D13.4: Validation of the OPTIDRILL system predictions using post-drilling logging and measurements

Deliverable No	D13.4
Work package No. and Title	WP13 OPTIDRILL system field-scale test and validation in a borehole
Version - Status	V1.0 – Final
Date of Issue	18/12/2024
Dissemination Level	PUBLIC
Filename	D13.4_v1.0_FINAL

This project has received funding from the European Union's Horizon 2020 research and innovation action under grant agreement No 101006964

DOCUMENT INFO

Authors

Author	Organization	e-mail
Matthew Arran	BGS	matarran@bgs.ac.uk
Ashadul Hoque	TVS	ashadul@technovativesolutions.co.uk
Andrew Kingdon	BGS	aki@bgs.ac.uk
Russell Parsons	BGS	rparsons@bgs.ac.uk

Document History

Date	Version	Editor	Change	Status
12/12/2024	0.0	Matthew Arran	Initiated	Draft
13/12/2024	0.1	Ashadul Hoque	Added economic impact estimates	Draft
16/12/2024	0.2	Matthew Arran	Added error analysis, conclusions, summary	Draft
16/12/2024	1.0	Andrew Kingdon	Edits and added annex	Draft
18/12/2024	1.1	Matthew Arran	Corrections	Draft
18/04/2025	1.2	Matthew Arran	Corrections	Final
20.04.2025	1.2	Shahin Jamali	Editorial review	Final

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

This deliverable reports on work conducted towards deliverable 13.4: validation of the OPTIDRILL system predictions, through examination over field tests of variability in predictions' accuracy. Due to issues encountered in other work packages, field testing was limited to the drilling of two shallow boreholes at a site in Bochum, Germany, as described in deliverable report 13.1. The machine-learning-based models from work packages 7, 8, and 9 were applied to predict both the Rate Of Penetration (ROP) during drilling and the class of lithology being drilled (Lithology), and also to identify anomalous datapoints that were potentially indicators of drilling problems.

Here, we describe analysis of results from these field tests, linking errors in predictions to the datasets from which models make these predictions. Specifically, we compare the phase spaces of the datasets used to train the ROP and Lithology prediction models to the phase spaces of the datasets to which the models were applied. We examine the generalisation error that arises from differences between these phase spaces and, by fitting Generalised Additive Models, link variation in subsets of input parameters to variability in both the magnitude of ROP prediction errors and the likelihood of erroneously identifying drilling problems. This quantifies influences on the generalisation error and indicates where additional data may decrease this error.

The analysis demonstrates that model predictions' accuracy varied significantly over each field tests and between the two tests, with generalisation error increasing as the data with which predictions were made differed more from the data with which models were trained. Variation in accuracy was substantial even over short intervals of drilling in similar conditions. Consequently, the accuracy observed in field tests is very unlikely to be representative of accuracy in different drilling conditions, using different equipment at much greater depths.

Nevertheless, we use the errors in ROP prediction in field tests, with various simplifying assumptions, to estimate the potential economic impact of the OPTIDRILL system's parameter-recommendation subsystem, based on the ROP prediction model. This suggests that the system could reduce the cost of drilling a 5-km deep geothermal well by between 2.4% and 6.0%.

Achieving such cost reductions, and those associated with effective lithology prediction, will require new model training for new drilling conditions, and hence the development of a corpus of standardised geothermal well drilling data which can be used for this training. For further development of drilling problem detection, a native-digital standard for propagation of Daily Drilling Reports is essential. Annex 1 contains recommendations for geothermal drilling datasets and data standards, so that future geothermal drilling programmes are informed by standardised benchmark datasets.

1. INTRODUCTION

Work package number 13 deals with the validation in a field-scale test of the OPTIDRILL drilling advisory system. Deliverable report 13.1 describes the field tests, conducted at Fraunhofer IEG, Bochum, while deliverable reports 13.2 and 13.3 examine the accuracy of the machine-learning-based systems developed in work packages 7, 8, 9, and 11, for predicting rates of penetration (ROP) and lithology classes, identifying drilling problems, and recommending drilling parameters. This report examines the variation of accuracy over each field test, indicating the reliability of the systems when applied in new circumstances and suggesting means of improving that reliability. The report further estimates the potential economic impact of applying the parameter-recommendation system to geothermal well drilling, if accurate.

Due to challenges encountered in other work packages, the field tests differed from those envisaged in the original project proposal. Firstly, acoustic emission data were not available as inputs for the model training or development of the ROP and lithology-prediction models. Therefore, these attributes could not be used in the derived drilling recommendation system. Secondly, the system for drilling problem identification was developed using legacy data from UK onshore hydrocarbon exploration, with depth resolution different from that of the field tests. Thirdly, a lack of data prevented development of any advisory system for drilling by jetting. Finally, boreholes were chosen for system testing based on opportunity and the drilling opportunities available, and did not represent the full range of drilling conditions to which the drilling advisory system was intended to be applied. Consequently, the average accuracy of system predictions during the field tests is unlikely to be representative of such prediction accuracy under other conditions, making it essential to understand that variability.

Understanding of accuracy’s variability requires quantitative models that are both general and interpretable. Complex physical systems connect lithology, drilling parameters, rates of penetration, and drilling problems. These involve strong non-linearities and multiple interactions, and the machine-learning-based systems used to model aspects of these systems include correspondingly complex, non-linear interactions. Therefore, discrepancies between the outputs of the physical and machine-learning-based systems will also depend non-linearly on multiple interacting parameters, and models for these inaccuracies must be sufficiently general to be capable of describing such dependence. In order for accuracy’s variability to be understood, rather than just described, and to ensure that results are transferable to contexts that may be quite different, such a model must also be interpretable, indicating the individual importance and effects of small subsets of parameters. One class of models that offers this balance of generality and interpretability is that of Generalised Additive Models (GAMs).

1.1 Generalised Additive Models

Motivated by the theory of smooth function approximation, a GAM supposes that an underlying quantity y is a multivariate function of inputs $\boldsymbol{x}$, that exists in the form:

$$y = f(\boldsymbol{x}) = \Phi\left(\sum_{j=1}^{n} \phi_j(x_j)\right),$$

for smooth, monotonic function Φ and smooth functions ϕ_j (Hastie & Tibshirani, 1990). Such models can represent an extremely broad class of functions f, including strong nonlinearities, while permitting the effects of individual parameters to be interpreted, via the functions ϕ_j. Permitting functions of the form $\phi_{j,k}(x_j, x_k)$ in the sum further extends the class of functions f that can be represented, to include interactions between pairs of input parameters, while retaining interpretability of these interactions.

To implement a GAM in practice involves using one of many available software packages. The function Φ, a statistical model for errors in y, and the input parameters of concern are specified, including any interaction terms. The functions ϕ are estimated statistically, by supposing each a weighted sum of a specified class of basis functions, fitting to observations for a range of different basis sizes, and selecting the optimum basis size using a criterion such as the generalized cross-validation statistic.

2. ANALYSIS OF VARIATION IN PREDICTION ERROR

2.1 Data

The drilling analysis was conducted using three datasets:

- files of the Geostar data (in CSV format) with which the first rate-of-penetration and lithology prediction models, based on supervised machine learning, were trained
- a file of the field test data (in EXCEL format) to which the first supervised-learning models were applied, together with the unsupervised-learning-based drilling-problem-identification model, and with which the second supervised-learning models were trained
- a file of the field test data (in EXCEL format) to which the second supervised-learning models were applied, alongside the drilling-problem-identification model.

