Application of artificial intelligence methods for identifying and predicting complications in the construction of oil and gas wells: problems and solutions

A.D. Chernikov¹, N.A. Eremin¹,²*, V.E. Stolyarov¹, A.G. Sboev³, O.K. Semenova-Chaschina¹,
L.K. Fitsner⁴

¹Oil and Gas Research Institute of the Russian Academy of Sciences, Moscow, Russian Federation
²National University of Oil and Gas «Gubkin University» (Gubkin University), Moscow, Russian Federation
³National Research Center «Kurchatov Institute», Moscow, Russian Federation

Abstract. This paper poses and solves the problem of using artificial intelligence methods for processing Big volumes of geodata from geological and technological measurement stations in order to identify and predict complications during well drilling. Digital modernization of the life cycle of wells using artificial intelligence methods helps to improve the efficiency of drilling oil and gas wells. In the course of creating and training artificial neural networks, regularities were modeled with a given accuracy, hidden relationships between geological and geophysical, technical and technological parameters were revealed. The clustering of Big data volumes from various sources and types of sensors used to measure parameters while well drilling has been carried out. Artificial intelligence classification models have been developed to predict the operational results of the well construction. The analysis of these issues is carried out, and the main directions for their solution are determined.

Keywords: artificial intelligence, machine learning methods, geological and technological research, neural network model, regression model, construction of oil and gas wells, identification and prediction of complications, prevention of emergency situations


Introduction

In a period of increasing competition in the energy market, the task of radically rethinking their activities and approaches to ensuring its efficiency comes to the fore for domestic oil and gas producing and service companies (Abukova et al., 2017; Muslimov, 2017; Dmitrievsky et al., 2019; Dmitrievsky et al., 2020a). Solving this problem requires focusing attention on the key factors affecting the operating activities of companies, the most important of which is the introduction of automation of production processes based on the use of artificial intelligence (AI) systems. Artificial intelligence and machine learning, or computational intelligence, are science and technology aimed at creating intelligent tools, devices, complexes and systems. Its application for solving complex problems in the oil and gas industry is becoming more and more popular and acceptable from an economic point of view (Bobb, 2018; Diakonov et al., 2017; Eremin, 1994; Ivlev et al., 2018; Kabanikhin et al., 2018; Kaznacheev et al., 2016; Djamaluddin et al., 2019).

Artificial intelligence methods are being developed and implemented around the world in an increasing number of applications due to the ability to detect physically hidden processes and phenomena, predictive potential and flexibility. Table 1 shows the application of various artificial intelligence methods in the design and construction of wells based on the analysis of published foreign data (Dmitrievsky et al., 2019; Eremin et al., 2020; Lind et al., 2013; Loermans, 2017; Pichugin et al., 2013; Development of a high-performance automated system..., 2019; Abu-Abed, Khabarov, 2017; Alotaibi et al., 2019; Chen, Guestrin, 2016; Gurina et al., 2019; Kanfar et al., 2020; Kohonen, 1990; Liu et al., 2008; Mayani et al., 2020; Noshi, Schubert, 2018; Rakichinsky, Sledkov, 2014; Singh et al., 2019).

