D13.1: Report on field testing of the OPTIDRILL system

Deliverable No	D13.1
Work package No. and Title	WP13: Task 13.1 Field testing of the OPTIDRILL system in a real borehole
Version - Status	V1.0 - Final
Date of Issue	01/12/2024
Dissemination Level	PUBLIC
Filename	OptiDrill D13.1 Report on field testing of the OPTIDRILL

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
Kevin Mallin	Geolorn Ltd	kevinmallin@geolorn.co.uk
Pufinji Obene	OLIVERIS Tech Incubator Ltd	pufinjiobene@oliveris.com
Henning Knauer	Fraunhofer IEG	henning.knauer@ieg.fraunhofer.de

Document History

Date	Version	Editor	Change	Status
1/12/24	0.1	Kevin Mallin	Document creation	Draft
14/12/24	0.2	Pufinji Obene/Pufinji Obene	Document review	Draft
21/03/25	Rev 0.1	Kevin Mallin	Document revision	Draft
25/03/25	Rev 0.1	Henning Knauer	Added contents from D13.2 Added concise decription of AI developments	Draft
20/04/25	V1.0	Shahin Jamali	Final revision	Final

TABLE OF CONTENTS

LIST OF FIGURES

EXECUTIVE SUMMARY

Deliverable D13.1 focused upon the validation of the OptiDrill Drilling Advisory system in a field-scale operation. This involved bringing together all the work carried out in WP's 4, 11 & 12, into the operational system.

The original plan was to use a drilling contractor willing to run the in-hole sensors and a proprietary 'wired' drill pipe, such as the NOV Intellisense™ system. Due to some delays in achieving all the work required to bring the system to field ready status, this proved to be an insurmountable problem.

The project, therefore, decided to test the system in two formats; sensors placed on the rotary head of the drill rig and connected to the drilling advisory system, to test the drilling data transfer and interpretation module, and then to place the data sensor subs into a drill string assembly and record the drilling data on memory sticks.

Both testing regimes proved very successful, validating that the sensors, the advisory system and the sensor subs, worked in an operational environment, transferring data in real-time. The, obvious, next step is to combine the entire system with wired drill pipe, which will be achieved either through EIC grant funding, or co-operation with commercial operations.

The sensor sub testing highlighted the difficulties of keeping power available, but battery advances will alleviate this and with energy harvesting systems becoming more available, this issue will be quickly overcome.

Overall the OptiDrill system to identify lithologies, predict ROP and provide an advisory system to the drilling crew, has proven a great success, whilst identifying the additional developments required to provide a fully commercial system.

Deviations from the Actions Described in the Grant Agreement (for D13.1)

D. No.	Action according to GA	Action performed / Deviations
D13.1	A suitable project will be identified at an early stage of the OPTIDRILL project and the operator/contractor will be approached to run the system (possibly with a third-party drill string, such as NOV's Smart Pipe).	It was not possible to engage a contractor operating NOV's IntelliSense™ or any other proprietary wired drill string during the project timeline. Instead, the project team opted for a two-part testing regime: (i) real-time surface testing of the OptiDrill advisory software integrated with a Fraunhofer drill rig and (ii) field testing of battery-powered sensor subs installed within the Bottom Hole Assembly (BHA), recording data on memory sticks for post-analysis. These methods preserved the deliverable's core validation aims while addressing commercial and technical constraints.
	Testing and validation of OPTIDRILL system in a real borehole.	A real borehole was used at the Fraunhofer IEG site as part of the PUSH-IT EU project and a separate site in the UK. While the advisory system and hardware components were tested in operational drilling environments, full integration with wired drill pipe systems was not achieved. Nonetheless, the advisory system's modules (ROP prediction, lithology classification, and anomaly detection) were successfully validated during the drilling of multiple wellbores.

	Combined operation of hardware and software for real-time system feedback.	Full real-time integration between hardware (sensor subs) and the advisory software system was not feasible due to the absence of downhole connectivity infrastructure. Instead, real-time operation was demonstrated using surface sensors and simulated data flow, while subsystems operated in a standalone mode for downhole testing. This modular approach, although a deviation, allowed separate validation of software and sensor functionality, ensuring flexibility for future integration.

Addendum – Positive Deviations / Additional Achievements

D. No.	Action according to GA	Action performed / Deviations
D13.1	Not explicitly required in GA.	Expanded multi-environment testing: In addition to the expected single drilling site validation, testing was conducted at two separate field sites — the Fraunhofer IEG facility in Germany and a high-impact, air-powered DTH drilling site in the UK. This allowed for broader environmental validation, testing both real-time surface system feedback and downhole sensor durability under extreme conditions. The UK trials, in particular, introduced an additional layer of complexity and robustness by validating blind lithology prediction using single-IMU setups — a methodology not required by the GA but critical for real-world applicability.

1. INTRODUCTION

This deliverable reports on the testing of the OptiDrill system in field-scale test scenarios, to evaluate the robustness and efficacy of the system in operational environments.

All drilling operations, by their nature, place huge demands on equipment, whether it is utilised within the wellbore or as part of the surface monitoring systems. Downhole, there is pressure, temperature, aggressive fluids, torque, cyclical stressing and constantly varying stress and strain of the entire drill string. Surface systems face similar challenges, particularly related to weather conditions, corrosive fluids/substances, incorrect handling and extended operating periods.

It was, therefore, vital to design all components within the OptiDrill system (indeed any system) that can withstand the demands placed upon it. For the OptiDrill system, it was also necessary to ensure that downhole components could be easily integrated with current commercial drill strings and Bottom Hole Assemblies (BHA). This particularly applied to thread connections, which are generally configured in accordance with American Petroleum Institute (API) or International Association of Drilling Contractors (IADC) standards. These standards are universally adopted for well drilling operations, and whilst the tool joint nomenclature may vary, there is some compatibility between the thread configurations.

The data sensor subs manufactured for Geolorn used the following reference:

API – 3 ½" I.F. / IADC – NC38,

along with X-Over Subs allowing them to be used with:

API – 2 3/8" I.F. / IADC – NC26 drill string and BHA's, during field scale testing.

There are a wide range of tool joint configurations used in drilling, meaning that any downhole components need to be designed to meet variable drill string and BHA set-ups; the configuration chosen for the OptiDrill sensor subs were based upon the most likely availability of drilling operations at the time of eventual testing.

Initially, the testing of the sensor subs and the drilling advisory system were planned to be as an integrated operation, using a commercially available connected drill pipe, such as NOV's IntelliSense, but it was not possible to find any contractors that were running or willing to run such drill strings. Within the Horizon 2020 Geo-Drill project (Project I.D. 815319) it was demonstrated that connection from a sensor downhole was established with surface systems by means of connected tool-joints and wired drill pipe. However, this was only achieved in a drilling simulator and not in a field drilling operation.

It was, therefore, decided to test the system in two configurations:

1. With sensors placed on the rotary power swivel of a drill rig and then connected to the drilling advisory system (testing in Bochum) in real-time and,
2. With the manufactured sensor subs placed into the drill string, and the data generated recorded on a memory stick and analysed post-drilling (testing in UK).

The UK testing was to ensure that the data sensor subs could withstand the rigours associated with drilling operations; particularly make-up and break-out using an Iron Roughneck, torque and vibration, and high velocity cuttings/water. It tested both the mechanical properties of the subs, including sealing from water ingress and the ability of the sensors, batteries and memory sticks to withstand the harsh conditions.

The Bochum testing was aimed at demonstrating how data could be transferred, analysed and displayed in real-time, using the drilling advisory system developed within the project.

The drilling advisory system is based on several software modules that are using artificial intelligence methods for the prediction and optimization of different aspects of the drilling operation. The advisory system gatheres and processes drilling data that is received during the ongoing drilling operation. The processed data is on the one hand used for the monitoring of the drilling operation and on the other hand for making predictions that provide valuable information to the operator and thereby help to increase process awareness and overall efficiency. The individual AI-based software modules have the following objectives:

- Prediction of the ROP and provision of recommendations for process parameters settings for the optimization of the drilling operation
- Prediction of the drilled lithology based of the MWD data

- Detection of anomalous data that might indicate drilling problem
- Jetting performance prediction based on acoustic emission and jetting process data

State of the art machine learning methods in combination with extensive historical drilling datasets, the project partners' domain knowledge and drilling expertise have been applied to develop the software modules within individual work packages. The developed machine and deep learning models are mostly based on supervised learning algorithms such as artificial neural networks, however unsupervised learning methods have been applied as well.

After all individual software modules were developed, evaluated and validated they were unified within the final OptiDrill drilling advisory system within work package 11. The prototype is a purely software based system that handles data of a given format from an incoming data stream and provides data visualizations for monitoring purposes, as well as prediction for the respective objectives listed above based on the data that is processed. The system prototype was successfully integrated with the Fraunhofer IEG's mobile drill rig and the interface was adapted to the drill rig controller. It has to be mentioned that the sensor system that was developed in parallel could not be integrated with the final prototype and both systems were finalized as stand alone solutions. Due to the unavailability of the downhole sensor data, also the jetting performance prediction module was not included in the final OptiDrill prototype.

Before the actual field testing the prototype was tested within work package 12 with the actual drill rig, as well as with extensive raw historical drilling data derived from MWD logs of the same rig in order to ensure proper functionality of all system components. Issues identified during these tests were solved and optimizations were implemented to ensure a seamless and successful testing of the system during the planned field tests that are reported on within this and the following reports of work package 13.

2. FIELD TESTING OF THE OPTIDRILL SOFTWARE AND HARDWARE AT THE IEG IN GERMANY

The OptiDrill field test conducted in Bochum in November 2024 marked an important step in evaluating the sensor system's and the drilling advisory system's performance during a complex borehole excavation. The objective was to drill to a depth of approximately 120 meters to access a concrete bunker, passing through a variety of soil and rock layers. The test provided valuable insights into sensor and software performance under real-world conditions and offered critical lessons for future system optimizations.
The test setup included two sensors, identified as sub 57 and sub 59, attached to the central spindle of the drilling equipment. The sensors were orientated with the x-axis aligned along the drilling direction, the y-axis tangential to the drill's axis, and the z-axis normal to the drilling plane. This arrangement whilst not perfect as detailed below from a systems identification standpoint, still allowed for semi-precise measurement of the forces and dynamics experienced during drilling. See Figure 2 and Figure 2.

Alongside the testing of the sensors the OptiDrill drilling advisory system was deployed and extensively tested for the whole duration of the drilling campaign, delivering very insightful results into the system's performance and application in a real-world scenario.

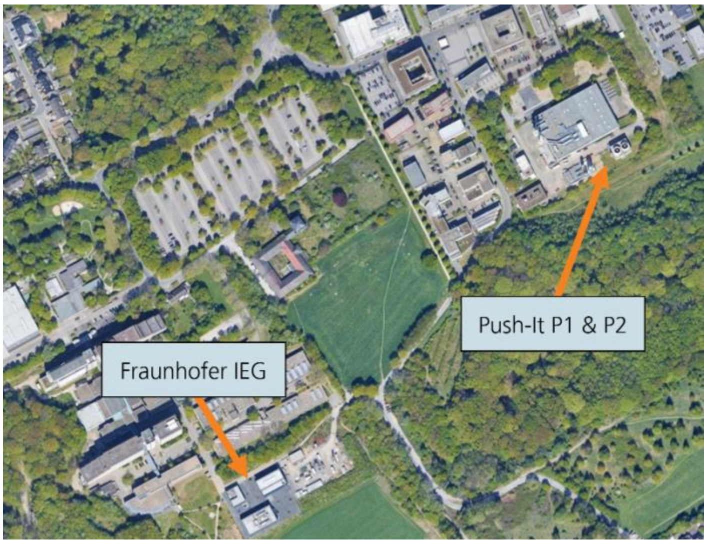

Figure 1: Location of the field test site next to the Fraunhofer IEG

2.1Drilling Campaign

As a validation site a location close to the Fraunhofer IEG in Bochum was chosen. The OptiDrill project team had the opportunity to join the drilling campaign of another EU project called Push-It[1] (Piloting Underground Storage of Heat In Geothermal Reservoirs). The drilling campaign of this project is lead by another competence centre of the Fraunhofer IEG. The objective of the drilling campaign was to drill 4 wellbores in total, all targeting an abandoned and sealed coal mining shaft that was converted into a bunker

Figure 2: Drill rig at the field test site in Bochum

during the second world war. The old underground building complex, that lies directly underneath a heating plant, is planned to be used as a thermal underground storage within the Push-It project. Two of the four planned wellbores were finished until end of November. The targeted depth was planned to be between 115 and 120m, depending on when the ceiling of the abandoned underground building would be encountered and penetrated,

and the borehole diameter is 324mm (12.75”). The drill rig used for the drilling campaign was the Fraunhofer owned 150 ton rig from *Hütte Bohrtechnik* which can be seen set up at the drill site on the figure above. The interfaces of the OptiDrill system prototype have been developed for this particular drill rig and had been integrated and tested beforehand. The drill rig controller software was updated by the manufacturer to enable the seamless data transfer from the drill rig to the OptiDrill system over an ethernet cable.

The location of the drill site that was chosen for the field testing of the OptiDrill system can be seen in the figure above. The site lies in close proximity to the Fraunhofer IEG and also just around 600m from the site of a former drilling campaign that has been used as one of the development datasets for the machine learning module development. This is important since historical drilling data has to be available for the training of the supervised learning based machine learning models of the OptiDrill system.

[1] PUSH-IT – Piloting Underground Storage of Heat In geoThermal reservoirs

2.2 OptiDrill Software System Field Testing

The final OptiDrill Drilling Advisory system prototype had been integrated into the Fraunhofer *Hütte* drill rig and tested extensively. The results from the testing were used to perform a few final optimizations of the software system that were implemented into the final version of the prototype.

The OptiDrill software system has been testing during the complete drilling operation of the first two wellbores of the Push-It project planned for 2024. The first wellbore was drilled in late September from the 25th to the 30th and the second wellbore around one month later from the 4th of November until the 13th of November. Both wellbores were finished without any major issues. However, since there was a significant deviation in the drill path trajectory of the first wellbore, leading to the targeted bunker being missed, some changes to the drilling equipment and some limitations to the drilling process parameter ranges had to be made after finishing the first wellbore. Due to the deviation in the drill path, which was not noticed during the drilling operation and first found out after a deviation survey after the drilling was finished, the total depth of the first wellbore is deeper than the original target depth. It was decided to drill deeper in case the bunker was actually located deeper than it was calculated based on the limited documentation that was available from the time of its construction. The following table gives an overview of the two wellbores the OptiDrill system has been tested on.

