Please fill in the name of the event you are preparing this manuscript for.	Gas & Oil Technology Showcase and Conference 2019
Please fill in your 6-digit SPE manuscript number.	SPE-123456-MS
Please fill in your manuscript title.	AI-based ESP Optimal Control Solution to Optimize Oil Flow Across Multiple Wells

Please fill in your author name(s) and company affiliation.

Given Name	Surname	Company
Dmitry	Krikunov	Zyfra Oy
Semen	Kosyachenko	Zyfra Oy
Dmitry	Lukovkin	Zyfra Oy
Alexander	Kunchinin	JSC Slavneft-Megionneftegaz
Roman	Tolmachev	Zyfra Oy
Roman	Chebotarev	Zyfra Oy

Abstract

Objectives/Scope: This paper describes the results and operation principles of the developed AI-based solution for optimization of operating modes of electrical submersible pumps (ESPs). The solution allows exploration and production companies to increase the oil flowrate from ESPs without additional capital investments. The solution is based on optimal adjustments of ESP modes in a certain range of motor frequencies and control types.

Methods, Procedures, Process: The hypothesis behind the proposed method is that there is a certain ESP operation mode in each period of time which ensures the maximum daily oil and gas flowrates in the long term. The proposed solution uses statistical methods to ensure data sufficiency and stability and employs custom algorithms to conduct problem-specific data preprocessing. The "well-reservoir" model is the combination of physical (i.e. the ESP pump) and machine learning (ML) models. The ML part of the solution was designed for the short- and long-term predictions of well-reservoir system response. However, the model faced constraints such as current frequency, load and amperage.

Results, Observations, Conclusions: The following results were delivered: data collection and preprocessing pipeline was built, a hybrid physical-ML prediction model was designed and trained, and a numeric optimization model which uses outputs of the hybrid prediction model was designed to recommend multi-well operation modes. The solution has been successfully deployed and tested for 6 months in 500 oil wells in Western Siberia. The designed ESP optimal control solution enhanced the efficiency of oil extraction by boosting oil well production rates by 1.5% with no additional capital investments, as proven during the production phase. Thus, the proposed "reservoir-well-pump" system modeling approach showed efficiency and superiority over the previous control methods used by oilfield operators.

Novel/Additive Information: To the authors' best knowledge, the present paper is the first work covering research, development and production testing of hybrid physical-ML methods in the optimal control of ESPs with respect to optimization of oil production. Application of known statistical methods to the problem of provision of reliable data to the model is additive to the current body of knowledge.

Introduction

The renewable energy industry is growing rapidly, however, oil has been the world's major commercial energy source for many decades and the consensus view is that it will maintain this leading role well into the 21st century. Oil currently heads the list of energy resources with 31% share. Worldwide oil consumption increased by 1.3% in 2018 and there is a possibility that it will reach 100 million bpd, an all-time high, in 2019 (Exarheas 2019), as reported by IEA (International Energy Agency).

The extraction of crude oil to the surface is a complex and energy-intensive process. Globally, more than 95% of oil wells work with deficient oil reservoir pressure and utilize artificial lift (Editorial 2012).

The effective operation of the downhole equipment is critically important due to the growing demand for oil as an energy resource and the widespread use of an artificial lift. The paper describes operation principles of the developed recommendation solution for the optimization of electrical submersible pumps (ESPs) operating mode.

Generally, ESPs with variable speed drives use control stations that regulate the pump current frequency. Modern ESPs can be adjusted to a variety of current frequencies, from 40 Hz to 60 Hz. A change of the pump current frequency leads to the pump change of impellers' rotation speed. An adjustment in the pump impellers' rotation velocity changes the kinetic energy of extracting liquids in the diffuser. A change in the kinetic energy leads to the change in the pump pressure and flowrate.

However, there are different factors which make determination of the most efficient operation mode rather difficult. Some of these factors are formation pressure changes over time, re-injection of produced water rate and others.