The main advantages of artificial intelligence systems are (Yurchenko, Kryukov, 2018; Kanfar et al., 2020; Li et al., 2019; Gurina et al., 2019; Kohonen, 1990; Liu et
<table>
<thead>
<tr>
<th>Operational Result</th>
<th>Application</th>
<th>Applied AI Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well design</td>
<td>Bit selection</td>
<td>Artificial neural network (ANN)</td>
</tr>
<tr>
<td></td>
<td>Preliminary estimate of the deviation</td>
<td>Generalized regression neural network</td>
</tr>
<tr>
<td></td>
<td>Casing failure prediction</td>
<td>Backpropagation neural network</td>
</tr>
<tr>
<td></td>
<td>Cement Quality / Performance Assessment</td>
<td>Artificial neural network (ANN)</td>
</tr>
<tr>
<td></td>
<td>Offshore drilling platform selection</td>
<td>Hybrid (Backpropagation Neural Network)</td>
</tr>
<tr>
<td></td>
<td>Geosteering</td>
<td>Use case-based machine learning (CBR systems)</td>
</tr>
<tr>
<td>Methodological determination of</td>
<td>Bottom hole assembly monitoring</td>
<td>Artificial neural network (ANN)</td>
</tr>
<tr>
<td>optimal characteristics</td>
<td>Bit wear control</td>
<td>Artificial neural network (ANN)</td>
</tr>
<tr>
<td></td>
<td>Stuck and Load Prediction</td>
<td>Artificial neural network (ANN)</td>
</tr>
<tr>
<td></td>
<td>Vibration control</td>
<td>Artificial neural network (ANN)</td>
</tr>
<tr>
<td></td>
<td>Cleaning the wellbore from cuttings</td>
<td>Backpropagation neural network / multiple linear regression</td>
</tr>
<tr>
<td>Wellbore stability</td>
<td>Hydraulic shock monitoring, loss and leak rate</td>
<td>Artificial neural network (ANN)</td>
</tr>
<tr>
<td>Decision support in</td>
<td>Monitoring and troubleshooting</td>
<td>Backpropagation Neural Network / (Artificial Neural Network-GA) hybrid</td>
</tr>
<tr>
<td>problematic situations</td>
<td>Real time drilling risk assessment</td>
<td>Use case-based machine learning (CBR systems)</td>
</tr>
<tr>
<td>Recognition of troubles, risk</td>
<td>Drilling equipment condition</td>
<td>Artificial neural network (ANN)</td>
</tr>
<tr>
<td>assessment</td>
<td>Determination of permissible operations according to drilling conditions</td>
<td>Use case-based machine learning (CBR systems)</td>
</tr>
</tbody>
</table>

Tab. 1. Application of artificial intelligence methods in the design and construction of wells (Dmitrievsky et al., 2019; Eremin et al., 2020; Lind et al., 2013; Loermans, 2017; Pichugin et al., 2013; Development of a high-performance automated system..., 2019; Alotaibi et al., 2019; Chen, Guestrin, 2016; Gurina et al., 2019; Kanfar et al., 2020; Kohonen, 1990; Liu et al., 2008; Mayani et al., 2020; Noshi, Schubert, 2018; Rakichinsky, Slekov, 2014; Singh et al., 2019).

al., 2008; Mayani et al., 2020; Noshi, Schubert, 2018; Singh et al., 2019):
1. Ability for self-learning, as well as evolutionary development and self-organization;
2. Great potential for accurate analysis of Big historical and industrial databases in order to reveal hidden correlations and unknown patterns compared to traditional methods;
3. Ability to model complex nonlinear processes without any form of establishing a relationship between input and output variables;
4. High efficiency in forecasting, diagnostics, monitoring, condition control and identification of equipment and production processes;
5. Higher predictive accuracy of results than physical and simulation models using linear or nonlinear multiple regression;
6. Ultra-high performance of the neural network after training due to the use of massive parallelism of information processing;
7. Ability to learn from datasets in real time, without writing a program, which is often more cost-effective and practical, especially when changes become critical;
8. Possibility of rapid development using already existing standard software applications, and the necessary specificity can be incorporated into them in the learning process.

Distinctive characteristics of modern AI systems are not only their ability to learn from experience, but also to improve themselves during operation, which is an integral part of the so-called cognitive computing, which dramatically increases the efficiency of decision-making processes when working with big data.

Wells are the main part of fixed assets in the developed oil and gas fields. During well construction, an average of 20–25% of the construction time is spent on dealing with complications and emergencies. The cost of drilling wells tends to rise, and drilling complications are increasingly undesirable. Reducing the loss of working time to eliminate complications and their consequences
is one of the main opportunities for increasing the productivity factor during well construction. The main types of complications are: sticking of the drill string as a result of debris and collapse of unstable rocks, narrowing of the wellbore by crumbling rocks, losses of drilling mud, and gas, oil, and water inflow (kicks). The share of these complications is up to 85% of their total number recorded during the development of oil and gas fields. The types of major complications under consideration lead to long, costly downtime and significant unproductive costs for their elimination and elimination of consequences. The share of the costs of eliminating complications and the emergencies caused by them can be up to 25% of the cost of well construction. Timely prevention of complications and accidents during drilling is an extremely important and urgent task and requires the creation of a set of methods for their early detection using modern artificial intelligence and machine learning systems.