Table 1: OptiDrill field testing overview

	P1	P2
Start of drilling	25th of September	4th of November
End of drilling	30th of September	13th of November
Average ROP [m/h]	15	7
Total days of drilling	4	8
Total hours of run-time [h]	17	41
Total footage recorded [m]	127	102
Final Depth [m]	150	115

In total the OptiDrill software system was run for almost 60 hours of drilling operations that covered around 230m of drilled footage. Looking at the total hours of operation and the average ROP values it can be seen that the second wellbore was drilled significantly slower and thus took more time to finish although it was

Figure 3: Push-It drill site

shallower than the first wellbore. This was caused due to the changes of equipment and the limitations concerning the drilling process parameters mentioned before. The second wellbore successfully hit the bunker at the expected depth and no significant deviation in the drill path trajectory was encountered.
The picture (Figure 3) above gives an impression of the drill site of the Push-It project that was chosen for the field testing of the OptiDrill system. The OptiDrill Drilling Advisory system running on a rugged outdoor laptop was setup underneath the tent on the right in safe distance from any machinery or other hazards present at the drill site. The laptop was connected to the drill rig controller via an ethernet cable during the whole time of the drilling operation.
Figure 4 shows the OptiDrill Drilling Advisory software during the drilling operation with the drill rig in the background. On the screen of the laptop the drilling process parameters received from the drill rig, as well

Figure 4: OptiDrill Drilling Advisory system running on the outdoor laptop during the field test

as the outputs of the machine learning modules for the ROP prediction and optimization, the drilled lithology prediction and the drilling problem detection are displayed.
For the field testing of the OptiDrill software system the complete OptiDrill Drilling Advisory System prototype, as it has been reported on in the deliverables from work package number 11, has been used. The system comprises the monitoring and data transfer system developed by the project partner TVS and the three machine learning based frameworks from the work packages 7, ROP prediction and optimization developed by IEG, 8, drilled lithology prediction developed by IEG, and 9, drilling problem detection developed by BGS.
For the training of the machine learning based models the data from the Geostar 2 project has been used. This dataset originates from a former Fraunhofer IEG drilling campaign which was drilled with the same drill rig that has been used in the field testing. It consists of 12 wellbores from the same location that cover a depth of around 150m. The dataset is time-based with one measurement every three seconds, consequently the resolution of the dataset is quite high with measurements available for almost every centimetre of drilled depth. The map below shows the locations of the Geostar 2 drill site and the drill site of the Push-It project, chosen as field test site of the OptiDrill system. The two locations are only around 600m apart and share very similar geological properties. The name of the first wellbore of the field testing is P1.
The Geostar 2 dataset contains around 90000 rows of measured drilling process parameters and covers a total of around 1600 m of drilled depth. The classes of lithologies that are represented in the development dataset comprise claystone, siltstone, sandstone, coal, and claystone/siltstone, which basically described thin intermittent layers of both types of lithology.

Multiple different predictive models were developed for the ROP and lithology prediction modules. For the field testing the best performing models that were reported on in the respective deliverables from work packages 7 and 8 were chosen, being 1D convolutional neural networks and transformer models.

For the ROP prediction a transformer model with additional convolutional layers was selected. It consists of an embedding layer, followed by a positional encoding layer, a multi-headed attention encoding layer, convolutional blocks, a flatten layer, and dense layers. The convolutional blocks consist of 1D convolutions, batch normalization, 1D max pooling, and dropout layers. The inputs to the model include the measured depth, weight on bit, torque, RPM, pump pressure, mud flow in, and the previous rate of penetration. The model takes sequences of length 30 of all of these parameters, covering around 0.3 metres of drilled depth, as inputs and outputs the predicted ROP value.

The drilled lithology prediction model chosen for the field testing is a 1D convolutional neural network model. It consists of convolutional blocks, containing 1D convolutions, batch normalization, and dropout layers followed by a flatten layer and dense layers. The input data to the drilled lithology prediction model is very similar to the input data for the ROP prediction model. The only differences are that the measured depth is discarded from the inputs and instead of the precious ROP the current ROP is used. The drilled lithology model has been trained to predict the four classes claystone, siltstone, sandstone, and claystone/siltstone allowed for semi-precise measurement of the forces and dynamics experienced during drilling.

3. Drilling Advisory System First Field Tests Results and Evaluation

In this section the results and findings from the first round of field testing are reported on. In the first section the data used for training the machine learning models for the ROP and the drilled lithology prediction is compared with the data that has been recorded during the field test. In the second section of this chapter the machine learning model outputs are evaluated and compared with the results achieved during the model training and validation.

3.1 Data Comparison

The following plot shows an overview over the drilling process parameters used as inputs to the machine learning models plotted over the measured depth of the well. The plots show one larger gap starting at around 75 metres, which was caused by a minor operating error of the drill rig, leading to negative WOB values being recorded for one whole drill pipe. Since negative WOB values are dropped during the data preprocessing, no data has been recorded for the next four metres after adding this drill pipe.
Due to significant differences in the diameters of the wellbores of the Geostar 2 development dataset,

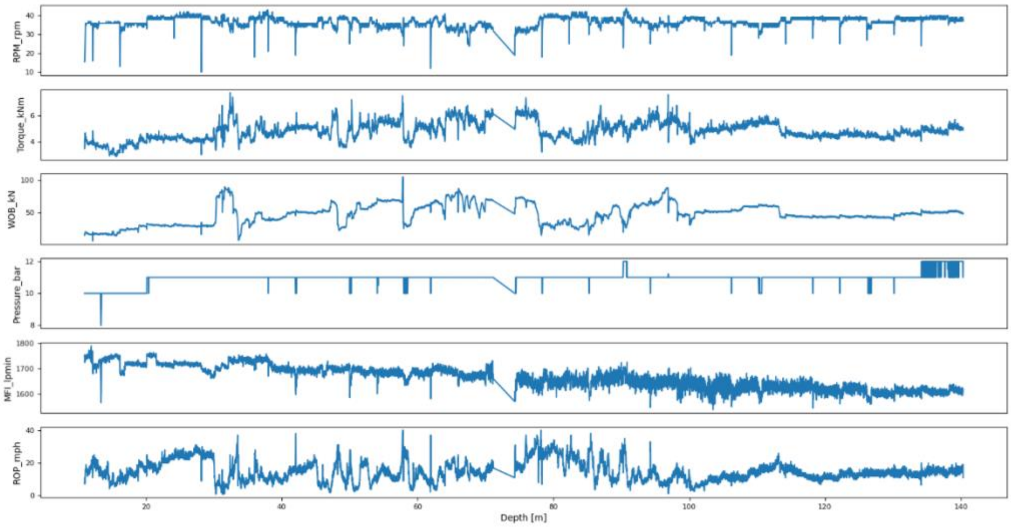

Figure 5: Plots of the drilling process parameters from the P1 wellbore processed and saved by the OptiDrill system

which was 187 millimetres, and the Push-It P1 wellbore, being 324 millimetres, the ranges of most of the drilling process parameters differ considerably. In the following two tables statistical overview over the drilling process parameters used as input to the machine learning modules are given.

Table 2: Statistical overview Geostar 2 development dataset

	Depth_m	RPM_rpm	TRQ_kNm	WOB_kN	P_P_bar	MFI_lpmin	ROP_mph
mean	77.47	92.22	2.43	33	5.76	653.98	21.2
std	35.41	8.03	0.33	10.32	0.7	84.26	7.61
min	15.02	63	1.28	0.1	2	196	1
25%	46.81	86	2.22	26	5	596	16
50%	77.8	92	2.41	32.86	6	637	21
75%	107.34	98	2.64	39.34	6	712	27
max	140.32	140	4.53	91.13	9	1280	40

	Depth_m	RPM_rpm	TRQ_kNm	WOB_kN	P_P_bar	MFI_lpmin	ROP_mph
mean	77.5	36.76	4.79	48.01	10.95	1663.9	13.49

Table 3: Statistical overview Push-It P1 dataset

std	38.01	2.44	0.66	15.57	0.35	44.04	5.43
min	11	10	2.85	5.4	8	1539	1
25%	44.12	36	4.37	39.63	11	1625	10
50%	79.98	37	4.76	48.46	11	1664	13
75%	110.71	39	5.2	59.06	11	1698	16
max	140.32	44	7.75	105.07	12	1789	40

The first table shows the statistical properties of the drilling process parameters of the 12 wells from the Geostar 2 dataset for the same depth as the P1 dataset. The second table shows the same values for the P1 wellbore dataset. Comparing the mean values of the parameters it can be seen that for each parameter the values in the two datasets are quite different. While the mean of the RPM in the Geostar 2 dataset lies at around 92, in the P1 dataset it lies around 38. The mean of the torque is 2.4 kNm in the Geostar 2 dataset and 4.8 in the P1 dataset. Very similar differences can be seen for all of the other parameters. These differences are all caused by the fact that a significantly larger drill bit diameter was used for the P1 wellbore than it had been used for the Geostar 2 wellbore. Having a larger drill bit diameter limited the RPM on the drill rig used at the IEG, led to higher torque and WOB values, lower ROP values, and required more drilling fluid circulation leading to higher pump pressure and mud flow values.
For the preprocessing of the model input data during the field testing of the OptiDrill system the min max scaler fitted to the Geostar 2 dataset was used. The significant differences in the statistical properties of the datasets, especially the min and max values in this case have a significant impact on the scaled values that are fed into the machine learning models as inputs. In this case some parameter ranges are significantly smaller in the inference dataset (P1), such as the RPM and the ROP, while others are larger, such as the torque, WOB, pressure and mud flow values. This can lead to several issues affecting the machine learning model performance:

- **Poor predictions:** The machine learning models might generate inaccurate predictions because they were trained on data with different statistical properties. If the data during inference falls outside of the range of the data the models were trained on, the models will extrapolate rather than interpolate, leading to unreliable results. Predictions can be expected to deviate stronger from the true values than it would be expected based on the results achieved on the development data. The performance metric scores achieved during validation of the models might not be reproducible and should be expected to be poorer on the inference data.
- **Bias:** The machine learning models will be biased towards the data they were trained on, causing errors in the predictions. E.g. since the ROP values in the training data were much higher on average, the ROP prediction model is likely to predict the ROP more optimistic with higher ROP values.
- **Model Robustness and Generalization:** Since the model was trained on a dataset with different statistical properties its robustness and ability to generalize might be compromised, having a negative impact on its overall validity on the inference data.

If this is the case, which is to be expected, there are basically two main countermeasures that can be takes to improve the model performance. The models using supervised learning techniques can be retrained on the new data gathered during the drilling operation or the data could be processed using a scaler that is more fitting to the inference data.
In the following section the machine learning model outputs will be examined using the same metrics that have been used for the model development. The performance metrics of the ROP prediction and the drilled lithology prediction models can then be compared to the values achieved during the validation of the models on the development dataset.
The outputs of the machine learning models have been recorded during the OptiDrill software runtime and will be discussed in the following sections. It has to be noted that at this point we will analyse the model outputs in the same manner as it has been done in the deliverables from the respective work packages. The

evaluation that is reported on at this point serves the general assessment and performance rating of the developed models and their outputs. Based on the findings from this first evaluation further optimizations will be implemented in the OptiDrill machine learning models integrated in the drilling advisory system. A more thorough analysis of the machine learning model outputs will be performed by the project partner BGS in the subsequent deliverable 13.4.

3.2 ROP Prediction and Optimization Module

The ROP prediction and optimization module is based on a machine learning regression model that predicts the ROP value from a set of sequences of drilling input parameters and the concept of mechanical specific energy (MSE). The heart of the whole framework is an artificial neural network that is capable of predicting the ROP value. Using this predictive model a number of fictitious drilling process parameter scenarios are analysed by altering the values of a set of controllable process parameters, such as the RPM or the WOB. The respective ROP values for these fictitious scenarios are predicted and subsequently the MSE values, used as a criterion to rate the efficiency of the drilling process, is calculated. Based on these calculation recommendations on how to set the controllable drilling process parameters are made and displayed in the OptiDrill software. The most critical component in this framework is the machine learning based model that predicts the ROP value, since the MSE value is directly dependant on it. Errors in the ROP prediction have a direct and significant impact on the MSE value and thereby on the selection of the system's recommendations. We can directly calculate and analyse the errors the ROP prediction model made during the field test using common performance metrics. We have calculated the mean absolute error (MAE), the root mean squared error (RMSE), and the R^2-Score based on the recorded actual ROP values and the respective predictions of the model. Furthermore, we have also evaluated the ROP predictions made on the real process parameters with regard to their sign, whether they are positive, meaning the predictions are more optimistic, or negative, meaning the predictions are more pessimistic. The following table shows the values of the calculated error metrics plus the numbers of instances with positive and negative errors, as well as the according mean error values.

Table 4: ROP prediction model errors evaluation

ROP Errors	MAE [m/h]	2.24
	RMSE [m/h]	3.11
	R2-Score	0.72
	Instances with positive error	4805
	Average positive error [m/h]	1.97
	Instances with negative error	3169
	Average negative error [m/h]	-2.66

The error metric values of the ROP prediction model lie at 2.24 m/h for the MAE, at 3.11 m/h for the RMSE, and at 0.72 for the R^2-Score. In comparison to that, on the validation dataset from the Geostar 2 data, on which the model had been trained, the MAE had a value of 1.52 m/h, the RMSE had a value of 2.01 m/h, and the R^2-Score has a value of 0.89. Comparing the performance metric values it is obvious that the metrics score achieved during the field testing are significantly worse than those achieved during validation on the development dataset. However, the metric values still lie in a acceptable range, given the fact that the training and inference data differ quite significantly. The following figure shows a plot of the actual ROP values recorded during the drilling operation in blue and the values predicted by the machine learning

Table 5: Statistical overview over the different ROP values

	Actual ROP	Predicted ROP	Optimised ROP
Average ROP [m/h]	13.31	13.44	14.43
Standard deviation [m/h]	5.74	4.3	4.44
Min. ROP [m/h]	1	2.27	2.54

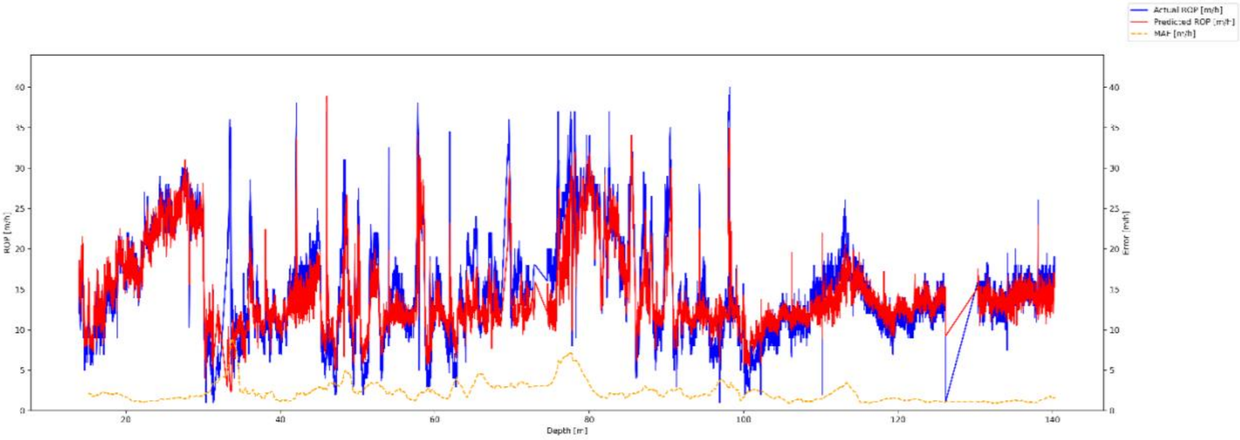

Figure 6: Actual ROP in blue and predicted ROP in red plotted over the measured depth. The MAE for a window of 100 data points is plotted in orange.

model in red. Alongside the two graphs the MAE values for a rolling window of 100 predictions is plotted in orange.

