streamline assisted EnKF which is basically damping or zeroing those cross-covariance at remote points will be automatically performed by the cross-covariance if the ensemble size is large enough.

Summarizing the results from this experiment we found that the advantages of the SL EnKF become clear while analyzing the static variable because the permeability field is more reliable than the standard EnKF; it seems there is not clear advantages if we compare the error in the production data predictions. Both techniques present similar level of results.

Because the updated permeabilities in standard EnKF and the SL EnKF produces similar forecast and the SL EnKF has a more realistic updated permeability it is logical to conclude that the SL EnKF is a superior methodology when compared to the standard EnKF (at least for a moderated to small ensemble size)



(a) (b)

Figure 45. RMS error from the mean of the ensemble. For this synthetic case the reference permeability is known, thus the RMS error can be computed. Similar result can be inferred from the multidimensional variance plot map. (a) RMS error for the natural logarithm of the permeability after each assimilation step. (b) RMS error for the natural logarithm of the permeability for the standard EnkF and the streamline assisted EnKF for different number of members in the ensemble.

4.4.3 The Goldsmith field case study

Section 4.1.3 gives a complete description of the field characteristic; the section also shows that the standard EnKF has serious limitations with this model because some overshooting/undershooting problems as wells as the limitation to work with a non-Gaussian permeability distribution. These limitation suggest that the standard EnKF should no be recommended for this type of fields. Luckily the SL EnKF seems to behave better when applied to non-Gaussian distribution; it will be shown that this characteristic makes the SL EnKF more suitable for non-Gaussian distribution than the standard EnKF. Our purpose now is to study the behavior of the SL EnKF with this field case.

In this section we demonstrate the feasibility of the approach for field studies by application to a large-scale 3-D example. We repeated the same history matching but this time we conditioning the covariance matrix using the streamline path information. To select the grid blocks we used those streamlines arriving to a producer that have brokenthrough at a given time (see equation 4.6). Figure 46 shows the final regions selected by all the streamlines arriving at producer P-6 Notice how at early assimilation steps a broad region is selected. As expected grid blocks are clustered around the producer at any time. The shape of the region can change based on the injector schedule and/or breakthrough times

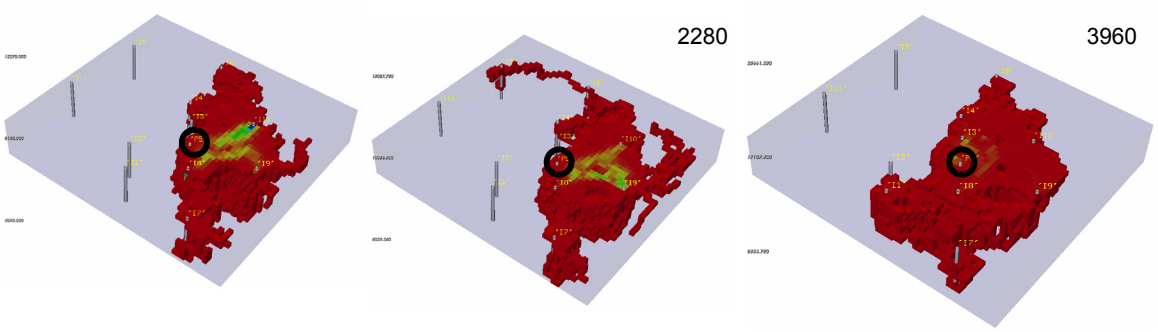


Figure 46. Regions selected by the streamlines arriving to producer P6 (black circle) at different times; (right) time 1680 days, (center) 2280 days (left) 3960 days. Region change with time because the schedule changes and the pressure field changes. This information was used to condition the covariance matrix.

To compare the benefits of using the standard EnKF and the SL EnKF we compared the RMS error of the dynamic variables. The water cut RMS error for both cases are similar. The real advantage of using the SL EnKF becomes evident during the analysis of the final permeability. Figure 47 shows the updated permeability for some of the members in the ensemble. Notice how the SL EnKF tends to preserve the initial permeability distribution. Furthermore the strong overshooting problems presented while using the standard EnKF fully disappeared.

An analysis of the water cut spread before and after assimilation reveals that the SL EnKF is able to match the observation; Figure 48 shows the initial and updated spread for wells P1, P7 and P8, observation in red.