Derivation of the Geostar dataset is described in the reports for work packages 7 and 8, while the derivation of field test data, including of advisory system predictions, are described in the report for deliverable 13.1. Importantly, field tests were conducted close to the sites from which Geostar data were collected, with similar drilling and measurement equipment, but with a larger bore drill bit and consequently lower rates of drill string rotation. The second field test was similar to the first, but with drill collars added to the drill string, increasing the weight immediately above the bit and so stabilising drilling. In neither test were any significant drilling problems encountered.

Datasets include:

- drill bit downhole depth, typically recorded at 1 cm intervals (Depth, in metres),
- drill string rotation rate (RPM, in rotations per minute)
- rate of penetration (ROP, in metres per hour)
- weight on bit (WOB, in kilonewtons)
- torque applied to the bit (TRQ, in kilonewton-metres)
- drilling fluid pumps' pressure (P_P, in kilopascals)
- pump flow rate (MFI, in litres per minute)
- for the first field test, the true lithology class, inferred after the end of operations (Lithology)

Also recorded for each depth are the machine-learning-based models' predictions for:

- ROP (ROP_pred, in metres per hour), from supervised learning
- lithology (Lit_pred), from supervised learning
- whether the datapoint is anomalous (Anomalous), from unsupervised learning

2.2 Phase spaces

Figure 1 to Figure 4 indicate the phase spaces of the three datasets, via projections onto related subsets of parameters. This permits, for each of the two supervised-learning models, the values of measurements with which the model was trained to be compared to the values to which the model was applied.

Whilst data from Geostar wells and data from the first field test include similar values of Depth and WOB, TRQ in the first field test was generally higher than that for Geostar well drilling, and RPM generally lower These are both a direct consequence of the larger drill bit size. In addition, values of P_P and MFI were higher in the first field test than almost all values recorded during Geostar well drilling. Consequently, there is no overlap between the phase space of the Geostar wells and that of the first field test. Predicted ROP for the first field test are therefore necessarily extrapolations, even though ROP lies within the range of Geostar drilling. For Lithology prediction, meanwhile, the 'Claystone/Sandstone' class was unobserved during Geostar well drilling, so that its prediction requires further extrapolation.

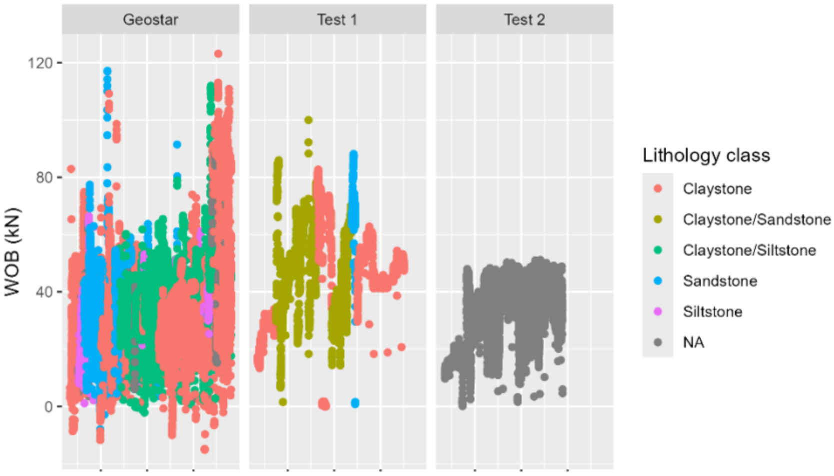

Figure 1: Variation of weight on bit and lithology class with downhole depth

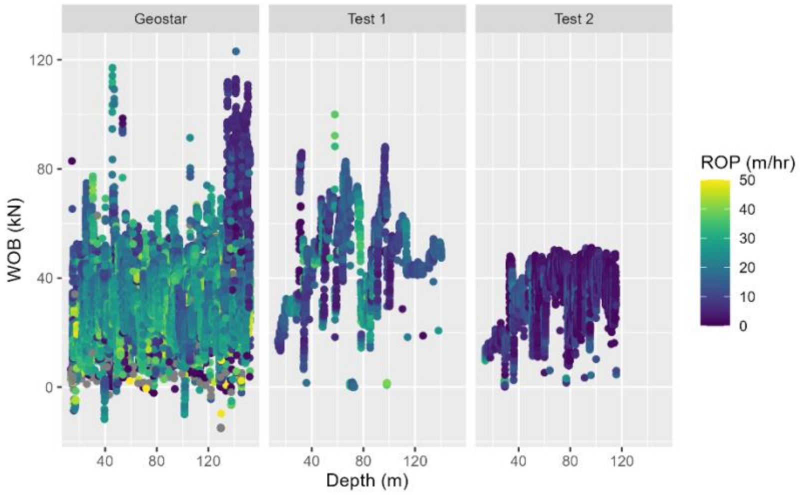

Figure 2: Variation of weight on bit and rate of penetration with downhole depth

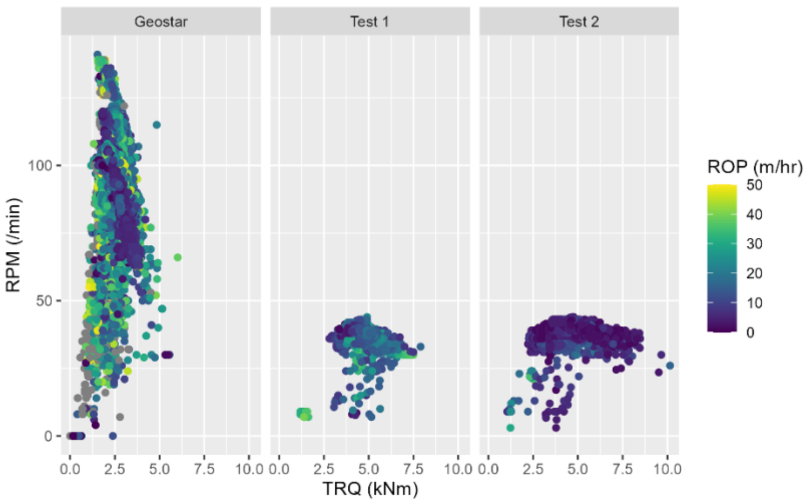

Figure 3: Variation of torque, rotation rate, and rate of penetration

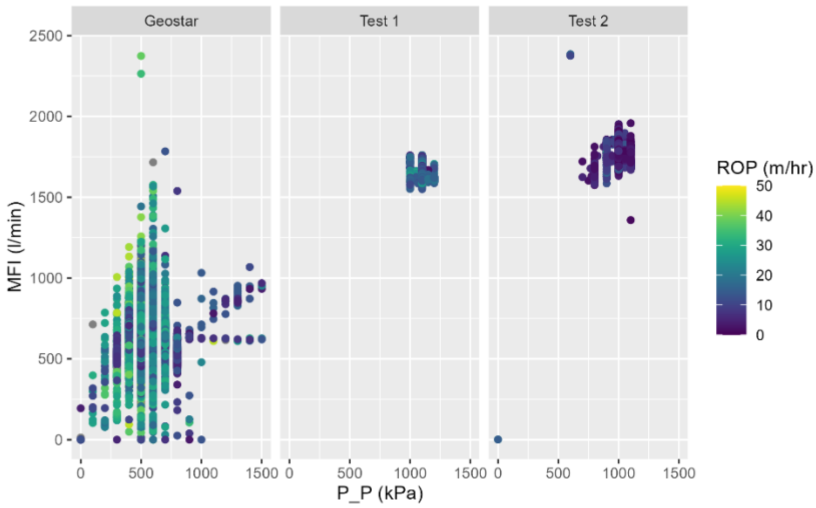

Figure 4: Variation of pump pressure, pump flow rate, and rate of penetration

The phase spaces of measurements from both the first and second field tests overlap, but also differ significantly. TRQ, P_P, and MFI resulted in generally similar values, but were more variable during the second field test than the first. WOB values were generally lower, though within the range of the first field test, while RPM values were similar and Depth values mostly identical. Correspondingly, ROP values in the second field test were within the range of those in the first field test, if generally lower. The absence of true lithology data for the second field test prevents comparison with the first field test.

