

Accepted Manuscript

Geoenergy

An Isolation Forest-based system for drilling problem identification in geothermal wells

Matthew I. Arran, Andrew Kingdon, Kevin Mallin, Henning Knauer & Shahin Jamali

DOI: <https://doi.org/10.1144/geoenergy2026-002>

To access the most recent version of this article, please click the DOI URL in the line above. When citing this article please include the above DOI.

This article is part of the Geothermal energy for decarbonization: advancing heat and power solutions in the UK and beyond collection available at:

<https://www.lyellcollection.org/topic/collections/geothermal-energy>

Received 6 January 2026

Revised 28 September 2026

Accepted 29 September 2026

© 2026 UKRI. The British Geological Survey. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Published by The Geological Society of London for GSL and EAGE. Publishing disclaimer: <https://www.lyellcollection.org/publishing-hub/publishing-ethics>

Supplementary material at <https://doi.org/10.6084/m9.figshare.c.8754041>

Manuscript version: Accepted Manuscript

This is a PDF of an unedited manuscript that has been accepted for publication. The manuscript will undergo copyediting, typesetting and correction before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Although reasonable efforts have been made to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record once published for full citation and copyright details, as permissions may be required.

An Isolation Forest-based system for drilling problem identification in geothermal wells

Abbreviated title: Isolation Forests for drilling problem detection

Matthew I. Arran^{1*}, Andrew Kingdon², Kevin Mallin³, Henning Knauer⁴, Shahin Jamali⁴

1. British Geological Survey, The Lyell Centre, Research Avenue South, Edinburgh EH14 4AP, UK
2. British Geological Survey, Keyworth, Nottingham NG12 5GG, UK
3. Geolorn Ltd., 20 Barnton Street, Stirling, FK8 1NE, UK
4. Fraunhofer IEG, Am Hochschulcampus 1, 44801 Bochum, Germany

M.I.A. <https://orcid.org/0000-0001-5711-1033>

A.K. <https://orcid.org/0000-0003-4979-588X>

K.M. <https://orcid.org/0009-0009-4085-9139>

H.K. <https://orcid.org/0009-0000-0910-8593>

S. J. <https://orcid.org/0000-0002-8924-9181>

Abstract

Geothermal energy systems often require drilling to multi-kilometre depths to access and extract heat. This is an expensive process, and the risk of unforeseen drilling problems can significantly affect the long-term economic sustainability of such projects. Moreover, the low availability of drilling data from geothermal development impedes efforts to improve drilling processes and mitigate this risk. Here, we demonstrate that publicly accessible legacy data from hydrocarbon exploration and development drilling can be used to develop an effective, generalisable system for the early identification of drilling problems. The system, based on the Isolation Forest algorithm, identifies problems that interrupted drilling at a rate significantly greater than random chance (detection rate 29% compared to 8.6% baseline, $p=0.012$), even when applied to new measurements in new geological contexts. Whilst there are significant limitations of such a system, which we discuss, our work highlights the utility of legacy data holdings in facing contemporary challenges.

Introduction

Geothermal district heating is a major opportunity for decarbonisation. Domestic energy use accounted for 26% of total energy use in the UK in 2023, with over 60% of this used for space heating (Department for Energy Security and Net Zero, 2024) and significant additional energy used for space heating of industrial and commercial premises. Most of this heating need is currently met by individual natural gas boilers (Ministry of Housing, Communities and Local Government, 2025). District heating schemes, powered by heat derived from deep geothermal energy, have the potential to heat entire neighbourhoods and so to replace these sources of atmospheric carbon (Cariagi, 2025; Government Office for Science, 2024).

However, achieving sustainable and sufficient geothermal supply requires drilling into permeable strata at depths below the Earth's surface that may reach several kilometres (Abesser & Walker, 2022), imposing high drilling costs that can be exacerbated by operational drilling problems. Drilling to significant depth involves challenges that are not only social and political (Hymans & Uchikoshi, 2022; Missaghieh-Poncet & de Sartre, 2026), but also technical, requiring an approximately 50-m-high rig, specialist equipment, and a crew of around a dozen personnel (Eden Geothermal, 2026). Mid-range estimates of the associated costs, based on internal BGS data and in 2024/25 prices, are £1.3 million per onshore borehole (Human Economics, 2025). Costs increase to around £1.7 million per km depth for boreholes between 1 and 2 km deep, corresponding to an estimated £2 million to £4 million per MW of thermal energy (Arup, 2021). And problems encountered during drilling operations are responsible for a significant proportion of these costs, being typically responsible for around 10 days of delay in 100 days of deep hydrocarbon exploration drilling (Hossain, 2015). The minimisation and mitigation of drilling problems is therefore critical in the delivery of economically sustainable geothermal boreholes.

Problems encountered during drilling remain important once drilling is complete. Geothermal heat and electricity supply are forecast to be relatively expensive in capital development costs compared to other renewable energy sources (Department for Energy Security and Net Zero, 2025), with cost recovery periods measured in decades. This implies a premium on long-term well integrity and the avoidance of well damage.

The OptiDrill project sought new systems for the early identification and mitigation of drilling problems, by leveraging continued improvements in the data available to drilling operators (OptiDrill, 2026). When surface or downhole measurements are monitored, changes in measured values can indicate a wide variety of problems (Hafezi, 2023). Increases in data availability allow systems to detect these indicators without the user-specified thresholds of e.g. Mao & Zhang (2019) and Elmgerbi & Thonhauser (2022). For example, Zhong, Salehi, & Johnson Jr. (2022) reviews 49 papers in which machine-

learning-based models are trained to detect stuck pipe, lost circulation, well kick, and various other drilling problem types.

Machine-learning based identification models can be broadly classified as supervised or unsupervised, and contrasted with physics-based models (Kutz, 2013). In physics-based models, the data generating process is described by theoretically derived governing equations, permitting deductive identification of the underlying system state. In supervised learning, the patterns linking the data and the underlying state are themselves deduced, using a representative 'training dataset' in which each datapoint is accompanied by an identifying label. In unsupervised learning, it is the patterns in the data alone that are deduced, without any reference to the underlying system.

Systems based on unsupervised machine learning can be applied to new drilling conditions and new sets of drilling data, without the need for physics-based models of complex drilling processes and without the risk of prior training becoming outdated. Physics-based models, while applicable in a broad range of conditions, cannot describe every aspect of complex drilling processes and, to describe any aspect, require many unknown physical parameters to be estimated (Erge & van Oort, 2022; Kaneko, et al., 2018). Models based on supervised machine-learning, meanwhile, rely on consistent training datasets in which drilling problems are labelled, and are applicable only to situations that resemble those from which these datasets were drawn (Ilyas, et al., 2019; D'Amour, et al., 2022). With drilling problems rare and datasets evolving rapidly, any such model's applicability is likely to be limited. In contrast, models based on unsupervised machine-learning can learn arbitrarily complex patterns in the data to which they are applied, without specification of parameters or of training data.