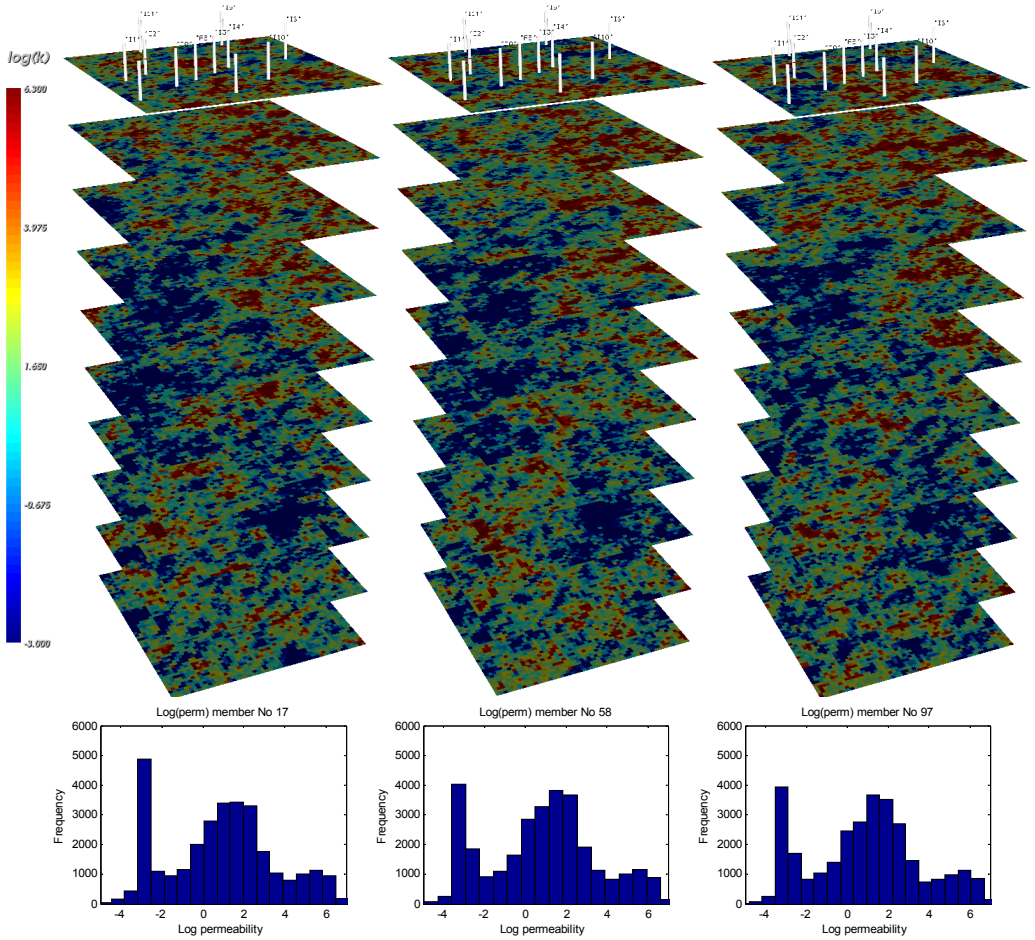


Figure 47. Updated permeability map using the SL EnKF after water cut assimilation; below each map it shows the corresponding histograms. Notice the SL EnKF is able to preserve the prior density function; this is a required characteristic of any good history matching technique.

Finally, recall from the discussion in section 4.1.3 about the changes in the models proposed by the standard EnKF; It was found that the changes were posted indiscriminate of the position of the wells and the injector. An analysis of the changes proposed by the SL EnKF (see Figure 49) reveals that changes are small and localized in zones were it is most likely the flow of water may be occurring. Notice how towards the stagnation points, close to the boundaries and the corners of the model, the value of the change proposed by SL EnKF are zero or close to zero. This behavior is a good indicator of the wise, discriminatory and preferential type of changes in the permeability field when using the SL EnKF approach.

