(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 January 2022 (13.01.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/010358 A1

(51) International Patent Classification:
E21B 49/00 (2006.01) *G01V 9/00* (2006.01)

(21) International Application Number:
PCT/NO2021/050158

(22) International Filing Date:
02 July 2021 (02.07.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2010337.0 06 July 2020 (06.07.2020) GB

(71) Applicant: **EQUINOR ENERGY AS** [NO/NO]; Forusbeen 50, 4035 Stavanger (NO).

(72) Inventors: **YANG, Tao**; Forusbeen 50, 4035 Stavanger (NO). **KOPAL, Margarete Maria**; Forusbeen 50, 4035 Stavanger (NO). **ARIEF, Ibnu Hafidz**; Forusbeen 50, 4035 Stavanger (NO). **YERKINKYZY, Gulnar**; Forusbeen 50, 4035 Stavanger (NO). **ULEBERG, Knut**; Forusbeen 50, 4035 Stavanger (NO).

(74) Agent: **LEES, Gregory**; Dehns, St. Bride's House, 10 Salisbury Square, London Greater London EC4Y 8JD (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*

(54) **Title:** RESERVOIR FLUID PROPERTY ESTIMATION USING MUD-GAS DATA

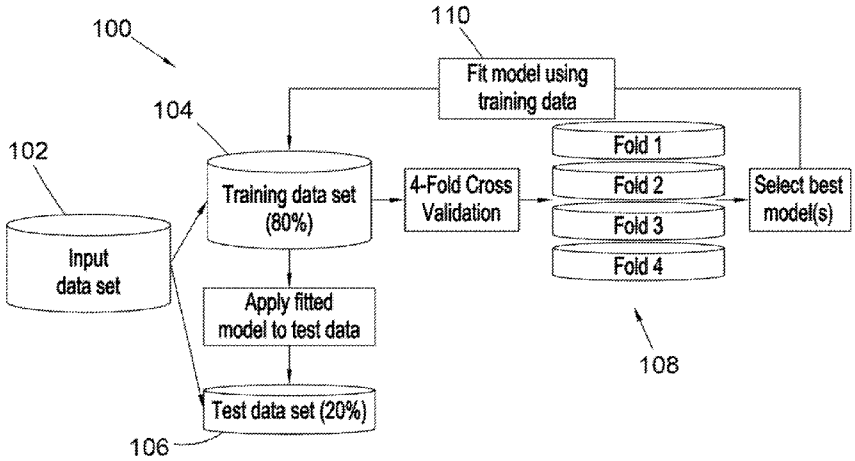


Fig. 2

(57) **Abstract:** A method is disclosed for generating a machine learning model to predict a reservoir fluid property, such as gas-oil ratio or density, based on standard mud-gas and petrophysical data. It has been found that this model predicts these reservoir fluid properties with an accuracy that is close to that which can be achieved using advanced mud-gas data. This is advantageous, as than standard mud-gas data and petrophysical data is much more readily available than advanced mud-gas data.

WO 2022/010358 A1

RESERVOIR FLUID PROPERTY ESTIMATION USING MUD-GAS DATA

The present disclosure relates to a logging technique for use whilst drilling a borehole, and particularly to a technique utilising mud-gas data to predict reservoir fluid properties.

Drilling fluid is a fluid used to aid the drilling of boreholes into the earth. The main functions of drilling fluid include providing hydrostatic pressure to prevent formation fluids from entering into the well bore, keeping the drill bit cool and clean during drilling, carrying out drill cuttings, and suspending the drill cuttings while drilling is paused and when the drilling assembly is brought in and out of the hole.

Drilling fluids are broadly categorised into water-based drilling fluid, non-aqueous drilling fluid, often referred to as oil-based drilling fluid, and gaseous drilling fluid. Liquid drilling fluids, i.e. water-based drilling fluid or non-aqueous drilling fluid, are commonly referred to as “drilling mud”.

Mud-gas logging entails gathering data from hydrocarbon gas detectors that record the levels of gases brought up to the surface in the drilling mud during a bore drilling operation.

Conventionally, mud-gas logging is used to identify the location of oil and gas zones as they are penetrated, which can be identified by the presence of hydrocarbon gas in the mud system. This may be used to provide a general indication of the type of reservoir, as well as to determine where to take downhole fluid samples for more detailed analysis of the fluid composition. The presence of hydrocarbon gas may be detected, for example, with a total gas detector.

Once the presence of hydrocarbon gas is detected, its composition may be examined for example with a gas chromatograph.

The most common gas component present is usually methane (C_1). The presence of heavier hydrocarbons, such as C_2 (ethane), C_3 (propane), C_4 (butane) and C_5 (pentane) may indicate an oil or a "wet” gas zone. Heavier molecules, up to about C_7 (heptane), may also be detectable, but are typically present only in very low concentrations. Consequently, the concentrations of these hydrocarbons are often not recorded.

There are two types of mud-gas data that can be collected, which are sometimes referred to a “standard” mud-gas logging, and “advanced” mud-gas logging. The equipment for standard mud gas logging and advanced mud gas logging are different.

For a standard mud gas system, the degasser does not usually have heating or uses constant volume gas separation. There is also only one sampling point of mud sample ("out") and therefore it is not suitable for recycling correction. The measured gas composition is usually referred standard mud-gas data, which is not directly comparable to the actual C_1 to C_5 composition of the reservoir fluid sample.

For an advanced mud gas system, the degasser has heating and usually uses a constant volume for gas separation. There are two sampling points of mud samples ("out" and "in"), and therefore it is possible to perform recycling correction. The measured gas composition is usually referred advanced mud-gas data.

When generating advanced mud-gas data, in order to make the data closely correspond to the actual reservoir fluid C_1 to C_5 concentrations, two correction processes are applied to the "raw" mud-gas data from the advanced mud gas logging system.

Firstly, a recycling correction is made to eliminate contamination of the gas by gases originating from previous injections of the drilling mud. This correction is applied based on a separate mud-gas measurement that was taken before the drilling mud was injected into the drilling string.

