INTELLIGENT SOLUTIONS, INC.

OILFIELD DATA MINING

IMPLEMENTING THE PARADIGM SHIFT IN ANALYSIS & MODELING OF THE OILFIELD

55 TARA PLACE
MORGANTOWN, WV 26050
USA

OILFIELD DATA MINING

OVERWHELMED WITH DATA BUT STARVED OF KNOWLEDGE

Intelligent Solutions, Inc. is a pioneer in application of Artificial Intelligence and Data Mining (AI&DM) in the upstream oil and gas industry. With more than 18 years of direct experience in building smart applications for the industry using AI&DM, Intelligent Solutions, Inc. is the premier provider of smart products, services and training. ISI's contribution to the art and science AI&DM in the petroleum & natural gas engineering is documented in more than 100 technical articles in conference proceedings and peer-reviewed journal articles.

Due to the public disclosure of many of our technologies in related technical articles that have been done with the aim of promotion of AI&DM in the industry and dissemination of useful knowledge among the professionals in the industry, ISI's technology is widely copied[1] and used by major service companies (in their products and services) and by many consulting outfits.

DATA MINING

Data mining is the process of extracting hidden patterns from data. As more data is gathered in the oilfield, data mining is becoming an increasingly important tool to transform this data into information. Data Mining is commonly used in a wide range of applications, such as marketing, fraud detection and scientific discovery. Use of data mining in the oilfield dates back to the early 1990s. Data mining can be applied to data sets of any size. However, while it can be used to uncover hidden patterns in data that has been collected, obviously it can neither uncover patterns which are not already present in the data, nor can it uncover patterns in data that has not been collected.

Data mining has been defined as "The nontrivial extraction of implicit, previously unknown, and potentially useful information from data." It uses artificial intelligence, machine learning, statistical and visualization techniques to discover and present knowledge in a form which is easily comprehensible to humans.

It is becoming more and more evident that an oilfield operator can create strategic advantages over its competitors by making use of data mining to get important insights from the data it collects. Different oilfield operators collect and analyze data differently. So an oilfield operator can easily gain competitive values over others by using data mining.

To make better decisions one needs to discover and understand the underlying patterns involved in the particular operation from the data. For example, it's not enough for a production engineer to know just the amount of oil and/or gas production from a field and the amount of CAPEX and OPEX for company in this highly competitive business environment. To increase recovery and achieve higher production the production engineer has to search for answers to the questions like:

- Which service company should be used more often for better results?

[1] Unfortunately many times without properly referencing the source of the technology.

- How to balance the quality of an intervention job with the cost?
- How to select the best candidate wells for stimulation?
- What would be the best stimulation design for a particular well?

Data mining has the potential to answer many such questions. Data mining can help oilfield operators to have useful insights into their operations from the data they have collected over the years and make better decisions to achieve new heights.

ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) is the area of computer science focusing on creating machines that can engage in behaviors that humans consider intelligent. The ability to create intelligent machines has intrigued humans since ancient times and today with the advent of the computer and 60 years of research and development into AI techniques, the dream of smart machines is becoming a reality. Researchers are creating systems which can mimic human thought, understand speech, beat the best human chess-player, compete and win "Jeopardy[2]" contests and countless other feats never before possible.

Artificial Intelligence (AI) is a combination of computer science, physiology, and philosophy. AI is a broad topic, consisting of different fields, from machine vision to expert systems. The element that the fields of AI have in common is the creation of machines that can "think". In order to classify machines as "thinking", it is necessary to define intelligence. To what degree does intelligence consist of, for example, solving complex problems, or making generalizations and relationships? And to what degree does intelligence consist of perception and comprehension?

Research into the areas of learning, of language, and of sensory perception has aided scientists in building intelligent machines. One of the most challenging approaches facing experts is building systems that mimic the behavior of the human brain, made up of billions of neurons, and arguably the most complex matter in the universe. Perhaps the best way to gauge the intelligence of a machine is British computer scientist Alan Turing's test. He stated that a computer would deserve to be called intelligent if it could deceive a human into believing that it was human. May be the most recent winning competition in the U.S. television show called "Jeopardy" by a IBM made computer called "Watson" can be counted as a successfully passed the Turing test.