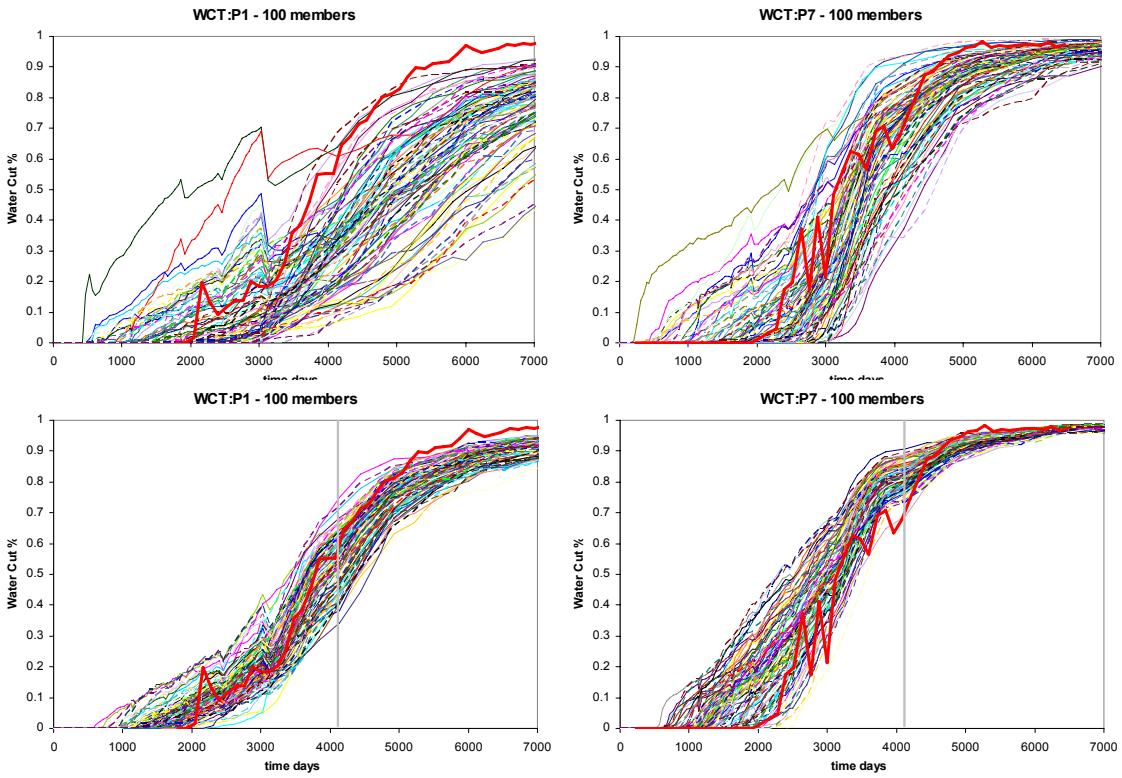


Figure 48. Initial (top row) and final, after SL EnKF (bottom row) water cut spread at wells P1, P7 100 members in the ensemble; after assimilation the spread was greatly decreased around the observation shown in red; gray line around 4000 days shows the time up to which the information was assimilated.

4.5 Parallel Implementation of the Algorithm and Scalability Analysis

Most of the examples presented in this thesis run with an acceptable performance in terms of time. Synthetic cases with grid sizes of 50x50x1 and 100 members in the ensemble typically took from 2-3 hours to assimilate 25-40 observation times, in a single CPU computer. On the other hand, while running the field case example on single CPU computer, it would take from 20 to 25 hours to assimilate the observation data for an ensemble of 100 members. Table 4 shows the total running time expend when assimilating data for the Goldsmith field case with 100 members in the ensemble. The test were run in a 128 CPU shared memory computer at the Texas A&M University

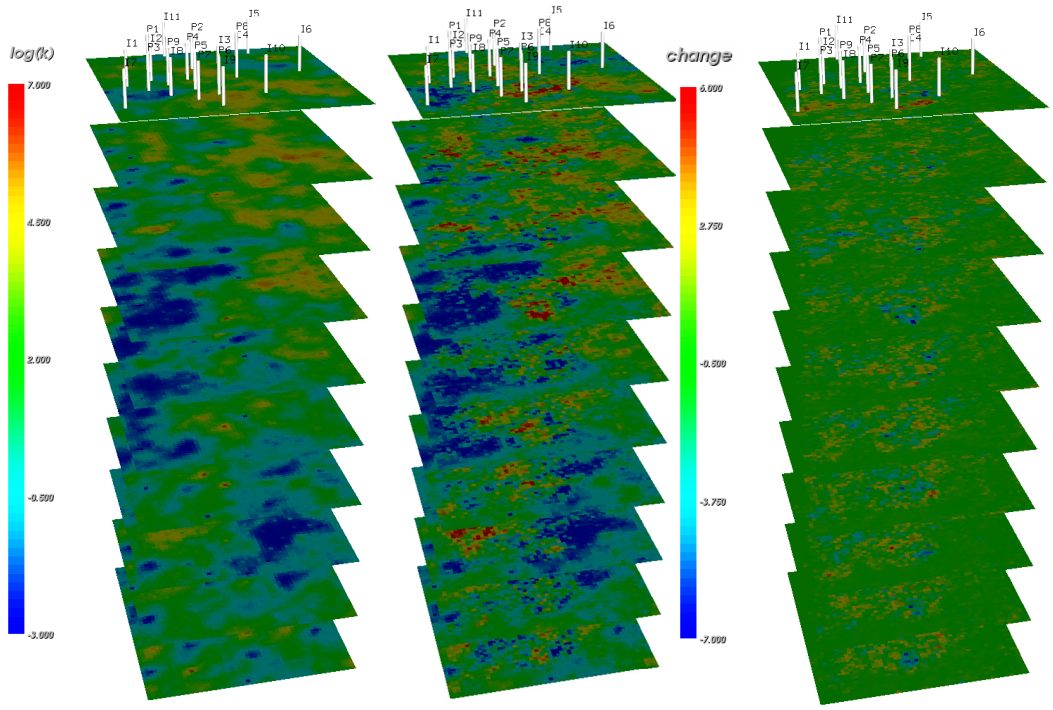


Figure 49. (Left) Initial mean from the 100 members ensemble; (center) mean from the 100 members ensemble after SL EnKF update.; (right) changes in the permeability it shows the changes have discriminately occur in few zones around the reservoir field. The changes are small, realistic and localized as compared with the one proposed by the standard EnKF (see Figure 20).