Taking into account the complexity of operations performed in the development of oil and gas fields, the presence of uncertainties associated with geological and geophysical and external conditions, artificial neural networks (ANN) and machine learning methods can be classified as effective tools in the construction of an automated system for preventing complications and emergencies during construction oil and gas wells (AS POAS) (Yurchenko, Kryukov, 2018) (Figure 1).

The adoption of the necessary measures to prevent accidents is possible with reliable prediction of their occurrence based on the analysis of the results of measurements of the parameters of technological processes of well construction. The automated system must perform software processing of measurement results in real time, predict the occurrence of possible complications and issue warning messages. Moreover, in most cases, the occurrence of complications during well construction is determined by a complex set of geological, geophysical and technological parameters and cannot be detected as a result of visual observations by the operator.

For the effective functioning of the AS POAS, taking into account the specifics of scenarios for the occurrence of various types of complications, it must include an integrated complex of AI technologies, which, as a rule, combines auxiliary machine learning methods and classification neural network models. In this case, the architecture of the system should be open at all levels of the organization: structural, functional, data organization and interface (Bakanov et al., 2009). The decisive factor for the construction of an AS POAS based on modern technologies of artificial intelligence is the collection and organization of information, the formation of an integrated database of technical, technological and geological and geophysical data.

**Organization of storage and preparation of data in an automated system for preventing complications and emergencies during the construction of oil and gas wells**

Currently, the international open standard WITSML (Wellsite Information Transfer Standard Markup Language), based on open Internet standards (W3C, SOAP, WSDL, XML), is widely used to exchange data between various services and organizations operating in the oil and gas industry: open interface of application

---

**Fig. 1. Functional diagram of the automated system for preventing complications and emergencies (Yurchenko, Kryukov, 2018)**
programs (Standards Software Development Kit (SDK), Open Subsurface Data Universe). The following data sources were used during the research: the open Dataset of Equinor for the field (https://data.equinor.com/dataset/Volve) – data on the development of 16 wells, archived data of geological and technological studies of domestic companies-developers of West Siberian oil and gas fields basin – 25 wells, and the Central Russian oil and gas basin – 32 wells. As a result of the analysis, data on 38 complications of various types were confirmed and processed. To increase the efficiency of predicting complications during data processing, machine learning methods were used to identify abnormal deviations of parameters from the standard operating modes of drilling equipment.

To expand the area of initial data and its clustering, specially prepared simulation data were used, formed from the results of modeling typical situations of occurrence of complications of specified types on a drilling simulator (Arkhipov et al., 2020; Dmitrievsky et al., 2020). The preparation of initial data for constructing models of neural network calculations consists of the formation and marking of sets of temporary/deep data (WITSML Realtime drilling data) and data of drilling logs (WITSML Daily drilling reports) in WITSML format (WITSML Data Standards), containing information about complications. Such sets can be generated both using the available information for a specific well, and based on archived data containing information about previously drilled wells with similar geological characteristics.

To work with data in WITSML format and form initial sets for constructing models for detecting and predicting complications, a data preparation software module was developed, consisting of a set of service procedures and a client part (Figure 2).

The data preparation module provides the following procedures:
- viewing and preliminary analysis of WITSML Realtime drilling data for each of the wells and selection of wells for use in further calculations;
- interactive parsing of the data structure of drilling logs Daily Drilling Reports WITSML Data;
- viewing records by lithology for each well;
- selection of records for abnormal and emergency situations according to specified criteria.

When performing procedures, the Energistic object data model is used. The information is stored in the form of linked tables that reflect the XML structure of drillReports objects in accordance with the WITSML 1.4.1 specification. File storage and a database based on the MS SQL Server database management system (DBMS) are used to store data.

