HARNESSING DATA SCIENCE FOR RESERVOIR CHARACTERIZATION & FIELD DEVELOPMENT DECISION-MAKING

WHILE IMPROVING THE INPUT DATA ON WHICH IT WILL FEED

PREMIER
OILFIELD GROUP

HARNESSING DATA SCIENCE FOR RESERVOIR CHARACTERIZATION AND FIELD DEVELOPMENT DECISION-MAKING

WHILE IMPROVING THE INPUT DATA ON WHICH IT WILL FEED

PREMIER
OILFIELD GROUP

"*It's difficult to imagine the power that you're going to have when so many different sorts of data are available.*"

- Tim Berners-Lee

The upstream oil and gas industry generates a lot of data, and data science promises to unlock tremendous value through its analysis. In our opinion, the data-driven oil and gas industry is here. Multi-disciplinary, high-resolution models are being used to make field development decisions. However, subsurface data are often sparse, difficult to interpret, and scattered across different storage media and locations. We urgently need new sources of data and more efficient, more accurate measurements to reduce uncertainties and make our models more effective.

In this paper we discuss faster, lower-cost approaches to core analysis, drill cuttings as an under-appreciated source of rock property data, and new approaches to generating and sharing data that can lead to transformative performance for the companies willing to adopt and embrace them.

DATA, DATA EVERYWHERE…

Without question, the oil and gas industry generates a *lot* of data. New petabytes are recorded and transmitted each day from myriad sensors connected to everything from seismic surveying equipment to brownfield beam pumps. And yet, in contrast to life-changing data science applications in the consumer sector, oil and gas companies remain largely in the dark about what much of our data is telling us.

We can point an accusing finger in various directions.

The data we gather is still sparse relative to the scale of subsurface heterogeneity we're trying to understand. A petabyte of data might seem unfathomably large to those of us who grew up with computers boasting *kilobytes* of RAM, but it doesn't come close to fully describing the idiosyncrasies of several cubic miles of naturally varying strata.

Many of the measurements are made from the surface, far from the interval of interest, or by instruments confined to narrow boreholes that can only investigate a few feet into the surrounding rock. The gaps in between these measurements must be filled using geostatistics, which when stretched too far amounts to little more than educated guesswork.

The data are difficult to interpret, especially when they are inferred from proxy measurements and only calibrated to a scattering of direct measurements made on samples recovered from the subsurface. Although we're very good at building correlations and working with analogs, this lack of ground-truth measurements introduces serious uncertainties into our models and simulations.

There are difficulties from a data science point of view, too. Learning algorithms rely on thousands of training runs for which the outcome is known and has been classified by a human operator. In the B2C environment, a similar event or transaction can be observed millions of times and the outcome of each instance logged and classified. In short order, a deep learning algorithm can spot complex patterns within the data and use them to predict how future transactions will unfold.

Financial and internet data – two industries at the leading edge of data science - usually comprise a very large number of small data sets generated at high frequency. Oil and gas data, on the other hand, are often both space and time dependent and arrive in a multitude of complex formats, such as arrays and images. We measure large and complex sets of data related to a very *limited* number of transactions – and the outcome of each transaction can be difficult to measure and characterize.

For example, we can assemble a detailed collection of geological data to describe where a new well will be drilled but the resulting performance of that well – something we would dearly love to predict with accuracy – is difficult to quantify, and the overall number of wells being drilled and characterized (each of which represents a transaction) is, in the data science world, very low.

Finally, because many oil and gas decisions are made comparatively slowly – typically hours or days compared to financial transactions that take microseconds – our industry has prioritized data processing speed over data access. This means that oil and gas data is scattered across different storage media and locations, some of which are not easy to connect and integrate in real time.

Does this mean that data science offers no value to reservoir characterization and field development? On the contrary, of course it does. We need to think carefully about how and when it can best be applied, as well as how to improve the input data on which it will feed.

EVOLUTION OF DECISION MAKING

In reality, the data-driven oil and gas industry is already here. Even though data science isn't able to live up to all of the hype – much of it blatant extrapolation from far easier-to-predict sectors – data-driven decision making is marching inexorably across the oil and gas patch.

Barely three decades ago, most field development ideas were tested directly in the field – by "turning to the right", as the old drilling expression goes. Learning was almost entirely empirical.

As computers became available to academic institutions and businesses, single-well simulations became possible. Many simplifications had to be made for them to run but, nevertheless, options could be screened, and some decisions made before testing in the field.

Throughout the 1990s and early 2000s, rapid improvements in processing speed and power enabled more complex, higher-resolution modeling and simulation. Physical processes could be represented explicitly, without requiring excessive simplification. Larger models featuring many layers or multiple wells could be constructed and used to investigate more complex behavior.

