Sand Production and Control Benchmarking through Unstructured Data Analysis with Machine Learning in the North Sea

Introduction

Sand production has been serving as a bottleneck to the oil and gas industry, contributing to disruption of daily production operations, casing deformation, erosion of well tubing, pipelines, and surface equipment, expediting to significant non-productive time (NPT) costing millions of dollars in loss annually. Conventional areal studies for sand production will only be limited to few wells and heavily dependent on data availability, human-based interpretation, and time constraints. To administer a holistic basin study for sand production comprising of hundreds of wells with conventional manual method is considered complex and time consuming, hence sand mitigation best practices are typically derived from localized reservoir and production engineering data only, and knowledge is organically built through accumulated expert experience in the area over multiple years of operatorship. Information and reports may be derived from millions of pages of legacy well reports, documentations, or files from over 40 years ranging from digitized medium to hand-written reports in countless formats.
A sustainable strategy of data-driven basis shall be leveraged to address the knowledge and information management issues utilizing the latest advancement of Machine Learning (ML) and data analytics to maximize the potential of unstructured data. Utilizing the intuitive data-driven approach, the paper will highlight the areal causation of sand production based on geological characteristics and the best practices of sand control commenced by 8 operators in Norwegian Basins practically informative for future exploration wells to be developed nearby current wells. The study first creates a relationship between the causation of sand production versus the sand control practices implied and best practices are derived from the practices of multi-wells.

Methodology

Unstructured data in nature is significantly sophisticated to be manually interpreted and skimmed through by human-intervention. An intuitive approach embedded with Deep Convolutional Neural Network (DCNN) for autonomous image recognition and Natural Language Processing (NLP) for texts and entity processing and recognition has been a pioneering enterprise-scaled platform in managing unstructured data (Hernandez et al., 2019). It will be capable to ingest *"big data"* for the case study comprising of 70,000 files with 490,000 pages and 430,000 images inclusive of 361 wells over 5 basins in Norway. 40 years of unstructured data for sand production case study is consolidated approximately within 16 days of study period.

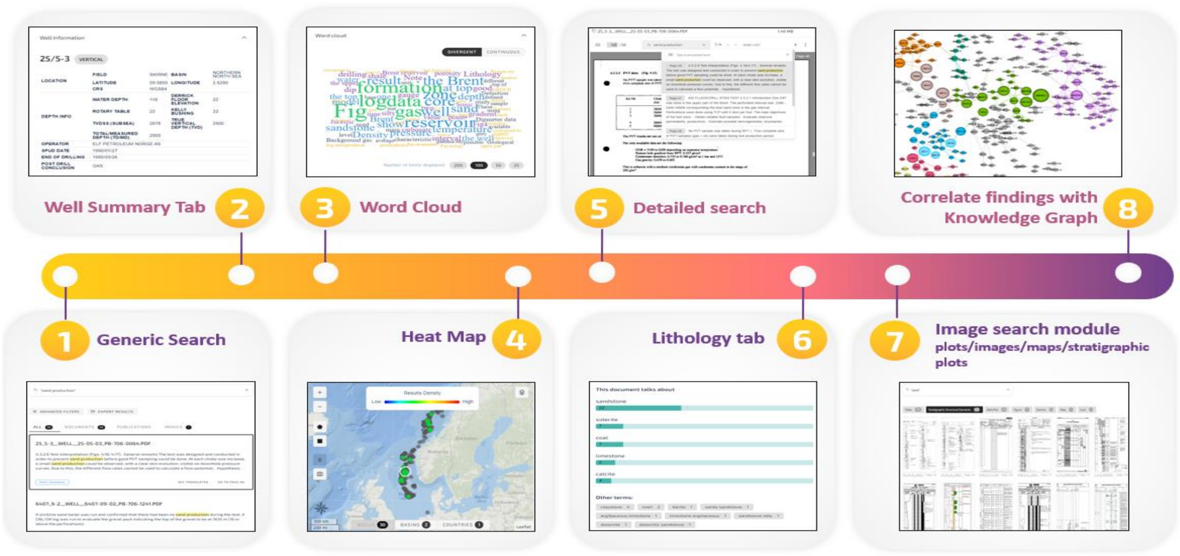


Figure 1 *Data-driven case study research strategy for sand production in Norwegian Basins.*

