SMART CFD PROXY

Reproducing CFD Results Accurately & at High Speed

Shahab D. Mohaghegh¹, Mehrdad Shahnam², Amir Ansari¹,
Ali Takbiri¹, Ebrahim Fathi¹, and Jean F. Dietiker²

¹West Virginia University
²DOE-NETL

August 2016
Outline

• Smart Proxy
  • Engineering Application of Big Data Analytics

• Original Application in Reservoir Modeling
  • Tracking Pressure & Saturation Changes at the Grid Block Level

• Application to CFD
  • Multi-Phase Fluidized Bed (Proof of Concept)
Computational Science

• Numerical solution to complex, multi-physic, non-linear, partial differential equation emerged, when
  • Engineering models became too complex to be solved analytically,
  • High speed computers became widely available.
**Computational Footprint**

- Our models soon increased in size such that use of super computers and CPU clusters became a necessity to solve serious problems.

- Proxy models are used to address the extensive computational cost of the numerical simulation models.
Proxy Models

• Proxy models attempt to behave like the original models but with lower computational cost.

• The current paradigm for developing proxy models is to simplify the original model.

• Existing proxy models are either:
  • Statistics-based (response surfaces – curve fit at specific locations)
  • Physics-based (Reduced Physics)
  • Coarse Models (low resolution)
Paradigm Shift in Proxy Modeling

- In this new paradigm of developing proxy models,
  - The proxy model represents the entire system.
  - Physics of the model is **not** simplified.
  - Model resolution is **not** reduced.

- In this new paradigm, data (generated by CFD), and not the physics, will be the basis for the proxy model.
**Paradigm Shift in Proxy Modeling**

- Smart Proxy uses the pattern recognition capabilities of Artificial Intelligence and Data Mining (Big Data Analytics) to **LEARN** and then **MIMIC** the behavior of the original CFD.

- Smart Proxy is trained using massive amount of data that is generated from (only a handful of) CFD models runs and validated using blind CFD runs.
Paradigm Shift in Proxy Modeling

• A single run of the original model is treated as a treasure trove of data and information that can be used to train and calibrate a new, smart model.

• This technology is the manifestation of “BIG DATA ANALYTICS” in engineering problem solving, specifically, in the numerical modeling arena.
Surrogate Reservoir Modeling - SRM

Smart Proxy Model for high speed replication of the results of numerical reservoir simulator
Numerical Model Characteristics

- Study Completed: 2005
- Implemented: 2006
- Field Results Published: 2014 (SPE 170664)

- Size of the Full Field Model: 1MM Grid Blocks
- Formation Type: Naturally Fractured Carbonate
- Number of Wells: 167 (Horizontal)
- Simulation Run: 10 Hours on 12 parallel CPU
Field Production Optimization Study
Field Production Optimization Study
Black Oil Simulation
Green Field
6.5 MM Grid Blocks
Oil Rate (bbls/yr)

Cum. Oil Production (MMbbls)

Oil Rate for well: VRT0019; BHP = vars; Qmax = 15M

SRM
Powers

Black Oil Simulation
Green Field
6.5 MM Grid Blocks
Otway CO$_2$ Seq. Project, Australia

Permeability Maps

History Matched Perm

High perm

Low perm
Otway CO₂ Seq. Project, Australia

Pressure Distribution
Otway CO₂ Seq. Project, Australia

Water Saturation Distribution
SACROC CO₂ EOR Project, Texas

The Coarse Reservoir Model
SACROC CO₂ EOR Project, Texas

Water Saturation Distribution
**SACROC CO₂ EOR Project, Texas**

**CO₂ Mole Fraction Distribution**

Actual Data: Realization # 8, Layer= 18, 100 yrs after Injection, Feature= Pressure (psi)

SRM Data: Realization # 8, Layer= 18, 100 yrs after Injection, Feature= Pressure (psi)

Actual Data: Realization # 9, Layer= 18, 100 yrs after Injection, Feature= Pressure (psi)

SRM Data: Realization # 9, Layer= 18, 100 yrs after Injection, Feature= Pressure (psi)

Actual Data: Realization # 10, Layer= 18, 100 yrs after Injection, Feature= Pressure (psi)

SRM Data: Realization # 10, Layer= 18, 100 yrs after Injection, Feature= Pressure (psi)
SACROC CO₂ EOR Project, Texas