Engineers or industrial software calculate operating modes recommendations at this moment **(Seems to be too much of Russian English).** Typically, they use old-fashioned and basic algorithms without any possibility of non-evident patterns recognition. In that regard, the operating mode calculated recommendation are not always optimal **(What?).**

Thus, the authors hypothesized potential improvement in optimization techniques using machine learning (ML) technology. For this, the objectives for solution development include reduction of reservoir feedback uncertainty and more precise evaluation of oil well production using ML technology in order to increase the oil production without capital investments.

Pilot area

The developed AI-based recommendation system (RS) was implemented at one of the onshore Siberian oil fields in 2018. The oil field is in the plateau period state (average water cut is 42%) and includes 700 wells (480 producing and 220 injection wells). Of the total producing wells, 360 wells work at a constant mode (i.e. constant engine frequency of ESPs), while the remaining 120 wells work at the variable mode (i.e. engine frequency of ESPs is constant at 50 Hz, but the working periods alternate with non-working periods). The average flowrate of producing wells is 56 tons of crude oil per day. In this oil field, 6 oil reservoirs are developed with the depth of 2600-3200 meters.

Solution overview

The process starts with the data preparation module to remove inaccurate measurements, sensor surges, etc. Auxiliary aggregate parameters are also devised in this stage. One of the main parameters is the true produced flowrate, which is particularly important for wells operating in short-term production mode.

The data assurance module then checks whether the data are sufficient for subsequent construction of algorithms and ML models. It also examines the wells to determine the accuracy and resiliency of the data for devising recommendations. Recommendations can be produced only for wells whose operation is described in some detail and whose operating mode can be regarded as stable. The criterion for data sufficiency is the availability of all critically important parameters (CIP) at a given frequency; the CIPs include motor temperature, current, intake pressure, cable insulation resistance, etc. The stability of the data is assessed using basic time sequence prediction techniques. If the variance in assessments by these methods using historical values does not exceed the specified limits, the well is deemed to be operating in a stable mode.

The third module simulates the 'well-formation' system. The element of greatest interest within this module builds a predictive model for fluid flowrate and the well CIPs following any alteration to the well's control mode. The predictive model consists of two parts: an individual forecast for each well and an overall forecast for all wells in the field. Several ML models are used, including random forest, linear regression, and others. The resulting forecast for each prediction parameters for each well is calculated as a linear weighted sum of the individual and overall forecasts with a derived confidence factor. The confidence factor is calculated based on data sufficiency for the description of the well's operation.

Data processing

To solve the task of ESP operation optimization, the available parameters from sensors (Table 1) were used. The period of observation and information storage of the current oil field is more than 10 years. However, during the research, it was found that the inclusion of measurements older than the last 1.5 years does not significantly improve the recommendation system. Thus, these measurements were not employed in the RS learning process.

Parameter	**Measurement frequency**
Engine frequency (Hz)	Once in a minute
Engine temperature (℃)	
Engine active power (kW)	
Annulus pressure (kg/m^2)	
Tubing pressure (kg/m^2)	
Current (A)	
Voltage (V)	
Manifold pressure (kg/m^2)	
Engine pressure (kg/m^2)	
Head (meters of the water column) (m)	
ESP on/ESP off	Once in 5 minutes
Flowing level (m)	Once in an hour
Crude oil density (kg/m^3)	Once in a day
Water cut (%)	
Crude oil production (m^3/d)	
Rated engine power (kW)	Once in a half year or rare
Reservoir pressure (kg/m^2)	
Rated engine productivity (m^3/d)	
Well depth (m)	
Well prolongation (m)	

Table 1. Parameters used in the solution

As shown, **Ошибка! Источник ссылки не найден.** includes measured as well as calculated parameters that were achieved by mathematical modeling of downhole equipment operation. Specifically, annulus pressure was calculated using the following model:

$$p_{annulus} = \frac{(w + (100 - w) \cdot \rho/100) \cdot 1000 \cdot 9.81 \cdot D}{101325} \tag{1}$$

where w – water cut, ρ – crude oil density, and D – well depth. (WC – water cut usually)