For the automated selection of data on complications in accordance with the specified criteria (the presence of specified keywords, characteristic changes in technological parameters, etc.), an operator interface has been developed (Figure 3). For clarity, the records displayed on the screen containing information on various types of complications are highlighted in color: “Sticking” – in red, “Fluid loss” – in purple, “Kick” – in green. For the convenience of the analysis, a procedure for the graphical presentation of parametric information contained in the Realtime Drilling files is implemented (Figure 4).

As a result of automated data collection and preparation, repositories are created for untagged (there is no corresponding contextual information, data of drilling logs, etc.) and marked up according to the results of the examination of geological and technological research data, configuration arrays (files) are formed for the formation and training of models, as well as test arrays for their validation, various types of geological and geophysical, technological and contextual information are structured and stored, forming in their totality an integrated AS POAS database.

**Structural organization of an automated system for preventing complications and emergencies**

The stages of the technological cycle of functioning of the AS POAS are shown on the right side of the diagram (Figure 1), according to which three main stages can be distinguished:
- preparation of data and information support for the work of experts to highlight possible complications in unlabeled data;
- formation, training and validation of neural network models and models of machine learning methods, prepared by experts and on unlabeled data sets;
- processing and analysis of real-time drilling data with predicting the possibility of occurrence of complications of the specified types: “Sticking”, “Fluid loss” and “Kick”, formation and display of the appropriate warning messages and recommendations on the prevention of emergency situations on the driller operator’s screen.

The module for the formation and training of models is implemented in the Python language (Keras: The Python Deep Learning library, LightGBM. Python API) and provides the preparation of models used for predicting and preventing emergency situations in drilling support systems. The module implements the functions of assembling classification neural network models.

The generated topology of the neural network of the AS POAS consists of three main layers:
- the first layer is a Multilayered perceptron (MLP);
- then there is a recurrent layer, consisting of 4 neurons of a controlled recurrent unit (Gated Recurrent Units, GRU);
Fig. 2. Block diagram of the data preparation module

Fig. 3. Interface of the block for selecting information related to complications

Fig. 4. An example of a graphical presentation of parametric information on selected drilling problems
- the output layer for solving the classification problem consists of two neurons with the softmax activation function.

The structural diagram of the classification neural network model for identifying and predicting complications of the AS POAS is shown in Figure 5.

Models are formed and trained in accordance with the specified configuration files, which allows you to change the hyperparameters of the models without making changes to the module code. The trained models are used as output data, which are saved as separate files with their own name in the hdf5 format and include the following structure:
- the topology of the model, which allows you to reproduce the trained model;
- customized model weights;
- the state of the optimizer.

The real-time data processing module ensures the integration of models into the AS POAS and performs the following functions:
- loading of trained models of prediction of emergency situations and preprocessing parameters in accordance with configuration files;
- transfer of the obtained vectors of parameters to the module for processing real data and obtaining the predicted values of the models from the accumulated window of parameters;
- logging (recording system information) of the models.

The operator interface of the AS POAS for a detailed parametric analysis of the causes of complications of the “Fluid loss” type during the drilling operation is shown in Figure 6.

On the left side of the screen, the timeline and the technological operations and modes being performed are displayed, and on the right, a graphical representation of the change in time parameters to identify the specified types of complications.

A simplified interface has been developed for the driller operator with automatic determination of the predicted probabilities of complications and displaying warning messages and alarms on the screen in case of exceeding a predetermined threshold (Figure 7).

On the left side of the screen, the values of technological parameters are displayed in real time, with the possibility of selection by the operator, and on the right, the values of the probabilities of complications, calculated according to the predicted and actual parameters of geological and technological studies (GTI). Warnings about the possibility of complications are displayed on the operator’s screen in the form of arrow indicators, as well as time scales for the probability of occurrence of complications of specified types with color alarms: green when there is no threat and red when the threat probability is greater than 0.5.