An Isolation Forest is an unsupervised machine-learning algorithm for identifying anomalous datapoints that, in the case of drilling data, may be associated with drilling problems. Introduced by Liu, Ting, & Zhou (2008), the algorithm recursively partitions a given dataset, choosing at each stage a random non-unit subset of datapoints, a random feature, and a random threshold, then splitting the subset into two new subsets according to whether datapoints' values exceed the threshold. This partition is equivalent to a binary tree, with each branch node a split and each leaf node a datapoint. Furthermore, the number of splits required to isolate an anomalous datapoint, equal to the path length to the datapoint in the corresponding tree, is expected to be small. By generating many such isolation trees and calculating the average path length to each datapoint, as illustrated in Fig. 1a, the algorithm classifies as anomalous the datapoints requiring the smallest number of splits. On datasets for which subsequent datapoints may be similarly anomalous, like drilling data, the Isolation Forest algorithm outperforms comparable unsupervised anomaly-detection algorithms, at lower computational cost (Bouman, et al., 2024). Consequently, it has been used by Fernandes Jr., Komati, & de Souza Gazolli (2023) to detect potentially

problematic anomalies during production from hydrocarbon wells, and by Altindal, et al. (2024) to detect problems in well drilling.

In this paper, we develop and test an Isolation Forest-based system for drilling problem identification in new drilling conditions, from new sets of drilling data. This goes beyond the work of Altindal et al. (2024), whose system used training data from 13 sections of a well to identify drilling problems in an intermediate 14th section, and is therefore not generalisable to new and different wells. Such generalisability is essential when aiming to improve geothermal drilling, for which relevant and accessible data are limited.

Data for improving geothermal drilling

Improving geothermal drilling requires an understanding of both the rocks being drilled and the performance of the drilling process through those rocks. Such understanding, in turn, requires quantitative data. Quantitative rock property data are supplied by core and cutting analysis and by geophysical logging; digital logging data has been routinely acquired during drilling operations since the 1970s. Quantitative drilling data, meanwhile, can be collected from two sources: instrumentation at the well head, which records weight on bit and drillstring rotation speed, and the Measurement While Drilling (MWD) sub built into the Bottom Hole Assembly (BHA). Quantitative understanding of geothermal drilling requires a combination of digital data from these two data sources, which has been available to drillers since the early 2000s.

Unfortunately, publicly accessible data from purely geothermal operations are as yet rare to non-existent, particularly in the UK. Since the 40-year-old resource assessment of Downing and Gray (1986), the only geothermal wells drilled to depths over 1 km are those of the Eden Geothermal Project (Eden Geothermal, 2026) and United Downs Deep Geothermal Project (Reinecker, et al., 2021). Data from the wells that have been drilled are not widely available, due to commercial sensitivity and a lack of regulatory data deposit requirements (Abesser, et al., 2023).

In contrast, there are plentiful, relevant, openly available drilling data from UK legacy hydrocarbon exploration and development. The drilling process is fundamentally the same for both geothermal and hydrocarbon use cases (Allahvirdizadeh, 2020), and the United Downs boreholes in Cornwall, UK demonstrate that similar types of drilling data and geophysical data can be collected during the drilling process (Ledingham, et al., 2019). Although there are potential plays that are specific to geothermal energy, such as Cornwall's deep granites, many of the potential recoverable heat resources are in deep sedimentary basins that have already been subject to hydrocarbon exploration. The UK benefits from the open release of all legacy oil and gas industry data via BGS (British Geological Survey, 2026), the UK Onshore Geophysical Library (UK Onshore Geophysical Library, 2026), and the North Sea Transition Authority's UK National Data Repository (NSTA NDR) (North Sea Transition Authority, 2026). Systems based on

unsupervised learning can therefore be developed with these legacy hydrocarbon data before application to new geothermal borehole drilling.

Legacy hydrocarbon data are served in a variety of formats, dependent upon the age and origins of the data. Through time, the UK hydrocarbon industry has progressively standardised the data supplied, but there remain major issues with non-standardised nomenclature. By the late 1970s, geophysical logging data was typically supplied on tape, in Log Information Standard (LIS) or Digital Log Information Standard (DLIS) file formats, and the digital Log ASCII Standard (LAS) format (Canadian Well Logging Society, 2017) has been typical since the mid-1980s. Routine capture of digital drilling parameters is much more recent, with digital data only available from the 2000s or later. Even now, the format of the digital MWD data that are available is not standardised.

Methods

Our methodology is illustrated in Fig. 1, with Fig. 1a and Fig. 1b illustrating components of our system for drilling problem detection, and Fig. 1c the workflow by which we use legacy hydrocarbon data to tune and test that system. We draw measurement-while-drilling (MWD) data and lists of drilling problems from the UK Onshore Geophysical Library (UKOGL) (UK Onshore Geophysical Library, 2026); design an Isolation-Forest-based system for drilling problem detection; tune the system's hyperparameters using data from 6 geographically dispersed wells; and test the system using data from 20 wells within the Wytch Farm oil field.

Datasets

The legacy data used in this work were taken from the drilling of UK onshore hydrocarbon exploration and development wells, and deposited in the UKOGL. These data holdings represent over 1,000 wells, including over 95% of all hydrocarbon wells drilled into the United Kingdom continental shelf and approximately 50% of all onshore UK hydrocarbon wells ever drilled. From these many wells, we identified those that sampled stratigraphy of potential relevance to deep geothermal drilling: deeply buried sedimentary basins similar to those near large British cities, such as the Permo-Triassic sandstones in the vicinity of Birmingham, Manchester and Liverpool. Of these, 26 wells were identified for which borehole data holdings include both digital MWD data and reports from which labels of drilling problems could be derived: either daily drilling reports (DDRs) or end-of-well reports (EOWRs).

MWD data record drilling-relevant measurements such as the downhole depth of the drill bit (MD); the equivalent circulating density (ECD) and total flow rate (TFLO) of drilling fluid; this fluid's pressure in the standpipe (SPP); the torque (TRQ) and weight (WOB) applied to the drill pipe or drill bit; and the bit's rotation rate (RPM) and rate of penetration (ROP). The availability and measurement details of these data differ over

the UKOGL dataset, as does the form in which they are recorded: single or multiple files in the LAS format (Canadian Well Logging Society, 2017); in other structured ASCII formats; in Microsoft Excel formats (Microsoft Corporation, 2023); or in less easily readable formats that are out of this paper's scope. Data may be indexed by either the time of measurement or by measured depth MD, and may include intervals of non-drilling time or of drilling through a previous well section's cement shoe, alongside intervals of fresh drilling.

DDRs and EOWRs include text descriptions of drilling activities and brief explanations of any drilling interruptions or measurements of concern, as recorded by a drilling operator or supervisor during the drilling operation process. The presentation of information is not standardised, with records from different wells including varying levels of detail and varying technical acronyms or shorthand. For example, the second DDR of well Lybster 1 uses "Circ round hi vis sweep, work string to pooh" to denote the pumping of a specialised, high-viscosity drilling fluid down the borehole, to remove debris, and the rotation and lifting of the drill pipe column attached to the drill bit, in order to facilitate their extraction. EOWRs' descriptions, in particular, are often highly simplified summaries of actual drilling operations, derived from DDRs' near-real-time descriptions. In general, any problem that interrupted drilling is recorded, but details include the depth at which drilling was interrupted rather than inferences of the depth at which the problem first arose. Descriptions may include the observed indicators of the problem, the type of problem inferred to have occurred, or the remedial action taken.

For this study, we required wells that sample the relevant stratigraphies described above, for which multiple, distinct, digital MWD measurements were available, and for which drilling problems could be manually identified from the DDRs and EOWRs. We required a set of these wells to be sufficiently diverse for hyperparameter tuning to be broadly representative, and the remainder to be sufficiently similar for tests over them to be consistent. Table 1 lists the 26 wells we chose to satisfy these requirements, with six geographically dispersed and twenty in the large Wytch Farm oil field, while Fig. 2 indicates their locations and stratigraphies.