Table 4. EnKF Running time for the Goldsmith field case.

No CPU	Time hours	Speedup
1	25.057	1
2	14.353	1.745824
4	7.855	3.190048
6	5.725	4.376971
8	4.659	5.378309
10	3.871	6.472804
12	3.622	6.917561
14	3.339	7.504576
16	3.171	7.901629

Realistic reservoir models nowadays consist of hundred even millions of cells, thus only one forward run might take several hours. Given the current complexity of real reservoir cases; it is likely to think that attempt to assimilate production data using the EnKF in a single CPU computer would be prohibitory expensive to implement. Fortunately parallel computer are common and cheap nowadays. It is quite common to

find computer clusters even in small companies. Computer with parallel capabilities have given a further boost to the application of parallel computing in forward models and in history matching.

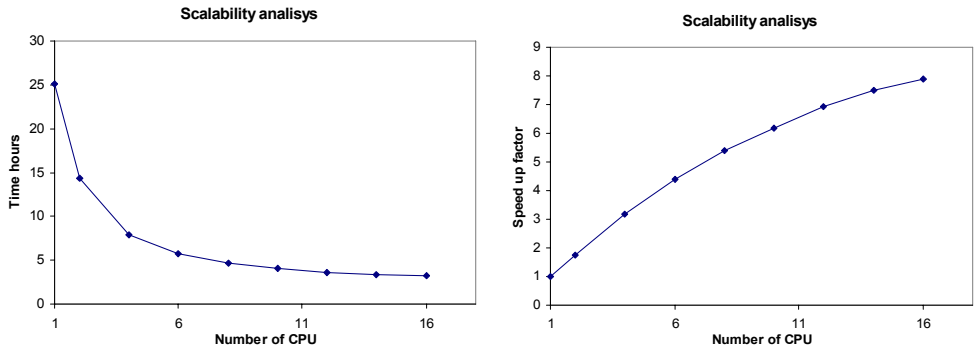


Figure 50. (a) total running time for the Goldsmith field case as a function of the number of processors. (b) Speed up factor for the same dataset; notice front the speed-up factor analysis the current implementation could in theory decrease the total running time by a factor of eight when using 16 CPUs.

The current EnKF computer code was designed following an object like structure; it is quite general, can handle very complex schedules changes, it can run with FRONTSIM and ECLIPSE, can handle two and three-phase problems; it is highly configurable and can be run for any configuration in the state vector, user can select the structure and number of variables in the state vector by modifying a few keywords in the EnKF input file. It was programmed using FORTRAN 95 and has been compiled under Windows and UNIX using the Intel FORTRAN compiler version 9.0. The parallel capabilities make use of the Menssage Parsing Interface (MPI). The parallel version of the code has been tested in a multiple CPU shared memory computer running UNIX only.

In a nutshell the EnKF is Monte Carlo technique thus parallelization of the code can be easily archive. Several authors have recognized the feasibility of parallel computation for any EnKF analysis. During the development of the present research project a parallel version of the EnKF was done. The more simplistic approach to parallelize any EnKF implementation will require running each forward model in parallel. Further parallelization can be done by parallelizing the Kalman update equation this approach is discussed by several authors[29,28].

In our implementation each member is forwarded in a different CPU (node). Result from each members are written, by the simulator in binaries files, to a common network file system available to each CPU; after each forward run, information is retrieved by the master node with a minimal level of interprocessor communication.

We used the Message Parsing Interface MPI to parallelize our code. Tests were run in 128 CPU shared memory super computer at Texas A&M University. Our test demonstrate that the EnKF scales very well. See Figure 50. Nevertheless, on the scalability plot, it was expected to have a quasi-linear behavior but our code does not. This may be explained by the fact that our code uses a shared memory computer, therefore all the CPU share the same hard drive. While the simulator is running, in each different CPU, it writes information to the hard drive. Then, each CPU has to compete with the others when writing to disk. This reduces the expected speed for more than 16 CPUs, but still reduces the running time by a factor of 9.

It is clear from our experiment that with the current computer power the feasibility of using the SL EnKF with moderate to large reservoir models is possible

CHAPTER V

CONCLUDING REMARKS AND RECOMMENDATIONS

5.1 Conclusions

A novel approach to three-phase flow using continuous model updating using the EnKF and the selective flow path information from Streamlines is presented. The power and utility of our proposed approach is demonstrated using both synthetic and field examples. The synthetic case includes matching of bottom hole pressure, water cut and gas oil ratios and is used to validate the method. The field example consists of 31 producers and 11 injectors. The permeability changes were found to be reasonable and geologically realistic. Some specific conclusions from this paper are summarized below.