The second figure shows the same contents but for the section of the wellbore starting from 80 metres to 100 metres of measured depth. It serves the purpose of giving a more detailed view illustration of the ROP models predictions. This particular section was chosen due to the reason that it shows variation in the range of the predicted ROP and errors made, but any other section could have been chosen as well.
Both plots show that the ROP prediction model is prediction lower and higher ROP values without showing any overly significant tendency to one direction. Looking at the MAE plots it can be observed that there are some peaks with higher MAE values at around 35 and 78 metres depth. Looking at the remaining rows of table 3 it can be seen that the model does have a tendency of predicting the ROP more optimistically, with around 60% of the predictions being higher than the actual ROP value and 40% of the predictions being lower.

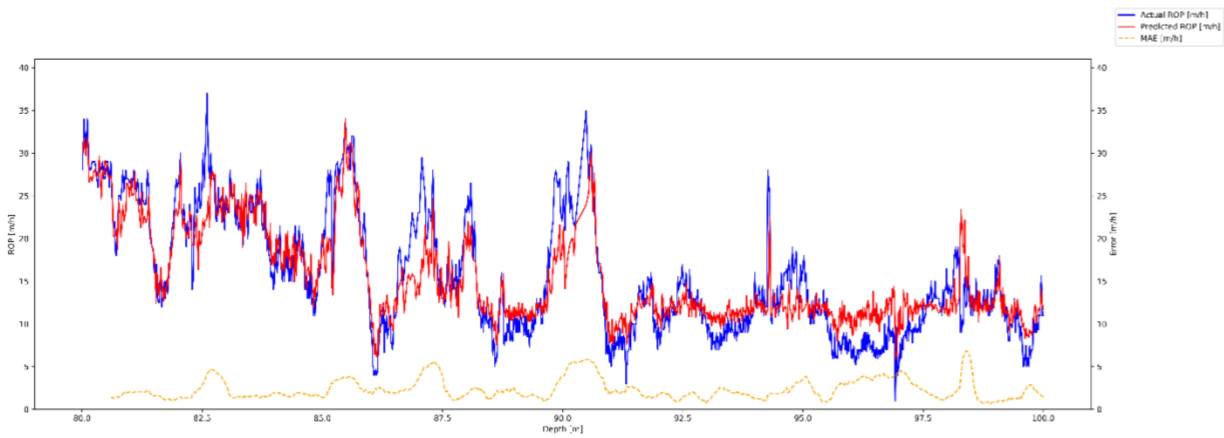

Figure 7: Actual ROP in blue and predicted ROP in red plotted over the measured depth for the section from 80m to 100m. The MAE for a window of 20 data points is plotted in orange.

The table above shows a statistical overview of the actual ROP values recorded during the drilling operation, the predicted ROP values based on the measured drilling process parameters, and the fictitious optimised ROP values based on the systems recommendations. Comparing the actual with the predicted ROP values the average ROP values are almost identical, while the standard deviation is lower for the prediction. Also, it looks like the model is struggling to predict very low and very high ROP values. This is most likely caused by the fact that these values are in general underrepresented in the development dataset. These observations could also be made during the model development and validation and were to be expected. The last column, showing the statistical properties of the optimised ROP values, shows that the average ROP value is higher than that of the actual ROP. Looking at the percentiles it can be seen that only the value for the first percentile of the optimised ROP is higher than that of the actual ROP. This could indicate that the highest potential for optimising the drilling operation, while also increasing the ROP value, lies in the lower half of the ROP range, which is plausible.

The figure above shows the actual ROP values in blue plotted over the measured depth alongside the optimised, expected ROP values in green. The plots show that the optimised ROP values are on average slightly higher than the actual ROP values.

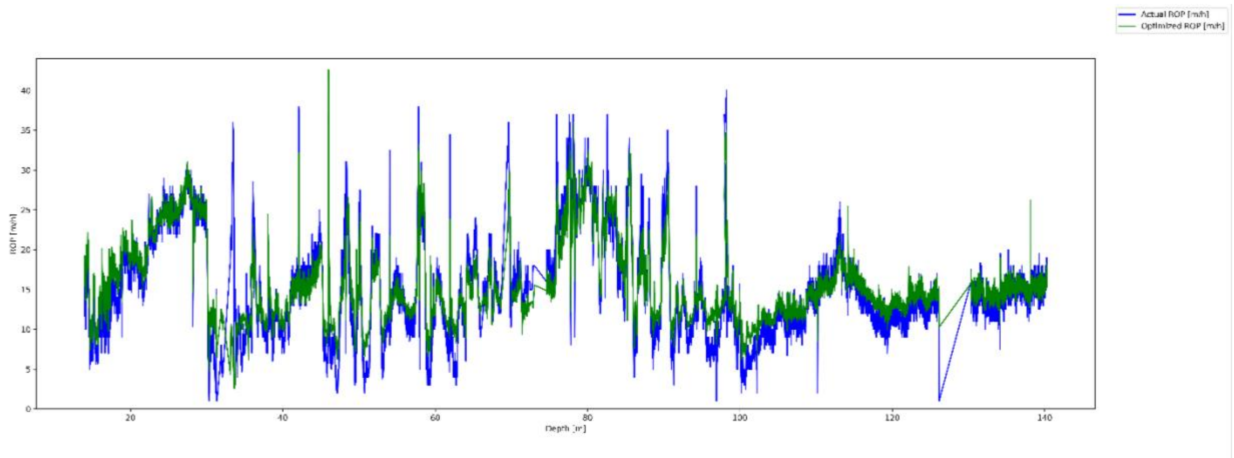

Figure 8: Actual ROP in blue and optimised (expected ROP) plotted in green over the measured depth

Table 7: Actual and optimised MSE values

	Actual MSE	Optimised MSE
Average MSE [N/mm2]	74.61	48.90
Standard deviation [N/mm2]	45.10	14.75
Min. MSE [N/mm2]	1.12	1.02
Max. MSE [N/mm2]	726.70	233.67

Table 6: ROP and MSE optimisation overview

ROP Optimisation	
Average ROP enhancement [m/h]	1.12
Average ROP enhancement [%]	18.88
Increased ROP values	5852
Decreased ROP values	2119
Constant ROP values	14

MSE Optimisation	
Average MSE enhancement [N/mm2]	25.73
Average MSE enhancement [%]	27.35
Decreased MSE values	7791
Increased MSE values	194
Constant MSE values	0

The tables above give an overview of the MSE and ROP values in context of the optimisation based on the controllable drilling process parameters. The first table compares the actual with the optimised MSE values. It shows that the average MSE values is significantly lower than the average actual MSE value. Furthermore, the standard deviation of the optimised MSE is also significantly lower than that of the actual MSE, which could indicate that the optimisation could lead to a more consistent and stable drilling process.
The other two tables give an overview of the impact of the optimisation on the ROP and MSE values. While based on the approach used for the optimisation, we expect the optimised MSE to be lower or at least equal to the actual MSE, we do not necessarily expect the optimised ROP to be higher than the actual ROP value. Looking at the MSE values in table 6 we can see that in almost all instances the optimised MSE value was decreased. Only in around 194 instances the optimised MSE value was higher than the actual MSE value. This occurs in instances when the outputs of the ROP prediction models are inaccurate and lead to the recommendation of a fictitious drilling process parameter scenario that is actually, based on the model's predictions, worse than the current process parameter settings. When we take a look at the table with the ROP optimisation overview, we can see that in the majority of the cases, in about 73%, the optimised ROP is increased, while in the remaining instances the ROP value is decreased except for a minority of only 14 cases in which it is not changed.

Table 8: Recommendations for the controllable drilling process parameters by the OptiDrill system

	Process Parameter		
Recommended Action	**RPM**	**WOB**	**MFI**
Increased	0	3596	3941
Decreased	7985	4091	3769
Constant	0	298	275

The table above gives an overview of the actions recommended by the OptiDrill system. You can see that for the RPM in all instances it was recommended to lower the current value, while for the other two controllable parameters it is more balanced between increasing and decreasing the current value.
It has to be noted at this point that the ROP prediction and optimization framework is completely based on the machine learning model's prediction accuracy. The error the model is making while predicting the ROP value based on the drilling process parameters has a direct impact on the calculated MSE values, which is used for picking the most favourable drilling process parameter scenario and selecting the recommendation for changing the controllable drilling process parameters. Apart from that, the limitations of the ROP optimisation framework described in the respective deliverable from work package number 7 also influence the MSE calculation. It must be considered that the calculated MSE values displayed here are expected to contain a certain error.

3.3 Drilled Lithology Prediction Module

The drilled lithology prediction module which was developed in work package number 8 and reported on in the respective deliverables consists of a machine learning based classification model that predicts the most likely drilled lithology from a number of learning lithology types based on a set of sequences of drilling process input parameters which is very similar to those used for the ROP prediction model. As already mentioned before in this report the drilled lithology prediction model was trained on the Geostar 2 dataset in which 4 different lithological classes are represented. These classes comprise *claystone*, *siltstone*, *sandstone*, and *claystone/siltstone*. Since the drill site of the Geostar 2 is not far from the location of the Push-It P1 wellbore, it was assumed that the same or very similar lithology classes would be encountered.

Table 9: Drilled lithology prediction model outputs from the first field test in Bochum

The following table shows the drilled lithology model outputs from the first round of field testing and their evaluation.

Lithology	Predicted Classes	Claystone	7965
		Sandstone	20
		Siltstone	0
		Claystone/Siltstone	0
	Actual Classes	Claystone	3949
		Claystone/Sandstone	3649
		Sandstone	387
	Predictions	True	3951
		False	4034

Looking at predicted classes in the table above it can be seen that from the four classes the model was trained on basically only one class, *claystone*, was predicted. Apart from that class only *sandstone* was predicted in 20 instances. The other two classes were not predicted at all. Comparing the predicted classes with the actual classes it also becomes clear that the classes the model was trained on do not match completely with the classes that were identified during the logging of the wellbore. The following figure shows the lithology log that was created by the geologists involved in the Push-It project. The lithology classes identified based on the logging and the cuttings analysis are shown in the lithology log on the right. The predominant class is *claystone*, which almost covers more than half of the depth of the wellbore with about 71 metres (from 12 to 140 metres measured depth). The second largest class is claystone/sandstone, which is a mix of both rock types. It covers about 52 metres of depth. The third class present is *sandstone*, however this class is quite underrepresented and covers only 5 metres of depth. Both classes *claystone* and *sandstone* are present in both the development and the inference datasets. The classes *siltstone* and *claystone/siltstone* are only present in the development dataset, while the class *claystone/sandstone* is only present in the inference dataset. This makes is difficult to apply the same metrics that were used during the model development. However, looking at the drilled lithology model predictions it becomes clear that almost only one class, which is *claystone*, was predicted. Apart from that *sandstone* was predicted in only 20 instances. This could quite likely be due to the significant difference in the input parameters ranges that was described in the first section of this chapter leading to the outputs of the lithology prediction model falling almost only into a single category.

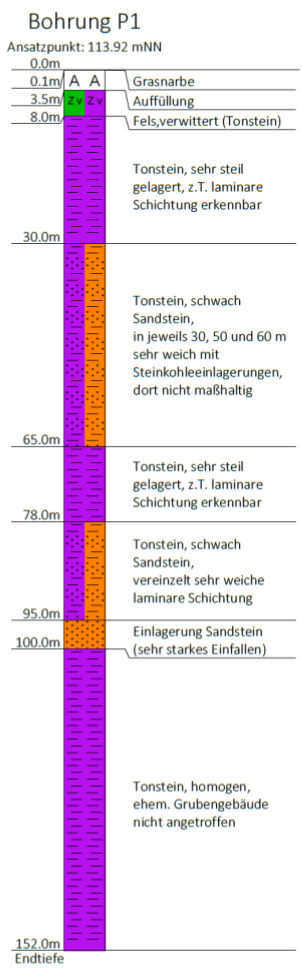

Figure 9: Lithology log Push-It P1 wellbore (created by Stefan Klein, CC Bergbaufolgenutzung, IEG)

The confusion matrix below shows the predictions that the drilled lithology model made during the field test. Since the classes the model was trained on and the classes that were actually determined differ and the model basically only predicted one class, the confusion matrix does only offer very limited new insights into the model's predictive performance. Even calculating the model's accuracy, which in this case would lie at around 50%, only gives very limited reliable information in this case.

3.4 Drilling Problem Detection Module

The drilling problem detection module that was developed within work package number 9 by the project partner BGS and reported on in the respective deliverables is based on an unsupervised learning approach. The isolation forest algorithm is used to detect anomalies in the drilling data that might indicate that an abnormal event is happing which might be a caused by a drilling problem. The model used for the anomaly

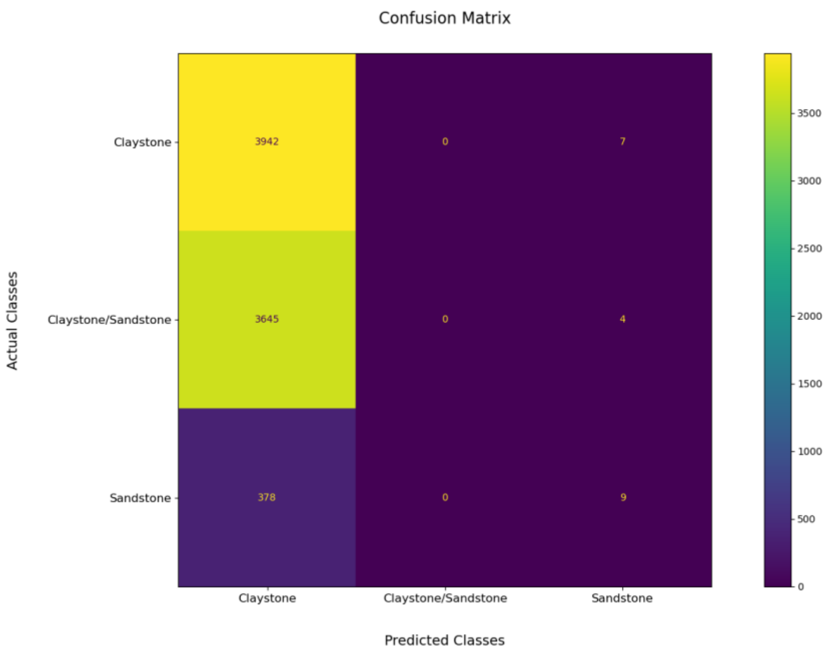

Figure 10: Confusion matrix of the model predictions for the Push-It P1 wellbore

detection is retrained during the drilling operation on the data that is being recorded and the same input data is used for making a prediction as for the other two machine learning based models. The model's output is a Boolean, *True* represents a detected anomaly and possible drilling problem, and *False* represents normal data that is not classified as anomalous. The following table gives an overview of the predictions of the drilling problem detection model.

