Surrogate Reservoir Model for Generating Pressure, Saturation and Compositional Characteristics at the Grid-Block Level; SACROC field, Scurry County, Texas

David Moreno
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation
Moreno, David, "Surrogate Reservoir Model for Generating Pressure, Saturation and Compositional Characteristics at the Grid-Block Level; SACROC field, Scurry County, Texas" (2014). Graduate Theses, Dissertations, and Problem Reports. 431.
https://researchrepository.wvu.edu/etd/431

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Surrogate Reservoir Model for Generating Pressure, Saturation and Compositional Characteristics at the Grid-Block Level; SACROC field, Scurry County, Texas

David Moreno
Thesis submitted
To the Benjamin M. Statler College of Engineering and Mineral Resources
At West Virginia University

In partial fulfillment of the requirements for the degree of
Master of Science
In
Petroleum and Natural Gas Engineering

Approved by
Prof. Sam Ameri
Dr. Shahab D. Mohaghegh, Ph.D.
Dr. Khashayar Aminian, Ph.D.
Dr. Ali Takbiri Borujeni, Ph.D.
Dr. Ebrahim Fathi, Ph.D

Morgantown, West Virginia, 2014
Keywords: Reservoir Simulation, Surrogate model, Proxy model, Artificial Intelligence, CO2 Enhanced oil recovery, water alternating gas, SACROC.
Abstract

Surrogate Reservoir Model for Generating Pressure, Saturation and Compositional Characteristics at the Grid-Block Level;

SACROC field, Scurry County, Texas

David Moreno

Nowadays, due to advancements in data acquisition technologies in oil and gas industry more data are available for generating reservoir simulation models. This leads to high fidelity reservoir simulation models which are highly complex and computationally expensive. The conventional reservoir management studies require hundreds realizations of the simulation models. As S. Gencer (2007) described, the reservoir simulation trend is towards “more: more users, more models, more cells, more wells, more cases, more data and more integration” [1]. In order to enhance the reservoir model descriptions, more computational power would have to be designed and engineered to keep up with our modeling needs; hence, creating an unsustainable cyclical process. Therefore, even with the advancements in the computational powers, the industry cannot take advantage of the full potential of these full-field reservoir simulation models.

Many studies have tried to create alternative methods in order to replicate the performance of full-field reservoir simulation models and at the same time decrease the cost of operation. Traditional proxy models, such as statistical based approaches, are examples of these studies. The degree of success, particularly practical aspects, for these approaches remains to be argued.

As an alternative to traditional proxy modeling methods, the objective of this study is to investigate the feasibility of use of a fast intelligent approximation of the numerical simulation model. This replica will accurately reproduce dynamic reservoir properties of complex full-field numerical simulation models in matter of seconds. A Grid-based Surrogate Reservoir Model (GSRM) is developed based on data-driven and Artificial Intelligence techniques. This technology is able to learn from the provided examples of the reservoir simulation model. The robustness of this technology is validated by testing it on non-seen instances. Finally the trained and validated GSRM will produce the results of full-field simulation models accurately and in a very short time (seconds).

This concept will be proven by building a GSRM of CO2 injection–EOR numerical model of SACROC field, Scurry County, Texas. The SACROC model (CMG GEM) in use was previously generated and history matched by the Petroleum Engineering & Analytics Research Lab - PEARL - at West Virginia University, it is based on a comprehensive geological study that includes 3D seismic surveys and well logs; in order to generate the GSRM this model is to be ran using multiple injection scenarios that will create an appropriate solution space so we can comprehend and grasp its behavior using artificial intelligence.
Objective

The objective of this study is to assess the feasibility of using state of the art data-driven proxy models to effectively replicate dynamic reservoir properties of complex full field numerical models, by building and validating a grid block level Surrogate Reservoir Model (GSRM) of a macro-scale complex full field CO2 injection –EOR numerical model.
Dedication

To Astrid, without you, this achievement would not be possible.

To my Family, for backing me up in every step of the way.

To Daniel, thanks for bringing me into this and carrying me all the way. Thanks for being my Brother.

Gracias totales...
# Table of Contents

Abstract ...................................................................................................................................................... ii  
Objective .................................................................................................................................................. iii  
Dedication ........................................................................................................................................ iv  
Table of Contents ...................................................................................................................................... v  
Tables index ................................................................................................................................................ vii  
Figures index ........................................................................................................................................ vii  
Equation index ........................................................................................................................................ xxiii  
Chapter 1 Introduction .............................................................................................................................. 1  
  Problem Statement ................................................................................................................................ 1  
  SACROC Unit Field Background .......................................................................................................... 5  
    Field History ....................................................................................................................................... 5  
    Geologic Description ............................................................................................................................ 6  
  Numerical Reservoir Simulation Model (PEARL WVU) ........................................................................ 8  
Chapter 2 - Literature Review .................................................................................................................. 11  
  Reservoir Engineering: CO2 –EOR (WAG) ......................................................................................... 11  
  Artificial Neural Networks .................................................................................................................... 12  
    Definition ........................................................................................................................................ 12  
    Supervised Learning ........................................................................................................................... 15  
    General Regression Neural Networks (GRNN) ................................................................................... 17  
    Back propagation Neural Networks ................................................................................................... 19  
  Reservoir Simulation ............................................................................................................................... 21  
  Proxy Reservoir models at grid block level ........................................................................................... 23  
    Grid Surrogate Reservoir Model (GSRM) ........................................................................................ 24
Tables index

Table 1 - Initial oil composition in the upscaled geocellular model (Evaluation and Design of a CO2 Miscible Flood Project - SACROC Unit, Kelly Snider Field, 1973) ................................................................. 10

Table 2 Parameters of biological and artificial neural networks. [22] .......................................................... 15

Table 3 - Independent Variables (Scenario Experimental Design) ......................................................... 27

Table 4 - Dependent Variables (Scenario Experimental Design) ............................................................. 27

Table 5 - Scenario Design .......................................................................................................................... 28

Table 6 - Initial Features extracted from Simulator extracted Data .......................................................... 32

Table 7 - Secondary Features Generated from Simulator extracted Data ................................................. 33

Table 8 - Averaging methods used for different features in Layer Lumping process ............................... 34

Table 9 - Intersection with Well databases Features generated description ............................................. 35

Table 10 - R2 coefficient of all NN's generated ......................................................................................... 43

Table 11 - GSRM all cases output: Statistics analysis & Uncertainty Assessment summary (P10, P50, and P90) .................................................................................................................. 77

Figures index

Figure 1 Data collection processes in Reservoir Management. [2] ............................................................ 1

Figure 2 GOCAD 2009 times for three large reservoir models in HPC facility [4] ................................. 2

Figure 3 SACROC unit, Canyon Reef formation reservoir map [11] ......................................................... 5

Figure 4. SACROC Unit at the Horseshoe Atoll in west Texas and structural contours map of top of carbonate reef modified from Stafford . Contours are m scale [11]. ................................................................. 7

Figure 5. A structural and stratigraphic cross-section of profile A-A’, located within the SACROC northern platform [14]. See Figure 4 for a location of profile A-A’. ................................................................. 7

Figure 6 - Well logs representing SACROC northern Platform [11] .............................................................. 7

Figure 7 - Geological model for this study (Axes scale in m) - Left: Porosity Distribution map, Right: Permeability distribution map [6]........................................................................................................ 9

Figure 8 - Property distribution Histograms. Left: Porosity, Right: Permeability [6] ............................... 9
Figure 9 - Reservoir model well distribution.................................................................10

Figure 10- Schematic diagram of a water-alternating-gas (WAG) miscible CO2 EOR operation [18] ........12

Figure 11 Schematic of a Generic feed-forward neural network with a single hidden layer. ..................13

Figure 12 Schematic of a biological neuron [22] ........................................................................14

Figure 13 Schematic of artificial neuron [22].............................................................................15

Figure 14 Supervised learning process [22].................................................................................16

Figure 15 - Schematic of a General Back propagation Neural Network [25].....................................20

Figure 16 - Backpropagation Neural Network Algorithm schematic .............................................20

Figure 17 - CPU time consumption of each process in a reservoir simulator [4] ..............................23

Figure 18 - General Methodology Flowchart..................................................................................25

Figure 19 - *.INC CMG output simulator example (Screenshot) ......................................................30

Figure 20 - *.XLSX / *.NPY Python Console output example (Screenshot) ...................................30

Figure 21 – Dataset matrixes transformations trough the data management process. .......................32

Figure 22 - Tier system definition .................................................................................................36

Figure 23 - Sampling methods comparison. ....................................................................................38

Figure 24 - Input classification for Neural Network Architecture design. ........................................39

Figure 25 - KPI performance Analysis Screenshot and Neural Net design (ISI-IDEA 2013). ..............39

Figure 26 - Neural Network Training Screenshot and Live results analysis (ISI-IDEA 2013) ..........40

Figure 27 - Cascading Design. .......................................................................................................41

Figure 28 - Final input selection.......................................................................................................42

Figure 29 - Best Training case prediction, Pressure Model, Year 1 (2006). .................................44

Figure 30 - Best Training case prediction, SO Model, Year 1 (2006).............................................45

Figure 31 - Best Training case prediction, SW Model, Year 1 (2006). ..........................................46

Figure 32 - Best Training case prediction, CO2 Model, Year 1 (2006). .........................................47

Figure 33 - Best Training case prediction, Pressure Model, Year 5 (2010). .................................48

Figure 34 - Best Training case prediction, SO Model, Year 5 (2010) ............................................49
Figure 35 - Best Training case prediction, SW Model, Year 5 (2010)................................. 50
Figure 36 - Best Training case prediction, CO2 Model, Year 5 (2010)........................................ 51
Figure 37 - Worst Training case prediction, Pressure Model, Year 1 (2006)............................ 52
Figure 38 - Worst Training case prediction, SO Model, Year 1 (2006)........................................ 53
Figure 39 - Worst Training case prediction, SW Model, Year 1 (2006)........................................ 54
Figure 40 - Worst Training case prediction, CO2 Model, Year 1 (2006)........................................ 55
Figure 41 - Worst Training case prediction, Pressure Model, Year 5 (2010)............................ 56
Figure 42 - Worst Training case prediction, SO Model, Year 5 (2010)........................................ 57
Figure 43 - Worst Training case prediction, SW Model, Year 5 (2010)........................................ 58
Figure 44 - Worst Training case prediction, CO2 Model, Year 5 (2010)........................................ 59
Figure 45 - Best Blind case prediction, Pressure Model, Year 1 (2006)................................. 60
Figure 46 - Best Blind case prediction, SO Model, Year 1 (2006)........................................ 61
Figure 47 - Best Blind case prediction, SW Model, Year 1 (2006)........................................ 62
Figure 48 - Best Blind case prediction, CO2 Model, Year 1 (2006)........................................ 63
Figure 49 - Best Blind case prediction, Pressure Model, Year 5 (2010)................................. 64
Figure 50 - Best Blind case prediction, SO Model, Year 5 (2010)........................................ 65
Figure 51 - Best Blind case prediction, SW Model, Year 5 (2010)........................................ 66
Figure 52 - Best Blind case prediction, CO2 Model, Year 5 (2010)........................................ 67
Figure 53 - Worst Blind case prediction, Pressure Model, Year 1 (2006)............................ 68
Figure 54 - Worst Blind case prediction, SO Model, Year 1 (2006)........................................ 69
Figure 55 - Worst Blind case prediction, SW Model, Year 1 (2006)........................................ 70
Figure 56 - Worst Blind case prediction, CO2 Model, Year 1 (2006)........................................ 71
Figure 57 - Worst Blind case prediction, Pressure Model, Year 5 (2010)............................ 72
Figure 58 - Worst Blind case prediction, SO Model, Year 5 (2010)........................................ 73
Figure 59 - Worst Blind case prediction, SW Model, Year 5 (2010)........................................ 74
Figure 60 - Worst Blind case prediction, CO2 Model, Year 5 (2010)........................................ 75
Figure 113- GSRM Results, SO, year-case: 2008-3 ................................................................. 144
Figure 114- GSRM Results, SW, year-case: 2008-3 ................................................................. 145
Figure 115- GSRM Results, CO2, year-case: 2009-3 ................................................................. 146
Figure 116- GSRM Results, PRESS, year-case: 2009-3 ............................................................. 147
Figure 117- GSRM Results, SO, year-case: 2009-3 ................................................................. 148
Figure 118- GSRM Results, SW, year-case: 2009-3 ................................................................. 149
Figure 119- GSRM Results, CO2, year-case: 2010-3 ................................................................. 150
Figure 120- GSRM Results, PRESS, year-case: 2010-3 ............................................................ 151
Figure 121- GSRM Results, SO, year-case: 2010-3 ................................................................. 152
Figure 122- GSRM Results, SW, year-case: 2010-3 ................................................................. 153
Figure 123- GSRM Results, CO2, year-case: 2006-4 ................................................................. 154
Figure 124- GSRM Results, PRESS, year-case: 2006-4 ............................................................ 155
Figure 125- GSRM Results, SO, year-case: 2006-4 ................................................................. 156
Figure 126- GSRM Results, SW, year-case: 2006-4 ................................................................. 157
Figure 127- GSRM Results, CO2, year-case: 2007-4 ................................................................. 158
Figure 128- GSRM Results, PRESS, year-case: 2007-4 ............................................................ 159
Figure 129- GSRM Results, SO, year-case: 2007-4 ................................................................. 160
Figure 130- GSRM Results, SW, year-case: 2007-4 ................................................................. 161
Figure 131- GSRM Results, CO2, year-case: 2008-4 ................................................................. 162
Figure 132- GSRM Results, PRESS, year-case: 2008-4 ............................................................ 163
Figure 133- GSRM Results, SO, year-case: 2008-4 ................................................................. 164
Figure 134- GSRM Results, SW, year-case: 2008-4 ................................................................. 165
Figure 135- GSRM Results, CO2, year-case: 2009-4 ................................................................. 166
Figure 136- GSRM Results, PRESS, year-case: 2009-4 ............................................................ 167
Figure 137- GSRM Results, SO, year-case: 2009-4 ................................................................. 168
Figure 138- GSRM Results, SW, year-case: 2009-4 ................................................................. 169
Figure 191- GSRM Results, CO2, year: 2008-7
Figure 192- GSRM Results, PRESS, year: 2008-7
Figure 193- GSRM Results, SO, year: 2008-7
Figure 194- GSRM Results, SW, year: 2008-7
Figure 195- GSRM Results, CO2, year: 2009-7
Figure 196- GSRM Results, PRESS, year: 2009-7
Figure 197- GSRM Results, SO, year: 2009-7
Figure 198- GSRM Results, SW, year: 2009-7
Figure 199- GSRM Results, CO2, year: 2010-7
Figure 200- GSRM Results, PRESS, year: 2010-7
Figure 201- GSRM Results, SO, year: 2010-7
Figure 202- GSRM Results, SW, year: 2010-7
Figure 203- GSRM Results, CO2, year: 2006-8
Figure 204- GSRM Results, PRESS, year: 2006-8
Figure 205- GSRM Results, SO, year: 2006-8
Figure 206- GSRM Results, SW, year: 2006-8
Figure 207- GSRM Results, CO2, year: 2007-8
Figure 208- GSRM Results, PRESS, year: 2007-8
Figure 209- GSRM Results, SO, year: 2007-8
Figure 210- GSRM Results, SW, year: 2007-8
Figure 211- GSRM Results, CO2, year: 2008-8
Figure 212- GSRM Results, PRESS, year: 2008-8
Figure 213- GSRM Results, SO, year: 2008-8
Figure 214- GSRM Results, SW, year: 2008-8
Figure 215- GSRM Results, CO2, year: 2009-8
Figure 216- GSRM Results, PRESS, year: 2009-8
Figure 217- GSRM Results, SO, year-case: 2009-8 ................................................................. 248
Figure 218- GSRM Results, SW, year-case: 2009-8 ................................................................. 249
Figure 219- GSRM Results, CO2, year-case: 2010-8 ................................................................. 250
Figure 220- GSRM Results, PRESS, year-case: 2010-8 .............................................................. 251
Figure 221- GSRM Results, SO, year-case: 2010-8 ................................................................. 252
Figure 222- GSRM Results, SW, year-case: 2010-8 ................................................................. 253
Figure 223- GSRM Results, CO2, year-case: 2006-9 ................................................................. 254
Figure 224- GSRM Results, PRESS, year-case: 2006-9 .............................................................. 255
Figure 225- GSRM Results, SO, year-case: 2006-9 ................................................................. 256
Figure 226- GSRM Results, SW, year-case: 2006-9 ................................................................. 257
Figure 227- GSRM Results, CO2, year-case: 2007-9 ................................................................. 258
Figure 228- GSRM Results, PRESS, year-case: 2007-9 .............................................................. 259
Figure 229- GSRM Results, SO, year-case: 2007-9 ................................................................. 260
Figure 230- GSRM Results, SW, year-case: 2007-9 ................................................................. 261
Figure 231- GSRM Results, CO2, year-case: 2008-9 ................................................................. 262
Figure 232- GSRM Results, PRESS, year-case: 2008-9 .............................................................. 263
Figure 233- GSRM Results, SO, year-case: 2008-9 ................................................................. 264
Figure 234- GSRM Results, SW, year-case: 2008-9 ................................................................. 265
Figure 235- GSRM Results, CO2, year-case: 2009-9 ................................................................. 266
Figure 236- GSRM Results, PRESS, year-case: 2009-9 .............................................................. 267
Figure 237- GSRM Results, SO, year-case: 2009-9 ................................................................. 268
Figure 238- GSRM Results, SW, year-case: 2009-9 ................................................................. 269
Figure 239- GSRM Results, CO2, year-case: 2010-9 ................................................................. 270
Figure 240- GSRM Results, PRESS, year-case: 2010-9 .............................................................. 271
Figure 241- GSRM Results, SO, year-case: 2010-9 ................................................................. 272
Figure 242- GSRM Results, SW, year-case: 2010-9 ................................................................. 273
Figure 269- GSRM Results, SO, year-case: 2007-11
Figure 270- GSRM Results, SW, year-case: 2007-11
Figure 271- GSRM Results, CO2, year-case: 2008-11
Figure 272- GSRM Results, PRESS, year-case: 2008-11
Figure 273- GSRM Results, SO, year-case: 2008-11
Figure 274- GSRM Results, SW, year-case: 2008-11
Figure 275- GSRM Results, CO2, year-case: 2009-11
Figure 276- GSRM Results, PRESS, year-case: 2009-11
Figure 277- GSRM Results, SO, year-case: 2009-11
Figure 278- GSRM Results, SW, year-case: 2009-11
Figure 279- GSRM Results, CO2, year-case: 2010-11
Figure 280- GSRM Results, PRESS, year-case: 2010-11
Figure 281- GSRM Results, SO, year-case: 2010-11
Figure 282- GSRM Results, SW, year-case: 2010-11
Figure 283- GSRM Results, CO2, year-case: 2006-12
Figure 284- GSRM Results, PRESS, year-case: 2006-12
Figure 285- GSRM Results, SO, year-case: 2006-12
Figure 286- GSRM Results, SW, year-case: 2006-12
Figure 287- GSRM Results, CO2, year-case: 2007-12
Figure 288- GSRM Results, PRESS, year-case: 2007-12
Figure 289- GSRM Results, SO, year-case: 2007-12
Figure 290- GSRM Results, SW, year-case: 2007-12
Figure 291- GSRM Results, CO2, year-case: 2008-12
Figure 292- GSRM Results, PRESS, year-case: 2008-12
Figure 293- GSRM Results, SO, year-case: 2008-12
Figure 294- GSRM Results, SW, year-case: 2008-12
Figure 373- GSRM Results, SO, year-case: 2008-16