1. The streamline covariance localization appears to eliminate and/or reduce previously reported problems when using the standard EnKF implementation. Some of the reported problems are: overshooting of the reservoir parameters, limitations of the standard EnKF to work with non-Gaussian distribution.
2. Comparison of the standard EnKF and the streamline EnKF regarding the RMS error in the reservoir static parameters (permeability, porosity, etc) the SL EnKF presents much better accuracy; this characteristic becomes more evident while using a reasonable small number of members in the ensemble.
3. When comparing the standard EnKF and the streamline EnKF regarding the uncertainty and the RMS error in the forecasted production data; both approaches present similar lever of accuracy.
4. Overall, the results presented here suggest that the covariance localization may provide dramatic improvements to the quality of the ensemble from the standard EnKF or its variants. It is likely that the cost of localizing the

covariance will be significant less that the cost of generating a large enough ensemble for the error to be similar.

5. The proposed approach allows the use of small number of members in the ensemble; this is critical for application of the EnKF in real field cases. Futhermore, the parallel implementation of our code scale very favorably with respect to the number of grid blocks, and the number of members in the ensemble making the approach suitable for history matching detailed geologic models.
6. The use of travel time inversion instead of amplitude inversion seems to slightly decrease under/overshooting problems.
7. Although it is possible to use travel time inversion instead of amplitude inversion the use of travel time inversion requires to rerun the simulator from time zero for every new assimilation; this greatly decrease the efficient of the EnKF algorithm and could make it prohibitive expensive to implement with large reservoir models.
8. The use of the normal score transform seems to work around the limitation of the EnKF to work with Gaussian distribution only; when using the normal score transform the posterior PDF preserved the prior PDF distribution. Nevertheless, for some cases the posterior permeability seems to have some unrealistic distribution.
9. This work demonstrates that the EnKF can be efficiently implemented on massively parallel computer with a minimal level of interprocessor communications. In this study, a serial model is used and each ensemble member is time stepped on a separate processor.

5.2 Recommendations

- Some ideas for future research are: It is well known that the computation of the sensitivity matrix is sometimes too expensive; thus an possible way to compute it could be using equation 3.23; which suggest that the effect of the sensitivity matrix over C_M can be approximated using the cross covariance $\mathbf{C}_{M^s,d}$; can this

ideas be used in conjunction with a quasi-Newton algorithm to find a MAP estimate[‡‡] ?

- The scaling problem and/or sequential EnKF when using different types of observation has been studied in detail in the weather forecasting scene[6,39]. However, to our knowledge, this issued has not been mentioned in any EnKF History Matching paper. As a suggestion it may interesting to study in more detail the impact of assimilating different types of measurement using a sequential algorithm and the standard in history matching and write a paper about it. The current code supports sequential processing, nevertheless we did not have enough time to check for bugs while using this option.
- Kalman filter can be used to assimilate time lapse seismic information or 4D information. It will be nice to include this capability in the current code.
- Although the code was designed to support to include several type of static variables (i.e permeability, porosity) we did not debug this option in detail to check for possible bugs. It will be nice to try several examples where we updated the permeability and porosity at the same time. And debug the result looking for possible bug when this option is active.
- It will be nice to modify the current design of the parallel implementation of the code, such that the code could be run in a distributed memory computer. This will in theory increase the scalability of the code.

[‡‡] Computing the effect of **G** over a vector o a matrix to find the MAP estimate was proposed by Chu, Komara and Schatzinger[38] in the technique known as the dual loop.

REFERENCES

1. Vasco, D.W., Yoon, S., and Datta-Gupta, A.: "Integrating Dynamic Data into High-Resolution Reservoir Models Using Streamline-Based Analytic Sensitivity Coefficients," *SPE Journal* (December 1999) 389.
2. Emanuel S. and Milliken W. J.: "History Matching Finite Difference Models with 3D Streamlines," paper SPE 49000 presented at the 1998 Annual Technical Conference and Exhibition, New Orleans 27-30 September.
3. Sen, M., Datta-Gupta, A., Stoffa, P., Lake, L. W. and Pope, G. A.: "Stochastic Reservoir Modeling Using Simulated Annealing and Genetic Algorithm," *SPE Formation Evaluation*, (1995) **10** (1), 49-55.
4. Sen, M., Datta-Gupta, A., Stoffa, P., Lake, L. W. and Pope, G. A., "Stochastic Reservoir Modeling Using Simulated Annealing and Genetic Algorithm," *SPE Formation Evaluation*, 1995 **10** (1), 49-55.
5. Evensen, G.: "Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte Carlo Methods to Forecast Error Statistic," *J. of Geophysical Research*, (1994), 162, 143.
6. Nævdal, G., Mannseth, I., and Vefring, E.H.: "Near-Well Reservoir Monitoring Through Ensemble Kalman Filter," paper SPE 75235 presented at the 2002 Improve Oil Recovery Symposium, Tulsa, Oklahoma, 13-17 April.
7. Nævdal, G. *et al*.: "Reservoir Monitoring and Continuous Model Updating Using Ensemble Kalman Filter," paper SPE 84372 presented at the 2003 Annual Technical Conference and Exhibition, Denver, 8-5 October.
8. Gu, Y. and Oliver, D. S.: "History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter," paper SPE 89942 presented at the 2004 Annual Technical Conference and Exhibition, Houston, 26-29 September.
9. Gao, G., Zafari, I., and Reynolds, A.C.: "Quantifying Uncertainties for the PUNQ-S3 Problem in a Bayesian Setting with RML EnKF," paper SPE 93324 presented at the 2002 Improved Oil Recovery Symposium, Tulsa, Oklahoma, 13-17 April.

