GEOMECHANICAL RESERVOIR MODEL CALIBRATION AND UNCERTAINTY ASSESSMENT FROM MICROSEISMIC DATA

A Dissertation

by

MOHAMMADALI TARRAHI

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

DOCTOR OF PHILOSOPHY

Chair of Committee,	Eduardo Gildin
Co-Chair of Committee,	Yousef Jafarpour
Committee Members,	Michael J. King
	Richard L. Gibson
Head of Department,	A. Daniel Hill

May 2015

Major Subject: Petroleum Engineering

ABSTRACT

Hydraulic stimulation of low permeability rocks in unconventional reservoirs has been observed to trigger microearthquakes (MEQs). Triggering of the MEQ events has been linked to the pore pressure, temperature, and in-situ stress variations which result in crack initiation. The resulting clouds of micro-seismic events are believed to carry information about the underlying coupled flow, geomechanics, and thermal processes and hence rock hydraulic and geomechanical property distributions. We develop a probabilistic framework called stochastic seismicity-based reservoir characterization (SSBRC) to integrate microseismic events to infer reservoir property distributions. To model the geothermal reservoir stimulation, a fully coupled thermo-poroelastic finite element method (FEM) model has been developed to handle the coupled process of heat transport, fluid flow, and rock deformation. To simulate the stimulation process, an alternate simplistic approach is also acquired based on a major hypothesis that MEQ events are triggered by an increase in pore pressure. Based on this hypothesis, the distribution of the resulting microseismicity clouds can be viewed as monitoring data that carry important information about the spatial distribution of rock permeability. We apply the ensemble Kalman filter (EnKF) to integrate the resulting continuous seismicity map to estimate hydraulic and geomechanical property distributions. We demonstrate that the standard application of the EnKF with such large correlated datasets can result in substantial loss of ensemble spread. We investigate three alternative implementation methods to mitigate this issue.

We first present the methodology proposed for MEQ data integration with the EnKF, followed by a number of examples of applying SSBRC to both forward

modeling methods to illustrate the uncertainty underestimation effect when the standard EnKF is applied to large-scale seismicity density map data. We then discuss the proposed methods for improving the uncertainty quantification results and illustrate the effectiveness of these methods by applying them to a number of numerical examples. We also apply and extend the proposed microseismic data integration method to unconventional reservoir with horizontal well and multistage hydraulic fractures to characterize the reservoir and induced fractures.

We also investigate the effect of variogram model uncertainty in the EnKF performance and propose a modified EnKF algorithm to handle the uncertainty in variogram parameters. We also develop a computationally efficient data assimilation procedure by employing a pseudo forecast method and geological model clustering method along with EnKF. By a set of numerical experiments, we show how the proposed fast history matching method is successful in preserving the ensemble spread and expediting the integration procedure.

TABLE OF CONTENTS

Page

ABSTRACT ii

TABLE OF CONTENTS iv

LIST OF FIGURES vii

LIST OF TABLES xviii

1. INTRODUCTION 1

2. GEOMECHANICAL FORWARD MODELING METHODOLOGY 11

2.1 Pore Pressure Diffusion Forward Model 12
2.2 Geomechanical Forward Model 15
2.2.1 Governing Equations 17
2.2.2 Rock Damage Model and Microseismic Events 19
2.2.3 Finite Element Method for Thermo-Poroelasticity 22
2.3 Continuous Microseismicity Interpretation with Kernel Density Estimation 25
2.3.1 Consideration of Uncertainty in Event Occurrence Time . . . 31
2.4 Discussion of Physical Relationship of Induced Microseismic Events and Rock Properties 36
2.5 Motivating Real Field Microseismic Data and Event Triggering Mechanism 42

3. INVERSE MODELING APPROACH IN GEOMECHANICAL RESERVOIR MODELS 48

3.1 Estimation with Ensemble Kalman Filter 50
3.2 Improved Uncertainty Quantification 59
3.2.1 Inflated Observation Error Variance 60
3.2.2 Reduced-Order Projection 61
3.2.3 Coarse-Scale Microseismicity Density Map 64

4. NUMERICAL EXPERIMENTS FOR GEOMECHANICAL RESERVOIR MODELS 66

4.1 Description of Experimental Setup: Pore Pressure Diffusion 66
4.2 Results and Discussion 68
4.2.1 Experiment 1: 2D Heterogeneous 68

4.2.2 Experiment 2: 3D Heterogeneous . . . 72
4.2.3 Experiment 3: Resolving Spread Underestimation . . . 79
4.2.4 Experiment 4: Resolving Spread Underestimation . . . 87
4.2.5 Joint Parameter Estimation in Pore Pressure Diffusion Model 93
4.3 Description of Experimental Setup: Geomechanical Model . . . 93
4.4 Results and Discussion . . . 97
4.4.1 Experiment 1: 2D Homogeneous . . . 98
4.4.2 Experiment 2: 2D Heterogeneous Tensile Strength . . . 103
4.4.3 Experiment 3: 2D Heterogeneous Elastic Modulus . . . 107
4.4.4 Experiment 4: Improving Experiment 3 . . . 114
4.4.5 Experiment 5: 2D Heterogeneous Permeability . . . 124
4.4.6 Experiment 6: 3D Heterogeneous Elastic Modulus . . . 125
4.4.7 Joint Parameter Estimation in Geomechanical Model . . . 132
4.4.8 Application of SSBRC to KTB Site . . . 139

5. APPLICATION AND EXTENSION TO HYDRAULICALLY FRACTURED RESERVOIRS . . . 142

5.1 Application of SSBRC to Horizontal Well with Multistage Hydraulic Fractures . . . 142
5.1.1 Estimating Permeability Distribution in Hydraulically Fractured Reservoir . . . 142
5.1.2 Estimating Permeability Distribution with the Assumption of Unknown Hydraulic Fractures Characteristics . . . 151
5.1.3 Estimating Hydraulic Fracture Characteristics . . . 154
5.2 Assessment of Welltest Techniques for Hydraulic Fracture Characterization and its Integration with MEQ Data . . . 162
5.3 Discussion and Recommendations for Real-time Application of SSBRC 166

6. ASSESSING THE PERFORMANCE OF THE ENSEMBLE KALMAN FILTER FOR SUBSURFACE FLOW DATA INTEGRATION UNDER VARIOGRAM UNCERTAINTY . . . 173

6.1 Introduction . . . 173
6.2 Methodology . . . 179
6.2.1 Ensemble Kalman Filter . . . 179
6.2.2 Prior Structural Model: Variogram . . . 182
6.2.3 Experimental Setup . . . 182
6.2.4 Motivating Example . . . 185
6.3 EnKF Performance Under Variogram Uncertainty . . . 189
6.3.1 Direct Estimation of Variogram Model Parameters . . . 189
6.3.2 Permeability Estimation Under Uncertain Variograms . . . 197
6.3.3 Application to SPE10 Model (Model C) . . . 204
6.4 Discussion and Summary . . . 209

7. USING EFFICIENT PSEUDO-FORECASTS AND GEOLOGICAL MODEL CLUSTERING WITH ENSEMBLE KALMAN FILTER FOR FAST HISTORY MATCHING 212

7.1 Introduction 212
7.2 Inverse Problem Methodology Development 215
7.2.1 Ensemble Kalman Filter 215
7.2.2 Efficient Ensemble Kalman Filter with Pseudo-Forecasts and Geological Model Clustering 216
7.3 Numerical Experiments 221
7.3.1 1D Reservoir Model with Homogeneous Permeability 222
7.3.2 1D Reservoir Model with Heterogeneous Permeability 225
7.3.3 2D Reservoir Model 229
7.3.4 PUNQ-S3 Model 232
7.3.5 SPE10 Model 236
7.4 Discussions and Summary 239

8. SUMMARY AND CONCLUSION 240

REFERENCES 245

LIST OF FIGURES

FIGURE Page

1.1 Proposed framework of SSBRC for EGS (*Ghassemi and Jafarpour*). . 6

2.1 Overall workflow of the pore-pressure diffusion forward model that relates permeability distribution to microseismicity density observations. 12

2.2 Overall workflow of the geomechanical forward model that relates hydraulic and geomechanical reservoir property distributions to microseismicity density observations. 13

2.3 Overall workflow diagram of the fully coupled geomechanical forward model. 24

2.4 Schematic illustration of converting discrete microseismic event measurements to continuous seismicity density observations in one dimension using kernel density estimation method. The crosses on the x axis show the reconstruction of the discrete microseismic events, while the short symmetric curves display the corresponding Gaussian kernels used to represent each event as a density function with maximum value at the location of the discrete events. The red line shows the density of the microseismic events in space as a continuous observation to be used in the EnKF. 27

2.5 Illustrating the continuous seismicity interpretation in 2D. (a) the individual events in separate plots that make the seismicity cloud of (c). (b) corresponding density maps of single events of (a) that are made by putting 2D Gaussian kernel at the location of each events. (d) the seismicity density map corresponds to the cloud in (c) which is obtained by adding (averaging) all individual densities of (b). . . . 29

2.6 Seismicity cloud evolution in 6 different time steps. 30

2.7 Seismicity density maps created by KDE method corresponding to seismicity clouds of Figure 2.6. 31

2.8 The schematic of consideration of event occurrence time in 2D model using KDE method. The well location is indicated by x mark in the middle of the field. The MEQ clouds are incremental clouds. Blue dots illustrate the events that happen in each time step. Red Gaussian curves ($s_t(t)$ in Equation (2.29)) show the density of each event in time. For simplified and clear illustration the Gaussian kernel is placed only for first two time steps. 33

2.9 The schematic of consideration of event occurrence time in 2D model using KDE method along with time step refinement. The well location is indicated by x mark in the middle of the field. The MEQ clouds are incremental clouds. Blue dots illustrate the events that happen in each time step. Red Gaussian curves $s_t(t)$, show the density of each event in time. For simplified and clear illustration the Gaussian kernel is placed only for one of the time steps as well as the time step refinement. 35

2.10 Effect of cold water injection in stress state change and rock failure. . 37

2.11 A schematic illustration to show the coupled effect of pore pressure change and stress state change on how the failure can happen (showing two different initial stress states). 39

2.12 Tornado diagram of sensitivity analysis of MEQ events with respect to playing parameters. 41

2.13 A plan view of Geysers geothermal field. The triangles are the locations of the seismic monitoring stations and the blue boxes are the locations of some of the injection wells 43

2.14 Induced microseismic events in 1994 experiment (pink section of the well is the open hole section) 46

2.15 Induced microseismic events in 2000 experiment (pink sections of the well are the open hole sections) 47

3.1 Full forward model relating geomechanical parameter distribution to microseismicity density map. 57

3.2 EnKF procedure (in each integration time step). 58

4.1 Microseismicity cloud generation in a two-dimensional model with heterogeneous permeability: (a) snapshots of diffusive pore pressure distributions at different time steps $p(\mathbf{u}, t)$, (b) the true log-permeability distribution, (c) spatially uncorrelated Gaussian (white noise) rock criticality $C(\mathbf{u})$, and (d) the cloud of microseismic events generated by comparing rock criticality with pore pressure distributions at different time steps. 70

4.2 Estimating heterogeneous permeability by SSBRC. The evolution of the estimated ensemble mean through integration steps. 71

4.3 SSBRC performance measures in estimating 2D heterogeneous permeability field. 72

4.4 Evolution of the standard deviation (Std) map of permeability ensemble through integration steps. 73

4.5 Microseismicity cloud generation in a three-dimensional model with heterogeneous permeability: (a) snapshots of diffusive pore pressure distributions at different time steps, (b) the true log-permeability distribution, (c) spatially uncorrelated Gaussian (white noise) rock criticality, and (d) the cloud of microseismic events generated by comparing rock criticality with pore pressure distributions at different time steps. 75

4.6 Estimating 3D heterogeneous permeability by SSBRC. The evolution of the estimated ensemble mean through integration steps. 76

4.7 SSBRC performance measures in estimating 3D heterogeneous permeability field. 77

4.8 Evolution of the standard deviation (Std) map of permeability ensemble through integration steps. 78

4.9 The SSBRC estimation results with inflated observation Std for a 2D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread. 81

4.10 The SSBRC estimation results with observation projection approach for a 2D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread. 84

4.11 Continuous interpretation of MEQ cloud on different sizes of coarse-scale grid configuration. (a) discrete MEQ cloud, (b)-(i) coarse seismicity density maps on different coarse grid systems. 86

4.12 Effect of coarse microseismicity density dimension on the performance of SSBRC in heterogeneous permeability estimation experiment (Initial RMSE = 1.4478). 87

4.13 The SSBRC estimation results with coarse-scale seismicity density approach for a 2D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread. 88

4.14 The SSBRC estimation results with inflated observation Std for a 3D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread. 90

4.15 The SSBRC estimation results with observation projection approach for a 3D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread. 92

4.16 The SSBRC estimation results with coarse seismicity density for a 3D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread. 94

4.17 The observed MEQ data, true permeability and criticality maps, initial ensemble means, and final estimated maps (middle row is permeability and last row is criticality) 95

4.18 True seismicity cloud (tensile failures) of homogeneous parameter experiment; $P_{ini} = 10$ (MPa), $S_{H,max} = 25$ (MPa), $S_{h,min} = 15$ (MPa), Inj. Rate = 12.5 (Lit/s), $E = 10$ (GPa), $k = 0.005$ (md), $T_0 = -4$ (MPa). The black cross shows the injection well location. 99

4.19 Damage factor and enhanced permeability distribution to show the induced fracture and its propagation. 100

4.20 Results of homogeneous elastic modulus E, estimation; plots (a) to (g) show the evolution of ensemble histogram at each integration steps. Plot (h) shows the RMSE of estimation (True $E = 10$ GPa). 101

4.21 Results of homogeneous permeability k, estimation; Plot (a) shows the $RMSE$ of estimation. Plots (b) to (e) show the evolution of ensemble histogram in integration steps (true $k = 0.005$ md). 102

4.22 Results of homogeneous tensile strength T_0, estimation; Plot (a) shows the $RMSE$ of estimation. Plots (b) to (e) show the evolution of ensemble histogram in integration steps (true $T_0 = -4$ MPa). 103

4.23 True model specification. $S_{h,min} = 20$ (MPa), $S_{H,max} = 15$ (MPa), $P_{ini} = 10$ (MPa), Inj. Rate = 15 (Lit/s), $k = 0.005$ (md), Young's modulus E, with Normal PDF (spatially random distribution) $\mu_E = 50$ (GPa), $\sigma_E = 15$ (GPa). Black cross at the center shows the injection well location. 104

4.24 Tensile strength T_0, estimation results in terms of estimated ensemble mean map, RMSE and ensemble spread. 106

4.25 Evolution of a single tensile strength T_0, realization and its corresponding MEQ cloud throughout integration steps. 107

4.26 Evolution of tensile strength T_0, ensemble standard deviation map throughout integration steps. 108

4.27 True model setup; true Young's modulus distribution and the corresponding MEQ cloud. 109

4.28 Showing the induced fracture propagation through damage factor and enhanced permeability distribution. 110

4.29 Young's modulus E, estimation results of standard SSBRC with typical range of observation error Std (the reference experiment). . . 111

4.30 Young's modulus E, estimation results of SSBRC with typical range of observation error Std (the reference spread experiment with 1000 realizations), its true model is shown in Figure 4.27. 113

4.31 Comparing the EnKF experiments with 100 (small ensemble) and 1000 (large ensemble) realizations 113

4.32 Young's modulus E, estimation results of SSBRC with increased observation error Std, its true model is shown in Figure 4.27. 115

4.33 Young's modulus E, estimation results of SSBRC with projection approach, its true model is shown in Figure 4.27. 117

4.34 Young's modulus E, estimation results of SSBRC with projection approach with the consideration of uncertainty in event occurrence time along with adaptive time step refinement. 119

4.35 Microseismicity density maps (with different sizes) for final seismicity cloud. (b-e) show coarse seismicity density maps. (f) reference seismicity density map (black cross at the center shows the injection well location). 120

4.36 Effect of coarse microseismicity density dimension on the performance of SSBRC in Young's modulus estimation experiment (initial Young's modulus $RMSE = 6.8134$ GPa). 121

4.37 Results of SSBRC with coarse seismicity density map $N_{X,red} = 10$, for Young's modulus estimation. Its true model is shown in Figure 4.27. . 123

4.38 True model setup; true k distribution and the corresponding MEQ cloud. 124

4.39 Permeability k, estimation results of SSBRC with projection approach, its true model is shown in Figure 4.38. 126

4.40 Injection rate scenario. 127

4.41 True model setup; true Young's modulus and T_0 distribution and the corresponding MEQ cloud. 128

4.42 Final (at 6^{th} integration step) true MEQ density maps on different coarse size meshes. 129

4.43 Young's modulus estimation results: evolution of estimated ensemble mean through integration steps. 131

4.44 Estimation performance measures. 132

4.45 Evolution of the standard deviation map of Young's modulus ensemble. 133

4.46 Joint Young's modulus (top) and permeability (bottom) estimation with reduced-order projection approach. 135

4.47 True specifications of 3D model for joint estimation; true tensile strength and permeability distribution, Young's modulus map, injection rate scenario and the corresponding MEQ cloud. 137

4.48 Results of joint tensile strength (top) and permeability (bottom) estimation in 3D model with projection approach. 138

4.49 The 2D configuration of KTB microseismic cloud which is used as the integration data here (the pink cross is the injection well) (events chosen from a depth interval of the 3D MEQ cloud) 140

4.50 Estimating permeability distribution form MEQ cloud in Figure 4.49 140

4.51 Ensemble spread through integration steps 141

5.1 The configuration of the multistage hydraulic fractures and the horizontal well in the reservoir. 145

5.2 True model configuration (top view), the location of the fractures and the reservoir permeability distribution. 145

5.3 The top view of the two fracture stages to illustrate the refinement configuration. 146

5.4 The distributions of true log-permeability (implemented hydraulic fractures as in Figure 5.1 are not shown here) and criticality. The criticality or critical pore pressure values come from a Gaussian random distribution with mean and standard deviation of 5850 (psi) and 550 (psi), respectively. 146

5.5 Evolution of the pore pressure distribution in the reservoir model during the hydraulic stimulation. 147

5.6 Evolution of the MEQ cloud in the reservoir model during the hydraulic stimulation. 148

5.7 Initial and final estimated ensemble mean. 149

5.8 Evolution of estimation RMSE and ensemble spread throughout the integration steps. 150

5.9 Evolution of the Std map throughout the integration steps. 150

5.10 Evolution of ensemble mean throughout the integration steps. 153

5.11 Evolution of estimation RMSE and ensemble spread throughout the integration steps. 154

5.12 The true configuration and geometries (half-length and height) of the multistage hydraulic fractures (8 stages) and the horizontal well in the reservoir. 157

5.13 The distribution of matrix permeability (implemented hydraulic fractures as in Figure 5.12 are not shown here). Both Log-permeability and permeability (md) values are shown in this figure. Final pore pressure distribution and MEQ cloud are also shown. 159

5.14 Evolution of hydraulic fractures geometries (half-length and height) ensembles through integration steps. Blue stars depict the realizations and red circles show the true hydraulic fracture geometry value. . . . 160

5.15 Evolution of estimation RMSE of half-length and height ensembles throughout the integration steps. 161

5.16 The log-log plot of pressure change and pressure derivative versus time (diagnostic plot) and the associated linear flow part (matrix permeability = 600 nd, $C_D = 13889$). 164

5.17 The log-log plot of pressure change and pressure derivative versus time (diagnostic plot) and the associated linear flow part (matrix permeability = 0.02 md, $C_D = 417$). 165

5.18 Comparison of final estimated ensemble mean for different SSBRC setups in Table 5.5. 170

5.19 The overall workflow of the proposed geomechanical forward model by combining Eclipse and MATLAB. 171

6.1 Reservoir well configuration for model A with reference Lperm models and saturation profiles. (a) 13-spot well configuration. (b) Reference Lperm RM1 and corresponding saturation profiles. (c) Reference Lperm model RM2 and corresponding saturation profiles. 184

6.2 Ensemble Kalman filter (EnKF) estimation results for log permeability with correct and incorrect initial variogram models. (a) (top) Reference log permeability model RM1 and (middle) a sample and (bottom) the ensemble mean of the estimated log permeabilities for example 1. (top) Reference log permeability model RM1 and (middle) a sample and (bottom) the ensemble mean of the log permeability estimation results for example 2. (c) Sample water flow rate forecast with the initial and final permeability ensemble for example 1. (d) Sample water flow rate forecast with initial and final permeability ensemble for example 2. The estimation results show a sample log permeability realization and the corresponding log permeability ensemble mean. The number of ensemble replicates is $N = 100$. . . . 187

6.3 Estimation results for experiments A1 and A2. (top) The histogram of the estimated variogram model parameters with the EnKF for reference model RM1 in experiment A1. The true values of the variogram parameters are $\theta = 45°$, $a_{\max} = 500$, and $a_{\min} = 100$. (bottom) The histogram of the estimated variogram model parameters with the EnKF for reference model RM2 in experiment A2. The true values of the variogram parameters are $\theta = 45°$, $a_{\max} = 500$, and $a_{\min} = 100$. The initial variogram model parameters are described with the uniform distributions $\theta \sim \mathcal{U}(0°, 180°)$, $a_{\max} \sim \mathcal{U}(100\text{ m}, 1000\text{ m})$, and $a_{\min} \sim \mathcal{U}(10\text{ m}, 1000\text{ m}) \leq a_{\max}$. The EnKF is used to directly update the variogram parameters at the update steps. The number of ensemble replicates is $N = 100$. 192

6.4 True log permeability model and saturations for the one-dimensional example (model B): (a) rough sketch of model domain and well configuration, (b) true log permeability model with $a = 700$ m, and (c) oil saturation profiles in time. 193

6.5 Direct variogram estimation results for the one-dimensional variogram range (experiment B1): (a) estimated histogram of the variogram range, (b) the root-mean-square error (RMSE) of the estimated range values, (c) the ensemble spread of the estimated range values, (d) cross correlation between the flow response and variogram range, (e) cross correlation between the flow response and log permeability values generated using the correct variogram range, and (f) cross correlation between the flow response and log permeability values generated using random variogram ranges. 195

6.6 Grid block log permeability estimation results for the one-dimensional example (experiment B2) : (a) estimated log permeability ensemble for selected time steps, (b) RMSE of the estimated log permeabilities in time, (c) ensemble spread of the estimated log permeabilities in time, (d) initial and final ensemble water production forecasts for sample production wells, and (e) initial and final ensemble water production forecasts for sample production wells. 198

6.7 Sensitivity of the EnKF performance in estimating log permeability to errors in the variogram parameters (experiments A3A5): (a) sensitivity of the EnKF to direction of major continuity θ (experiment A3), (b) sensitivity of the EnKF to maximum range amax (experiment A4), and (c) sensitivity of the EnKF to maximum range amin (experiment A5). A sample and the ensemble mean of the log permeability for the (top) initial and (bottom) final ensemble are shown in each case. The number of ensemble replicates is $N = 100$. 200

6.8 The log permeability estimation results for experiment A6 with the EnKF and where all three variogram parameters are assumed to be uncertain and follow the uniform distributions $\theta \sim \mathcal{U}(0°, 180°)$, $a_{\max} \sim \mathcal{U}(100\text{ m}, 1000\text{ m})$, and $a_{\min} \sim \mathcal{U}(10\text{ m}, 1000\text{ m}) \leq a_{\max}$. The estimation results show two sample log permeability realizations, the corresponding log permeability ensemble mean, and ensemble variance. The number of ensemble replicates is $N = 100$. 205

6.9 (a) Field setup and well configuration, (b) the true log permeability, and (c) the final saturation profile for the top layer of the SPE10 model (model C). 206

6.10 Estimation results for experiment C. (b) The RMSE (left), ensemble spread (middle), and scatterplot (right) of the log permeability estimates. The initial ensemble is generated from a variogram model with parameters sampled from the following uniform distributions $\theta \sim \mathcal{U}(0^\circ, 180^\circ)$, $a_{\max} \sim \mathcal{U}(100 \text{ m}, 3000 \text{ m})$, and $a_{\min} \sim \mathcal{U}(10 \text{ m}, 3000 \text{ m}) \leq a_{\max}$. The number of ensemble replicates is $N = 100$. 207

6.11 Estimation results for experiment C. (a) Sample and ensemble mean of the log permeability updates in time. 208

7.1 Water injection rate scenario. 222

7.2 Histograms of final ensemble mean of all trial experiments for each setup (red line shows the true permeability value). 224

7.3 Water injection rate scenario. 225

7.4 Comparison of Lperm RMSE of 4 experiments. 227

7.5 Lperm ensemble estimation results (initial and final ensemble in all 4 experiments). 228

7.6 Comparison of ensemble forecast for all 4 experiments (water rate); vertical blue line indicates the end of history matching period. 230

7.7 Comparison of final (after 12 integration steps) estimated ensemble mean for all 3 experiments. 231

7.8 Comparison of permeability estimation RMSE and ensemble spread of three experiments. 231

7.9 Reservoir model configuration and the true porosity maps of PUNQ-S3 model in all layers. 233

7.10 Estimation results in terms of estimated ensemble mean (porosity of layer 1). 234

7.11 Estimated ensemble variance map (porosity of layer 1). 235

7.12 Performance comparison in terms of ensemble forecasts of final cumulative gas and water production by boxplot. 236

7.13 SPE10 model and well configuration and specifications. 237

7.14 Comparison of the estimated ensemble mean and ensemble variance map for the three experiments. 238

LIST OF TABLES

TABLE Page

3.1 Parallel EnKF algorithm for SSBRC. 56

4.1 Sensitivity of the SSBRC performance to different ranges of observation Std (standard SSBRC with seismicity density on the original fine grid) in estimating 2D heterogeneous permeability. 80

4.2 SSBRC with observation projection approach. Sensitivity of the performance of projection approach with respect to kernel bandwidth Std (σ_h) and truncation number (n_{trunc}). 83

4.3 Results of jointly estimating permeability and criticality from MEQ data. 96

4.4 Rock/fluid properties used in simulations. 97

5.1 Reservoir model parameter values. 144

5.2 True geometries (half-length and height) of 8 hydraulic fracture stages. 156

5.3 The matrix permeability statistics of all four experiments (all values are in md). 158

5.4 The estimation results of all four experiments (initial Xf RMSE = 164.74 ft, initial H RMSE = 53.05 ft). 161

5.5 Performance comparison of SSBRC with different ensemble sizes N_e, and number of integration steps. 168

6.1 General simulation and assimilation information[a]. 183

7.1 Efficient EnKF algorithm using pseudo forecasts and clustering for fast history matching. 217

7.2 The final estimation results of all trial experiments. 224

7.3 Computation time comparison of the performed experiments (machine specifications: Intel Xeon CPU 3.07 GHz, 6.0 GB RAM). 233

1. INTRODUCTION

The production of geothermal energy from tight and low-permeability reservoirs is achieved by hydraulic stimulation of the rock through borehole injection in natural and/or man-made fractures, and is referred to as enhanced geothermal systems (EGS). Stimulation of a rock mass to create permeable zones is a process that involves fracture initiation and/or activation of discontinuities such as faults and joints due to pore pressure and in-situ stress perturbations. Hydraulic stimulation of rock is typically accompanied by multiple microseismic events [1, 2, 3, 4, 5, 6], which are believed to be associated with rock failure in shear, and shear slip on new or pre-existing fracture planes [7, 8]. While the true nature and source of such events remains to be fully understood [9, 10], shear and tensile failures in the rocks are considered as the primary mechanisms for triggering microearthquake (MEQ) events [11, 12, 13]. Even though the source of MEQ events is mostly concluded to be shear failure, non-shearing or tensile failures are also reported as the triggering factors [10, 11, 12, 14, 15]. The rock may experience failure through shear and tensile fracture in intact rock material or bedding planes and remobilisation of pre-existing fractures and consequently the strain energy, released due to stress drop associated with failure, generates the P- and S-waves which will be detected to place failure locations as MEQ events [11, 16, 17].

The characteristics of microseismic events such as their locations, spatial patterns of distribution, and temporal relations between the occurrence of seismicity and reservoir activities are often studied for enhanced geothermal systems (EGS). The microseismic signals contain information about the triggering source locations and have been used to understand the hydraulic fracturing process [18, 19, 20].

Detection and interpretation of microseismic events is useful for estimating the reservoir permeability, the stimulated volume and fracture growth, as well as the geometry of the geological structures and the in situ stress state [21, 22, 23]. In geothermal and unconventional resources, the coupled process of fluid flow, rock deformation, heat transfer and chemical interactions along with rock damage models and failure criteria explain cloud of microseismicity events while hydraulic and geomechanical reservoir properties are physical parameters in governing equations therefore the spatio-temporal distribution of MEQ events are expected to reveal important information about reservoir parameter distributions [24, 25, 26].

In this work to simulate reservoir stimulation by hydraulic fracturing and to model induced microseismicity, a fully-coupled thermo-poro-mechanical finite element model with damage mechanics is utilized [27, 28, 29, 30]. The model considers stress-dependent permeability, and convective heat transport in the thermo-poroelastic formulation. Rock damage is reflected in the alteration of elastic modulus and permeability. Coupled rock deformation and fluid flow as in geothermal reservoir stimulation is described in Biot's poroelastic theory [31] while thermal and chemical effects can also be significant in this context [32]. The influence of fluid flow and temperature change around the wellbore on the stress variations in the reservoir can be described using thermo-poroelasticity which in this model is computed based on non-linear rock behavior with rock failure consideration. Since damage evolution causes significant permeability alteration [33], some damage and permeability models are also proposed to relate permeability change, micro-crack and void evolution [34, 35, 36, 37] that *Tang et al.* rock damage model is used in this model [38, 39]. To trace rock failure and the progress of damage in geothermal and petroleum reservoir operation different failure criteria are proposed [40]. In this model to simulate induced microseismicity events, Mohr-Coulomb failure criterion

with tension cut-off is used to model shear and tensile failures in the reservoir.

Some alternative mechanisms for triggering microearthquake (MEQ) events have been proposed in the literature. Among the existing hypotheses, pore pressure relaxation is widely studied in the literature [2, 7, 41, 26, 42, 43, 44]. In this work to model the reservoir stimulation process and MEQ activities, a simplistic forward model only based on pore pressure diffusion is also used. The pore pressure hypothesis postulates that a rise in fluid pressure in a reservoir increases the pressure in the connected pore space of the rock, thereby increasing the pore pressure and decreasing the effective normal compressional stress on the rock surfaces. In critical locations of the rock the fall in the compressional stress can result in sliding along some of the preexisting cracks. The pore pressure relaxation hypothesis is supported by several observations [45, 46]. The spatiotemporal distribution of MEQ events has been observed to have signatures of a diffusion-like process (including a forward triggering front and back front of seismicity waves [4, 47, 48]) consistent with the diffusive nature of pore pressure distribution. Other observations supporting the above hypothesis are related to the ellipsoid-shaped seismicity clouds after normalization of the event coordinates by their occurrence time [47] and the spatial density of MEQ events [46]. If pore pressure diffusion can be used to explain the spatiotemporal signatures of microseismic event clouds, the MEQ events are expected to reveal important information about the distribution of hydraulic diffusivity or permeability in reservoirs. This concept has been exploited in the development of seismicity-based reservoir characterization (SBRC) methods of *Shapiro et al.* [49, 50, 45, 46] and *Rothert and Shapiro* [51] where hydraulic diffusivity or permeability is estimated from analysis and integration of injection-induced MEQ monitoring measurements (passive seismic monitoring). The SBRC analysis involves solving the parabolic equation of pore pressure diffusion in the rock mass and

comparing the distribution of pore pressure with the rock criticality distribution to identify locations that undergo failure. In this context, rock criticality at a given location refers to the minimum pore pressure required to trigger a seismic event. Adopting the pore pressure relaxation hypothesis in this study as an alternate simplistic approach, the initiation of MEQ events can be associated with pore pressure and stress values that exceed the rock criticality. Hence, the distribution of MEQ observations in the reservoir can be correlated with pore pressure distribution, which is in turn related to hydraulic properties of the reservoir rock. Therefore, the MEQ events are viewed as a new source of monitoring measurements that, after interpretation into prior descriptive models, are expected to reveal important information about the distribution of rock flow properties.

Reservoir characterization by MEQ data can be substantially improved by developing more sophisticated forward geomechanics-based reservoir models with damage mechanics and inversion algorithms that offer important estimation properties (features). Having pore pressure relaxation assumption, using criticality as failure criterion and only estimating permeability without uncertainty quantification are the limitations of the SBRC approach which we resolve by proposing stochastic seismicity-based reservoir characterization (SSBRC). Additionally the estimation process can be automated in an iterative or sequential manner and can be implemented to account for the uncertainties in the prior models or the observed microseismic events. Also more importantly, through a stochastic inversion approach, a mechanism can be included for quantitative assessment of the quality of the solution obtained and for rigorous characterization of solution uncertainty. The need for such a stochastic inversion approach is imperative in light of the significant uncertainties that exist in describing the spatial distribution of rock physical properties. As we will discuss soon, an outstanding challenge in applying state-of-the-art inversion method

for MEQ data integration is the discrete nature of these events, which does not fit into most of the conventional estimation methods that are designed for assimilating continuous data.

Since most established model calibration algorithms are designed to integrate continuous measurements, inverting the discrete microseismic events calls for development of inversion methods to handle discrete data types. In some cases, however, it may be possible to equivalently interpret discrete data sets as continuous measurements (through a simple conversion) that can be readily processed using well established inversion techniques. One way to model the MEQ data is to consider the density of these discrete events. This interpretation leads to combining (counting) the discrete microseismic events at each location in the reservoir and interpreting the results as the distribution of "seismicity density". This conceptual framework is followed in this work by taking advantage of kernel density estimation (KDE) methods [52, 53]. The KDE methods are common for smoothing data and estimating nonparametric probability density functions [52, 53]. We will use this approach to convert discrete MEQ events into a map of seismicity density as a continuous representation of the data. The continuous representation of MEQ data can then be used with a data integration technique to estimate the relevant reservoir properties. The data integration of our choice in this work is the ensemble Kalman filter (EnKF) [54, 55]. We evaluate the feasibility of using the EnKF for estimating reservoir hydraulic and geomechanical property distributions such as permeability, Young's modulus, tensile strength and Cohesion from the KDE-based continuous representation of microseismic measurements. The stochastic inversion framework of SSBRC is shown in Figure 1.1.

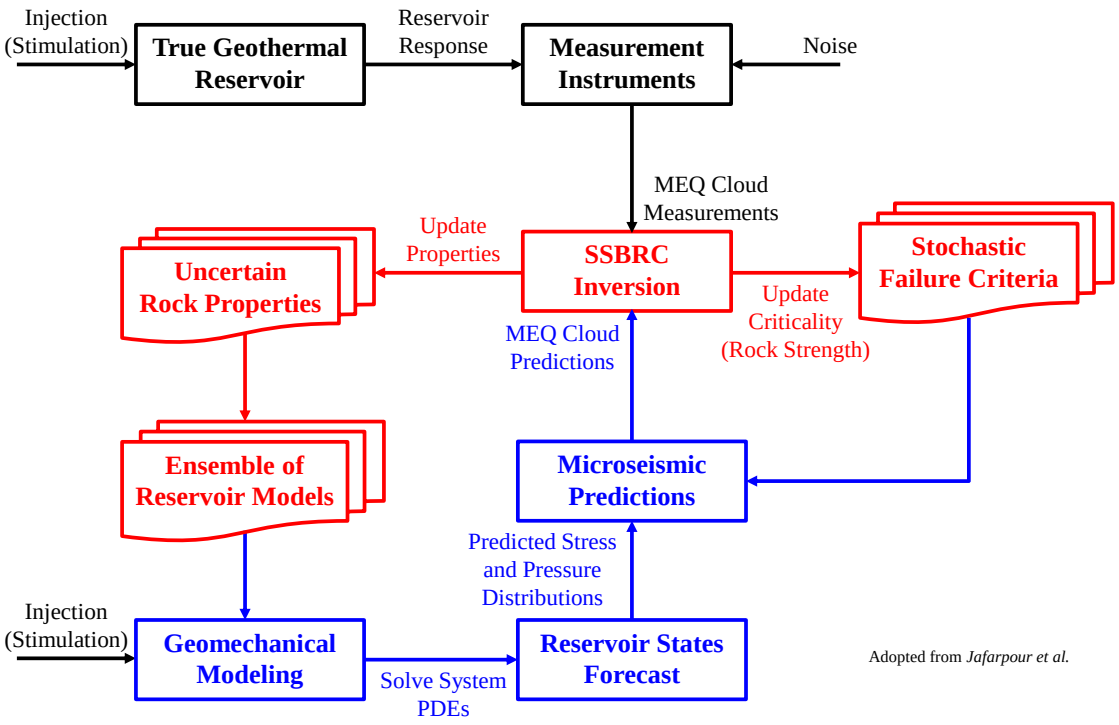


Figure 1.1: Proposed framework of SSBRC for EGS (*Ghassemi and Jafarpour*).

The EnKF has been widely established as a practical data integration method for large-scale nonlinear dynamical systems and has been received favorably by the scientific and research community in a range of applications including hydrology [56], meteorology and oceanography [57, 58, 59], groundwater model calibration [60, 61, 62, 63], and oil reservoir characterization [64, 65, 66, 67, 68]. *Evensen* [69] reviews the EnKF formulation and its wide range of applications. *Ehrendorfer* [70] presents a review of important issues that are encountered in implementing the EnKF. Despite the existing limitations in operational implementation of the EnKF for more complex (non-Gaussian) and challenging large-scale problems, this approach has become popular as a promising approximate nonlinear estimation method in several applications. In this work, we propose SSBRC by applying the

EnKF for MEQ data integration and evaluate its performance using several numerical experiments. A parallel EnKF algorithm will also be implemented to speed up the computations.

Generating seismicity density maps on the same grid system or mesh structure that is used for describing the geomechanical property distributions artificially increases the data resolution and, hence, the number of data that will be assimilated during the EnKF update step. Large scale datasets, such as 4D seismic data [71, 72], and particularly in this study, high resolution seismicity density maps, can exhibit spurious spatial correlations in the observed data and create unrealistic correlations between rock properties and microseismic data, thereby can degrade the performance of the EnKF update and lead to underestimated solution uncertainty or ensemble collapse [56, 73, 74]. We will first show the estimation results for the SSBRC approach using the standard EnKF algorithm to illustrate the underestimation of ensemble spread, and then will propose alternative implementation methods to resolve this issue and improve SSBRC uncertainty quantification. A straightforward approach to overcome ensemble spread underestimation is to increase the observation error artificially by using a large variance for the observation noise. We can also reduce the number of observations either by using a spectral projection (spectral dimension reduction) approach or by coarsening the seismicity density map (spatial dimension reduction). In projection approach, the ensemble of perturbed observations are projected to a reduced subspace that is defined by the leading left singular vectors of the observation matrix. This step is aimed at de-correlating the original observations of the seismicity map. The EnKF update is then used to assimilate the resulting low-dimensional description of the data. In the second approach, we use a coarse grid system for interpretation of the seismic events. This approach is very similar to the original SSBRC implementation except that it uses a coarse grid system in

KDE-based continuous seismicity interpretation to make lower resolution density maps. Then the reduced dimension or coarse seismicity density maps are used in the EnKF update equation. In this work, we will develop the practice of permeability inference from discrete MEQ data with EnKF [24, 25] and then extend it to geomechanical parameters by using a coupled geomechanical forward model and subsequently we will improve the uncertainty quantification.

We also verify and extend the applicability of SSBRC to hydraulically fractured reservoirs. SSBRC method also will be utilized to infer the reservoir properties and hydraulic fractures characteristics by integrating the MEQ data. In unconventional and tight formations where the hydraulic fracturing is the prominent method of production enhancement, microseismic observations can be used to characterize both reservoir and induced fractures. We will present several numerical examples of unconventional reservoir with horizontal well and multiple transverse hydraulic fractures to verify the promising estimation performance of SSBRC for reservoir characterization and microseismic fracture mapping. Some improvements and modifications are also recommended along with SSBRC method to improve its computational efficiency and make it a real-time procedure.

The EnKF initial ensemble is commonly constructed by assuming a known geological continuity model such as a variogram. However, geologic continuity models are derived from incomplete information and imperfect modeling assumptions, which can introduce a significant level of uncertainty into the produced models. Neglecting this important source of uncertainty can lead to systematic errors and questionable estimation results. We investigate the performance of the EnKF under varying levels of uncertainty in the variogram model parameters. We first attempt to directly estimate variogram model parameters from flow data and show that the complex and nonunique relation they have with the flow data provides little sensitivity for

an effective inversion with the EnKF. We then assess the performance of the EnKF for estimation of permeability values under uncertain and incorrect initial variogram parameters and show that any bias in specifying variogram parameters tends to persist throughout the EnKF analysis even though locally reasonable permeability updates may be obtained near observation points. More importantly, we show that when variogram parameters are specified probabilistically to account for the full range of structural variability in the initial permeability ensemble, the EnKF update results are quite promising. The practical implications of the results are significant for designing the EnKF for realistic ensemble model calibration problems where the level of uncertainty in the initial ensemble is usually not known a priori.

The EnKF implementation relies on flow predictions with a relatively large number of model realizations, which in the case of realistic reservoir models can be computationally prohibitive. When a small number of model realizations are used the statistics computed for the EnKF model updating step become inaccurate and can lead to inaccurate results. Therefore, several localization methods have been introduced to account for statistical errors due to limited ensemble sizes. While these practical considerations have been useful, they do not address the core issue that a larger ensemble size is needed to accurately compute the required update statistics. We propose to use a large ensemble of models to improve the calculation of ensemble statistics while using a fast approximate forecast method to reduce the computational cost of the EnKF. The forecast for each realization in this case is derived from linearization around a representative or similar realization for which full simulation is performed. We use adjoint model generated gradient or an ensemble-calculated gradient approximation as tangent linear model for the linearization purpose. In the forecast step of our implementation we perform full forecasts for very few realizations to compute ensemble-based gradients. We then

perform approximate (pseudo) forecasts for the remaining models in the ensemble by linearizing about representative models for which full forecasts are computed. The representative realizations for full forecast in each step are selected through a well-established clustering procedure.

The remainder of this dissertation is outlined as follows. We begin with Section 2, which covers an overview of both forward modeling methodologies; first, the pore-pressure diffusion reservoir model and second, the geomechanical reservoir model, and then induced seismicity modeling followed by continuous seismicity interpretation method. Then the SSBRC inverse modeling approach based on EnKF is explained and the proposed methods for improving uncertainty quantification are presented in Section 3. Next in Section 4, we present and discuss the results of applying the proposed approach to a series of geomechanical reservoir examples. In Section 5, application of SSBRC to hydraulically fractured reservoirs is presented. The performance of EnKF under variogram uncertainty is presented in Section 6 and the improved EnKF with pseudo forecast and clustering is reported in Section 7. We close the work with Section 8, which includes general remarks about the presented formulations and procedures, their advantages, limitations and possible future extensions.

2. GEOMECHANICAL FORWARD MODELING METHODOLOGY*

The proposed SSBRC method aims at inferring spatially distributed reservoir properties by integrating MEQ monitoring data with EnKF. As mentioned earlier, we integrate microseismicity density maps as observed data. We use two forward modeling approaches for simulating the reservoir stimulation process. A complete forward model is required to relate the permeability distribution of the reservoir to seismicity density maps which includes pore-pressure diffusion reservoir model, criticality as failure criterion and KDE-based continuous seismicity interpretation. The main steps involved in the implementation of the pore-pressure diffusion overall forward model are schematically shown in Figure 2.1.

Similarly in the case of geomechanical reservoir model, a comprehensive forward model is needed to relate the reservoir hydraulic and geomechanical property distributions to seismicity density maps which includes FEM-based coupled geomechanical simulator, failure criteria, damage and permeability model and finally KDE-based continuous seismicity interpretation. The main steps involved in the implementation of the geomechanical overall forward model are schematically shown in Figure 2.2. The reservoir property distributions such as permeability, Youngs modulus, tensile strength and cohesion are used as input parameters to the coupled FEM simulator which involves the coupled process of rock deformation, fluid flow and heat transfer. The FEM simulator predicts the stress distributions in the reservoir and where the rocks effective stress exceeds its strength (prescribed by failure criteria) is used to establish the triggering mechanism and to predict microseismic events

*Part of this chapter is reprinted with permission from "Inference of permeability distribution from injection-induced discrete microseismic events with kernel density estimation and ensemble Kalman filter" by Mohammadali Tarrahi and Behnam Jafarpour, 2012. Water Resources Research, 48, Copyright 2012 by John Wiley and Sons.

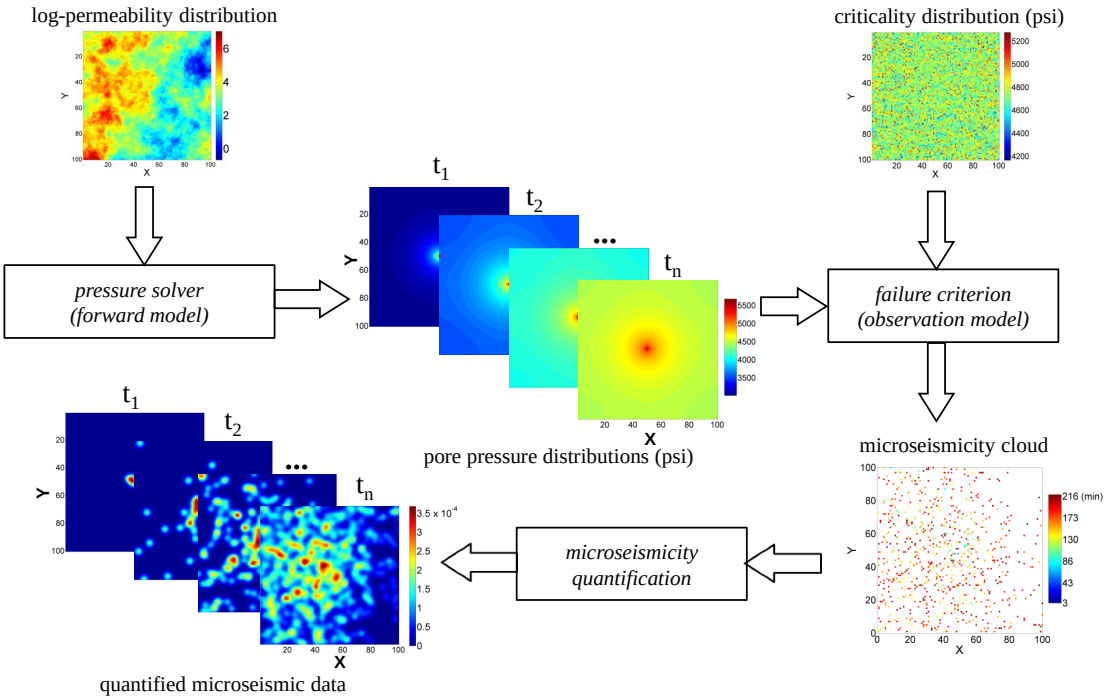


Figure 2.1: Overall workflow of the pore-pressure diffusion forward model that relates permeability distribution to microseismicity density observations.

[75, 30]. These predictions are then converted into continuous seismicity densities, using the KDE method. Rock damage and permeability models are also used to update the elastic modulus and permeability values to model the rock degradation during hydraulic stimulation.

The details of each of the steps involved in forward model are discussed next.

2.1 Pore Pressure Diffusion Forward Model

The Frenkel-Biot equations [76] in a homogeneous isotropic saturated poroelastic medium, identifies three waves (P, S and a dissipative slow wave named Frankel-Biot) propagating from a source to an observation point. Field evidence suggests that evolution of MEQ events is a relatively slow process that is likely to be associated,

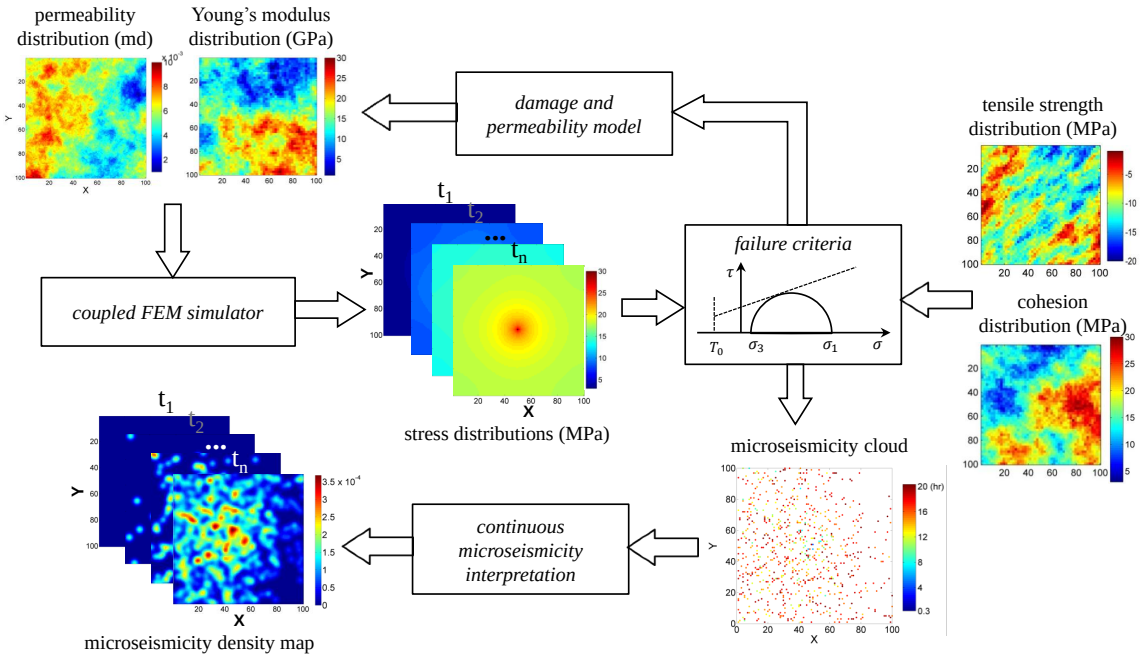


Figure 2.2: Overall workflow of the geomechanical forward model that relates hydraulic and geomechanical reservoir property distributions to microseismicity density observations.

at least in part, with the Frankel-Biot wave. The pore pressure variation in the slow wave can be described by a simple diffusion equation for a homogeneous isotropic porous medium. Recent studies (see [26] and references therein) have proposed a similar diffusion equation for describing the spatiotemporal evolution of pore pressure relaxation in heterogeneous anisotropic porous media.

We follow the same approach in this work. We also assume the rock failure occurs at locations in the reservoir where the pore pressure exceeds rock criticality. To implement this failure criterion with the pore pressure relaxation assumption, we numerically solve the diffusive pore pressure equation for a heterogeneous reservoir. This forward model relates rock hydraulic conductivity distribution to pore pressure distribution, which is directly related to the rock failure mechanism

and the distribution of microseismic events (data). For single phase flow with slightly compressible fluid, the pore pressure equation is expressed as

$$\nabla \cdot \Big(m\left(\mathbf{u}\right) \nabla p\left(\mathbf{u}, t\right) \Big) = \phi \mu c \frac{\partial p\left(\mathbf{u}, t\right)}{\partial t} \tag{2.1}$$

where $\mathbf{u}$ denotes the location in space, t is time, ϕ is porosity, c is total compressibility, and μ represents viscosity. In Equation (2.1), the spatial distribution of permeability in space is denoted by $m(\mathbf{u})$ while the spatiotemporal distribution of pore pressure is represented with $p(\mathbf{u}, t)$. We solve the diffusion equation in (2.1) for each permeability realization $m(\mathbf{u})$ using a finite difference-based commercial fluid flow simulator [77]. Under the assumptions described above and heterogeneity and anisotropy of hydraulic permeability, Equation (2.1) can be derived from the Frenkel-Biot equations in the low-frequency range [78, 79].

To generate the corresponding microseismicity clouds for each permeability model realization, we apply the failure criterion used by *Shapiro et al.* [45] using the predicted pore pressure distributions to generate the spatiotemporal distribution of seismicity events. An important property of the EnKF inversion is that the forward and observation models can be quite general with varying level of complexity such as the fully coupled geomechanical simulators and damage models that will be introduced next. However, the update equation is designed for continuous random variables (parameters and observations). As a result, the updates can be applied under various forms of event triggering mechanisms and failure criteria as long as the random variables representing the states, parameters and measurements are continuous. With this forward model we consider the pore pressure as the MEQ triggering mechanism and investigate the feasibility of estimating permeability distribution from microseismic observations using the EnKF.

In the pore pressure diffusion method the failure criterion simply is comparing the pore pressure and criticality. This failure criterion is used to generate the MEQ cloud. Denoting rock critical stress as $C(\mathbf{u})$ and the pore pressure distribution at location $\mathbf{u}_i$ and time t_j as $p(\mathbf{u}_i, t_j)$, we assume that a microseismic event (failure) is triggered at location $\mathbf{u}_e$ and time t_e if

$$p(\mathbf{u}_i, t_j) > C(\mathbf{u}_i) \Rightarrow (\mathbf{u}_e, t_e) = (\mathbf{u}_i, t_j) \tag{2.2}$$

From Equation (2.2), at each time step t the comparison between predicted pore pressure and criticality at different locations in the reservoir identifies the distribution of seismicity clouds. Figure 2.1 illustrates this procedure schematically for a two-dimensional heterogeneous permeability model and a given random criticality distribution. Note that following [26], we have represented the rock criticality at each grid block as an uncorrelated random variable. This assumption can be easily relaxed and a spatially correlated random field (possibly correlated with other rock physical properties) can be considered as a criticality distribution.

2.2 Geomechanical Forward Model

The geomechanics-based reservoir modeling is constructed by combining governing equations of thermo-poroelasticity and rock damage model, and then numerical implementation as a finite element program. The theory of thermo-poroelasticity (or porothermoelasticity) is developed by combining the influence of thermal stress and differential solid/fluid expansion to rock stresses and fluid diffusion. Whereas the coupled pressure and stress problems in the porous media (coupled process of rock deformation and fluid flow) is described with *Biots* consolidation theory [31, 80], a poro-thermoelastic approach combines the theory of heat conduction with poroelastic constitutive equations.

Thermo-poroelasticity can be used to assess the influence of fluid flow and temperature change on the stress variations in the reservoir. This influence is often computed assuming a linear elasticity with constant mechanical and transport rock properties which has limitations in predicting the real behavior of the reservoir rock. Generally, the strain-stress behavior of rocks in triaxial tests shows hardening and post-peak softening that depends on the rock type, pore pressure, stress conditions, and temperature [40]. The continuum damage mechanics approach can capture the hardening and softening behavior of the rock [81] which was first introduced by *Kachanov* and since has been developed and applied by many researchers [82, 83, 84, 38, 85, 33] who have investigated inelastic behavior caused by crack initiation, microvoid growth, and fracture propagation. Also, the evolution of rock damage in the presence of poroelastic and thermo-poroelastic effects has been considered. *Selvadurai* studied damage in poroelastic brittle rock [33]. His results showed a significant permeability alteration caused by damage evolution in consolidation problems. *Hamiel et al.* developed a model with a time dependent damage variable, porosity, and material properties [86]. They proposed different rock behavior with degradation and healing within the framework of the poroelastic theory. *Tang et al.* proposed a damage and permeability model based on experimental strain-stress observations and permeability measurements [38, 85]. This model which corresponds to the brittle rock failure behavior is implemented in a finite element model and is used to alter elastic modulus and permeability in the elastic and damage phase to simulate rock degradation and also hydraulic fracture propagation. Injection induced microseismic events are also simulated by shear and tensile failures made by Mohr-Coulomb criterion with tension cut-off.

Both two and three dimensional coupled finite element models are implemented and verified by different numerical simulations.

2.2.1 Governing Equations

The governing equations include the constitutive and transport laws. The constitutive equations of thermo-poroelasticity have been developed by *McTigue* [87] and *Palciauskas and Domenico* [88]. Using the geomechanics sign convention of compression positive, the constitutive equations for total stress, pore pressure, and temperature are:

$$\dot{\sigma}_{ij} = 2G\dot{\varepsilon}_{ij} + \left(K - \frac{2G}{3}\right)\dot{\varepsilon}_{kk}\delta_{ij} + \alpha\dot{p}\delta_{ij} + \gamma_1\dot{T}\delta_{ij} \tag{2.3}$$

$$\dot{\zeta} = \alpha\dot{\varepsilon}_{ii} + \beta\dot{p} - \gamma_2\dot{T} \tag{2.4}$$

where σ_{ij} and ε_{ij} are the total stress and strain tensors, p and T are the pore pressure and temperature respectively. α is the Biot coefficient, ζ is the variation of fluid contents (pore volume), δ_{ij} is the Kronecker delta, K is bulk modulus, and G is the shear modulus; γ_1, γ_2 and β are given by:

$$\beta = \frac{\alpha - \phi}{K_s} + \frac{\phi}{K_f} \tag{2.5}$$

$$\gamma_1 = K\alpha_m \tag{2.6}$$

$$\gamma_2 = \alpha\alpha_m + (\alpha_f - \alpha_m)\,\phi \tag{2.7}$$

where ϕ is the porosity, α_m and α_f are the thermal expansion coefficients of solid and fluid, respectively. The bulk moduli of the solid material and the fluid are indicated by K_s and K_f respectively.

Fluid flow in porous rock is governed by Darcys law, and heat conduction obeys

Fouriers law, so that:

$$J^f = -\rho_f \frac{k}{\eta} \nabla p \tag{2.8}$$

$$J^T = -k^T \nabla T \tag{2.9}$$

where ρ_f is fluid mass density, k and η are the permeability and viscosity, respectively, and k^T is the thermal conductivity. J^f and J^T represent fluid and heat flux respectively.

The balance of force and continuity for the fluid mass are given by:

$$\sigma_{ij,j} = 0 \tag{2.10}$$

$$\frac{\partial \zeta}{\partial t} = -\frac{1}{\rho_f} \nabla J^f \tag{2.11}$$

By substituting the constitutive equations into the balance equations given by Equations (2.10),(2.11), we obtain the field equations for the rock deformation and fluid flow, namely Equations (2.12),(2.13). The conservation of energy with Fouriers law yields the field equation for the temperature distribution:

$$\left(K + \frac{G}{3}\right) \nabla \left(\nabla \cdot u\right) + G\nabla^2 u + m\left(\alpha \nabla p + \gamma_1 \nabla T\right) = 0 \tag{2.12}$$

$$\alpha \left(\nabla \cdot \dot{u}\right) + \beta \dot{p} - \frac{k}{\eta} \nabla^2 p - \gamma_2 \dot{T} = 0 \tag{2.13}$$

$$\dot{T} + \nu \left(\nabla T\right) - c^T \nabla^2 T = 0 \tag{2.14}$$

where u is the displacement, $m = [1, 1, 0]^T$ for two dimensional problems and $m = [1, 1, 1, 0, 0, 0]^T$ for three dimensional cases. In Equation (2.14), c^T represents the thermal diffusivity and we consider convective heat transfer because of cooling effects that are from increased flow velocity in damage phase. This fluid velocity is coupled

with pore pressure variations in Darcys law:

$$\nu = -\frac{k}{\eta}\nabla p \tag{2.15}$$

2.2.2 Rock Damage Model and Microseismic Events

We use a model of rock damage and permeability as proposed by *Tang et al.* based on experiments [38, 85]. The model is modified herein to consider continuous stress relaxation from the peak stress to the residual strength (softening regime). According to this model, the strain-stress behavior of rock can be divided into an elastic phase and a damage phase. In the elastic phase there is no damage in the rock, whereas the rock begins to fail by crack initiation, crack-growth and void-growth when the stress conditions reach the failure level i.e., they satisfy the failure criterion. To account for the rock failure due to the change of the stress conditions in the rock, we adopt a robust elastic damage constitutive model. According to the stress level, the damage model could be classified into four phases: (1) elastic phase; (2) damage phase due to compress shear or tensile; (3) cracked phase; and (4) crack close phase. An elastic-damage mechanics represents the rock degradation by expressing the damage in terms of a reduction on the elastic modulus as the damage proceeds:

$$E = (1 - d)\, E_0 \tag{2.16}$$

where d represents the internal damage variable describing the amount of degradation (crack initiation, micro-void growth and crack propagation) and E_0 is the initial rock modulus. If damage occurs by compressive stress, the damage variable is defined in

terms of strain, peak stress, and residual compressive strength of the rock as:

$$d = 1 - \frac{\left[\left(\frac{f_{cr}-f_c}{\varepsilon_{cr}-\varepsilon_c}\right)(\varepsilon-\varepsilon_c)+f_c\right]}{E_0\varepsilon} \qquad (\varepsilon_c < \varepsilon < \varepsilon_{cr}) \tag{2.17}$$

$$d = 1 - \frac{f_{cr}}{E_0\bar{\varepsilon}} \qquad (\varepsilon > \varepsilon_{cr}) \tag{2.18}$$

where f_{cr} and ε_{cr} are the residual compressive strength and strain, and f_c and ε_c are the peak stress and strain in compressive field. $\bar{\varepsilon}$ is the equivalent strain.

If damage evolves in a tensile stress field, the damage variable is defined using the residual tensile strength of the rock as:

$$d = 1 - \frac{f_{tr}}{E_0\bar{\varepsilon}} \qquad (\varepsilon > \varepsilon_{tr}) \tag{2.19}$$

where f_{tr} and ε_{tr} are the residual tensile strength and strain, respectively.

To trace the progress of damage under tensile stress, we introduced a tension cut-off, T_0, for tensile failure because the Mohr-Coulomb failure criterion was developed based on shear failure and it often overestimates the stress state for rock failure. The Mohr-Coulomb failure criterion for shear failure can be described as,

$$F = \frac{\sigma_1+\sigma_3}{2}\sin\phi_f - \frac{\sigma_1-\sigma_3}{2} - c_f\cos\phi_f \tag{2.20}$$

where σ_1 and σ_3 are the maximum and minimum principal stresses, respectively; ϕ_f and c_f represent the friction angle and cohesion, respectively. The rock failure criteria are expressed as $F < 0$ in Equation (2.20) for shear failure and $\sigma_3 > -f_t$ for tensile failure (f_t or T_0 represents rock tensile strength).

The rock permeability model used also considers altered permeability in the

elastic and damage phase [38, 85]:

$$k = k_0 e^{-\beta_d\left(\frac{\sigma_{ii}}{3}-\alpha p\right)} \qquad (d=0) \tag{2.21}$$

$$k = \zeta_d k_0 e^{-\beta_d\left(\frac{\sigma_{ii}}{3}-\alpha p\right)} \qquad (0<d\leq 1) \tag{2.22}$$

where k_0 is the initial permeability, ζ_d and β_d are material constants determined empirically. Here ζ_d ($\zeta_d > 1$) indicates permeability increase by damage. The parameter β_d $[Pa^{-1}]$ in the exponent term is the control parameter for stress sensitivity of permeability in porous rock.

In the implementation of damage model in finite element method, as the elements may change from one phase to another, an iterative method is used to update the elastic modulus and permeability by taking into account such phase changes until no change occurs at each time step. Therefore, to find the permeability at each Gaussian point (nodes in the finite element configuration), iterative method is used to update the permeability in which the new permeability values on each element are compared with the previous step until a convergence criterion is satisfied. Then, we update the permeability and march to the next times step.

In this study we skip the process of modeling of MEQ data acquisition and MEQ event locating (analyzing the P and S waves received by geophones to determine the MEQ location) [17, 89], and we assume that the final discrete MEQ data cloud is created already and we use directly the spatio-temporal MEQ cloud (location and occurrence time of events) in SSBRC method to estimate reservoir hydraulic and geomechanical parameters. As the monitoring data in SSBRC method, we recognize both shear and tensile failures as microseismicity events [11, 12, 13], however it is reported that tensile failures are not easily recorded and MEQ events are mostly shear failures [10, 11, 12, 14, 15].

2.2.3 Finite Element Method for Thermo-Poroelasticity

We develop both 2D and 3D finite element program with full poro-thermoelasticity capability [90]. The 2D (3D) model uses eight-node quadrilateral (hexahedron) elements for the displacements u, and four-node (eight-node) elements for the pore pressure p, and temperature T to improve numerical resolution of displacements. The following variables are approximated using Galerkins method for u, p, and T:

$$u = N_u \tilde{u} \tag{2.23}$$

$$p = N_p \tilde{p} \tag{2.24}$$

$$T = N_T \tilde{T} \tag{2.25}$$

where the shape functions for the displacement, pore pressure and temperature are N_u, N_p, and N_T, respectively, and nodal variables for displacements, pore pressure and temperature are $\tilde{u}$, $\tilde{p}$ and $\tilde{T}$, respectively. Weak formulations are obtained by substituting Equations (2.23), (2.24), (2.25) to the field Equations (2.12), (2.13), (2.14). For discretizing the time domain, the Crank-Nicolson type approximation is applied. The model has been verified using analytical solutions for stress, pore pressure, temperature, and concentration of solute distributions around a wellbore [27]. In this work, we consider injection rate boundary condition. This boundary condition has a type of pressure gradient (Neumann boundary condition) so that tractions and fluid flux do not have unique solution. Therefore, we use an iterative method for the traction at the boundary until its value converges. Specifically, we first obtain the pore pressure value at the wellbore under the assumption of no traction for the first iteration, and iteratively computed the tractions using obtained pore pressure values until it satisfies the convergence limit. Full description of coupled

chemo-thermo-poroelastic finite element formulation and verifications are published by *Zhou and Ghassemi (2009)* for the 2D case [90]. In convective heat transfer computation, we applied Streamline-Upwind/Petrov-Galerkin (SUPG) method to avoid numerical oscillation [91].

Theory of damage mechanics has been implemented in the finite element code described above. According to the experimental results [38, 86, 92], stresses show a rapid decrease which is related to the softening regime after rock failure. We first performed a number of simulations (of triaxial compression tests) to find the material parameters for the residual strength which determines the level of softening in the damage phase. By comparing the strain-stress behavior and permeability change with experimental data in [38, 86, 92], we selected the optimum parameters for permeability ζ_d, β_d, and the cohesive strength in Mohr-Coulomb failure criterion.

The overall workflow of the fully coupled geomechanical forward reservoir simulator is shown in Figure 2.3. This diagram illustrates the procedure of simulating the coupled process of rock deformation, fluid flow and heat transfer and combining the governing equations of thermo-poroelasticity and rock damage model (failure criteria and, damage and permeability model) and then numerical implementation as a finite element program.

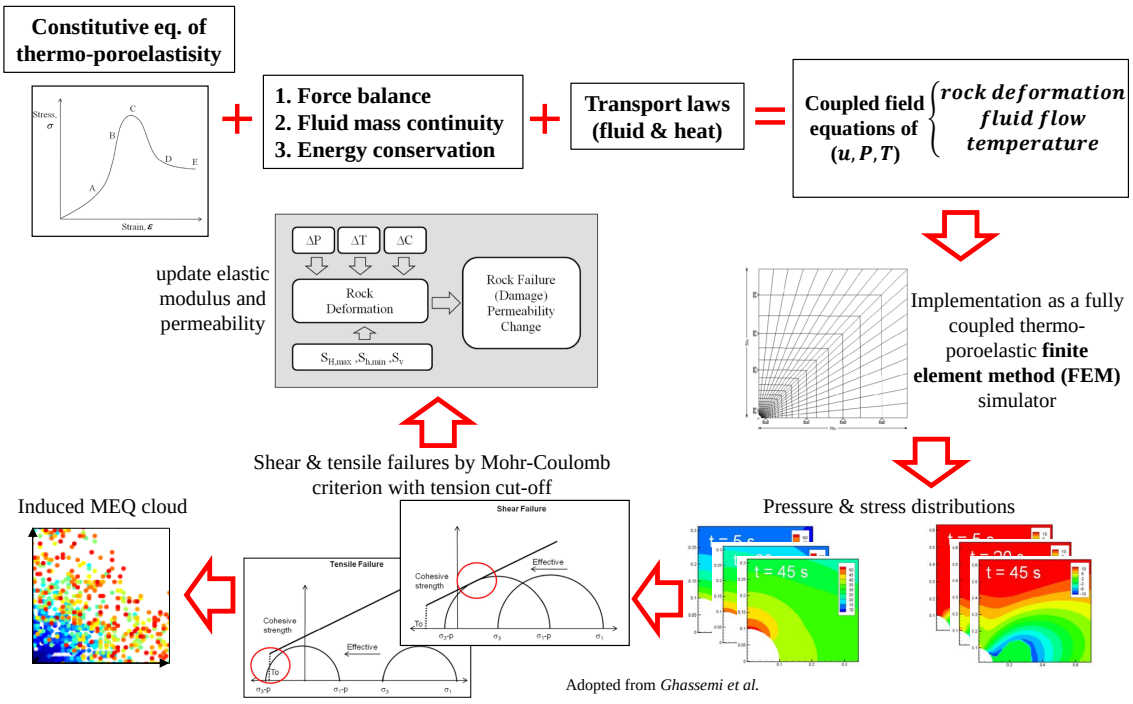


Figure 2.3: Overall workflow diagram of the fully coupled geomechanical forward model.

In this work, we use a 2D FEM model with point source injection by quadrilateral regular mesh with $50 \times 50 = 2500$ elements (10000 nodes) and the reservoir size of $500m \times 500m$. The developed 2D and 3D coupled FEM model is capable of handling spatially distributed parameters (heterogeneous distributions) such as permeability, Youngs modulus, tensile strength, cohesion and friction angle. We assign the spatially distributed parameters to the nodes therefore 2D model parameter dimension is 10000. We also use a 3D fully coupled FEM model with point source injection by hexahedron regular mesh with $30 \times 30 \times 15 = 13500$ elements (108000 nodes) and the reservoir size of $750m \times 750m \times 370m$. So for the 3D model the parameter dimension is 108000.

The verification and illustration of the developed geomechanical forward model

under different conditions (e.g. different far-field stress regimes) and investigating the role of various mechanisms in rock fracture and distributed damage evolution during stimulation are presented with numerical examples by *Lee and Ghassemi* [27, 28, 29, 30].

2.3 Continuous Microseismicity Interpretation with Kernel Density Estimation

An important property of the EnKF inversion is that the forward and observation models can be quite general with varying level of complexity. However, the update equation is designed for continuous random variables (parameters and observations). As a result, the updates can be applied under various forms of event triggering mechanisms and failure criteria as long as the random variables representing the states, parameters and measurements are continuous.

As described before, for geomechanical forward model the MEQ events are determined by shear or tensile failures generated using the specified failure criterion (Mohr-Coulomb model with tension cut-off) at the nodes of the FEM mesh, consequently each MEQ event has its associated location $\mathbf{u}_e$ and occurrence time t_e ($\mathbf{u}$ and t denote general location and time respectively). Therefore at each time step t evaluating the failure criterion at different locations (nodes of the FEM mesh) in the reservoir identifies the distribution of seismicity clouds. Similarly for the pore pressure diffusion forward model, the location and occurrence time of the events are identified by comparing the pore pressure and criticality at each grid block of the reservoir model.

In practice, the discrete microseismic events identify the location of the passive seismic sources and are often generated through seismic source inversion methods. The raw seismic data (collected either from surface or borehole geophones) are

inverted to map the location of seismic sources and characterize the associated uncertainty. In this work, however, we skip the seismic source inversion part and assume that, after seismic data analysis, the map of observed source (event) locations is available.