Figure 374- GSRM Results, SW, year-case: 2008-16

Figure 375- GSRM Results, CO2, year-case: 2009-16

Figure 376- GSRM Results, PRESS, year-case: 2009-16

Figure 377- GSRM Results, SO, year-case: 2009-16

Figure 378- GSRM Results, SW, year-case: 2009-16

Figure 379- GSRM Results, CO2, year-case: 2010-16

Figure 380- GSRM Results, PRESS, year-case: 2010-16

Figure 381- GSRM Results, SO, year-case: 2010-16

Figure 382- GSRM Results, SW, year-case: 2010-16

Figure 383- GSRM Results, CO2, year-case: 2006-23

Figure 384- GSRM Results, PRESS, year-case: 2006-23

Figure 385- GSRM Results, SO, year-case: 2006-23

Figure 386- GSRM Results, SW, year-case: 2006-23

Figure 387- GSRM Results, CO2, year-case: 2007-23

Figure 388- GSRM Results, PRESS, year-case: 2007-23

Figure 389- GSRM Results, SO, year-case: 2007-23

Figure 390- GSRM Results, SW, year-case: 2007-23

Figure 391- GSRM Results, CO2, year-case: 2008-23

Figure 392- GSRM Results, PRESS, year-case: 2008-23

Figure 393- GSRM Results, SO, year-case: 2008-23

Figure 394- GSRM Results, SW, year-case: 2008-23

Figure 395- GSRM Results, CO2, year-case: 2009-23

Figure 396- GSRM Results, PRESS, year-case: 2009-23

Figure 397- GSRM Results, SO, year-case: 2009-23

Figure 398- GSRM Results, SW, year-case: 2009-23
Figure 399 - GSRM Results, CO₂, year-case: 2010-23 ................................................................. 430
Figure 400 - GSRM Results, PRESS, year-case: 2010-23 ............................................................ 431
Figure 401 - GSRM Results, SO, year-case: 2010-23 ................................................................. 432
Figure 402 - GSRM Results, SW, year-case: 2010-23 ................................................................. 433

Equation index

Equation 1 - Total sum of squares (observed output) .................................................................... 16
Equation 2 – Regression sum of squares (simulated output) ........................................................ 17
Equation 3 - R-square general definition ..................................................................................... 17
Equation 4 - General Regression Neural Network algorithm [23] ............................................. 18
Petroleum engineering as of today, due to the contribution of several disciplines (Mechanical, Electrical, Chemical, Geoscientists) and a critical advancement of technology in the electronics field over the past decades, has augmented their petroleum reservoir descriptions in orders of magnitudes when to data volumes comes. An enormous amount of multidisciplinary data is acquired throughout the life of a reservoir.

This process starts with exploration and continues through discovery, and is followed by delineation, development, production, and finally, abandonment. A realistic reservoir description, which is characterized by utilizing this data, is of vital importance for successful management of the reservoir. [2]

The fundamental classification of reservoir data according to their AOR and impact on final volume of data, according to Satter-2006, include the following:

- Microscopic: Microphotographs showing sand grains and pores from the cuttings or core slices
- Macroscopic: Core data providing basic rock properties such as porosity, permeability, capillary pressure and relative permeability, etc.
- Megascopic: Data from the Drainage area around the well, such as porosity, and fluid saturation from well logs.
- Gigascopic: Data from the entire reservoir, such as seismic surveys, well production and Injection, well pressure tests.

The collection of this data is of great importance, the processes involved in this step are as follows:

![Figure 1 Data collection processes in Reservoir Management. [2]](image-url)
This data is prepared and represented via commercial geological simulators and applied in complex numerical simulation processes, in order to accomplish high fidelity Geological Simulation Models, more often having millions of cells with large quantity of data inherent within each one of these particles; the computers used for these type of models are multi-cluster specialized processing facilities that represent the associated high costs of today’s simulation industry.

Later on, these models are upscaled/coarsened to be more manageable, this is a common practice in reservoir simulation for reducing the model size to improve computational efficiency. As part of the standard simulation workflow, engineers upscale geological models to simulation models to reduce requirement on computational resources. Coarsening is also an attractive approach for reduced order modeling. Reduced order modeling plays a crucial role in uncertainty quantification, history matching, optimization and other applications where a large number of simulation models need to be evaluated [3]; a fundamental downside to this proxy modeling method is that it compromises the non-linearity of reservoir dynamics.

These final Reservoir Simulation models, often composed by millions of cells, take enormous amounts of CPU time and power to make single runs; the following chart is a benchmarking comparison for GOCAD 2009 (Saudi Aramco’s in-house simulator) with the use of a HPC (High Performance Computing Facilities):

<table>
<thead>
<tr>
<th>Task</th>
<th>106 M</th>
<th>425 M</th>
<th>850 M</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D Visualization</td>
<td>10 frame/sec</td>
<td>1 frame/sec</td>
<td>0.01 frame/sec</td>
</tr>
<tr>
<td>Processing</td>
<td>2 hours</td>
<td>5 hours</td>
<td>12 hours</td>
</tr>
<tr>
<td>Export</td>
<td>1 hour</td>
<td>4.5 hours</td>
<td>7.5 hours</td>
</tr>
</tbody>
</table>

Figure 2 GOCAD 2009 times for three large reservoir models in HPC facility [4].

Therefore Numerical simulation of petroleum reservoirs is a data-intensive and computational resources consuming activity, the trend in reservoir simulation is towards “more: more users, more models, more cells, more wells, more cases, more data and more integration” [5], for our technological advancement more specialized HPC’s would have to be designed and engineered in order to keep up with our simulation needs as an industry.

The reliance on supercomputing facilities and parallel processing techniques (HPC-High Performance Computing facilities) to perform uncertainty analyses, sensitivity analyses and optimization analyses, that take hundreds of runs to perform using high performance simulations is impractical, since the
computational times exceed in all cases the physical time of the process occurring, rendering virtually useless the numerical models in predicting the course of events or aiding a reservoir management team on making “real-time or near real-time decisions” [1].

Under this premise, instantaneous or near instantaneous simulation results are needed in order to better manage reservoirs in real-life, alternative methods surface to mitigate the long processing hours and impracticality of simulation itself, these alternative methods are now possible due to the emergence of intelligent systems and new computing techniques.

It is possible to build Surrogate Intelligent Models (SIM’s) that can mimic functionalities of complex simulators in real-time; conventional approaches are based in a geo-statistical focus that is extensively proven to be able to interpolate between simulation scenarios, but not so able to effectively extrapolate for further simulation scenarios, rendering forecasting capabilities useless in the conventional approaches.

One such method that is used quite often is Response Surfaces, they are statistical interpolations (based on fitting some type of pre-determined model – linear or quadratic –) of model responses to different geological, geophysical and petro-physical realizations. Another method that has been used more in other industries is Reduced order Models. Reduced order Models are approximations of full three dimensional numerical simulation models that essentially approach an analytical model for tractability [6]. Both of this technologies had been discussed extensively and usually discarded as impractical, due to their dimensionality dependence problems, described often as the dimensionality curse [7] of proxy modeling due to the large solution space that has to be generated from, often hundreds, reservoir simulations.

As an alternative to traditional methods, the objective of this study is to discuss and assess the feasibility of the application of Artificial Neural Networks (ANNs) as proxy models for reservoir simulation of full field complex numerical models. Fast and reliable estimation of reservoir’s behavior can be achieved by training ANNs to mimic the behavior and response of the reservoir model to changes in input parameters. These ANNs, which are called grid block level Surrogate Reservoir Models (GSRM), are prototypes of the full-field reservoir simulation models. An intelligent simulation of the numerical simulation that could effectively reproduce dynamic reservoir properties of complex full field numerical models, based on data-driven technologies and Artificial Intelligence techniques. The high speed of GSRM in generating results is an important feature that enables us to perform optimization processes more efficiently and would potentially bring back practicality to these types of studies in the industry.
This concept will be proven by building a grid block level Surrogate Reservoir Model (SRM) of a macro-scale complex full field CO2 injection–EOR numerical model of SACROC field, Scurry county, Texas and validating our experiment using blind runs verification; the SACROC model (CMG GEM-Generalized Equation of State Model) in use was previously generated and history matched by the Petroleum Engineering & Analytics Research Lab - PEARL - at West Virginia University, it is based on a comprehensive geological study that includes 3D seismic surveys and well logs; in order to generate the GSRM this model is to be ran using multiple injection scenarios so we can comprehend and grasp its behavior using artificial intelligence.

The GSRM development of complex CO2-EOR water alternating gas fields has never been performed; more information on GSRM development for other type of fields (Ex: CO2 sequestration and Storage), which is a recently developed technology, can be found in literature (S. H. Mohaghegh)(e. a. Mohaghegh, Design Optimum Frac Jobs Using Virtual Intelligence Techniques)(e. a. Mohaghegh, Development of Surrogate Reservoir Models (SRM) for Fast Track Analysis of Complex Reservoirs).

This work presents a method for the construction of this grid block level Surrogate Reservoir Model (SRM) for complex field configurations numerical Reservoir simulation models, the methodology here is in the proof of concept stage, further analysis is required to improve this technique and to identify its strengths and limitations in more detail.
SACROC Unit Field Background

The field subject of this study is the northern platform of the Scurry Area Canyon Reef Operators Committee (SACROC) unit, located in Scurry County, Texas. SACROC unit was formed by original operators and royalty owners of the Scurry County Kelly-Sneider Field operations group in January 1951, in order to enhance ultimate oil recovery of the reservoir and study the effects of pressure maintenance.

The Reservoir is the Canyon Reef formation located in the Horseshoe Atoll, one of the largest subsurface limestone reef mounds in the world. As shown in figure 3, The atoll is 175 miles long, stretching beneath western Kent and western Scurry counties before turning south under Borden and Howard. Although several fields produce from the Canyon Reef formation in the Horseshoe Atoll, Kelly-Snyder is the most prolific [9]. Its cumulative production at the beginning of 2013 was approximately 1.4 billion barrels of oil [10].

Field History

The Kelly-Sneider Field, discovered in 1948, is one of the major oil reservoirs in the US, having approximately 2.75 billion bbl. of oil originally in place. The early performance of this reservoir suggested that only 20% of its OOIP was recoverable due to solution gas drive characterization as its primary production mechanism. Once the SACROC unit was formed, a massive waterflood program was started in the center-line row of wells along the longitudinal axis of the reservoir.

Later on, in 1968, a technical committee recommended that carbon dioxide (CO2) be used to miscibly displace the oil in the non-water flooded portion of the reservoir and that a pattern injection program be
developed in this area to implement the slug process and improve ultimate oil recovery. CO2 injection begun at early 1972. [12]

To predict the CO2 miscible flood performance and incremental reserves, the Engineering committee developed a layered calculation procedure. The model considered that the CO2 would be injected as a slug and would be driven through the reservoir by a trailing water flood (Water Alternating Gas method - WAG). [13]

Geologic Description

SACROC unit is located in SE section of the Horseshoe Atoll in the Midland basin of West Texas, comprises an area of 356 Km2, and elongated anticline with a hinge (E-W) of approximately 40 Km and limbs(N-S) of approximately 3-15 Km (in its wider point).

SACROC is a carbonate reef complex composed of great amounts of bedded bioclastic limestone and thin shale beds that represent the Canyon, Straw and Cysco formations of the Pennsylvanian period (Carboniferous Period), and the Wolfcamp Series of the lower Permian period [14]. The Cisco and Canyon Formation are the ones selected for most of the CO2 injection, these are mostly composed of limestone, minor amounts of sand, anhydrite and shale are present locally [15]. The Wolfcamp shale acts as a seal (confining unit) above the Canyon and Cisco Formations, this shale was studied with X-ray diffraction compositional analysis which determined that is mostly illite and quartz with minor carbonate, pyrite, feldspar [15]. Based on this mineralogical analysis and on the measured vertical permeability [16] suggested to be low enough to build a capillary barrier that hampers buoyancy-driven migration of CO2. Furthermore, their mineralogical and isotopic (oxygen and carbon) analyses suggested that the carbonates in the Wolfcamp Shale Formation appear to be derived from primarily diagenetic processes, supporting that CO2 is effectively trapped in the Cisco and Canyon Groups. [11].
Figure 4. SACROC Unit at the Horseshoe Atoll in west Texas and structural contours map of top of carbonate reef modified from Stafford. Contours are m scale [11].

Figure 5. A structural and stratigraphic cross-section of profile A-A’, located within the SACROC northern platform [14]. See Figure 4 for a location of profile A-A’.

Figure 6 - Well logs representing SACROC northern Platform [11]
Numerical Reservoir Simulation Model (PEARL WVU)

The initial high resolution geocellular model describes the Cysco and Canyon formations and has a staggering amount of 9,450,623 blocks, its approximate size is of 13,000 ft wide and 26,000 ft long. The top of the geocellular model describes the top configuration of the Cisco Formation, which is below approximately 3,600 ft depth, and the bottom of the model describes the bottom configuration of the Canyon Formation, which is below a depth of approximately 4,200 ft. This model encompasses a broad range of studies performed to the SACROC unit, such as but not limited to, core data, well-log data, sedimentologic and stratigraphic interpretations, analysis of seismic attributes and analysis of rock physics data; This initial high resolution geocellular model detailed the heterogeneity to help better understand the subsurface migration of CO2 injection. This model was provided by the US Department of Energy (DOE) and the National Energy Laboratory (NETL) under Research contract, and is used prior authorization of the project principal investigator, Dr. Shahab D. Mohaghegh.

The top section of this model was selected, prepared, upscaled and history matched for this study by the WVU Petroleum Engineering & Analytics Research Laboratory (PEARL WVU) [6], this processes were performed such that natural heterogeneity depiction was preserved. This section consists of 3,138,200 (149 x 100 x 221) fine scale grid blocks; as a result of areal up scaling, each grid block was changed from almost 10,000 ft² to about 14,000 ft².

The final reservoir model was developed using commercial reservoir simulator CMG's GEM (Generalized Equation of State Model); this numerical simulator is Developed by Reservoir simulation software powerhouse CMG (Computer Modelling Group) and is described as a multi-dimensional, finite-difference, isothermal compositional simulator that can simulate three-phase (oil, water, gas) and multicomponent fluids.

