

砂層型メタンハイドレートフォーラム 2021

貯留層評価におけるAIアルゴリズム 適用可能性検討

MH21-S研究開発コンソーシアム (MH21-S)

石油資源開発株式会社 岡野祥之

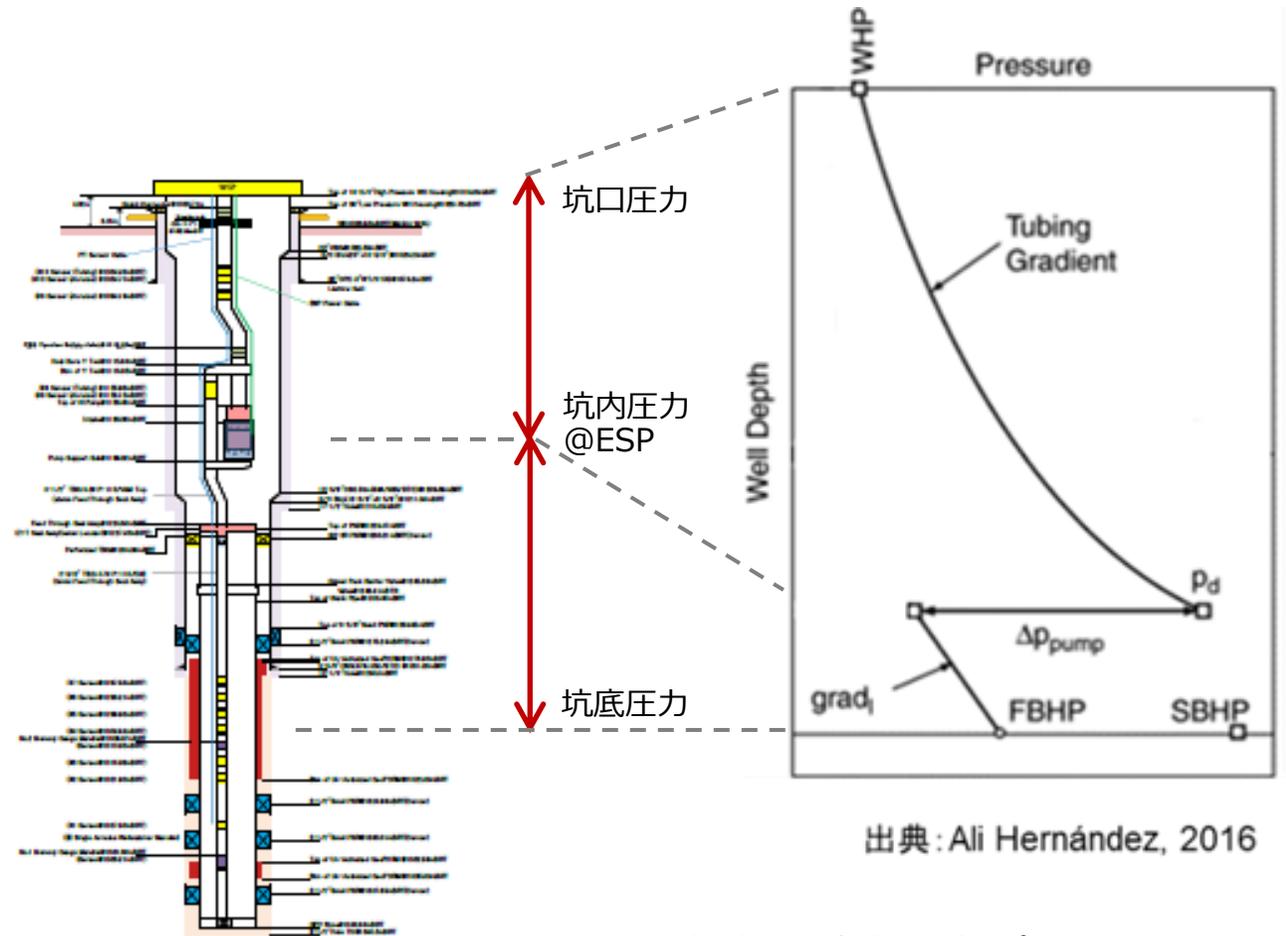
2021年12月1日 (水)