Accuracy and f1 score metrics were used to calculate the accuracy. The Accuracy score was calculated as the ratio of the number of moments in which the reference and predicted marks coincided to the total number of moments. To calculate the f1 score, the number of points correctly assigned (TP) to it, incorrectly assigned (FP) and incorrectly unassigned (FN) was first calculated for each class. After that, the total value of accuracy was calculated, equal to TP/(TP + FP), and completeness – TP/(TP + FN). Moreover, each example was taken with a weight depending on the representativeness of the class. The choice of quality metrics was based on the composition of the data used and the methods used for their processing.

Based on the results of testing the classification neural network model, the following predictive accuracy of

Fig. 5. Block diagram of the classification neural network model for predicting drilling problems
various types of complications was obtained: “Kick” – 96%; “Fluid loss” – 79%; “Sticking” – 87%.

**Problems and main directions of their solution**

One of the main problems in the development and implementation of AI systems in the domestic oil and gas industry is the problem of data availability, organization of their collection, structuring, storage and distribution to consumers. The main obstacles in solving this problem are departmental barriers and the protectionism of such oil and gas companies - operators of the fields.

Currently, the oil and gas industry has made significant strides in improving drilling performance by adding high-tech downhole tools and sensors, redefining classic drilling procedures, and utilizing state-of-the-art surface rig systems. Progress in optimizing the construction of oil and gas wells based on the use of constantly available historical and operatively obtained geological, geophysical and technological data turned out to be insignificant. Equipping drillers and engineers with specific and fast solutions based on the implementation of artificial intelligence technologies for modeling and processing field data in real time is now the key to increasing operational efficiency and reducing costs in the construction of oil and gas wells, ensuring operational and environmental safety.

The main development vectors in this direction are the following:
- creation of modern interactive environments to ensure the collection, systematization and analysis of...
all operational information in real time and providing on this basis proactive management of the process of construction of wells (fields);
- automation of production processes based on the introduction of artificial intelligence systems;
- creation and implementation of new AI tools for remote monitoring and management of operational activities;
- use of integrated cross-functional performance indicators for AI systems and the company’s overall performance, which allow optimizing all stages of their operational activities.

The main vectors of development in this direction are the following:
- creation of modern interactive environments to ensure the collection, systematization and analysis of all operational information in real time and providing on this basis proactive management of the process of construction of wells (fields);
- automation of production processes based on the introduction of artificial intelligence systems;
- creation and implementation of new AI tools for remote monitoring and management of operational activities;
- use of integrated cross-functional performance indicators for AI systems and the company’s overall performance, which allow optimizing all stages of their operational activities.

Due to the scale and complexity of this task, it cannot be solved without the introduction of modern methods of artificial intelligence and innovative information technologies with direct participation in projects of it and service companies, as well as specialized scientific organizations.

An example is the activity of Equinor, which became one of the founders of the OSDU (Open Subsurface Data Universe™) initiative, a global collaboration between most of the world’s largest operators and service companies in defining standards for the open data architecture for subsurface resources, creating open data banks of geological and technological information generated from the design and construction of wells to their support at all stages of the life cycle. When designing new wells and fields, accumulated data integrated on the basis of cloud technologies is used.

The increasing use of artificial intelligence methods to improve the efficiency of oil and gas well construction leads to an exponentially growing number and greater specialization of artificial neural network models that are configured to solve various target objectives: development planning, optimization of technological modes, forecasting various types of drilling complications (drilling column sticking, lost circulation, kicks of reservoir fluid, bit wear, etc.) of oil and gas wells.

Currently, in the field of application of information technologies in the oil and gas industry, the image of universal information systems has developed – a single digital platform with the ability to create an API programming interface for interconnection with the combined resources of the developer company and consumers of different levels. In contrast, in the field of AI, there are no unified approaches for combining specialized systems, methods and solutions (ANN, machine learning methods, decision support systems, expert systems), based on a single digital AI platform, which allows working with Big amounts of unstructured data.

At the same time, the main problematic issue is the integration of specialized models of artificial neural networks and machine learning methods into a single system that provides an effective solution to a given set of problems under conditions of a priori uncertainty associated with specific geological and geophysical, technical and technological conditions and factors. With regard to the problem of implementing a systematic approach when introducing AI methods to solve problems of increasing the efficiency of construction of oil and gas wells, the question arises: on what basis is it possible to combine such heterogeneous models as forecasting complications that differ in nature: sticking, lost circulation, gas, oil and water kicks, etc.