Secondly, an extraction efficiency correction step is applied to increase the concentration of intermediate components (from C_2 to C_5), such that the concentration of these components, relative to the C_1 concentration, more closely resemble the relative compositions of a corresponding reservoir fluid sample. The extraction efficiency correction is applied based on the type of drilling mud used for the borehole.

In the past, the advanced mud gas data would have been examined to estimate certain fluid properties of the reservoir fluid using broad, empirical correlations between the advanced mud-gas composition and certain fluid properties of the reservoir fluid. For example, extremely dry gas reservoirs should comprise mostly C_1 and not much C_{2+}, e.g. with each of the C_1/C_2, C_1/C_3, C_1/C_4 and C_1/C_5 ratios (for the raw mud-gas data) being greater than 50. Wet gas reservoirs will often have ratios between 20 and 50, and oil reservoirs will have ratios between 2 and 20.

Recently, an advanced machine learning model has been developed, making it possible to predict reservoir fluid properties much more accurately from

the advanced mud-gas data, even where those properties are dependent upon the oil part (C_{7+}) of the fluid which is not measured by the mud-gas data.

Details of how such a machine learning model was trained to determine a gas-oil ratio of the reservoir fluid based on the advanced mud-gas data can be found in the paper Tao Yang et. al. (2019), "A Machine Learning Approach to Predict Gas Oil Ratio Based on Advanced Mud Gas Data". Society of Petroleum Engineers. doi:10.2118/195459-MS

Advantageously, this model can be used to generate a substantially continuous log of the respective reservoir fluid property. This was not previously possible, and in the past, it was necessary to rely on downhole fluid samples. Furthermore, the model allows reservoir fluid property predictions to be made at a very early stage of the drilling process and without needing to interrupt the drilling process, as might be required to take downhole fluid samples or the like.

This model has been found to be very useful, but is limited in that it requires the availability of advanced mud-gas data. A need exists for a technique that can be used when advanced mud-gas data is not available.

The present invention provides a method of generating a model for predicting at least one property of a fluid at a sample location within a hydrocarbon reservoir, comprising:

providing a training data set comprising input data and target data, the input data comprising mud-gas data and petrophysical data for each of a plurality of sample locations, and the target data comprising the at least one property of the fluid for each of the plurality of sample locations; and

generating a model using the training data set such that the model can be used to predict the at least one property of the fluid at the sample location based on measured mud-gas data and measured petrophysical data for the sample location,

wherein a drilling fluid recycling correction has not been applied to the mud-gas data.

It is a commonly held belief within the oil and gas industry that petrophysical data provides only a qualitative indication of a reservoir fluid. The data usually predicts lean gas with good certainty but has reduced accuracy when used to distinguish rich gas condensate and oil. However, it has been identified that, by supplementing standard mud-gas data with petrophysical data, it is possible to provide an estimation of certain reservoir fluid properties with an accuracy that is close to the accuracy that can be achieved using advanced mud-gas data alone.

This is particularly advantageous where it is desirable to generate fluid property logs for large amounts of existing and new wells because standard mud-gas data and petrophysical data are collected for almost all the wells, including both exploration and production wells. Whereas, the additional cost of collecting advanced mud-gas data, and particularly of having the two sets of mud-gas analysis tools required to perform the recycling correction, means that it is often only collected when drilling some exploration wells. The number of wells with advanced mud gas data only represents a small portion of the total wells with standard mud-gas data and petrophysical data.

Additionally, the above technique allows for reservoir fluid property logs to be generated for new wells at a reduced cost, as it does not require the additional costs associated with collecting advanced mud-gas data. Importantly in this regard is that petrophysical data can be collected as a substantially continuous log, similar to mud-gas data. This contrasts with downhole fluid sample data, which requires interruption of the drilling process, adding significant additional costs to the drilling process.

In some embodiments, the input data may not comprise downhole fluid sampling data, and the model may not require downhole fluid sampling data as an input to predict the at least one property of the fluid at the sample location.

The method is preferably a computer-implemented method, and generating the model may comprise instructing a machine learning algorithm to generate the model using the training data set such that the model can be used to predict the at least one property of the fluid at the sample location based on measured mud-gas data for the sample location.

The at least one property is preferably a property influenced by the oil-related components of the fluid. That is to say, a property that not solely the product of the gaseous hydrocarbons within the fluid, whose composition can be predicted based on the mud-gas data.

The at least one property may comprise a density of the fluid at the sample location. It will be appreciated that the density may be calculated either at atmospheric conditions or reservoir conditions (e.g. taking into account the oil formation volume factor).

The at least one property may comprise a gas-oil ratio. That is to say, a ratio between the quantity of gaseous hydrocarbon and the quantity of liquid hydrocarbon, which is normally determined at surface conditions. The gas-oil ratio

is preferably a volume ratio. The gas-oil ratio may be a single-flash gas-oil measurement. However, any suitable gas-oil measurement may be used.

The at least one property may comprise a saturation pressure of the fluid at the sample location. That is to say, the pressure at which a secondary phase will appear with pressure depletion.

The at least one property comprises a formation volume factor of the fluid at the sample location. That is to say, the ratio of the volume of the fluid at reservoir (in-situ) conditions to the volume of the fluid at surface conditions.

The at least one property may comprise a concentration of a hydrocarbon within the fluid at the sample location. The hydrocarbon may be a hydrocarbon that is not included within the mud-gas data. For example, the hydrocarbon may be a C_{7+} hydrocarbon. That is to say, the hydrocarbon may be a C_7 hydrocarbon or may be a hydrocarbon heavier that C_7, e.g. a C_8 or heavier hydrocarbon. The hydrocarbon may be a hydrocarbon that is substantially an oil at reservoir conditions. The concentration of the hydrocarbon may be an absolute concentration (e.g. a molar concentration), or may be a relative concentration (e.g. a ratio compared to C_1), or may be an otherwise normalised concentration.

The reservoir may be a gas reservoir, a multiphase reservoir or an oil reservoir.

The at least one property for each sample location may be determined from reservoir fluid properties data associated with the sample location. The reservoir fluid properties data may comprise measured composition data for a fluid at the sample location. The reservoir fluid properties data may contain the composition of C_1 to C_{7+} hydrocarbons at the sample location, and preferably C_1 to C_{20+} hydrocarbons, and more preferably C_1 to C_{36+} hydrocarbons at the sample location. As used herein, the "C_{X+}" notation should be understood as meaning C_X or heavier hydrocarbons.