10. Zafari, M., and Reynolds, A.C.: "Assessing the Uncertainty in Reservoir Description and Performance Prediction with the Ensemble Kalman Filter," paper SPE 95750 presented at the 2005 Annual Technical Conference and Exhibition, Dallas, 9-12 October.
11. Skjervheim, J.A. *et at*.: "Incorporating 4D Seismic Data in Reservoir Simulation Model Using Ensemble Kalman Filter," paper 95789 presented at the 2005 Annual Technical Conference and Exhibition, Dallas, 9-12 October.
12. Wen, X.H., and Chen W.H.: "Real-Time Reservoir Model Updating Using Ensemble Kalman Filter," paper SPE 92991 presented at the Reservoir Simulation Symposium. Houston, 31 January – 2 February.
13. Gu, Y., and Oliver, D.S.: "The Ensemble Kalman Filter for Continuous Updating of Reservoir Simulation Models," *J. of Energy Resources Technology* (March 2006), 79.
14. Evensen, G.: "The Ensemble Kalman Filter: theoretical formulation and practical implementation," *Ocean Dynamics* (2003), **53** pp 343-367.
15. Houtekamer, P.L. and Mitchell, H.L.: "Data Assimilation Using an Ensemble Kalman Filter Technique," *J. Monthly Weather Review* (1998), **126** pp 796.
16. Burgers, G., Leeuwen, P., and Evensen, G.: "Analysis Scheme in the Ensemble Kalman Filter," *J. Monthly Weather Review* (1998), **126**, 1719-1724.
17. Epstein, E.S.: "Stochastic Dynamic Prediction," *Tellus, Ser*.(1969) **A 21**, 739-759.
18. Tarantola, A.: *Inverse Problem Theory and Method for Model Parameter Estimation*, SIAM, Philadelphia, (2005).
19. Scales, J. Smith, M.L. and Treitel S: *Introduction to Geophysical Inverse Theory*, Samizdat, Golden (2001).
20. Backus, G: Inference From Inadequate and Inaccurate Data, *Proc National Academic of Science*. (1970) **65**, 1-298.
21. Anderson, J.L.: "A Local Least Squares Framework for Ensemble Filtering," *J. Monthly Weather Review*, (2003) **131**, 634-642.
22. Lorentzen, R.J. *et al.*: "Underbalanced and Low-head Drilling Operations: Real Time Interpretation of Measured Data and Operation Support," paper SPE 71384 presented at the 2001 Annual Technical Conference and Exhibition, New Orleans, 30 September – 3 October.

23. Brouwer, D.R. *et al*.: "Improved Reservoir Management Through Optimal Control and Continuous Model Updating," paper SPE 90149 at the 2004 Annual Technical Conference and Exhibition, Houston, 26-29 September.
24. Anderson, B.D., and Moore, J.B.: *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, N.J. (1979).
25. Reichele, R.H., McLaughlin, D.B., and Entekhabi, D.: "Hydrologic Data Assimilation with the Ensemble Kalman Filter" *J. Monthly Weather Review* (2002), **130**, 103-114.
26. Kamen, E.W., and Su, J.K.: *Introduction to Optimal Estimation*, Springer, London ; New York (1999).
27. Brown, R.G., and Hwang, Y.C.: *Introduction to Random Signals and Applied Kalman Filtering : With MATLAB Exercises and Solutions*, Wiley, New York:, (1997).
28. Houtekamer, P.L. and Mitchell, H.L.: "A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation," *J. Monthly Weather Review* (2001), **129** 123.
29. Keppenne, CL.: "Data Assimilation into a Primitive-Equation Model with a Parallel Ensemble Kalman Filter," *J. Monthly Weather Review* (2000), **128**, 1971-1981.
30. Aziz, K. and Settari, A.:, *Petroleum Reservoir Simulation*, Applied Science Publishers, Essex, England (1979).
31. Dake, L.P.: *Fundamentals of Reservoir Engineering*, Springer, Amsterdam (1978).
32. Deutsch, C.V., and Journel, A.:, *GSLIB Geostatistical Software Library and User's Guide*, Oxford U. Press, New York (1992).
33. Journel, A.G., and Alabert, F.G.: "New Method for Reservoir Mapping," *JPT*. (February 1990), 212.
34. He, Z., and Datta-Gupta, A.: "Streamline-based Production Data Integration Under Changing Field Conditions," *SPE Journal*, (December 2002), 7 (4).
35. Cheng, H., Datta-Gupta, A. and He Z.: "A Comparison of Travel Time and Amplitude Inversion for Production Data Integration into Geologic Models: Sensitivity, Non-linearity and Practical Implications," *J. SPE*, (2005) 10(1), March,75-90.
36. Hamill, T.M., and Whitaker, J.S.: "Distance-Dependent Filtering of Background Error Covariance Estimate in an Ensemble Kalman Filter," *J. Monthly Weather Review* (2001), **129** pp 2776.

