# The Integration of Data Management and Geological Modelling in a Geothermal SubsurfaceTeam

**Chris Hanton[1], Alexander Edwards[1], Ian Richardson[2], Brennan Williams[3], and Jeremy O'Brien[3]**

**[1] Ikon Science Ltd, 1 The Crescent, Surbiton, London, KT6 4BN, United Kingdom,**

**[2] Mercury Energy Ltd, 283 Vaughan Road, Rotorua 3010, New Zealand,**

**[3] Seequent Ltd, 20 Moorhouse Ave, Christchurch 8011, New Zealand**

**Keywords**

*Data Management, Integration, Subsurface Modelling, Auditability, Consistency, Standardization, QA/QC*

## ABSTRACT

The lifecycle of a geothermal energy resource can be broadly considered as four phases; exploration, feasibility, development and operation. Each of these phases has its own subsurface challenges, for example, resource delineation, creation of a conceptual model, understanding reservoir dynamics and optimizing production. To solve these challenges, we aim for best practice including the collection, management and interpretation of data feeding into the conceptual and dynamic modelling of the geothermal resource.

The key to understanding the subsurface is through the acquisition and interpretation of relevant data. In the geothermal industry large amounts of data are acquired during the full lifecycle of a field. The primary limitation of the utilization of these data are not the acquisition processes themselves, but how these data are internally managed and distributed to those who require it to make informed decisions.

This paper will discuss the modern approaches to data management and utilization with focus on the geothermal industry. First the key challenges faced with the amount of data provided from the wellbore will be outlined. Secondly insight into the solutions that exist in adjacent domains such as oil and gas will be explored and finally, with the aid of a case study, it will be shown how this technology has successfully been implemented in the geothermal domain.

## 1. Introduction

Four broad phases in the lifecycle of a geothermal resource are shown in Figure 1. The objectives of each phase differ and consequently the subsurface data and modelling requirements and workflow also differ.

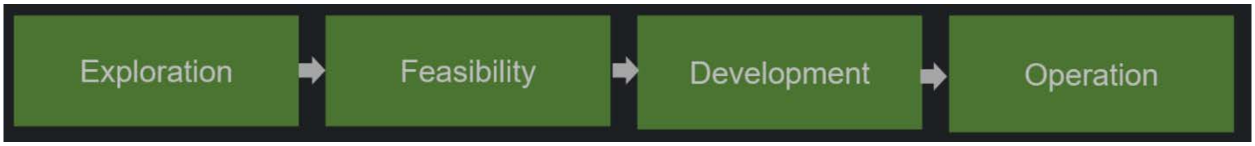


**Figure 1: Lifecycle of the geothermal resource.**

Subsurface data comes in many varieties and formats, such as regional geophysical studies, continuous wellbore log traces or pore-scale measurements from core. From the wellbore alone, hundreds of types of measurements can be made over the lifetime of a well, from drilling, completion and production phases.

The range and variety of data can provide vital insights into the nature of the subsurface, assisting with optimized well placement and improving field economics. However, the pure quantity and variation of data received often makes it difficult to manage; resulting in reduced utilization levels and directly impacting the quality of the work able to be undertaken by the geoscientist.

## 2. Data Challenges

The authors have identified 7 steps where bottlenecks in the handling of data are evident. Each of these has a fundamental impact on the efficiency and accuracy of the engineer's workflow. The steps are displayed in Figure 2 and described in detail below.


**Figure 2: Seven steps of handling data in the oil and gas industry.**

Location – Locating data has been estimated to consume up to 40% of a geoscientists time (Hawtin & Lecore, 2011). Data can reside in several disconnected repositories, both within the organization (e.g. databases, file stores, applications etc.) or outside (e.g. with vendors or public repositories) meaning the user has to trawl countless archives to ensure they have all the relevant data to hand.

Conversion – Much legacy data are stored in a format unsuitable to the modern geoscience workflows. Decades of data remains locked in unsuitable formats such as pdfs of log plots or hand typed reports. In order to utilize this data, it must first be digitized into a format allowing it to be read by today's software packages.

Aggregation – Data must be brought together into a single area of interpretation, whether that is to be loaded into an application, a singular database or a data lake. This aggregation process exposes the user to duplicate data of the same or differing vintages. The creation of rules to handle this conflicting data are vital to providing geoscientists with high quality data they can trust.

Control – Data are only useful if it can be considered high quality. With significant quantities of data being produced during drilling and operations (in oil and gas an estimated 15TB of data are produced during drilling (Dittrick, 2016)), processes must be implemented to ensure that poor quality data are identified and treated before being used in the any interpretation or geological modelling process.

Storage – A single repository for wellbore data are often desirable for users as it removes the chances of duplicate data, minimizes the time taken to locate data and increases accessibility. If multiple repositories are required it is vital these repositories communicate to ensure a flow of data between them.

Distribution – Storage of data in a stand-alone single repository, or connected repositories is not desirable if the data are still hard to access. Modern systems rely on data being provided to multiple applications and processes automatically so geoscientists can access data in environments they are familiar with.

Utilization – Data are only valuable if it can be used to improve the understanding of the subsurface. The previous 6 steps ensure that high quality data are available to the geoscientist with minimal effort, what is equally important is that any added value from interpretations is captured and added to the expanding database.

Overcoming the bottlenecks associated with these 7 steps is key to enabling digital transformation of an organization by allowing efficient usage of the data available.