Table 10: Anomaly detection model outputs

Anomaly Detection	Predictions	7941
	Anomalous	1685
	Not anomalous	6256

The table shows that out of a total of 7941 predictions in 1685 instances, accounting to around 20%, the model inputs were classified as anomalous. Based on the nature of the isolation forest algorithm is it to be expected that more data points are classified as anomalous as there are actual anomalies present in the data. The algorithm has a parameter that specifies the percentage of most anomalous data points that will be identified by the model. In the case of the first field test that was conducted no actual drilling problems were encountered, that could be correlated to the model's predictions. Within deliverable 13.4 the model outputs will be analysed more thoroughly.

4. OptiDrill Sensor Sub Tests

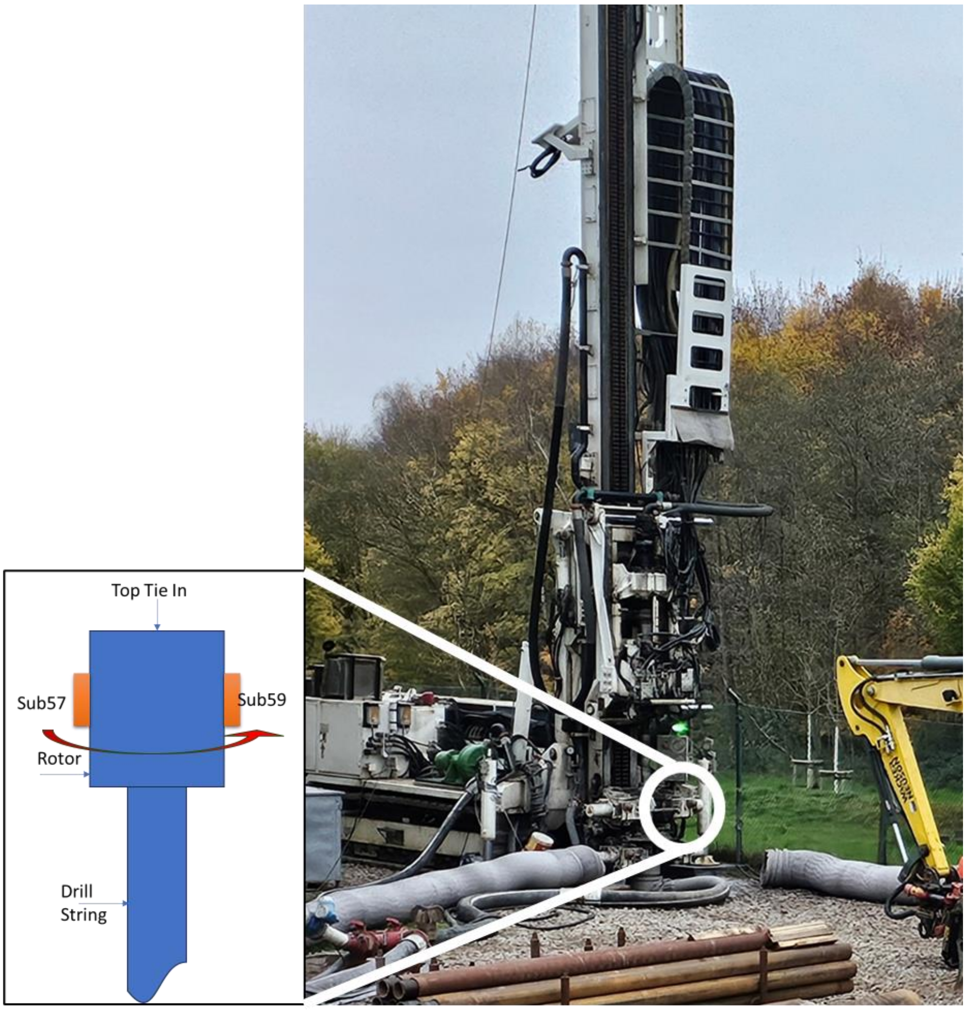

Figure 11: Setup of OptiDrill System, sensor subs attached to rotary drill head in Bochum. Two sensor subs, identified as sub 57 and sub 59 attached to the power swivel.

On the first day, the drilling operation commenced at 12:47. The team encountered challenges with time synchronization between the drill rig and the sensor data, necessitating manual alignment through photo documentation. Despite these challenges, sensor sub 57 successfully recorded data throughout the day. However, sensor sub 59 unexpectedly stopped functioning after 4 hours and 15 minutes, leading to incomplete data capture and the absence of high-G data.

The second day of the test presented further challenges with the sensor system. Both sensors experienced intermittent data logging and battery management issues. These problems required manual interventions to restart the sensors periodically. While sensor sub 57 was operational intermittently, sensor sub 59 ceased recording entirely after the first 30 minutes of the session. Unfortunately, neither sensor was able to capture high-G data on this day, which impacted the scope of the collected dataset.

On the third and final day, significant improvements were observed following OLV implementing a firmware and logging software update. The updated system addressed the earlier issues, resulting in both sensors operating reliably throughout the day. IMU data was successfully captured, and the sensors demonstrated their capability to function as intended. At approximately 12:15, the drill reached the concrete bunker, with full penetration achieved by 13:20. The success of this phase of the operation validated the improvements made to the sensor system and its resilience in field conditions. Additional drill rig data has been analysed to corroborate and refine these results.

While the field test highlighted areas requiring further development, such as time synchronization, high-G data logging, and power management, the final day validated the OptiDrill sensor system's capability to deliver reliable data for in a real-world drilling environment. The test(s) provided an essential foundation for enhancing sensor performance and operational integration in future projects, aligning with the overarching goals of the OptiDrill initiative.
With day 3 start time of 11:31:14 estimate the data packet number at 12:25 (drill reached the concrete bunker) and 13:20 (with full penetration achieved). See Figure 12

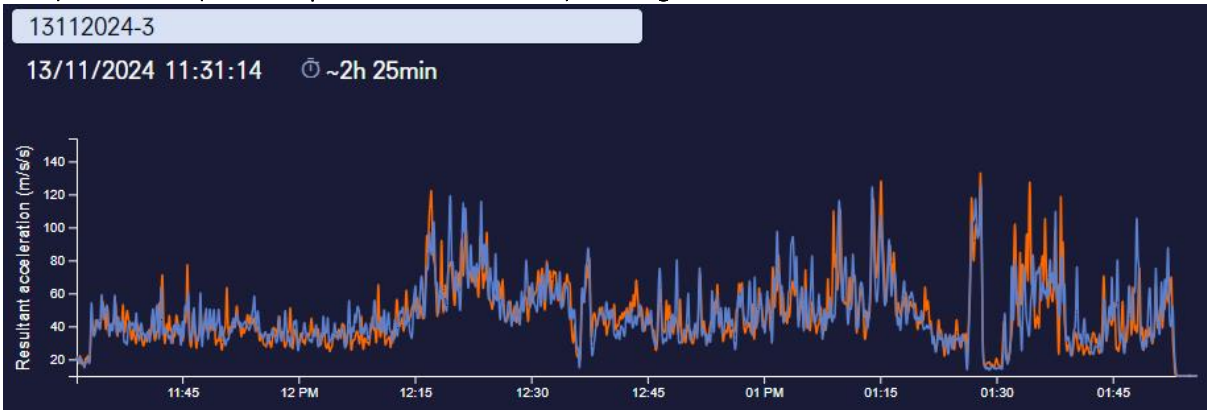

Figure 12: Onsight Data Review Using Mobile Phone Apps developed

4.1 Data Stream and Packetisation for Analysis

The data stream was divided into packets, each containing 16,384 data points for each of the 40 measured and evaluated parameters (see Figure 13). These parameters were calculated using the microprocessors within the sensor sub-architecture. This approach ensures that, when connectivity protocols are implemented in the future, only advisory data needs to be transmitted to the drill rig operator. The advisory data will include key parameters such as:

1. Tangential Forces
2. Axial Force
3. Rate of Penetration (ROP)
4. Rotational Speed
5. Specific Rock Energy (SRE)
6. System Identification Metrics (Gain, Phase, and Coherency Functions)

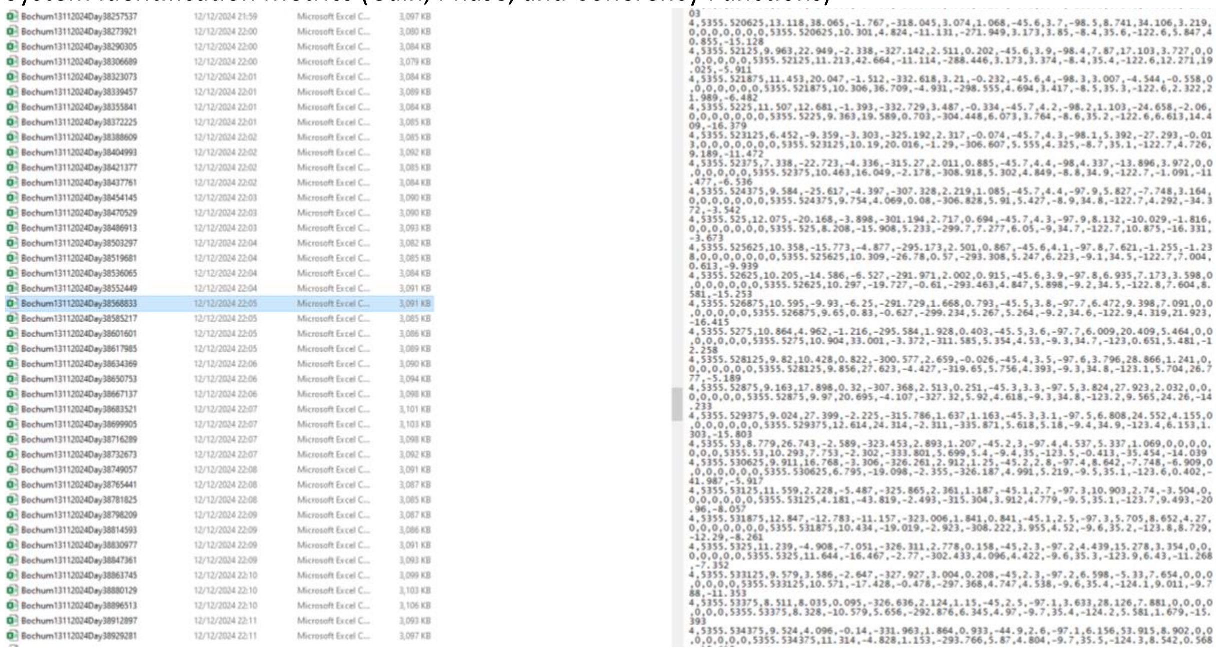

Figure 13: OptiDrill Data Packet Structure for Sensor Subs.

The ultimate goal is to use these parameters to provide real-time recommendations to the drill rig operator, enabling adjustments at the top tie-in to maximize ROP. This novel technique allows low-latency advisory data to be viewed in real time while logging all raw data in packet form for more detailed analysis after the drilling operation is complete.

4.2 Data Arrangement and System Identification

For analysis and visualization, the logged data packets are structured into 3D matrices as follows:

1. SRE Data:
   - Sensor sub 57: 849×16,384849 \times 16,384849×16,384 (packets × data points)
   - Sensor sub 59: 849×16,384849 \times 16,384849×16,384 (packets × data points)
2. System Identification:
   - Gain, Phase, and Coherency functions are derived by using the SRE data from sensor sub 57 as input and the SRE data from sensor sub 59 as output.
   - To compute these metrics, the 16,384 data points in each packet are divided into 32 non-overlapping segments of 512 data points. This segmentation ensures smoothing and improves the reliability of the derived transfer functions.
3. Output Metrics:
   - Gain_SPE: 849×256849 \times 256849×256 (packets × frequency points)
   - Phase_SPE: 849×256849 \times 256849×256 (packets × frequency points)
   - Coherency_SPE: 849×256849 \times 256849×256 (packets × frequency points)

4.3 Event Identification Using Difference Equations

To detect significant changes in the drilling process, difference equations are applied to successive packets. This technique is used to identify anomalies or spikes in the data, which may correspond to lithological transitions or key drilling events. Two specific events from the field test are analyzed:

- 12:25 (Bunker Reached): Corresponds to **Packet Number 315**
- 13:20 (Full Penetration Achieved): Corresponds to **Packet Number 653**

These specific packets, along with their surrounding data, are analyzed in detail to evaluate variations in Gain, Phase, and Coherency metrics. This analysis provides valuable insights into changes in rock properties and drilling performance. By leveraging system identification techniques in conjunction with high-resolution sensor data, the approach enhances operational decision-making and supports comprehensive post-drill analysis.

However, due to the placement of the sensors equidistant from the Bottom Hole Assembly (BHA), the system's capability is limited to identifying conditions where the Gain is low (e.g., less than 5 dB) at low frequencies. These conditions typically correspond to scenarios of low Rotations Per Minute (RPM) and/or low Weight on Bit (WOB). This limitation emphasizes the need to interpret the results in the context of the specific sensor configuration and operational parameters.

4.4 Data Review

Figure 14 provides a detailed visualisation of SRE data measured using two sensor subs (Sub57 and Sub59), along with associated parameters during drilling tests conducted on November 13, 2024, starting at 11:31:14.