2.3 Errors

2.3.1 Errors in rate-of-penetration prediction

As the differences between the phase spaces of the three datasets suggest, the two models' predictions for the rate of penetration are subject to generalisation error. Figure 5 demonstrates the error that arises when TRQ and RPM values, used by each of the two models to predict ROP, lie outside the range of values with which the corresponding model was trained. Specifically, in the first field test, the lowest values of TRQ and highest values of RPM are the furthest outside the range of values in the Geostar dataset used for model training, and are associated with significant underestimates of ROP. In the second field test, meanwhile, RPM values higher than those observed in the first field test were associated with significant overestimates of ROP, and much lower RPM and TRQ values were associated with ROP underestimates.

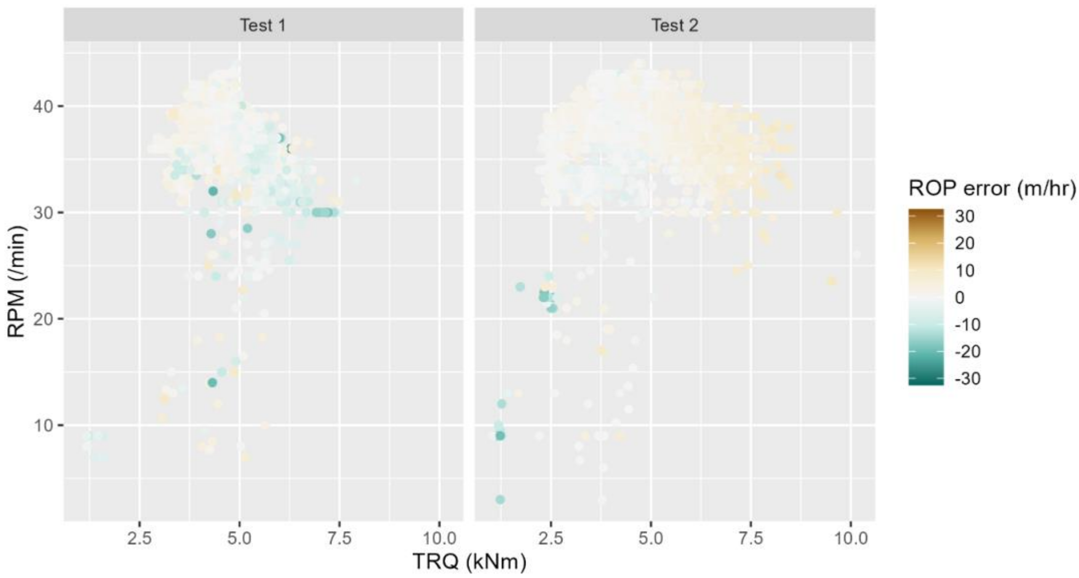

Figure 5: Relationship between generalisation for TRQ and RPM, and error

The degree of generalisation required for ROP prediction can also be quantified and linked to the error. Normalising a training dataset's datapoints of WOB, TRQ, RPM, P_P, and MFI values, using the maximum and minimum values within the dataset, the mean of the distances between a normalised datapoint and its three nearest neighbours represents the datapoint's proximity to others, and the 99th percentile of the mean distances represents the degree of generalisation required during training. The degree of generalisation required for ROP to be predicted for a given datapoint is then calculated by applying the same normalisation, and calculating the mean of the distances between the normalised datapoint and the normalised training dataset's three nearest datapoints. Figure 6 indicates the high degree of generalisation required in the first field test and the consequent generalisation error, together with increases in error in the second field test where the degree of generalisation required is higher.

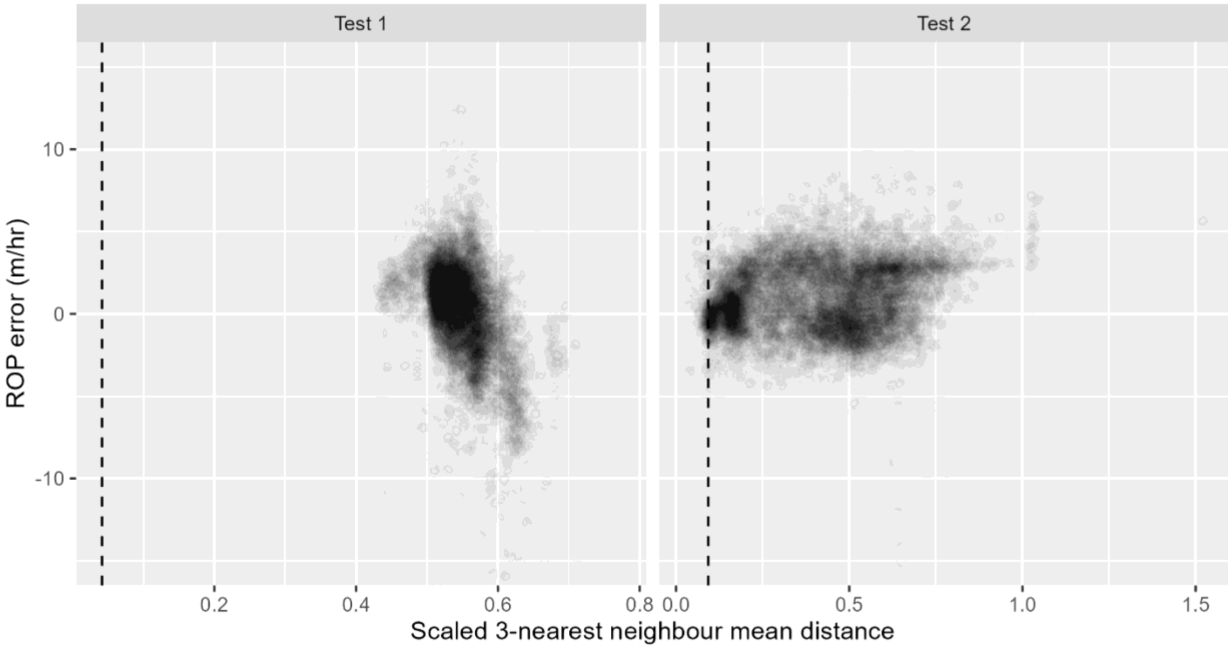

Figure 6: Relationship between error in ROP predictions and the degree of generalisation required, quantified by the mean of distances to the three nearest neighbours in the training dataset, after normalisation. Dashed lines represent the 99th percentile of corresponding means within the training dataset.

