Automatic Well Interference Identification and Characterization: A Data-Driven approach to Improve Field Operation

Mohammad Rahman and Gian Gioria, IPCOS Inc.; Sathish Sankaran and Diego Molinari, Anadarko

Abstract

Inter-well communication in unconventional reservoirs has received huge attention due to its significant effects on well production. Though it has long been a known side effect of hydraulic fracturing, well interference has become more prominent and frequent as the industry moves to larger completion designs with closer well spacing and infill drilling. Fracturing of infill wells ("child" wells) directly places the older adjacent producing wells ("parent" wells) at risk of suffering premature change in production behavior. Some wells may never fully recover and, in worst cases, permanently stop producing after taking severe frac hits.

This paper presents an automatic data-driven workflow developed to identify inter-well interference events and their impact on EUR (estimated ultimate recovery) based on changes in the well productivity trend. The innovative approach of the workflow is the ability to automatically analyze interference using the complete production history for all wells in a field, using routinely collected data and without introducing human bias in the derivation of the results, instead applying a consistent criteria. The final result is a comprehensive collection of all well interference events occurred in a field, which may be used as a training set for statistical and machine learning based models aiming at predicting such events.

First, the automatic identification of anomalies in the well behavior was developed and criteria set to label the interference events. Next, probabilistic simulations are run to forecast multiple scenarios to quantify the impact of a well interference event reported in terms of change in cumulative oil production. Finally, every event is analyzed in the overall context of field operations, in an attempt to present possible causes which may explain the change of production behavior.

Introduction

Over the past decades, hydraulic fracturing has advanced and using horizontal multi-stage multi-cluster fracturing has become a common process, especially in the ultra-low permeability unconventional resources. This multi-stage horizontal well stimulation has become more economic, boosting an intensive development of unconventional basins with tighter well spacing and larger completions job size. This has caused concerns in the industry as operators have repeatedly observed that an existing producer well can communicate with tightly-spaced, newly-completed offset wells through fractures (i.e. inter-well, fracture-driven interference,
more commonly called as "frac hits"). The existing producer well may be "knocked offline" or otherwise be affected by varying degrees of fracture-driven interference (FDI), which can significantly affect its production (Klenner et al. 2018). Hence, as part of determining the optimal field development plan, it is crucial to understand the severity of this interference with neighboring wells and how it correlates with well placement, completion design and development timing.

In this work, we are concerned with inter-well communication events, where an existing offset producing well (frequently termed as "parent" well) is affected by the pumping of a hydraulic fracturing treatment and/or start-up of a new neighbor well (called the "child" well). FDI can be strong enough to generate adverse effects such as damaging production tubing, casing, and even wellheads (Jacobs 2017 b). In other cases, they may just re-pressurize the drainage area of the parent well and thus boost its productivity.

As field development in unconventional fields mature, FDI are becoming a rule rather than the exception because of infill drilling, downsizing of well spacing and the tendency to pump larger frac jobs, with higher proppant and fluid intensities, shorter stages and more clusters per stage. Operators are drilling new wells closer together, and closer to older wells, all in an effort to drain as much of the available reservoir volume as possible. For example, most shale wells drilled in 2010 were at least 1000 ft apart; today well spacing may include tighter ranges from 550 ft to 250 ft (Thomson et al. 2016 a). Individual fracture stages are being placed tens of feet apart instead of hundreds, as was the norm a few years back (Thomson et al. 2016 b). In addition, companies are pumping two or three times the volume of water and sand than that of the parent wells were initially completed. The higher volumes of water and the increased amounts of sand injected are thought to be creating longer fractures than desired. These long fractures can extend into depleted reservoir zones, where large volumes of oil, gas, and water have been already produced. These low pressure depleted zones turn into a sink during fracturing operation that attract fracture direction and extension from the newly fractured offset well, causing low productivity (Jacobs 2017 a).

