Feasibility of Gas Well Restimulation
In New York State

prepared for

Gas Technology Institute
Chicago, Illinois

and

New York State Energy Research and Development Authority
Albany, New York

Ronald J. MacDonald P.G.
Senior Petroleum Geologist
Project Leader

Charles M. Boyer II
Principal Consultant
Project Auditor

Joseph H. Frantz, Jr. P.E.
Operations Manager Eastern U.S.
Division Manager

July 2001

Holditch - Reservoir Technologies Consulting Services Schlumberger

Table of Contents

List of Tables

List of Figures

1 Executive Summary

Schlumberger Holditch – Reservoir Technologies Consulting Services (H-RT), with subcontract support provided by Advanced Resources International, Inc. (ARI), and Intelligent Solutions, Inc. (IS), performed a study to develop a restimulation candidate selection process for various wells operated by Belden & Blake Corporation located in Chautauqua County, New York State.

One hundred fifty nine wells producing from the Lower Silurian Medina Group's Grimsby and Whirlpool sandstones located in Chautauqua County, NY have been evaluated using three different methods to quantify the restimulation potential of these wells. H-RT performed a Moving Domain Analysis™ (MDA™), IS applied a virtual intelligence neural network, and ARI utilized type curve analysis techniques. These differing approaches and their results are discussed throughout this report.

The primary objective of this study is to evaluate production data and determine which, if any wells have incremental production potential via a rework or additional lease compression. Most of the wells in this study have approximately twenty-three years of production history.

Utilizing our MDA™ and recovery factor calculations, we have determined that the following wells are good candidates for recompletion, **Table 1**. These candidate wells have 5-year cumulative production values significantly lower than expected based upon the median production value of the wells within its domain. This list is sorted by the descending difference in production volume between a well's actual 5-year cum versus the median of 5-year cumulative production for all wells within its domain. Wells with higher differences should have higher incremental recoverable reserves through restimulation. Recovery factors were also considered based upon EUR's from decline curves, 80-acre maximum drainage, and net pay (height, porosity, and extrapolated Sw).

Table 1 – H-RT list of restimulation candidates

Well_Name		5 yr Cumulative Gas (Mscf)	Domain Median (Mscf)	Difference (Mscf)	Recovery Factor (%)
MOTRYNCZUK, PAUL	#382	44,580	89,907	45,327	23%
SHEPARD, GEORGE	#297	20,063	64,631	44,568	40%
SUPPO, PETER	#073	8,613	40,065	31,452	20%
STARR, HERB	#151	8,106	35,658	27,552	2%
CRANDALL, RICHARD	#327	15,300	41,900	26,600	6%
DEAN, LUTHER	#017	6,242	28,133	21,891	4%
RAYNOR, WARD	#386	10,515	31,959	21,444	
BROWN, CHARLES	#028	37,280	57,661	20,381	92%
BERGER, CARL	#288	6,694	23,222	16,528	6%
LONGHOUSE, ARTHUR	#064	11,892	28,250	16,358	25%
BEARDSLEY, JOHN	#134	12,446	25,525	13,079	9%
JOHNSTON, CHARLES	#304	17,100	29,362	12,262	7%
BARKER, JAMES	#338	5,280	16,392	11,112	86%
DARBY, LEON	#336	7,174	17,892	10,718	14%
YONKERS, FRANK	#137	12,680	22,448	9,768	17%
HOWARD, VELMA	#320	3,959	11,901	7,942	27%
WELLMAN, DONALD	#120	5,115	11,023	5,908	5%

Intelligent Solution’s list of recompletion candidates shown below in **Table 2**.

Table 2 – IS list of top 25 restimulation candidates

Final Ranking				
		Ranking		
Rank	Well Name	Last Month Rate	Reservoir Quality	Three-Input System
1	Suess, George #430	1	2	4
2	Wills, William #325	13	3	2
3	Winchell, Francis #303	16	5	3
4	Palmer, Lonnie #074	17	8	1
5	Martin, David #300	5	1	21
6	Augustinians of the Assum #103	6	14	14
7	McLarney, Jane #102	19	13	5
8	Suppo, Jeter #073	15	18	9
9	Scholl, Mary A. #115	14	21	12
10	Smith, Warren #069	23	11	13
11	Colt, Alvin #024	11	20	22
12	Smith, M. #1 #245		6	8
13	Cranston, Claude #311		4	11
14	Village of Brocton #299		17	7
15	Dubois, Florence #079		16	10
16	Cornell, Gordon #423		12	19
17	Chilcott, Eugene #365	20		18
18	Barber #2	24		15
19	Van Dette, Albert #356		22	17
20	Bemus, Cecile #214		23	24
21	Crandall, Richard #326		25	25
22	Straight, Frank #140			16
23	Miller, Morris #433			20
24	Josephson, Walfred #329			21
25	Zook, Marvin #276			23

ARI's list of candidate wells is shown below in **Table 3.**

Table 3 – ARI list of restimulation candidates

a) Rankings by Incremental Recovery Due to Restimulation

			TC Results					Restim	Inc.
Well No.	Well Name	Formation	k (md)	Xf (ft)	A (acres)	5 Yr Cum (Bcf)	P 5/01 (psia)	5 Yr Cum (Bcf)	5 Yr Cum (MMcf)
16	Bemus, Cecile, #214	Comingled	0.150	-	74	0.010	380	0.021	10.5
115	Raynor, Ward, #323	Comingled	0.056	5	62	0.013	518	0.022	9.4
150	Wellman, Donald #120	Comingled	0.007	-	21	0.003	716	0.011	7.6
146	Van Dette, Albert #356	Comingled	0.070	10	110	0.012	579	0.020	7.5
117	Reno, Norman #277	Comingled	0.055	7	131	0.011	467	0.019	7.3
87	Marrano, Anthony #389	Comingled	0.023	25	73	0.007	570	0.015	7.2
110	Powell, Irving #077	Comingled	0.066	15	137	0.009	406	0.017	7.2
63	Furmanek, Aloysious #144	Comingled	0.090	9	293	0.013	426	0.019	6.1
34	Chilcott, Eugene #365	Comingled	0.325	20	104	0.022	283	0.028	6.0
149	Webster Castle Inn #015	Grimsby	0.140	15	104	0.015	365	0.021	6.0

b) Rankings by Incremental Recovery Due to Added Lease Compression

			TC Results					Restim	Inc.
Well No.	Well Name	Formation	k (md)	Xf (ft)	A (acres)	5 Yr Cum (Bcf)	P 5/01 (psia)	5 Yr Cum (Bcf)	5 Yr Cum (MMcf)
154	Wills, William #325	Comingled	0.150	15	62	0.002	366	0.028	25.4
97	Miller, Morris #433	Comingled	0.210	35	109	0.005	279	0.027	21.3
94	Mikula, Joseph #152	Grimsby	0.075	25	259	0.012	471	0.027	15.0
140	Sutton, David #315	Comingled	0.140	25	268	0.017	440	0.032	14.7
112	Przybylski, Leonard #113	Comingled	0.110	15	255	0.009	390	0.021	12.5
75	Josephson, Walfred #329	Grimsby	0.090	15	109	0.005	423	0.017	11.3
145	Van Dette, Albert #339	Comingled	0.065	13	77	0.008	462	0.019	10.7
153	Wilkens, Roy #343	Comingled	0.100	17	133	0.013	407	0.021	7.8
52	Dennison, Wilber #062	Comingled	0.090	20	133	0.008	359	0.016	7.7
82	Lanford, C. #240	Comingled	0.105	14	43	0.004	303	0.011	7.4

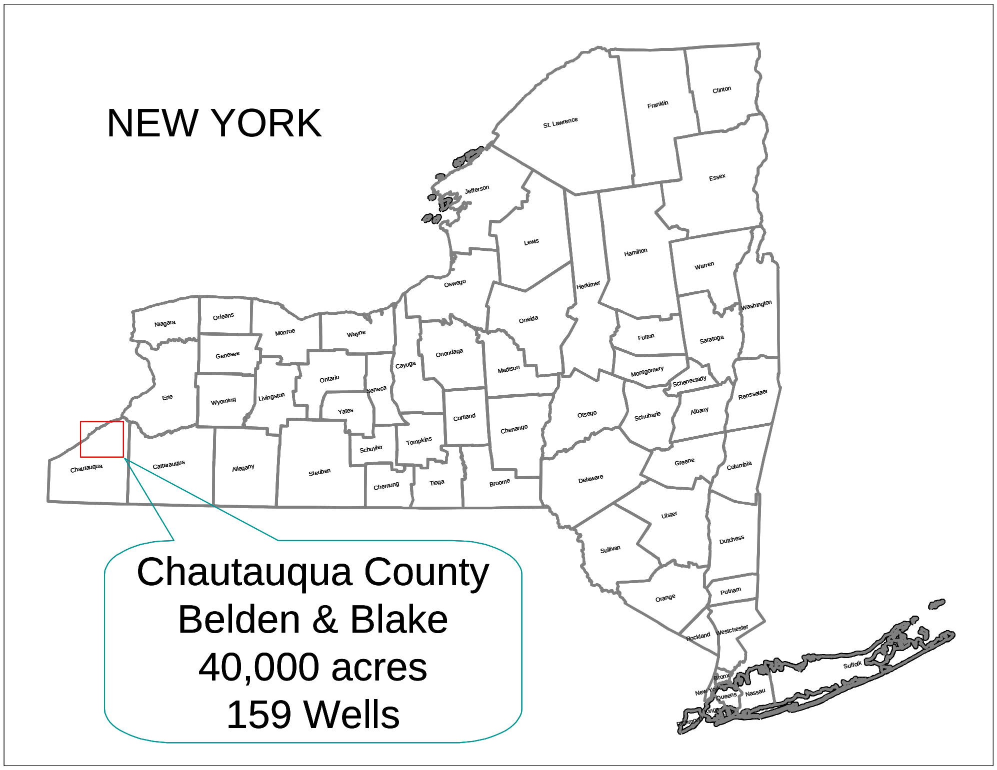
2 Introduction

Historical development of the study area and its relatively shallow reservoirs began in the late 1880's and the wells were produced naturally, or with only minimal nitroglycerine stimulation. Production is from the lower Silurian Medina group's Grimsby and the Whirlpool formations, which consists of interbedded sandstones, siltstones, and shales. The Medina group is further subdivided into the Grimsby sandstone, Cabot Head shale, and Whirlpool sandstone that unconformably overlies the Upper Ordovician Queenston shale. These formations provide a stratigraphic play with production primarily related to porosity and permeability variations within the reservoir. Local variations within the structural setting may influence production in some areas, but to a considerably lesser extent. All of the wells were stimulated similarly (e.g. proppant amounts, nitrogen volumes, fluid rate, etc.) in the Lower Silurian Medina group Grimsby and Whirlpool sandstones, with perforations ranging from 2,300 to 3,400 feet deep. Typical treatments utilized gelled water with nitrogen assist, averaging 623 barrels of fluid, 5,470 pounds of proppant, and 18 perforations.

In 1996, the Gas Research Institute (GRI) [currently the Gas Technology Institute (GTI)] authorized an evaluation to study potential benefits from the restimulation of existing natural gas wells. This analysis estimated that over one trillion cubic feet of natural gas reserves and up to $500 million in benefits could potentially be derived from successful recompletion programs conducted in the tight gas reservoirs of South Texas, Mid-Continent, and Rocky Mountain regions of the United States. Case studies have subsequently shown that restimulation benefits may also be significant in low permeability unconventional reservoirs similar to the Grimsby and Whirlpool sandstones in New York State.

Belden & Blake Corporation (B&B) worked with Advanced Resources International, Inc. (ARI) in 1998 to determine the magnitude of reserve and production enhancement opportunities that existed in their wells situated in this area of New York, **Fig. 1**. The New York State Energy Research and Development Authority (NYSERDA) and GTI jointly sponsored a feasibility study modeled after the 1996 GRI evaluation. Methodology used included production data analysis, artificial neural networks, type curve construction, and reservoir modeling. These tools have helped to identify restimulation candidates for Belden & Blake's Chautauqua County, NY Medina Group field. Our evaluation applies and enhances the restimulation feasibility methods developed by GRI in 1996 to these formations, and was requested jointly by NYSERDA, GTI, and B&B.

Figure 1 – Study area location map



3 Conclusions

Based on the worked performed by H-RT and our subcontractors, we conclude that:

1) A significant number of recompletion candidates are available for Belden & Blake's review.

2) Varying methodologies of identifying recompletion candidates often identify a core group of the same wells, thus providing added confidence in the selection of these wells for recompletion.

4 Recommendations

Based on the results of this study, H-RT makes the following recommendations:

1) Belden & Blake should review the three lists of recompletion candidates, determine what restimulation options are available, and estimate their associated costs.

2) Perform economic analyses on the selected recompletion candidates based upon expected stimulation costs, anticipated production increase and additional operating costs, if any.

3) Upon finalization and execution of an actual recompletion, estimate future production rates and reserves.

5 Discussion

Work performed on this project was undertaken by H-RT as the lead contractor, with IS and ARI providing sub-contract support. Discussion of our approach and result obtained are presented below. The approach and results of IS and ARI are presented in Appendices 1 and 2 to this report.

5.1 Database Construction

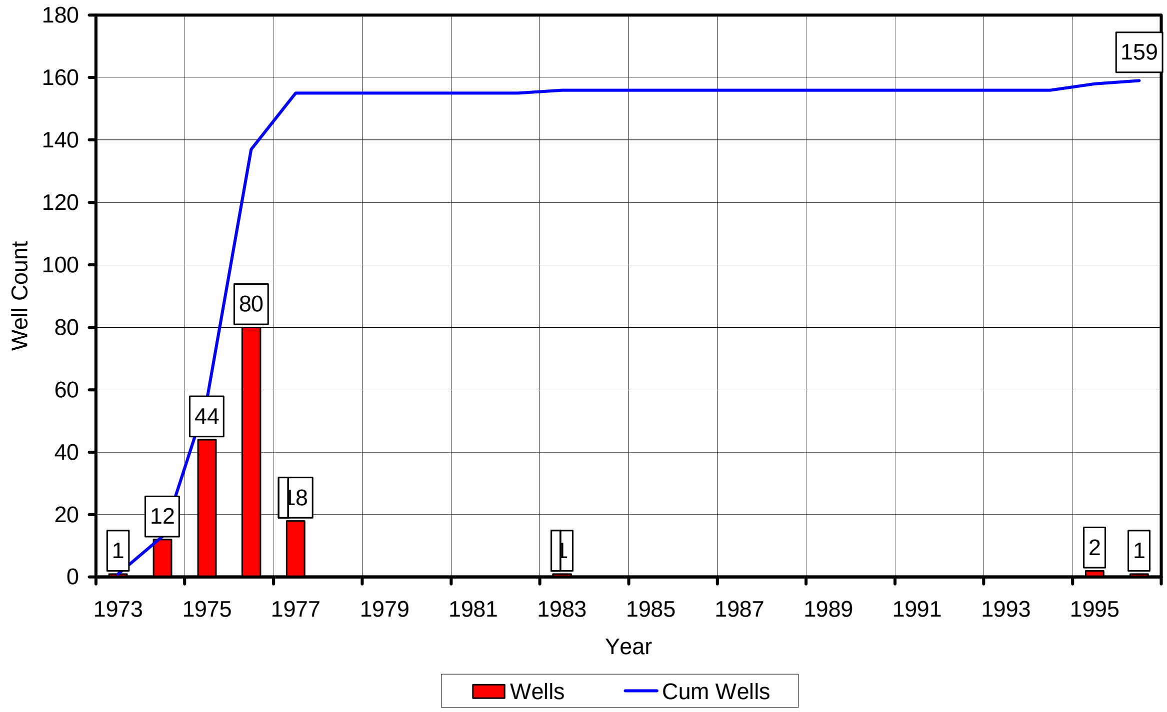
In 2000, Belden & Blake provided ARI with production data, completion methods, and well information. ARI estimated net pay thickness, created a preliminary database, and presented it to H-RT for use in the study. We organized and uploaded this information into various Microsoft ACCESS™ databases and EXCEL™ spreadsheets, and then reviewed and quality checked it for integrity. **Table 4** shows examples of information used for this study.

Table 4 – Examples of data used for this study

Well Permit Number	Latitude and Longitude
KB Elevation	Well Number
Well Name	Operator
Spud Date	Completion Date
Date of First Production	Total Depth
Completed Intervals and Names	Net Pay
Stimulation Type (by stage and total)	Cumulative Production
Average Sand (by stage and total)	Best 12-Month Gas Production
Injection Rate by Stage	Shut-In Pressure
Production Data by Month	Number of Stimulation Stages
Best Production Month	Average Treating Pressure by Stage
Average Flowing Pressure (best 6- and 12-months)	Average Sand Concentration (by stage and total)
	Sand and Volume of Fluid Used in Stimulation (by stage and total)

Fig. 2 shows that initial drilling dates for these wells are between 1973 and 1996, though 154 wells have online dates between 1974 and 1977 (12 wells in 1974, 44 wells in 1975, 80 wells in 1976, and 18 wells in 1977).