Fitting Generalised Additive Models permits errors in ROP predictions to be linked to specific changes in inputs, although with caveats. GAMs for the errors in each machine-learning-based model were fitted using the R package mgcv (Wood, 2017), using the model

$$\widehat{ROP} - ROP = \phi_0 + \phi_1(WOB) + \phi_2(TRQ, RPM) + \phi_3(P_P, MFI) + \epsilon$$

for constant ϕ_0, scalar function ϕ_1 and tensor functions ϕ_2, ϕ_3 composed of cubic regression splines, and independent, identically distributed errors ϵ. Since errors will not have been independent, with autocorrelation in $\widehat{ROP}$ and ROP, we sought to avoid overfitting by constraining the maximum number of degrees of the scalar and tensor functions to be 10 and 25, then selecting the optimal number of degrees of freedom using generalised cross validation. GAMs were fitted separately for the results of the first and second field tests, with the corresponding fitted values for ϕ_0 0.13±0.03 and 1.024±0.016, respectively, while the fitted functions are shown in Figure 7. After fitting, the GAMs explained 41% and 49% of the variance of $\widehat{ROP} - ROP$, respectively. However, the incorrect error model makes estimates of standard error inaccurate, decreasing the reliability with which results can be transferred. Furthermore, the GAM is informative only about the fairly narrow range of values for which data is available.

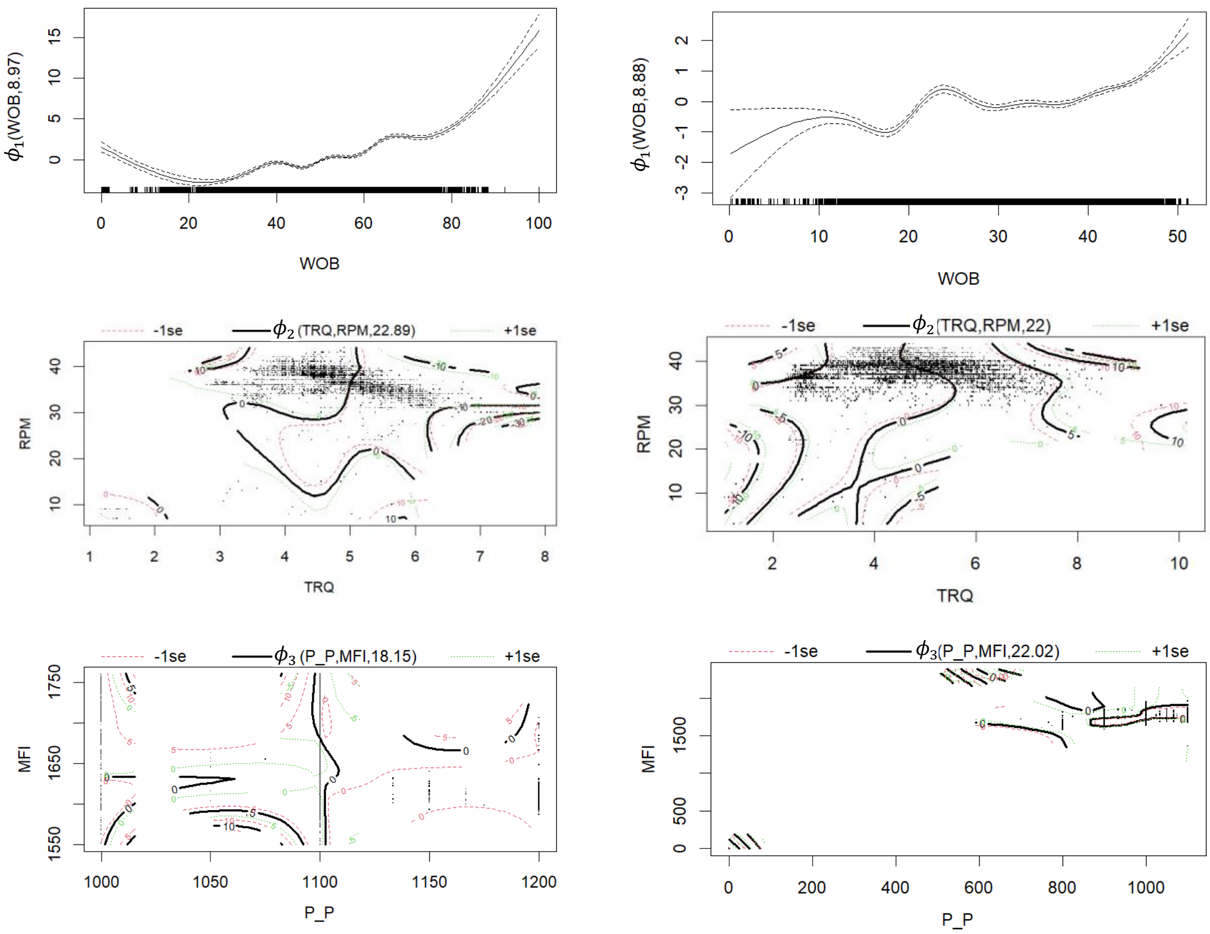

Figure 7: Components of the best-fit GAMs for model prediction errors. In the top panels, solid lines indicate function values, dashed lines standard errors, and lines on the x-axis the values of datapoints. In the remaining panels, black contours indicate function values, coloured lines standard errors, and dots datapoints' values.

The fitted GAMs are not easy to interpret, but broadly confirm the intuition developed earlier: errors grow in magnitude as the datapoints with which each machine-learning-based model makes predictions differ more from the datapoints with which the model was trained. The most significant effects are of variation in TRQ and RPM, with error contributions changing most rapidly in the direction away from the training dataset's phase space: with increasing TRQ and decreasing RPM for the first field test, and with increasing TRQ in the second field test. The next most significant effect is of WOB in the first field test, with high contributions to error at the high WOB values that are rare in the Geostar dataset. The effect of P_P and MFI, meanwhile, is hard to determine, with little variation in their values over the field tests. Better understanding of how pump pressures and flow rates affect rate-of-penetration prediction error would require a wider exploration of the parameter space during drilling.

2.3.2 Errors in lithology class prediction

Figure 8 demonstrates the correspondence between the true lithology classes and those predicted by the machine-learning-based model, for the single field test for which true lithology classes are available.

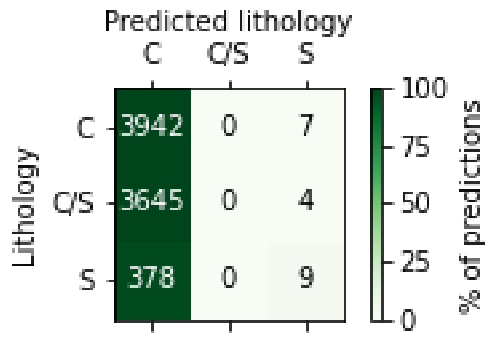

Figure 8: Confusion matrix for lithology prediction.
C stands for Claystone and S for Sandstone.

By far the most frequent error is the identification of Claystone/Sandstone as Claystone, with no Claystone/Sandstone correctly identified. This is an inevitable consequence of the lithology prediction model being based on supervised machine learning, predicting categorical classes rather than the underlying geology, and having been trained with the Geostar dataset, in which no Claystone/Sandstone lithology class appeared. A model based on supervised learning can only make predictions based on the data used to train it, while a model trained to predict only categorical classes is unable to infer any connection between the class 'Claystone/Sandstone' and the classes 'Claystone' or 'Sandstone'. Consequently, the absence of the Claystone/Sandstone class from the training data prevents the lithology prediction model from predicting it.