Frac hits have become a top concern now in the unconventional business because they can affect several wells on a pad, along with those on nearby pads too. Though the immediate effect of frac-hits have been largely studied in the past years (Ajisafe et al. 2017; Sun et al. 2017; Cao et al. 2017; Jacobs 2017 b), the long-term impact of fracture-induced communication between wells is still only partially understood. Many wells can establish pressure communication between each other even if no apparent frac-hit was observed during completion of the "child" well. In other cases, the communicating fractures close after the completion treatment is over, stopping the communication between two wells (Esquivel and Blasingame 2017; Awada et al. 2015). For this reason, it is important to establish a procedure that is able to reliably identify well interference events that have a long-lasting impact on production, through a unified approach that goes beyond the classic frac-hit identification. In this paper, an advanced analytics approach is presented to automatically identify interference events and quantify their effect on production.

Objectives

The objective of this study is to develop a workflow to automatically identify whether a well production behavior has been significantly changed by an interference event.

It would be beneficial to describe the characteristics of a suitable workflow. The identification method should be automatable, consistent and scalable for all wells. The procedure should be rooted in physics and yet data-driven, in that routinely measured variables are used for continuous analysis. The workflow should be able to find all possible well interference events, even the ones which didn't catch the eye of the operators. The focus is to derive a complete collection of well interference events and their impact on production.

The change in production behavior is sustained and considerable, in such a way that the expected EUR (Estimated Ultimate Recovery) is altered compared to the original behavior. The estimated impact is quantified using a probabilistic production decline analysis and reported in terms of change in cumulative
produced oil volume for each identified event. Every event would be analyzed in the overall context of field operations, in an attempt to find a possible cause for the change of production behavior.

**Methodology**
The methodology applied in this study can be seen in Figure 1. The workflow is composed of the following main modules:

1. Identify interference events
2. Quantify the impact
3. Label the event

![Well Interference Workflow](image)

**Identify interference events**
This module of the workflow includes data preparation and identification of interference event. The main objective is to automatically find anomalies in the well behavior that can be labelled as possible interference events.

Analytics were developed to detect anomalies in well behavior. Filtering is a fundamental step in the preparation of the data for the automatic processing. As the ensuing algorithm will try to find anomalies in the expected production behavior, it is imperative that noise and isolated outliers are not detected as an anomaly, but they must be removed to retain the main production trend. The configured filter removes points that differ from the average behavior within a certain threshold, which is usually a multiplier of the overall standard deviation of the sample from its reference behavior.
In the following sections, an overview of the selected variables used for anomaly detection is presented. These variables are deemed reliable and the authors of this paper consider anomalies in their behavior as markers of well interference phenomena.

**Productivity index – total liquid.** The production behavior of a well is usually monitored by observing the oil rate vs. time decline. However, this approach is very sensitive to surface operations (e.g. a choke change might increase the oil production, even though the output is expected to decline). For this reason, the Productivity Index \((PI)\) has been chosen as a good representation for the true production decline of a well over time. In the data analytics terminology, PI was chosen as the selected "feature for subsequent analysis.

\[
PI_i = \frac{q_i}{p_{res} - BHP}
\]

where:
- \(q_i\): Daily rate of phase \(i\) [bbl/d]
- \(p_{res}\): Reservoir pressure [psia]
- BHP: Bottomhole pressure [psia]

In this approach, the \(p_{res}\) is kept constant over time and equal to the initial estimated reservoir pressure, *de-facto* excluding the effect of depletion in the reservoir, so this paper uses a pseudo-productivity index as an approximation for the true PI. In this way, the productivity index can be considered a "pressure-normalized rate". The PI is in principle not sensitive to operational changes on the surface (e.g. choke change), so it can be considered as a more reliable indication of the well production decline (Molinari et al. 2019).

To reduce the influence of the phase allocation (e.g. increase in water production, incorrect production allocation and faulty phase-cut measurements), while accounting for the production behavior of a well, the PI is measured on total liquid (oil and water) rate, especially as the method was developed for a field producing high water cuts. In this way, a misallocation of water and oil will not impact the overall estimation of PI. Gas is not taken into consideration, as gas production is heavily affected by the depletion of the reservoir (higher GOR are expected at later stage of production). Furthermore, gas is much more mobile than the liquid phases, so is less sensitive to changes in the reservoir behavior due to interference events.