産出試験データの特徴

生産予測にAIアルゴリズムを適用できないか？

メタンガス生成から計量までの長い道のり： 貯留層（生成・流動）～坑内～ 地上生産施設（計量）

✓ 人為操作（バルブ、ESPなど）、ノイズが多い



出典: Ali Hernández, 2016

ESP坑井の流動圧カプロフィール

2020年度検討業務

業務名：砂層型メタンハイドレート貯留層評価におけるAIアルゴリズム適用に向けた課題
解決策の検討業務

実施期間：2020年7月22日～2021年2月26日

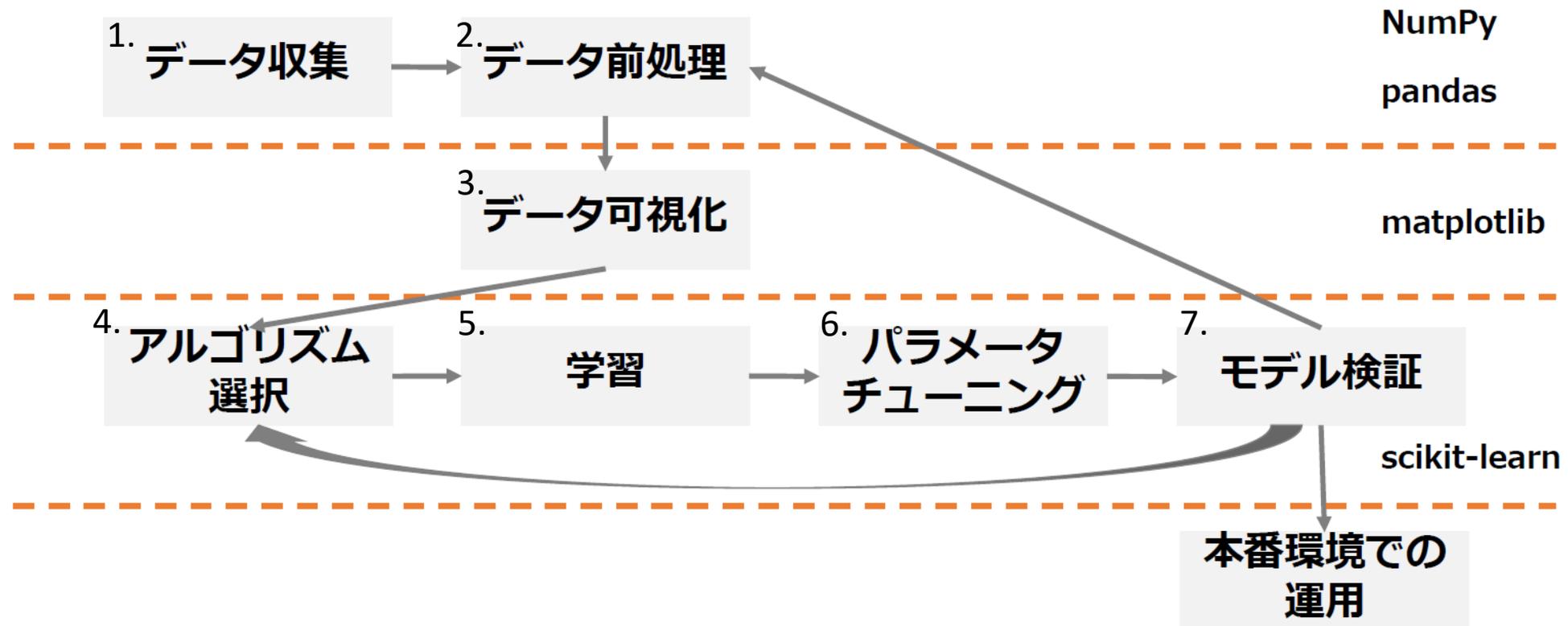
業務内容：

AIアルゴリズム（機械学習）に関する知見、貯留層工学や産出テストのエンジニアリング
サイドの知見を活かして、

1. 砂層型メタンハイドレート貯留層評価へのAI アルゴリズム適用にかかわる課題を抽出・整理する。
2. ガス産出レートを予測できるモデルを作成し、検証手順を取り纏める。