Besides, the pump head model was implemented (Lea et al. 2019):

$$H_{\text{curr}} = H_{\text{def}} \cdot \left(\frac{f_{\text{curr}}}{f_{\text{def}}}\right)^2 \tag{2}$$

where, H_{def} – head obtained with default ESPs frequency (ESP characteristic), f_{curr} – current ESPs frequency, and f_{def} – default ESP frequency.

It should be noted that the cooperation of the data scientists and the domain experts played a great role in this project. The domain experts provided simple and high accuracy models of physical processes related to the operation of the wells. These models significantly increased the RS accuracy.

We can see that most of parameters affect the predicted parameter with a certain delay. As such, to achieve better prediction quality, a lag parameter was generated from one of the initial parameters. It is the same parameter as the initial, but the value of the lag-parameter is equal to the value of the initial parameter in the previous time moment (shifted by the lag), such that $x_{lag}(t) = x(t - t_{lag})$. Such kinds of parameters lead the statistical model to a better "understanding" of how predicted parameters are related to x parameters. Furthermore, smooth-lag parameters were used. The smooth-lag parameters are the lag-parameter smoothed by the selected time base: $x_{sm\_lag}(t) = \underset{[t-t_{sm}, t+t_{sm}]}{\text{mean}} x_{lag}(t)$.

During the project, the data mismatch problem was also solved. In some periods, there were non-zero oil flowrates for some wells when the pump was switched off. Clearly, that was an obvious mistake in recording the pump state. The ML prediction model of the ESPs state was implemented for these periods. This model was trained on the selected data when there was no mismatch between the parameters. Thus, function f can be developed to evaluate the mismatch between key parameters in "mismatching" periods:

$$d = f(x) \tag{3}$$

where, d denotes the predicted value for the chosen parameter, x - the parameters used in the prediction model f.

Further, handling the historical data, the predicted value d can be compared with the historical value d_{real}, and the predicted value d can be employed instead of d_{real} in periods of large mismatches between d and d_{real}:

$$d_{real} = \begin{cases} d_{real} & , |d_{real} - d| < s \cdot \text{std}(d_{real}) \\ d & , |d_{real} - d| > s \cdot \text{std}(d_{real}) \end{cases} \tag{4}$$

where, $\text{std}(d_{real})$ is the standard deviation of d_{real} values in an observation period, and s is the acceptable mismatching constant based on the domain expert's physical model.

The proposed approach is not always applicable, and it can be difficult to build a predictive model for certain parameters. In such kind of cases with abnormal values, these periods were excluded from consideration. Moreover, the periods when the key parameters were not observed were also excluded.

Furthermore, the RS implementation was associated with insufficient frequency of the crude oil flowrate measurements. In the field under consideration, group metering stations are used to measure the flowrates. In this case, the flowrate meter is not related to specifically one well but measures flow from several wells. The oil flowrate measurement periods are not associated with the pump condition. The meter measures oil flowrate from the specific well for 1-2 hours, then it switches to another well, and so on. Let us imagine that the well works constantly but would be stopped for a period of meter measurement, so the zero daily flow rate for this well would be obtained; clearly, this is a mistake. To eliminate this mistake, the calculation of a "true" (calculated) flowrate was implemented. The algorithm used for this calculation is quite complex and requires a separate report. So, a simplified model is used in the case when only one measurement per day was made and the engine frequency of ESPs was kept constant:

$$q_{real} = q_{meas} \cdot \frac{\Delta t_{meas}^{on}}{\Delta t_{meas}} \cdot \frac{24}{t_{24h}^{on}} \tag{5}$$

where, q_{real} is the «true» daily oil flowrate, q_{meas} is the measured daily flowrate, Δt_{meas} is the time period of flowrate measurement, Δt_{meas}^{on} is the duration of period when a pump is switched on during oil flow measurement time , and t_{24h}^{on} is the duration of period when a pump is switched on during last 24 hours.