Pressure Distribution

Simulator (CMG) Results

Actual Data: Realization # 9, Layer= 18, 9 years after Injection, Feature= CO₂ Mole Fraction %

SRM Results

Actual Data: Realization # 10, Layer= 18, 9 years after Injection, Feature= CO₂ Mole Fraction %

Absolute Error

Actual Data: Realization # 10, Layer= 18, 100 years after Injection, Feature= Absolute Error %

Actual Data: Realization # 7, Layer= 18, 9 years after Injection, Feature= CO₂ Mole Fraction %

SRM Data: Realization # 9, Layer= 18, 9 years after Injection, Feature= CO₂ Mole Fraction %

SRM Data: Realization # 10, Layer= 18, 9 years after Injection, Feature= CO₂ Mole Fraction %

SRM Data: Realization # 7, Layer= 18, 9 years after Injection, Feature= CO₂ Mole Fraction %

Absolute Error

Realization # 7, Layer= 18, 100 years after Injection, Feature= Absolute Error %
SACROC CO$_2$ EOR Project, Texas

High Resolution Model
SACROC CO$_2$ EOR Project, Texas

High Resolution Model
SACROC CO₂ EOR Project, Texas

High Resolution Model
Can Smart Proxy be successfully applied to CFD?
A Multi-phase CFD Model

- Initial pressure is 1 atm. in the entire domain
  1 atm. = 101,325 Pa

<table>
<thead>
<tr>
<th>Grid Classification</th>
<th>I, j, k - Cell size</th>
<th>No. of Cell</th>
<th>No. of Grid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coarse</td>
<td>8 x 48 x 8 (15 mm)</td>
<td>3,072</td>
<td>3,969</td>
</tr>
<tr>
<td>Medium</td>
<td>12 x 72 x 12 (10 mm)</td>
<td>10,368</td>
<td>12,337</td>
</tr>
<tr>
<td>Fine</td>
<td>18 x 108 x 18 (6.6 mm)</td>
<td>34,992</td>
<td>39,349</td>
</tr>
<tr>
<td>Very Fine</td>
<td>27 x 162 x 27 (4.4 mm)</td>
<td>118,098</td>
<td>127,792</td>
</tr>
</tbody>
</table>
Three different physics

Early-time
Time step 200

Mid-time
Time step 1000

Late-time
Time step 4000
Early-time
Time step 200

Five cross-sections used to show details in 2-D
Mid-time
Time step 1000
Late-time

Time step 4000

Gas Fraction at time step=4000

Layer=1

Layer=7

Layer=14

Layer=21

Layer=27

K = 1

K = 14

K = 21

K = 27
Data from the CFD Runs

Spatio-Temporal Database Generated to be used by the Smart CFD Proxy

Key factors of multi-phase gas-solid system

<table>
<thead>
<tr>
<th>Factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume fraction ($\varepsilon$)</td>
<td>Gas Density ($\rho$)</td>
</tr>
<tr>
<td>Velocity vector of gas ($u, v, w$)</td>
<td>Particle diameter ($d$)</td>
</tr>
<tr>
<td>Velocity vector of solid ($u, v, w$)</td>
<td>Maximum packing volume fraction ($\varepsilon^*$)</td>
</tr>
<tr>
<td>Pressure field of gas ($P$)</td>
<td>Location to the boundaries</td>
</tr>
<tr>
<td>Pressure field of solid ($P$)</td>
<td>Location to the interface ($x, y, z$)</td>
</tr>
<tr>
<td>Time ($t$)</td>
<td>Viscosity ($\mu$ and $\lambda$)</td>
</tr>
<tr>
<td>Viscosity ($\mu$ and $\lambda$)</td>
<td>Gravity force ($g$)</td>
</tr>
</tbody>
</table>

Data from the CFD Runs
Displaying Smart CFD Proxy

• Results are displayed in 5 different cross-sections:

• \( K = 1, 7, 14, 21, 27 \)
Thank you
Part One – Non-Cascading Results

• In Non-Cascading scheme at every time-step actual CFD results are used as input and the trained Smart Proxy forecasts the results for the next time-step.

• No error accumulation takes place in this scheme.