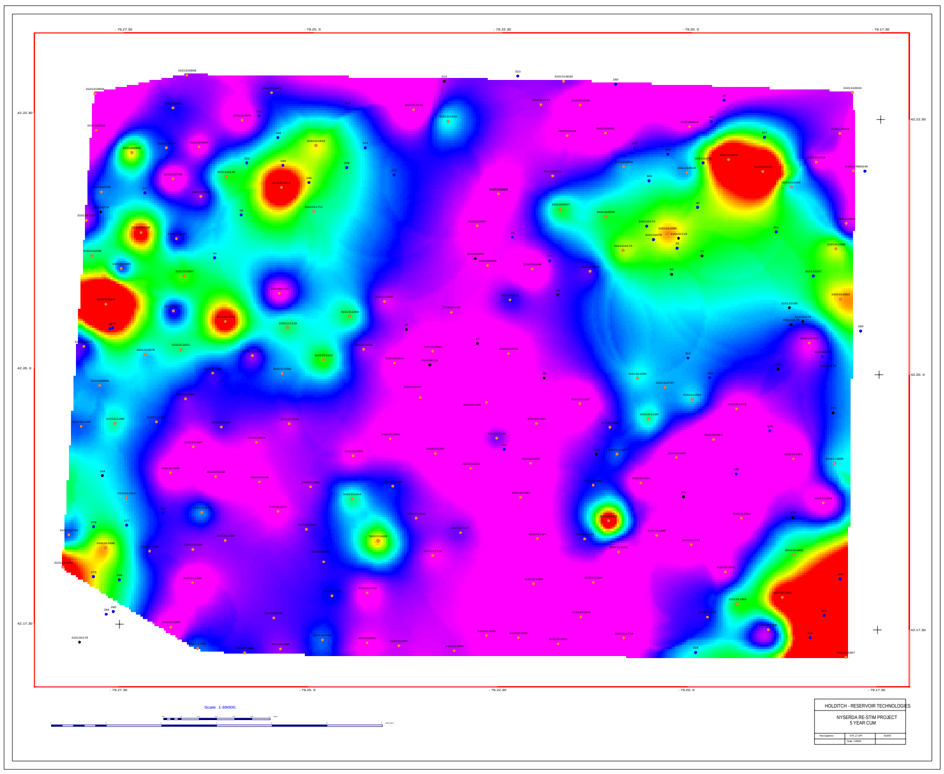
Figure 2 – Study wells versus year drilled



5.2 Production Indicator and Recovery Factors

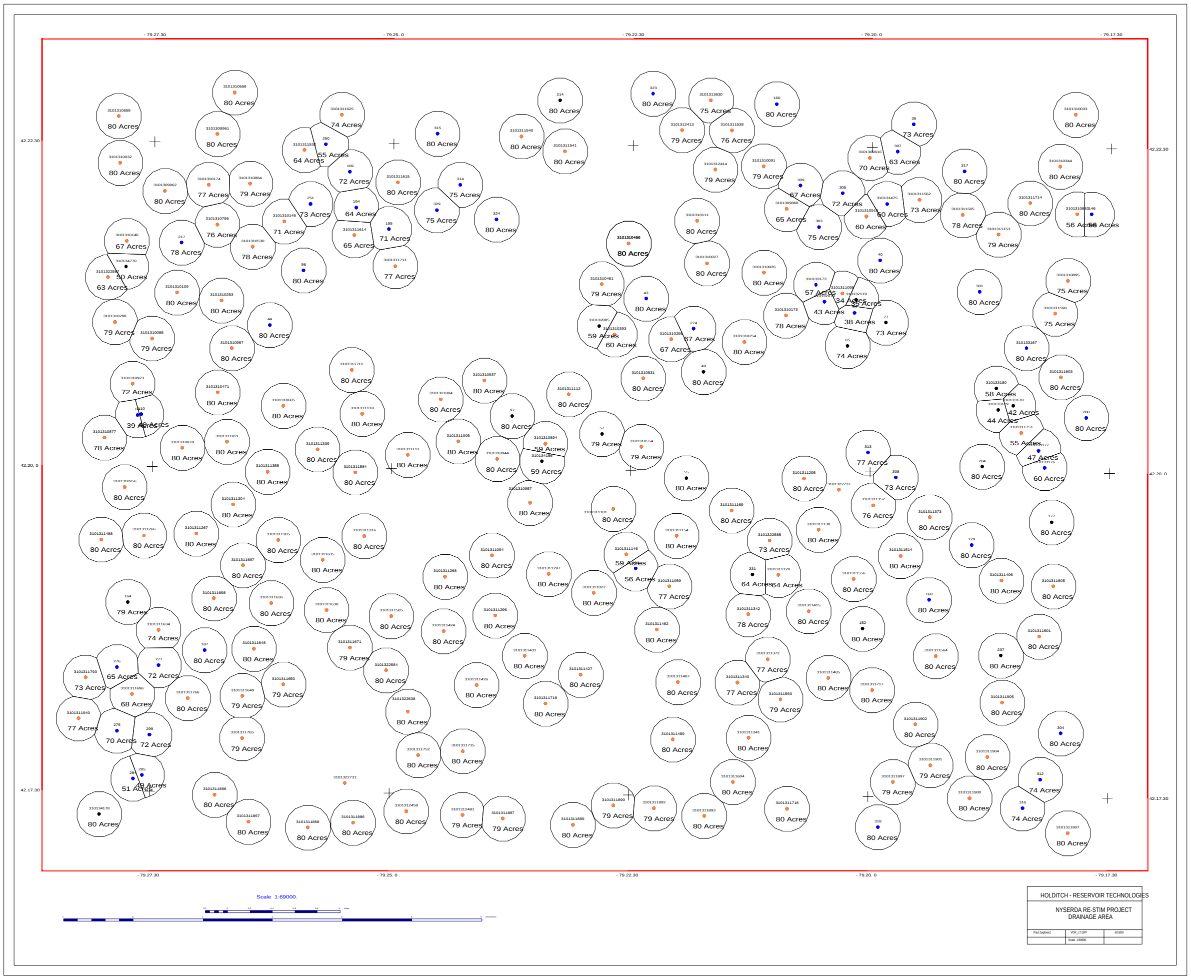
To investigate the usefulness of different primary production indicators in identifying potentially recoverable incremental reserves, H-RT evaluated 5-year cumulative gas, 10-year cumulative gas, and highest twelve-month (q_{12}) volumes relative to dates of first production (DOFP). Early in our study it became apparent that the 5-year cumulative was the most helpful in recognizing restimulation candidates with the use of MDA**.** **Fig. 3** is a color-filled contour map of these 5-year values. Note that the areas in red are regions of higher volumes.

Figure 3 – Color-filled contour map of 5-year cumulative gas production



For recovery efficiency calculations, we used Belden & Blake's EUR as the best measure of ultimate well performance, which assumed an 80-acre maximum drainage area per well, and net sand height, porosity, and extrapolated Sw values. From these assumptions, we determined that 17,312 acres are being drained with an original gas-in-place of 50 Bscf. Of this total gas, 15.6 Bscf is recoverable resulting in a recovery efficiency of 31.3%. **Fig. 4** is a map of the study well locations with an 80-acre maximum spacing ring surrounding each well.

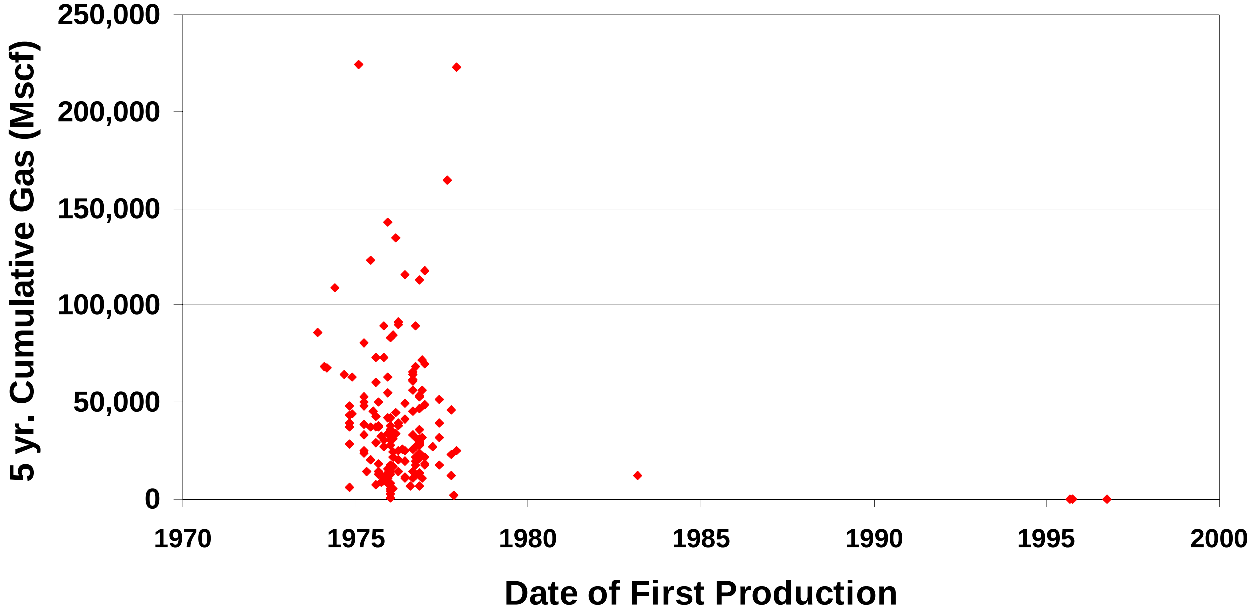
Figure 4 – Study well locations and 80-acre maximum drainage boundaries



Moving Domain Analysis™ (MDA™) is a technique to rapidly process production data and completion information and aids in determining reserve distributions and infill potential. It essentially analyzes a mosaic of overlapping localized studies (domains) and requires only minimal data such as latitude/longitude of each well, and monthly production information. Each of the 159 wells has been compared with offset wells in its vicinity based upon the 5-year cumulative production indicator. We have statistically determined that this is a reliable short-term indicator of long-term well performance and is influenced by reservoir quality, completion, and operating conditions.

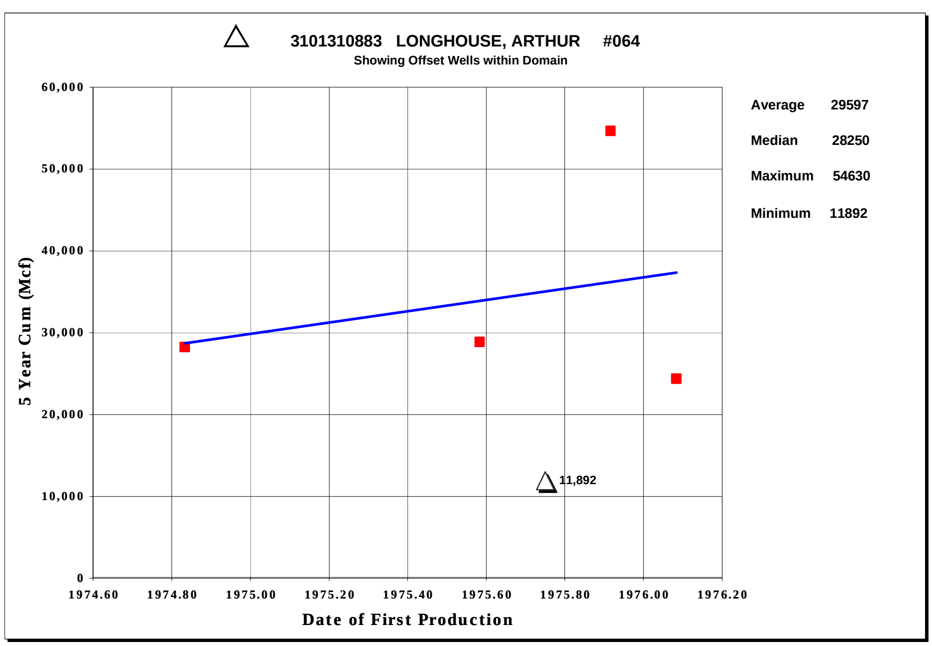
Fig. 5 is a chart of date of first production versus 5-year cumulative production for the wells in the study, and shows that these values range from zero to about 225 million standard cubic feet (MMscf), though most are 75 MMscf or less.

Figure 5 – Date of first production versus 5-year cumulative production



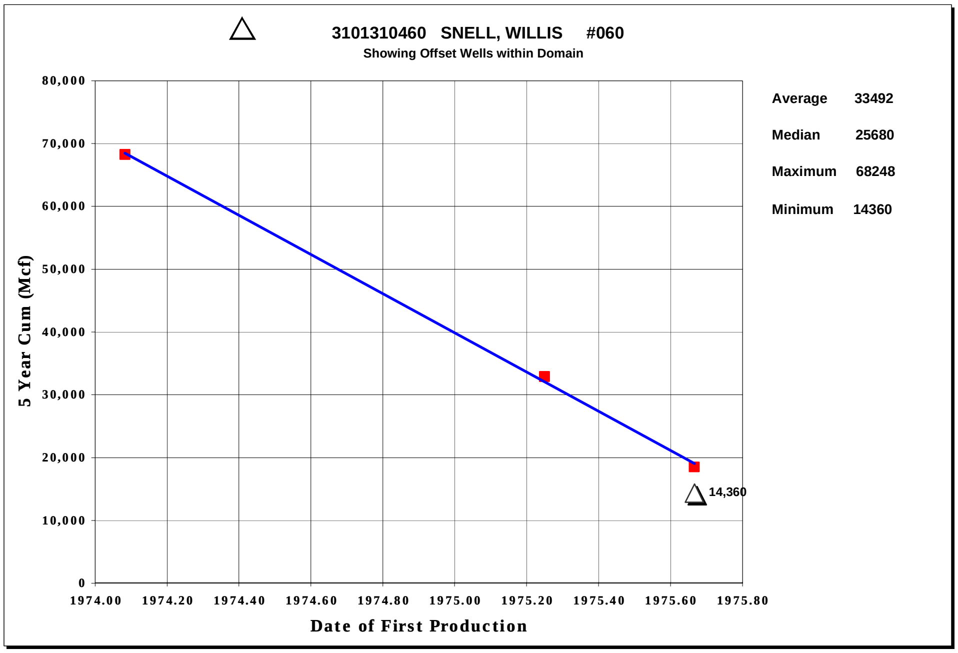
During the MDA™ process, we also plotted 5-year cumulative on a chart's y-axis, versus dates of first production on an x-axis. We did this for each well (target well) and all offsets within its domain. This enabled us to identify target wells that were underperforming relative to their offsets. **Fig. 6** is an example of a well producing at a rate lower than we would expect based upon wells in its domain, and is potentially a candidate for recompletion due to underperformance.

Figure 6 – Example of underperforming well



Depletion can also be recognized with the use of MDA and viewing 5-year cumulative figures versus DOFP. An indication that depletion may be occurring within a domain is a negative slope of a best-fit linear regression trend line of this data. **Fig. 7** shows depletion over time. Note that 5-year values become lower as DOFP's move forward in time.

Figure 7 – Example of depletion based on 5-year cumulative production



5.3 Stimulation Candidates

H-RT used a Moving Domain Analysis™ to calculate each wells expected 5-year cumulative based upon the wells within its domain versus its actual 5-year production history. Wells below the norm may be restimulation candidates. **Table 5** is a list showing our top candidates.

Table 5 – H-RT list of restimulation candidates

Well_Name	5 yr Cumulative Gas (Mscf)	Domain Median (Mscf)	Difference (Mscf)	Recovery Factor (%)
MOTRYNCZUK, PAUL #382	44,580	89,907	45,327	23%
SHEPARD, GEORGE #297	20,063	64,631	44,568	40%
SUPPO, PETER #073	8,613	40,065	31,452	20%
STARR, HERB #151	8,106	35,658	27,552	2%
CRANDALL, RICHARD #327	15,300	41,900	26,600	6%
DEAN, LUTHER #017	6,242	28,133	21,891	4%
RAYNOR, WARD #386	10,515	31,959	21,444	
BROWN, CHARLES #028	37,280	57,661	20,381	92%
BERGER, CARL #288	6,694	23,222	16,528	6%
LONGHOUSE, ARTHUR #064	11,892	28,250	16,358	25%
BEARDSLEY, JOHN #134	12,446	25,525	13,079	9%
JOHNSTON, CHARLES #304	17,100	29,362	12,262	7%
BARKER, JAMES #338	5,280	16,392	11,112	86%
DARBY, LEON #336	7,174	17,892	10,718	14%
YONKERS, FRANK #137	12,680	22,448	9,768	17%
HOWARD, VELMA #320	3,959	11,901	7,942	27%
WELLMAN, DONALD #120	5,115	11,023	5,908	5%

IS applied a virtual intelligence neural network to identify wells with restimulation potential utilizing neural nets, genetic algorithms, engineering expertise, and fuzzy logic. Parameters such as 5-year cumulative, reservoir quality index, last production rate, were processed.