The final reservoir model key characteristics are as follows:

- Dimensions of 14,000 ft wide, 10,000 ft long and up to 800 ft thick.
- Reservoir initially under saturated, WOC at 4500 ft, and initial water saturation below WOC is of 36%.
- Initial Reference pressure at 4300 ft depth is of 3122 psi.
- Upper boundary is set as a no flow boundary. The eastern, western, and northern boundaries are treated as no-flow boundaries, because the Wolfcamp Shale Formation meets these boundaries. The bottom boundary is also designated as a no-flow boundary because of the Strawn Formation below the Cisco and Canyon Formations [6].
Since south of the reservoir is connected to the rest of its sections, a flow boundary was simulated by adding an area at its border that follows the same geology description trend.

Porosity distributions used in this final model were determined from both seismic survey and wire log analyses. Permeability distributions were calculated from seismically calibrated porosity values using empirical equations derived from rock fabric classification [6].

The following figure will illustrate the reader about the structure and distribution of Porosity and Permeability properties throughout the reservoir, as follows:

Figure 7 - Geological model for this study (Axes scale in m) - Left: Porosity Distribution map, Right: Permeability distribution map [6]

Figure 8 - Property distribution Histograms. Left: Porosity, Right: Permeability [6]
There are 127 wells in total, from which 87 are producers and 40 are water alternating gas injectors, production and injection rates are the main constraint of the model. The maximum bottom-hole pressure in the injectors is set at 5,000 psi. The wells are distributed as follows:

![Grid Top (ft)](image)

**Figure 9 - Reservoir model well distribution.**

The oil phase of this reservoir is regarded as a mixture of 11 different gas components and is summarized in the following table:

**Table 1 - Initial oil composition in the upscaled geocellular model (Evaluation and Design of a CO2 Miscible Flood Project - SACROC Unit, Kelly Snider Field, 1973)**

<table>
<thead>
<tr>
<th>Oil Composition</th>
<th>Mol</th>
<th>Molecular Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO2</td>
<td>0.0032</td>
<td>44.01</td>
</tr>
<tr>
<td>N2</td>
<td>0.0083</td>
<td>28.01</td>
</tr>
<tr>
<td>C1(Methane)</td>
<td>0.2865</td>
<td>16.04</td>
</tr>
<tr>
<td>C2(Ethane)</td>
<td>0.1129</td>
<td>30.07</td>
</tr>
<tr>
<td>C3(Propane)</td>
<td>0.1239</td>
<td>44.10</td>
</tr>
<tr>
<td>I-C4(I-Butane)</td>
<td>0.0136</td>
<td>58.12</td>
</tr>
<tr>
<td>N-C4(N-Butane)</td>
<td>0.0646</td>
<td>58.12</td>
</tr>
<tr>
<td>I-C5(I-Pentane)</td>
<td>0.0198</td>
<td>72.15</td>
</tr>
<tr>
<td>N-C5(N-Pentane)</td>
<td>0.0251</td>
<td>72.15</td>
</tr>
<tr>
<td>FC6(Hexane)</td>
<td>0.0406</td>
<td>86.00</td>
</tr>
<tr>
<td>C7+(Heptanes plus)</td>
<td>0.3015</td>
<td>275.00</td>
</tr>
</tbody>
</table>
Reservoir Development Oil recovery techniques can be divided into three distinct phases: Primary, Secondary and Tertiary; Primary recovery is the first stage of hydrocarbon production, in which natural reservoir energy (gas drive, water drive or gravity drainage) displaces hydrocarbons from the reservoir into the wellbore and to the surface, during primary recovery, only a small percentage of the initial hydrocarbons in place are produced, typically around 10% for oil reservoirs [17]; Natural Reservoir pressure declines because of production, reducing the pressure difference which is the main drive of production at this point.

Under those circumstances is necessary to maintain or enhance the Reservoir pressure hence commencing the Secondary recovery processes to improve the field productive life, injecting and external fluid such as water or gas into the reservoir wells located so that fluid has communication with production wells in order to displace oil and drive it into the production wellbore. The most common secondary recovery techniques are waterflooding (injection to production zone in order to create sweep effect) and gas injection (injected into gas cap in order to maintain pressure). This phase reaches its limits when the production and disposal of injecting fluid is such that is no longer economical. The successive use of primary recovery and secondary recovery in an oil reservoir produces about 15% to 40% of the original oil in place. [17]

With much of the easy-to-produce oil already recovered from U.S fields, producers have developed and used different oil recovery enhancement methods using sophisticated techniques that will ultimately alter the original properties of oil, its purpose is not only to restore formation pressure but also to improve oil displacement trough out the reservoir. These techniques offer prospects for ultimately producing 30 to 60 percent, or more, of the reservoir’s original oil in place [17]; there are three major types of EOR operations are Chemical flooding, thermal recovery and miscible displacement (carbon dioxide [CO2] injection or hydrocarbon injection) , latter one being the main operation performed in the field under study in this work.

During CO2-WAG EOR operations, the field is surveyed in order to calculate, according to pressure and temperature, the most suitable conditions in which the CO2 would be compressed and injected such that is miscible with residual oil present in the reservoir, the interfacial tension between this two phases disappear and becomes a single homogenous phase, swelling the oil and reducing its viscosity in order to improve the efficiency of the displacement process. CO2-WAG (Water alternating gas) processes are
known to provide better sweep efficiency and reduce gas channeling from injector to producer in EOR CO2 flood operations [9].

![Schematic diagram of a water-alternating-gas (WAG) miscible CO2 EOR operation][18]

**Figure 10- Schematic diagram of a water-alternating-gas (WAG) miscible CO2 EOR operation [18]**

**Artificial Neural Networks**

**Definition**

In computer science and related fields, Artificial Neural Networks (ANN’s), are computational models inspired by Biological Neural Networks (in particular the brain), which are capable of performing machine learning as well as pattern recognition, both the most prominent features in a Biological Neural Network.

As Biological Neural Networks are composed by neurons, Artificial Neural Networks are composed by artificial neurons or nodes (programming elements that mimic the properties of biological neurons) that are interconnected or functionally related, and in similarity with the Biological architecture, are programmed to perform a specific function. Artificial neural networks may either be used to gain an understanding of biological neural networks, or to solve artificial intelligence problems without necessarily creating a model of a real biological system [19].

[20] Provided a formal definition of Artificial Neural Networks (ANN’s) as follows:

“A neural network is a massively parallel distributed processor made up of simple processing units that has a natural propensity for storing experiential knowledge and making it available for use, It resembles the brain in two respects:
1. Knowledge is acquired by the network from its environment through a learning process.
2. Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge. “

Like other machine learning methods, systems that learn from data, neural networks have been used to solve a wide variety of tasks that are hard to solve using ordinary rule-based programming, including computer vision and speech recognition [21].

ARNN’s provide capabilities and useful properties such as:

- Non-linearity.
- Input-Output mapping.
- Adaptivity.
- Evidential Response.
- Contextual information.

- Fault tolerance.
- Very-Large-Scaled integrated (VLSI) technology implementability.
- Uniformity.
- Neurobiological Analogy.

An Artificial Neural Network consists of an input layer, an output layer and one or more hidden layers. A schematic of a generic feed forward Artificial Neural Network (ARNN’s) with one hidden layer is shown in figure 11.

![Figure 11 Schematic of a Generic feed-forward neural network with a single hidden layer.](image)

The input layer is responsible for receiving input parameters, therefore the number of input neurons should be equal to inputs parameters in our dataset, used for training porpoises. The output of the system is obtained via output layer, therefore, as in the input layer the output neurons should be the same quantity as outputs for our network. Hidden Neurons perform calculations and converts the input into a numerical internal vector, each neuron in the network is connected to other neurons in neighboring
layers. Each connection in the network has an inherent weight which will be referred as synaptic weight, which describes the strength of the signal that the neuron sends to another neuron. Finally the ARNN can be described as a function $X(x,w)$ where $x$ is an example of behavior of our system, and $w$ is a set of weights associated to it. Machine learning, according to these circumstances, would be the procedure of finding the optimum combination of $w$ values to minimize the error over a dataset prediction.

The basic training of an ARNN’s, called supervised learning, uses training set data (multiple examples of inputs and outputs of a given system) in multiple training cycles using the total error of the given data in comparison with output prediction of each training cycle to modify the “synaptic weights”, representing the “experience” of the system. The training of the network is repeated for many examples in the set, until the network reaches a steady state, where there are no further significant changes in the synaptic weights. This is the most popular learning paradigm, which will be covered in later sections in more detail.

Neural networks have a built-in capability to adapt their synaptic weights to changes in the surrounding environment. For example, a network trained to operate in a specific environment, can easily be retrained to deal with minor changes in the operating environmental conditions.

A schematic of a biological neuron is shown in figure 12.

![Figure 12 Schematic of a biological neuron](image)

As the figure shows, a Biological Neuron is connected to several other Biological neurons through its axons and dendrites. A synapse connects an axon to a dendrite. Given a signal, a synapse might increase (excite) or decrease (inhibit) electrical potential. A neuron fires when its electrical potential reaches a threshold. This is for the Artificial Neurons called the “activation function”.

There are two basic types of activation functions, one continuous and one discreet, sigmoid function and threshold function respectively.

A schematic of an artificial neuron is shown in figure 13.
Biological neurons are five to six orders of magnitude slower than silicon logic gates (Artificial neuron signal firing), however the brain makes up for its relatively slow rate of operation by having large amounts of neurons with massive interconnections between them. It has the capability of organizing neurons so as to perform certain computations many times faster than the digital computer. The following table will provide a comparison within the parameters of biological and artificial neural networks:

Table 2 Parameters of biological and artificial neural networks. [22]

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Biological NN</th>
<th>Artificial NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of neurons</td>
<td>$10^{10}$</td>
<td>'000</td>
</tr>
<tr>
<td>Number of synapses</td>
<td>$60 \times 10^{12}$</td>
<td>'000</td>
</tr>
<tr>
<td>Speed [s/op]</td>
<td>$10^4$</td>
<td>$10^{-9}$</td>
</tr>
<tr>
<td>Energy [J/op]</td>
<td>$10^{-16}$</td>
<td>$10^{-6}$</td>
</tr>
</tbody>
</table>

Artificial neurons can be linear or nonlinear. A neural network, made up of an interconnection of nonlinear neurons, is itself nonlinear. Nonlinearity is a highly important property, especially if the underlying physical mechanism responsible for generation of the input signal (e.g., Reservoir Simulation) is inherently nonlinear [21].

Supervised Learning

There are two major learning paradigms in artificial neural network training. These are Supervised and Unsupervised Learning.

Supervised learning as described before, uses a training dataset (multiple examples of inputs and outputs of a given system) that provides the network with the desired response for each training example. This desired response would be the optimum action to be performed by a training set. The network synaptic weights are adjusted due to the generated error of each training sample, this error is calculated from the
difference between the desired response and the actual response of the network. These adjustments are performed iteratively until the network can reproduce all the outputs with reasonable accuracy. At this stage, the network does not need any more supervision and it becomes independent.

**Figure 14 Supervised learning process [22]**

These are the mechanics of how the knowledge of the environment is transferred to the network through training and stored in a form of synaptic weights, representing the experience of the network. The trained network can predict outcomes of the environment completely independently.

The inputs presented to a network in supervised learning are a set of pairs \((x, y), x \in X, y \in Y\), where \(x\) is an example of the environment and pertains to the observation group \(X\), \(y\) is the correspondent actual (or simulated in this work) output that pertains to the results group \(Y\), and the objective is to find a function \(f: X \rightarrow Y\) that matches the examples of the environment with their correspondent outputs.

Error in the output of this function is measured by calculating the Determination Coefficient (R2), therefore providing a measure of how well the model predicts the future outcomes for a given input data. For \(\hat{y}_n\) simulated outputs and \(y_n\) observed outputs of a system, R2 is calculated using the following equations:

\[
SS_{total} = \sum_{i=1}^{n} (y_i - \bar{y})^2, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \quad \{y_i \rightarrow \text{observed output} \quad \bar{y} \rightarrow \text{Mean of observed outputs}\}
\]

Equation 1 - Total sum of squares (observed output)

\[
SS_{regression} = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} \hat{y}_i \quad \{\hat{y}_i \rightarrow \text{simulated output} \quad \bar{y} \rightarrow \text{Mean of simulated outputs}\}
\]
The most general definition of $R^2$ is:

$$R^2 = 1 - \frac{SS_{\text{regression}}}{SS_{\text{total}}}, \quad (0 \leq R^2 \leq 1)$$

**Equation 3 - R-square general definition**

The Determination Coefficient $R^2$ is a relative good yardstick to measure how well an ARNN will predict outputs for a given new input data, accordingly to its extremes lower and upper acceptable numerical bounds it represents no correlation and full correlation with prediction correspondently.

However, $R^2$ does not provide information on whether:

- The independent variables are the true cause of the changes in the dependent variables,
- Omitted variable bias exists,
- The most appropriate set of independent variables has been selected
- There is co-linearity present in the data, or
- The model might be improved by using transformed versions of the existing set of independent variables.

Due to applicability and the constant evolution of this concepts the ARNN’s vary on their structure and porpoise widely, the Networks structures used in this work are established and extensively proven, are as follows: General Regression Neural Networks (GRNN’s) and Backpropagation Neural Networks (BPN’s).

This network types will be described in following sections of this work.

**General Regression Neural Networks (GRNN)**

The general regression neural network (GRNN) is a one-pass learning algorithm with a highly parallel structure. Is a memory based network that provides estimates of continuous variables and converges to the underlying function (linear or non-linear); it is shown that, even with sparse data in a multidimensional measurement space, the algorithm provides smooth transitions from one observed value to another. [23]

Is a type of application of Probabilistic Neural Network, therefore its construct is based heavily in statistical data and not in “real machine learning”. It extracts characteristics from a training sample to come to knowledge about the underlying function in a one-step approach, hence its strength is not in the selection
of weights itself but in the function associated to its neuron. Since is a one-pass algorithm is really fast to generate and study, many researchers use it because of the ease of its application.

The algorithmic form can be used for any regression problem in which an assumption of linearity is not justified. The weighting factors computed by the GRNN are unequal, as a modification of a basic Radial basis function (RBF), and based on the distances between the sample input and each member of the training population.

Given the training sample \((X_n, Y_n)\) and the prediction point \(X\), The general algorithm of GRNN is the following for the predicted output \(Y\):

\[
Y(X) = \frac{\sum_{i=1}^{n} Y_i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}
\]

\[
D_i^2 = (X - X_i)^T (X - X_i)
\]

*Equation 4 - General Regression Neural Network algorithm [23]*

Observations:

- **Neuron Activation Function** -> \(\exp\left(-\frac{D_i^2}{2\sigma^2}\right)\) : acts like a synaptic weight for the input.
- **Di^2** -> Euclidean distance between Training sample and Point of Prediction: Measure of how well represented is the training sample portrayed by the point of prediction.
- **X** -> Point of prediction , **Y** -> Predicted point
- **Xi** -> training sample , **Yi** -> output of training sample

The GRNN limitations are the following:

- Prediction for extreme values is not possible.
- Depends heavily on the \(\sigma\) (smoothing factor) factor selection, there are various methods not covered in this work for the determination of this factor (Ex: Wiggle method, holdout method).
- Forecast capabilities (Extrapolation) using GRNN in a multidimensional model (or even slightly complex one-dimensional model) is very limited.
- Its construct is more based in statistical data of inputs than in real machine-learning.
Back propagation Neural Networks

Back propagation Neural Networks (BPN's) are often used with feed forwards networks, is used to calculate the most optimized synaptic weights values in order to accomplish a target function that can resemble as closely as possible to observed values of a training set.

The method usually used to calculate the weight changes is the gradient descent. Training begins with an arbitrary set of weights. A series of computations (iterations) is done in which the calculated output is compared with the known values, adjusting the weights in such a way that the difference between the calculated values and the target function is minimized (Smith, 1993).

This process is described by [24] as follows:

“With each iteration the hidden layer passes information through based on values of the weights in memory and the output values are calculated. The output nodes are then informed of the difference between the actual and target values. Each output neuron determines in which direction its weights must be adjusted to reduce the error and propagates the information to the hidden layer, which in turn determines in which direction its weights must be changed. At the hidden layer level the weights are adjusted in such a way as to reduce the error across the full set of output neurons thus minimizing the error in the network. For each iteration there is thus a forward pass followed by a backward pass during which error information is propagated backward from the output neurons to the hidden neurons.”

Is a normal practice to partition the Training set data (input data) for its use in the network development as follows:

- 80% of the data for Training purposes,
- 10% of the data for Calibration purposes,
- 10% of the data for Verification and Validation purposes.