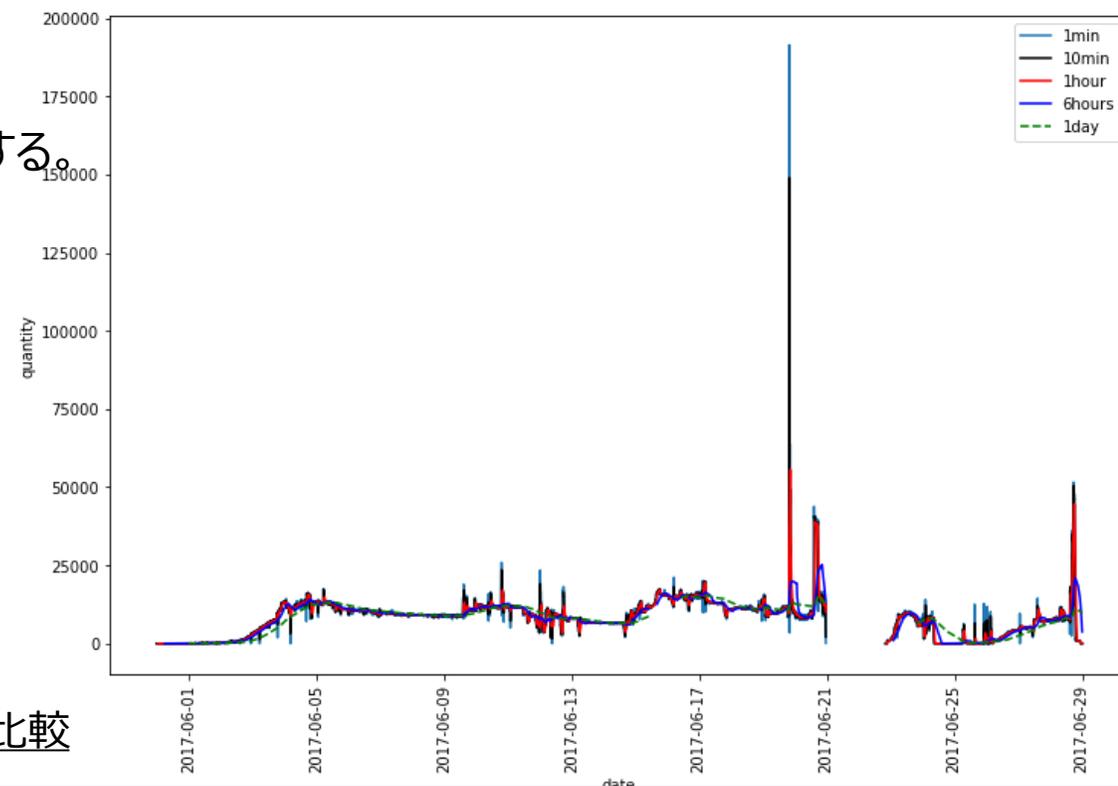
予測モデル作成手順

一般的な機械学習のワークフローに則って、ガス産出レートの予測モデル作成をすすめ、検証手順を取り纏めた。



検証手順：予測モデルを作成する前に（その1）

1. データ収集：取得した時系列データを機械学習で扱える形式に再構築する。
 - ✓ 欠損があるか確認する。
2. データ前処理：ノイズの多いデータの傾向をつかむために、必要な処理を行う。
 - ✓ 人的操作等に起因する外れ値を処理する。
 - ✓ ノイジーなデータの傾向をつかむため、移動平均値を採用する。
 - ✓ データの欠損期間を学習期間から除く。
 - ✓ 必要に応じて欠損データの推計や補完を行う。