The next most frequent error is the identification of Sandstone as Claystone. The lithology prediction model performs better than random at Sandstone identification, predicting this class for 2% of actual Sandstone compared to 0.2% of actual Claystone and 0.1% of Claystone/Sandstone, but Claystone is predominant in the Geostar dataset used for training, so the model preferentially identifies this lithology class. As a result, the model correctly predicts Sandstone too rarely for a GAM to provide information about the factors that influence such predictions.

2.3.3 Errors in drilling problem identification

Since no significant drilling problems were encountered during field testing, the errors in drilling problem identification are the 'false positives': identifications of an anomalous datapoint as an indicator of a potential drilling problem. Variability in error is therefore variability in the probability of identifying datapoints as anomalous, as indicated in Figure 9.

The drilling problem identification model is based on unsupervised machine learning, and so no external training dataset determines variation in the frequency with which datapoints are identified as anomalous. Instead, the model identifies as anomalous those datapoints that are easily isolated from all previous datapoints from lesser depths in the same borehole. Datapoints with extreme values or in hitherto unexplored regions of the parameter space will therefore be identified as anomalous by design, and the variability in identification rates that is of interest is that with Depth.

To examine variability in identification rates, the fitted GAM specified that a given datapoint is identified as anomalous with probability dependent on constant ϕ_0 and scalar function s composed of thin-plate regression splines:

$$p = \{1 + \exp[-\phi_0 - s(\text{Depth})]\}^{-1}$$

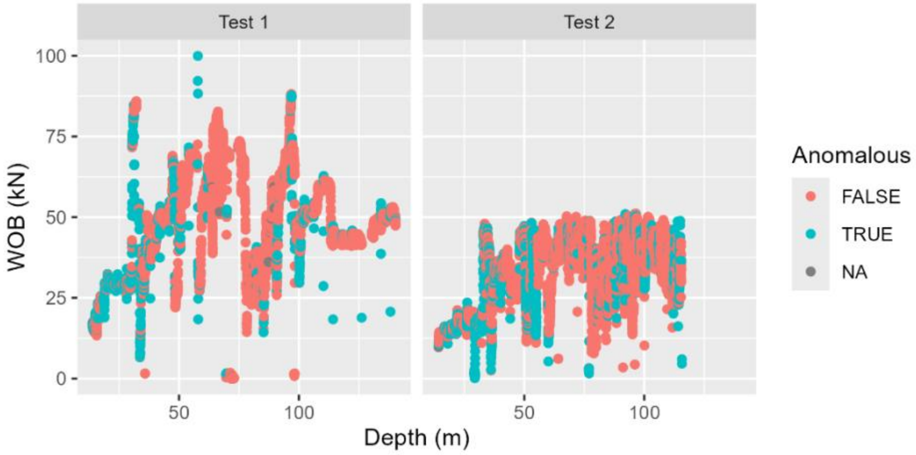

Figure 10: Datapoints identified as anomalous, with Depth and WOB

Again, the fit was conducted with R package mgcv (Wood, 2017) and with the maximum number of degrees of freedom constrained, to 10, but now a single GAM was fitted to the results of both field tests, since the unsupervised model was identical in each case. The best-fit intercept was $\phi_0 = -1.39 \pm 0.02$, while Figure 10 represents the best-fit function $s(\text{Depth})$.

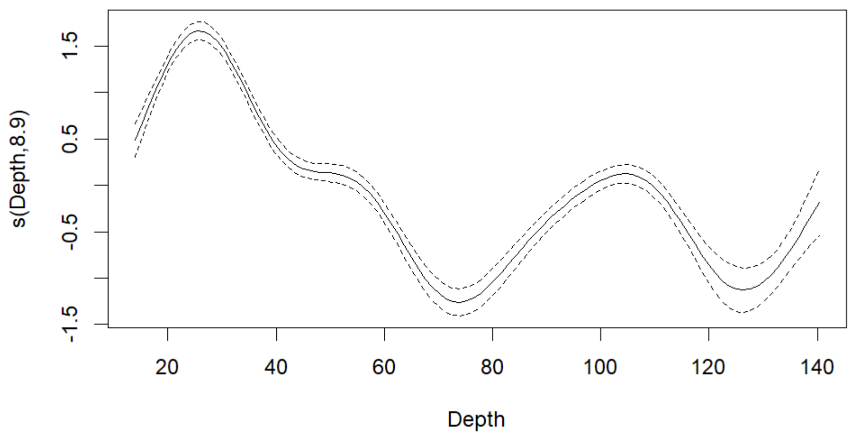

Figure 9: Logit influence of Depth on the probability of a datapoint being identified as anomalous, in the best-fit GAM

The results of the GAM fit should again be approached with caution, but it is clear that the machine-learning-based model is more likely to identify datapoints as anomalous at shallow depths, and less likely thereafter. This results from the growth with depth of the information available to the model: any region of parameter space appears anomalous when observed for the first time, but does not after repeated occurrences. And as drilling continues, the probability of a new region of parameter space being observed in the normal course of drilling will decrease. Consequently, the machine-learning-based model is more likely to falsely identify datapoints as potential drilling-problem indicators at shallow depths.

2.4 Implications for economic impact estimates

The previous section demonstrates that model predictions' accuracy varied significantly over the two field tests, with generalisation error increasing as the data with which predictions are made differed more from the data with which models were trained. Variation in accuracy was substantial even over short intervals of drilling in similar conditions. Consequently, the accuracy observed in field tests is very unlikely to be representative of accuracy in different drilling conditions, using different equipment at much greater depths. To estimate the potential economic impact of the OPTIDRILL advisory system, however, it is necessary to make extreme simplifying assumptions.

3. ECONOMIC IMPACT ESTIMATES

The OPTIDRILL project developed a drilling advisory system utilising machine learning methods to predict ROP, lithology, drilling problems, and unite those methods under one system for drilling process optimisation and intelligent decision making. The drilling advisory system advises drillers on optimum drilling parameters to improve efficiency. The drilling advisory system recommended drilling parameters aim to reduce mechanical specific energy (MSE), i.e., the most efficient drilling in terms of energy consumed. MSE is a key performance metric representing the energy required to remove a unit volume of rock (Xiao, Liu, & Tan, 2019). It provides a measure of the efficiency of the drilling process, helping operators assess and optimise the performance of the drill bit and overall drilling system. Optimising MSE minimises wasted energy during drilling, reducing fuel usage and carbon footprint. MSE minimisation helps in maximising ROP (rate of penetration) and minimising wear on tools, which minimises costly replacements and delays. It also reduces non-productive time (NPT) through better drilling efficiency, lowering operational expenses, including rig rental, fuel, and personnel costs.

A hypothetical greenfield geothermal power plant was used to evaluate the economic performance of OPTIDRILL drilling advisory system. The goal of the economic analysis was to evaluate drilling OPTIDRILL optimisation technology. Therefore, we have kept other sub-systems of the geothermal plant the same. Also, we did not consider transmission lines in the analysis.

Table 1: Economic analysis hypothetical power plant

	Case study
Resource temperature	200°C
Well depth	5 km
Conversion technology	Double flash
No of production wells	10
Project life	30 years
Capacity factor	95%
Transmission line length	0 km

The rate of penetration (ROP), i.e., how fast we are drilling a wellbore and bit lifetime, depends on rock formation and drilling depth. The economic benefits of new drilling technologies come from drilling faster or replacing worn-out drill bits less often. Table 2 presents the rock formations to 5500 meters in the case study locations. The upper 1000 meters in the Upper Rhine Graben is different soft rock types; the rest is hard rock. Table 3 presents rock types, ROP and bit life collated from literature review. OPTIDRILL demo encountered soft rocks Claystone and Sandstone. Hence, we lack data on the ROP improvement in e.g. Granite from the drilling advisory system. Therefore, the analysis considers ROP improvement in line with the Bochum demonstration. Though we expect drill bit lifetime also improved due to MSE optimisation, we have assumed same bit lifetime for the analysis as the data gathered from the Bochum demonstration was insufficient to evaluate drill bit lifetime improvement.