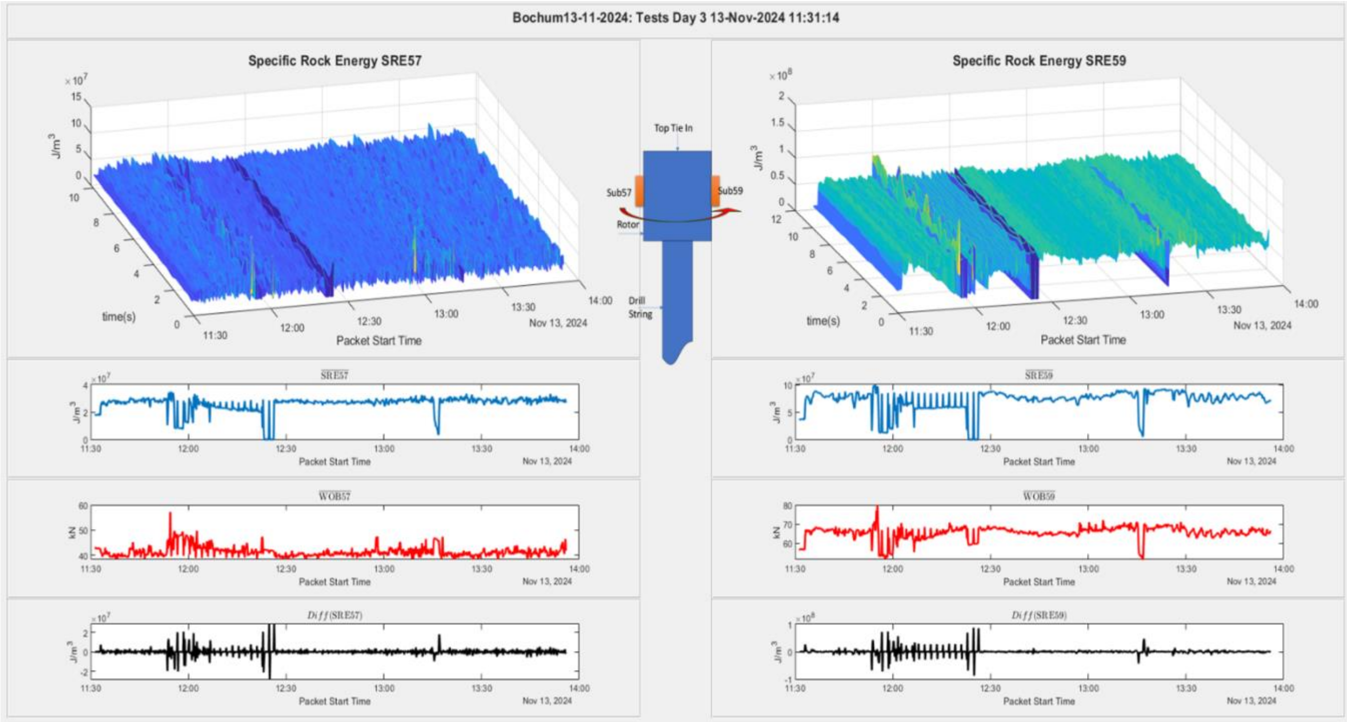

Figure 14: SRE data evaluation

4.5 Observations of SRE from Figure 14

1. Top Left (3D Surface Plot: Specific Rock Energy SRE57):
   - The 3D plot on the left shows the Specific Rock Energy (SRE) measured by Sensor Sub57 over time.
   - The horizontal axes represent packet start time and time (seconds), while the vertical axis represents Specific Rock Energy (J/m³).
   - There is a clear variation in SRE values during the drilling process, with noticeable trends and fluctuations that may correspond to changes in drilling conditions or lithology.
2. Top Right (3D Surface Plot: Specific Rock Energy SRE59):
   - The 3D plot on the right displays the SRE measured by Sensor Sub59 over the same time period.
   - Similar to SRE57, the horizontal axes represent packet start time and time (seconds), with the vertical axis showing SRE values.
   - The differences in energy profiles between Sub57 and Sub59 indicate potential variations in dynamic forces, rotational effects, or system identification outputs.
3. Middle Section (Time-Series Data for SRE57 and SRE59):
   - The middle row contains two line plots, showing the time-series evolution of SRE for Sub57 (left) and Sub59 (right) as a function of packet start time.
   - Both plots capture fluctuations in SRE values over the drilling period, highlighting periods of stability, spikes, and drops, which may indicate changes in rock properties or drilling dynamics.
4. Second Row (Weight on Bit - WOB):
   - The second row plots Weight on Bit (WOB) data in red, synchronized with the packet start time.
   - The WOB fluctuations appear to align with the variations in SRE, suggesting a correlation between axial loading conditions and rock energy measurements.
5. Third Row (Derivative of SRE):
   - The third row shows the derivative of SRE (dSRE/dt) for both Sub57 and Sub59.

- The sharp spikes and changes in the derivative reflect transitions or disturbances in the drilling process, potentially corresponding to lithological boundaries or equipment-related dynamics.

In summary both subs show similar likeness as expected and potential 12:25 (Bunker Reached) in the viscinity of Packet Number 315 and 13:20 (Full Penetration Achieved) in the viscinity of Packet Number 653 are clearly observed.
Individual packet observations are shown

4.6 Evaluation of Gain, Coherency, Angular Velocity (ω), and Specific Rock Energy (SRE) Analysis.

The analysis examines Gain (dB), Coherency, Angular Velocity (ω, in rad/s), and Specific Rock Energy (SRE) across transitional and non-transitional regions. Observations are derived from Sensor Subsystems 57 and 59, with references to packet numbers and corresponding figures.
At 12:15, during bunker reach, packet number 310 (Figure 15), packet number 315 (Figure 16), and packet number 330 (Figure 17) capture sensor responses. Figure 15 (packet 310) reflects the pre-transition state, where Gain and Coherency are stable, and SRE remains consistent, indicating minimal disturbance in the system. Angular Velocity (ω) at this stage is steady, suggesting rotational dynamics are not significantly impacting the observed signals. However, in Figure 16 (packet 315), Gain exhibits notable fluctuations, Coherency decreases, and SRE rises, pointing to increased energy demands as the system interacts with more resistant rock formations during the transition. These changes are accompanied by minor variability in ω, potentially amplifying signal disturbances. Figure 17 (packet 330) demonstrates post-transition stabilization, with Gain and Coherency normalizing as SRE decreases, indicating a return to more manageable energy conditions. Angular Velocity also steadies, reflecting improved operational stability.
At 13:20, packet number 620 (Figure 18), packet number 637 (Figure 19), and packet number 657 (Figure 20) highlight system behavior. Figure 18 (packet 620) shows pre-penetration conditions, where Gain and Coherency begin to vary slightly, and SRE starts to increase as rock resistance intensifies. Angular Velocity is stable, playing a secondary role in signal behavior. In Figure 19 (packet 637), representing the critical penetration phase, Gain shows significant variability, Coherency decreases sharply, and SRE peaks, reflecting the maximum energy demand needed to penetrate the formation. During this phase, any fluctuations in ω exacerbate instability by influencing how signals interact with the high SRE conditions. In Figure 20 (packet 657), post-penetration responses show stabilized Gain, improved Coherency, and reduced SRE as the system acclimates to the fully penetrated environment. Angular Velocity returns to a steady state, supporting uniform sensor performance.
The relationship between Angular Velocity, Gain, and SRE is most evident in transitional phases, such as those represented by packets 315 and 637. Variability in ω during these phases magnifies the effects of increased SRE, leading to pronounced fluctuations in Gain and reduced Coherency. This dynamic interplay highlights the challenges of maintaining signal integrity under high-energy-demand conditions. Conversely, in non-transitional areas such as packets 310, 330, 620, and 657, a steady ω correlates with stable SRE, consistent Gain, and higher Coherency, reflecting a more controlled and predictable environment.
In conclusion, Angular Velocity plays an amplifying role in transitional zones by intensifying the effects of rising SRE on Gain and Coherency. In non-transitional areas, steady ω and lower SRE contribute to system stability, underscoring the interconnected nature of these variables in both dynamic and steady-state operations. This analysis provides critical insights into the influence of SRE and rotational dynamics on sensor performance.

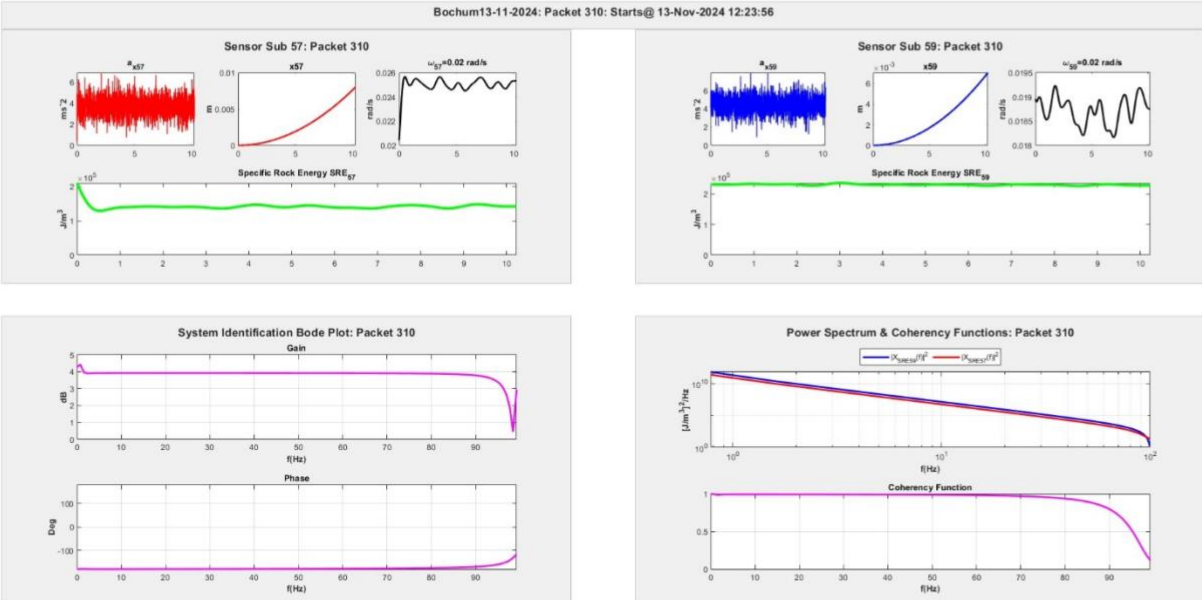

Figure 15: Packet number 310 for Sensor sub 57 and 59 along with their analysed Gain, Phase and Coherency function.

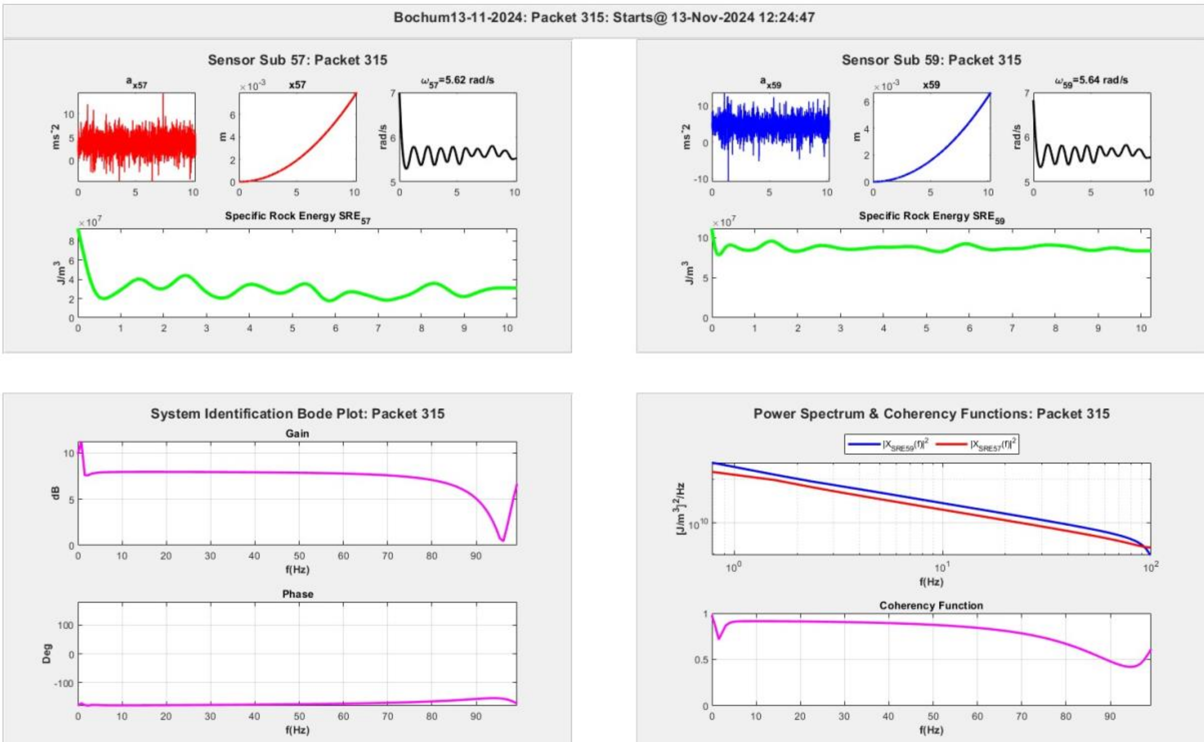

Figure 16: Packet number 315 for Sensor sub 57 and 59 along with their analysed Gain, Phase and Coherency function.

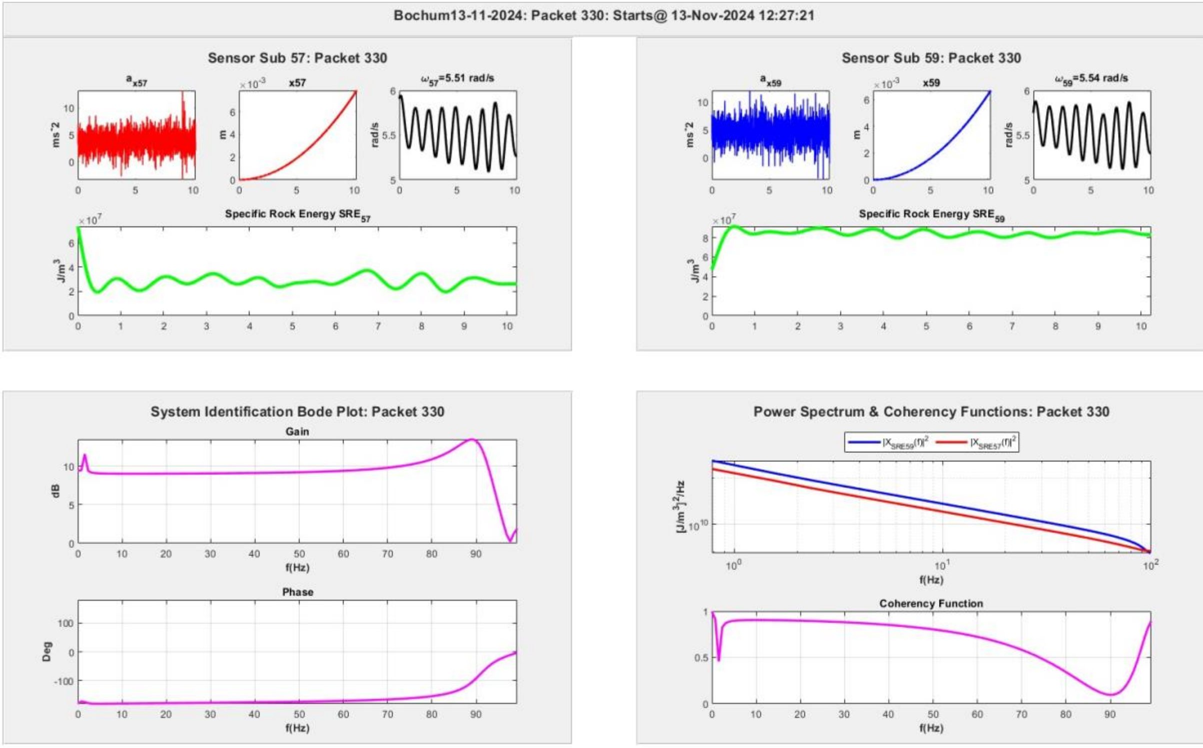

Figure 17: Packet number 330 for Sensor sub 57 and 59 along with their analysed Gain, Phase and Coherency function.