General consensus of opinion is that back-propagation is the best general-purpose model and probably the best at generalization. The following figure depicts how a general operation schematic for a Back propagation neural network should be:
Figure 15 - Schematic of a General Backpropagation Neural Network [25]

Given the training sample \((X_n, Y_n)\), the prediction point \((X)\) and set of initial synaptic weights \((w_i)\), the general basic algorithm of Backpropagation Neural Network is the following for the predicted output \((Y)\):

1. \[ e_j = Y_j \ast (1 - Y_i) \ast \sum (e_k \ast w_{jk}) \rightarrow \text{error calculation backprop} \]

2. \[ w_{ij} = w_{ij}^i + LR \ast e_j \ast X_i \rightarrow \text{weight adjustment} \]

Figure 16 - Backpropagation Neural Network Algorithm schematic
Observations:

- \( w_{ij} \), New weight calculated
- \( w'_{jk} \), Previous weight
- \( k \), sum of errors terms for each neuron
- \( \text{LR} \), Learning Rate
- \( e_{ij} \), Error term
- \( Y_i \), Observed output
- \( Y_j \), Calculated output

Backpropagation Neural Networks are very flexible, suitable for pattern recognition and decision making problems. Training process has a high dependence to order of the training samples used. The process can be highly parallel, therefore cutting down processing times (Processing times include the randomization of input order of relatively huge amount of data needed to perform training). Is based in pure machine learning (gradient descent, momentum factor), doesn’t use methods that are based in statistical methods.

Reservoir Simulation

In this section we will briefly discuss the governing equations for a water-oil flow numerical reservoir simulator that doesn’t account for compositional flow nor thermal considerations. [26]

Oil-Water Flow Equations

Consider incompressible oil-water flow with no mass transfer between phases when capillary pressure effects are ignored. Mass balance equation can be expressed as:

\[
\phi \frac{\partial S_j}{\partial t} + \nabla \cdot \mathbf{U}_j = q_j ,
\]

- \( j \), phase (j=0 for oil and w for water)
- \( \phi \), porosity
- \( S_j \), Saturation phase of \( j \)
- \( \mathbf{U}_j \), phase Darcy velocity
- \( q_j \), source term

The phase velocity is related to the pressure gradient through Darcy’s law:

\[
U_j = -\frac{\rho_j k_{ij}}{\mu_j} \nabla \varphi_j , \quad \varphi_j = \frac{p}{\rho_j} - gD
\]
• \( p \), pressure
• \( D \), depth
• \( g \), gravitational acceleration
• \( \rho_j \), phase density
• \( \mu_j \), phase viscosity
• \( k_{rj} \), relative permeability to the phase \( j \)
• \( k \), absolute permeability
• \( \nabla \varphi_j \), phase potential

Applying a standard finite volume discretization, the system can be written as:

\[
T^{n+1}X^{n+1} + D(X^{n+1} - X^n) + Q^{n+1} = R^{n+1}, \quad F^{n+1} = T^{n+1}X^{n+1}
\]

• \( n \), time level variable
• \( F \), flow term
• \( X \), designates the vector primary unknowns (pressure and saturation in each grid block)
• \( T \), transmissibility matrix
• \( D \), Accumulation matrix
• \( Q \), source/sink terms
• \( R \), Residual Vector

This system is not linear and has to be solved using Newton’s method as follows:

\[
J\delta = -R
\]

• \( J \), designates a Jacobian matrix
• \( \delta \), represents the solution update; i.e. \( X^{n+1,k+1} = X^{n+1,k} \), where \( k \) designates the iteration level.

Other methods for non-linear solving problems that are worth mentioning are Gauss-seidel method, Point-over successive relaxation method (PSOR) and Linear-over successive relaxation method (LSOR).

Computational tasks in a Commercial reservoir simulation can be divided into a few main categories, and they take different CPU processing times percentage according to their size and node/processor configuration, as described in the following figure for GigaPOWERS (Saudi Aramco in-house Simulation Software):
Proxy Reservoir models at grid block level

Proxy Reservoir models are accurate depictions of Reservoir simulations, their objective is to provide fast approximated solutions that substitutes large numerical simulation models in order to assist managerial decisions on a timely manner.

Most commonly used proxy models in the industry are Response surfaces (Approximates the problem using statistical techniques) and Reduced Order Models (ROM) (Approximates the solution space of the problem using different techniques), these simplifications can only provide, at best, limited responses at well locations [27], limiting the Reservoir Engineer to the scenarios contemplated within the solution space, reproducing probable outputs that will ultimately run useless for a managerial decision unless expanding of solution space takes place. [8].

The term “Proxy” means “to act on behalf of another”.

Reduced Order Models (ROM) are a type of proxy models that reproduces the Reservoir simulation into a lower dimensional space, giving us a solution within a grossly approximated solution space. Using multiple mathematical approaches (Radial Basis Function NN’s, Proper Orthogonal Decomposition-POD, and interpolation methods) researchers are successful in “reducing” models and obtaining good results, within
their solution space, while using hundreds of reservoir simulation results, running useless this type of analysis due to the complexity of their structure and application.

**Grid Surrogate Reservoir Model (GSRM)**

Grid Surrogate Reservoir models (SRM) are a type of proxy model that based on Artificial Intelligence technology and data mining techniques can learn, accurately replicate and expand our knowledge of a reservoir using a commercial Reservoir compositional numerical simulator (CMG-GEM) to generate a vast dataset that gives neural networks enough knowledge over its different scenarios outputs (experimental design) for this purpose.

Necessity of SRM’s resides in the fact that massive potentials of existing numerical reservoir simulation models go unrealized because of the long CPU times needed to make single runs, as of the complicated partial differential equations involved and the computational fluid dynamics large quantities of data.

In order to perform history matching processes, optimization or quantification of uncertainties in a numerical reservoir model (analyses that require large number of simulation runs to perform) idle time between scenarios statement – study – result processes increases exponentially for every simulation run performed, henceforth making impractical the managerial decisions over an asset because of the timeframe in which they are proposed.

SRM’s gives us the would-be response of the numerical reservoir simulator nor in days or hours, but in seconds, because of the simplicity of the resultant mathematical functions associated with them, that are manufactured via highly parallel processes that establish a target function based on real data inputs to sculpt a build-to-perform model, a tuned Neural Network.

While hundreds of simulations are required to build other Reservoir proxy models approaches, a GSRM requires only a small amount of simulation runs due to a unique an innovative method of data sampling for the generation of required spatio-temporal database.

SRM at grid block level and well level has been applied numerous times with outperforming results each time, we can safely conclude that the efficiency of this technology has been proven for different cases and scenarios. [1] [27] [28]

The uniqueness of this study is that a GSRM development for a Water Alternating Gas (WAG) CO2-EOR field multi-compositional Reservoir Simulation Model (CMG GEM), one of the most complex type of simulation in the industry, has not been done before.
Chapter 3 – Methodology

General methodology

In this chapter, the methodology of this work is presented in detail. The objective of this work, as mentioned earlier, is to assess the feasibility of using state of the art data-driven proxy models to effectively replicate dynamic reservoir properties of complex full field numerical models, by building and validating a grid block level Surrogate Reservoir Model (SRM) of a macro-scale complex full field CO2 injection –EOR numerical model. The methodology presented in this study is detailed in the following figure:

Figure 18 - General Methodology Flowchart
The dynamic reservoir properties to replicate field wide using the Surrogate Reservoir Model are the following:

- Pressure (PRESS)
- Oil Saturation (SO)
- Water Saturation (SW)
- Global CO2 mole fraction (CO2)

The timeframe chosen for this study is of 5 years, 2006 to 2010; since we cannot access the source code of the Numerical simulator (CMG GEM) in order to convert its outputs from Data in a convoluted Non-human readable data format, Yearly datasets will be deconvoluted, extracted and organized from simulator outputs for each scenario stipulated, on the latter we will be performing various analyses and performing data mining techniques in order to obtain “not apparent data” that will be useful in revealing hidden patterns to the Neural Network architecture design. Later on, these Neural Networks will be studied in a standalone fashion to determine their individual quality, then concatenated in a Cascading feed-forward design in order to obtain a fully-fledged Grid based Surrogate model (GSRM) for the former Numerical Reservoir model.

Analyses will be performed on the GSRM generated in order to assess the feasibility of use of this technology for the optimization of CO2-EOR WAG processes in a full complex field Numerical Simulator.

Commercial and Non-Commercial Software employed

For the different stages of this study we will be using multiple software’s, described as follows:

- Computer Modeling Group (CMG), General Equation of State Model (GEM), Commercial Numerical Simulator software.
- Intelligent Solutions Inc. (ISI), Intelligent Data Evaluation & Analysis (IDEA), Neural Network Developer and Data analysis software.
- Microsoft EXCEL 2013, Spreadsheet software.
- Python ver2.7, Programming language.
- Microsoft Visual Basic 6, Programming language.
Experimental design of CMG simulator Scenarios

In order to generate a solution space that would allow the GSRM to generate sufficient response for the reservoir model under study, we generated a set of 17 CMG simulation runs, 13 of this simulation results were used to build a spatio-temporal dataset that would in further steps be used for training purposes, and 4 of this simulation runs were used for verification purposes, that is, for comparison of the output of the CMG simulation runs vs. the GSRM runs under the same scenario defined by its independent and dependent variables. This process is called blind verification process.

The independent variables taken in consideration for the experimental design and their correspondent ranges are:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range of values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gas Injection rate (GV)</td>
<td>10 – 80 (MMscf/day)</td>
</tr>
<tr>
<td>Gas injection Time (GT)</td>
<td>6-24 months</td>
</tr>
<tr>
<td>Water injection rate (WV)</td>
<td>10-50 (Mbbl/day)</td>
</tr>
<tr>
<td>Water injection times (WT)</td>
<td>6-24 months</td>
</tr>
</tbody>
</table>

The dependent variables taken in consideration for the experimental design and their correspondent ranges are:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Range of values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water slug Volume (WSV)</td>
<td>Injected Slug of water volume (WV*WT)</td>
<td>0.1-2% reservoir pore volume</td>
</tr>
<tr>
<td>Gas slug Volume (GSV)</td>
<td>Injected Slug of gas volume (GV*GT)</td>
<td>0.1-2% reservoir pore volume</td>
</tr>
<tr>
<td>Water alternating Gas</td>
<td>Water/Co2 injection ratios at reservoir conditions (WV/GV)</td>
<td>0.5 - 4</td>
</tr>
</tbody>
</table>
The selection or sampling of this multidimensional space must be done in a methodical fashion, ensuring that every parameter (independent or dependent) is statistically well represented in order to ensure that the solution space generated for the GSRM will comprehensively portray the CMG simulations responses in fewer experiments.

To perform this sampling we chose the Latin Hypercube sampling method, this is the generalization of the Latin square method (Euler method) to an arbitrary number of dimensions, whereby each scenario selected is the only in every multidimensional axis-aligned hyperplane that contains it, ensuring this way that our solution space will contain an unique probable answer for each type of response under the parameters drawn. The scenarios designed are as follows:

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>GV (MMscf/day)</th>
<th>GT (months)</th>
<th>GSV (MMscf)</th>
<th>WV (Mbbl/day)</th>
<th>WT (months)</th>
<th>WSV (Mbbl)</th>
<th>Ratio</th>
<th>WAG</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.883721</td>
<td>12</td>
<td>118.6047</td>
<td>50</td>
<td>24</td>
<td>1200</td>
<td>10.11765</td>
<td>4</td>
<td>Training</td>
</tr>
<tr>
<td>2</td>
<td>13.17829</td>
<td>8</td>
<td>105.4264</td>
<td>40</td>
<td>20</td>
<td>800</td>
<td>7.588235</td>
<td>3</td>
<td>Training</td>
</tr>
<tr>
<td>3</td>
<td>10.87209</td>
<td>6</td>
<td>65.23256</td>
<td>30</td>
<td>11</td>
<td>330</td>
<td>5.058824</td>
<td>2</td>
<td>Training</td>
</tr>
<tr>
<td>4</td>
<td>13.34302</td>
<td>24</td>
<td>320.2326</td>
<td>45</td>
<td>18</td>
<td>810</td>
<td>2.529412</td>
<td>1</td>
<td>Training</td>
</tr>
<tr>
<td>5</td>
<td>14.37632</td>
<td>22</td>
<td>316.2791</td>
<td>25</td>
<td>16</td>
<td>400</td>
<td>1.264706</td>
<td>0.5</td>
<td>Training</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>6</td>
<td>60</td>
<td>10</td>
<td>8</td>
<td>80</td>
<td>1.333333</td>
<td>0.527132</td>
<td>Training</td>
</tr>
<tr>
<td>7</td>
<td>70</td>
<td>13</td>
<td>910</td>
<td>50</td>
<td>24</td>
<td>1200</td>
<td>1.318681</td>
<td>0.521339</td>
<td>Training</td>
</tr>
<tr>
<td>8</td>
<td>30</td>
<td>7</td>
<td>210</td>
<td>30</td>
<td>22</td>
<td>660</td>
<td>3.142857</td>
<td>1.242525</td>
<td>Training</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>14</td>
<td>700</td>
<td>42</td>
<td>23</td>
<td>966</td>
<td>1.38</td>
<td>0.545581</td>
<td>Training</td>
</tr>
<tr>
<td>10</td>
<td>80</td>
<td>6</td>
<td>480</td>
<td>45</td>
<td>24</td>
<td>1080</td>
<td>2.25</td>
<td>0.889535</td>
<td>Training</td>
</tr>
<tr>
<td>11</td>
<td>23</td>
<td>8.0358</td>
<td>184.8234</td>
<td>49.116</td>
<td>19.1418</td>
<td>940.16</td>
<td>5.086849</td>
<td>2.01108</td>
<td>Training</td>
</tr>
<tr>
<td>12</td>
<td>80</td>
<td>24</td>
<td>1920</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>Training</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>24</td>
<td>1200</td>
<td>-</td>
<td>-</td>
<td>Training</td>
</tr>
<tr>
<td>14</td>
<td>11.18272</td>
<td>10</td>
<td>111.8272</td>
<td>45</td>
<td>22</td>
<td>990</td>
<td>8.852941</td>
<td>3.5</td>
<td>Blind</td>
</tr>
<tr>
<td>15</td>
<td>11.06977</td>
<td>7</td>
<td>77.48837</td>
<td>35</td>
<td>14</td>
<td>490</td>
<td>6.323529</td>
<td>2.5</td>
<td>Blind</td>
</tr>
<tr>
<td>16</td>
<td>9.883721</td>
<td>8</td>
<td>79.06977</td>
<td>50</td>
<td>6</td>
<td>300</td>
<td>3.794118</td>
<td>1.5</td>
<td>Blind</td>
</tr>
<tr>
<td>17</td>
<td>30</td>
<td>16</td>
<td>480</td>
<td>50</td>
<td>14</td>
<td>700</td>
<td>1.458333</td>
<td>0.57655</td>
<td>Blind</td>
</tr>
</tbody>
</table>

The total field volumes of Water/CO2 were divided amongst the water alternating gas injector wells, and taking in consideration each well rate constraint, pressure constraint and history of injection until 2002, we were able to assign a suitable rate for each one of them in each scenario.
Data Deconvolution and Extraction

Deconvolution is an algorithm-based process used to reverse the effects of convolution on recorded data from the numerical simulator step-based operations outputs; due to the nature of this data and its large size (5GB per case -> approx. 200GB total) a Console Application was developed on Python language in order to achieve in an efficient manner the Deconvolution, Extraction and storage of this data in a friendly format set of large matrixes designed to provide quick quality checking characteristics.

CMG’s GEM software keywords involved are the following,

For dynamic properties:

- **CO2**: Represents Global CO2 mole fraction per block.
- **PRESS ALL**: Represents Pressure in psi per block.
- **SO**: Represents Oil Saturation in percentage (%) per block.
- **SW**: Represents Water Saturation in percentage (%) per block.

For static properties:

- **X-dist**: Represents the distance of model origin to center of the block in X axis for non-orthogonal grid.
- **Y-dist**: Represents the distance of model origin to center of the block in Y axis for non-orthogonal grid.
- **PERM**: Represents the permeability in md per block.
- **PORO**: Represents the effective porosity in % per block.

For well and field production:

- **d-PRO CO2, d-PRO WAT, d-PRO OIL**: Represents field wide production data for different components.
- **d-INJ CO2, d-INJ WAT**: Represents field wide injection data for different components.
- **PRO CO2, PRO WAT, PRO OIL**: Represents individual well production data for different components.
- **INJ CO2, INJ WAT**: Represents individual well injection data for different components.
Numerical Simulator output *.INC files are of the following characteristics:

Data obtained after deconvoluted and extracted *.NPY (Numeric Python Compressed Binary file) and *.XLSX (MS-EXCEL 2013 file) files are of the following characteristics:

The 17 base Datasets generated (28 features extracted for 3,600,000 blocks) are spatio-temporal datasets, meaning that they represent in a block per block basis all the features extracted in a dynamic (time-dependent) fashion. These datasets will serve different purposes in the whole process, described as following:
• 13 datasets (Scenarios: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13) will be used for training and verification purposes.
• 4 datasets (Scenarios: 14, 15, 16, 23) will be used solely for blind run verification purposes.