## 3. Comparisons Between the Oil and Gas and Geothermal Industries

### *3.2 Data Types*

The wellbore data acquired in both geothermal and oil and gas domains are typically similar. A series of exploration and appraisal wells are drilled with the focus on the acquisition of data. Much of this data are ‘static’ with an emphasis on logs, cuttings and core samples that enable an in depth understanding of the reservoir and aquifer. Later wells focus on ‘dynamic’ production data and geochemistry data, allowing models to be updated and refined to match historical production rates and analyse geochemical trends. To improve efficiency, any system that is designed to utilize data should be comfortable with both dynamic and static data.

### *3.3 Data Utilization*

Often in the geothermal industry data are spread around different disciplines and offices with no structured approach to its storage or version control. This can often lead to confusion about which data are the ‘actual’ raw file, which is the interpretation or if either or both exist. This then leads to ambiguity around the data being used in downstream applications including geochemical analysis, well planning, 3D modelling, reservoir simulation and conceptual model updates or reviews. Given data are at the core of all these applications which have significant decision making and cost implications.

With both the geothermal and oil and gas markets focusing on return of investment and maximizing opportunities, there is a growing appetite for data driven analytics (Hanton & Miller, 2018). One example of this is the use of 'machine learning' based algorithms, designed to increase value by being able to deal with much larger quantities of data than traditional modelling methods and unveil potential missed opportunities. Due to the largely 'hands-off' approach that these methods use to utilize the data, a clean set of conditioned data are fundamental. This, itself has driven a redefinition of what data management is, and how it should be utilized in the industry.

## 4. Integration of Data Management and Resource Modelling

As energy companies and organizations have focused on efficiency and economics, data management systems have been revolutionized as places to not only store data, but also to provide direct access to users and processes. This has required a new breed of data management tool, one which is accessible, intuitive and connected. Figure 3 shows a schematic of a modern data management solution.

The schematic can largely be split up into 3 parts:

Aggregation and Standardization – Utilization of automated processes to locate, aggregate and load data. Business rules are automatically implemented to ensure data are cleaned and standardized in line with company internal policy and data conflicts are resolved.

Storage and Accessibility – Storage of clean data in a secure repository that can be hosted either internally or remotely on a cloud-like infrastructure. Accessibility to the data should be provided through an intuitive tool to allow users to find their own data.

Distribution – Any repository for data should be connected to other processes and applications as required. This allows high quality data to be automatically provided to end users as required.

The combined effect of this workflow is to consolidate and clean data before providing it directly to end users in formats and applications that they are comfortable with and able to use quickly.

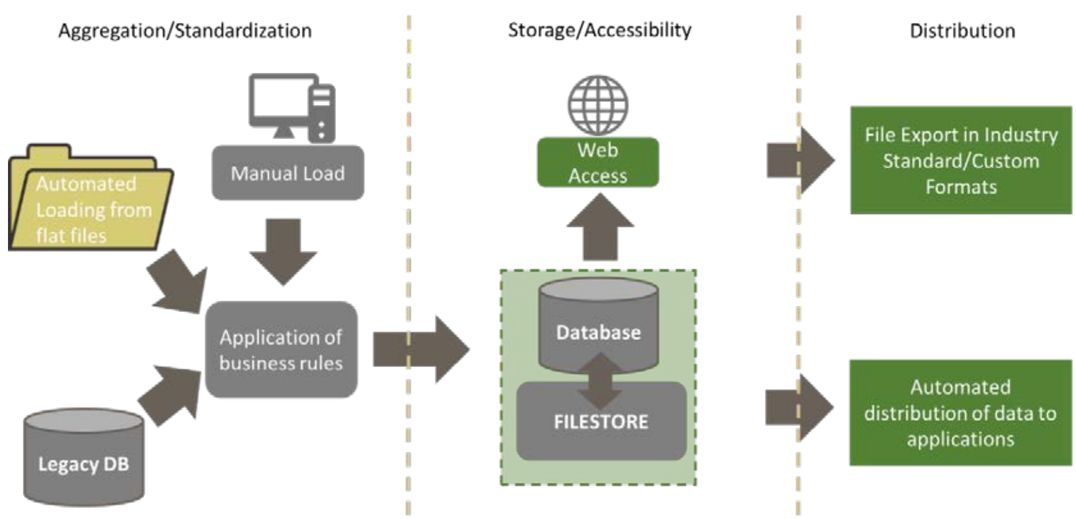


**Figure 3: Schematic of modern data management solution.**

## 5. Case Study – Technology Cross-Over

Beginning in 2018, Mercury Energy Ltd., New Zealand engaged in a project to evaluate a data management solution developed in the oil and gas industry with the aim being, if the evaluation was successful, to then follow through to the implementation of a new data management solution to provide a single location for a range of datatypes from the subsurface for the four geothermal fields they operate in New Zealand. The solution needed to cover data collected from well drilling and logging, through to well operation and abandonment. Both spatial and temporal data are stored and visualized, with geologic, geochemical, production, engineering, well and reservoir monitoring data all available in a single platform.

The challenges that Mercury Energy faced prior to the engagement are those common throughout the geothermal industry. These challenges include a lack of certainty on the availability, location quality, applicability and importance of the data in the organization. Well data can often be duplicated by being stored in multiple locations, file formats and data storage systems or worse, slightly different versions of the data may be available in different locations. The amount of data to be taken through the implementation process depends on factors such as the number of geothermal fields, the number of wells, the age of the wells and how diligent the company has been in collecting the appropriate data at appropriate time intervals. When as asset or field has been transferred through different ownerships, all these issues can become exacerbated. This made finding data time-consuming as multiple locations and systems were searched for the required information. This also resulted in duplication of information and datasets, which added to concerns with version control and subsequently the quality of retrieved data.