• This scheme, as the first step, explores the learning capabilities of the Smart Proxy.
Early Time Results

• Smart Proxy is generated for the early time in the system when the bubble behavior is reasonably calm and non-chaotic.
Results of Layer 1 – Gas Fraction

• Model was trained using Time-step 101 as input and Time-step 102 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 102 through Time-step 126

• Results generated by the Smart Proxy are compared with actual CFD output.
Results of Layer 27 – Gas Fraction

- Model was trained using Time-step 101 as input and Time-step 102 as output.
  - Trained on 75% of data and validated using the remaining 25%

- Trained model was deployed in forecast mode and generated Time-step 102 through Time-step 126

- Results generated by the Smart Proxy are compared with actual CFD output.
Multiple bubbles in the system

• The process is now repeated when several bubbles have already been formed in the system.
Results of Layer 1 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4001 through Time-step 4040

• Results generated by the Smart Proxy are compared with actual CFD output.
Gas Fraction of Layer=1 at time step=4006

CFD Result

Smart Proxy Result

Error Distribution (%)
Results of Layer 7 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4001 through Time-step 4040

• Results generated by the Smart Proxy are compared with actual CFD output.
Results of Layer 14 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4001 through Time-step 4040

• Results generated by the Smart Proxy are compared with actual CFD output.
Gas Fraction of Layer=14 at time step=4014
Results of Layer 21 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4001 through Time-step 4040

• Results generated by the Smart Proxy are compared with actual CFD output.
Results of Layer 27 – Gas Fraction

- Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  - Trained on 75% of data and validated using the remaining 25%

- Trained model was deployed in forecast mode and generated Time-step 4001 through Time-step 4040

- Results generated by the Smart Proxy are compared with actual CFD output.
Part Two – Cascading Results

• In Cascading scheme at every time-step results from the Smart Proxy are used as input and the trained Smart Proxy forecasts the results for the next time-step. (Smart Proxy receives feedback from itself)
  • Models are developed for all output variables in CFD such as Void Fraction, and gas/solid Pressure and Velocity.

• In this scheme the error has the opportunity to accumulate.

• Limits of the information content of each time-step is explored in this scheme.
Early Time Results

• Smart Proxy is generated for the early time in the system when the bubble behavior is reasonably calm and non-chaotic.
Results of Layer 1 – Gas Fraction

• Model was trained using Time-step 101 as input and Time-step 102 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 101 through Time-step 105

• Results generated by the Smart Proxy are compared with actual CFD output.
Results of Layer 7 – Gas Fraction

• Model was trained using Time-step 101 as input and Time-step 102 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 101 through Time-step 105

• Results generated by the Smart Proxy are compared with actual CFD output.
Gas Fraction of Layer=7 at time step=104

CFD Result

Smart Proxy Result

Error Distribution (%)
Results of Layer 14 – Gas Fraction

• Model was trained using Time-step 101 as input and Time-step 102 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 101 through Time-step 105

• Results generated by the Smart Proxy are compared with actual CFD output.
Gas Fraction of Layer=14 at time step=104

Comparison of CFD Result and Smart Proxy Result with Error Distribution (%).
Results of Layer 21 – Gas Fraction

• Model was trained using Time-step 101 as input and Time-step 102 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 102 through Time-step 126

• Results generated by the Smart Proxy are compared with actual CFD output.
Gas Fraction of Layer=21 at time step=103

CFD Result

Smart Proxy Result

Error Distribution (%)

y-division

x-division

y-division

x-division

y-division

x-division
Results of Layer 27 – Gas Fraction

- Model was trained using Time-step 101 as input and Time-step 102 as output.
  - Trained on 75% of data and validated using the remaining 25%

- Trained model was deployed in forecast mode and generated Time-step 101 through Time-step 105

- Results generated by the Smart Proxy are compared with actual CFD output.
Gas Fraction of Layer=27 at time step=103
Multiple bubbles in the system

- The process is now repeated when several bubbles have already been formed in the system.
Results of Layer 1 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4002 through Time-step 4014

• Results generated by the Smart Proxy are compared with actual CFD output.
Gas Fraction of Layer=1 at time step=4010

CFD Result

Smart Proxy Result

Error Distribution (%)
Results of Layer 7 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4002 through Time-step 4010

• Results generated by the Smart Proxy are compared with actual CFD output.
Gas Fraction of Layer=7 at time step=4006