The available seismic observations, however, are of discrete nature since they only identify the seismic status (active or inactive) of a node in the FEM reservoir model. The discrete nature of MEQ events introduces a difficulty in implementing inversion methods that are designed for continuous problems. For gradient-based methods, the discrete form of MEQ observations complicates the calculation of their gradients with respect to unknown parameters. On the other hand, while the EnKF does not require gradient information explicitly, by construction it is formulated for estimation of continuous variables and observations. To address this issue, we interpret the MEQ events as continuous measurements using the kernel density estimation method. KDE is often used for nonparametric approximation of continuous probability density functions (PDFs). The general idea is to convert the discrete MEQ data (and their predictions) into a smooth and continuous seismicity density map. For this purpose, at each time step, we replace each MEQ event/source with a Gaussian kernel function centered at the event location. By adding up the kernels, we construct a continuous function over the model mesh that represents the spatial density of the MEQ events. The procedure for implementing the KDE method is illustrated in Figure 2.4 for a one-dimensional example. We note that the procedure in Figure 2.4 can be easily applied to two and three dimensional problems. Mathematically, the continuous seismicity density map can be written as

$$s\left(\mathbf{u}\right) = \frac{1}{n_{MEQ}} \sum_{i=1}^{n_{MEQ}} K_i^{MEQ}\left(\mathbf{u}\right) \tag{2.26}$$

where $K_i^{MEQ}(\mathbf{u}) = N(\mathbf{u}_{MEQ}, \Sigma)$ is a Gaussian kernel, n_{MEQ} is the number of MEQ events at each time step, $\mathbf{u}_{MEQ}$ denotes the location coordinate of the MEQ events (center of the individual Gaussian kernels) and Σ is the covariance matrix of the Gaussian kernel. The continuous map $s(\mathbf{u})$ represents the seismicity density at all locations (all nodes of the FEM mesh) in the reservoir and constitutes the observations for the EnKF update.

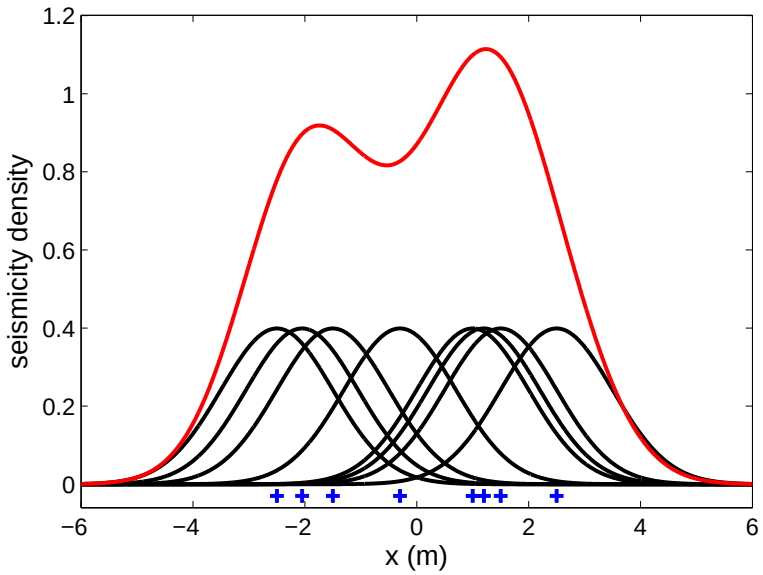


Figure 2.4: Schematic illustration of converting discrete microseismic event measurements to continuous seismicity density observations in one dimension using kernel density estimation method. The crosses on the x axis show the reconstruction of the discrete microseismic events, while the short symmetric curves display the corresponding Gaussian kernels used to represent each event as a density function with maximum value at the location of the discrete events. The red line shows the density of the microseismic events in space as a continuous observation to be used in the EnKF.

The Gaussian kernel has the form

$$K^{MEQ}(\mathbf{u}) = N(\mathbf{u}_{MEQ}, \Sigma) = \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}}\left(-\frac{1}{2}(\mathbf{u}-\mathbf{u}_{MEQ})^T \Sigma^{-1}(\mathbf{u}-\mathbf{u}_{MEQ})\right) \quad (2.27)$$

in which the covariance matrix can be specified either globally for all events or locally (or separately) for individual events. The covariance matrix for the kernel determines the shape, size and orientation of the Gaussian ellipsoid centered at the microseismic event location. In this work, we select an isotropic Gaussian kernel for quantification of the microseismic events and the uncertainty in the MEQ locations. Additionally the assigned kernels of all events are assumed exactly the same. Therefore the covariance matrix Σ, is diagonal and all diagonal members (two members for 2D and three members for 3D) are equal. Therefore, to specify the Gaussian kernel or the quantification specifications we only need to determine one single parameter which is the (isotropic) bandwidth h, of the smoothing kernel. The kernel covariance matrix is as follows,

$$\begin{aligned} &\text{In 2D case: } \Sigma = \begin{bmatrix} h^2 & 0 \\ 0 & h^2 \end{bmatrix} = h^2\mathbf{I}_{2\times 2} \\ &\text{In 3D case: } \Sigma = \begin{bmatrix} h^2 & 0 & 0 \\ 0 & h^2 & 0 \\ 0 & 0 & h^2 \end{bmatrix} = h^2\mathbf{I}_{3\times 3} \end{aligned} \quad (2.28)$$

where $\mathbf{I}_{2\times 2}$ and $\mathbf{I}_{3\times 3}$ are the identity matrices of the specified dimensions.

We note that, in practice, the values of the bandwidth used for the kernel functions are identified from the uncertainty in locating the seismic sources from the raw surface or borehole seismic data. In this work, a simple sensitivity analysis

revealed that selecting a bandwidth parameter as large as twice the dimension of each element in the FEM mesh leads to reasonable results. One advantage of the KDE is that it also provides a convenient procedure to account for the spatial uncertainty in the location of the events.

For better illustration, the procedure for implementing the KDE method for a two-dimensional example is shown in Figure 2.5.

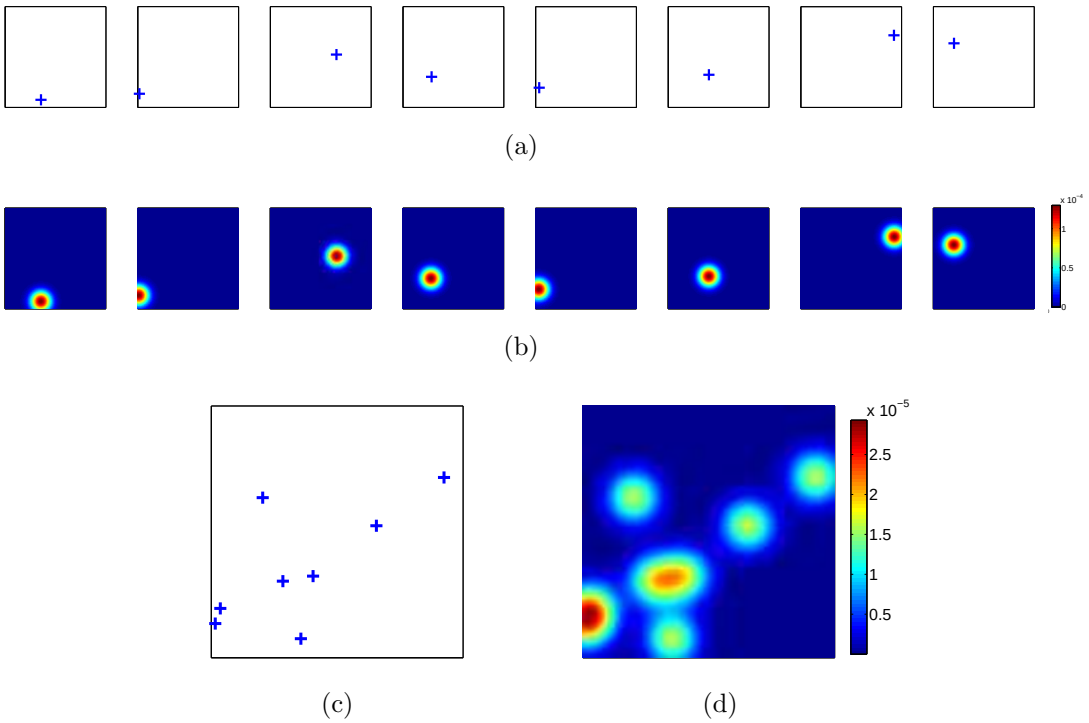


Figure 2.5: Illustrating the continuous seismicity interpretation in 2D. (a) the individual events in separate plots that make the seismicity cloud of (c). (b) corresponding density maps of single events of (a) that are made by putting 2D Gaussian kernel at the location of each events. (d) the seismicity density map corresponds to the cloud in (c) which is obtained by adding (averaging) all individual densities of (b).

Taken from a 2D geomechanical example (experiment 2 in Section 4.4.2), Figure

2.6 shows the evolution of the MEQ cloud at 6 different time steps (called integration time steps) and Figure 2.7 illustrates the corresponding seismicity density maps created by KDE-based continuous interpretation through these time steps.

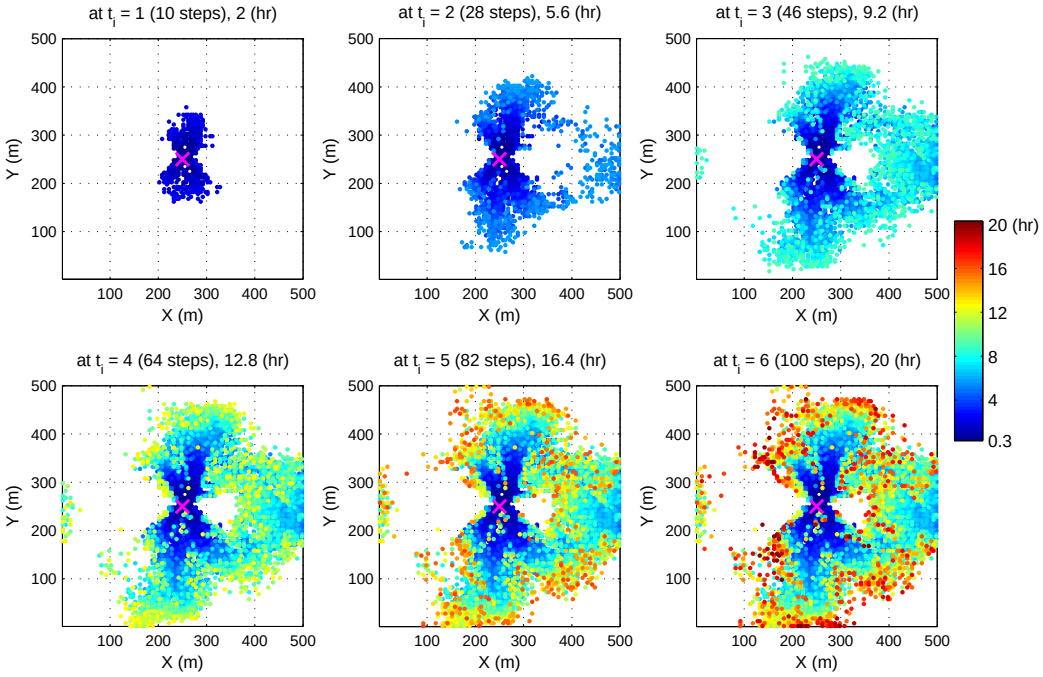


Figure 2.6: Seismicity cloud evolution in 6 different time steps.

Quantification of microseismicity cloud on the original FEM mesh (or the original grid block configuration for pore pressure diffusion model) that the reservoir parameters are assigned to, results in a high resolution seismicity density and redundancy in observations. As we discuss in Section 3.2, high dimensional observation or measurement in the EnKF update leads to uncertainty underestimation and bias in the estimation results. Therefore in Section 3.2.3, we propose to quantify microseismicity on a coarse mesh (with larger elements or grid

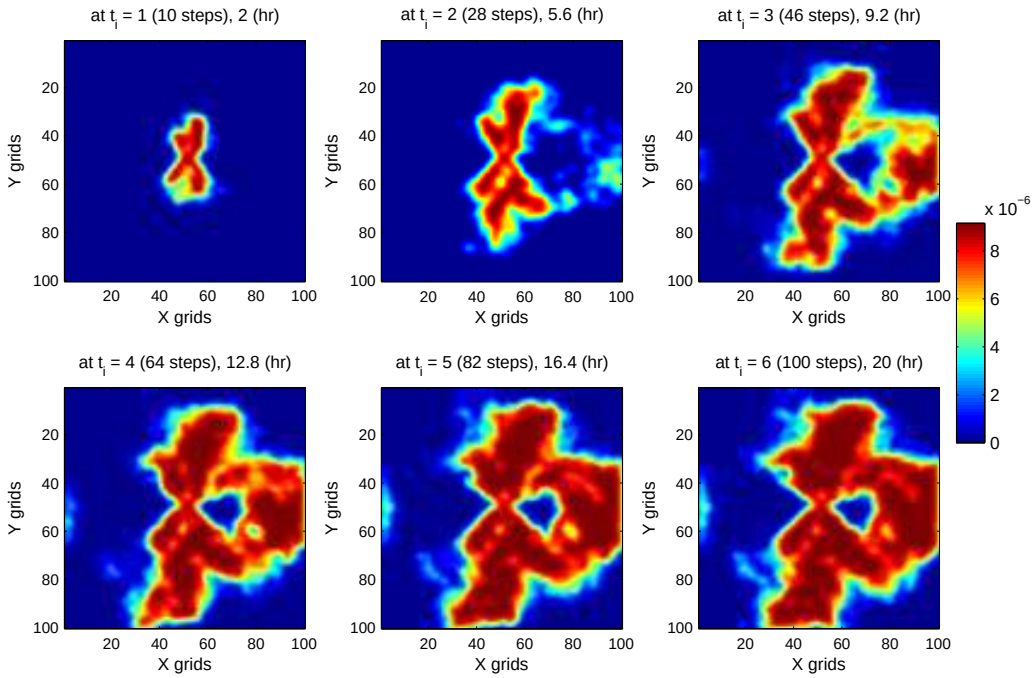


Figure 2.7: Seismicity density maps created by KDE method corresponding to seismicity clouds of Figure 2.6.

blocks) which results in a coarse-scale seismicity density map that reduces the number of measurements.

2.3.1 Consideration of Uncertainty in Event Occurrence Time

As resulted from seismic source location process both event location and occurrence time are uncertain. In general the MEQ event attributes are location in space and its associated uncertainty, occurrence time and its associated uncertainty, magnitude and confidence. We also include the effect of uncertainty in occurrence time of events, as we already consider the location uncertainty by KDE method we can also extend KDE to consider occurrence time uncertainty. Therefore we generalize the idea of MEQ density map both in space and time which represents the density (probability of occurrence) of each event in space and time, i.e. at the center

of each Gaussian kernel probability of happening of an event is maximum and the occurrence probability reduces as we become farther from the center of the kernel.

To incorporate occurrence time uncertainty along with location uncertainty, we propose firstly to place a 1D Gaussian kernel with the bandwidth of h_t on each event with occurrence time of t_{MEQ} so the time density (partial occurrence) of each event in each time is determined then we use these newly created MEQ clouds to make the MEQ density maps by considering the location uncertainty data which discussed previously. The schematic of this procedure is shown in Figure 2.8.

Since we consider the event occurrence time uncertainty, the associated Gaussian kernel to make the event density in time is a one dimensional function of time. Microseismicity density of each individual event in time is defined as

$$\begin{aligned} s_t(t) &= N(t_{MEQ}, h_t) \\ N(t_{MEQ}, h_t) &= \frac{1}{h_t\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-t_{MEQ}}{h_t}\right)^2} \end{aligned} \tag{2.29}$$

where the Gaussian kernel bandwidth h_t, represents the event occurrence time uncertainty bound and the event occurrence time is indicated by t_{MEQ}.

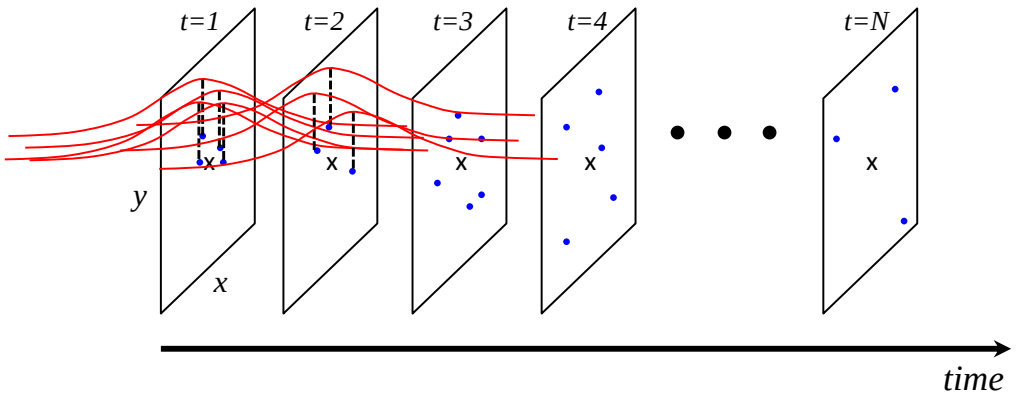


Figure 2.8: The schematic of consideration of event occurrence time in 2D model using KDE method. The well location is indicated by x mark in the middle of the field. The MEQ clouds are incremental clouds. Blue dots illustrate the events that happen in each time step. Red Gaussian curves ($s_t(t)$ in Equation (2.29)) show the density of each event in time. For simplified and clear illustration the Gaussian kernel is placed only for first two time steps.

Since we accumulate the events at each integration step the occurrence time uncertainty of event should not have significant effect and this approach of data integration is not very sensitive to occurrence time uncertainty. If we follow the approach of integrating incremental MEQ events (events that happened between two integration steps) or we determine more integration times then the effect of occurrence time uncertainty is considerable in the estimation procedure.

The implemented approach (as the standard event time uncertainty consideration approach) works for relatively large time uncertainty intervals while the actual event occurrence time uncertainty interval seems to be significantly lower. Therefore we also present a slightly modified approach to handle smaller time uncertainties through refining the time steps in the simulation process.

The event occurrence time uncertainty interval is much less than the interval considered in the standard approach therefore we need to modify the current

approach to be able to handle small time uncertainty intervals. For the 1D Gaussian time Kernels with small bandwidth (time uncertainty) to have effect on the previous and next time steps the time steps should be refined. So we propose to refine the time steps in the simulation procedure. We can both do time step refinement at all time steps (that makes the simulation time really lengthy by increasing the simulation time by many folds because of so many new refined time steps) or only at integration steps which we call adaptive time step approach. To implement a more time efficient approach we propose to apply the time step refinement only at integration steps (adaptive time step refinement) which still increases the simulation time by a few folds but it is still more computationally efficient than the full time step refinement. Generally the proposed (adaptive) time step refinement approach is able to handle any small time uncertainty interval. Here in this study we implement time step refinement approach to handle the event time uncertainties (1D time Gaussian kernel bandwidth) in the order of seconds and milliseconds [93, 94, 95, 96].

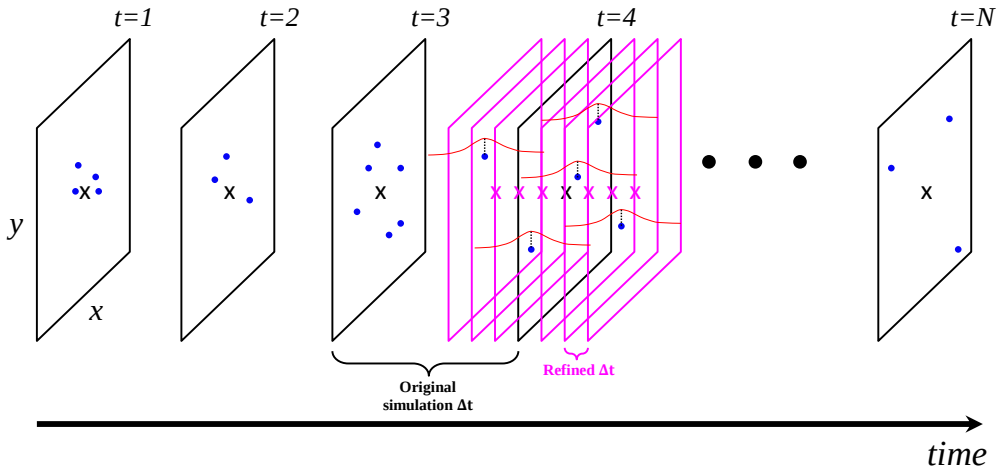


Figure 2.9: The schematic of consideration of event occurrence time in 2D model using KDE method along with time step refinement. The well location is indicated by x mark in the middle of the field. The MEQ clouds are incremental clouds. Blue dots illustrate the events that happen in each time step. Red Gaussian curves $s_t(t)$, show the density of each event in time. For simplified and clear illustration the Gaussian kernel is placed only for one of the time steps as well as the time step refinement.

In Figure 2.9 the schematic of adaptive time step refinement approach is presented. The refinement of original simulation time step to refined time steps are performed only at integration time steps. With this approach we are able to handle very small time uncertainties in the SSBRC method.

In the process of microseismic source inversion to detect the location and occurrence time of MEQ events, a main ingredient is earth velocity model which can be considered as the major source of uncertainty. Since the location and time of MEQ events are related through velocity model, the uncertainty can be addressed via either location or time uncertainty. Thus the incorporating of uncertainty (mainly the uncertainty in velocity model) in SSBRC procedure can be performed through either event location uncertainty consideration or event occurrence time uncertainty

[93, 97, 98].

The numerical example will be presented with consideration of event occurrence time uncertainty in Section 4.4.4.2 and as we expected by adding more uncertainty to the inversion procedure, the estimated property distribution loses a bit of its similarity to true property distribution comparing to the case with only event location uncertainty and also estimation RMSE will increase slightly.

2.4 Discussion of Physical Relationship of Induced Microseismic Events and Rock Properties

Fracture reactivation or rock failure occurs when the change in pressure and/or temperature moves the stress state closer to the failure threshold or criteria. The microseismic event can be the result of intact rock failure (crack or fracture initiation) or the reactivation of any kind of naturally-occurring or preexisting discontinuity or weakness such as a natural fracture, fault, joint, or weak bedding plane [99]. This also should be noted that in fact there is no strictly intact rock and all rocks have in some way natural weaknesses. In this section we discuss the physical relationship of the MEQ observations and rock properties and the basis for the SSBRC method capability to infer rock properties from MEQ data.

In general, change in stress state can be the results of pore pressure change (injection/production) and/or temperature change (rock expansion/contraction). For instance cold injection causes condensing in the rock (produces tensile stresses) which consequently shifts stress state (Figure 2.10) to tensile threshold and leads to tensile failures. On the contrary hot water injection or warming the reservoir produces compressive stress and as long as the Mohr circle size is constant there will not be any failure and it will stabilize the reservoir. However the Mohr circle will become larger eventually and leads to shear failure. In general compressive

stress stabilizes the reservoir. Generally thermal effects are very slow and have less geomechanical effects and for instance in the case of thermal EOR cases failure might happen after months or years.

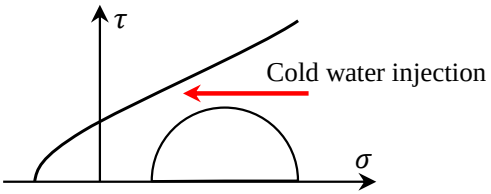


Figure 2.10: Effect of cold water injection in stress state change and rock failure.

Young's modulus (elastic modulus) is one of the important playing rock properties in rock failure and MEQ observations. A rock with high Young's modulus is a competent rock (e.g. granite) which needs much more stress change for making the same displacement comparing to a low Young's modulus rock (e.g. sandstone). Hence for injecting the same amount of water (same pore pressure change) to the reservoirs with the same porosity there will be much more MEQ events (because of more stress) in granite rock comparing to sandstone. High Young's modulus rock resists against pore pressure change and produces higher stress which leads to more MEQ events. Higher (lower) Young's modulus results in larger (smaller) Mohr circle and this increases (decreases) the probability of touching the failure criterion. Change in pore pressure has a significant effect on MEQ generation so the injection rate affects the rock failure considerably. Permeability has indirect effect on stress change (through pore pessure change) so MEQ sensitivity with respect to permeability is not significant. Pore pressure change both causes shift of Mohr circle and change of its size (comes from coupling of pore pressure and stress).

Pore pressure increase shifts the Mohr circle to the left and because of the interchangeable effect of pore pressure and stress (through Navier–Stokes equations), the principal stresses also change and leads to the change of Mohr circle size. This fully coupled effects of pore pressure and stress leads to stress state change and probably rock failure if Mohr circle is close enough to the failure thresholds.

In the presented numerical experiments with the proposed SSBRC method, we aim to represent the random preexisting distribution of natural weaknesses, fractures and cracks in the rock by a rock property distribution so we followed the approach in [26] where the criticality distribution (critical pore pressure) is considered a spatially random distribution. Therefore in our application we considered the rock strength properties (tensile strength and cohesion) to be spatially random to represent the random distribution of preexisting natural weaknesses. Consequently the resulted MEQ cloud which is generated by the direct effect of rock strength property through failure criteria will show somehow a random spatial behavior and may not directly follow the shape of the elastic modulus however it certainly carries the physical effect of elastic modulus or other rock properties values. The random representation of rock strength properties leads to spatially random propagation or widely scattered MEQ events. The randomness of MEQ cloud comes from the randomness of the considered rock strength distribution (e.g. Tensile strength map).

Rock strength properties like tensile strength T_0, cohesion C, and friction angle ϕ (or friction factor) which are directly present in the failure criteria can be easily related to the MEQ distribution. For instance wherever T_0 (absolute) value is low (with the convention of compression being positive, T_0 is always negative so low T_0 value means T_0 close to zero), with a very small pore pressure increase (Mohr circle shifted to left) or stress increase (Mohr circle becomes larger) Mohr circle will touch Tensile failure criterion line and failure will happen, while wherever T_0 has higher

(absolute) value the tensile criterion line is far from Mohr circle and change in position or size of Mohr circle will not cause crossing the tensile failure line. Therefore MEQ distribution clearly and directly guides us in inferring T_0 in the reservoir (Figure 2.11).

In the case of cohesion, lower cohesion takes down the shear failure line so change in Mohr circle easily makes it touch the shear failure line and produce MEQ event while higher cohesion bring the shear failure line up and makes it difficult for Mohr circle to touch it so higher cohesion will result in less shear damage or failure. (Figure 2.11)

In the case of friction angle, by reducing the friction angle (or friction factor) the shear failure line will become more flat and makes it easier for Mohr circle to touch it so lower friction angle will result in more MEQ events, on the contrary higher friction angle will results in less MEQ event, therefore MEQ event distribution (its density and location of events) will help us identify the friction angle or factor values.

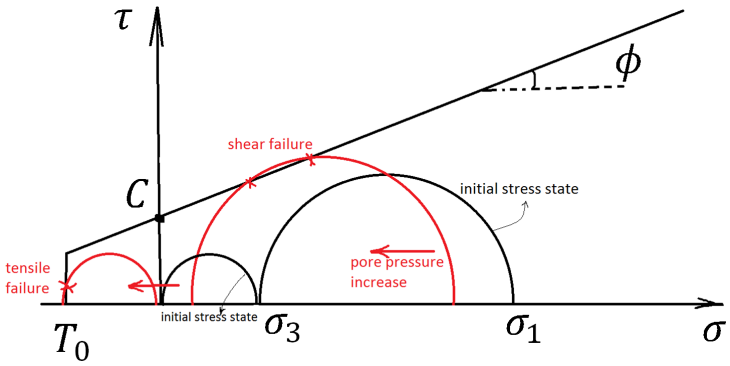


Figure 2.11: A schematic illustration to show the coupled effect of pore pressure change and stress state change on how the failure can happen (showing two different initial stress states).

In the case of Young's modulus E, Higher E means more competent rock, with same injection rate or pore pressure increase or injected fluid volume more stress will build up in higher E (hard) rock comparing to low E (soft) rock so higher E rock will probably touch failure criteria faster and produce more failures or MEQ events, therefore the location, time and density of MEQ can be informative about the elastic modulus of the rock. Based on constitutive equations of thermo-poroelasticity presented in Equations (2.3),(2.4), change in pore pressure (resulting from fluid injection) results in stress change in the reservoir.

In the case of permeability, high permeability makes the pore pressure front travel faster in the reservoir and consequently become higher so high permeability will result in earlier events away from injection point (not very dense) and low permeability does not let pore pressure go much farther and results in pore pressure build up leading to late events (more dense), therefore this MEQ characteristics help us in inferring permeability distribution in the reservoir.

The conventional model (versus the coupled model) for brittle failure caused by pore pressure change assumes that total stresses are independent of pore pressure. Therefore failure type (shear or tensile) in a conventional (uncoupled) model of brittle failure is determined fully by the initial maximum differential stress magnitude, with high initial differential stresses inducing shear failure and low differential stresses leading to tensile failure [100].

The rock strength properties like Tensile strength and Cohesion have much more noticeable effect on MEQ cloud than Young's modulus and permeability because they have very straightforward and direct relation with happening a failure in a point (through failure criteria which involve Tensile strength and Cohesion). To best show the effect of each parameter in MEQ cloud we performed an extensive sensitivity analysis that the final results are shown in Figure 2.12 as a Tornado diagram. As

shown the parameters are ordered in the diagram based on their effect on the number of generated MEQ events. For sure the most effective parameter is injection rate which controls the amount of pore pressure disturbance in the field. The least effective parameter as expected is permeability. The second most effective parameter as anticipated in Tensile strength which through failure criterion has direct effect on happening a failure. The third most effective parameter is Youngs modulus which its increase leads to more failures and vice versa. To generate the Tornado diagram, first the model is run for a set of reference parameters then for each parameter the rest of playing parameters are kept fixed and we change the target parameter by ± 10, ± 25, and ± 50.

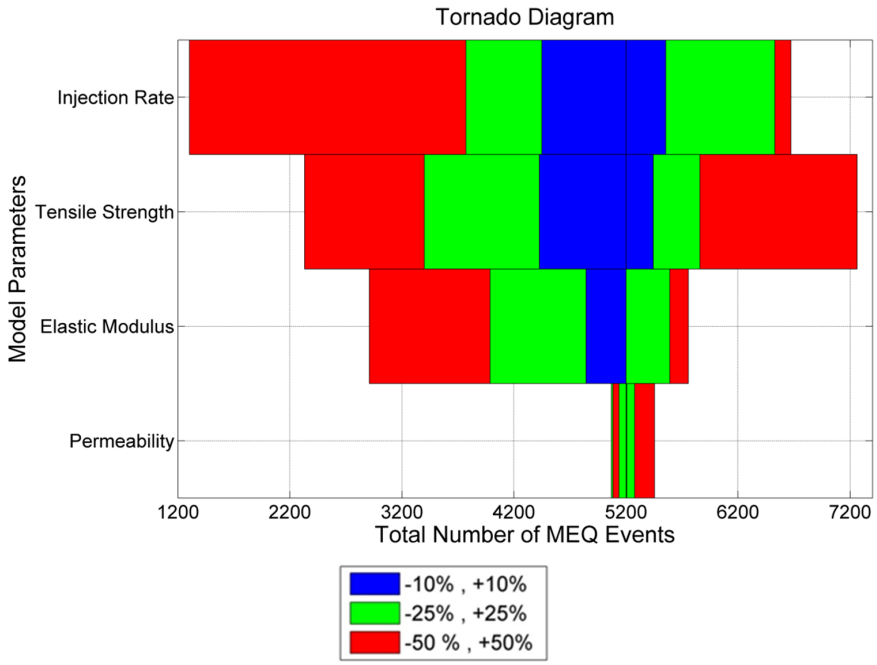


Figure 2.12: Tornado diagram of sensitivity analysis of MEQ events with respect to playing parameters.

2.5 Motivating Real Field Microseismic Data and Event Triggering Mechanism

Firstly the developed geomechanical reservoir simulator (by *Lee* et al) as well as the developed data inversion or parameter estimation method (SSBRC method) are intended for enhanced geothermal systems (EGS). The enhanced geothermal systems and geothermal water circulation systems (utilizing one or two wells for injecting cold water and producing hot water or steam) are not hydraulically fractured or stimulated in multiple stages (multiple stage completion) as in wells in unconventional oil and gas reservoirs or plays hence the usual multiple stage separation of induced MEQ events in hydraulically fractured oil or gas wells is not typical in geothermal systems. Examples of real field propagated and scattered microseismic cloud in EGS as well as oil and gas reservoirs are presented next.

The induced seismicity cloud in the Geysers geothermal field [1] (the world's largest geothermal field, located in the Mayacamas Mountains north of San Francisco, California) is shown in Figure 2.13 where the MEQ events are very diffused and also covered the domain.

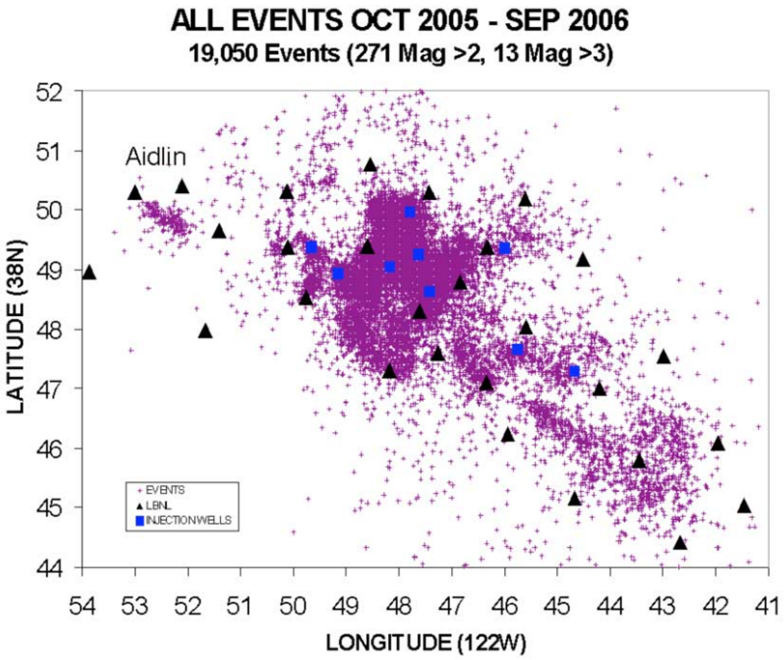


Figure 2.13: A plan view of Geysers geothermal field. The triangles are the locations of the seismic monitoring stations and the blue boxes are the locations of some of the injection wells [1]

In the performed experiments the MEQ clouds more look like (and are intended to model) the observed MEQ cloud in geothermal systems where the resulted MEQ cloud is very diffused and scattered unlike the multi-stage hydraulically fractured unconventional oil and gas reservoirs. Also because of existing highly interconnected natural fractures in the geothermal reservoirs we usually see a very scattered MEQ cloud which is the result of a very complex and interconnected fracture network. In the geomechanical coupled reservoir model the resulting damaged elastic modulus and permeability basically represent the newly created (induced) fractures, activated fractures and the fracture propagation, therefore the updated (enhanced) permeability or damage factor (D) actually reveals the fracture propagation structure in the reservoir.

Also it should be noted when the boundary stress anisotropy is low (very close $S_{H,max}$ and $S_{h,min}$), the resulted MEQ will not elongate in any specific direction and MEQ will be very diffused which is mostly the case in the performed numerical experiments.

In the performed numerical experiments the purpose is to show the capability of SSBRC method to infer rock properties from the observed MEQ cloud so the presented examples are essentially focused on (and modeled as) the areas around the well and where the MEQ data is available because where MEQ is not observable estimation performance of the distributed and heterogeneous rock properties by SSBRC method would not be very promising.

Causes of diffused or scattered MEQ cloud:

- Highly stressed initially, critically stressed reservoir: reservoir initially is in the critical condition very close to failure and a very slight disturbance will cause in extensive MEQ cloud
- Extensive natural fractures and preexisting weaknesses in the reservoir
- In geothermal context there is only one well or at most two wells. There is no staging and the fracture initiation point is only one section. There are not stages along the well and it does not look like the typical hydraulic fracturing MEQ data. In geothermal application stimulation is like a single stage hydraulic fracturing and where in multi-stage hydraulic fracturing each stage is assigned a single color (as time parameter) while the resolution of MEQ time is much higher while in the geothermal application there is one stage so the time color coding has more resolution
- Boundary stresses are very close (small stress anisotropy): this also leads to

very scattered MEQ data and not elongated in a specific direction

- By the way the presented hydraulic fracturing example is the procedure of hydraulic injection to all stages in the same time (can be done stage by stage; to make it more realistic)

The monitored microseismic cloud in the German Continental Deep Drilling site (KTB) in two fluid injection experiments in 1994 and 2000 are shown in Figure 2.14 and Figure 2.15 which demonstrate very scattered and diffused MEQ events [49, 2]. It should be noted that in 2000 there was a second injection point - due to leakage (well damage) - that fluid entered the reservoir at 5.3/5.4 km and triggered most of the events (Figure 2.15).

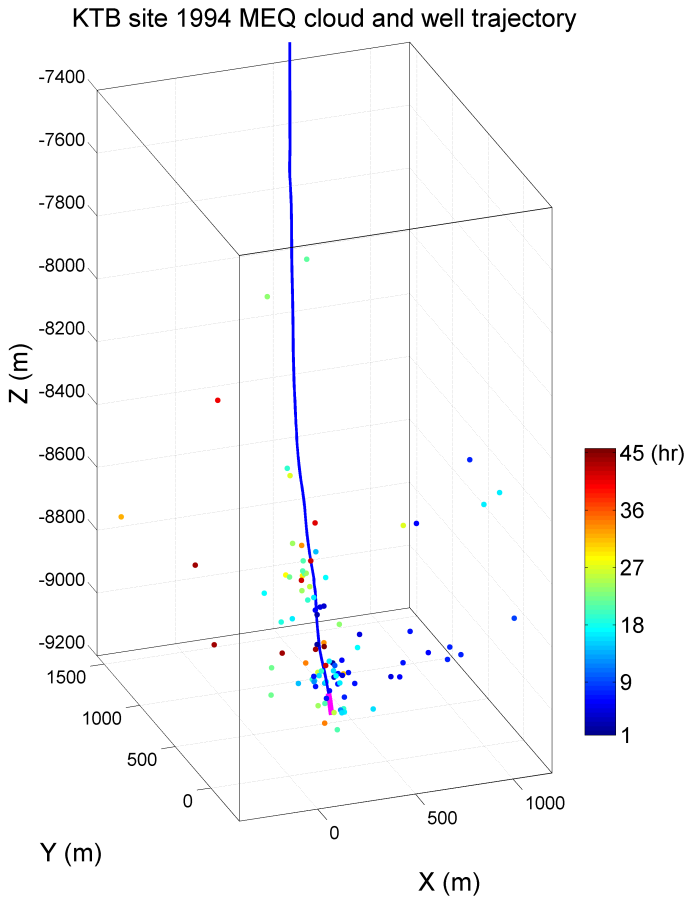


Figure 2.14: Induced microseismic events in 1994 experiment (pink section of the well is the open hole section)

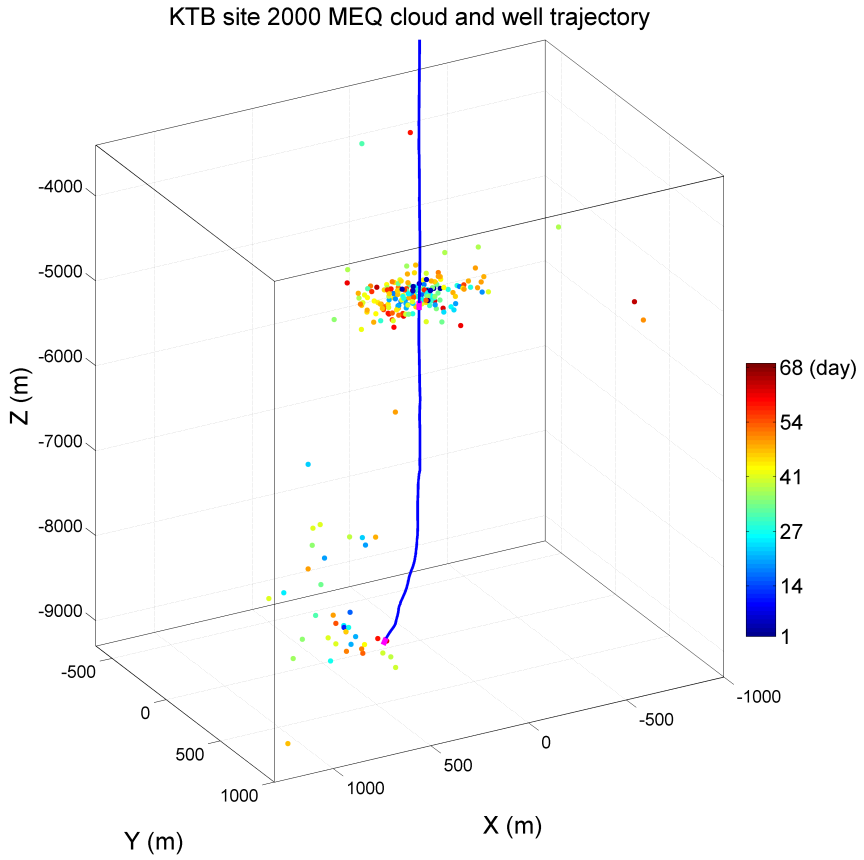


Figure 2.15: Induced microseismic events in 2000 experiment (pink sections of the well are the open hole sections)

3. INVERSE MODELING APPROACH IN GEOMECHANICAL RESERVOIR MODELS*

We develop an inverse modeling approach called SSBRC method by adopting EnKF to integrate seismicity density observation for inference of hydraulic and geomechanical parameters. By applying SSBRC to the pore pressure diffusion forward model we infer permeability distribution and by employing SSBRC for the geomechanical forward model we are able to estimate permeability, Youngs modulus, tensile strength and cohesion distribution. We also propose some methods to improve the uncertainty quantification and estimation performance of the SSBRC method. The proposed EnKF-based inversion framework begins by generating an ensemble of N_e prior realizations of the reservoir parameter model (e.g. permeability or Youngs modulus) based on prior information (e.g., using geostatistical simulation methods). These realizations are used in the pore pressure diffusion forward model or the coupled FEM forward model to perform a Monte Carlo simulation to predict the pore pressure or stress distributions in the reservoir and then failure criterion is used to predict microseismic events. These MEQ cloud predictions are then converted into continuous seismicity densities, using the KDE method, and used in the EnKF update equation. The major computational cost of the method is related to the forecast or prediction step; however, EnKF can be conveniently parallelized to speed up the computations. The computations are implemented using the pore pressure diffusion forward model or the FEM geomechanical forward model and a parallel EnKF algorithm with MATLABs Parallel Computing Toolbox [101]. In addition,

*Part of this chapter is reprinted with permission from "Inference of permeability distribution from injection-induced discrete microseismic events with kernel density estimation and ensemble Kalman filter" by Mohammadali Tarrahi and Behnam Jafarpour, 2012. Water Resources Research, 48, Copyright 2012 by John Wiley and Sons.

an ensemble of perturbed observations is generated using a zero-mean Gaussian error distribution with a variance value that is obtained through sensitivity study. The main steps involved in the implementation of the proposed SSBRC method are summarized as follows. (The details of SSBRC algorithm is also depicted in Table 3.1)

1. Convert discrete microseismic data (measurements) into quantified continuous seismicity density maps using KDE method (Section 2.3).

2. Generate an ensemble of rock property models from available prior information (e.g., using geostatistical simulation techniques). Repeat steps 3 – 5 until all measurements are processed.

3. For the prediction step, using the developed FEM geomechanical numerical simulator (the pore pressure diffusion forward model), forecast the stress distribution (pore pressure distribution) for each member of the most recently updated ensemble realizations (Section 2.1 or 2.2). This step is implemented in parallel.

4. Predict the microseismic events for each realization by the failure criterion (Section 2.2.2) and convert the results into seismicity density maps using the KDE approach (Section 2.3).

5. Use the EnKF analysis equation with the seismicity density observations from (1) to update the ensemble of reservoir property models (Section 3.1).

The details of each of these steps are discussed next.

3.1 Estimation with Ensemble Kalman Filter

The classical Kalman filter [102] is a sequential state estimation method for characterization of the first and second statistical moments of the states posterior distribution. Hence, the filter fully characterizes the posterior distribution of linear state-space systems that are characterized with jointly Gaussian distributions [102, 103]. The implementation of the filter involves two steps: (1) a forecast step, in which a linear state propagation model is used to predict the mean and covariance of the states at the next time step; and, (2) an analysis step that updates the mean and covariance of the states using the dynamic observations and the forecast states mean and covariance. These two steps are repeated sequentially until all observations are assimilated.

For nonlinear dynamical systems, the EnKF provides a practical approximation of the Kalman filter that has been successfully applied to many applications ranging from hydrology, meteorology and oceanography to groundwater and oil reservoir model calibration (see, [55, 64, 69] and references therein). The sequential formulation of the EnKF distinguishes a forecast (or prior) PDF for the states (augmented vector of geomechanical reservoir parameter and continuous seismicity response $\mathbf{x}_t$) $p\left[\mathbf{x}_t|\mathbf{y}_{0:t-1}\right]$, conditioned on all measurements $\mathbf{y}_{0:t-1}$ taken through time $t-1$, and an updated (or posterior) density $p\left[\mathbf{x}_t|\mathbf{y}_{0:t}\right]$ conditioned on all measurements $\mathbf{y}_{0:t}$ (continuous seismicity response maps) taken through time t. To compute the cross covariance between predicted observations and parameters, the original state vector is augmented with uncertain model parameters (e.g. permeability, Youngs modulus and tensile strength distribution) and predicted measurements [55]. This state augmentation approach can be used to update states and parameters simultaneously. Alternatively, one can only update the uncertain parameters and

derive the updated states by solving the flow equations (or coupled flow, heat and geomechanics equations) with the updated parameters. This is the approach taken in this work. The measurements $\mathbf{y}_t$ consist of seismicity density map $s_t(\mathbf{u})$, defined in Equation (2.26) that represent microseismic measurements in space at time t.

Since the general multivariate PDFs and their statistical moments are difficult to characterize, the EnKF uses a Monte Carlo approximation approach by sampling an initial set of realizations from the high-dimensional prior PDF of the uncertain properties to form an ensemble of reservoir states (and/or parameters). These property maps are then used to generate an ensemble of state and measurement predictions that can be used to compute a sample (prior) covariance matrix for the EnKF update step as described below.

The forecast step in the EnKF can be written as

$$\mathbf{x}^j_{t|t-1} = f_t\left(\mathbf{x}^j_{t-1|t-1}, \mathbf{z}_{t-1}, \mathbf{w}^j_{t-1}\right) \qquad j = 1, ..., N_e \tag{3.1}$$

where $\cdot|t$ represents conditioning on observations up to time t; $\mathbf{z}_{t-1}$ is a vector of known (nonrandom) time-dependent boundary conditions and controls (such as injection rate); and $\mathbf{w}^j_{t-1}$ is a vector of random variables that accounts for modeling errors. The function $f_t(\cdot,\cdot,\cdot)$ represents the state propagation equation from time $t-1$ to time t. The notations j and N_e are used to indicate the realization index and total number of realizations, respectively. In our application, Equation (3.1) represents the solution of the coupled thermo-poroelastic equations (or pore pressure diffusion equation) that describes the time evolution of pore pressure, stress and temperature distributions for each individual realization j of the ensemble reservoir parameter. At time steps when MEQ data are available, the EnKF analysis equation is used to update the reservoir property realizations using the gain matrix

and the misfit between predicted and observed seismicity density maps for each realization. At the update step we use an augmented state vector consisting of spatially distributed reservoir property (parameters to estimate) and realizations of the predicted continuous seismicity density map. After each update we apply a confirmation step [65] by forecasting the future states and predictions from the initial time step with the updated parameters. We repeat the sequence of prediction and update steps until all measurements are integrated.

For a model with N_b nodes (or grid blocks), each reservoir parameter realization $\mathbf{m}^j$ and its corresponding microseismicity density response $\mathbf{s}^j$ are vectors of size $N_b \times 1$. In this work, the reservoir property models are jointly Gaussian random fields that are generated using the sgsim [104] geostatistical simulation technique. The augmented state vector for this case is of the form

$$\mathbf{x}^j = \begin{bmatrix} \mathbf{m}^j \\ \mathbf{s}^j \end{bmatrix} \qquad j = 1, ..., N_e \quad \Rightarrow \quad \mathbf{X} = \begin{bmatrix} \mathbf{x}^1 & \mathbf{x}^2 & \cdots & \mathbf{x}^{N_e} \end{bmatrix} \tag{3.2}$$

The EnKF analysis equation that is used to update each reservoir property realization can be expressed as

$$\begin{aligned} \mathbf{x}_u^j &= \mathbf{x}_f^j + \mathbf{K}\left(\mathbf{y}^j - \mathbf{H}\mathbf{x}_f^j\right) \\ \mathbf{K} &= \mathbf{C}_\mathbf{x}^e \mathbf{H}^T \left(\mathbf{H}\mathbf{C}_\mathbf{x}^e \mathbf{H}^T + \mathbf{C}_\mathbf{d}\right)^{-1} \end{aligned} \tag{3.3}$$

where $\mathbf{K}$ is the Kalman gain matrix and the subscripts u and f denote updated and forecast quantities while the superscript e indicates ensemble calculated statistics. The notations $\mathbf{C}_\mathbf{x}^e$ and $\mathbf{C}_\mathbf{d}$ represent the states sample covariance and observation covariance matrices, respectively. The measurement matrix $\mathbf{H}_{N_b \times 2N_b} = [\mathbf{0}_{N_b \times N_b} | \mathbf{I}_{N_b \times N_b}]$, where $\mathbf{0}_{N_b \times N_b}$ and $\mathbf{I}_{N_b \times N_b}$ are zero and identity matrices of the specified dimensions, respectively, acts as a selection operator that extracts the

predicted measurement components from the augmented state vector. The notation $\mathbf{y}^j$ is used to represent the j^{th} realization of the perturbed observations. The states sample covariance $\mathbf{C}_{\mathbf{x}}^e$ can be computed from the ensemble of state vectors

$$\begin{aligned} \mathbf{C}_{\mathbf{x}}^e &= \frac{1}{N_e - 1} \sum_{j=1}^{N_e} \left(\mathbf{x}_f^j - \bar{\mathbf{x}}_f^j\right) \left(\mathbf{x}_f^j - \bar{\mathbf{x}}_f^j\right)^T \\ \bar{\mathbf{x}}_f^j &= \frac{1}{N_e} \sum_{j=1}^{N_e} \mathbf{x}_f^j \end{aligned} \tag{3.4}$$

where $\bar{\mathbf{x}}_f^j$ is used to denote the ensemble mean of the forecast states (that is, the reservoir property distribution from the previous step and the corresponding microseismic response forecasts). In the EnKF implementations, the covariance matrix in Equation (3.4) need not be constructed explicitly and the update can be applied using its low-rank representation though a compact SVD implementation. The covariance matrix in Equation (3.4) contains the covariance information about the reservoir parameter field as well as the cross covariance information between the reservoir parameter and (microseismic) measurements. It is the latter cross covariance that allows the estimation of uncertain geomechanical reservoir parameter distributions from microseismic observations. This relation bears similarity with the use of covariance and cross covariance in the kriging/simulation [104, 105] and cokriging/cosimulation [106, 107, 108, 104] methods, respectively. Note that in Equation (3.3), the term $\left(\mathbf{y}^j - \mathbf{H}\mathbf{x}_f^j\right)$ is the misfit between the j^{th} perturbed observation and prediction, which in this case represents the observed and predicted continuous map of seismicity density. Several remarks regarding the update equation for our problem will follow.

In addition to nonlinearity in the forward coupled geomechanics and flow model, a complexity of the measurement model in our application is the nonlinear failure

criterion (i.e., hard truncation) that is used to convert the continuous stress distributions to discrete microseismic events. The Gaussian kernel that we apply to convert the MEQ predictions to continuous maps of seismic density makes the data more amenable to processing with the EnKF. However, the relationship between the magnitude of stress and the resulting seismicity map remains complex.

Considering the dynamic alteration of reservoir geomechanical parameters (Young's modulus and permeability) in our developed coupled FEM simulator based on the damage and permeability model (in Section 2.2.2), in the SSBRC inverse modeling approach we estimate geomechanical properties of the intact (initial or undamaged) rock. It is also consistent with our EnKF data integration approach with confirmation step [65] that for each EnKF analysis step we run the forward model from the beginning (initial state) by updated or estimated intact rock properties. Therefore to obtain the properties of the final stimulated or damaged reservoir rock we just need to run the geomechanical simulator with the estimated intact rock properties.

In our EnKF implementation, to perturb the observations, we add an uncorrelated realization from a Gaussian random noise, with a specified observation covariance matrix $\mathbf{C_d}$, to the value of the observed quantities. We assume that the observation error standard deviation (Std) is proportional to the value of the observed quantity and compute the diagonal elements of the observation error matrix as

$$\sigma_k^2 = \left(\sigma_{min} + \frac{(\sigma_{max} - \sigma_{min})\,(y_k - y_{min})}{(y_{max} - y_{min})} \right)^2, \qquad k = 1, 2, ..., N_b \tag{3.5}$$

where σ_k^2 is the observation variance at the k^{th} node or grid block (the k^{th} diagonal entry for the observation covariance matrix), σ_{max} and σ_{min} are the minimum and maximum standard deviations specified for the observations, respectively. The

notation y_k represents the observed seismicity density at location k while y_{min}, y_{max} represent the minimum and maximum observed values of the seismicity density, respectively. The realization j of the perturbed observation at location k, can then be written as

$$\begin{aligned} y_k^j &= y_k + \varepsilon_k^j, \qquad j = 1, 2, ..., N_e \\ \varepsilon_k^j &\sim N\left(0, \sigma_k^2\right) \end{aligned} \tag{3.6}$$

In this work, we assume an uncorrelated Gaussian observation error with zero mean and standard deviation obtained from Equation (3.5). We note that other methods for generating the perturbed observations may also be considered. In particular in Section 3.2.2 as one of the methods to improve uncertainty quantification of SSBRC, we first generate perturbed observations by perturbing the kernel bandwidth and given the large dimension and the spatial correlation that may exist between the observation errors, we then propose to assimilate the resulting observations in a low-rank subspace defined by the left singular vectors of the ensemble observations perturbations matrix in a similar way to *Kepert* [109].

The given SSBRC formulations here are based on the assumption of estimating one of the reservoir properties distribution (assuming one property unknown and the rest known) which can be simply extended to simultaneously estimating more than one reservoir property distribution by only augmenting reservoir parameter vectors in the EnKF state vector.

To generate the ensemble of reservoir parameter realizations, we used a variogram-based geostatistical simulation method with specified variogram parameters. The sgsim algorithm [104] was used to implement the geostatistical simulations. In real applications, the number of realizations is typically determined through a trade-off between available computational resources and the desired statistical accuracy in computing the required sample statistics. For large-scale

problems where the number of realizations is limited, practical considerations such as localization or local analysis [110] have been proposed to avoid inaccurate updates due to spurious (nonphysical) correlations and to reduce the possibility of an ensemble collapse. As the results of the numerical experiment in Section 4.2 and 4.4 show, the proposed SSBRC method with standard EnKF results in severe ensemble spread underestimation which will be resolved by proposed methods in section 3.2. In the examples that follow, we implement the EnKF algorithm with $N_e = 100$ and do not apply any localizations.

The detailed steps of SSBRC method with parallel EnKF algorithm is shown in Table 3.1.

Parallel EnKF Pseudo Code
1: Generate N_e initial parameter ($\mathbf{m}^j$) realizations
2: Generate perturbed observations ($\mathbf{y}^j$) from true observation (based on $\mathbf{C}_d$)
3: **For** $t_i = 1$ **to** t_N **do** (*integration time steps*)
4: **Par-For** $j = 1$ **to** N_e **do** (*run in parallel on different available cores*)
5: Initialize the geomechanical simulator
6: Write the j^{th} realization ($\mathbf{m}^j$) as the reservoir parameter
7: Run simulator from beginning until current integration time (corresponds to t_i)
8: Generate the corresponding seismicity cloud ($\mathbb{S}^j$)
9: Use KDE to convert seismicity cloud ($\mathbb{S}^j$) to seismicity density ($\mathbf{s}^j$)
10: **End Par-For**
11: Calculate $\bar{\mathbf{x}}_f^j$, $\mathbf{C}_{\mathbf{x}}^e$, $\mathbf{K}$
12: Update realizations by EnKF analysis equation
13: **End For**

t_i = integration time step index which corresponds to integration time
t_N = the total number of integration time steps (in this study = 6)
Par-For = parallel **For** loop which executes its underlying commands in parallel
$\mathbb{S}^j$= the simulated seismicity cloud corresponding to each $\mathbf{m}^j$

Table 3.1: Parallel EnKF algorithm for SSBRC.

The full forward model simulator that relates the hydraulic or geomechanical parameter distribution to the microseismicity density map consists of geomechanical simulator (relating geomechanical parameter distribution to microseismicity cloud) and microseismicity continuous interpretation (relating microseismicity discrete cloud to microseismicity continuous density map). Figure 3.1 shows the schematic of the full forward model. In the propagation (Monte Carlo simulation) step of the EnKF procedure we run the full forward model for all ensemble members (geomechanical parameter samples). For the pore pressure diffusion forward model that relates permeability distribution to MEQ cloud the same full forward model including MEQ continuous interpretation is constructed.

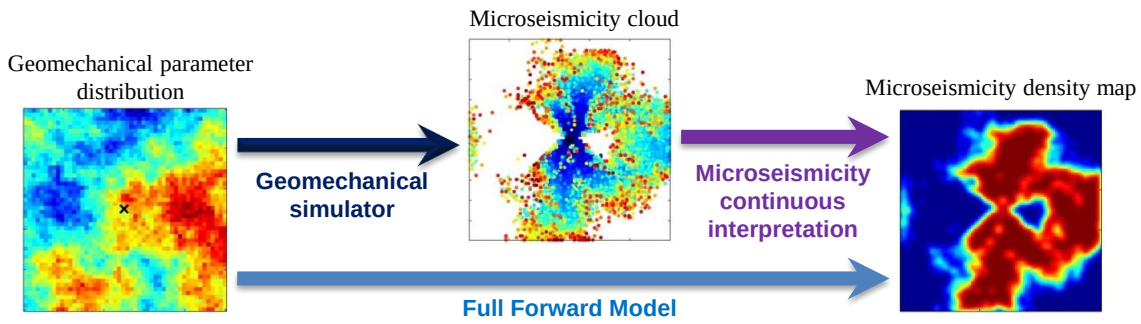


Figure 3.1: Full forward model relating geomechanical parameter distribution to microseismicity density map.

Figure 3.2 shows the schematic of the parallel EnKF pseudo code. We perform full forward model of the parameter ensemble (propagation step) in parallel. Since running the forward model is completely independent for different parameter samples, we are able to run some samples simultaneously on the available cores of the machine. It should be noted that Figure 3.2 represents the estimation procedure in each integration step. For instance in the first integration step, the first column of

Figure 3.2 is the initial ensemble of parameters and after integration of the first true seismicity density map, we will obtain the first updated ensemble (the last column of Figure 3.2) which will be the input ensemble (the first column of Figure 3.2) for the second (next) integration step.

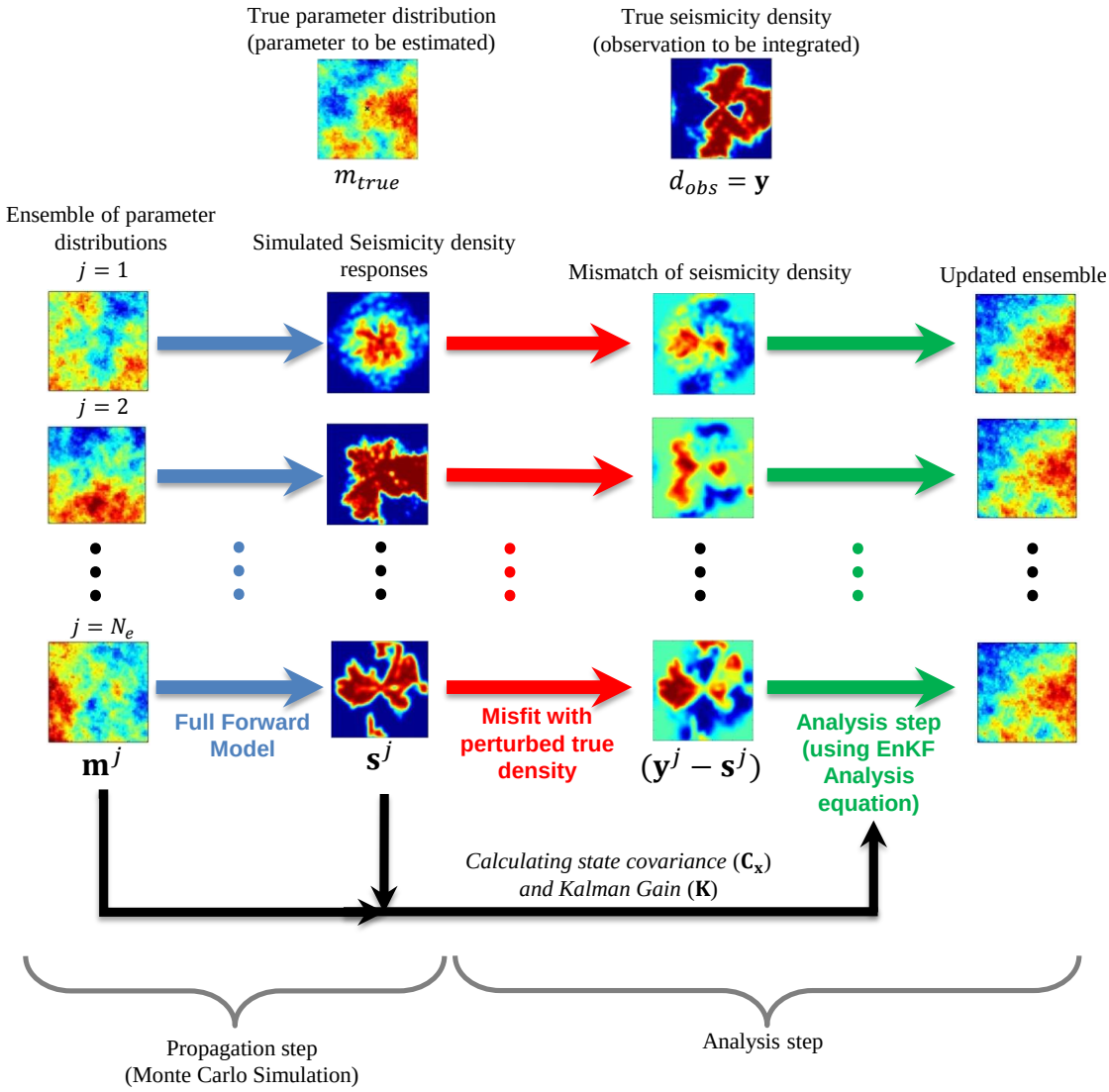


Figure 3.2: EnKF procedure (in each integration time step).

3.2 Improved Uncertainty Quantification

In this section we focus on quantification of uncertainty that is a key concept of application of the EnKF data assimilation approach. It is well known that for large scale datasets, such as seismic data [72] and high resolution spatial map of seismicity density, application of the standard EnKF without taking into account the spatial correlation in the observations can lead to underestimated solution uncertainty or ensemble spread. Underestimation of ensemble spread is not favorable as it can introduce unrealistic confidence in potentially inaccurate future predictions and decreases the likelihood of capturing the true behavior of the reservoir. High dimensional observation also leads to expensive computational load in updating scheme of EnKF [55]. In this situation, severe underestimation of the prediction uncertainty can results in biased forecasts and an ensemble collapsing into a single realization.

The standard SSBRC implementation was based on generating seismicity density maps on the same grid system or mesh structure that was used for describing the hydraulic or geomechanical property distributions (Section 2.3). A byproduct of this implementation is that it artificially increases the data resolution and, hence, the number of data that will be assimilated during the EnKF update step. As another significant byproduct of this preprocessing step, the resulting maps can exhibit spurious spatial correlations in the observed data and create unrealistic correlations between rock properties and microseismic data, thereby degrading the performance of the EnKF update.

We first show the estimation results for the SSBRC approach using the standard EnKF algorithm to illustrate the underestimation of ensemble spread (several numerical examples which are reported in Section 4.2 and 4.4 confirm it), and

then propose three methods to resolve this issue. The first and simplest approach to overcome ensemble spread underestimation is to increase the observation error artificially by using a large variance for the observation noise. In the other two methods, we reduce the number of observations first by using a spectral projection (spectral dimension reduction) approach and second by coarsening the seismicity density map (spatial dimension reduction). In projection approach, the ensemble of perturbed observations are projected to a reduced subspace that is defined by the leading left singular vectors of the observation matrix. This step is aimed at decorrelating the original observations of the seismicity map. The EnKF update is then used to assimilate the resulting low-dimensional description of the data. In the second approach, we use a coarse grid system for interpretation of the seismic events. This approach is very similar to the original SSBRC implementation except that it uses a coarse-scale grid system or mesh structure in KDE-based continuous seismicity interpretation in Equations (2.26),(2.27) to make lower resolution density maps. Then the reduced dimension or coarse seismicity density maps are used in the EnKF update equation. Additionally in general, observation space dimension reduction (either spectral or spatial) improves the computational efficiency of the analysis step of the EnKF.

We present the estimation results for these three methods following an introductory example to illustrate underestimation of uncertainty when the standard EnKF is used with large-scale seismicity density maps as observed data (in Section 4.2 and 4.4).

3.2.1 Inflated Observation Error Variance

The simple way to reduce the underestimation of ensemble variance is to increase observation error variance. The effect of this remedy is equivalent to damping the

EnKF updates, the term $\mathbf{K}\left(\mathbf{y}^j - \mathbf{H}\mathbf{x}_f^j\right)$ in Equation (3.3) or reducing the weight of observations (or Kalman gain, $\mathbf{K}$). To do so, we increase σ_{min} and σ_{max} in Equation (3.5) which leads to diminishing the effect of seismicity density observation in the EnKF update and finally results in preserving the ensemble spread during data assimilation. While the level of noise considered goes beyond most practical situations, this provides a simple way to improve the underestimation effect as shown in Section 4.2 and 4.4. We also summarize the results from several experiments with increasing level of observation error variance. From the performed experiment results it is clear increasing the observation error variance can improve the underestimation effect observed in the standard EnKF. However, a main drawback of this approach is that it is not clear how the introduction of significant noise to the observed quantities will generally impact the quality of the EnKF update beside the damping effect. Hence, we consider two other alternative methods that do not corrupt the observations.

3.2.2 Reduced-Order Projection

The EnKF updating scheme particularly with high-dimensional observation can suffer from the problem known as filter divergence [111], resulting from rank issues [112] and estimation uncertainty [58]. To reduce these problems, *Evensen* [55, 112] introduced dimension reduction techniques in an EnKF setting. *Skjervheim et al.* [72] also suggested an alternative EnKF updating using well-known dimension reduction techniques.

In this section we first propose a new method for generating an ensemble of perturbed observations then we use a dimension reduction method to reduce the number of observations integrated in EnKF update equation. The proposed method of perturbing seismicity observation is completely different than the typical procedure

of adding Gaussian random noise to the observation (in Section 3.1). To exhibit and finally capture the spatial correlation or redundancy of observation in more efficient fashion, we propose to generate each perturbed observation realization by perturbing the bandwidth of Gaussian kernel. In the standard observation perturbation method explained in Section 3.1, we perturb the observed seismicity density map however here we use the observed seismicity cloud and generate perturbed observations by perturbing the kernel bandwidth in KDE based quantification.

After converting the discrete microseismic measurements to continuous seismicity density maps using Gaussian kernels, the resulting observations exhibit strong spatial correlations. Hence, this correlation (or redundancy) should either be taken into account during the update or should be removed from the data. To remove the correlations in the observations, we project the ensemble of perturbed observations onto a low-dimensional subspace defined by the leading left singular vectors of the observation matrix [109]. During the EnKF update, we use the transformed observations (after projection to the mentioned subspace) for data assimilation. To implement the update, the predicted observations must also be projected onto the same subspace. The procedure is described below.

In the projection approach, at each integration step we first choose a kernel bandwidth h, and its standard deviation σ_h, and then to make each perturbed observation realization, we individually perturb kernel bandwidth h, for each MEQ event of the true observation using a Gaussian distribution as below:

$$h_i^j \sim N\left(h, \sigma_h^2\right); \qquad j = 1, 2, ..., N_e; \qquad i = 1, 2, ..., n_{MEQ} \tag{3.7}$$

where N_e and n_{MEQ} are the number of realizations and the number of MEQ events at the specified integration step, respectively. Superscript j, and subscript

i, indicate realization index and MEQ event index, respectively. Therefore, the j^{th} perturbed observation realization $\mathbf{y}^j$ is made of a set of perturbed bandwidths h_i^j $(i = 1, 2, , n_{MEQ})$ as follows:

$$\mathbf{y}^j = \frac{1}{n_{MEQ}} \sum_{i=1}^{n_{MEQ}} K_i^{MEQ}\left(\mathbf{u}; h_i^j\right) \tag{3.8}$$

Equation (3.8) is the same as Equation (2.26) but for the Gaussian kernels $K_i^{MEQ}\left(\mathbf{u}; h_i^j\right)$, which have different bandwidths. Afterwards we make the perturbed observation ensemble as:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}^1 & \mathbf{y}^2 & \cdots & \mathbf{y}^{N_e} \end{bmatrix} \tag{3.9}$$

To project the observation to a lower dimension space we take the SVD of $\mathbf{Y}$ to obtain the matrix of eigenvectors which is the projection matrix $\mathbf{U}$. Columns of $\mathbf{U}$ are eigenvectors spanning the space made by $\mathbf{Y}$. A finite number of the leading left singular vectors of $\mathbf{Y}$ form a low-dimensional subspace defined by columns of the matrix $\mathbf{U}$ that accurately approximate each observation realization. Since $\mathbf{Y}$ has a maximum rank of N_e, the maximum dimension of the transformed observations is N_e. To reduce the dimension of the observation space, non-leading columns of $\mathbf{U}$ can be truncated. The truncation number n_{trunc} varies in the range 1 to N_e. The projected perturbed observation ensemble $\mathbf{Y}_p$ is calculated as

$$\begin{aligned} \mathbf{Y}_p &= \mathbf{U}_{trunc}^T \mathbf{Y} \\ \mathbf{U}_{trunc} &= \begin{bmatrix} \mathbf{u}^1 & \mathbf{u}^2 & \cdots & \mathbf{u}^{n_{trunc}} \end{bmatrix} \\ \mathbf{U} &= \begin{bmatrix} \mathbf{u}^1 & \mathbf{u}^2 & \cdots & \mathbf{u}^{N_e} \end{bmatrix} \end{aligned} \tag{3.10}$$

where $\mathbf{u}^j$ is the j^{th} eigenvector. We also apply the same projection to the predicted seismicity density maps. The ensemble of predicted seismicity density is made by

the same kernel band h.