**Water Oil Ratio (WOR).** The literature review suggests that a sudden and sustained change in water production might be an indication of well interference due to fracture induced communication between a stimulated well ("child well") and an existing neighboring well ("parent well") (Klenner et al. 2018). This type of interference is caused by an aggressive hydraulic fracturing of a well next to tightly-spaced existing wells, resulting in the parent well proportionally producing more water than prior to the event. This type of interference has an observable impact on the production decline of a well, which is visible also on the PI liquid behavior. However, it has been decided to incorporate the analysis of the WOR profile along with the PI Total Liquid to reinforce and improve the automatic detection of interference events.

**Time Online.** The evolution of the two production parameters listed above are monitored on a certain time scale. The long-lasting interference events that we are interested in capturing tends to disclose themselves over a few days, where production profiles are consistently altered. At this stage, there is no reason to further refine the analysis to process values at a higher measurement resolution. Having daily values for PI and WOR is considered enough to detect significant changes in behavior and reduce the effect of measurement noise. Furthermore, unconventional fields may lack high-frequency data (e.g. coming from real time transmitters) for all wells, so it was decided to rely on daily values generated as part of production allocation. This would ensure the applicability of the workflow on any field, no matter what level of instrumentation is available.

The possible shut-in periods in the life of a well cause the selected production parameters to be zero for some time. Those zeros on the timeline affect the capability of automatically detecting the interference events and, therefore, must be removed. The valid timelines to be considered are then reduced to either the
"Time Online" or "Material Balance Time" ($t_{mbc}$). "Time Online" is the cumulative amount of days that the well was flowing. In other words, the total amount of days since the start of production, minus the shut-in days. "Material balance time" is defined as the ratio of cumulative production to instantaneous rate (Sun, 2015).

Plotting the PI of known cases of interference on material balance time produced a particular observation: if the effect of interference is (temporarily) positive (meaning the well produces higher liquid rate for a comparable drawdown), the material balance time "rolls back". Effectively, the momentarily higher production resulting from a positive interference event causes the material balance time to shrink, effectively "rejuvenating" the well as it gets a boost of energy from the frac hit. This effect produces overlapping production decline curves for the same moment in time (Figure 2). This observation is in line with the expectations: the same well behaves completely differently after sustaining an interference event (a frac-hit in this case), resulting in two distinct decline profiles on Material Balance Time scale. However, such production profiles are not suitable for automatic processing, as they would require complex clustering algorithms to discover the distinct decline curves which may be more difficult to implement, though possible.

If the same case is reported on a "Time Online" scale, then the production profile is easier to be processed by an automatic algorithm (Figure 3), which can now leverage on the sequentially distinct occurrence of anomalies in the decline behavior with no overlaps and no gaps. Eventually, "Time Online" was chosen in this work as the most appropriate timeline for the production parameters.
Anomaly detection

After removing outliers, the workflows analyzes the profiles of Total Liquid PI and WOR vs Time Online in search of anomalies. Both these variables are ever-changing through time, so finding an anomaly in their timeline requires advanced data analytics techniques, mixed with consolidated principles of decline curve analysis. The key principle is to identify an expected behavior that would help the algorithm to highlight anomalies as compared to that baseline.

The algorithm is based on an iterative approach that aims at finding the best solution that fits the decline curve around its anomalies. The algorithm is presented below, step-by-step, in its fundamental stages. Starting with a first guess, the algorithm needs to measure the anomalies of the signal, compared to an expected behavior. The expected behaviors for the signals are: hyperbolic decline for total liquid PI (Figure 4) and cumulative average for WOR (Figure 5). While using a hyperbolic decline sounds intuitive for PI decline, the choice of using cumulative average for WOR needs to be explained.
In this case, the reservoir has a significant amount of source water and the field behavior shows a very constant or stable WOR. The safest approach is to use a moving line to be used as a baseline that slowly follows the WOR trend, so that sudden changes of WOR are promptly visible, but long-term changes are properly hidden. Based on extensive empirical observations and fine-tuning, cumulative average seems to be a reasonable choice (Figure 5).