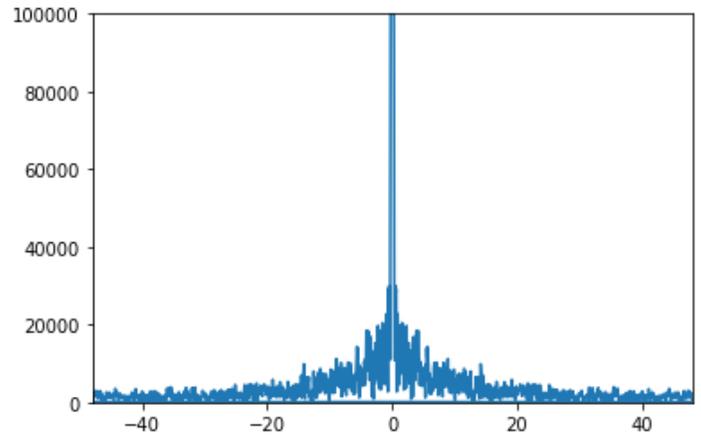
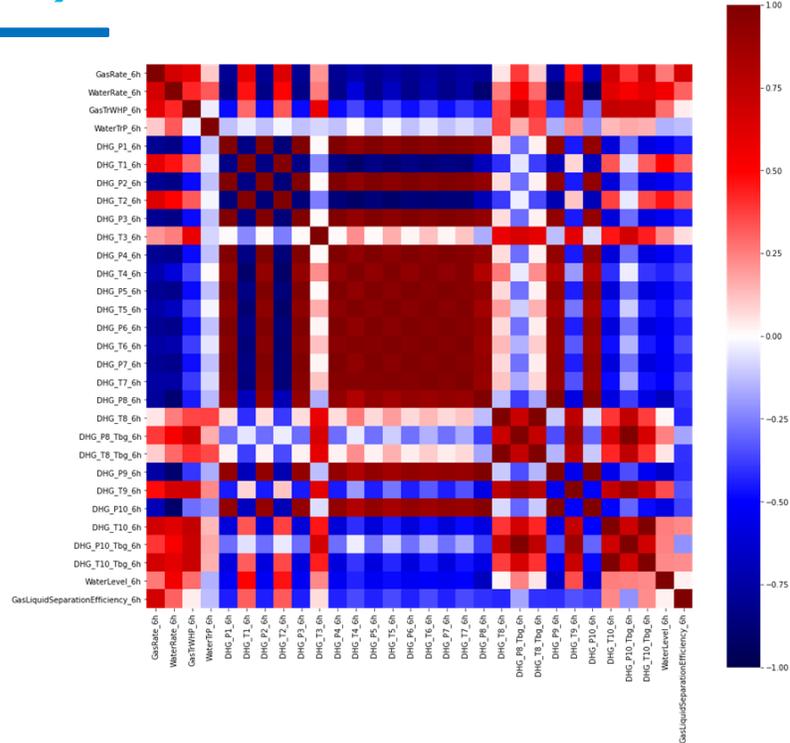


ガス産出レートの移動平均化・比較

検証手順：予測モデルを作成する前に（その2）

3. データ可視化・分析：データ値の分布や周期性の有無等を確認し、データの特徴を把握する。

- ✓ ヒストグラムでデータの数値の分布や、高速フーリエ変換により周期性の有無等を確認し、データの特徴を把握する。
- ✓ ヒートマップ等を活用してデータどうしの関連性を確認し、同じ傾向を持つ（相関性が高い）ものが重複している場合は、多重共線性（説明変数間で相関係数の絶対値が大きくなるときに起こるやっかいな現象）により上手く回帰モデルを構築できなくなることがあるため、代表的なデータをピックアップして説明変数に採用する。



ヒートマップ（元データどうしの関連性）
 赤色：正の相関、青色：負の相関
同じ傾向を持つ（相関性が高い）ものが重複している

ガス産出レート（6月5日～6月17日）の高速フーリエ変換
 横軸は1/daysなので、0にピークがたっている結果は、無限遠にしか周期性がない、すなわち**周期性がない**ことを意味する。