Step (1) is a generic deep search of the sand production scope across the whole corpus. Step (2) leads to discovering all wells significant parameters based on the files search results within well summary tab. Step (3) portrays an early insight of the general idea of the document with word cloud. Step (4) heat map resembles the wells distribution in a GIS map based on colour density of corpus search frequency mentioned in the documents. To uncover more questions along the research process and to obtain more in-depth information, step (5) is conducted in an iterative manner. Step (6) is to relate the lithology distribution by lithology count within each well document to the previous detailed search parameters. Step (7) aims to find more representable images to support the case study through the automatically classified images through DCNN in the reports. After all significant parameters are obtained, well-to-well relationship is studied to get more details on further causation and best practices of sand production management.
The features of this intuitive knowledge management platform transform voluminous unstructured data into structured data that are ready to be consumed and utilised for production enhancement case study. Four main ML analytical tools embedded in the platform are as presented below (Baillard et al.,2021):

- Expeditious and intelligent search module by keyword-basis searching through hundreds of thousands of pages of texts and texts embedded inside images.
- Autonomous extraction of images from documents and image segregation into respective image classes of tables, figures, well plots and maps with DCNN image detection algorithm.
- Knowledge graph with contextual well name relationship portraying connectivity of 'related corpuses' to understand well-to-well relationship as described in their respective document corpus.
- Heat Map illustrates the density of keywords by colour gradient on wells based on the search results on a map. Polygonal or square filter feature enable selective wells to be screened out for users narrowed search interest.

Case Study: Preliminary Study on Sand Production and Developing Sand Control Benchmark of Norwegian Basin with Unstructured Data

The study was conducted extensively throughout approximately 361 wells consuming16 days of study period covering the analysis and interpretation of sand production trends, causation and best practices in Norwegian Basins. A total of 8 operators were participating in exploration phase of reservoirs especially in Voring and Northern North Sea basins.

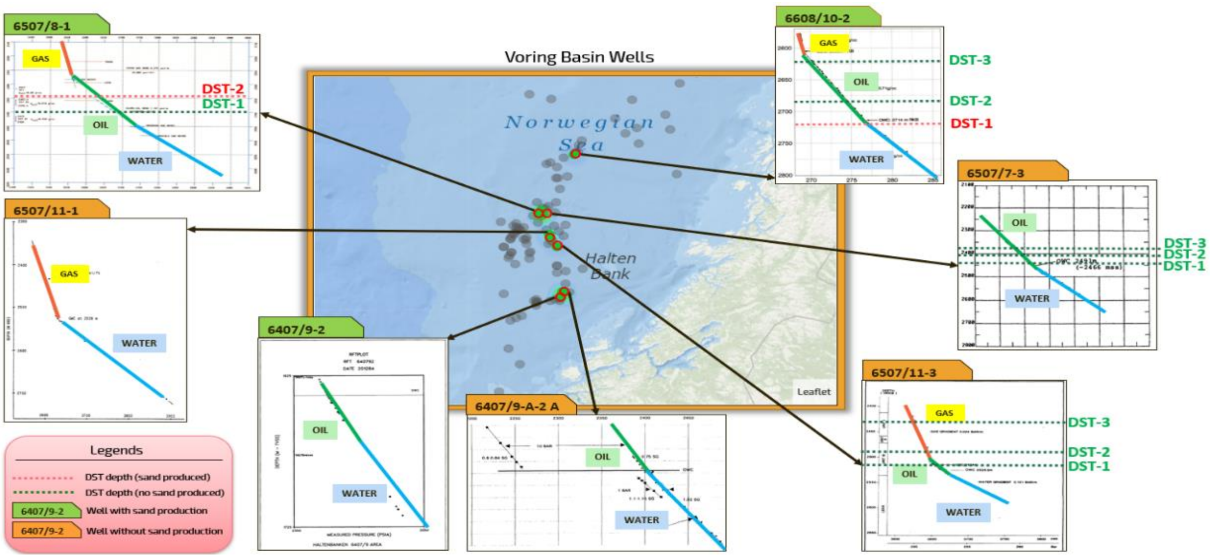
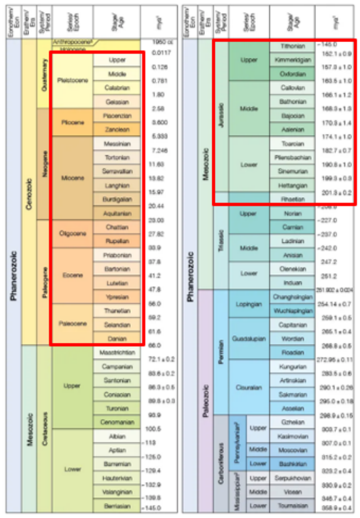


