AI for Oilfield Service Companies in 2025[1]

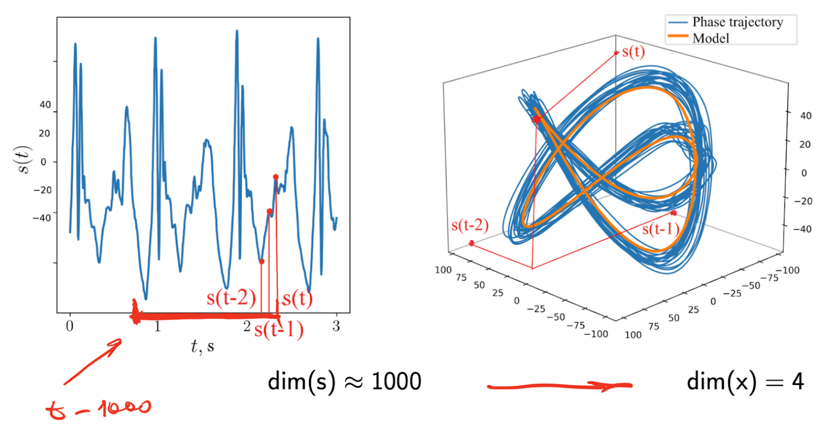
A brief review of the challenges and tools that bring benefits to oil service companies

The petroleum industry generates and stores not only oil and gas but also vast amounts of data to monitor, control, and optimize production processes. In response to evolving demands and regulatory requirements, the industry now faces significant challenges in effectively analyzing and modeling this data. While computer science offers robust tools for processing multimedia data such as text, audio, and video, there remains a notable gap in tools designed for handling data from the diverse array of sensors used to monitor subsurface, surface, and operational parameters throughout exploration, drilling, production, and maintenance. The following list outlines proposed tools tailored specifically for working with oilfield sensor data.

We focus on key areas in oilfield services that demand advanced data analysis tools: 1) flow assurance and 2) carbon capture and storage combined with enhanced oil recovery. These domains present specific data challenges, including 1) the presence of large volumes of time series data from sensors, often noisy, incomplete, and cross-correlated; 2) the critical role of spatial context in analyzing these time series; and 3) the need for multi-modeling approaches due to the diversity and heterogeneity of the sensor types.

Extraction Monitoring and Early Malfunction Forecasting

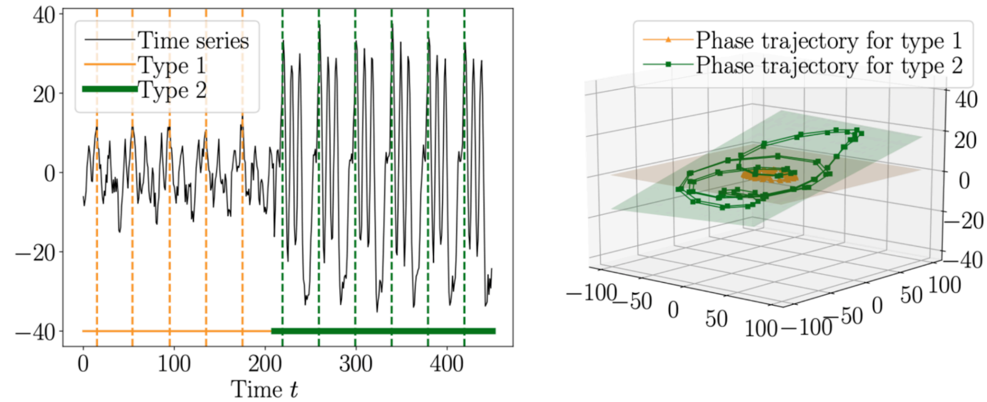
The problem of forecasting equipment malfunctions is the change-point detection within the state space of a time series. This state space encompasses phase trajectories representing normal operating conditions, which can be compared against those indicative of malfunctions. To construct these state-space models, we employ singular spectrum analysis. It effectively defines the system’s phase trajectories and captures its dynamic behavior over time.



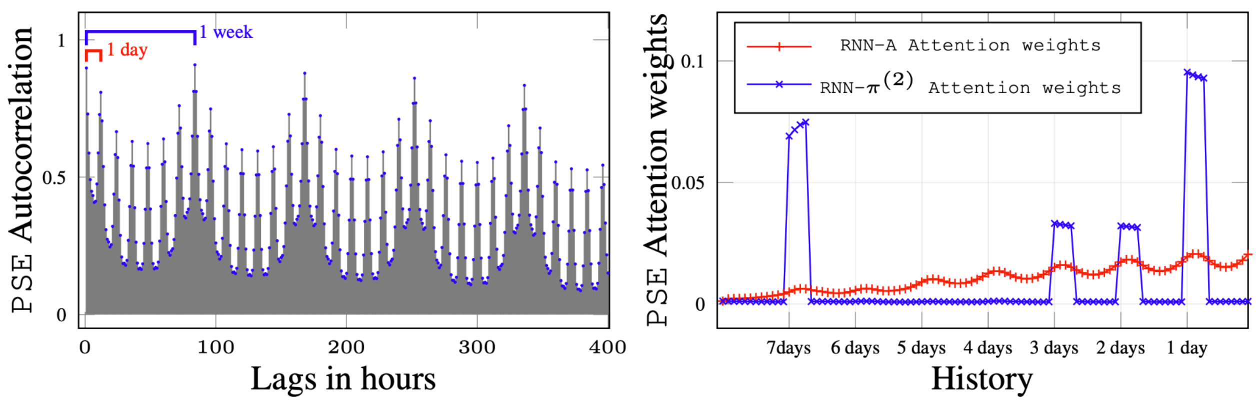
The figure shows how a time segment is represented in the state space. The sequence of the time segments defines the phase trajectory. The dimensionality reduction reveals dependencies in the orange model.