CFD Result

Smart Proxy Result

Error Distribution (%)

y-division

x-division
Results of Layer 14 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4002 through Time-step 4024

• Results generated by the Smart Proxy are compared with actual CFD output.
Results of Layer 21 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4002 through Time-step 4020

• Results generated by the Smart Proxy are compared with actual CFD output.
Results of Layer 27 – Gas Fraction

• Model was trained using Time-step 4001 as input and Time-step 4002 as output.
  • Trained on 75% of data and validated using the remaining 25%

• Trained model was deployed in forecast mode and generated Time-step 4002 through Time-step 4012

• Results generated by the Smart Proxy are compared with actual CFD output.
Paradigm Shift in Proxy Modeling

- **Very Large Amounts of Data** is carefully extracted from a handful of runs of the original model.

- The process of data extraction from several runs of the original model is **engineered** such that it can be used to train, calibrate and validate a smart, new model.
Paradigm Shift in Proxy Modeling

• The Smart Proxy model is developed using MACHINE LEARNING that is now the main paradigm for treating and handling “BIG DATA”.

• The Smart Proxy learns to mimic the behavior of the original model with all its complexities and intricacies.

• Once developed, the Smart Proxy Model runs at very low computational cost (thousands of simulation runs in minutes).
The Spatio-Temporal Database

Tier One
Neighbor Cells

<table>
<thead>
<tr>
<th>Cell$\text{ijk}$ at time step $n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_g$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tier Cell$\text{ijk}$ at time step $n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_g$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cell$\text{ijk}$ at time step $(n+1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_g$</td>
</tr>
</tbody>
</table>

Number of inputs $6 + 9 + 6 \times 9 = 69$
Number of outputs = 9
Optimize the number of inputs

EP_self → \( W_1 \) → Hidden → Output
V_g_self → \( W_2 \) → Hidden → Output
V_s_self → \( W_3 \) → Hidden → Output
P_g_self → \( W_4 \) → Hidden → Output
P_s_self → \( W_5 \) → Hidden → Output
EP_top → Hidden → Output
Tornado chart by adding all the weights

Tornado Chart of Neural Network Weights

[Graph of Tornado chart showing weights]
**Eliminating 14 parameters**

<table>
<thead>
<tr>
<th>By parameter</th>
<th>By location</th>
</tr>
</thead>
<tbody>
<tr>
<td>ug-west</td>
<td>vg-self</td>
</tr>
<tr>
<td>ug-north</td>
<td>vs-bottom</td>
</tr>
<tr>
<td>ug-south</td>
<td>ug-north</td>
</tr>
<tr>
<td>ug-east</td>
<td>wg-north</td>
</tr>
<tr>
<td>vg-self</td>
<td>vs-bottom</td>
</tr>
<tr>
<td>vg-west</td>
<td>ug-south</td>
</tr>
<tr>
<td>wg-north</td>
<td>ws-west</td>
</tr>
<tr>
<td>us-north</td>
<td>vs-west</td>
</tr>
<tr>
<td>ws-west</td>
<td>ug-west</td>
</tr>
<tr>
<td>vs-west</td>
<td>vg-west</td>
</tr>
<tr>
<td>vs-bottom</td>
<td>ug-east</td>
</tr>
<tr>
<td>distance to south</td>
<td>distance to south</td>
</tr>
<tr>
<td>distance to east</td>
<td>distance to east</td>
</tr>
<tr>
<td>distance to west</td>
<td>distance to west</td>
</tr>
</tbody>
</table>
Comparison between RMSE

- RMSE when 70 parameters were used
- RMSE when 56 parameters were used
- RMSE when 35 parameters were used

RMSE

Time Step

0 500 1000 1500 2000 2500 3000 3500 4000 4500
The Tier System
The Tier System

Tier One
Interface: Plane

Tier Two
Interface: Line

Tier Three
Interface: Point
**Tier One - Subsystem**

- **Tier One Interface: Plane**
- **Tier One Two Subsystems**
  - Tier $1_{\text{Input}}$ and Tier $1_{\text{Output}}$
- **Tier One Three Subsystems**
  - Tier $1_{\text{EW}}$, Tier $1_{\text{SN}}$, and Tier $1_{\text{BA}}