Project Title: Interwell Connectivity and Diagnosis Using Correlation of
Production and Injection Rate Data in Hydrocarbon Production

ANNUAL REPORT

Reporting Period Start Date: June 2003

Reporting Period End Date: June 2004

Principal Authors: Jerry L. Jensen, Tx A&M University
Dr. Larry W. Lake, UT-Austin
Dr. Thang D. Bui, Texas A&M University
Ali Al-Yousef, UT-Austin
Pablo Gentil, UT-Austin

August, 2004

DOE Contract No. DE-FC26-03NT15397

Submitting Organization:
Texas Engineering Experiment Station
Texas A&M University, College Station, TX 77843-3116

Subcontractor: Larry W. Lake
Petroleum and Geosystems Engineering
University of Texas, Austin, TX 78712

DISCLAIMER

“This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.”

ABSTRACT

This report details much of the progress on inferring interwell communication from well rate fluctuations. The goal of the project was to investigate the feasibility of inferring reservoir properties through weights derived from correlations between injection and production rates.

We have focused on and accomplished the following items:

1. We have identified two possible causes for the source of negative weights. These are colinearity between injectors, and nonstationarity of be production data.
2. Colinearity has been addressed through ridge regression. Though there is much to be done here, such regression represents a trade-off between a minimum variance estimator and a biased estimator.
3. We have applied the ridge regression and the original Albertoni procedure to field data from the Magnus field.
4. The entire procedure (with several options) has been codified as a spreadsheet add-in.
5. Finally, we have begun, and report on, an extension of the method to predicting oil rates.

Successful completion of these items will constitute the bulk of the final year's report.

TABLE OF CONTENTS

DISCLAIMER	ii
ABSTRACT.....	iii
TABLE OF CONTENTS.....	iv
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF FIGURES	vii
INTRODUCTION	1
RESULTS AND DISCUSSION: TASK 2- RELAX ASSUMPTION OF HOMOGENEOUS RESERVOIR	6
MLR Diffusivity Filter.....	6
Comparisons between MLR Weights and Streamlines	6
Alternative Methods to Analyze Production Data	10
Cross Correlation.	10
Neural networks.....	10
Summary of Results on Task 2	12
RESULTS AND DISCUSSION: TASK 3- INTERPRET AND EXPLOIT OCCURRENCE OF NEGATIVE WEIGHTS	13
Simulation Models	13
Analysis of Simulation Results	14
Base case (Case 0)	14
Effect of Well Placement	15
Effect of Injection Rate.....	17
Effect of Injection Rate Variability.....	18
Effect of Reservoir Compressibility	18
Effect of Unbalanced Total Injection-Production Rate	19
Case with fluid support	20
Case with fluid loss	20
Effect of Unequal Bottom Hole Pressure of the Producers	21
Effect of Changing Bottom Hole Pressure.....	22
Effect of Autocorrelation of Injection Rates.....	22
Effect of Reservoir Heterogeneity	25
Detecting Collinearity	28
The Condition Number and Condition Index	28
Variance Inflation Factor and Multiple Determination Coefficient	29
Remedies for Collinearity	30
Ridge Regression	31
The Modified Ridge Regression Method.....	37

Summary of Results on Task 3	42
RESULTS AND DISCUSSION: TASK 4 – TEST AND EXTEND TOLERANCE FOR VARIABLE COMPRESSIBILITY CASES	44
Relative permeability and the b exponent.....	45
The 5x4 Homogeneous Synfield.....	45
5x4 Synfield - The b Exponent	47
Pseudo Allocation Models and the Oil Model.....	49
The 5x4 Faulted Homogeneous Synfield.....	50
RESULTS AND DISCUSSION: TASK 5 – TEST OF METHOD ON LARGE DATASET	55
Procedure of Data Analysis	58
The Producer: M30ZA5.....	59
The Producer: M36ZB7	65
Magnus Summary:	71
CONCLUSIONS.....	82
RECOMMENDATIONS	83
REFERENCES	84

LIST OF TABLES

TABLE 1 – PARAMETER VALUES FOR THE BASE CASE SIMULATION MODEL.....	14
TABLE 2: CENTERED-MLR ANALYSIS.....	33
TABLE 3 MLR ANALYSIS	61
TABLE 4: THE MLR ANALYSIS USING THE RANK DATA.....	64
TABLE 5: MLR ANALYSIS USING THE ACTUAL DATA AND THE RANK OF THE DATA	68
TABLE 6: COMPARISON OF MLR WEIGHTS BEFORE AND AFTER THE MODIFIED RIDGE METHOD	70

LIST OF FIGURES

Figure 1 - Streamlines at time step 25, 30, 35, 40 (months) – Case0 homogeneous.....	8
Figure 2 – Streamline ratio versus MLR weights for homogeneous case, 9-spot pattern.	9
Figure 3 – Streamline at timesteps 30, 35, 40 for the Bloque I simulation case.....	9
Figure 4 - Streamline ratio and MLR weight for Bloque I simulation case	10
Figure 5 – Comparison of the NN influence and MLR weights – case heterogeneous reservoir. 11	
Figure 6– NN influence and MLR weights for the case of changing the BHP of P1.	12
Figure 7 – Reservoir configuration of the standard 9-spot pattern. Distances are in feet. This is very similar to Albertoni. What is different?	13
Figure 8 – MLR weights map for the base case: symmetric.....	15
Figure 9 – Asymmetric 9-spot pattern to investigate the effect of well placement.	16
Figure 10 – MLR weights map for the case of asymmetric well placement.	16
Figure 11 – MLR weights as a function of interwell distance. What is the interwell distance relative to in the last figure? Would it make sense to divide the weights by the distances? 17	
Figure 12 – MLR weights of the unequal injection rates as compared to the base case.	17
Figure 13 – MLR weights of the unequal rate variability, as compared to the base case (0).....	18
Figure 14 – MLR weights of the high compressibility case versus base case.....	19
Figure 15 – MLR weights of the high compressibility and large injection rate variability case. I like this result but Pizarro did this also. Can you think of a dimensionless group that would indicate when compressibility effects are negligible? Like $c \text{ time delta } P$	19
Figure 16 – MLR weights for the case of extra fluid support.....	20
Figure 17 – MLR weights of the case of losing injection fluid.	21
Figure 18 – MLR weights for the case of producers having unequal BHP's. Interesting result. Same trend holds.....	21
Figure 19 – MLR weights for the case of changing producer BHP.....	22
Figure 20 – MLR weight for the case of injection rate autocorrelation -1	23
Figure 21 – MLR weights for the case of injection rate autocorrelation - 2.....	24
Figure 22 – MLR weights for the case of injection rate autocorrelation - 3.....	24
Figure 23 – MLR weights map for the case of heterogeneity.....	25
Figure 24 – MLR weights for heterogeneous case compared to the base case.....	26
Figure 25 – MLR weights for the case where h_i perm between one well pair does not include producer grid block.	26
Figure 26: The ridge trace plot.....	34
Figure 27: MDC Vs. The ridge parameter	34
Figure 28: Weight Standard deviation Vs. The ridge parameter	35
Figure 29: R of the system Vs. The ridge parameter	36
Figure 30: MLR weights compared to simple weights.....	36
Figure 31: The ridge weights compared to simple weights	37
Figure 32: The modified ridge trace plot	39
Figure 33: MDC Vs. The modified ridge parameter.....	39
Figure 34The weight standard deviation Vs. The modified ridge parameter	40
Figure 35: R of the system Vs. The modified ridge parameter.....	41
Figure 36: The modified ridge weights compared to simple weights.....	41
Figure 37: WOR vs. Cumulative Liquid Fluids produced by well P01. A constant power law relationship can be observed.....	46

Figure 38: Oil rate vs time for producer P01. Regression results compare favorably with data. Note that first points were omitted from the regression.	46
Figure 39: Oil rate regression results vs Oil Rate data for producer P01. Regression results vs. data. Lie on a 45 degree line.	47
Figure 40: WOR regression results vs WOR data for producer P01. Regression results vs. data. lie on a 45 degree line.	47
Figure 41: Relative permeability data input for the 5x4 homogeneous Synfield case.	48
Figure 42: Relative permeability ratio and corresponding power law regression for the 5x4 homogeneous Synfield case. Exponent value is 1.74.	48
Figure 43: Results of applying the power law oil model to the <i>Model 4b-non centered</i> allocation. Individual oil contribution from each of the three neighboring injectors to producer P01 are inferred. Left figure: linear graph; right figure shows semilog graph.	49
Figure 44 Graph showing inferred WOR evolution for the volume of fluids supported by each of the three injectors supporting producer P01.	50
Figure 45 : Results of applying the power law oil model to the <i>Model 4b-non centered</i> allocation. Oil rate regression vs. oil rate data lie on 45 degree line. The oil regression shows very good agreement except for the initial part of the history (high oil rates) where WORs are very low.....	50
Figure 46: Graph on the left shows a more complex WOR behavior with an early time trends transitioning into a late time trend. On the right, graphic representation of the results of Model4b on P01 suggesting a lack of connection between the left and right blocks.	51
Figure 47: Regression using late time information. On the left WOR and WOR regressed against cum fluids produced. On the right, plot of Qo and Qo regressed vs. time. Late time (aprox>2000d) agreement is good.	52
Figure 48: Regression using early time data only. A quality of the match is lost after aprox. 1000d.....	53
Figure 49: Results for WOR I1, WOR I2 and WOR I3, their combined result into WOR Pred and comparison WOR data. By combining three simple power law relationship a match to a more complex behavior can be attempted.	54
Figure 50: Combined results of Model4b-noncentered model and power law oil models. The combined WOR behavior of P01's three injectors (see Fig PG49) provide the capability of integrating early and late time behaviors resulting in an overall improvement on the quality of the regression.	54
Figure 51 Magnus field structure map.....	56
Figure 52 Injection and poduction history.	57
Figure 53The normal probability plot for production data.....	58
Figure 54 The normal probability plot for injection data.	58
Figure 55The crestal area.....	59
Figure 56Time series of M30ZA5 and nearby injectors.....	60
Figure 57The simple correlation coefficient analysis.....	61
Figure 58 MLR weights after applying the successive elimination method.....	62
Figure 59The modified ridge weights.....	62
Figure 60Time series of injection-production rates rank.....	63
Figure 61: The spearman rank correlation coefficient.....	64
Figure 62: The simple correlation coefficient using the rank of the data.	64
Figure 63: MLR weight after applying the successive elimination method.....	65

Figure 64 The modified ridge weights.....	65
Figure 65: Time series rates using actual data.....	66
Figure 66: Time series rates using the rank of the data	67
Figure 67: Correlation coefficient analysis using ranks	67
Figure 68: The modified ridge regression for the constrained MLR.....	69
Figure 69: The original estimated weights of the rank of the data	71
Figure 70: The modified weights of the rank of the data.....	71
Figure 71: The image representation of estimated weights using actual data	72
Figure 72: The image representation of stability of estimated weights	73
Figure 73: The image representation of weights using the rank of the data	73
Figure 74: The image representation of stability of weights using the rank of the data.....	74
Figure 75 A graphical representation of actual data weights.....	75
Figure 76: A graphical representation of rank data weights	76
Figure 77: The compass diagram for the actual weights	77
Figure 78: The compass diagram for the rank weights.....	78
Figure 79 The rose histogram for the direction of actual data weights (# bins= 5).....	79
Figure 80: The rose diagram for the direction of actual data weights (#bins< 20).....	79
Figure 81: The rose histogram for the direction of rank data weights (#bins = 5)	80
Figure 82: The rose histogram for the direction of rank data weights (#bins < 20)	80

INTRODUCTION

The multiple linear regression (MLR) method consists of correlating the production rate q_j of a producer j to the injection rates of the injectors $i_1, i_2, \dots, i_I$ according to the multivariate linear equation (Albertoni and Lake, 2003):

$$q_j = \beta_0 + \sum_{i=1}^I \beta_{ij} i_i, \dots \text{(Eq. 1)}$$

The coefficients β_{ij} in the Eq. 1 are the MLR weights. For each producer, the MLR weights are calculated by minimizing the sum of squared error between the calculated and the observed production rates for a time period T .

$$SSE = \sum_{t=1}^T (q_t - \hat{q}_t)^2, \dots \text{(Eq. 2)}$$

The goodness of the production-injection rate correlation is measured by R^2 :

$$R^2 = 1 - \frac{SSE}{SSY}, \dots \text{(Eq. 3)}$$

where $SSY = \sum_{t=1}^T (q_t - \bar{q})^2$.