---

[1] By Vadim Strizhov, Jun 13, 2025.



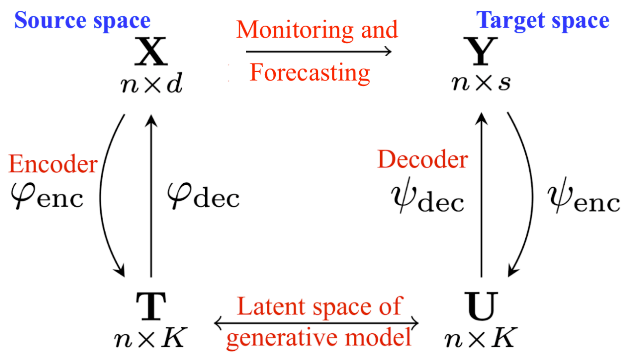
This figure illustrates a change-point (yellow to green) in the sensor time series. It indicates a shift in the operational regime or the onset of a malfunction. The corresponding change is evident in the state-space representation (right panel). To align with classical modeling approaches in the oil and gas industry, we utilize controlled differential equations. These models capture the continuous topology of neural network structures and leverage automatic differentiation techniques. The combination of state-space analysis and controlled differential equations offers a powerful framework for modeling heterogeneous sensor signals.



For long-term monitoring, an operational schedule introduces periodical changes in the network load.

Flow assurance for optimization of production

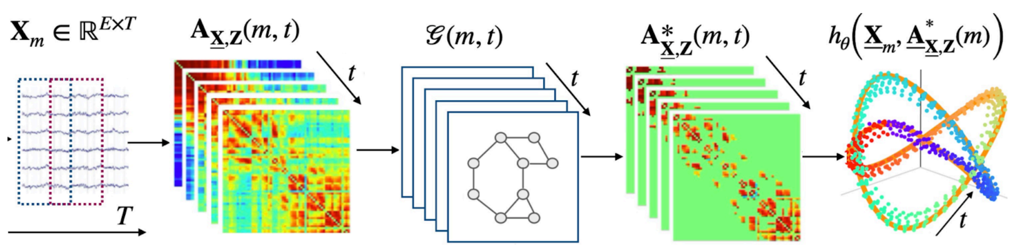
Flow assurance aims to ensure the uninterrupted transport of hydrocarbons from the reservoir to the delivery point. This process involves two domains of sensor measurements: one at the source and the other at the target of each production stage. The system state at any given step is inherently dependent on the state of the preceding step. We denote these measurement sets as the source and the target. The proper model of flow assurance analysis is called the canonical correlation analysis. In deep learning, it is called a sequence-to-sequence model. The advantages of this analysis are 1) it assembles the entire production sequence into a cohesive analytical framework, 2) it boosts the quality of the monitoring, and early malfunction forecasting.



The figure shows the monitoring and early forecasting model. The source space collects phase trajectories of sensor measurements at early stages of production. The target space represents measurements at the current stage. The encoder and decoder extract dependencies in both spaces. The latent space analyzes and classifies the mismatch between two spaces.

Spatial measurement and physics-informed learning

Oil production equipment generates spatial and temporal measurements that are closely tied to underlying physical models. These measurements are collected from various sources, including surface sensors, vibration sensors, multiphase flow meters, hydroacoustic sensors, and weather monitoring systems. To effectively process and analyze this complex data, we employ geometric deep learning and physics-informed models. Spatial time series in this context exhibit high variance and strong cross-correlation, driven by their mutual influence. To model relations between time series and track changes in these relations, we use Riemannian geometry methods, connecting the curvature of space and the graph structure.



The figure shows the composition of models to forecast a set of time series and changes in their cross-dependencies. First, we construct the state space of the time series and the cross-dependency matrix. It changes in time. Second, we extract and prune the graph of clusters and reconstruct the behavior to select sensor measurements with significant forecasting power.

Conclusion

Oilfield service companies generate large volumes of measurement data to monitor and control the production process. Much of this data remains underutilized, despite its potential to optimize operations and enhance production quality and efficiency. Currently, there is a noticeable gap in data analysis tools within this sector – one that urgently needs to be addressed.