The mud-gas data of the training data set may comprise measured mud-gas data for the sample location, i.e. measured standard mud-gas data for the sample location.

The measured mud-gas data may be indicative of a composition of gases released from drilling fluid used whilst drilling through the sample location (i.e. passing through a drill bit performing the drilling). The measured mud-gas data may be indicative of a concentration of at least C_1 to C_4 gaseous hydrocarbons, and

preferably at least C_1 to C_5 gaseous hydrocarbons, that was released from drilling mud.

As discussed above, the mud-gas data preferably does not comprise advanced mud-gas data, i.e. where the mud-gas data has not been corrected so as to correspond to the gaseous hydrocarbon composition of the fluid at the sample location.

A drilling fluid recycling correction refers to correcting the mud-gas data to remove errors due to gases released from previous drilling operations, such as due to recycling of the drilling fluid. Typically, this would require a reference mud-gas data measurement collected before injection of the drilling mud into the drilling string.

Optionally, an extraction efficiency correction has not been applied to the mud-gas data.

An extraction efficiency correction refers to correcting the mud-gas data (C_1 to C_5) to closely correspond to reservoir fluid composition due to different hydrocarbons components have different abilities to vaporize from the drilling mud.

Where an extraction efficiency correction has not been applied, the training data may additionally comprise drilling mud compositional data. This may allow the machine learning model to correct the data within the model generated by the machine learning algorithm.

Alternatively, an extraction efficiency correction may have been applied to the standard mud-gas data. Often, the type of drilling mud used for a well is known and in many cases the extraction efficiency corrections can be estimated by Equation of State (EOS) simulation or approximated by testing. Consequently, even when using standard mud-gas data, it may be possible to retrospectively apply an extraction efficiency correction to standard mud-gas data.

Optionally, the mud-gas data was collected without the use of heating. Whilst standard mud-gas data may use heating; heating has often not been used when collecting mud-gas data. Consequently, where it is desirable to utilise the model to examine existing wells, a model trained using mud-gas data collected without the use of heating is particularly useful.

The petrophysical data may comprise any one or more of: bulk density, neutron porosity, resistivity data, acoustic data, natural gamma ray, nuclear magnetic resonance data, as well as slowing down time and gamma ray

spectroscopy data from pulsed neutron measurements, and the like. Optionally, the petrophysical data may comprise two or more of these data types.

Generating the model may comprise: training a machine learning algorithm with a first subset of the training data set; and testing the machine learning algorithm with a second, disjoint subset of the training data set. The first subset preferably comprises at least 50% of the samples of the training data set. The second subset preferably comprises at least 10% of the samples of the training set.

Viewed from a second aspect, the present invention provides a computer-based model for predicting at least one property of a fluid at a sample location within a hydrocarbon reservoir based on measured mud-gas data and measured petrophysical data for that sample location, the computer-based model having been generated by the method above.

Viewed from third aspect, the present invention provides a tangible computer-readable medium storing the computer-based model.

Viewed from a fourth aspect, the present invention provides a method of predicting a value of a property of a fluid at a sample location within a hydrocarbon reservoir, the method comprising: receiving measured mud-gas data and measured petrophysical data for the sample location; and predicting the value of the property of the fluid at the sample location by supplying the measured mud-gas data and the measured petrophysical data to the computer-based model.

The method may further comprise determining a quality for the measured mud-gas data and/or the measured petrophysical data.

The method may further comprise generating an indication of confidence associated with the predicted value of the fluid property. The indication of confidence may be a numerical indication, but other indications may be used, such as colour indications (e.g. red/yellow/green), or word indications (e.g. “good” / “poor”).

The indication of confidence may be based on the quality of the measured mud-gas data and/or the measured petrophysical data.

For a single data point, the indication of confidence may be reduced by one or more of a C_1, C_4 or C_5 concentration that is below a respective predetermined threshold.

Where standard mud-gas data is taken at a series of locations at different depths, the indication of confidence may be reduced by fluctuations of a component

concentration of the mud-gas data greater than a threshold amplitude within a predetermined depth range.

Where standard mud-gas data and/or the petrophysical data is taken at a series of locations at different depths, the indication of confidence may be reduced by the missing of a predetermined number of preceding measurements or over a predetermined depth range.

Viewed from a fifth aspect, the present invention provides a method of predicting a value of a fluid property of a fluid along a length of a well through a hydrocarbon reservoir, the method comprising: predicting a value of a fluid property of a fluid at a plurality of sample locations along a length of a well using the method above.

The method may comprise: displaying, using an electronic display screen, a graph plotting the predicting value of the fluid property against a location of the respective sample location for each of the plurality of sample locations along the length of the well.

The method may further comprise: indicating, using the electronic display screen, an indication of confidence associated with one or more of the predicted value. For example, the indication of confidence may be illustrated numerically, verbally, chromatically or iconographically.

All of the method described above, i.e. the methods of the first, fourth and fifth aspects may be performed in any suitable and desired way and on any suitable and desired platform. In a preferred embodiment the methods are each a computer-implemented method, e.g. the steps of the method are performed by processing circuitry.

The methods in accordance with the present invention may be implemented at least partially using software, e.g. computer programs. It will thus be seen that when viewed from further aspects the present invention provides computer software specifically adapted to carry out the methods described herein when installed on a data processor, a computer program element comprising computer software code portions for performing the methods described herein when the program element is run on a data processor, and a computer program comprising code adapted to perform all the steps of a method or of the methods described herein when the program is run on a data processing system.

The present invention also extends to a computer software carrier comprising such software arranged to carry out the steps of the methods of the

present invention. Such a computer software carrier could be a physical storage medium such as a ROM chip, CD ROM, DVD, RAM, flash memory or disk, or could be a signal such as an electronic signal over wires, an optical signal or a radio signal such as to a satellite or the like.

It will further be appreciated that not all steps of the methods of the present invention need be carried out by computer software and thus from a further broad embodiment the present invention provides computer software and such software installed on a computer software carrier for carrying out at least one of the steps of the methods set out herein.