That trend has continued, leveraging ever-larger clusters of computing power made available to all through the advent of cloud computing. This has facilitated the coupling of various models and simulations across the industry's traditional silos: petrophysics, geology, geomechanics, drilling, completion and stimulation, inflow modeling, static reservoir modeling, and dynamic reservoir simulation. This multi-disciplinary approach has enabled complex, iterative modeling and decision making across the full seismic-to-production workflow.

We have reached a point where modeling and simulation capabilities can simultaneously ingest almost all the data we have gathered and rigorously represent our understanding of the many physical processes taking effect in the dynamic subsurface. Our ability to create a virtual representation of the assets we are developing allows us to test almost any concept and make almost any decision without leaving the comfort of our $800 ergonomic chair.

Now we face a new problem: the results produced by our multi-disciplinary, high-resolution models are inherently too uncertain for us to make reliable, effective decisions. To solve this challenge, we must develop more efficient and more accurate measurements, identify new sources of data, and explore new ways of making that data accessible – all while keeping costs in check to ensure we increase the overall value of information.

LOW-COST APPROACHES

One of the great businesses that we were fortunate to integrate into Premier Oilfield Group started life as a post-graduate idea developed at the University of Texas' Bureau of Economic Geology. Its developers were looking for ways of characterizing core samples much more quickly and cheaply than traditional sample-based methods.

Cutting and analyzing a core has long been the *de facto* approach to obtaining representative, direct rock property measurements. However, it is also known to be time consuming and expensive. Coring the well takes much longer than regular drilling and isn't always successful; the core can get jammed, damaged, or lost during retrieval. Characterizing the core, picking depths at which to subsample, drilling plugs at those points, and running a suite of lab tests can take several months. The total price tag will often exceed a half-million dollars.

The team that later became MUD Geochemical turned to x-ray fluorescence (XRF) measurements to help simplify and accelerate part of the process. An XRF measurement captures the spectrum of light emitted when elements are illuminated with x-rays and fluoresce. Element-to-mineral models can then be used to interpret the material composition indicated by the recorded spectrum.

Hand-held XRF devices are used across many industries to measure the elemental composition of material samples. One of the more exotic examples is for establishing provenance of paintings. Those claiming to have been painted by Vincent Van Gogh, for example, should feature pigments with mineral compositions very similar to those observed in other known works by the Grand Master.

At MUD, the technology was tested and refined for the mineralogical characterization of rock cores and cuttings. The team developed proprietary element-to-mineral correlations for translating XRF spectra into lithological facies and reduced the time taken to make each measurement from a couple of minutes to just a few seconds.

Using this technology, we can scan a core at high-resolution (typically, every inch) and quickly determine where facies transitions or boundaries occur. Subsamples can then be targeted at the most appropriate depths to properly characterize the different facies that have been encountered. This is a faster process and leads to a more efficient analysis program than traditional approaches.

NEW SOURCES OF DATA

Making core analysis faster and cheaper is valuable but cores are only collected from less than one percent of wells drilled. This produces a very sparse dataset of ground truth rock property measurements across an area of interest. While a higher percentage of wellbores are logged, miles of intervening reservoir must still be modeled using interpolation techniques tied back to that same handful of physical measurements.

Drill cuttings are an often-overlooked source of high-density rock samples. Although most wells are mud logged for operational geology purposes, such as picking casing points or landing zones, the cuttings samples are frequently discarded without performing any further analysis.

Geochemical analysis at the wellsite is sometimes used to assist with directional drilling. However, detailed characterization typically requires transporting the samples to a central laboratory where they can be properly inspected, separated from extraneous material, and processed to ensure a consistent and representative measurement.

At Premier, we encourage our clients to secure and archive cuttings samples from *every* well they drill. Even if the cuttings won't be analyzed right away, they represent a rich, dense sample set from which lateral heterogeneity can be measured and used to fill in the gaps between wells where whole core has been cut and evaluated.

We combine proprietary sample preparation techniques with the rapid, high-resolution XRF measurements developed by MUD and a state-of-the-art x-ray diffraction (XRD) lab to match mineralogical signatures from cuttings to chemofacies observed in offset cores.

Information from cuttings samples expedited from the wellsite for fast-turnaround analysis can be used to optimize completion and stimulation of the well being drilled. For example, geomechanical properties correlated with the chemofacies identified along a horizontal well can be used to adjust stimulation stage boundaries. The objective is for every fracture initiation point within a stage to encounter rock with similar geomechanical characteristics, increasing the probability of successful fracture initiation at each point.

Castleton Resources was one of the first operating companies to latch onto this cost-effective approach. They were able to eliminate high-end logging runs from their East Texas Haynesville wells while simultaneously reducing stimulation costs and increasing well productivity through cuttings-based engineered completion designs. At a recent Society of Petroleum Engineers' workshop, they reported saving upward of $500,000 per well, which represents a greater than ten-fold return on investment.