Problem

The main purpose of this work is to increase the volume of lifted crude oil, avoiding the risk of well drying and preventing well accidents.

To solve this problem, it is necessary to determine acceptable limits of safe and efficient functioning of wells. The conducted statistical investigations and the domain experts' knowledge allowed us to develop a table of restrictions (**Ошибка! Источник ссылки не найден.**). Following the specified limits guaranteed safety and constant oil production process in our case.

Parameter	Limit
Input pressure	> 28 (kg/m^2) (Pa or bar is better)
Engine temperature	< 110 (℃)
Isolation resistance	> 8 000 Om
Tubing pressure	< $40 \cdot 10^4$ (kg/m^2)
Current	< 200A

Table 2. CIPs and their respective limits

The ESPs' engine frequency is the only control parameter in this maximization task that can variate between 40-60 Hz.

Hence, the task can be reformulated as follows: for each ESP, it is necessary to determine the maximum engine frequency at which all the parameters from Table 2 will stay within their limits.

The given formulation assumes the construction of a predictive model for each parameter from Table 2. During this project, several attempts were made to build such predictive models. Unfortunately, all of them were unsuccessful. The general problem of predictive model construction is an unknown reservoir state at a perforation area. The reservoir pressure is usually estimated twice in a year; furthermore, the reservoir pressure cannot fully describe crude oil flowing from the reservoir to the wellbore.

Reservoir pressure processing

Let us look at the simple measurement model of the "reservoir pressure" parameter:

$$p_{meas}(t) = p_{real}(t) + \delta\tilde{p}(t) + \delta\bar{p}, \tag{6}$$

where, $p_{meas}(t)$ is the measured value of reservoir pressure at the moment t, $p(t)$ is the true reservoir pressure at the moment t, $\delta\tilde{p}(t)$ - the high-frequency (random) part of a measurement error, and $\delta\bar{p}$ - the constant (systematic) part of a reservoir pressure measurement error. The $\delta\bar{p}$ part is the main error part that changes slowly, but because the reservoir pressure measurements are rare, it is difficult to estimate the exact reservoir pressure at any given time. In order to get rid of this component of error, the approach "transition from the absolute measurement to the incremental measurement" was implemented. This method consists of choosing a certain time base Δt, where $\delta\bar{p}$ can be considered as a constant:

$$\Delta p_{meas} = p_{meas}(t) - p_{meas}(t+\Delta t) = \\ p_{real}(t) + \delta\tilde{p}(t) + \delta\bar{p} - p_{real}(t+\Delta t) - \delta\tilde{p}(t+\Delta t) - \delta\bar{p} \tag{7}$$

Since $\delta\bar{p}$ is constant within the selected time base, it is possible to reduce this component, so we get:

$$\Delta p_{meas} = p_{real}(t) - p_{real}(t+\Delta t) + 2 \cdot \delta\tilde{p} \tag{8}$$

As we can see, we managed to eliminate the constant error component. Although the high-frequency component has doubled, it is not significant because this part is very small. It is important to choose the correct time base Δt. On one hand, too small Δt leads to an increase in $\delta\tilde{p}$ error; on the other hand, a big Δt value leads to an additional error related to the difference between $\delta\bar{p}(t)$ and $\delta\bar{p}(t+\Delta t)$.

The transition from the absolute values of parameters to their increments for all other parameters was implemented in the same way.