The Console Application for deconvolution, extraction and organization of data from CMG output simulator files source code will be released for public usage and is provided for readers study in Appendix section A.
Dataset Management

Data extracted was conformed into matrixes of large quantity of items, in this section we will explain the base datasets transformation into analyzed, tuned and manageable final datasets that will definitively provide us with sufficient information to build a Grid Based Surrogate Reservoir Model; the transformation evolution through the different processes performed is as follows:

![Dataset matrixes transformations through the data management process.](image)

This process is going to be explained in detail in the following sub-sections of this work.

Model Initial features

The initial features extracted from Simulator output data are the following:

<table>
<thead>
<tr>
<th>Initial Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>case</td>
<td># of Scenario from were the data is extracted</td>
</tr>
<tr>
<td>I</td>
<td>index for block</td>
</tr>
<tr>
<td>J</td>
<td>index for block</td>
</tr>
<tr>
<td>K</td>
<td>index for block</td>
</tr>
<tr>
<td>X</td>
<td>Euclidian Distance of model origin to center of the block in X axis for non-orthogonal grid. (ft)</td>
</tr>
<tr>
<td>Y</td>
<td>Euclidian Distance of model origin to center of the block in Y axis for non-orthogonal grid. (ft)</td>
</tr>
<tr>
<td>Grid Top</td>
<td>Euclidian Distance of model origin to Average of top corners (4) in Z axis for non-orthogonal grid. (ft)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initial Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERM</td>
<td>Represents the permeability per block (md)</td>
</tr>
<tr>
<td>PORO</td>
<td>Represents the effective porosity per block (%)</td>
</tr>
<tr>
<td>CO2</td>
<td>Represents Global CO2 mole fraction per block. (%) per time step (Year1-Year5)</td>
</tr>
<tr>
<td>PRESS</td>
<td>Represents Pressure per block (psi) per time step (Year1-Year5).</td>
</tr>
<tr>
<td>SO</td>
<td>Represents Oil Saturation in percentage (%) per block per time step (Year1-Year5).</td>
</tr>
<tr>
<td>SW</td>
<td>Represents Water Saturation in percentage (%) per block per time step (Year1-Year5).</td>
</tr>
</tbody>
</table>

Table 6 - Initial Features extracted from Simulator extracted Data.
Additional secondary features were generated from initial extraction features cross-validation with pre-defined structure of the Reservoir simulation model, described as following:

<table>
<thead>
<tr>
<th>Secondary Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>patch</td>
<td>#unique identifier for each block</td>
</tr>
<tr>
<td>Grid Thickness</td>
<td>Computation of spatial property Grid Top. (ft)</td>
</tr>
<tr>
<td>Z center</td>
<td>Computation of spatial property Grid Top. (ft)</td>
</tr>
<tr>
<td>Boundary</td>
<td>Boolean (0/1) value that indicates if a block is a boundary block.</td>
</tr>
<tr>
<td>NULL</td>
<td>Boolean (0/1) value that indicates if a block is a Null simulator [pinch out] block.</td>
</tr>
<tr>
<td>DistNFB</td>
<td>Computation of spatial property X and Y to No flow Boundary, distance. (ft)</td>
</tr>
</tbody>
</table>

Table 7 - Secondary Features Generated from Simulator extracted Data.

The base datasets have a total 34 features x 3,600,000 blocks matrix dimension, quality checking procedures were performed with initial raw data in order to assure no errors existed in the deconvolution and extraction process. Statistical analyses were performed in order to insure non erratic behavior of data and correlation of variables and features.

Layer Lumping

Layer lumping process was applied to the massive datasets generated because of the number of blocks present could not be handled by the available software (ISI-IDEA) in order to ensure proper processing and management of data.

As described in Chapter 1 of this work under Numerical Reservoir Model description section, the Reservoir numerical model estimates 36 layers for the Reservoir and 100x110 blocks per layer in i and j direction respectively; The selection of lumped layers was designed in concordance with the predominating layer structure present in the Numerical reservoir model for 5 different layer intervals, described as the following:

- Layer 1: Layer 1 to Layer 15.
- Layer 4: Layer 29 to Layer 32.
- Layer 5: Layer 33 to Layer 36.

The lumping process was designed such that different properties were properly averaged using distinct methods according to their nature, described as follows:
Table 8 - Averaging methods used for different features in Layer Lumping process.

The Layer Lumped datasets generated are of 34 features x 55,000 blocks; quality checking procedures were performed with initial raw data and random sampled data in order to assure no errors existed in the layer lumping calculation process. Statistical analyses were performed in order to insure non erratic behavior of data and correlation of variables and features.

**Data Mining**

Data Mining is the computational process of discerning arrangements and patterns inherent to large datasets involving methods such as intersection of different databases, cluster analysis based in spatial properties, anomaly data detection (outliers) and association rule mining (dependency determination). As defined by Wikipedia (Online Encyclopedia): “The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use”.

In order to generate “not apparent Data” that will be useful in revealing hidden patterns within the data and fully grasp the complexity and heterogeneity of the field under study in the Neural Network architecture design process, we introduced the following databases to intersect the base datasets:

- Well Location, Production and Operational Constraints database.
- Field Production database.
This intersection generated the following features:

<table>
<thead>
<tr>
<th>Intersection Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1, p2, p3</td>
<td>First, Second and Third closest production well unique identifier for each block based on Euclidian distance.</td>
</tr>
<tr>
<td>g1, g2, g3</td>
<td>First, Second and Third closest CO2 injection well unique identifier for each block based on Euclidian distance.</td>
</tr>
<tr>
<td>w1, w2, w3</td>
<td>First, Second and Third closest water injection well unique identifier for each block based on Euclidian distance.</td>
</tr>
<tr>
<td>dp1, dp2, dp3</td>
<td>First, Second and Third closest Euclidian distance to production well for each block (Ft)</td>
</tr>
<tr>
<td>dg1, dg2, dg3</td>
<td>First, Second and Third closest Euclidian distance to CO2 injection well for each block (Ft)</td>
</tr>
<tr>
<td>dw1, dw2, dw3</td>
<td>First, Second and Third closest Euclidian distance to Water injection well for each block (Ft)</td>
</tr>
<tr>
<td>bhp p1, bhp p2, bhp p3</td>
<td>First, Second and Third closest bottomhole pressure Constraint for production wells based on Euclidian distance for each block (PSI)</td>
</tr>
<tr>
<td>q g1, q g2, q g3</td>
<td>First, Second and Third closest Rate Constraint for CO2 injection wells based on Euclidian distance for each block (SCF/day)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intersection Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>q w1, q w2, q w3</td>
<td>First, Second and Third closest Rate Constraint for Water injection wells based on Euclidian distance for each block (Bbls/day)</td>
</tr>
<tr>
<td>P block, G block, W block</td>
<td>Boolean (1/0) values for each block in wether they represent a production, CO2 injection or water injection well</td>
</tr>
<tr>
<td>Well Cumulative production of CO2, Water and Oil. SIMULATOR BASED</td>
<td>For the First, Second and Third production wells, the cumulative in a time step based fashion of component produced. (Scf, Bbls)</td>
</tr>
<tr>
<td>Field cumulative (FieldCum) production of CO2, Water and Oil. SIMULATOR BASED</td>
<td>For the First, Second and Third production wells, the cumulative in a time step based fashion of component injected. (Scf, Bbls)</td>
</tr>
<tr>
<td>Field cumulative (FieldCum) injection of CO2 and Water. SIMULATOR BASED</td>
<td>For the First, Second and Third injection wells, the cumulative in a time step based fashion of component injected. (Scf, Bbls)</td>
</tr>
<tr>
<td>Field cumulative (FieldCum) injection of CO2 and Water. CONSTRANT BASED</td>
<td>For the Field, cumulative production of component produced (Scf, Bbls).</td>
</tr>
<tr>
<td>Field cumulative (FieldCum) production of CO2 and Water. SIMULATOR BASED</td>
<td>For the Field, cumulative production of component produced (Scf, Bbls).</td>
</tr>
<tr>
<td>Field cumulative (FieldCum) injection of CO2 and Water. SIMULATOR BASED</td>
<td>For the Field, maximum cumulative production of component injected (Scf, Bbls).</td>
</tr>
</tbody>
</table>

Table 9 - Intersection with Well databases Features generated description.

In order to generate features from a proper cluster analysis based in spatial properties, we generated definitions of a block based tier system composed of 8 tiers, averaging properties for each group of block for each tier; Based on each block i, j, k indexes, the description of each tier is as follows:

- Tier 1: Average of Immediate adjacent blocks in same layer.
- Tier 2: Immediate inferior layer block.
- Tier 3: Immediate superior layer block.
- Tier 4: Average of Immediate non-adjacent blocks in same layer.
- Tier 5: Average of Immediate adjacent and non-adjacent blocks in inferior layer
- Tier 6: Average of Immediate adjacent and non-adjacent blocks in superior layer
- Tier 7: Average of Tier 1, Tier 2, Tier 3 values
- Tier 8: Average of Tier 4, Tier 5, Tier 6 values
The following figure will illustrate this definitions:

![Figure 22 - Tier system definition](image)

This tier design generated 8 new features per dynamic and static property per time step studied (160 total new features), giving our models a proper spatial analysis and placement in order to better understand the distribution of such properties throughout the field.

The data mined datasets generated are of 357 features x 55,000 blocks; quality checking procedures were performed with initial raw data and random sampled data, in order to assure no errors existed in the database intersection and tiering system calculation process. Statistical analyses were performed in order to insure non erratic behavior of data and correlation of variables and features.
Neural Network Architecture Design and GSRM development

Neural Networks have the capability of performing Machine learning, as explained in former chapters, they read each block of the simulation as a different case to learn and try to replicate, consequently, a single scenario dataset contains 357 features for 55,000 cases of study in neural network terms.

Data was analyzed and structured into large spatio-temporal datasets with an important quantity of features available to be considered into our inputs selection for Neural Network Architecture design and training, this said, not all the inputs generated are going to be useful in “teaching” the machine about the behavior of this Reservoir.

Data will be unified, methodically sampled, timely parsed thus conforming Training spatio-temporal datasets for our neural networks in order to achieve the goals outlined, this reduction takes place granted that the computing powers and the tools used for this Research have data management limitations such that our Training spatio-temporal datasets have to be limited to a maximum frame of 32,000 cases of study; Consequence of this computing power and tools data management limitations the actual final Training spatio-temporal dataset will be only 22.34% of the total data available for all the cases of study.

Is a normal practice to randomly partition the Training spatio-temporal dataset (input data) for its use in the network development, as follows:

- 80% of the data for Training purposes,
- 10% of the data for Calibration purposes,
- 10% of the data for Verification and Validation purposes.

Once this data is sampled, inputs are selected and data is partitioned for different purposes in behalf of better understatements of the neural network and its grasp with the behavior of the reservoir; a Cascading design is going to be generated in order to create Neural Network models that will be trained and later deployed, thus conforming a Grid Based Surrogate Reservoir Model.

This process is going to be explained to detail in the following sub-sections of this work.

Data Sampling

Data sampling is the process were we select the data within spatio-temporal datasets, previously generated for each case of study, as aforementioned the product Sampled Training spatio-temporal dataset will only represent approximately 22% of the whole available data pool.

Three data sampling methods were developed and selected for this study, described as follows:
• Method A: 10% of data was selected from well locations, 90% of data was randomly selected from whole data pool.

• Method B: The data was ranked according to their pressure values (high pressure to low pressure) and Probability Distribution function (PDF) of the data was studied, data was then divided in uneven pressure ranges in order to acquire for the selected bin-ranges approximately the same % of total data present (20% of data for each bin).

• Method C: The data was ranked according to their pressure values (high pressure to low pressure) and Probability Distribution function (PDF) of the data was studied, data was then divided in even pressure ranges (450 psi difference) in order to acquire for the selected bin-ranges the data available.

The following figure will explain the main differences of the sampling methods here explained, as follows:

![Figure 23 - Sampling methods comparison.](image)

The data was sampled for the different methods and studied using multiple General Regression neural networks (GRNN) analyses, using a basic input selection, in order to estimate their potential to generate good Backpropagation models. Once this analyses are performed the best sampling method (under these circumstances) is selected and employed to generate all the Training spatio-temporal datasets.

Once the Training spatio-temporal datasets are generated, they will be parsed for the time steps within the time frame of this study (2006 to 2010) adding the features for the past time step in each of them, hence generating 5 different Training datasets with dimensions of 141 features for 32,000 blocks (cases).
Input Selection and Neural Network Training

The features in the Training spatio-temporal Datasets are classified into 8 main categories, described as follows:

<table>
<thead>
<tr>
<th>Group</th>
<th># of Features</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP ID</td>
<td>5</td>
<td>Contains the unique ID # features for each case.</td>
</tr>
<tr>
<td>GROUP: CONSTRAINT</td>
<td>38</td>
<td>Contains the constraint parameters for each case.</td>
</tr>
<tr>
<td>GROUP: STATICS</td>
<td>18</td>
<td>Contains the values of Permeability and Porosity (and its tiers) for each block.</td>
</tr>
<tr>
<td>GROUP: DYNAMICS</td>
<td>72</td>
<td>Contains the dynamic values of Pressure, Global Co2 mole fraction, Water saturation and Oil saturation for timestep.</td>
</tr>
<tr>
<td>GROUP : FIELD DATA (Simulator Data Generated)</td>
<td>2</td>
<td>Contains the Cumulative production and injection of the whole field, based in simulator results.</td>
</tr>
<tr>
<td>GROUP : WELL DATA (Simulator Data Generated)</td>
<td>2</td>
<td>Contains the Cumulative production and injection based on the first to third closest well per condition, based in simulator results.</td>
</tr>
<tr>
<td>GROUP : FIELD DATA (Constraint Generated)</td>
<td>2</td>
<td>Contains the Cumulative production and injection of the whole field, based in constraint data.</td>
</tr>
<tr>
<td>GROUP : WELL DATA (Constraint Generated)</td>
<td>2</td>
<td>Contains the Cumulative production and injection based on the first to third closest well per condition, based in constraint data.</td>
</tr>
</tbody>
</table>

Figure 24 - Input classification for Neural Network Architecture design.

Multiple Group of Inputs are selected in order to strategy the best output architecture possible for the Neural Network design, these data is imported into the SRM Software for preprocessing, inputs/outputs are selected for each property to replicate and a proprietary technique known as “KPI Performance Analysis” is used to analyze the data an rank the input parameters based on their level of influence for the specified output.

Figure 25 - KPI performance Analysis Screenshot and Neural Net design (ISI-IDEA 2013).

The input selection criteria is the aided by the “KPI Performance Analysis”, giving us a hint of whether parameters should be included or discarded from the study. The vast majority of inputs are discarded and
the top performing inputs that make more sense are selected. (Based on Reservoir Engineering Knowledge)

The training of the Neural Network models then is commenced and continues until either maximum specified number of iterations has been reached or the estimation error has reached the allowable error.

Figure 26 - Neural Network Training Screenshot and Live results analysis (ISI-IDEA 2013).
Cascading Design

Cascading is the arrangement of the Neural Networks input selection in order to emulate a feed forward succession of data, thus creating a Grid Based Surrogate reservoir model that would only need initial inputs to replicate responses for pre-defined time steps.

The cascading design generated for this study is as follows:

![Figure 27 - Cascading Design.](image)

The initial input data is the dynamic properties and constraint features for the former time step (t-1), these will be input for the pressure model of time step (t), its output will be input for the oil saturation model of time step (t), its output will be input for the water saturation model of time step (t), its output will be input for the Global CO2 mole fraction model of time step (t) and finally all this outputs are gathered and organized once more as inputs for model in the future time step (t+1). This process is repeated until we reach model pertaining to Year 2010 output, since we reached our final prediction time step.

There are 20 models in total to be cascaded, 4 per year, one for each dynamic property to replicate; the runtime of this application is estimated to be around 20-30 minutes, this because of the intensive data management that is taking place in each step in order to arrange inputs and export outputs pertaining to every stage of the process.
Chapter 4 - Results and Discussion

In this section we will present the results of the applied Grid Based Surrogate model (GSRM); these are portrayed and analyzed in an attempt to measure the feasibility of use of this technology in the proxy modeling of CO2 flooding Enhanced oil recovery fields that use Water alternating Gas (WAG) Techniques.

Initially, results for the different methodology stages will be presented and analyzed, in order to set the sufficient frame to present the final results for Verification and Validation of the models using Training inputs and Blind inputs, this will ultimately lead us to the study of the error probability distribution function (PDF) of each property set of models and of the GSRM in general, thus giving us the appropriate tools to determine a conclusion that will ultimately define this study.