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}^1 & \mathbf{s}^2 & \cdots & \mathbf{s}^{N_e} \end{bmatrix} \tag{3.11}$$

where $\mathbf{s}^j$ is the j^{th} simulated seismicity density map which is made of j^{th} simulated seismicity cloud $\mathbb{S}^j$. The projected ensemble of predicted seismicity density maps is then calculated as:

$$\mathbf{S}_p = \mathbf{U}_{trunc}^T \mathbf{S} \tag{3.12}$$

The resulting observed and predicted data for the EnKF update step are $\mathbf{Y}_p$ and $\mathbf{S}_p$. It should be noted that projection method has two tuning parameters; kernel bandwidth standard deviation σ_h, and truncation number n_{trunc}, and the observation error standard deviation parameters σ_{min} and σ_{max} are no longer needed.

Application of this method to improve uncertainty quantification performance of SSBRC is shown in Section 4.2 and 4.4 by numerical examples. As we will see, reduced-order projection of seismicity density observation both preserves the ensemble spread and improves the computational efficiency.

3.2.3 Coarse-Scale Microseismicity Density Map

Another approach to reduce the dimension of the seismicity map is to use a coarse scale description. This approach uses a coarser scale grid system or mesh structure to quantify the seismicity observations. The number of observations is equal to the dimension of seismicity density map. To reduce the dimension of the seismicity density map (number of observations to integrate) which is the major reason of the ensemble spread underestimation, we can build the continuous function of seismicity density on a coarser mesh or grid system $\mathbf{u}_{red}$, instead of the original FEM fine mesh (or original fine grid block configuration for pore pressure diffusion forward model) $\mathbf{u}$, in Equations (2.26),(2.27). So we only need to evaluate continuous seismicity density

map on a new coarser grid system. In this work (in 2D experiments) the original mesh configuration of model is square with 100 nodes (or grid blocks) at each side N_X, that results in 10000 nodes (or grid blocks) or seismicity density observation values at u. The coarse-scale mesh (or grid system) is assumed to have $N_{X,red}$ nodes (or grid blocks) at each side which leads to $N_{X,red}^2$ total nodes (grid blocks) or seismicity density observations at $\mathbf{u}_{red}$. In SSBRC with coarse-scale seismicity density, we use a typical range of 5% to 10% for observation error standard deviation.

The result of sensitivity analysis of SSBRC performance with respect to different grid sizes (different number of observation) is presented in Section 4.2 and 4.4 The results demonstrate that while the estimation quality in terms of reservoir parameter map is not affected, the estimation variance is severely underestimated when a large number of correlated observations in a high resolution map is used. The results suggest that the information content of the high resolution map does not provide significant additional details in estimating the reservoir geomechanical parameter.

4. NUMERICAL EXPERIMENTS FOR GEOMECHANICAL RESERVOIR MODELS*

In this chapter we present several numerical examples to show that the distribution of the MEQ events (their source locations) can be used to infer the spatial distribution of the reservoir parameter field. In this work, we have assumed that an interpretation of the microseismic data (through seismic source inversion) in some preprocessing step provides a spatial map of the seismic event locations and then we use the proposed KDE-based continuous interpretation to generate the seismicity density map. Therefore for a reservoir model with N_b nodes (or grid blocks or Gaussian points), at each update step, a vector of N_b observations of seismicity density values is assimilated. The dimension of reservoir parameter vector is also N_b.

In this section we present the results of applying SSBRC to both pore pressure diffusion forward model and geomechanical forward model. The estimation results of homogeneous and heterogeneous 2D and 3D reservoir models in different settings are presented. We first present the application of standard SSBRC and its estimation performance and how it leads to ensemble spread underestimation and then we apply the proposed methods of improved uncertainty quantification along with SSBRC to resolve the issue of spread underestimation.

4.1 Description of Experimental Setup: Pore Pressure Diffusion

In this section we present the results of SSBRC application to single phase pore-pressure diffusion forward model (finite difference numerical modeling with

*Part of this chapter is reprinted with permission from "Inference of permeability distribution from injection-induced discrete microseismic events with kernel density estimation and ensemble Kalman filter" by Mohammadali Tarrahi and Behnam Jafarpour, 2012. Water Resources Research, 48, Copyright 2012 by John Wiley and Sons.

Eclipse [77]) and we also demonstrate the results of applying improved uncertainty quantifications methods for resolving ensemble spread underestimation. We present three sets of experiments covering a two-dimensional (2D) homogeneous and heterogeneous reservoir model, and a three-dimensional (3D) heterogeneous reservoir model. For the 2D example, we consider the estimation of a homogeneous and a heterogeneous permeability model and show that the distribution of the MEQ events can be used to infer the uniform permeability value and the spatial distribution of the permeability field. Our second experiment is based on a 3D reservoir configuration with a heterogeneous permeability model. In these experiments, one water injection well is located at the center of the field and the boundaries are closed to flow (noflow boundary conditions). The injection-induced MEQ events for this injection well are used to estimate the permeability in the reservoir. The 2D examples consist of 100×100 discretized models, leading to $N_b = 10000$ grid blocks. In this work, we have assumed that an interpretation of the microseismic data (through seismic source inversion) in some preprocessing step provides a spatial map of the seismic event locations. Therefore, at each update step, a vector of 10000 observations of seismicity density values is assimilated. In homogeneous 2D model we estimate one single parameter (the value of uniform permeability) from 10000 seismicity observations and in heterogeneous 2D model we estimate 10000 parameters (spatial permeability distribution) from 10000 seismicity density observations. In the 3D example, the reservoir is discretized into a $50 \times 50 \times 30$ ($N_b = 75000$) grid configuration. Also in this case, one injection well is located at the center of the domain, which is perforated throughout the entire thickness of the formation. The source locations of the MEQ events throughout the 3D domain are used to estimate the heterogeneous permeability distribution.

4.2 Results and Discussion

We present and discuss the results of applying our methodology to the experiments described above. We present the results in terms of the estimated property maps and the ensemble statistics prior to and after data integration. As is common in ensemble data assimilation, we use the evolution of reservoir parameter estimation root-mean-square error ($RMSE$) and the ensemble spread (Sp) as performance measures. These measures are computed in each integration step using the following equations:

$$RMSE\left(\mathbf{m}\right) = \frac{1}{N_b}\sum_{i=1}^{N_b}\sqrt{\frac{1}{N_e}\sum_{j=1}^{N_e}\left(m^{i,j} - m^i_{true}\right)^2} \tag{4.1}$$

$$\begin{aligned} Sp\left(\mathbf{m}\right) &= \frac{1}{N_b}\sum_{i=1}^{N_b}\sqrt{\frac{1}{N_e}\sum_{j=1}^{N_e}\left(m^{i,j} - m^i_{mean}\right)^2} \\ m^i_{mean} &= \frac{1}{N_e}\sum_{j=1}^{N_e} m^{i,j} \end{aligned} \tag{4.2}$$

where N_b is the number of parameters (same as number of nodes or grid blocks here), N_e is the number of realizations and $m^{i,j}$ is the i^{th} parameter of realization j. We plot the ensemble spread as a percentage of the initial ensemble spread.

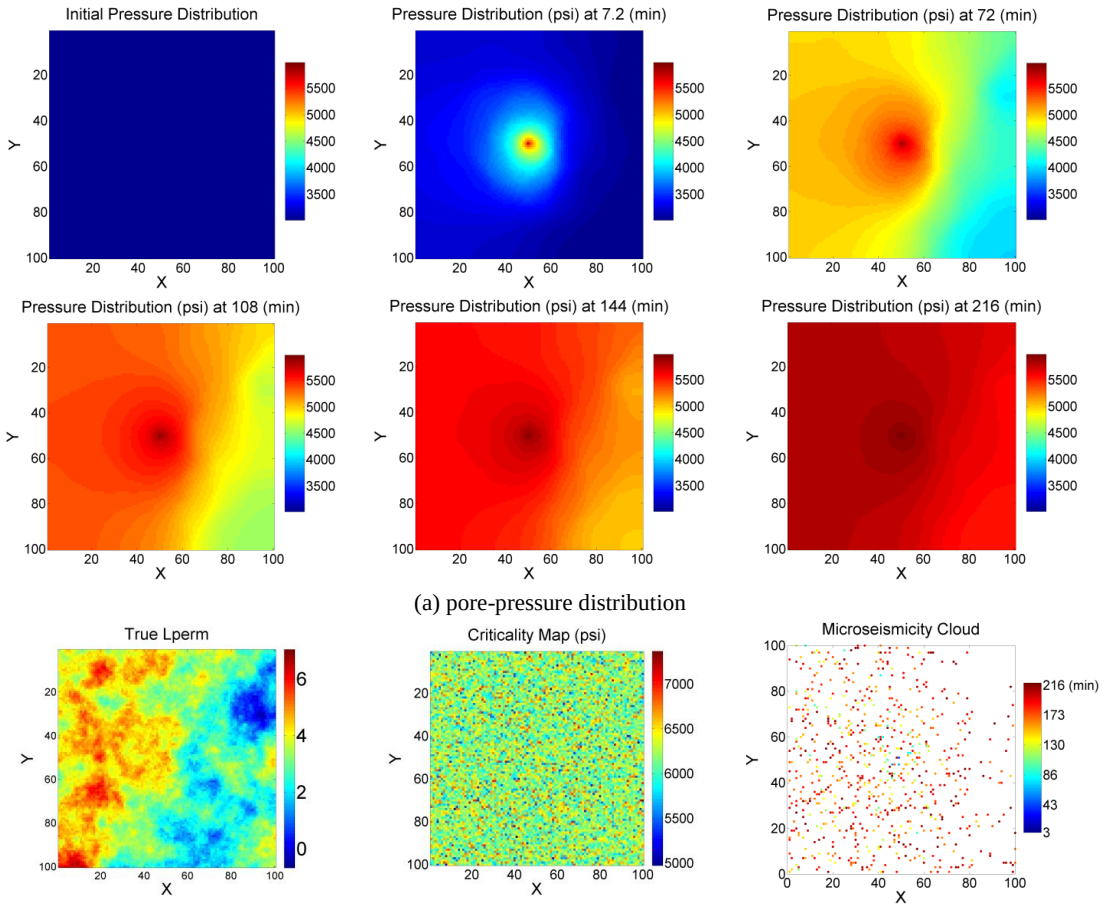
4.2.1 Experiment 1: 2D Heterogeneous

In this section, we apply SSBRC to a 2D pore pressure diffusion reservoir model with heterogeneous permeability distribution. The 2D model has a 100×100 grid block structure (10000 grid blocks) and the permeability field is a heterogeneous parameter. There is only on water injection well at the center of the field. The true log-permeability distribution is shown in Figure 4.1b. In this experiment the dimension of the parameter to estimate and seismicity density observation are both

equal to 10000. An uncorrelated Gaussian random distribution is also assumed for criticality map shown in Figure 4.1c. Other specifications of the true model are illustrated in Figure 4.1 including the pore pressure distribution in the field through stimulation period and the MEQ cloud observation which is generated by comparing pore pressure distributions and the criticality map. Using SSBRC we assimilate MEQ cloud observation, Figure 4.1d, (after quantifying by KDE method) by EnKF update equation to infer heterogeneous log-permeability field, Figure 4.1b. The stimulation period is 216 (min) and we specify 6 integration steps to assimilate MEQ observations. The MEQ density map is constructed on the original fine grid system and the observation error Std range is the typical range of 5% to 10%.

In the following figures in this section, the extensive estimation results of SSBRC are presented. Figure 4.2 shows the result of updating the permeability ensemble in terms of ensemble mean throughout 6 integration steps. It is seen that the estimated ensemble mean is becoming more similar to the true permeability map in Figure 4.2a by assimilating MEQ observations in time. The initial ensemble (almost uniform initial ensemble mean in Figure 4.2b) is completely uninformative about true permeability map, Figure 4.2a, however after MEQ data assimilation the final ensemble mean, Figure 4.2h, is very similar to the true map. Therefore SSBRC proves its suitability for heterogeneous permeability estimation from MEQ data.

The performance measures of SSBRC estimation procedure are also shown in Figure 4.3. The continuous reduction of estimation RMSE, Figure 4.3a, shows the success of SSBRC in estimating permeability distribution. The final RMSE value is almost half of initial RMSE. The ensemble spread is plotted in Figure 4.3b which shows that almost 90% of the initial ensemble spread is lost through estimation procedure. This very low final ensemble spread (10%) shows severe ensemble spread underestimation. This uncertainty quantification issue will be resolved by applying



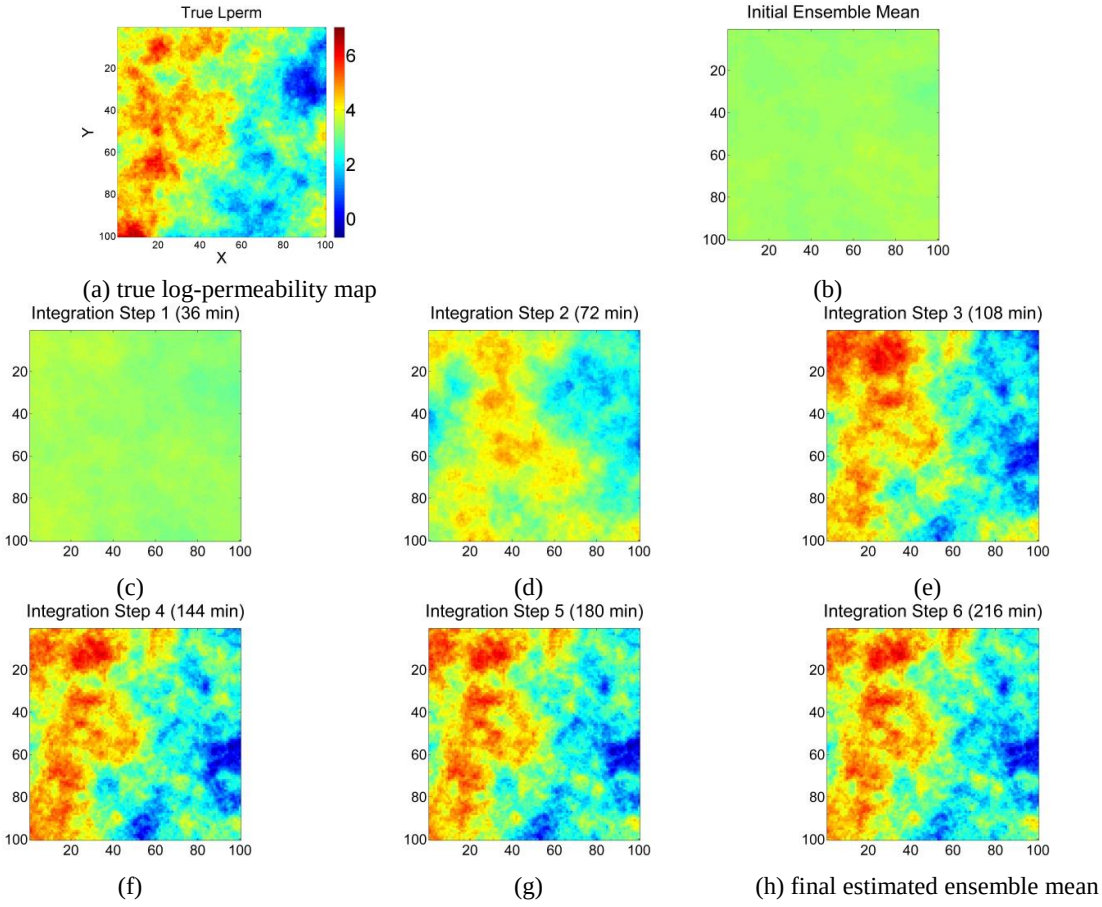
(a) pore-pressure distribution

(b) true log-permeability map

(c) criticality map

(d) MEQ cloud

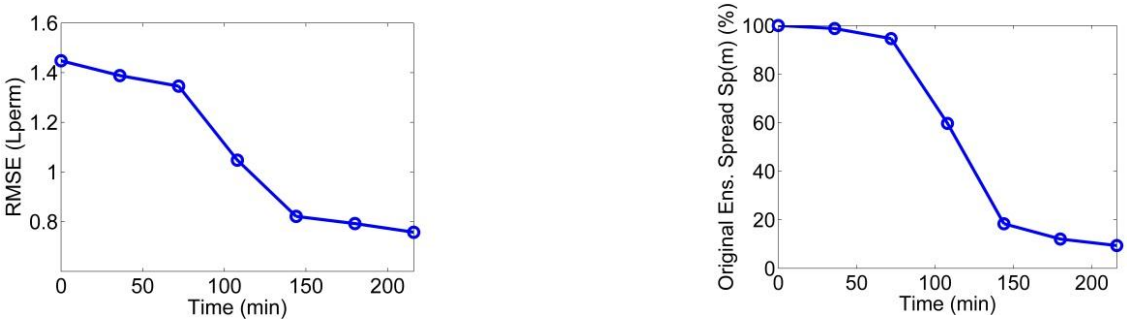
Figure 4.1: Microseismicity cloud generation in a two-dimensional model with heterogeneous permeability: (a) snapshots of diffusive pore pressure distributions at different time steps $p(\mathbf{u}, t)$, (b) the true log-permeability distribution, (c) spatially uncorrelated Gaussian (white noise) rock criticality $C(\mathbf{u})$, and (d) the cloud of microseismic events generated by comparing rock criticality with pore pressure distributions at different time steps.



(a) true log-permeability map (b) (c) (d) (e) (f) (g) (h) final estimated ensemble mean

Figure 4.2: Estimating heterogeneous permeability by SSBRC. The evolution of the estimated ensemble mean through integration steps.

the proposed methods of Section 3.2 and the new estimation results will be presented in experiment 3, Section 4.2.3.



(a) estimation RMSE (b) ensemble spread

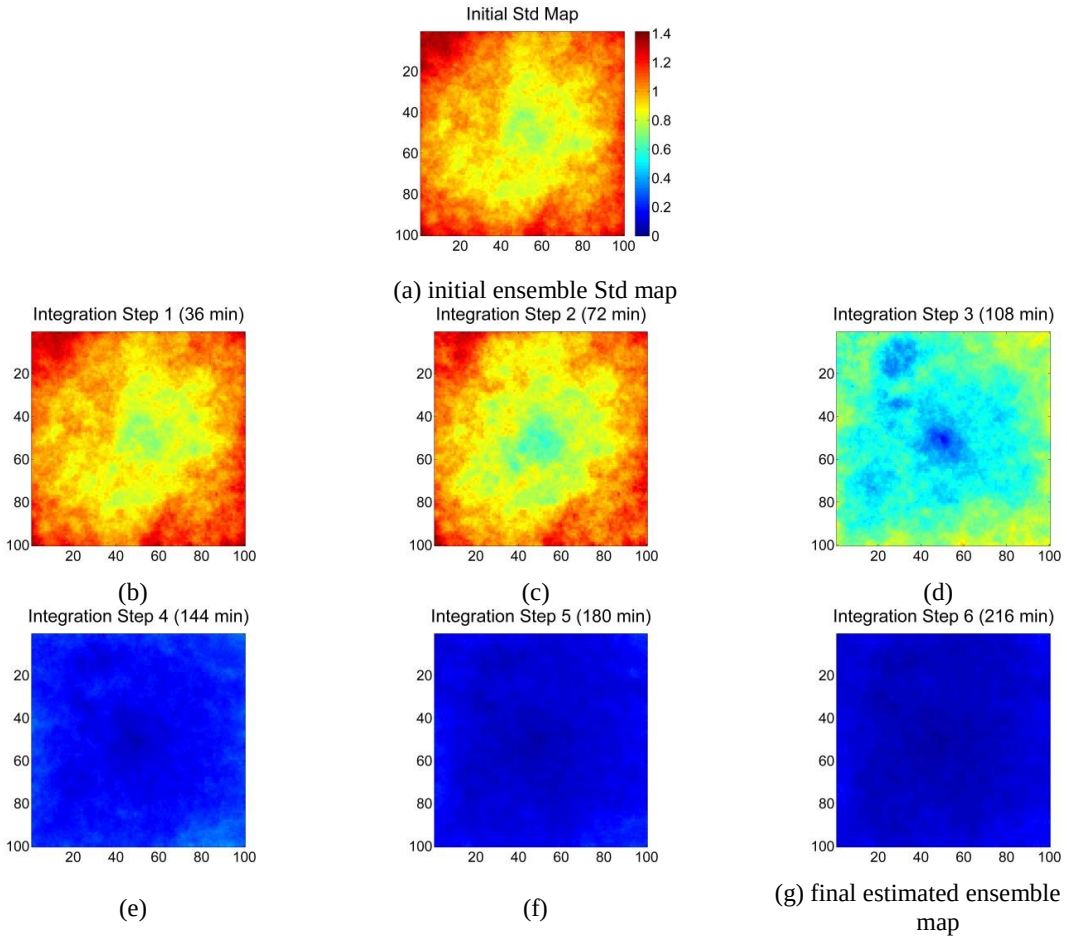
Figure 4.3: SSBRC performance measures in estimating 2D heterogeneous permeability field.

The other representation of ensemble spread is the standard deviation (Std) map of the ensemble which its evolution in integration steps is shown in Figure 4.4. By assimilating the MEQ data, the spread of the permeability ensemble decreases that results in lowering the ensemble Std map. Figure 4.4 shows constant reduction of ensemble spread. As Figure 4.4g represents the final estimated ensemble Std is very low that again demonstrates the ensemble spread underestimation of SSBRC with high dimensional observations (MEQ density map on original fine grid system).

We also performed examples of 2D homogeneous case (not presented here) and SSBRC showed successful performance in estimating homogeneous parameter from MEQ data as well.

4.2.2 Experiment 2: 3D Heterogeneous

In this section we apply SSBRC to a 3D pore pressure diffusion reservoir model with heterogeneous permeability distribution. The 3D model has a $50 \times 50 \times 30$ grid



(a) initial ensemble Std map

(b)

(c)

(d)

(e)

(f)

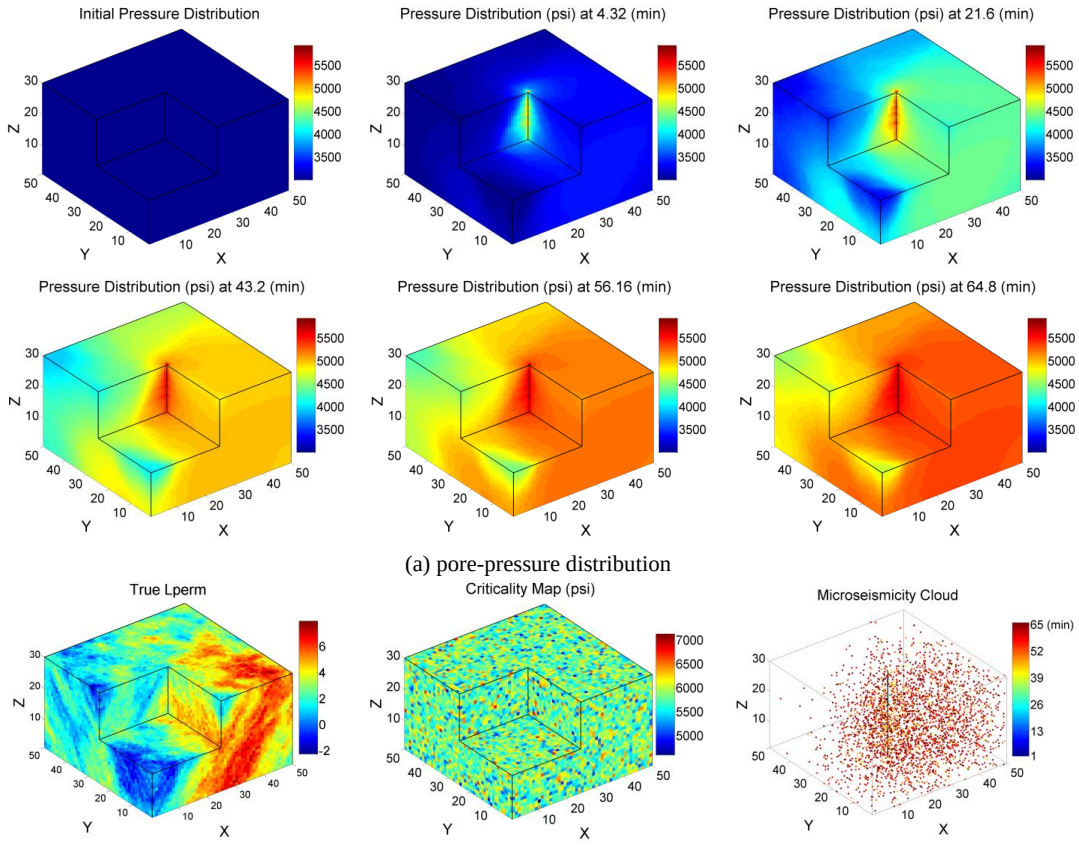
(g) final estimated ensemble Std map

Figure 4.4: Evolution of the standard deviation (Std) map of permeability ensemble through integration steps.

block structure (75000 grid blocks) and the permeability field is a heterogeneous parameter. There is only on water injection well at the center of the field which is perforated throughout the entire thickness of the formation. The true log-permeability distribution is shown in Figure 4.5b. In this experiment the dimension of the parameter to estimate and seismicity density observation are both equal to 75000. An uncorrelated Gaussian random distribution is also assumed for criticality map shown in Figure 4.5c. Other specifications of the true model are illustrated in Figure 4.5 including the pore pressure distribution in the field throughout stimulation period, Figure 4.5a, and the MEQ cloud observation, Figure 4.5d, which is generated by comparing pore pressure distributions and the criticality map. Using SSBRC we assimilate MEQ cloud observation, Figure 4.5d, (after quantifying by KDE method) by EnKF update equation to infer 3D heterogeneous log-permeability field, Figure 4.5b. The stimulation period is 65 (min) and we specify 6 integration steps to assimilate MEQ observations. The MEQ density map is constructed on the original fine grid system and the observation error Std range is the typical range of 5% to 10%.

In the following figures in this section the extensive estimation results of SSBRC are presented. Figure 4.6 shows the result of updating the permeability ensemble in terms of ensemble mean through 6 integration steps. It is seen that the estimated ensemble mean is becoming more similar to the true permeability map in Figure 4.6a by assimilating MEQ observations in time. The initial ensemble (almost uniform initial ensemble mean in Figure 4.6b) is completely uninformative about true permeability map, Figure 4.6a, however after MEQ data assimilation the final ensemble mean, Figure 4.6h, is very similar to the true map. Therefore SSBRC proves its suitability for 3D heterogeneous permeability estimation from MEQ data.

The performance measures of SSBRC estimation procedure are also shown in



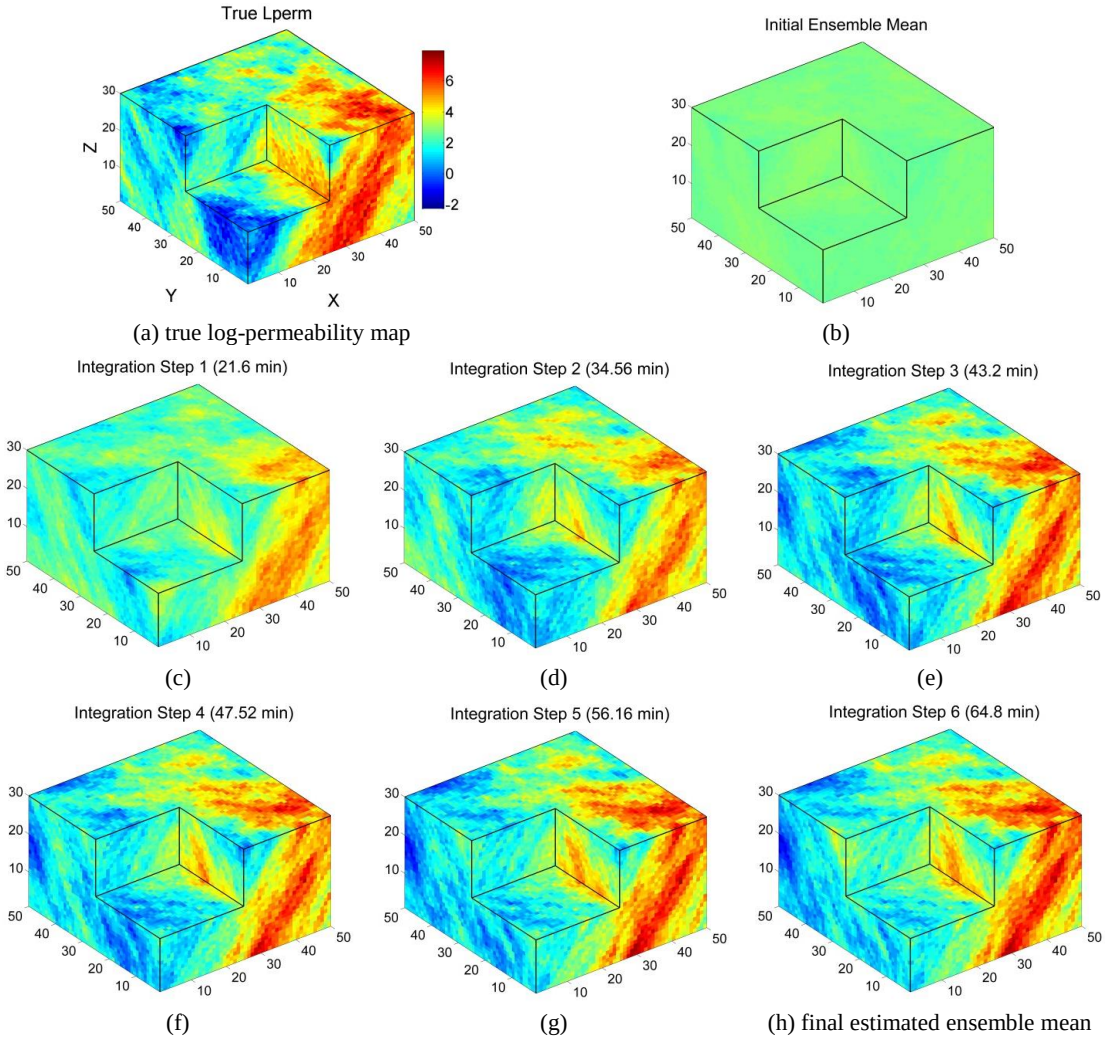
(a) pore-pressure distribution

(b) true log-permeability map

(c) criticality map

(d) MEQ cloud

Figure 4.5: Microseismicity cloud generation in a three-dimensional model with heterogeneous permeability: (a) snapshots of diffusive pore pressure distributions at different time steps, (b) the true log-permeability distribution, (c) spatially uncorrelated Gaussian (white noise) rock criticality, and (d) the cloud of microseismic events generated by comparing rock criticality with pore pressure distributions at different time steps.



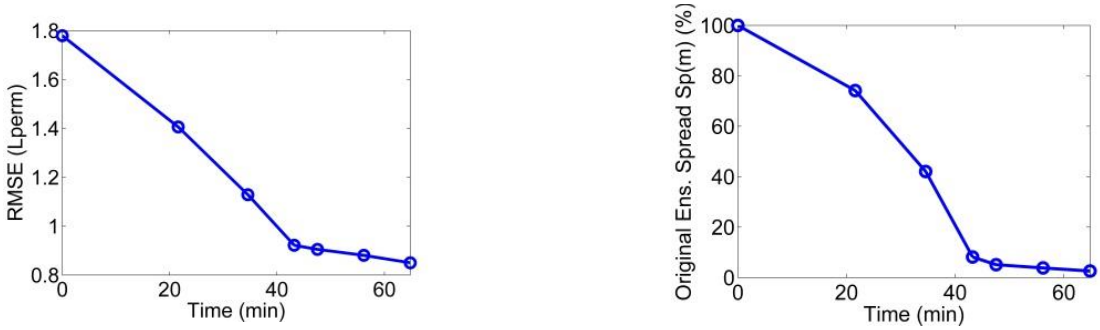
(a) true log-permeability map (b)

(c) (d) (e)

(f) (g) (h) final estimated ensemble mean

Figure 4.6: Estimating 3D heterogeneous permeability by SSBRC. The evolution of the estimated ensemble mean through integration steps.

Figure 4.7. The continuous reduction of estimation RMSE, Figure 4.7a, shows the success of SSBRC in estimating 3D permeability distribution. The final RMSE value is less than half of initial RMSE. The ensemble spread is plotted in Figure 4.7b which shows that 97% of the initial ensemble spread is lost throughout the estimation procedure. This very low final ensemble spread (3%) shows severe ensemble spread underestimation and ensemble collapse which is due to erroneous and spurious correlation of very high resolution seismicity density observation and permeability distribution. As discussed before, very high dimensional seismicity density observation (75000 seismicity observations) which is made on the original fine grid system introduces high amount of redundancy and spurious correlation to EnKF update and consequently results in ensemble collapse. This uncertainty quantification issue will be resolved by applying the proposed methods of Section 3.2 and the new estimation results will be presented in experiment 4, Section 4.2.4.

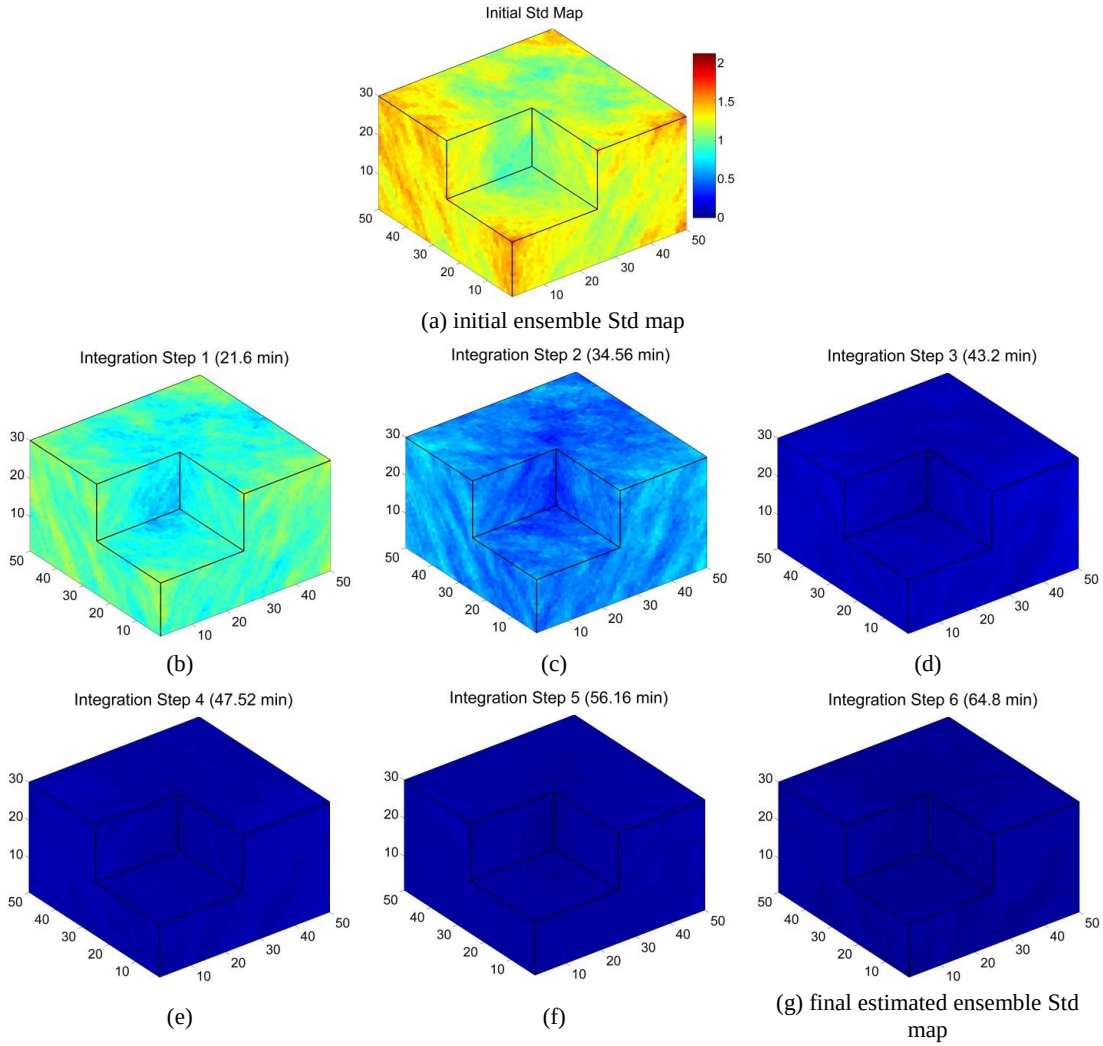


(a) estimation RMSE (b) ensemble spread

Figure 4.7: SSBRC performance measures in estimating 3D heterogeneous permeability field.

The other representation of ensemble spread is the standard deviation (Std) map of the ensemble which its evolution in integration steps is shown in Figure 4.8. By

assimilating the MEQ data, the spread of the permeability ensemble decreases that results in lowering the ensemble Std map. Figure 4.8 shows continuous reduction of ensemble spread. As Figure 4.8g shows the final estimated ensemble Std is extremely low that again demonstrates the ensemble spread underestimation of SSBRC with high dimensional observations (MEQ density map on original fine grid system).



(a) initial ensemble Std map

(b)

(c)

(d)

(e)

(f)

(g) final estimated ensemble Std map

Figure 4.8: Evolution of the standard deviation (Std) map of permeability ensemble through integration steps.

4.2.3 Experiment 3: Resolving Spread Underestimation

In this set of experiments, we apply the three methods of improving uncertainty quantification in Section 3.2 to the experiment 1 in Section 4.2.1 (reference experiment) to resolve the issue of ensemble spread underestimation.

4.2.3.1 Inflated Observation Error Variance

In this section to resolve the ensemble spread underestimation issue of SSBRC in experiment 1 Section 4.2.1, we apply inflated observation error variance method. To investigate the effect of observation Std range, SSBRC results with different Std intervals are shown in Table 4.1. The first row of Table 4.1, Test # 1, shows the results of reference model that is experiment 1 in Section 4.2.1. These results show improvement of final ensemble spread, characterized by $Sp(\mathbf{m})$, (preserving more ensemble spread) by increasing the observation error Std. It can be seen in Table 4.1 column $Sp(\mathbf{m})$ that with increasing observation Std we can improve final ensemble spread from 10 % to 50 %. The Test # 4 from Table 4.1 (with 100 to 200 % observation Std range) is chosen as a representative experiment and its estimation results are shown in Figure 4.9.

Test #	σ_{min} (%)	σ_{max} (%)	Final RMSE (Lperm)	Final Spread (%) $Sp(\mathbf{m})$
1 (Reference)	11	16	0.7576	9.39
2	34	90	0.7067	21.31
3	82	144	0.7078	32.03
4	100	203	0.7339	37.90
5	111	187	0.7394	41.09
6	143	227	0.7305	41.71
7	255	385	0.8075	51.42

Initial RMSE = 1.4478

Table 4.1: Sensitivity of the SSBRC performance to different ranges of observation Std (standard SSBRC with seismicity density on the original fine grid) in estimating 2D heterogeneous permeability.

In Figure 4.9 for concise illustration, only initial and final (at the sixth integration step) estimated maps are shown. Figure 4.9b shows initial permeability ensemble mean, ensemble Std map and an individual permeability sample and Figure 4.9c represents the final estimated ensemble mean, Std map and an individual sample after assimilating all MEQ observations. SSBRC is very successful in inferring the true permeability since the final estimated maps are very similar to the true permeability distribution. Additionally SSBRC along with inflated observation error Std results in preserving ensemble spread through estimation procedure which is shown by high Std map of Figure 4.9c (middle plot). Estimation RMSE, Figure 4.9d, and ensemble spread, Figure 4.9e, prove successful estimation and ensemble spread improvement, respectively.

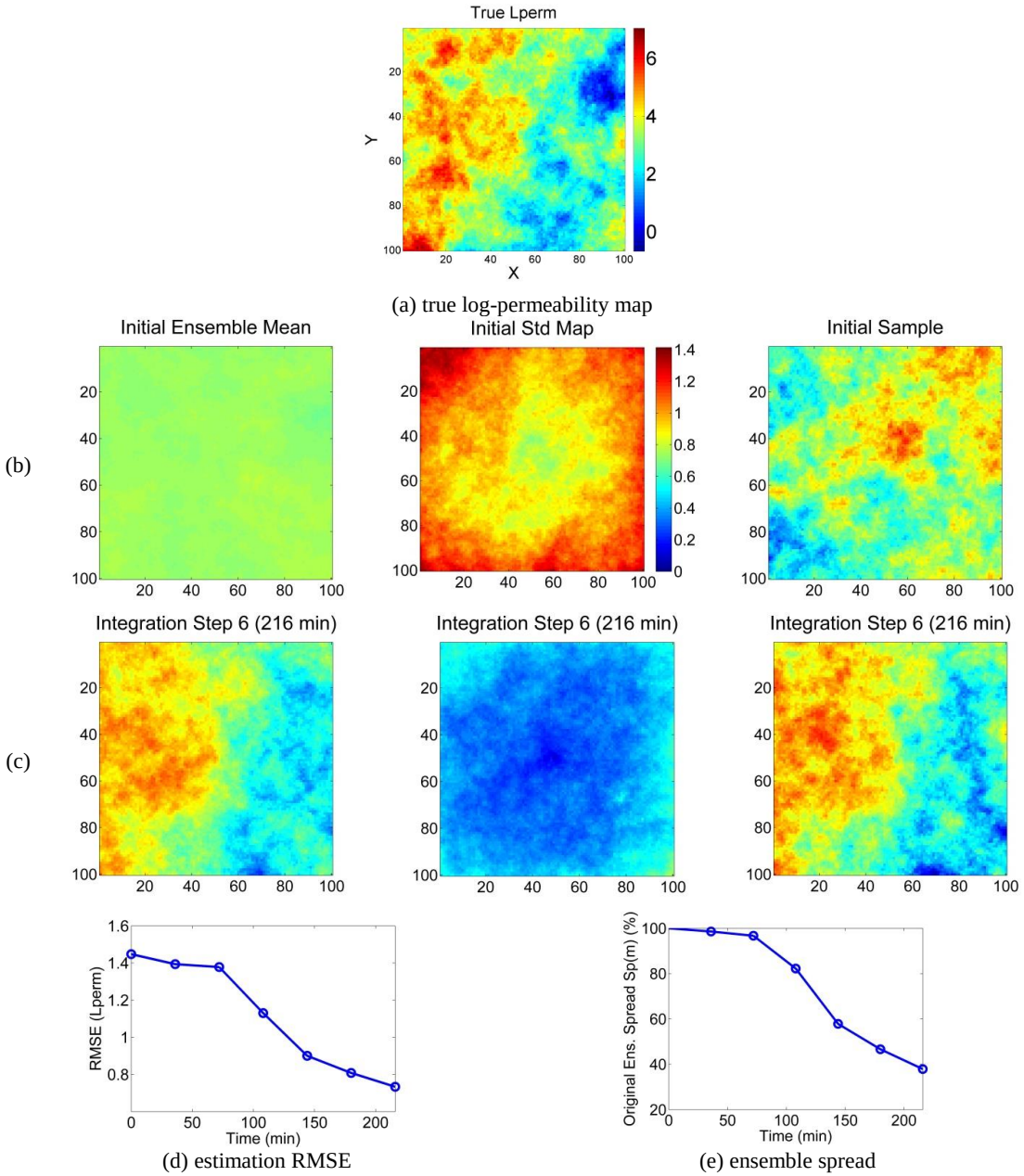


Figure 4.9: The SSBRC estimation results with inflated observation Std for a 2D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread.

4.2.3.2 Reduced-Order Projection

The results of applying reduced-order projection along with SSBRC for different values of kernel bandwidth Std, σ_h, and truncation number, n_{trunc}, to experiment 1, Section 4.2.1, are presented in Table 4.2. The reference experiment (experiment 1, Section 4.2.1) suffers from severe ensemble spread underestimation with only 10 % final spread however Table 4.2 shows promising improvement of ensemble spread towards 40 to 80 %. The estimation RMSE of projection approach as shown in Table 4.2 is not as low as reference experiment. It is clear from Table 4.2 column n_{trunc} that reduced-order projection approach lowered the number of observations from 10000, in the reference experiment, to 25, 50 and 100, by spectral dimension reduction. We choose Test # 4 from Table 4.2 as the representative experiment and its estimation results are shown in Figure 4.10. It should be noted that for this experiment five integration steps are considered. As it is seen in Table 4.2 Test # 4, the truncation number, n_{trunc}, is 100 which means by applying reduced-order projection we reduced the number of observations from 10000, in the reference experiment, to 100 in the improved experiment.

Test #	σ_h (%)	n_{trunc}	Final RMSE (Lperm)	Final Spread (%) $Sp(\mathbf{m})$
1	25	100	1.02	50.61
2	25	50	0.9975	68.62
3	25	50	0.9682	64.84
4	25	100	0.9546	41.89
5	10	100	1.1382	28.41
6	25	25	1.1427	76.31

Initial RMSE = 1.4478

Table 4.2: SSBRC with observation projection approach. Sensitivity of the performance of projection approach with respect to kernel bandwidth Std (σ_h) and truncation number (n_{trunc}).

The final estimated ensemble mean and individual sample (left and right plots of Figure 4.10c) demonstrate significant similarity with the true permeability map, Figure 4.10a, and the estimation $RMSE$ curve, Figure 4.10d, is continuously decreasing. Therefore the estimation performance of SSBRC with projection approach is promising. As Figure 4.10e shows the final ensemble spread is 40 %. The final Std map, Figure 4.10c (middle plot), also demonstrates high values which confirms successful application of reduced-order projection for resolving ensemble spread underestimation.

4.2.3.3 Coarse-Scale Microseismicity Density Map

In this section preserving ensemble spread and improving uncertainty quantification is performed through reducing the number of observations by interpreting the discrete MEQ cloud on a coarse-scale grid system instead of

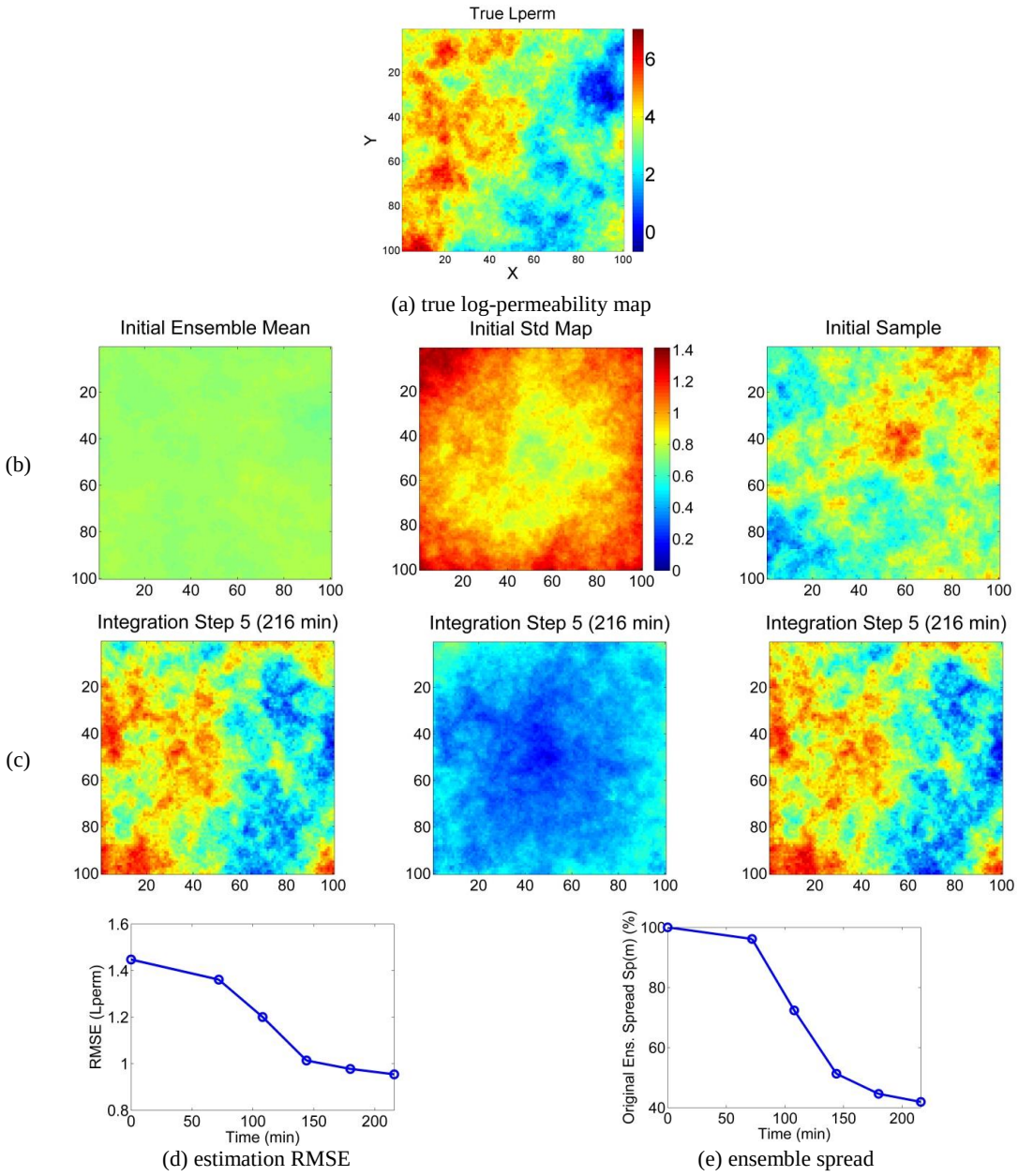


Figure 4.10: The SSBRC estimation results with observation projection approach for a 2D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread.

original fine grid configuration. The reference experiment with ensemble spread underestimation in the one in section 4.2.1. The dimension of microseismicity density map is considered in terms of number of grid blocks in x direction (the field configuration is square). For coarsening of the seismicity density map, we specify different sizes of $N_{X,red}$ equal to 10, 15, 20, 30, 50, 70 and 90, that decreases the number of observations to 100, 225, 400, 900, 2500, 4900 and 8100, respectively where in the reference experiment by $N_{X,red} = N_X = 100$, there are 10000 observations. In Figure 4.11, coarse-scale microseismicity density maps generated on different coarse grid configurations are shown. It is clear from Figure 4.11 that by coarsening the seismicity density map we are able to reduce the redundancy of MEQ density and consequently reduce the number of observations. In this experiment where permeability is heterogeneous the parameter dimension is also 10000.

The results of sensitivity analysis of SSBRC performance with respect to coarse seismicity density size are given in Figure 4.12. All these experiments are performed with the typical range of observation error Std (5 % to 10 %). As we can see estimation RMSE is not very sensitive to size of the coarse grid system so the determining factor in choosing the appropriate size is ensemble spread. Figure 4.12 clearly shows that by increasing the dimension of the coarse grid system (rising the number of seismicity density observations) the ensemble spread will decrease. Therefore using coarse-scale grid system for generating seismicity density map greatly helps in preserving ensemble spread and avoiding ensemble collapse. As the representative experiment, we choose coarse grid system of 10×10 ($N_{X,red} = 10$) from Figure 4.12 and its estimation results are shown in Figure 4.13.

Figure 4.13 presents the SSBRC estimation results with coarse-scale seismicity density map. The final estimated maps of permeability (Figure 4.13c) are very similar to the true map (Figure 4.13a) and the estimation RMSE curve is continuously

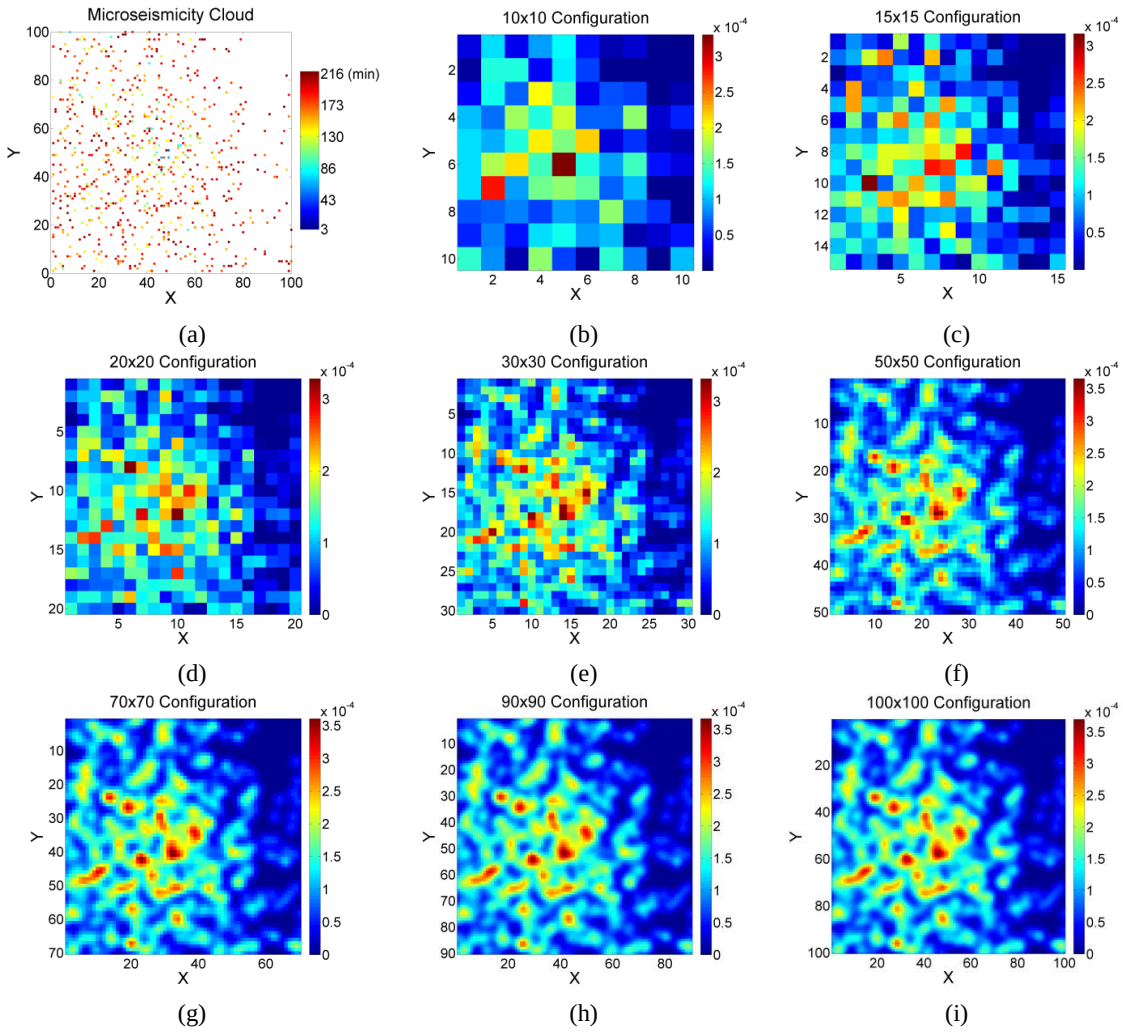


Figure 4.11: Continuous interpretation of MEQ cloud on different sizes of coarse-scale grid configuration. (a) discrete MEQ cloud, (b)-(i) coarse seismicity density maps on different coarse grid systems.

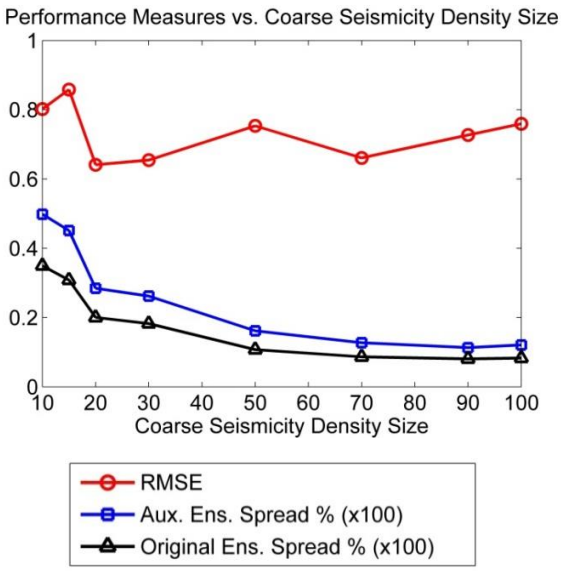


Figure 4.12: Effect of coarse microseismicity density dimension on the performance of SSBRC in heterogeneous permeability estimation experiment (Initial RMSE = 1.4478).

decreasing which confirms promising estimation performance of SSBRC. Utilizing coarse-scale seismicity density resolves the ensemble spread underestimation issue and increases the final ensemble spread from 10 % in the reference experiment to 40 % in this improved experiment (Figure 4.13e). Comparing the final ensemble Std map of Figure 4.13c (middle) to the same map of reference experiment, Figure 4.4g, indicates how effective coarse seismicity density approach is in improving uncertainty quantification of SSBRC.

4.2.4 Experiment 4: Resolving Spread Underestimation

In this set of experiments, we apply the three methods of improving uncertainty quantification in Section 3.2 to the experiment 2 in Section 4.2.2 (reference experiment) to resolve the issue of ensemble spread underestimation and ensemble collapse.

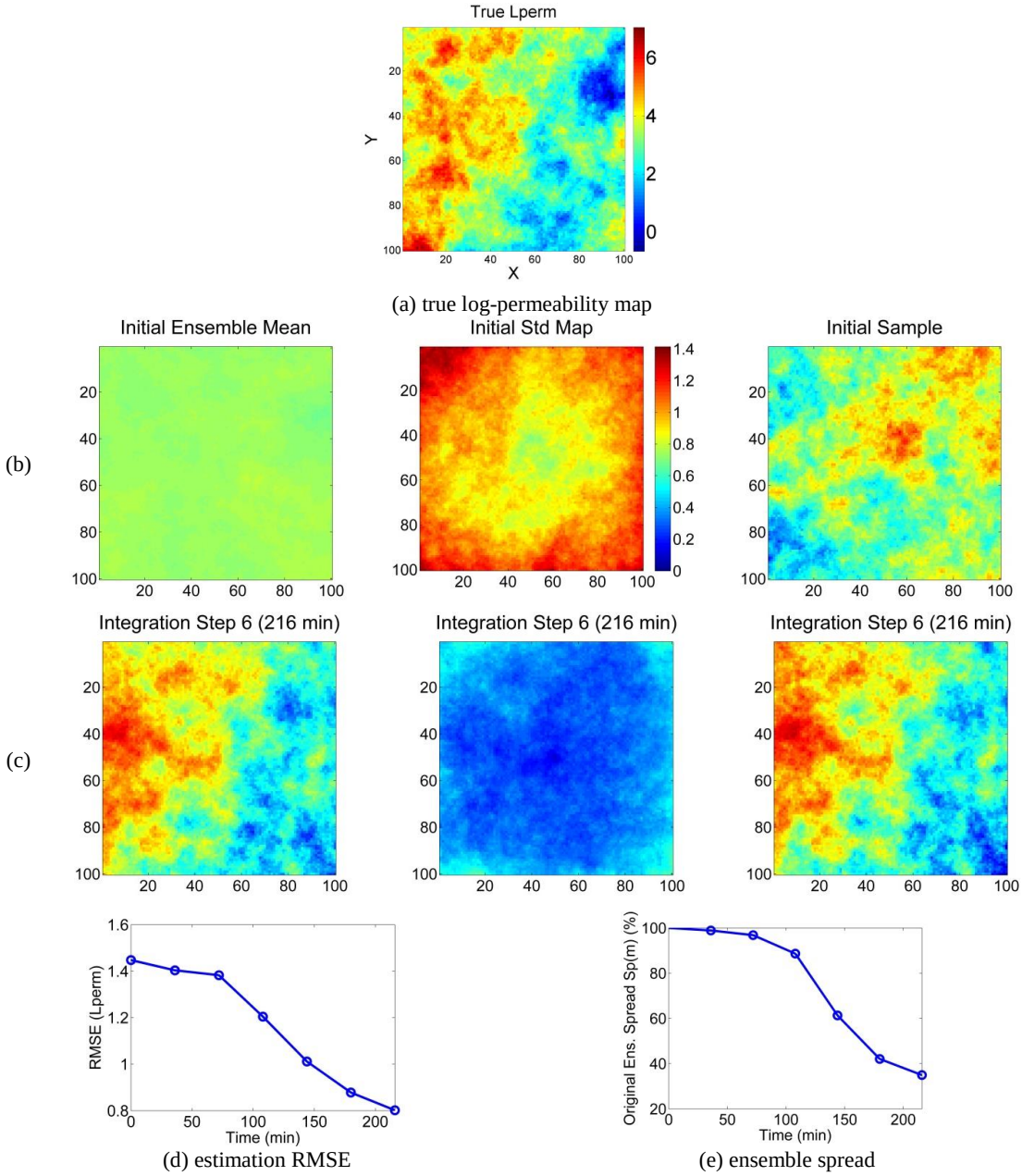


Figure 4.13: The SSBRC estimation results with coarse-scale seismicity density approach for a 2D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread.

4.2.4.1 Inflated Observation Error Variance

To resolve ensemble collapse problem with experiment 2 in Section 4.2.2, in this section, we artificially increase observation error standard deviation range. We investigated the effect of different observation Std range on SSBRC performance. These results show improvement of final ensemble spread (preserving more ensemble spread) by increasing the observation error Std. An experiment with 100 to 200 % observation Std range is chosen as a representative experiment and its estimation results are shown in Figure 4.14. We see that final ensemble spread increased from 3 % to almost 20 %.

In Figure 4.14 for concise illustration, only initial and final (at the sixth integration step) estimated maps are shown. Figure 4.14b shows initial permeability ensemble mean, ensemble Std map and an individual permeability sample and Figure 4.14c represents the final estimated ensemble mean, Std map and an individual sample after assimilating all MEQ observations. SSBRC is very successful in inferring the true permeability since the final estimated maps are very similar to the true permeability distribution. Additionally SSBRC along with inflated observation error Std results in preserving ensemble spread and avoiding ensemble collapse through estimation procedure which is shown by high Std map of Figure 4.14c (middle plot).

4.2.4.2 Reduced-Order Projection

The results of applying reduced-order projection along with SSBRC for kernel bandwidth Std, σ_h, of 25% and truncation number, n_{trunc}, of 100 to experiment 2, Section 4.2.2, are presented here.Its estimation results are shown in Figure 4.15. The truncation number, n_{trunc}, is 100 which means by applying reduced-order projection we reduced the number of observations from 75000, in the reference experiment, to 100 in the improved experiment.

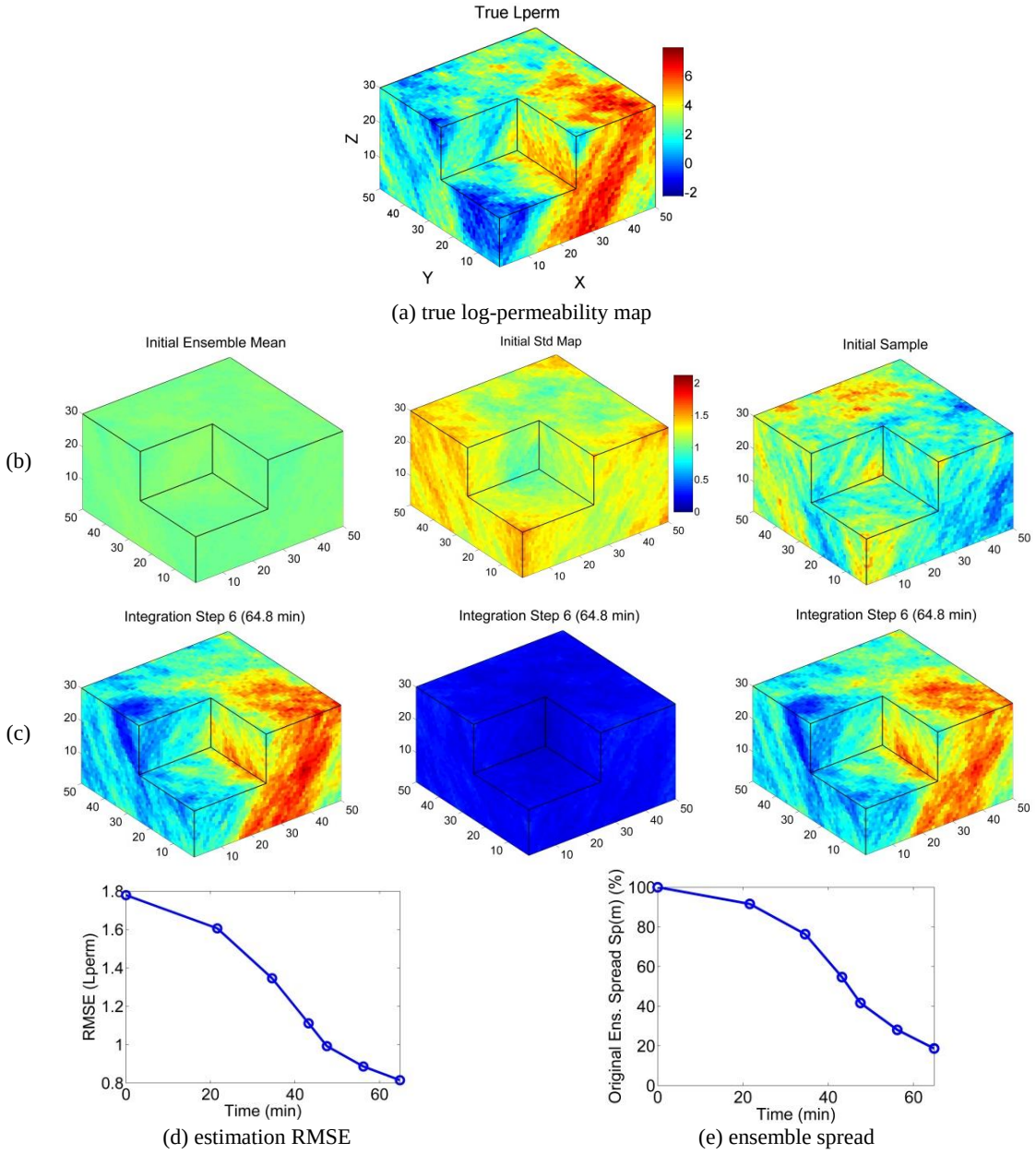


Figure 4.14: The SSBRC estimation results with inflated observation Std for a 3D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread.

The final estimated ensemble mean and individual sample (left and right plots of Figure 4.15c) demonstrate significant similarity with the true permeability map, Figure 4.15a, and the estimation $RMSE$ curve, Figure 4.15d, is continuously decreasing. Therefore the estimation performance of SSBRC with projection approach is promising. As Figure 4.15e shows the final ensemble spread is 20 %. The final Std map, Figure 4.15c (middle plot), also demonstrates high values which confirms successful application of reduced-order projection.

4.2.4.3 Coarse-Scale Microseismicity Density Map

In this section preserving ensemble spread and improving uncertainty quantification is performed through reducing the number of observations by interpreting the discrete MEQ cloud on a coarse-scale grid system instead of original fine grid configuration. The reference experiment with ensemble spread underestimation in the one in Section 4.2.2. These SSBRC experiments with coarse seismicity density are performed with the typical range of observation error Std (5 % to 10 %). The coarse grid system size (obtained from investigating different coarse grid sizes) is $10 \times 10 \times 5$ which results in the coarse seismicity density dimension of 500. In fact we discretize the original 3D field to $10 \times 10 \times 5$ configuration to generate seismicity density instead of using the original fine grid system ($50 \times 50 \times 30$) which is used for numerical reservoir simulation. Therefore the number of observations is reduced from 75000 in the reference experiment to 500 in the improved experiment. Final ensemble spread value clearly shows that by coarsening the grid system the ensemble spread will increase. Therefore using coarse-scale grid system for generating seismicity density map greatly helps in preserving ensemble spread and avoiding ensemble collapse. the coarse grid system experiment results are shown in Figure 4.16.

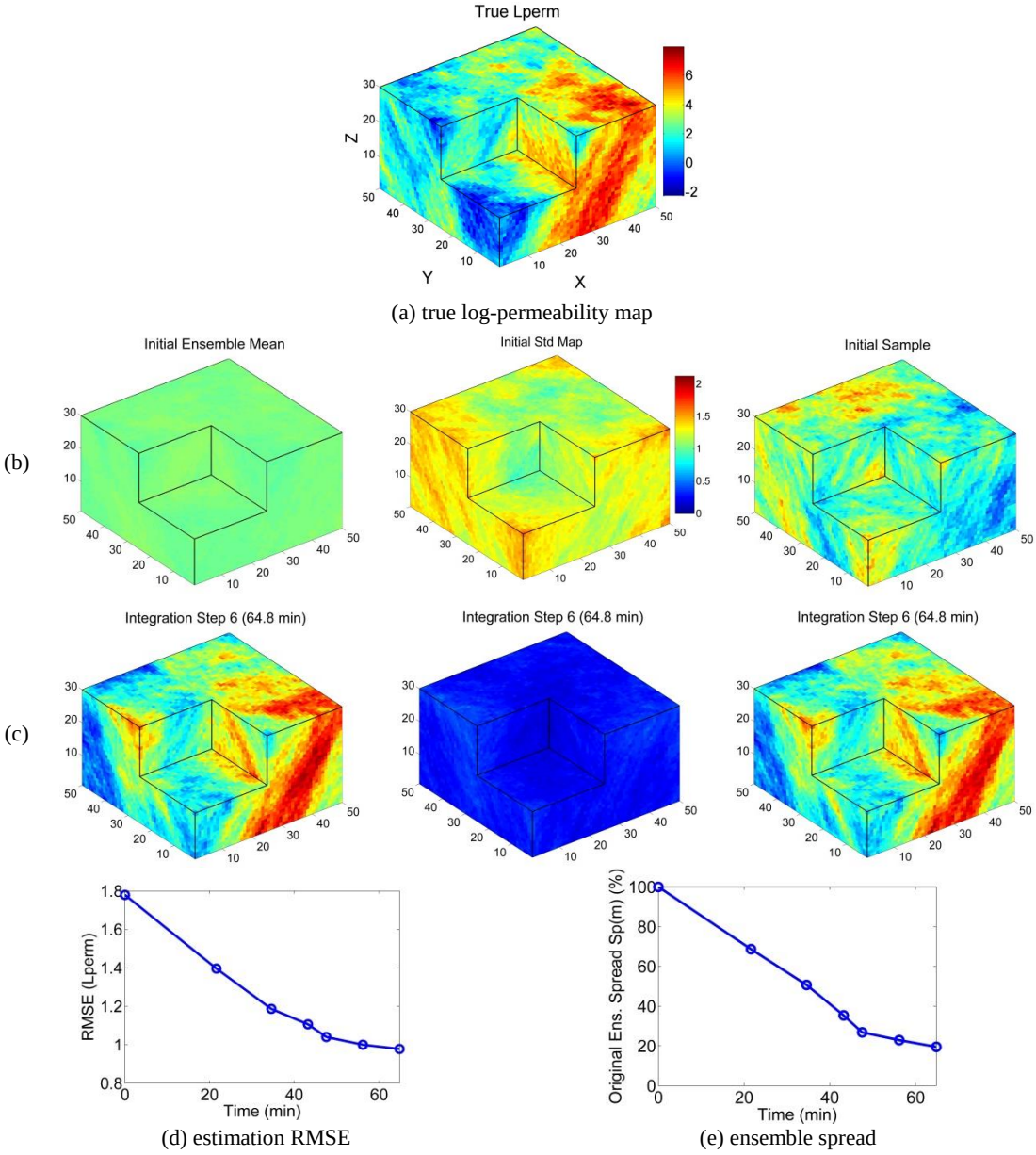


Figure 4.15: The SSBRC estimation results with observation projection approach for a 3D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread.