For both PI and WOR signals, the residuals are measured from how data deviates from the respective reference expected behavior. The residuals are processed using a 2nd-order Savitzky–Golay filter and then taken as absolute values (to equally highlight positive and negative anomalies) standard smoothing and normalization techniques. Residuals are plotted below to show how a large anomaly can be already guessed by looking at the picture at around 800-1000 days of production (Figure 6).
These two normalized residual profiles are combined with a weighting factor. The resulting combined residuals trend is the summation of the PI residual + Weighting Factor * WOR residuals (Figure 7). The weighting factor value for WOR residuals is derived from fine-tuning.

From the combined residual profile, the local maxima are picked. Every spike in the residual is a possible anomaly and, therefore, a possible well interference event occurred in the history of the analyzed well.

The algorithm iteratively cycles through all the highlighted anomalies, trying to fit two new hyperbolic decline curves (as independent segments) through the split PI vs. Time Online dataset, one before the anomaly and one after (Figure 8).
The fitting quality of the two proposed decline curves is measured against the fit quality of the base case (i.e. only one decline curve). Classic statistical parameters are used to measure the quality of the fit: $R^2$, total error, error per data point (on both the "before" and "after" datasets).

The new curves are compared to the fitting quality of the single first guess curve. If the two new curves are comparably better than the original fit, then the well decline behavior is better described by the two curves rather than just one. This is equivalent of saying than the decline behavior has changed as a result of the anomaly, therefore we can safely assume that the anomaly constitutes a well performance changing event, which may potentially be caused by well interference (Figure 9). The Appendix provides the mathematical detail of how these comparisons are carried out.

Alternatively, if the fitting quality of the new curves is not a significant improvement compared to the single curve case, it can be inferred that the original curve had sufficient descriptive capability to represent the well decline and no significant changes have occurred over the detected anomaly. In other words, the anomaly in PI and WOR behavior cannot be considered a relevant event.
In this example, only one step in the anomaly detection algorithm has been described. The algorithm is capable of further analyzing the well production history to find additional anomalies (not just one). The procedure stops when none of the new proposed curves improves the descriptive quality of the old curves. In other words, when none of the anomalies caused a significant change in PI and therefore no more events can be detected.

In summary: anomalies in PI and WOR history are considered markers for possible well interference events. The algorithm evaluates each anomaly by establishing whether the PI decline is better described by one or two hyperbolic curves. If two hyperbolic curves are better suited for the description of well behavior around an anomaly, then we can infer that an event changed the production decline. This event is a possible well interference occurrence. Another example is presented below (Figure 10):
Quantification of the impact

In order to quantify the impact of all the possible events in terms of lost/gained oil after the event, forecasts are generated using the two decline behaviors computed "before" the event and "after" the event, and their profiles are compared to assess if there was a long term production degradation. The decline showed before the event will not occur anymore, as the well behavior is permanently altered by the event. However, a forecast is still computed using the "before" conditions to serve as a baseline scenario. A second forecast is computed using the "after" conditions to represent the predicted future behavior of the well after the event. The difference in oil production between the two scenarios quantifies the impact of the well interference event.

Considering all the uncertainties inherent to the measured data, a deterministic "traditional" forecast, based on decline curves might not be sufficiently accurate. The authors of this paper decided to use a probabilistic forecast: a range of possible forecast scenarios are generated to quantify the impact of a well interference event, through the Data Space Inversion (DSI) technique (Jiang et al. 2018). Data-space inversion (DSI) is a data-space approach that creates forecasted distributions through sampling the randomized prior distribution of relevant quantities (PI decline in this case). The sampling works through a randomized maximum likelihood method applied to the conditional distribution of priors given observed data (Sun and Durlofsky, 2017). In other words, the posteriors are the most likely realizations of randomized priors, given the observed data.

The resulting forecast profiles are model-free, they cannot be expressed as explicit functions and they are simply a collection of data points in the future. To complete the analysis, the DSI procedure has its counterpart in the traditional Model Inversion (also known as model history matching), where prediction is done on model parameters based on previous measured observations. In Model Inversion, the resulting predictions are given in terms of model parameters, which can be plugged into functions and obtain functional forms for the forecasts.