For the six geographically dispersed wells, plus four of the Wytch Farm wells, the first stage of DDR interpretation was the production of drilling activity charts, presenting on-site activity for each hour of drilling operations and including summaries of DDR descriptions. The Supplementary Material includes an example, for well Lybster 1. With productive drilling time clearly labelled, problems that interrupted drilling could be easily identified, and categorised either as drilling problems or as other issues: surface equipment failures or logistical delays, for example. For the remaining Wytch Farm wells, drilling problems were identified directly from DDRs or EOWRs. For all wells, we listed the downhole depths MD at which drilling problems occurred; an example of such a list is in the Supplemental Material.

Algorithm for Drilling Problem Detection

The system for drilling-problem detection involved three stages: the preprocessing of MWD data; the application of the Isolation Forest algorithm to assign datapoints anomaly scores; and the consequent identification of anomalous incidents.

Given MWD data like that described above, the first stage of anomaly detection was cleaning and rescaling. Datapoints for which a feature takes negative or excessively duplicated values were judged erroneous and removed from the dataset, with the threshold for excessive duplication a value frequency 20 times higher than that of a uniformly distributed variable with the same interquartile range and measurement precision. Also removed from the dataset were any records associated with non-drilling time or the drilling of cement well shoes. Such records, determined using records of the initial and terminal measured depths of each well section, are mechanically anomalous with respect to the drilling of rock but not associated with drilling problems, and thus were not relevant to these analyses. For each well and each feature, the remaining values were then rescaled using a Yeo-Johnson transformation (Yeo & Johnson, 2000), reducing skewness and thereby permitting comparison of anomalousness across features with different distributions.

Once a dataset was clean and rescaled, the Isolation Forest algorithm incorporated within scikit-learn (Pedregosa, et al., 2011) was applied to the combination of measured depth with each feature individually and to the combination of all features. So that the random selection of threshold feature values was based on the observed ranges of values, each algorithm application fit 100 isolation trees using subsamples of the dataset: 256-datapoint subsamples for individual features and 128×2^K -datapoint subsamples for the combination of all features, where K denotes the number of features. Using these sets of isolation trees, each datapoint was assigned $K + 1$ anomaly scores, with more anomalous datapoints having lower anomaly scores.

Finally, from a well's datapoints' anomaly scores, the problem-detection system sought to identify an average of r anomalous incidents per unit distance of MWD data, by identifying anomalous datapoints with measured depths within depth interval ΔMD as belonging to the same incident. A point was considered anomalous if one or more of its anomaly scores was in the lowest c th quantile of all datapoints' corresponding anomaly scores, with c chosen to ensure that there were the required number of distinct anomalous incidents. Each incident was then considered to extend from a measured depth $\Delta MD / 2$ before the first anomalous point that belonged to it, to a measured depth $\Delta MD / 2$ after the last such anomalous point. A simplified example of this identification is illustrated by Fig. 1.

To simulate our system's application for real-time drilling-problem detection, these three stages were applied after each 5 m interval of new drilling, to all MWD data taken

at the well up to that point, and the output was any anomalous incident within that most recent 5 m. The number of relevant features available, K , differed between wells, and the implementation required the specification of two hyperparameters: the target rate with depth r at which anomalous incidents were to be detected, and the minimum depth interval ΔMD of each such incident.

Hyperparameter tuning

The hyperparameters r and ΔMD were tuned to optimise simulated real-time detection of the drilling problems in the six geographically dispersed UK onshore hydrocarbon exploration wells: Balcombe 2, Hykeham 1, Kirby Misperton 5, Lester's Lane 1, Lybster 1, and Willows 1. The effectiveness of this hyperparameter tuning does not rely on the comprehensiveness of these drilling problems, but only on the representativeness of their average rate of occurrence and extent, which allows us to determine appropriate values of r and ΔMD .

Specifically, for each 5-m interval of recorded drilling in each of the six wells, we applied our system to data from that interval and shallower depths, with minimum depth intervals of anomalous incidents $\Delta MD = 0.5$ m, 1 m, 2 m, and 5 m, and with the proportions of points anomalous

$$c = \left(1 - c_0^{\frac{1}{K+1}}\right), \quad (1)$$

for K the number of features and with $c_0 = 0.00, 0.01, 0.02, \dots 1.00$. The features used were the subset of MD, ECD, TFLO, TRQ, WOB, RPM, and ROP that were available, as listed in Table 1. The values of c were chosen to recover good coverage of potential method specificity, even with each datapoint assigned $K + 1$ anomaly scores, each of which could lead the datapoint to be considered anomalous. For each combination of hyperparameter values, each DDR-or-EOWR-derived drilling problem in the 5-m interval was regarded as correctly identified if the associated depth was within a system-identified anomalous incident.

For each value of ΔMD , the system's implementation for all recorded drilling, for various c , permitted calculation of the receiver operating characteristic (ROC) described by Fawcett (2006): the proportion of all true drilling problems identified (sensitivity) as a function of the proportion of all depths identified as within an anomalous incident ($1 - \text{specificity}$). Including the additional point (1, 1), corresponding to the identification of all depths as anomalous, the optimum value of ΔMD was taken to be that which maximised the area under the ROC curve: $\Delta MD = 1$ m (see Fig. 3). With this value of ΔMD , the optimum value of r was taken to be that, among those resulting from the above c , which maximised the positive likelihood ratio sensitivity / ($1 - \text{specificity}$): $r = 0.0032$ / m. Higher values of r would increase the drilling problem detection rate, but at the cost of a higher number of false positives.

Testing

The testing data were taken from the Wytch Farm oil field: a set of three reservoirs into which over 100 onshore wells have been drilled, from 13 sites in Dorset, England. These Wytch Farm wells were chosen to be geographically and geologically distinct from the wells used for hyperparameter tuning, as Fig. 2 shows, and so to ensure the applicability of test results to new drilling contexts. The large number of wells provide examples of different drilling problems experienced within the same geological context, offering the opportunity to compare the rates at which different types of problem are detected. To verify the problem-detection system's applicability to new measurements, the standpipe pressure (SPP) was included in analyses alongside previous features (measured depth MD, rate of penetration ROP, total pump flow rate TFLO, surface weight on bit WOB, drill bit rotation rate RPM, drilling fluid equivalent circulating density ECD, and drilling torque TRQ).

To simulate the real-time drilling problem detection for which the system is intended, MWD data were processed in the same manner as for hyperparameter tuning. Specifically, for each well, test points were defined at 5 m intervals between the least and greatest measured depths at which MWD data was available. For each test point, the tool was provided with the available MWD data from all fresh drilling at lesser depths in the well, as if applied when drilling reached that point, and identified anomalous incidents within these data. Any new anomalous incidents, restricted to the 5 m depth interval since the last test point, were recorded, and their depth ranges compared to the depths of drilling-problem-related drilling interruptions.

Results

Our drilling-problem-detection system rapidly identified anomalous incidents in MWD data, with a standard laptop's 3 GHz CPU requiring at most 5 s of computation for 5 m of new drilling. As Fig. 4 demonstrates, these anomalous incidents corresponded weakly to the depths of drilling-problem-related drilling interruptions, which we take to be the depths of drilling problems. Most anomalous incidents did not include a drilling interruption, with anomalies instead corresponding to erroneous measurements, extremes of natural variation, or changes in drilling conditions. Near the start of drilling, in particular, new datapoints were likely to fall outside the range of previous variation, and so to be identified as anomalous. Furthermore, many drilling interruptions did not lie within any anomalous incident, with no anomalies apparent within the available MWD data. Equivalent results, for each feature of each Wytch Farm well's MWD data, are in the Supplementary Material.