DATA-INTENSIVE SCIENCE, THE FOURTH PARADIGM

History of science and technology can be divided into several eras (Hey, 2009). It all started with experimental science at the early age of science. Several hundred years ago the theoretical branch of science emerged and gave rise to theories such as Newton's laws of motion, Kepler's laws of planetary motion and Maxwell's laws of electrodynamics, optics and electric circuits. The last several decades have been the age of computational science where fast computers have provided the means

[2] A television game show based on general knowledge of the contestants.

for simulation and modeling in areas such as computational fluid dynamics, meteorological and climatological sciences, aerospace and hydrocarbon reservoir simulations, to name a few. According to Jim Gray[3], the legendary American computer scientist, we have now entered the new age of *escience* or *data-intensive science* where massive amounts of data can be collected from physical phenomena and or simulations and new models can be built based on these data.

Moving from each of the above ages of science to the next required a paradigm shift on how we observe, interact, model and attempt to control the phenomena around us. It is now time for another paradigm shift into the fourth paradigm that is the *data-intensive science*.

ARTIFICIAL INTELLIGENCE & DATA MINING IN THE E&P INDUSTRY

Since the early 1990s Intelligent Solutions, Inc. has been focused on promoting intelligent technologies in the oil and gas industry by developing state of the art solutions and tackling important and challenging problems. This is evident from more than 100 technical articles that we have contributed to the art and science of application of AI&DM in the exploration and production industry. Following are examples of the types of problems that are now solved regularly using the technologies that have been pioneered by Intelligent Solutions, Inc.

INTELLIGENT WELL STIMULATION

Intelligent Solutions, Inc. has been able to successfully develop predictive models for well stimulation in some of the most notoriously complex fields in the world. Our Intelligent Well Stimulation technology makes it possible to:

- Build and validate predictive well stimulation models by coupling reservoir quality to the stimulation design.
- Design optimum stimulation or workover jobs for individual wells based on their location (reservoir quality) in the field.
- Identify the Best Stimulation Practices based on historical data.
- Identify the best candidate wells for stimulation and workover at any given budget period.

INTELLIGENT WELL STIMULATION – CASE STUDY

In a recent study ISI's Intelligent Well Stimulation technology was applied to a complex mature carbonate reservoir in Mexico. This onshore asset includes 15 fields with a large number of well stimulation jobs performed by four separate service companies. The inconsistency of well stimulation results and the apparent lack of correlation between stimulation outcome with job design, well location, geology and reservoir characteristics, cost of the stimulation and the service company that

[3] Jim Gray: (1944-2007) Legendary American computer scientist received the Turing Award for seminal contributions to computer science.

performs the stimulation had reduced the stimulation and workover operations in this asset to a random, hit and miss process. The objective of ISI's study was to make sense of the well stimulation and workover operations in this asset and provide a predictive model based on the historical data. The predictive model was to identify the best candidate wells for stimulation in each budget period while identifying the optimum stimulation design for the stimulation operation. The fact that record keeping of each service company was different and that they would report the stimulation design and implementation in a non-standard fashion and lack of robust data gathering and data maintenance by the operator created special challenges for the ISI team during the implementation of this project.