The reservoir quality index is equal to:

$$\frac{(\text{net porosity} * \text{net pay thickness})_{Medina} + 2(\text{net porosity} * \text{net pay thickness})_{Whirlpool}}{\text{Number of perforations}}$$

The Whirlpool has been given more weight than the Medina in this equation. IS's recompletion candidates are shown in **Table 6**. Further detail is presented in their attached final report.

Table 6 – IS list of top 25 restimulation candidates

Final Ranking				
		Ranking		
Rank	Well Name	Last Month Rate	Reservoir Quality	Three-Input System
1	Suess, George #430	1	2	4
2	Wills, William #325	13	3	2
3	Winchell, Francis #303	16	5	3
4	Palmer, Lonnie #074	17	8	1
5	Martin, David #300	5	1	21
6	Augustinians of the Assum #103	6	14	14
7	McLarney, Jane #102	19	13	5
8	Suppo, Jeter #073	15	18	9
9	Scholl, Mary A. #115	14	21	12
10	Smith, Warren #069	23	11	13
11	Colt, Alvin #024	11	20	22
12	Smith, M. #1 #245		6	8
13	Cranston, Claude #311		4	11
14	Village of Brocton #299		17	7
15	Dubois, Florence #079		16	10
16	Cornell, Gordon #423		12	19
17	Chilcott, Eugene #365	20		18
18	Barber #2	24		15
19	Van Dette, Albert #356		22	17
20	Bemus, Cecile #214		23	24
21	Crandall, Richard #326		25	25
22	Straight, Frank #140			16
23	Miller, Morris #433			20
24	Josephson, Walfred #329			21
25	Zook, Marvin #276			23

Decline curve analysis was used by ARI to recognize underperforming wells and its list is presented in **Table 7**. ARI used type curve matching to distinguish wells that potentially offer incremental reserves. Since none of the study wells had been recompleted, production profiles of wells that previously had artificial lift systems installed were evaluated and used as models for representative type curves.

Table 7 – ARI list of restimulation candidates

a) Rankings by Incremental Recovery Due to Restimulation

			TC Results					Restim	Inc.
Well No.	Well Name	Formation	k (md)	Xf (ft)	A (acres)	5 Yr Cum (Bcf)	P 5/01 (psia)	5 Yr Cum (Bcf)	5 Yr Cum (MMcf)
16	Bemus, Cecile, #214	Comingled	0.150	-	74	0.010	380	0.021	10.5
115	Raynor, Ward, #323	Comingled	0.056	5	62	0.013	518	0.022	9.4
150	Wellman, Donald #120	Comingled	0.007	-	21	0.003	716	0.011	7.6
146	Van Dette, Albert #356	Comingled	0.070	10	110	0.012	579	0.020	7.5
117	Reno, Norman #277	Comingled	0.055	7	131	0.011	467	0.019	7.3
87	Marrano, Anthony #389	Comingled	0.023	25	73	0.007	570	0.015	7.2
110	Powell, Irving #077	Comingled	0.066	15	137	0.009	406	0.017	7.2
63	Furmanek, Aloysious #144	Comingled	0.090	9	293	0.013	426	0.019	6.1
34	Chilcott, Eugene #365	Comingled	0.325	20	104	0.022	283	0.028	6.0
149	Webster Castle Inn #015	Grimsby	0.140	15	104	0.015	365	0.021	6.0

b) Rankings by Incremental Recovery Due to Added Lease Compression

			TC Results					Restim	Inc.
Well No.	Well Name	Formation	k (md)	Xf (ft)	A (acres)	5 Yr Cum (Bcf)	P 5/01 (psia)	5 Yr Cum (Bcf)	5 Yr Cum (MMcf)
154	Wills, William #325	Comingled	0.150	15	62	0.002	366	0.028	25.4
97	Miller, Morris #433	Comingled	0.210	35	109	0.005	279	0.027	21.3
94	Mikula, Joseph #152	Grimsby	0.075	25	259	0.012	471	0.027	15.0
140	Sutton, David #315	Comingled	0.140	25	268	0.017	440	0.032	14.7
112	Przybylski, Leonard #113	Comingled	0.110	15	255	0.009	390	0.021	12.5
75	Josephson, Walfred #329	Grimsby	0.090	15	109	0.005	423	0.017	11.3
145	Van Dette, Albert #339	Comingled	0.065	13	77	0.008	462	0.019	10.7
153	Wilkens, Roy #343	Comingled	0.100	17	133	0.013	407	0.021	7.8
52	Dennison, Wilber #062	Comingled	0.090	20	133	0.008	359	0.016	7.7
82	Lanford, C. #240	Comingled	0.105	14	43	0.004	303	0.011	7.4

6 References

Frantz, J. H., Jr., Spivey, J. P., Voneiff, G. W., and Jacot, H., "Practical Production Data Analysis for the Appalachian Basin," paper SPE 37347 presented at the 1996 Eastern Regional Meeting, Columbus, Ohio, October 23-25.

Hopkins, C. H., Frantz, J. H., Jr., Tatum, C. L., and Hill, D. G., "Screening Restimulation Candidates in the Antrim Shale," paper SPE 29172 presented at the 1994 Eastern Regional Conference and Exhibition, Charleston, West Virginia, November 8-10.

7 Appendices

Appendix 1

Final Report

Feasibility of Gas Well Restimulation

In New York State

Intelligent Solutions, Inc.

FINAL REPORT

Feasibility of Gas Well Restimulation in New York State

Prepared by:
Shahab D. Mohaghegh, Ph.D.
Intelligent Solutions, Inc.
Morgantown, West Virginia

Prepared For:
S. A. Holditch Reservoir Technology
Pittsburgh, Pennsylvania

Project Funded by:
NYSERDA/GRI Advanced Drilling and Production Technology Partnership Program

September 2000

Table of Contents

BACKGROUND

In 1996, GRI commissioned an evaluation of the benefits that could be realized from restimulation of existing gas wells. This analysis suggested that over 1 Tcf of reserves and up to $500 million in benefits could be derived from successful restimulation programs in the tight gas reservoirs of South Texas, Mid-continent, and Rocky Mountain Regions of the United States. Additionally, case studies have shown that restimulation benefits are greatest in low permeability, unconventional resources such as the tight Medina and Whirlpool sands of southwest New York State. This project is designed to apply the methodology employed in GRI restimulation projects being performed in the Frontier sand of Wyoming, Mesaverde sand of Colorado, Wilox/Lobo sand of South Texas, and Cotton Valley sands of East Texas to the Medina/Whirlpool sands of western New York.

This report summarizes the efforts by Intelligent Solutions, Inc. to apply the intelligent systems approach to the Medina/Whirlpool sands of western New York, and consequently select and rank the restimulation candidates in this formation.

The report will start by a brief explanation of the methodology developed by the Intelligent Solutions, Inc. for restimulation candidate selection, and then will cover the application of this methodology to the Medina/Whirlpool sands of western New York.

METHODOLOGY

This section summarizes the intelligent systems approach developed by Intelligent Solutions Inc. for the restimulation candidate selection. Figure 1 is a flowchart that represents the process developed for this methodology. This figure is used as a road map to explain the intelligent systems methodology.

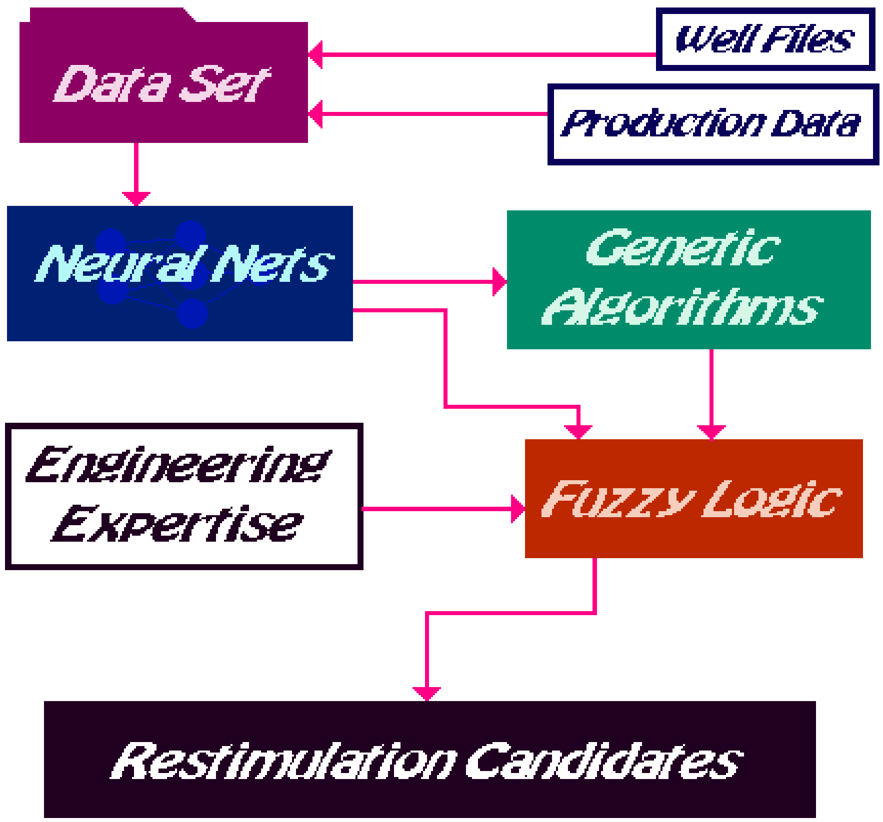


Figure 1. Virtual intelligence approach for restimulation candidate selection.

Three steps are involved for selecting restimulation candidates using virtual intelligence techniques. Before starting the process, however, a data set that includes all the relevant available data for the formation being studied should be compiled. The data usually includes four major categories. First the general information about each well. This can include the location coordinates of the wells, an indication of the depth, and the date the well was put into production. Second category addresses reservoir quality. Information in this category may include net pay, porosity, saturation, permeability (if available) and any kind of pressure indicator. If the reservoir is consists of several layers, it helps to have the above information on a per layer (zone) basis. The third category is the stimulation-related data. This data may include information on the type and amount of proppant that has been used, type and amount of fracturing fluid, perforation density and information on the number of zones and layers involved in each frac job in the event of a layered reservoir. The last category includes mainly production data. This data is used to calculate five-year cumulative production or EUR for each well. The five-year cumulative production or EUR would be the production indicator to which the above parameters are correlated.

The data that needs to be compiled can usually be found in well files and publicly available databases. Once the data set has been compiled, the next step is to apply the virtual intelligence approached. The first step of this process calls for the use of neural networks.

STEP 1: Neural Network Analysis

Neural networks are used to build a representative model of well performance in the particular reservoir being studied. The data is used as input-output pair to train the neural network. The first three categories – well information, reservoir quality and stimulation related data - are used as input and are coupled with the fourth category – production data – as output.

Since it is impractical to model such a complex process using the conventional modeling techniques – mathematical modeling – neural networks can provide a valuable insight into the intricacies of interaction of the formation with the hydraulic fracturing designs and implementations. Once a reasonably accurate and representative neuro-model of the stimulation processes has been completed for the formation under study, more analysis can be performed. These analyses may include the use of the model in order to answer many “what if” questions that may rise. Furthermore, the model can be used to identify the best and worst completion/stimulation practices in the field.

The ideal situation for using a neural model in restimulation candidate selection is when some restimulation jobs have been performed and some known results exists. These results can be used during the model building process to calibrate the accuracy of the model. In the case that such restimulation jobs do not exist, a surrogate method should be used in an attempt to deduce the maximum information from the available data.

This brings us to the second step of the methodology. Now that we have a representative model of the stimulation process for the formation being studied, how can we use it to identify in which wells restimulation potential exist? Once the neural model has identified the best practices, each hydraulic fracture treatment can be tested to examine if the stimulation job that had been performed was the best design for that particular well at the time it was implemented. The degree of departure from the optimum design is translated to the missed production opportunity, which in turn can be used as a proxy for restimulation potential.

STEP 2: Genetic Optimization

Genetic algorithms are used to perform this section of the analysis. The neural networks developed in the first step are used as the "fitness function" for the genetic algorithm routines.

The process of identifying the missed production opportunities - because of less than optimum hydraulic fracturing treatments - is as follows. The neuro-model developed in the first section of the methodology is able to provide an output (e.g., five-year cum.) based on the input to the network, namely, stimulation design, well information and reservoir quality for each particular well. Among these input categories only stimulation design parameters are controllable. Well information and reservoir quality is obviously beyond the engineer's control. Therefore, the genetic algorithm is set to search among all the possible combinations of the stimulation parameters and identify the most optimum combination. The most optimum combination of stimulation parameters are defined as the combination that for any particular well (based on the well information and reservoir quality) provides the highest output (five-year cumulative production - 5YCum). The difference between the 5YCum from the optimum stimulation treatment and the actual 5YCum produced by the well is interpreted as the production potential that may be recovered by restimulation of that well. This incremental production is the surrogate variable that was mentioned in the previous section.

This analysis concludes the second part of the methodology. Furthermore, the candidate selection process is not entirely based on the outcome of the genetic algorithms.

STEP 3: Fuzzy Decision Support System

The third and final step of the restimulation candidate selection methodology incorporates a fuzzy decision support system. This fuzzy expert system uses the information provided by the neural networks and genetic algorithms. The expert system then augments those findings with information that can be gathered from the expert engineers who have worked on that particular field for many years in order to select the best restimulation candidates. Keep in mind that the information provided to the fuzzy expert system may be different from formation to formation and from company to company. This part of the methodology provides the means to capture, maintain and use some valuable expertise that will remain in the company even when engineers are transferred to other sections of the company where their expertise is no longer readily available. The fuzzy expert system is capable of incorporating natural language to process information. This capability provides maximum efficiency in using the imprecise information in less than certain situations. A typical rule in the fuzzy expert system that will help engineers in ranking the restimulation candidates can be expressed as follows:

If the well shows a ***high*** potential for an increase in 5YCum, **And** has a ***moderate*** pressure, **And** has a ***low*** proppant volume for the net pay completed, **Then** this well is a ***good*** *candidate* for restimulation.

A truth-value is associated with every rule in the fuzzy expert system developed for this methodology. The process of making decisions using fuzzy subsets using the parameters and relative functional truth-values as rules provides the means of using approximate reasoning in making decisions. This process is known to be one the most robust methods in developing high-end expert systems in many industries.

RESULTS AND DISCUSSIONS

S. A. Holditch Reservoir Technology was in charge of gathering, compiling and organizing the available data into a database format. Once the data was received from S. A. Holditch Reservoir Technology some preliminary data processing was carried out in order to develop a better understanding of the nature of the available data. This part of the data processing mainly included conventional statistical analysis. Also some preprocessing of the data was required to prepare it for use by the remaining of the analysis.

Table 1 shows the data extracted from the database to be used for the analysis in this study.

Category	Parameter	Definition
General	API_Number	Well API number
	WELL NAME	Well name
Location	X	x coordinate of the well
	Y	y coordinate of the well
	UG_Top	Depth to the top of Upper Grimsby
	WP_Top	Depth to the top of Whirlpool
Reservoir Quality	Net_PhiH_TG	Net porosity, net pay product for all of Grimsby
	Net_PhiH_WP	Net porosity, net pay product for all of Whirlpool
Completion	MEDINA PERFS TOP	Depth to the top of the Medina perforations
	MEDINA PERFS BOTTOM	Depth to the bottom of the Medina perforations
	NUMBER OF PERFS MEDINA	Number of perforations in Medina
	PERF DIAMETER	Diameter of perforations
	WHIRLPOOL PERFS TOP	Depth to the top of the Whirlpool perforations
Stimulation	BBL WATER MEDINA	Barrels of water used during stimulation
	MMCF N2 MEDINA	MMCF of Nitrogen used during the stimulation
	SKS SAND MEDINA	Sacks of sand used during the stimulation
	ISIP MEDINA	Initial Shut-In Pressure
	AVE TREATING RATE MEDINA	Average treatment rate
	TREATING PRESS MEDINA	Average treatment pressure
Production	Prior_Production_Gas	Cumulative gas produced
	Remaining_Gas	Remaining gas to be produced
	Ultimate_Gas	Estimated Ultimate Recovery
	Gas_Months_Prod	Number of months the well was on production
	Gas_Last_1_Months	Last one month of gas production
	Gas_Last_3_Months	Last three months of gas production
	Gas_Last_6_Months	Last six months of gas production
	Gas_DOFP	Date of the first gas production
	Gas_DOLP	Date of the last gas production
OUTPUT	Gas_Best_12_Months	Best 12 months of gas production
	Gas_5-Yr_Cum	Five year cumulative production
	Gas_10-Yr_Cum	Ten year cumulative production

Table 1. Data used for the analysis.