However, the proportion of drilling problems that were detected using the problem-detection system is greater than would be expected by random chance, to a statistically

significant extent. Across all 31 km of drilling for which data were analysed for testing, only 5 anomalous incidents were identified in every 100 m of drilling, and the combined depth ranges of all anomalous incidents covered 8.6% of the total depth range. Yet of the 17 drilling-problem-related interruption depths within the range of MWD data, 5 lay within an identified anomalous incident. If $P = 0.086$ is the probability of randomly detecting each drilling problem, the probability of detecting at least $d = 5$ of $n = 17$ drilling problems is

$$p = \sum_{k=d}^n \binom{n}{k} P^k (1 - P)^{n-k} = 0.012. \quad (2)$$

Consequently, the null hypothesis H_0 , that the drilling problem detection tool is no better than random selection, must be rejected at the 5% significance level.

There is no evidence for differences between the chances of detection of drilling problems of different type. Table 2 lists all the drilling-problem-related drilling interruptions inferred from DDRs and EOWRs, demonstrating that no type of drilling problem is consistently detected or consistently undetected. Quantitatively, we calculate the numbers of drilling problems that are detected across three broad categories: 3 of 9 high losses of drilling fluid, 1 of 5 low rates of penetration, and 1 of 4 motor failures. For each category, the proportion of drilling problems that are detected is consistent with the 29% chance of detection implied by the cross-category results. This may, however, be a consequence of the limited number of drilling interruptions in the dataset.

Discussion

Limitations of tests

Whilst the tests described in this paper provide a guide to the effectiveness of the machine-learning-based system in detecting drilling problems, the system's practical effectiveness is likely to differ significantly from the results described here. This generalization error emerges from three limitations of the legacy dataset available for testing:

1. the limited number of drilling problems described within the dataset,
2. the need to infer the depths at which drilling problems were encountered, and
3. the impossibility of examining the counterfactual case, in which the tool was applied during drilling.

Within the entirety of the dataset used for testing, only 17 drilling problems were identified from DDRs and EOWRs. Such reports provide unambiguous descriptions of only those drilling problems that were sufficiently severe to interrupt drilling and, while it is these problems that are of most concern to drilling operators, they are generally rare. In the Wytch Farm oil field the experience accumulated over repeated drilling, in similar geology, reduced the number of severe drilling problems encountered to a very small number of incidents per borehole. Consequently, the number of drilling problems in this analysis was sufficient to show the problem-detection system to be effective, but not to identify variation in its effectiveness. Expansion of our analysis to other wells, such as the offshore wells within the NSTA dataset, would require substantially larger drilling problem datasets. These, in turn, require either manual effort or the development of Natural Language Processing tools that can parse drilling problems within DDRs and EOWRs.

More problematically, the true depths at which drilling problems were encountered were not described in DDRs or EOWRs, and therefore had to be inferred from the depths at which an interruption in drilling was attributed to a drilling problem. These two sets of depths are likely to differ significantly, as operators may interrupt drilling due to the perception of a problem where one does not exist, and may not interrupt drilling until long after a progressive or subtle problem has first emerged. For example, Fig. 5a shows that the drilling of well A13 continued for over 100 m between a substantial drop in standpipe pressure, identified by the problem-detection tool as anomalous, and a drilling interruption due to uncontrollable losses. Such occasions, in which the system could have allowed drillers to act long before they are documented to have acted, are strong indicators of the tool's potential value. But we cannot objectively distinguish them from false positives and so cannot count them in our retrospective assessment. Similarly, we cannot count successful detections of problems that went unnoticed, or were deemed too minor to interrupt drilling.

Relatedly, records of what actually happened, in the problem-detection system's absence, provide limited information about what would have happened, had the system been applied. The identification of anomalous incidents might have provided drilling operators with information that improved the drilling process, even when these incidents did not correspond to a problem that interrupted drilling. On the other hand, the identification of anomalous incidents during normal drilling might have distracted drilling operators, impeding drilling, or resulted in unnecessary interruptions. The system would only have been useful if it allowed drillers to take action that they did not take without it and that improved the drilling process more than it impeded it. Effective application of a problem-detection system would require not only that the system is effective in detecting problems, but also that it is effectively integrated into the drilling process.

Limitations of MWD data for drilling problem detection

Even a drilling problem detection tool that is both effective and effectively integrated will be subject to the limitations of MWD data.

All electronic sensors are sensitive to issues when operating in a hostile subsurface environment. Sensors operating on a drill string are at a risk of water ingress, heat, pressure and physical damage by vibration and abrasion. Data transfer from subsurface sensors is also problematic. In current drilling operations, data transfer during drilling is reliant upon low bandwidth acoustic pulses, with full data available only through physical download once drilling has ceased. Connectors are always at risk of creating leakage pathways and live connections are likely to be subject to disruption from the drilling process.

The effects of these measurement and data transfer issues are dropouts in data or spikes of unrealistic data. Drilling effects can decrease the accuracy of sensor recording, with effects such as mudcake on sensors lowering sensor resolution, and some drilling events going unrecorded. For example, Fig. 5b shows no evidence in well F25's RPM data that the bit stuck at a measured depth of 1,805 m. In consequence, the problem-detection system cannot be considered a diagnostic tool but only, if it provides data useful in overcoming drilling difficulties, an indicative tool.

Hydrocarbon and geothermal development drilling

Although we expect the system we have developed to be applicable to geothermal development drilling, based on its effectiveness on data from hydrocarbon drilling, there are differences in drilling problem detection requirements between the two use cases. Each hydrocarbon well presents an opportunity to encounter very large and financially remunerative volumes of hydrocarbons; a common aim of the drilling process is to drill as fast as possible into the potentially productive horizon. A

successful well might be worth millions of dollars immediately after a hydrocarbon column is encountered, so the quality of the derived borehole created is largely separate from the financial payback. In contrast, if a geothermal well provides access to warm or hot water, in permeable strata at depth, then the payback period is measurable in decades. Since the largest expense in geothermal development is the cost of drilling the well, there is a strong incentive to maximise the integrity of this well, ensuring a robust borehole that lasts for this extended payback period (Lukawski, et al., 2014).

We expect that risks to long-term well stability will be associated with anomalies similar to those considered in this paper, associated with drilling interruptions. But further research is required to validate the association between the anomalies detected by our system and the drilling problems that impair long-term well stability, especially as the completion of geothermal wells differs from that of oil or gas wells. Consistent problem reporting and increased data availability for geothermal well drilling would aid such research.

Conclusions

The availability of datasets from the drilling of geothermal development wells is limited, but the UK's North Sea Transition Authority holds plentiful, openly available datasets from the analogous drilling of hydrocarbon exploration and development wells. From the holdings associated with UK onshore wells, we extracted a dataset of digital measurement-while-drilling data and records of the depths at which drilling problems interrupted drilling.

Using this dataset, we developed a broadly applicable system for the real-time detection of drilling problems, from measurement-while-drilling data. The system uses unsupervised machine learning, and the Isolation Forest algorithm in particular, to identify intervals of anomalous data. It therefore relies neither on the theoretical understanding necessary for a physics-based system, nor on the prior, equivalent, fully representative, and labelled datasets necessary for supervised machine learning. The system has only two hyperparameters, which we tuned using six geographically dispersed wells, maximising the positive likelihood ratio of detecting depths at which drilling problems interrupted drilling: the probability that such a depth lies within an anomalous interval, divided by the overall proportion of depths within anomalous intervals.