In the current workflow, both Data Space and Model Inversion have been implemented, though only the DSI technique has been tested and run extensively. This module of the workflow is applied for each well.
with at least one event. The algorithm loops over every well with events. The DSI forecast computation is run separately on the two datasets around the event (before and after the event).

The following section presents the DSI steps applied in this specific case. The probabilistic forecast is applied to the Productivity Index computed on the oil phase only, referred as PI Oil, as we are interested in estimating the future well performance on oil production.

**Generate priors.** Prior decline curves aims at representing every possible expected outcome of the well decline. In other words, any possible forecast can be derived from these priors. Any possible forecasted value must be enclosed in the range of priors, so a wide range of possibilities are included, to ensure that the uncertainty is bounded in the full range of priors.

Practically, priors are generated by inducing exaggerated perturbation of the best fit decline parameters, in order to generate huge variability in the decline curves. As it can be seen from Figure 11, a multitude of decline curves are generated spanning the measured data. Here, 1000 priors are generated for each forecast (i.e. 2000 priors for each event, 1000 before and 1000 after the event).

![Prior distribution before and after the event](image)

Prior curves have no specific meaning, other than defining the range of possible values out of which the probabilistic forecast will be derived. Note how the priors extend into the future, beyond the observed data (in Blue) (Figure 11)

**Compute posteriors.** Posterior distributions are priors that are most likely to be realized, given the observed data and a certain measurement error (Sun, 2017). In this case, the measurement error is not known upfront. Typically, it should come from the instrumentation used to generate the data.

In this case, the measurement error is assumed to be proportional to the standard deviation of the difference between the observed data and a moving average running through them (assumed to be a "reference" line). The idea is to correlate the measurement error (necessary for DSI procedure) to the variability of PI values compared to an assumed reference behavior.

Using the DSI procedure, 20 posteriors are generated for each dataset (i.e. 40 posteriors in total, 20 before and 20 after) over the fully extended timeline (observed data + forecast) (see Figure 12). The posteriors are
a collection of data points, not derived from an explicit functional form. They may look like decline curves, because they are computed from priors generated using modified hyperbolic decline equation.

Sample relevant forecasts. The twenty forecast profiles need to be reduced to a relevant dataset suitable for comparison and more interpretable. Therefore, the P90, P50 and P10 profiles are sampled from the posterior forecast distributions. The sampling is done for both distributions: before and after the event.

At this point, we have a probabilistic forecast for the PI Oil in both scenarios: following the original decline before the event and the new behavior after the event. However, the forecast is only derived for the productivity index of the well, and it must be converted to an oil production prediction.

Calculate \( \Delta p \). The easiest way to translate a forecasted PI into a forecast oil production is to derive an estimate for the drawdown delta pressure \( \Delta p \).

The measured oil rate and the PI Oil time trend after the event are acquired to calculate the estimated time trend of drawdown \( \Delta p \).

\[
\Delta p(t) = \frac{q_{oil}(t)}{PI_{oil}(t)}
\]

This drawdown \( \Delta p \) is an approximated glimpse into the future after the event. There are better methods that could be used for predicting the BHP and drawdown of a well, but the dissertation on those is beyond the scope of this paper. Nevertheless, the implementation of a robust drawdown prediction algorithm is a possible improvement to this methodology. As a strong simplification, we assumed that the drawdown \( \Delta p \) will remain constant for the next 2 years after the event, and it's equal to the average drawdown \( \Delta p \) that was applied to the well after the event (Figure 14).

Calculate cumulative oil. For each event, the cumulative oil volume at that date is acquired. The PI Oil is forecasted daily after the event and \( \Delta p \) is assumed constant. Therefore, we can compute the daily forecast values for oil rate:

\[
q_{oil,\text{forecast}}(t) = \Delta p \times PI_{oil,\text{forecast}}(t)
\]
Since the oil rate is given daily in barrels per day, we can easily compute the cumulative produced oil by simply summing the values of oil rate. The resulting cumulative at date \( t_{\text{end}} \) is computed as follows:

\[
\text{CumOil}(t_{\text{end}}) = Q_{\text{oil, event}} + \sum_{t_{\text{event}}}^{t_{\text{end}}} q_{\text{oil, forecast}(t)}
\]

The cumulative oil is computed for the three sampled forecast profiles: P10, P50 and P90.