It is vitally important that the implementation of a data management solution needs to be run as a proper project with allocated and dedicated resources. It also requires buy-in by all the stakeholders involved. These stakeholders include the members of the sub-surface asset team (geochemists, geophysicists, geologists, engineers), IT, asset team management, exploration and production management, and information management. Ideally, there should be a dedicated data manager role or, as in Mercury’s case, an internal champion, and this person should not necessarily be a member of the asset team. The project is not just a technical implementation; it also a change management project which requires individuals to buy into the culture of data management and the organization to build processes about how they interact with data to facilitate ease of use. This can come with hurdles as people change their habits and there may be pain points as people transition to a new way of operating.

To improve stakeholder engagement, a vision of the project outcome needs to be articulated with regular communication and consultation as to project progress and ensuring that the project deliverables are a true reflection of the stakeholder requirements. Planning for quick win milestones along the project time frame can also be used to maintain project momentum and keep stakeholders engaged. For this project one method of achieving this was breaking down the data cleansing and loading steps into discrete packets of work, so as not to overwhelm stakeholders with the pure volume of data handled in the project.

The pilot project also provided an opportunity to showcase the solution to stakeholders to build engagement and to identify potential project champions who could assist in the project implementation later.

To evaluate the proposed solution, we ran a pilot project using a small three well subset, the selection of the wells was important as they needed to have all the types of data you might expect to encounter across any of the assets. We worked with the solution vendor to create a data structure and set of standards and to confirm that the database solution supported each type of data, the types of plots and analysis we required, what the potential bottlenecks or roadblocks might be and how these could be solved.

Following the success of the pilot project, Mercury Energy moved into an implementation phase.

Prior to implementation it was essential to ensure that the data to be used was of enough quality. It was an opportunity for Mercury to scrutinize our data and our processes which is something that we would recommend to any other organisation. With a large amount of data being available it made sense to carry out an initial scan or audit of the data. This enabled us to allocate an initial "level of importance" to data and to identify data that was either redundant, obsolete or trivial and data that was missing or required some remediation. Reviewing and preparing data for a new system also provides the opportunity to improve the quality and consistency of the data; for example, using a consistent projected coordinate system, reconciling naming conventions (Well-1 vs W1 vs W_1 etc.) or adding additional metadata. This formed the basis of a QA/QC process for inputting data into the system and being able to access it.

The solution being implemented may be replacing an existing legacy system with an underlying database or a simpler data source such as sets of Excel or CSV files. If there is a large amount of data, then the more automated the transfer process can be the better. The data management solution in this case included the concept of a "data loader" which essentially is a robust template based process enabling the batch input of well data even when there is some variability amongst the wells, e.g. missing columns, columns in different orders or with slightly different names. The "data loader" concept also provides a facility for future data delivered in other formats by external contractors to be imported.

Mercury Energy also worked with both the database vendor and the geological modelling solution vendor to identify and prioritise geothermal specific features that needed to be implemented to meet our needs. Examples of this included geochemistry ternary diagrams (Figure 4), some basic decline curve analysis functionality including exponential, hyperbolic and harmonic regressions (Figure 5), well correlation and downhole well data summary plots (Figure 6) and improved support for time-dependent data and multi-axis plots (Figure 7). As with any software solution, this is an on-going effort with the aim being to have scheduled roundtable meetings to guide future development; for example, a direct link between the database and the geological modelling software and improved data presentation tools for reporting purposes. The advantage of this is that with the data now stored in one place the plots created and data analysed can be trusted.