Applying the system to separate data from twenty separate wells, from a single large oil field, our system detected problems that interrupted drilling at a rate significantly greater than random (29%, $p = 0.012$), with around 5 false identifications per 100 m of drilling and a low computational cost. This test indicates that our system can be successfully applied to new drilling, in new geological and operational contexts.

However, drilling problems were too rare within our dataset to permit comparison of different drilling problems' detection rates, and results from retrospective records of drilling interruptions cannot be directly transferred to real-time detection of drilling problems. Further work is required to establish the applicability of our results to geothermal development drilling, especially given the limitations of measurement-while-drilling data. Such work could also examine other unsupervised machine-learning algorithms for drilling problem detection.

Acknowledgements

This paper was published with the permission of the Executive Director, British Geological Survey (NERC, UKRI). We're also grateful for the comments and suggestions of Dr. Arka Dyuti Sarkar, Dr. Richard Jones, and one anonymous reviewer.

Author contributions

MIA: conceptualisation, data curation, formal analysis, investigation, methodology, software, visualisation, writing – original draft; AK: conceptualisation, funding acquisition, project administration, resources, supervision, writing – review & editing; KM: data curation, validation, writing – review & editing; HK: validation, writing – review & editing; SJ: funding acquisition, project administration.

Funding information

This work was partially supported by the OptiDrill project (<https://www.optidrill.eu/>) which received funding from the Building a Low-Carbon, Climate Resilient Future: Secure, Clean And Efficient Energy (H2020-Lc-Sc3-2018-2019-2020) programme.

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

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The datasets analysed during the current study are available in the UK Online Geophysical Library (UK Onshore Geophysical Library, 2026).

References

- Abesser, C. et al., 2023. *The case for deep geothermal energy - unlocking investment at scale in the UK*, London: Department for Energy Security and Net Zero.
- Abesser, C. & Walker, A., 2022. *Geothermal energy*, POSTbrief 46, London: UK Parliament. <https://doi.org/10.58248/PB46>
- Allahvirdizadeh, P., 2020. A review on geothermal wells: Well integrity issues. *Journal of Cleaner Production*, **275**, 124009, <https://doi.org/10.1016/j.jclepro.2020.124009>
- Altindal, M. C., Nivlet, P., Tabib, M., Rasheed, A., Kristiansen, T. G. & Khosravanian, R., 2024. Anomaly detection in multivariate time series of drilling data. *Geoenergy Science and Engineering*, **237**, 212778, <https://doi.org/10.1016/j.geoen.2024.212778>
- Arup, 2021. *Deep Geothermal Energy - Economic Decarbonisation Opportunities for the United Kingdom*, London: The Association for Renewable Energy and Clean Technology (REA).
- Bouman, R., Bukhsh, Z. & Heskes, T., 2024. Unsupervised anomaly detection algorithms on real-world data: how many do we need?. *Journal of Machine Learning Research*, **25**, 1–34, <https://doi.org/10.48550/arXiv.2305.00735>
- British Geological Survey, 2026. GeoIndex (onshore), <https://www.bgs.ac.uk/map-viewers/geoindex-onshore> [Accessed June 2026].
- Canadian Well Logging Society, 2017. *LAS Version 2.0: A Digital Standard for Logs*, Update February 2017, https://cwls.org/wp-content/uploads/2017/02/Las2_Update_Feb2017.pdf [Accessed June 2026].
- Cariagi, C., 2025. Geothermal district heating in Aarhus, Denmark officially starts operations, <https://www.thinkgeoenergy.com/geothermal-district-heating-in-aarhus-denmark-officially-starts-operations/> [Accessed December 2025].
- D'Amour, A. et al., 2022. Underspecification Presents Challenges for Credibility in Modern Machine Learning. *Journal of Machine Learning Research*, **23**(226), 1–61, <https://doi.org/10.48550/arXiv.2011.03395>
- Department for Energy Security and Net Zero, 2024. *Energy consumption in the UK (ECUK) 2024*, <https://www.gov.uk/government/statistics/energy-consumption-in-the-uk-2024> [Accessed September 2025].
- Department for Energy Security and Net Zero, 2025. *UK geothermal energy review and cost estimations*, <https://www.gov.uk/government/publications/uk-geothermal-energy-review-and-cost-estimations> [Accessed September 2025].

Downing, R. A. & Gray, D. A., 1986. Geothermal resources of the United Kingdom. *Journal of the Geological Society*, **143**, 499–507, <https://doi.org/10.1144/gsjgs.143.3.0499>

Eden Geothermal, 2026. Drilling and operations, <https://www.edengeothermal.com/the-project/drilling-and-operations> [Accessed June 2026].

Elmgerbi, A. & Thonhauser, G., 2022. Holistic autonomous model for early detection of downhole drilling problems in real-time. *Process Safety and Environmental Protection*, **164**, 418–434, <https://doi.org/10.1016/j.psep.2022.06.035>

Erge, O. & van Oort, E., 2022. Combining physics-based and data-driven modeling in well construction: Hybrid fluid dynamics modeling. *Journal of Natural Gas Science and Engineering*, **97**, 104348, <https://doi.org/10.1016/j.jngse.2021.104348>

Fawcett, T., 2006. An introduction to ROC analysis. *Pattern Recognition Letters*, **27**(8), 861–874, <https://doi.org/10.1016/j.patrec.2005.10.010>

Fernandes Jr., W., Komati, K. S. & de Souza Gazolli, K. A., 2023. Anomaly detection in oil-producing wells: a comparative study of one-class classifiers in a multivariate time series dataset. *Journal of Petroleum Exploration and Production Technology*, **14**, 343–363, <https://doi.org/10.1007/s13202-023-01710-6>

Government Office for Science, 2024. Future of the subsurface: geothermal energy generation in the UK (annex), <https://www.gov.uk/government/publications/future-of-the-subsurface-report/future-of-the-subsurface-geothermal-energy-generation-in-the-uk-annex> [Accessed June 2026].

Hafezi, S., 2023. *Real-time detection of drilling problems & issues during drilling by listing & using their signs both on the surface and downhole*, Master's thesis: Norges teknisk-naturvitenskapelige universitet. <https://hdl.handle.net/11250/3082757>

Hossain, M. E., 2015. Drilling Costs Estimation for Hydrocarbon Wells. *Journal of Sustainable Energy Engineering*, 3–32, <https://doi.org/10.7569/jsee.2014.629520>

Human Economics, 2025. *Economic valuation of BGS' National Geological Repository: final summary report*, London: Human Economics. <https://nora.nerc.ac.uk/id/eprint/540081>

Hymans, J. E. C. & Uchikoshi, F., 2022. To drill or not to drill: determinants of geothermal energy project siting in Japan. *Environmental Politics*, **31**(3), 407–428, <https://doi.org/10.1080/09644016.2021.1918493>

Ilyas, A., Santurkar, S., Tsipras, D., Engstrom, L., Tran, B. & Madry, A., 2019. Adversarial Examples Are Not Bugs, They Are Features. NeurIPS, Vancouver, 125–136, <https://doi.org/10.48550/arXiv.1905.02175>

Kaneko, T., Wada, R., Ozaki, M. & Inoue, T., 2018. Combining Physics-Based and Data-Driven Models for Estimation of WOB During Ultra-Deep Ocean Drilling. ASME, Madrid, V008T11A007, <https://doi.org/10.1115/OMAE2018-78229>

Kutz, J. N., 2013. *Data-Driven Modeling & Scientific Computation*. 1st ed. Oxford: OUP.