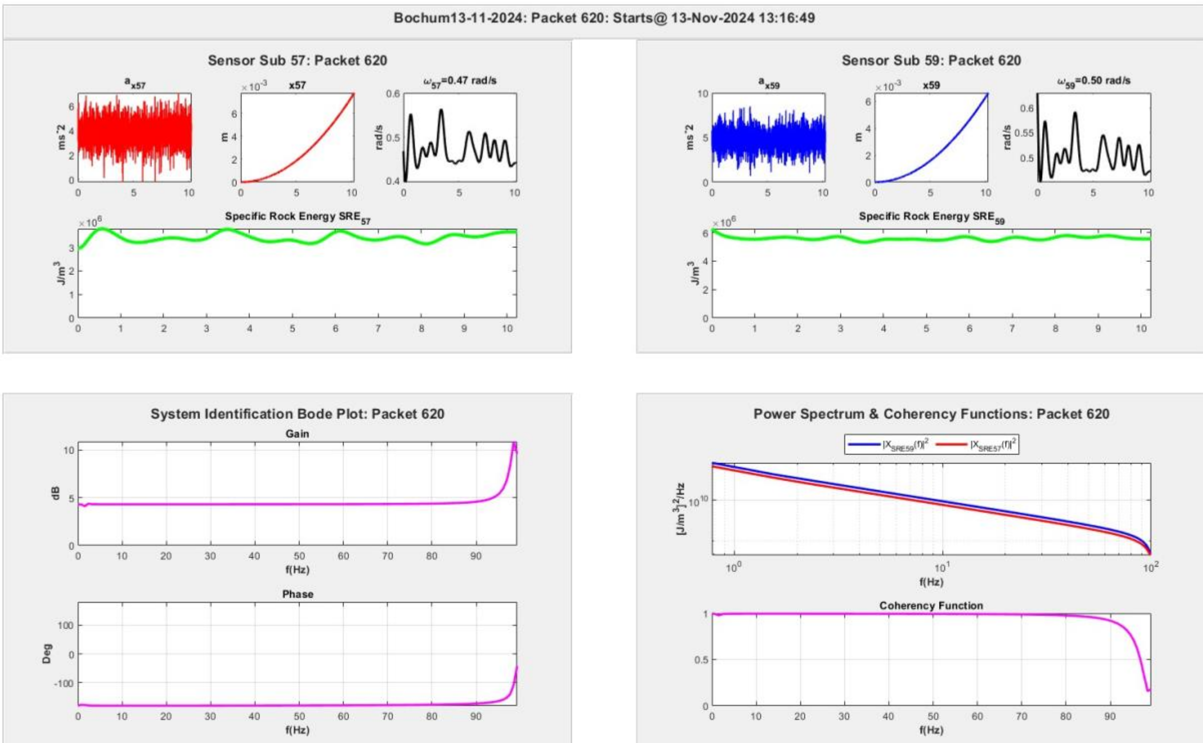

Figure 18: Packet number 620 for Sensor sub 57 and 59 along with their analysed Gain, Phase and Coherency function.

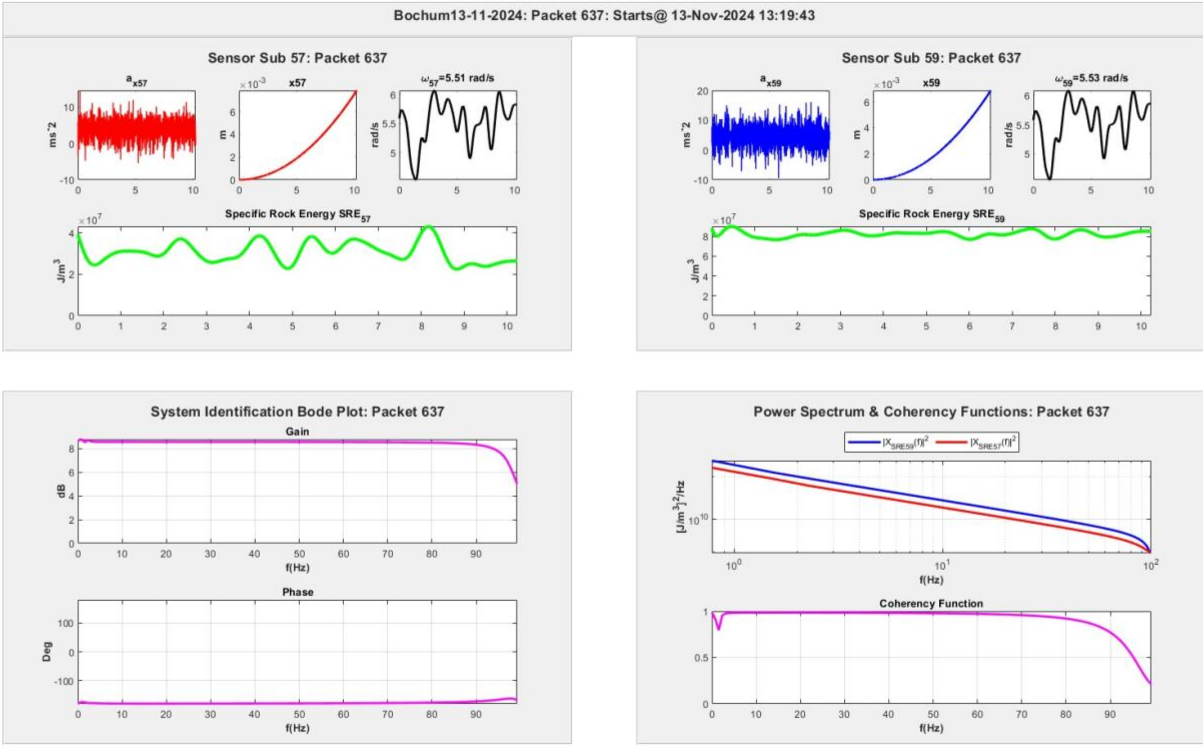

Figure 19: Packet number 637 for Sensor sub 57 and 59 along with their analysed Gain, Phase and Coherency function.

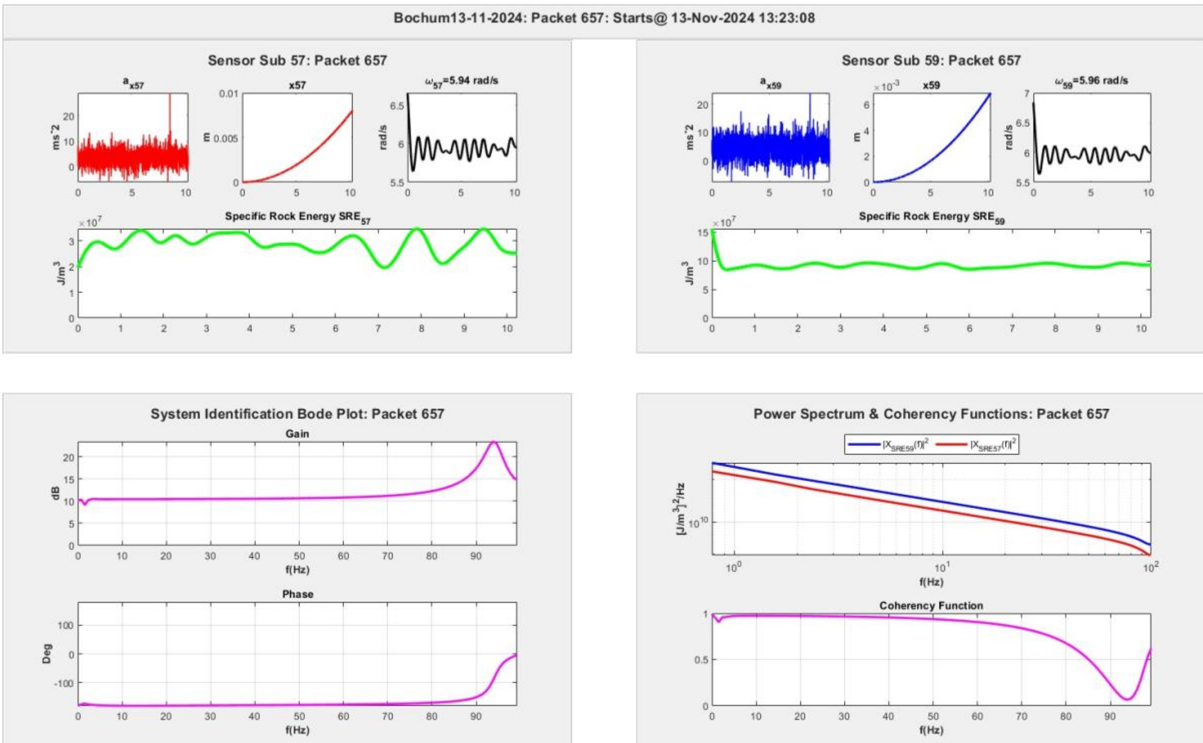

Figure 20: Packet number 657 for Sensor sub 57 and 59 along with their analysed Gain, Phase and Coherency function.

4.7 Summary of Differences in Measurement and Calculation

Force (F) measurements at the drill bit interface accurately represent tangential forces acting directly on the rock, whereas top tie-in sensors capture forces inflated by upstream mechanical losses. Similarly, rotational speed (ω) at the drill bit reflects true cutting dynamics, while upstream measurements are skewed by torsional oscillations and stick-slip effects. The rate of penetration (R), directly measurable at the bit, is often delayed or inaccurately inferred from top tie-in data. Consequently, SRE accuracy is higher at the drill bit interface but overestimated when derived from top tie-in data due to compounded mechanical losses.

4.8 Implementation of the Sensor System: A Justified Deviation

The OptiDrill sensor system was intentionally developed as a standalone, battery-operated module rather than being directly integrated into the OptiDrill software for real-time advisory data evaluation. This decision was guided by a critical assessment of the prohibitive costs associated with implementing real-time connectivity for downhole sensors at the Bottom Hole Assembly (BHA). While the feasibility of such connectivity had already been proven in the H2020 GeoDrill project, it was deemed financially and logistically unfeasible for the current project. Instead, OptiDrill opted for a battery-powered system, as successfully demonstrated in Deliverables D4.1 and D4.2, which enabled precise, modular, and cost-effective data acquisition.

The implemented system captures high-resolution data from accelerometer, gyroscope, and magnetometer sensors (Sub57 and Sub59), positioned a fixed distance apart along the drill string. This configuration allows the independent recording of critical parameters such as accelerations, angular velocities, and magnetic field variations during drilling operations. By avoiding real-time data transfer, the focus shifted toward ensuring data accuracy and reliability in standalone conditions while maintaining operational flexibility across various environments.

4.8.1 Justification for Deviation from the Original Plan

Originally, the OptiDrill sensor system was envisioned to be tightly integrated into the software framework, providing real-time data for decision-making. However, during the project's development phase, it became evident that the high cost of implementing real-time communication infrastructure—especially for sensors located at the BHA—posed a significant barrier. Additionally, the computational burden associated with processing real-time sensor data would have required a more complex and expensive software architecture.

The decision to transition to a standalone, battery-operated system was therefore both pragmatic and strategic, offering several key advantages:

1. **Cost Efficiency**: The battery-powered design eliminated the need for expensive real-time communication systems while ensuring high-quality data acquisition. This allowed the project to focus resources on refining the precision and reliability of the sensor hardware.
2. **Modularity and Flexibility**: Decoupling the sensor hardware from the OptiDrill software framework resulted in a modular system adaptable to a variety of drill string configurations and operational environments.
3. **Thorough Experimental Testing**: As demonstrated in D4.1 and D4.2, the standalone design facilitated controlled experimental drilling tests. The sensor data was stored locally for later analysis, enabling high-resolution integration of SENSOR outputs with surface equipment data, including torque, Weight on Bit (WOB), and Rotations Per Minute (RPM).

4.8.2 Benefits and Future Applications

The OptiDrill sensor system, as implemented, serves as an independent add-on tool for drilling operations, providing modular, precise, and cost-effective data acquisition. Although it does not offer real-time advisory capabilities, the system's design ensures that its data can complement OptiDrill software workflows. This modularity circumvents the financial and logistical barriers associated with fully integrated real-time systems while maintaining the ability to deliver high-quality insights for operational decision-making.

5. UK Field Scale Testing

The tests undertaken in the UK, during January 2025, were carried out by Geolorn and Oliveris for a water well drilling operation in South Wales, in conjunction with Apex Drilling Services, to whom the project are very grateful

This site was chosen due to the fact that a large number of offset wells have been drilled, so whilst we lacked wireline geophysical data, we had a great deal of analogue data, on which to validate the drilling generated data, recorded by the sensor subs.

The drill rig was a Fraste XL, with load sensing hydraulic system, and a Compair Compressor delivering compressed air at 30m^3/minute, at an output pressure of 25Bar. With frictional losses the standpipe pressure was 23Bar.

The drilling operation used a high pressure air hammer system, with the wellbore flushing being achieved with the air exhausted from the hammer. The drilling rig itself, is designed with a load-sensing hydraulic system, which allows drilling parameters such as Weight on Bit and Torque to be accurately set and maintained throughout the drilling operation.

A CRI 6” (nominal size) hammer was used with a 165mm bit, using semi-ballistic inserts. The hammer consumes 27.5m^3/min of air when operating at 23Bar and the 18Kg piston cycles at 1,738 strokes/min, travelling a distance of 102mm /stroke. The exhaust air flushes the wellbore clean.

Down the Hole Hammers (DTH), require a much lower Weight on Bit (WOB) than conventional rotary drilling methods and, therefore, much lower torque inputs. Rotation speeds are also lower. The rates of penetration (ROP) achieved with hammers is generally much higher than rotary drilling, but the inefficiencies of compressing air (input energy for compressors) needs to be balanced against the operating time, to get a clearer picture of overall costs.

High velocity air (in this case ~ 2,500m/min) in the annulus can also result in destabilising the formations, although there is limited blocking of fractures and pores, that are often associated with rotary drilling that utilise fluids and viscosifying agents (e.g. bentonite or polymers), which can result in the reduction of permeable flow into a well, which is a key requisite for geothermal applications.

DTH Hammers rely on impact energy being transferred to the bit from the piston cycling within the hammer sleeve. This impact can cause high vibration, although the effects are localised to the hammer and near drill string, which in the field trials included the sensor subs.

Figure 21: Preparing the data sensor subs to run in hole

Figure 22: Activation of the memory recording system by BLUETOOTH

Figure 23: Data sensor subs installed into BHA

The well was cased through the drift deposits and weathered rock to a depth of ~8.0m, sealing into competent rock. The data sensor drill string assembly was picked-up and ran in hole. With drilling commencing at 8.0m below ground level.
The three controllable drilling parameters – WOB, Torque & RPM, were set by the driller and the load sensing hydraulic system maintains the set-values until changed. This approach is particularly suitable for DTH Hammer drilling, especially in predominately homogenous formations, such as the mudstones encountered.
Each drill pipe section was 3.0m and all three parameters were recorded along with the time taken to drill each pipe and the air pressure in the standpipe. The details are shown in Figure 24

OptiDrill Sensor Subs Field Testing - 15/01/2025

Start Depth (BGL)	Finish Depth	WOB (Kgs)	Torque (KgM)	RPM	Time (Mins:Secs) (Time per 3.0m Pipe)	ROP (m/hr)	Formation
8.00m	11.00m	1,485	231	38	04:10	43.2	
11.00m	14.00m	1,485	231	38	04:07	43.7	
14.00m	17.00m	1,485	231	38	04:15	42.4	Mudstone
17.00m	20.00m	1,485	231	38	04:15	42.4	
20.00m	23.00m	1,485	231	38	03:45	48.0	
23.00m	26.00m	1,485	231	38	03:36	50.0	
26.00m	29.00m	1,485	231	38	05:11	34.7	
29.00m	32.00m	1,485	231	38	04:21	41.4	
32.00m	35.00m	1,485	231	38	03:18	54.5	
35.00m	38.00m	1,485	231	38	03:19	54.5	
38.00m	41.00m	1,485	231	38	04:36	39.1	Sandstone with mudstone bands. Some fracture zones and high water inflow.
41.00m	44.00m	1,485	231	38	04:15	42.4	
44.00m	47.00m	1,485	231	38	04:29	40.1	
47.00m	50.00m	1,485	231	38	04:41	38.4	
53.00m	56.00m	1,485	231	38	04:15	42.4	
56.00m	59.00m	1,485	231	38	04:58	36.2	
59.00m	62.00m	1,485	231	38	03:37	49.8	
62.00m	65.00m	1,485	231	38	04:29	40.1	
65.00m	68.00m	1,485	231	38	05:16	34.2	
END OF HOLE							
				Totals	1h:20m:33s	Average: ~46.15m/hr	

Hammer Details	
Bit Size	165mm
DTH Hammer	6" Nominal
Compressor Air Volume	30m3/min
Air Pressure (output)	25Bar
Standpipe Pressure	23Bar
Piston Weight	18.00Kgs
Piston Stroke	102mm
Blows/minute	1,738
Air Consumption	27.5m3/minute

Figure 24: Details of drilling parameters recorded – the highlighted times show when fracture zones were encountered.