The MLR method requires multiple steady states to determine the weights, β . Under some strict conditions (constant producer bottom hole pressures, steady states), the physical meaning of these weights can be interpreted as the change of the production rate at a particular well per unit change of the injection rate at an injector, given that the rates of other injectors remain constant. Under these conditions, the weight should not be negative or exceed one (i.e., $0 \leq \beta \leq 1$). Minimizing the SSE in Eq. 4 results in the following equations in matrix notation.

$$\bar{Y} = \bar{X} \bar{\beta} + \bar{\varepsilon} \quad \text{(Eq. 4)}$$

where :

$${}_{(n \times 1)} \bar{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$$

$${}_{(n \times p)} \bar{X} = \begin{bmatrix} 1 & x_1^{(1)} & x_1^{(2)} & \cdot & \cdot & \cdot & x_1^{(p-1)} \\ 1 & x_2^{(1)} & x_2^{(2)} & \cdot & \cdot & \cdot & x_2^{(p-1)} \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_n^{(1)} & x_n^{(2)} & \cdot & \cdot & \cdot & x_n^{(p-1)} \end{bmatrix}$$

$${}_{(p \times 1)} \bar{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \beta_{p-1} \end{bmatrix}$$

$${}_{(n \times 1)} \bar{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix}$$

As shown above, there are four matrices needed to express the MLR model. The vector Y represents the observations of the dependent variable Y_i where the dependent variable in our system is the production rate of a particular producer. Each column in the X matrix represents the rates for a particular independent variable, which in this case is an injection rate. The first column is to account for the intercept in the MLR model. The vector β is a vector of unknown coefficients or weights to be estimated from solving the above linear system of equations. The vector ε is the vector of random errors where each value reflects the deviation between the observed Y value and the estimated value through the linear model. The random error should be characterized by a normal distribution with zero mean and constant variance to insure the normality of the dependent variable Y .

Using the minimum error variance as an optimization criterion generates a system of normal equations the solution of which gives a solution for the desired coefficients or weights. The inferences about regression weights are straightforward. The mathematical representation of the solution is as follow:

The normal equation:

$$X'Xb = X'Y \quad (\text{Eq. 5})$$

The MLR estimator:

$$\bar{b} = (X'X)^{-1} (X'Y) \quad (\text{Eq. 6})$$

The minimum sum of squares of the residuals:

$$SSE = (Y - bX)' (Y - bX) \quad (\text{Eq. 7})$$

where:

$(p \times p) (X'X) \equiv \text{Variance-Covariance Matrix of injectors,}$

$(p \times 1) (X'Y) \equiv \text{Covariance of injectors-producer}$

The variance-covariance matrix of unbiased Coefficient:

$$\sigma^2(\bar{b}) = \sigma^2 (X'X)^{-1} \quad (\text{Eq. 8})$$

The variance-covariance matrix of estimated Coefficient:

$$s^2(\bar{b}) = MSE (X'X)^{-1} \quad (\text{Eq. 9})$$

where:

MSE is the mean squares error

With some manipulations, the matrix $X'X$ is symmetric where the diagonal elements are the variances of injectors and the off-diagonal elements are covariances of injector-injector pairs. The elements of the vector $X'Y$ are covariances of injector-producer pairs. If the first column in X is omitted and all treated variables in the normal equation are centered using their means, the

matrix $X'X$ will be the variance-covariance matrix of injectors with no need for matrix manipulations. Occasionally we will use centered variables scaled to their standard deviation if there are significant variations in magnitude among the input variables. In this case the resultant $X'X$ matrix will be simply the injector-injector simple correlation matrix where all the diagonal elements are 1 and off-diagonal elements are simple correlation coefficients of injector-injector pairs. In both cases, the estimated weights are different but it is straightforward to convert from one to another using the following relation:

$$b_i = \frac{S_y}{S_i} b'_i \quad (\text{Equation 10})$$

where:

b_i : estimated weight using centered moment

b'_i : estimated weight using centered-scaled moment

S_y : the standard deviation of dependent variable, producer

S_i : the standard deviation of the input variable, injector

Generally, $X'X$ has an inverse and then the normal equations have a solution as given in Eq. (6). When the treated input variables are totally independent, the solution tends to have the least sensitivity to the input data and satisfies the system. The MLR estimation procedure is a good tool if $X'X$, when in the form of injector-injector simple correlation matrix, is nearly a unit matrix (Albertoni, 2002).

In the independent case, the estimated weights, called simple weights, using centered-scaled moment are simply equal to the corresponding simple correlation coefficient of injector-producer pairs. Simple weights are characterized by minimum variability as can be estimated by Eq. (9). This solution is really consistent with simple correlation coefficient of each injector with producer where both they have the same sign and magnitude.

Intuitively, we expect that the larger the MLR weight, the greater the degree of influence that injector i_i has on producer q_j . Thus, the MLR weights can be used to represent the “interwell connectivity” between injector and producer, similar to the cross correlation coefficient between producer and injector (Albertoni, 2002). Quantitatively, the model (Eq. 1) indicates that the flow rate at a specific producer can be directly attributed to a proportion, given by the MLR weight, of an injection well’s injection rate.

Other applications of the MLR weights are to infer the permeability between an injector and producer well pair and identify variations of permeability throughout the swept reservoir. Use of the values of the MLR weights to infer the reservoir permeability and heterogeneity requires some caution, however. The analyses of Muskat (1982, Chap. 9) and others, clearly shows that factors other than permeability, such as the interwell distances and well locations, affect the weights.

Nonetheless, analysis of simulated data gives some promising results (e.g., Albertoni and Lake, 2003). The MLR weights seem insensitive to the injection rates and their variability. For the

case of constant working condition of the producer and strong permeability anisotropy, the magnitude of the MLR weights may be able to indicate this permeability anisotropy or a permeability barrier in the reservoir. For the case of more complex heterogeneity, however, the MLR weights give less satisfactory results in terms of representing the permeability heterogeneity in the reservoir. Upon applying the MLR method to actual field data, a significant number of weights can be negative and/or abnormally large. A proposal to eliminate the negative weights by simply removing those injectors from the MLR model (Eq. 1) have been made. By doing so, the connectivity between injector and producer is assumed to be zero, and the model accommodates this change by adjusting the remained positive weights (Albertoni, 2002). It is expected, however, that removing the injectors with negative weights without fully investigating their significance may lead to incorrect conclusions about the nature of interwell connectivity.

EXPERIMENTAL

No experimental procedures were involved in this project.

RESULTS AND DISCUSSION: TASK 2- RELAX ASSUMPTION OF HOMOGENEOUS RESERVOIR

We investigate modifications and alternatives to the MLR approach advocated by Albertoni and Lake (2003). The diffusivity filter performance is discussed. We recommend that, when needed, the diffusivity filter may require a longer time window than that currently applied. We compare the MLR weights to those obtained using streamline models and find that the correspondence is poor in some cases. The MLR weights might be better interpreted as “influence factors” rather than weights, since significantly non-zero weights can be obtained for cases where there is no flow between injection and production wells. Finally, this study also finds some limited advantage to using a nonlinear correlation method (neural networks) instead of MLR to analyze production and injection data.

MLR Diffusivity Filter

The diffusivity filter, as proposed by Albertoni and Lake¹, is used to account for the transient effects of changes of the injection rate. From pressure transient theory, changing the injection rate at the producer will create a pressure disturbance at the injector location. This pulse travels in the reservoir with the speed determined by the reservoir diffusivity coefficient. Once this disturbance arrives at a producer, the production rate responds accordingly as the system returns to steady-state condition (Muskat, 1982, p. 621-630).

The time for the pressure pulse traveling a distance r can be estimated using a line-source solution in an infinite radial homogeneous reservoir. For the reservoir properties used in this study, the travel time for the distance between any injector-producer pair is on the order of several days. Therefore, with monthly sampling rate, there is no significant delay in the effect of injector on producer. As a result, applying the MLR diffusivity filter only slightly improves the R^2 of the correlation. With daily sampling, the situation is different. Because the delay could be up to 4-5 days in this case, applying the diffusivity filter might significantly improve the correlation coefficient.

Since the diffusivity filter coefficient within its time window sums to unity, a longer time window used in calculating the diffusivity filter may be required. However, a wider time windows reduces the role of each time step in the correlation and decreases the sensitivity coefficient of the diffusivity filter. Similar results can be achieved in cases of daily data by shifting the production data with time until the correlation between injection and production rates yields the highest R^2 . The magnitude of the time lag between the production data and injection data in this case can be used to help infer the interwell connectivity.

Comparisons between MLR Weights and Streamlines

Injector-producer connectivity may be thought of as relating to streamlines connecting injectors with producers. The more streamlines from a particular injector land on a given producer, the greater should be the connectivity for that pair. If a producer has N lines in total and n of those lines are from one injector, we would expect the connectivity of that pair of wells to be given as the ratio n/N . Do the MLR weights give the same result as n/N ?

We used Frontsim, a streamline simulation to generate the streamline distributions. An example is shown in Figure 1 for the case of homogeneous standard 9-spot pattern. There is no streamline connecting the long distance injector-producer pairs (e.g., I1-P3, I1-P4, and I2-P2 in Figure 1). We calculated the average streamline ratio for each producer-injector pair as the ratio of the number of streamlines from injector i to the total number of streamlines landing at producer j .

For the case of equal average injection rates (Figure 1), there are two groups of streamline ratios: one for the center injector, and one for the corner injectors. By changing the average injection rate, however, the streamline ratio changes accordingly, whereas the MLR weights are relatively injection rate independent. The plot of the average streamline ratio versus the MLR weights for the base case of homogeneous standard 9-spot pattern is shown in Figure 2. For the short distance well pairs, the MLR weights seem to be proportional to the average streamline ratio for this homogeneous case. For the long distance pairs, however, the two methods disagree.

The lack of long-distance streamlines between, for example, I1 and P3 is because the nearest injectors I2, I3, and I5 raise the pressure P in the region around P3. P is sufficiently great that regions with $\nabla P = 0$ are created and stop any flow from I1. This behavior has been studied by Muskat (1982) who terms the effect “mutual interference” (p. 529) and “shielding” (p. 530-556). So why does the MLR method give non-zero β 's? The answer can again be found in Muskat (p. 529). Injection at distant wells may still influence the pressure field and thus affect the streamline paths of the nearby injectors. An extreme example would be if I1 were to become a producer. Some of the I3 streamlines would be pulled away from P3 and land on I1. Varying production at I1 would thus influence production at P3 without any streamlines connecting I1 and P3. So a non-zero β may reflect one of two fundamentally different mechanisms: a) an injector-producer pair shares one or more streamlines, making the injector directly responsible for a portion of the production; and b) an injector is “modulating” the pressure field and influencing the contributions other injectors make to production at a particular well.

These results suggest:

1. The MLR weights should be considered as “influence factors” rather than connectivity measures;
2. The MLR weights may not reflect the amount of injected fluid which appears at a producers.
3. Interpretation of β_{ij} as reflecting the interwell permeability level may prove difficult as the influence of injector j on producer i may not involve flow through the region connecting the pair.

Streamlines represent the images of the possible fluid travel paths while MLR weights represent how the system is reacting to changes of injection and production rates.

Our study on other reservoir configurations and rate schemes supports the above analysis. For all our simulation cases, the average streamline ratio varies between 0 and 1. The streamline connecting long distance well pair is rare and only occurs at the peripheries of the reservoir.

We ran one case with a more general distribution of the well placement and injection rates. The reservoir is homogeneous and the injection rate for this simulation case is taken from real field data. Figure 3 shows the example of the streamline distribution and Figure 4 shows the average

streamline ratio versus the MLR weights. In general, the average streamline ratio and MLR weights show no correlation. For this homogeneous case, the majority of the well pairs show the average streamline ratio of 0 or 1. The MLR weights range between -0.16 and 0.62. We believe that the negative MLR weights are resulting from the injection rate autocorrelation.

Figure 1 - Streamlines at time step 25, 30, 35, 40 (months) – Case0 homogeneous.

Figure 2 – Streamline ratio versus MLR weights for homogeneous case, 9-spot pattern.

Figure 3 – Streamline at timesteps 30, 35, 40 for the Bloque I simulation case

Figure 4 - Streamline ratio and MLR weight for Bloque I simulation case

Alternative Methods to Analyze Production Data

Cross Correlation.

The Spearman rank cross correlation of injector-producer has been used to identify the connection or the lack of connection between wells (Soeriawinata and Kelkar, 1999). Applying this method on simulation data suggests that it is not able to represent the relative magnitude of the injector-producer influence. Within a homogeneous reservoir compartment, the cross correlation may show a strong long range correlation and weak short range correlation. We expect that this method is also strongly affected by the injection rate autocorrelation. This autocorrelation may cause a far field injector-producer influence that may be wrongly interpreted as a channel between wells or geomechanical changes in the field.

Neural networks.

Neural networks (NN's) have been proposed to correlate production and injection rates (Panda and Chopra, 1998). We reexamine the use of NN's on simulated data to correlate the production and injection rates.