The present invention may accordingly suitably be embodied as a computer program product for use with a computer system. Such an implementation may comprise a series of computer readable instructions, which may be fixed on a tangible, non-transitory medium, such as a computer readable medium, for example, diskette, CD ROM, DVD, ROM, RAM, flash memory or hard disk. It could also comprise a series of computer readable instructions transmittable to a computer system, via a modem or other interface device, over either a tangible medium, including but not limited to optical or analogue communications lines, or intangibly using wireless techniques, including but not limited to microwave, infrared or other transmission techniques. The series of computer readable instructions embodies all or part of the functionality previously described herein.

Those skilled in the art will appreciate that such computer readable instructions can be written in a number of programming languages for use with many computer architectures or operating systems. Further, such instructions may be stored using any memory technology, present or future, including but not limited to, semiconductor, magnetic or optical, or transmitted using any communications technology, present or future, including but not limited to optical, infrared or microwave. It is contemplated that such a computer program product may be distributed as a removable medium with accompanying printed or electronic documentation, for example, shrink wrapped software, pre-loaded with a computer system, for example, on a system ROM or fixed disk, or distributed from a server or electronic bulletin board over a network, for example, the Internet or World Wide Web.

Certain preferred embodiments of the present disclosure will now be described in greater detail, by way of example only and with reference to the accompanying drawings, in which:

Figure 1 is a schematic illustration of a mud-gas analysis tool; and

Figure 2 illustrates a workflow for a machine learning algorithm to generate a first model for predicting a gas oil ratio using a training data set;

An exemplary standard mud-gas analysis tool 20 is shown schematically in Figure 1.

The tool 20 is coupled to a flow line 10 containing drilling mud returned from a borehole of a well. As discussed above, the drilling mud may be water-based mud or oil-based mud.

The tool 20 comprises a sampling probe 22 disposed with respect to the flow line 10 so as to collect a sample 24 of the drilling mud from the flow line 10. The drilling mud sample 24 is preferably a continuous sample, i.e. such that a portion of the flow of drilling mud within the flow line 10 is diverted through the mud-gas analysis tool 20.

The drilling mud sample 24 is supplied to a gas-separation chamber 26 where at least a portion of the gas carried by the drilling mud is released. The sample of drilling mud may optionally be heated by a heater 28 upstream of the gas-separation chamber 26. Heating the drilling mud sample 24 helps to release the gas from the drilling mud sample 24. Typically, the mud sample 24 is not heated and the temperature are typically ranging from 10°C to 60°C. However, in some implementations, heating is used to 80°C to 90°C.

The released gas 30 is directed from the separation chamber 26 to a gas analysis unit (not shown), while the degassed mud 32 is returned to the flow line 10 or to another location for re-use.

The gas analyser may comprise a total gas detector, which may provide a basic quantitative indication as to how much gas is being extracted from the drilling mud by the tool 20. Total gas detection typically incorporates either a catalytic filament detector, also called a hotwire detector, or a hydrogen flame ionization detector.

A catalytic filament detector operates on the principle of catalytic combustion of hydrocarbons in the presence of a heated platinum wire at gas concentration below the lower explosive limit. The increasing heat due to

combustion causes a corresponding increase in the resistance of the platinum wire filament. This resistance increase may be measured through the use of a Wheatstone bridge or equivalent detection circuit.

A hydrogen flame ionization detector functions on the principle of hydrocarbon molecule ionization in the presence of a very hot hydrogen flame. These ions are subjected to a strong electrical field resulting in a measurable current flow.

The gas analysis device may additionally or alternatively comprise an apparatus for detailed analysis of the hydrocarbon mixture. This analysis is usually performed by a gas chromatograph. However, several other detecting devices may also be utilised including a mass spectrometer, an infrared analyser or a thermal conductivity analyser.

A gas chromatograph is a rapid sampling, batch processing instrument that provides a proportional analysis of a series of hydrocarbons. Gas chromatographs can be configured to separate almost any suite of gases, but typically oilfield chromatographs are designed to separate the paraffin series of hydrocarbons from methane (C_1) through pentane (C_5) at room temperature, using air as a carrier. The chromatograph will report (in units or in mole percent) the quantity of each component of the gas detected.

A carrier gas stream 34, commonly comprising air, may be supplied to the separation chamber 26 and mixed with the released gas 30 to form a gas mixture 36 that is supplied to the gas analysis unit. The carrier gas stream 34 provides a continuous flow of carrier gas in order to provide a substantially continuous flow rate of the gas mixture 36 from separation chamber 26 to the gas analysis unit. Additionally, in the case of a gas analyser comprising a combustor, the use of air as the carrier gas may provide the necessary oxygen for combustion.

In some arrangements, the tool 20 may be configured to detect and/or remove H_2S from the gas to prevent adverse effects that could influence hydrocarbon detection.

In some embodiments, non-combustibles gases, such as helium, carbon dioxide and nitrogen, can be detected by the gas analyser in conjunction with the logging of hydrocarbons.

The following technique seeks to utilise a machine learning algorithm to produce a model that accurately estimates certain properties relating to the

reservoir fluid, in particular the gas-oil ratio and the density of the reservoir fluid, based on the standard mud-gas data and other petrophysical data.

Figure 2 illustrates a workflow 100 for training the machine learning algorithm in order to generate a model for prediction of a gas-oil ratio of a reservoir based on measured standard mud-gas data.

In the following example, an input data set 102 is used as a training data set and comprises data relating to a plurality of reservoir samples.

The input data set comprises reservoir fluid properties data from a large number of reservoir fluid samples. Reservoir samples may be obtained, for example, by downhole fluid sampling. However, other techniques could also be used, for example by taking a sample of well fluid after the well has been completed.

The reservoir fluid properties data should include at least hydrocarbon composition data, which may be either in the form of direct measurements of the concentration of each hydrocarbon component within the sample, typically covering C_1 to C_{36+} hydrocarbons. In some embodiments, the concentration data may be in the form of relative data (e.g. as a ratio of compositions of different hydrocarbons) or may be otherwise normalised. The reservoir fluid properties data may optionally also include concentrations of one or more other constituents within the well.