Initial input selection contemplated Constraint based and Simulator output based data for Well and Field features in training spatio-temporal datasets; although Simulator output based data performed significantly better than Constraint based data, in the Key performance analysis and in our r-square values for each model, these features had to be discarded due to the cascading design chosen for this work and the maximum amount of models (NN's) to be generated. Thereupon our final input selection for the GSRM design proposed, is as follows:

<table>
<thead>
<tr>
<th>GROUP ID</th>
<th>PRESS models</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP CONSTRAINT</td>
<td>PRESS (t-1)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP CONSTRAINT</td>
<td>PRESS (t-1)</td>
<td>T7</td>
</tr>
<tr>
<td>GROUP CONSTRAINT</td>
<td>PRESS (t)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP CONSTRAINT</td>
<td>PRESS (t)</td>
<td>T7</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>SO (t-1)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>SW (t-1)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>CO2 (t-1)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>CO2 (t)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>CO2 (t-1)</td>
<td>T7</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>CO2 (t)</td>
<td>T7</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>PRESS (t-1)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>PRESS (t-1)</td>
<td>T7</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>PRESS (t)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>PRESS (t)</td>
<td>T7</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>SO (t-1)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>SW (t-1)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>CO2 (t-1)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>CO2 (t)</td>
<td>T1</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>CO2 (t-1)</td>
<td>T7</td>
</tr>
<tr>
<td>GROUP ID</td>
<td>CO2 (t)</td>
<td>T7</td>
</tr>
</tbody>
</table>

Figure 28 - Final input selection.
This input selection ensures that cascading takes place during our GSRM execution by feeding forward the outputs of each model as inputs of the following one. This input selection design was used to train NN's for approximately 4000 epochs (learning cycle) using 32,000 cases with features ranging from 25 to 37, depending on the property to be modeled.

Aforementioned, determination coefficient R2 is a relative good yardstick to measure how well the NN will predict outputs for a given new data; The R2 coefficients of the Neural Networks (models) generated for each property, for each time step, are the following:

<table>
<thead>
<tr>
<th></th>
<th>PRESS</th>
<th>SO</th>
<th>SW</th>
<th>CO2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>0.84</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Year 2</td>
<td>0.91</td>
<td>0.99</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>Year 3</td>
<td>0.81</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Year 4</td>
<td>0.46</td>
<td>0.99</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>Year 5</td>
<td>0.82</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 10 - R2 coefficient of all NN's generated.

This R2 coefficients obtained, were the best coefficients that could be obtained for each property modeled with our past input selection and data available, the low R2 coefficient for pressure model in Year 4 (2009) suggests that improvement could take place. In further chapters we will give recommendations for future researchers.

The generated models were concatenated using the past mentioned cascading design, the Best Training Verification Case prediction distribution maps for the first and last year of simulation and correspondent histograms are as follows:
For Year 1 (2006):

Figure 29 - Best Training case prediction, Pressure Model, Year 1 (2006).
Figure 30 - Best Training case prediction, SO Model, Year 1 (2006).
Year: 2006 - Case: 6 - Property: SW

Figure 31 - Best Training case prediction, SW Model, Year 1 (2006).
Figure 32 - Best Training case prediction, CO2 Model, Year 1 (2006).
For Year 5 (2010):

Figure 33 - Best Training case prediction, Pressure Model, Year 5 (2010).
Figure 3.4 - Best Training case prediction, SO Model, Year 5 (2010).
Figure 35 - Best Training case prediction, SW Model, Year 5 (2010).
Figure 36 - Best Training case prediction, CO2 Model, Year 5 (2010).
The Worst Training Verification Case prediction distribution maps for the first and last year of simulation and correspondent histograms are as follows:

Figure 37 - Worst Training case prediction, Pressure Model, Year 1 (2006).
Figure 38 - Worst Training case prediction, SO Model, Year 1 (2006).
Figure 39 - Worst Training case prediction, SW Model, Year 1 (2006).
Figure 40 - Worst Training case prediction, CO2 Model, Year 1 (2006).
• For Year 5 (2010):

Figure 41 - Worst Training case prediction, Pressure Model, Year 5 (2010).
Figure 42 - Worst Training case prediction, SO Model, Year 5 (2010).
Figure 43 - Worst Training case prediction, SW Model, Year 5 (2010).
Figure 44 - Worst Training case prediction, CO2 Model, Year 5 (2010).
The Best blind Validation prediction distribution maps for each year and correspondent histograms are as follows:

**Figure 45 - Best Blind case prediction, Pressure Model, Year 1 (2006).**
Figure 46 - Best Blind case prediction, SO Model, Year 1 (2006).
Figure 47 - Best Blind case prediction, SW Model, Year 1 (2006).
Figure 48 - Best Blind case prediction, CO2 Model, Year 1 (2006).
• For Year 5 (2010):

**Figure 49** - Best Blind case prediction, Pressure Model, Year 5 (2010).
Figure 50 - Best Blind case prediction, SO Model, Year 5 (2010).
Figure 51 - Best Blind case prediction, SW Model, Year 5 (2010).
Figure 52 - Best Blind case prediction, CO2 Model, Year 5 (2010).
The Worst blind Validation prediction distribution maps for the first and last year of simulation and correspondent histograms are as follows:

Figure 53 - Worst Blind case prediction, Pressure Model, Year 1 (2006).
Figure 54 - Worst Blind case prediction, SO Model, Year 1 (2006).
Figure 55 - Worst Blind case prediction, SW Model, Year 1 (2006).
Figure 56 - Worst Blind case prediction, CO2 Model, Year 1 (2006).
For Year 5 (2010):

Figure 57 - Worst Blind case prediction, Pressure Model, Year 5 (2010).
Figure 58 - Worst Blind case prediction, SO Model, Year 5 (2010).
Figure 59 - Worst Blind case prediction, SW Model, Year 5 (2010).
Figure 60 - Worst Blind case prediction, CO2 Model, Year 5 (2010).
As we can observe in this plots, the critical error points in the error maps correspond to the injection wells location blocks, those that produce the highest change in their properties during injection periods; as the GSRM tends to normalize the results within closest points in the distribution during the training calibration process, is almost impossible to grasp these changes without further data treatment.

We can observe a good correlation of simulator results and GSRM results for almost every year of simulation, error histogram data, for both training verification and blind validation processes, indicates that the larger percent of data is present within acceptable thresholds of error (≤10%), thus providing a good prediction of the reservoir model in general.

The following figure will provide Probability Error histograms; the population of this study contains all the active reservoir block error prediction calculations values for each property in each time step:

![Error Histograms for all prediction blocks](image)

**Figure 61 – GSRM all cases output: Error Histogram**

This analysis depicts the Probability density function of the error produced by the GSRM model; after analyzing these charts we can infer that the GSRM predicts in a very accurate fashion (error<10%) the CMG numerical model outputs.
The Error probability distribution is an important method that measures the quality of the data obtained via GSRM model, uncertainty assessment for the possible outputs of this surrogate model are as follows:

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>So</th>
<th>Sw</th>
<th>CO2</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>P10</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>P50</td>
<td>1.77%</td>
<td>0.47%</td>
<td>0.44%</td>
<td>0.53%</td>
<td>1%</td>
</tr>
<tr>
<td>P90</td>
<td>5.76%</td>
<td>2.97%</td>
<td>2.99%</td>
<td>4.61%</td>
<td>4%</td>
</tr>
<tr>
<td>max</td>
<td>31%</td>
<td>60%</td>
<td>30%</td>
<td>75%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Uncertainty Assessment of the output data interpretation is as follows:

- For pressure models deployment and application:
  - 90% of chance having an average error $\leq 5.76\%$ for prediction results.
  - 50% of chance having an average error $\leq 1.77\%$ for prediction results.
  - 10% of chance having an average error of 0 % for prediction results.

- For Oil Saturation (SO) models deployment and application:
  - 90% of chance having an average error $\leq 2.97\%$ for prediction results.
  - 50% of chance having an average error $\leq 0.47\%$ for prediction results.
  - 10% of chance having an average error of 0 % for prediction results.

- For pressure models deployment and application:
  - 90% of chance having an average error $\leq 2.99\%$ for prediction results.
  - 50% of chance having an average error $\leq 0.44\%$ for prediction results.
  - 10% of chance having an average error of 0 % for prediction results.

- For pressure models deployment and application:
  - 90% of chance having an average error $\leq 4.61\%$ for prediction results.
  - 50% of chance having an average error $\leq 0.53\%$ for prediction results.
  - 10% of chance having an average error of 0 % for prediction results.
The following figure will assess the efficiency of the GSRM in runtime terms, as follows:

<table>
<thead>
<tr>
<th>Model</th>
<th>CMG GEM</th>
<th>GSRM (Python + VB6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hours</strong></td>
<td>18.4</td>
<td>0.43</td>
</tr>
</tbody>
</table>

As the former figure depicts, there is a runtime reduction of approximately 98%, from 18.4 hours runtime to 27 min per run, this means that during a single scenario run for the CMG numerical simulator we can perform approximately 43 scenario runs using the GSRM generated.

All the results generated by this study will be present in the Appendix section B of this work.
Concluding Remarks

The major conclusions of this work can be summarized as follows:

- A novel Grid block level Surrogate Reservoir Model (GSRM) to replicate dynamic responses from a macro-scale complex full field CO2 injection –EOR numerical model was generated and validated; as portrayed in the results section of this study, the accurate results and the short runtime of this GSRM in replicating the numerical simulator outputs (From training and blind scenarios) makes this technology suitable and practical for the industry complex model proxy simulation needs.

- GSRM predictive quality is highly dependable in data mining methods involved in the Training dataset generation process, therefore dataset generation is the most critical step for success in any GSRM generation study.

- Artificial Intelligence based generated models, such as GSRM, are objective-defined proxy models not suitable to use in another type of study that is not contemplated in their solution space and training inputs selection, that said, the GSRM model generated in this study would be a good tool to study the WAG field design effect on dynamic responses of the reservoir but not a good tool to study different geological reservoir realizations effect on dynamic responses.

- In the making of this GSRM, software applications were generated to deconvolute and extract data from CMG numerical simulator software outputs, this applications will be useful for future WVU researchers in the area of Reservoir simulation and proxy modeling.

- The GSRM generated large amounts of data of this SACROC Reservoir simulation model (200GB+), this data will be useful for future WVU researchers in the area of proxy modeling.
Recommendations and Future Work

The recommendations for future researchers are as follows:

• Runtime of GSRM is significantly lower than from CMG numerical simulator (about 98% faster), this runtime is mostly for the massive data arrangement routines (coded in Python language); Runtime could be critically improved by using other programming languages that are native to the user operating system (such as VB.NET in Windows 7).

• In order to evolve the GSRM generated into not only a tool that could investigate the effects of WAG field design effects on Reservoir models dynamic responses but a tool that could analyze dynamic responses given simulator dynamic rates, is necessary to include well and field data from simulation rates enhancing the actual inputs that only consider constraint based well and field data.

• In order to enhance the GSRM accuracy is necessary to investigate further input combinations selections and sampling methods, since this are the most critical steps in obtaining accurate results.
References


Appendix section A – Console Application Source Code

```python
#!/usr/bin/env python

import os
import csv
import time
import gmtk, strftime

def reader_stastics(filepath, filename):
    if "Poreability" in filename:
        reshaper="no"
    elif "Porousity" in filename:
        reshaper="no"
    else:
        reshaper="yes"

    p=filename.strip(".inc")
    ins = open(filepath)
    Km =
    array = []
    final_array[]
    times[]
    time_count=0
    for line in ins:
        lines=line.strip("/n").split()
        if "TIME" in line:
            time_count+=1
            time.extend([line[3],time_count])
        elif "RESULTS" in line:
            continue
        elif "Grid" in line:
            continue
        elif p in line:
            continue
        elif "PERMII" in line:
            continue
        elif "POR" in line:
            continue
        elif "K" in line:
            continue
```
```python
def reader_dynamic(filepath, filename):
    ins = open(filepath, 'r')
    X = []
    array = []
    final_array = []
    time = []
    time_count = 0
    for line in ins:
        lines = line.strip('/n').split()
        if "TIME" in line:
            time.append(int(line[3]))
            if time_count <= 2:
                final_array = np.array(array, dtype='f4')
                array = []
            elif time_count > 2:
                array = np.array(array, dtype='f4')
                final_array = np.vstack((final_array, array))
                array = []
        elif "RESULTS" in line:
            continue
        elif "ALL" in line:
            continue
        elif "X" in line:
            X = line[6]
            if X == 1:
                X = 1
                continue
        else:
            array.extend((lines))
    ins.close()
    array = np.array(array, dtype='f4')
    final_array = np.vstack((final_array, array))
```
array=[]
final_array=final_array.T
filename.strip(".inc")
headers=list
headers=[(filename+"="+str(1)), (filename+"="+str(2)), (filename+"="+str(3)), (filename+"="+str(4)), (filename+"="+str(5)), (filename+"="+str(6)),
(filename+"="+str(7)), (filename+"="+str(8)), (filename+"="+str(9)), (filename+"="+str(10)),
(filename+"="+str(11)), (filename+"="+str(12)), (filename+"="+str(13)), (filename+"="+str(14)),
(filename+"="+str(15)), (filename+"="+str(16)), (filename+"="+str(17))]

return final_array, headers, time

#===============================================================================================
#===============================================================================================
def timelist(p,filepath, filename):
    filepath=osp.path.join(filepath+'\'+filename+'.inc')
    ins = open(filepath)
    K=0
    array = []
    final_array=[]
    time=[]
    time_count=0
    for line in ins:
        lines=line.strip("/n").split()
        if "TIME" in line:
            time_count=time_count+1
            time.append((lines[3]))
    ins.close()
    time=np.array(time, dtype='f4')
    return time

#===============================================================================================
#===============================================================================================
def indexing(tier, I, J, K):
    array=[]

    try:
        for item in tier:
            x=item[0]
            y=item[1]
            z=item[2]
            index=(x-1)+(y-1)+(z-1)
            array.insert(0, index)
    except:
        try:
            x=tier[0]
            y=tier[1]
            z=tier[2]
            tier=[(x-1)+(y-1)+(z-1)]
        except:
            tier=[]
return tier

if len(tier)>3:
    c=c-1
    for x_item in tier[:,0]:
        c=c+1
        if x_item<1:
            try:
                tier=np.delete(tier, c, 0)
                c=c-1
            except:
                pass
        elif x_item>x:
            try:
                tier=np.delete(tier, c, 0)
                c=c-1
            except:
                pass
    c=c-1
    for y_item in tier[:,1]:
        c=c+1
        if y_item<1:
            try:
                tier=np.delete(tier, c, 0)
                c=c-1
            except:
                pass
        elif y_item>y:
            try:
                tier=np.delete(tier, c, 0)
                c=c-1
            except:
                pass
    c=c-1
    for z_item in tier[:,2]:
        c=c+1
        if z_item<1:
            try:
                tier=np.delete(tier, c, 0)
                c=c-1
            except:
                pass
        elif z_item>z:
            try:
                tier=np.delete(tier, c, 0)
                c=c-1
except:
    pass
else:
    x_item = tier[0]
    y_item = tier[1]
    z_item = tier[2]
    if x_item > x or x_item == 1:
        tier = []
    elif y_item > y or y_item == 1:
        tier = []
    elif z_item > z or z_item == 1:
        tier = []
    return tier

def tiering(I, J, K):
    # tier 1
    tier1 = np.array([[I-1, J, K],
                      [I+1, J, K],
                      [I, J-1, K],
                      [I, J+1, K]], dtype='int')
    # tier 2 - above
    tier2 = np.array([[I, J, K+1]], dtype='int')
    # tier 3 - below
    tier3 = np.array([[I, J, K-1]], dtype='int')
    # tier 4
    tier4 = np.array([[I+1, J+1, K],
                      [I+1, J-1, K],
                      [I+1, J+1, K],
                      [I+1, J-1, K]], dtype='int')
    # tier 5
    tier5 = np.array([[I-1, J, K+1],
                      [I, J-1, K],
                      [I, J+1, K],
                      [I-1, J, K]], dtype='int')
    # tier 6
    tier6 = np.array([[I-1, J, K-1],
                      [I, J-1, K],
                      [I, J+1, K],
                      [I-1, J, K]], dtype='int')
    # tier 7
    tier7 = np.vstack((tier1, tier2, tier6))
# tier8
tier8=np.vstack((tier4,tier5,tier6))

tier1=indexing(limia(tier1,100,110,36),100,110,36)
tier2=indexing(limia(tier2,100,110,36),100,110,36)
tier3=indexing(limia(tier3,100,110,36),100,110,36)
tier4=indexing(limia(tier4,100,110,36),100,110,36)
tier5=indexing(limia(tier5,100,110,36),100,110,36)
tier6=indexing(limia(tier6,100,110,36),100,110,36)
tier7=indexing(limia(tier7,100,110,36),100,110,36)
tier8=indexing(limia(tier8,100,110,36),100,110,36)

return tier1,tier2,tier3,tier4,tier5,tier6,tier7,tier8

#////////////////////////////////////////////////////////////////////////////////////////////////////////
#////////////////////////////////////////////////////////////////////////////////////////////////////////
#////////////////////////////////////////////////////////////////////////////////////////////////////////
def define_tiers(I,J,K):