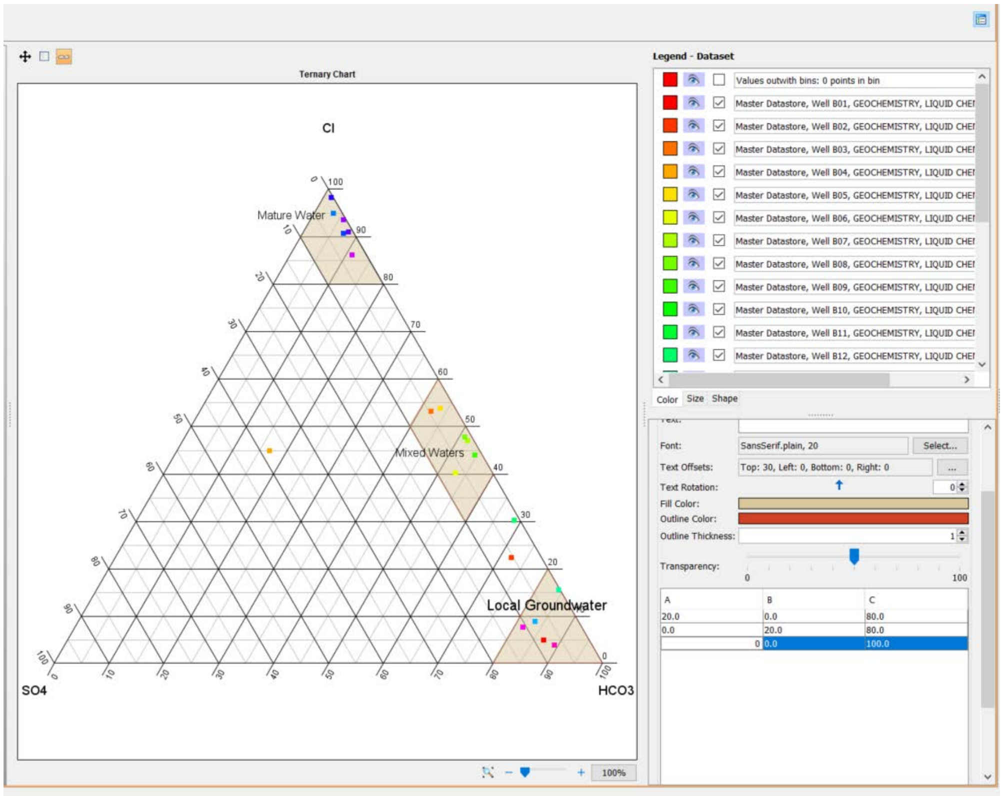


**Figure 4: An example of a Cl-SO4-HCO3 ternary diagram plotted straight from the database.**

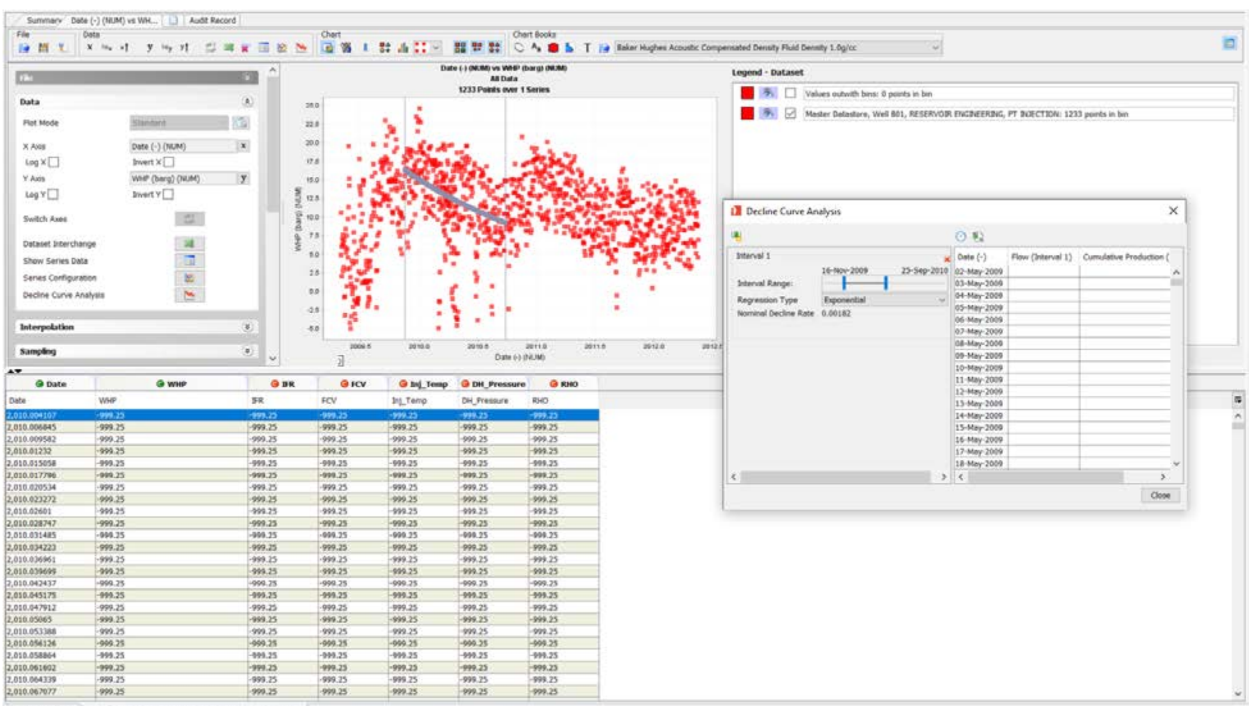


**Figure 5: An example of utilizing the decline curve functionality from production data in the database.**

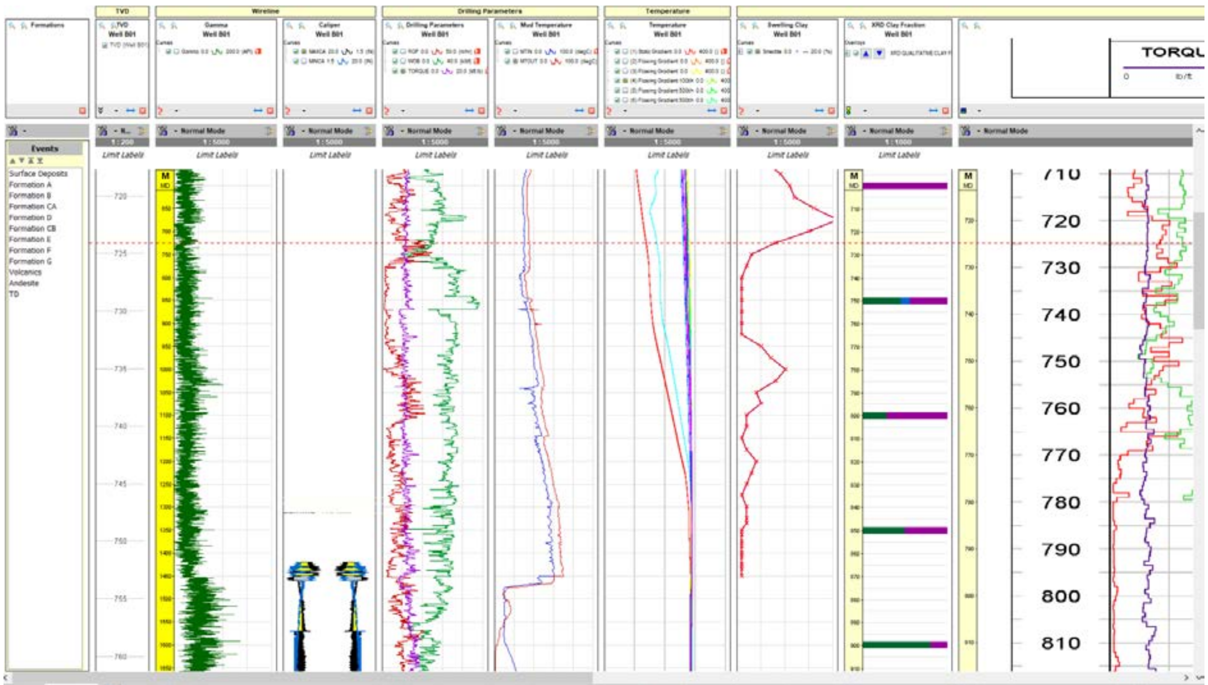


**Figure 6: Well summary plot with gamma, caliper, drilling parameters (Rate of penetration, ROP; Weight on Bit, WOB and Torque), PTS data (Pressure, Temperature and Spinner), Electron Beam Techniques (e.g. MEB) and PDF drilling reports.**

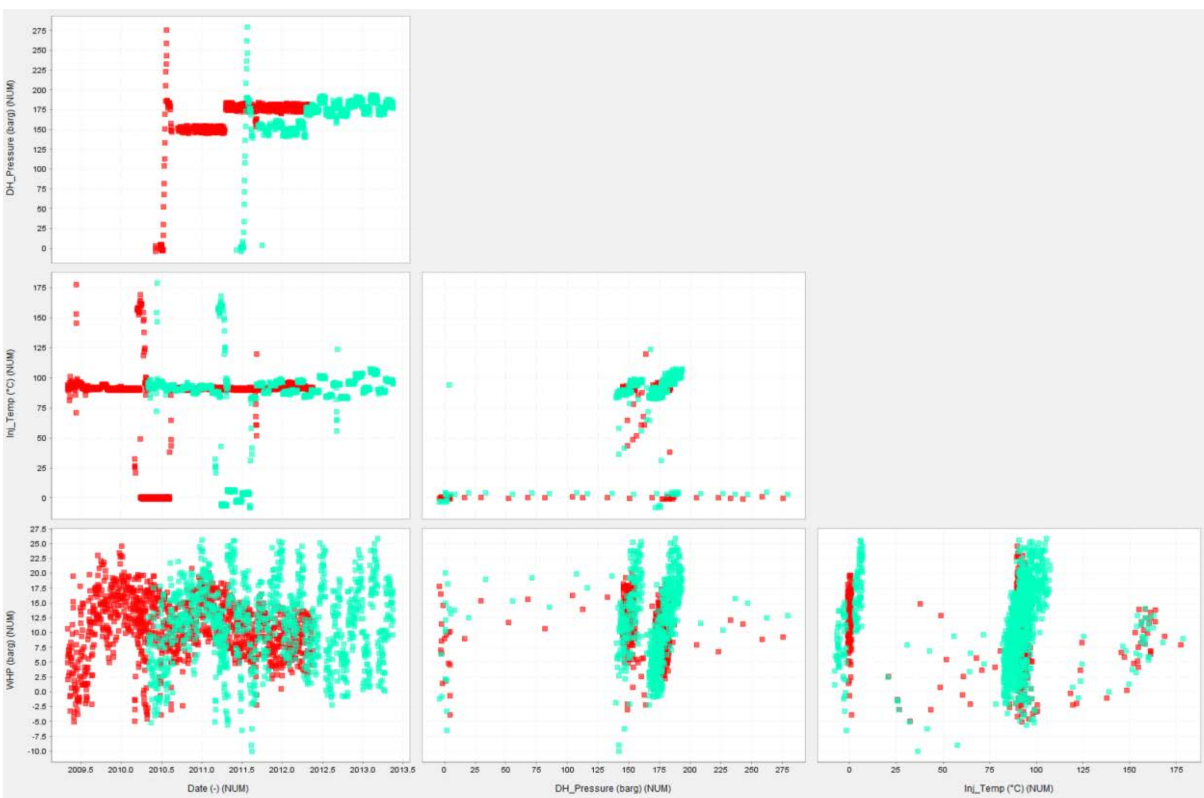


**Figure 7: Multiple plots comparing date, well head pressure, temperature and down hole pressure for two injection wells.**

The underlying database of the data management solution is an SQL server database located on a local server. The workflow for the link between the geological modelling software and the database requires the following steps:

- Connect to the database and provide username and password authentication (Figure 8).
- Select the desired project or working set (a project is a geothermal asset or field; a working set can be a subset of a project).
- Select the wells we want to import data for.
- Select the interval data tables (also known as datasets) we want to import (Figure 8).
- Import the selected well headers, surveys and data (Figures 9, 10 and 11).
- Visualise and model with the data in the geological modelling software (Figures 12 and 13).

Interval data are defined at "from, to" intervals down the well trajectory. Examples include categorical data such as lithological units or numeric data. Additional data such as numeric point data, for example, temperature or resistivity logs can also be imported, as can feed zone data and structural data.

When new data are available, either for new wells, or additional data for existing wells, these data can be updated using the link. The geological modelling software enables models to be dynamically updated when data or interpretation the model is dependent on has been updated.