For each producer, a network consisting of 6 input nodes and 4 hidden nodes was built. The network was trained on the given injection-production rates. After the network was trained, we used it to estimate the injector-producer influence by changing the injection rate at one input node while keeping others at their average value. The ratio of the change at the output node to the change at the input node is termed the NN influence. This value was compared with the corresponding MLR weight.

Figure 5 shows the plot of the NN- derived influence versus the MLR weights for the base case. Figure 6 shows the similar plot for the case of changing BHP of the producers. The NN influence and the MLR weights show insignificant differences.

The most important problems of using NN's are the possibility of overtraining and the nonlinearity in extracting the input-output causal relationship. Our analysis shows that, for the case of changing BHP, the NN influence strongly depends on the magnitude of the change at the input node and the level of injection rate at other injectors. Thus, NN's may lead to a better production-injection correlation but may be too rate-dependent. Because we are seeking a coefficient that should be rate independent to represent the interwell connectivity, NN's may not suit for the purpose of interpreting the reservoir characteristics.

Figure 5 – Comparison of the NN influence and MLR weights – case heterogeneous reservoir.

Figure 6– NN influence and MLR weights for the case of changing the BHP of P1.

Summary of Results on Task 2

1. The role of the diffusivity filter in MLR analysis should be viewed differently for different production data reporting schemes. For data that were collected on a monthly basis, the diffusivity filter only slightly improves the R^2 of the correlation. For data collected on a daily basis, the role of the diffusivity filter increases significantly.
2. MLR weights and streamline-determined connectivities may disagree. This is because the streamlines show steady-state hydraulic connections between wells, while the MLR weights show how injection perturbations affect production.
3. The MLR weights are better interpreted as influence factors rather than measures of (hydraulic) connectivity because non-zero weights may obtain in cases where there is no direct hydraulic connection between wells.
4. Interpreting MLR weights as measures of interwell permeability may be problematic because non-zero weights may arise for cases where there is no flow between an injector and producer.

The nonlinear correlation such as neural network can improve the R-squared of the correlation, compared to the MLR method. However, the result may be more rate dependent and more difficult to interpret.

RESULTS AND DISCUSSION: TASK 3- INTERPRET AND EXPLOIT OCCURRENCE OF NEGATIVE WEIGHTS

We present an investigation of using MLR for interpreting the interwell connectivity. We built and ran different numerical simulation models to investigate the actual meaning of the MLR weights in interpreting the reservoir properties. In particular, the reasons for the negative and abnormally large positive MLR weights sometimes obtained during application of the method to field data are investigated. The results of the study suggest that the two most important factors in cases having negative and/or abnormally large weights are a) the correlation of rates among several injection wells and b) changes in the working bottom hole pressure of the producer. These results suggest that the MLR method may benefit from a different, more robust, way to calculate interwell correlations.

Simulation Models

We used Eclipse 100 for different reservoir and well situations to generate the production data used in the MLR analysis. To aid comparisons, a base case with a standard 9-spot pattern, consisting of 5 injectors and 4 producers, was created (Figure 7). The main features for the base case are: homogeneous reservoir, small compressibility, random injection rates of equal magnitude and variability, constant and equal bottom hole pressure for producers, and injection and production rates are reported on the monthly basis. The reservoir parameters for the base case are given in Table 1.

Figure 7 – Reservoir configuration of the standard 9-spot pattern. Distances are in feet.

TABLE 1 – PARAMETER VALUES FOR THE BASE CASE SIMULATION MODEL.

Parameter	Value
Porosity, fraction	0.18
Horizontal permeability, mD	40
Vertical permeability, mD	4 2D or 3D?
Oil compressibility, 1/psi	5×10^{-6}
Water compressibility, 1/psi	1×10^{-6}
Rock compressibility, 1/psi	1×10^{-6}
Model dimension	31 x 31 x 3
Grid size, ft	80 x 80 x 20

From the base case (case 0), modified cases are run to investigate the following effects on the MLR weights:

1. interwell distance and well placement
2. injection magnitude
3. injection rate variability
4. reservoir compressibility
5. unbalanced total injection-production rate
6. unequal bottom hole pressure of producer
7. changing bottom hole pressure of producer
8. autocorrelation of injection rates
9. reservoir heterogeneity

Analysis of Simulation Results

The results of using MLR to analyze the simulated data are presented. With the exception of the last case, the “ideal” outcome is for all the weights to be equal since the models are homogeneous. We find, however, that many other factors besides heterogeneity affect the weights.

Base case (Case 0)

All 5 injectors have injection rates randomly generated within the range from 500 to 600 bbl/d for 60 months. The rates are independent of each other and there is no autocorrelation within each rate. This resulted in the average injection rate of ~550 bbl/d, with an average standard deviation of ~27 bbl/d. Several sets of injection rates were used. Figure 8 shows the resulting weights for one of the rate schemes. In general, there are three groups of weights for different interwell distance and well placement: 1) short distance for corner injectors ($\beta \sim 0.34-0.37$); 2) short distance for center injectors ($\beta \sim 0.25$); and 3) long distance for corner injectors ($\beta \sim 0.14-0.18$). This suggests that, since there are only two injector-producer distances for this pattern, both the interwell distance and the well placement affect the MLR weights.

Figure 8 – MLR weights map for the base case: symmetric

Effect of Well Placement

The reservoir used for investigating the effect of the well placement is the modified 9-spot pattern with asymmetrical well placement (Figure 9). With the same injection rates as in the symmetric well pattern, the MLR weights for the asymmetric well pattern (Figure 10) show more variability. A plot of the MLR weights for both symmetric and asymmetric well patterns versus the interwell distance (Figure 11) suggests that both interwell distance and well placement in the reservoir have an effect on the magnitude of the MLR weights. In general, the MLR weight decreases with the distance. At the same interwell distance, the MLR weight for a corner injector is about 50% larger than the weight for the center injector.

Figure 9 – Asymmetric 9-spot pattern to investigate the effect of well placement.

Figure 10 – MLR weights map for the case of asymmetric well placement.

Figure 11 – MLR weights as a function of interwell distance.

Effect of Injection Rate

Simulation cases with a higher average injection rate for one of the injectors were run. The reservoir is the symmetric 9-spot pattern (Figure 7). In injector I1 the monthly rate is randomly generated within the range from 700 to 800 bbl/d, resulting in the average value of 752 bbl/d and standard deviation of 28 bbl/d. The injection rates of all other injectors and the reservoir properties are the same as for the base case. The plot of the MLR weights versus the base case (Figure 12) suggests that the MLR weights are insensitive to the relative magnitude of the injection rate.

Figure 12 – MLR weights of the unequal injection rates as compared to the base case.

Effect of Injection Rate Variability

Simulation cases with a higher standard deviation of rate for one injector were run. For example, the injection rate of well I1 was randomly generated within the range from 400 to 700 bbl/d, resulting in the average of 550 bbl/d and standard deviation of 98 bbl/d. The injection rates of all other injectors and the reservoir properties were the same as for the base case. The resulting MLR weights are shown in Figure 13, comparing the weights to those from the base case. The results suggest that the relative variability of the injection rate may have a small effect on the MLR weights. In all cases, the injection rate variability, however does not appear to significantly affect the MLR weights.

Figure 13 – MLR weights of the unequal rate variability, as compared to the base case (0).

Effect of Reservoir Compressibility

Two cases with a higher compressibility (5 times higher than in the base case) were run: the first case (case 1_0) is run with the injection rates taken from the base case and the second case (case 1_3) is run with the injection rates taken from a case with high injection rate variability. Figure 14 shows the comparison of the MLR weights for case 1_0 and the base case. Figure 15 shows the comparison of the MLR weights for case 1_3 and the base case. The results show that the effect of increased reservoir compressibility is not significant when the injection rates and their variabilities are at the same level. As the injection rate variability increases however, the clear separation between the three groups of weights no longer exists. Nonetheless, significant deviations of the MLR weights from the base case are not observed.

Figure 14 – MLR weights of the high compressibility case versus base case.

Figure 15 – MLR weights of the high compressibility and large injection rate variability case.

Effect of Unbalanced Total Injection-Production Rate

Several cases were run to investigate the effect on the MLR weights when a well is present but neglected or when the region of the reservoir used in the analysis is receiving or losing fluid to other regions. We add either one injector or producer in our 9-spot pattern in the simulation but will not include this well in the MLR analysis. In one case, the injection rates for well I6 (at grid block 31, 10), not clear where this well is the phantom injector, are randomly generated with the average of 300 bbl/d and standard deviation of 30 bbl/d. In the other case, the producer P5 (at

grid block 31, 10) works at the same constant bottom hole pressure as other producers. The typical results are as follows.\

Case with fluid support

The resulted MLR weights are substantially more scattered compared to the base case (Figure 16). The biggest change occurred for the long-range correlations, far from the point of getting fluid: I4-P3, I4-P1, and I1-P3. Injector I6 is close to wells P3 and P1. By not including I6 in the model, the correlation weights for the producer most affected by outer source of support increase, especially for long range injectors, to compensate the amount of fluid injected by the well I6.

Figure 16 – MLR weights for the case of extra fluid support

Case with fluid loss

In contrast to the case of getting extra support, the MLR weights in this case decrease, especially for the weights from the short range injectors to the producers most affected by the fluid drainage, i.e., close to the point of drainage (Figure 17). In Figure 17, the weight for well pair I2-P3, I2-P1 decreases at a greater extent, compared to other pairs.

These results reinforce the analysis of Albertoni (2002), who observed that exclusion of wells or poor identification of representative units can have a large effect on the calculated weights.

Figure 17 – MLR weights of the case of losing injection fluid.

Effect of Unequal Bottom Hole Pressure of the Producers

In these cases, producers work at constant but unequal bottom hole pressure (BHP) throughout the whole simulation period. BHP's of 250, 350, 400, and 500 psi for wells P1, P2, P3, and P4 respectively, were used. The unequal BHP's led to different average production rates for the producers: ~800, 720, 670, 590 bbl/d for P1, P2, P3, P4, respectively, with the injection rate of the base case. The resulting MLR weights are more scattered, compared to the base case (Figure 18). The most obvious change is for the weights of the central injector. The weight between this injector and the producer with low BHP (i.e. largest production rate) is higher than between this injector and the producer with high BHP (i.e. smallest production rate). Similar patterns are observed for other injector-producer weights. The general observation is that the producer with higher average rate will have larger MLR weights from all injectors, compared to the base case.

Figure 18 – MLR weights for the case of producers having unequal BHP's.

Effect of Changing Bottom Hole Pressure

Changing the BHP during the analysis period for one producer leads to the appearance of negative weights in the model. In this simulation case, the bottom hole pressure of the producer P1 decreases from 500 psi to 300 psi at 912 days of simulation. At 1247 days, its BHP increases back to 500 psi for the duration of the simulated period (60 months total). Figure 19 shows the MLR weights versus the base case. There are three negative weights in the model and several weights that are large. The smallest MLR weight is -0.81 and the largest weight is 0.93. Overall, the MLR weights become unpredictable and do not associate with the distance and geometric well placement. R^2 values of the MLR correlation in this case are very low, especially for the well with changing bottom hole pressure (R^2 is ~0.4, 0.5, 0.5, 0.6 for P1, P2, P3, and P4 respectively).

Figure 19 – MLR weights for the case of changing producer BHP.

Effect of Autocorrelation of Injection Rates

Several cases were run to investigate the effect of autocorrelation of the injection rate on the MLR weights. Two simulation cases were designed to capture different scenarios of the autocorrelation: autocorrelation between two injectors, autocorrelation between one injector and a group consisting of several injectors. We present the results of three cases where the injection rate of injector I2 correlates with the injection rates of other injectors according to the equation:

- $I2 = 0.1 * (I1 + I3 + I4 + I5),$
- $I2 = 0.1 * (I1 + I3 + I4 + I5) + 220,$ and
- $I2 = 2700 - (I1 + I3 + I4 + I5)$

The resulting MLR weights were plotted versus the MLR weights of the base case (Figures. 20-22). The results suggest that the negative degree of the correlation does not significantly affect

the MLR weights (Figure 22) whereas, the positive correlation of one injector with other injectors can significantly affect the MLR weights. In all cases of injection rate inter-correlation, only one case shows the presence of slightly negative MLR weights (Figure 20). This case is characterized by a small injection of I2 that correlates with the rate of all other injectors. This case is also characterized by several abnormally large positive MLR weights. The general observation is that the effect of the autocorrelation on the MLR weights is controlled by the degree of the autocorrelation and the relative magnitude of the injector. The smaller the injector rate of the well in the correlation, the wilder the MLR weights can be.

Albertoni's study (2002) suggests that the size of the negative weights can be reduced by increasing the number of data points in the model. It is expected that the longer time period used in the analysis, the smaller chance there will be of having autocorrelation among injection rates.