Figure 4.16 presents the SSBRC estimation results with coarse-scale seismicity density map. The final estimated maps of permeability (Figure 4.16c) are very similar to the true map (Figure 4.16a) and the estimation RMSE curve is continuously decreasing which confirms promising estimation performance of SSBRC. Utilizing coarse-scale seismicity density resolves the ensemble spread underestimation (ensemble collapse) issue and increases the final spread from 3 % in the reference experiment to 25 % in this improved experiment (Figure 4.16e). Comparing the final ensemble Std map of Figure 4.16c (middle) to the same map of reference experiment, Figure 4.8g (where ensemble collapse happened), indicates how effective coarse seismicity density approach is in improving uncertainty quantification of SSBRC.

4.2.5 Joint Parameter Estimation in Pore Pressure Diffusion Model

In the previous examples following *Shapiro's* work, criticality is assumed known. This assumption is also relaxed here. We investigated jointly estimating permeability and criticality in the pore pressure diffusion model to relax the assumption of having a spatially random criticality distribution. In this experiment, pore pressure diffusion model is the forward model. Permeability and criticality are assumed to have the same variogram parameters and are uncorrelated. Both permeability and criticality distribution are assumed unknown and we infer them both from MEQ observation. The results of joint estimation are presented in the following Table 4.3 and Figure 4.17. For this example to preserve ensemble spread projection and coarse scale methods are applied along SSBRC. As reported in Table 4.3 and shown in Figure 4.17, SSBRC is successful in inferring both permeability and criticality distributions.

4.3 Description of Experimental Setup: Geomechanical Model

In this section we apply the SSBRC method to the fully coupled geomechanical forward reservoir simulator. In these set of experiments we assimilate MEQ cloud

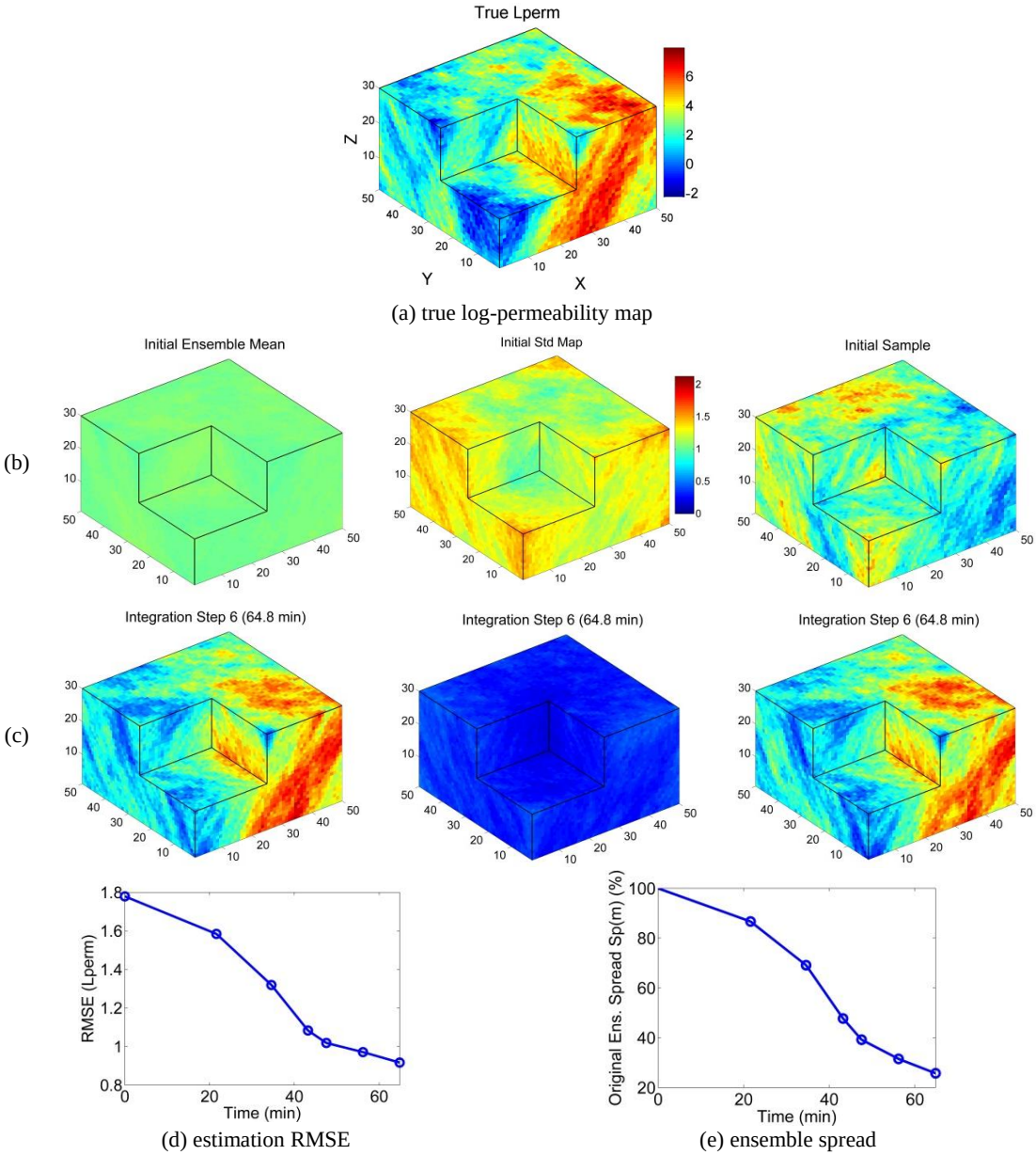


Figure 4.16: The SSBRC estimation results with coarse seismicity density for a 3D heterogeneous permeability model: (a) the true log permeability model, (b) initial log-permeability ensemble mean (left), standard deviation map (middle), and an individual realization (right), (c) final log permeability ensemble mean (left), standard deviation (middle), and individual realization (right) after six update steps, and time evolution of (d) the log permeability RMSE and (e) normalized ensemble spread.

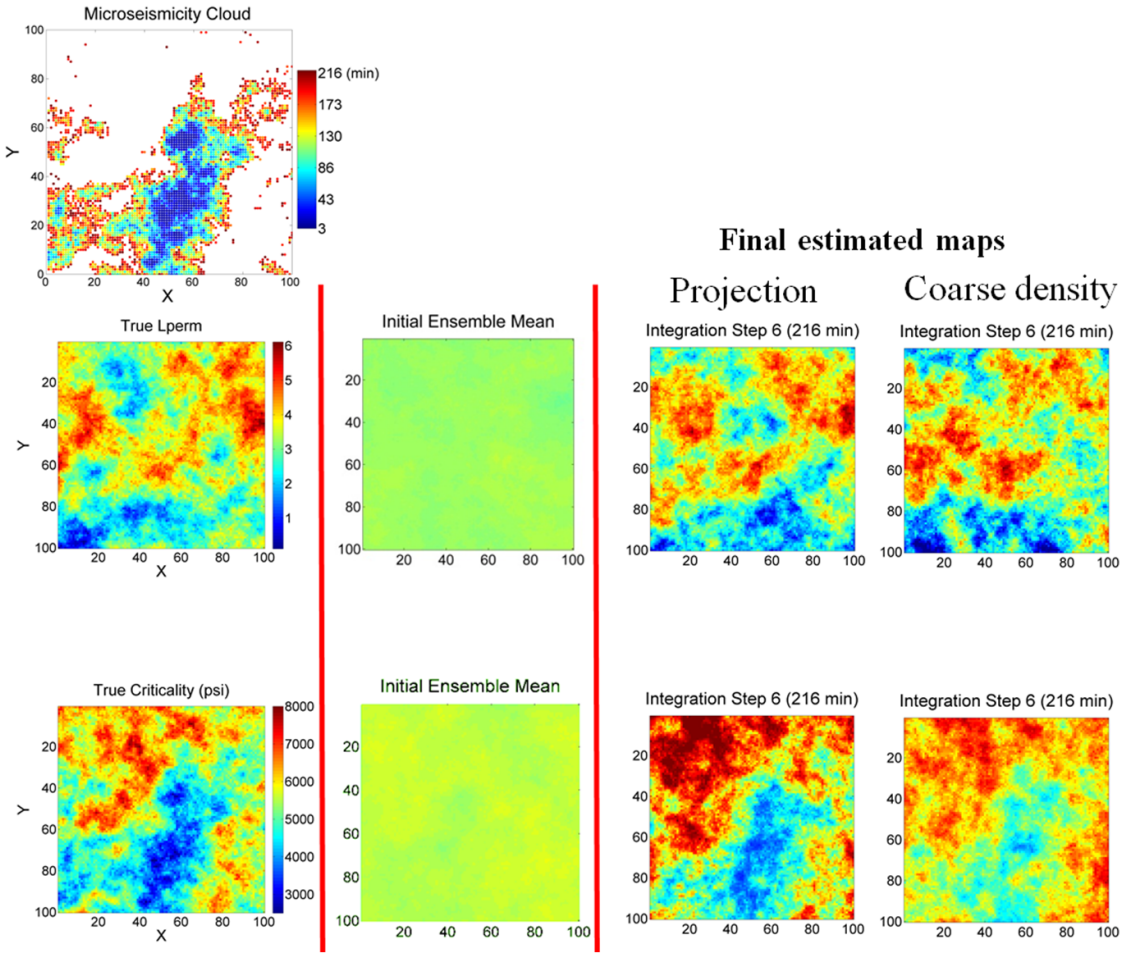


Figure 4.17: The observed MEQ data, true permeability and criticality maps, initial ensemble means, and final estimated maps (middle row is permeability and last row is criticality)

Experiment	*Final Lperm RMSE*	*Final Crit. RMSE (psi)*	*Final Lperm Spread %*	*Final Crit. Spread %*
Projection	1.0234	833.37	61.58	25.03
Coarse density	0.8083	732.88	21.88	10.08

Initial Lperm RMSE = 1.3130
Initial Crit. RMSE = 1243.53 (psi)

Table 4.3: Results of jointly estimating permeability and criticality from MEQ data.

(after converting it to continuous seismicity density map) in EnKF analysis equation to infer hydraulic (permeability) and geomechanical (Young's modulus or elastic modulus, tensile strength, Cohesion) parameters of the reservoir. The experiments in this section are performed on both homogeneous and heterogeneous 2D model and also on heterogeneous 3D model. We use a 2D FEM model with point source injection by quadrilateral regular mesh with $50 \times 50 = 2500$ elements (10000 Gaussian points) and the reservoir size of $500m \times 500m$. The developed coupled FEM model is capable of handling spatially distributed parameters (heterogeneous distributions) such as permeability, Young's modulus, tensile strength, cohesion and friction angle. We assign the spatially distributed parameters to the nodes therefore 2D model parameter dimension is 10000. There is one water injection well with constant injection rate at the center of the field. The fluid is assumed single phase and the boundaries are closed to flow (No flow boundary condition). In 2D model, the stress boundary condition or far-field stress regime is specified by maximum horizontal stress $S_{H,max}$, minimum horizontal stress $S_{h,min}$, and the initial field pressure P_{ini}. It should be noted that our simulations are performed in isothermal reservoir conditions however the forward model is capable of handling thermal effects. The rock and fluid properties are reported in Table 4.4. In the 2D forward simulation, time step length is 720 (s) and there are 100 time steps. We also consider this 2D geomechanical

model with homogeneous parameter and perform estimation experiment with SSBRC to infer a single parameter (e.g. the homogeneous Young's modulus) from MEQ observations.

Fluid density, ρ_f	1000 $\frac{kg}{m^3}$
Fluid viscosity, η	10^{-3} Pa.s
Drained Poisson's ratio, ν	0.22
Undrained Poisson's ration, ν_u	0.46
Porosity, ϕ	0.30
Material constant, ζ_d	20
Material constant, β_d	10^{-7}

Table 4.4: Rock/fluid properties used in simulations.

In this work, we also use a 3D fully coupled FEM model with point source injection by hexahedron regular mesh with $30 \times 30 \times 15 = 13500$ elements (108000 Gaussian points) and the reservoir size of $750m \times 750m \times 370m$.

To do EnKF data integration, we consider six integration steps. In each SSBRC estimation experiment, we assume the spatial distribution of one reservoir property unknown (the parameter to estimate) and the rest of the properties are assumed known. We can also estimate more than one reservoir property distribution simultaneously from MEQ cloud.

4.4 Results and Discussion

In this section we present the SSBRC estimation results with the geomechanical forward model. For the 2D model we first show the results of standard SSBRC

that leads to ensemble spread underestimation and then the results of improved uncertainty quantification with SSBRC are presented. The estimation results of SSBRC with 3D geomechanical model are presented only by incorporating improved uncertainty quantification methods.

4.4.1 Experiment 1: 2D Homogeneous

To demonstrate the applicability of the SSBRC method for geomechanical model, we first apply it to a homogeneous parameter estimation problem. As the simplest experiment, we consider all parameters to be constant (homogeneous or uniform over the whole field, i.e., spatially invariable) and then we consider one of the parameters to be unknown and set out to estimate the unknown parameter (which is a scalar). To show the ensemble we simply use a histogram. In this set of experiments we use the standard SSBRC with high resolution seismicity density which is generated on the original fine FEM mesh. The observation error Std range (σ_{min}, σ_{max}) is also chosen as 15 % to 50 % which is inflated to some extent comparing to typical range of observation Std (5 % to 10 %).

We consider integrating tensile microseismicity events to estimate the homogeneous reservoir parameters. We consider Young's modulus (E), permeability (k) and tensile strength (T_0) to be homogeneous. A single true geomechanical model is considered while in different estimation experiments different parameters are estimated. We consider three different settings for estimating three different parameters (Young's modulus, permeability and tensile strength). For instance in the first setting we assume k and T_0 known and E unknown then we integrate seismicity data to estimate homogeneous Young's modulus E. By the same procedure we setup the other two settings to estimate k and T_0. Figure 4.18 shows the true microseismicity cloud for the homogeneous parameter estimation experiment. The

geomechanical model specifications are also indicated in Figure 4.18. The result resembles a conventional hydraulic fracture propagation.

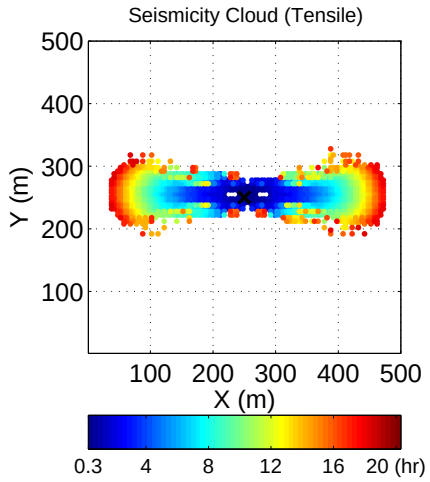
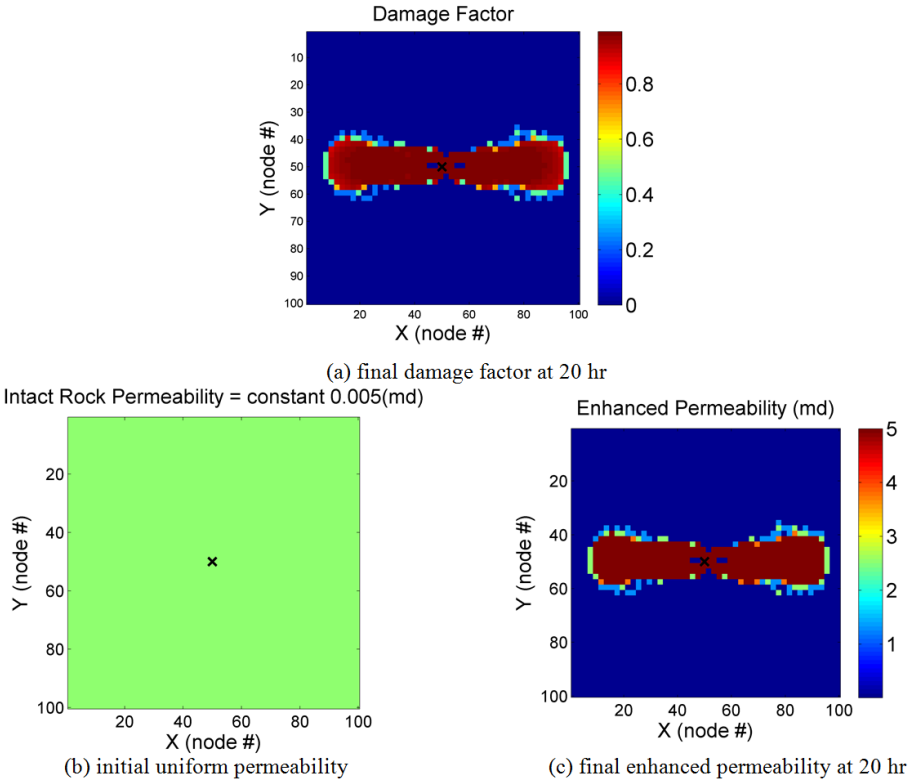


Figure 4.18: True seismicity cloud (tensile failures) of homogeneous parameter experiment; $P_{ini} = 10$ (MPa), $S_{H,max} = 25$ (MPa), $S_{h,min} = 15$ (MPa), Inj. Rate = 12.5 (Lit/s), $E = 10$ (GPa), $k = 0.005$ (md), $T_0 = -4$ (MPa). The black cross shows the injection well location.

Damage factor and enhanced permeability distributions of this example are used to demonstrate the induced fracture and its propagation in Figure 4.19. The enhanced permeability region can be viewed as a highly interconnected fracture network region and the stimulated reservoir volume (SRV).

First we consider estimating the homogeneous Young's modulus (E). To show the estimation procedure we plot the histogram of the ensemble at each integration step. Moreover, as a measure of estimation procedure performance we show the evolution of Root Mean Square Error ($RMSE$) of estimation through time (integration time steps). Figure 4.20 shows the estimation results of homogeneous E by integration of tensile events of Figure 4.18. The red vertical line in the histogram shows the true



(a) final damage factor at 20 hr

(b) initial uniform permeability

(c) final enhanced permeability at 20 hr

Figure 4.19: Damage factor and enhanced permeability distribution to show the induced fracture and its propagation.

value of E (10 GPa) which we are trying to estimate. In all experiments we use 100 realizations or samples of E i.e., the EnKF ensemble size is equal 100. Figure 4.20 shows the evolution of the ensemble histogram from (a) to (g) through integration time steps. It can be seen that the EnKF process is very effective in estimating the unknown parameter. The initial histogram (a) is very wide but by integration of seismicity data the histogram narrows around the true E value and from integration step 3 to the end all the realizations or samples of E almost coincide with the true value of E. In plots (d) to (e) only the red vertical line is visible meaning that all the samples are equal true E value which shows the almost perfect estimation of E. The estimation error of E ($RMSE$), Figure 4.20h, goes to almost zero which indicates a perfect estimation of the unknown parameter.

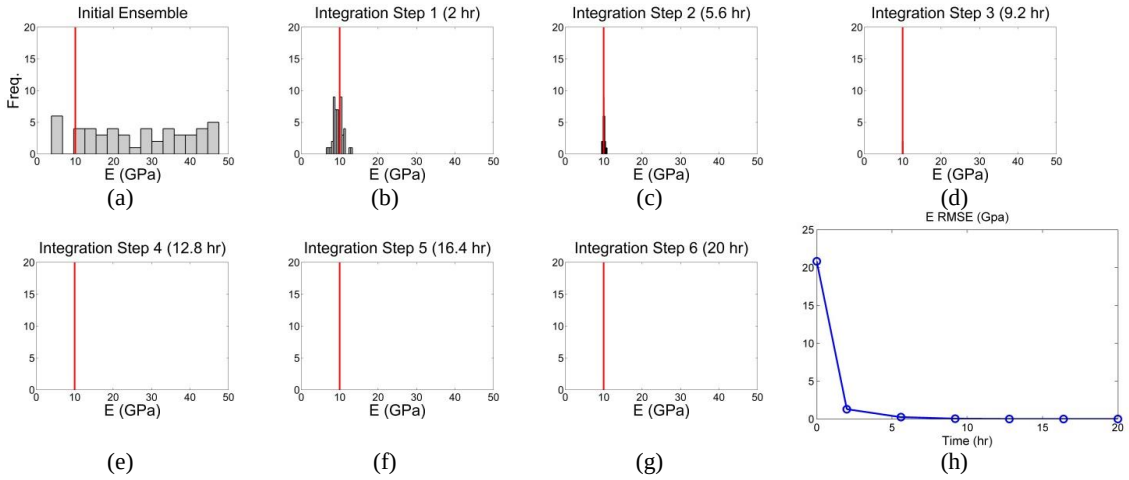


Figure 4.20: Results of homogeneous elastic modulus E, estimation; plots (a) to (g) show the evolution of ensemble histogram at each integration steps. Plot (h) shows the RMSE of estimation (True $E = 10$ GPa).

Figure 4.21 shows the estimation results of homogeneous permeability k, by integration of tensile events of Figure 4.18. The red vertical line in the histogram

shows the true value of k (0.005 md) which we are trying to estimate. Figure 4.21 just represents the histogram of permeability ensemble in some specific integration steps to make the representation more concise. In this estimation example we again see at integration step 3 (Figure 4.21d) the histogram fully narrows down to the true k value which means almost perfect estimation of unknown permeability value. Figure 4.21a, the permeability $RMSE$, also shows the estimation error almost goes to zero.

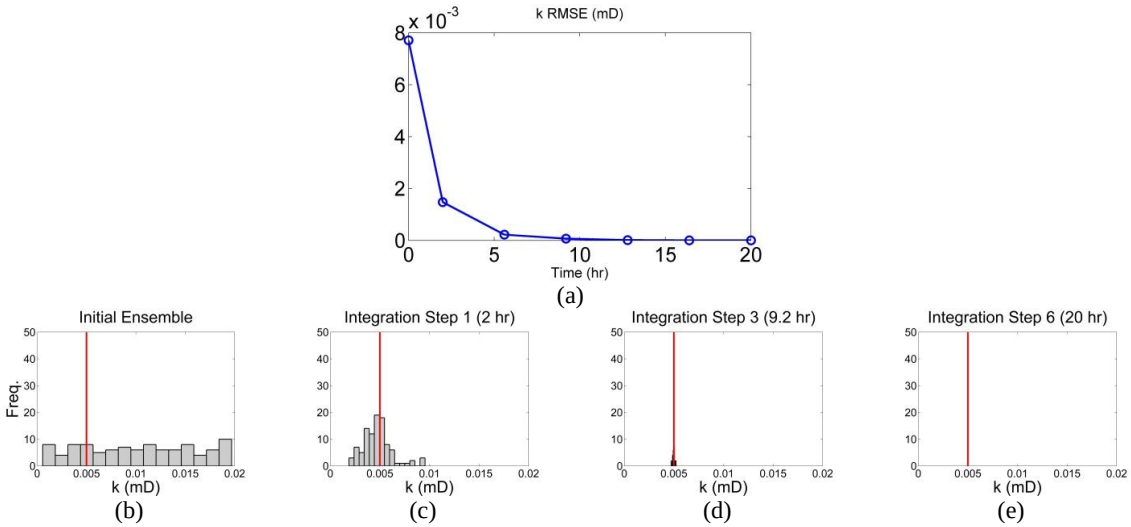


Figure 4.21: Results of homogeneous permeability k, estimation; Plot (a) shows the $RMSE$ of estimation. Plots (b) to (e) show the evolution of ensemble histogram in integration steps (true $k = 0.005$ md).

In the next example setting we assume tensile strength T_0, to be the unknown parameter and estimate it by seismicity data integration. Again the true model is shown in Figure 4.18 and we estimate homogeneous T_0 by integrating tensile events. The true tensile strength value is equal to -4 (MPa). Figure 4.22 illustrates the estimation results including both T_0 estimation error and T_0 ensemble histogram

evolution. Confirming the previous homogeneous parameter estimation results, we can see the unknown homogeneous T_0 is estimated perfectly.

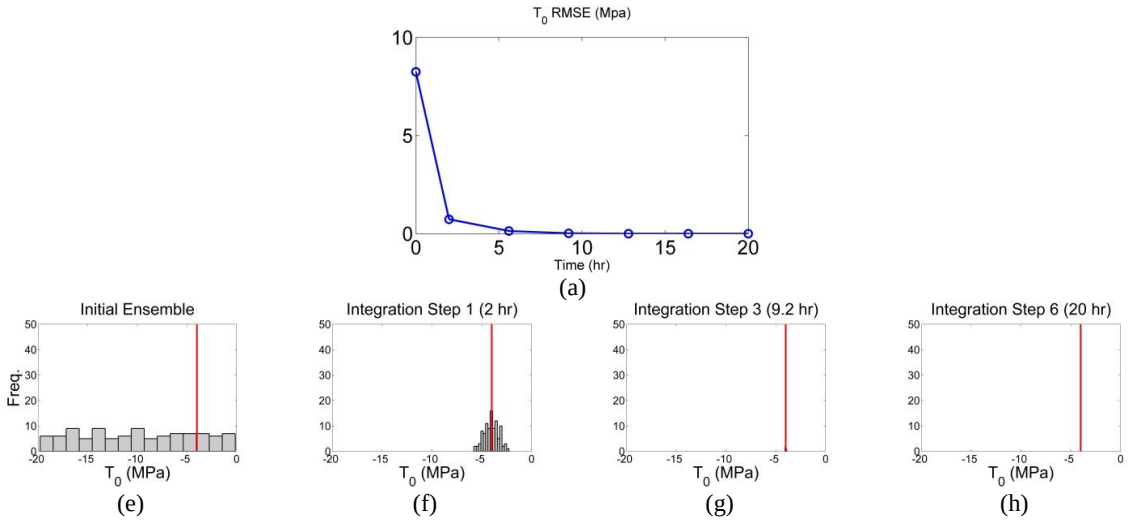


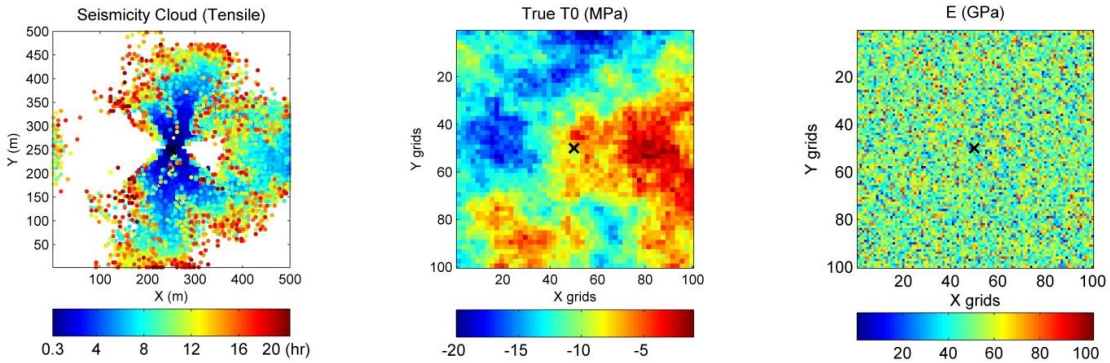
Figure 4.22: Results of homogeneous tensile strength T_0, estimation; Plot (a) shows the $RMSE$ of estimation. Plots (b) to (e) show the evolution of ensemble histogram in integration steps (true $T_0 = -4$ MPa).

Based on these three examples for estimating homogeneous geomechanical reservoir parameters, we see the promise of using EnKF to effectively infer unknown reservoir parameters using MEQ data.

4.4.2 Experiment 2: 2D Heterogeneous Tensile Strength

In this experiment we estimate heterogeneous tensile strength T_0, distribution from MEQ monitoring data. The true model specifications are shown in Figure 4.23. Figure 4.23a shows the microseismicity cloud (due to only tensile failure) generated by the geomechanical forward model for a given distribution of tensile strength T_0, (Figure 4.23b). In this setup, elastic modulus E, is assumed to be spatially random distributed and known. In this experiment permeability k, is constant in space. In

the estimation process all the parameters except T_0 are assumed known and the unknown (assumed) T_0 is estimated. We generate Young's modulus E, distribution from a normal probability distribution by specific mean and standard deviation which is shown in Figure 4.23 (μ_E = mean of normal distribution = 50 GPa, σ_E = standard deviation of normal distribution = 15 GPa).



(a) true microseismicity observation (b) true T_0 map (c) E map

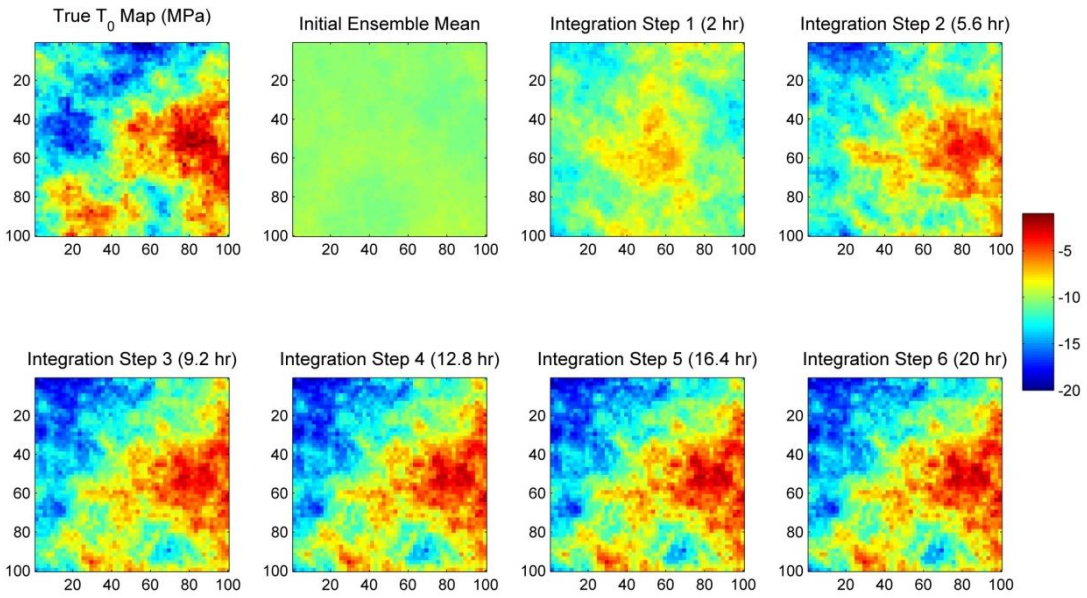
Figure 4.23: True model specification. $S_{h,min}$ = 20 (MPa), $S_{H,max}$ = 15 (MPa), P_{ini} = 10 (MPa), Inj. Rate = 15 (Lit/s), k = 0.005 (md), Young's modulus E, with Normal PDF (spatially random distribution) μ_E = 50 (GPa), σ_E = 15 (GPa). Black cross at the center shows the injection well location.

In this example the standard SSBRC method with MEQ density map on the original fine mesh is used for data integration. The observation error Std range (σ_{min}, σ_{max}) is also assumed 20 % to 55 % which is somewhat inflated comparing to the typical range of observation Std. Figure 4.24a shows the estimation results by illustration of ensemble mean evolution throughout integration steps. In this EnKF procedure we use 100 random realizations of T_0 distributions as the initial ensemble. Before using any observation (seismicity data) the mean of the initial ensemble does not have any spatial trend or feature because it is made of 100

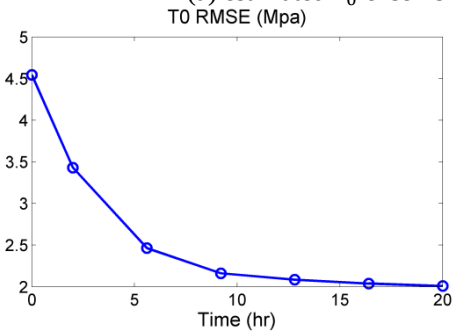
random realizations (as we can see in Figure 4.24a, initial ensemble mean is almost uniform and non-informative). By subsequent integration of seismicity observations we can see that the ensemble mean is developing some special features in it, and finally the ensemble mean becomes very similar to the true tensile strength T_0, map. Figure 4.24b shows the estimation error evolution in time which is $RMSE$ of T_0 and is decreasing with time which means the estimation error is decreasing and the estimated map is becoming closer to true map. The tensile strength ensemble spread evolution throughout integration steps is shown in Figure 4.24c and it clearly shows severe ensemble spread underestimation of standard SSBRC where the final spread in only 3 %.

Figure 4.25a shows the evolution of an individual tensile strength realization throughout MEQ data integration procedure in 3 integration steps (out of 6). In Figure 4.25b, the corresponding MEQ clouds of the intermediate estimated T_0 maps are shown. As shown in Figure 4.25a, in terms of the estimated parameter, the estimated T_0 map becomes increasingly similar to the true T_0 map (shown in Figure 4.23b) throughout integration steps. In terms of the predicted observation, it is clear from Figure 4.25b that the corresponding predicted MEQ cloud of the intermediate estimated T_0 map develops high similarity with the true MEQ cloud (presented in Figure 4.23a) during assimilation procedure. This result confirms the effectiveness of the proposed SSBRC method for reservoir parameter estimation by MEQ data integration in simultaneously appropriate adjustment of the reservoir property and honoring the observed MEQ data.

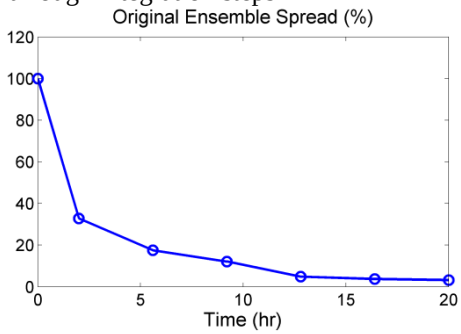
The initial ensemble represents the uncertainty in the parameter before integrating the seismicity observation. Having made of random samples the spread or uncertainty in initial ensemble is high. In the integration procedure by incorporating new data the samples are corrected and the new estimated ensemble will have



(a) estimated T_0 ensemble mean evolution through integration steps

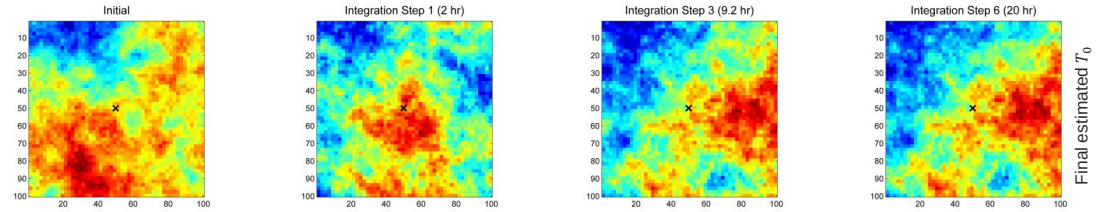


(b) T_0 RMSE evolution in time

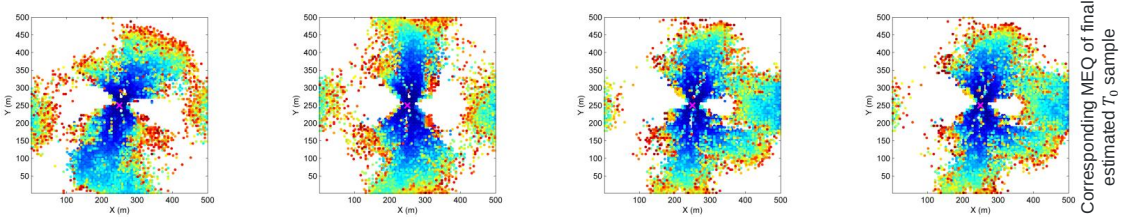


(c) T_0 ensemble spread

Figure 4.24: Tensile strength T_0, estimation results in terms of estimated ensemble mean map, RMSE and ensemble spread.



(a) evolution of one of the T_0 samples by MEQ data integration (T_0 sample estimation evolution)



(b) corresponding simulated MEQ for each estimated T_0 sample in each integration step

Figure 4.25: Evolution of a single tensile strength T_0, realization and its corresponding MEQ cloud throughout integration steps.

smaller uncertainty or standard deviation (Std). So the ensemble spread changes by seismicity integration. As we integrate more data, the uncertainty in the ensemble decreases which results in reduction of the ensemble spread. To analyze the ensemble spread, the standard deviation (Std) of the ensemble is calculated. Figure 4.26 shows the evolution of the standard deviation of tensile strength T_0, ensemble and as it is seen, the uncertainty in the ensemble is reduced. Uncertainty reduction in the ensemble is what we expect from the EnKF method. However as we discussed before the standard SSBRC with high resolution seismicity density map leads to severe ensemble spread underestimation.

4.4.3 Experiment 3: 2D Heterogeneous Elastic Modulus

In this experiment we present the application of standard SSBRC to estimate the heterogeneous distribution of Young's modulus E, from the MEQ cloud observation. To generate the seismicity density map, the original fine mesh is used in this experiment. The synthetic true Young's modulus distribution (parameter to

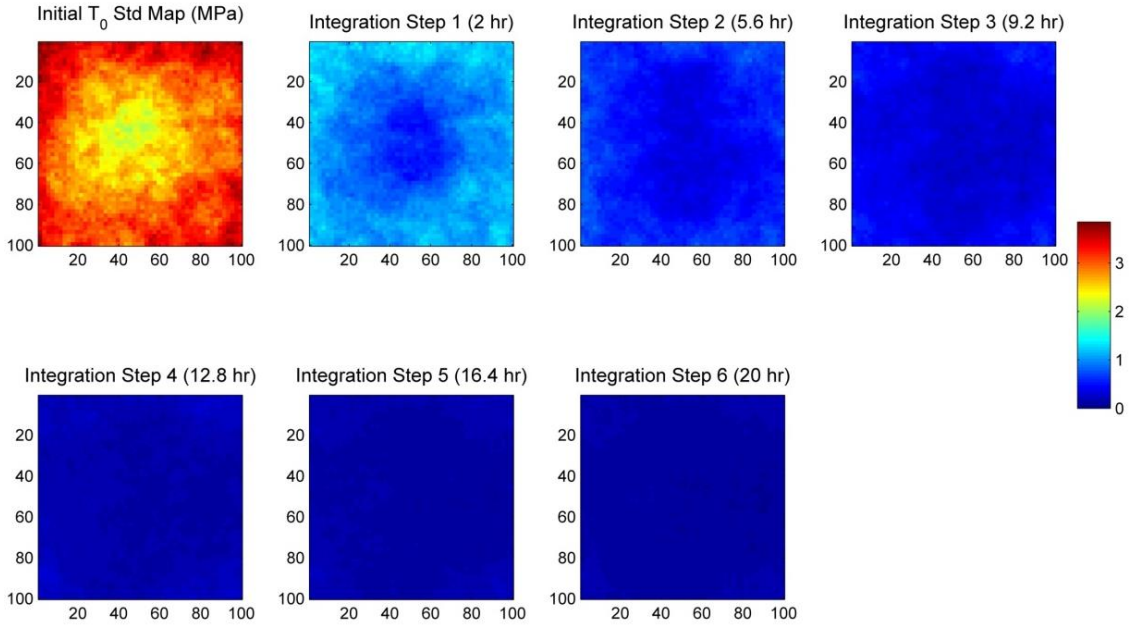
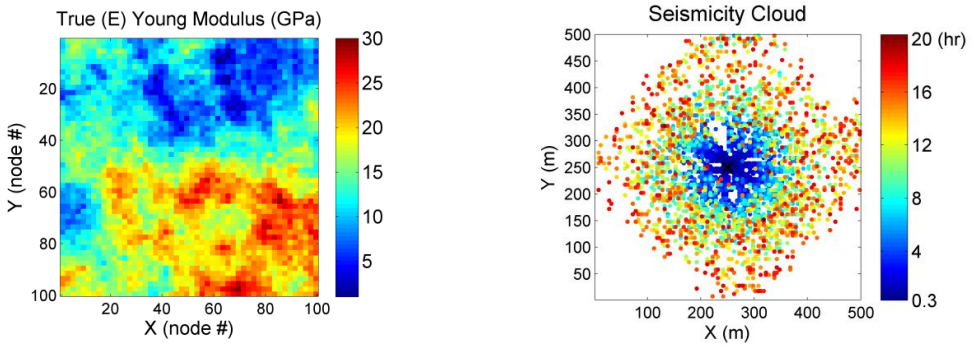


Figure 4.26: Evolution of tensile strength T_0, ensemble standard deviation map throughout integration steps.

estimate) and the corresponding seismicity cloud as the observed data are shown in Figure 4.27. The far-field stress boundary conditions $S_{H,max}$, $S_{h,min}$ and P_{ini}, are all assumed 10 (MPa) to investigate only the effect of hydraulic stimulation (injection-induced stress) in the reservoir. Fluid injection rate is equal 12.5 (Lit/s) and permeability is homogeneous in the field and is equal 0.005 (md). Tensile strength T_0, is assumed spatially random distributed and its values come from a Uniform PDF in the range of 5 (MPa) to 30 (MPa). In this experiment only tensile failures (as MEQ events) are considered.

To illustrate the induced fracture and its propagation, the damage factor and enhanced permeability at the end of hydraulic injection are recorded. The damage factor distribution as well as the enhanced permeability distribution show the induced fracture network and its propagation in the field. An example is presented in Figure



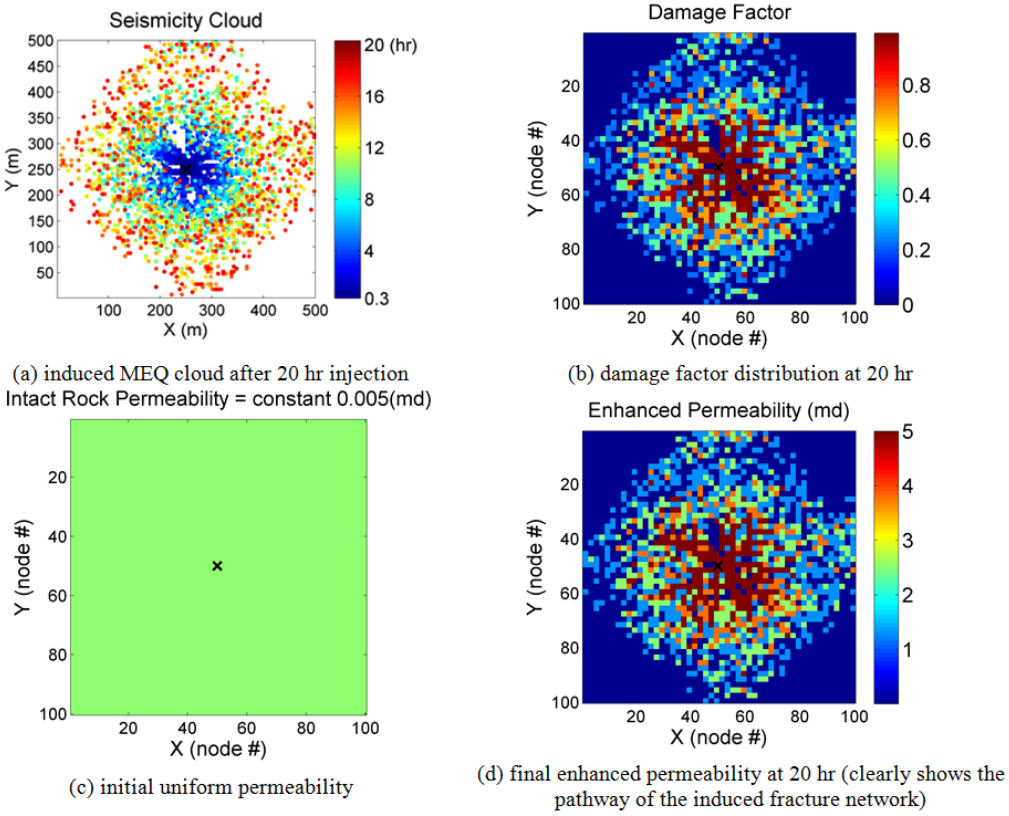
(a) True Young's modulus (E) map

(b) True microseismicity cloud (tensile events)

Figure 4.27: True model setup; true Young's modulus distribution and the corresponding MEQ cloud.

4.28 where high damage factor regions as well as high enhanced permeability regions (dark red regions) clearly show the induced fracture and its propagation (which looks like a complex fracture network).

This experiment was performed with the standard EnKF algorithm by a typical range of 5 % to 10 % for observation error standard deviation. Estimation results are shown in Figure 4.29. The evolution of the estimated ensemble mean of Young's modulus realizations, an estimated individual realization and the standard deviation of the ensemble in integration steps are shown in Figure 4.29a,b,c, respectively. The estimated Young's modulus maps in Figure 4.29a,b, tend to identify the major high- and low-value regions of elastic modulus in the reservoir. The final estimated maps (at 6^{th} integration step) of Figure 4.29a,b, are very similar to the true Young's modulus map. However, the problem with these estimation results is a very low final ensemble spread and ultimately collapsing the ensemble to a single realization (final ensemble spread is only 3 % which means the ensemble loses 97 % of its spread through data assimilation procedure). The experiment of Figure 4.29 is the reference experiment that we will improve its results (ensemble spread) by our



(a) induced MEQ cloud after 20 hr injection

(b) damage factor distribution at 20 hr

(c) initial uniform permeability

(d) final enhanced permeability at 20 hr (clearly shows the pathway of the induced fracture network)

Figure 4.28: Showing the induced fracture propagation through damage factor and enhanced permeability distribution.

proposed approaches.

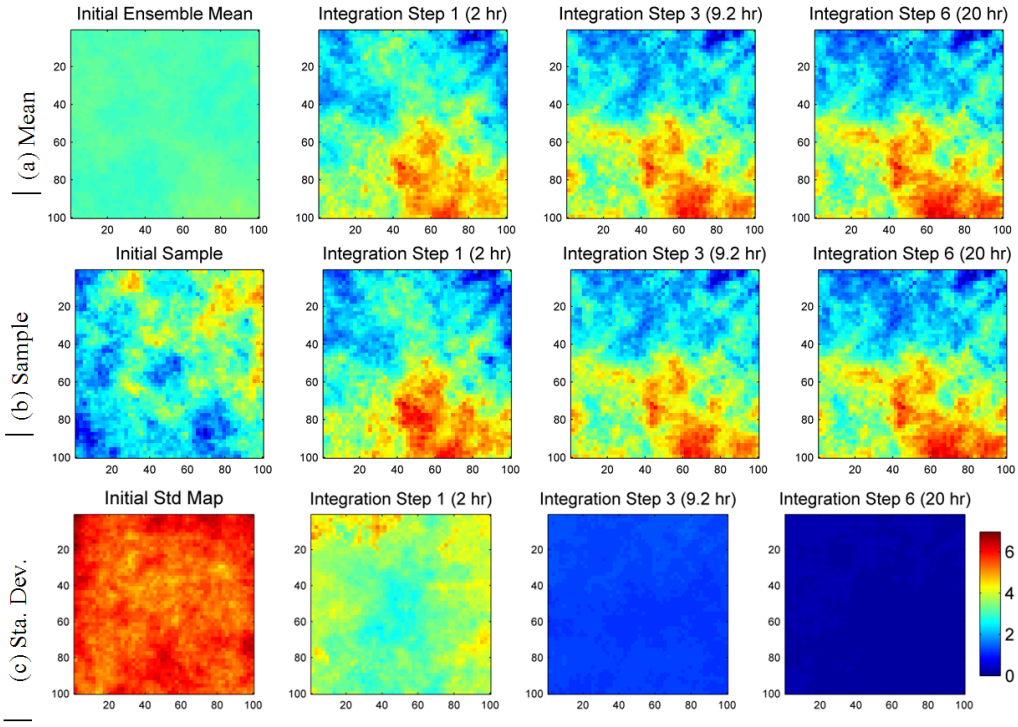


Figure 4.29: Young's modulus E, estimation results of standard SSBRC with typical range of observation error Std (the reference experiment).

In the performed experiments, as the typical pitfall with EnKF, the ensemble collapse might be because of the small ensemble size (100 realizations). To demonstrate that the issue of ensemble collapse is mostly arising from high dimensional observation, we perform an experiment with a large ensemble size (1000 realizations) to address the small ensemble size situation. This experiment is performed with the unmodified or standard EnKF with the typical range of observation error Std and will be called the reference spread experiment because it shows what the value of ensemble spread would be with a large enough ensemble

that removes the effect of small ensemble size. We will show that increasing the size of ensemble will result in larger ensemble spread which demonstrates small ensemble size is not solely the reason of ensemble spread underestimation and large number of correlated observations also contribute to this underestimation. Figure 4.30 shows the results of reference spread experiment. This figure shows that larger ensemble can result in greater final ensemble spread and the main cause of spread underestimation in this problem is the very high-dimensional observation. The final estimated elastic modulus maps of Figure 4.30a,b are acceptably close to the true elastic modulus map. Final ensemble spread is 22 % which is considerably higher than the final spread in the reference experiment in Figure 4.29 (only 3 %). The final Std map in 4.30c (at integration step 6) also demonstrates that large ensemble size will result in higher ensemble spread and therefore the main cause of ensemble spread underestimation in SSBRC is having too many observations.

This experiment with a large ensemble shows that by increasing the number of realizations the final spread will become larger but still it is much lower than the expected final spread for instance in applications where production data (well bottom hole pressure and fluid rate data) is assimilated and this unusually small final ensemble spread resulted from the nature of the observation in this application which is very high dimensional and correlated. Figure 4.31 shows the comparison of performance measures for small and large ensemble. As shown in Figure 4.31 the estimation RMSE is the same for both small and large ensemble and also this figure demonstrates that the final ensemble spread with large ensemble is considerably higher than small ensemble case.

The experiments in this section (and many others with the same setup but different reference reservoir property) shows promising estimation performance of the SSBRC however the problem of ensemble spread underestimation is evident. In

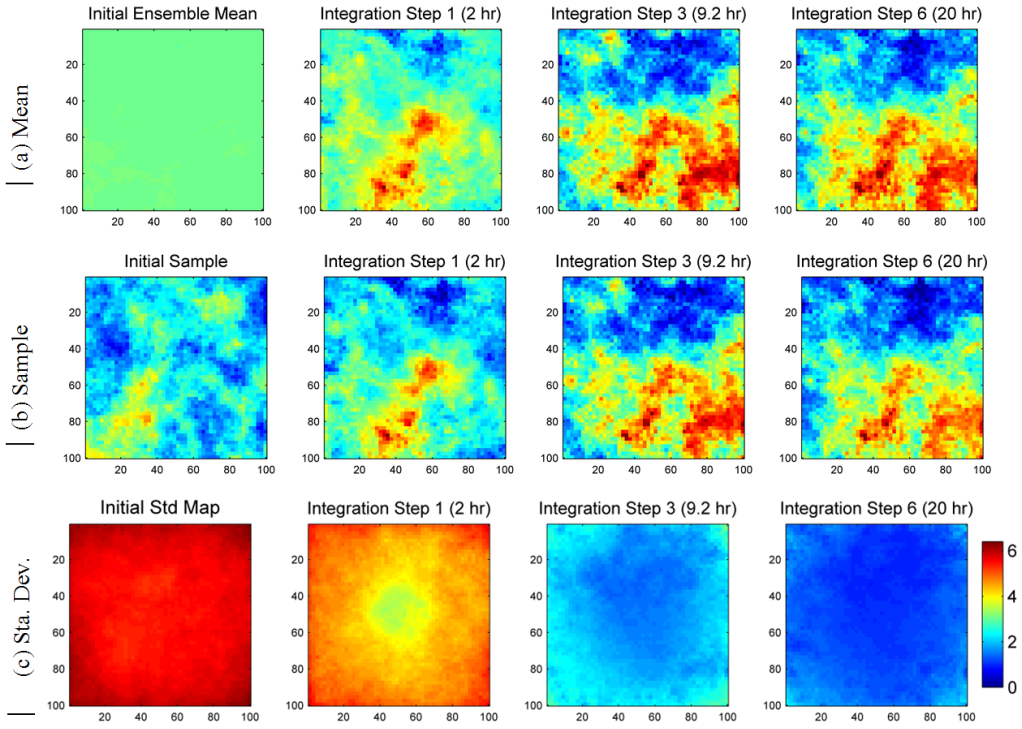


Figure 4.30: Young's modulus E, estimation results of SSBRC with typical range of observation error Std (the reference spread experiment with 1000 realizations), its true model is shown in Figure 4.27.

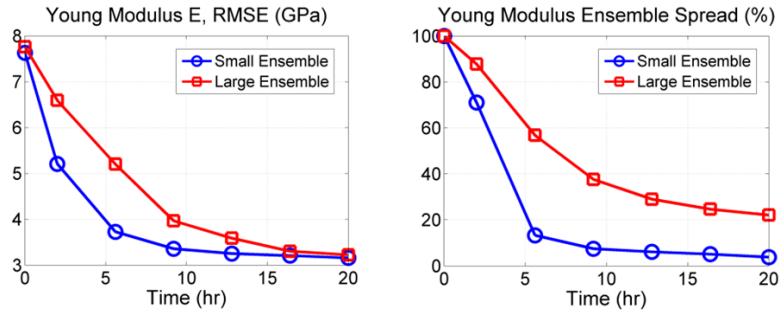


Figure 4.31: Comparing the EnKF experiments with 100 (small ensemble) and 1000 (large ensemble) realizations

the next set of experiments we apply the proposed methods for improving uncertainty quantification to the example in Figure 4.29 (as the reference experiment) to resolve ensemble spread underestimation issue.

4.4.4 Experiment 4: Improving Experiment 3

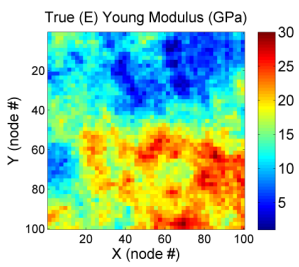
In this set of experiments, we apply the three methods of improving uncertainty quantification in section 3.2 to the experiment 3 in Section 4.4.3 (reference experiment) to resolve the issue of ensemble spread underestimation.

4.4.4.1 Inflated Observation Error Variance

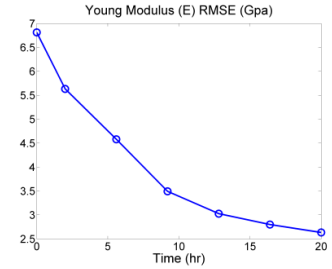
The simplest way of avoiding spread underestimation or ensemble collapse is adding large amount of noise to observation which can be done by specifying high observation standard deviation (Std) through increasing σ_{min} and σ_{max} in Equation (3.5). Figure 4.32 shows the estimation results of SSBRC with inflated observation error Std range of 100 % - 200 %.

The decreasing trend of estimation $RMSE$ in Figure 4.32b shows promising estimation performance of SSBRC even with greatly increased Std range of observation. Figure 4.32d,e,f, represent the evolution of ensemble mean, an individual realization and standard deviation map throughout integration steps. As shown by final estimated Young's modulus maps (at 6^{th} integration step) in Figure 4.32d,e, SSBRC is successful in capturing the trends of true elastic modulus of Figure 4.32a. More importantly ensemble spread underestimation problem of SSBRC is also resolved as can be seen in Figure 4.32c that shows the ensemble spread of the estimated parameter.

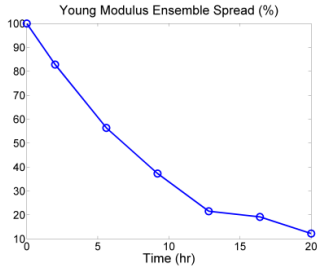
In general in SSBRC, it is the correlation of MEQ events with the reservoir parameter (elastic modulus in this case) that is exploited by the EnKF update to reconstruct the trend in the reservoir property map.



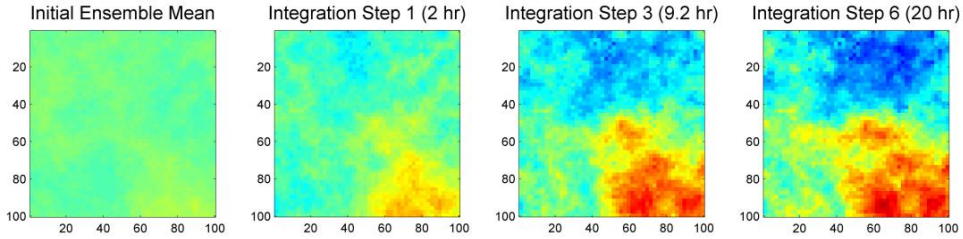
(a) true Young's modulus map



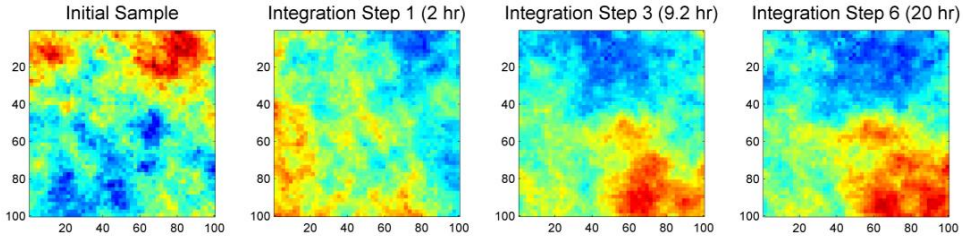
(b) estimation RMSE



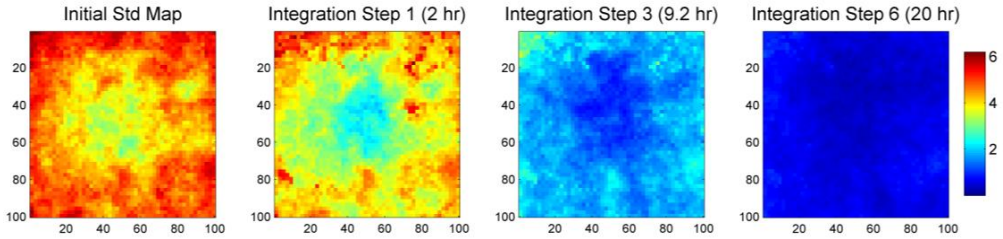
(c) ensemble spread



(d) evolution of ensemble mean



(e) evolution of an individual realization



(f) standard deviation map evolution

Figure 4.32: Young's modulus E, estimation results of SSBRC with increased observation error Std, its true model is shown in Figure 4.27.

4.4.4.2 Reduced-Order Projection

In this section we apply the projection method of Section 3.2.2 to experiment 3 in Section 4.4.3 to improve the final estimated ensemble spread. As mentioned earlier, projection method has two tuning parameters; kernel bandwidth standard deviation σ_h, and truncation number n_{trunc} that in this example are chosen to be 25 % and 100, respectively. It should be noted that the number of observations is decreased from 10000 to 100, by projection method. The proposed observation projection approach is successful in preserving the ensemble spread and resolving ensemble spread underestimation of SSBRC through observation dimension reduction. However, the resulted estimation $RMSE$ with projection approach is not as low as the standard SSBRC method. Detailed estimation results are shown in Figure 4.33.

We reduced the number of observations from 10000 (in the reference experiment) to 100 (n_{trunc}) which resolves the observation redundancy issue of SSBRC. Figure 4.33 shows Young's modulus E, estimation results by SSBRC with projection approach. The final estimated Young's modulus distributions of Figure 4.33, (at integration step 6) are very similar to the true E distribution. The main advantage of projection approach in preserving ensemble spread and preventing ensemble collapse is shown by final ensemble Std map in Figure 4.33b (at integration step 6). Figure 4.33d shows that final ensemble spread is 58 % which is significantly improved comparing to the reference experiment (Figure 4.29) by final ensemble spread of only 3 %.

As explained and formulated in Section 2.3.1 there is also uncertainty in event occurrence time that should be involved in the data integration process. Here the results of SSBRC with the consideration of event occurrence time uncertainty is presented. Examples determined that the time kernel bandwidth h_t equal to three times of time step should be a suitable choice. The projection approach parameters

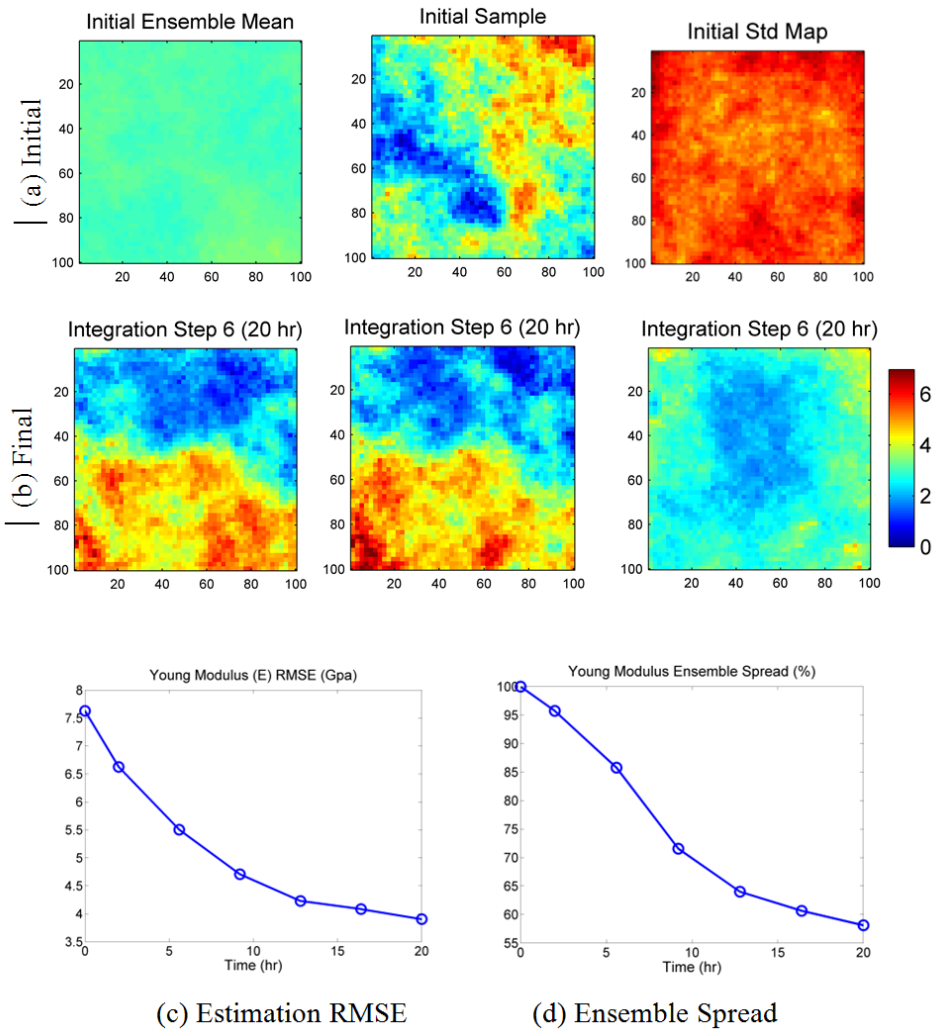


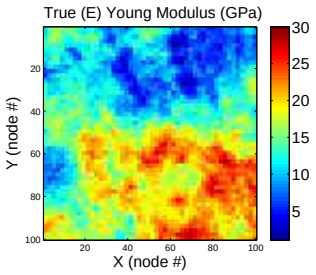
Figure 4.33: Young's modulus E, estimation results of SSBRC with projection approach, its true model is shown in Figure 4.27.

of n_{trunc} and σ_h are 100 and 25 %, respectively. While the total stimulation time is 20 (hr) consisting of 100 time steps of 720 (s), the standard time uncertainty handling approach is capable of dealing with event occurrence time uncertainty (1D Gaussian time Kernel bandwidth, h_t) of 10 % of the time step (72 sec) to three times of the time step (example with standard time uncertainty approach is not presented here). In the standard approach time step refinement is not employed. An example is presented in Figure 4.34 where SSBRC with projection approach along with adaptive time step refinement is employed to estimate Young's modulus distribution from MEQ clouds. In this example the time uncertainty interval or 1D time Guassian kernel bandwidth is set to 3 (s) while the method could handle any given small time uncertainty interval.

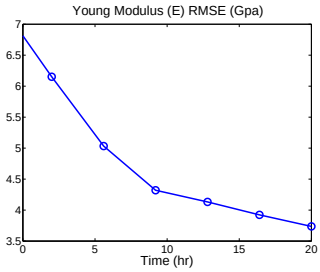
Figure 4.34 presents the results of SSBRC with time step refinement for handling event occurrence time uncertainty. The estimated maps and estimation RMSE as well as ensemble spread show promising performance of SSBRC in estimating Young's modulus while there is uncertainty in occurrence time. Comparing to the case with only consideration of location uncertainty, it seems that the consideration of time uncertainty does not change the estimation results considerably, and in the case of considering time uncertainty, implementing either large or small time uncertainty does not have significant effect on the estimation results.

4.4.4.3 Coarse-Scale Microseismicity Density Map

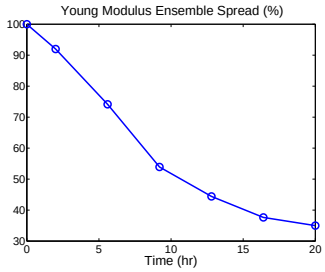
The number of observations is equal to the dimension of seismicity density map. To reduce the dimension of the seismicity density map which is the major reason of the ensemble spread underestimation we build the seismicity density map on a coarse mesh or grid system instead of the original fine mesh. So we only need to evaluate continuous seismicity density map on a new coarse grid system. In 2D case studies



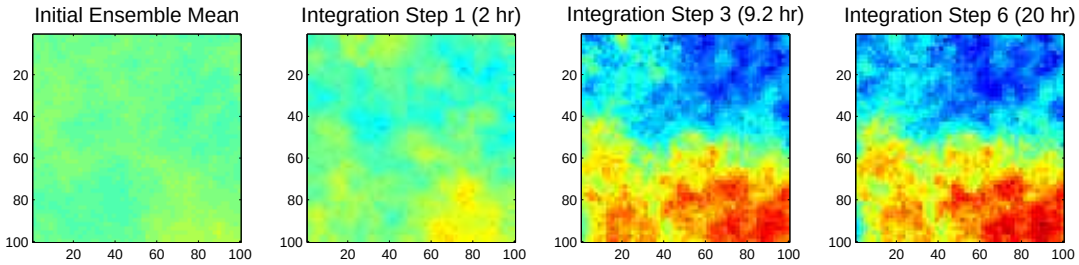
(a) true E map



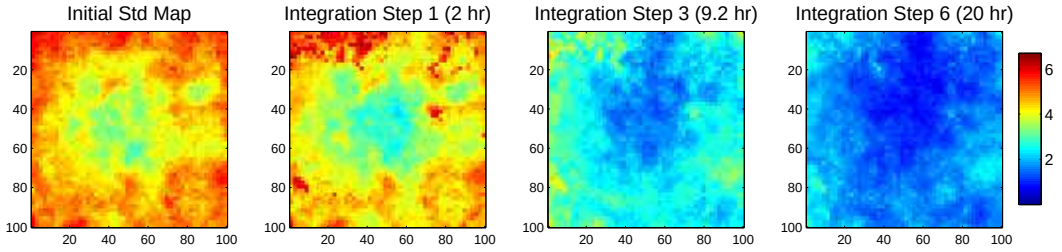
(b) estimation $RMSE$



(c) ensemble spread



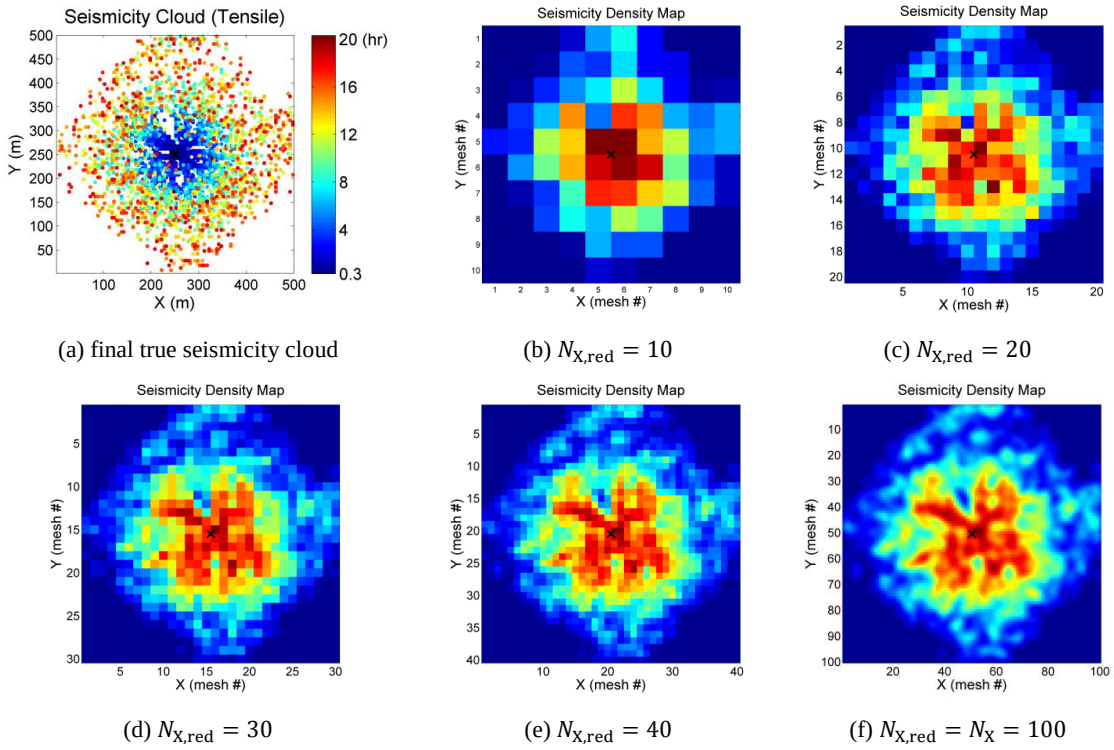
(d) evolution of ensemble mean



(e) standard deviation map evolution

Figure 4.34: Young's modulus E, estimation results of SSBRC with projection approach with the consideration of uncertainty in event occurrence time along with adaptive time step refinement.

where the node configuration size is 100×100 ($N_b = 10000$, $N_X = 100$) we choose the coarse seismicity density mesh size e.g. equal to 10×10 ($N_b = 100$ seismicity observations, $N_{X,red} = 10$) which reduces the dimension of the observation vector by orders of magnitude. It should be noted that the range of the observation error Std in coarse seismicity density approach is the typical range of 5 % to 10 % which leads to ensemble collapse in the regular approach i.e. reference experiment in Figure 4.29 (where the size of seismicity density mesh is equal to the size of parameter mesh).