Therefore, the main direction for solving this problem in the oil and gas industry is the aggregation of heterogeneous software algorithmic complexes (SAC) for AI into a single system. The aggregation of heterogeneous SAC AI is understood as their integration into a self-learning system based on unified AI self-organization algorithms that form a single Smart environment (platform) in the information and control space of the oil and gas industry technological processes.

This paradigm of a self-organizing AI-System, as the latest concept of dynamic adaptation to the conditions of a specific oil and gas production, will allow to ensure the integration of promising oil and gas technologies based on the implementation of a Smart platform for aggregating heterogeneous SAC for AI. The development of new Smart AI technologies for the oil and gas industry is planned to be implemented as part of the creation of an Integrated Center for Oil and Gas Technologies based on an aggregated artificial intelligence system, the creation of which will allow moving to a qualitatively new technological level of solving the entire complex of problems in the oil and gas industry.

Conclusion
In the course of the research, the structure and parameters of the optimal configuration of models of neural networks and machine learning methods were
determined, an experimental sample of a software complex was developed, designed to ensure the functioning of an automated system for preventing complications and emergencies during the construction of oil and gas wells.

The research made it possible to identify a number of common problematic issues in the implementation of artificial intelligence technologies in the oil and gas industry and to determine the main directions for their solution within the framework of the development and implementation of a single integrated digital AI platform and improvement of methods for streamlining processing of Big volumes of geological-geophysical and real-time technological data.

Acknowledgments

The article was prepared as part of the work of the Federal Target Program «Research and Development in Priority Areas of Development of the Scientific and Technological Complex of Russia for 2014–2020» on the topic: «Development of a high-performance automated system for preventing complications and emergencies during the construction of oil and gas wells based on permanent geological and technological models of fields using artificial intelligence technologies and industrial blockchain to reduce the risks of geological exploration, incl. on offshore projects» under the Agreement with the Ministry of Science and Higher Education of the Russian Federation on the allocation of a subsidy in the form of a grant dated November 22, 2019 No. 075-15-2019-1688.

References


Development of a high-performance automated system for preventing troubles and emergencies during the construction of oil and gas wells based on constantly operating geological and technological models of fields using artificial intelligence technologies and industrial blockchain to reduce the risks of geological exploration, including on offshore projects. (2019). Report. Oil and Gas Research Institute of the Russian Academy of Sciences. (In Russ.)


About the Authors

Alexander D. Chernikov – Leading Researcher, Cand. Sci. (Engineering), Oil and Gas Research Institute of the Russian Academy of Sciences
3, Gubkin st., Moscow, 119333, Russian Federation

Nikolai A. Eremin – Dr. Sci. (Engineering), Chief Researcher, Deputy Director, Oil and Gas Research Institute of the Russian Academy of Sciences; Professor, National University of Oil and Gas «Gubkin University»
3, Gubkin st., Moscow, 119333, Russian Federation
E-mail: ermn@mail.ru

Vladimir E. Stolyarov – Deputy Head of the Analytical Center for Technology Forecasting in the Oil and Gas Industry, Institute of Oil and Gas Problems of the Russian Academy of Sciences
3, Gubkin st., Moscow, 119333, Russian Federation

Alexander G. Sboev – Leading Researcher, Cand. Sci. (Engineering), National Research Center «Kurchatov Institute»
1, Ak. Kurchatov pl., Moscow, 123098, Russian Federation

Olga K. Semenova-Chashchina – Leading Engineer, Cand. Sci. (Engineering), Oil and Gas Research Institute of the Russian Academy of Sciences
3, Gubkin st., Moscow, 119333, Russian Federation

Leonid K. Fitsner – Leading Engineer, Cand. Sci. (Engineering)
Oil and Gas Research Institute of the Russian Academy of Sciences
3, Gubkin st., Moscow, 119333, Russian Federation

Manuscript received 21 July 2020;
Accepted 10 September 2020;
Published 30 September 2020