The well was drilled to a total depth of 68.00m below ground level, where drilling was terminated, as the water inflow was sufficient to maintain the design output of the well – 1.5litres/second. Upon completion of the drilling, the drill string was withdrawn and the sensor subs were broken down from the BHA and the memory sticks recovered. The data recorded was taken away to download and analyse.
In addition to the downhole conditions, the sensor subs also had to cope with the Iron Roughneck (Figure 25)

5.1 Initial Test Outcome

During these trials, the sensor units were securely housed within specially fabricated subs, positioned in the Bottom Hole Assembly (BHA) directly above a Down-The-Hole (DTH) hammer, which was operated using high-pressure air. Unlike previous tests, which primarily focused on conventional drilling, this setup introduced significantly higher impact loads.

Approximately one-third into the testing sequence, IMU59 ceased functioning. A subsequent teardown investigation revealed that the failure was due to a short circuit caused by internal battery damage. Specifically, the fluid electrolyte within the lithium battery allowed the anode and cathode to come into contact under severe vibration, resulting in a short. Upon replacing the battery, IMU59 resumed normal operation.

Figure 25: Iron Roughneck and Heavy Duty Pipe Slips. The gripping and rotating action to make-up and break-out tool joints can be very harsh on components.

To prevent recurrence, a design change is planned whereby future iterations of the IMU will utilize solid-state electrolyte batteries, which are more resilient to high-impact environments. For the analysis presented here, only data from IMU57 was used. See Figure 26.

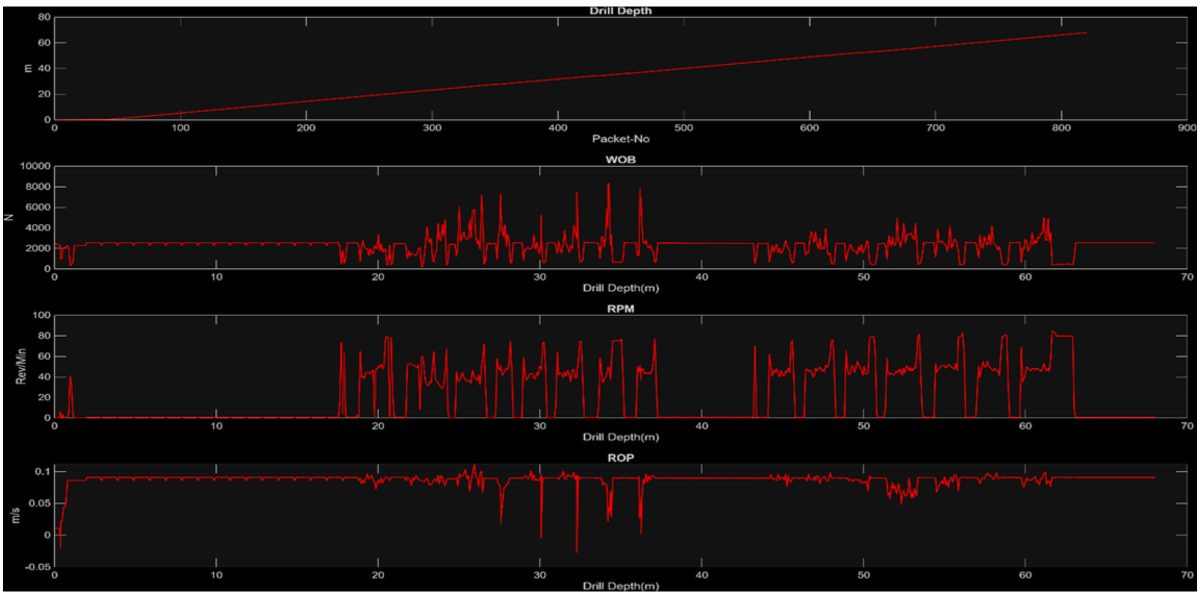

Figure 26: Evaluated with Sensor IMU57. From top to Bottom : Drill Depth .vs. Packet-No, Estimated WOB at Drill Depth, Estimated RPM at defined Drill Depth, Estimated ROP at defined Drill Depth.

Although IMU59 ceased functioning approximately one-third into the testing sequence, the data collected from IMU57 remains a reliable and representative source for continued analysis. As illustrated in the field setup, IMU57 was positioned directly above IMU59 within the drill string. While the two IMUs were not placed symmetrically, they were nonetheless aligned within the same structural axis of the BHA and experienced the same operational drilling cycle, including exposure to axial and torsional vibrations, bit-rock interactions, and variations in lithological strata. The vertical placement of IMU57, being slightly higher in the BHA stack, does not compromise its ability to capture the critical dynamic behaviours of the system. Given the mechanical continuity of the assembly, transmitted forces and vibrations propagate uniformly along the drill string. Thus, the variations in Gain and SRE values key features extracted for system identification and lithology classification remain consistent and meaningful even when derived from a single sensor sub.

Importantly, the data analysis method focuses on time- and frequency-domain trends (e.g., ΔGain and ΔSRE) derived from packet-to-packet variations. These are inherently robust against absolute positional differences, especially when the sensors are part of the same mechanically coupled system. As such, IMU57's data effectively captures the evolving system behaviour and is well suited for continued AI training, classification, and performance assessment. The use of this single-sensor approach is also aligned with real-world deployment scenarios, where redundancy may not always be available, and reinforces the robustness of the developed predictive framework.

5.2 Introduction to Blind Lithology Prediction.

In Work Package 4, specifically Tasks 4.1 and 4.2, OLV developed a novel framework for rock lithology classification by harnessing advanced deep learning methods, with a focus on convolutional neural networks (CNNs) such as ResNet, VGG, and GoogleNet. This approach revolved around the transformation of complex geophysical signals including Gain, Coherence, Phase, and Specific Rock Energy (SRE) difference into structured RGB image representations, enabling the use of pretrained image recognition networks for rock-type discrimination.

To further strengthen the predictive capability and generalizability of the model, OLV introduced principles from the field of System Identification (SID). In this paradigm, both frequency-domain features (e.g., Bode-based Gain and Phase) and time-domain features (e.g., ΔSRE) are traditionally used to characterize the dynamic behaviour of physical systems. By repurposing these signals within a CNN-based classification context, we established a cohesive and interpretable framework that treats rock prediction not simply as pattern recognition, but as a system-level modelling problem rooted in geophysical principles.

This hybrid approach—merging system identification theory with data-driven learning enabled us to extract deeper insights into rock formation Behavior and interactions during drilling, while simultaneously enhancing classification accuracy. It also facilitates operational flexibility: once trained using either laboratory-derived or field-acquired drilling datasets, the model can be applied in situ to predict lithology in real-time, contingent on the availability of annotated reference data per rock type.

Our longer-term goal is to assemble a comprehensive training dataset encompassing diverse lithologies for which critical mechanical properties are known specifically Uniaxial Compressive Strength (UCS), Hardness, Bulk Density, and Rock Quality Designation (RQD). By summarizing these properties for each rock category and correlating them with ΔGain and ΔSRE (derived from sequential time-frequency data packets captured via a single IMU sensor), we propose that a single CNN architecture (such as the modified GoogLeNet model developed in Task 4.2) can be trained to recognize one rock type at a time.

This enables implementation of a "one-vs-rest" classification strategy also referred to as class-specific probability modelling in which the model is trained to detect one lithology against all others. When deployed across the full dataset, the trained network can then be evaluated to determine how uniquely and confidently it identifies the target rock class amidst others. This framework not only improves classification performance but also provides a scalable approach for adaptive drilling intelligence in geothermal and other subsurface exploration domains.

5.2.1 Comparing Training with Successive Gain Differences vs. Successive SRE Differences

In our scenarios when one of the IMUs becomes non-functional, ensuring the robustness of our evaluation is crucial. To investigate whether comparable training performance can be achieved using data from a single IMU in this case IMU57, we analyse the impact of training the model separately with Gain Difference (ΔGain) and Successive SRE Differences (ΔSRE). Again here we limit our analysis to the two best classified rock types; Mudstone and Sandstonebut only evaluate the probability scores for Mudstone.

Figure 27 presents a comparative visualisation of ΔGain and ΔSRE signals from IMU57 recorded at three different drilling packet intervals (Packet 104, 667, and 741), during field operations in Wales. Each row corresponds to a distinct time segment in the drilling process, showcasing both the frequency-domain signal (ΔGain, in red) on the left and the time-domain signal (ΔSRE, in blue) on the right. At Packet 104, the ΔGain plot shows a nearly flat frequency response across the spectrum, suggesting a relatively uniform or low-energy interaction between the bit and the formation, which is consistent with the smooth and low-variation ΔSRE trace, indicating stable drilling and minimal rock disruption.

In contrast, Packet 667 exhibits a highly dynamic ΔGain profile with strong low-frequency energy that decays across the spectrum, indicating high energy transmission and resonance at lower frequencies. This correlates with a corresponding ΔSRE trace showing large amplitude variations and increasing energy over time, implying aggressive bit-rock interaction or a shift to a harder or more fractured rock layer. At Packet 741, the ΔGain spectrum intensifies, displaying high-frequency content and sustained oscillations, which reflects substantial vibrational input likely due to tough drilling conditions. The corresponding ΔSRE signal reaches its highest magnitude of energy fluctuations, with significant amplitude and noise throughout the packet duration. This suggests intense mechanical interaction and possibly the presence of high-resistance rock features or a transition zone. Overall, this figure effectively highlights the sensitivity of both ΔGain and ΔSRE as indicators of changing subsurface lithology and drilling dynamics. The complementary patterns observed in the frequency and time domains demonstrate how drilling response can be monitored in real-time, offering a valuable diagnostic tool for rock classification and system identification.

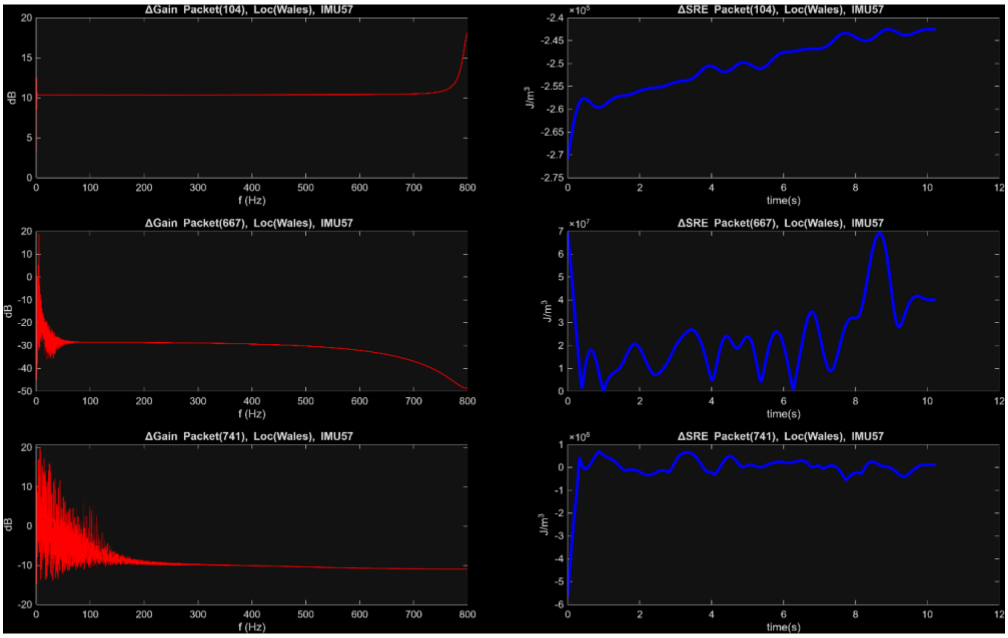

Figure 27: Examples of ΔGain and ΔSRE at Packet-No 104, 667 and 741.

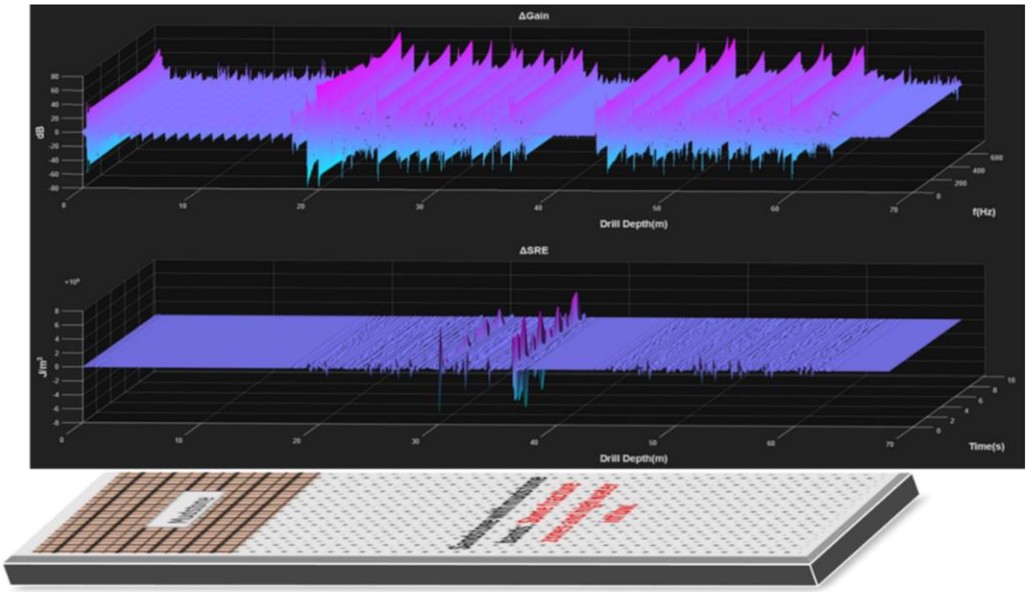

Figure 28: ΔGain (Top), ΔSRE(Middle) Matrices and Schematic of known Rock Lithology (Bottom)

Figure 28 above presents a comprehensive visualization of how ΔGain (differences in signal gain across frequency) and ΔSRE (differences in specific rock energy over time) can be used to detect lithological transitions during drilling. The top 3D plot illustrates ΔGain as a function of drill depth and frequency, where distinct shifts in gain values indicate changes in subsurface rock characteristics. The middle plot shows ΔSRE across drill depth and time, highlighting fluctuations in energy required for penetration, which also signal changes in rock type. These data-driven metrics, drawn from IMU sensor readings, capture both frequency and time-domain features, providing a dual-perspective analysis of drilling behavior. The bottom schematic links these observations to known lithological zones—Mudstone, Sandstone, and Limestone emphasizing the correlation between observed signal behavior and geological transitions. Together, the figure supports the concept that by monitoring ΔGain and ΔSRE, it is possible to identify rock formations in real time, enabling intelligent classification through machine learning models such as those employed in the OptiDrill project.