Please note that the above data was extracted from the database after the completion of the preliminary statistical analysis. This does not necessarily mean that all of the parameters shown in the Table 1 were used as input to the neural network. As a matter of fact after further analysis some of the parameters shown in Table 1 were eliminated from the list of parameters used in the final neural network predictive model.

STEP 1: Neural Network Analysis

In this section the result of neural network analysis will be presented. The analysis started with identifying the most influential category of the input parameters in the data set. The parameters in the data set were divided into five categories such as location, reservoir quality, completion, stimulation, and production. The results of this analysis will indicates the chances for the success of a restimulation program. The higher the influence of a category of parameters, the higher will be the chances for a successful restimulation program if the parameters of that category are altered in a positive manner. It should be indicated that the word "success" is being used in a relative context. **Figure 2** shows the neural network analysis for identification of most influential category in this data set. This figure demonstrates that the influence of completion and stimulation parameters are limited when compared to reservoir quality and production categories.

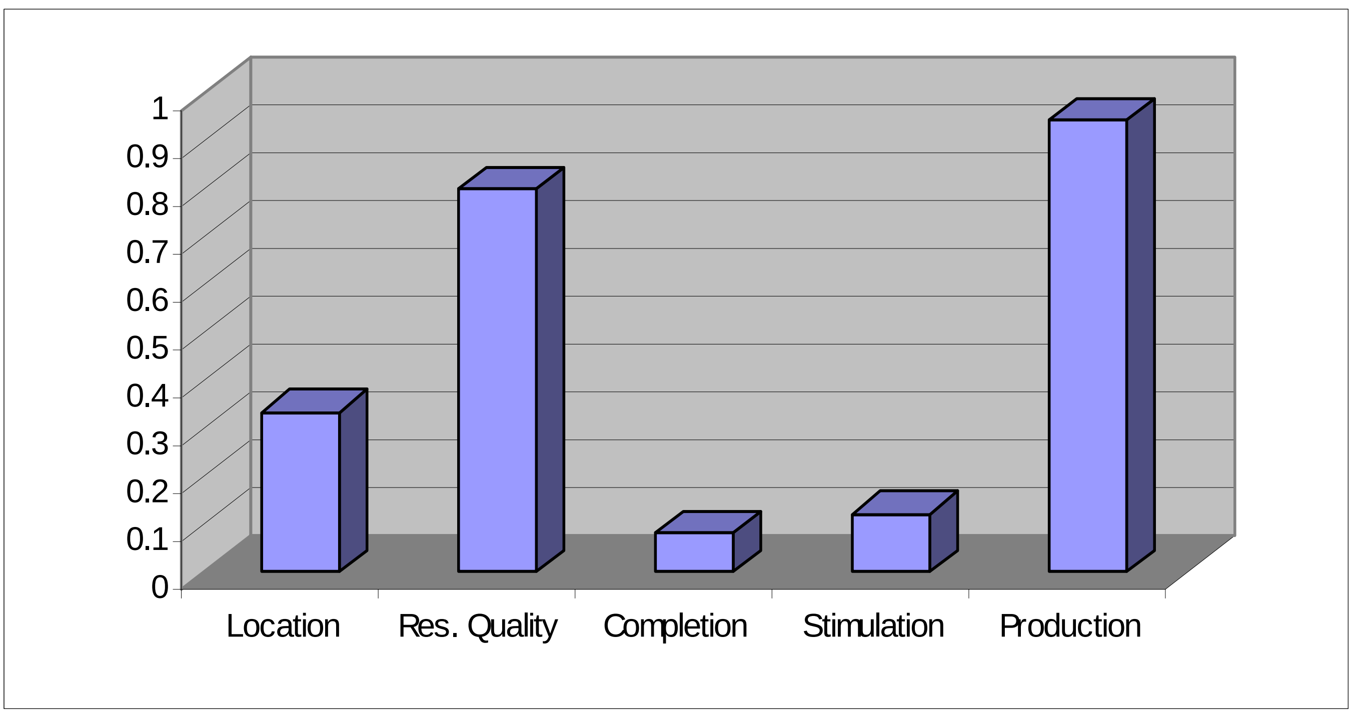


Figure 2. Influence of the categories in the data set.

This may lead to interpretations such that the stimulation practices in this field are not as effective as one would like them to be. It may also be concluded that a different set of stimulation practices should be tried in this field. It should be noted that almost all of the jobs in this field have been water fracs (at least those in our data set). Since the parameters in the stimulation category are the only set of parameters that can be altered during the optimization process (they are the only controllable parameters), the results in Figure 2 indicates that the improvements realized by the optimization process would also be limited.

Two production indicators, Best 12 months of production, Five year cumulative production, were used as the output for training of the neural networks. The input parameters were chosen from those shown in Table 1. Figures 3 and 4 show the results of the neural model building process.

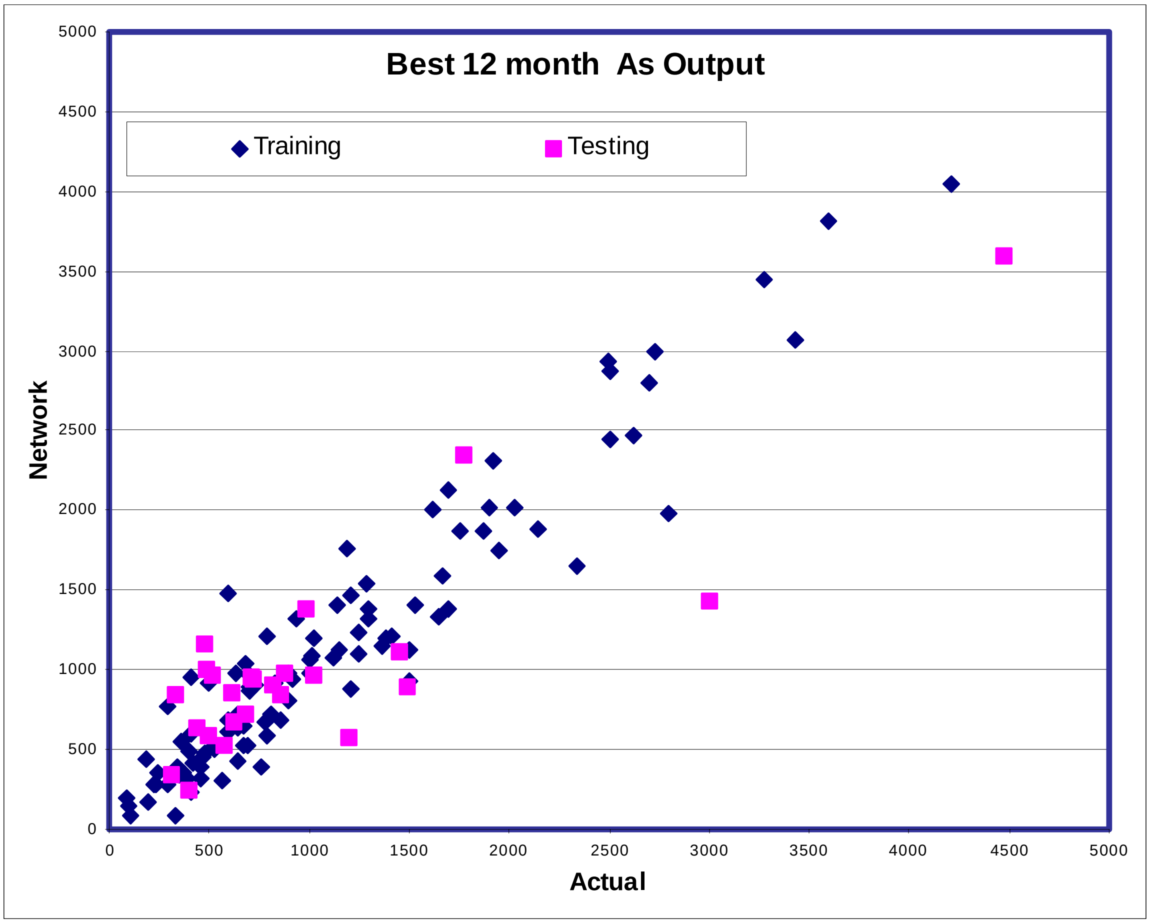


Figure 3. Training and verification data for best 12 months production indicator.

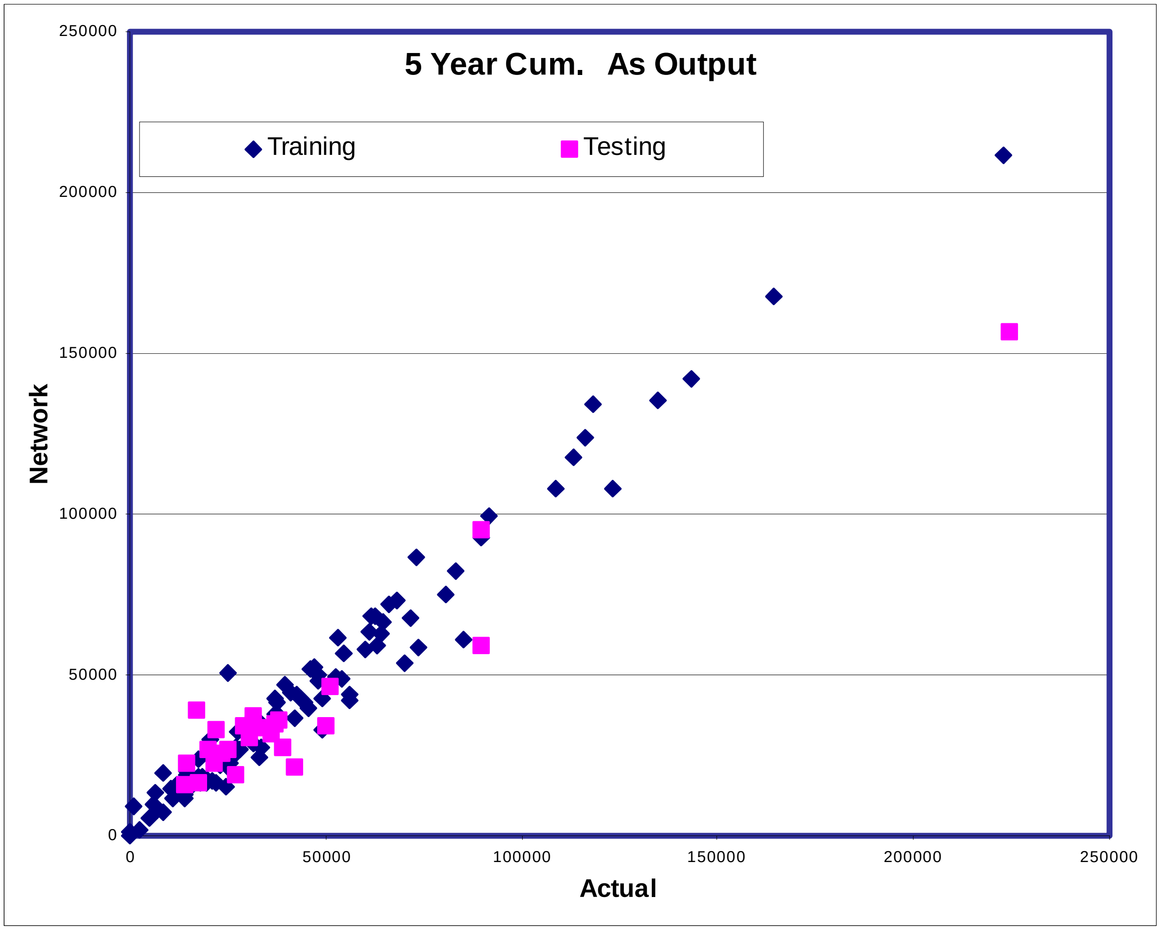


Figure 4. Training and verification data for five year cumulative production indicator.

It can be seen from these two figures that five year cumulative is a better production indicator than the best 12 months production to be used for our analysis. This conclusion is simply based on the correlation coefficient of the above two figures. The correlation coefficients for Figures 3 and 4 are shown in the Table 2 below.

	Training			Testing	
Patterns processed:	103			25	
Output:	**Best 12**	**5 Yr. Cum.**		**Best 12**	**5 Yr. Cum.**
Correlation coefficient r:	**0.9694**	**0.9833**		**0.8492**	**0.9546**

Table 2. Correlation coefficients for Figures 3 and 4.

STEP 2: Genetic Optimization

The first step in the genetic optimization process is the identification of controllable parameters. In this field the controllable parameters consist mainly of the stimulation parameters. These parameters are:

- Amount of water used in the frac job
- Amount of Nitrogen used in the frac job
- Amount of sand used in the frac job
- Average Treating Rate
- Treating Pressure

During the genetic optimization process the neural network that was developed during the step 1 of the process is used as the fitness function. The minimum and maximum for each of the above parameters are identified and the search and optimization process looks for the ideal combination of these parameters that results in the highest five year cumulative production using the neural network as the model that provides that output (five year cumulative production) when presented with inputs.

This process is repeated for each individual well. The pre-optimization five year cumulative production is then subtracted from the post-optimization five year cumulative production and the difference is suggested to be the "potential five year cumulative production". The idea is that the "potential five year cumulative production" is the incremental production that would have been recovered if the optimized stimulation job would have been implemented on the well, and therefore would be the recoverable production upon restimulation.

A software application was developed to perform the operations explained in steps 1 and 2 for the Medina/Whirlpool sands of western New York. Figures 5 and 6 show two screen shots of this software application. Figure 5 shows the screen shot for the neural predictive model. This application interfaces the database developed for this project. It provides all the available information for each well as the well is selected in the list box. The application includes a run-time version of the developed neural model and provides an answer for five year cumulative production each time the network is fired.

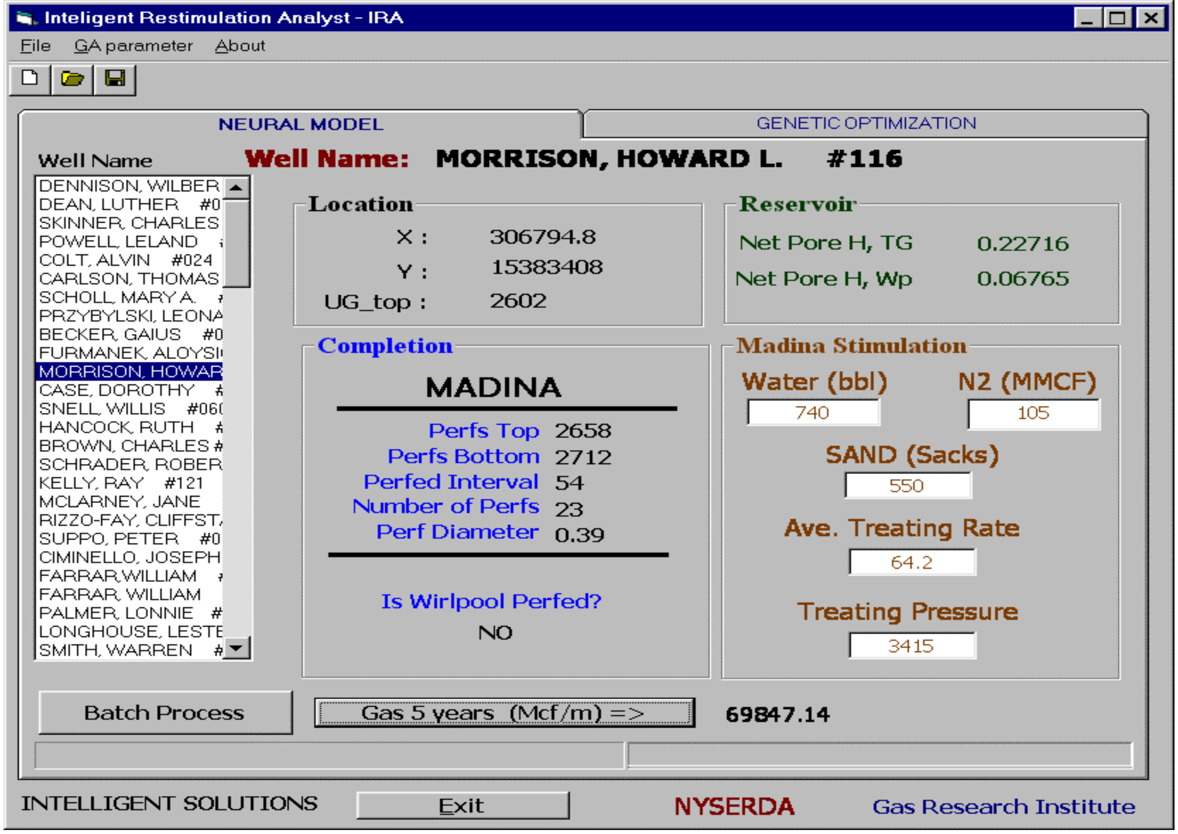


Figure 5. Neural predictive model interface.

Figure 6 show the genetic optimization interface. It also shows the process of evolution of the best solution (stimulation detail) as the process completes.