The reservoir fluid properties data may include one or more derived properties of the reservoir fluid sample. Such derived properties may include the target property to be determined by the machine-learning algorithm, e.g. a gas-oil ratio in this case. Other derived properties may include a density of the fluid.

The reservoir fluid properties data is sometimes referred to as PVT data, as it is commonly obtained in a pressure-volume-temperature (PVT) laboratory, where researchers will employ various instruments to determine reservoir fluid behaviour and properties from the reservoir samples.

The input data set 102 further comprises measured standard mud-gas data for each PVT sample at the same reservoir depth. The measured standard mud-gas data comprises measured hydrocarbon composition data for gas released from the drilling fluid from the sample location.

It will be appreciated that there is a lag-time between the drill bit passing through the sample location, and when the mud reaches the surface and is analysed. However, workers in this field will be familiar with the procedures for

calculating the lag time to determine the depth to which the mud-gas sample corresponds. Therefore, this will not be discussed in detail.

The composition data for the mud-gas preferably comprises data for at least C_1 to C_4 hydrocarbons, and preferably at least C_1 to C_5 hydrocarbons (as is the case in the present example). In some cases, concentrations for up to C_7 or greater hydrocarbons may be included.

The composition data may be stored either as a direct measurement of concentration (e.g. measured in ppm or similar units), or alternatively as a relative concentration (e.g. as a proportion of another hydrocarbon, usually C_1). In some embodiments, the composition data may be normalised.

The measured standard mud-gas data is “raw” mud-gas data, i.e. it has not been corrected for recycling or extraction efficiency. This is important as the use of “raw” mud-gas data will allow the subsequent model to be utilised more widely, where advanced mud-gas data is not available.

The input data set 102 further comprises measured petrophysical data for each PVT sample at the same reservoir depth. The petrophysical data may comprise any one or more of: bulk density, neutron porosity, resistivity data, acoustic data, natural gamma ray, nuclear magnetic resonance data, as well as slowing down time and gamma ray spectroscopy data from pulsed neutron measurements, and the like.

The input data set 102 comprises target data and input data for each sample that passed the screening. The target data corresponds to the desired output of the model. The input data corresponds to the data that will be input into the eventual model.

The target data in this example is a gas-oil ratio, and in this example is the single-flash gas-oil measurement of the sample. As discussed above, this data is stored as part of the reservoir properties data within the initial data set. Alternatively, other measurements of gas-oil ratio may be used, or a gas-oil ratio may be derived from the reservoir composition data, i.e. based on the concentrations of the various hydrocarbons.

The input data is standard mud-gas data, i.e. data indicative of the composition of gases released from the drilling fluid from the sample location, and at least one type of petrophysical data, e.g. bulk density, and neutron porosity.

As mentioned above, the measured mud-gas data comprises “raw” mud-gas data, i.e. it has not been corrected for recycling or extraction efficiency.

Whilst it is not possible to apply a recycling correction after collection of the data, nor is it possible to account for the lack of heating (if heating was not used), it may be possible to apply a retrospective extraction efficiency correction. This is because the composition of the drilling mud for a particular well is usually known, and the extraction efficiency correction factors for that particular drilling fluid can be estimated either from EOS simulation or may be approximated by experiment. Results show that the temperature dependent extraction efficiency correction far outweighs recycling corrections.

Consequently, the mud-gas data used for the input data set 102 preferably comprises standard mud-gas data where an extraction efficiency correction has been applied.

Next, a model generation is performed, in which a model is generated and validated based on the input data set 102.

The input data set 102 is first divided into a training data set 104 and a testing data set 106. The input data set 102 is preferably curated such that at least the testing data set 106 contains data that spans the various classes of the input data set 102 as a whole (e.g. dry gas reservoirs, wet gas reservoirs, oil reservoirs).

Typically, at least 50% of the input data set 102 should be used for training, and at least 10% of the input data set 102 should be used for testing. Common ratios include 50:50, 70:30, 75:25, 80:20, 90:10. However, it will be appreciated that other divisions may be used instead.

Generally the larger the training data set, the more accurate the model will be. However, if too small a test data set is used (or indeed if no test data set is used) then it is not possible to confidently verify the accuracy of the model, e.g. making it difficult to detect an over-fitted model (only accurate for the specific training data).

To generate a model, a machine learning algorithm is provided with the training data set 104, and a set of training parameters to control the machine learning algorithm.

In one example, Gaussian Process Regression and Random Forest were found to be best performing models. However, it will be appreciated that any suitable algorithm may be used, such as Universal Kriging, KMean or Elastic Net algorithms. Those operating within this field will be familiar with the procedures for selecting and utilising a machine learning algorithm. Therefore, this will not be discussed in detail.

Model validation 108, e.g. cross-validation, may then then be performed. During the model validation 108, the model is tested to determine how well it predicts new data that was not used in estimating the model, in order to flag problems such as over fitting or selection bias. Model validation 108 is an optional step.

Cross-validation involves partitioning the training data set 104 into complementary subsets, performing the model fitting using one subset of the training data set 104, and validating the analysis on the other subset of the training data set 104. To reduce variability, most methods use multiple rounds of cross-validation, performed using different partitions, and the validation results are combined (e.g. averaged) over the rounds to give an estimate of the model's predictive performance (e.g. a mean average prediction error, MAPE).

In this example K-fold cross-validation, and particularly 4-fold cross-validation is used. In K-fold cross-validation, the training data 104 is separated in K disjoint subsets (in this case, four), known as "folds". Then, cross-validation is performed by training the model on all of the data except for one fold, and validating the trained model using the fold that was not used for training. The best model is then selected as the model having the best predictive performance, e.g. the lowest MAPE.

A first testing step 110 is then performed, in which the model is tested using the training data set 104 as a whole.

A second testing step 112 is then performed, in which the model is tested using the test data set 106. As discussed previously, this is a curated set of data that is broadly representative of the data as a whole, and was not used during the generation of the model.