37. Cheng, H. *et al*.: "Fast History Matching of Finite-difference Models Using Streamline-derived Sensitivities," *SPE Reservoir Evaluation and Engineering*, (October 2005), **8** (5), 426-436.
38. Jimenez, E., *et al*.: "Spatial Error and Convergence in Streamline Simulation," paper 92873-MS presented at the 2005 Reservoir Simulation Symposium, The Woodlands, Texas, 31 January-2 Feburary.
39. Floris, F.J.T *et al*.: "Methods for Quantifiying the Uncertanty of Production Forecast: A Comparative Study," *Petroleum Geoscience* (2001) **7**, 87.
40. Baker, J.W., Cuypers, M., and Holden, L.: "Quantifying Uncertainty in Production Forecasts: Another Look at the PUNQ-S3 Problem," *J. SPE* (December 2001), 433-441.
41. Datta-Gupta, A. and King, M. J.: "A Semianalytic Approach to Tracer Flow Modeling in Heterogeneous Permeable Media," *Advances in Water Resources*, (1995) **18** (1), 9-24.
42. Chu, L; Komara, M; and Schatzinger, R.A.: "Efficient Technique for Inversion of Reservoir Properties Using Iteration Method," paper SPE 36512 at the 1996 Annual Technical Conference and Exhibition, Denver, 6-9 October.

APPENDIX

Here our goal is to understand how conditioning the covariance matrix using the streamline flow path information will be advantageous. For that purpose we will follow the same approach presented by Hamill and Whitaker[17]. In fact the proof showed here was provided by Dr Hamill, personal communications. The idea is to understand the basic effects of what will happen to the state vector if the covariance matrix contains an error in the cross-covariance calculations. For the sake of simplicity we will use the simplest system possible. A single observation and Gaussian statistics. Let us assume our random state vector $\mathbf{Y}=\{M_1,D_2\}^t$ represent the unknown true model. Prior information tell us that the mean from the prior distribution looks like $\mathbf{y^p}=\{m^p,d^p\}^t$ Our prior covariance matrix is given by equation A.1

$$C_{\Psi}^{p} = \begin{bmatrix} \sigma_m^2 & c_{md} \\ c_{md} & \sigma_d^2 \end{bmatrix} \quad \text{..... (A.1)}$$

Since we have no more information available the best we can do is to represent our state vector distribution as $P(Y) = N(y^p, C_{\Psi}^{p})$. Were N indicates that the distribution in normal with mean y^p and covariance matrix C_{ψ}^{p}. Now assume a new observation d_{obs} becomes available. This observation has some error; the error in the observation is defined by $d_\varepsilon = N(0,C_D)$

Using a Bayesian framework it is easy to show that the updated posterior estimate and its variance will be given by

$$\begin{Bmatrix} m^a \\ d^a \end{Bmatrix} = \begin{Bmatrix} m^p \\ d^p \end{Bmatrix} + \begin{Bmatrix} \dfrac{c_{md}}{\sigma_d^2 + C_D} \\ \dfrac{\sigma_d^2}{\sigma_d^2 + C_D} \end{Bmatrix} \left(d_{obs} - d^p\right) \text{..... (A.2)}$$

$$C_{\Psi}^{a} = \begin{bmatrix} \sigma_m^2 & c_{md} \\ c_{md} & \sigma_d^2 \end{bmatrix} + \begin{bmatrix} \frac{c_{md}^2}{\sigma_d^2 + C_D} & \frac{c_{md}\sigma_d^2}{\sigma_d^2 + C_D} \\ \frac{c_{md}\sigma_d^2}{\sigma_d^2 + C_D} & \frac{\sigma_d^4}{\sigma_d^2 + C_D} \end{bmatrix} \quad \ldots\ldots \text{(A.3)}$$

Since we want to understand why condition the covariance matrix in certain cells can be beneficial. Let us assume there is an error, or noise $\varepsilon_c \sim N(0,\tau_{md})$ in the estimation of the cross covariance calculation c_{md}. Therefore our inaccurate covariance matrix $\hat{C}_{\Psi}^{p}$ will be given by

$$\hat{C}_{\Psi}^{p} = \begin{bmatrix} \sigma_m^2 & c_{md} + \varepsilon_c \\ c_{md} + \varepsilon_c & \sigma_d^2 \end{bmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(A.4)}$$