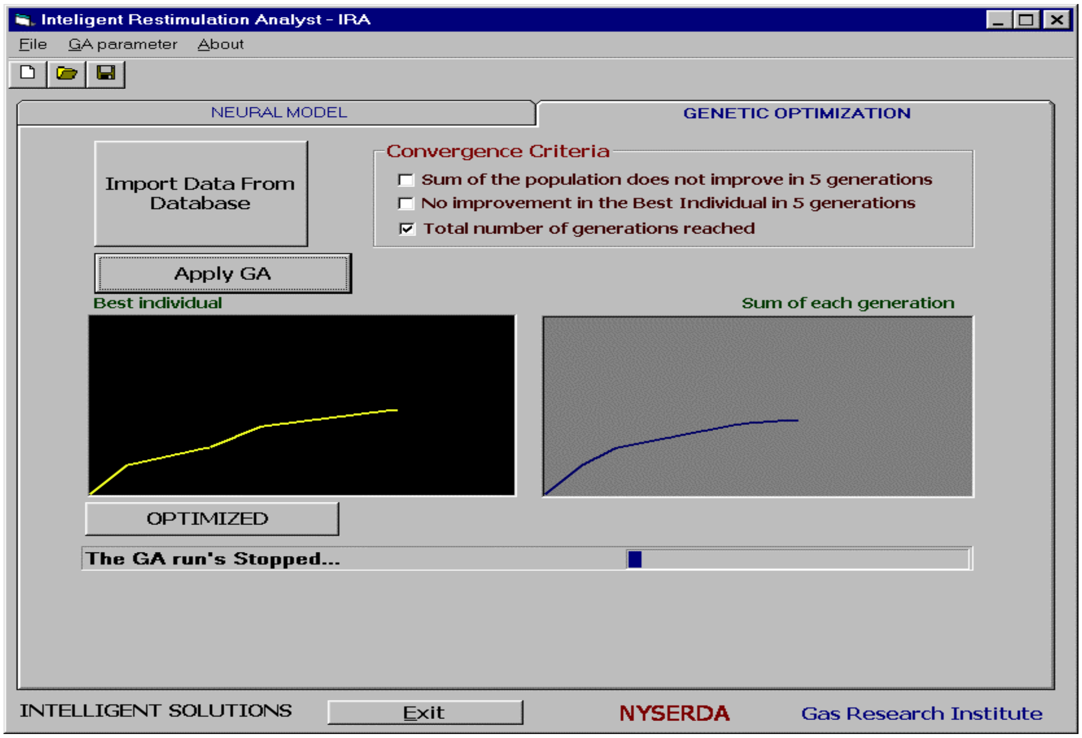


Figure 6. Genetic optimization interface.

STEP 3: Fuzzy Decision Support System

In a recent project sponsored by Gas Technology Institute (GTI) it was found that last month production of a well is a good indicator of the well being a restimulation candidate. Therefore it was decided to use this indicator along with two other parameters as the inputs to the fuzzy decision support system. The following three parameters were used in this section of the study:

1. Potential five-year cumulative production from steps 1 and 2.
2. Reservoir quality.
3. Last month production rate.

The reservoir quality was calculated using the following relationship:

$$reservoir\ quality\ index = \frac{Net(\phi * h)_{Madina} + 2\ Net(\phi * h)_{Whirlpool}}{No.\ of\ \ Perforations}$$

The above equation was designed in order to give more weight to the Whirlpool formation.

Three different fuzzy decision support systems were designed. The first and second fuzzy systems had two inputs and one output while the third fuzzy system had three inputs and one output.

First Fuzzy System

The inputs to this fuzzy system included potential five-year cumulative production and last month production rate.

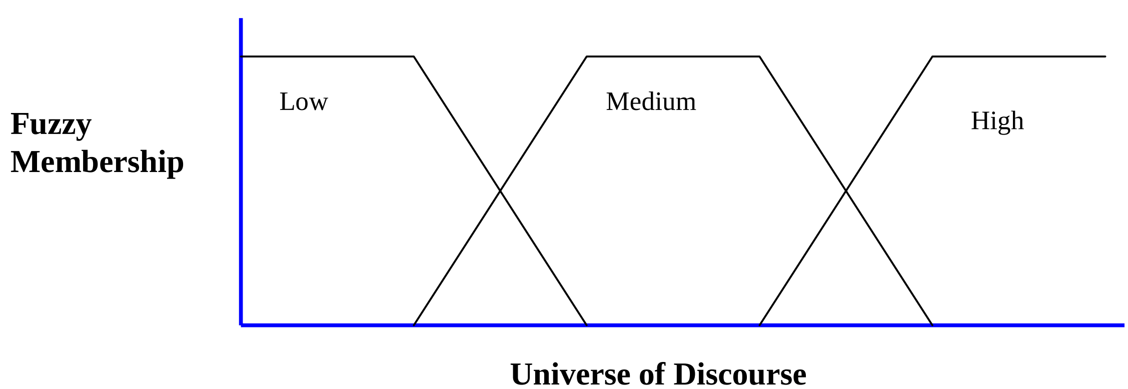


Figure 7. A typical set of fuzzy sets for the input or output parameters.

The output for this fuzzy system is the restimulation candidacy. Figure 7 shows the typical fuzzy sets for each one of the parameter. For the two input parameters the fuzzy sets included low, medium and high. The fuzzy sets for the output were "The well is not a candidate", "The well may be a candidate", and "The well is a candidate". Using two input parameters each having three fuzzy sets requires the fuzzy system to have nine fuzzy rules. The nine fuzzy rules for this system is shown in figure 8.

Incremental				
	Low	No (VT)	No (T)	May Be (T)
	Medium	No (T)	May Be (T)	Yes (T)
	High	May Be (T)	Yes (T)	Yes (VT)
		Low	Medium	High
		Last Month Rate		

Figure 8. Set of nine fuzzy rules used for the first fuzzy system.

A typical fuzzy rule in the above figure can be written as:

If the potential five year cumulative production for this well is ***Low*** **and** the Last month production rate is ***Low*** **then** the well ***is not*** a restimulation candidate. This rule in the above figure is qualified by the approximate reasoning parameter "very true" (VT). The approximate reasoning in this fuzzy system is constructed as shown in Figure 9.

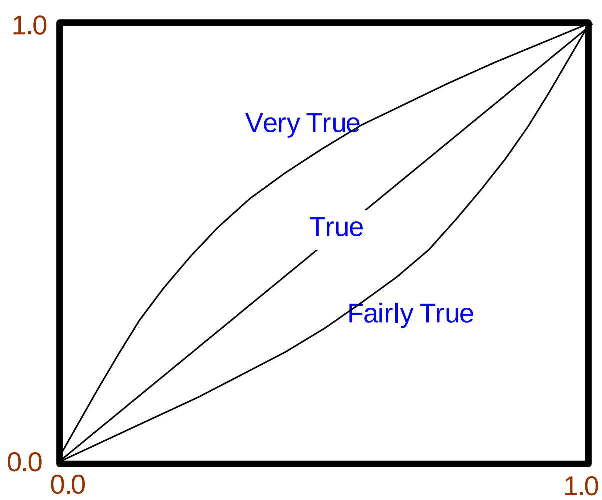


Figure 9. The approximate reasoning algorithm used in the fuzzy systems.

The above approximate reasoning algorithm provides qualification for each rule distinguishing them from one another. Two different fuzzy rules may have the same outcome but based on the input values one might be more or less true than the other. The approximate reasoning algorithm is designed to account for such qualifications.

The fuzzy sets, rules and inference engine were embedded in a Windows application to assist in implementation of the fuzzy system for the Madina/ Whirlpool sands of western New York. Figure 10 shows the interface of this application.

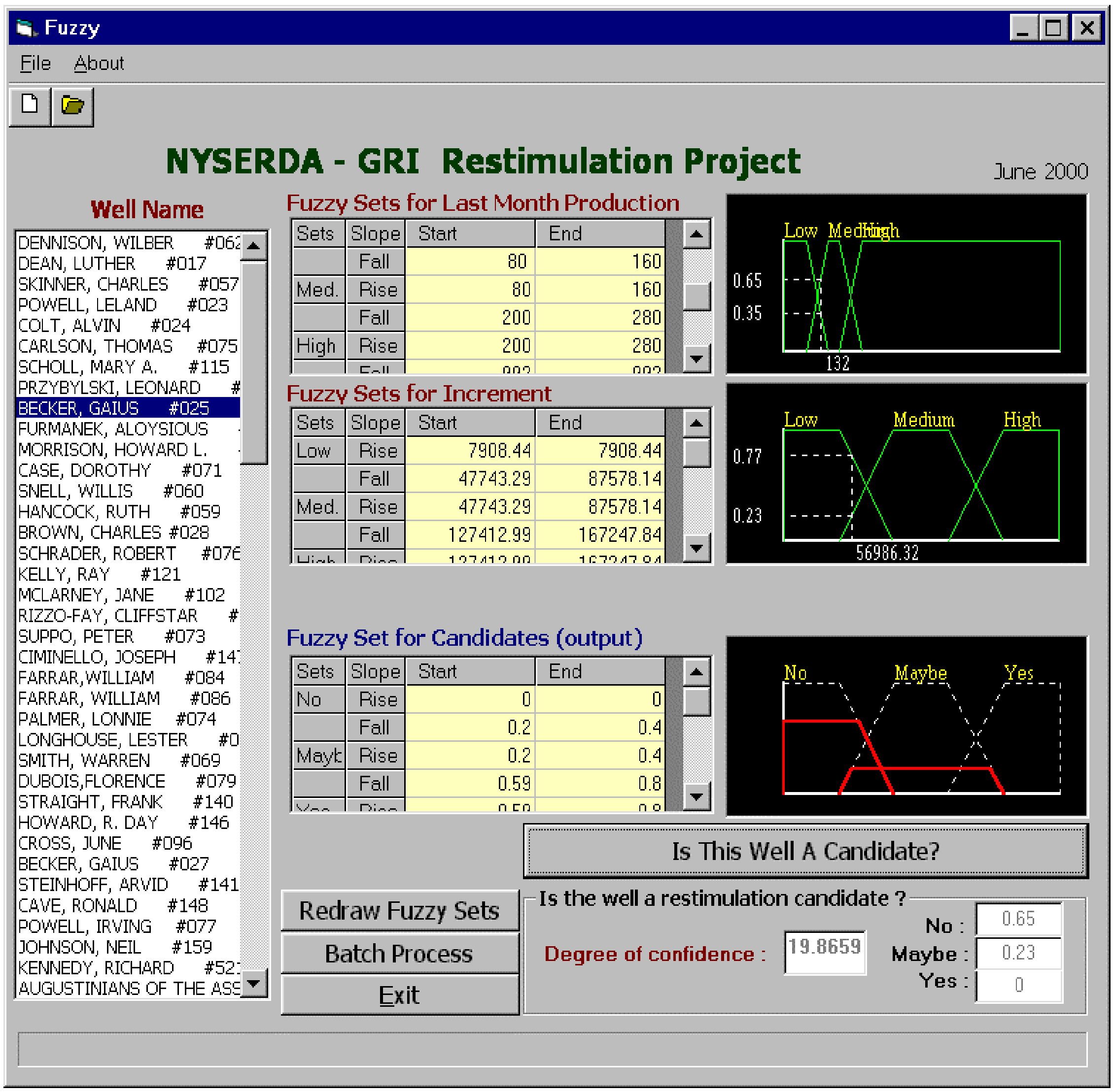


Figure 10. The application interface for the first fuzzy systems.

Using this application and applying the fuzzy system to every individual well in the study and then ranking them based on the system output provided a list that shows the top restimulation candidates. Figure 11 shows the top 25 candidates selected using the first fuzzy system with two inputs.

Rank	Well Name
1	SUESS, GEORGE #430
2	PRZYBYLSKI, LEONARD #113
3	MIKULA,JOSEPH #153
4	SUTTON, DAVID #315
5	MARTIN, DAVID #300
6	AUGUSTINIANS OF THE ASSUM #103
7	WILKENS, ROY #343
8	DORMAN, G. & M. #3 #373
9	CROSS, LINDA KAHLE #388
10	LONGHOUSE ESTATE, INC. #24
11	COLT, ALVIN #024
12	POWELL, LELAND #023
13	WILLS, WILLIAM #325
14	SCHOLL, MARY A. #115
15	SUPPO, PETER #073
16	WINCHELL, FRANCIS #303
17	PALMER, LONNIE #074
18	FURMANEK, ALOYSIOUS #144
19	MCLARNEY, JANE #102
20	CHILCOTT, EUGENE #365
21	BECKER, GAIUS #027
22	MIKULA, JOSEPH #152
23	SMITH, WARREN #069
24	BARBER #2
25	DORMAN, G. & M. #2 #349

Figure 11. Top 25 restimulation candidates from the first fuzzy system.

Second Fuzzy System

Same approach as the first fuzzy system was used to develop a second fuzzy system. The second fuzzy system also had two inputs and one output. The inputs of this fuzzy system included potential five year cumulative production and reservoir quality index. Figure 12 shows the fuzzy rules used in this fuzzy system.

Incrementa	Low	No (VT)	No (T)	May Be (T)
	Medium	No (T)	May Be (T)	Yes (T)
	High	May Be (T)	Yes (T)	Yes (VT)
		Low	Medium	High
		Reservoir Quality Index		

Figure 12. Set of nine fuzzy rules used for the second fuzzy system.

Using the same approximate reasoning algorithms shown in Figure 9 and the application shown in Figure 10 a new set of restimulation candidates were selected. Figure 13 shows the list of the candidates generated using the second fuzzy system.

1	MARTIN, DAVID #300
2	SUESS, GEORGE #430
3	WILLS, WILLIAM #325
4	CRANSTON, CLAUDE #311
5	WINCHELL, FRANCIS #303
6	SMITH, M. #1 #245
7	PRZYBYLSKI, LEONARD #113
8	PALMER, LONNIE #074
9	MONETTE, CONSTANCE #427
10	FURMANEK, ALOYSIOUS #144
11	SMITH, WARREN #069
12	CORNELL, GORDON #423
13	MCLARNEY, JANE #102
14	AUGUSTINIANS OF THE ASSUM #103
15	DORMAN, G. & M. #3 #373
16	DUBOIS,FLORENCE #079
17	VILLAGE OF BROCTON #299
18	SUPPO, PETER #073
19	MIKULA,JOSEPH #153
20	COLT, ALVIN #024
21	SCHOLL, MARY A. #115
22	VAN DETTE, ALBERT #339
23	BEMUS, CECILE #214
24	VAN DETTE, ALBERT #356
25	CRANDALL, RICHARD #326

Figure 13. Top 25 restimulation candidates from the second fuzzy system.

Third Fuzzy System

The third fuzzy system was designed with three inputs and one output. The input parameters were included potential five year cumulative production, last month production and reservoir quality index. Each of these input parameters had three fuzzy sets; low, medium, and high. This resulted in 27 fuzzy rules. Figure 14 shows the 27 fuzzy rules that was used in the third fuzzy system. As can be seen in this figure all the rules are qualified using the same approximate reasoning demonstrated in Figure 9.

	Reservoir Quality Low			Reservoir Quality Medium			Reservoir Quality High		
Last Month Rate									
High	No (T)	May Be (FT)	May Be (FT)	May Be (T)	Yes (FT)	Yes (FT)	Yes (T)	Yes (VT)	Yes (VT)
Med.	No (VT)	No (T)	May Be (FT)	No (FT)	May Be (T)	Yes (FT)	May Be (VT)	Yes (T)	Yes (VT)
Low	No (VT)	No (VT)	No (T)	May Be (FT)	No (FT)	May Be (T)	May Be (VT)	May Be (VT)	Yes (T)
	Low	Med.	High	Low	Med.	High	Low	Med.	High
	Incremental			Incremental			Incremental		

Figure 14. The 27 fuzzy rules used in the third fuzzy system.

A new interface for the third fuzzy system was developed. This interface is shown in Figure 15.