Alternatively, a more robust method to estimate the covariances of the rates may solve this problem (Chilés and Delfiner, p. 64-70).

Figure 20 – MLR weight for the case of injection rate autocorrelation -1 .

Figure 21 – MLR weights for the case of injection rate autocorrelation - 2.

Figure 22 – MLR weights for the case of injection rate autocorrelation - 3.

Effect of Reservoir Heterogeneity

A streak of one simulation grid in width with high permeability (400 mD) is introduced between injector I4 and producer P4. One case (Case H0) with the injection rates taken from base case is run. The MLR weights are shown in Figure 23. The special feature of this case is that the weights for producer P4 are higher, compared to the base case, whereas the MLR weights for other producers are smaller than the base case values (Figure 24). It is worth noting that because P4 is completed in the grid with higher permeability, its average production rate is significantly higher than the other producers (1280 bbl/d compared to 540, 450, and 510 bbl/d for P1, P2, and P3 respectively). As a result, all the MLR weights for this well are significantly greater than the MLR weights for other producers. The MLR weights on Figure 24 form two groups: one is for P4 and the other is for P2, P2, and P3. This case and the case of unequal BHP suggest that the MLR weights may strongly depend on the producer rate. The MLR weights for the well with high rate are larger compared to the wells with lower rates.

Figure 23 – MLR weights map for the case of heterogeneity.

Figure 24 – MLR weights for heterogeneous case compared to the base case.

If a wider band of high permeability (of 4 grid cells in width) is introduced into the model, but does not include the cell where well P4 is completed, then the average production rates of the producers are not significantly different (670, 630, 680, and 800 bbl/d respectively for P1, P2, P3, P4) than the base case. The plot of the resulted MLR weights versus base case shows a characteristic high weight for the pair I4-P4 but a slightly low weight for the pair I4-P2 (Figure 25).

Figure 25 – MLR weights for the case where hi perm between one well pair does not include producer grid block.

Collinearity - As the above section indicated, one of the reasons behind the negative weight is the lack of independence of the injector stimuli. This phenomenon is called collinearity. The next few sections investigate ways to suppress collinearity.

A simple system consisting of a producer and two injectors can be used to illustrate collinearity. The mathematical representation of the estimated weights using centered-scaled moments is as follows:

The MLR system of normal equations:

$$\begin{bmatrix} r_{11} & r_{12} \\ r_{12} & r_{22} \end{bmatrix} \bullet \begin{bmatrix} b'_1 \\ b'_2 \end{bmatrix} = \begin{bmatrix} r_{Y1} \\ r_{Y2} \end{bmatrix} \quad (11)$$

The MLR estimator:

$$\begin{bmatrix} b'_1 \\ b'_2 \end{bmatrix} = \frac{1}{r_{11}r_{22} - r_{12}^2} \begin{bmatrix} r_{22} & -r_{12} \\ -r_{12} & r_{11} \end{bmatrix} \bullet \begin{bmatrix} r_{Y1} \\ r_{Y2} \end{bmatrix} \quad (12)$$

The estimated weights :

$$b'_1 = \frac{r_{22}r_{Y1} - r_{12}r_{Y2}}{r_{11}r_{22} - r_{12}^2} \quad (13) \quad b'_2 = \frac{r_{11}r_{Y2} - r_{12}r_{Y1}}{r_{11}r_{22} - r_{12}^2} \quad (14)$$

$\because r_{11} = r_{22} = 1$

The estimated weights become:

$$b'_1 = \frac{r_{Y1} - r_{12}r_{Y2}}{1 - r_{12}^2} \quad (15) \quad b'_2 = \frac{r_{Y2} - r_{12}r_{Y1}}{1 - r_{12}^2} \quad (16)$$

The variance of estimated weights:

$$\sigma^2(b'_1) = MSE \left(\frac{1}{1 - r_{12}^2} \right) \quad (17) \quad \sigma^2(b'_2) = MSE \left(\frac{1}{1 - r_{12}^2} \right) \quad (18)$$

or

$$\sigma^2(b'_1) = MSE \times VIF_1 \quad (19) \quad \sigma^2(b'_2) = MSE \times VIF_2 \quad (20)$$

where:

MSE is the mean squares error

VIF_i is the variance inflation factor of i^{th} estimated weight

In Eq. 11, there are three matrices needed to express the MLR system of normal equations. The first matrix is the injector-injector simple correlation matrix that can be estimated using the well rates of injectors. The vector, consisting b'_1 and b'_2 , represents the weights to be estimated. The elements, r_{Y1} and r_{Y2} , are the simple correlation coefficients of injector-producer well pairs that can be estimated using the well rates of the producer and nearby injectors.

The estimated weights, as depicted in Eqs. 15 and 16, are functions of the correlation coefficient of injector-producer pairs (r_{Y1} and r_{Y2}) and also functions of the correlation coefficient of injector-injector pairs (r_{12}). If injectors are independent r_{12} is zero, the weights are simply equal to the corresponding correlation coefficient of injector-producer pairs; these are referred to as simple weights. In Eqs. 15 and 16, the denominator is always positive and tends to increase the estimated weight. The numerator can be positive or negative where the estimated weight can have different sign from the corresponding simple weight. The discrepancies between the sign of the estimated weight and the sign of simple weight come from increasing the interaction (as r_{12} increases) among injectors. If injectors are independent, the estimated weights are equal to

the simple case weights where the discrepancies disappear. Also, the variability of the estimated weights is a function of the correlation coefficient among injectors as seen in Eqs. 17 and 18. This shows that the interactions between injectors also make the estimated weight less stable. The term variance inflation factor, VIF, comes from the fact that the variance of an estimated weight, as shown in Eq. 19 or 20, is inflated by this factor.

All of these observations suggest that MLR weights are more error-prone when the linear dependencies among injectors increase significantly, leading to have non-physical negative weights.

The collinearity creates severe problems in MLR or any linear estimation procedure. MLR attempts to separate the effect of each injector while holding other injectors constant, however, this is not possible when there is strong collinearity. The estimated MLR weights tend to vary widely from one sample data to another using the same injectors; they also become extremely unstable. The instability in regression results is reflected in very large standard errors for the estimated weights, as seen in the producer-two injectors case.

Perfect collinearity occurs when two injectors are identical. In this case, the injector-injector covariance matrix $X'X$, or the injector-injector correlation matrix does not have an inverse; it is a singular matrix where at least one or two of the singular values of the matrix are zeros. Perfect collinearity is rare in practice, however, a near-perfect or a strong collinearity is common in observational studies where two or more input variables are highly correlated. In this case, the $X'X$ matrix has an inverse but the estimated solution will be very sensitive to the input data so that a different sample data will give a different solution. This implies that there are near redundancies among the input variables (injectors); essentially the same information is being provided in more than one way.

Detecting Collinearity

The injector-injector simple correlation matrix can help to detect collinearity in the input data. It can not show the whole picture of the collinearity when more than two injectors are linearly correlated. However, it is a good starting tool to examine the collinearity in the input data.

There are many tools to detect collinearity in the observational data. some of these tools are:

The Condition Number and Condition Index

The condition number of a matrix X is defined as the ratio of the largest singular value to the smallest singular value (Belsley, 1991).

$$K = \left[\frac{\omega_{\max}}{\omega_{\min}} \right]^{1/2} \quad (\text{Eq. 21})$$

where:

K : the condition number

$\omega_{\max}$: the maximum singular value

$\omega_{\min}$: the minimum singular value

The condition number provides a measure of the sensitivity of the linear regression solution to small changes in input data. A large condition number indicates that a near-singularity is causing the matrix to be poorly conditioned. The condition number of injector-injector correlation matrix is one when all injectors are independent where all singular values are equal to one.

The condition number is extended to define the condition index for each principle component of the input matrix. The condition index is defined as

$$\delta_k = \left[\frac{\omega_{\max}}{\omega_k} \right]^{1/2} \quad (22)$$

where:

δ_k : the condition index of k^{th} principle component

$\omega_{\max}$: the maximum singular value in the system

ω_k : the singular value of k^{th} principle component

Belsley et al. (1980) suggest that condition indices around 10 indicate weak collinearity. Condition indices of 30 to 100 indicate moderate to strong collinearity and indices larger than 100 indicate strong collinearity.

Variance Inflation Factor and Multiple Determination Coefficient

Another common measure of collinearity is the variance inflation factor (VIF) for each estimated weight. The variance inflation factors are computed from the injector-injector simple correlation coefficient matrix $X'X$. Thus, the input variables (injectors) are centered and standardized to unit length. The diagonal elements of $(X'X)^{-1}$, the inverse of $X'X$, are the variance inflation factors. As stated in Eq. 23, the variance of each estimated weight is equal to the mean squares error multiplied by its corresponding variance inflation factor VIF when centered-scaled moments are implemented,

$$s^2(b'_i) = \text{MSE} * \text{VIF}_i \quad (23)$$

where:

$s^2(b'_i)$: the variance of i^{th} estimated weight using centered-scaled moments

MSE: the mean squares error of the system

VIF_i : the variance inflation factor of i^{th} estimated weight

The term variance inflation factor comes from the fact that the variance of an estimated weight, as shown in Eq. 23, is directly proportional to this factor. When all injectors are totally independent, the variance of each estimated weight is simply equal to the mean squares error of the system where the diagonal elements of the inverse injector-injector correlation matrix are all equal to one. However, when there is collinearity among injectors, the variance inflation factor of each weight tends to be greater than one. Thus, the variance of each weight is inflated by this factor. In the case of one producer and two injectors, VIF is simply equal to $(1 / (1 - r_{12}^2))$ as depicted in Eq. 17.

The link between the variance inflation factor and the multiple determination coefficient is through the relationship,

$$\boxed{\text{VIF}_i = \frac{1}{1-R_i^2} \quad (24)}$$

where R_i^2 is the multiple determination coefficient from the regression of i^{th} injector against the other injectors in the system. If there is a strong collinearity involving i^{th} injector and the other injectors, R_i^2 will be near one and VIF_i will be large. If i^{th} injector is orthogonal to the other injectors, R_i^2 will be zero and VIF_i will be one.

The variance inflation factors are simple diagnostics tool for detecting overall collinearity problems. Because of their simple computation, they are widely used in multiple regression procedures. However, they will not identify the source of the collinearity in the data.

Other diagnostics tools are able to detect and identify the source of the collinearity in certain cases. However, these tools follow lengthy procedures and require huge computational efforts. In this study, therefore, using condition number, multiple determination coefficient, and variance inflation factor are sufficient for detecting collinearity in the injection-production data.

Remedies for Collinearity

The only true solution to eliminate collinearity is to collect a new data in a manner to avoid this problem. This could be possible in controlled statistical experiments, which are usually designed to avoid collinearity. However, in all observational studies such as our case, we are restricted to deal with data as it exists.

There are many methods proposed to suppress the effect of collinearity. The most common approach is model re-specification. The idea of this method is to combine similar variables in groups or one variable can be chosen to represents other similar variables where the effects of collinearity disappear or suppress. In this case, high correlations among input variables are preferable to insure the reliability of this method. However, this method is not practical in our case where our objective is to determine individually the connectivity for injector-producer pairs.

Another common approach to collinearity is the use of some variable selection technique such as forward elimination method, backward elimination method or both. Similarly, the idea is to have less number of input variables to decrease the effect of collinearity. The problem with this technique is that the order of entry of the input variables is the criterion by which the importance of input variable is determined. If two injectors are highly correlated, where both of them carry the same information to the system, only the one that enters first will be statistically significant and accordingly enters the regression model.

Introduce additional information conveyed to the statistical procedure will significantly improve the inference about connectivity. Pressure data when it combines with the injection-production

data will definitely improve the inference procedure about injector-producer connectivity. However, this method depends on gathering information from other sources, which could be not available or accessible.

Biased estimation, mainly ridge regression, is one of the most practical remedies for collinearity. Ridge regression was developed for the purpose of trying to gain more precise or stable weights in the presence of the collinearity. This method enables MLR to determine more stable weights without the need to replace or lump injectors; it tries to reduce the sensitivity of the solution to the input data by making the injector-injector covariance matrix less singular. In this study, the collinearity problems in Magnus field data are investigated using ridge regression method.

Ridge Regression

Hoerl and Kennard (1970a, 1970b) introduced ridge regression as a tool to produce more stable weights with good mean squared error properties in the regions where input data suffer from collinearity. In MLR, the only optimization criterion or constraint is that the estimated weights must give the minimum error variance. There is no constraint or control about the sensitivity of solutions to the input data. Thus, the sensitivity of the MLR solutions to input data is not addressed in the usual multiple linear regression procedures such as MLR.