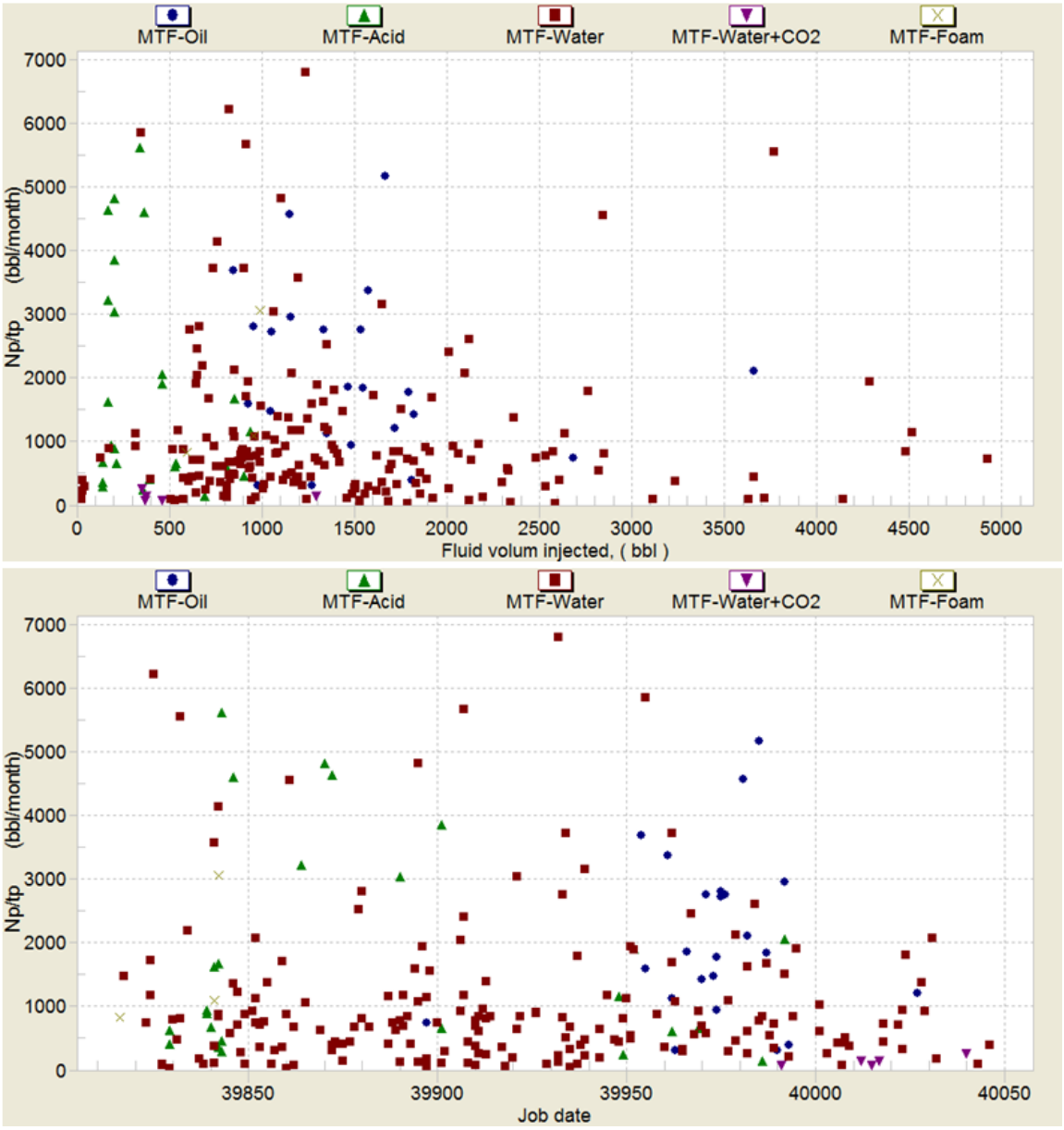

Figure 1. *Lack of apparent pattern in the results of stimulation jobs in this asset are clear from these figures. The plot on the top shows the normalized production versus fluid volume injected and the plot on the bottom shows the normalized production versus job date for five different classes of injected fluids.*

Figure 1 demonstrates the lack of apparent correlations and patterns in the results of stimulation jobs reported by four different service companies in this asset. In this figure normalized production for each well is plotted against injected fluid volume (top) and job date (bottom) for five different fluid types used during the stimulation of wells in this asset.

Figure 2 shows the comparison between ISI's data driven modeling technology using IMprove™ software application with multiple regressions that is usually used for statistical analysis. In this figure actual normalized production from the field (y axis) is plotted against the model prediction (x axis). It shows that using ISI's predictive modeling technology can significantly improve the chances of building successful predictive models. While multiple regression model generates a model with predictive capability of 0.29 (1.0 being perfect predictions), ISI's model provides predictive capability of 0.89. From this figure it is evident that the statistical analysis grossly underestimates the results of the stimulation jobs and is never able to predict normalized production of over 3000 bbls/month while ISI's predictive model provides reasonable prediction of well stimulation results of up to 6000 bbls/month.