We are particularly interested in what will happen to the model parameter m. Thus, move our attention only to the parameter variable m. If the error ε_c is uncorrelated with the errors m and d; them it can be proved that the mean and the variance of the posterior estimate will be given by

$$m^a = m^p + \frac{c_{md} + \tau_{md}}{\sigma_d^2 + C_D}(d_{obs} - d^p) \quad \ldots\ldots\ldots\ldots \text{(A.5)}$$

$$Var(m^a) = \sigma_m^2 - \frac{c_{md}^2}{\sigma_d^2 + C_D}\left(1 - \frac{\tau_{md}^2}{c_{md}^2}\right) \quad \ldots\ldots\ldots \text{(A.6)}$$

It is easy to notice that the posterior estimate of m^a will tend to overshoot/undershoot over a sequence of many updates. Notice that when the ratio (τ_{md}/c_{md}) is greater that 1.0, the posterior m^a is typically degraded, its variance increase instead of decrease, by assimilating observation d_{obs}.

Given the true covariance matrix C_{ψ}^{p} with variances $\sigma_m^2=1$, $\sigma_d^2=\eta^2$ and true crosscovariance c_{md} then the true correlation coefficient will be given by $\rho = c_{md}/(\sigma_m\sigma_d)$,

using this definitions Hamill and Whitaker[17] showed that the variance τ_{md} of the error in the calculation of the covariance can be approximately

$$Var(\varepsilon_c) = \tau_{md} \approx \frac{1}{n}(1+\rho^2)\eta^2 \quad \text{..... (A.7)}$$

More details about the derivation of this expression can be found in Hamill and Whitaker[17]. Now since it is well known that the correlation function decreases with increasing distance from the observation it is interesting to study how the ratio (τ_{md}/c_{md}) changes as the ensemble size and the correlation ρ changes. Figure 51 shows the corresponding values of the ratio (τ_{md}/c_{md}).

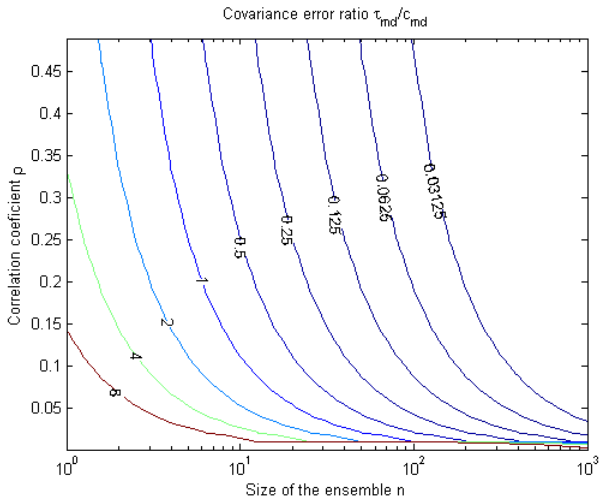


Figure 51. Covariance error ratio for different values of correlation coefficient and size of the ensemble.

The figure shows that for a small number of members in the ensemble, say 30 those remote grid blocks were the correlation coefficient is lower than 0.15 will have a covariance error ratio higer than one. Equation A.6 shows that when this ratio is greater that one, assimilation of the observation is not longer helpful; thus condition the covariance matrix using the streamline path information may be beneficial.

VITA

Elkin Rafael Arroyo Negrete
Petroleum Engineering Department
3116 TAMU
College Station, TX USA, 77843
Ph: (979) 847-8797
earroyo@numerica.com.co

EDUCATION

Bachelor of Science. Mechanical Engineering. Universidad Industrial de Santander, Colombia. March 1998

Master of Science. Petroleum Engineering. Texas A&M University. December 2006

PROFILE

Mechanical engineer with seven years of combined experience in structural finite element analysis, seismic processing and computer programming. Specific knowledge in seismic processing, geostatistics, reservoir and streamline simulation, finite element analysis, integrated reservoir characterization, and software development. Currently, involved in research concerning stochastic reservoir characterization and streamline simulation. Special interest in the development and application of high-resolution numerical schemes for reservoir simulation, dynamic data inversion using Monte Carlo technique, assessment of uncertainty, integration of 4D seismic, software design and development.

EXPERIENCE

Texas A&M University: Research Assistant. 2004-2006

Numerica Ltda- Colombia: Numerical Software Developer 2002-2004

Numerica Ltda- Colombia: Front End Seismic Processor. 2000-2002

Numerica Ltda- Colombia: Finite Element Application Engineer 1999-2000

Ecopetrol-Colombia: Front End Seismic Processor. 1998-1999