Ledingham, P., Cotton, L. & Law, R., 2019. The United Downs deep geothermal power project. Proceedings of the 44th workshop on geothermal reservoir engineering, 11-13 February, Stanford, CA, USA, 513–523.

Liu, F. T., Ting, K. M. & Zhou, Z.-H., 2008. Isolation Forest. 2008 Eighth IEEE International Conference on Data Mining Proceedings, 15-19 December, Pisa, Italy, 413–422, <https://doi.org/10.1109/ICDM.2008.17>

Lukawski, M. Z., Anderson, B. J., Augustine, C., Capuano, L. E., Beckers, K. F., Livesay, B. & Tester, J. W., 2014. Cost analysis of oil, gas, and geothermal well drilling. *Journal of Petroleum Science and Engineering*, **118**, 1–14, <https://doi.org/10.1016/j.petrol.2014.03.012>

Mao, Y. & Zhang, P., 2019. An Automated Kick Alarm System Based on Statistical Analysis of Real-Time Drilling Data. Abu Dhabi International Petroleum Exhibition & Conference Proceedings, November 11–14, Abu Dhabi, UAE, SPE-197275-MS, <https://doi.org/10.2118/197275-MS>

Microsoft Corporation, 2023. Standards Support, https://learn.microsoft.com/en-us/openspecs/standards_support [Accessed June 2026].

Ministry of Housing, Communities and Local Government, 2025. English Housing Survey 2023 to 2024: low carbon technologies in English homes - fact sheet, <https://www.gov.uk/government/statistics/english-housing-survey-2023-to-2024-low-carbon-technologies-in-english-homes-fact-sheet> [Accessed June 2026].

Ministry of Housing, Community & Local Government, 2021. *English Housing Survey Energy report, 2019-20*, https://assets.publishing.service.gov.uk/media/620e3bbb8fa8f54911e22180/Energy_Report_2019-20.pdf [Accessed September 2025].

Missaghieh-Poncet, J. & de Sartre, X. A., 2026. Unlocking the subsurface through knowledge controversy: how earthquakes induced by deep geothermal drilling have politicized the subsurface environment. *The Extractive Industries and Society*, **25**, 101780, <https://doi.org/10.1016/j.exis.2025.101780>

North Sea Transition Authority, 2026. UK National Data Repository, <https://www.nstauthority.co.uk/data-and-insights/data/uk-national-data-repository/> [Accessed June 2026].

OptiDrill, 2026. OptiDrill: Optimisation of Geothermal Drilling Operation with Machine Learning, <https://www.optidrill.eu/> [Accessed June 2026].

Pedregosa, F. et al., 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830, <https://doi.org/10.48550/arXiv.1201.0490>

Reinecker, J., Gutmanis, J., Foxford, A., Cotton, L., Dalby, C. & Law, R., 2021. Geothermal exploration and reservoir modelling of the United Downs deep geothermal project, Cornwall (UK). *Geothermics*, **97**, 102226, <https://doi.org/10.1016/j.geothermics.2021.102226>

UK Onshore Geophysical Library, 2026. UK Onshore Geophysical Library, <https://ukogl.org.uk/> [Accessed January 2026].

Yeo, I. & Johnson, R. A., 2000. A new family of power transformations to improve normality or symmetry. *Biometrika*, **87**(4), 954–959, <https://doi.org/10.1093/biomet/87.4.954>

Zhong, R., Salehi, C. & Johnson Jr., R., 2022. Machine learning for drilling applications: A review. *Journal of Natural Gas Science and Engineering*, **108**, 104807, <https://doi.org/10.1016/j.jngse.2022.104807>

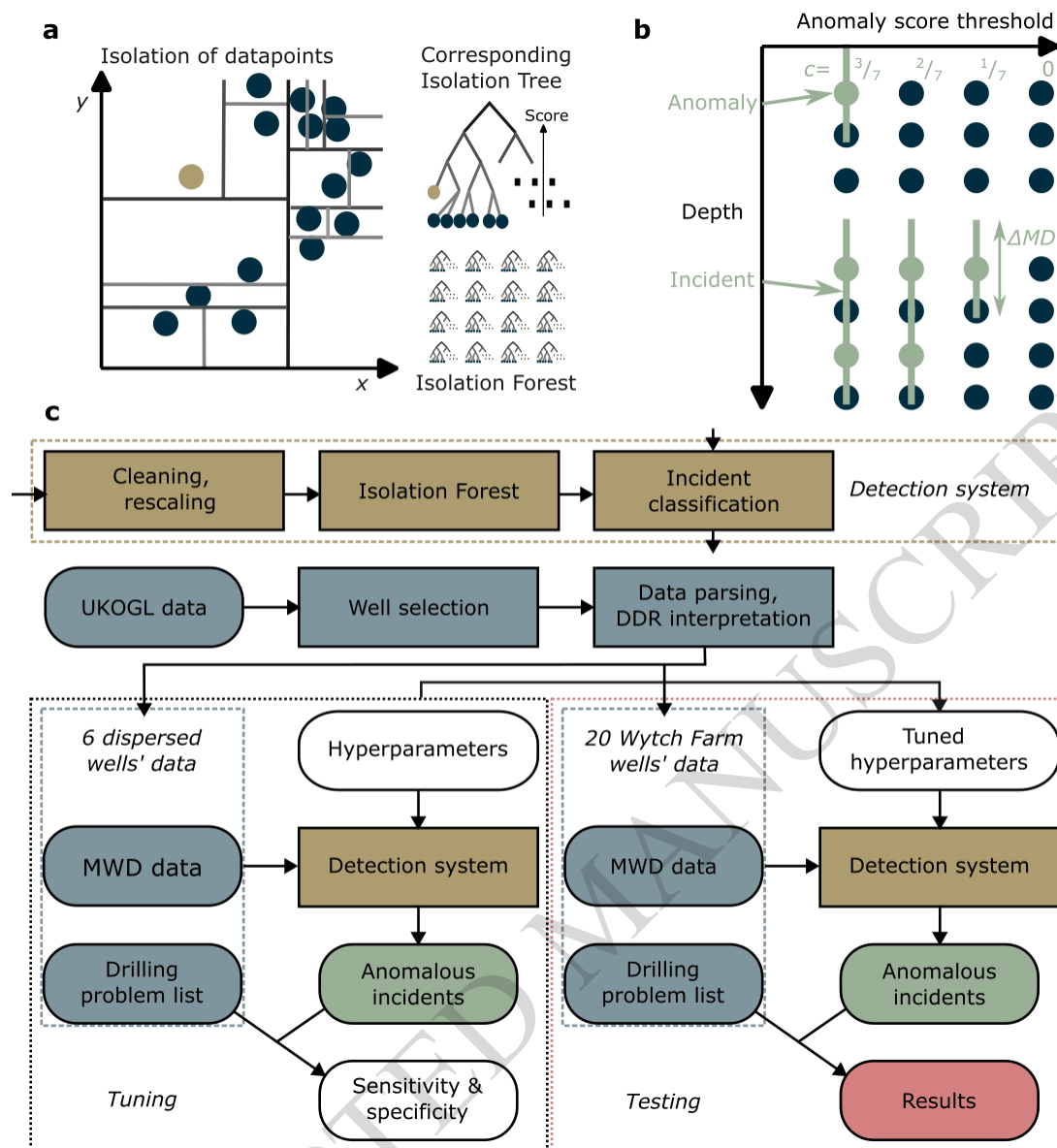


Fig. 1: Schematics of our detection system's components and our workflow for its tuning and testing. (a) Isolation Forest anomaly score calculation. (b) Classification of anomalous incidents (green) from anomaly scores, given hyperparameters c and ΔMD . (c) Information flow within our drilling problem detection system (gold); our derivation of datasets (blue); our tuning of system hyperparameters (white); and our calculation of test results (pink).