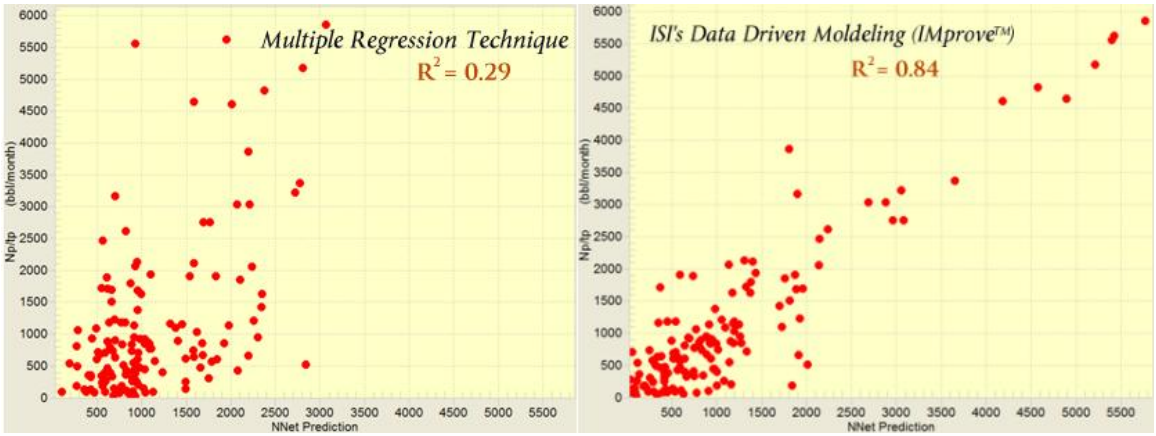

Figure 2. *Comparing ISI's data driven modeling technology using IMprove™ software application with multiple regressions that is used for statistical analysis. In both cases above, actual normalized production from the field (y axis) is plotted against the model prediction (x axis). Perfect prediction would result in* R^2 *of 1.0.*

Once a predictive model is trained and validated using a blind data set, it is time for using the predictive model as a tool for managing the production optimization in the field. The production optimization is a multi-step process that includes identification of candidate wells for the stimulation for the current budgetary period and design of optimum stimulation jobs for the candidate wells. While all the wells in the field are being evaluated in order to identify the best candidate wells for the stimulation it is necessary to perform uncertainty analysis. Since reservoir characteristics have been coupled with stimulation practices during the model building process, and since they are inherently uncertain by nature, therefore, uncertainty analysis becomes a necessary part of the analysis.

Please remember that inclusion of reservoir characteristics in the predictive model was a necessary step during the model building process. This necessary coupling of reservoir characteristics with stimulation practices addresses the important issue of dependence of well productivity to reservoir characteristics as well as the stimulation practices and modeling the impact of both of these sets of parameters in the predictive model. This is due to the fact that sometimes the

unresponsiveness of a well to stimulation practices may have nothing to do with the stimulation and is a result of reservoir characteristics. On the other hand, since reservoir characteristics have been included in the model and since they are inherently uncertain values, uncertainty analysis becomes an important part of the overall analyses for the identification of best candidate wells.

IMprove™ (ISI's software product for stimulation modeling and analysis) includes a module that accommodates uncertainty analysis using Monte Carlo simulation technique. As shown in Figure 3, multiple parameters (usually reservoir characteristics) can be identified as parameters for the uncertainty analysis and Monte Carlo simulation can be performed in order to quantify the impact of the uncertainties that are associated with these parameters on the well productivity. In Figure 3 well productivity is shown in the "x" axis while relative frequency (left) and cumulative probability (right) are shown in the "y" axis. In this figure it is evident that this particular well has the capability of producing between 1,700 to 2,200 bbls/month (the range with the highest probability) while a minimum of less than 1,000 bbls/month and a maximum of 4,500 bbls/month are the limits of the productivity for this well. Similar analyses can be performed for all the wells in the field, each taking only a few seconds to be completed.