Prediction model

Assuming the considerations above, we used model (*9*) for building predictions instead of the classical model (3).

$$d(t) = d(t-\Delta t) + \Delta d_{pred} = d(t-\Delta t) + f_{\Delta}(x, \Delta x, \Delta f), \tag{9}$$

where, $d(t)$ is the predicted value, $d(t-\Delta t)$ is the value of the same parameter in time $t-\Delta t$, Δd_{pred} is the prediction of changing parameter value, and $f_{\Delta}(x, \Delta x, \Delta f)$ is the predictive function that implements the prediction depending on current state x, last changing - Δx and changing engine frequency - Δf.

From the entire history of measurements, the moments when the ESP engine frequency changed were selected to build a predictive model f_{Δ}. The ML methods - supervised learning - were used to implement the predictive model. The measurement value before switching was used as x, and the value parameters after switching were d.

Unfortunately, there were not enough changepoints for every ESP in history to ascertain how engine frequency depends on key parameters. The common predictive model was built for these cases, so the full predictive model consists of two parts:

Common model – It was learned using all changepoints for all ESPs. This model is based on the random forest method (Breiman 2001), with a lot of branches and a small depth for each tree.

Individual model – It was learned using only current ESP data. The linear regression model (Murphy 2012, chap. 7) was used in this case.

The final prediction was built as follows:

$$d = d(t-\Delta t) + \alpha \cdot f_{\Delta}(x, \Delta x, \Delta f)_{ind} + (1-\alpha) \cdot f_{\Delta}(x, \Delta x, \Delta f)_{comm} \tag{10}$$

where, $f_{\Delta}(x, \Delta x, \Delta f)_{ind}$ is the individual prediction, $f_{\Delta}(x, \Delta x, \Delta f)_{comm}$ is the common prediction, and α – the confidence of individual prediction ($\alpha \in [0,1]$). The algorithm of α calculation is not given here because of its complexity. Let us just say that α is close to 1 if we are confident in individual model prediction, and it is 0 otherwise.

Recommendation system

In this section, the working procedure of the recommendation is described. First, the prediction model builds a prediction for each parameter from Table 2 for each well. The prediction contains an error. Let us try to estimate this error by selecting all the moments when the ESP frequency changed. Let us consider equation (8) one more time but with the prediction error:

$$d(t+\Delta t) = d(t) + \Delta d_{pred}(t+\Delta t) + \delta\Delta d_{pred} \tag{11}$$

where, $d(t+\Delta t)$ is the parameter value after engine frequency changing, $d(t)$ is the parameter value before engine frequency changing, $\Delta d_{pred}(t+\Delta t)$ is the prediction of changing value after engine frequency changing, and $\delta\Delta d_{pred}$ is the prediction error.

It is important for $d(t) + \Delta d_{pred}(t+\Delta t) + \delta\Delta d_{pred}$ to stay within the specified range (**Ошибка! Источник ссылки не найден.**) in order to maintain safe functioning of mining equipment. Consequently, it is necessary to evaluate $\delta\Delta d_{pred}$. Let us take all engine frequency switches and build a prediction of the input pressure changes and take all differences between Δd_{real} and Δd_{pred} that consist of a prediction error - $\delta\Delta d_{pred}$. This will result in the histogram in Figure 1.