**Compare cumulative oil in different cases.** In order to have a descriptive quantification of the event impact on the production of the well, we need to compare the forecasted cumulative oil volumes using the forecast generated on the "before event" and "after event" datasets. Three different scenarios are computed, based on the different sampled forecast taken from the "before" and the "after" event datasets. Table 1 provides clarity on the rationale and the definition of the delta forecasts.

<table>
<thead>
<tr>
<th>Before event</th>
<th>After event</th>
<th>P10</th>
<th>P50</th>
<th>P90</th>
</tr>
</thead>
<tbody>
<tr>
<td>P10</td>
<td></td>
<td></td>
<td></td>
<td>Optimistic ¹</td>
</tr>
<tr>
<td>P50</td>
<td></td>
<td></td>
<td></td>
<td>Neutral</td>
</tr>
<tr>
<td>P90</td>
<td></td>
<td></td>
<td>Pessimistic ²</td>
<td></td>
</tr>
</tbody>
</table>

The forecasted profiles are combined from the point of view of the event: was the event beneficial or detrimental for the well productivity? For example: when the lowest possible oil output (P10) from the "before" scenario is compared to the best possible outcome (P90) of the "after" scenario, the overall forecast scenario is labelled as "Optimistic". The three forecast scenarios shown in Table 1 (pessimistic, optimistic and neutral) are used to classify the probable impact of an identified well interference event.

The following figure (Figure 15) shows the cumulative forecast in the "neutral" case (P50 vs. P50). It can be observed how the interference event had dramatically increased the production on the short term, as it was already visible from Figure 13.
Figure 13—P90, P50 and P10 from posterior distributions

Figure 14—Bottomhole drawdown pressure over well history, with highlighted time of well interference event (dashed line) and average value after the event (green line)
For further comparison, the "pessimistic" case is also reported (Figure 16):

It can be observed how the beneficial effect of the interference event is much smaller than the previous case, though still on the positive side. However, the predicted decline after the event in the P10 case (i.e. worst outcome) shows a faster degradation compared to the base case (i.e. the "red" line) so that the positive gain is slightly reducing with time.

**Event Labeling**

We analyzed other contextual field information, in an attempt at finding the possible causes of the event, which is likely due to well interference but may also be originated by other causes. The workflow searches for parent-child relationships between wells using commonly available datasets. It derives a wide range of causes and present them to the user for additional investigation.

Several datasets are mined for additional information:
• Completion schedule
• Production start-up schedule
• Well interference event list
• Surface, heel and toe locations for each well (longitude, latitude and depth)
• Operator's daily comments, recorded as part of daily well inspection round

From each of the datasets above, additional information can be correlated to each event.

Completion schedule. The date of an event is cross-checked with the completion schedule, to find out if any frac-related activity was occurring in the vicinity of the "parent" well. This check is rather intuitive and it typically returns the list of wells which were completed nearby at the time of the well interference event. These wells are all possible "children" wells.

Start-up schedule. Similar to the previous step, the date of the event is cross-checked with the start-up schedule, to find out if well was put on production in the vicinity of the "parent" well. In fact, starting up a well might reveal that pressure communication is actually occurring between two wells, even though no frac-hits were noticed at the time of completion. This check typically returns the list of wells which were started-up nearby at the time of the well interference event. These wells are all possible "children" wells.

Well interference event list. Similarly, the event date is cross-checked against the list of other well interference events, to discover if similar events have occurred in the same area. This check might lead to the discovery of concurrent events that may have either a common cause or might be the result of mutual interference.

Well Locations. Well location database is necessary to compute the distance between "parent" and "child" well. In fact, well spacing is believed to be a major factor in determining well interference events (Klenner, et al. 2018) Therefore, computing spacing between the "parent" and the possible "child" well is one key and useful data to be included in the description of a well interference event.