***Figure 2** Reservoir Performance Tests (RFT) for Voring Basin Wells*

Sand production was reported in 7 wells and most discoveries of sanding were reported during Drill-stem Testing (DST) and reported in completion reports and drilling program reports as in Figure 2. A few wells were reported to experience little to no sanding issues however providing adequate information in the scope of best practices and recommendations.

Formation Age	Depositional Environment	Formation	Interval (m RKB)	Description	Well Name
Upper Toarcian	Marginal marine – deltaic – inner shelf	Tofte	2668.0 – 2719.5	Upper part – clean sandstone, fine grained, well sorted and friable Lower part – medium to coarse –grained, moderately sorted, friable to loose grains.	6608/10-2
Sinemurian / Pleinsbachian – Early Toarcian		Aldra	2386.5 – 2559.0	Moderately sorted, fine to very fine grain size, predominantly friable except the tight zones with hard cementation.	6507/8-1
Upper Jurassic		Froya	1651 - 1657	Soft tertiary clay stones with silt and sand layers	6407/9-2
Paleocene - Danian		Top Heimdal	2158.5 – 2169.5	Sandstones, friable to very friable, moderately sorted, subrounded to rounded	25/5-5
Late Paleocene		Top Heimdal	2258 - 2267	Traces of sands (+/-1% at 5000 STB/day rate)	25/8-8 S
Upper Jurassic		Top Heimdal	2391.4 – 2398.4	N/A	25/10-8



Geologic Time Scale

***Figure 3** Chronostratigraphic evaluation of Sand Prone Wells*

Referring to Figure 3, chronostratigraphic reports denote sufficient reasoning of sand production occurrence throughout Norwegian basins as most sanding issues are prone in younger formations or tertiary aged formations from Upper Jurassic, Late Paleocene, Danian, Early Toracian, Sinemurian and Upper Toracian. Cuttings and core samples obtained from different stratigraphy depths of sanding prone formations; Tofte, Aldra, Froya and Top Heimdal were mostly described with friable, loosely grained, traces of sands, soft tertiary claystone, little to no cementation and fine-grained characteristics dominantly originating from sandstone, carbonate and claystone-sandstone mix lithology. Marginal marine and deltaic depositional environments leads to less cementation and loosely grained deposited sand characteristics.