The model has been found to predict a gas-oil ratio of the reservoir fluid based on C_1 to C_5 standard mud-gas data and the petrophysical data with MAPE that is close to that achieved using a model based on C_1 to C_5 advanced mud-gas data.

Understanding the quality of the measured mud-gas data is important before performing a fluid property (e.g. gas-oil ratio) prediction because the mud-gas data quality will significantly impact prediction accuracy. The following characteristics of the mud-gas data values have been identified as indicating low quality or unreliable data:

- Large fluctuations of a component within a small depth range.

- First observations after missing measurements.
- C_1 content below a given threshold.
- C_4 or C_5 content below a given threshold.

To quantify the quality of the mud-gas data, the inventors derived a quality control metric (QC metric) which ranged from 0 to 1. High quality mud-gas data would have QC metric value close to 1. If one or more of the above factors are found, then the QC metric would be reduced. Low-quality mud-gas data was indicated by QC metric close to 0. A single numeric quality measure between 0 and 1 can be plotted side-by-side with a predicted fluid property log (as will be discussed below) to visualize the confidence level associated with each prediction, based on mud-gas data quality.

Samples having a higher QC metric correspond closely, whilst samples having a lower QC metric have poor correspondence. Thus, these factors provide a useful indication of the accuracy of a prediction of the gas-oil ratio.

Mud-gas data and petrophysical data are both generated continuously during the drilling process. Therefore, by applying the machine learning model to the mud-gas data and petrophysical data, it is possible to provide, at an early stage of the well installation procedure, a continuous log for the well bore of the predicted reservoir property, e.g. gas-oil ratio or fluid density. This is something that has not been possible previously until much later in the process.

Whilst the above examples have been described in the context of a gas-oil ratio as the target reservoir fluid property, the same technique may also be employed to create a model for estimating other reservoir fluid properties of the reservoir fluid at a sample location, based on measured mud-gas data. Exemplary reservoir fluid properties include a fluid density of the reservoir fluid, either a stock tank oil density or a live reservoir density, a saturation pressure of the reservoir fluid, and a formation volume factor of the reservoir fluid.

Furthermore, a similar technique may be used to train a model to estimate the reservoir fluid composition and corresponding C_{7+} fraction properties. This is advantageous, as this information can be used to for an equations of state (EOS) model calculation. The EOS model for a particular fluid is an expression that describes the relationship between pressure, temperature and volume of the fluid and can be used to predict the phase behaviour of the fluid in order to derive further properties thereof.

It is normally considered necessary to know at least the following properties of the fluid in order to determine the equations of state:

1) The absolute composition of each of the C_1 to C_6 hydrocarbons and the absolute composition of the C_{7+} hydrocarbons combined;
2) The average hydrocarbon density of the C_{7+} hydrocarbons; and
3) The average hydrocarbon molecular weight of the C_{7+} hydrocarbons.

When determining the equations of state for a fluid, the C_{7+} hydrocarbons are usually grouped together because these hydrocarbons usually remain in the liquid/oil phase. A standard C_{7+} characterisation method can split the C_{7+} into multiple pseudo components for EOS calculation.

Although individual fluid property models (like density and GOR) were developed in the first examples, it will be appreciated that a physical model could be generated that would calculate all fluid properties. The EOS model approach in the second example demonstrates a good solution for predicting all reservoir fluid properties.

Whilst preferred embodiments have been described above, it will be appreciated that these have been provided by way of example only, and the scope of the invention is to be limited only by the following claims.

CLAIMS

1. A method of generating a model for predicting at least one property of a fluid at a sample location within a hydrocarbon reservoir, comprising:

providing a training data set comprising input data and target data, the input data comprising mud-gas data and petrophysical data for each of a plurality of sample locations, and the target data comprising the at least one property of the fluid for each of the plurality of sample locations; and

generating a model using the training data set such that the model can be used to predict the at least one property of the fluid at the sample location based on measured mud-gas data and measured petrophysical data for the sample location,

wherein a drilling fluid recycling correction has not been applied to the mud-gas data.

2. A method according to claim 1, wherein generating the model comprises instructing a machine learning algorithm to generate the model using the training data set.

3. A method according to claim 1 or 2, wherein the at least one property comprises a property influenced by the oil-related components of the fluid.

4. A method according to any preceding claim, wherein the at least one property comprises one or more of:

a density of the fluid at the sample location;

a gas-oil ratio of the fluid at the sample location;

a saturation pressure of the fluid at the sample location;

a formation volume factor of the fluid at the sample location;

a concentration of C_{7+} hydrocarbons within the fluid at the sample location.

5. A method according to any preceding claim, wherein the mud-gas data of the training data set comprises measured standard mud-gas data for the sample location.

6. A method according to claim 5, wherein an extraction efficiency correction has been applied to the mud-gas data of the training data set.

7. A method according to claim 5, wherein an extraction efficiency correction has not been applied to the mud-gas data of the training data set, and wherein the training data comprise drilling mud compositional data.

8. A method according to claim 5, 6 or 7, wherein the measured mud-gas data was collected without the use of heating.

9. A method according to any preceding claim, wherein the petrophysical data comprise one or more of:

bulk density;
neutron porosity;
resistivity data;
acoustic data;
natural gamma ray;
nuclear magnetic resonance data; and
gamma ray spectroscopy data.

10. A computer-based model for predicting at least one property of a fluid at a sample location within a hydrocarbon reservoir based on measured mud-gas data and measured petrophysical data for that sample location, the computer-based model having been generated by a method according to any preceding claim.

11. A tangible computer-readable medium storing a computer-based model according to claim 10.

12. A method of predicting a value of a property of a fluid at a sample location within a hydrocarbon reservoir, the method comprising:

receiving measured mud-gas data and measured petrophysical data for the sample location; and

predicting the value of the property of the fluid at the sample location by supplying the measured mud-gas data and the measured petrophysical data to a computer-based model according to claim 10.

13. A method of predicting a value of a fluid property of a fluid along a length of a well through a hydrocarbon reservoir, the method comprising:

predicting a value of a fluid property of a fluid at a plurality of sample locations along a length of a well using a method according to claim 12 for each sample location.