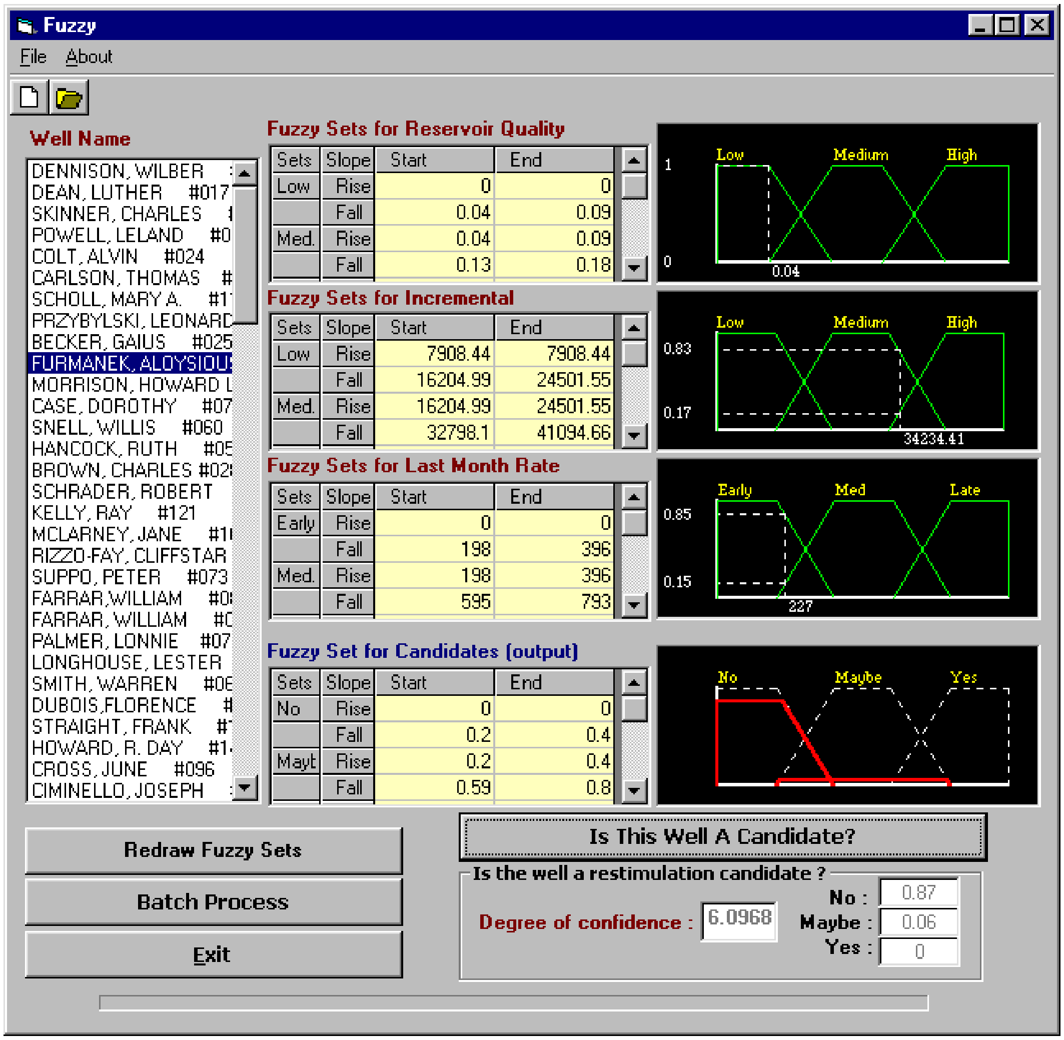


Figure 15. The application interface developed for the third fuzzy system.

Using the third fuzzy system a new set of restimulation candidates were selected. This list is shown in Figure 16.

Rank	Well Name
1	PALMER, LONNIE #074
2	WILLS, WILLIAM #325
3	WINCHELL, FRANCIS #303
4	SUESS, GEORGE #430
5	MARTIN, DAVID #300
6	MCLARNEY, JANE #102
7	VILLAGE OF BROCTON #299
8	SMITH, M. #1 #245
9	SUPPO, PETER #073
10	DUBOIS,FLORENCE #079
11	CRANSTON, CLAUDE #311
12	SCHOLL, MARY A. #115
13	SMITH, WARREN #069
14	AUGUSTINIANS OF THE ASSUM #103
15	BARBER #2
16	STRAIGHT, FRANK #140
17	VAN DETTE, ALBERT #356
18	CHILCOTT, EUGENE #365
19	CORNELL, GORDON #423
20	MILLER, MORRIS #433
21	JOSEPHSON, WALFRED #329
22	COLT, ALVIN #024
23	ZOOK, MARVIN #276
24	BEMUS, CECILE #214
25	CRANDALL, RICHARD #326

Figure 16. Top 25 restimulation candidates from the third fuzzy system.

CONCLUSION

Upon completion of the analysis using three different fuzzy systems three different lists of candidates were developed. The last step in this analysis is to resolve the three separate lists into a single list and recommend the final list of the restimulation candidates.

Figure 17 shows the resolved list of the candidates.

FINAL RANKING				
		Ranking		
Rank	Well Name	Last Month Rate	Reservoir Quality	Three-Input System
1	SUESS, GEORGE #430	1	2	4
2	WILLS, WILLIAM #325	13	3	2
3	WINCHELL, FRANCIS #303	16	5	3
4	PALMER, LONNIE #074	17	8	1
5	MARTIN, DAVID #300	5	1	21
6	AUGUSTINIANS OF THE ASSUM #103	6	14	14
7	MCLARNEY, JANE #102	19	13	5
8	SUPPO, PETER #073	15	18	9
9	SCHOLL, MARY A. #115	14	21	12
10	SMITH, WARREN #069	23	11	13
11	COLT, ALVIN #024	11	20	22
12	SMITH, M. #1 #245		6	8
13	CRANSTON, CLAUDE #311		4	11
14	VILLAGE OF BROCTON #299		17	7
15	DUBOIS,FLORENCE #079		16	10
16	CORNELL, GORDON #423		12	19
17	CHILCOTT, EUGENE #365	20		18
18	BARBER #2	24		15
19	VAN DETTE, ALBERT #356		22	17
20	BEMUS, CECILE #214		23	24
21	CRANDALL, RICHARD #326		25	25
22	STRAIGHT, FRANK #140			16
23	MILLER, MORRIS #433			20
24	JOSEPHSON, WALFRED #329			21
25	ZOOK, MARVIN #276			23

Figure 17. Top 25 restimulation candidates from the virtual intelligence analysis.

APPENDIX A

Artificial Neural Networks

In a typical neural data processing procedure, the data set is divided into two separate groups called the training and the test sets. The training set is used to develop the desired network. In this process (depending on the paradigm that is being used), the desired output in the training set is used to help the network adjust the weights between its neurons or processing elements (supervised training.) Once the network has learned the information in the training set and has "converged," the test set is applied to the network for verification. It is important to note that, although the user has the desired output of the test set, the network has not seen it. This is to ensure the integrity and robustness of the trained network. In order to clarify the actual functionality of a neural system, a short discussion on the mechanics and components of artificial neural network follows. Our experience with neural networks have shown that one will get <u>some</u> sort of results by treating neural a network as a black box, where one inputs the data, trains the network and gets some output. It has been our observation that a fundamental understanding of the theory and application of virtual intelligence in general and neural networks specifically is essential in achieving meaningful results and repeatable outcomes.

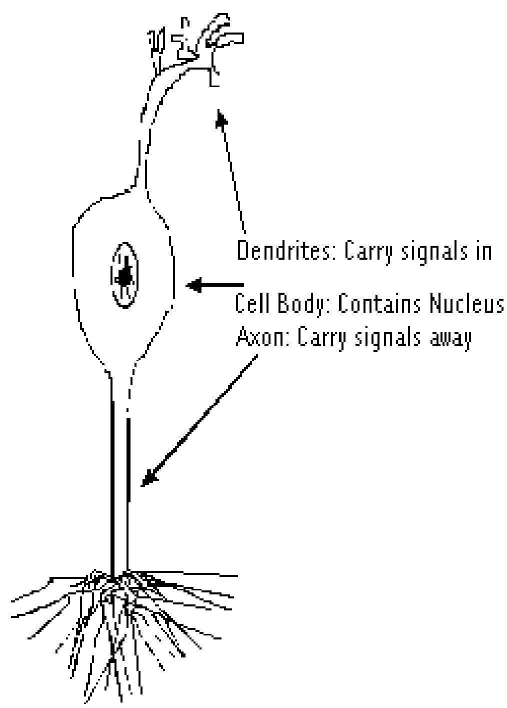


Figure A-1. **Three parts of a typical nerve cell.**

A biological neuron is a nerve cell with all of its processes. Neurons are one of the main distinguishing features of animals. Figure A-1 is a bipolar neuron, which means it has two processes. The cell body contains the *nucleus.* Leading into the nucleus is one or more dendrites. These branching, tapering processes of the nerve cell, as a rule, conduct impulses toward the cell body. The *axon* is the nerve cell process that conducts impulses away from the cell body. Bundles of neurons, or nerve fibers, form *nerve structures*. In a simplified scenario, nerves conduct impulses from receptor organs (such as eyes or ears) to effector organs (such as muscles or glands). The point between two neurons in a neural pathway, where the termination of the axon of one neuron comes into close proximity with the cell body or dendrites of another, is called a *synapse.* At this point, a microscopic gap, the relationship of the two neurons is one of contact only. The impulse traveling in the first neuron initiates an impulse in the second neuron. Signals come into the synapses. These are the inputs. They are "weighted." That is, some signals are stronger than others. Some signals excite (are positive), and others inhibit (are negative). The effects of all weighted inputs are summed. If the sum is equal to or greater than the threshold for the neuron, then the neuron fires (gives output). This is an "all-or-nothing" situation. Either neuron fires or it doesn't fire.

Within the last few years, hardware improvements have made computer simulation of artificial neural network possible. Although it may seem strange to simulate a parallel process on a sequential machine, there have been many benefits. It has bought time for the real objective of implementing neural networks in hardware, and it has illuminated problems in earlier models. Simulations have allowed us to better understand and improve the technology, and to tell in advance how well a particular neural network will perform in a given application. In addition to simulations, analog neural network circuits have been built and tested.

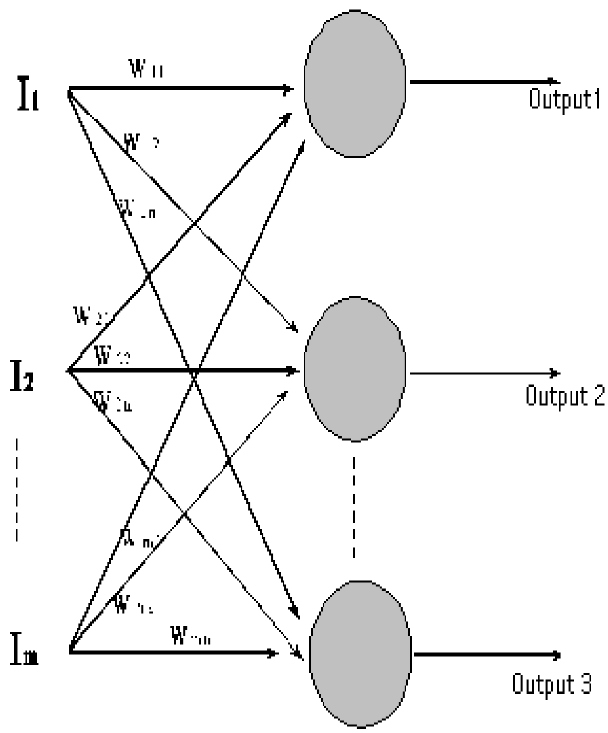


Figure A-2. Inputs can be connected to many nodes.

In neural computing the artificial neuron is called a *Processing Element* or *PE* for short. The word *node* is also used for this simple building block, which is represented by circles in Figure A-2. These artificial neurons bear only a modest resemblance to the real thing. They are barely a first order approximation of biological neurons. Neurons in the human brain perform at least 150 different processes, where as Processing Elements model approximately three of those processes. The PE handles several basic functions. It must evaluate input signals and determine the strength of each one. Next, it must calculate a total for the combined input signals and compare that total to some threshold level. Finally, it must determine what the output should be. Just as there are much input (stimulation levels) to a neuron, there should be many input signals to a PE. All of them should come into PE simultaneously. In response, a neuron either "fires" or "doesn't fire," depending on some *threshold* level. The PE will be allowed a single output signal, just as in a biological neuron - much input, one output.

In this project the collected data will be transformed into fuzzy sets (explained below) and then a neural network is trained using the transformed data. The neural model is built to discover any possible patterns in the data. Once the network is trained and tested it will be the predictive model, which is one of the objectives of this study.

APPENDIX B

Genetic Algorithms

In order to understand genetic algorithms evolutionary computation should first be defined.

Evolutionary Computation

Evolutionary Computation is an umbrella term used to describe computer-based problem solving systems which use computational models of some of the known mechanisms of evolution as key elements in their design and implementation. A variety of evolutionary computation have been proposed.

The major ones are: Genetic Algorithms, Evolutionary Programming, Evolution Strategies, Classifier Systems, and Genetic Programming. They all share a common conceptual base of simulating the evolution of individual structures via processes of selection, mutation, and reproduction. The processes depend on the perceived performance of the individual structures as defined by an environment. More precisely, Evolutionary Computation maintain a population of structures, that evolve according to rules of selection and other operators, that are referred to as "search operators", (or genetic operators), such as recombination and mutation. Each individual in the population receives a measure of it's fitness in the environment. Reproduction focuses attention on high fitness individuals, thus exploiting the available fitness information. Recombination and mutation perturb those individuals, providing general heuristics for exploration. Although simplistic from a biologist's viewpoint, these algorithms are sufficiently complex to provide robust and powerful adaptive search mechanisms.

Biological Basis

To understand Evolutionary Computation, it is necessary to have some appreciation of the biological processes on which they are based. Firstly, we should note that evolution (in nature or anywhere else) is not a purposive or directed process. That is, there is no evidence to support the assertion that the goal of evolution is to produce Mankind. Indeed, the processes of nature seem to boil down to a haphazard generation of biologically diverse organisms. Some of evolution is determined by natural selection or different individuals competing for resources in the environment. Some are better than others. Those that are better are more likely to survive and propagate their genetic material.

In nature, we see that the encoding for genetic information (GENOME) is done in a way that admits asexual reproduction. Asexual reproduction typically results in offspring that are genetically identical to the parent. (Large numbers of organisms reproduce asexually; this includes most bacteria which many biologists hold to be the most successful species known.) Sexual reproduction allows some shuffling of chromosomes, producing offspring that contain a combination of information from each parent. At the molecular level what occurs (wild oversimplification alert!) is that a pair of almost identical chromosomes bump into one another, exchange chunks of genetic information and drift apart. This is the recombination operation, which is often referred to as CROSSOVER because of the way that biologists have observed strands of chromosomes crossing over during the exchange.

Recombination happens in an environment where, among other things, the selection of who gets to mate is a function of the fitness of the individual, i.e. how good the individual is at competing in its environment. Some Evolutionary Computational technics use a simple function of the fitness measure to select individuals (probabilistically) to undergo genetic operations such as crossover or asexual reproduction (the propagation of genetic material unaltered). This is fitness-proportionate selection. Other implementations use a model in which certain randomly selected individuals in a subgroup compete and the fittest is selected. This is called tournament selection and is the form of selection we see in nature when stags rut to vie for the privilege of mating with a herd of hinds.

Much Evolutionary Computation research has assumed that the two processes that most contribute to evolution are crossover and fitness based selection/reproduction. As it turns out, there are mathematical proofs that indicate that the process of fitness proportionate reproduction is, in fact, near optimal in some senses.

Evolution, by definition, absolutely requires diversity in order to work. In nature, the primary source of diversity is mutation. In Evolutionary Computation, a large amount of diversity is usually introduced at the start of the algorithm, by randomizing the GENEs in the population. The importance of mutation, which introduces further diversity while the algorithm is running, therefore continues to be a matter of debate. Some refer to it as a background operator, simply replacing some of the original diversity which has been lost, while others view it as playing the dominant role in the evolutionary process.

It cannot be stressed too strongly that an evolutionary algorithm (as a simulation of a genetic process) is not a random search for a solution to a problem (highly fit individual). Evolutionary Computations use stochastic processes, but the result is distinctly non-random search.

What's a Genetic Algorithm?

Genetic Algorithm is a model of machine learning which derives its behavior from a metaphor of one of the mechanisms of evolution in nature (namely, hard selection). This is done by the creation within a machine of a population of individuals represented by chromosomes, in essence a set of character strings that are analogous to the base-4 chromosomes that we see in our own DNA. The individuals in the population then go through a process of selection (evolution). Genetic Algorithms are used for a number of different application areas. An example of this would be multidimensional optimization problems in which the character string of the chromosome can be used to encode the values for the different parameters being optimized. In practice, therefore, we can implement this genetic model of computation by having arrays of bits or characters to represent the chromosomes. Simple bit manipulation operations allow the implementation of crossover, mutation and other operations. Although a substantial amount of research has been performed on variable length strings and other structures, the majority of work with Genetic Algorithms is focussed on fixed-length character strings. We should focus on both aspects of fixed-lengthiness and the need to encode the representation of the solution being sought as a character string, since these are crucial aspects that distinguish genetic programming, which does not have a fixed length representation and there is typically no encoding of the problem.

When the Genetic Algorithm is implemented it is usually done in a manner that involves the following cycle: Evaluate the fitness of all of the individuals in the population. Create a new population by performing operations such as crossover, fitness- proportionate reproduction and mutation on the individuals whose fitness has just been measured. Discard the old population and iterate using the new population.

One iteration of this loop is referred to as a generation. There is no theoretical reason for this as an implementation model. Indeed, we do not see this punctuated behavior in populations in nature as a whole, but it is a convenient implementation model. The first generation (generation 0) of this process operates on a population of randomly generated individuals. From there on, the genetic operations, in concert with the fitness measure, operate to improve the population.