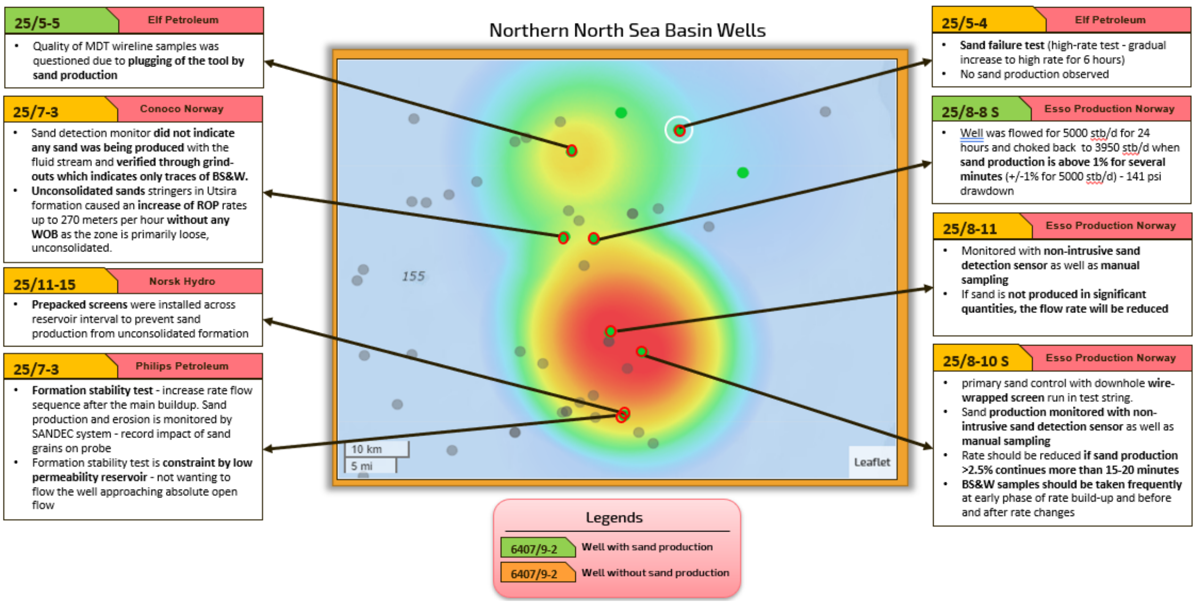


***Figure 4** Heat Map for sand production best practices – Voring Basin*

Reservoir parameters analysis was conducted to observe the sanding occurrence trends with respect to porosity, permeability, skin and perforation shots for each wells experiencing sand production issue. Stimulated wells with negative skin value, significant high permeability and porosity values are arbitrarily associated to sanding issues. However, a few wells do highlight these characteristics but no or little sanding occurred, and assumptions were made that inter-grain cementation are intact or sanding probably will occur soon in the later phases of the reservoir as sand production onset is distinct in each well. Perforation design does not lead primarily towards sand production as comparison has been made for wells with the same perforation shots with significantly different reservoir parameters for comparative analysis. Pore pressure abnormalities attained from Knowledge Graph module possibly causes sand production problems specifically reported in well 25/5-5 and 25/6-3 within Heimdal formation interval, leading to depleted pressure gradient trends in both wells. Analysing the causation of sanding creates an understanding in relation to the sand control practices conducted for each of the wells. Wells describing sand production issues, sand control mitigation methods and other relatable descriptions of sand production were intensively analysed as Figure 4 above.

Conclusion

Managing unstructured data into an intuitive structured data with embedded end-to-end ML advanced technology made it possible to interpret, analyze and make decisions with regards to handling *"big data"* and derive sand production causation and best practices across 490,000 pages of public documents inclusive of 361 wells and 2 Norwegian basins in total. The novel approach serves as a holistic study of sand management focused on unstructured *"big data"* which combines multiple digitalization techniques currently applied in the petroleum industry. Maximizing the potential of underutilized unstructured data leads to opening of vast opportunities for enhancement of production in existing oil and gas wells, and reduces investment in the drilling of newer, more expensive wells, in alignment with a re-use, reduce, up-cycle mentality, towards sustainable energy transition for the industry.

Acknowledgment

This paper utilizes the data from the Norwegian Petroleum Directorate (NPD) open dataset. Disclaimer of those interpretations from the study are from investigation and analysis of the authors alone.

References

Acock, A. & ORourke, T. & Shirmboh, D. & Alexander, J. & Andersen, G. & Kaneko, T. & Venkitaraman, A. & Lopez de Cardenas, Jorge & Nishi, M. & Numasawa, M. & Yoshioka, K. & Roy, A. & Wilson, A. & Twynam, Allan. (2004). Practical approaches to sand management. Oilfield Review. 16. 10-27.

Baillard, F., & Hernandez, N. (2021). A Case Study of Understand Bonaparte Basin using Unstructured Data Analysis with Machine Learning Techniques. EAGE Annual.

Hernandez, M., & Baillard, F. (2019). An effective G&G exploration strategy inspired by a wolfpack. Force workshop.

Hernandez, N., Lucañas, P., Graciosa, J., Mamador, C., & Panganiban, I. (2019). Automated Information Retrieval from Unstructured Documents Utilizing a Sequence of Smart Machine Learning. EAGE Workshop on Big Data and Machine Learning for E&P Efficiency 25 - 27 February.

Tiab, D. D., Erle C. (2012). Petrophysics - Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties (3rd Edition) - 9.36 Porosity as Strength Indicator to Evaluate Sand Production. Elsevier.

Kim, S.H., Sharma, M. M., and Harvey J. F. (2011). A Predictive Model for Sand Production in Poorly Consolidated Sands. International Petroleum Technology Conference. Doi: https://doi.org/10.2523/IPTC-15087-MS.