(a) final true seismicity cloud

(b) $N_{\mathrm{X,red}} = 10$

(c) $N_{\mathrm{X,red}} = 20$

(d) $N_{\mathrm{X,red}} = 30$

(e) $N_{\mathrm{X,red}} = 40$

(f) $N_{\mathrm{X,red}} = N_{\mathrm{X}} = 100$

Figure 4.35: Microseismicity density maps (with different sizes) for final seismicity cloud. (b-e) show coarse seismicity density maps. (f) reference seismicity density map (black cross at the center shows the injection well location).

Figure 4.35 shows the seismicity density maps with different sizes (based on final seismicity cloud of Young's modulus estimation problem from tensile

seismicity events in Section 4.4.3). Different microseismicity density map sizes produces different number of observations and help in reducing the number of observations to resolve the observation space redundancy and ultimately ensemble spread underestimation in EnKF process. The dimension of microseismicity density map is considered in terms of number of nodes in x direction (the field configuration is square). There are 10000 parameters to estimate. For coarsening of the seismicity density map, we specify different sizes of $N_{X,red}$ equal to 10, 20, 30, 40, that decreases the number of observations to 100, 400, 900, 1600, respectively where in the reference experiment by $N_{X,red} = N_X = 100$, there are 10000 observations.

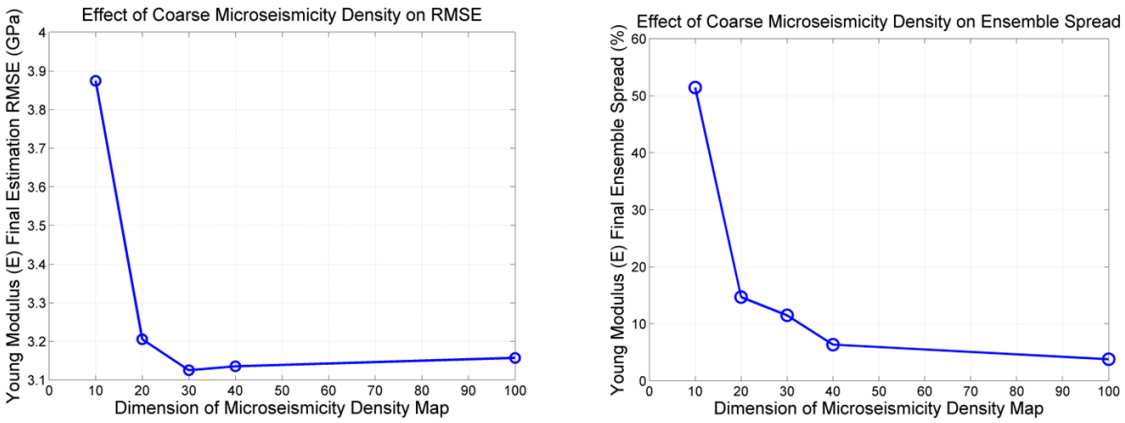


Figure 4.36: Effect of coarse microseismicity density dimension on the performance of SSBRC in Young's modulus estimation experiment (initial Young's modulus $RMSE$ = 6.8134 GPa).

The effect of different dimensions of microseismicity density map on the performance of SSBRC process is shown in Figure 4.36. Figure 4.36 shows both final estimation $RMSE$ and final (original and auxiliary) ensemble spread of these five experiments. Based on Figure 4.36, it is seen that there is an optimal size of the seismicity density map which results in the minimum $RMSE$ (at $N_{X,red} = 30$).

It should be noted that this coarse seismicity density map size is optimal in $RMSE$ sense however we are looking for both low $RMSE$ and high spread simultaneously. Figure 4.36 shows the final ensemble spread (both original Young's modulus ensemble and auxiliary ensemble) versus size of seismicity density map. As we already expected, by increasing the size of the density map there will be higher number of observations which consequently results in more underestimation of ensemble spread. As the result, the improvement of the ensemble spread is obviously seen in Figure 4.36 (spread curve) by decreasing the size of the seismicity density map (or using coarse seismicity density map). To choose the appropriate size of the density map, we are looking for fairly low RMSE and high spread in Figure 4.36, respectively. Based on the obtained results, the density size of 10 to 30 (in a 100×100 field) for coarse-scale mesh can be a suitable choice.

The estimation results of SSBRC with coarse seismicity density map $N_{X,red} = 10$, are shown in Figure 4.37 which represents the evolution of ensemble mean, an individual realization and standard deviation map throughout integration steps. The final estimated ensemble mean and an individual realization (at integration step 6) show appropriate performance of SSBRC in capturing the unknown true Young's modulus distribution. Considering the performance of SSBRC with coarse density in terms of ensemble spread, final standard deviation map of elastic modulus ensemble clearly shows appropriate amount of ensemble spread at the final integration step. From Figure 4.37d, we also see that the final ensemble spread is 52 % which shows promising performance of coarse seismicity density method in preserving the ensemble spread.

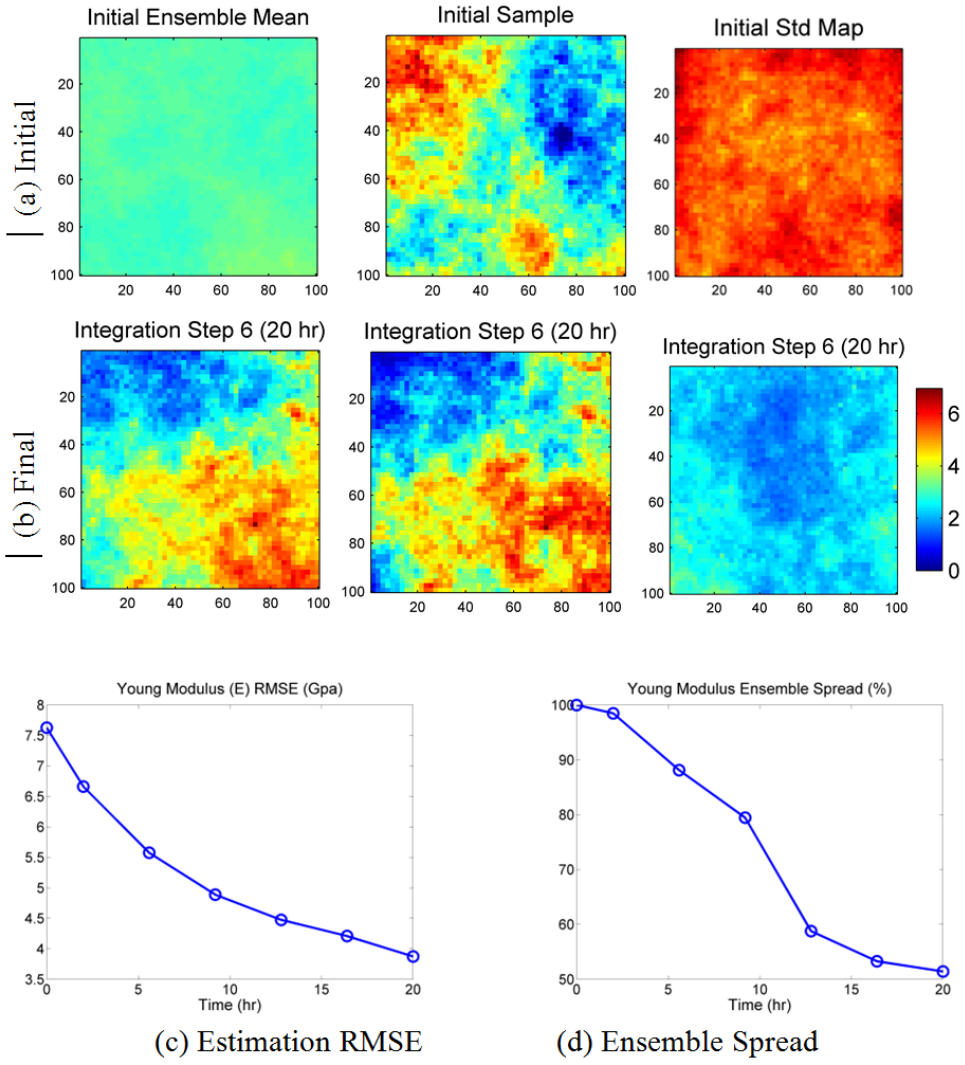
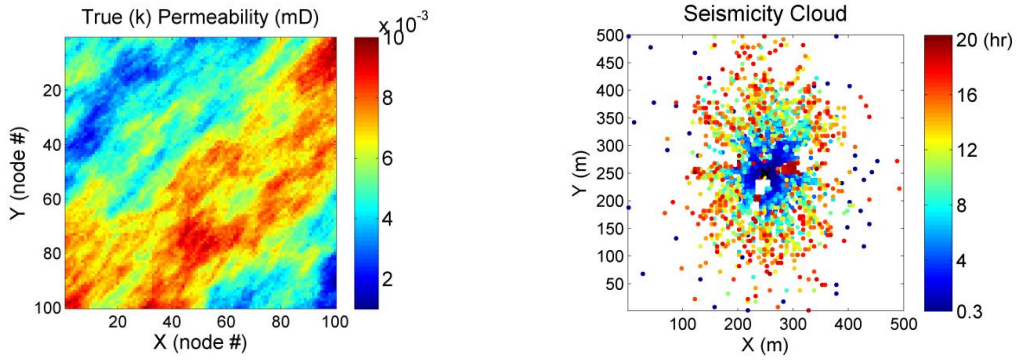


Figure 4.37: Results of SSBRC with coarse seismicity density map $N_{X,red} = 10$, for Young's modulus estimation. Its true model is shown in Figure 4.27.

4.4.5 Experiment 5: 2D Heterogeneous Permeability

In this experiment we estimate permeability distribution from MEQ cloud (shear failures). In this example we apply the projection method of Section 3.2.2 along with SSBRC to estimate permeability distribution from MEQ cloud and to improve the final estimated ensemble spread. The true setup of the model is shown in Figure 4.38. The far field stress regime (boundary condition) is assumed $[S_{min},\ S_{max},\ P_{ini}] = [20,\ 15,\ 10]$ (MPa). There is one injection well at the center of the model. Fluid injection rate is equal 12.5 (Lit/s) and Young's modulus E, is homogeneous in the field and equal 40 (GPa). Cohesion C, is assumed spatially random distributed with Normal PDF by 20 (MPa) and 7 (MPa) as its mean and Std, respectively. Friction angle ϕ, is assumed homogeneous and equal 25°.



(a) True permeability (k) map

(b) True microseismicity cloud

Figure 4.38: True model setup; true k distribution and the corresponding MEQ cloud.

As mentioned earlier, projection method has two tuning parameters; kernel bandwidth standard deviation σ_h, and truncation number n_{trunc} that are set equal to 25 % and 100, respectively. As we can see from this test, the permeability estimation

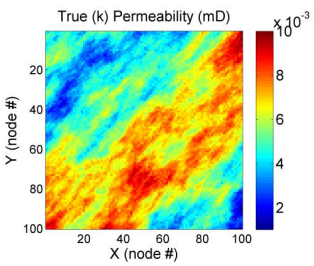
$RMSE$ is fairly low and as the objective of the projection method, ensemble spread is greatly preserved (final ensemble spread is 70 %). It should be noted that, in this experiment, the number of observations is decreased from 10000 to 100, by projection method. Therefore the proposed observation projection approach is successful in preserving the ensemble spread and resolving ensemble spread underestimation of SSBRC through observation dimension reduction.

Figure 4.39 shows permeability k, estimation results by SSBRC with projection approach. The final estimated permeability distributions of Figure 4.39d,e, are reasonably similar to the true k distribution in Figure 4.39a. The main advantage of projection approach in preserving ensemble spread and preventing ensemble collapse is shown by final ensemble Std map in Figure 4.39f. Figure 4.39c shows that final ensemble spread is 70 %.

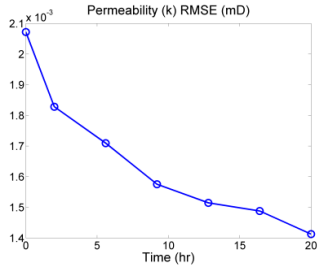
4.4.6 Experiment 6: 3D Heterogeneous Elastic Modulus

In this section, we apply SSBRC to the developed 3D fully coupled thermo-poroelastic finite element method (FEM) model. We present the results of estimating 3D Young's modulus distribution from MEQ cloud using SSBRC. In this experiment we use coarse-scale seismicity density method to preserve the ensemble spread through update steps.

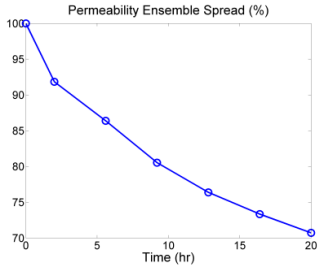
In the 3D model the injection well as a point source is at the center of the field. Permeability distribution is homogeneous in the field and equal to 0.5 (md). We suppose there are only tensile failures, so in order to avoid any shear failures, the cohesion as a homogeneous property in the field is assumed a very high value of 1000 (MPa). In this experiment, the far-field stress boundary conditions including initial reservoir pressure P_{ini}, minimum horizontal stress $S_{h,min}$, maximum horizontal stress $S_{H,max}$ and vertical stress S_v, are equal 10 (MPa), 13 (MPa), 15 (MPa) and



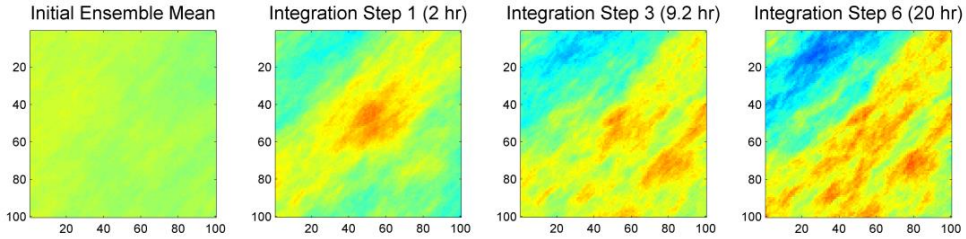
(a) true permeability map



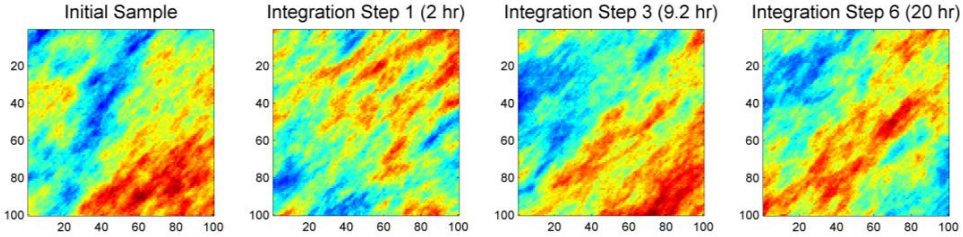
(b) estimation RMSE



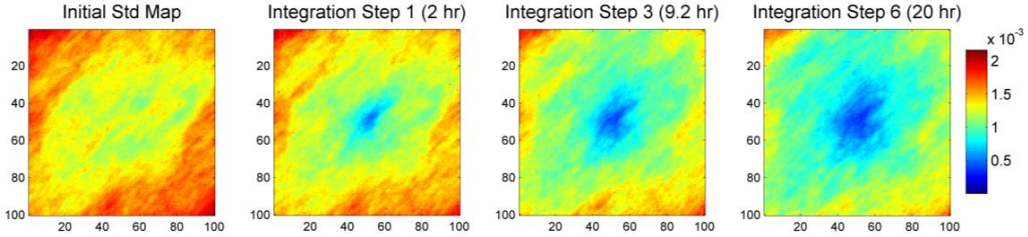
(c) ensemble spread



(d) evolution of ensemble mean



(e) evolution of an individual realization



(f) standard deviation map evolution

Figure 4.39: Permeability k, estimation results of SSBRC with projection approach, its true model is shown in Figure 4.38.

18 (MPa), respectively. In each forward simulation, time step length is 3600 (s) and there are 72 time steps so the stimulation period is 72 (hr). The injection rate scenario is shown in Figure 4.40.

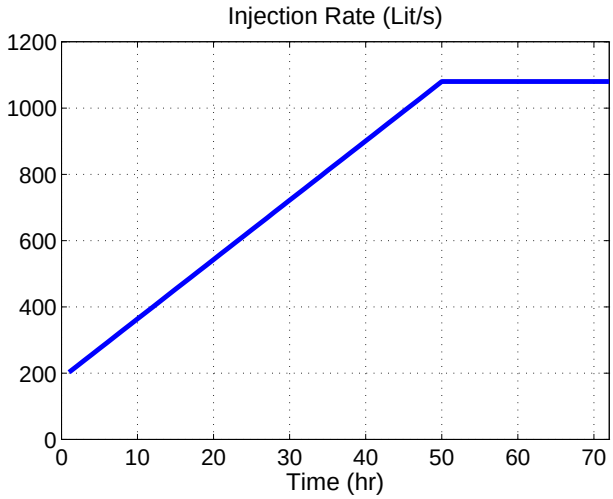
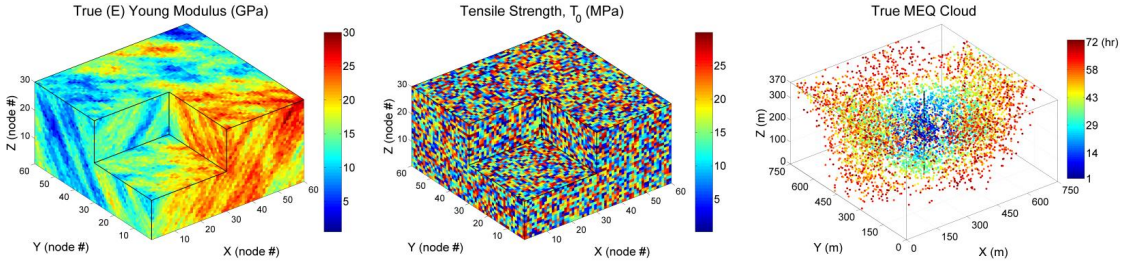


Figure 4.40: Injection rate scenario.

The true Young's modulus distribution (as the parameter to estimate), the assumed tensile strength (failure criteria) and the MEQ cloud response (the true monitoring MEQ data to be integrated in EnKF) are shown in Figure 4.41. In Figure 4.41c, the microseismic events are tensile failures as mentioned previously. By SSBRC stochastic estimation algorithm, we estimate the 3D Young's modulus distribution (Figure 4.41a) which is assumed unknown from discrete MEQ cloud observation (Figure 4.41c).

The vertical black line in MEQ cloud plots (e.g. Figure 4.41c) indicates the location of the injection well and the injection point (point source) is at the center of the model (at the middle of the black line). In SSBRC estimation procedure



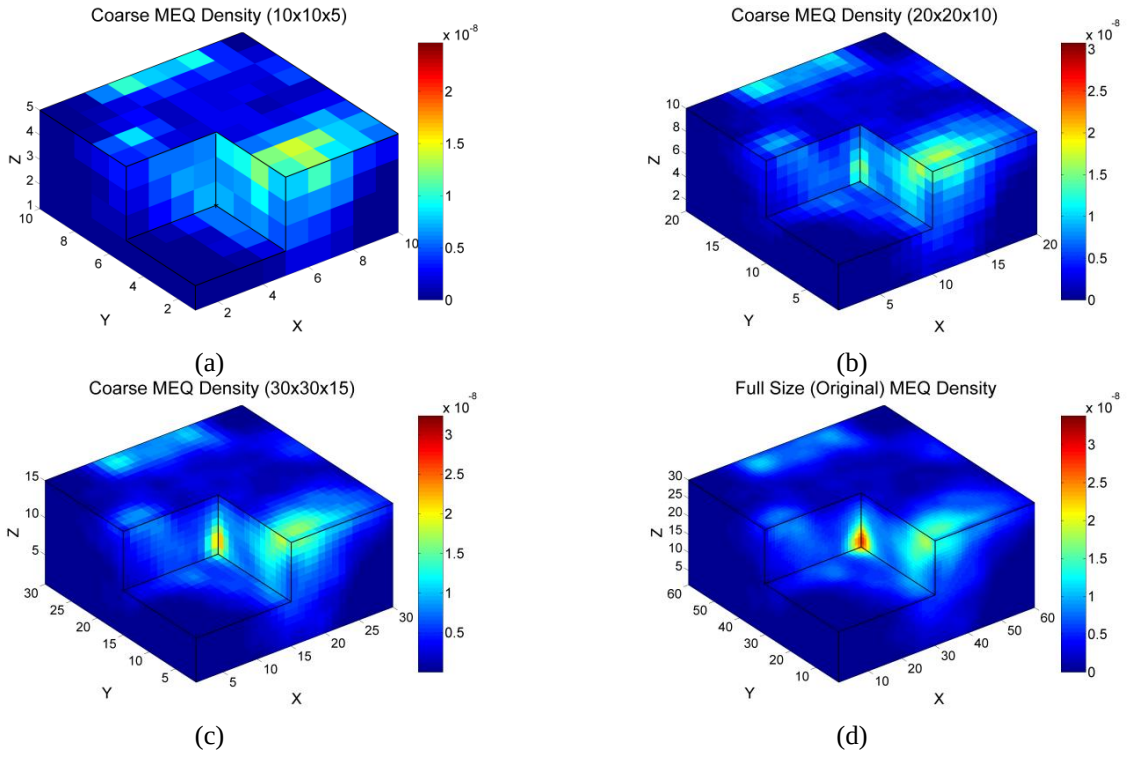
(a) true Young's modulus distribution (b) T_0 distribution (c) true microseismicity cloud

Figure 4.41: True model setup; true Young's modulus and T_0 distribution and the corresponding MEQ cloud.

we consider 6 integration time steps. We sequentially integrate the monitoring MEQ data at each integration step by EnKF to estimate the Young's modulus distribution. In this example we consider using coarse-scale seismicity density method to reduce the number of observations.

The final true MEQ cloud is converted to MEQ density map with different sizes of coarse-scale mesh and the resulted density maps are shown in Figure 4.42. From the investigation of Figure 4.42 and performing the EnKF analysis equation after one integration step with different sizes of coarse scale mesh we chose $20 \times 20 \times 10$ mesh as the appropriate size for coarse-scale continuous MEQ interpretation.

For the EnKF algorithm, the ensemble size (number of realizations) is assumed 100. We also considered 6 integration steps regularly specified in 72 (hr) of total forward run simulation. Initially to make sure we will resolve ensemble spread underestimation of SSBRC we utilize coarse-scale seismicity density method. Since the forward 3D model is very computationally expensive to run, it was not convenient to run SSBRC with coarse-scale density for different sizes of coarse mesh. Therefore to find the appropriate size of coarse mesh, we investigated different sizes of coarse-scale seismicity density map after one integration step only. So we propagated



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

Figure 4.42: Final (at 6^{th} integration step) true MEQ density maps on different coarse size meshes.

the initial ensemble till the first integration step and performed the first EnKF analysis equation with different coarse mesh configuration then we decided about the optimal size based on the best values of $RMSE$ and Spread. By applying coarse-scale seismicity density method in SSBRC for this experiment, we estimate or update 108000 parameters (Young's modulus value at each node) from 4000 observation values (coarse seismicity density map dimension for $20\times20\times10$ mesh configuration) at each integration time step. In the estimation procedure we assume all the parameters of the true model are known but the Young's (elastic) modulus distribution. The true model setup is shown in Figure 4.41.

As the first estimation result we show the estimated mean of elastic modulus ensemble through integration steps in Figure 4.43. While the initial ensemble mean (Figure 4.43b) is almost homogeneous and completely uninformative about the true elastic modulus (Figure 4.43a), as we can see in Figure 4.43b-h by marching through integration steps and integrating MEQ data the estimated ensemble mean becomes more similar to the true Young's modulus distribution (Figure 4.43a) and ultimately the final estimated ensemble mean (Figure 4.43h) captures the trends and features of the true parameter distribution.

The estimation performance of SSBRC is also represented by the root mean squared error ($RMSE$) of estimated elastic modulus ensemble through time (Figure 4.44a). Continuous reduction of estimation $RMSE$ from 5.5 (GPa) to 2 (GPa) through integration steps confirms successful application of SSBRC in characterization of 3D geomechanical models by MEQ data. Uncertainty quantification performance of SSBRC estimation procedure is shown by the ensemble spread which should not reduce to a very low value. Both the original Young's modulus ensemble spread and auxiliary ensemble spread are shown in Figure 4.44b,c, respectively. The ensemble spread percentage is defined as the ratio of the ensemble

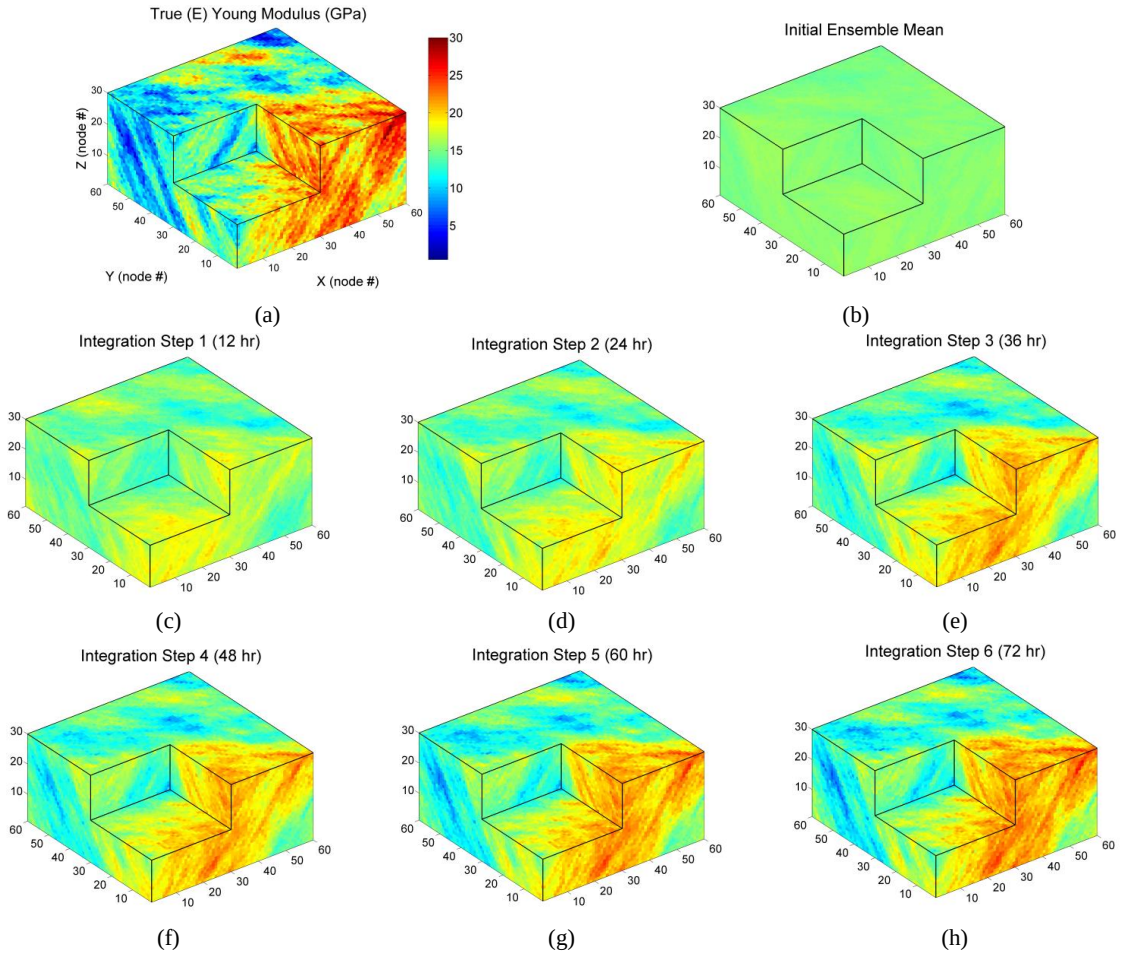
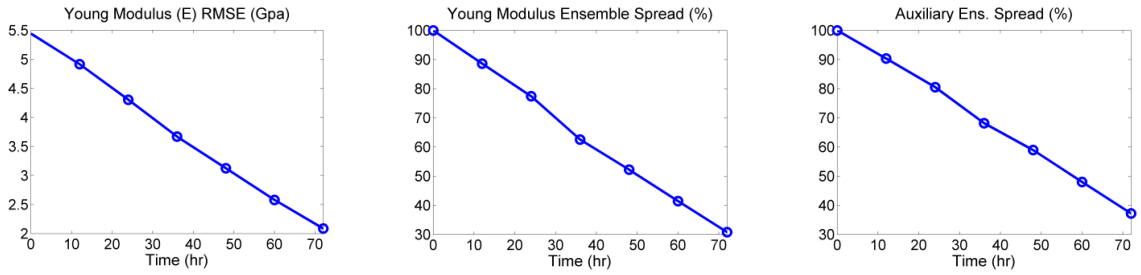


Figure 4.43: Young's modulus estimation results: evolution of estimated ensemble mean through integration steps.

spread at each integration step to the initial ensemble spread. As can be seen the final ensemble spread is 30 % which is a promising in EnKF estimation procedure and it confirms that ensemble collapse did not happen and the ensemble spread underestimation issue is resolved.



(a) (b) (c)

Figure 4.44: Estimation performance measures.

As mentioned before the uncertainty quantification performance of SSBRC is evaluated by ensemble spread or ensemble standard deviation. Therefore the evolution of the ensemble standard deviation map is shown in Figure 4.45 which represents the reduction of the ensemble spread throughout the integration steps that means the realizations are becoming more and more similar and in the same time they are getting close to the true elastic modulus distribution. Reduction of ensemble spread or standard deviation shows that the initial guesses are becoming more similar to true value.

4.4.7 Joint Parameter Estimation in Geomechanical Model

To this point we assumed only one reservoir property is unknown and the rest of the reservoir properties are known. In this section to relax this assumption we present a 2D and a 3D model example to show the suitability of SSBRC to jointly estimate

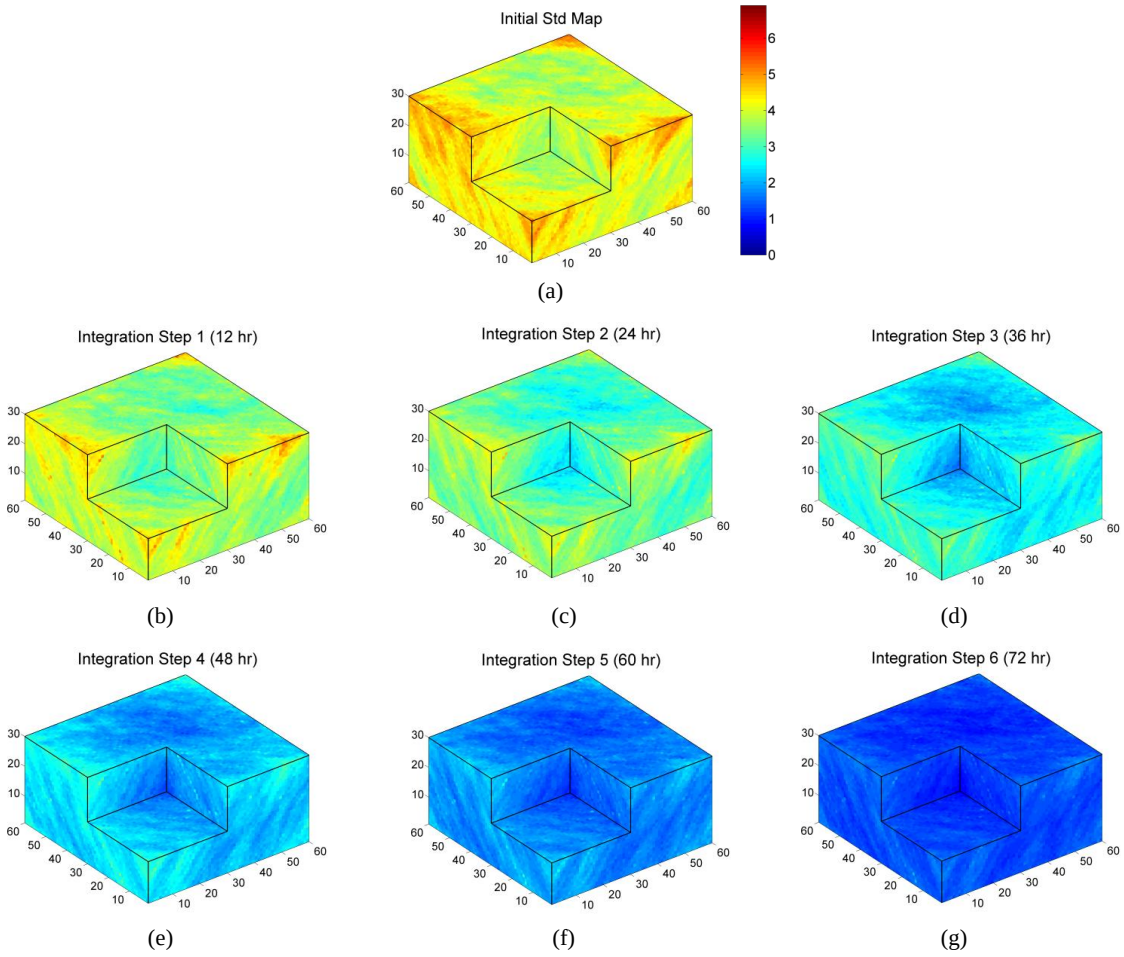


Figure 4.45: Evolution of the standard deviation map of Young's modulus ensemble.

more than one parameter of the reservoir. Assuming that more parameters are unknown or some of the parameters are uncertain will for sure add more uncertainty to the integration procedure and consequently the jointly estimated results are not as good as individual estimation.

4.4.7.1 Joint Estimation of Elastic Modulus and Permeability in 2D Model

Here we present the application of SSBRC to jointly estimate the heterogeneous distribution of Young's modulus E, and permeability k, from the MEQ cloud observation. The synthetic true Young's modulus and permeability distributions (parameters to estimate), tensile strength map and the corresponding seismicity cloud as the monitoring data are shown in Figure 4.46a,b,c,d. The far-field stress boundary conditions $S_{H,max}$, $S_{h,min}$ and P_{ini}, are equal to 13, 11 and 10 (MPa), respectively. Fluid injection rate is equal 12.5 (Lit/s). Tensile strength T_0, is assumed spatially random distributed and taken from a Normal PDF with mean and standard deviation of -15 and 5 (MPa), respectively. In this experiment only tensile failures (as MEQ events) are considered. For parameter ensemble generation, Young's modulus and permeability ensembles are assumed uncorrelated but with the same variogram parameters.

We apply SSBRC along with projection approach to this example (kernel bandwidth Std and truncation number are 15 % and 100, respectively). Detailed estimation result is shown in Figure 4.46. The Young's modulus and permeability estimation results confirm the promising performance of SSBRC with projection method in parameter estimation as well as preserving ensemble spread.

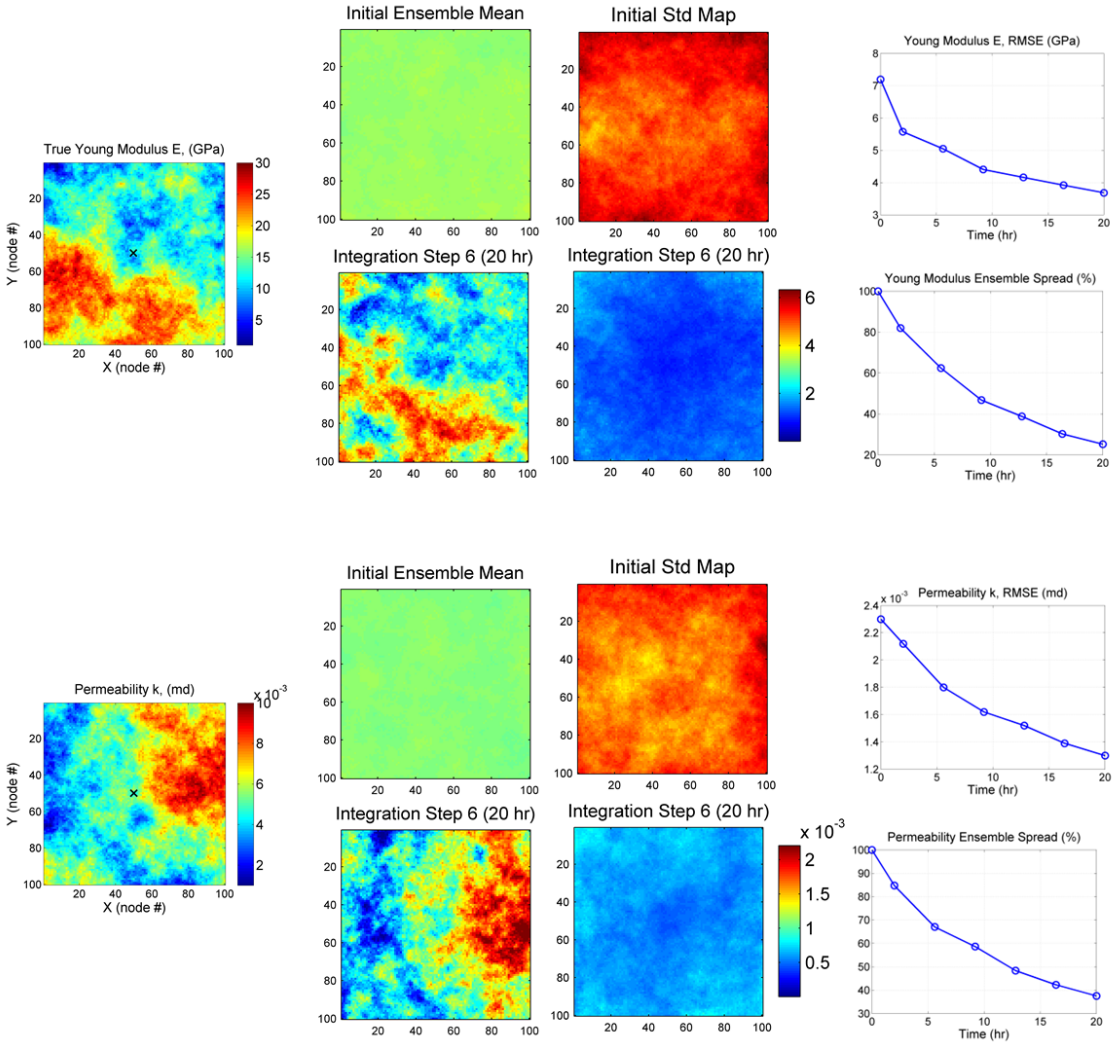


Figure 4.46: Joint Young's modulus (top) and permeability (bottom) estimation with reduced-order projection approach.

4.4.7.2 Joint Estimation of Tensile Strength and Permeability in 3D Model

In this 3D example we present the application of SSBRC to jointly estimate the heterogeneous distribution of tensile strength T_0, and permeability k, from the MEQ cloud observation. The synthetic true tensile strength and permeability distributions (parameters to estimate), Young's modulus map, injection rate scenario and the corresponding seismicity cloud as the monitoring data are shown in Figure 4.47. The far-field stress boundary conditions S_v, $S_{H,max}$, $S_{h,min}$ and P_{ini}, are equal to 18, 15, 13 and 10 (MPa), respectively. Young's modulus E, is assumed spatially random distributed and taken from a Normal PDF with mean and standard deviation of 50 and 20 (GPa), respectively. In this experiment only tensile failures (as MEQ events) are considered. For parameter ensemble generation, tensile strength and permeability ensembles are assumed uncorrelated but with the same variogram parameters.

We apply SSBRC along with projection approach to this example with kernel bandwidth standard deviation σ_h, and truncation number n_{trunc} of 25 % and 100, respectively. In fact we reduce the dimension of the observation from 108000 to 100. Detailed estimation result is shown in Figure 4.48. The tensile strength estimation results in Figure 4.48a-h and permeability estimation results in Figure 4.48i-p confirm the promising performance of SSBRC with projection method in parameter estimation as well as preserving ensemble spread.

The results of the above experiments (and many others with similar well setup but different reference reservoir property maps that were not included) indicate that the EnKF can be used to successfully infer hydraulic and geomechanical reservoir parameter distributions from continuous interpretations (through KDE) of the discrete MEQ monitoring measurements. This outcome has important

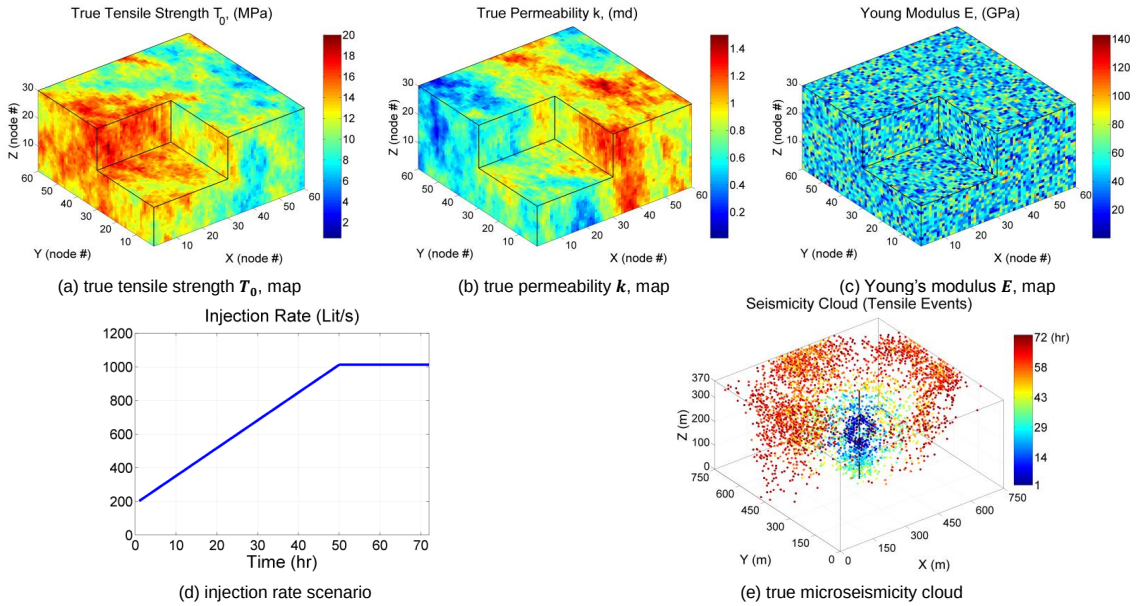


Figure 4.47: True specifications of 3D model for joint estimation; true tensile strength and permeability distribution, Young's modulus map, injection rate scenario and the corresponding MEQ cloud.

implications for characterization of subsurface reservoirs from MEQ events as an emerging monitoring technology in several important energy and environmental applications. While simple and easy to implement, the EnKF proves to be an effective model calibration tool for nonlinear problems where the optimality requirements of the original Kalman filter update equation, namely jointly Gaussian states and measurements and linear state-space model assumptions, are not strictly met. While the examples illustrated in this work clearly show the feasibility of applying the EnKF to constrain different reservoir parameter distributions based on microseismic event locations, we did not consider the seismic analysis step that is required to provide the MEQ sources locations. In addition to event locations, other information about the seismic source may be extracted from the raw seismic data (e.g., the magnitude and confidence of events) and be used to further constrain rock property distributions.

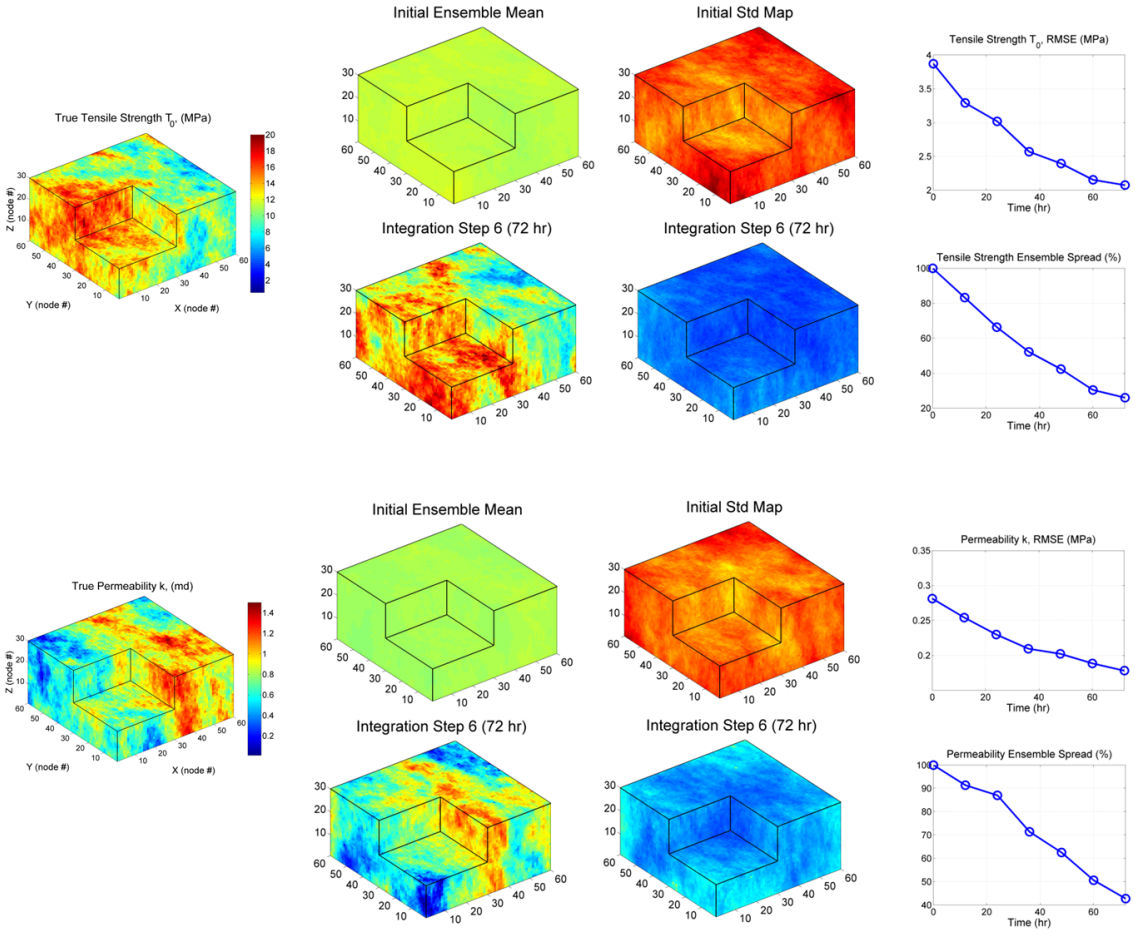


Figure 4.48: Results of joint tensile strength (top) and permeability (bottom) estimation in 3D model with projection approach.

An important aspect that was not considered in this study is the presence, initiation and propagation of fractures in the rock during the hydraulic fracturing process. In general, microseismic events can carry important information about the location and geometrical attributes of the fractures, which can be exploited for fracture model calibration purposes.

4.4.8 Application of SSBRC to KTB Site

To show the application of SSBRC to real field data, the monitored microseismic cloud in the German Continental Deep Drilling site (KTB) in fluid injection experiment in 2000 is utilized. For simplicity the 2D configuration of the monitored MEQ data is chosen as the observed data (shown in Figure 4.49). Our purpose is to estimate the permeability distribution from this 2D MEQ cloud (integration data). As the forward model, we employed the pore pressure diffusion forward model and we assumed criticality distribution is known. We discretized the field to 100x100 grids. We chose 6 integration steps and used 100 realizations (some assumptions made about the spatial correlation of the permeability distribution; variogram model and its parameters). SSBRC along with the reduced projection method is used to avoid ensemble collapse. The estimation results in terms of ensemble mean and ensemble Std maps are given in Figure 4.50. Ensemble spread is also shown in Figure 4.51 showing that the ensemble spread is preserved because of using reduced order projection method.

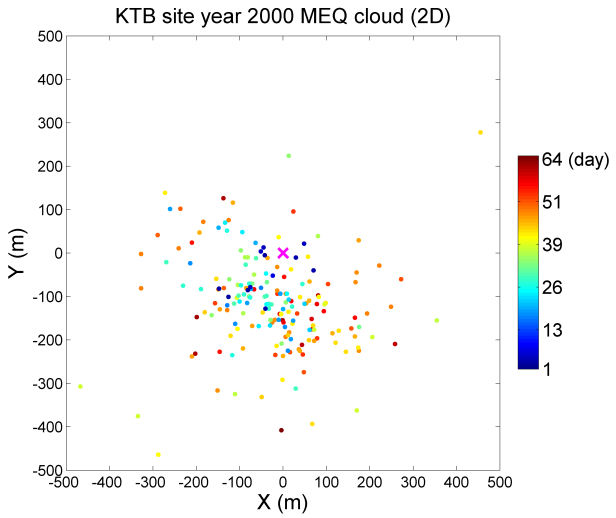
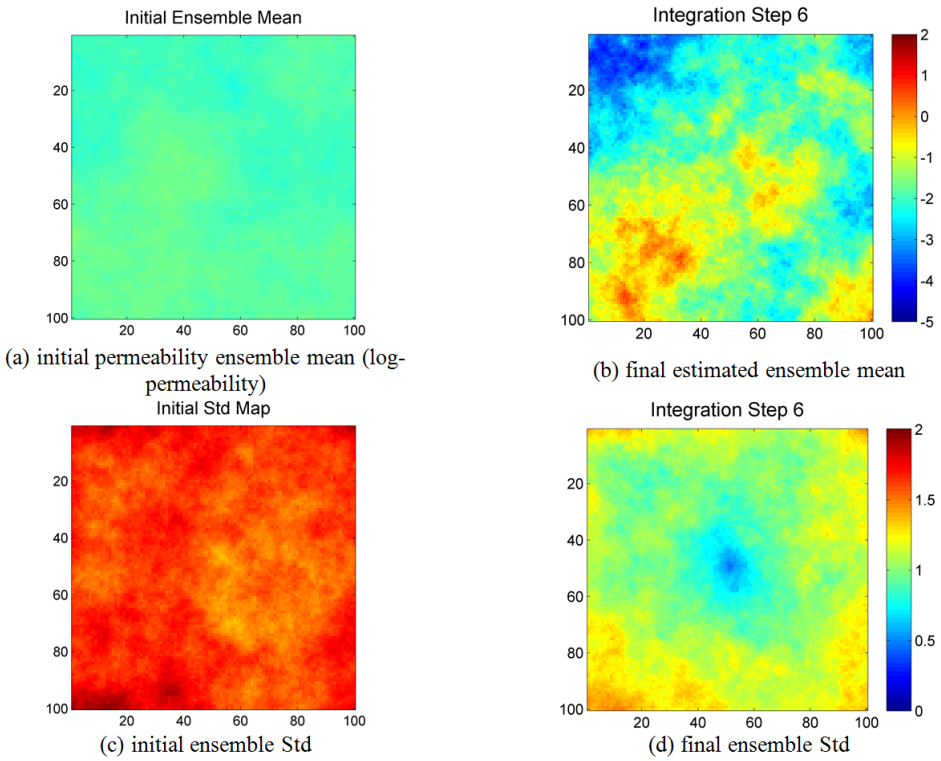


Figure 4.49: The 2D configuration of KTB microseismic cloud which is used as the integration data here (the pink cross is the injection well) (events chosen from a depth interval of the 3D MEQ cloud)



(a) initial permeability ensemble mean (log-permeability)

(b) final estimated ensemble mean

(c) initial ensemble Std

(d) final ensemble Std

Figure 4.50: Estimating permeability distribution form MEQ cloud in Figure 4.49

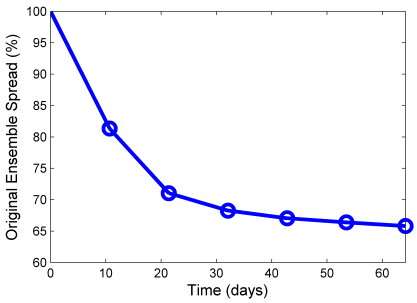


Figure 4.51: Ensemble spread through integration steps

5. APPLICATION AND EXTENSION TO HYDRAULICALLY FRACTURED RESERVOIRS

5.1 Application of SSBRC to Horizontal Well with Multistage Hydraulic Fractures

The application of SSBRC for reservoir and hydraulic fractures characterization by MEQ data integration is shown in this section with presenting a set of different numerical examples.

5.1.1 Estimating Permeability Distribution in Hydraulically Fractured Reservoir

In this section we aim at illustrating the utility of SSBRC to infer reservoir properties from MEQ cloud in hydraulically fractured reservoirs. To be able to realistically model the treatment phase in a reservoir with horizontal well and multistage hydraulic fractures, a reservoir simulator with coupled fluid flow and geomechanics effects, fracture propagation model, rock damage and permeability model capabilities is required. In this experiment to present a preliminary example and show the applicability of SSBRC we set up an unsophisticated reservoir model based on some assumptions to be able to utilize the available reservoir model tools for the mentioned purpose. The experiment is based on the pore pressure diffusion assumption in Section 2.1 and no fracture propagation model is utilized which implies that the hydraulic fractures are created in the very first moments of fluid injection and the induced MEQ events are triggered because of change in the stress and pore pressure distribution around the fracture after its creation. To model this process Eclipse [77] is utilized and as the failure criterion criticality or critical pore pressure is assigned to each grid block. The reservoir, well and hydraulic fractures configuration

is schematically shown in Figure 5.1.

The specifications of the reservoir model are as follows:

- A horizontal well with 8 transverse hydraulic fractures (regularly spaced).
- Horizontal well length is equal to reservoir length.
- Stimulation period is 48 (hr).
- Three phase fluid system (oil, water and gas).
- There are three periods of fluid flow in hydraulically fractured reservoir: treatment (hydraulic stimulation), flow back and production. In here we simulate the treatment period.
- The injection pressure is specified (water is injected).
- *Shapiro* assumptions and model setup are followed [26].
- Eclipse (E100) is used for reservoir simulation [77].
- Local grid refinement is used in the grids with hydraulic fracture.
- Spatially random criticality is used (taken from a Normal or Gaussian PDF).
- Permeability distribution is the target parameter to estimate.

Reservoir configuration	$100 \times 30 \times 30$
Grid size	$30(\text{ft}) \times 30(\text{ft}) \times 5(\text{ft})$
Reservoir size	$3000(\text{ft}) \times 900(\text{ft}) \times 150(\text{ft})$
Initial reservoir pressure	4000 (psi)
Reservoir depth	6000 (ft)
Fracture height	65 (ft)
Number of fractures	8
Fracture half-length	150 (ft)
Fracture width	2 (in)
Fracture permeability	10000 (md)
Injection pressure	10000 (psi)
Dimensionless fracture conductivity C_D	0.28

Table 5.1: Reservoir model parameter values.

To realistically model fluid flow in and around the hydraulic fractures (because of the high contrast of permeability in fracture and matrix) the local grid refinement is performed in the grid where the hydraulic fracture is present. The size of the refined grids (refinement perpendicular to fracture plane; in x direction) are smaller in the middle of the grid (the refined grid at the middle is the fracture grid which has the high permeability) and are larger in the two sides as shown in Figure 5.3. The local grid refinement configuration in each direction is as follows:

- x direction: 1 grid to 11 refine grids.
- y direction: 1 grid to 2 refine grids.
- z direction: 1 grid to 1 refine grids (no refinement).

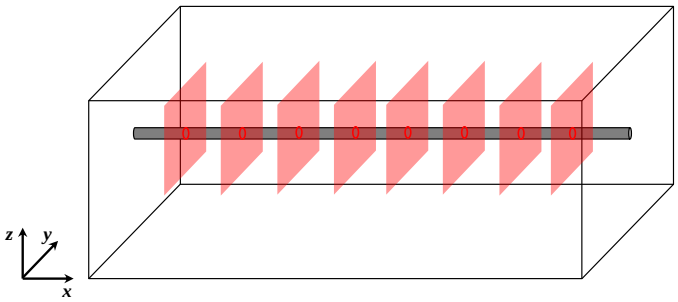


Figure 5.1: The configuration of the multistage hydraulic fractures and the horizontal well in the reservoir.

The top view of the reservoir including the permeability (in md) distribution and the location of the stages is presented in Figure 5.2.

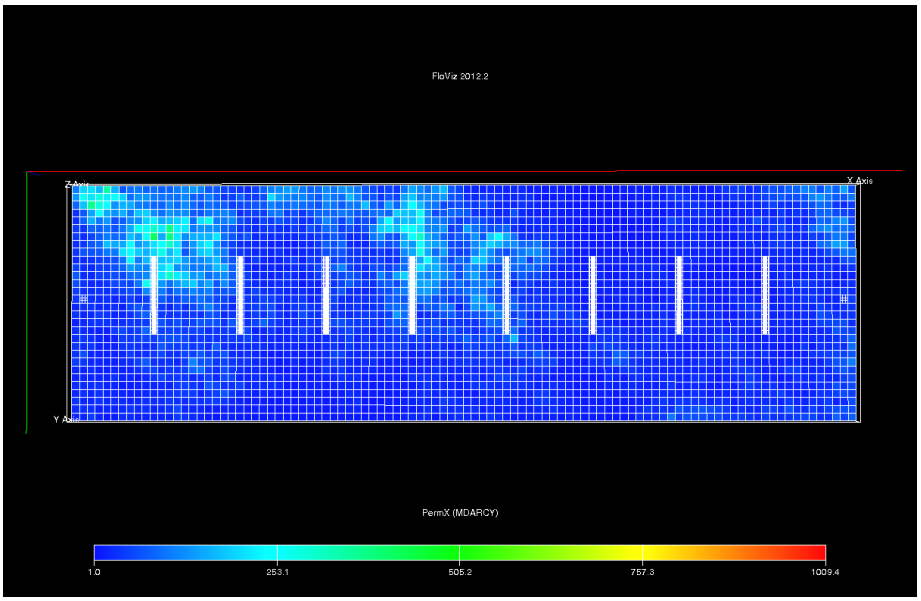


Figure 5.2: True model configuration (top view), the location of the fractures and the reservoir permeability distribution.

Figure 5.3 shows the configuration and the refinement of two stages.

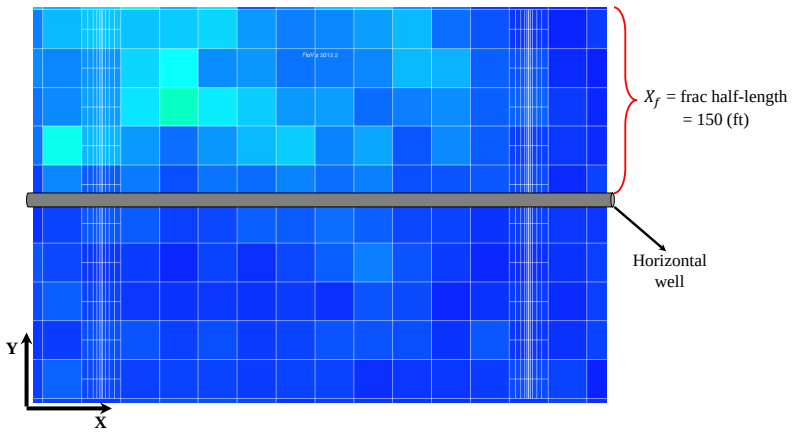
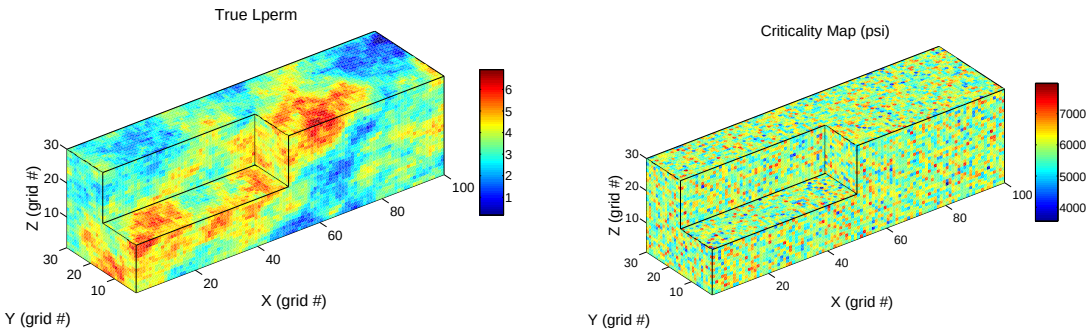


Figure 5.3: The top view of the two fracture stages to illustrate the refinement configuration.

The true permeability distribution as well as the criticality map are presented in Figure 5.4.



(a) true log-permeability distribution

(b) criticality distribution

Figure 5.4: The distributions of true log-permeability (implemented hydraulic fractures as in Figure 5.1 are not shown here) and criticality. The criticality or critical pore pressure values come from a Gaussian random distribution with mean and standard deviation of 5850 (psi) and 550 (psi), respectively.

The pore pressure distribution in the reservoir during the hydraulic injection is presented in Figure 5.5. As expected higher pore pressure occurs in the vicinity of

the fracture stages since the pressure front travels faster in these regions. It is clear in Figure 5.5d that the pore pressure distribution reflects the features (low/high values) of permeability distribution of Figure 5.4a.

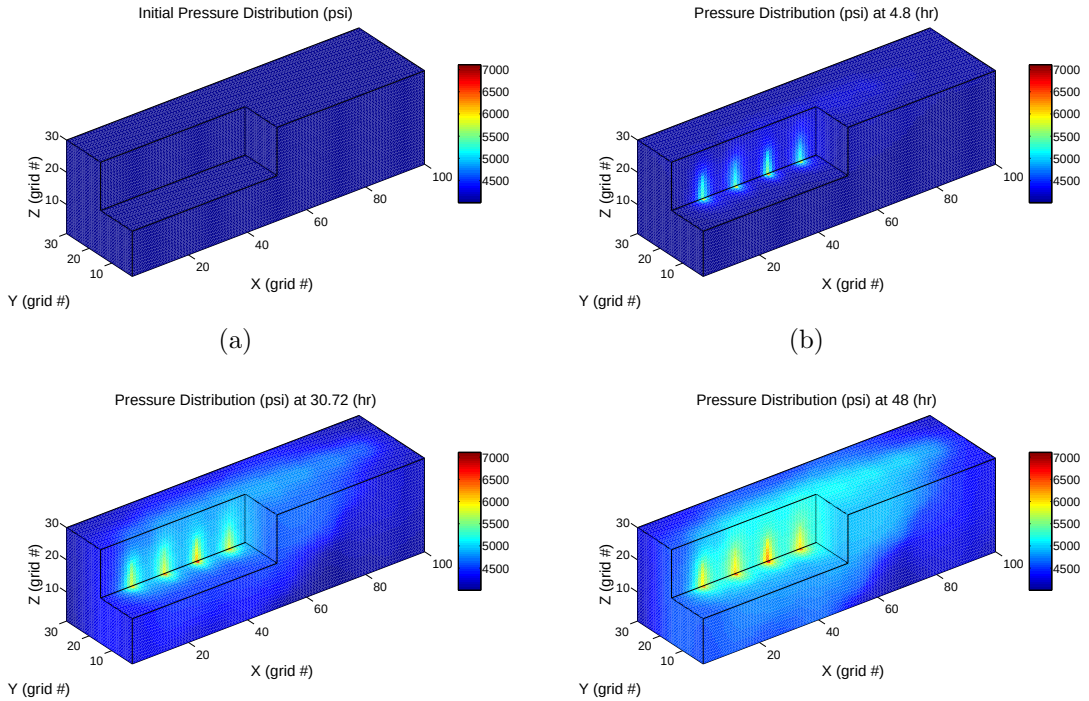
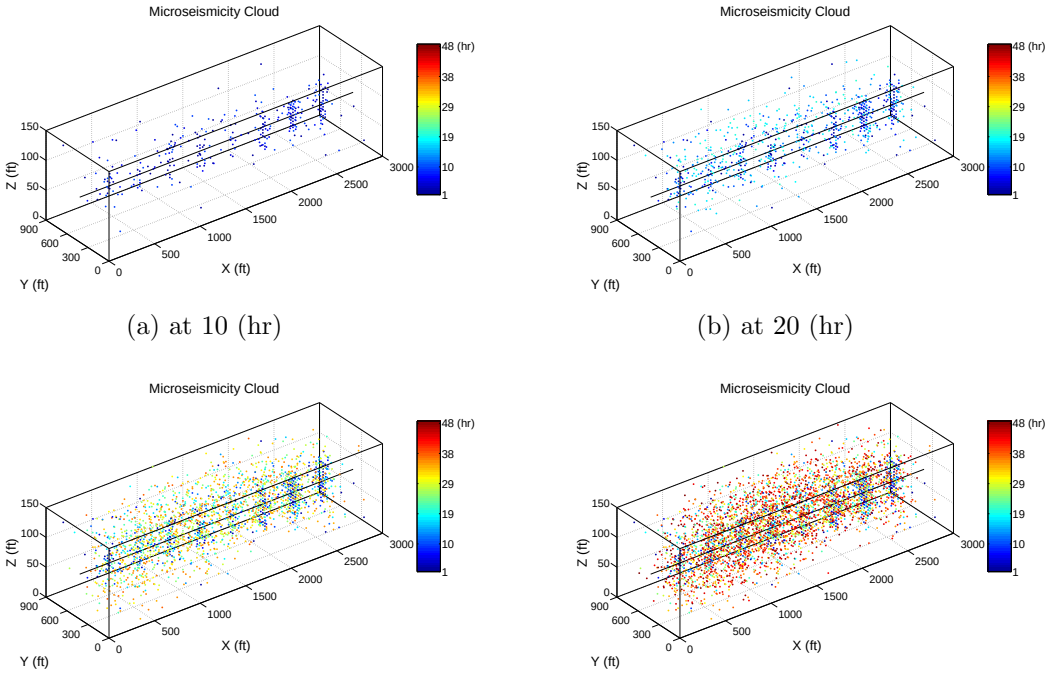


Figure 5.5: Evolution of the pore pressure distribution in the reservoir model during the hydraulic stimulation.

Figure 5.6 shows the evolution of the microseismicity cloud throughout the hydraulic injection. As we expected, the high density of the MEQ events along the hydraulic fracture stages (resulted from the high pore pressure in these regions) is clearly distinguishable in Figure 5.6a,b.

In this example to estimate permeability distribution from the spatio-temporal MEQ cloud we apply SSBRC approach along with the reduced-order projection



(a) at 10 (hr)

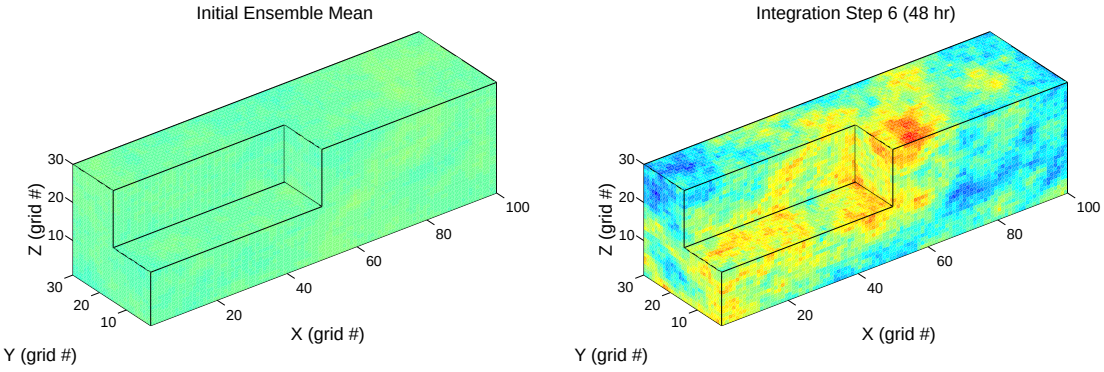
(b) at 20 (hr)

(c) at 38 (hr)

(d) final MEQ cloud at 50 (hr)

Figure 5.6: Evolution of the MEQ cloud in the reservoir model during the hydraulic stimulation.

method to avoid ensemble collapse. In this experiment 100 permeability realizations is used as always and n_{trunc} and σ_h (the reduced-order projection parameters) are 100 and 25 %, respectively. The estimation result of SSBRC in terms of estimated ensemble mean is shown in Figure 5.7. The final estimated ensemble mean in Figure 5.7d shows acceptable similarity with the true permeability map in Figure 5.4a.



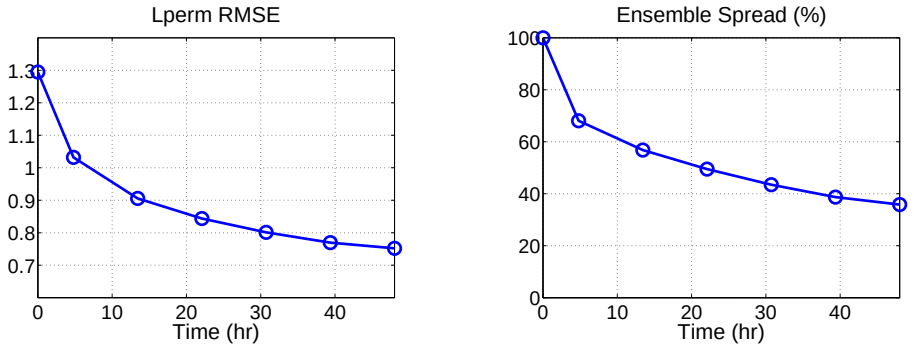
(a) (b)

Figure 5.7: Initial and final estimated ensemble mean.

The performance measures of SSBRC method in terms of estimation RMSE and ensemble spread are presented in Figure 5.8. The continuous decrease of RMSE in Figure 5.8a confirms the promising estimation capability of SSBRC and the final ensemble spread of almost 40 % in Figure 5.8b illustrates that the reduced-order projection method helps significantly in preserving the ensemble spread throughout the integration procedure.

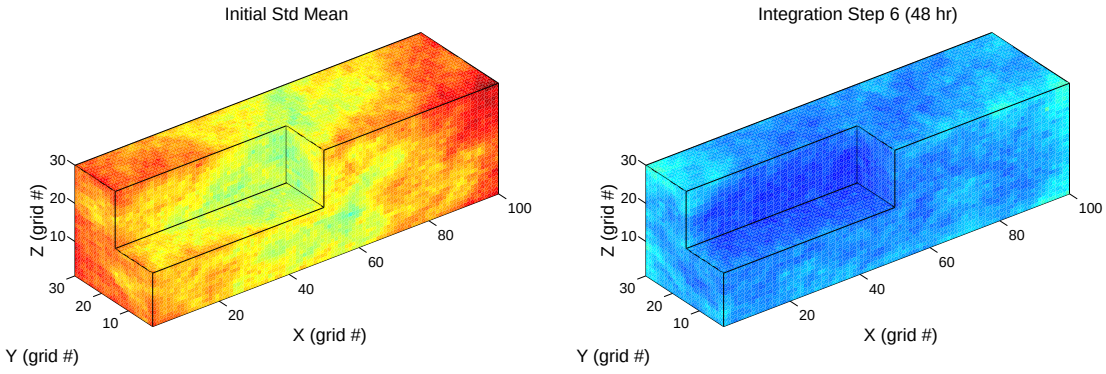
Figure 5.9 shows the initial and final standard deviation map which again demonstrates the promising final ensemble spread because of utilizing the projection approach.

This experiment clearly shows the utility of SSBRC to estimate rock properties



(a) (b)

Figure 5.8: Evolution of estimation RMSE and ensemble spread throughout the integration steps.