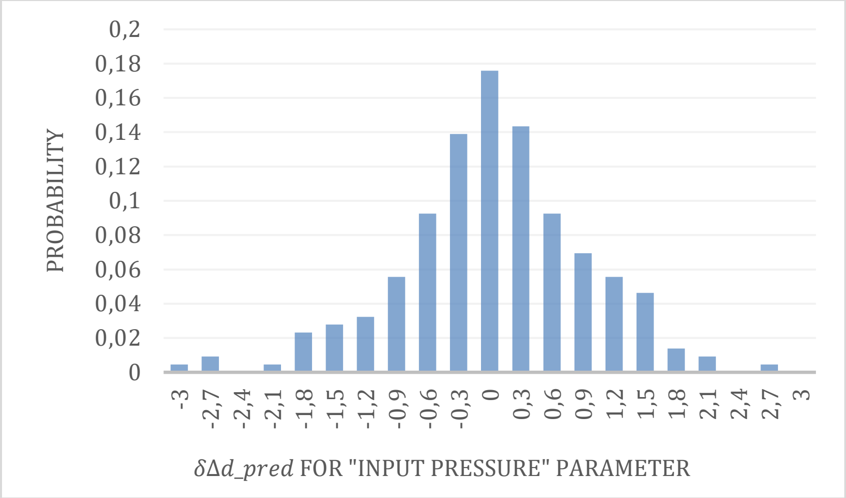

Figure 1. Distribution of the prediction error

On analyzing input pressure, the cases when $\delta\Delta d_{pred} < 0$ are the most interesting. In such cases, there is a risk to cross the bottom limit ($28 \cdot 10^4$ kg/m^2) (make bars), and it can lead to well drying. Let us look at the left tail of the histogram in Figure 1. It can be seen (honestly, it is impossible to see but could be calculated) that the built prediction model works quite good for the input pressure parameter. The prediction model with a probability of 99% does not produce error of more than $2.4 \cdot 10^4$ kg/m^2(make bars). So, we know if an input pressure predicted value $d(t) + \Delta d_{pred}(t + \Delta t)$ will exceed $30.4 \cdot 10^4$ kg/m^2(make bars), then the limit for the input pressure will be accomplished with probability of more than 99%.

All the other key parameters were processed using the same logic. Note that the values of the key parameters do not change immediately. The engine temperature, for example, changes in 2-3 hours after the engine frequency change. After that time, it becomes asymptotic and remains almost constant until the next engine frequency change. Moreover, input pressure keeps changing in 5-10 days after ESP engine frequency change and becomes constant only after that. So, it is important to choose true Δt to get a correct evaluation of $\delta\Delta d_{pred}$.

Thus, we have a method to estimate an interval of the values of each CPI for the given engine frequency with a specified confidence. Then we perform a grid search for engine frequency values within the range from 40 to 60 Hz with a step of 1 Hz for each CPI. As a result, we obtain a maximum value of the ESP engine frequency which can be used with a given confidence that none of the key parameters exceed its operational limits.

Unfortunately, ESP cannot be effective enough in wells with low reservoir pressure. In some cases, the electricity cost spent on crude oil recovery can be more than the crude oil cost. To avoid inefficient energy consumption, the following criterion is applied:

$$\left(V_{pred}^{oil} - V_{curr}^{oil}\right) \cdot (1 - w) \cdot z_{oil} - \left(E_{pred} - E_{curr}\right) \cdot z_{electricity} > 0, \tag{12}$$

where, V_{pred}^{oil} is the prediction of daily production of crude oil, V_{curr}^{oil} is the current daily production of crude oil, z_{oil} is the oil price, w is the water cut, E_{pred} is the prediction of daily energy consumption that will be used in predicted mode, E_{curr} is the current energy consumption, and $z_{electricity}$ is the energy unit cost.

Using this criterion, RS gives the recommendation only if the increase in electricity consumption does not exceed the increase in profit from crude oil production.

Results

The real example of recommendations proposed by RS can be seen in **Ошибка! Источник ссылки не найден.**.

Well id	Current engine frequency	Recommended engine frequency	Current crude oil production	Predicted crude oil production	Recommendation	Comment
…	40.02	40.02	27.78	27.78	Optimal mode installed	
…	49.17	48.17	205.76	196.19	Decrease engine frequency	Alert!!! «input pressure to low». risk of well drying out
…	52	54	316.65	348.89	Increase engine frequency	
…	50.04	51	38.66	42.98	Increase engine frequency	
…	50.07	50.07	48.18	47.57	Optimal mode installed	
…	49.08	50	50.97	46.29	Increase engine frequency	

Table 2. Example of RS recommendations (exported to spreadsheet file).