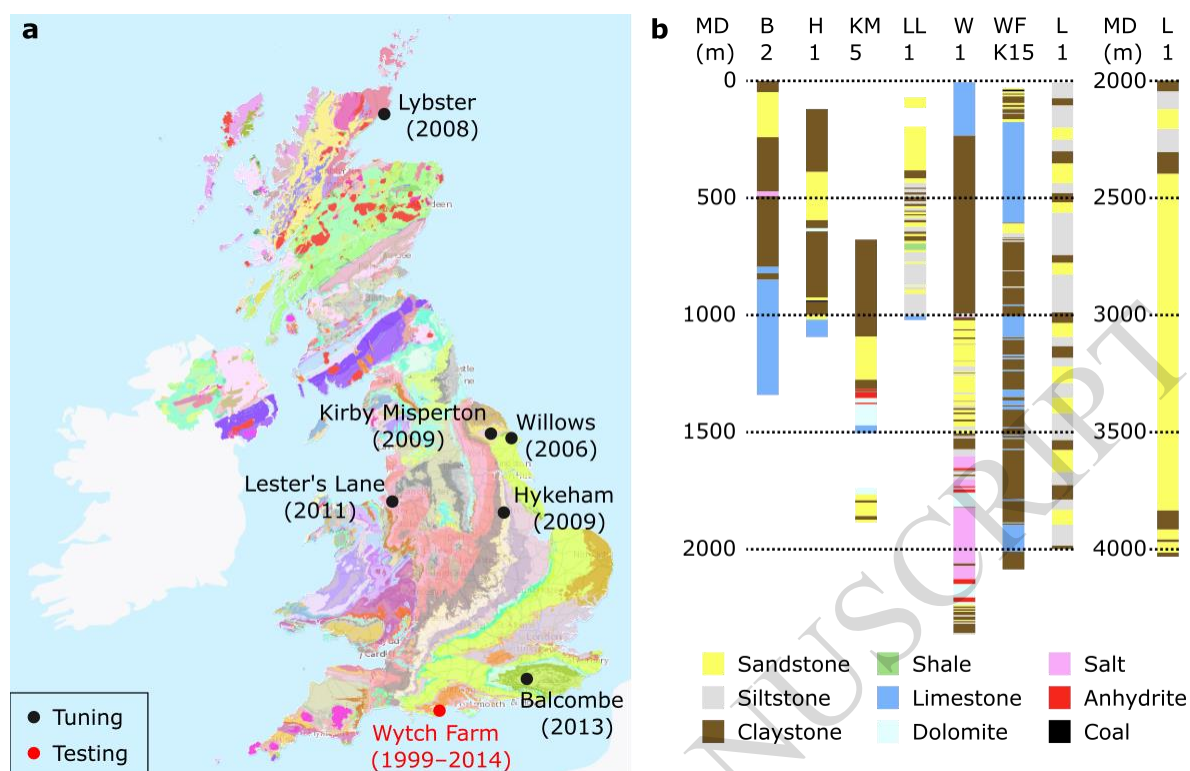


Fig. 2: Context of drilling data. (a) Locations and years at which drilling started, in the context of British bedrock geology, excluding Shetland (BGS © UKRI). (b) Simplified well stratigraphy, drawn from available lithology logs, with each well indicated by its initials. Wytch Farm's stratigraphy is represented by Wytch_Farm_K15's and, for reasons of scale, Lybster_1's stratigraphy is in two sections. The geographically dispersed wells were used for hyperparameter tuning and the Wytch Farm wells for testing.

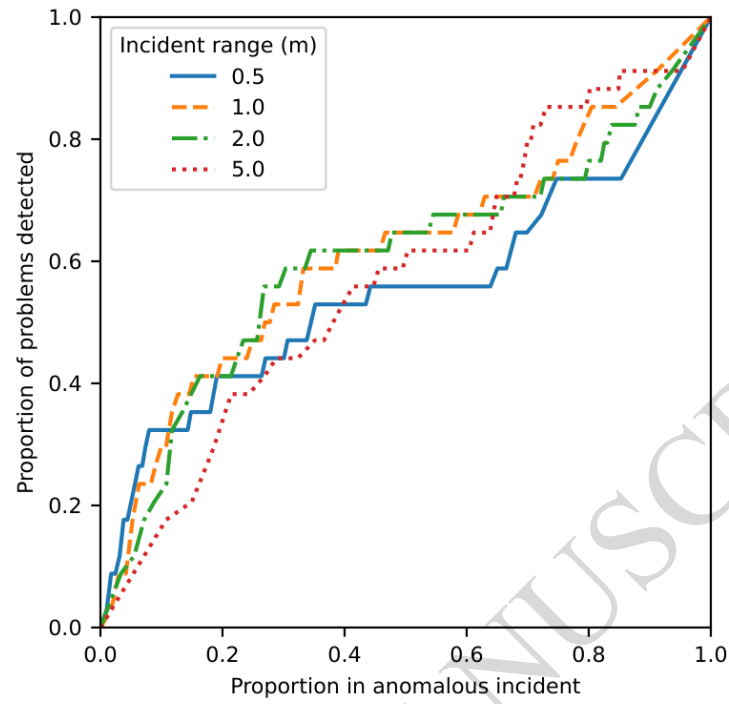


Fig. 3: ROC curve for simulated real-time identification of drilling problems.