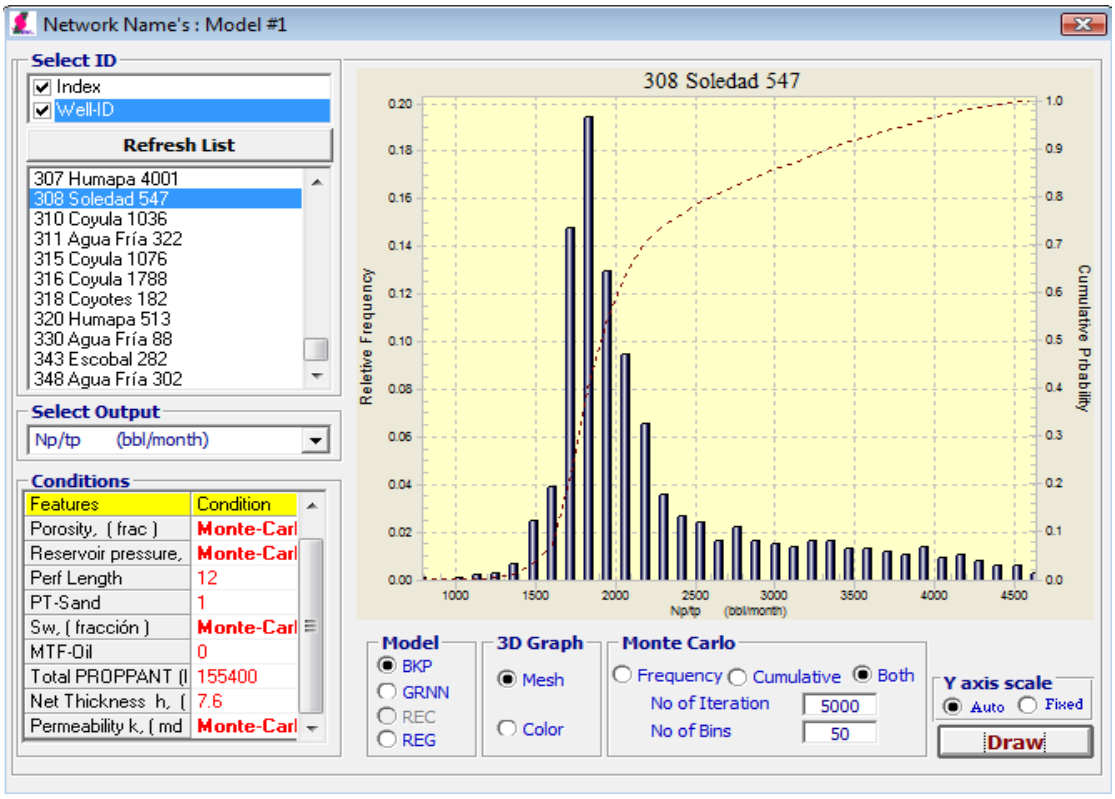

Figure 3. *Uncertainty analysis using Monte Carlo simulation technique in IMprove™.*

Upon completion of the uncertainty analysis for all the wells and consequently identification of best candidate wells through a well designed workflow, optimum stimulation design needs to be performed for each of the candidate wells. IMprove™ provide a module that can help engineers in designing optimum stimulation design for each of the wells. It should be noted that as expected, the optimum stimulation design is not unique for all wells and will differ from well to well accommodating the unique geology and reservoir characteristics (stresses in the field) of each

individual well. Meaning that a "one size fit all" approach that is commonly applied by many service companies may not necessarily be the best approach to increasing production from wells in a field.

The module in IMprove™ that accommodates this functionality is called the Type Curve module. In this module impact of two different parameters (at a time) involved in the predictive model can be shown on the normalized well productivity as type curves. The type curves can be generated for each well separately in a fraction of a second. For example in Figure 4 normalized well productivity (y axis) is plotted against injection rate (x axis) for each well when the formation porosity is changed from 5% to 8% for two different wells (graphs on the top). These curves clearly display the sensitivity of the stimulation job's outcome on formation porosity while all other parameters such as stimulation design are kept constant. Graphs on the bottom of this figure show normalized well productivity (y axis) that is plotted against injection rate (x axis) for different wells as a function of total amount of proppant used in the stimulation jobs (from 200,000 lbs to 350,000 lbs). Type curves are used as an effective tool for designing best possible stimulation jobs for each individual well.