As you can see in **Ошибка! Источник ссылки не найден.**, the RS gives recommendation not only to increase engine frequency but also to decrease engine frequency to avoid risky situations (e.g. engine overheating, etc.) for different wells. In particular, the second row of Table 3 contains the recommendation to decrease the frequency of ESP in order to prevent the risk of well drying.

It is worth mentioning that the proposed RS could be applied (and is applied in the real life) not only for ESPs that work in constant mode but also for ESPs that work in the short-term well operation (STWO) mode (Kuzmichev 2012). In the STWO mode, the ESP engine always works with the same constant frequency (50 Hz) for some time t_w, it then stops for time t_s, then it starts working again for time t_w, and so on. In this case, RS does not operate the engine frequency, but it operates the t_w and t_s length. The extension of the period t_w, instead of the period t_s, also makes it possible to increase the oil flow rate.

Finally, the system generates two types of recommendations:

- Increase or decrease motor operating frequency (or in ratio of operating and idling time in the case of short-term operation) to boost the production rate.
- Reduction in frequency to reduce the risk of well shutdown (drying out, overheating of motor, etc.).

Outcomes of the testing in production

The designed ESP optimal control solution has been implemented and tested in the pilot area (described above). The proposed RS was deployed into production in the fall of 2018. It was proven to be feasible and efficient, enhancing the efficiency of oil extraction by boosting crude oil well production rates by 1.5% in average without additional capital investments.

Discussion and next steps

The solution has been developed to test hypotheses and does not aim to be exhaustive. At this moment, the solution has a few simplifications that significantly limit the possibility of issuance of recommendations and implementation for each oil well.

First, the operation mode recommendation can be issued for the oil wells with stable measured operational parameters in the current release of the solution; unstable wells are out of the solution scope.

Secondly, for now, the solution does not consider the mutual influence of the wells in different well clusters. Well clusters are connected by means of the infield pipeline. An increase or decrease in the flow line pressure in the outlet of one of the wells affects the oil flowrate of wells in the same well cluster and other clusters connected to the same infield pipeline.

Thirdly, each of the recommended operating modes for the ESP must be approved by the geologist. This approval is carried out manually, which severely limits the solution effectiveness in the number of recommendations implemented.

Finally, there is an unsolved issue of the equivalent water injection calculation and realization.

All of the complexities above prevent one of the usage of the proposed RS for ESP optimal control in the automatic mode. Close attention to the solution recommendations and additional work from the oil and gas company employees are needed.

Further development of the solution is aimed at eliminating restrictions in the solution implementation by means of creating additional software modules.

Team, which the authors belong to, is developing the well monitoring module in order to increase the number of wells with recommendations. The monitoring module core is the algorithm of head and rate correction. It will allow differentiation of the reasons for instability in wells' measured parameters, geological and mechanical, related to the ESP current state and nominal characteristics degradation automatically. It is also planned to use head and rate degradation model in the interest of ESP failure prediction.

Optimization algorithm is also being developed for infield pipeline pressures balance calculation. The calculation algorithms are based on the existing well-known pipeline calculation methods, with the only difference being that the algorithm allows fast recalculation according to the optimizer needs. This is the attempt to switch from the local oil well optimization to the global oilfield optimization considering oil treating limits.

Solution testing, along with the additional modules, is planned for 2020. Oilfields with surplus of water are in the scope of testing to ensure water injection for the additional production equivalent compensation. Furthermore, we plan to clarify the existing relationships between the injection and well productivity using AI, which is currently impossible due to low data variability.

Conclusion

The proposed prediction and recommendation systems have shown high accuracy and reliability throughout the entire testing period. The authors relate this result not only to the high quality of the ML models but also to the close interaction between the data scientists and the domain experts.

Also, during the implementation of the RS, there was an increase in the reliability of ESP functioning. The potential cost of ESP failure is extremely higher than one warning to stop or correct the ESP engine frequency. This result can be connected to the recommendations to reduce engine frequency in critical cases.