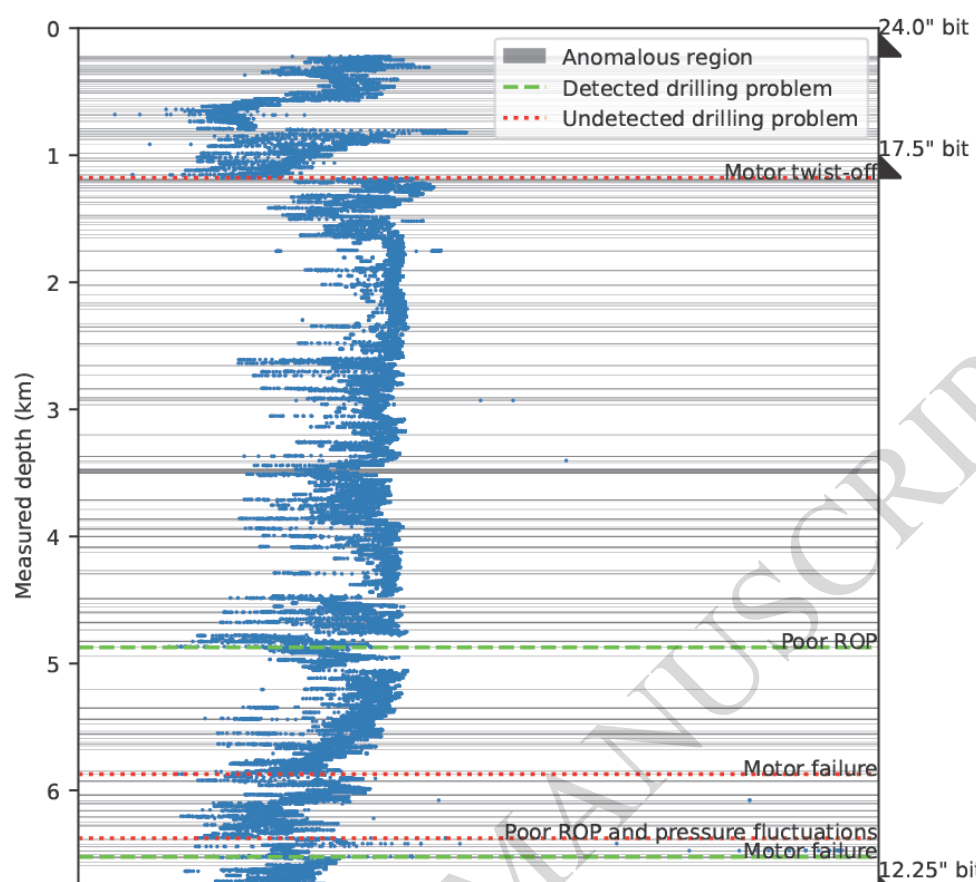


Fig. 4: Correspondence between drilling problems and anomalous incidents, for Wytch Farm well M18. The depths of system-identified anomalous incidents and of problem-related drilling interruptions are indicated in the main plot, with each problem labelled by its description in the EOWR, while the depths of well section changes are indicated at right. Rate of Penetration datapoints are provided for context.

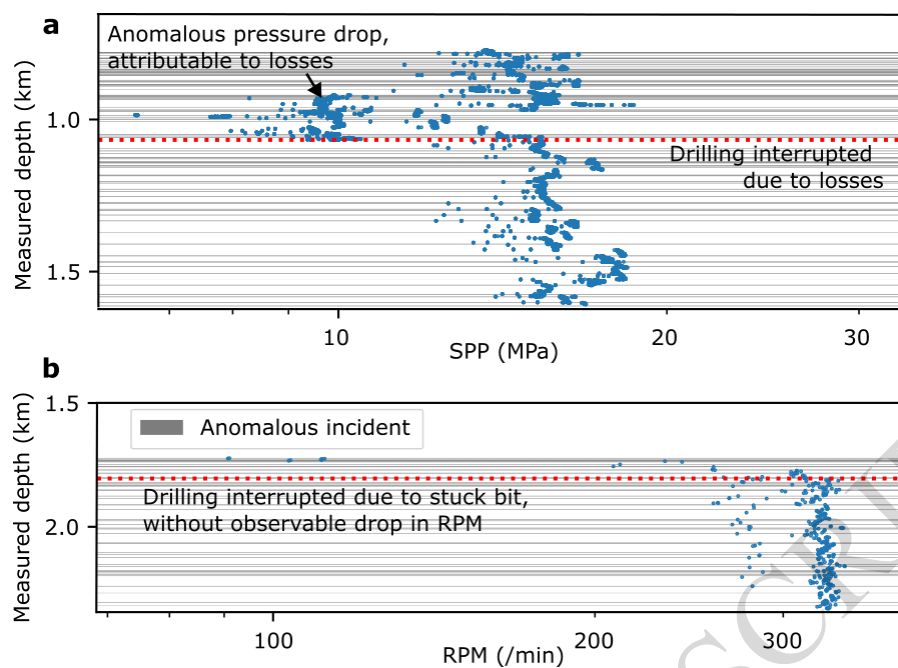


Fig. 5: Potential disparities between the depths of drilling problems and of drilling interruptions. Points indicate standpipe pressure in well A13 (a) and bit rotation rate in well F25 (b), grey regions anomalous incidents, and red lines drilling interruptions, labelled by the corresponding problem description. The horizontal scales are logarithmic.

Table 1: MWD data in this study, with feature acronyms defined above. Wytch Farm wells were reserved for testing.

Well designation	Datapoints	Medians of MWD data, where used in analyses							
		MD	WOB	RPM	TRQ	ECD	TFLO	SPP	ROP
		(km)	(kN)	(/min)	(kNm)	(kg/l)	(l/s)	(MPa)	(m/hr)
Balcombe_2	2701	0.42	51.2	107	3.91	1.27	37.3		9.5
Hykeham_1	3960	0.62	41.8	205	0.98		48.8		12.7
Kirby_Misperton_5	4314	1.31	37.3			1.22	19.4		8.9
Lesters_Lane_1	1003	0.57	37.2	83	15.25	1.05	24.3		5.4
Lybster_1	13298	2.07	205.1	49	19.73		51.8		2.8
Willows_1	2231	1.29	85.0	99			32.1		8.7
Wytch_Farm_A12	26022	2.47	27.8	45	6.82		15.9	22.3	9.5
Wytch_Farm_A12Y	9170	3.31	21.2	0	0.36		15.1	22.5	5.3
Wytch_Farm_A13	106778	2.09	44.3	99	7.95		17.8	19.4	14.5
Wytch_Farm_A13Z	7960	2.37	25.0	100	6.97		16.4	27.7	5.7
Wytch_Farm_B11	873	1.85	22.7	261	0.02		15.0	21.5	15.6
Wytch_Farm_B11Y	51857	2.71	27.4	50	0		15.7	20.6	9.9
Wytch_Farm_B11Z	119122	2.43	32.2	49			16.2	21.5	14.3
Wytch_Farm_F22Y	10000	4.44	47.2	263	13.88				12.0
Wytch_Farm_F24	9409	2.36	36.0	35	0.60	1.24	16.5	23.2	13.1
Wytch_Farm_F25	1909	2.68	35.5	326	7.23	1.28	16.9	19.5	15.4
Wytch_Farm_G4	7379	1.26	42.5	107	6.08		33.6	17.2	17.5
Wytch_Farm_K1Y	1250	1.95	71.6	49	2.14			22.5	12.3
Wytch_Farm_K1Z	688	1.51	45.8	0	0.11			22.5	10.8
Wytch_Farm_K15	2096	1.07	52.8	115	10.06	1.11	35.6	13.6	19.9
Wytch_Farm_K16	6321	0.43	92.9	0	0		51.6	10.3	6.3
Wytch_Farm_K16Y	21239	2.87	36.5	119	10.71		13.2	16.9	7.1
Wytch_Farm_K16Z	154780	1.47	52.4	113	9.76		33.5	15.2	10.9
Wytch_Farm_L14	806	1.85	31.8	269	0.02	1.24	15.8	19.3	11.0
Wytch_Farm_M11	5948	1.51	88.5	80					46.8
Wytch_Farm_M18	43733	3.54	93.4	102					35.8

Table 2: Details of drilling-problem-related drilling interruptions

WELL	MEASURED DEPTH (m)	DESCRIPTION IN DDR OR EOWR	DETECTED?
A13	1067	Uncontrolable [sic] losses	X
F22Y	3949	Failed mud motor	X
F22Y	4118	Mud losses, LCM pill blocking BHA	X
F22Y	4408	Dynamic losses at 50bbls/hr	X
F22Y	4688	Low ROP, 46bbls/hr losses	X
F22Y	4862	Pressure losses, drill pipe backed out	X
F22Y	5109	No progress	X
F24	3072	Sudden downhole mud loss of 30 bbls	✓
F25	1805	Bit stuck	X
G4	1673	POOH to cure losses	✓
K16Y	2898	Dynamic losses observed - 105bbl/hr	✓
K16Y	2900	Losses seen increasing to 75bbl/hr dynamic	X
M18	1179	Motor twist-off	X
M18	4873	Poor ROP	✓
M18	5872	Motor failure	X
M18	6376	Poor ROP and pressure fluctuations	X
M18	6522	Motor failure	✓