The details of how spacing is computed between two wells are beyond the scope of this paper. As a quick reference, spacing is computed as Euclidean distance between the lateral sections of the wells, assumed as straight lines.

Operator's daily comments. Even though the comments may be unstructured and they might not follow a specific standard or common conventions, they are still a valuable insight into the field conditions and operations occurring around the time of the event and they can provide important information for event labeling. These comments are free text entered by field personnel as part of routine well inspection.

Using a simple algorithm that counts how many times certain keywords appear in the comments, the workflow derives a possible explanation of what the cause of the well interference event was. The results of this module have highlighted how sometimes a sudden change in production might be the result of just a long shut-in period that leads to sustained flush production. Therefore, some of the identified possible events were actually not related to well interference.

The final result is a list of wells that were close enough in both space and time to have possibly caused the analyzed event, split by cause type:

1. Nearby fracture (i.e. a frac-hit)
2. Nearby start-up (child wells whose start-up have caused a perturbation in the reservoir)
3. Nearby event (i.e. concurrent mutual production interference)
In addition, for each event the algorithm suggests the most likely cause and the wells that might have triggered it. The overall final result is a catalog of all the possible events, along with the estimated impact and an indication of the possible causes.

Results
The above methodology has been applied to all producing wells in a large operated US Onshore field. The probabilistic forecast scenarios (P10, P50 and P90) were generated to compute the impact on future well performance, along with the estimated likely causes of the event.

While well interference often has a negative connotation, one of the key findings of this study is that not all such events were detrimental to parent well productivity. However, this could be achieved at the cost of child well's expected productivity, which is not the focus of this work.

The end result of the example well that has been presented in earlier sections, are shown in Table 2 (well names have been anonymized).

Table 2—Example of well interference event complete analysis results

<table>
<thead>
<tr>
<th>Well</th>
<th>Well A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event Date</td>
<td>8/12/2017</td>
</tr>
<tr>
<td>Water Increase</td>
<td>Yes</td>
</tr>
<tr>
<td>Likely cause</td>
<td>Offset frac</td>
</tr>
<tr>
<td>Nearby frac</td>
<td>Well B, Well D</td>
</tr>
<tr>
<td>Nearby event</td>
<td>None</td>
</tr>
<tr>
<td>Nearby start-up</td>
<td>None</td>
</tr>
<tr>
<td>ΔCumulative Oil³</td>
<td>Pessimistic +21.6 kbbbl, Neutral +36.6 kbbbl, Optimistic +53.0 kbbbl</td>
</tr>
<tr>
<td>Δqoil Neutral</td>
<td>+187%, +106%, +60%, +28%, +9%</td>
</tr>
<tr>
<td>Δqoil Pessimistic</td>
<td>+186%, +94%, +41%, +5%, -17%</td>
</tr>
<tr>
<td>Δqoil Optimistic</td>
<td>+190%, +119%, +77%, +47%, +30%</td>
</tr>
</tbody>
</table>

Well A has experienced an event that changed its performances around 12th Aug, 2017. The event was associated with a strong increase in water production. As per operator's comments, the well was shut-in due to an offset frac-job in the proximity of the well. The only wells that were reported to be fracked in the previous month and in the proximity of the analyzed well were: Well B and Well D. The first one is more likely to have caused the event. No other events or well start-up were reported in the proximity of the well. The Cumulative Oil production is estimated to have increased by about 36,000 bbls in the "neutral" case (P50 vs P50) over a two-year horizon (Figure 15), so the interference event has had a positive effect. Most likely, the child well improved the stimulation in the parent well. We can see that the oil rate after the event is consistently higher than the base case scenario forecast (if the event had not happened). In the first period, the oil rate is nearly 3 times higher than the base case scenario. Over the next period the effect tends to disappear and, after 2 years, the oil rate is just 9% higher than the base case. In the "pessimistic" case, the oil rate is actually 17% less than the base case 2 years after the interference event, highlighting how the uncertainties in the forecast can potentially change the sign of the prediction.