The ridge regression procedure attempts to portray the sensitivity of the estimates to collinearity in input data (Neter, 1996). It essentially shrinks the weights by applying a penalty on their size where the uncertainty of estimated weights is minimized. This is accomplished by increasing the main diagonal elements of injector-injector simple correlation matrix $X'X$ to produce more precise weights and reduce the mean square error. It does so by adding a constant C , called the ridge parameter, to the main diagonal elements of $X'X$ matrix. Following the addition of the parameter, MLR or least squares method applies where more precise weights are estimated. The mathematical representation of this procedure is

The Ridge Regression estimator:

$$b_{ridge} = (R_{XX} + C I)^{-1} (R_{XY}) \quad (25)$$

The minimum sum of squares of the residuals:

$$SSE = (Y - b_{ridge} X)' (Y - b_{ridge} X) \quad (26)$$

The variance-covariance matrix of estimated Coefficient:

$$s^2(b_{ridge}) = MSE (R_{XX} + C I)^{-1} R_{XX} (R_{XX} + C I)^{-1} \quad (27)$$

where :

$(p \times p) (R_{XX}) \equiv$ injector-injector correlation matrix

$(p \times 1) (R_{XY}) \equiv$ injector-producer correlation vector

b_{ridge} : the ridge estimated weights using centered-scaled moments

C : the ridge parameter

I : the identity matrix

MSE :the mean squares error

The ridge regression can be also applied using centered moment. The normal linear system using injection-production notations can be represented as follows

$$\begin{bmatrix}
 \sigma_{11}^2 (1+C) & \sigma_{12}^2 & \sigma_{13}^2 & \cdot & \cdot & \cdot & \sigma_{1p}^2 \\
 \sigma_{21}^2 & \sigma_{22}^2 (1+C) & \sigma_{23}^2 & \cdot & \cdot & \cdot & \sigma_{2p}^2 \\
 \sigma_{31}^2 & \sigma_{32}^2 & \sigma_{33}^2 (1+C) & \cdot & \cdot & \cdot & \sigma_{3p}^2 \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \sigma_{p1}^2 & \sigma_{p2}^2 & \sigma_{p3}^2 & \cdot & \cdot & \cdot & \sigma_{pp}^2 (1+C)
 \end{bmatrix}
 \cdot
 \begin{bmatrix}
 \lambda_1 \\
 \lambda_2 \\
 \lambda_3 \\
 \cdot \\
 \cdot \\
 \cdot \\
 \lambda_p
 \end{bmatrix}
 =
 \begin{bmatrix}
 \sigma_{y1}^2 \\
 \sigma_{y2}^2 \\
 \sigma_{y3}^2 \\
 \cdot \\
 \cdot \\
 \cdot \\
 \sigma_{yp}^2
 \end{bmatrix}
 \quad (28)$$

Where:

C = the ridge parameter, taking values (0, 0.1, 0.2, 0.3,, 1.0)

λ_i = the estimated weight for the i^{th} injector

The ridge regression is a biased estimator in the sense that the minimum error variance constraint is relaxed to some extent to obtain more stable weights. The right trade-off between bias and variance depends on selecting the appropriate value of ridge parameter. The best value gives the lowest minimum error variance. However, increasing values are associated with lower mean squares error up to a point, when the mean squares error starts to increase.

The central problem with ridge regression is finding the appropriate value of C . Several ad hoc methods have been proposed, but none of these can provide a sound justification for determining the right trade-off between bias and variance. The most convenient method is the ridge trace plot in which plots values of estimated weights against possible values for the ridge parameter C . The value of C at the point where weights stabilize is usually considered the optimum.

The procedure for ridge regression can be as follows:

- 1 Apply MLR with computing the required statistical measures
- 2 Detect collinearity using an appropriate tool such as condition number, variance inflation factor, or multiple determination coefficient
- 3 Apply the ridge parameter for the diagonal elements of injector-injector covariance matrix
- 4 Re-apply MLR and compute the ridge weights with required statistical measures
- 5 Keep detecting collinearity with the appropriate tool
- 6 Repeat the above procedure with different values of ridge parameter

Below in this report an example from the Magnus field is used to illustrate the idea of ridge regression. The MLR analysis and the required statistical inferences are shown in Table 2. The

linear model consists of intercept and four injectors. The statistical measures of producer M30ZA5 are listed below the name of the producer. The intercept is indicated by Lamda0. The first column gives the name of injectors; the second column gives the distance of each injector from the producer M30ZA5; the average rate and the variance of each injector are listed in the next two columns. The estimated weights and their variances are shown in the next two columns indicated Weights and std. The last column, indicated λ_0 , is shown the simple weights of the system if injectors are independent. Also, the injector-injector simple correlation matrix and the injector-producer simple correlation vector are listed.

TABLE 2: CENTERED-MLR ANALYSIS

Centered MLR:

Time Window:

3/20/1998
10/9/1998

M30ZA5:

Correl

0.835

MSE

1.29E+07

Predict.

0.697

std

3596.59

Lamda0

18022.65

Lamda0 StD

--

Avrage q

32381.66

StD q

6464.61

StD q estim

5395.43

A sign of strong collinearity

Inj.Name	Dist.	Average Rate	inj.Std	Weight	std	Part Coef	λ_0
M34C3	1980.87	43342.9	9829.8	-0.291	0.047	16.318%	0.105
M06B3	2057.05	24277.4	5238.3	1.151	0.078	52.076%	0.830
F2	2372.31	21413.9	6892.7	-0.145	0.060	2.828%	-0.218
M31ZC4	2523.34	6306.7	3293.4	0.340	0.116	4.137%	1.185

	inj-inj r				Pro-inj r
	M34C3	M06B3	F2	M31ZC4	M30ZA5
M34C3	1.00	0.70	0.60	0.25	0.16
M06B3	0.70	1.00	0.27	0.52	0.67
F2	0.60	0.27	1.00	-0.35	-0.23
M31ZC4	0.25	0.52	-0.35	1.00	0.60

The discrepancies between estimated weights and simple weights are clear for injector M34C3. This is usually a strong sign of collinearity and this can be easily detected by examine the simple correlation matrix for M34C3. M34C3 is highly correlated with M06B3 and F2 where it has very low correlation with the producer. However, the simple correlation matrix of injectors, as stated above, does not show the whole picture of the collinearity in the data.

The ridge trace plot, Figure 26, shows the trends in weights with ridge parameter. Clearly, the MLR estimates are altered significantly as the ridge parameter increases. Generally, as ridge parameter increases, weights tend to be smaller. The most important observation is that the MLR negative weight of M34C3 tends to become less negative until it ultimately reaches zero. This eliminates the discrepancies between the M34C3 estimated weight and simple weight, which proves that the collinearity is the main cause for non-stable negative weights for this case.

Figure 26: The ridge trace plot

Collinearity in the data is examined by the multiple determination coefficient MDC for each injector as shown in Figure 27. Injectors that are highly correlated are indicated by high MDC values when the ridge parameter is zero. This means the collinearity is really a major problem in the examined data. The MDC tends to decrease as the ridge parameter increases, which means the collinearity of the system is suppressed with the ridge regression as is expected.

Figure 27: MDC Vs. The ridge parameter

The stability of weights and the mean square error of the system are represented by Figure 28, and 29 respectively. The weights become more stable with increasing the ridge parameter, Figure 28. At the same time, the R-squared of the system decreases gradually with ridge parameter, Figure 29. This indicates that the estimation becomes biased with ridge regression. The best value of the ridge parameter is that gives the lowest mean square error (the least reduction in R of the system) and the least variance of estimated weights (Gibbons and McDonald, 1984). This is the criterion to select the right value for the ridge parameter; however, implementation is difficult because the minimum value of mean square error (the highest value of R of the system) is not shown in Figure 29. What is shown is that the mean square error monotonically increases (R of the system decreases) with the ridge regression, so that there is no minimum value of mean square error where this criterion can be applied. Therefore, the appropriate value of ridge parameter could be selected as that value at which the weights are stabilized with the minimum increase in mean square error or reduction in the R of the system (Belsley, 1991 and Neter, 1996). From Figure 26, $C = 0.4$ of ridge parameter is that at which weights are relatively stabilized with minimum increase in mean square error.

Figure 28: Weight Standard deviation Vs. The ridge parameter

Figure 29: R of the system Vs. The ridge parameter

Figure 30: MLR weights compared to simple weights

Figure 31: The ridge weights compared to simple weights

A graphical representation of MLR weights and ridge weights are in Figures 30 and 31, respectively. They are compared to simple weights, the base case. If estimated weights and simple weights have the same sign, the points will be in the first or third quadrant of the plots indicated the consistency. If they are in different sign, points will be in the second or fourth quadrant of the plots indicated the discrepancy. There are two different types of estimated negative weights. When the estimated weights and their corresponding simple weights are negative, these weights tend to be always negative even in the presence of collinearity. This is the case for injector F2 where its weight is always in the fourth quadrant of both figures. However, when estimated and simple weights have different sign, this discrepancy tends to disappear after reducing the collinearity. Similarly, this could happen for the positive weights also. Therefore, we have to distinguish between these two types. If the weights are in different sign from that of simple weights, they are referred to as non-stable weights due to collinearity. If they have the same sign, they are referred to as stable weights.

The Modified Ridge Regression Method

The simple weights of independent case do not appear with ridge regression method. The reason is that the ridge method applies extra weights on the main diagonal elements of injector-injector correlation matrix where all singular values of the matrix will shift a way from zero. This makes the MLR solution to be less sensitive to the collinearity problem in the input data. However, this does not insure the link to independent case.

A modification could be proposed to insure this link so that simple weights appear at the end of this method. This could be done by applying less weight on the off-diagonal elements of injector-injector correlation matrix (or covariance matrix). It does so by subtracting the constant C from the off-diagonal elements of the matrix. Following the subtraction of the parameter, MLR applies where more precise weights are estimated. The normal equation of the modified method is as follows

$$\begin{bmatrix}
 \sigma_{11}^2 & \sigma_{12}^2(1-C) & \sigma_{13}^2(1-C) & \cdot & \cdot & \cdot & \sigma_{1p}^2(1-C) \\
 \sigma_{21}^2(1-C) & \sigma_{22}^2 & \sigma_{23}^2(1-C) & \cdot & \cdot & \cdot & \sigma_{2p}^2(1-C) \\
 \sigma_{31}^2(1-C) & \sigma_{32}^2(1-C) & \sigma_{33}^2 & \cdot & \cdot & \cdot & \sigma_{3p}^2(1-C) \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \sigma_{p1}^2(1-C) & \sigma_{p2}^2(1-C) & \sigma_{p3}^2(1-C) & \cdot & \cdot & \cdot & \sigma_{pp}^2
 \end{bmatrix} \bullet \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_p \end{bmatrix} = \begin{bmatrix} \sigma_{y1}^2 \\ \sigma_{y2}^2 \\ \sigma_{y3}^2 \\ \cdot \\ \cdot \\ \cdot \\ \sigma_{yp}^2 \end{bmatrix} \quad (\text{Eq. 29})$$

Where: $C = (0, 0.1, 0.2, 0.3, \dots, 1.0)$
 λ_i = the estimated weight for the i^{th} injector

When the value of ridge parameter is one, simple case definitely will appear. In terms of collinearity, this procedure eliminates explicitly the linear dependencies among injectors by reducing the off-diagonal correlation coefficients in the matrix; thus more stable weights will be estimated. For mean square error, all inferences about statistical measures are related directly to the simple case where a better criterion for selecting the best value of ridge parameter would reveal. This new procedure is referred to as the modified ridge method.

To examine the significance of this procedure, let us use the same example from Magnus field data. The modified ridge trace plot, Figure 32, shows different trends, compared to the ridge method, for the weights with the parameter. When C is zero, weights are simply the original weights of the system. When C is one, weights are the simple case solution where injectors are totally independent. This representation explicitly indicates the link between the original system and the simple case system where it helps to examine the causes of weights deviations from independent case. The weight trend of M34C3 explains clearly the discrepancies between the original and simple weights. This was not clear using the ridge method, Figure 26.

Figure 32: The modified ridge trace plot

Collinearity in the data is presented by Figure 33. The new procedure forces MDC trend of all weights to end at zero MDC, reflecting the independent among injectors. In the ridge method, all injectors reached to zero MDC at different values of ridge parameter, Figure 27. Also, negative MDC trends appeared because there is no upper bound for the ridge parameter in previous method. In the new procedure, MDC trend monotonically decreases where it is non-linear. However, it is always linear in the ridge method. This indicates that collinearity is treated differently in both cases. Singular values are linearly shifted by the same increment (the ridge parameter) when the parameter is added to the main diagonal elements of the injector-injector matrix. This explains the linear trend of MDC. In the new procedure, singular values are linearly shifted but towards the singular values of independent case.