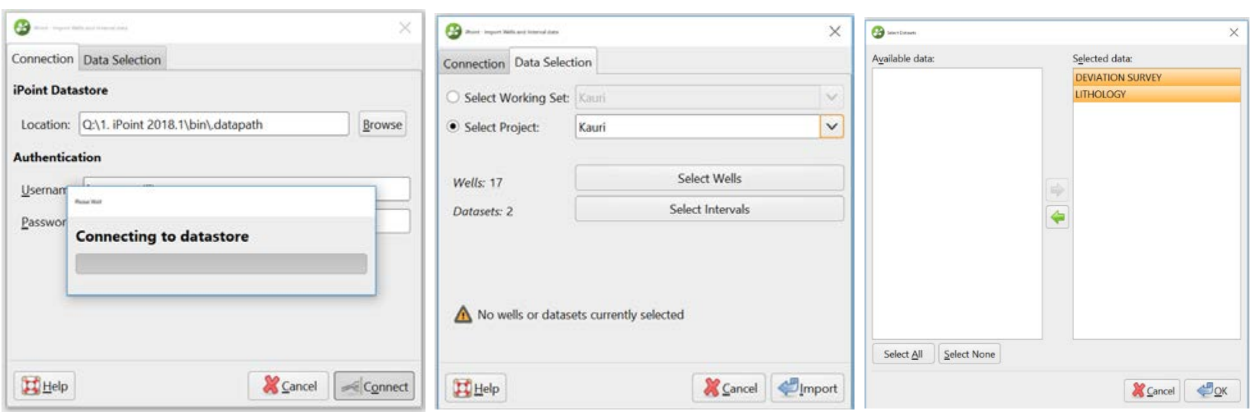


**Figure 8: Connecting from the geological modelling software to the database, selecting the project, wells and dataset tables to be imported.**