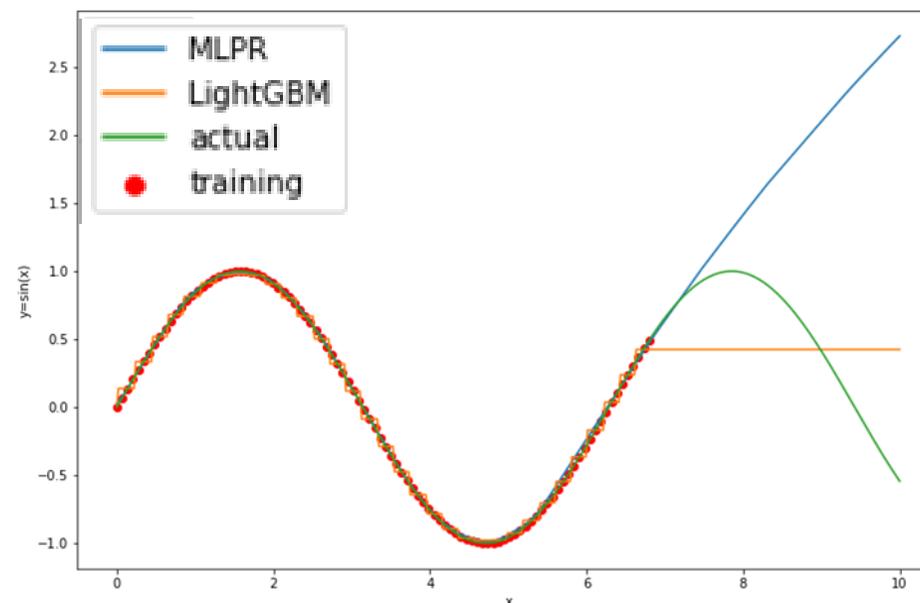
検証手順：予測モデル作成にあたって

4. アルゴリズム選択：データ分析をもとに、目的関数の特性とアルゴリズムの適性を考慮した上で選択する。

- ✓ データ分析をもとに、目的変数の周期性や説明変数のラグ性（遅延効果など）を組み込むかどうかにより時系列回帰とするか、あるいは時系列性を考慮しない回帰にするか選択する。
- ✓ ガス生産レートなどの減退傾向の予測は、外挿領域の予測になる可能性がある。外挿になってもある程度予測できる可能性が高いニューラルネットワーク（多層パーセプトロン回帰（MLPR））が適する可能性が高い。

Sinカーブの学習と予測の比較

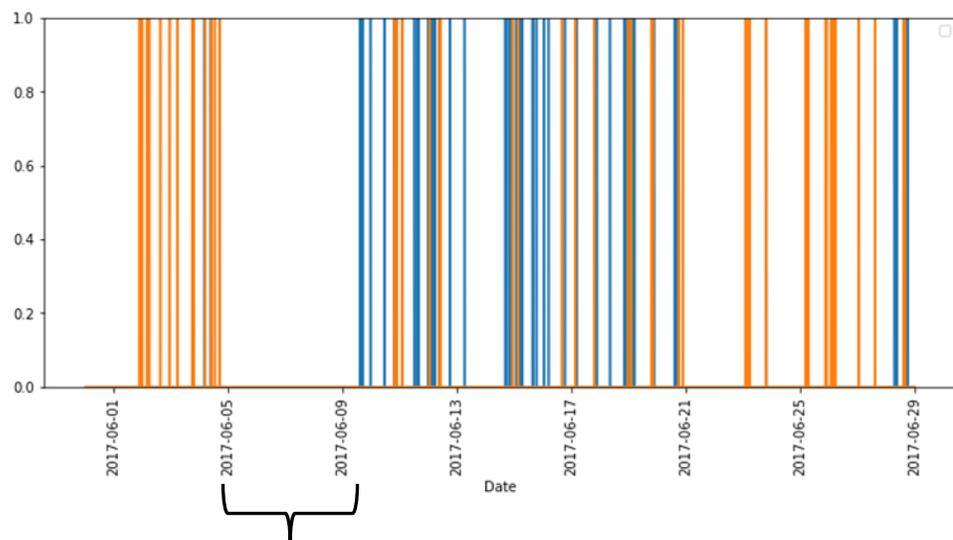
LightGBMとMLPRを使ってSinカーブを学習し予測させたところ、LightGBM（橙色実線）の方は、学習データから外れた途端に一直線に伸びて行き、外挿領域を予測するのは難しいことがわかる。