(a) (b)

Figure 5.9: Evolution of the Std map throughout the integration steps.

in the hydraulically fractured reservoirs (multistage fractures) with horizontal well. While the presented experiment is based on the assumption of pore pressure relaxation hypothesis and it lacks a fracture propagation model, by relaxing these assumptions and making the reservoir model more physically realistic, SSBRC applicability conditions will not change and it still will have promising utility in inferring reservoir properties from MEQ data. This experiment is presented as a preliminary example of applying SSBRC to a hydraulic fracture case with horizontal well that can be extended to any complex reservoir model with planar or complex fracture network and natural fractures. SSBRC can also be extended to be used for characterizing hydraulic fractures (estimating hydraulic fracture geometry and conductivity) and inferring the structure of the induced fracture network. It is clear to do so one needs to acquire more involved reservoir simulators with geomechanical effects, natural fracture handling capability, fracture propagation model and the interaction between natural and induced fractures.

5.1.2 Estimating Permeability Distribution with the Assumption of Unknown Hydraulic Fractures Characteristics

In this section we present the results of employing SSBRC for characterization of hydraulically fractured reservoir. In the previous section an example was presented for estimating matrix permeability estimation from MEQ data while we assumed that all the hydraulic fracture characteristics are known. In this section we present the same example but we assume that the fractures are also unknown and we try to estimate the reservoir permeability distribution to see if SSBRC can infer fracture locations by assigning high permeability to fractured regions of the reservoir.

The true model specifications are presented in the previous example. There are 8 equally spaced hydraulic fractures with the same geometries and conductivity, shown

in Figure 5.1. The previous example showed the results of matrix permeability distribution estimation by integration MEQ data in a hydraulically fractured reservoir by the assumption of known hydraulic fractures. In that example the hydraulic fractures locations and characteristics were known and they were already implemented in all the realizations and the matrix permeability distribution was the only unknown. While in the new experiment we do not implement hydraulic fractures in the initial realizations in advance. We assume that the hydraulic fractures and matrix permeability distribution are unknown and we integrate MEQ data to infer permeability distribution and consequently the hydraulic fractures (their locations and permeabilities). So the purpose of this experiment is to investigate if the SSBRC approach can reveal the hydraulic fractures location and permeability through MEQ integration. In case of successful application the signature of hydraulic fractures should be seen as extremely high permeability (narrow rectangular) regions around the well (represented by high permeability grid blocks).

As one of the major pitfalls of EnKF, it tends to smooth out the distribution of estimated parameter and the dynamic range of the estimated parameter will always be underestimated. Therefore inferring extremely high permeability regions as hydraulic fractures when dynamic range of permeability is much lower is not feasible in EnKF context.

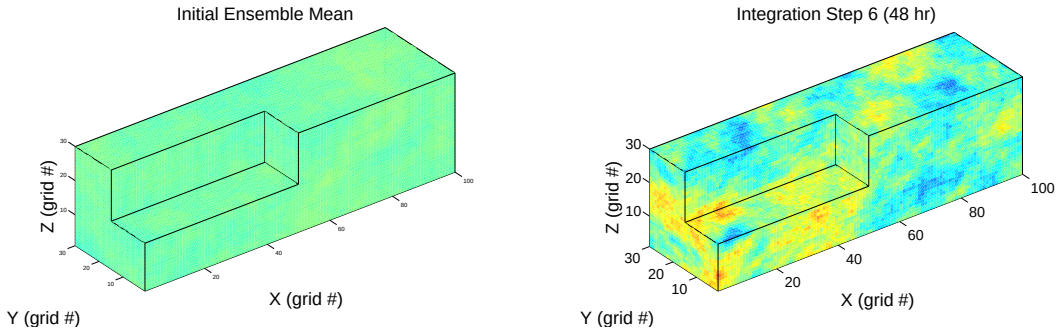
The hydraulic fractures are explicitly implemented in the true model through local grid refinement (explained in previous example) and highly increasing permeability values at fracture grid blocks while within EnKF procedure in the permeability realizations the hydraulic fractures are not implemented (as high permeability regions in permeability distribution) and it is expected that MEQ integration would detect the hydraulic fractures regions.

In Figure 5.4 (true permeability and criticality distribution) the hydraulic

fractures are not shown while they are implemented as high permeability narrow rectangular regions around the wellbore however for the ensemble of permeability distributions hydraulic fractures are not implemented.

The stimulation time is 48 (hr). The pore pressure distribution during the stimulation is shown in Figure 5.5. It is clear that the pore pressure goes high at the location of the stages.

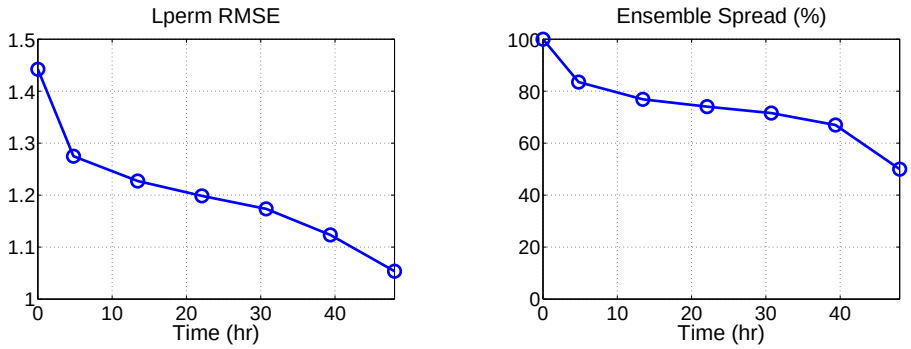
The evolution of the MEQ cloud during stimulation process is shown in Figure 5.6. We can see that the MEQ events are mostly gathered around the fracture stage locations.



(a) (b)

Figure 5.10: Evolution of ensemble mean throughout the integration steps.

Figure 5.10 shows the initial and final estimated ensemble mean. The final estimated permeability map is not similar to the true permeability distribution and also there is no sign of high permeability regions (fracture stages) which implies the unsuccessful application of SSBRC to estimate reservoir permeability distribution (matrix and fracture in the same time) from MEQ data in this setup (using pore pressure relaxation assumption).



(a) (b)

Figure 5.11: Evolution of estimation RMSE and ensemble spread throughout the integration steps.

While in Figure 5.11a, estimation RMSE shows continuous reduction, the estimated map is not close to the true map and the fracture locations are not detected.

Overall if we assume that the fractures characteristics as well as matrix permeability distribution are unknown and try to estimate hydraulic fractures through permeability estimation (finding high permeability regions around the well as stages), SSBRC does not show successful performance.

5.1.3 Estimating Hydraulic Fracture Characteristics

In this section we employ SSBRC to infer hydraulic fractures geometries (fracture half-length and height) from MEQ data while we assume the matrix permeability is known. We applied SSBRC to a wide range of matrix permeability such as relatively high matrix permeability (could be resulted from preexisting natural fractures and rock weaknesses) which results in a very scattered MEQ cloud (representing complex fracture network) and also extremely low matrix permeability (to represent unconventional shale resources) which leads into grouped, elongated and well-shaped MEQ cloud (representing planar fractures). We also performed examples with medium matrix permeability range (to cover a wide spectrum of matrix permeability

values) to demonstrate the utility of the method for a wide range of matrix permeability. We aim to verify the successful performance of SSBRC in hydraulic fracture characterization for any possible range of matrix permeability which means any type of MEQ cloud shape ranging from well-shaped clouds (placed along the planar hydraulic fractures) to very scattered MEQ cloud (in case of complex fracture network).

The well known approach of fracture characterization from MEQ cloud is recognized as hydraulic fracture mapping [113, 114, 115] to find the planar fracture characteristics (geometries like half-length, height and orientation) from microseismic cloud which is a direct or explicit solution to this problem. This method works by fitting line segments and planes to the scattered events in each stage. Complex fracture network also can be inferred by hydraulic fracture mapping methods where a set of parallel and perpendicular line segments which construct the complex fracture network are fitted to the MEQ cloud. Here we propose to infer hydraulic fractures characteristics by inversion or inverse modeling from MEQ data using SSBRC method. Hydraulic fracture mapping methods are straightforward to use for elongated and well-shaped MEQ clouds but it is problematic for scattered MEQ clouds.

Fracture Stage #	True Fracture Half-length (ft)	True Fracture Height (ft)
1	180	120
2	350	90
3	260	100
4	100	92
5	240	100
6	180	40
7	75	130
8	410	140

Table 5.2: True geometries (half-length and height) of 8 hydraulic fracture stages.

As shown in Figure 5.12, we assume 8 equally spaced hydraulic fractures with different geometries (presented in Table 5.2). The fracture permeability is also set at 10000 (md) as in the previous example.

In the performed examples we estimate hydraulic fractures geometries by integrating MEQ data using SSBRC method while assuming known high and low matrix permeability. There are 8 hydraulic fractures and each with two geometry parameters (half-length and height) therefore we estimate 16 parameters and we do not assume any correlation between parameters. The ensemble size is 100 and Uniform PDF is used to generate initial realizations. Based on the reservoir configuration and size the half-length samples are coming from interval $[5, 450]$ (ft) and height samples are generated from interval $[5, 150]$ (ft).

To summarize the matrix permeability ranges of all four experiments, the important statistics and ranges of matrix permeabilities are given in Table 5.3. As

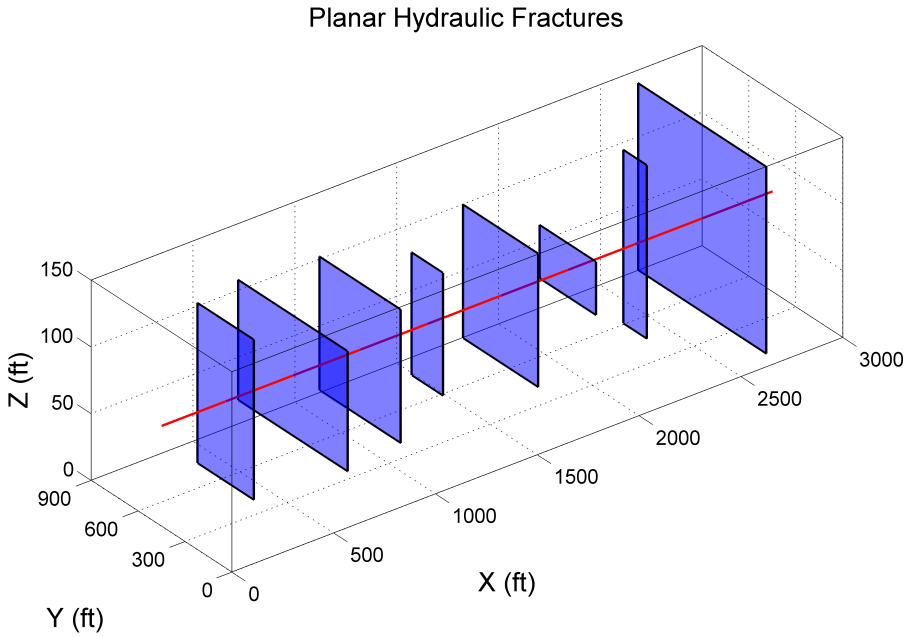


Figure 5.12: The true configuration and geometries (half-length and height) of the multistage hydraulic fractures (8 stages) and the horizontal well in the reservoir.

presented in this table, the chosen matrix permeability ranges well cover the possible reservoir permeability values in wide range of applications. It should be noted again that the matrix permeability statistical distribution is Log-normal. Detailed estimation results of only one of the experiments (Medium # 2) are reported here.

Matrix perm. range	Mean	Median	Mode	Min.	Max.	90% quantile
Low	6.52×10^{-4}	4.44×10^{-4}	4.76×10^{-4}	1.98×10^{-5}	0.0169	0.0014
Medium #1	0.026	0.018	0.0193	8.02×10^{-4}	0.7	0.05
Medium #2	1.06	0.7271	0.7797	0.03	27	2.2
High	39	27	28	1.1873	1009	81

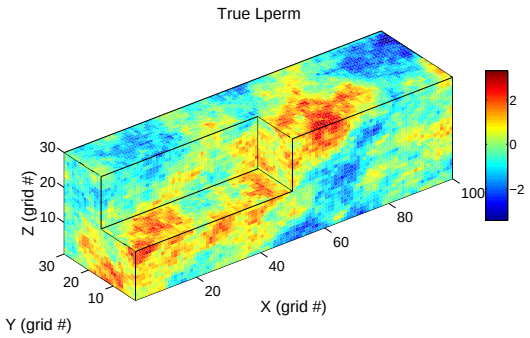
Table 5.3: The matrix permeability statistics of all four experiments (all values are in md).

The MEQ clouds even for the very low matrix permeability case (clear sign of hydraulic fractures in MEQ cloud) might seem relatively noisy. As explained before the SBRC approach is followed here and the critical pore pressure or criticality map is assumed spatially random therefore the resulting MEQ cloud has some signatures of the random criticality map.

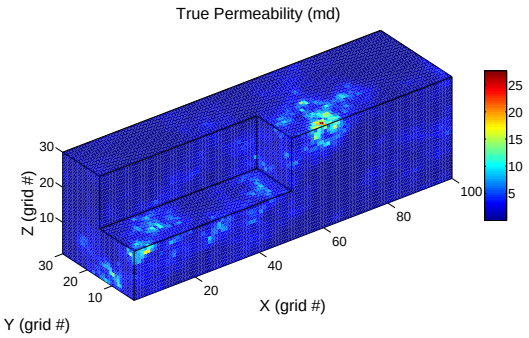
Here we consider a medium range matrix permeability (Medium # 2) distribution with the previously defined hydraulic fracture setup. The matrix permeability distribution, final pore pressure and MEQ cloud are shown in Figure 5.13. As shown in this Figure 5.13d, the MEQ cloud hardly has sign of the planar fractures and shows very scattered behavior however it is not as scattered as the high matrix permeability case.

The estimation results are presented in Figure 5.14. As seen in Figure 5.14a, the fracture half-length ensemble tend to converge to the true values during integration steps. Figure 5.14b also shows that the fracture height ensembles become closer to true height values through integration steps.

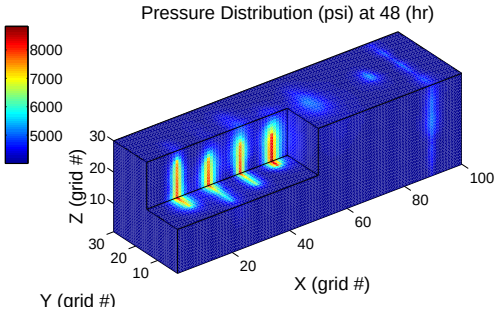
Figure 5.15 shows performance of SSBRC in estimating fracture half-length



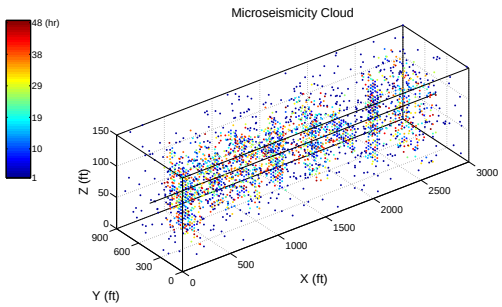
(a) matrix log-permeability distribution



(b) matrix permeability distribution

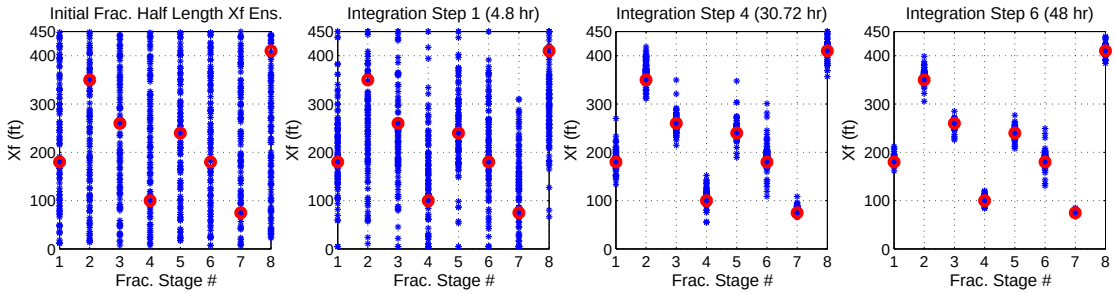


(c) Final pore pressure

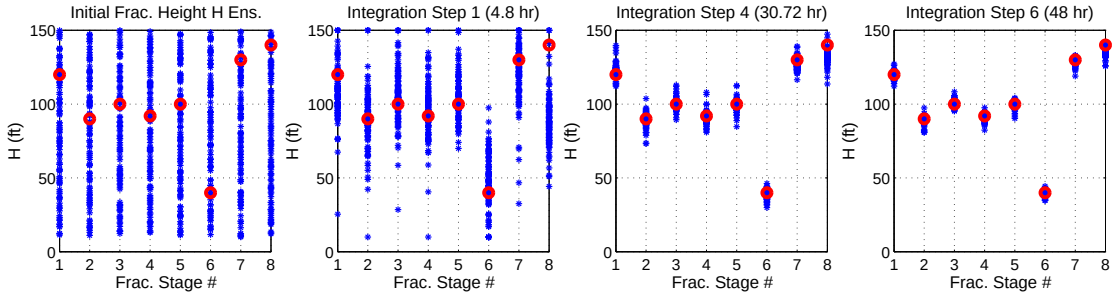


(d) Final MEQ cloud

Figure 5.13: The distribution of matrix permeability (implemented hydraulic fractures as in Figure 5.12 are not shown here). Both Log-permeability and permeability (md) values are shown in this figure. Final pore pressure distribution and MEQ cloud are also shown.



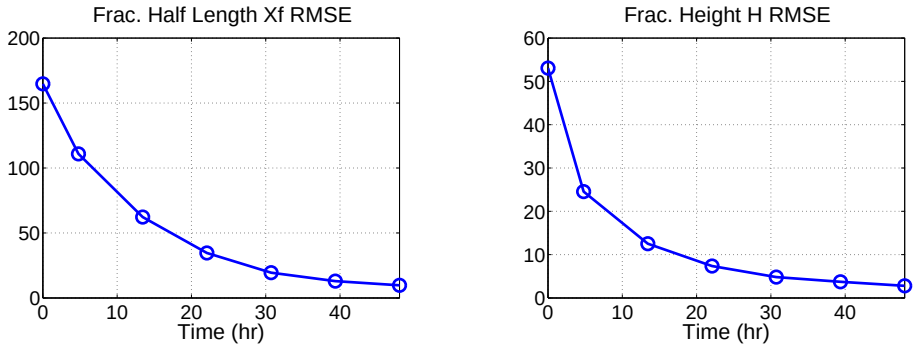
(a) hydraulic fracture half-length estimation



(b) hydraulic fracture height estimation

Figure 5.14: Evolution of hydraulic fractures geometries (half-length and height) ensembles through integration steps. Blue stars depict the realizations and red circles show the true hydraulic fracture geometry value.

and height in terms of estimation RMSE. Continuous reduction of both estimated ensemble half-length and height confirms the promising performance of SSBRC in inferring fracture geometries.



(a) (b)

Figure 5.15: Evolution of estimation RMSE of half-length and height ensembles throughout the integration steps.

The hydraulic fractures characteristics estimation results with SSBRC for different matrix permeabilities are summarized in Table 5.4.

Matrix perm. range	Xf RMSE (ft)	H RMSE (ft)
Low	9.3	3.39
Medium #1	7.17	2.88
Medium #2	9.73	2.79
High	42.49	6.13

Table 5.4: The estimation results of all four experiments (initial Xf RMSE = 164.74 ft, initial H RMSE = 53.05 ft).

5.2 Assessment of Welltest Techniques for Hydraulic Fracture Characterization and its Integration with MEQ Data

Well test can be utilized to help estimating effective hydraulic fracture length and conductivity, for low-complexity fractures, and effective matrix surface area for high-complexity (complex fracture networks) situations [116]. Fracture type curves can be used to determine the flow regimes in the hydraulically fractured reservoirs (based on the dimensionless fracture conductivity and drainage area) e.g. linear or bilinear flow. Although some hydraulic fracture information as well as drainage volumes can be obtained from well test, it is impossible to assess fracture or SRV geometry. Microseismic fracture mapping as the existing technology as well as the proposed SSBRC method can be used to provide critical information regarding fracture geometry (height, length and azimuth) and complexity, and also can be used to provide SRV limits and shape. The integration of well test (RTA) and microseismic fracture mapping can provide a promising framework to efficiently characterize hydraulic fractures and reservoir as well as ultimately plan field development and optimize hydraulic fracturing strategy.

Numerical experiments are done and well test analysis is applied to their results to illustrate the outcome of PTA application for reservoir characterization. The performed examples (to illustrate application of PTA) include single hydraulic fracture fully penetrating the formation (fracture height is equal to the formation thickness). Fracture half-length is assumed to be 200 (ft). We performed two numerical examples with different formation permeabilities of 600 (nd) and 0.02 (md). Our case is effectively acting as infinite conductivity fracture. (fracture permeability $k_f = 10000$ md, dimensionless fracture conductivities C_D are also reported next). The linear flow regime (half slope section in pressure derivative curve) is analyzed

as following:

$$\Delta p' = \frac{1}{2} m_{lf} \sqrt{\Delta t} \Rightarrow m_{lf} = \frac{2\Delta p'}{\sqrt{\Delta t}} \tag{5.1}$$

$$x_f \sqrt{k} = \left(\frac{4.064 \Delta q B}{m_{lf} h} \right) \left(\frac{\mu}{\phi c_t} \right)^{\frac{1}{2}} \tag{5.2}$$

where x_f is hydraulic fracture half-length (ft) and k is matrix or formation permeability (md). m_{lf} can be obtained by the given formula using any point on the half slope section of the pressure derivative curve. In the performed examples (drawdown test) the linear flow regime is clear in the log-log plot of pressure change and pressure derivative versus time. Using PTA specifically for linear flow regime in hydraulic fracturing treatment, we can estimate $x_f \sqrt{k}$ (fracture half-length and formation permeability provided which data is available).

Based on the given formulation in the half slope part of diagnostic plot of Figure 5.16, the estimated $x_f \sqrt{k}$ is 5.73 (the true value is 4.90) and if we assume the formation permeability is known the estimated fracture half-length is 234 (ft).

Based on the given formulation in the half slope part of diagnostic plot of Figure 5.17, the estimated $x_f \sqrt{k}$ is 31.6 (the true value is 28.28) and if we assume the formation permeability is known the estimated fracture half-length is 223 (ft).

Well testing is a model based fracture diagnostic method that requires accurate permeability and reservoir pressure estimates. Microseismic mapping techniques (microseismic data integration; by microseismic data here we mean the microseismic events location and time obtained after microseismic processing and locating the MEQ source) can determine hydraulic fracture length, height, azimuth, dip and asymmetry (generally hydraulic fractures geometry and configuration) and also in case of complex fracture network can determine SRV extent, shape and volume. While well testing techniques may determine (based on the available flow regime

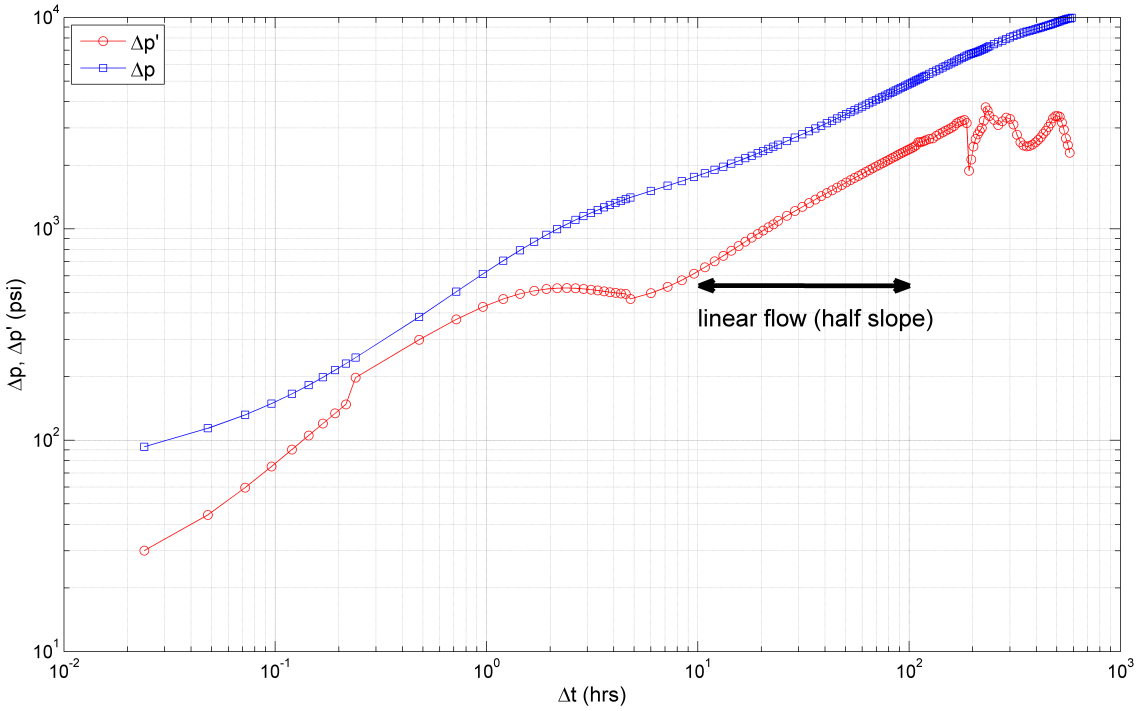


Figure 5.16: The log-log plot of pressure change and pressure derivative versus time (diagnostic plot) and the associated linear flow part (matrix permeability = 600 nd, C_D = 13889).

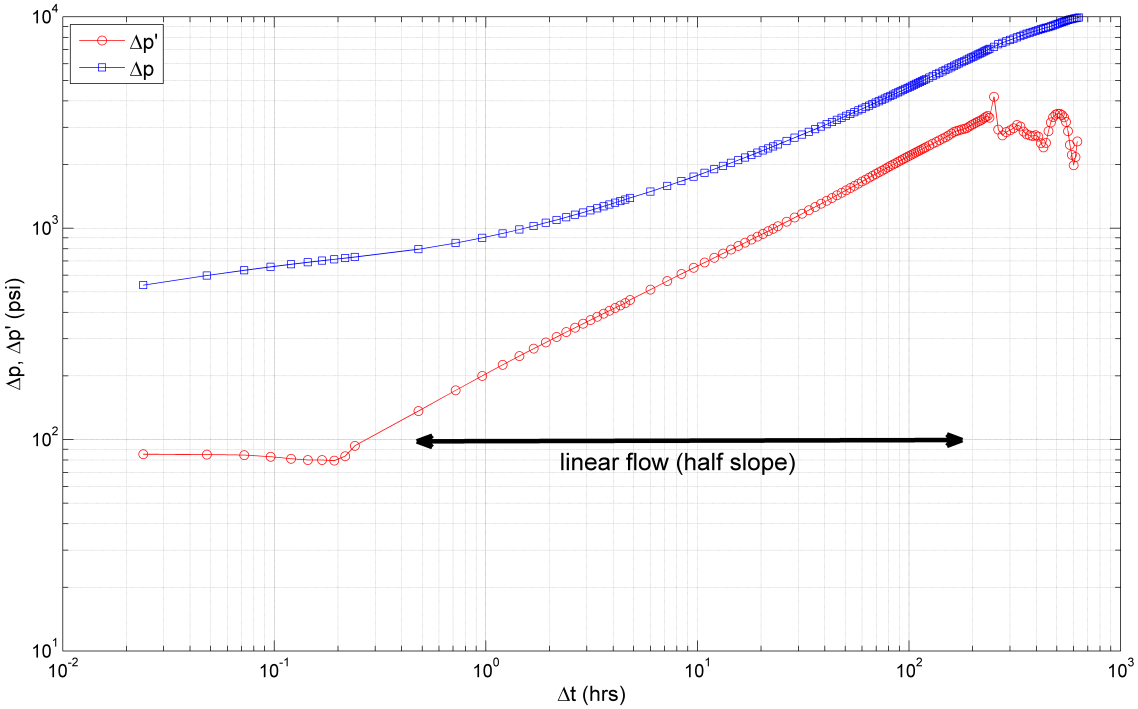


Figure 5.17: The log-log plot of pressure change and pressure derivative versus time (diagnostic plot) and the associated linear flow part (matrix permeability = 0.02 md, $C_D = 417$).

in acquired data) fracture length, width and conductivity [117]. We need to obtain matrix permeability independently from other sources of information such as pre-fracture well test [118, 119]. Well testing technique is limited to estimating a single stage fracture half-length (fully penetrating) while microseismic data integration (e.g. SSBRC) can infer half-length and height (and orientation; in general fracture geometry) of each fracture in multi-stage hydraulic fracturing.

As presented in detail in previous section (for different ranges of matrix permeability), SSBRC is able to estimate geometries (half-length and height) of multi-stage hydrauilic fractures. It is also shown that if the hydraulic fracture characteristics are available (from another source of information e.g. tiltmeter), MEQ data integration through SSBRC can determine the distributed matrix permeability (estimating the heterogeneous matrix permeability distribution).

5.3 Discussion and Recommendations for Real-time Application of SSBRC

The utility of SSBRC in characterizing the reservoir based on the MEQ observations will be of great advantage if it can be applied to a reservoir in the real-time fashion. We accomplished applying SSBRC to different reservoir simulators and demonstrated its capability to estimate hydraulic and geomechanic reservoir distributed parameters from MEQ data. The pore pressure diffusion model is not very physically involved which leads to less computational complexity and makes it a much faster option. On the other hand the geomechanical forward model is a fully coupled reservoir model and is very computationally expensive. In performing inversion with SSBRC by geomechanical reservoir simulator as the forward model it might seem that the total simulation time is not feasible in some cases. For instance if the SSBRC simulation time is comparable to the real field stimulation

time then the inversion procedure can hardly be considered as real-time. So to make the whole SSBRC procedure faster we recommend to implement some variations of the EnKF-based integration approach and we also recommend to utilize faster reservoir simulators. To utilize the fast pore pressure diffusion model for achieving to a real-time SSBRC procedure we need to incorporate or translate the reservoir geomechanical properties in the criticality or critical pore pressure [120].

The available geomechanical forward reservoir simulator is in the executable form and to speed up the SSBRC process using this reservoir model all the improvements need to be done in the SSBRC inversion algorithm. Having access to the geomechanical model source code we can though make the simulator faster by:

- Removing some of the physical effects e.g. removing temperature effect, assumption of isothermal process.
- Using coarser finite element mesh (less nodes) in geomechanical model.
- Performing the process in iterative or uncoupled fashion.

The improvements that need to be done in regards to the SSBRC algorithm in the way to reduce the required number of forward simulations can be as follows:

- Reducing the number of integration steps.
- Reducing the ensemble size.
- Utilizing more CPU in the parallel inversion process.
- Using sequential EnKF instead of EnKF with confirmation step.
- A combination of these approaches.

Test #	Ensemble size N_e	# of integration steps	CPU run time (hr)	Final RMSE (%)	Final Spread (%) $Sp(\mathbf{m})$
1 (Reference)	100	6	42.9	54.72	54.34
2	50	6	23.3	60.53	55.65
3	25	6	12.06	54.42	44.49
4	100	3	26.1	63.92	64.79
5	50	3	13.8	66.99	66.37
6	25	3	6.5	61.57	62.15
7	100	1	13.2	77.82	84.70
8	50	1	7.0	80.19	87.19
9	25	1	3.7	79.53	88.05

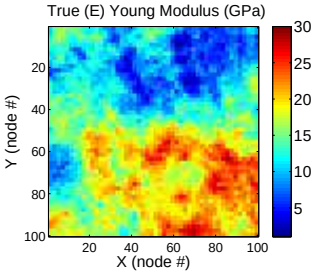
Table 5.5: Performance comparison of SSBRC with different ensemble sizes N_e, and number of integration steps.

For the experiment that its true model is shown in Figure 4.27, the performance comparison of SSBRC with different number of realizations and integration steps is presented in Table 5.5. The results clearly show that reducing the number of integration steps and ensemble size significantly reduces the computational load of the SSBRC procedure. As presented in Table 5.5, with a fixed number of integration steps, ensemble size does not have a significant effect on RMSE and ensemble spread values. In Table 5.5 since we are using reservoir property ensemble with different sizes, to be consistent we present the percentage of the final RMSE to the initial RMSE. It should be noted that the real-field stimulation period is 20 (hr) as indicated in the MEQ cloud in Figure 4.27. In all these experiments to avoid ensemble collapse, projection approach is utilized and in each test n_{trunc} and σ_h are equal to the ensemble

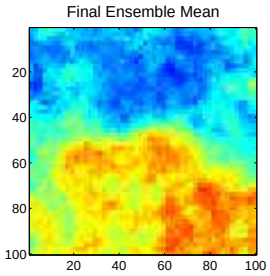
size of the test and 25 %, respectively. To run the experiments, a same machine with 4 CPUs in parallel was utilized. The final estimated ensemble mean of all the performed tests are shown in Figure 5.18. In terms of final estimated ensemble mean, the test with 3 integration steps and 25 realizations (Test # 6 in Table 5.5) resulted in the promising estimated map in Figure 5.18g comparing to true map in Figure 5.18a as well as the reference test result in Figure 5.18b while it is almost 7 times faster than the reference test.

The pore pressure diffusion model and geomechanical model seem to be at the two extremes of the physics-based reservoir simulations. The pore pressure diffusion model (using a commercial reservoir simulator Eclipse [77]) only has the single phase fluid flow effect and it loosely includes geomechanical effects and failure criteria by incorporating criticality or critical pore pressure (low-order physics-based model). In contrast the geomechanical reservoir model [27, 28, 29, 30] (which we have access to its executable file) handles the fully coupled effects of fluid flow, temperature and rock deformation as well as geomechanical failure criteria and damage and permeability model (high-order physics-based model) that makes this model very slow to run. To make use of both advantages of this two modeling approaches we recommend to utilize a reservoir simulator with both features i.e. a fast commercial reservoir simulator with uncoupled fluid flow and geomechanics effects. We recommend and introduce a novel approach to combine Eclipse (using its uncoupled or iterative geomechanics module) with MATLAB (to implement the geomechanical failure criteria, damage and permeability model) to build a rather computationally inexpensive reservoir model which though has the geomechanical effects. The overall workflow of the proposed approach is presented in Figure 5.19.

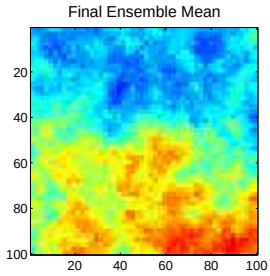
Therefore we can use Eclipse with its geomechanics module along with MATLAB to develop a faster geomechanics-based reservoir simulator. We first reproduce



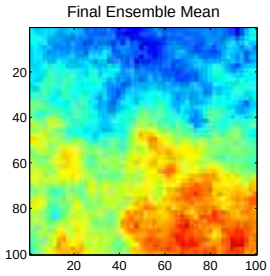
(a)



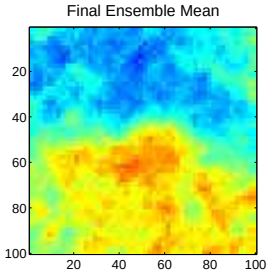
(b) int. step = 6, $N_e = 100$



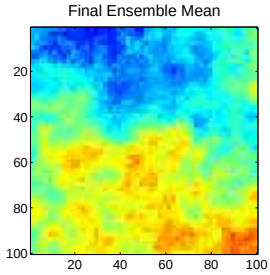
(c) int. step = 6, $N_e = 50$



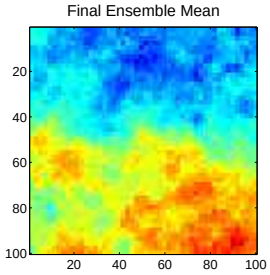
(d) int. step = 6, $N_e = 25$



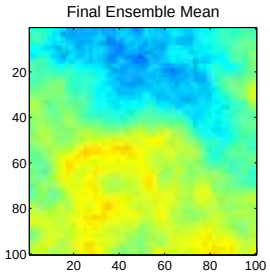
(e) int. step = 3, $N_e = 100$



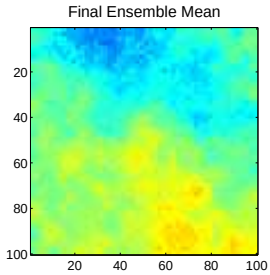
(f) int. step = 3, $N_e = 50$



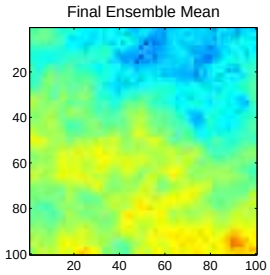
(g) int. step = 3, $N_e = 25$



(h) int. step = 1, $N_e = 100$



(i) int. step = 1, $N_e = 50$



(j) int. step = 1, $N_e = 25$

Figure 5.18: Comparison of final estimated ensemble mean for different SSBRC setups in Table 5.5.

pore pressure and principal stress distributions by Eclipse then we can read these distributions into MATLAB. We set up the failure criteria, damage and permeability model in MATLAB to generate MEQ events and update the permeability and elastic modulus values because of the failure, then we feed back the updated permeability and elastic modulus to Eclipse to move forward to the next time step. This developed model by combining Eclipse and MATLAB is more physically realistic that pore pressure relaxation model and faster than the fully coupled FEM model.

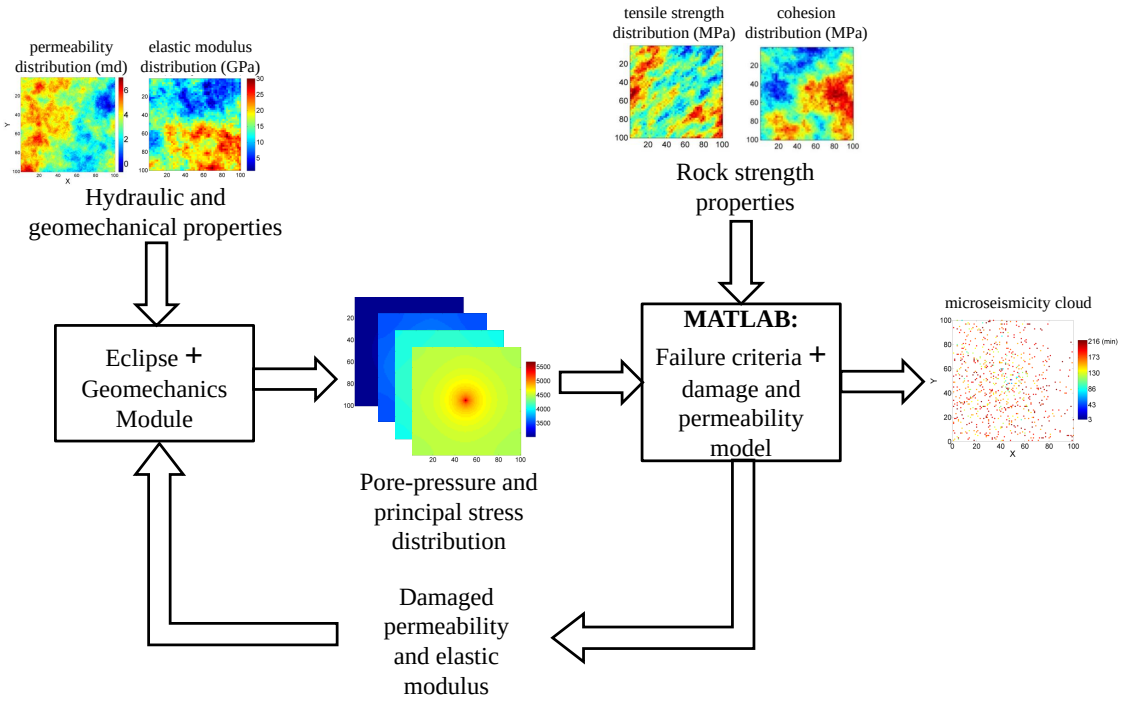


Figure 5.19: The overall workflow of the proposed geomechanical forward model by combining Eclipse and MATLAB.

Utilizing Eclipse and its geomechanics module combined with Matlab offers the following advantages:

- Implementing horizontal, deviated or any complex well trajectory.

- Multiple well and multiple stage configuration.
- Implementing stimulated reservoir volume (SRV).
- Simulating all three phases of treatment, flow back and production.
- Explicitly implementing the pre-existing anomalies in the reservoir (faults and natural fractures).
- Capability of natural fracture modeling.
- Capability to integrate various types of data (production data, MEQ data, geological and geomechanical data).
- Handling large number of grid blocks.

6. ASSESSING THE PERFORMANCE OF THE ENSEMBLE KALMAN FILTER FOR SUBSURFACE FLOW DATA INTEGRATION UNDER VARIOGRAM UNCERTAINTY*

6.1 Introduction

The uncertainty in the spatial distribution of subsurface hydraulic properties can lead to unreliable predictions of fluid flow displacement behavior, which can adversely impact the strategies for the development of subsurface hydrological and energy resources. The complexity associated with heterogeneity and spatial variability of subsurface transport properties at several scales, together with the high cost and lack of convenient access to the subsurface environment for direct sampling, leads to significant uncertainty and systematic errors in subsurface characterization studies [121, 122, 123]. Characterization of geological heterogeneity from various sources of measurements has been extensively studied by researchers in the hydrogeology and petroleum engineering community, among others. Inference of flow-related heterogeneous rock properties, such as permeability, from flow measurements leads to a nonlinear dynamic inverse problem that is inherently underdetermined, i.e., has more unknowns than can be uniquely resolved by available measurements [124, 125, 126, 121, 122, 127]. In general, it is possible to find several distinct models that describe the available data equally well but provide different predictions for future flow displacement behavior. To account for the ill-posed nature of the problem, uncertainty assessment and quantification are an important part of any parameter estimation approach used to characterize heterogeneous subsurface properties.

*Reprinted with permission from "Assessing the performance of the ensemble Kalman filter for subsurface flow data integration under variogram uncertainty" by Behnam Jafarpour and Mohammadali Tarrahi, 2011. Water Resources Research, 47, Copyright 2011 by John Wiley and Sons.

The problem of uncertainty quantification is closely tied to selection (sampling) of relevant model parameters $\mathbf{m}$ using the existing measurements $\mathbf{y}$, the likelihood model $p(\mathbf{y}|\mathbf{m})$, and the prior model $p(\mathbf{m})$ (here $p(\cdot)$ represents the probability density function (pdf) and $p(\mathbf{y}|\mathbf{m})$ denotes the pdf of measurements $\mathbf{y}$ conditioned on a specified set of parameters $\mathbf{m}$). From a probabilistic perspective, the model identification problem can be formulated as characterization of the pdf, $p(\mathbf{m}|\mathbf{y})$, by conditioning a prior model density $p(\mathbf{m})$ on available measurements $\mathbf{y}$. The Bayesian approach [128] provides an elegant framework for conditioning model parameter distributions on available data through the likelihood function, i.e., $p(\mathbf{y}|\mathbf{m})$. In practice, however, except for a very limited number of simplified cases, complete characterization of the conditional density $p(\mathbf{m}|\mathbf{y})$ is not feasible [128]. Therefore, it is common to either estimate representative point statistics of the conditional distribution (e.g., mean or mode), with a related estimation error metric, or to approximately describe the conditional distribution $p(\mathbf{m}|\mathbf{y})$ by estimating N sample realizations from it, i.e., $\{\mathbf{m}\}_{1:N} \in p(\mathbf{m}|\mathbf{y})$. Rigorous sampling from an often high-dimensional and possibly complex (e.g., multimodal) conditional density is nontrivial [128, 129], which becomes only more complicated in this case where the conditional distribution is unknown.

To deal with the nonuniqueness of the solution in subsurface inverse problems and to quantify the corresponding uncertainties, stochastic inversion methods have become increasingly popular for groundwater model calibration [130, 131, 132, 133, 134, 135, 136, 60, 137, 62]. Conditional simulation techniques either update individual realizations from a prior distribution to generate multiple samples that can be regarded as realizations from the conditional distribution [130, 131, 132, 133] or collectively process an ensemble of realizations to generate several conditional realizations at once [54, 60, 137, 62] (see also [61] and [62] for further discussion).

The former approach is known to suffer from high computational complexity, while efficient implementations of the latter approach have only been developed for relatively simple schemes, e.g., the ensemble Kalman filter (EnKF), which is based on a linear second-order update equation [54, 138]. In this work, we focus on ensemble methods and, in particular, assess the performance of the ensemble Kalman filter under variogram structural uncertainty.

The EnKF model calibration procedure provides a systematic mechanism for approximating the uncertainty in model predictions using a finite set of realizations. As an ensemble data assimilation approach, the filter was originally introduced to extend the optimal linear Kalman filter for state estimation in nonlinear systems [54]. It has been widely used for state and parameter estimation in several branches of engineering and science, including meteorology and oceanography [58, 54, 57], hydrology [56, 60], and oil reservoir characterization [135, 136, 137, 67]. For subsurface flow model calibration, the method is mainly used as a parameter estimation tool to identify the spatial distribution of aquifer hydraulic properties from flow data. Under correctly specified variogram models, the performance of EnKF in estimating underground hydraulic properties has been shown to be encouraging [67]. This success can be largely attributed to the strong correlations in the spatial distribution of hydraulic properties and their significant cross correlation with the flow response. The former implies that the true dimension of the underlying parameter space is far less than its nominal dimension. This important fact may be used to explain why a small ensemble of permeability models that is derived from a correct variogram model can effectively capture the main spatial variability in the seemingly high-dimensional parameter space. In addition, the permeability values directly appear in the flow equations and leave a strong and explicit signature on both local and global flow behavior. Thus, the permeability field tends to be strongly

correlated with the flow data, which can be effectively exploited by the second-order Kalman update equation.

In practice, however, variogram models are generated by combining geologic expertise with qualitative information (e.g., outcrop maps and regional geology) and limited static measurements. Hence, the functional form and parameters of variogram models can carry significant uncertainty. In a broader context, one could also consider the uncertainty in the type of geologic continuity and the conceptual model used to describe the main variability in subsurface physical properties. In many cases, however, the geologic history of the site and exploratory surveys (well logs, seismic, and core analysis) can reveal important information about the formation type. In this work, we assume variogram-based modeling of subsurface heterogeneity without considering the uncertainty in the type of conceptual geologic continuity model. The uncertainty in the geologic continuity is clearly an important topic that deserves future research attention. The main focus of this work is on the effect of variogram uncertainty on the EnKF performance.

To date, very few studies in the literature have considered the application of the EnKF for estimating hydraulic properties of geologic formations under incorrectly specified individual variogram parameters [60, 137, 62]. However, to our knowledge, a systematic evaluation of the EnKF update under uncertain variogram models and direct estimation of the variogram model parameters is not available. In [60], The authors applied the EnKF to sequentially update groundwater model parameters and states such as hydraulic conductivity and pressure head, respectively. They used a two-dimensional example to demonstrate the capability of EnKF and its sensitivity to different factors, including the number of realizations, measurement times, and the initial ensemble. The authors found that EnKF provides an efficient approach for obtaining satisfactory estimation of the hydraulic conductivity field

from dynamic measurements. They also studied the effect of incorrect variogram models on the estimation results and applied the EnKF to a three-dimensional model. However, a stochastic treatment of the variogram model parameters to account for the uncertainty in the variogram model parameters or to estimate the parameters of the variogram model from dynamic flow data was not considered.

In [137], the authors studied the performance of EnKF for continuously updating an ensemble of permeability models to match real-time multiphase production data in oil reservoirs. They avoided the physical inconsistency between model parameters and states by resolving the flow equations from the previous assimilation step to the current step using the updated permeability models (the authors called this step a confirming option). While in [137] improved results are reported due to the confirming option, in [61] little or no improvement was found as a result of applying the restart option. In [137] the sensitivity of using a different number of realizations in the EnKF is also investigated and found that a relatively large number of realizations are needed to obtain stable results, particularly for a reliable uncertainty assessment. The sensitivity of the estimation results to using different covariance functions is also considered in [137]. A systematic analysis to deal with possible uncertainty in the variogram parameters is not included in [137]. In [139], a combination of laboratory-based aquifer tracer experimentation and a bias-aware EnKF is presented to demonstrate that systematic modeling errors in source loading dynamics and the spatial distribution of hydraulic conductivity pose severe challenges to groundwater transport forecasting under uncertainty.

Given the importance of the variogram model parameters (or parameters controlling global continuity in other spatial models) and the significant uncertainty associated with them, it is imperative to acknowledge and incorporate variogram uncertainty in applying EnKF to realistic problems. Intuitively, the flow data

are expected to carry important information about the structural attributes of the hydraulic conductivity field and hence can be used to constrain variogram parameters. In this work, we report the results of our investigation of the EnKF performance for nonlinear dynamic flow data integration under uncertain variogram models. The main contributions of this work can be summarized as (1) assessing the feasibility of directly estimating structural variogram parameters from flow data using the EnKF, (2) investigating the sensitivity of the EnKF performance to errors in individual variogram parameters, (3) evaluating the EnKF performance when variogram model parameters are unknown, and (4) illustrating the advantage of overestimating the uncertainty in the variogram model and initializing the EnKF with a diverse ensemble. The results reported in this work provide important insights about effective design and implementation of the EnKF in realistic settings where one is usually in doubt about the uncertainty in the geologic continuity and hence the global flow displacement patterns.

We begin our discussion in Section 6.2 with a brief overview of the EnKF method and variogram modeling, followed by a description of the experimental setup for the water-flooding numerical examples and a set of two-dimensional examples to motivate the significance of variogram model uncertainty in model calibration with EnKF. Section 6.3 presents two sets of experiments in which EnKF is used (1) to directly infer variogram parameters and (2) to estimate permeability maps under highly uncertain initial variogram parameters. A final example with a single layer of the SPE10 model is presented to assess the validity of the discussed method under more realistic settings. The conclusions drawn from these experiments and their implications in application of the EnKF under variogram uncertainty are presented in Section 6.4.

6.2 Methodology

6.2.1 Ensemble Kalman Filter

For large-scale nonlinear state-space models, derivation of the states' forecast mean and error covariance under a nonlinear state transition function is not trivial. Linearized versions of the Kalman filter, such as the extended Kalman filter (EKF), have been used to address nonlinear error propagation for small- to medium-scale problems and when nonlinearities are mild [111]. However, the additional cost of linearization makes the EKF-type methods impractical for large-scale or strongly nonlinear problems. In [54] an alternative ensemble-based solution is provided by using the Monte Carlo forward simulation for error propagation under nonlinear dynamical models, which has found widespread application in ensemble-based data assimilation across several disciplines [54, 58, 56, 135, 136, 67].

The EnKF proceeds by nonlinear forecasting of the state density using a Monte Carlo simulation with a finite number of samples from the prior density as follows:

$$\begin{aligned} \mathbf{x}_{t|t-1}^{j} &= f_t\left(\mathbf{x}_{t-1|t-1}^{j}, \alpha_{t-1}, \mathbf{w}_{t-1}^{j}\right), \quad j = 1, 2, ..., N \\ \mathbf{y}_{t|t-1}^{j} &= h_t\left(\mathbf{x}_{t|t-1}^{j}\right) + \mathbf{v}_t^{j}, \quad j = 1, 2, ..., N \end{aligned} \tag{6.1}$$

where $j = 1, 2, ..., N$ is the jth replicate of the state vector, α_{t-1} denotes a vector of nonrandom and generally timedependent model input and control variables, $\mathbf{w}_t^j$ and $\mathbf{v}_t^j$ are the jth realizations of model and measurement errors, respectively, and the functions $f_t(\cdot, \cdot)$ and $h_t(\cdot)$ represent the nonlinear state transition function and measurement operator that relates the states to measured quantities, respectively. The forecast ensemble is then used to approximate the forecast mean and covariance that are needed in a slightly modified LLS update equation, written for individual samples in the ensemble. The ensemble version of Equation (6.1) is written as [138]

$$\hat{\mathbf{x}}_{t|t}^{j} = \mathbf{x}_{t|t-1}^{j} + \mathbf{P}_{t|t-1}^{e}\mathbf{H}_{t}^{T}\left(\mathbf{H}_{t}\mathbf{P}_{t|t-1}^{e}\mathbf{H}_{t}^{T} + \mathbf{R}_{t}\right)^{-1}\left(\mathbf{y}_{t}^{j} - \mathbf{H}_{t}\mathbf{x}_{t|t-1}^{j}\right), \quad j = 1, 2, ..., N \tag{6.2}$$

where the superscript e is used to indicate ensemble approximated statistics.

The system of Equations in (6.1) and (6.2) together with the initial ensemble of states $\mathbf{x}_{0|0}^{j}$, define the ensemble Kalman filter recursion for a nonlinear problem. It is important to note that instead of updating the first- and second-order moments (as done in the original Kalman filter), the ensemble form of the filter updates individual samples (or their perturbations from the mean depending on the implementation), which can then be used to approximate any desired (update) statistics. At this point, a few implementation remarks are in order.

First, for a parameter with physical bounds $s_{\min} \leq s \leq s_{\max}$, we can apply the following pair of inverse (before update) and forward (after update) error function transforms:

$$s_T = \text{erf}^{-1}\left(\frac{2s - (s_{\min} + s_{\max})}{s_{\max} - s_{\min}}\right) \Leftrightarrow s = \frac{(s_{\min} + s_{\max})}{2} + \frac{(s_{\max} - s_{\min})}{2}\text{erf}(s_T) \tag{6.3}$$

where s_T denotes the transformed (unbounded) variables. This transformation tends to impart (univariate) Gaussian properties on the estimation parameters. In the past, we have applied this transform to saturation states and obtained acceptable results [66]. In this work, we apply this transformation to variogram parameters. Second, it is common to update the predicted measurements by including them in the state vector, which facilitates the calculation of cross correlations between measurements and states. As a result, the measurement matrix Ht becomes a

simple selection operator that extracts the observation part of the new state vector. Moreover, the EnKF formulation as presented in Equations (6.1) and (6.2) is only a state estimation approach. In this work, however, we apply the EnKF to estimate the static parameters (i.e., permeability or variogram parameters) of a subsurface flow model. Application of the EnKF to parameter estimation is easily accomplished by augmenting the state vector with uncertain parameters. That is, to estimate the parameter vector $\mathbf{m}$, we define a new (augmented) state vector composed of the original states and parameters to be estimated; that is, $\mathbf{x}_t^{\text{aug}} = [\mathbf{x}_t \quad \mathbf{m}_t]^T$. After this simple modification, the analysis equation (Equation (6.2)) is applied to $\mathbf{x}_t^{\text{aug}}$ to update both states and parameters. Note that the forecast for time-invariant parameters is the same as the updated parameters at the previous analysis step, i.e., $\mathbf{m}_{t+1|t} = \mathbf{m}_{t|t}$.

Last, substituting the theoretical covariances with a low-rank sample approximation introduces sampling errors that can lead to rank deficiency, ill-conditioning and inaccuracies of the matrices involved in the update equation. Determination of the ensemble size is problem specific and, in general, depends on the dimensionality and complexity of the problem and, in practice, is limited by available computational resources [140]. For a linear model in which the uncertainties are described with jointly Gaussian statistics, in the limit of an infinite number of samples, the solution converges to the MMSE estimate given by the Kalman filter [141, 55]. However, when a nonlinear model is used to propagate the state vector, the forecast states are not likely to have a Gaussian distribution even if the distributions of the initial state and model errors are Gaussian. As a result, the EnKF update becomes suboptimal even when an infinitely large ensemble size is used. Nonetheless, one can still apply the simple Kalman filter (KF) update form to find the LLS estimate regardless of the prior and measurement distribution types. In this work,

we have used the square root filter of [112, 73] to implement the EnKF updates. In all experiments, an ensemble size of $N = 100$ is used.

6.2.2 Prior Structural Model: Variogram

A traditional approach for representing spatially variable properties such as permeability or porosity is through (second-order) stationary random processes characterized by a variogram model $\gamma(\mathbf{h})$, or equivalently a covariance model $C(\mathbf{h})$, where the correlation in the field is a function of distance only (for a specific direction). A stationary variogram function describes the structural variability through point-to-point (or two-point) correlation in space [104]. Among the most commonly used parametric variogram models for describing spatial rock properties are the spherical $\gamma_{\text{Sph}}(\mathbf{h}; a, c_1)$, Gaussian $\gamma_{\text{G}}(\mathbf{h}; a, c_1)$, and exponential $\gamma_{\text{E}}(\mathbf{h}; a, c_1)$ models [104]. The parameters a and c_1 are used to denote the range and sill of a one-dimensional variogram, respectively. In one-dimensional problems, these variogram functions are completely characterized by their (effective) range or characteristic correlation length a, which represents the smallest distance at which the variogram function takes its (effective) maximum value, i.e., the variance of the random process. For two-dimensional variograms the major direction of continuity θ and the anisotropy ratio, i.e., $a_{\max}/a_{\min}$, are introduced as two new variables [104]. The underlying assumption is that the direction of minimum continuity is perpendicular to the direction of maximum continuity.

6.2.3 Experimental Setup

In Section 6.3, several numerical water-flooding experiments are used to illustrate the estimation performance of the EnKF under variogram uncertainty. We refer to the three models used in this chapter as model A (two-dimensional 64×64 model), model B (one-dimensional 100×1 model), and model C (layer 3 of the SPE10 model,

60×220). The detailed description of simulation and data integration parameters for each case is summarized in Table 6.1. The bottom hole pressure at the injection ports and water and oil rates at the production ports are measured before being assimilated. In each example, the measurements were obtained every 3 months by running a forward flow simulation with a (synthetic) reference log permeability field. The first set of experiments is performed using symmetric 13-spot well configurations, consisting of nine water injection wells and four oil production wells (Figure 6.1a) in a 64×64 (two-dimensional) domain, which is model A.

	Model A (64×64)	Model B (1-D)	Model C (SPE10-L3)
	General Simulation Inputs		
Phases	oil/water	oil/water	oil/water
Simulation time	12×3 months	12×3 months	12×3 months
Grid system	$64 \times 64 \times 1$	$100 \times 1 \times 1$	$60 \times 220 \times 1$
Cell dimensions	$10 \times 10 \times 10$	$10 \times 10 \times 10$	$10 \times 10 \times 10$
Rock porosity	0.20 (constant)	0.20 (constant)	0.20 (constant)
Initial oil saturation	0.90 (uniform)	0.90 (uniform)	0.90 (uniform)
Initial pressure	3000 psi (uniform)	3000 psi (uniform)	3000 psi (uniform)
Geostatistical simulation	SGSIM	SGSIM	SGSIM
Injection volume	1 PV	1 PV	1.5 PV
Number of injectors	9	1	15
Number of producers	4	2	8
Injection well constraints	water flow rate	water flow rate	water flow rate
Production well constraints	pressure	pressure	pressure
	Data Assimilation Parameters		
Observation interval	3 months	3 months	3 months
Observation at injection wells	pressure	pressure	pressure
Measurement noise STD	50 psi	10%–20%[b] (mean 30 psi)	10%–20%[b] (mean 45 psi)
Observation at production wells	oil and water rate	oil and water rate	oil and water rate
Measurement noise STD	100 STBD	10%–20%[b] (8 STBD)	10%–20%[b] (105 STBD)

[a]PV, pore volume; STD, Standard Deviation; psi, pounds per square inch (6895 Pa); STBD, stock tank barrel per day.
[b]Of the dynamic range.

Table 6.1: General simulation and assimilation information[a].

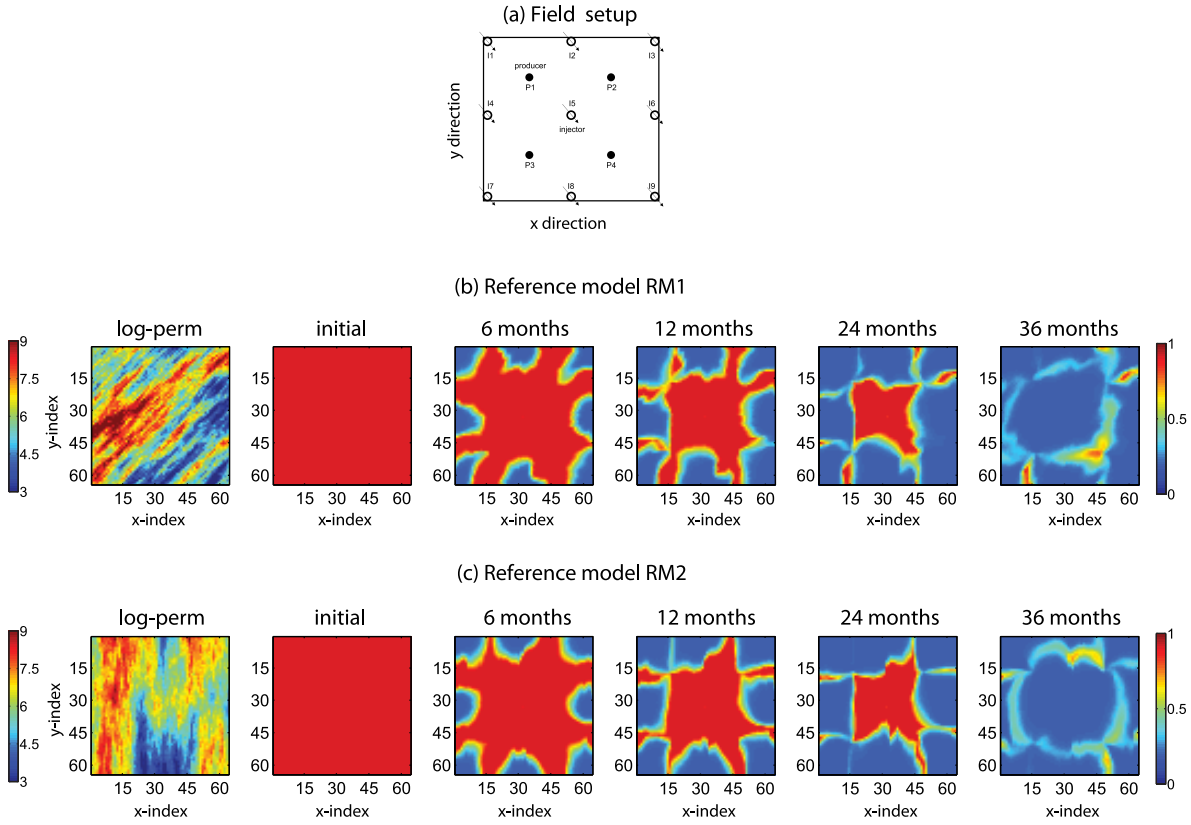


Figure 6.1: Reservoir well configuration for model A with reference Lperm models and saturation profiles. (a) 13-spot well configuration. (b) Reference Lperm RM1 and corresponding saturation profiles. (c) Reference Lperm model RM2 and corresponding saturation profiles.

6.2.4 Motivating Example

A common practice in geostatistical reservoir modeling is to estimate the variogram model from available static data (log, core, and time-lapsed seismic) and use it to construct multiple realizations of property maps that are subsequently conditioned on dynamic flow measurements. Throughout this inversion process, it is typically assumed that the variogram model is known. However, since the variogram model is constructed from limited data and given that it has a significant global effect on the description of the spatial property distribution, it is critical to acknowledge and account for the uncertainty in the variogram model parameters. In particular, model calibration with the EnKF is significantly helped (or misled) by assuming a known correct (or incorrect) variogram model. For instance, the EnKF update may have difficulty in compensating for incorrect structural assumptions. The first two examples in this section are used to motivate the topic and the experiments that follow in the remainder of the this chapter. Throughout this chapter, the variogram type is assumed to be known, the nugget effect is ignored, and the uncertainty in the structural model is only introduced through specification of the main variogram parameters, i.e., direction of maximum continuity as well as the maximum and minimum ranges for the two-dimensional variogram models.

In example 1 of this section, a synthetic reference log permeability field is estimated from production measurements using the EnKF. The reference model (RM1) in this case is generated using the SGSIM algorithm and an exponential variogram model with a maximum continuity direction of $\theta = 45^\circ$ and maximum and minimum ranges of 500 and 100 m, respectively (i.e., $\theta = 45^\circ$, $a_{\max} = 500$ m, and $a_{\min} = 100$ m). Figure 6.1b shows the synthetic reference log permeability model and snapshots of the corresponding saturation plots at specified times. For this example,

the initial ensemble of log permeabilities is also derived from the same variogram model with different random seed numbers and by conditioning the realizations on the hard data at well locations. Hence, the initial ensemble is structurally consistent with the reference log permeability model (through honoring the variogram) and locally captures the variability around the wells (due to conditioning on hard data). Figure 6.2a shows samples from the initial log permeability ensemble. The EnKF updated permeabilities are shown for 3, 18, and 36 months. As can be seen from these results, when the initial ensemble is structurally consistent with the reference model RM1, the EnKF can use the measurements to identify the location and orientation of the main features in the reference field. The ensemble forecast of the flow response for sample production wells is shown in Figure 6.2c. The final updated models tend to provide more accurate predictions of the flow that envelope the true values (suggesting little bias in the flow predictions).

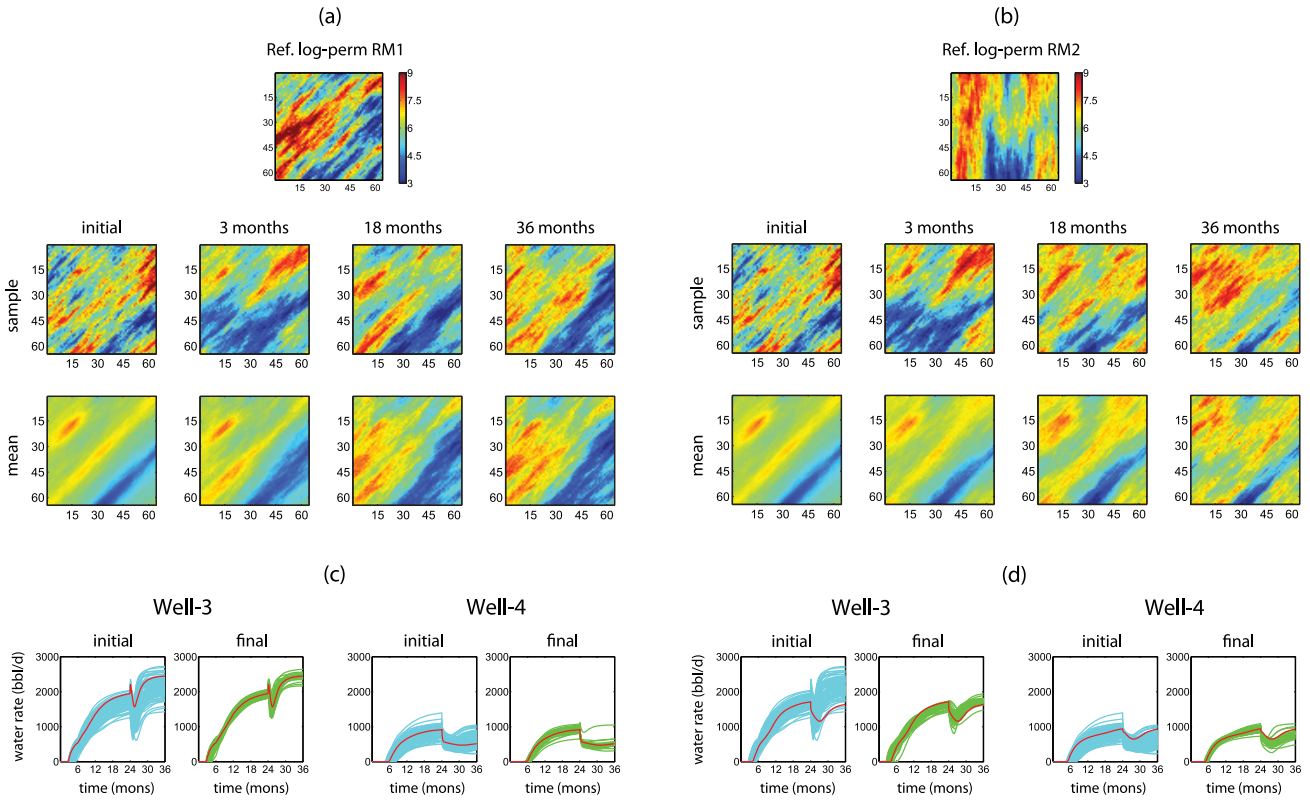


Figure 6.2: Ensemble Kalman filter (EnKF) estimation results for log permeability with correct and incorrect initial variogram models. (a) (top) Reference log permeability model RM1 and (middle) a sample and (bottom) the ensemble mean of the estimated log permeabilities for example 1. (top) Reference log permeability model RM1 and (middle) a sample and (bottom) the ensemble mean of the log permeability estimation results for example 2. (c) Sample water flow rate forecast with the initial and final permeability ensemble for example 1. (d) Sample water flow rate forecast with initial and final permeability ensemble for example 2. The estimation results show a sample log permeability realization and the corresponding log permeability ensemble mean. The number of ensemble replicates is $N = 100$.

Example 2 replicates example 1, with the only exception being that the reference log permeability in this case (reference model RM2) is generated from a variogram model with parameters $\theta = 90°$, $a_{\max} = 500$ m, and $a_{\min} = 100$ m. Note that the two ranges are the same as in the previous model (RM1) and are consistent with the initial ensemble. In this case, the observations are generated by running the multiphase flow simulation forward with this reference model. The initial ensemble and all other experimental conditions remain identical to example 1, which means that now the direction of continuity in the initial ensemble ($\theta = 45°$) is inconsistent with the reference model RM2 ($\theta = 90°$). The updated log permeability results are shown in Figure 6.2b. The results suggest that a bias in specifying the variogram continuity direction tends to persist during the EnKF updates in time and can significantly degrade the inversion performance. As can be observed in Figure 6.2b, the EnKF updates are not able to compensate for the specified error in the structure of the variogram model. An important observation is that the ensemble variance (spread) is reduced after each update in both examples; however, this reduction of variance is not a measure of filter update accuracy and only reflects the effect of observations on bringing the ensemble members closer to each other and increasing the confidence in model predictions. The flow response in Figure 6.2d shows a bias in the predictions even in this case where the given well configuration significantly constrains the flow pattern and well responses.

Since the error in the variogram parameters can significantly degrade the performance of the EnKF, it is necessary to account for the uncertainty in the variogram model and represent it in the initial ensemble. Moreover, if the uncertainty in the variogram model is significant, it is important to know which variogram parameter(s) has more impact on the EnKF performance. Furthermore, one would ideally like to use the production data as additional information to improve the

description of the variogram model. Therefore, we also consider application of the EnKF to directly updating variogram parameters.

6.3 EnKF Performance Under Variogram Uncertainty

In this section we explore the performance of the EnKF when the variogram model parameters are unknown or highly uncertain. We follow two alternative methods to deal with unknown variogram parameters: (1) directly estimating the variogram parameters and (2) estimating permeability values after assuming uniform noninformative distributions for variogram parameters in generating the initial ensemble. The objective is to find out whether the flow data reveal structural information about the permeability field and, if so, to evaluate the performance of EnKF in inferring the structural information from the measurements.