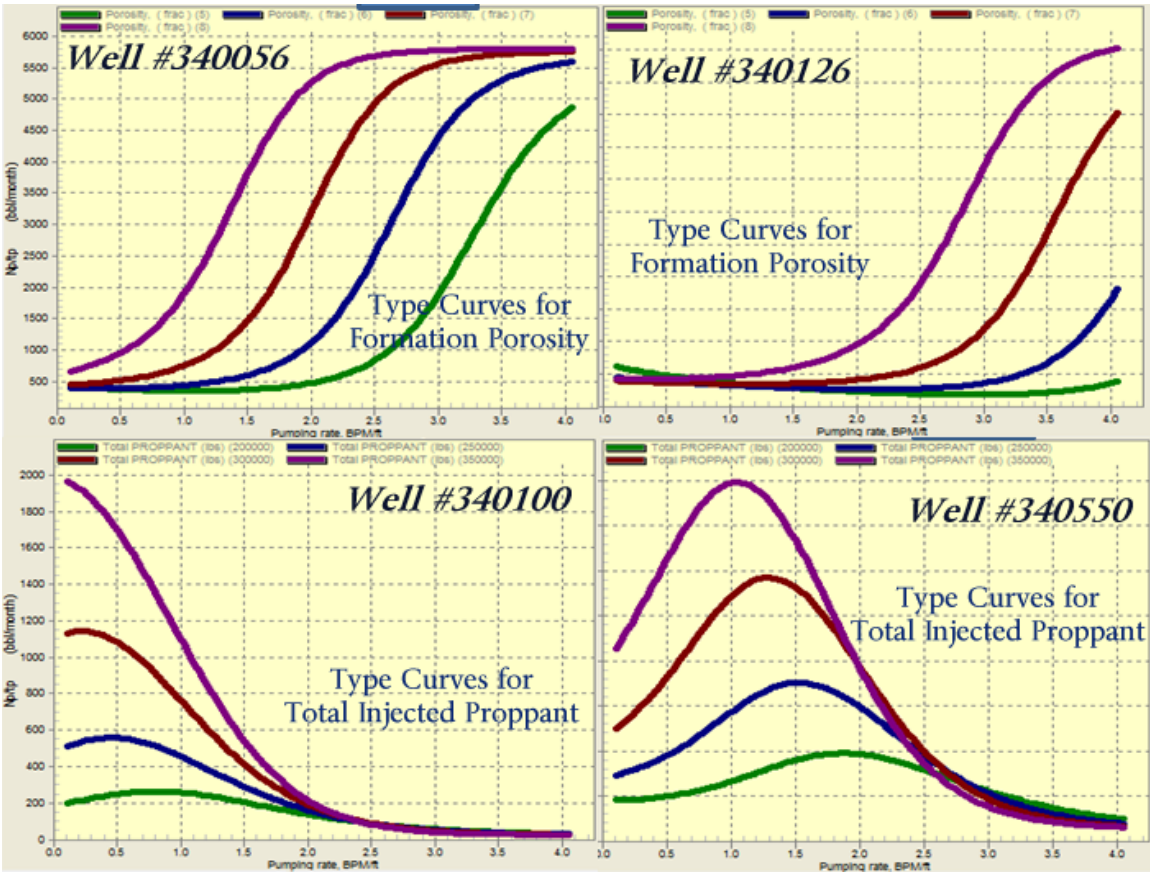

Figure 4. *Generating multiple type curves for each well in IMprove™. Impact of formation porosity, total injected proppant and injection rate on the well productivity.*

This technology has effectively been used to help many operators make the most of their stimulation budget. This technology has been turned into a complete workflow and is currently available in ISI's software product called IMprove™.

REFERENCES

- Tony Hey, et.al. 2009, The Fourth Paradigm; (Data-Intensive Scientific Discover), Microsoft Research.