KJ_indexes=[]
for K in xrange(1,136+1):
    for J in xrange(1,110+1):
        for I in xrange(1,100+1):
            KJ_indexes.append([I,J,K])
KJ_indexes=np.array(KJ_indexes,dtype='int')
for item in KJ_indexes:
    i=item[0]
    j=item[1]
    k=item[2]
    tier1,tier2,tier3,tier4,tier5,tier6,tier7,tier8=indexing(i,j,k)
    item=np.append(item,(tier1,tier2,tier3,tier4,tier5,tier6,tier7,tier8))
final_array=np.vstack((final_array,item))

return final_array

#////////////////////////////////////////////////////////////////////////////////////////////////////////
#////////////////////////////////////////////////////////////////////////////////////////////////////////
#////////////////////////////////////////////////////////////////////////////////////////////////////////
def single_index(x,y,z,I,J,k):
    index=(x-1)+(y-1)+(z-1)+(I-1)*(y-1)+(1)*(J-1)+(1)*(k-1)
    return index

#////////////////////////////////////////////////////////////////////////////////////////////////////////
#////////////////////////////////////////////////////////////////////////////////////////////////////////
#////////////////////////////////////////////////////////////////////////////////////////////////////////
def whave(array,axis,weights):

c=0
result=[]
thickness=sum(weights)
for col in xrange(0,len(array[1])):
    val=np.array([,col])
    if col in range(0,7+1):
        ave=np.average(val,0)
    else:
        c=0
```python
ave = val[-1]

elif col == 3:
    NA_indexes = np.where(val == 0)
    val = np.delete(val, NA_indexes, 0)
    inst_weights = np.delete(weights, NA_indexes, 0)
    ave = (len(val)) / float(sum(1 / val))

elif col == 10:
    NA_indexes = np.where(val == 0)
    val = np.delete(val, NA_indexes, 0)
    inst_weights = np.delete(weights, NA_indexes, 0)
    ave = np.average(val, 0, inst_weights)

elif col in range(0, 27 + 1):
    NA_indexes = np.where(val == 0)
    val = np.delete(val, NA_indexes, 0)
    inst_weights = np.delete(weights, NA_indexes, 0)
    ave = np.average(val)

elif col in range(27 + 1, 78 + 1):
    NA_indexes = np.where(val == 0)
    val = np.delete(val, NA_indexes, 0)
    inst_weights = np.delete(weights, NA_indexes, 0)
    ave = np.average(val, 0, inst_weights)

result.append(ave)

zblock = array[:, 8].min() + thickness / 2
result.insert(8, zblock)
result = np.array(result)
result[8] = thickness

return result

# ///////////////////////////////////////////////////
# ///////////////////////////////////////////////////
# ///////////////////////////////////////////////////
# ///////////////////////////////////////////////////

def lumper((layer1), (layer2), (layer3), (layer4), (layer5), array):

layers = [layer1, layer2, layer3, layer4, layer5]
index_array = []
new_array = []
temp_array = []
layer_count = 0
for layer in layers:
    counter = 0
    slice_index = single_index(106, 110, layer[1], 100, 110, 36)
    slice_index = single_index(1, 1, layer[0], 106, 110, 36)
    temp_arr = array[slice_index:1:slice_index + 1, :]
    layer_count = layer_count + 1
    new_array = []
    for j in range(1, 110 + 1):
        for i in range(1, 100 + 1):
            counter = counter + 1
            index_array = []
            NA_indexes = []
            inst_array = []
            for k in range(1, layer[1] - layer[0] + 2):
                indexes = single_index(i, j, k, 100, 110, 36)
```
index_array.insert(0,indexes)
inst_array=temp_array[index_array,:]
junk_array=temp_array[index_array,:]
NA_indexes=np.where(inst_array[:,0]==0)
inst_array=np.delete(inst_array,NA_indexes,0)

try:
    inst_array=inst_array[~np.isnan(inst_array[:,0]),:]
    inst_array[0,1]=
    inst_array[2]=layer_count
except:
    inst_array=np.hstack((junk_array[-1,:],np.zeros(72)))
    inst_array[:,:,layer_count
if counter==1:
    new_array=inst_array
    inst_array=[]
else:
    new_array=np.vstack((new_array,inst_array))
    last error!!!!!!! debugging OJO

print "temporal array"
np.array.shape, layer

if layer_count==1:
    result_array=new_array
else:
    result_array=np.vstack((result_array,new_array))

print result_array.shape
print result_array[:,:100]
print strftime(\"%Y-%m-%d %H:%M:%S\", gmttime())

return result_array

# def well_dict(filepath,filename_list,cases,cases,times,welref):

field_header=['time1', 'time2', 'time3', 'time4', 'time5', 'time6', 'time7', 'time8', 'time9',
'time10', 'time11', 'time12', 'time13', 'time14', 'time15', 'time16', 'time17']
array_constraint_per_well=np.zeros([254,85])
array_constraint_per_well[array.constraint_per_well==0]=999
start_pos_list=list(range(0,85,17))
total_wells=list()
cols_header=list()

for num_file,files in enumerate(filename_list):
    array_file = [line.strip().split() for line in open(os.path.join(cases+"\"cases+
"\"Well\"\"count*"\"\"files))]
    listA_wells=list()
    flagger=False
    for line in array_file:"
```python
if flagger==True:
    if str(line[0]) not in lista_wells:#-----< added 3/10/2014 for stop repeating
double wells!!!! because f restart file phenomena
    lista_wells.append(str(line[0]))
flagger=False

if "WELL" in line:
    flagger=True

well_index=list()
for i in range(0,len(lista_wells)):
    for j in range(0,len(wellREF)):
        if wellREF[j]==lista_wells[i]:
            well_index.append(j)

total_wells.append(lista_wells)
print total_wells

flagger=False

i=0
while i < len(array_file)-1:
    if flagger==True:
        i=i+1
    flagger=False
    if array_file[i]==[]:
        array_file.pop(i)
    try:
        line=np.array(array_file[i],dtype=np.float)
    except:
        array_file.pop(i)
    flagger=True

i=i+1
array_file=np.array(array_file,dtype=np.float)
times=np.array(times,dtype=np.float)
index_list=list()
for count, item in enumerate(times):
    index_list=np.where(array_file[:,0]==item)[0][:]

index=np.array(index)
for i in range(0,len(index)*5):
    index=index[range(0,len(index),5)]

if "Prod" in files:
    index=np.array(0: len(index)-10)
array_constraint_per_well[well_index,start_pos[num_file]+count]=array_file[index,1]
cons_header.append("-T"+str(count+i))

for item in sublist in total_wells for item in sublist:
    total_wells=total_wells

print total_wells

print len(total_wells)-10
c_header=",".join([c_header])
return array_constraint_per_well,cons_header
```

---
def constrainter(case_constraint, gen_constraint, constraint_headers):
    c=0
    headers=list([])
    for col in range(6, 17+1):
        c=c+1
        data_col=gen_constraint[:, col]
        indexes=case_constraint[:, 6]
        a=list()
        for item in data_col:
            indx=np.array(np.where(indexes==item))
            a.append(indx)
        b=list()
        for item in a:
            try:
                val=item[0][0]
                b.append(val)
            except:
                val=0
                b.append(val)

    print(col)
    values=case_constraint[b, 1]
    values=np.array(values)
    if c==1:
        bhp_rates=values
    else:
        bhp_rates=np.vstack((bhp_rates, values))

    constraint_array=mp.hstack(((gen_constraint, bhp_rates.T)))
    print(constraint_array.shape)
    c_header='\,'.join(constraint_headers)
    return constraint_array, constraint_headers

def constrainter_final(gen_constraint, cons2_basearray, constraint_array1, c_header, cons2_headers):
    final_headers=list()
    for col_con, item in enumerate(gen_constraint[:, range(9, 11+1)].T):
        constraint_array2_1b=mp.zeros((55000, 51))
        constraint_array2_1b[constraint_array2_1b==0]=999
        for row_con, well in enumerate(item):
            if well<=999:
                constraint_array2_1b[row_con]=cons2_basearray[well-1, range(0, 50+1)]
        if col_con==0:
            constraint_array2=constraint_array2_1b
        elif col_con==1:
            constraint_array2=np.hstack((constraint_array2, constraint_array2_1b))
        else:
            constraint_array2=np.hstack((constraint_array2, constraint_array2_1b))
        for i in range(0, 50+1):
            final_headers.append(c_header[9+col_con]+"+"+cons2_headers[1])
    for col_con, item in enumerate(gen_constraint[:, range(12, 14+1)].T):
        constraint_array2_1b=mp.zeros((55000, 17))
Appendix B – GSRM output results.

Figure 63 - GSRM Results, CO2, year-case: 2006-1
Figure 64 - GSRM Results, PRESS, year-case: 2006-1
Figure 65 - GSRM Results, SO, year-case: 2006-1
Year: 2006 - Case: 1 - Property: SW

Figure 66- GSRM Results, SW, year-case: 2006-1
Figure 67 - GSRM Results, CO2, year-case: 2007-1
Figure 68 - GSRM Results, PRESS, year-case: 2007-1
Figure 69 - GSRM Results, SO, year-case: 2007-1
Figure 70- GSRM Results, SW, year-case: 2007-1
Figure 71- GSRM Results, CO2, year-case: 2008-1
Figure 72 - GSRM Results, PRESS, year-case: 2008-1
Figure 73: GSRM Results, SO, year-case: 2008-1
Figure 74- GSRM Results, SW, year-case: 2008-1
Figure 75: GSRM Results, CO2, year-case: 2009-1
Figure 76 - GSRM Results, PRESS, year-case: 2009-1
Figure 77 - GSRM Results, SO, year-case: 2009-1
Figure 78- GSRM Results, SW, year-case: 2009-1
Figure 79 - GSRM Results, CO2, year-case: 2010-1
Figure 80- GSRM Results, PRESS, year-case: 2010-1
Figure 81- GSRM Results, SO, year-case: 2010-1
Figure 82 - GSRM Results, SW, year-case: 2010-1
Year: 2006 - Case: 2 - Property: CO2 Results

Figure 83 - GSRM Results, CO2, year-case: 2006-2
Figure 84: GSRM Results, PRESS, year-case: 2006-2
Figure 85- GSRM Results, SO, year-case: 2006-2
Figure 86- GSRM Results, SW, year-case: 2006-2
Figure 87 - GSRM Results, CO2, year-case: 2007-2
Figure 88- GSRM Results, PRESS, year-case: 2007-2
Figure 89 - GSRM Results, SO, year-case: 2007-2
Figure 90- GSRM Results, SW, year-case: 2007-2
Figure 91 - GSRM Results, CO2, year-case: 2008-2
Figure 92 - GSRM Results, PRESS, year-case: 2008-2
Figure 93- GSRM Results, SO, year-case: 2008-2
Figure 94- GSRM Results, SW, year-case: 2008-2
Figure 95 - GSRM Results, CO2, year-case: 2009-2
Figure 96 - GSRM Results, PRESS, year-case: 2009-2
Figure 97: GSRM Results, SO, year-case: 2009-2
Figure 98- GSRM Results, SW, year-case: 2009-2
Figure 99- GSRM Results, CO2, year-case: 2010-2
Figure 100 - GSRM Results, PRESS, year-case: 2010-2
Figure 101 - GSRM Results, SO, year-case: 2010-2
Figure 102- GSRM Results, SW, year-case: 2010-2
Figure 103 - GSRM Results, CO2, year-case: 2006-3
Figure 10.4 - GSRM Results, PRESS, year-case: 2006-3
Figure 105 - GSRM Results, SO, year-case: 2006-3
Figure 106- GSRM Results, SW, year-case: 2006-3
Figure 107 - GSRM Results, CO2, year-case: 2007-3
Figure 108- GSRM Results, PRESS, year-case: 2007-3
Figure 109 - GSRM Results, SO, year-case: 2007-3
Figure 110 - GSRM Results, SW, year-case: 2007-3
Figure 111 - GSRM Results, CO2, year-case: 2008-3
Figure 112 - GSRM Results, PRESS, year-case: 2008-3
Figure 113- GSRM Results, SO, year-case: 2008-3
Figure 114 - GSRM Results, SW, year-case: 2008-3
Figure 115: GSRM Results, CO2, year-case: 2009-3
Figure 116 - GSRM Results, PRESS, year-case: 2009-3
Figure 117 - GSRM Results, SO, year-case: 2009-3
Figure 118- GSRM Results, SW, year-case: 2009-3
Figure 119 - GSRM Results, CO2, year-case: 2010-3
Figure 120- GSRM Results, PRESS, year-case: 2010-3
Figure 121- GSRM Results, SO, year-case: 2010-3
Figure 122: GSRM Results, SW, year-case: 2010-3
Figure 123 - GSRM Results, CO2, year-case: 2006-4
Figure 124 - GSRM Results, PRESS, year-case: 2006-4
Figure 125 - GSRM Results, SO, year-case: 2006-4
Figure 126 - GSRM Results, SW, year-case: 2006-4
Figure 127: GSRM Results, CO2, year-case: 2007-4
Figure 128 - GSRM Results, PRESS, year-case: 2007-4
Figure 129 - GSRM Results, SO, year-case: 2007-4
Figure 130 - GSRM Results, SW, year-case: 2007-4
Figure 131 - GSRM Results, CO2, year-case: 2008-4
Figure 132 - GSRM Results, PRESS, year-case: 2008-4
Figure 133: GSRM Results, SO, year-case: 2008-4
Year: 2008 - Case: 4 - Property: SW

Figure 134: GSRM Results, SW, year-case: 2008-4
Figure 135: GSRM Results, CO2, year-case: 2009-4
Figure 136 - GSRM Results, PRESS, year-case: 2009-4
Figure 137: GSRM Results, SO, year-case: 2009-4
Figure 138- GSRM Results, SW, year-case: 2009-4
Figure 139- GSRM Results, CO2, year-case: 2010-4
Figure 140 - GSRM Results, PRESS, year-case: 2010-4
Figure 141 - GSRM Results, SO, year-case: 2010-4
Figure 142: GSRM Results, SW, year-case: 2010-4
Figure 143: GSRM Results, CO2, year-case: 2006-5
Year: 2006 - Case: 5 - Property: PRESS

Figure 144 - GSRM Results, PRESS, year-case: 2006-5
Figure 145 - GSRM Results, SO, year-case: 2006-5
Figure 147 - GSRM Results, CO2, year-case: 2007-5
Figure 148 - GSRM Results, PRESS, year-case: 2007-5
Figure 149- GSRM Results, SO, year-case: 2007-5
Figure 150 - GSRM Results, SW, year-case: 2007-5
Figure 151 - GSRM Results, CO2, year-case: 2008-5
Figure 152 - GSRM Results, PRESS, year-case: 2008-5
Figure 153 - GSRM Results, SO, year-case: 2008-5
Figure 154 - GSRM Results, SW, year-case: 2008-5
Figure 155: GSRM Results, CO2, year-case: 2009-5
Figure 156 - GSRM Results, PRESS, year-case: 2009-5
Figure 157: GSRM Results, SO, year-case: 2009-5
Year: 2009 - Case: 5 - Property: SW

Figure 158 - GSRM Results, SW, year-case: 2009-5
Figure 159 - GSRM Results, CO2, year-case: 2010-5
Figure 160 - GSRM Results, PRESS, year-case: 2010-5
Figure 161 - GSRM Results, SO, year-case: 2010-5
Figure 162 - GSRM Results, SW, year-case: 2010-5
Figure 163: GSRM Results, CO2, year-case: 2006-6
Figure 164 - GSRM Results, PRESS, year-case: 2006-6
Figure 165 - GSRM Results, SO, year-case: 2006-6
Year: 2006 - Case: 6 - Property: SW

Figure 166- GSRM Results, SW, year-case: 2006-6
Figure 167: GSRM Results, CO2, year-case: 2007-6
Figure 168 - GSRM Results, PRESS, year-case: 2007-6
Year: 2007 - Case: 6 - Property: SO

Figure 169 - GSRM Results, SO, year-case: 2007-6
Figure 170 - GSRM Results, SW, year-case: 2007-6
Figure 171- GSRM Results, CO2, year-case: 2008-6
Figure 172 - GSRM Results, PRESS, year-case: 2008-6
Figure 173 - GSRM Results, SO, year-case: 2008-6
Figure 174- GSRM Results, SW, year-case: 2008-6
Year: 2009 - Case: 6 - Property: CO2

Figure 175: GSRM Results, CO2, year-case: 2009-6
Figure 176 - GSRM Results, PRESS, year-case: 2009-6
Figure 177 - GSRM Results, SO, year-case: 2009-6
Figure 178 - GSRM Results, SW, year-case: 2009-6
Figure 179- GSRM Results, CO2, year-case: 2010-6
Figure 180 - GSRM Results, PRESS, year-case: 2010-6
Figure 181 - GSRM Results, SO, year-case: 2010-6
Figure 182 - GSRM Results, SW, year-case: 2010-6
Year: 2006 - Case: 7 - Property: CO2

Figure 183 - GSRM Results, CO2, year-case: 2006-7
Figure 184 - GSRM Results, PRESS, year-case: 2006-7
Year: 2006 - Case: 7 - Property: SO