The implementation of RS based on ML approaches does not require additional equipment (high-performance servers, sensors, etc.) and can be easily implemented on any available technical base. It is worth noting that the proposed RS can also be adapted to the rod pumps.

The developed RS can be considered as a component of the "smart oil and gas field" concept. According to Gartner, implementation of the smart oil and gas field concept might help energy companies to reduce costs by 5% and boost the crude oil production by 2% (McAvey and Cushing 2019). Close projections were provided by CERA (Cambridge Energy Research Associates), according to CERA, an implementation of the smart oil and gas field concepts could reduce operating costs by 1 to 6% and cut oil well downtime by 1 to 4% (Wood 2007).

Nomenclature

D = well depth, m

E_{pred} = prediction of daily energy consumption that will be used in predicted mode, kW·h

E_{curr} = current energy consumption, kW·h

f_{curr} = current ESP frequency, Hz

f_{def} = default ESP frequency, Hz

H_{curr} = current pump head, m

H_{def} = pump head obtained with default ESP frequency, m

$p_{annulus}$ = pressure, kg/m^2

p_{meas} = measured reservoir pressure, kg/m^2

p_{real} = real reservoir pressure, kg/m^2

q_{real} = true daily oil flow rate, m^3
q_{meas} = measured daily oil flow rate, m^3
t_{24h}^{on}= duration of period when a pump is switched on during last 24 hours, h
V_{pred}^{oil} = prediction of daily production of crude oil, m^3
V_{curr}^{oil} = current daily production of crude oil, m^3
z_{oil} = the oil price, \$/m^3
$z_{electricity}$ = \$/kWh
w = water cut, %
$\delta\tilde{p}(t)$ = high-frequency (random) part of a pressure measurement error, kg/m^2
$\delta\bar{p}$ = constant (systematic) part of a reservoir pressure measurement error, kg/m^2
Δt_{meas}^{on} = duration of period when a pump is switched on during oil flow measurement time, h
Δt_{meas}= total oil flow measurement time, h
ρ = crude oil density, kg/m^3

References

Breiman, Leo. 2001. “Random Forests.” *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Editorial. 2012. “2012 Forecast: U.S. Oil Well Counts Rise in All Regions.” *World Oil Magazine*. https://www.worldoil.com/magazine/2012/february-2012/special-focus/2012-forecast-us-oil-well-counts-rise-in-all-regions.

Exarheas, Andreas. 2019. “IEA Cuts Oil Demand Growth Forecast | Rigzone.” 2019. https://www.rigzone.com/news/iea_cuts_oil_demand_growth_forecast-15-may-2019-158835-article/.

Kuzmichev, Nikolay P. 2012. “Short-Term Well Operation [STWO] Is Unique Approach to Reduction of Effects of complicating factors [in Petroleum Production].” *Exposition Oil Gas* 4 (22): 56–59. https://cyberleninka.ru/article/n/kratkovremennaya-ekspluatatsiya-skvazhin-unikalnyy-sposob-borby-s-oslozhnyayuschimi-faktorami.

Lea, James F., Lynn Rowlan, James F. Lea, and Lynn Rowlan. 2019. *Electrical Submersible Pumps*. *Gas Well Deliquification*. Gulf Professional Publishing. https://doi.org/10.1016/b978-0-12-815897-5.00012-3.

McAvey, Rich, and Simon Cushing. 2019. “Oil and Gas Industry Innovation Through Digital Technology Primer 2019.” 2019. https://www.gartner.com/en/documents/3900966/oil-and-gas-industry-innovation-through-digital-technolo.

Murphy, Kevin P. 2012. *Machine Learning: A Probabilistic Perspective*. The MIT Press.

Wood, Tony. 2007. “The Connected Oilfield.” https://www.cisco.com/c/dam/en_us/about/ac79/docs/wp/Connected_Oilfield_0629b.pdf.