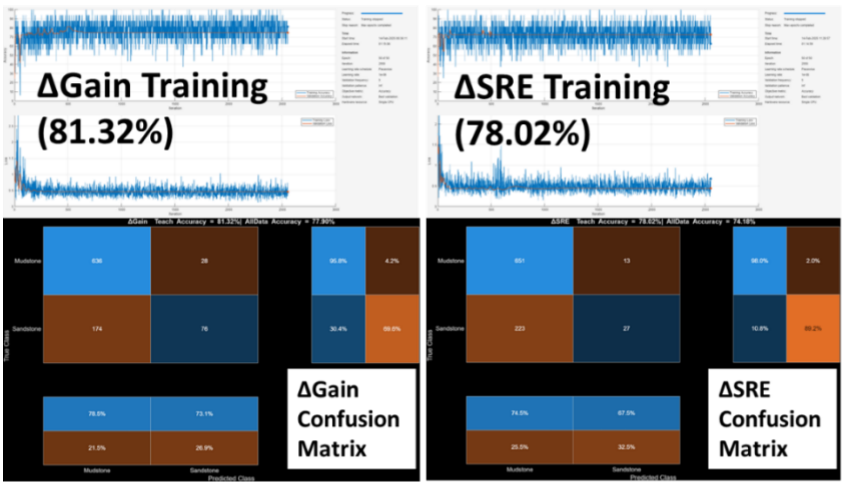

Figure 29: Top Row: Training Performance for ΔGain (Left) and ΔSRE (Right). Bottom Row: Confusion Matrices for ΔGain (Left) and ΔSRE (Right).

Figure 29 presents a comparative analysis of rock lithology classification performance using two distinct input feature types: ΔGain and ΔSRE. The left panel displays results from training the CNN model using

ΔGain (difference in gain across successive packets), achieving a training accuracy of 81.32%. The accompanying confusion matrix shows that Mudstone is classified with high accuracy (95.8%), while Sandstone is more frequently misclassified, with a lower correct prediction rate (30.4%). However, the model still demonstrates an overall trend of effective learning from ΔGain inputs, particularly in distinguishing Mudstone.

The right panel showcases the training outcomes using ΔSRE (difference in specific rock energy between successive drilling time intervals), yielding a slightly lower training accuracy of 78.02%. In this case, Mudstone again shows strong classification performance (98.0% accuracy), while Sandstone continues to pose classification challenges, with a true positive rate of only 10.8%. The confusion matrix further reveals that Sandstone samples are frequently misidentified as Mudstone, indicating an overlap in the temporal energy signature characteristics of these rock types when using SRE as the input.

Despite the classification challenges with Sandstone in both models, the training curves indicate steady convergence, with loss consistently decreasing across epochs, affirming model stability. The results collectively suggest that ΔGain offers a better discriminative capacity than ΔSRE in identifying lithologies, particularly when distinguishing between Mudstone and Sandstone. This reinforces the idea that frequency-domain derived features may offer greater sensitivity to lithological differences than purely time-domain metrics, particularly in high-impact drilling environments.

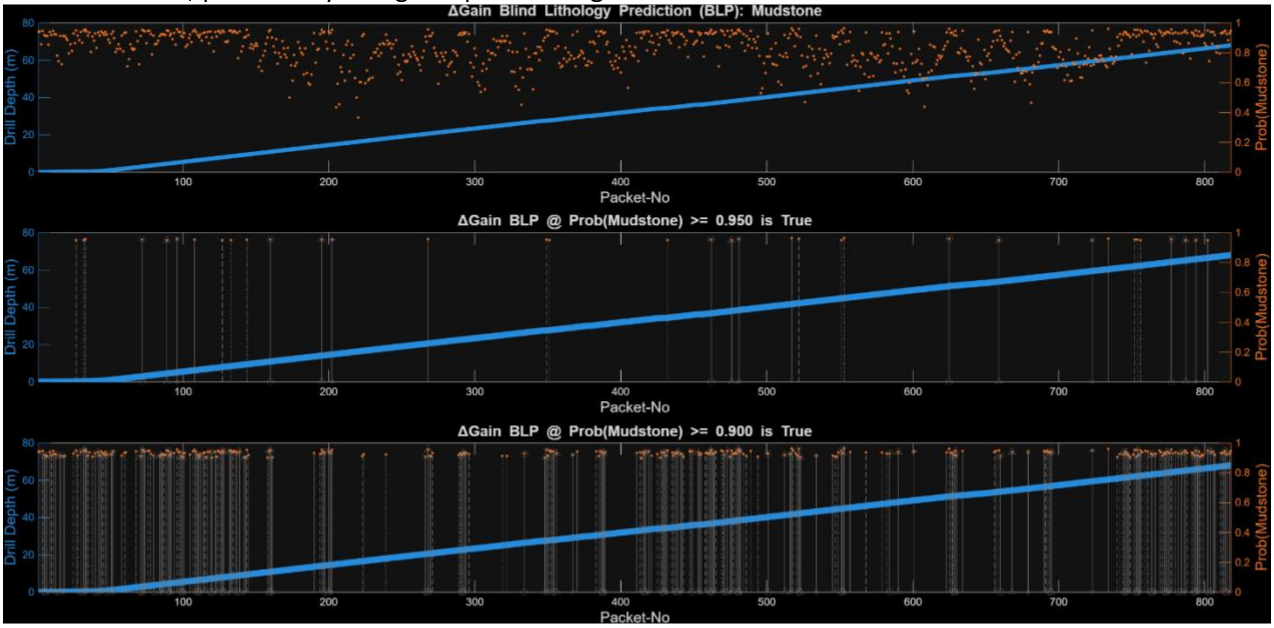

Figure 30: Predicting Mudstone Probability using Laboratory ΔGain Teach Data

Figure 30 ΔGain-Based Classification: This shows a probabilistic output derived from a model trained on the ΔGain feature extracted from IMU sensor data in the laboratory. When applied to field drilling data, the model estimates the likelihood of each data packet corresponding to Mudstone. The resulting probability trace overlays this output against the depth of drilling or time progression, revealing zones with a high or low probability of Mudstone presence. The higher probability regions align closely with the known Mudstone intervals provided on page 1, thereby validating the model's ability to generalize from controlled lab conditions to unpredictable field environments.

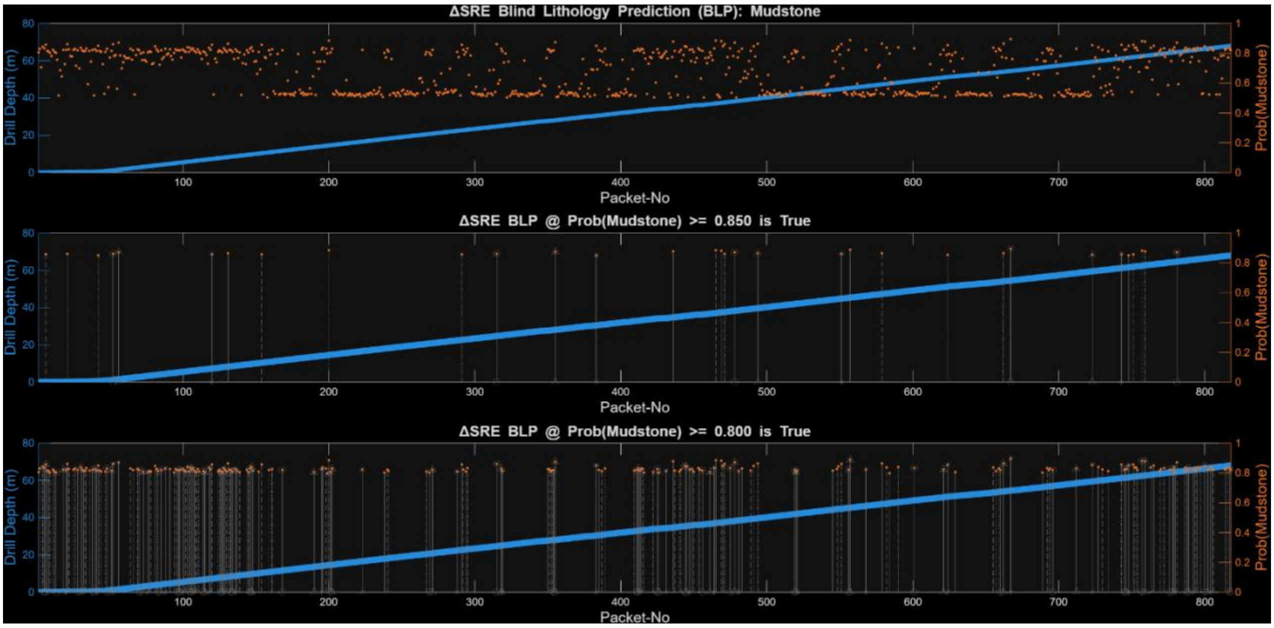

Figure 31: Predicting Mudstone Probability using Laboratory ΔSRE Teach Data

Figure 31 ΔSRE-Based Classification Similarly, illustrates the predictions from a model trained on the time-domain derived ΔSRE (Successive Rock Energy Difference) values. These capture changes in energy demand during drilling, indirectly reflecting rock hardness and consistency. The ΔSRE-based classifier also outputs probability values for Mudstone presence across the drilling timeline. This trace again aligns with the established lithology from page 1, though it may capture slightly different nuances due to its time-domain basis. Together, these models provide complementary views of subsurface geology.

5.2.2 Methodological Novelty and Impact

This two-pronged approach leveraging both frequency-domain (ΔGain) and time-domain (ΔSRE) data is notably innovative. Traditionally, lithology is determined through core sampling or basic parameter monitoring such as Rate of Penetration (ROP) or torque. The OptiDrill methodology introduces a data-driven, sensor-informed alternative that allows continuous, in-situ lithology prediction with a high degree of granularity. This enables real-time lithological classification, which not only reduces reliance on time-consuming sampling methods but also enhances operational decision-making.

In terms of drilling efficiency, especially in scenarios like geothermal or hard rock environments, identifying transitions between lithologies such as Mudstone and Sandstone can help operators adapt drilling parameters (e.g., weight on bit, fluid pressure, bit type) on the fly. Such adjustments directly affect ROP, bit wear, and overall energy usage, thus improving both performance and cost-effectiveness.

Application to Lunar and Martian Missions The broader implications of this method extend to extraterrestrial drilling. On the Moon or Mars, where traditional sample return or human intervention is limited, onboard AI models capable of blind lithology prediction based on sensor data could revolutionize planetary exploration. These models can infer the subsurface composition in real time, guiding autonomous drilling systems, optimizing power allocation, and improving mission planning. By applying ΔGain and ΔSRE data captured from lightweight IMUs, future missions could classify regolith or bedrock types remotely, enhancing the scientific return while ensuring safer and more efficient subsurface access.

6. CONCLUSION

The first field test of the OptiDrill system at the Bochum site has provided valuable insights into the performance and applicability of the drilling advisory system developed within the project. Over the course of four days, the system successfully monitored and analyzed approximately 127 meters of drilled depth, using advanced machine learning frameworks to predict and optimize the rate of penetration, identify drilled lithological classes, and detect anomalies in the data that could potentially indicate drilling problems. Despite the challenges posed by the significant differences in drilling parameters between the training dataset and the field test conditions, the results indicate that the OptiDrill system can offer meaningful predictions and optimization recommendations. The ROP prediction model outputs had a mean absolute error (MAE) of 2.24 m/h, which, while higher than the validation dataset, remained within acceptable limits given the contextual variances. When the new scaler was applied to the data preprocessing and the model outputs were reconstructed, the MAE could even be lowered to 1.44 m/h, reaching a similar level as during the model validation on the Geostar 2 dataset. The optimization of the ROP values showed an theoretical enhancement potential of nearly 19%, indicating the system's capability to improve drilling efficiency.These enhancements remain theoretical. Also, since the optimization is based on the ROP prediction model, the error made by the model has a direct impact on the results calculated for the optimized ROP and MSE values.

The drilled lithology prediction module faced limitations, primarily due to discrepancies between the training and inference datasets. The model predominantly predicted claystone during the field test, reflecting potential biases from the training data and the unsuitable scaler that was used for the preprocessing. During a second evaluation of the model using a scaler fit to the actual drilling data of the field test, the model's predictive behavior and performance could be improved. However, the results were still not satisfying and showed some limitations of the system. This underscores the necessity for further model retraining and adaptation to enhance its robustness and generalization capabilities across new geological settings.

Overall, the findings from this field test highlight both the promise of the OptiDrill system and the areas for improvement. For the second round of the field testing of the OptiDrill system the findings from the first round will be of great value. The models will be retrained on the new dataset from the first wellbore to improve the model performance. Especially the drilled lithology model's predictive capabilities will be refined by retraining it on the classes from the first round of drilling. The ROP prediction model already delivered very good predictions that already matched those that were achieved during validation on the historical dataset the model was trained on.

The sensor sub testing campaign, conducted across both German and UK field sites, confirmed the viability of OLIVERIS's modular sensor system in capturing high-fidelity drilling data under operational conditions. By shifting from real-time transmission to a high-capacity memory-based approach, the system circumvented the cost and complexity of BHA connectivity, focusing instead on robust signal acquisition for SRE and SID analysis. The packetised data architecture enabled fine-grained interpretation of drilling events, transitions, and lithological changes.

Notably, the UK trials introduced the concept of "blind" lithology prediction using single-sensor input. Models trained on laboratory ΔGain and ΔSRE features were applied to field data with strong alignment to known rock types, proving the capability of frequency- and time-domain fusion techniques to improve in-situ classification. The results demonstrate that a single IMU sensor, properly deployed and interpreted, can support intelligent drilling decisions in real time or retrospectively.

These findings affirm the novel integration of sensor-based system identification with AI-driven lithology classification as a transformative approach for drilling operations—one with implications that extend beyond geothermal applications to broader industrial and even planetary exploration contexts.

6.1 Future developments

The next steps are to continue working with industry to test the data subs in a downhole environment, through the increase in battery life and dfata collection reliability, as well as ensure that the subs can survive the rigours of drilling. The project will search out additional funding opportunities, to fully develop and test the system for commercial operations and in tegration into drilling systems.