14. A method according to claim 13, further comprising:

displaying, using an electronic display screen, a graph plotting the predicting values of the fluid property against a location of the respective sample location for each of the plurality of sample locations along the length of the well.

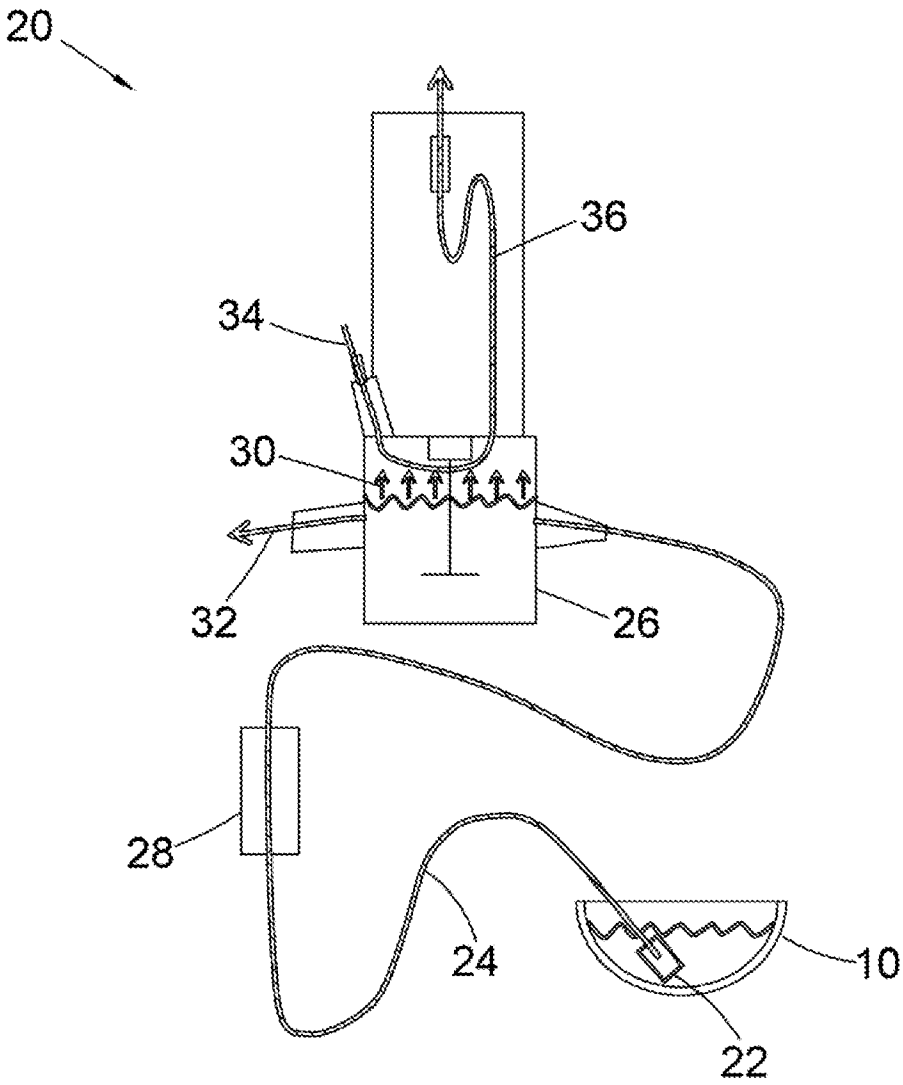


Fig. 1

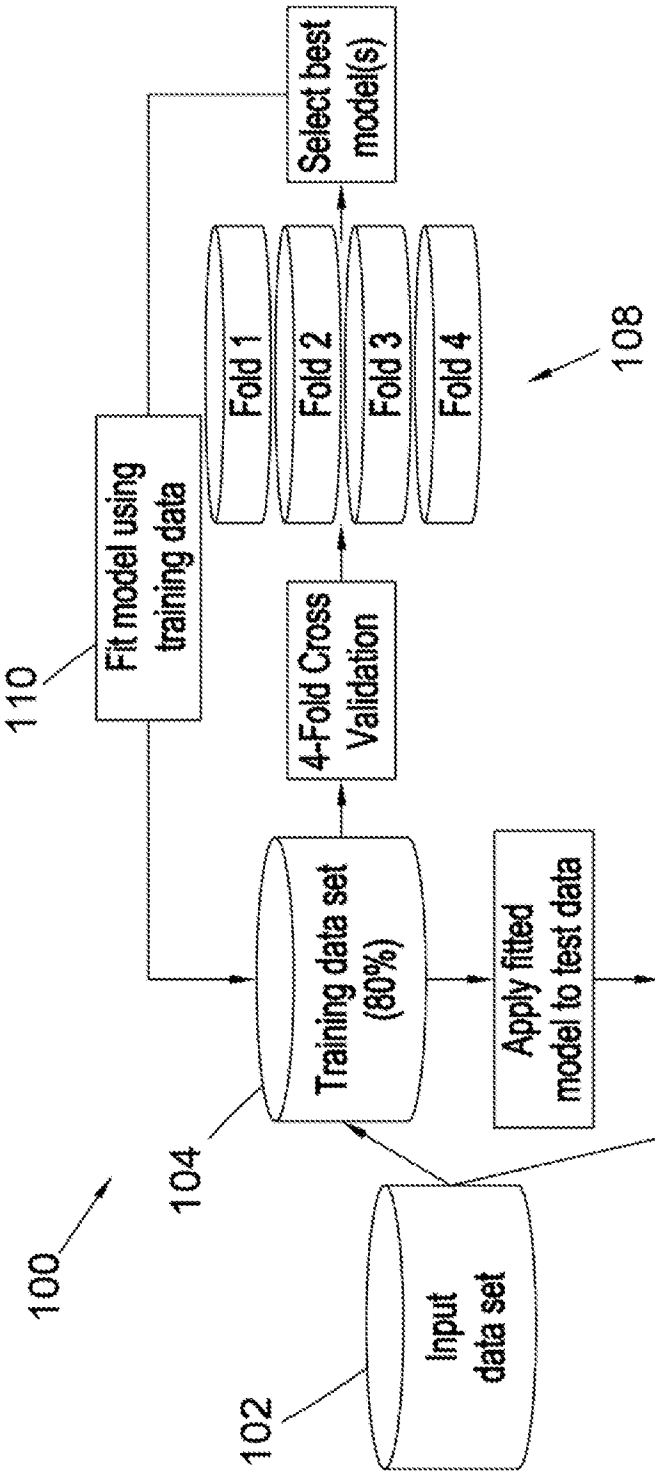

Input data set
102
100
Training data set (80%)
104
Test data set (20%)
106
Apply fitted model to test data
4-Fold Cross Validation
Fit model using training data
110
Fold 1
Fold 2
Fold 3
Fold 4
108
Select best model(s)