Table 2: Rock formation in the case study (Genter, et al., 2016)

Rock formation	Depth (meter)
Tertiary and Jurassic sediments	750
Keuper (dolomite, shales, or claystone)	50
Shelly limestone	200
Coloured sandstone	400
Granite	4100
Basalt rock	4500

Table 3: Drilling performance data (Baujard, et al., 2017) (Hackett, Blankenship, & Robertson-Tait, 2020) (Thorhallsson, Matthiasson, Gislason, Ingason, & Palsson, 2003) (Deliverable report D9.1)

Rock type	Drilling technology		ROP (meter/hr)		Drill bit lifetime (meter)	
	Existing technology	OPTIDRILL	Existing technology	OPTIDRILL	Existing technology	OPTIDRILL
Tertiary and Jurassic sediments	Tricone	OPTDRILL drilling advisory system	4.5	ROP improvement in line with Bochum demonstration	2000	Same as existing technology
Keuper (dolomite, shales, or claystone)						
Shelly limestone						
Coloured sandstone						
Granite	PDC		7.16		164	

Wellbore drilling cost depends upon the speed at which the drill bit penetrates rocks (ROP) and the cost and time spent replacing damaged drilling components like the drill bits. Services like drill rigs are rented daily, and faster drilling reduces such costs. Therefore, drilling technology able to penetrate rocks faster and dill longer without replacing components will spend more time on actual drilling and reduce overall drilling costs. ROP and bit life is dependent on the rock formation being drilled. Progress is slower in harder rocks, and lifetime is also lower. Therefore, our wellbore cost model must consider rock formation encountered in every section of the wellbore and the expected ROP and life there.

The levelised cost of energy (LCOE) methodology was used in the economic evaluation. LCOE estimates the representative cost of generating electrical power from a plant over its lifetime and is used to compare different methods of electricity generation and is the average revenue per unit of electricity that would be required for a power plant to break even. Table 4 summarises the results of the economic impact evaluation. It considers predicted ROP from the OPTIDRLL drilling advisory system in addition expected ROP which is calculated based on the recommended drilling process parameters. Expected ROP value is theoretical and based on the model predictions. We have used the following formula to calculate the expected ROP (Teale, 1965):

$$MSE = \frac{WOB}{Bit\ Area} + \frac{2\pi * RPM * Torque}{Bit\ Area * ROP}$$

The economic impact evaluation shows that with OPTIDRILL drilling advisory system, wellbore drilling costs reduce by 2.43% - 6.03%, and drilling time reduces by 3.31% - 8.22% contributing to the reduction of LCOE (1.69% - 4.19%). This reduction comes from increased drilling speed resulting in less time to do actual drilling. Due to the shallow nature of the demonstration drilling at Bochum, Germany we could gather data on the ROP improvement only, not lifetime improvement due to less wear. Therefore, though we expect equipment lifetime to improve due to less wear, we considered ROP improvement form the OPTIDRILL drilling advisory system only.

Table 4: Summary of the economic impact evaluation

		SOA	Predicted ROP		Estimated ROP	
LCOE	*€/ kWh*	0.220	0.217	-1.69%	0.211	-4.19%
Drilling	*€/ kWh*	0.093	0.091	-2.43%	0.088	-6.03%
Drilling time	*Day/well*	39.83	38.51	-3.31%	36.56	-8.22%
Well Cost	*€million per well*	16.32	15.92	-2.43%	15.34	-6.03%

4. CONCLUSION

Despite challenges in other work packages, the OPTIDRILL drilling advisory system was successfully applied in two field tests. In each of these tests, the system provided predictions for the rate of penetration during drilling (ROP) and the lithology class being drilled (Lithology) derived using models based on supervised machine learning; identification of potential drilling problems derived using a model based on unsupervised machine learning; and recommendations for drilling parameters derived from the ROP prediction model.

The accuracy of model predictions varied significantly over the field tests, in ways that corresponded to the data available to each model. In both tests, the phase space of the data with which the ROP prediction model was trained did not cover the phase space of the data to which the model was applied, and the model predictions were subject to higher generalisation error when data were from new regions of the phase space. In the first test, for which actual lithology was recorded, the Lithology prediction model failed to identify a significant lithology class - a direct consequence of this class being absent from the data with which the model had been trained - and overpredicted the lithology class that was most frequent in the training data. Meanwhile, the drilling problem identification model was more likely to identify datapoints as anomalous at shallow depths, when less data was available to indicate the standard range of parameters during normal drilling.

Because of the variability of model prediction accuracy within the field tests, estimates of accuracy in other conditions are unreliable. The magnitudes of ROP prediction errors differed substantially even within each field test, so that changes in drilling conditions are likely to change the error magnitude substantially. Consistent accuracy in such a model’s application will depend on extensive, representative training data, especially for the Lithology prediction model, with specific requirements described in the Annex, section 5. The difficulty of gathering representative data is indicated by the variation between parameters’ phase spaces even when drill sites were nearby, drilling equipment similar, and drilling measurements the same. Furthermore, the system’s accuracy in the observational setting that was explored during the field tests may differ from its accuracy when used to make interventions: the behaviour of drilling operators may lie in its own phase space, with generalisation error when the drilling recommendation system leads to behaviour outside that normal behavioural range.

However, simplifying assumptions can be made to estimate the potential economic impact of the OPTIDRILL advisory system. Considering only improvements in ROP due to the recommendation system, assuming that model predictions for these improvements are correct, and taking the improvements predicted during field tests to be representative, the advisory system could reduce the cost of drilling a 5-km deep geothermal well by between 2.4% and 6.0%.

5. ANNEX: Optidrill recommended data capture for safety and enhanced drilling

5.1 Geological constraints on the delivery of the OptiDrill Project

Geothermal heat from igneous provinces and sedimentary basins presents significant opportunities for heat extraction in many European countries. OptiDrill seeks to evaluate mechanisms for optimising drilling of deep geothermal wells to support both electrical power generation and to support district heating networks. Take-up of geothermal opportunities have been slow across Europe. Many of these prospects are theoretical, rather than tested due to the high costs in developing geothermal wells to test heat reserves at significant depth.

Minimising the costs of drilling through optimising the processes for geothermal resource extraction is a complex business opportunity with long payback periods. Geothermal heat extraction in basins requires drilling through several vertical kilometres of over burden to hopefully encounter hot fluids in permeable strata at significant depths which have extensive reserves of heat which will be continuously refreshed. The cost of rig rental for rigs capable of drilling to thousands of metres depth may reach several €100000 per week, without the immediate pay off opportunities that can be found if significant oil reserves are uncovered by drilling operations. Minimising interruptions and lost-time incidents to the drilling process is therefore highly important to delivering effective cost control of the drilling process. Therefore, examining prior drilling in the region into similar strata will allow important insights to be derived.