THE CUTTINGS MOTHERLODE

In late 2017, as Premier Oilfield Group was busy integrating its world-class data generation laboratories, we made another prescient acquisition. The Midland International Sample Library – now renamed the Premier Sample Library (PSL) – houses an incredible collection of drill cuttings and core samples dating back to the 1940s. The compelling story of how we saved it from imminent demise is the subject of another article[1]. Suffice to say the premises needed some TLC but the collection itself is in remarkable condition.

The library is home to over fifty million samples from an estimated two hundred thousand wells – most of them onshore the United States, and many within areas of contemporary interest like the Permian Basin. It gives us the ability to produce high-density rock property datasets that can be used to reduce the uncertainty in all manner of subsurface models and simulations.

New samples are donated to the library each week, many of them from horizontal wells. This provides invaluable insight into lateral facies changes and reservoir heterogeneity. Instead of relying only on the sparse core measurements discussed earlier, our 3D reservoir models can now be populated with superior geostatistical distributions conditioned with data from hundreds of sets of horizontal well cuttings.

Boxes of Historical Drill Cuttings Stored at the Premier Sample Library.

[1] https://www.linkedin.com/pulse/turning-80-years-drilling-history-valuable-data-today-matt-bell/

DATA SHARING

At Premier Oilfield Group, we believe that generating and sharing rock and fluid data is the key to making more efficient and more effective field development decisions. The Premier Sample Library is a prime example of that belief in action.

We have created a processing team that works around the clock to prepare and characterize library samples using a suite of laboratory measurements. Since it would take years to work through the entire collection, we have prioritized wells that will give us broad data coverage across basins of greatest industry interest. Over time, guided by our clients' data needs, we will expand that coverage and create even higher-density data sets.

In an ideal world, cuttings samples from every new well drilled would be contributed to the PSL collection, preserving that information for future generations of explorers and developers.

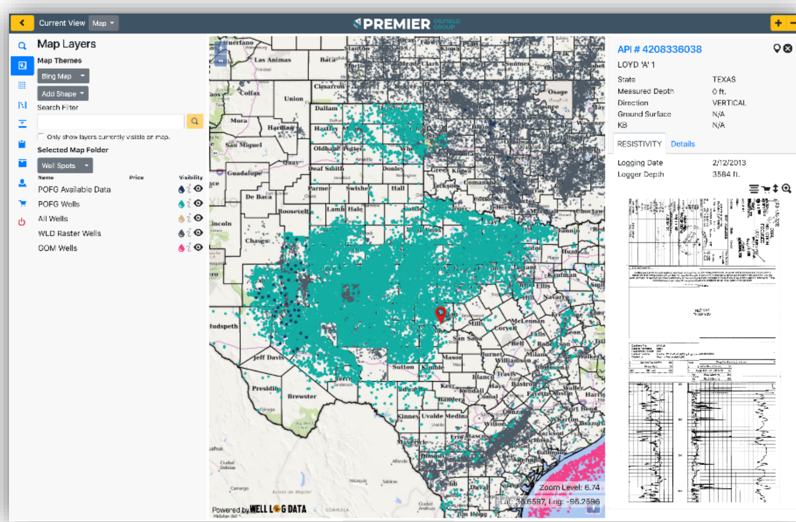
We have already helped many clients study previously overlooked intervals by reaching back in time to test samples from wells passing through to historically more prolific horizons. Thanks to their predecessors' rock squirreling habits, these clients have access to an otherwise unobtainable set of data.

To encourage today's operators to become library benefactors, we index and store donated cuttings for free. This applies to their existing in-house cuttings and core collections, as well as samples coming from newly-drilled wells. Please contact us to discuss transferring stored materials or to set up a donation program with your mudlogging service provider.

An even greater benefit accrues to those who join us in both storing and analyzing their samples. In return for contributing both the physical materials and the data derived from them, we can offer analytical pricing that is significantly more attractive than market rates. Limited time exclusivity periods can be negotiated to ensure the well owner benefits fully from the data before it is made available to others.

MAKING DATA MORE ACCESSIBLE

Arguably the most incredible aspect of discovering and restoring the Premier Sample Library is how few people were even aware of its existence. The facility didn't have an internet connection. Requests to locate and pull sample boxes for viewing were handled by phone and fax. The entire collection was catalogued on index cards – many of them hand-written - in old-school metal filing drawers.



Premier's datastak™ online marketplace allows subscribers to search and purchase many types of rock property data.

Following an intense program of scanning, digitization, well identification and location, we are now able to produce a searchable, GIS-based index for a large part of the collection. Many of the remaining boxes are hand-labeled with long-since-acquired operating companies and non-unique well names but forensic work will continue in an attempt to match them with API-recognized locations.