After validating the algorithm results with operations staff, it was confirmed that the event was actually a frac-hit. The workflow was able to discover this event in a matter of few seconds, without having any prior knowledge. Additional 368 such events were discovered by the workflow after performing a full scale
analysis of the entire field (about 528 wells, with varying length of production history). The analysis of the entire field took about 20 minutes on a regular laptop.

While the accuracy of the outcomes is still under review, and not every detected event is a real interference event, it is indisputable that the results brought attention on some cases that were unknown to the operators and definitely improved the asset team's understanding of the phenomena and its pervasiveness. These results provide an excellent starting point to perform a field wide analysis and validate the suspected events.

A cleaned-up version of the resulting dataset was used to generate a training dataset for predictive modeling of ΔEUR. The full dataset (inputs and outputs) was fed into a standard machine learning algorithm, which resulted in roughly 80% accuracy in predicting the ΔEUR on "parent" wells after a child has been completed nearby. However, a full dissertation on this machine learning generated model is beyond the scope of this paper and the evaluation is still ongoing.

**Conclusion**

Building a catalog of well interference events that are automatically computed, and continuously as more data is acquired, is by itself a valuable data source for subsequent analysis to understand the leading causes and optimize field development (well placement, completions design, targeting, sequencing etc.).

The main contributions of this workflow are to:

- Detect well interference events based on a novel PI-based approach with no prior knowledge, mitigating human bias and potentially implemented as a recommendation system for interference.
- Derive valuable information from routinely collected data, with no need for extensive (and expensive) instrumentation campaigns.
- Provide the engineers a comprehensive collection of well interference events with several information already pre-processed for them, significantly saving time, especially in developments with large well counts.

**Reference:**


8. Jiang, S., Sun, W., Durlofsky L.J. (2018), A Data-Space Approach for Well Control Optimization under Uncertainty, 16th European Conference on the Mathematics of Oil Recovery, ECMOR 2018


Appendix:

Comparison technique between two new fitted curves vs a single curve

The comparison's goal is to determine whether the quality of fit is better using two curves or just one curve. Assume that the following parameters are known for each of the three fitting curves (the two new and one old):

- Coefficient of determination: $R^2$
- Sum of squares of residuals: $SS_{res}$
- Number of samples in the dataset: $N$
- Average square of the residuals per sample: $avgS_{res} = \frac{SS_{res}}{N}$

The fitting parameters above are combined in the case of the two new curves to facilitate the readability and comparison. The "before" and "after" subscripts are added to better distinguish that the parameters are referred to two new datasets derived from a unique dataset split at the time of the event:

$\text{Combined } R^2 = \frac{R^2_{before}N_{before} + R^2_{after}N_{after}}{N_{before} + N_{after}}$

$\text{Combined } SS_{res} = SS_{res, before} + SS_{res, after}$

$\text{Combined } avgS_{res} = \frac{avgS_{res, before} + avgS_{res, after}}{2}$

The fit quality provided by the two new curves is considered better if all the four conditions below are True:

- Combined $SS_{res} << \text{Original } SS_{res}$ OR Combined $R^2 > \beta \times \text{Original } R^2$
- Combined $SS_{res} < \text{Original } SS_{res}$
- Combined $avgS_{res} < \text{Original } avgS_{res}$ OR Combined $R^2 > \text{Original } R^2$
- $R^2_{before} > 0.5$ AND $R^2_{after} > 0.5$

$\beta$ is a "bias" parameter. It has been introduced in the calculation in order to facilitate the validity check of the first condition. It's set to 0.935 (manually adjusted) for the first iteration, then progressively increased.

This means that the first anomaly can be labelled as an event (i.e. fitting two curves is better than using just one) even though the Combined $R^2$ is not necessarily better than the original $R^2$. This "nudge" towards accepting new curve fit is progressively reduced as the iterations go on. It's initially set to "permissive mode" (i.e. smaller than 1) to compensate for the fact that the two split datasets resulting from the first split are still likely to contain several significant anomalies and their $R^2$ might not have a high value. Therefore, they might need a little "nudge" to pass the validation and their evaluation is "biased".

Generally, the two new curves should reduce the overall error or improve the Coefficient of determination $R^2$ or a combination of these two improvement directions.