Given the low numbers of geothermal wells available to inform this process, it makes sense to examine data from the other deep subsampling processes in the subsurface. The oil industry also locates reserves of fluid at depth in permeable so presents a useful analogue. In countries like the UK, they provide the only ubiquitous dataset of drilling and geological data into significant depths that encounter strata of relevance to geothermal prospectivity. Where geothermal boreholes are drilled in Europe targeting deep sedimentary strata, likely candidate for development as geothermal opportunities are strata which are productive for oil at depths of 1.5 to 2.5 kilometres but in deeper parts of the sedimentary basins. The presence of recoverable oil or gas indicate that there are permeable zones from which fluids/ gases can be recovered. The extension of these same strata into the deepest parts of the sedimentary basin are therefore amongst the best possible locations for identifying zones of maximum potential permeability sufficient to deliver continually refreshed hot fluids.

Recent changes in European government data release regulations from the oil industries, and also coal and minerals industries, means that are now publicly accessible datasets of historic “legacy” drilling and geological data. The UK and Netherlands have particularly enlightened data regulation with wholly open release. For geographic reasons this report concentrates upon the data made available under UK data release regulations but is applicable to other data in other locations.

Assessment of subsurface properties and the associated drilling parameters is therefore much easier than it would have been in the pre-public data release period, or in jurisdictions where such data are not routinely available. However, the efficacy of work undertaken within the OptiDrill project has been hampered by the lack of standardisation or data available for such analysis. This report summarises best practise and the nature of optimum legacy datasets for future studies, thereby enabling the insights developed using OptiDrill to be used most effectively.

5.2 Drilling data

During the process of drilling a geothermal well significant amounts of data are generated, whether this be data derived from the rig itself, or logging data created by specialist tools lowered into the well or incorporated in the drill string's borehole head assembly (BHA), to characterise the lithology and fluids present.

The rig data serves two purposed, parameters such as weight on bit (WOB) pump strokes (RPM), standpipe pressure (SPP) and torque (TRQ) provide an insight into the performance of the bottom hole assembly (BHA) allowing the driller to optimise the parameters to achieve better drilling performance.

The second use of the data is around safety, mud weight plays a critical role in maintaining hole condition while also preventing unwanted fluids from entering the borehole (Kick) and while drilling is closely monitored for volume changes and gas content.

The second data stream is the downhole logging of the wellbore, this is generally defined by the client, and is designed to answer specific questions about the downhole properties such as porosity, temperature, net to gross, stress field and permeability that has been encountered by the well. Often this data isn't available until the well has been completed and the rig moved off location. However, increasingly some of this data is captured real time using specialist tool contained within the BHA, providing valuable insights such as well trajectory, casing point decisions and potential section that might be viable for coring that can be transmitted to the well head in real time using acoustic pulses through the mud column.

There is a third data stream that sits between the rig data and the downhole logging, Mudlogging is primarily a geological service, but also provides additional back up sensors to the ones already installed on the rig, with the most important on being gas detection. Unlike the rig data, the mudlogging data is a constant monitoring service and produces a continuous digital copy of the data combined with a description of the geology derived from the drill cuttings taken from the shale shakers.

The abundance of data that is recorded is captured in a variety of ways., and ultimately is used to define best practices for follow on well drilling campaigns. At the end of operations, and within 6 months, under current UK (and Norwegian) regulations it is a requirement that all data be uploaded to a publicly searchable online database (unless commercially sensitive). Under UK regulations almost all such drilling and geological data is openly released.

5.3 Rig data – surface measurements

The drilling rig monitors all aspects of the mechanical operations, from the mud pump pressures and strokes, through to the torque and weight on bit, mud pumped in and mud returned. all are served electronically on screens to the drillers, and other remote staff. Monitoring of these parameters can be an early indication of downhole issues such as drill pipe washouts, stick slip, rotational torque. While this data is vital during drilling in real time, its greater value comes when plotted as continuous logs along with some of the mudlogging data. As a minimum, the WOB, ROP TRQ and SSP should be plotted along with GR, Mud gas both total and chromatographic and a simplified lithology column along with cuttings descriptions.

To optimise further drilling in the development, all measurements taken at the rig should be stored digitally, traditionally this is a service that is offered by the mudlogging company and is detailed below in the CPI log plot section.

These outputs are typically similar from one contractor or operating company to another allowing their easy ingestion into drilling data tools.

5.4 Daily drilling reports

A major problem within the OptiDrill project has been the lack of consistency experienced in Daily Drilling Reports (DDR) requiring them to be simplified and standardised before any ingestion into data tools to be used for understanding their mechanics and logistics of drilling processes des. This section describes typical attributes of such a report but which are not usually followed exactly.

Data captured by the report should include, but not necessarily limited to the following:

- General: Rig name, well name, location spud date and operator
- Crew details: number of crew, shift details and positions
- Drilling: progress made in 24 hours and current depth
- Type of drill pipe, both currently in the hole and stored in the racking system
- Casing data, both in the hole and stored at the well site
- Mud properties in the hole and mud chemicals available on board
- Current BHA configuration
- Geological : formations encountered along with formation tops
- Operations: a brief overview of the last 24 hours
- Equipment: general condition, planned maintenance
- Safety: any incidents recorded, planned emergency drills, environmental spills and issues, including how to address them.
- Forward: the following 24 hour plans and operations

The DDR is designed to be an all-encompassing resource that can be used by multiple disciplines. Contractors need to be able to follow the rig operations in preparation to deliver both staff and materials in timely manner. Safety officers will need to plan safety drills, often related to a specific up coming operation, such as well testing or running casing. The daily drilling report should be distributed to anyone who has a business interest in the well. An example of a standard DDR is included in Appendix 1.

5.5 Mudlogging

The mudlogging unit at the rig site is a high-tech environment harvesting data from the drilling process along with physical data in the sense of rock cutting from the drilling process. The duties of the mudlogger are to compile a concise record of the subsurface and any major drilling events that may occur, such as kicks or losses in the drilling mud system, variations in stand pipe pressure or increase in mud gas levels.

In terms of safety the two most important aspects of the mudlogging are mud gas monitoring and the monitoring of the overall volume of mud within the closed drilling system. The data that is recorded in the mudlogging unit is displayed and stored as a continuous plot. As with the DDR the mudlog should be updated and distributed on a 24 hour period, and will form part of the daily meeting between the wellsite, client and logging contractors.

Data contained on the mud log should include

- Well metadata
    - Well name, rig name, contractor supplying the service,
- Well contsruction metadata
    - Casing points, Bit information after each bit change or casing point
- Orientation data
    - Hole inclination and azimuth as text data plotted at set intervals
- Drilling parameters
    - Continuous parameters, WOB, ROP, RPM, Torque, SPP
- Drilling fluid parameters

- Mud temperature in and mud temperature out
    - Mud weight in and mud weight out
    - Gas values from the mud stream, both total and C1 – C5 breakdown
- Geological attributes
    - Simplified lithology column.
- Geomechanical data
    - Formation integrity tests (FIT) or leak off tests (LOT)

The mud log plot is used by engineers and geoscientists alike; for geoscientists including the Gamma Ray log as part of the mud log data stream can allow stratigraphy to be followed in real time. Understanding these data in context of the lithostratigraphy allows for better planning of geological data, as the upcoming lithologies are known and operations can be planned accordingly.

Post-well operations, the mudlog forms the only data source that provides a continuous depth record of the rig data and the lithology encountered.

Post-well operations the final mud log is usually delivered to the client in digital format and will be a record of all data from the initial spudding of the well to the TD. It should typically include all the information stated above.