Figure 33: MDC Vs. The modified ridge parameter

The variance of each weight and the mean square error of system are depicted by Figures 34 and 35, respectively. The mean square error of the system increases (R of the system decreases) as in the case with ridge method. The trend of weight variance is different where there is a minimum for each weight, as seen in Figure 34. This minimum value gives the criterion to select the appropriate value of the modified ridge parameter so that a stable estimate with a minimum increase in the mean square error can be determined.

The variance of estimated weight is function of two components: mean square error and variance inflation factor VIF. With decreasing collinearity, the mean square error tends to increase gradually whereas VIF tends to decrease steeply so that the dominant component is VIF. As a result, variances of weights tend to decrease with eliminating collinearity by the ridge method. In the modified ridge method, the mean square error tends to increase gradually to a point and then start to amplify sharply whereas VIF tends to decrease steeply to a point and then start to decrease very slowly. Thus, a minimum value should appear in the variance of the estimated weight.

Figure 34 The weight standard deviation Vs. The modified ridge parameter

Figure 35: R of the system Vs. The modified ridge parameter

From Figure 34, the minimum value gives the optimum value of modified ridge parameter which is equal to 0.5 in this case. Based on this value, the most stable weights with the least increase in mean square error can be determined from Figure 32.

Figure 36: The modified ridge weights compared to simple weights

Figure 36 represents the modified ridge weights compared to the simple case weights. If weights are assumed random variables characterized by normal distributions, estimated weights and their

variances provided by modified ridge method give the characteristics of these distributions. The distribution characteristics are

	M34C3	M06B3	F2	M31ZC4
Weights	-0.022	0.751	-0.218	0.802
Std	0.033	0.062	0.046	0.095

According to the characteristics of M34C3, 95 % probability that the weight to have a value between $-0.022+2*0.033$ and $-0.022- 2*0.33$. This means the weight random variable is symmetric around the zero so that it can be safely assumed to be equal to zero. Therefore, M34C3 injector should be excluded from the model since it has no significant impact on the producer. Because there are no discrepancies between F2 estimated weight and its corresponding simple weight and there is no negative connectivity, the weight is referred to as stable negative weight where it should be excluded from the model. Therefore, M06B3 and M31ZC4 are only in contact with the producer.

It would be interesting if M06B3 and M31ZC4 weights, estimated in the presence of F2 and M34C3, are compared to their weights without exist the others. The new weights and their variances, estimated using the modified ridge method, are as follow

	C	M06B3	M31ZC4
ALL WELLS	0.5	0.751 +/- 0.062	0.802 +/- 0.095
KEY WELLS	0.5	0.681 +/- 0.063	0.904 +/- 0.1

Both systems provide similar weights characteristics insuring the reliability and stability of the estimated solution using the modified ridge method.

Summary of Results on Task 3

1. The MLR weights are not sensitive to the injection rate and the injection rate variability.
2. The MLR weights strongly depend on the condition of constant BHP. Changing the BHP of one producer during the analyzed time leads to the presence of both negative and the abnormally large positive weights. The R^2 of the correlation in these cases is usually small.
3. The MLR weights are strongly affected by the autocorrelation between injectors or between one injector and a group of injectors. Large positive weight may result in this case. The R^2 in this case may still be very high. A different way to estimate covariances may diminish this problem.

4. The average rate of the producer affects the interwell MLR weights. Wells with large average rate will have larger interwell MLR weights from all injectors
5. Under certain conditions, including steady state, constant producer BHP's, and relatively equal average production rates, the MLR weights may indicate the degree of reservoir connectivity between injector and producers.
6. When the producer average rates are not equal, the MLR weights may indicate the connectivity, but relatively for each producer. Comparing the MLR weights for different producer in this case may not be justified.

RESULTS AND DISCUSSION: TASK 4 – TEST AND EXTEND TOLERANCE FOR VARIABLE COMPRESSIBILITY CASES

Work in this area has moved in another direction, namely extension to prediction of oil rates.

Oil prediction is based on a power law relationship between WOR and cumulative water injected (Eq. 30). The approach taken is to superimpose this relationship onto the allocations determined by the weights.

We use this relationship for each producer and replace the cumulative water injected by total fluids produced. Hence the oil prediction model builds on the total fluids produced model and further, relies on its accuracy. The predicted water-oil-ratio is

$$\text{WOR} = a \cdot W_i^b \quad (\text{Eq. 30})$$

where W_i : Cumulative water injected into well i

$$W_i \approx N_p + W_p = Q_p$$

In the case we use one of the MLR methods where we obtain a total liquid rate prediction without an explicit information regarding the origin of the injected water, we can linearize the expression by taking logs:

$$\log(\text{WOR}) = A + b \cdot \log(W_i) \quad (\text{Eq. 31})$$

We now proceed to develop a linear regression on this relationship to find A and b . The squared error is:

$$(\text{LOG}(\text{WOR}) - \text{LOG}(\text{WOR})^*)^2 =$$

$$\text{LOG}^2(\text{WOR}) - 2 \text{LOG}(\text{WOR}) \text{LOG}(\text{WOR})^* + \text{LOG}^2(\text{WOR})^*$$

$$\log^2(\text{WOR}) - 2 \cdot \log(\text{WOR}) \cdot \log(\text{WOR})^* + \log^2(\text{WOR})^*$$

$$\log^2(\text{WOR}) - 2 \cdot \log(\text{WOR}) \cdot (A + b \cdot \log(W_i)) + (A + b \cdot \log(W_i))^2$$

Taking the expectation:

$$E[\log^2(\text{WOR}) - 2 \cdot \log(\text{WOR}) \cdot (A + b \cdot \log(W_i)) + (A + b \cdot \log(W_i))^2] \quad (\text{Eq. 32})$$

We now minimize this expectation by taking the derivatives with respect to A and b and equating them to zero.

$$\frac{\partial E[\dots]}{\partial A} = 0$$

$$\frac{\partial E[\dots]}{\partial b} = 0$$

This yields:

$$-2 \cdot E[\log(WOR)] + 2 \cdot A + 2 \cdot b \cdot E[\log(Wi)] = 0$$

$$-2 \cdot E[\log(WOR) \cdot \log(Wi)] + 2 \cdot A \cdot E[\log(Wi)] + 2 \cdot b \cdot E[\log^2(Wi)] = 0$$

In matrix form:

$$\begin{pmatrix} 1 & E[\log(Wi)] \\ E[\log(Wi)] & E[\log^2(Wi)] \end{pmatrix} \begin{pmatrix} A \\ b \end{pmatrix} = \begin{pmatrix} E[\log(WOR)] \\ E[\log(WOR) \log(Wi)] \end{pmatrix} \quad (\text{Eq. 33})$$

Relative permeability and the b exponent

The work by Yortsos, et al. (1999) shows that if the relative permeability oil to water ratio can be approximated by a power law relationship such as

$$k_{ro}/k_{rw} \approx (1 - S_w - S_o)^m \quad (\text{Eq. 34})$$

then, under appropriate conditions, this slope of the WOR curve can provide information on the *exponent* of the power-law dependence of the oil's relative permeability on saturation.

For late time behavior the WOR would follow:

$$\log(WOR) \approx \frac{m}{1-m} \log(t) + H \quad (\text{Eq. 35})$$

where H is a constant and t is a dimensionless injection time expressed as pore volumes injected.

By comparison to: Eq. 31

$$b = \frac{m}{m-1} \quad (\text{Eq. 36})$$

The 5x4 Homogeneous Synfield

This is a *well-behaved* case where we can see that the power law relationship holds constant for the whole range of WOR.

Figure 37: WOR vs. Cumulative Liquid Fluids produced by well P01. A constant power law relationship can be observed.

The regression interval has to be after breakthrough and care has to be taken to avoid initial period where small fluctuations in WOR generate relatively large fluctuations in the $\log(\text{WOR})$.

As we can see in the graphs (Figures. 38, 39, 40) the regression results for producer 1 are very good. Similar results are obtained for the other three producers.

Figure 38: Oil rate vs time for producer P01. Regression results compare favorably with data. Note that first points were omitted from the regression.

Figure 39: Oil rate regression results vs Oil Rate data for producer P01. Regression results vs. data. Lie on a 45 degree line.

Figure 40: WOR regression results vs WOR data for producer P01. Regression results vs. data. lie on a 45 degree line.

5x4 Synfield - The b Exponent

As mentioned before under favorable conditions the value of the regressed b exponent might contain information on the relative permeability. This was tested for the 5x4 Homogeneous case to assess whether further research was warranted.

The following is the relative permeability set input into the simulator for the 5x4 homogeneous Synfield case (Figure 41, Figure 42)

Figure 41: Relative permeability data input for the 5x4 homogeneous Synfield case.

Figure 42: Relative permeability ratio and corresponding power law regression for the 5x4 homogeneous Synfield case. Exponent value is 1.74.

The value for m is 1.74. Applying the relationship given by:

$$b = \frac{m}{m-1} \quad (\text{Eq. 37})$$

$$b = \frac{1.74}{1.74-1} = 2.35$$

This value compares favorably with the value of 2.52 regressed previously. More research will be needed under heterogeneous setups to better interpret these results. It is suspected that this value should characterize some kind of average around the producer.

Pseudo Allocation Models and the Oil Model

Pseudo allocation models are based on the use of regression models with weights associated only to neighboring injectors to the producer. The motivation for developing the pseudo allocation Models is shown in this section with an example using one more time the 5x4 homogeneous synfield.

Assuming that the pseudo allocation models provide information on how the water is being allocated from the injectors to it's neighboring producers, then a WOR-CumProducedFluids power law can be applied to the individual volumes contributed by each injector to the producer.

The results are shown for the Producer 1 using the Model 4b-noncentered allocation. The results show similar declining oil rates for the volumes being swept by injectors 2 and 3 and scaled down behavior for the volume swept by injector 3. The oil rates seem to be converging.

Figure 43: Results of applying the power law oil model to the *Model 4b-non centered* allocation. Individual oil contribution from each of the three neighboring injectors to producer P01 are inferred. Left figure: linear graph; right figure shows semilog graph.

Figure 44 Graph showing inferred WOR evolution for the volume of fluids supported by each of the three injectors supporting producer P01.

Figure 45 : Results of applying the power law oil model to the *Model 4b-non centered* allocation. Oil rate regression vs. oil rate data lie on 45 degree line. The oil regression shows very good agreement except for the initial part of the history (high oil rates) where WORs are very low.

The 5x4 Faulted Homogeneous Synfield

The oil rate for P01, in this case, exhibits a more complex behavior. This is because of the dissimilar flow geometries and breakthrough times that I02 and I05 that feed P01 have.

Figure 46: Graph on the left shows a more complex WOR behavior with an early time trend transitioning into a late time trend. On the right, graphic representation of the results of Model4b on P01 suggesting a lack of connection between the left and right blocks.

Based on this one might attempt a piece wise regression or alternatively only regress the early behavior or the late time behavior. Since late time behavior is usually more appealing for predictions purposes, we will show an example of the latter first:

Figure 47: Regression using late time information. On the left WOR and WOR regressed against cum fluids produced. On the right, plot of Q_o and Q_o regressed vs. time. Late time (aprox >2000d) agreement is good.

Early time behavior can also be matched as shown in Fig 48.

Figure 48: Regression using early time data only. A quality of the match is lost after aprox. 1000d.

As we have seen this more complex behavior of the WOR curve cannot be matched with a single two parameter power law relationship. However, a combination of these might prove more appropriate.

This is another motivation for developing Pseudo Allocation Models. In order to illustrate this an example follows which uses a Model 4b allocation. As mentioned in the previous section, note, that since the model was constructed assuming absence of faults, the allocation to P01 is from I01, I02, I03 and although volumes from I01 is negligible and from I03 are small (which actually absent) the model lacks allocation from I05 by construction (see Figure 46). This is physically incorrect and results from ignoring the presence of the fault at the time of constructing the model. The example is still appropriate to show the results of combining the response from two injectors (I02, I03) each with separate WOR power law relationships.

Figure 49: Results for WOR I1, WOR I2 and WOR I3, their combined result into WOR Pred and comparison WOR data. By combining three simple power law relationship a match to a more complex behavior can be attempted.

Figure 50: Combined results of Model4b-noncentered model and power law oil models. The combined WOR behavior of P01's three injectors (see Fig PG49) provide the capability of integrating early and late time behaviors resulting in an overall improvement on the quality of the regression.

RESULTS AND DISCUSSION: TASK 5 – TEST OF METHOD ON LARGE DATASET

The technique in the report was applied to Magnus field data in the North Sea. This section presents a brief description of the field. The general procedure followed to use the proposed method is illustrated. Finally, a discussion of application results is provided.