Fig. 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NO2021/050158

A. CLASSIFICATION OF SUBJECT MATTER

E21B 49/00 (2006.01)i; G01V 9/00 (2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

E21B; G01V

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

DK, NO, SE, FI: Classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPODOC, WPI, FULL TEXT: ENGLISH, GERMAN, FRENCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2008/0147326 A1 (ELLIS, L.) 19 June 2008 (2008-06-19) Summary of the invention, paras. [0042]-[0050] and the claims	1-14
A	WO 2009/142873 A1 (SCHLUMBERGER CA LTD) 26 November 2009 (2009-11-26) Whole document	1-14
A	US 2014/0025301 A1 (STORM JR, B. H. et al.) 23 January 2014 (2014-01-23) Whole document	1-14
A	WO 2009/035918 A1 (SCHLUMBERGER CA LTD) 19 March 2009 (2009-03-19) Whole document	1-14
A	WO 2014/178883 A1 (HALLIBURTON ENERGY SERV INC) 06 November 2014 (2014-11-06) Whole document	1-14
A	US 2016/0168985 A1 (BETANCOURT-POCATERRA, S. S. et al.) 16 June 2016 (2016-06-16) Whole document	1-14

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
01 September 2021	**01 September 2021**

Name and mailing address of the ISA/XN	Authorized officer
Nordic Patent Institute **Helgeshoj Allé 81,** **2630 Taastrup** **Denmark** Telephone No. **+45 43 50 85 00** Facsimile No. **+4543508008**	**VON UBISCH, Petter** Telephone No. **+45 43 50 85 00**

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2004/0133531 A1 (CHEN, D. et al.) 08 July 2004 (2004-07-08) Whole document	1-14
A	US 2015/0081265 A1 (KAUERAUF, A. I. et al.) 19 March 2015 (2015-03-19) Whole document	1-14
A	WO 2015/047249 A1 (HALLIBURTON ENERGY SERVICES INC) 02 April 2015 (2015-04-02) Whole document	1-14
P	EP 3789580 A1 (SCHLUMBERGER TECHNOLOGY BV) 10 March 2021 (2021-03-10) Whole document	1-14
P	WO 2020/185094 A1 (EQUINOR ENERGY AS) 17 September 2020 (2020-09-17) Whole document	1-14

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/NO2021/050158

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2008/0147326 A1	19 June 2008	WO 2009059132 A1	07 May 2009
WO 2009/142873 A1	26 November 2009	EP 2304174 A1	06 April 2011
		US 2011088895 A1	21 April 2011
		US 8714246 B2	06 May 2014
		US 2014238670 A1	28 August 2014
		GC 4057 A	30 April 2016
		EP 2304174 A4	23 September 2015
US 2014/0025301 A1	23 January 2014	WO 2014014727 A2	23 January 2014
		WO 2014014727 A3	04 December 2014
		CA 2879610 A1	23 January 2014
		GB 2521775 A	01 July 2015
WO 2009/035918 A1	19 March 2009	US 2009071239 A1	19 March 2009
		US 2010132450 A1	03 June 2010
		CA 2698598 A1	19 March 2009
		MX 2010002699 A1	30 April 2010
		EP 2208167 A1	21 July 2010
		RU 2010114583 A	20 October 2011
		BR PI0816685 A2	17 March 2015
WO 2014/178883 A1	06 November 2014	CA 2908626 A1	06 November 2014
		CN 105121779 A	02 December 2015
		GB 2528398 A	20 January 2016
		DE 112013007027T T5	28 January 2016
		US 2016123141 A1	05 May 2016
US 2016/0168985 A1	16 June 2016	EP 3032026 A1	15 June 2016
		WO 2016093842 A1	16 June 2016
		MX 2015016991 A1	06 October 2016
		US 10184334 B2	22 January 2019
		MX 361877 B	18 December 2018
US 2004/0133531 A1	08 July 2004	WO 2004063769 A2	29 July 2004
		AU 2003299896 A1	10 August 2004
		GB 2411503 A	31 August 2005
		BR 0317938 A	29 November 2005
		MX PA05007294 A	01 October 2005
		AU 2003299896 A8	10 November 2005
		GB 2411503 B	13 December 2006
		MX 254626 B	13 February 2008
		WO 2004063769 A3	17 February 2005
		US 8374974 B2	12 February 2013
		CA 2512640 C	31 December 2013
US 2015/0081265 A1	19 March 2015	EP 2848964 A2	18 March 2015
		EP 2848964 A3	27 January 2016
		US 10083258 B2	25 September 2018
WO 2015/047249 A1	02 April 2015	US 2015260703 A1	17 September 2015
		CA 2921151 A1	02 April 2015
		AU 2013402073 A1	03 March 2016
		NO 20160238 A	12 February 2016
		GB 2534703 A	03 August 2016
		IN 201617004867 A	05 August 2016
		AU 2013402073B B2	15 December 2016

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/NO2021/050158

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
				GB	2534703	B	23 August 2017
				BR	112016003320	A2	01 August 2017
				GB	2550483	A	22 November 2017
				US	10001465	B2	19 June 2018
				GB	2550483	B	12 September 2018
				CA	3028914	A1	02 April 2015
				CA	2921151	C	19 February 2019
				CA	3028914	C	28 April 2020
				GC	7541	A	30 April 2018
EP	3789580	A1	10 March 2021	US	2021062650	A1	04 March 2021
WO	2020/185094	A1	17 September 2020	GB	2582294	A	23 September 2020
				GB	2582294	B	14 April 2021