6.3.1 Direct Estimation of Variogram Model Parameters

In this section, instead of estimating log permeability maps, we use the EnKF to update the three variogram parameters θ, $a_{\max}$, and $a_{\min}$ (in two-dimensional examples) by augmenting them to the state vector. Hence, we generate an ensemble of three variogram model parameters using the following wide (noninformative) distributions: $\theta \sim \mathcal{U}(0^\circ, 180^\circ)$, $a_{\max} \sim \mathcal{U}(100 \text{ m}, 1000 \text{ m})$, and $a_{\min} \sim \mathcal{U}(10 \text{ m}, 1000 \text{ m}) \leq a_{\max}$. The uniform distribution on variogram parameters allows for including a broad range of uncertainty on variogram parameters without any preference (bias) on the expected value of these parameters. It is important to note that while these parameters are not Gaussian, the transformation applied to them through Equation (6.3) before the update step imparts univariate Gaussian properties on them.

Random samples from these variogram model parameters are used to generate an ensemble of conditional (to hard data) log permeability fields that are used during

the forecast steps to predict the ensemble of states and flow measurements. The cross correlations between the flow predictions and variogram parameters are calculated and used in an EnKF analysis step to update the ensemble of states and variogram parameters. The updated variogram parameters are then used to generate a new ensemble of log permeability maps for the next prediction step, and the sequence of forecast and update steps are repeated until all measurements are assimilated. It is important to note that the new permeability fields that are generated from the updated variogram models are not conditioned on the flow data (discussed in Section 6.3.1.2). We also note that similar to the log permeability fields, the variogram parameters do not change during the forecast steps. Furthermore, to keep the variogram direction of anisotropy within its natural constraints $0^\circ \leq \theta \leq 180^\circ$, we used a similar transformation pair as in Equation (6.3) with appropriate bounds of $\theta_{\min} = 0^\circ$, $\theta_{\max} = 180^\circ$. Another constraint to honor during the updates is $a_{\min} \leq a_{\max}$ to avoid shifting anisotropy direction. We implemented this constraint by updating $a_{\max}$ and the inverse of the anisotropy ratio $0 \leq r \leq 1$, with the latter constraint being imposed in a similar fashion using Equation (6.3).

6.3.1.1 Two-Dimensional Example

Figure 6.3 (top) shows the histogram of the variogram parameters after selected update steps in experiment A1 for reference model RM1. Noting that the reference field is generated with variogram model parameters $\theta = 45^\circ$, $a_{\max} = 500$ m, and $a_{\min} = 100$ m, the EnKF update results for $\theta = 45^\circ$ and $a_{\max} = 500$ m are poor. To better evaluate these results, a similar example using the reference model RM2 is also shown in Figure 6.3 (bottom) (experiment A2). The reference model RM2 is constructed from a variogram model with parameters $\theta = 90^\circ$, $a_{\max} = 500$ m, and $a_{\min} = 100$ m. Also in this case, it is evident that the EnKF update provides

estimates that are far from the true values. The results clearly indicate that the EnKF is not able to correctly update the three parameters of the variogram. This outcome may seem surprising since the EnKF is used to update only three parameters from the dynamic flow data at several locations. In particular, the results seem to contradict the promising outcomes that are often obtained when the EnKF is used to update thousands of grid block log permeability parameters. The filter seems to have failed in estimating the three parameters in this case. It appears that the success of EnKF in parameter estimation depends on the nature and complexity of the problem and, more importantly, on the strength of linear correlation between observations and unknown parameters. The results of other experiments that we performed to estimate variogram parameters were also inconsistent (not shown). What can partly explain this poor performance is the complex and nonunique relation between variogram parameters and flow data, which is hard to describe with the linear second-order correlations that are used in the EnKF update. A particular issue that makes the problem rather challenging is solution nonuniqueness since a single variogram model can generate many permeability realizations that share similar structures but have completely different spatial distribution and hence flow response. In the same manner, incorrect combinations of variogram parameters may result in flow responses that better match the data than some of the realizations generated with correct variogram parameters. Next, we investigate this important issue by estimating the variogram range in a set of simple one-dimensional examples.

6.3.1.2 One-Dimensional Example

To remove the effect of the interaction between the three variogram parameters, in this section we evaluate the performance of the EnKF in estimating a single range parameter in a one-dimensional (1-D) problem. We have also considered (not shown)

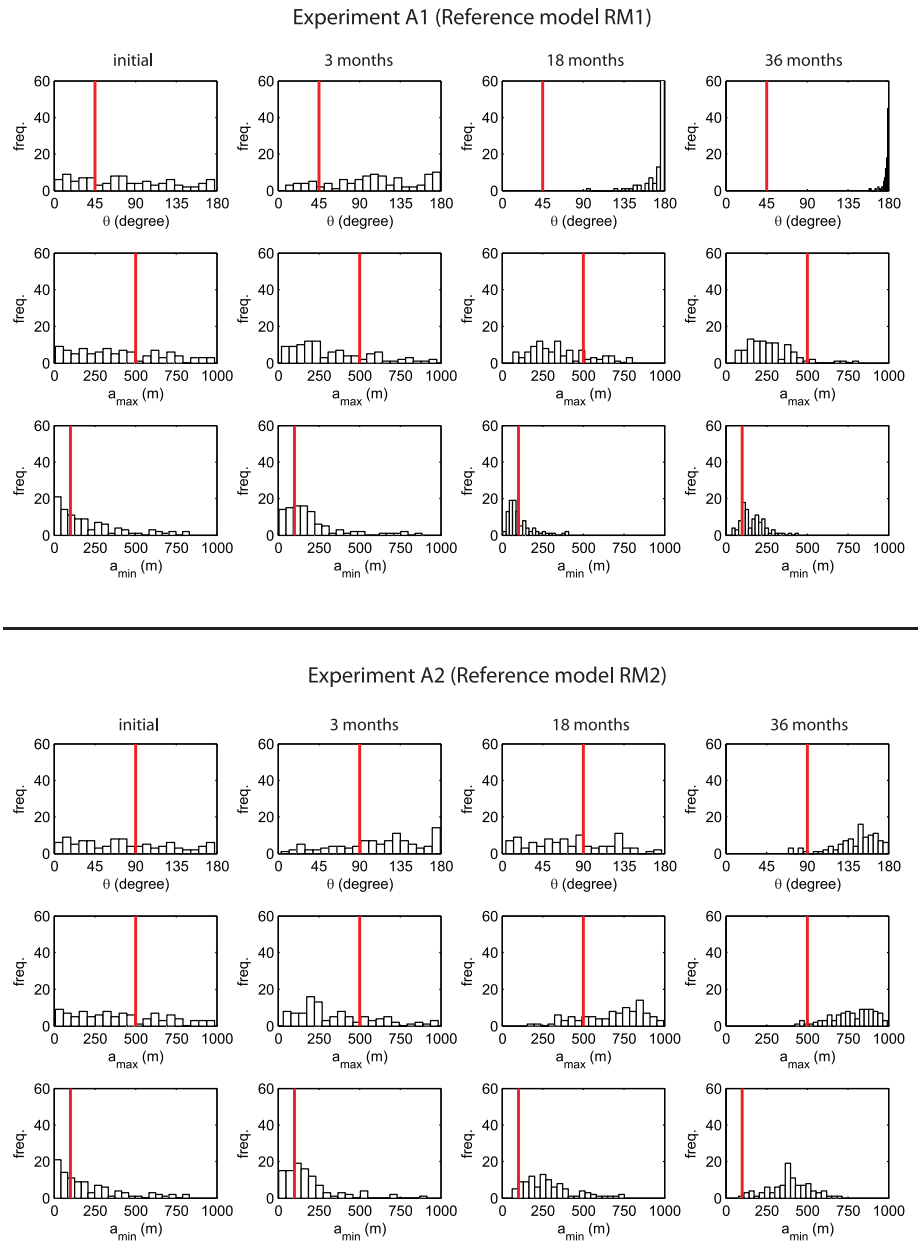


Figure 6.3: Estimation results for experiments A1 and A2. (top) The histogram of the estimated variogram model parameters with the EnKF for reference model RM1 in experiment A1. The true values of the variogram parameters are $\theta = 45°$, $a_{\max} = 500$, and $a_{\min} = 100$. (bottom) The histogram of the estimated variogram model parameters with the EnKF for reference model RM2 in experiment A2. The true values of the variogram parameters are $\theta = 45°$, $a_{\max} = 500$, and $a_{\min} = 100$. The initial variogram model parameters are described with the uniform distributions $\theta \sim \mathcal{U}(0°, 180°)$, $a_{\max} \sim \mathcal{U}(100\text{ m}, 1000\text{ m})$, and $a_{\min} \sim \mathcal{U}(10\text{ m}, 1000\text{ m}) \leq a_{\max}$. The EnKF is used to directly update the variogram parameters at the update steps. The number of ensemble replicates is $N = 100$.

estimating only one single parameter in a two-dimensional setting where the other two variogram parameters are fixed at their true values (where similar conclusions are drawn). However, we present a set of one-dimensional examples to completely remove the effect of the other two parameters. Figure 6.4 shows the experimental setup (Figure 6.4a), the true one-dimensional permeability map (Figure 6.4b), and the saturation profiles corresponding to the true permeability map (Figure 6.4c). There is one injector in the center and two producers at the two ends of the domain. The simulation domain is discretized into a total number of 100 grid blocks.

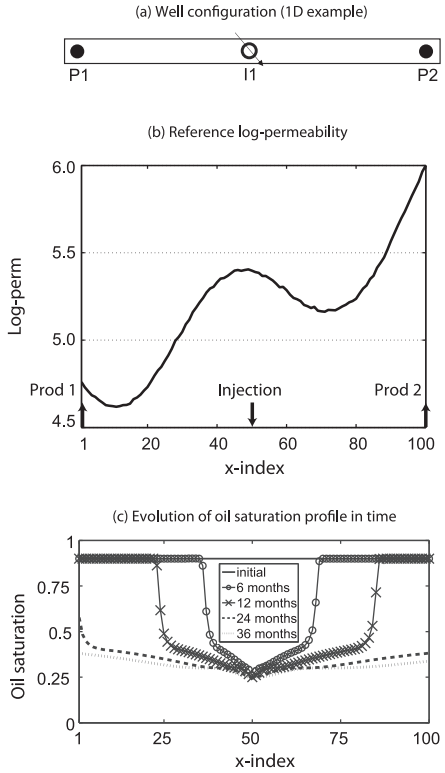


Figure 6.4: True log permeability model and saturations for the one-dimensional example (model B): (a) rough sketch of model domain and well configuration, (b) true log permeability model with $a = 700$ m, and (c) oil saturation profiles in time.

Figure 6.5 provides the summary of the estimation results where the true value of the variogram range is 700 m. The variogram type is Gaussian and features a smooth large-scale variability across the domain. The initial permeability ensemble in this experiment was obtained by conditioning each realization on the hard data. The initial realizations are generated from a random sample of the variogram range distribution a $a \sim \mathcal{U}(100 \text{ m}, 1000 \text{ m})$. Figure 6.5a shows the evolution of the estimated variogram range histogram after selected update steps. The true value is shown with the red bar. It is clear that the estimated variogram parameters are not acceptable. Figures 6.5b and 6.5c show the RMSE and spread (S) for the estimated variogram range as a function of time.

The spread and RMSE measures become very similar when the estimated ensemble mean approaches the true parameter. Clearly, the updates do not provide correct estimates of the variogram range. To show that this outcome does not result from insufficient information in the data, we have also provided the results when the EnKF is used to update the grid block log permeability field in the next experiment. However, before proceeding with the next set of examples, we point out that to make our previous results statistically significant, we repeated a modified version of the above experiment 100 times, each time with a new true variogram range and the corresponding set of observations. The results (included in Appendix) show that the EnKF updates do not provide consistent estimates and the overall performance of the filter in updating a single variogram range is not satisfactory. One way to explain this behavior is to examine the correlation between the flow response and variogram parameters (shown in Figure 6.5d). From Figure 6.5d, it can be seen that among the types of well responses the pressure data seem to have larger correlation coefficients with the range; however, the computed correlations in Figure 6.5d are not significant (relative to similar correlations with the permeability distribution).

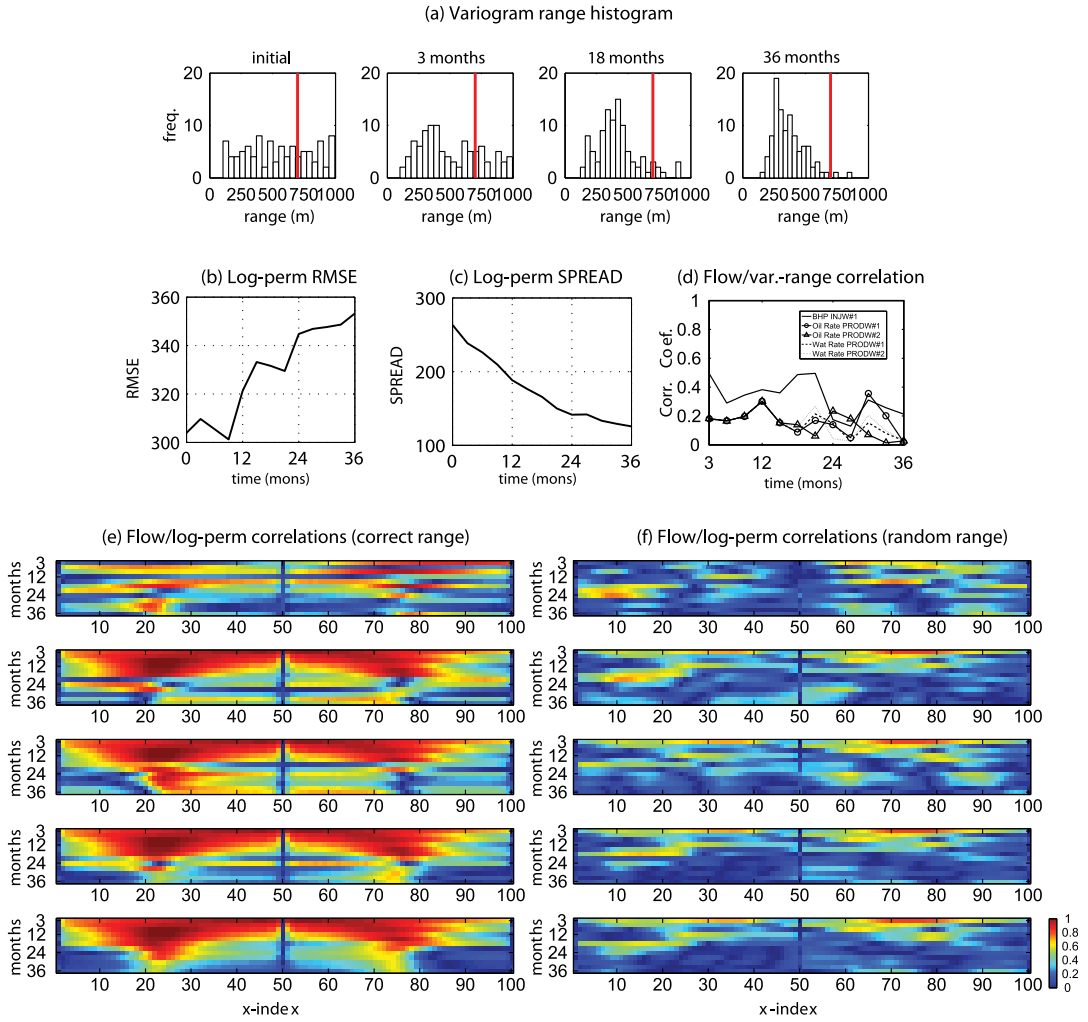


Figure 6.5: Direct variogram estimation results for the one-dimensional variogram range (experiment B1): (a) estimated histogram of the variogram range, (b) the root-mean-square error (RMSE) of the estimated range values, (c) the ensemble spread of the estimated range values, (d) cross correlation between the flow response and variogram range, (e) cross correlation between the flow response and log permeability values generated using the correct variogram range, and (f) cross correlation between the flow response and log permeability values generated using random variogram ranges.

Plotted in Figures 6.5e and 6.5f are the correlations of the dynamic well response variables with grid block permeabilities when the initial permeability ensemble is generated using the correct and randomly specified variogram ranges, respectively. The correlations are shown in time (y axis) for each grid block permeability value (shown on the x axis). The flow responses have significantly larger and physically more meaningful correlations with the grid block permeability than they do with the variogram range. Clearly, this correlation is stronger and more accurate when the correct variogram range is used. Since the relationship between variogram parameters and flow response is rather complex and nonunique, the linear content of the flow information (i.e., correlations) does not seem to provide the resolving power necessary for the EnKF to estimate variogram parameters. (The EnKF update equation only exploits the linear information in the measurements; thus, it appears to have difficulty updating variogram parameters from the weak correlations between variogram parameters and flow response.)

In estimating variogram parameters, we used the updated saturation and pressure fields to initialize the EnKF forecast for the next step, which introduces inconsistency between states and parameters. We tested two alternative approaches to assess possible drawbacks of this assumption and potentially improve the results. The first approach is to forecast the states and observations, after each update step, from the initial time step. This would ensure (at the cost of extra computation) that the states and parameters are consistent; however, the spatial distribution of the permeability would not be conditioned on the flow data. In the second approach, both variogram parameters and permeabilities were updated simultaneously; the log-permeability models were updated to improve the quality of the state forecasts for the next time step. Neither of the above schemes, however, provided results significantly different from what was reported above.

Figure 6.6 shows the estimation results for the one-dimensional example when the EnKF is used to update grid block log permeability values. The initial ensemble and other experimental conditions remain the same as in the previous example. Figure 6.6a shows the evolution of the log-permeability ensemble (and its mean) throughout the updates. The results clearly show improvements in the accuracy of the reconstructed permeability fields. Figures 6.6b and 6.6c show the RMSE and spread, respectively, of the log-permeability values that can be used to quantify the performance of the EnKF. The initial and final oil and water production plots for the two production wells are shown in Figures 6.6d and 6.6e. The results indicate that even with three wells, the grid permeability values can be successfully reconstructed if the initial uncertainty in the ensemble covers the full range of variability in the structural continuity model.

6.3.2 Permeability Estimation Under Uncertain Variograms

An alternative approach to deal with unknown variogram parameters is to use the regular EnKF with a very diverse initial log permeability ensemble by assuming noninformative variogram model parameters. Hence, in this section we use the same assumptions and distributions as in Section 6.3.1, but now we apply the EnKF to update the grid block log permeability parameters. Next, we present a brief sensitivity analysis to study the influence of incorrectly specified variogram parameters of the two-dimensional model (model A) on the EnKF performance.

6.3.2.1 Experiment A3: Sensitivity to Variogram Direction of Major Continuity

In the first set of experiments, the sensitivity of the EnKF to the direction of maximum continuity is studied by changing the parameter θ while fixing the other two variogram parameters at their correct values (i.e., $a_{\max} = 500$ m and $a_{\min} = 100$

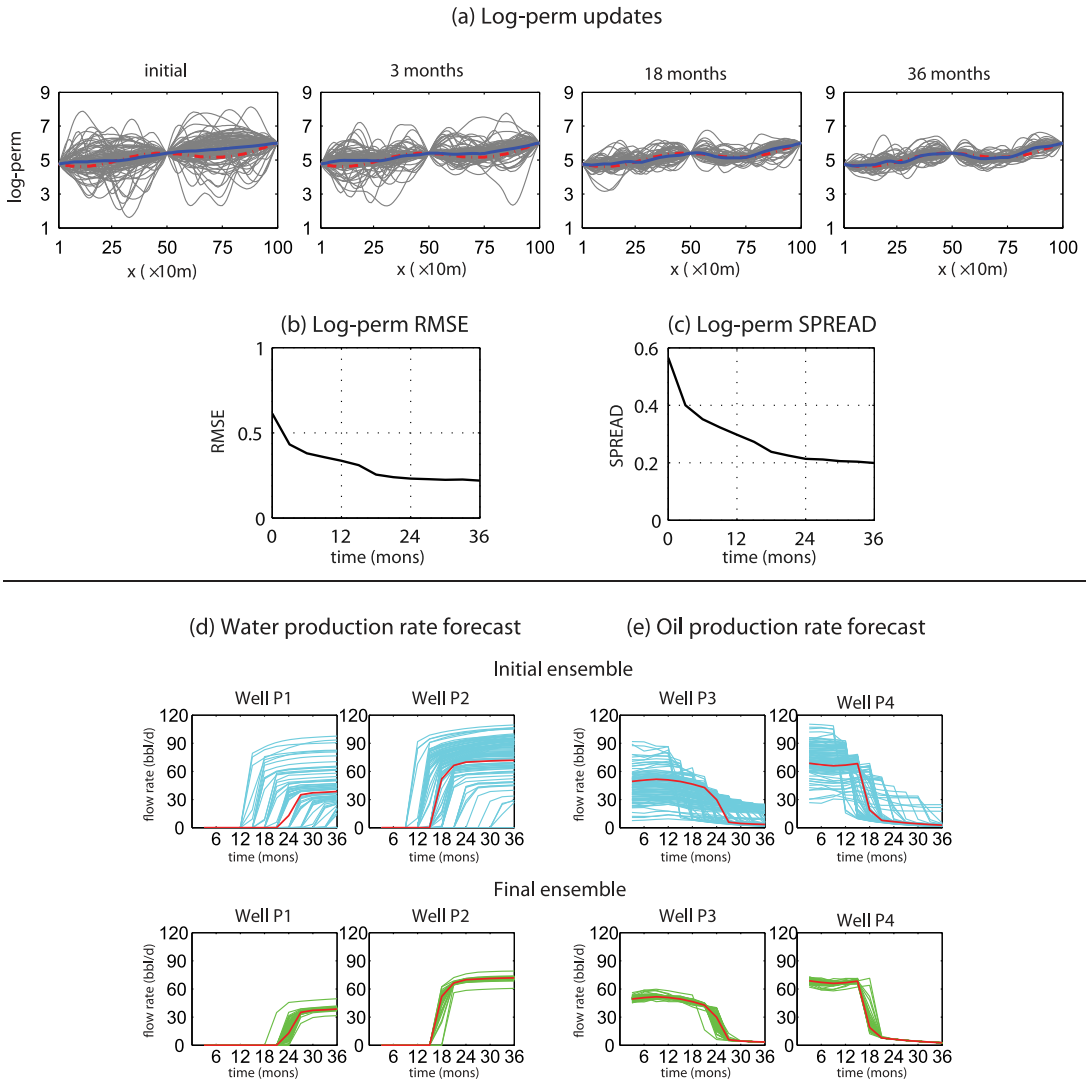


Figure 6.6: Grid block log permeability estimation results for the one-dimensional example (experiment B2) : (a) estimated log permeability ensemble for selected time steps, (b) RMSE of the estimated log permeabilities in time, (c) ensemble spread of the estimated log permeabilities in time, (d) initial and final ensemble water production forecasts for sample production wells, and (e) initial and final ensemble water production forecasts for sample production wells.

m).

Experiments A3-1 to A3-4 are used to evaluate the performance of the EnKF when $\theta = 0°$, $30°$, $90°$, $135°$, respectively. In these experiments, deviation of θ from its correct value is considered, and its effect on the performance of the EnKF is examined. For brevity, we only show the initial and final mean of the log permeability plots and report the main observations. The results are summarized in Figure 6.7a. In all of these experiments the initial ensembles of log permeabilities are conditioned on the hard data from the 13 wells. Hence, the regions with high and low permeability values (left and right parts of the domain) are evident in the initial ensemble. However, the features in the updated log permeability mean are not sharp, and their exact location and orientation are not consistent with those in the reference model. A qualitative comparison between the reference and updated map of the mean log permeability indicates that while the EnKF inversion has been able to apply reasonable local updates, the bias in the specified continuity direction is not corrected. The estimated log permeability fields in these examples tend to preserve their initial direction of major continuity. This is particularly evident for $\theta = 0°$, $\theta = 90°$, and $\theta = 135°$, which represent larger deviations. The example with $\theta = 30°$, however, suggests that the EnKF updates are not severely affected by a small bias in specifying the variogram direction of major continuity.

It can be argued that specifying a highly inaccurate variogram continuity direction is a rather pessimistic representation of prior knowledge. The fifth column in Figure 6.7a shows the results in which the direction of major continuity is assumed uncertain and described with a noninformative distribution $\theta \sim \mathcal{U}(0°, 180°)$. In this case, the EnKF is initialized with a wide ensemble that covers the full range of possible structural continuity directions. While this assumption still seems too pessimistic, it is intentionally adopted to illustrate the performance of the EnKF

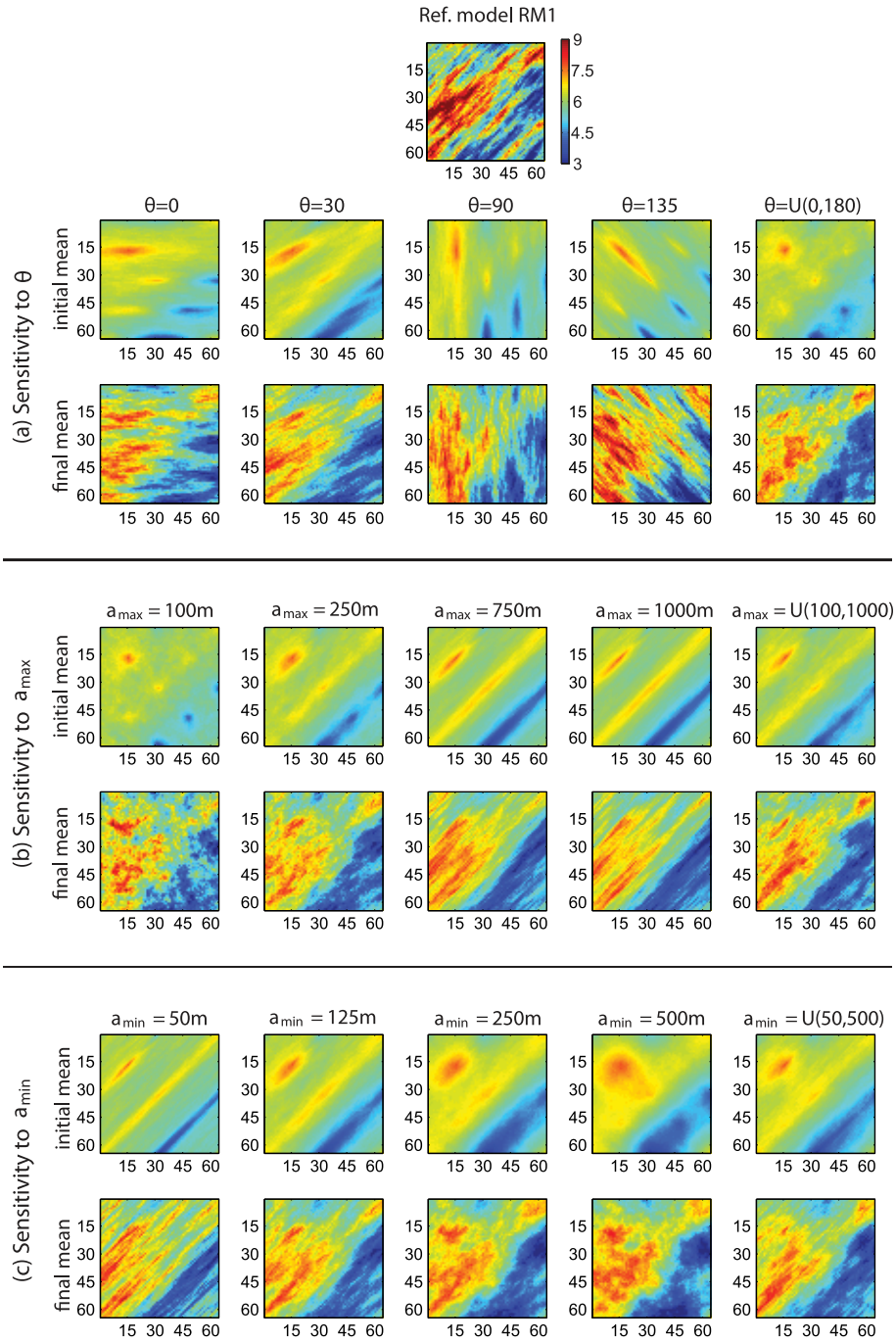


Figure 6.7: Sensitivity of the EnKF performance in estimating log permeability to errors in the variogram parameters (experiments A3A5): (a) sensitivity of the EnKF to direction of major continuity θ (experiment A3), (b) sensitivity of the EnKF to maximum range amax (experiment A4), and (c) sensitivity of the EnKF to maximum range amin (experiment A5). A sample and the ensemble mean of the log permeability for the (top) initial and (bottom) final ensemble are shown in each case. The number of ensemble replicates is $N = 100$.

in the extreme case of having no knowledge about the continuity direction. Apart from the initial log permeability ensemble, the EnKF implementation for updating the spatial log permeability fields is identical to its regular form discussed above. The updated mean for this case is shown in the fifth column of Figure 6.7a. As seen from Figure 6.7a, the approximate direction and location of the main high- and low-permeability trends are captured. The results suggest that including the full range of uncertainty in variogram anisotropy direction can be more advantageous to making a strict deterministic assumption about uncertain parameters. We note that in these experiments it was assumed that the other two variogram parameters are fixed at their correct values (this assumption will be relaxed in Section 6.3.2.4, Figure 6.8).

6.3.2.2 Experiment A4: Sensitivity to Variogram Maximum Range

A similar set of experiments to those in Section was conducted to assess the sensitivity of the EnKF updates to the variogram maximum range $a_{\max}$. In this case, the direction of anisotropy and the minimum range are fixed at their true values, i.e., $\theta = 45^\circ$ and $a_{\min} = 100$ m. All other experimental conditions are kept the same. Overall, the estimation results seem to be in good agreement with the global high- and low-permeability trends in the reference model. In experiments A4-1 to A4-4 the maximum range is incorrectly specified at 100, 250, 750, and 1000 m, respectively. Figure 6.7b displays the initial (Figure 6.7b, top) and final (Figure 6.7b, bottom) mean permeability estimation results for this case. For small $a_{\max}$ values, i.e., 100 and 250 m, the estimated log permeability fields do not represent the correct correlation length scales. On the other hand, large values of the range, i.e., $a_{\max} = 750$ m and $a_{\max} = 1000$ m, lead to persistent overestimation of the correlation length in the reference model. Consistent with the results in experiments

A3-1 to A3-4, the outcomes of experiments A4-1 to A4-4 indicate that an error in a prespecified variogram maximum range is carried over throughout the EnKF updates. It also appears that overestimation of the correlation length seems to have a smaller adverse effect than its underestimation.

Experiment A4-5 is performed to estimate the log permeability field when the maximum range of the variogram model is described with a uniform distribution $a_{\max} \sim \mathcal{U}(100 \text{ m}, 1000 \text{ m})$. This choice was made to ensure that the direction of anisotropy does not shift ; that is, the constraint $a_{\max} \geq a_{\min}$ was explicitly enforced in generating the variogram parameters. The estimated log permeabilities are shown in the fifth column of Figure 6.7b.

Overall, the log permeability updates have reconstructed the main features of the reference model without showing a particular structural trend.

6.3.2.3 Experiment A5: Sensitivity to Variogram Minimum Range

A suite of experiments is also carried out to evaluate the impact of the variogram minimum range $a_{\min}$ on the EnKF performance. Experiments A5-1 to A5-4 are used to update the initial ensemble of log permeability fields when the minimum range of the variogram is incorrectly specified as $a_{\min} = 25$, 50, 250, and 500 m, respectively. A summary of the estimation results is provided in Figure 6.7c. The effect of underestimating or overestimating the initial value of the minimum variogram range also seems to prevail throughout the assimilation. This is clearly observed by examining the increase in the average width of the diagonally oriented high-permeability feature as $a_{\min}$ increases in experiments A5-1 to A5-4. While the effect of incorrectly specified $a_{\min}$ is notable, the regions with low and high permeability values are clearly identified in all experiments. Experiment A5-5 shows the results when the variogram minimum range is uncertain and follows a uniform

distribution $a_{\min} \sim \mathcal{U}(50 \text{ m}, 500 \text{ m})$. The constraint $a_{\max} > a_{\min}$ is again explicitly enforced. The results are shown in the fifth column of Figure 6.7c. A comparison of the initial log permeability mean with the updated mean reveals that the EnKF updates have reconstructed the location and major trends in the reference model. In these examples, the effect of $a_{\min}$ on EnKF performance was not as pronounced as the other two parameters (in particular, the direction of anisotropy).

The results of experiments A3, A4, and A5 suggest that the performance of EnKF in estimating permeability values is more robust and reliable when the initial ensemble represents the full range of possible structural variability. When the uncertainty is not accounted for and deterministic variogram models are applied, the EnKF updates cannot correct the systematic errors from an incorrect structural assumption. Next, we use the EnKF to estimate the log permeability maps when all three variogram parameters are uncertain, which introduces additional diversity into the initial ensemble.

6.3.2.4 Experiment A6: Accounting for Full Variogram Uncertainty

In this section, uncertainty in all three parameters of a two-dimensional variogram model is accounted for. The initial ensemble of variogram parameters θ, $a_{\max}$, and $a_{\min}$ are generated independently and, with the probability distributions specified in section 3.1, to cover a wide range of structural variability. Each triplet of sample variogram parameter realizations is used to generate one replicate of the initial ensemble of log permeability maps, and $N = 100$ of these maps compose the initial ensemble used in the sequence of EnKF forecasts and updates.

Figure 6.8 summarizes the results. Even though a large variability is assumed in specifying the variogram parameters, the updated log permeability fields can identify the main permeability features in the reference model. It can be seen from

the updated results that the identified features do not have the sharp structural attributes of the reference model. While this can be, in part, due to a lack of sufficient measurements to convey such detailed information, the high level of variability in the initial ensemble can also lead to a "noisy" or "spurious" covariance model, as a contributor to this effect.

The results thus far appear to suggest that the EnKF is, in fact, an effective estimation approach for reconstruction of the spatial distribution of log permeability fields from fluid flow measurements under highly uncertain variogram models. In particular, the performance of the filter seems to be quite encouraging when the full range of structural uncertainty is incorporated in the initial ensemble. When the uncertainty in the prior structural model is not accounted for, the filter fails to correct the initial structural bias; however, even in that case the log permeability values near the observation points appear to be updated correctly.

6.3.3 Application to SPE10 Model (Model C)

In previous examples, we used small synthetic models for which we knew the correct values of the variogram parameters. As our last example, we consider application of the EnKF under highly uncertain variogram models to update a scaled version of layer 3 in the SPE10 model. Figure 6.9a shows the well configuration that we have used for this example. In Figures 6.9b and 6.9c the reference log permeability model and the corresponding saturation profile at the end of the simulation (36 months) are shown, respectively. Details about the simulation and assimilation information can be found in Table 6.1 (fourth column). The high-permeability trend in the reference model seems to be in two different directions, with small $\theta \simeq 20$ and large $\theta \simeq 160$ angles from the horizontal (i.e., x) axis, which can make the description of the permeability field with a single variogram structure

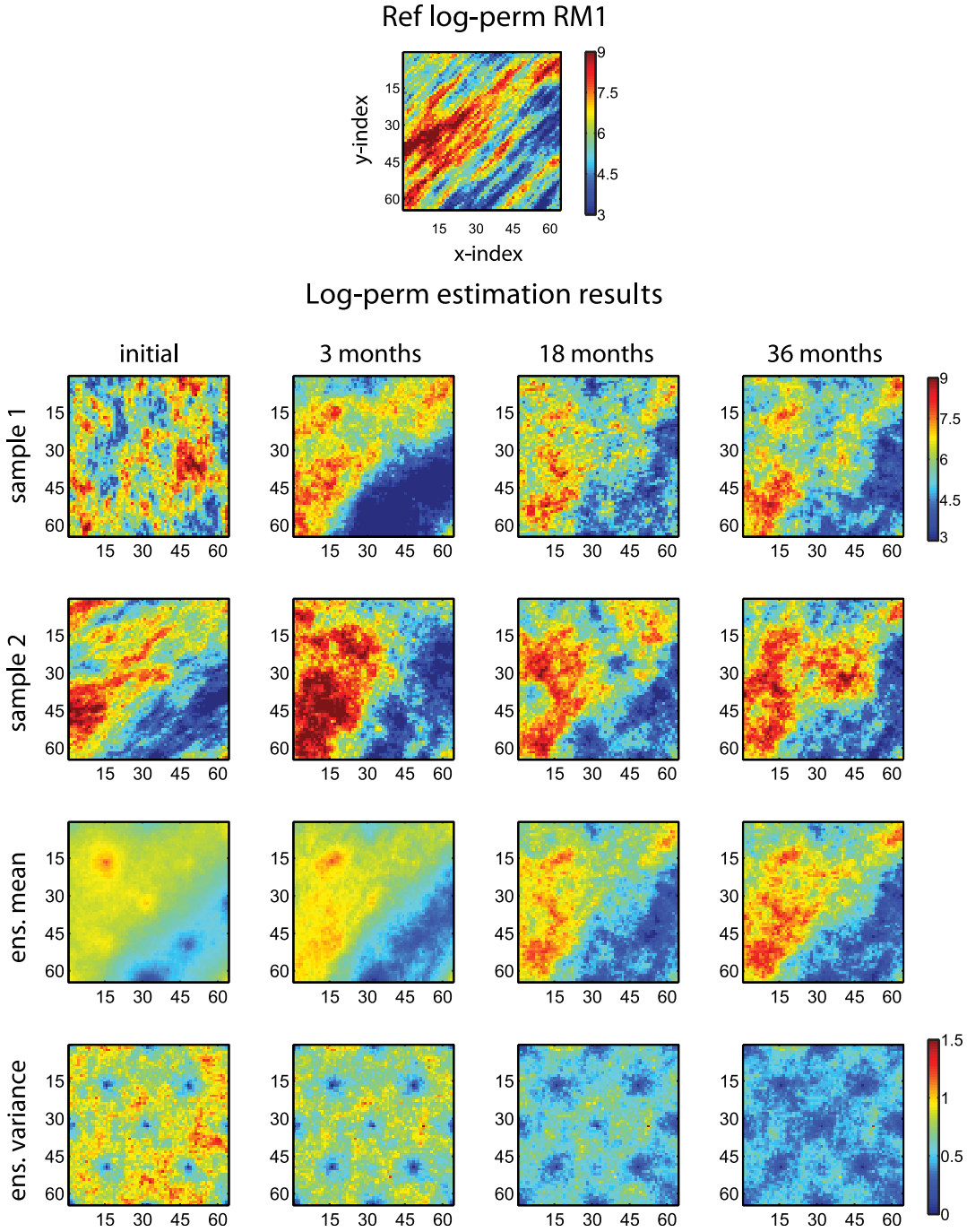


Figure 6.8: The log permeability estimation results for experiment A6 with the EnKF and where all three variogram parameters are assumed to be uncertain and follow the uniform distributions $\theta \sim \mathcal{U}(0°, 180°)$, $a_{\max} \sim \mathcal{U}(100\text{ m}, 1000\text{ m})$, and $a_{\min} \sim \mathcal{U}(10\text{ m}, 1000\text{ m}) \leq a_{\max}$. The estimation results show two sample log permeability realizations, the corresponding log permeability ensemble mean, and ensemble variance. The number of ensemble replicates is $N = 100$.

more difficult. The values of the true variogram model parameters are unknown, and the EnKF is initialized with a very wide ensemble of permeability models using the following variogram parameters: $\theta \sim \mathcal{U}(0°, 180°)$, $a_{\max} \sim \mathcal{U}(100 \text{ m}, 3000 \text{ m})$, and $a_{\min} \sim \mathcal{U}(10 \text{ m}, 3000 \text{ m}) \leq a_{\max}$.

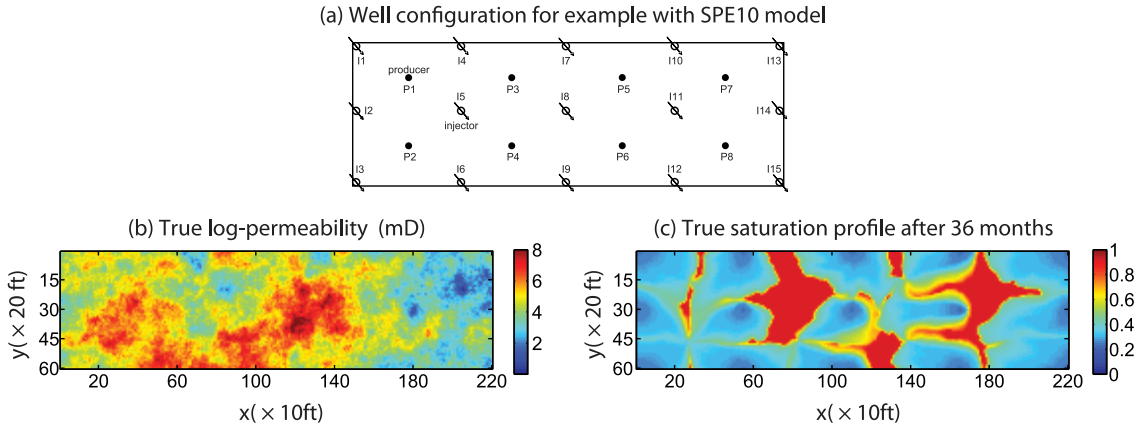


Figure 6.9: (a) Field setup and well configuration, (b) the true log permeability, and (c) the final saturation profile for the top layer of the SPE10 model (model C).

Figures 6.11 and 6.10 shows the estimation results. Displayed in Figure 6.11a is the time evolution of three log permeability samples (first through third columns) and the mean of the log permeability ensemble (fourth column). As can be verified from the first row of Figure 6.11a, the initial ensemble has very diverse and noninformative structures. The initial log permeabilities are generated without conditioning on the hard data to introduce even more variability. Figure 6.11a shows that the final ensemble of permeability maps is consistent with the reference model. Figure 6.10b contains the RMSE, ensemble spread, and scatter plots for the log permeability field. These plots are consistent with the updates to the permeability field and, combined with the log permeability results in Figure 6.11a, reveal significant (structural)

improvements after the EnKF updates. We also carried out this experiment using incorrect (biased) variogram parameters to generate the initial log permeability ensemble, which resulted in significantly degraded performance (not shown). Also not shown here are the results of updating variogram parameters from the flow data, which did not yield reliable updates, consistent with our observation in previous examples.

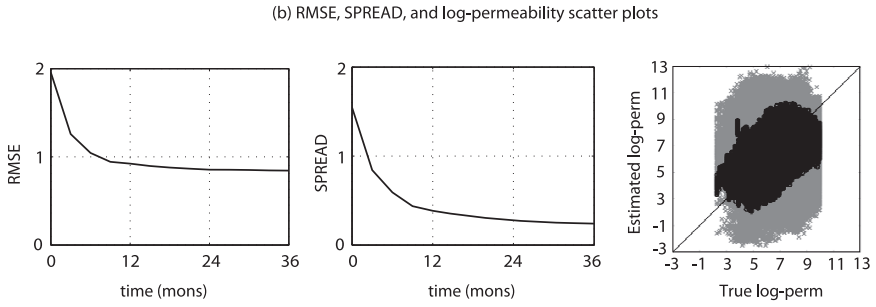


Figure 6.10: Estimation results for experiment C. (b) The RMSE (left), ensemble spread (middle), and scatterplot (right) of the log permeability estimates. The initial ensemble is generated from a variogram model with parameters sampled from the following uniform distributions $\theta \sim \mathcal{U}(0^\circ, 180^\circ)$, $a_{\max} \sim \mathcal{U}(100 \text{ m}, 3000 \text{ m})$, and $a_{\min} \sim \mathcal{U}(10 \text{ m}, 3000 \text{ m}) \leq a_{\max}$. The number of ensemble replicates is $N = 100$.

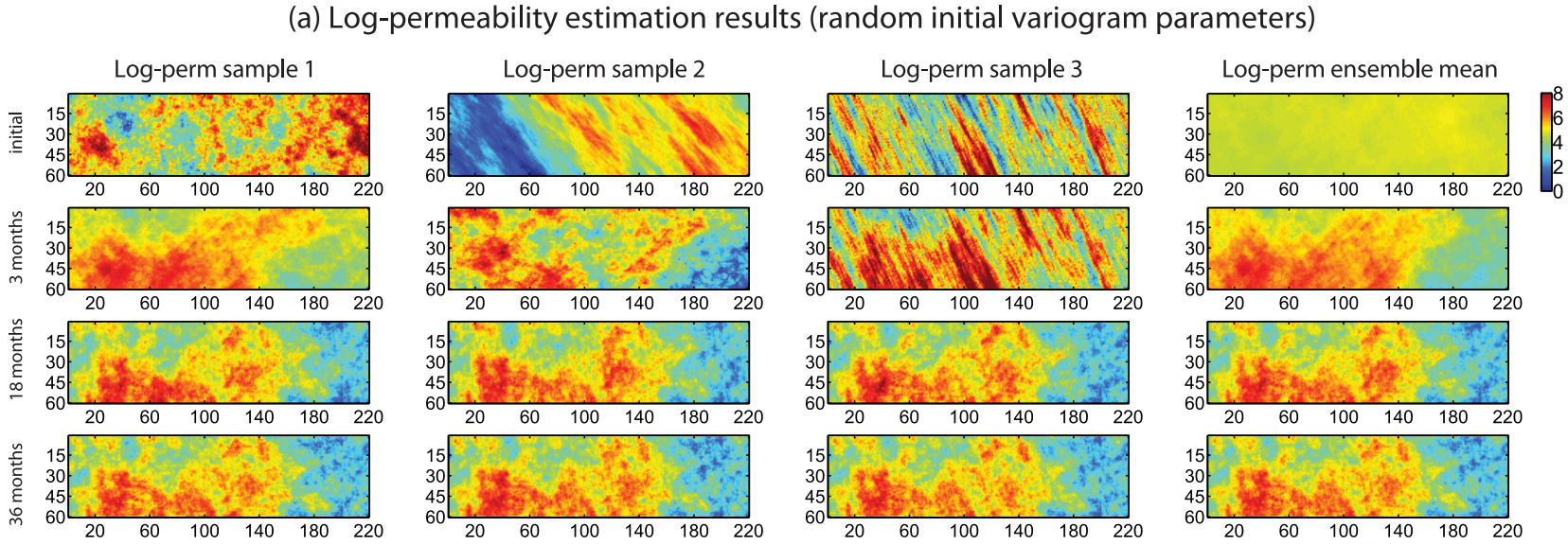


Figure 6.11: Estimation results for experiment C. (a) Sample and ensemble mean of the log permeability updates in time.

6.4 Discussion and Summary

Application of the EnKF to groundwater model calibration and history matching of oil reservoirs has been extensively investigated in recent years. In this chapter, we evaluated the performance of this approach under uncertain variogram models, which is essential for field applications where geologic continuity can only be known with significant uncertainty. Since the structural continuity dominates the global flow patterns in subsurface formations, to obtain reliable solutions and quantify the corresponding uncertainties in a meaningful manner, it is imperative to include the uncertainty about the geologic continuity model in the inversion. The view taken in this work is that the dynamic flow data can be used as an additional source of information to constrain the global variability in subsurface hydraulic properties and the resulting flow and transport behavior. While the type of geologic continuity, and the corresponding conceptual models to describe them, can also be viewed as a major source of uncertainty (for example, Gaussian versus non-Gaussian or fractal random fields), we limited our attention to Gaussian random functions and studied the effect of uncertainty in the variogram model parameters.

We presented a series of numerical experiments to evaluate the performance of the EnKF under significant uncertainty in the variogram model. The importance of the variogram model was first illustrated in a simple example to motivate the topic and the experiments that followed. The first set of experiments consisted of direct estimation of three unknown variogram parameters (i.e., anisotropy direction and maximum and minimum ranges) of a two-dimensional permeability model from flow measurements. The resulting EnKF estimates for these variogram parameters were discouraging, which was explained by the complexity and nonuniqueness of the relation between the flow data and variogram parameters. Additional

one-dimensional experiments were conducted and showed that the flow data had little sensitivity to the variogram range. We presented a different suite of experiments, when updating grid block permeability values, to investigate the sensitivity of the EnKF performance to errors in specification of individual variogram model parameters. Our sensitivity analysis suggests that a structural error (bias) introduced in generating the initial ensemble of permeabilities tends to carry over and persist during the EnKF update and forecast sequence. In particular, the results of our sensitivity study indicate that a large error in specifying variogram parameters, e.g., direction of major continuity, is not corrected throughout the EnKF analysis steps; however, reasonable local updates may be obtained near the measurements, implying that the situation may be improved under spatially dense measurement configurations. Other noteworthy observations from these experiments were the stronger sensitivity of the filter to errors in the direction of anisotropy and higher degradation of the EnKF performance when the maximum variogram range was underestimated rather than when it was overestimated.

In another set of experiments, we investigated the performance of the EnKF for permeability estimation when the initial ensemble of permeabilities was generated by overly uncertain (noninformative) variogram parameters (described via a set of wide independent distributions). The results from these experiments supported the hypothesis that permeability estimation with the EnKF can still be quite promising under significant uncertainty in the initial variogram parameters as long as the initial permeability ensemble accounts for the full range of variability in structural continuity. We explained the observed behavior by noting that the ensemble of updated permeabilities is obtained by combining the prior permeability ensemble with appropriate weights that are determined by the EnKF update equation. Thus, when the initial ensemble has incorrect structural continuity, the updated

permeability ensemble (being constructed by combining the prior permeability ensemble) cannot depart too far from the imposed incorrect initial structure. On the other hand, when the initial permeability ensemble contains structurally diverse realizations, the EnKF can identify the update weights such that after a few updates the new ensemble gradually reveals the correct continuity structure.

In conclusion, the experiments and analysis presented in this chapter suggest that in designing the EnKF initial ensemble it is important to account for the full range of uncertainty in the geologic continuity model. Although introducing significant uncertainty in the prior geologic model significantly reduces the confidence in the initial ensemble, the results in this chapter indicate that this initial uncertainty can be substantially reduced after the first few updates and the structural diversity in the initial ensemble tends to facilitate the reconstruction of the correct geologic continuity in the field. Consequently, in designing the filter for subsurface model calibration, it is safer to overestimate the uncertainty in the variogram model than to underestimate or disregard it at the risk of introducing persistent structural errors. While the flow data seem to contain information about the local distribution of subsurface hydraulic properties as well as the structural continuity of the important features in them, our results suggest that estimating variogram parameters from the complex relation between the flow data and variogram parameters requires more sophisticated inversion approaches than the second-order linear update in the EnKF approach. The findings of this work underscore the need to develop more advanced model calibration methods that can utilize the complex variogram structural information in the flow measurements to consistently improve the description of the geologic continuity model (global trends) as well as the spatial distribution of subsurface properties (local variabilities).

7. USING EFFICIENT PSEUDO-FORECASTS AND GEOLOGICAL MODEL CLUSTERING WITH ENSEMBLE KALMAN FILTER FOR FAST HISTORY MATCHING

7.1 Introduction

Ensemble methods present a practical framework for uncertainty assessment in estimating heterogeneous reservoir properties from production data. In particular, application of ensemble Kalman filter (EnKF) to history matching problems has received extensive attention, primarily because of its simplicity and successful performance in several applications. The EnKF implementation relies on flow predictions with a relatively large number of model realizations, which in the case of realistic reservoir models can be computationally prohibitive. When a small number of model realizations are used the statistics computed for the EnKF model updating step become inaccurate and can lead to inaccurate results. Therefore, several localization methods have been introduced to account for statistical errors due to limited ensemble sizes. While these practical considerations have been useful, they do not address the core issue that a larger ensemble size is needed to accurately compute the required update statistics.

High resolution and fidelity reservoir simulators e.g. ECLIPSE [77] are computationally expensive. So the major computational cost in history matching problems such as EnKF is the time used on propagating the forward model [142, 143]. Moreover to appropriately represent geological model we need to use large number of realizations. Therefore it is expensive to use EnKF with large ensemble size. Since the quality of the covariance estimates depends highly on the size of ensemble, reducing the size of ensemble to accelerate EnKF leads to

inaccurate estimates and ensemble collapse. To alleviate this difficulty previously covariance localization, Kalman gain regularization and improved initial sampling were proposed [144, 145, 146]. These methods work with a few realizations that hardly cover the parameter space. They try to correct the sample covariance or Kalman gain (calculated by limited realizations) that generally are not a good approximation of real covariance or Kalman gain. Although these methods add extra degree of freedom and expand parameter space they are not quite able to make a representative ensemble like what a large ensemble does. Our purpose here is improving the performance and computation time of EnKF using limited ensemble size (limited high fidelity forward model simulations) and more representative initial ensemble along with updating large enough auxiliary ensemble utilizing a fast or pseudo forecast method.

We propose to use a large ensemble of models to improve the calculation of ensemble statistics while using a fast approximate forecast method to reduce the computational cost of the EnKF. The forecast for each realization in this case is derived from linearization around a representative or similar realization for which full simulation is performed. We use adjoint model generated gradient [147, 148] or an ensemble-calculated gradient approximation as tangent linear model for the linearization purpose. In the forecast step of our implementation we perform full forecasts for very few realizations to compute ensemble-based gradients. We then perform approximate (pseudo) forecasts for the remaining models in the ensemble by linearizing about representative models for which full forecasts are computed. The representative realizations for full forecast in each step are selected through a well-established clustering procedure [101, 149, 150].

Our proposed method to reduce the computational time of EnKF is using a fast pseudo model for reservoir simulation that is model linearization around some

representative realizations. Model linearization and using model gradient were applied successfully to history matching through Extended Kalman Filter (EKF), Maximum Likelihood Ensemble Filter (MLEF) and Ensemble Randomized Maximum Likelihood Filter (EnRML) to handle model nonlinearity [151, 152, 153, 154]. For model gradient (sensitivities) we use either Adjoint model (e.g. implemented in ECLIPSE) or ensemble based gradient approximation. Latter method makes our procedure fully independent of forward simulator and does not need extra routine. Here we rectify the problem of inaccurate covariance by using large ensemble size to be more representative of uncertainty space. In each forecast step of EnKF we do full forecast for only a few realizations then calculate ensemble based gradient and we do pseudo forecast for the rest of them. The number of realizations with full forecast is comparatively very fewer than the typical number of realizations which we need to avoid ensemble collapse. The representative realizations for full forecast in each step is selected by clustering procedure consistent with pseudo forecast method. The geological model clustering is also a form of improved sampling which helps covering more parameter space. The (initial) ensemble representativeness (how representative or general the initial ensemble is) has important and determining effect in the performance and success of EnKF procedure [144], therefore an appropriate clustering or model selection method should be used to pick the best realization candidates as the ensemble representatives for high fidelity or full resolution reservoir simulation.

The reminder of this chapter covers first the standard ensemble Kalman filter approach for history matching then the proposed efficient EnKF with pseudo forecast and clustering method. Finally We validate the performance of the proposed method using several numerical experiments in 1D, 2D and 3D dimensional reservoirs including benchmark reservoir models of PUNQ-S3 and SPE10 and demonstrate its

applicability to large-scale model calibration problems.

7.2 Inverse Problem Methodology Development

7.2.1 Ensemble Kalman Filter

Ensemble Kalman filter's formulations and algorithm and its successful application in data assimilation problems were already extensively introduced and presented in the previous chapters. This section is more focused in presenting the proposed improvements and modifications to the standard EnKF to enhance its performance. The probability density function or the statistical characteristics of the reservoir parameter is represented by an ensemble of realizations in EnKF so to better represent the probabilistic characteristic the ensemble should be as large as possible. However due to computational limitations (expensive reservoir simulation run at EnKF propagation step) it is not possible to utilize a large ensemble of reservoir geological models in the history matching problem. Using small ensemble also leads to incorrect and biased estimation results and mostly leads to ensemble collapse. To resolve the issue with small ensemble size and also better represent the initial probability space as well as reduce the computational load of EnKF procedure we propose to perform reservoir simulation only for a few realizations and utilize an efficient pseudo forecast (proxy modeling) to approximate the production data for the rest of the ensemble members.

Through the EnKF analysis step the reservoir state vector becomes updated by integrating the available observations. To build the state vector, only the reservoir parameter and its corresponding observation or reservoir response are used. Therefore for solving this problem, the EnKF with confirmation option or restart EnKF is utilized [137] which is consistent with the proposed fast history matching approach. Consequently it is only needed to approximate the observations (reservoir

response e.g. fluid rates and BHPs) for the clusters members from the cluster representatives full forecast (high fidelity forecast using the reservoir simulator). And the reservoir state variables (e.g. saturations and pressure) are not needed in the integration procedure.

7.2.2 Efficient Ensemble Kalman Filter with Pseudo-Forecasts and Geological Model Clustering

The proposed efficient EnKF approach consists of two main ingredients: 1) pseudo forecast method, 2) geological clustering method. In the proposed algorithm to preserve the ensemble spread and avoid ensemble collapse we keep the typical number of realizations (no reduction in ensemble size) and to improve the computational efficiency we only perform numerical reservoir simulation for a small subset (cluster representatives) of original ensemble. For the rest of the ensemble we develop an efficient and fast pseudo forecast method (based on linearization of the reservoir simulator) to predict the production data. The linearization is performed around the representative realizations which their production forecasts are obtained from simulation. To best select the ensemble representatives we propose to apply a clustering method (based on the similarity of the permeability values) to the ensemble (initially and after each EnKF update) and pick the clusters' centroids as the representative realizations. The detailed proposed efficient EnKF algorithm is presented in Table 7.1.

7.2.2.1 Forward Model Linearization as an Efficient Pseudo-Forecast

Our proposed method for approximating the forward reservoir model (reservoir simulator) is the linearization (gradient-based pseudo forecast) around the cluster centroids. For this purpose we need the gradient of the production data with respect to the reservoir parameter. This gradient can be obtained through ensemble-based

Efficient EnKF Pseudo Code
1: Generate N_e initial parameter ($\mathbf{m}_j$) realizations (initial ensemble)
2: Generate perturbed observations ($\mathbf{y}_j$) from true observation (based on $\mathbf{C}_d$)
3: **For** $t = 1$ **to** t_N **do** (*integration time steps*)
4: Apply k-means clustering to N_e realizations and build k clusters
5: **For** $i = 1$ **to** k **do** (*full reservoir simulation for cluster representatives*)
6: Initialize the reservoir simulator
7: Write the i^{th} cluster's representative (centroid) ($\mathbf{m}_{c,i}$) as the reservoir parameter
8: Run reservoir simulator from beginning until current integration time (corresponds to t)
9: Record the corresponding reservoir response (production data)
10: **End For**
11: Calculate ensemble-based gradient G *
12: Calculate (approximate) the pseudo forecasts for the cluster members
13: Calculate ensemble statistics of $\bar{\mathbf{x}}_f^j$, $\mathbf{C}_{\mathbf{x}}^e$, and Kalman gain $\mathbf{K}$ using all realizations (N_e) data
14: Update all realizations by EnKF analysis equation
15: **End For**

t = integration time step index which corresponds to integration time

t_N = the total number of integration time steps

* If the reservoir simulator has adjoint model and provides gradient data, gradient data is recorded in step 9 and step 11 is removed.

Table 7.1: Efficient EnKF algorithm using pseudo forecasts and clustering for fast history matching.

approximation and directly from the simulator if the adjoint model is implemented in the reservoir simulator (e.g. Eclipse). Here we utilize both approaches which lead to promising estimation results for improved EnKF. The ensemble-based gradient is calculated from the production forecasts of centroids and this approach is more general which can be applied along with any reservoir simulator and does not need adjoint model as well as it is faster and does not have the over head time of adjoint model computation. We investigate both gradient-based pseudo forecast methods (adjoint provided directly by forward model and ensemble based gradient) in this study. The proposed pseudo forecast's computation time advantage over the run time of perfect forecast e.g. Eclipse is very significant.

To estimate the production data (or reservoir response) of the members (realizations) of each cluster, the following equation is applied:

$$\hat{g}(\mathbf{m}_j) \approx g(\mathbf{m}_c) + G(\mathbf{m}_j - \mathbf{m}_c) \tag{7.1}$$

where $g(\cdot)$ is the reservoir simulator response which produces production data from the input parameters e.g. permeability, $\mathbf{m}_j$ is a parameter realization member of a cluster, $\mathbf{m}_c$ is the centroid or representative realization of the corresponding cluster, and G is the gradient (sensitivity matrix) of reservoir response (production data) with respect to reservoir parameters (e.g. permeability). The high fidelity or full reservoir simulation response of the cluster centroid is represented by $g(\mathbf{m}_c)$ and $\hat{g}(\mathbf{m}_j)$ is the linear approximation of $g(\mathbf{m}_j)$ (actual reservoir response of member $\mathbf{m}_j$) around the cluster centroid response. In this study the reservoir responses or observations are injectors BHP and producers oil, water and gas rates and the parameter to estimate is permeability distribution. This approximation procedure is done in each integration step. If the observation is a p-dimensional vector (at each time step) and parameter is an n-dimensional vector then gradient matrix dimension is $p \times n$.

Ensemble-based gradient approximation is performed using ensemble-based (sample) parameter ($\mathbf{m}$) covariance matrix and cross-covariance matrix of reservoir parameter ($\mathbf{m}$) and reservoir response (d or g) [154]:

$$\mathbf{C}_{\mathbf{m},d} \approx \mathbf{C}_{\mathbf{m}} G^T \tag{7.2}$$

where $\mathbf{C}_{\mathbf{m},d}$ is the sample cross-covariance matrix of parameter and observation and its size is $n \times p$ and $\mathbf{C}_{\mathbf{m}}$ is the sample covariance matrix of parameter and its size is $n \times n$. The cross-covariance matrix (that needs the parameter and its corresponding

reservoir response) is calculated with the results of running the simulator for cluster representatives (centroids) and covariance matrix of parameters is calculated with the entire parameter ensemble. This matrix equation is solved for G using *gmres* method [101].

7.2.2.2 Clustering as Improved Sampling and Geological Model Selection Method

To pick the most representative realizations for high fidelity forecast (reservoir simulation), we propose to utilize clustering method. Clustering method alone (as an improved initial sampling approach) will improve EnKF performance [144] by introducing a more representative initial sample space. Effect of improved sampling and the importance of ensemble representativeness is also shown by a numerical example in Section 7.3.2 while comparing the results of two similar size ensembles, one wide spread, and the other narrow spread. We aim to construct an improved ensemble (smaller size) that is more representative of the initial ensemble space. We propose to use k-means clustering method [101, 149, 150] which is a well-established clustering or unsupervised learning method. Clustering is performed based on Euclidean distance of samples (grid block property values e.g. permeability values).

k-means clustering is vector quantization method, originally from signal processing which is popular for cluster analysis in data mining. k-means clustering aims to partition N samples into k clusters in which each sample belongs to the cluster with the nearest mean which is serving as a prototype or representative of the cluster.

Given a set of samples $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N)$ (e.g. an ensemble of permeability realizations), where each observation is an n-dimensional real vector, k-means clustering aims to partition the N samples into k ($\leq N$) sets (or clusters) $\mathbf{S} =$

$\{S_1, S_2, ..., S_k\}$ so as to minimize the within-cluster sum of squares (WCSS). In other words, its objective is to find:

$$\underset{\mathbf{S}}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \mu_i\|^2 \tag{7.3}$$

where μ_i is the mean of points or samples in S_i.

The most common algorithm uses an iterative refinement technique. Given an initial set of k means (initially they can be k randomly generated samples), the algorithm proceeds by alternating between two steps: [155]

Assignment step: Assign each sample to the cluster whose mean yields the least within-cluster sum of squares (WCSS). Since the sum of squares is the squared Euclidean distance, this is intuitively the "nearest" mean.

Update step: Calculate the new means to be the centroids of the samples in the new clusters. (Since the arithmetic mean is a least-squares estimator, this also minimizes the within-cluster sum of squares (WCSS) objective.)

The commonly used initialization methods are Forgy and Random Partition [156]. The Forgy method randomly chooses k samples from the data set and uses these as the initial means. The Random Partition method first randomly assigns a cluster to each observation and then proceeds to the update step, thus computing the initial mean to be the centroid of the cluster's randomly assigned points.

With this algorithm each cluster's representative is the clusters centroid which is the mean of the cluster members. Therefore there is no guaranty that the cluster centroid is a member of the cluster. Since we aim to perform reservoir simulation with one of the reservoir parameter realizations (namely as the centroid of each cluster), we pick the closest (Euclidean distance) member of each cluster to the cluster's centroid as the cluster's representative.

Clustering is performed at each update step, the ensemble representatives or cluster centroids are updated at each integration or update step i.e. a new set of clusters are made after each EnKF update.

Clustering method and measure are in fact compatible with pseudo forecast method. Clustering is based on the Euclidean distance of geological models and cluster the realizations based on similarity and closeness of pixel values which is compatible with the proposed pseudo forecast method that linearize the forward model around each cluster representative (or cluster centroid) and calculate (approximate) the forecast based on the difference of each (e.g. permeability) realization with its cluster representative.

7.3 Numerical Experiments

In this section, a series of 1D, 2D and 3D reservoir model examples are presented to verify the utility of the proposed EnKF method for fast history matching. A 1D reservoir model case with homogeneous permeability is presented with numerous trial simulations to demonstrate that the promising performance of the proposed method is statistically consistent.

In all performed experiments by the proposed method there is no rank deficiency (in all updates states) in cross covariance matrix of simulated observations and permeabilities (or the matrix of state vectors ensemble). It means by clustered EnKF method when the number of realizations is N_e and the number of clusters or full forecasts is k, then always the rank of state ensemble matrix is N_e. So by linearization of some realizations observations around centroids no rank deficiency in state ensemble matrix happens.

7.3.1 1D Reservoir Model with Homogeneous Permeability

In this example to show the statistically consistent performance of the proposed method, we set up a simple 1D homogeneous reservoir and repeat the EnKF experiments several times and compare and analyze the mean of the performed experiments. The reservoir model is 1D ($100 \times 1 \times 1$ grid configuration) with 2-phase fluid (oil and water). The permeability (the parameter to estimate) is homogeneous so in the estimation procedure we estimate a single value. There are two wells at the two ends of the reservoir (one injector and one producer). The total simulation (history matching) period is 180 days (6 regularly spaced integration steps). Production is at constant BHP and the water injection scenario is shown in Figure 7.1.

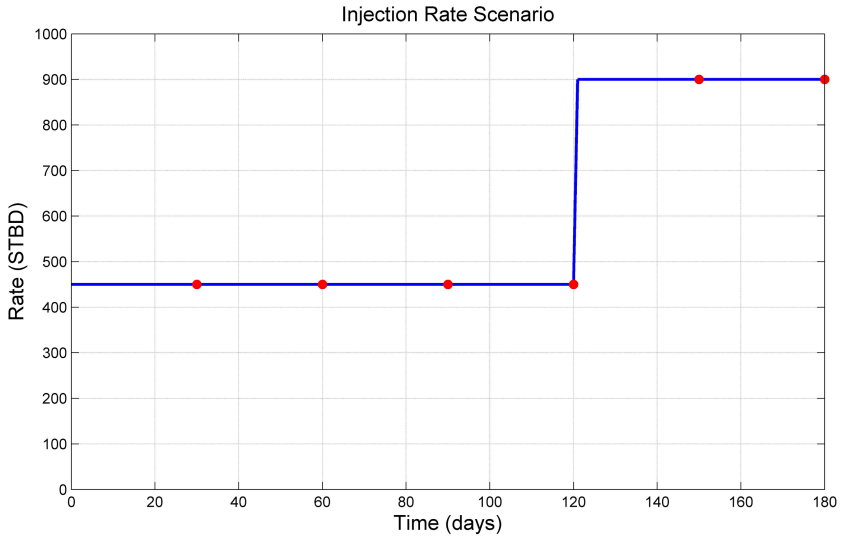


Figure 7.1: Water injection rate scenario.

EnKF with confirmation step is used i.e. in all integration steps, propagation of

realizations is performed from the beginning). At each integration step there are 3 observations (oil rate, water cut and injection BHP). In EnKF procedure here, we only have one parameter to estimate (homogeneous permeability field). This very simple model setup was chosen because the permeability estimation results (just one number) is completely representative of the estimation performance (no need for ensemble forecasts) and the ensemble based gradient estimation (calculation) is also very straight forward (just estimating one value for each observation which is its sensitivity with respect to homogeneous permeability). By this simple setup we also could run large number of experiments to show that our results are statistically consistent.

The homogeneous permeability of the true case is 150 (md). To make initial ensemble, we generate random numbers (each random number is a realization) from a Gaussian distribution with $\mu = 400$ (md) and $\sigma = 200$ (md). There are three different setups:

1. Small ensemble: simple EnKF with 3 realizations (ensemble size = 3)
2. Large ensemble: simple EnKF with 20 realizations (ensemble size = 20)
3. Clustered EnKF: improved EnKF by clustering and pseudo forecast with 20 realizations and 3 clusters (ensemble size = 20 and 3 full forecasts)

To show statistically consistent results we performed 100 trial experiments for each setup. The difference in each try or experiment is the chosen initial ensemble. We made an initial large ensemble of realizations and then for each try we randomly picked 20 realizations then ran all three mentioned setups. Totally 300 simulations (i.e. 300 EnKF experiments) were done.

Experiment setup (100 tries for each)	Mean of final ensemble means	Std of final ensemble means	Mean of final ensemble RMSE
1 (Small ensemble)	93.57	78.41	73.51
2 (Large ensemble)	147.18	4.97	4.42
3 (Clustered EnKF)	144.94	18.76	7.02

True permeability = 150 (md)

Table 7.2: The final estimation results of all trial experiments.

As shown in the Table 7.2, the result of clustered EnKF is significantly better than small ensemble setup and is slightly worse than large ensemble setup. Also obtaining these results from 100 experiments for each setup confirms that improvement by clustered EnKF method is generally promising. The histograms in Figure 7.2 are also indicating improved performance of clustered EnKF approach.

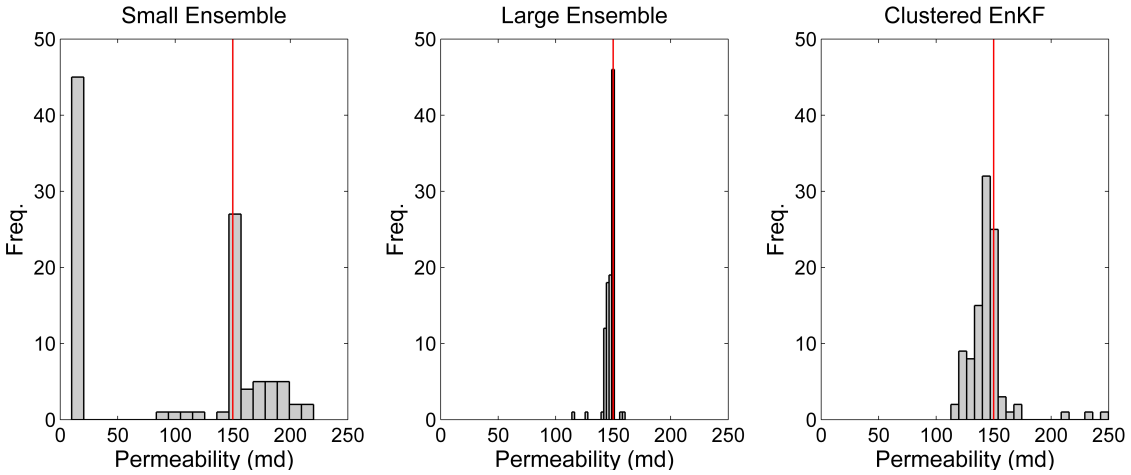


Figure 7.2: Histograms of final ensemble mean of all trial experiments for each setup (red line shows the true permeability value).