APPENDIX C

Fuzzy Logic

Fuzzy Logic has emerged as a profitable tool for the controlling of subway systems and complex industrial processes, as well as for household and entertainment electronics, diagnosis systems and other expert systems. Fuzzy Logic is basically a multi-valued logic that allows intermediate values to be defined between conventional evaluations like yes/no, true/false, black/white, etc. Notions like rather warm or pretty cold can be formulated mathematically and processed with the computer. In this way an attempt is made to apply a more human-like way of thinking in the programming of computers. Fuzzy Logic was initiated in 1965 by Lotfi A. Zadeh, professor of computer science at the University of California in Berkeley. Zadeh started Fuzzy Logic as a means to model the uncertainty of natural language. Zadeh says that rather than regarding fuzzy theory as a single theory, we should regard the process of ``fuzzification'' as a methodology to generalize ANY specific theory from a crisp (discrete) to a continuous (fuzzy) form. Thus recently researchers have also introduced "fuzzy calculus", "fuzzy differential equations", and so on.

Fuzzy Sets

Just as there is a strong relationship between Boolean logic and the concept of a subset, there is a similar strong relationship between fuzzy logic and fuzzy subset theory. In classical set theory, a subset U of a set S can be defined as a mapping from the elements of S to the elements of the set {0, 1},

U: S --> {0, 1}

This mapping may be represented as a set of ordered pairs, with exactly one ordered pair present for each element of S. The first element of the ordered pair is an element of the set S, and the second element is an element of the set {0, 1}. The value zero is used to represent non-membership, and the value one is used to represent membership. The truth or falsity of the statement:

x is in U

is determined by finding the ordered pair whose first element is x. The statement is true if the second element of the ordered pair is 1, and the statement is false if it is 0.

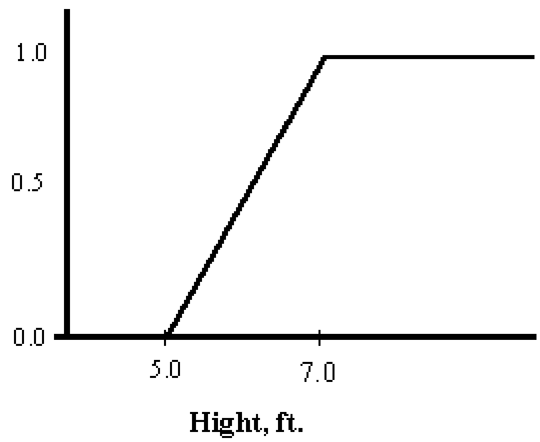
Similarly, a fuzzy subset F of a set S can be defined as a set of ordered pairs, each with the first element from S, and the second element from the interval [0,1], with exactly one ordered pair present for each element of S. This defines a mapping between elements of the set S and values in the interval [0,1]. The value zero is used to represent complete non-membership, the value one is used to represent complete membership, and values in between are used to represent intermediate DEGREES OF MEMBERSHIP. The set S is referred to as the UNIVERSE OF DISCOURSE for the fuzzy subset F. Frequently, the mapping is described as a function, the MEMBERSHIP FUNCTION of F. The degree to which the statement

x is in F

is true is determined by finding the ordered pair whose first element is x. The DEGREE OF TRUTH of the statement is the second element of the ordered pair. In practice, the terms "membership function" and fuzzy subset get used interchangeably. Let’s clarify these definitions with an example. Let's talk about people and "tallness". In this case the set S (the universe of discourse) is the set of people. Let's define a fuzzy subset TALL, which will answer the question "to what degree is person x tall?" Zadeh describes TALL as a LINGUISTIC VARIABLE, which represents our cognitive category of "tallness". To each person in the universe of discourse, we have to assign a degree of membership in the fuzzy subset TALL. The easiest way to do this is with a membership function based on the person's height.

$$\text{tall}(x) = \begin{cases} 0, & \text{IF} \quad \text{height}(x) < 5 \text{ ft.,} \\ (\text{height}(x)-5\text{ft.})/2 \text{ ft.,} & \text{IF} \quad 5 \text{ ft.} <= \text{height}(x) <= 7 \text{ ft.,} \\ 1, & \text{IF} \quad \text{height}(x) > 7 \text{ ft.} \end{cases}$$

Graph of the above statement would look like this:



Given this definition, here is some example values:

Person	Height	Degree of Tallness
Billy	3' 2"	0.00
Joe	5' 5"	0.21
Drew	5' 9"	0.38
Erik	5' 10"	0.42
Mark	6' 1"	0.54
Kareem	7' 2"	1.00

Expressions like "A is X" can be interpreted as degrees of truth, e.g., "Drew is TALL"=0.38. Membership functions used in most applications almost never have as simple a shape as tall(x). At minimum, they tend to be triangles pointing up, and they can be much more complex than that. Also, the discussion characterizes membership functions as if they always are based on a single criterion, but this isn't always the case, although it is quite common. One could, for example, want to have the membership function for TALL depend on both a person's height and their age (he's tall for his age). This is perfectly legitimate, and occasionally used in practice. It's referred to as a two-dimensional membership function, or a "fuzzy relation". It's also possible to have even more criteria, or to have the membership function depend on elements from two completely different universes of discourse.

Logic Operations

Now that we know what a statement like "X is LOW" means in fuzzy logic, how do we interpret a statement like

X is LOW and Y is HIGH or (not Z is MEDIUM)

The standard definitions in fuzzy logic are:

truth (not x) = 1.0 - truth (x)

truth (x and y) = minimum (truth(x), truth(y))

truth (x or y) = maximum (truth(x), truth(y))

Some researchers in fuzzy logic have explored the use of other interpretations of the AND and OR operations, but the definition for the NOT operation seems to be safe.

Note that if one plugs just the values zero and one into these definitions, one gets the same truth tables as one would expect from conventional Boolean logic. This is known as the EXTENSION PRINCIPLE, which states that the classical results of Boolean logic are recovered from fuzzy logic operations when all fuzzy membership grades are restricted to the traditional set {0, 1}. This effectively establishes fuzzy subsets and logic as a true generalization of classical set theory and logic. In fact, by this reasoning all crisp (traditional) subsets ARE fuzzy subsets of this very special type; and there is no conflict between fuzzy and crisp methods.

Appendix 2

Production Type Curve Analysis Methodology and Evolution for the Determination of Restimulation Production Potential in New York

Advanced Resources International, Inc.

Subject: Production type curve analysis methodology and evolution for the determination of restimulation production potential in New York.

Chautauqua County, New York Study – Background

In response to the success of the 1996 GTI evaluation of restimulating existing tight gas sand wells, the New York State Energy Research and Development Authority (NYSERDA) and GTI jointly sponsored a restimulation feasibility study patterned after the aforementioned program. The deterministic methodology employed production data analysis, artificial neural networks and engineering-based performance type curves as well as advanced reservoir modeling techniques to identify restimulation candidates for the Belden & Blake Corporation's Chautauqua County, New York Medina Group gas production field.

Chautauqua County, New York Study – Analysis Execution and Results

Chautauqua Study Data and Analysis: Belden & Blake had previously worked with ARI on a 1998 NYSERDA-sponsored project to help determine if natural gas production and reserve improvement opportunities existed in the Medina Group, located in Chautauqua County, New York[1,2]. As part of this work, a systematic geologic and engineering evaluation was performed, inclusive of well log petrophysical analysis, geological mapping and advanced production type curve analysis, which provided net sand, effective porosity, saturation as well as historic production and pressure data for this study.

However, following a review of the existing production database, Schlumberger - Holditch Reservoir Technologies (HRT) noted several discrepancies between a given production well's historic first date of production and the first record in the database. Therefore, each production history was reviewed to standardize this difference, resulting in the addition of up to one year of production history for a large proportion of the wells.

Further, several wells experienced unexplained decreases and increases in gas production throughout their history. So, at the request of ARI, Belden & Blake performed a critical review of six well files to characterize the nature of the aberrant production data. ARI selected the six wells (the George Sheppard 142, William Potkovick 376, Warren Smith 69, YMCA 523, Rizzo-Fay 151 and Roy Wilkens 343) such that they covered each of the completion types (Whirlpool, Grimsby or Commingled) while maintaining production behavior that was similar (in some fashion) to several other wells. Major findings of this effort, which were generally applied across the field, are as follows:

- Wells have not been recompleted during their life.
- Wells have not been refractured during their life.
- No additional compression seems to have been added to the field.
- Aberrant production increases were found to be due to flush production following shut-in periods, the installation of artificial lift methods (casing plunger, rabbit), swabbing or the use of soap.

- Aberrant production decline was found to be due to the lack of sufficient energy to lift reservoir fluids, sand production and, in one case, a casing leak. Subsequent remedial operations were able to increase production.

Type Curve Matching Methodology: While there were no recompletions, restimulations or lease compression changes to consider, ARI recognized that the installation of artificial lift methods, such as plunger lifts or rabbits, would effectively reduce the well's bottomhole flowing pressure. So, each of the historical semi-log production and pressure charts was reviewed in conjunction with available completion/wellbore data to identify those wells that were impacted by this production enhancement method. To account for its impact using production type curves, a match restart of the type curve was made at the onset of artificial lift installation, using a reduced flowing pressure consistent with the provided historical data. **Figure 1** shows a typical, restarted, production type curve match. **Appendix A** contains the complete set of *METEOR* inputs and results for the type curve matching effort, while **Figures 2, 3, 4 and 5** show the results in a histogram distribution format.

The median results were determined to be a permeability of 39 microdarcies, an infinite conductivity fracture half-length of 20 feet, a drainage area of 77 acres and an expected average reservoir pressure of 409 psia on May 2001. Although there is no data to verify the accuracy of the permeability and fracture half-length values, they appear within reason. The median drainage area is also a reasonable number. However, given the uncertainty of differential depletion between the Whirlpool and Grimsby formations and the recent experience of encountering reduced pressure at infill drilling locations, verification of this estimate is problematic.

Low net sand estimates for a small group of wells skewed the type curve results to larger than expected values of permeability, fracture half-length and drainage area. For example, the David Sutton #315 type curve match resulted in an estimated permeability of 0.14 md, a fracture half-length of 40 feet and a drainage area of 268 acres, based on 27 feet of net sand. Reviewing the distributions, these results were near the largest range for each of the distributions, indicating contributing sand may have been underestimated. In cases such as this, a Voronoi well spacing estimate could have provided a ceiling to the drainage area, resulting in an increased reservoir thickness to reduce the drainage volume.

Restimulation and Added Compression Incremental Recovery Methodologies: To determine the impact of restimulation on each of the wells, it was assumed that new drainage area would not be created during the restimulation of an existing well. Further, the well’s full drainage area would be at the May 2001 (the estimated restimulation date) average reservoir pressure, creating a more conservative assessment of the incremental restimulation values. Note that these premises assume that both the Grimsby and Whirlpool formations are depleted (with no differential depletion) and that no additional pay thickness would be encountered within the variable sequence of the Medina Group.

Further, a practical expectation of restimulated fracture half-length was determined to be 75 feet of infinite conductivity fracture half-length, which is equivalent to a skin factor of –5.0. So, if a predicted fracture half-length was already greater than 75 feet, no incremental recovery could be captured.

Figure 6 exhibits a typical restimulation candidate, the Eugene Chilcott #365 well. The predicted restimulation gas production stream is shown as a dark red line, yielding an estimated 5-year incremental production of 6.0 MMcf, which ranked it as the number nine candidate.

On the whole, the estimated 5-year incremental values were less than anticipated, ranging from no expected improvement to 10.5 MMcf. This was due, in part, to very little remaining differential pressure between the estimated average reservoir pressure and sandface flowing pressure. In many cases, the opportunity existed to determine the impact of a reduced sandface flowing pressure, which would act as added lease compression.

So, the impact of added lease compression was explored for each of the production wells. The lowest observed wellhead flowing pressure value, 50 psia, was assumed to be the best field practice. Therefore, only those wells producing at pressures larger than 50 psia were considered for improvement by this method.

Table 1 shows the top ten candidates for restimulation and added lease compression, while **Appendix B** contains each well's incremental computations for both applications. The results showed that the top ten lease compression candidates could garner an additional 27 MMcf of gas as compared to restimulation, suggesting it may be a more cost effective approach for the operator.

Discussion: Although the type curve matching results were quite reasonable, implementing quality review checks, such as a Voronoi well spacing comparison with drainage area, could have enhanced the outcome. Without an estimate of the Voronoi well spacing, this confirmation point was bypassed, resulting in a handful of wells with unreasonably large permeability, fracture half-length and drainage area values.

However, by assuming that the estimated drainage area and reservoir pressure values would be unchanged by restimulation, conservative expectations can be made of the incremental 5-year recovery. As shown in **Figure 6**, the expected production profile has a production decline consistent with the preceding matches. Also, the predicted peak gas rate, while an order of magnitude higher than the pre-restimulation rate, was found to be considerably less than the match immediately prior to it, as might be expected in a true "field application."

Although the expected incremental recoveries were found to be more realistic, especially in a qualitatively sense, the 5-year volumes were considerably less than what might have been expected. Therefore, it may be more cost effective for the operator to pursue production improvements by other means, such as added lease compression. As **Table 1** showed, the number one lease compression response was achieved at the William Willis #325 well, which predicted a 5-year recovery of about 25 MMcf by reducing the lease production pressure from 350 psia to 50 psia.

References

1 Pekot, L., Wozniak, D., Martin, J., "Tight Sand Evaluation Applied to the Medina Group of Chautauqua County, NY," SPE 57440, presented at the 1999 SPE Eastern Regional Meeting, Charleston, WV, 20-22 October, 1999.

2 Advanced Resources International, Inc., “Tight Sand Evaluation Applied to the Medina Group of Chautauqua County, NY,” Final Report for NYSERDA, Contract Number: 4715-ERTER-ER-98, 2000.