Magnus field is the most northerly field in the UK North Sea, with 14 platforms and two subsea producers, and 13 injectors of which 8 are subsea. The reservoir is an Upper Jurassic turbidite reservoir, with a high net to gross upper reservoir and a low net to gross lower reservoir. Figure 13 shows the map of the field where the reservoir is a south-east dipping tilted fault block, overlain by Cretaceous mudstones. Reservoir quality improves towards the crest of the structure. Production and injection wells were generally perforated in all reservoir intervals encountered. The vertical fault is the most significant physical barrier in the field holding back high pressure support up to 4000 psi from the crestal area. Other horizontal barriers marked on Figure 13, generate some pressure difference but they are thought to allow some sweep.

The waterflood is essentially peripheral intended for crestal zone pressure support. In early 90's, the field came off plateau because water breakthrough occurred in flank producers, scale problems, and cross flow in producers. When the actual complexity of the field appeared, the strategy of the waterflood management was devoted to implement the zonal water injection where producers and injectors were completed into certain layers to maintain production [6].

Figure 51 Magnus field structure map

The MLR and modified ridge methods have been tested using the large Magnus production data. The Magnus data consists of daily rate measurements from permanent downhole sensors. The field is characterized by wells having frequent shut-in periods. The injection and production history is depicted in Figure 52.

Figure 52 Injection and production history.

From the waterflood history, the best time period for data analysis is after implementing the zonal water injection where the ability to infer the connectivity among injector-producer pairs are more reasonable. Therefore, time period starting in early 1998 was selected for the analysis. There are over 230 points, which is equivalent to 7 months injection time. The crestal area is the target zone in the analysis with 7 injectors and 8 producers in Figure 51.

The departure from normality of the production and injection data is checked, as shown in Figure 53 and 54. The data in Figure 53 presents the oil production rate for all producers where the data are stacked on top of each other independently from time factor. The near straight line suggests the normality of Magnus production data. Similarly, the water injection rate for all injectors is a normal distribution as seen in Figure 54.

Figure 53 The normal probability plot for production data

Figure 54 The normal probability plot for injection data.

The MLR Technique is based on several assumptions (Albertoni and Lake, 2003). The general assumption is that within the period of time selected for the analysis, all the parameters in the field must be constant with the exception for the injection and production rates. The constant reservoir conditions are rarely attained in waterfloods. However, the selected period for Magnus field tends to be at satisfactory constant operating conditions.

Procedure of Data Analysis

The procedure implemented in this application includes:

Determine time delay between the producer and the nearby injectors
Apply MLR and the modified ridge method

We exhibit this procedure using different producers from the crestal area.

Figure 55 The crestal area

The Producer: M30ZA5.

The system is the producer M30ZA5 and the nearby injectors in the crest area, Figure 55. The time delay for the system can be examined using the time series of production and injection rates, Figure 56.

Figure 56 Time series of M30ZA5 and nearby injectors

The first set of plots shows the actual rates as function of time and the second set represents the actual rates after applying a smoothing filter. There appears to be a major signal applied by all injectors at the same time. The time lag of the system can be determined if major signals can also be seen in the producer rate profile. This signal can be observed easily in the M30ZA5 history, so that the time lag is about 20 days. It is impossible to determine which injector relatively in a good communication with the producer as the signal is sent by all injectors simultaneously. This could be a foreshadowing of the presence of colinearity in the data.

The change in operating condition of the producer can be examined using these plots. If there is major signal in the producer that cannot be explained by the injector histories, the only reason for this behavior is a sudden change in the operating conditions. Constant operating conditions, mainly bottom-hole pressure, are required for MLR analysis. Therefore, this kind of signals should be removed from the producer trend. For M30ZA5, no evidence of this exists.

The simple correlation coefficient analysis for each injector-producer pair will provide a detailed view about time delay for each injector-producer pair. Figure 57 depicts this analysis using both actual and smooth data. Injectors F2, M17C5, and F4ZB are always negatively correlated with the producer. This means that simple weights of these injectors are negative too, so they will be characterized by negative estimated weights. Thus, they should be excluded from the system.

Figure 57 The simple correlation coefficient analysis

On other hand, injectors M06B3, M31ZC4, and F4ZA are strongly positively correlated with the producer M30ZA5. Thus, this suggests that they are in communication with the producer. There is a sudden jump in the correlation coefficient trends around 20 days, which matches with the observation from time series analysis. The average time lag is 19 days.

If the MLR statistical measures improve significantly with the proposed time lag, this indicates this value represents to some extent attenuation in the field. Table 3 shows the results of MLR analysis with the best statistical measures, using the proposed time lag value.

TABLE 3 MLR ANALYSIS

Centered MLR:

Time Window: 3/20/1998
10/9/1998

Smooth window = 6
Shift = 19
No. of point= 204 days

M30ZA5:

Correl 0.855
Predict. 0.731
Lamda0 19203.52
Lamda0 StD --
Avrage q 32381.66
StD q 6464.61
StD q estim 5526.88

MSE 1.16E+07
std 3412.68

Inj.Name	Dist.	Average Rate	inj.Std	Weight	std	Part Coef	□□
F4ZA	1885.72	15659.9	5013.2	-0.121	0.105	0.681%	0.532
F4ZB	1956.73	1800.5	1676.4	-0.181	0.177	0.528%	0.003
M34C3	1980.87	43342.9	9829.8	-0.316	0.045	20.121%	0.105
M06B3	2057.05	24277.4	5238.3	1.183	0.083	50.821%	0.830
F2	2372.31	21413.9	6892.7	0.093	0.078	0.714%	-0.218
M31ZC4	2523.34	6306.7	3293.4	0.507	0.161	4.796%	1.185
M17C5	2935.20	33280.4	11316.2	-0.144	0.034	8.415%	-0.157

	inj-inj r							Pro-inj r
	F4ZA	F4ZB	M34C3	M06B3	F2	M31ZC4	M17C5	M30ZA5
F4ZA	1.00	-0.41	0.55	0.65	0.23	0.69	0.29	0.41
F4ZB	-0.41	1.00	-0.16	0.00	-0.06	-0.26	-0.01	0.00
M34C3	0.55	-0.16	1.00	0.70	0.60	0.25	0.44	0.16
M06B3	0.65	0.00	0.70	1.00	0.27	0.52	0.20	0.67
F2	0.23	-0.06	0.60	0.27	1.00	-0.35	0.75	-0.23
M31ZC4	0.69	-0.26	0.25	0.52	-0.35	1.00	-0.21	0.60
M17C5	0.29	-0.01	0.44	0.20	0.75	-0.21	1.00	-0.28

Collinearity is quite pronounced in the data, which can be seen from the simple correlation coefficient matrix. For eliminating injectors characterized by negative weights, Albertoni and Lake (2003) proposed the successive elimination method (SEM) which basically eliminates all injectors with negative weights. This method works only when all eliminated injectors are characterized by stable negative weights. Otherwise, when injectors have negative weights because collinearity, the heuristic method tends to eliminate highly potential injectors.

By applying the modified ridge method and removing injectors within significant weights, only M06B3 and M31ZC4 are in good communication with M30ZC4. The weights representations are shown in Figure 58 and 59.

Figure 58 MLR weights after applying the successive elimination method

Figure 59 The modified ridge weights

There are other statistical problems inherent in the linear estimation procedure of MLR. Large mean square error, insufficient variations in the input variables and short data are some of these problems. Using the rank of the data instead of actual data resolves some of these problems. Therefore, MLR and modified ridge method was investigated using the rank of Magnus field data.

For the same system, the time series of injection and production rates rank is shown in Figure 60. The first set of plots represents the rank of actual data for the system and the second set represents the rank of the actual data with smooth filter.

Figure 60 Time series of injection-production rates rank

Generally, using the rank of the data amplifies the small signals that are small in the actual data. Figure 61 shows the Spearman rank correlation coefficient analysis. The first plot shows the rank correlation for the rank of the data whereas the second plot shows the rank correlation for the rank of smoothed data. Only M06B3 and M31ZC4 have strong correlation with the producer. This matches with the result of inference connectivity using the actual data. Our experience with Spearman rank analysis is that it gives essentially the same insights as does analysis directly on the data.

Figure 62 demonstrates the result of simple correlation coefficient analysis which shows that the average time lag is also about 20 days.

Figure 61: The spearman rank correlation coefficient

Figure 62: The simple correlation coefficient using the rank of the data.

Table 4 shows the MLR analysis using the rank of the data. The simple weights of the system match their corresponding injector-producer simple correlation coefficients. When rank of the data is used, this makes all variables have the same variance. Therefore, using the rank of the data resolves the problem when injectors variances are different in order of magnitude. Also, the effect of collinearity tends to decrease because different data is implemented. The physical interpretation of weights in the rank analysis is, however, not clear.

TABLE 4: THE MLR ANALYSIS USING THE RANK DATA

Rank Centered MLR:				Smooth window = 6 Shift = 19 No. of point = 204 days			
Time Window:							
3/20/1998 10/9/1998							
M30ZA5:							
Correl	0.783	MSE	1.40E+03				
Predict.	0.613	std	37.36				
Lambda0	63.34						
Lambda0 Std	--						
Average q	102.50						
Std q	59.03						
Std q estim	46.23						
Inj.Name	Dist.	Average Rate	inj.Std	Weight	std	Part Coef	
F4ZA	1885.72	101.5	59.0	-0.174	0.073	2.843%	0.173
F4ZB	1956.73	80.3	56.5	-0.101	0.066	1.196%	0.031
M34C3	1980.87	101.2	58.7	-0.269	0.057	10.369%	-0.153
M06B3	2057.05	97.8	56.1	0.656	0.061	37.117%	0.705
F2	2372.31	100.3	58.3	-0.046	0.081	0.163%	-0.420
M31ZC4	2523.34	77.4	56.3	0.351	0.070	11.381%	0.587
M17C5	2935.20	100.9	58.6	0.054	0.071	0.299%	-0.343

	inj-inj r							Pro-inj r
	F4ZA	F4ZB	M34C3	M06B3	F2	M31ZC4	M17C5	M30ZA5
F4ZA	1.00	-0.61	0.22	0.27	0.20	0.54	0.10	0.17
F4ZB	-0.61	1.00	-0.22	0.05	-0.32	-0.24	-0.04	0.03
M34C3	0.22	-0.22	1.00	0.23	0.44	0.00	0.13	-0.15
M06B3	0.27	0.05	0.23	1.00	-0.29	0.49	-0.30	0.67
F2	0.20	-0.32	0.44	-0.29	1.00	-0.31	0.70	-0.42
M31ZC4	0.54	-0.24	0.00	0.49	-0.31	1.00	-0.38	0.56
M17C5	0.10	-0.04	0.13	-0.30	0.70	-0.38	1.00	-0.34

By applying the modified ridge method and removing injectors with insignificant weights, M06B3 and M31ZC4 are the only injectors in good communication with M30ZC4. The weights representations are shown in Figure 63 and 64.

Figure 63: MLR weight after applying the successive elimination method

Figure 64 The modified ridge weights

The Producer: M36ZB7

The system includes the producer M36ZB7 and the nearby injectors, Figure 55. The time series of the rates using both the actual data and the rank are in Figure 65 and 66, respectively. Using

the actual data, there is no major signal so that the time lag of the system could be determined. Thus, MLR analysis using the actual data cannot perform.

However, the rank of the data shows some interactions. This illustrates that the rank analysis of the data can provide insight about the reservoir even when the data analysis does not. The simple correlation coefficient using the rank of the data is depicted in Figure 67. The time lag is about 40 days.

Figure 65: Time series rates using actual data

Figure 66: Time series rates using the rank of the data

Figure 67: Correlation coefficient analysis using ranks

Table 5 shows the results of constrained MLR analysis, zero intercept, using positively correlated injectors at the indicated time lag. For the rank data, the weight estimated for M38C6 is negative. However, this injector is adjacent to a producer that is only few hundreds meters away. Also, it is highly correlated with M36ZB7 as seen in Figure 66. By examining the injector-injector simple correlation matrix, we see that M38C6 is highly correlated with injector M34C3.

TABLE 5: MLR ANALYSIS USING THE ACTUAL DATA AND THE RANK OF THE DATA

Centered ABMLR: S.E.M

Time Window:

2/13/1999
6/13/1999

M36ZB7:

Correl

0.865

MSE

1.30E+07

Predict.

0.585

std

3605.88

Lamda0

0.00

Lamda0 StD

--

Avrage q

20363.17

StD q

5585.11

StD q estim

2576.79

Inj.Name	Dist.	Average Rate	inj.Std	Weight	std	Part Coef	□□
M38C6	421.20	17946.2	2673.9	0.473	0.297	2.086%	0.557
M34C3	1440.76	55848.4	6413.1	0.213	0.095	4.003%	0.186
The Multiplier				285.10			

	inj-inj r		Pro-inj r
	M38C6	M34C3	M36ZB7
M38C6	1.00	0.92	0.79
M34C3	0.92	1.00	0.90

RankCentered ABMLR:

Time Window:

1/0/1900

1/0/1900

M36ZB7:

Correl

0.837

MSE

7.07E+02

Predict.