Figure 185: GSRM Results, SO, year-case: 2006-7
Figure 186- GSRM Results, SW, year-case: 2006-7
Figure 187: GSRM Results, CO2, year-case: 2007-7
Figure 188- GSRM Results, PRESS, year-case: 2007-7
Figure 189- GSRM Results, SO, year-case: 2007-7
Figure 190- GSRM Results, SW, year-case: 2007-7
Figure 19-1: GSRM Results, CO2, year-case: 2008-7
Year: 2008  - Case: 7  - Property: PRESS

Figure 192 - GSRM Results, PRESS, year-case: 2008-7
Figure 193 - GSRM Results, SO, year-case: 2008-7
Figure 194- GSRM Results, SW, year-case: 2008-7
Figure 195: GSRM Results, CO2, year-case: 2009-7
Figure 196 - GSRM Results, PRESS, year-case: 2009-7
Year: 2009 - Case: 7 - Property: SO

Figure 197 - GSRM Results, SO, year-case: 2009-7
Figure 198 - GSRM Results, SW, year-case: 2009-7
Figure 199- GSRM Results, CO2, year-case: 2010-7
Figure 200 - GSRM Results, PRESS, year-case: 2010-7
Figure 201 - GSRM Results, SO, year-case: 2010-7
Figure 202 - GSRM Results, SW, year-case: 2010-7
Year: 2006 - Case: 8 - Property: CO2

Figure 203 - GSRM Results, CO2, year-case: 2006-8
Figure 20.4 - GSRM Results, PRESS, year-case: 2006-8
Figure 205 - GSRM Results, SO, year-case: 2006-8
Figure 206- GSRM Results, SW, year-case: 2006-8
Figure 207- GSRM Results, CO2, year-case: 2007-8
Figure 208- GSRM Results, PRESS, year-case: 2007-8
Figure 209 - GSRM Results, SO, year-case: 2007-8
Figure 210 - GSRM Results, SW, year-case: 2007-8
Figure 211 - GSRM Results, CO2, year-case: 2008-8
Figure 212 - GSRM Results, PRESS, year-case: 2008-8
Figure 213 - GSRM Results, SO, year-case: 2008-8
Figure 214 - GSRM Results, SW, year-case: 2008-8
Figure 215: GSRM Results, CO2, year-case: 2009-8
Figure 216 - GSRM Results, PRESS, year-case: 2009-8
Figure 217 - GSRM Results, SO, year-case: 2009-8
Figure 218 - GSRM Results, SW, year-case: 2009-8
Figure 219 - GSRM Results, CO2, year-case: 2010-8
Figure 220 - GSRM Results, PRESS, year-case: 2010-8
Figure 221 - GSRM Results, SO, year-case: 2010-8
Figure 222 - GSRM Results, SW, year-case: 2010-8
Figure 223- GSRM Results, CO2, year-case: 2006-9
Year: 2006 - Case: 9 - Property: PRESS

Figure 224 - GSRM Results, PRESS, year-case: 2006-9
Figure 225- GSRM Results, SO, year-case: 2006-9
Figure 226- GSRM Results, SW, year-case: 2006-9
Year: 2007 - Case: 9 - Property: CO2

Figure 227 - GSRM Results, CO2, year-case: 2007-9
Figure 228 - GSRM Results, PRESS, year-case: 2007-9
Year: 2007 - Case: 9 - Property: SO

Figure 229: GSRM Results, SO, year-case: 2007-9
Figure 230 - GSRM Results, SW, year-case: 2007-9
Figure 231 - GSRM Results, CO2, year-case: 2008-9
Figure 232- GSRM Results, PRESS, year-case: 2008-9
Figure 233 - GSRM Results, SO, year-case: 2008-9
Figure 234 - GSRM Results, SW, year-case: 2008-9
Figure 235: GSRM Results, CO2, year-case: 2009-9
Figure 236 - GSRM Results, PRESS, year-case: 2009-9
Figure 237 - GSRM Results, SO, year-case: 2009-9
Figure 238 - GSRM Results, SW, year-case: 2009-9
Figure 239 - GSRM Results, CO2, year-case: 2010-9
Year: 2010 - Case: 9 - Property: PRESS

Figure 240 - GSRM Results, PRESS, year-case: 2010-9
Figure 241 - GSRM Results, SO, year-case: 2010-9
Figure 242- GSRM Results, SW, year-case: 2010-9
Figure 243 - GSRM Results, CO2, year-case: 2006-10
Figure 244 - GSRM Results, PRESS, year-case: 2006-10
Figure 245: GSRM Results, SO, year-case: 2006-10
Figure 246- GSRM Results, SW, year-case: 2006-10
Figure 247 - GSRM Results, CO2, year-case: 2007-10
Figure 248- GSRM Results, PRESS, year-case: 2007-10
Figure 249- GSRM Results, SO, year-case: 2007-10
Year: 2007 - Case: 10 - Property: SW

Figure 250 - GSRM Results, SW, year-case: 2007-10
Figure 251 - GSRM Results, CO2, year-case: 2008-10
Figure 252- GSRM Results, PRESS, year-case: 2008-10
Figure 253: GSRM Results, SO, year-case: 2008-10
Figure 254- GSRM Results, SW, year-case: 2008-10
Figure 255- GSRM Results, CO2, year-case: 2009-10
Figure 256 - GSRM Results, PRESS, year-case: 2009-10
Figure 257: GSRM Results, SO, year-case: 2009-10
Figure 258 - GSRM Results, SW, year-case: 2009-10
Figure 259 - GSRM Results, CO2, year-case: 2010-10
Figure 260 - GSRM Results, PRESS, year-case: 2010-10
Figure 261- GSRM Results, SO, year-case: 2010-10
Figure 262 - GSRM Results, SW, year-case: 2010-10
Figure 263- GSRM Results, CO2, year-case: 2006-11
Figure 264 - GSRM Results, PRESS, year-case: 2006-11
Figure 265- GSRM Results, SO, year-case: 2006-11
Figure 266 - GSRM Results, SW, year-case: 2006-11
Year: 2007 - Case: 11 - Property: CO2

Figure 267: GSRM Results, CO2, year-case: 2007-11
Year: 2007 - Case: 11 - Property: PRESS

Figure 268 - GSRM Results, PRESS, year-case: 2007-11
Figure 269 - GSRM Results, SO, year-case: 2007-11
Figure 270 - GSRM Results, SW, year-case: 2007-11
Figure 271- GSRM Results, CO2, year-case: 2008-11
Figure 272 - GSRM Results, PRESS, year-case: 2008-11
Figure 273- GSRM Results, SO, year-case: 2008-11
Figure 274- GSRM Results, SW, year-case: 2008-11
Year: 2009 - Case: 11 - Property: CO2

Figure 275 - GSRM Results, CO2, year-case: 2009-11
Figure 276- GSRM Results, PRESS, year-case: 2009-11
Figure 277- GSRM Results, SO, year-case: 2009-11
Figure 278- GSRM Results, SW, year-case: 2009-11
Figure 279 - GSRM Results, CO2, year-case: 2010-11
Figure 280- GSRM Results, PRESS, year-case: 2010-11
Year: 2010 - Case: 11 - Property: SO

Figure 281- GSRM Results, SO, year-case: 2010-11
Figure 282 - GSRM Results, SW, year-case: 2010-11
Figure 283 - GSRM Results, CO2, year-case: 2006-12
Figure 284 - GSRM Results, PRESS, year-case: 2006-12
Figure 285: GSRM Results, SO, year-case: 2006-12
Figure 286 - GSRM Results, SW, year-case: 2006-12
Year: 2007 - Case: 12 - Property: CO2

**Figure 287** - GSRM Results, CO2, year-case: 2007-12
Year: 2007 - Case: 12 - Property: PRESS

Figure 288 - GSRM Results, PRESS, year-case: 2007-12
Figure 289 - GSRM Results, SO, year-case: 2007-12
Figure 290 - GSRM Results, SW, year-case: 2007-12
Figure 291 - GSRM Results, CO2, year-case: 2008-12
Year: 2008 - Case: 12 - Property: PRESS

Figure 292 - GSRM Results, PRESS, year-case: 2008-12
Figure 293 - GSRM Results, SO, year-case: 2008-12
Figure 294 - GSRM Results, SW, year-case: 2008-12
Figure 295: GSRM Results, CO2, year-case: 2009-12
Figure 296 - GSRM Results, PRESS, year-case: 2009-12
Figure 297 - GSRM Results, SO, year-case: 2009-12
Year: 2009 - Case: 12 - Property: SW

Figure 298- GSRM Results, SW, year-case: 2009-12
Figure 299 - GSRM Results, CO₂, year-case: 2010-12
Figure 300 - GSRM Results, PRESS, year-case: 2010-12
Figure 301 - GSRM Results, SO, year-case: 2010-12
Figure 302- GSRM Results, SW, year-case: 2010-12
Figure 303 - GSRM Results, CO2, year-case: 2006-13
Year: 2006 - Case: 13 - Property: PRESS

Figure 304 - GSRM Results, PRESS, year-case: 2006-13
Figure 305: GSRM Results, SO, year-case: 2006-13
Year: 2006 - Case: 13 - Property: SW

Figure 306- GSRM Results, SW, year-case: 2006-13
Year: 2007 - Case: 13 - Property: CO2

Figure 307 - GSRM Results, CO2, year-case: 2007-13
Figure 308 - GSRM Results, PRESS, year-case: 2007-13
Figure 309- GSRM Results, SO, year-case: 2007-13
Year: 2007 - Case: 13 - Property: SW

Figure 310 - GSRM Results, SW, year-case: 2007-13
Figure 311 - GSRM Results, CO2, year-case: 2008-13
Figure 312 - GSRM Results, PRESS, year-case: 2008-13
Year: 2008 - Case: 13 - Property: SO

Figure 3.13- GSRM Results, SO, year-case: 2008-13
Year: 2008 - Case: 13 - Property: SW

Figure 314- GSRM Results, SW, year-case: 2008-13
Figure 315 - GSRM Results, CO2, year-case: 2009-13
Figure 316 - GSRM Results, PRESS, year-case: 2009-13
Figure 317 - GSRM Results, SO, year-case: 2009-13
Year: 2009 - Case: 13 - Property: SW

Figure 318 - GSRM Results, SW, year-case: 2009-13
Figure 319 - GSRM Results, CO2, year-case: 2010-13
Figure 320 - GSRM Results, PRESS, year-case: 2010-13
Figure 321 - GSRM Results, SO, year-case: 2010-13
Figure 322 - GSRM Results, SW, year-case: 2010-13
Figure 323 - GSRM Results, CO2, year-case: 2006-14
Figure 324 - GSRM Results, PRESS, year-case: 2006-14
Figure 325- GSRM Results, SO, year-case: 2006-14
Figure 326- GSRM Results, SW, year-case: 2006-14
Year: 2007 - Case: 14 - Property: CO2

Figure 327- GSRM Results, CO2, year-case: 2007-14
Figure 328- GSRM Results, PRESS, year-case: 2007-14
Figure 329 - GSRM Results, SO, year-case: 2007-14
Figure 330 - GSRM Results, SW, year-case: 2007-14
Year: 2008 - Case: 14 - Property: CO2

Figure 331 - GSRM Results, CO2, year-case: 2008-14
Figure 332- GSRM Results, PRESS, year-case: 2008-14
Figure 333- GSRM Results, SO, year-case: 2008-14
Figure 334 - GSRM Results, SW, year-case: 2008-14
Figure 335 - GSRM Results, CO2, year-case: 2009-14
Year: 2009 - Case: 14 - Property: PRESS

Figure 336- GSRM Results, PRESS, year-case: 2009-14
Figure 337- GSRM Results, SO, year-case: 2009-14
Year: 2009 - Case: 14 - Property: SW

Figure 338 - GSRM Results, SW, year-case: 2009-14
Figure 339 - GSRM Results, CO2, year-case: 2010-14
Figure 340 - GSRM Results, PRESS, year-case: 2010-14
Figure 3.41- GSRM Results, SO, year-case: 2010-14
Year: 2010 - Case: 14 - Property: SW

Figure 342 - GSRM Results, SW, year-case: 2010-14
Year: 2006 - Case: 15 - Property: CO2

Figure 343- GSRM Results, CO2, year-case: 2006-15
Figure 344- GSRM Results, PRESS, year-case: 2006-15
Figure 345: GSRM Results, SO, year-case: 2006-15
Figure 346 - GSRM Results, SW, year-case: 2006-15
Figure 347 - GSRM Results, CO2, year-case: 2007-15
Figure 3.48 - GSRM Results, PRESS, year-case: 2007-15
Figure 349- GSRM Results, SO, year-case: 2007-15
Figure 350- GSRM Results, SW, year-case: 2007-15
Year: 2008 - Case: 15 - Property: CO2

Figure 351- GSRM Results, CO2, year-case: 2008-15
Figure 352 - GSRM Results, PRESS, year-case: 2008-15
Figure 353- GSRM Results, SO, year-case: 2008-15
Figure 354- GSRM Results, SW, year-case: 2008-15
Year: 2009 - Case: 15 - Property: CO2

Figure 355 - GSRM Results, CO2, year-case: 2009-15
Figure 356- GSRM Results, PRESS, year-case: 2009-15
Figure 357 - GSRM Results, SO, year-case: 2009-15
Year: 2009 - Case: 15 - Property: SW

Figure 358 - GSRM Results, SW, year-case: 2009-15
Year: 2010 - Case: 15 - Property: CO2

Figure 359 - GS RM Results, CO2, year-case: 2010-15
Year: 2010 - Case: 15 - Property: PRESS

Figure 360 - GSRM Results, PRESS, year-case: 2010-15
Figure 361- GSRM Results, SO, year-case: 2010-15
Figure 362 - GSRM Results, SW, year-case: 2010-15
Figure 363: GSRM Results, CO2, year-case: 2006-16
Figure 364 - GSRM Results, PRESS, year-case: 2006-16
Figure 365- GSRM Results, SO, year-case: 2006-16
Year: 2006 - Case: 16 - Property: SW

Figure 366 - GSRM Results, SW, year-case: 2006-16
Year: 2007 - Case: 16 - Property: CO2

Figure 367 - GSRM Results, CO2, year-case: 2007-16
Figure 368 - GSRM Results, PRESS, year-case: 2007-16
Figure 369 - GSRM Results, SO, year-case: 2007-16
Figure 370 - GSRM Results, SW, year-case: 2007-16
Figure 371- GSRM Results, CO2, year-case: 2008-16
Figure 372- GSRM Results, PRESS, year-case: 2008-16
Year: 2008  Case: 16  Property: SO

Figure 373- GSRM Results, SO, year-case: 2008-16
Figure 374- GSRM Results, SW, year-case: 2008-16
Year: 2009 - Case: 16 - Property: CO2 Error Histogram

Figure 375- GSRM Results, CO2, year-case: 2009-16
Figure 376 - GSRM Results, PRESS, year-case: 2009-16
Figure 377: GSRM Results, SO, year-case: 2009-16
Figure 378- GSRM Results, SW, year-case: 2009-16
Year: 2010 - Case: 16 - Property: CO2

Figure 379 - GSRM Results, CO2, year-case: 2010-16
Figure 380- GSRM Results, PRESS, year-case: 2010-16
Figure 381- GSRM Results, SO, year-case: 2010-16
Figure 382- GSRM Results, SW, year-case: 2010-16
Year: 2006 - Case: 23 - Property: CO2

Figure 383 - GSRM Results, CO2, year-case: 2006-23
Figure 384: GSRM Results, PRESS, year-case: 2006-23
Year: 2006 - Case: 23 - Property: SO

Figure 385: GSRM Results, SO, year-case: 2006-23
Year: 2006 - Case: 23 - Property: SW

Figure 386- GSRM Results, SW, year-case: 2006-23
Figure 387 - GSRM Results, CO2, year-case: 2007-23
Figure 388- GSRM Results, PRESS, year-case: 2007-23
Figure 389- GSRM Results, SO, year-case: 2007-23
Figure 390- GSRM Results, SW, year-case: 2007-23
Figure 391 - GSRM Results, CO2, year-case: 2008-23
Figure 392 - GSRM Results, PRESS, year-case: 2008-23
Figure 393 - GSRM Results, SO, year-case: 2008-23
Figure 394- GSRM Results, SW, year-case: 2008-23
Figure 395- GSRM Results, CO2, year-case: 2009-23
Year: 2009 -Case: 23 - Property: PRESS

Figure 396- GSRM Results, PRESS, year-case: 2009-23
Figure 397 - GSRM Results, SO, year-case: 2009-23
Figure 398- GSRM Results, SW, year-case: 2009-23
Figure 399- GSRM Results, CO2, year-case: 2010-23
Figure 400 - GSRM Results, PRESS, year-case: 2010-23
Year: 2010 - Case: 23 - Property: SO

Figure 401 - GSRM Results, SO, year-case: 2010-23
Figure 402 - GSRM Results, SW, year-case: 2010-23