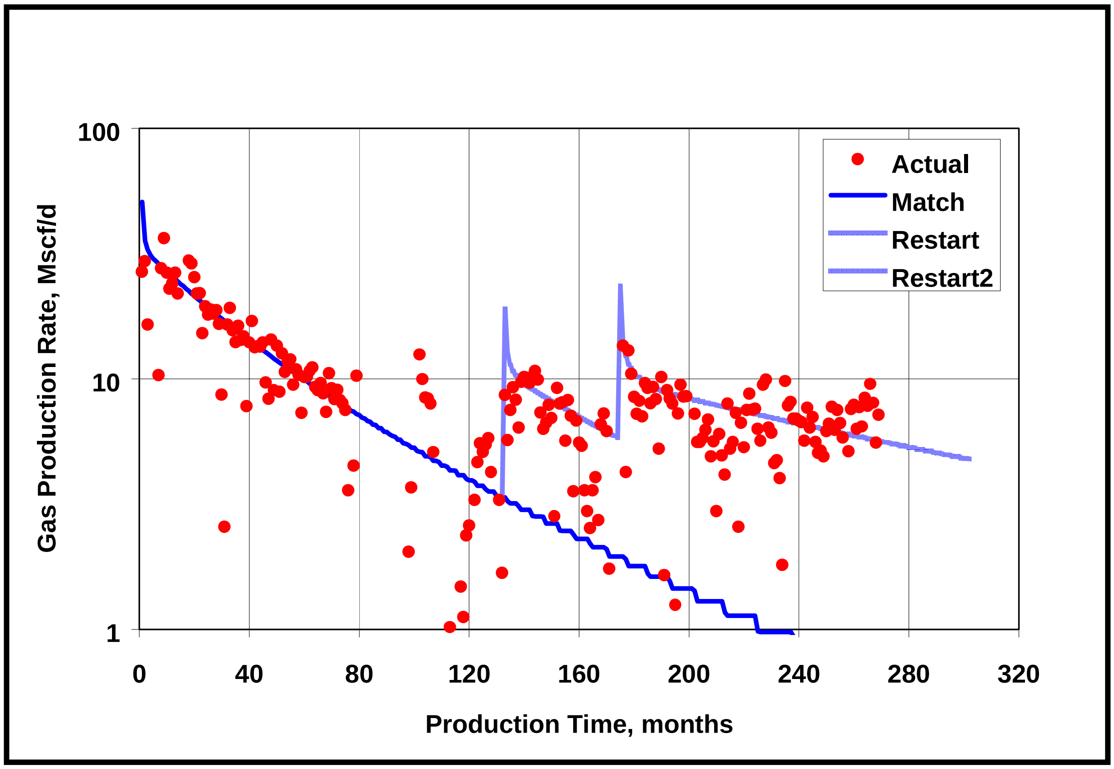


Figure 1 – Typical, Restarted, Production Type Curve Match

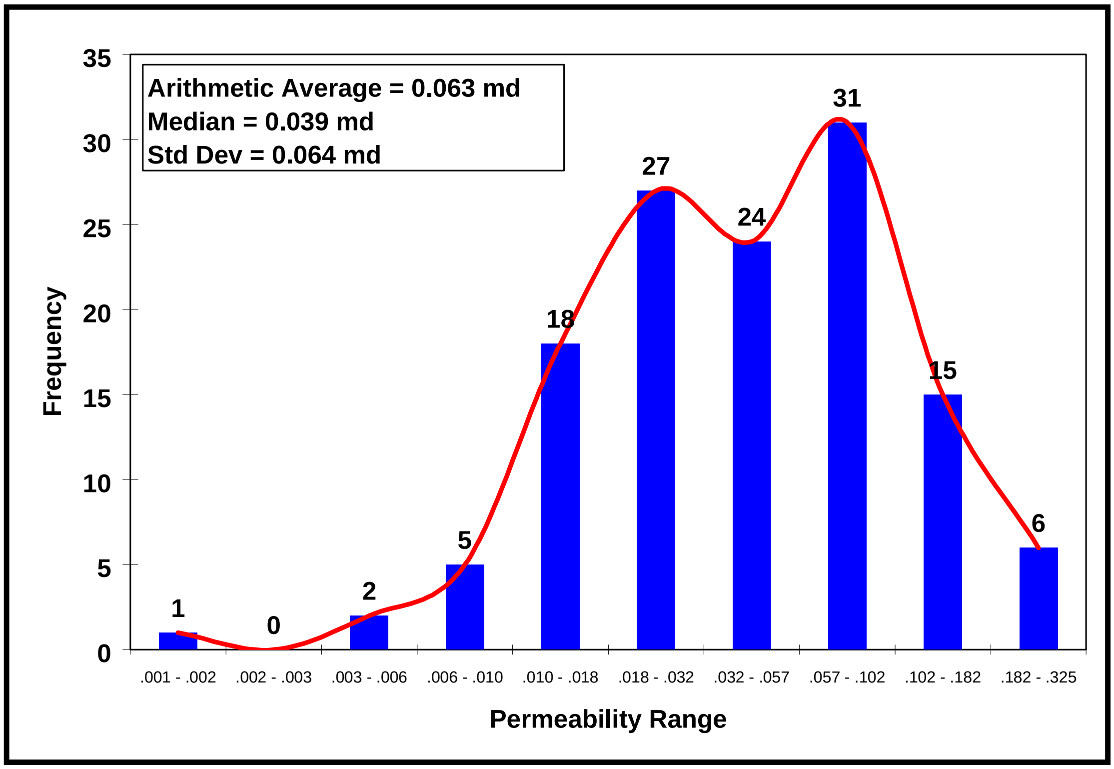


Figure 2 – Distribution of Estimated Permeability Results

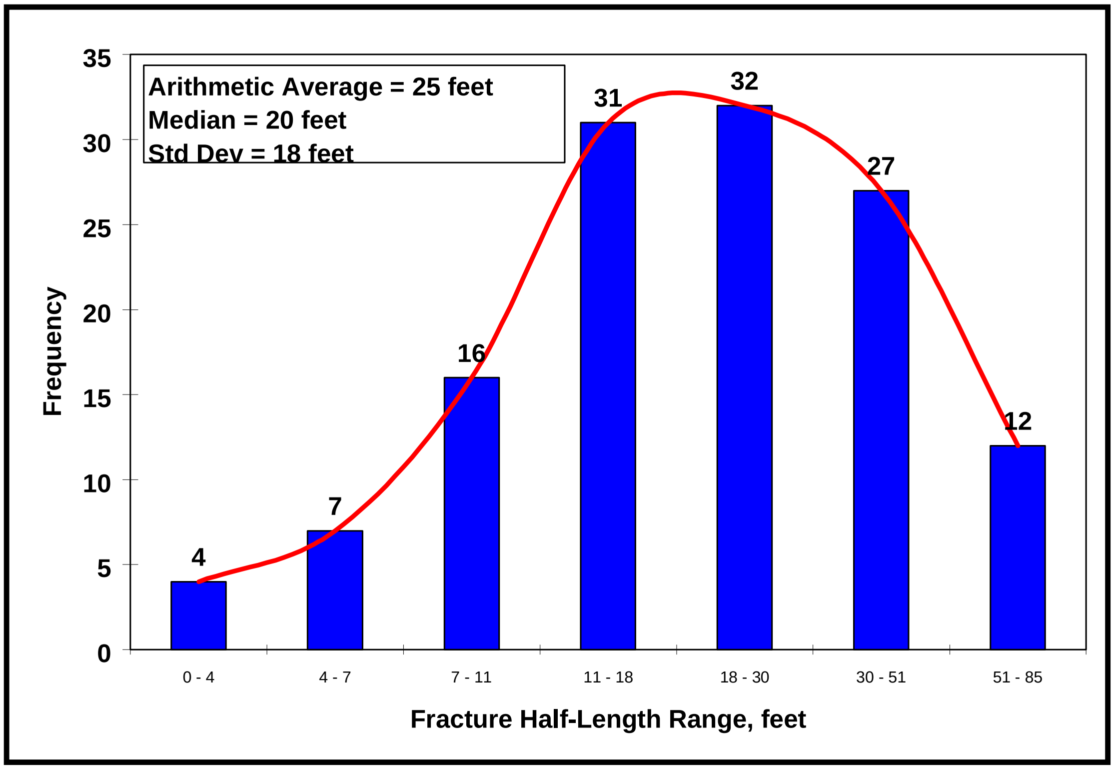


Figure 3 – Distribution of Estimated Fracture Half-Length Results

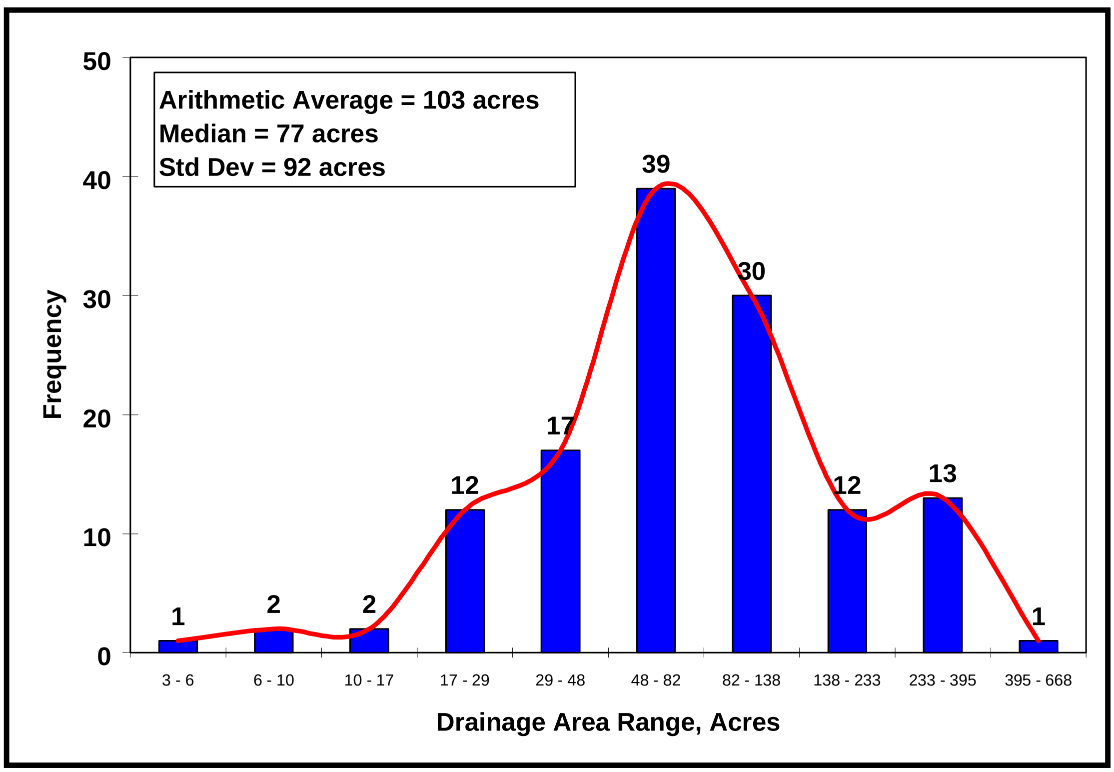


Figure 4 – Distribution of Estimated Drainage Area Results

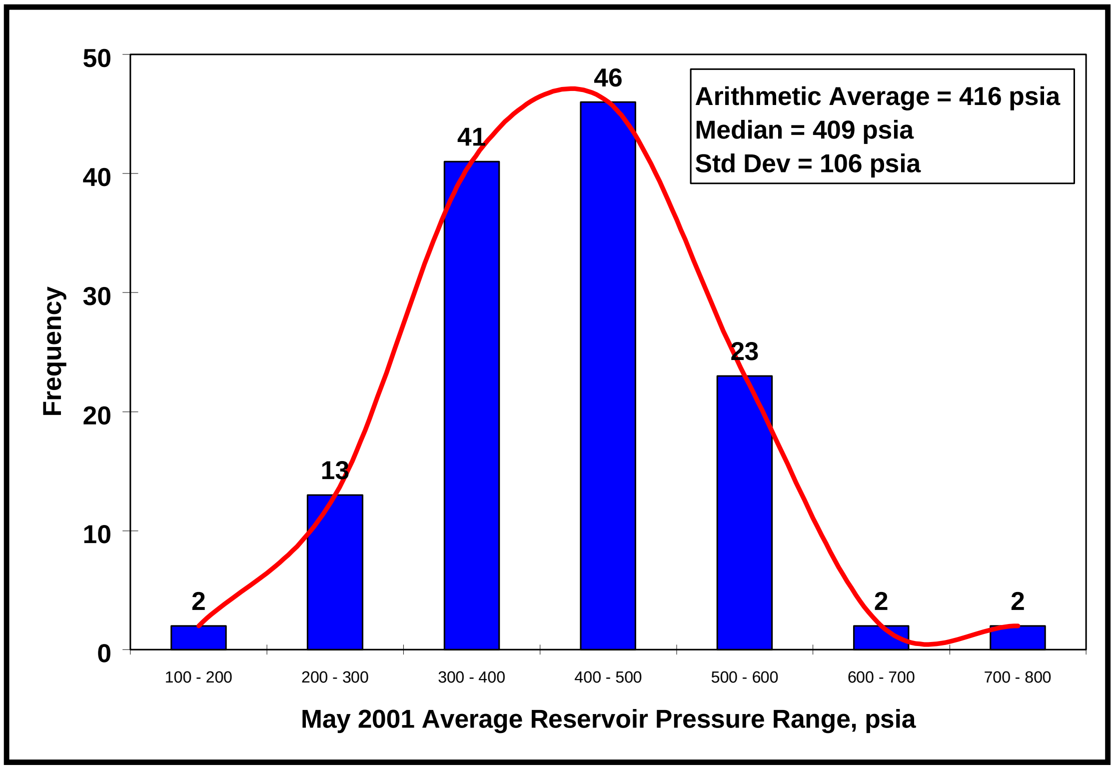


Figure 5 – Distribution of Estimated May 2001 Average Reservoir Pressure Results

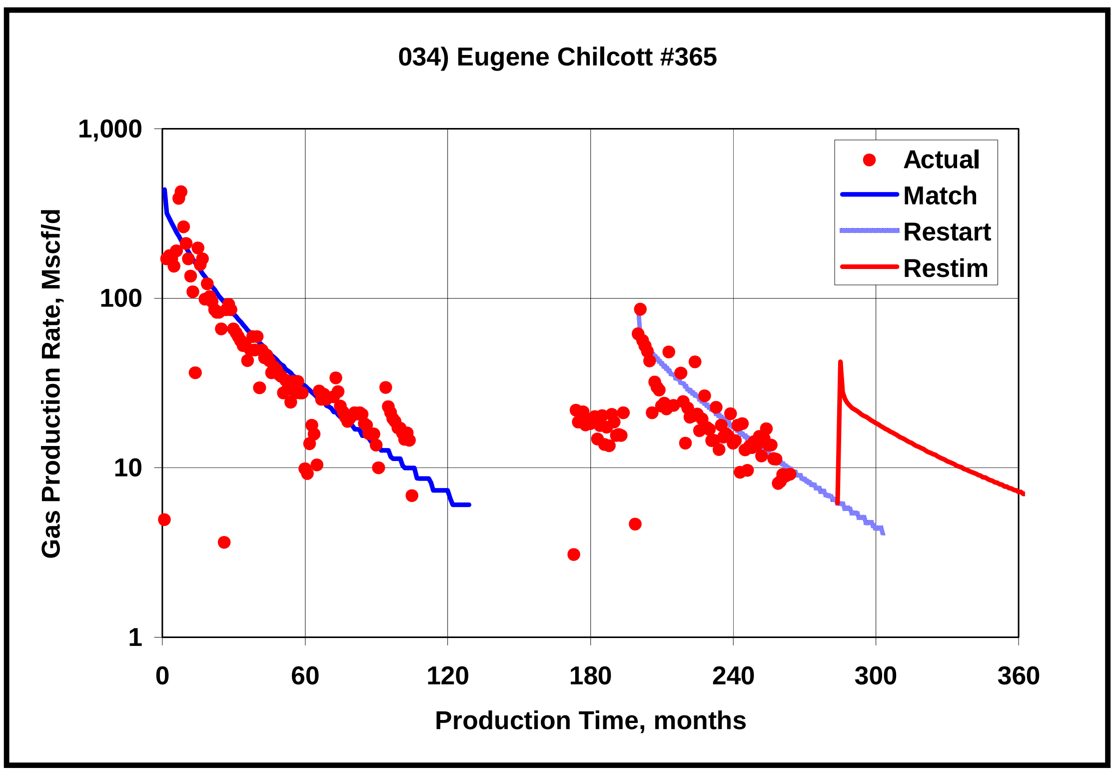


Figure 6 – Typical 5-Year Production Profile due to Restimulation

Table 1 – Predicted Top Ten Well Recoveries due to Restimulation and Added Lease Compression

a) Rankings by Incremental Recovery due to Restimulation

			TC Results					Restim	Inc.
Well #	Well Name	Formation	k md	Xf ft	A acres	5 Yr Cum Bcf	P, 5/01 psia	5 Yr Cum Bcf	5 Yr Cum MMcf
16	BEMUS, CECILE #214	Comingled	0.150	-	74	0.010	380	0.021	**10.5**
115	RAYNOR, WARD #323	Comingled	0.056	5	62	0.013	518	0.022	**9.4**
150	WELLMAN, DONALD #120	Comingled	0.007	-	21	0.003	716	0.011	**7.6**
146	VAN DETTE, ALBERT #356	Comingled	0.070	10	110	0.012	579	0.020	**7.5**
117	RENO, NORMAN #277	Comingled	0.055	7	131	0.011	467	0.019	**7.3**
87	MARRANO, ANTHONY #389	Comingled	0.023	25	73	0.007	570	0.015	**7.2**
110	POWELL, IRVING#077	Comingled	0.066	15	137	0.009	406	0.017	**7.2**
63	FURMANEK, ALOYSIOUS #144	Comingled	0.090	9	293	0.013	426	0.019	**6.1**
34	CHILCOTT, EUGENE #365	Comingled	0.325	20	104	0.022	283	0.028	**6.0**
149	WEBSTER CASTLE INN #015	Grimsby	0.140	15	104	0.015	365	0.021	**6.0**

b) Rankings by Incremental Recovery due to Added Lease Compression

			TC Results					Compression	Inc
Well #	Well Name	Formation	k md	Xf ft	A acres	5 Yr Cum Bcf	P, 5/01 psia	5 Yr Cum Bcf	5 Yr Cum MMcf
154	WILLS, WILLIAM #325	Comingled	0.150	15	62	0.002	366	0.028	**25.4**
97	MILLER, MORRIS #433	Comingled	0.210	35	109	0.005	279	0.027	**21.3**
94	MIKULA, JOSEPH #152	Grimsby	0.075	25	259	0.012	471	0.027	**15.0**
140	SUTTON, DAVID #315	Comingled	0.140	25	268	0.017	440	0.032	**14.7**
112	PRZYBYLSKI, LEONARD #113	Comingled	0.110	15	255	0.009	390	0.021	**12.5**
75	JOSEPHSON, WALFRED #329	Grimsby	0.090	15	109	0.005	423	0.017	**11.3**
145	VAN DETTE, ALBERT #339	Comingled	0.065	13	77	0.008	462	0.019	**10.7**
153	WILKENS, ROY #343	Comingled	0.100	17	133	0.013	407	0.021	**7.8**
52	DENNISON, WILBER #062	Comingled	0.090	20	133	0.008	359	0.016	**7.7**
82	LANFORD, C. #240	Comingled	0.105	14	43	0.004	303	0.011	**7.4**

Appendix A

Chautauqua County, New York Study Inputs and Results

NYSERDA Appendix A\Chautauqua Co, NY Type Curve Inputs and Results.xls

Key:

Assumed
No Petrophysical Data

Appendix B

Chautauqua County, New York Study Restimulation and Added Compression Results

NYSERDA Appendix B\Chautauqua County Restimulation Incrementals.xls