We are excited to have just launched the first version of our datastak™ online platform. This will finally make the PSL collection visible to everyone. Visitors can see detailed, depth-based information on sample availability and any measurements that have already been performed. Subscribers gain access to data purchase and manipulation tools and, as the platform develops, we will add click-through functionality to display underlying test results and images.

datastak™ subscriptions cater for everyone from individual geologists to multi-national corporations. We want the data that's generated to be as widely available and applicable as possible.

INTEGRATING DATA ACROSS BOUNDARIES

Another approach we're taking to help the industry maximize efficiency and effectiveness through data sharing is the facilitation of shared workspace projects.

Historically, projects sponsored by multiple operating companies have followed two models: *joint industry projects* (JIPs), where sponsors fund early-stage (typically pre-competitive) research and development work, and *consortia* where companies pay to join a closed-doors commercial group pursuing a program of work directed by an organizing company.

Premier Experts Work to Build Integrated Multi-Disciplinary Models

JIPs are notoriously slow and, with a few notable exceptions, seem to have fallen out of favor while discretionary funds for technology development remain constrained in a low commodity price environment.

Industry views of consortia vary widely, colored by individual experience. In many cases, participants describe misalignment between the consortium manager – often a service provider – and member companies, as well as a lack of collaboration between members, sometimes to the point of deliberately contributing less-than-ideal data or samples to avoid benefiting a competitor. Analytical work is sometimes carried out at test conditions deemed appropriate by the manager but not necessarily of interest to the members, resulting in mediocre satisfaction with the eventual outcome.

Premier Oilfield Group has adopted a different approach. Where we identify clients tackling similar challenges in adjacent acreage positions, we frequently observe that none of them possesses – or can even gather – enough information to fully describe the problem or produce a robust solution. Superior solutions can be found if they agree to share or exchange data – and indeed, data swaps are quite common in plays where acreage positions are well-established. Even better things happen when the companies collaborate on joint solutions, but this doesn't typically occur between competitors.

To facilitate the process, we create a shared workspace into which the interested companies contribute as much relevant data as possible. This data repository is only accessible to Premier's reservoir solutions team; the member companies do not see each other's confidential data. The founding members form a management committee that is responsible for agreeing and directing the technical scope of work, under guidance from and executed by Premier. This ensures that technical work is performed in areas, on samples, and under conditions that are considered most relevant and appropriate by the members.

Projects within our shared workspaces usually follow a seismic-to-simulation workflow. Geomodels and mechanical earth models are produced for the combined dataset and provided back to the members in a standard format (e.g. as a Petrel project). Completion and stimulation models are constructed, and history matched to stimulation records, helping to identify which of the many non-unique earth models best fits the data. This information is carried forward to dynamic reservoir simulation, history matched to production data, that is again provided back to members (e.g. as a CMG project).

The high-resolution, multi-disciplinary facies models we produce from these massive, rich data sets are inherently more robust that anything the member companies could derive on their own. They take them back to their area of interest and can more effectively interpret historical well behavior and tackle field development decisions.

The models are also used by Premier's reservoir solutions team to investigate pressing issues such as well spacing, parent-child well interference, fracture asymmetry, and the impact of well completion sequence on economic ultimate recovery.

Shared workspace projects seldom consist of a single phase of work. Throughout the model building process, we maintain an ongoing dialog around the uncertainties that remain. This often leads the member companies to acquire new samples on which to make measurements aimed at reducing a particularly significant area of uncertainty. For example, a member company might cut a new core in an area of sparse data, funded jointly by the group. This allows each member to benefit from the vital new data at a fraction of what it would cost to acquire it by themselves.

In our experience, trust and collaboration between shared workspace member companies increases steadily as the results and benefits become apparent.

CLOSING THOUGHTS

In today's data-driven oil and gas industry, we are absolutely convinced that generating and sharing relevant rock and fluid data is the key to more efficient and more effective field development. This applies throughout the E&P life cycle, from land evaluation to field development planning to enhanced oil recovery.

We believe drill cuttings offer an under-appreciated, rich source of data that should be harnessed by every operator. Cuttings can be archived at the Premier Sample Library for just the nominal cost of shipping, preserving their inherent data value for future analysis. Using our rapid, highthroughput analysis team, we can provide relevant-time input to engineered completion designs. Longer-term, we can generate laterally and vertically dense data sets to help clients construct more robust models on which to based superior decisions.

And, whether by donating samples and data to the library, transacting data through the datastak™ marketplace, or joining a shared workspace project, we see unquestionable value in sharing data to obtain a superior understanding of the subsurface. Hopefully these emerging business models will lead to transformative performance for the companies that adopt and embrace them.

To learn more, visit: https://pofg.com/