0.699

std

26.59

Lamda0

0.00

Lamda0 StD

--

Avrage q

88.31

StD q

48.04

StD q estim

42.09

Inj.Name	Dist.	Average Rate	inj.Std	Weight	std	Part Coef	□□
M38C6	421.20	87.6	39.8	-0.348	0.108	8.039%	0.303
M34C3	1440.76	79.8	40.2	1.168	0.092	57.506%	0.637
M06B3	1456.92	88.5	39.4	0.290	0.068	13.332%	0.124
The Multiplier				-0.90			

inj-inj r				Pro-inj r
	M38C6	M34C3	M06B3	M36ZB7
M38C6	1.00	0.77	0.60	0.56
M34C3	0.77	1.00	0.41	0.81
M06B3	0.60	0.41	1.00	0.42

This problem can be resolved by applying the modified ridge method for the constrained MLR where the intercept is forced to be zero all the time. The results of the ridge analysis are shown in Figure 68. The first plot shows trends of the estimated weight variances in which a minimum value does not occur. This behavior is mainly because the intercept is actually a function of the other weights in the system using the constrained MLR, so applying a constraint on intercept will affect the estimation of the other weights. The second plot represents the collinearity: M38C6 and M34C3 are highly correlated. The linear behavior of the MDC suggests that the variance inflation factor decreases steeply all the time. Thus, the expected minimum value in the weights variances will not be shown.

The best value of the modified ridge parameter should be selected so that it gives the most stable weights with minimum increase in the mean square error (minimum reduction in the R of the system). The best value is at C is equal to 0.4 as shown in Figure 68. At this value, the estimated weight for M38C6 is positive.

Figure 68: The modified ridge regression for the constrained MLR

Table 6 lists the constrained MLR weights before and after applying the modified ridge method. The results indicate that modified weights are the most stable with the minimum increase in mean square error.

TABLE 6: COMPARISON OF MLR WEIGHTS BEFORE AND AFTER THE MODIFIED RIDGE METHOD

Figure 69 and 70 show the original weights and the modified weights respectively where all modified weights are in the first quarter of the plot indicating the stability of the weights.

Figure 69: The original estimated weights of the rank of the data

Figure 70: The modified weights of the rank of the data

Magnus Summary:

Figure 71 depicts an image representation of estimated weights for all injector-producer pairs in the crest area. The scale of the weights is given in the color bar where by matching the color of each injector-producer pair in the plot with the corresponding color in the scale, the weight can be determined. Using the same representation for the confidence interval of estimated weights is

presented in Figure 72. For the producer M30ZA5, the supported injectors are M06B3 and M31ZC4 with weighting factors of 0.65 and 0.90 respectively. The STDs for the two injectors are ± 0.06 and ± 0.1 . These kinds of representation provide an easy way to read the statistical inferences of the applied method.

Similarly, Figure 73, and 74 show the same representation but with using the rank of the data. The rank weights tend to be larger and more stable than that of actual data because using the rank normalizes all wells to the same variance where the problem of insufficient variation of injectors is eliminated. Also, the rank data shows less collinearity among the injectors.

Figure 71: The image representation of estimated weights using actual data

Figure 72: The image representation of stability of estimated weights

Figure 73: The image representation of weights using the rank of the data

Figure 74: The image representation of stability of weights using the rank of the data

A graphical representation of injector-producer weights with Magnus field structure is shown in Figure 38. The size of the arrow issuing from an injector indicates the magnitude of the weight between this well and the producer to which it points. Similarly, Figure 39 shows the same representation using the data rank.

Figure 76: A graphical representation of rank data weights

The rank weights generally are larger than weights of actual data because the rank tends to give the same variance for all wells. Thus, the rank weights representation appears to give more additional information about the interaction among injectors and producers that can not be seen using the actual data.

The significance of estimated weights comes from our interpretation that they represent the connectivity of injector-producer pairs. Thus, if arrows presented in Figure 75 really capture this interpretation, they could be used to map different reservoir and geological features such as preferential permeability trends, transmissibility barriers, and reservoir heterogeneity. If this is the case in the Magnus field, Figure 75 or 76 can be used for further analysis to map these features.

M33C7 and M27B2 are the least supported producers in the crestal area, which could indicate low permeability trend or low preferential flow trend. This also indicates where additional injection wells could be located.

To examine the direction of the connectivity of Magnus field, a graphical representation, referred to as compass diagram, as shown in Figure 77 and Figure 78 can be used. In these figures, all arrows are referenced to the same point or origin with keeping their original direction. Clearly, the direction of the connectivity is mainly West to Northwest as indicated in diagrams.

Figure 77: The compass diagram for the actual weights

Figure 78: The compass diagram for the rank weights

To give more insight about the direction of flow in the field, a rose histogram was constructed where only the directions of the weights are used to construct the histogram. The rose histograms for actual data weights with bin no. of 5 and less than 20 are depicted in Figure 79 and 80 respectively. Similarly, Figure 81 and 82 are the rose histograms for rank data weights.

Figure 79 The rose histogram for the direction of actual data weights (# bins= 5)

Figure 80: The rose diagram for the direction of actual data weights (#bins< 20)

The main direction of field connectivity is W to N-W as shown in Figure 79 and 80. Similarly, the rose histogram using the direction of the rank data weights is the same as seen in Figure 81 and 82.

Figure 81: The rose histogram for the direction of rank data weights (#bins = 5)

Figure 82: The rose histogram for the direction of rank data weights (#bins < 20)

Inference connectivity for injector-producer pairs using the Magnus rate data should be validated to claim success for the method presented in this study. From a previous field review study for Magnus (Day et al., 1998), important observations regarding the waterflood system were mentioned. They can be summarized in the following

1. The producer M30ZA5 is supported by M06B3 and M31ZC4 where all completed in the same zone
2. No communication with injector F2 and producer M10A3
3. The producer M10A3 is supported by M06B3 and M31ZC4 where all completed in the same zone

4. The producer M24:B4 is supported by F2
5. M26:A4 is supported by M34C3

Figure 75 and 76 show that the producer M30ZA5 is mainly supported by M06B3 and M31ZC4. This matches the field observations. The proposed analysis doesn't suggest any connection between F2 and M10A3. The connection between M10A3 and injectors M06B3 and M31ZC4 could not be seen using the actual data. However, the connection appeared in the rank of the data, Figure 76. This is because some injectors are characterized by low variances where this is a typical problem with time series data. For the other observations, the proposed analysis provided relatively similar results.

CONCLUSIONS

Task 2

1. The role of the diffusivity filter in MLR analysis should be viewed differently for different production data reporting schemes. For data that were collected on a monthly basis, the diffusivity filter only slightly improves the R^2 of the correlation. For data collected on a daily basis, the role of the diffusivity filter increases significantly.
2. MLR weights and streamline-determined connectivities may disagree. This is because the streamlines show steady-state hydraulic connections between wells, while the MLR weights show how injection perturbations affect production.
3. The MLR weights are better interpreted as influence factors rather than measures of (hydraulic) connectivity because non-zero weights may obtain in cases where there is no direct hydraulic connection between wells.
4. Interpreting MLR weights as measures of interwell permeability is problematic because non-zero weights may arise for cases where there is no flow between an injector and producer.

The nonlinear correlation such as neural network can improve the R-squared of the correlation, compared to the MLR method. However, the result may be more rate dependent and more difficult to interpret.

Task 3

1. The MLR weights are not sensitive to the injection rate and the injection rate variability.
2. The MLR weights strongly depend on the condition of constant BHP. Changing the BHP of one producer during the analyzed time leads to the presence of both negative and the abnormally large positive weights. The R^2 of the correlation in these cases is usually small.
3. The MLR weights are strongly affected by the autocorrelation between injectors or between one injector and a group of injectors. Large positive weight may result in this case. The R^2 in this case may still be very high. A different way to estimate covariances may diminish this problem.
4. The average rate of the producer affects the interwell MLR weights. Wells with large average rate will have larger interwell MLR weights from all injectors
5. Under certain conditions, including steady state, constant producer BHP's, and relatively equal average production rates, the MLR weights may indicate the degree of reservoir connectivity between injector and producers.
6. When the producer average rates are not equal, the MLR weights may indicate the connectivity, but relatively for each producer. Comparing the MLR weights for different producer in this case may not be justified.

Task 4

Work in this area concerns prediction of oil rates. Oil prediction is based on a power law relationship between WOR and cumulative water injected. Tests on synthetic data are in progress.

Task 5

Analysis of the Magnus data set has begun and numerous issues are apparent. For example, there is substantial collinearity among injector rates. Use of ridge regression (developing under Task 3) may help. Further work is underway.

RECOMMENDATIONS

The following recommendations are made for further work:

1. Eliminate the autocorrelation of the injection rate or better use it in the analysis. One precaution is that the autocorrelation may not only exist between two injectors but also between a single injector and a group of injectors. Thus special treatment should be paid to all possible combinations of injection rate to detect the autocorrelation in screening the production data.
2. The violation of the constant working conditions of the producers is the strongest cause of abnormal MLR weights. Possible solutions to this problem could be: (1) select and use the time period during which no change of the well condition has occurred; and (2) incorporate the BHP change in the analysis. The first approach only requires the data screening before analysis, whereas the second approach requires further analytical investigation.
3. A simpler form of diffusivity filter is needed for analysis. The Ei-function based filter is too complex and nonlinear. The problem becomes unstable when we try to optimize the correlation with inclusion of the diffusivity coefficient for each producer-injector pair. A simple approximation of the pressure transient solution may help to eliminate this problem.
4. The analysis of MLR weights shows that they are affected by geometric well placement, magnitude of the injection rate autocorrelation, magnitude of the production rate, changing BHP of the producer, and reservoir heterogeneity. We can use simulation in a homogeneous reservoir to generate a set of data that incorporates the effects of well placement, injection rate, and the average production rate. By comparing the two sets of MLR weights (from simulated and real data) we may be able to eliminate these factors from the MLR weights.

REFERENCES

- Albertoni A. and Lake, L.: "Inferring Interwell Connectivity Only From Well-Rate Fluctuations in Waterfloods," SPEREE, Vol. 6, p. 6-16, 2003.
- Albertoni, A.: "Inferring interwell connectivity from well-rate fluctuations in waterfloods," M.S. Thesis, The University of Texas at Austin, May, 2002.
- Belsley, D., Kuh, E., and Welsch R., "Regression Diagnostics: Identifying Influential Data and Sources of Collinearity," Wiley, New York, 1980.
- Belsley, D., Kuh, E., and Welsch R., "Regression Diagnostics: Identifying Influential Data and Sources of Collinearity," Wiley, New York, 1980.
- Belsley, D.A., "Conditioning Diagnostics: Collinearity and Weak Data in Regression," Wiley, New York, 1991.
- Chilés, J-P., and Delfiner, P., Geostatistics, John Wiley, New York, 1999.
- Day, S., Griffin, T., and Martins, P., "Redevelopment and Management of the Magnus Field for Post-Plateau Production," SPE 49130, presented at the SPE Annual Technical Conference and Exhibition held in New Orleans, Louisiana, September 27-30, 1998.
- Dempster, A., Schatzoff, M., and Wermuth, N., "A Simulation Study of Alternatives to Ordinary Least Squares" Journal of the American Statistical Association, Volume 72 (March 1977), 77-91.
- Gibbons, D.I. and McDonald, G.C, "A Rational Interpretation of the Ridge Trace," Technometrics, 26 (November 1984), 339-346.
- Hoerl, Arthur, and Kennard, Robert, "Ridge Regression: Application to Non-Orthogonal Problems," Technometrics, 12 (February 1970b), 69-82.
- Hoerl, Arthur, and Kennard, Robert, "Ridge Regression: Biased Estimation for Non-Orthogonal Problems," Technometrics, 12(Febuary 1970a), 55-63.
- Muskat, M., The Flow of Homogeneous Fluids Through Porous Media, IHRDC Publications, Boston, 1982.
- Neter, J., "Applied Linear Statistical Models," Irwin, Chicago, 1996.
- Panda, M.N. and Chopra, A.K.: "An integrated approach to estimate well interactions," paper SPE 39563 presented at SPE India Oil and Gas Conference and Exhibition, New Delhi, India, February 17-19, 1998.
- Soeriawinata, T. and Kelkar, M.: "Reservoir management using production data," paper SPE 52224 presented at the SPE Mid-Continent Operation Symposium, Oklahoma City, Oklahoma, March 28-31, 1999.
- Yortsos, Y. C., Choi, Y, and Yang, Z., Analysis and Interpretation of Water/Oil Ratio in Waterfloods, SPEJ, Dec., 1999.