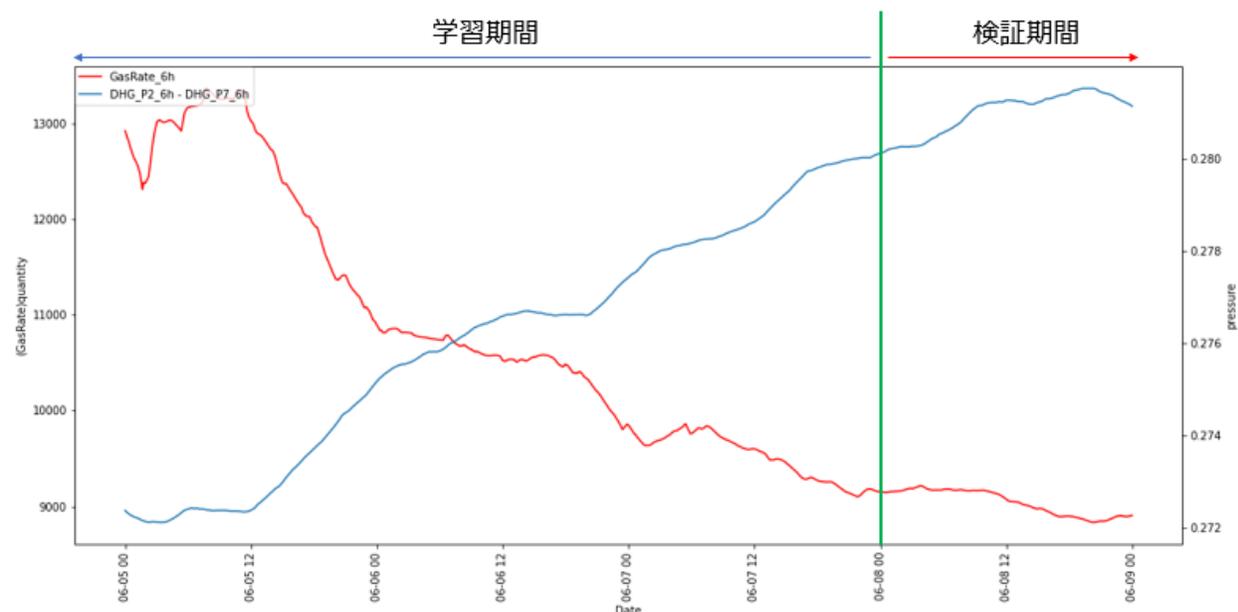
検証手順：上手く予測モデルを作成するには

5. 学習：目的等に応じて汎化性能（未知のデータに対する予測精度）の高い予測モデルが作成できるよう、**学習・検証期間**を適切に設定する。

- ✓ 産出テスト初期（坑内液面をコントロールしていた過渡期）を学習期間から除く。
- ✓ 人的操作に起因して坑内流動と地上で計測される目的変数（ガス生産レート）に関連性がない期間は学習期間に設定しない。



ニードルバルブ操作（橙）、オリフィスプレート交換（青）を行っていない期間を学習・検証期間に設定



ガス産出レートとセンサP2-P7（6時間移動平均）の推移
 学習期間：2017/6/5～2017/6/7、検証期間：2017/6/8

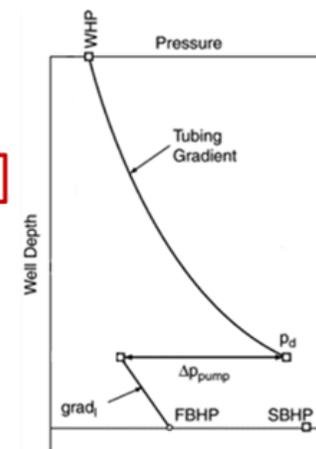
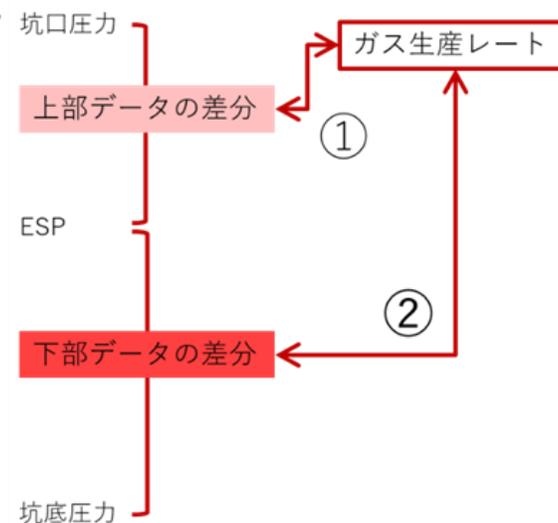
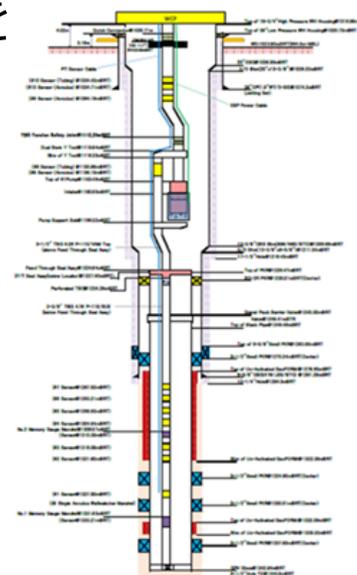
検証手順：モデルを作成したら