7.3.2 1D Reservoir Model with Heterogeneous Permeability

This example is a 1D reservoir model ($100 \times 1 \times 1$ grid configuration) with 5 wells (2 injectors and 3 producers). Constant BHP is set for producers and the water injection scenario is show in Figure 7.3. To generate initial ensemble, *sgsim* algorithm (unconditional simulation) is used. History matching is done for 8 integration steps (each step is 90 days). Production forecast is performed for 16 steps. The heterogeneous permeability is the parameter to estimate by integrating injector's BHP and producers oil and water rate.

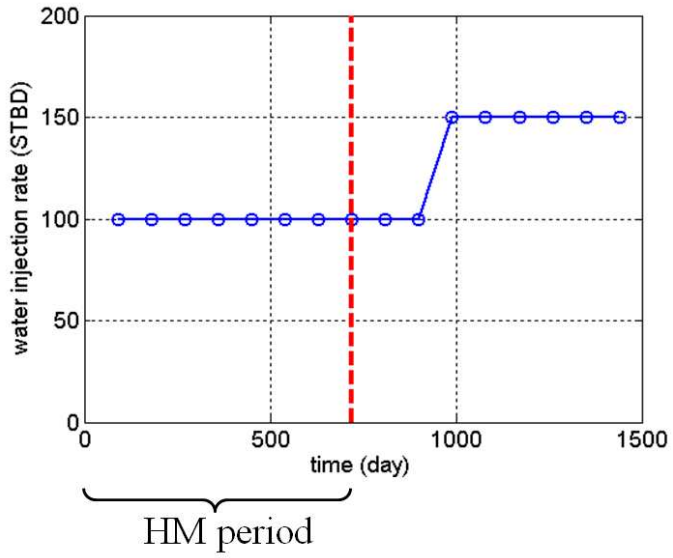


Figure 7.3: Water injection rate scenario.

In this example we also demonstrate the effect of initial improved sampling in enhancing the EnKF performance through providing more representative initial ensemble that covers more of possible initial parameter space than randomly selected initial ensemble. After generating 100 realizations we perform following experiments.

There are 4 setups for this model:

1. Small ensemble EnKF (10 realizations)
2. Small ensemble EnKF with improved initial sampling (10 realizations which are the centroids of 100 realizations. We perform clustering on 100 realizations and pick 10 most representative samples.)
3. Clustered EnKF (100 realizations with 10 clusters)
4. Large ensemble EnKF (100 realizations)

Experiment 1 is actually the simplest setup which results in the poorest performance. In experiment 1 we just run EnKF by 10 realizations picked randomly from initial 100 realizations. By an initial clustering of 100 realizations to 10 clusters then picking the centroids of the clustered we perform improved initial sampling to build a more representative initial ensemble even though the ensemble size is still small.

The estimation results of all 4 experiments in terms of Lperm estimation RMSE are shown in Figure 7.4. As expected, improved initial sampling resulted in less RMSE comparing to the random ensemble and the clustered EnKF RMSE is significantly lower than the small ensemble case and is comparable to large ensemble.

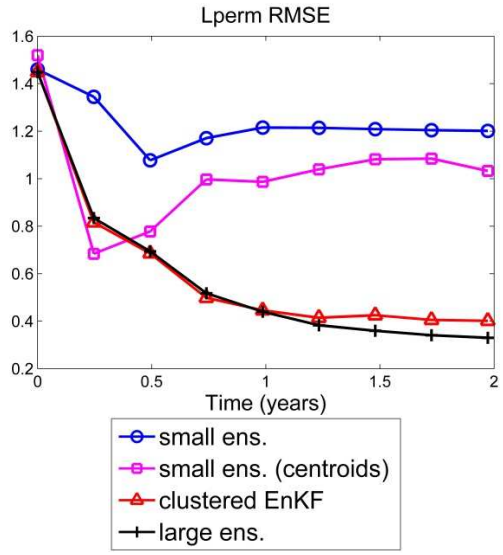


Figure 7.4: Comparison of Lperm RMSE of 4 experiments.

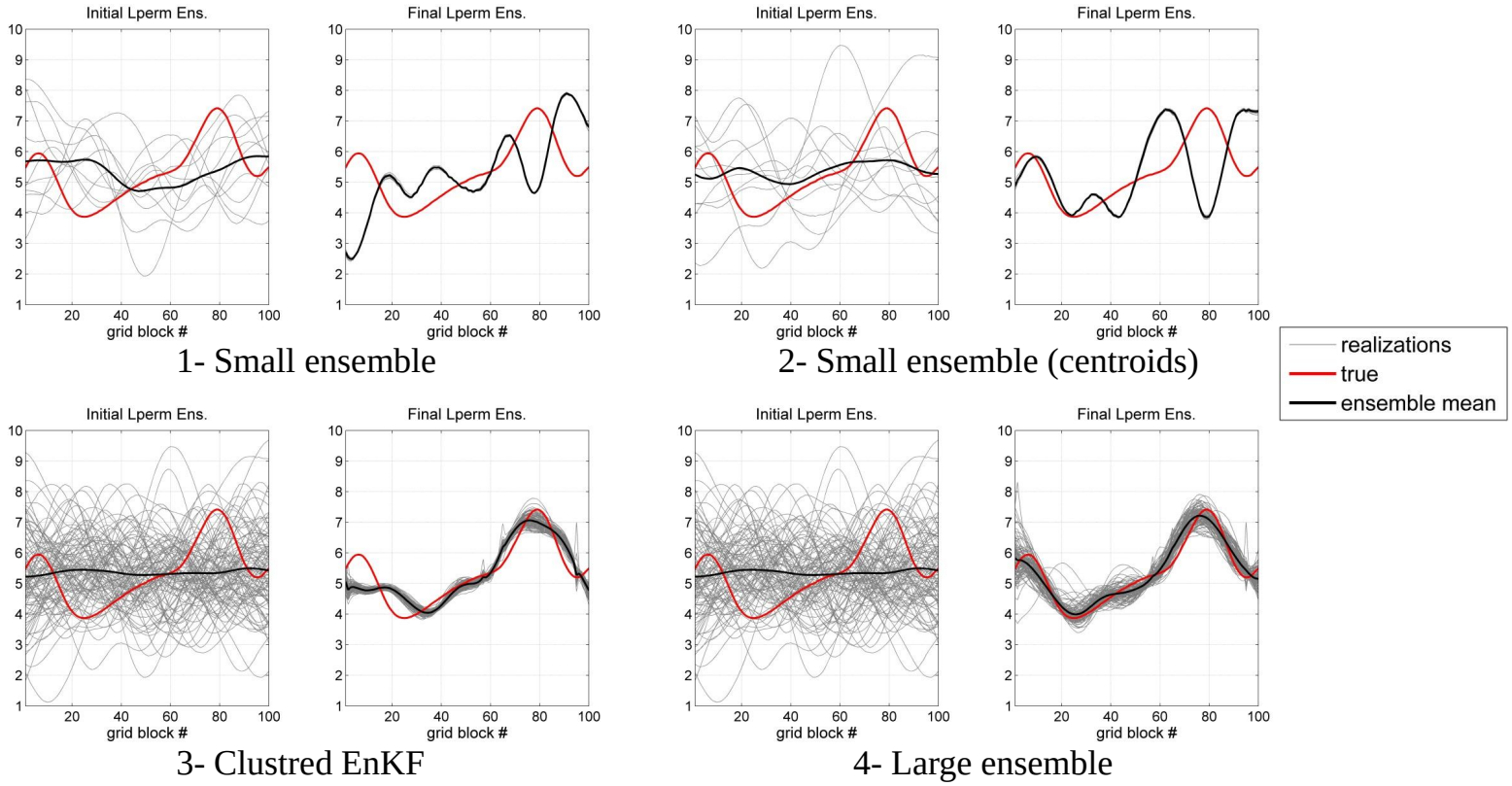


Figure 7.5: Lperm ensemble estimation results (initial and final ensemble in all 4 experiments).

The estimation results shown in Figure 7.5, also verifies the promising performance of improved initial sampling and the improved EnKF method.

Water rate of one of the production wells (initial and final ensemble forecast) is shown in Figure 7.6 for all 4 experiments which again verifies the utility of the proposed EnKF method for reservoir data assimilation.

7.3.3 2D Reservoir Model

The example presented in this section is a 2D 2-phase reservoir model with $100 \times 100 \times 1$ grid configuration. The well pattern is a single 9-spot (one injector and 8 producers) and unconditional simulation is used to generate the initial ensemble. The history matching period is 8 years (12 integration steps and each step is 8 months). The true Lperm distribution is also shown in Figure 7.7.

The goal here is to estimate the permeability distribution by integrating the production data. We perform three experiments as follows:

1. Small ensemble EnKF with 20 realizations
2. Clustered (improved) EnKF with 200 realizations and 20 clusters
3. Large ensemble EnKF with 200 realizations

Figure 7.7 shows the estimated results of these 3 experiments in terms of the final estimated mean. As shown the estimated permeability map of clustered EnKF approach is significantly closer to the true permeability map comparing to the small ensemble experiment. The final estimated mean of clustered EnKF is also comparable with the large ensemble case while its computation time is much less that verifies the applicability of the proposed method for fast history matching.

The estimation RMSE and ensemble spread of three experiments are shown in Figure 7.8. As it is shown by ensemble spread plots, the small ensemble EnKF almost

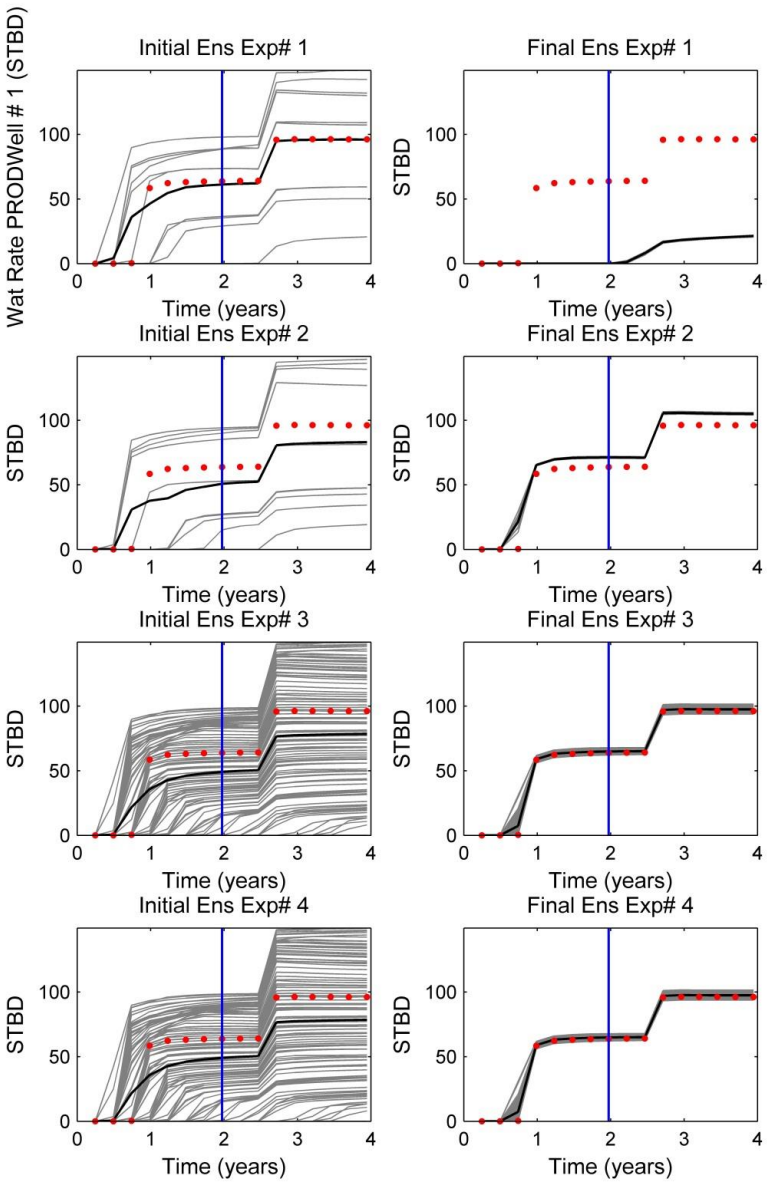


Figure 7.6: Comparison of ensemble forecast for all 4 experiments (water rate); vertical blue line indicates the end of history matching period.

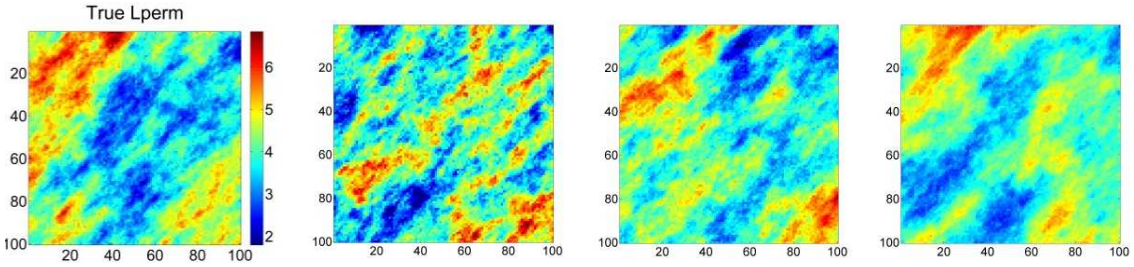


Figure 7.7: Comparison of final (after 12 integration steps) estimated ensemble mean for all 3 experiments.

collapsed. The RMSE of the small ensemble case also shows unacceptable result (not continuously decreasing) while the clustered EnKF RMSE is very close to the large ensemble experiment.

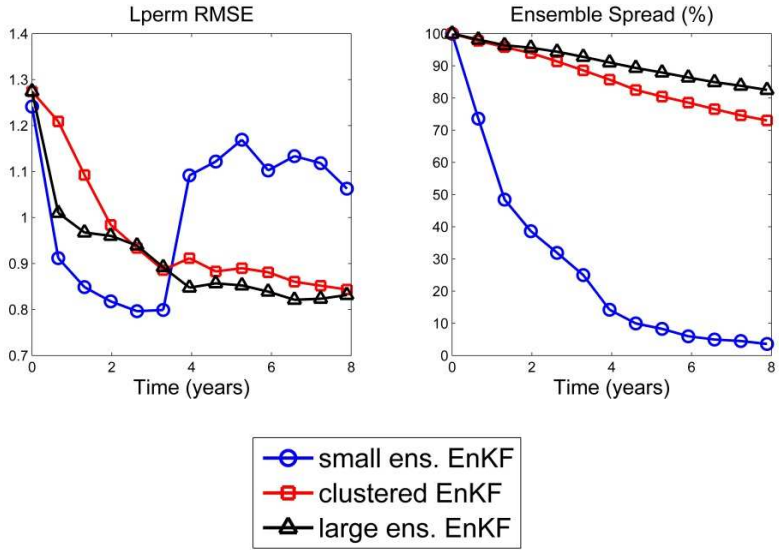


Figure 7.8: Comparison of permeability estimation RMSE and ensemble spread of three experiments.

7.3.4 PUNQ-S3 Model

In this section the proposed improved EnKF method is applied to the PUNQ-S3 reservoir model [157]. This is a 3D and 3-phase reservoir model with $19 \times 28 \times 5$ grid configuration. There are 1761 active grid blocks and 6 producers with no injectors. We are simultaneously estimating X, Y and Z permeabilities and also the porosity by integrating BHP, oil rate, water cut and GOR. We consider 8 years of history matching and 16.5 years of production forecast. The ensemble size is 200 and we consider 20 clusters (i.e. the number of full reservoir simulations or full forecasts). Therefore in each integration time we perform 20 full forecasts and 180 pseudo forecasts. For this example ensemble-based estimated gradient is used.

Again for this example we perform three experiment setups to show the utility of the proposed method:

1. Small ensemble EnKF with 20 realizations

2. Clustered (improved) EnKF with 200 realizations and 20 clusters

3. Large ensemble EnKF with 200 realizations

The true porosity map, the reservoir and wells configuration are shown in Figure 7.9.

As shown in Figure 7.10, the final estimated result of EnKF with small ensemble fails to detect the major features of the true porosity map while the improved EnKF is successful in capturing the shape and low/high value regions of the true map and has close estimation result to the large ensemble EnKF case.

The ensemble variance results of the experiments in Figure 7.11 also show the promising performance of improved (clustered) EnKF in preserving the ensemble

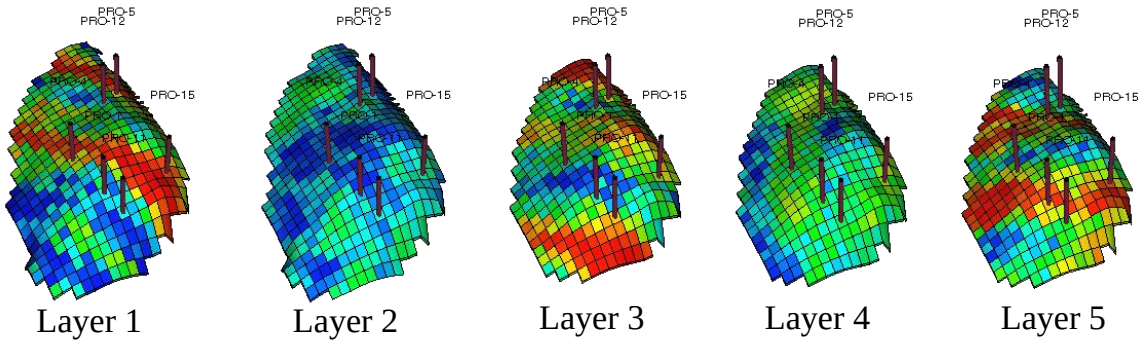


Figure 7.9: Reservoir model configuration and the true porosity maps of PUNQ-S3 model in all layers.

spread while it only performs 20 full reservoir forecasts. From Figure 7.11 it is also clear that small ensemble EnKF collapsed.

As presented in Table 7.3, the proposed improved EnKF approach significantly expedite the history matching process and makes the data integration procedure considerably faster.

Experiment	Computation time (hr)
EnKF (20 realizations)	0.39
EnKF (200 realizations)	3.87
EnKF (200 realizations, 20 clusters)	1.07

Table 7.3: Computation time comparison of the performed experiments (machine specifications: Intel Xeon CPU 3.07 GHz, 6.0 GB RAM).

To verify the promising performance of the proposed improved EnKF method in providing the proper production forecasts, the boxplots of ensemble forecasts are presented in Figure 7.12. It is clear from this figure that the improved EnKF provides

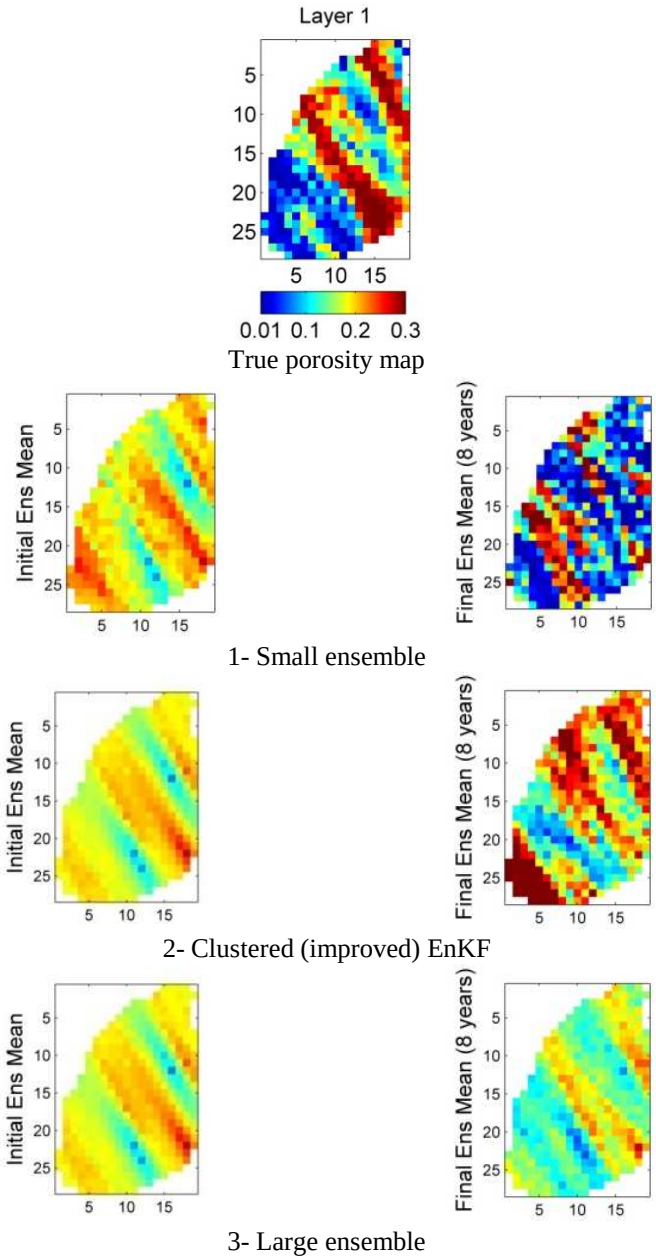


Figure 7.10: Estimation results in terms of estimated ensemble mean (porosity of layer 1).

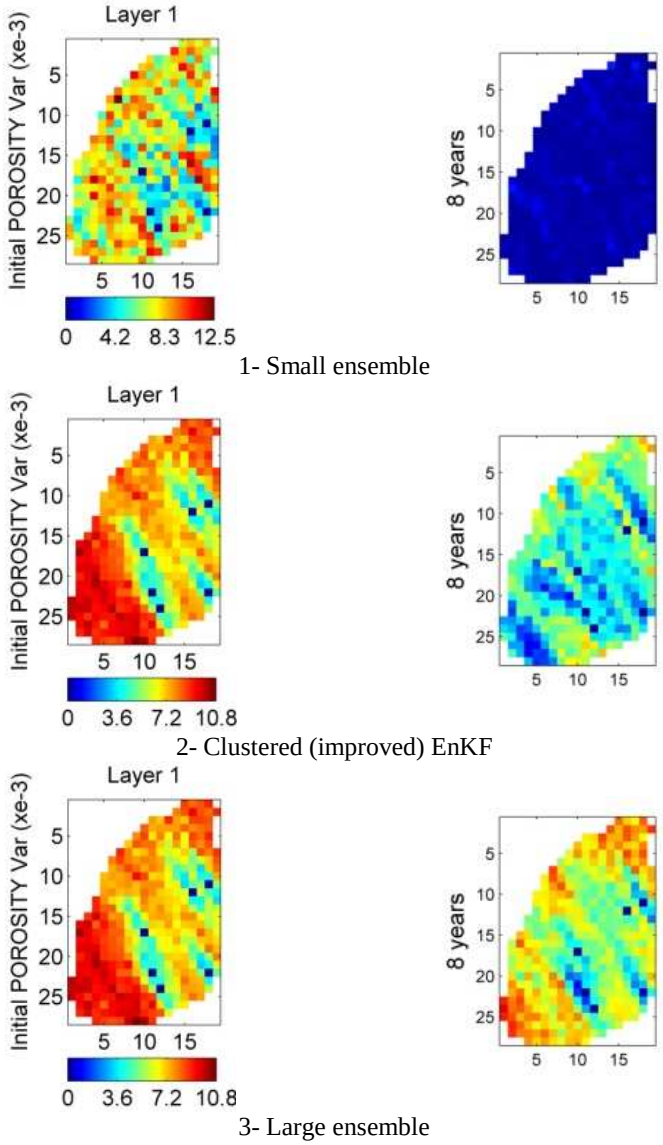


Figure 7.11: Estimated ensemble variance map (porosity of layer 1).

less biased and much closer forecasts comparing to the small ensemble EnKF. Also the results of improved EnKF is comparable with the large ensemble EnKF.

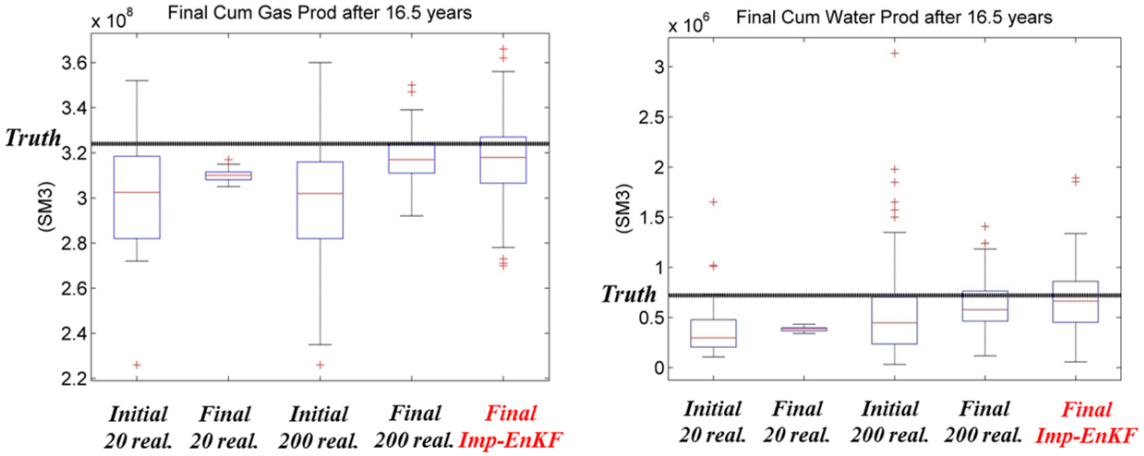


Figure 7.12: Performance comparison in terms of ensemble forecasts of final cumulative gas and water production by boxplot.

7.3.5 SPE10 Model

In this section we apply the proposed method to SPE10 model. We use 5 top layers of SPE10 model and put 23 wells in the model. The reservoir is 3-phase black oil model and the grid configuration is $60 \times 220 \times 5$ (66000 grid blocks). The unknown parameter to estimate in the permeability distribution and BHP, oil, water and gas rates are integrated. Similar to the previous examples, three sets of experiments are performed. The small and large ensemble sizes are 200 and 20, respectively. The clustered (improved) EnKF is done with 200 realizations and 20 clusters. To generate the ensemble of permeability realizations conditional simulation is used.

The reservoir and well configurations are shown in Figure 7.13. There are both water and gas injection wells in the reservoir. Again to show the promising estimation

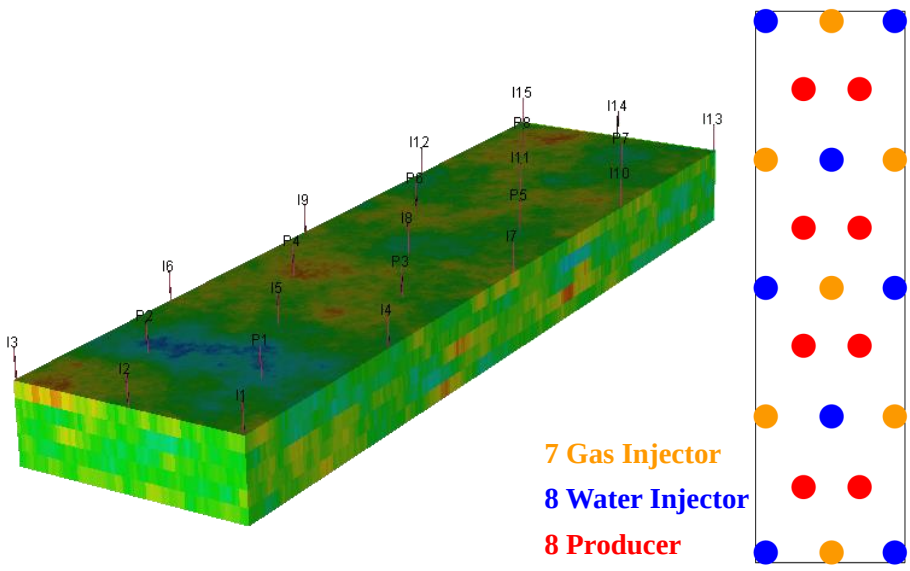


Figure 7.13: SPE10 model and well configuration and specifications.

performance of the proposed method, we conduct three experiments, first with small ensemble (20 realizations) and then with a large ensemble (200 realizations) and finally we perform the improved EnKF with 200 realizations and 20 clusters. The estimation results in terms of estimated ensemble mean and ensemble variance map are shown in Figure 7.14. True log-permeability distribution, initial ensemble mean and initial ensemble variance map are presented and also the final ensemble mean and final ensemble variance map of all three experiments are shown too. Firstly it is clear from the variance map of small ensemble experiment that ensemble collapse happened. The estimated ensemble mean of small ensemble case is also very different from the true map. The final ensemble mean mean of the improved EnKF is very well in agreement with the one from large ensemble experiment and both are close to the true map. The final ensemble variance map of improved EnKF experiment also proves that the proposed method is successful in preserving the spread of the ensemble.

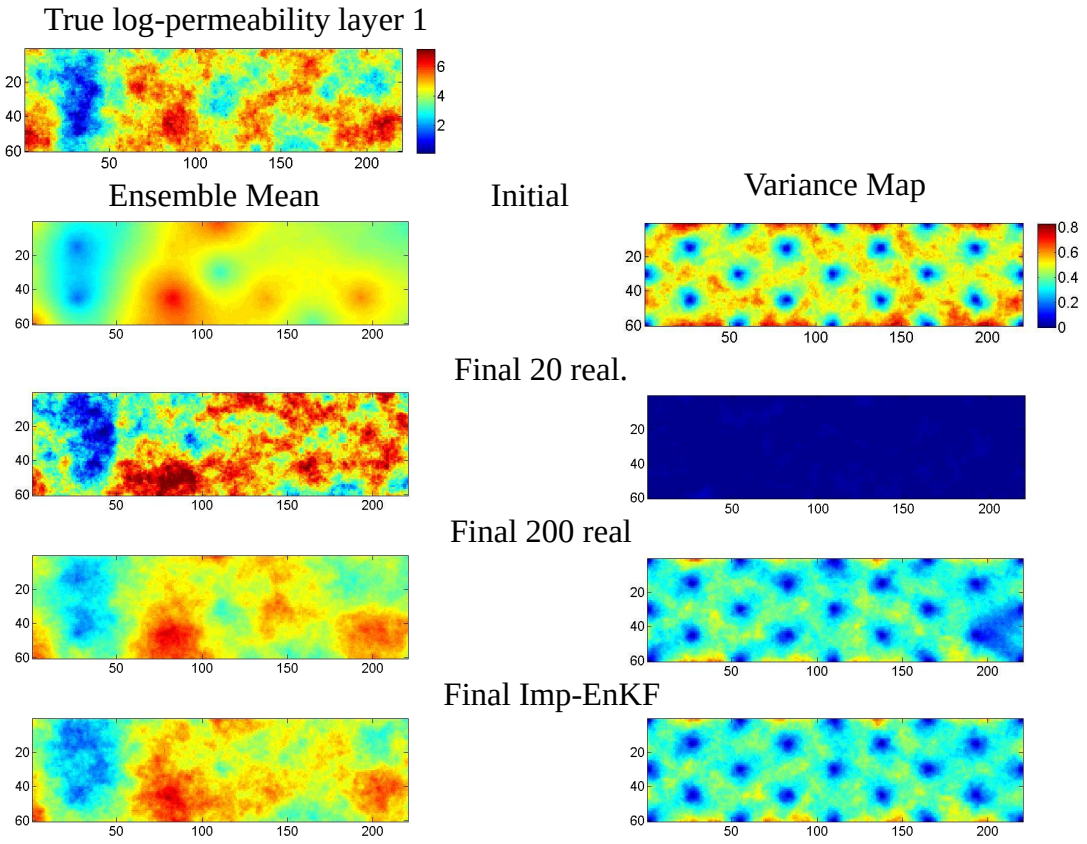


Figure 7.14: Comparison of the estimated ensemble mean and ensemble variance map for the three experiments.

Successful application of the proposed EnKF algorithm for SPE10 model proves the utility of this approach for fast history matching of realistic large-scale reservoirs.

7.4 Discussions and Summary

In this chapter to develop a computationally efficient EnKF-based history matching method, we propose to use efficient and fast pseudo forecasts along with geological model clustering in EnKF framework. The performing Monte Carlo simulation or propagation step of EnKF is the major computation load of EnKF procedure. Therefore to reduce the computation time we perform high fidelity reservoir forecast (reservoir simulation) for a few reservoir model realizations and for the rest of them we approximate the production data by a fast proxy method. We choose to linearize the reservoir forward model (pseudo forecast approach) around some representative samples which their forecasts results are obtained by reservoir simulation. For linearization we obtain the gradient of production data with respect to reservoir parameters either from adjoint model (implemented in reservoir simulator) or ensemble-based gradient calculation. A clustering method is also used to group the original ensemble to some subsets or clusters and choose the clusters' representatives for reservoir simulation. With ensemble-based gradient approximation method, this approach can be applied to any type of reservoir simulator (with any observation or parameter) for model updating. This approach showed significant improvement in computational load of EnKF procedure and made it considerably faster. We applied this approach along with Eclipse as the reservoir simulator to a suit of different reservoir models including two benchmark reservoir models and verified its promising estimation performance comparing to small and large ensemble EnKF.

8. SUMMARY AND CONCLUSION

Seismicity-based reservoir characterization is a promising approach for monitoring and improving reservoir performance in a number of important energy and environmental applications. We formulated an EnKF-based model calibration approach to integrate discrete MEQ events into the description of reservoir property distributions in both pore-pressure diffusion and fully coupled geomechanical forward model. Since the EnKF is a continuous estimation approach, we introduced a new interpretation of the MEQ event locations as a continuous seismicity density map that is amenable to assimilation with the EnKF. A main advantage of the EnKF to previously introduced SBRC methods is that it is a stochastic inversion that provides an ensemble of solutions to facilitate uncertainty assessment. Other important advantages of the EnKF are the ability to systematically incorporate uncertainty in models and observations, and its generality for application under any forward model, failure criteria, and MEQ event triggering mechanisms. In addition, the simple and versatile implementation of the EnKF allows for estimation of different types of parameters from various data types.

In this work, we developed a coupled geomechanical reservoir simulator with rock failure criteria and damage mechanics model, and focused on developing a framework called stochastic seismicity-based reservoir characterization (SSBRC) for automatic and robust integration of MEQ-type discrete data sets using the EnKF. We first developed 2D and 3D FEM fully coupled thermo-poro-elastic models with Mohr-Coulomb failure criterion (including tension cut-off) and, permeability and damage model to relate hydraulic and geomechanical reservoir parameters to discrete microseismicity cloud. An important property of the EnKF is that its sequential

update scheme provides different representations of unknown parameters after each update. By construction, the EnKF is designed to update time-varying states of a system. In forward geomechanical model that rock damage is reflected in the alteration of elastic modulus and permeability, field stress disturbances change the rock physical properties (parameters) with time, the EnKF-type sequential filtering techniques prove quite useful for estimation of dynamically varying parameters. We also set up 2D and 3D pore-pressure diffusion forward models using a finite difference based commercial reservoir simulator that relates permeability distribution to MEQ cloud distribution.

In pore-pressure forward model, by applying SSBRC we can infer permeability distribution from MEQ data. Considering the coupled flow and geomechanics-based forward model, SSBRC method is capable of inferring permeability, elastic modulus, tensile strength, cohesion and friction angle from MEQ event cloud. Here, both tensile and shear failures are considered as microseismicity events.

Using KDE to generate seismicity density map on the same fine grid system of pore pressure diffusion model or fine mesh of FEM model leads to high-dimensional and redundant observation. EnKF update with large number of correlated observations results in severe ensemble spread underestimation. We proposed three methods to preserve the ensemble spread and improve uncertainty quantification of SSBRC. We resolved this issue by either artificially adding large random noise to observation or reducing the number of observations by spectral and spatial dimension reduction. As the simplest method of avoiding ensemble collapse, we proposed to inflate observation error variance. We also proposed projecting the microseismic data onto a low-dimensional subspace that is defined by left singular vectors of the perturbed observations matrix, and lastly using coarse-scale continuous representation of the microseismic data. The proposed KDE approach for

transforming the discrete MEQ data in this work inevitably introduces some error into the estimation results. A more natural estimation approach for integration of MEQ data is one that does not convert the discrete events into continuous measurements. Developing discrete data integration algorithms can eliminate the discrete data quantification step and potentially lead to additional improvements in the estimation results.

SSBRC also showed a promising performance in inferring reservoir properties and also hydraulic fracture characteristics from MEQ data when applied to hydraulically fractured reservoirs. Reservoir matrix permeability distribution was successfully estimated from microseismic data by SSBRC in reservoir setting with multiple transverse induced fractures. SSBRC was also successfully employed for microseismic fracture mapping in place of the traditional methods.

Here, we adopted a continuum approach and did not include fracture systems in the estimation. In general, one may need to characterize fractures and fracture networks as part of the model calibration process. Further refinements of the workflow may also be possible by including seismic modeling as one of the components in the inversion framework that can help better characterize the MEQ events both in terms of their distribution and intensity. Analyses of the raw microseismic data can lead to additional information about the induced fractures and their properties.

The experiments and analysis presented in Section 6 suggest that in designing the EnKF initial ensemble it is important to account for the full range of uncertainty in the geologic continuity model. Although introducing significant uncertainty in the prior geologic model significantly reduces the confidence in the initial ensemble, the results indicate that this initial uncertainty can be substantially reduced after the first few updates and the structural diversity in the initial ensemble tends to facilitate the reconstruction of the correct geologic continuity in the field. Consequently, in

designing the filter for subsurface model calibration, it is safer to overestimate the uncertainty in the variogram model than to underestimate or disregard it at the risk of introducing persistent structural errors. While the flow data seem to contain information about the local distribution of subsurface hydraulic properties as well as the structural continuity of the important features in them, our results suggest that estimating variogram parameters from the complex relation between the flow data and variogram parameters requires more sophisticated inversion approaches than the second-order linear update in the EnKF approach. The findings of this work underscore the need to develop more advanced model calibration methods that can utilize the complex variogram structural information in the flow measurements to consistently improve the description of the geologic continuity model (global trends) as well as the spatial distribution of subsurface properties (local variabilities).

In Section 7 to develop a computationally efficient EnKF-based history matching method, we propose to use efficient and fast pseudo forecasts along with geological model clustering in EnKF framework. The performing Monte Carlo simulation or propagation step of EnKF is the major computation load of EnKF procedure. Therefore to reduce the computation time we perform high fidelity reservoir forecast (reservoir simulation) for a few reservoir model realizations and for the rest of them we approximate the production data by a fast proxy method. We choose to linearize the reservoir forward model (pseudo forecast approach) around some representative samples which their forecasts results are obtained by reservoir simulation. For linearization we obtain the gradient of production data with respect to reservoir parameters either from adjoint model (implemented in reservoir simulator) or ensemble-based gradient calculation. A clustering method is also used to group the original ensemble to some subsets or clusters and choose the clusters' representatives for reservoir simulation. With ensemble-based gradient approximation method, this

approach can be applied to any type of reservoir simulator (with any observation or parameter) for model updating. This approach showed significant improvement in computational load of EnKF procedure and made it considerably faster. We applied this approach along with Eclipse as the reservoir simulator to a suit of different reservoir models including two benchmark reservoir models and verified its promising estimation performance comparing to small and large ensemble EnKF.

REFERENCES

[1] E. L. Majer and J. E. Peterson, "The impact of injection on seismicity at the geysers, california geothermal field," *International Journal of Rock Mechanics and Mining Sciences*, vol. 44, no. 8, pp. 1079–1090, 2007.

[2] M. D. Zoback and H.-P. Harjes, "Injection-induced earthquakes and crustal stress at 9 km depth at the ktb deep drilling site, germany," *Journal of Geophysical Research*, vol. 102, no. B8, pp. 18 477–18, 1997.

[3] M. Fehler, L. House, W. S. Phillips, and R. Potter, "A method to allow temporal variation of velocity in travel-time tomography using microearthquakes induced during hydraulic fracturing," *Tectonophysics*, vol. 289, no. 1, pp. 189–201, 1998.

[4] P. Audigane, J.-J. Royer, and H. Kaieda, "Permeability characterization of the soultz and ogachi large-scale reservoir using induced microseismicity," *Geophysics*, vol. 67, no. 1, pp. 204–211, 2002.

[5] J. Rutledge and W. Phillips, "A comparison of microseismicity induced by gel-proppant-and water-injected hydraulic fractures, carthage cotton valley gas field, east texas," in *2002 SEG Annual Meeting*, 2002.

[6] J. T. Rutledge and W. S. Phillips, "Hydraulic stimulation of natural fractures as revealed by induced microearthquakes, carthage cotton valley gas field, east texas," *Geophysics*, vol. 68, no. 2, pp. 441–452, 2003.

[7] C. Pearson, "The relationship between microseismicity and high pore pressures during hydraulic stimulation experiments in low permeability granitic rocks,"

Journal of Geophysical Research: Solid Earth (1978–2012), vol. 86, no. B9, pp. 7855–7864, 1981.

[8] J. Albright and C. Pearson, "Acoustic emissions as a tool for hydraulic fracture location: Experience at the fenton hill hot dry rock site," *Old SPE Journal*, vol. 22, no. 4, pp. 523–530, 1982.

[9] C. I. Trifu, *The mechanism of induced seismicity*. Springer, 2002, vol. 159, no. 1-3.

[10] A. Dahi Taleghani and J. Lorenzo, "An alternative interpretation of microseismic events during hydraulic fracturing," in *SPE Hydraulic Fracturing Technology Conference*, 2011.

[11] W. Gale, K. Heasley, A. Iannacchione, P. Swanson, P. Hatherly, and A. King, "Rock damage characterization from microseismic monitoring," in *Proceedings of the 38th US Rock Mechanics Symposium, Washington, DC, 7–10 July*, 2001, pp. 1–313.

[12] G. Foulger, "Hengill triple junction, sw iceland 2. anomalous earthquake focal mechanisms and implications for process within the geothermal reservoir and at accretionary plate boundaries," *Journal of Geophysical Research: Solid Earth (1978–2012)*, vol. 93, no. B11, pp. 13 507–13 523, 1988.

[13] C. Tang, "Numerical simulation of progressive rock failure and associated seismicity," *International Journal of Rock Mechanics and Mining Sciences*, vol. 34, no. 2, pp. 249–261, 1997.

[14] J. Savage and L. Mansinha, "Radiation from a tensile fracture," *Journal of Geophysical Research*, vol. 68, no. 23, pp. 6345–6358, 1963.

[15] M. Cai, P. K. Kaiser, and C. D. Martin, "A tensile model for the interpretation of microseismic events near underground openings," in *Seismicity Caused by Mines, Fluid Injections, Reservoirs, and Oil Extraction.* Springer, 1999, pp. 67–92.

[16] C. Tang and P. Kaiser, "Numerical simulation of cumulative damage and seismic energy release during brittle rock failurepart i: fundamentals," *International Journal of Rock Mechanics and Mining Sciences*, vol. 35, no. 2, pp. 113–121, 1998.

[17] V. Oye and M. Roth, "Automated seismic event location for hydrocarbon reservoirs," *Computers & Geosciences*, vol. 29, no. 7, pp. 851–863, 2003.

[18] G. Foulger, B. Julian, D. Hill, A. Pitt, P. Malin, and E. Shalev, "Non-double-couple microearthquakes at long valley caldera, california, provide evidence for hydraulic fracturing," *Journal of Volcanology and Geothermal Research*, vol. 132, no. 1, pp. 45–71, 2004.

[19] L. Vandamme, S. Talebi, and R. Young, "Monitoring of a hydraulic fracture in a south saskatchewan oil-field," *Journal of Canadian Petroleum Technology*, vol. 33, no. 1, pp. 27–34, 1994.

[20] N. Warpinski, J. Uhl, and B. Engler, "Review of hydraulic fracture mapping using advanced accelerometer-based receiver systems," in *Proceedings, US DOE Natural Gas Conference, Houston, TX*, 1997.

[21] N. Warpinski, W. SL, and C. Wright, "Analysis and prediction of microseismicity induced by hydraulic fracturing," in *SPE Annual Technical Conference and Exhibition*, 2001.

[22] L. Gutierrez-Negrin and J. Quijano-Leon, "Analysis of seismicity in the los humeros, mexico, geothermal field," *Geothermal Resources Council Transactions*, vol. 28, pp. 467–472, 2003.

[23] R. Pine and A. Batchelor, "Downward migration of shearing in jointed rock during hydraulic injections," in *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, vol. 21, no. 5. Elsevier, 1984, pp. 249–263.

[24] M. Tarrahi and B. Jafarpour, "Inference of geothermal reservoir properties from micro-seismic events with ensemble kalman filter," in *AGU Fall Meeting Abstracts*, vol. 1, 2011, p. 1163.

[25] ——, "Inference of permeability distribution from injection-induced discrete microseismic events with kernel density estimation and ensemble kalman filter," *Water Resources Research*, vol. 48, no. 10, p. W10506, 2012.

[26] S. Shapiro, "Microseismicity: a tool for reservoir characterization," *Education tour Series CIS*, vol. 2, 2007.

[27] S. H. Lee and A. Ghassemi, "Thermo-poroelastic analysis of injection-induced rock deformation and damage evolution," in *Proceedings Thirty-Fifth Workshop on Geothermal Reservoir Engineering*, 2010.

[28] S. Lee and A. Ghassemi, "Thermo-poroelastic finite element analysis of rock deformation and damage," in *43th US Rock Mechanics/Geomechanics Symposium*, 2009.

[29] ——, "A three-dimensional thermo-poro-mechanical finite element analysis of a wellbore on damage evolution," in *44th US Rock Mechanics Symposium and 5th US-Canada Rock Mechanics Symposium*, 2010.

[30] ——, "Three-dimensional thermo-poro-mechanical modeling of well stimulation and induced microseismicity," in *45th US Rock Mechanics/Geomechanics Symposium*, 2011.

[31] M. A. Biot, "General theory of three-dimensional consolidation," *Journal of applied physics*, vol. 12, no. 2, pp. 155–164, 1941.

[32] A. Ghassemi and A. Diek, "Linear chemo-poroelasticity for swelling shales: theory and application," *Journal of Petroleum Science and Engineering*, vol. 38, no. 3, pp. 199–212, 2003.

[33] A. Selvadurai, "Stationary damage modelling of poroelastic contact," *International journal of solids and structures*, vol. 41, no. 8, pp. 2043–2064, 2004.

[34] L. Shiping, L. Yushou, L. Yi, W. Zhenye, and Z. Gang, "Permeability-strain equations corresponding to the complete stressstrain path of yinzhuang sandstone," in *International journal of rock mechanics and mining sciences & geomechanics abstracts*, vol. 31, no. 4. Pergamon, 1994, pp. 383–391.

[35] T. Kiyama, H. Kita, Y. Ishijima, T. Yanagidani, K. Aoki, and T. Sato, "Permeability in anisotropic granite under hydrostatic compression and triaxial compression including post-failure region," in *2nd North American Rock Mechanics Symposium*, 1996.

[36] F. Coste, A. Bounenni, S. Chanchole, and K. Su, "A method for measuring mechanical, hydraulic and hydromechanical properties during damaging in materials with low permeability," in *Hydromechanical and Thermohydromechanical Behaviour of Deep Argillaceous Rock: Theory and Experiments: Proceedings of the International Workshop on Geomechanics, Paris, France, 11-12 October 2000*. Taylor & Francis, 2002, p. 109.

[37] M. D. Zoback and J. D. Byerlee, "The effect of microcrack dilatancy on the permeability of westerly granite," *Journal of Geophysical Research*, vol. 80, no. 5, pp. 752–755, 1975.

[38] C. Tang, L. Tham, P. Lee, T. Yang, and L. Li, "Coupled analysis of flow, stress and damage (fsd) in rock failure," *International Journal of Rock Mechanics and Mining Sciences*, vol. 39, no. 4, pp. 477–489, 2002.

[39] S. H. Wang, T. H. Yang, L. G. Tham, C. A. Tang, and L. C. Li, "Simulation of multiple hydraulic fracturing in non-uniform pore pressure field," *Advanced Materials Research*, vol. 9, pp. 163–172, 2005.

[40] J. C. Jaeger, N. G. Cook, and R. Zimmerman, *Fundamentals of rock mechanics*. Wiley. com, 2009.

[41] A. Nur and J. R. Booker, "Aftershocks caused by pore fluid flow?" *Science*, vol. 175, no. 4024, pp. 885–887, 1972.

[42] J. B. Fletcher and L. R. Sykes, "Earthquakes related to hydraulic mining and natural seismic activity in western new york state," *Journal of Geophysical Research*, vol. 82, no. 26, pp. 3767–3780, 1977.

[43] M. Ohtake, "Seismic activity induced by water injection at matsushiro, japan," *J. Phys. Earth*, vol. 22, pp. 163–176, 1974.

[44] P. Talwani and S. Acree, "Pore pressure diffusion and the mechanism of reservoir-induced seismicity," *Pure and Applied Geophysics*, vol. 122, no. 6, pp. 947–965, 1984.

[45] S. A. Shapiro, S. Rentsch, and E. Rothert, "Fluid-induced seismicity: Theory, modeling, and applications," *Journal of engineering mechanics*, vol. 131, no. 9, pp. 947–952, 2005.

[46] ——, "Characterization of hydraulic properties of rocks using probability of fluid-induced microearthquakes," *Geophysics*, vol. 70, no. 2, pp. F27–F33, 2005.

[47] S. Shapiro, R. Patzig, E. Rothert, and J. Rindschwentner, "Triggering of seismicity by pore-pressure perturbations: Permeability-related signatures of the phenomenon," *Pure and applied geophysics*, vol. 160, no. 5-6, pp. 1051–1066, 2003.

[48] M. Parotidis, S. Shapiro, and E. Rothert, "Back front of seismicity induced after termination of borehole fluid injection," *Geophysical Research Letters*, vol. 31, no. 2, p. L02612, 2004.

[49] S. A. Shapiro, E. Huenges, and G. Borm, "Estimating the crust permeability from fluid-injection-induced seismic emission at the ktb site," *Geophysical Journal International*, vol. 131, no. 2, pp. F15–F18, 1997.

[50] S. A. Shapiro, E. Rothert, V. Rath, and J. Rindschwentner, "Characterization of fluid transport properties of reservoirs using induced microseismicity," *Geophysics*, vol. 67, no. 1, pp. 212–220, 2002.

[51] E. Rothert and S. A. Shapiro, "Microseismic monitoring of borehole fluid injections: Data modeling and inversion for hydraulic properties of rocks," *Geophysics*, vol. 68, no. 2, pp. 685–689, 2003.

[52] D. W. Scott, *Multivariate density estimation: theory, practice, and visualization*. Wiley. com, 2009, vol. 383.

[53] M. M. P. Wand and M. C. Jones, *Kernel smoothing*. Crc Press, 1995, vol. 60.

[54] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics," *Journal of*

Geophysical Research: Oceans (1978–2012), vol. 99, no. C5, pp. 10 143–10 162, 1994.

[55] ——, *Data assimilation: the ensemble Kalman filter*. Springer, 2009.

[56] R. H. Reichle, D. B. McLaughlin, and D. Entekhabi, "Hydrologic data assimilation with the ensemble kalman filter," *Monthly Weather Review*, vol. 130, no. 1, pp. 103–114, 2002.

[57] G. Evensen and P. J. Van Leeuwen, "Assimilation of geosat altimeter data for the agulhas current using the ensemble kalman filter with a quasigeostrophic model," *Monthly Weather Review*, vol. 124, no. 1, pp. 85–96, 1996.

[58] P. L. Houtekamer and H. L. Mitchell, "Data assimilation using an ensemble kalman filter technique," *Monthly Weather Review*, vol. 126, no. 3, pp. 796–811, 1998.

[59] H. Madsen and R. Canizares, "Comparison of extended and ensemble kalman filters for data assimilation in coastal area modelling," *International Journal for Numerical Methods in Fluids*, vol. 31, no. 6, pp. 961–981, 1999.

[60] Y. Chen and D. Zhang, "Data assimilation for transient flow in geologic formations via ensemble kalman filter," *Advances in Water Resources*, vol. 29, no. 8, pp. 1107–1122, 2006.

[61] H. Hendricks Franssen and W. Kinzelbach, "Real-time groundwater flow modeling with the ensemble kalman filter: Joint estimation of states and parameters and the filter inbreeding problem," *Water Resources Research*, vol. 44, no. 9, 2008.

[62] W. Nowak, "Best unbiased ensemble linearization and the quasi-linear kalman ensemble generator," *Water Resources Research*, vol. 45, no. 4, 2009.

[63] A. Schöniger, W. Nowak, and H.-J. Hendricks Franssen, "Parameter estimation by ensemble kalman filters with transformed data: Approach and application to hydraulic tomography," *Water Resources Research*, vol. 48, no. 4, 2012.

[64] N. v. Geir, L. Johnsen, S. Aanonsen, and E. Vefring, "Reservoir monitoring and continuous model updating using ensemble kalman filter," in *SPE Annual Technical Conference and Exhibition*, 2003.

[65] X.-H. Wen and W. Chen, "Real-time reservoir model updating using ensemble kalman filter," in *SPE reservoir simulation symposium*, 2005.

[66] B. Jafarpour and D. McLaughlin, "Estimating channelized-reservoir permeabilities with the ensemble kalman filter: the importance of ensemble design," *Spe Journal*, vol. 14, no. 2, pp. 374–388, 2009.

[67] S. Aanonsen, G. Naevdal, D. Oliver, A. Reynolds, and B. Valles, "Reservoir monitoring and continuous model updating using ensemble kalman filter," *SPE J*, vol. 14, no. 3, pp. 393–412, 2009.

[68] B. Jafarpour and M. Tarrahi, "Assessing the performance of the ensemble kalman filter for subsurface flow data integration under variogram uncertainty," *Water Resources Research*, vol. 47, no. 5, 2011.

[69] G. Evensen, "The ensemble kalman filter for combined state and parameter estimation," *Control Systems, IEEE*, vol. 29, no. 3, pp. 83–104, 2009.

[70] M. Ehrendorfer, "A review of issues in ensemble-based kalman filtering," *Meteorologische Zeitschrift*, vol. 16, no. 6, pp. 795–818, 2007.

[71] S. Aanonsen, I. Aavatsmark, T. Barkve, A. Cominelli, R. Gonard, O. Gosselin, M. Kolasinski, and H. Reme, "Effect of scale dependent data correlations in

an integrated history matching loop combining production data and 4d seismic data," in *SPE Reservoir Simulation Symposium*, 2003.

[72] J.-A. Skjervheim, G. Evensen, S. Aanonsen, B. O. Ruud, and T.-A. Johansen, "Incorporating 4d seismic data in reservoir simulation models using ensemble kalman filter," *SPE journal*, vol. 12, no. 3, pp. 282–292, 2007.

[73] P. Sakov and P. R. Oke, "A deterministic formulation of the ensemble kalman filter: an alternative to ensemble square root filters," *Tellus A*, vol. 60, no. 2, pp. 361–371, 2008.

[74] I. MYRSETH, "Robust ensemble kalman filter for high dimensional data."

[75] S. H. Lee and A. Ghassemi, "Threedimensional thermo-poro-mechanical modeling of reservoir stimulation and induced microseismicity in geothermal reservoir," in *Thirty Sixth Workshop on Geothermal Reservoir Engineering*, 2011.

[76] M. A. Biot, "Mechanics of deformation and acoustic propagation in porous media," *Journal of applied physics*, vol. 33, no. 4, pp. 1482–1498, 1962.

[77] "Eclipse reference manual and technical description," Schlumberger-GeoQuest, Houston, Tex, Tech. Rep., 2010.

[78] J. R. Rice and M. P. Cleary, "Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents," *Reviews of Geophysics*, vol. 14, no. 2, pp. 227–241, 1976.

[79] G. Van der Kamp and J. Gale, "Theory of earth tide and barometric effects in porous formations with compressible grains," *Water Resources Research*, vol. 19, no. 2, pp. 538–544, 1983.

[80] C. Cryer, "A comparison of the three-dimensional compaction theories of biot and terzaghi," *Quarterly Journal of Mechanics and Applied Mathematics*, vol. 16, pp. 401–412, 1963.

[81] S. Yuan and J. Harrison, "A review of the state of the art in modelling progressive mechanical breakdown and associated fluid flow in intact heterogeneous rocks," *International Journal of Rock Mechanics and Mining Sciences*, vol. 43, no. 7, pp. 1001–1022, 2006.

[82] L. Kachanov, "Time of the rupture process under creep conditions," *Isv. Akad. Nauk. SSR. Otd Tekh. Nauk*, vol. 8, pp. 26–31, 1958.

[83] J. Mazars, "A description of micro-and macroscale damage of concrete structures," *Engineering Fracture Mechanics*, vol. 25, no. 5, pp. 729–737, 1986.

[84] A. Simakin and A. Ghassemi, "Modelling deformation of partially melted rock using a poroviscoelastic rheology with dynamic power law viscosity," *Tectonophysics*, vol. 397, no. 3, pp. 195–209, 2005.

[85] T. Yang, L. Tham, C. Tang, Z. Liang, and Y. Tsui, "Influence of heterogeneity of mechanical properties on hydraulic fracturing in permeable rocks," *Rock mechanics and rock engineering*, vol. 37, no. 4, pp. 251–275, 2004.

[86] Y. Hamiel, V. Lyakhovsky, and A. Agnon, "Poroelastic damage rheology: Dilation, compaction, and failure of rocks," *Geochemistry, Geophysics, Geosystems*, vol. 6, no. 1, 2005.

[87] D. McTigue, "Thermoelastic response of fluid-saturated porous rock," *Journal of Geophysical Research: Solid Earth (1978–2012)*, vol. 91, no. B9, pp. 9533–9542, 1986.

[88] V. Palciauskas and P. Domenico, "Characterization of drained and undrained response of thermally loaded repository rocks," *Water Resources Research*, vol. 18, no. 2, pp. 281–290, 1982.

[89] Y. Cansi, "An automatic seismic event processing for detection and location: The pmcc method," *Geophysical research letters*, vol. 22, no. 9, pp. 1021–1024, 1995.

[90] X. Zhou and A. Ghassemi, "Finite element analysis of coupled chemo-poro-thermo-mechanical effects around a wellbore in swelling shale," *International Journal of Rock Mechanics and Mining Sciences*, vol. 46, no. 4, pp. 769–778, 2009.

[91] J. C. Heinrich and D. W. Pepper, *Intermediate finite element method: fluid flow and heat transfer applications*. Taylor & Francis, 1999.

[92] J.-A. Wang and H. Park, "Fluid permeability of sedimentary rocks in a complete stress–strain process," *Engineering geology*, vol. 63, no. 3, pp. 291–300, 2002.

[93] L. Eisner, P. M. Duncan, W. M. Heigl, and W. R. Keller, "Uncertainties in passive seismic monitoring," *The Leading Edge*, vol. 28, no. 6, pp. 648–655, 2009.

[94] D. K. Fagan, "Statistical clustering of microseismic event spectra to identify subsurface structure," Ph.D. dissertation, Boise State University, 2012.

[95] E. Auger, F. Aubin, and V. Rajic, "Microseismic data: Understanding the uncertainty," *Hart's E & P*, vol. 84, no. 11, 2011.

[96] A. M. Fish, "Microseismic velocity inversion and event location using reverse time imaging," Ph.D. dissertation, Colorado School of Mines, 2012.

[97] S. Maxwell, "Microseismic location uncertainty," *CSEG Recorder, April*, 2009.

[98] F. Scherbaum, "Combined inversion for the three-dimensional q structure and source parameters using microearthquake spectra," *Journal of Geophysical Research: Solid Earth (1978–2012)*, vol. 95, no. B8, pp. 12 423–12 438, 1990.

[99] M. R. Safari, O. Trevor, C. Queena, C. Hamed, N. Blair, M. Uno, C. Hawkes, *et al.*, "Effects of depletion/injection induced stress changes on natural fracture reactivation," in *47th US Rock Mechanics/Geomechanics Symposium*. American Rock Mechanics Association, 2013.

[100] M. R. Tingay, R. R. Hillis, C. K. Morley, R. C. King, R. E. Swarbrick, and A. R. Damit, "Present-day stress and neotectonics of brunei: Implications for petroleum exploration and production," *AAPG Bulletin*, vol. 93, no. 1, pp. 75–100, 2009.

[101] V. Matlab, "7.12. 0 (r2011a) the mathworks inc," *Natick, Massachusetts*, 2012.

[102] R. E. Kalman *et al.*, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[103] A. Gelb, *Applied optimal estimation*. The MIT press, 1974.

[104] C. Deutsch and A. Journel, "Geostatistical software library and users guide (gslib)," 1998.

[105] J. Vargas-Guzmán and T.-C. J. Yeh, "Sequential kriging and cokriging: Two powerful geostatistical approaches," *Stochastic Environmental Research and Risk Assessment*, vol. 13, no. 6, pp. 416–435, 1999.

[106] P. K. Kitanidis and E. G. Vomvoris, "A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations," *Water Resources Research*, vol. 19, no. 3, pp. 677–690, 1983.

[107] T.-C. J. Yeh, A. L. Gutjahr, and M. Jin, "An iterative cokriging-like technique for ground-water flow modeling," *Ground Water*, vol. 33, no. 1, pp. 33–41, 1995.

[108] W. Graham and D. McLaughlin, "Stochastic analysis of nonstationary subsurface solute transport: 2. conditional moments," *Water Resources Research*, vol. 25, no. 11, pp. 2331–2355, 1989.

[109] J. D. Kepert, "On ensemble representation of the observation-error covariance in the ensemble kalman filter," *Ocean Dynamics*, vol. 54, no. 6, pp. 561–569, 2004.

[110] T. M. Hamill, J. S. Whitaker, and C. Snyder, "Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter," *Monthly Weather Review*, vol. 129, no. 11, pp. 2776–2790, 2001.

[111] A. H. Jazwinski, "Stochastic processes and filtering theory. 1970."

[112] G. Evensen, "Sampling strategies and square root analysis schemes for the enkf," *Ocean Dynamics*, vol. 54, no. 6, pp. 539–560, 2004.

[113] M. K. Fisher, C. A. Wright, B. M. Davidson, N. P. Steinsberger, W. S. Buckler, A. Goodwin, E. O. Fielder, *et al.*, "Integrating fracture mapping technologies to improve stimulations in the barnett shale," *SPE Production & Facilities*, vol. 20, no. 02, pp. 85–93, 2005.

[114] M. J. Mayerhofer, E. P. Lolon, J. E. Youngblood, J. R. Heinze, *et al.*, "Integration of microseismic-fracture-mapping results with numerical fracture network production modeling in the barnett shale," in *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2006.

[115] C. Cipolla, C. Wright, *et al.*, "Diagnostic techniques to understand hydraulic fracturing: what? why? and how?" *SPE production & facilities*, vol. 17, no. 01, pp. 23–35, 2002.

[116] C. R. Clarkson, "integration of rate-transient and microseismic analysis for unconventional gas reservoirs: Where reservoir engineering meets geophysics," *CSEG Recorder*, 2011.

[117] L. Bennett, J. Calvez, D. Sarver, K. Tanner, W. Birk, G. Waters, J. Drew, G. Michaud, P. Primiero, L. Eisner, *et al.*, "2006," *The source for hydraulic fracture characterization: Oilfield Review*, vol. 17, pp. 42–57, 2005.

[118] H. Cinco-Ley and F. v. Samaniego, "Transient pressure analysis for fractured wells," *Journal of petroleum technology*, vol. 33, no. 9, pp. 1749–1766, 1981.

[119] H. Cinco-Ley and V. Samaniego, "Transient pressure analysis: finite conductivity fracture case versus damaged fracture case," Pet Mex, Tech. Rep., 1981.

[120] C. Langenbruch and S. Shapiro, "Geomechanical interpretation of pore pressures triggering microseismicity," in *73rd EAGE Conference & Exhibition*, 2011.

[121] G. De Marsily, A. Buoro, *et al.*, "40 years of inverse problems in hydrogeology," *Comptes Rendus de l'Academie des Sciences Series IIA Earth and Planetary Science*, vol. 329, no. 2, pp. 73–87, 1999.

[122] J. Carrera, A. Alcolea, A. Medina, J. Hidalgo, and L. J. Slooten, "Inverse problem in hydrogeology," *Hydrogeology journal*, vol. 13, no. 1, pp. 206–222, 2005.

[123] T.-C. J. Yeh, C. H. Lee, K.-C. Hsu, and Y.-C. Tan, "Fusion of active and passive hydrologic and geophysical tomographic surveys: The future of subsurface characterization," *Subsurface Hydrology: Data Integration for Properties and Processes*, pp. 109–120, 2007.

[124] W. W.-G. Yeh, "Review of parameter identification procedures in groundwater hydrology: The inverse problem," *Water Resources Research*, vol. 22, no. 2, pp. 95–108, 1986.

[125] J. Carrera, "State of the art of the inverse problem applied to the flow and solute transport equations," in *Groundwater Flow and Quality Modelling*. Springer, 1988, pp. 549–583.

[126] D. McLaughlin and L. R. Townley, "A reassessment of the groundwater inverse problem," *Water Resources Research*, vol. 32, no. 5, pp. 1131–1161, 1996.

[127] D. S. Oliver, A. C. Reynolds, and N. Liu, *Inverse theory for petroleum reservoir characterization and history matching*, 2008, vol. 1.

[128] A. Tarantola, *Inverse problem theory and methods for model parameter estimation*. siam, 2005.

[129] C. P. Robert and G. Casella, *Monte Carlo statistical methods*. Citeseer, 2004, vol. 319.

[130] A. Sahuquillo, J. Capilla, J. Gómez-Hernández, and J. Andreu, "Conditional simulation of transmissivity fields honoring piezometric data," *Hydraulic engineering software IV, Fluid flow modeling*, vol. 2, pp. 201–214, 1992.

[131] A. M. LaVenue, B. S. RamaRao, G. De Marsily, and M. G. Marietta, "Pilot point methodology for automated calibration of an ensemble of conditionally

simulated transmissivity fields: 2. application," *Water Resources Research*, vol. 31, no. 3, pp. 495–516, 1995.

[132] B. S. RamaRao, A. M. LaVenue, G. De Marsily, and M. G. Marietta, "Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 1. theory and computational experiments," *Water Resources Research*, vol. 31, no. 3, pp. 475–493, 1995.

[133] J. J. Gómez-Hernánez, A. Sahuquillo, and J. E. Capilla, "Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data–1. theory," *Journal of Hydrology(Amsterdam)*, vol. 203, no. 1, pp. 167–174, 1997.

[134] M. Sambridge and K. Mosegaard, "Monte carlo methods in geophysical inverse problems," *Reviews of Geophysics*, vol. 40, no. 3, pp. 3–1, 2002.

[135] R. J. Lorentzen, G. Nævdal, and A. C. Lage, "Tuning of parameters in a two-phase flow model using an ensemble kalman filter," *International Journal of Multiphase Flow*, vol. 29, no. 8, pp. 1283–1309, 2003.

[136] G. Nævdal, L. M. Johnsen, S. I. Aanonsen, E. H. Vefring, *et al.*, "Reservoir monitoring and continuous model updating using ensemble kalman filter," *SPE journal*, vol. 10, no. 01, pp. 66–74, 2005.

[137] X.-H. Wen, W. H. Chen, *et al.*, "Real-time reservoir model updating using ensemble kalman filter with confirming option," *Spe Journal*, vol. 11, no. 04, pp. 431–442, 2006.

[138] G. Evensen, "The ensemble kalman filter: Theoretical formulation and practical implementation," *Ocean dynamics*, vol. 53, no. 4, pp. 343–367, 2003.

[139] J. Kollat, P. Reed, and D. Rizzo, "Addressing model bias and uncertainty in three dimensional groundwater transport forecasts for a physical aquifer experiment," *Geophysical Research Letters*, vol. 35, no. 17, 2008.

[140] H. L. Mitchell, P. Houtekamer, and G. Pellerin, "Ensemble size, balance, and model-error representation in an ensemble kalman filter*," *Monthly weather review*, vol. 130, no. 11, pp. 2791–2808, 2002.

[141] G. Burgers, P. Jan van Leeuwen, and G. Evensen, "Analysis scheme in the ensemble kalman filter," *Monthly weather review*, vol. 126, no. 6, pp. 1719–1724, 1998.

[142] O. P. Lodoen, H. Omre, *et al.*, "Scale-corrected ensemble kalman filtering applied to production-history conditioning in reservoir evaluation," *SPE journal*, vol. 13, no. 02, pp. 177–194, 2008.

[143] J. He, P. Sarma, L. J. Durlofsky, *et al.*, "Use of reduced-order models for improved data assimilation within an enkf context," in *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2011.

[144] D. S. Oliver and Y. Chen, "Improved initial sampling for the ensemble kalman filter," *Computational Geosciences*, vol. 13, no. 1, pp. 13–27, 2009.

[145] Y. Chen and D. S. Oliver, "Cross-covariances and localization for enkf in multiphase flow data assimilation," *Computational Geosciences*, vol. 14, no. 4, pp. 579–601, 2010.

[146] Y. Zhang and D. S. Oliver, "Improving the ensemble estimate of the kalman gain by bootstrap sampling," *Mathematical Geosciences*, vol. 42, no. 3, pp. 327–345, 2010.

[147] P. Sarma, K. Aziz, L. J. Durlofsky, *et al.*, "Implementation of adjoint solution for optimal control of smart wells," in *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2005.

[148] P. Sarma, L. J. Durlofsky, K. Aziz, and W. H. Chen, "Efficient real-time reservoir management using adjoint-based optimal control and model updating," *Computational Geosciences*, vol. 10, no. 1, pp. 3–36, 2006.

[149] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14. California, USA, 1967, pp. 281–297.

[150] C. Ding and X. He, "K-means clustering via principal component analysis," in *Proceedings of the twenty-first international conference on Machine learning*. ACM, 2004, p. 29.

[151] C. K. Wikle and L. M. Berliner, "A bayesian tutorial for data assimilation," *Physica D: Nonlinear Phenomena*, vol. 230, no. 1, pp. 1–16, 2007.

[152] R. J. Lorentzen and G. Nævdal, "An iterative ensemble kalman filter," *Automatic Control, IEEE Transactions on*, vol. 56, no. 8, pp. 1990–1995, 2011.

[153] M. Zupanski, "Maximum likelihood ensemble filter: Theoretical aspects," *Monthly Weather Review*, vol. 133, no. 6, pp. 1710–1726, 2005.

[154] Y. Gu, D. S. Oliver, *et al.*, "An iterative ensemble kalman filter for multiphase fluid flow data assimilation," *Spe Journal*, vol. 12, no. 04, pp. 438–446, 2007.

[155] D. J. MacKay, *Information theory, inference, and learning algorithms*. Citeseer, 2003, vol. 7.

[156] G. Hamerly and C. Elkan, "Alternatives to the k-means algorithm that find better clusterings," in *Proceedings of the eleventh international conference on Information and knowledge management*. ACM, 2002, pp. 600–607.

[157] F. Floris, M. Bush, M. Cuypers, F. Roggero, and A. R. Syversveen, "Methods for quantifying the uncertainty of production forecasts: a comparative study," *Petroleum Geoscience*, vol. 7, no. S, pp. S87–S96, 2001.