For engineering aspects the data can be used to optimise casing points, select drill bits and optimise parameters to increase drilling rate based on WOB and SPP for future wells. The advantage of the mud log is in relating all the drilling parameters to an actual lithology complied by analysing the cuttings stream is a simple visual plot.

5.6 LWD / Downhole logging requirements

Natural Gamma ray (GR) – In modern well drilling to significant depth this almost always run in any BHA as part for the Measurement While Drilling (MWD) program, cheap, reliable and passive. Gamma ray Primarily used for correlation between wells. Historic archived Gr data can be used as a base line to identify scale build up if the well is re logged during its lifecycle.

Well orientation data – Data derived from the MWD tools also typically includes well orientation data, principally hole azimuth and deviation enabling well path to be fully reconstructed to high vertical and horizontal accuracy.

Caliper – This may be either mechanical or Ultrasonic. Caliper logs will give in real time an indication of borehole condition and where remedial action may be required. Washouts can make successful cementing harder to achieve. Formations that squeeze into the borehole will require additional reaming prior to running casing, identifying formations that either squeeze or washout can help in optimising further well drilling operations later in the development phase. Aside from hole condition the caliper can be used for determining stress direction enhancing potential fracture operations.

Spectral gamma ray (SGR) – similar to the GR tool, but only records radiation from 3 energy windows, Thorium, Potassium and Uranium. As with the regular Gr log, the SGR log is a valuable correlation tool, and can also give an insight into shale mineralogy. Open fractures and potential build up of salts or scale.

Neutron – secondary porosity log, useful in carbonates where porosity can be read directly from the real time log. An epithermal neutron tool will provide a baseline survey for future cased hole logging of fluids.

Density – primary tool for determining porosity and lithology; increasingly this can also be run as a image tool, primarily for aiding geosteering, but will also image fractures and sedimentary structures. Additional to the density measurement the density tool also calculates a Photo electric factor (PEF) for lithology identification.

Sonic – cross-dipole orientated sonic log provides a definitive indication of the stress field for potential artificial fracturing should the reservoir require stimulation. Porosity from the sonic log, unlike the density log only records the connected pore volume, this is useful in Vuggy carbonates where differentiation between the connected and unconnected pore volume may be large. In conjunction with the density log both compressional and shear sonic can be utilized to create a synthetic seismic trace to calibrate the lithology to the seismic.

Downhole drilling data - Downhole monitoring of WOB, TRQ, stick slip, and whirl (torsional energy build up in the drill pipe) and vibration. Provided real time will optimise drilling, when presented as part of the mudlog will give insights into improved drilling practice.

Temperature – Most LWD tools measure temperature as part of the measurement correction process. This meaurement is strongly affected by the lack of well equilibration unless run in a static water column that has equilibrated for an extended period and by the surface weather conditions that effects mud temperature. Though this will not be a true formation temperature it can be an indication of an increasing geothermal gradient. For an accurate formation temperature the well will require to be shut in for a period of time to reach thermal equilibrium, then logged with WL

Pressure – dedicated formation pressure tool is widely run in the oil and gas industry to confirm connectivity between various reservoir units. This measurement is not without significant risk of differential sticking, if pressure measurements are required it is better to run a dedicated wireline tool.

Core – in an undeveloped geothermal prospect coring of the reservoir is vital to provide a continuous record of the lithology for both sedimentological analysis and lab based studies, most notable being the development of a porosity permeability relationship, log calibration and confined and unconfined stress strain relationships. Dependant on the requirement, core can either be collected continuously using wireline retrieval inside the drill pipe, or if larger core diameter is required traditional coring can be employed, cutting 30ft barrels and tripping out the whole BHA for retrieval.

The above logging suites are considered the minimum data acquisition that would allow a full characterisation of the reservoir. As the development becomes more mature more specific questions may arise and the data logging can be adjusted to be more specific. With core data uncertainty to porosity permeability will be reduced, it may be possible to do without the neutron density logs, avoiding the risk of potential loss of a radioactive source downhole, replaced by the NMR tool.

Borehole imaging log suite: this provides understanding of the fracture conditions and lithological situation within the borehole and to enable identification of hole problems and geomechanical parameters such as borehole breakouts indicating in-situ stress anisotropy and orientation.

5.7 Fractured reservoirs – data acquisition

Hard rock fractured reservoirs pose a significant challenge to logging, rather than intergranular permeability, the reservoir will be produced through a set of open fractures, for these types of reservoir drilling has the associated risks of sudden gains/losses in the drilling mud, fractures are notoriously difficult to plug with lost circulation material. With this in mind, coupled with less requirement to characterise the host rock a much more basic LWD program would be sufficient. Once the well is drilled and under control, Wireline could be used to gather the main data acquisition program.

5.8 LWD Down hole logging – Fractured reservoir

Natural Gamma Ray – as part of the MWD deviation and azimuth assembly

Ultrasonic caliper – real time indication of borehole condition, fracture identification, orientation and if the fracture is open or closed.

Given the high temperature that are likely to be encountered and the potential for borehole instability due to losses or kicks the above data acquisition program will provide sufficient data to characterise the reservoir and inform future field development.

5.9 Conclusion

Geothermal boreholes are technically complex and expensive to deliver with a payback time measured in decades. The absence of a corpus of legacy data directly related to the specific drilling challenges of geothermal boreholes means that there is a limited pool of data available for use in planning of future wells. Whilst there is a continuity of practice for drilling of oil wells which can be used to train data for the geothermal industry, there is a need for development of a similar corpus of geothermal data. The increased depths and the need to ensure borehole stability to enable the long-term production of heated fluids means that boreholes need to be drilled with care. It is therefore essential that a native-digital standard for propagation of Daily Drilling Reports is established that allows easy data extraction so that these data can be easily extracted and prioritised for learning from.

6. REFERENCES

Baujard, C., Hehn, R., Genter, A., Teza, D., Baumgärtner, J., Guinot, F., . . . Steinlechner, S. (2017). Rate of Penetration of Geothermal Wells: a Key Challenge in Hard Rocks. *42nd Workshop on Geothermal Reservoir Engineering, Stanford University.* Stanford, California.

Genter, A., Baujard, C., Cuenot, N., Dezayes, C., Kohl, T., Masson, F., . . . Vidal, J. (2016). Geology, Geophysics and Geochemistry in the Upper Rhine Graben: the frame for geothermal energy use. *Proceedings European Geothermal Congress, EGC 2016.* Strasbourg: International Geothermal Association.

Hackett, L., Blankenship, D., & Robertson-Tait, A. (2020). Analysis of Drilling Performance Using PDC Bits, Fallon FORGE Well 21-31 . *45th Workshop on Geothermal Reservoir Engineering, Stanford University.* Stanford, California.

Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized Additive Models.* Chapman & Hall/CRC.

Teale, R. (1965). The Concept of Specific Energy in Rock Drilling. *International Journal of Rock Mechanics. 2.*, 57-73.

Thorhallsson, S., Matthiasson, M., Gislason, T., Ingason, K., & Palsson, B. (2003). Iceland Deep Drilling Project (IDDP): The challenge of drilling and coring into 350-500°C hot geothermal systems and down to 5 km. *Iceland Geothermal Conference.* Geothermal Association of Iceland.

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, Second Edition.* New York: Chapman and Hall/CRC.

Xiao, H., Liu, S., & Tan, K. (2019). Experimental Investigation of Force Response, Efficiency, and Wear Behaviors of Polycrystalline Diamond Rock Cutters. *Applied Sciences 9 (15)*.