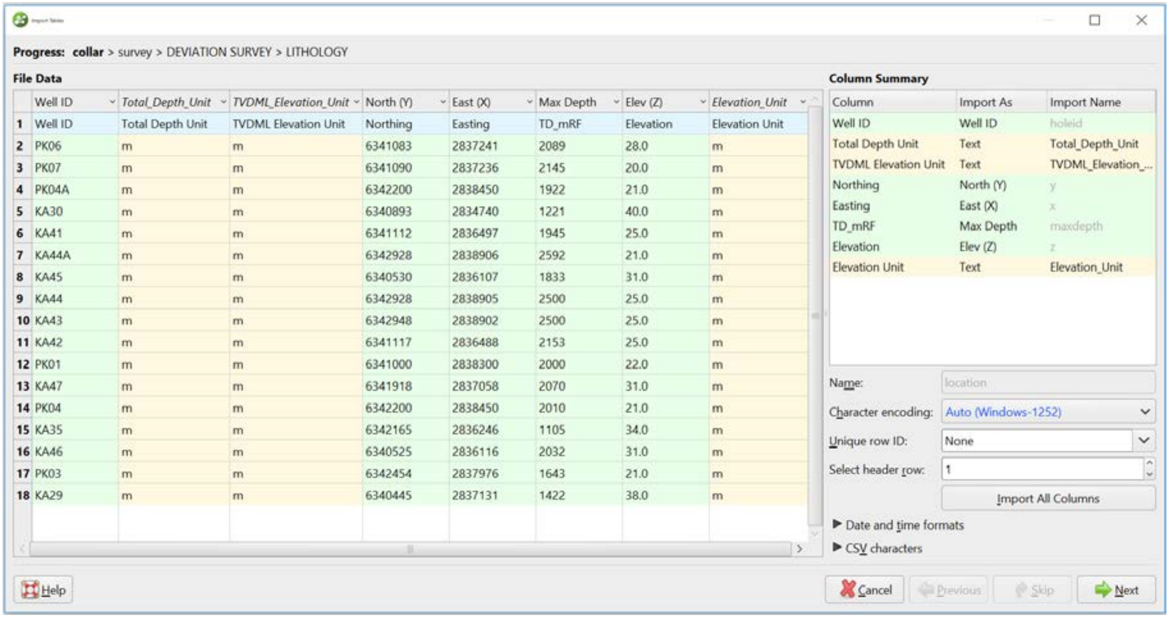


**Figure 9: Importing the well headers**

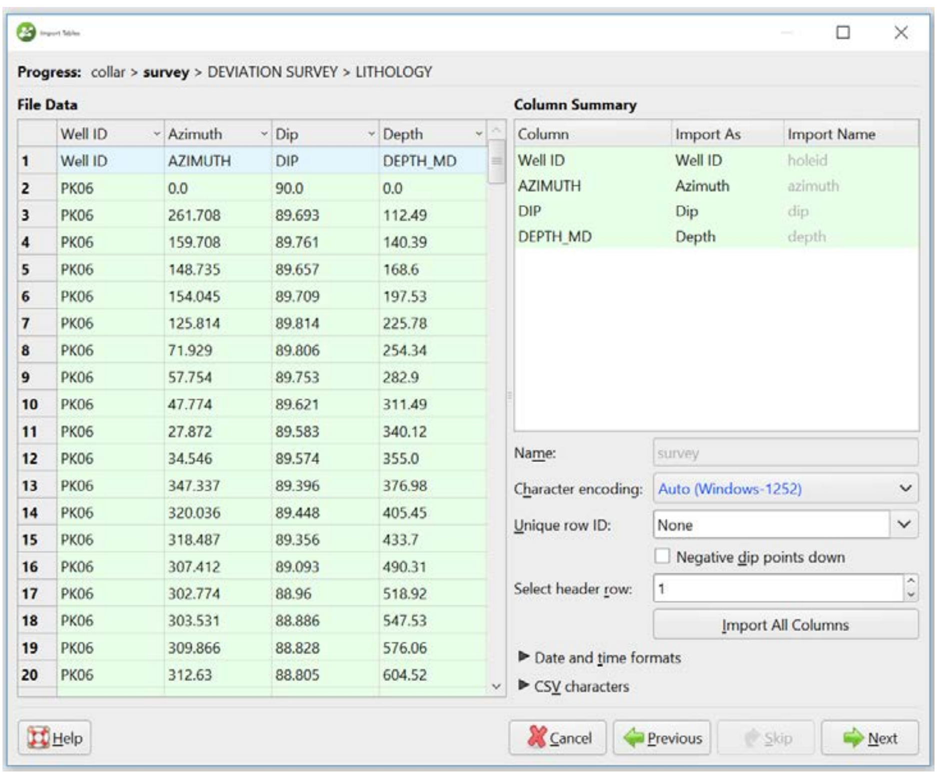


**Figure 10: Importing the well trajectories or surveys.**

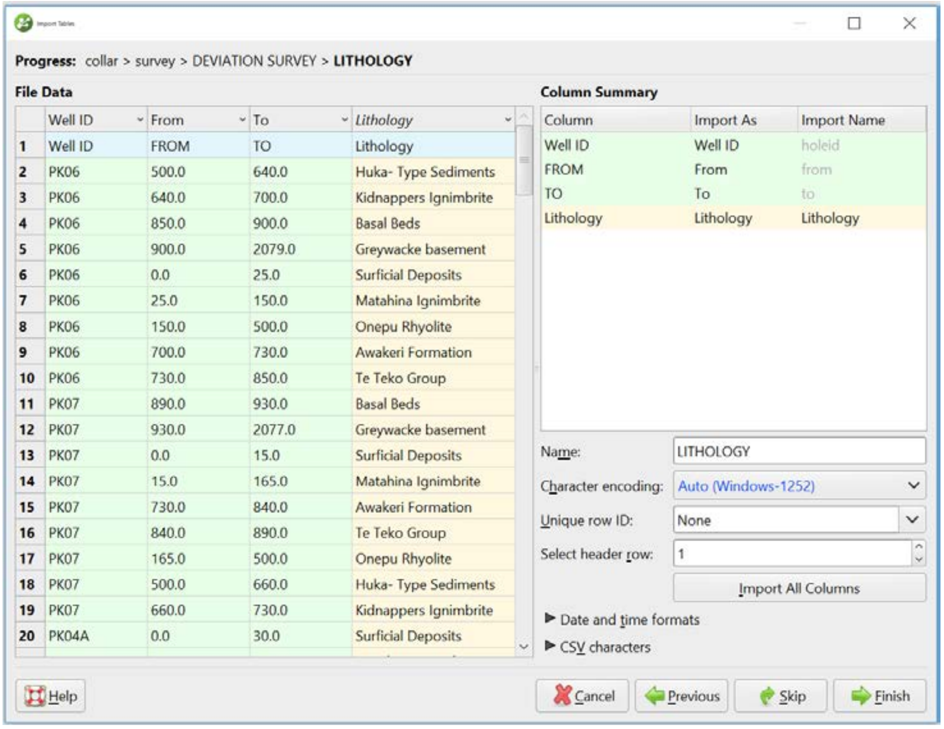


**Figure 11: Importing the selected data, for example, interval lithological data.**

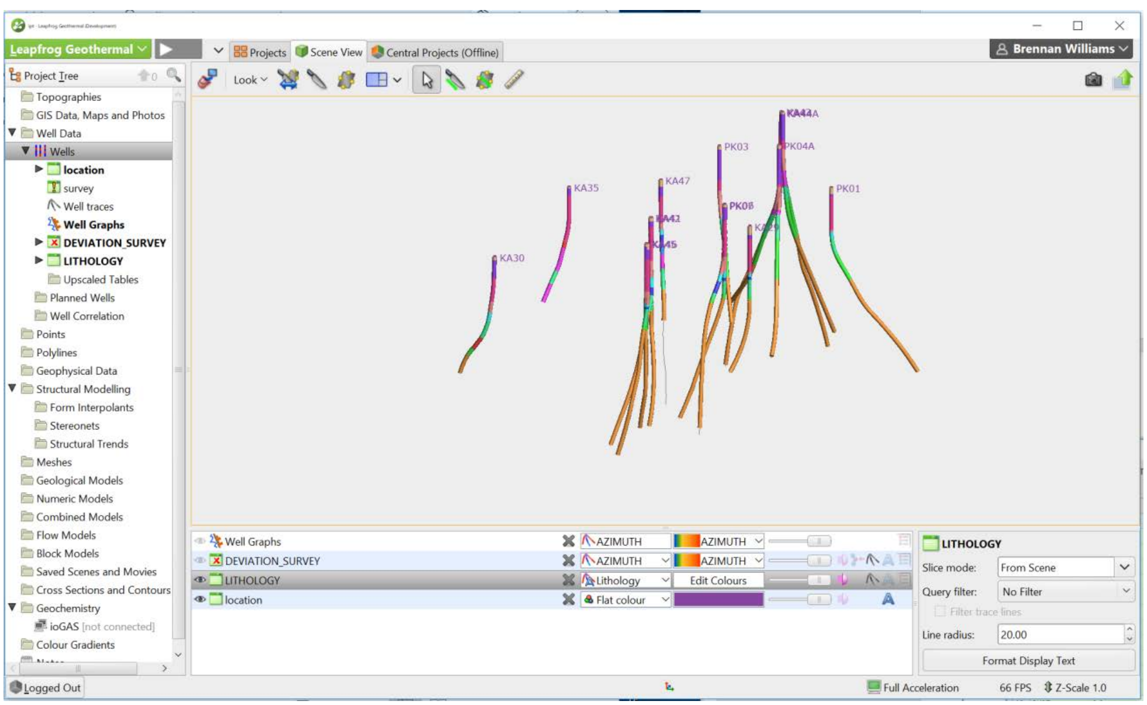


**Figure 12: Visualising the well trajectories and lithological interval data.**

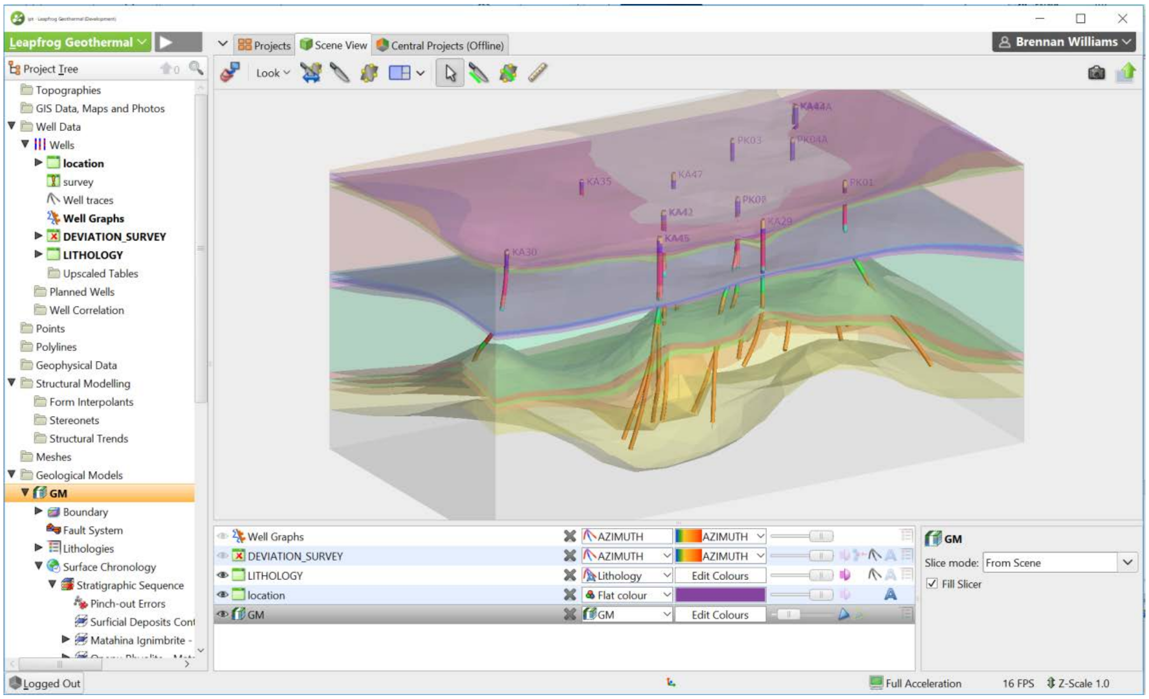


**Figure 13: An initial geological model built from the lithological data imported from the database.**

## 6. Conclusions

Mercury Energy have now implemented a data management solution that has been developed in the oil and gas industry since 2006 but optimized for the geothermal industry (Figure 4). The solution enabled them to store data from injection, production and monitoring wells along with plant and surface monitoring points in a single location. As part of the loading process several standardization and QC rules were applied, allowing Mercury Energy to build up a single definitive source of high quality data for their subsurface engineers.

As a result of this implementation Mercury's geoscientists and reservoir engineers are now able to locate, visualize and compare multiple data types and datasets in a single location. This has significantly reduced the amount of time spent on locating and processing data for both investigation work (trouble shooting and problem solving) and reporting requirements. Comparisons of wells, as well as comparisons between data types (e.g. geology, casing integrity, temperature and pressure) can now be quickly undertaken, whereas previously these data types were all stored separately, and data was manually copied into new datasets to undertake direct comparisons. Mercury Energy is now in the process of implementing the solution across the rest of their assets based off the success of the initial evaluation.

As the geothermal industry continues to pursue high productivity at low costs, the importance of data and it's handling within organizations only increases. By utilizing best practice workflows in the data management system for QA/QC, auditability, consistency and integration leads to improve decision making which can be realized at any stage of the geothermal lifecycle. Ultimately, by aggregation and the consolidation of data, underpinned by robust QC and standardization procedures can only lead to improved decision-making related to subsurface data.

**Acknowledgement**

The authors thank Mercury Energy, Seequent and Ikon Science for permission to publish the data and workflow.

## REFERENCES

Dittrick, P., “Data analytics optimize drilling technology.” *Oil and Gas Journal*, 114, (2016).

Hanton, C. and Miller, J. “Leveraging existing databases for advanced analytics of the future.” *23rd annual International Conference and Exhibition on Emerging Trends in Petroleum Data* (2018).

Hawtin, S. and Lecore, D. “The business value case for data management – A study.” *Oil and Gas UK Report* 051, (2011).