Forecasting Oil Production with Time-Series Foundation Models - a Benchmark Study against Classical Machine Learning Models

Objectives/Scope:

Recent advancements in time series forecasting with language-based foundation models like Chronos (Amazon), Moirai (Salesforce) and Uni-TS (MIT/Harvard) present exciting opportunities to significantly boost forecasting accuracy. They proved performance boost in other domains (example). In this paper we propose to study the application of the forementioned models in the context of oil production forecasting and benchmark their performance. We evaluate these methods in the oil and gas domain across a range of datasets, forecasting horizons, and historical data lengths and frequency. Our primary objective is to assess the effectiveness of Transformer-based deep learning models compared to the current state-of-the-art (SOTA) Machine Learning (ML) based methods used in oil production forecasting.

Methods, Procedures, Process:

The benchmark encompasses two primary approaches: Transformer-based models and established ML methods like ensemble models (RandomForest, XGBoost), ARIMA and Prophet serving as the baseline. We evaluate these models on various datasets, including publicly available Volve Oil Production dataset (daily and monthly) and a second real oil and gas production dataset. This study assesses the methods with regards to different important scenarios:

(i) Forecasting horizons length: short-term
(ii) Training data size
(iii) Data frequency: monthly

Results, Observations, Conclusions:

The study presents a qualitative and quantitative comparison between the performance of the language-based foundation models and the baseline ML approach. We evaluate the models performance on mainly 3 aspects, accuracy of the predictions using MAPE and MAE, prediction speed and model generalization under diverse conditions.

Novel/Additive Information:

This research focuses on a novel application of Transformer-based deep learning models for oil production forecasting. While ML methods are the current industry standard, this study explores the potential of Transformers to surpass their accuracy. The comparison between these approaches will provide valuable insights into the effectiveness of language-based foundation models for oil production forecasting. Additionally, the evaluation across diverse datasets and forecasting scenarios will contribute to a more comprehensive understanding of model generalizability and limitations.

Dataset:

- Volve dataset: https://www.kaggle.com/datasets/imranulhaquenoor/volve-well-production/data