7. **モデル検証**：検証期間の予測値と実績値の精度を比較する。精度が低い場合は、
 2. データ前処理の手順に戻り、特徴量エンジニアリングによる新たな説明変数の追加や、学習・検証期間の設定の変更等を実施し、再度予測モデル作成を実施する。

検証手順：予測精度を改善するには

8. 特徴量エンジニアリング：貯留層工学的知見や産出テストの操業条件に関するドメイン知識を用いて特徴量エンジニアリング（説明変数の追加など）を行う。これにより予測に効果的な説明変数を作り出し、モデル精度の向上を図る。

- ✓ 坑内に設置されたESPより下部の貯留層の流動挙動に近いデータ、坑内に設置されたESPより上部でESPの制御の影響が高いデータに分け、説明変数を設定する。
- ✓ 流動条件の違いによって期間を区切り、それぞれの期間での目的変数と説明変数の関連性を把握した上で特徴量エンジニアリングを行う。

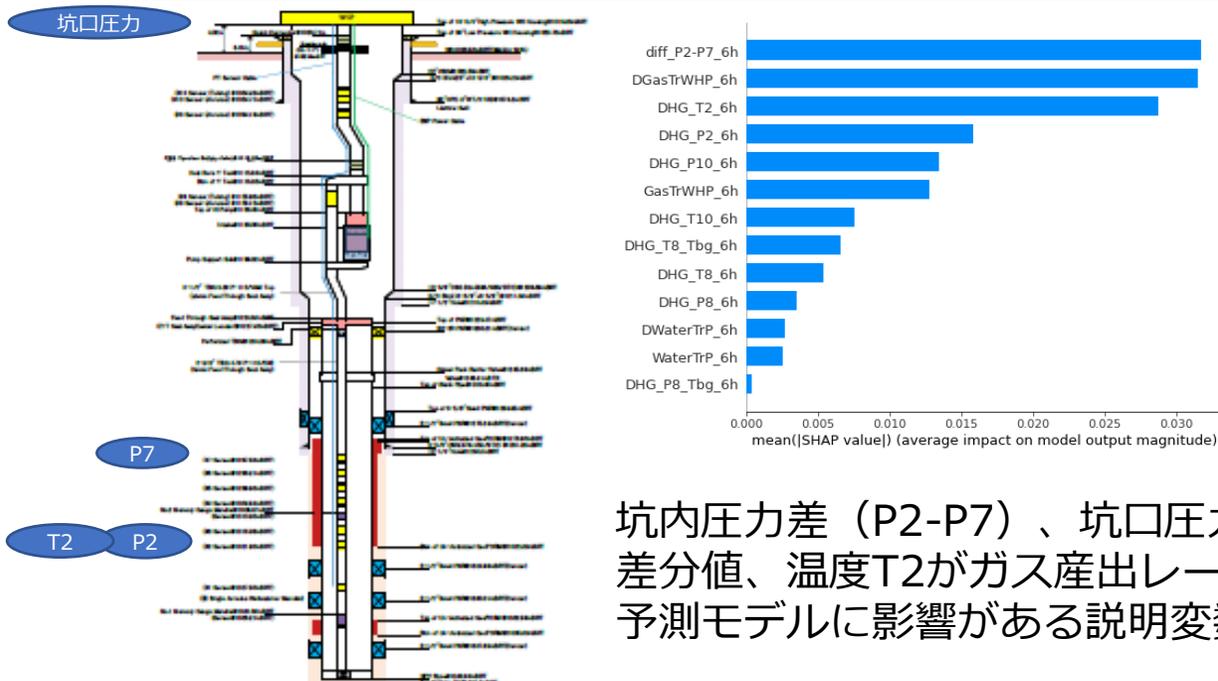


出典：Ali Hernández, 2016

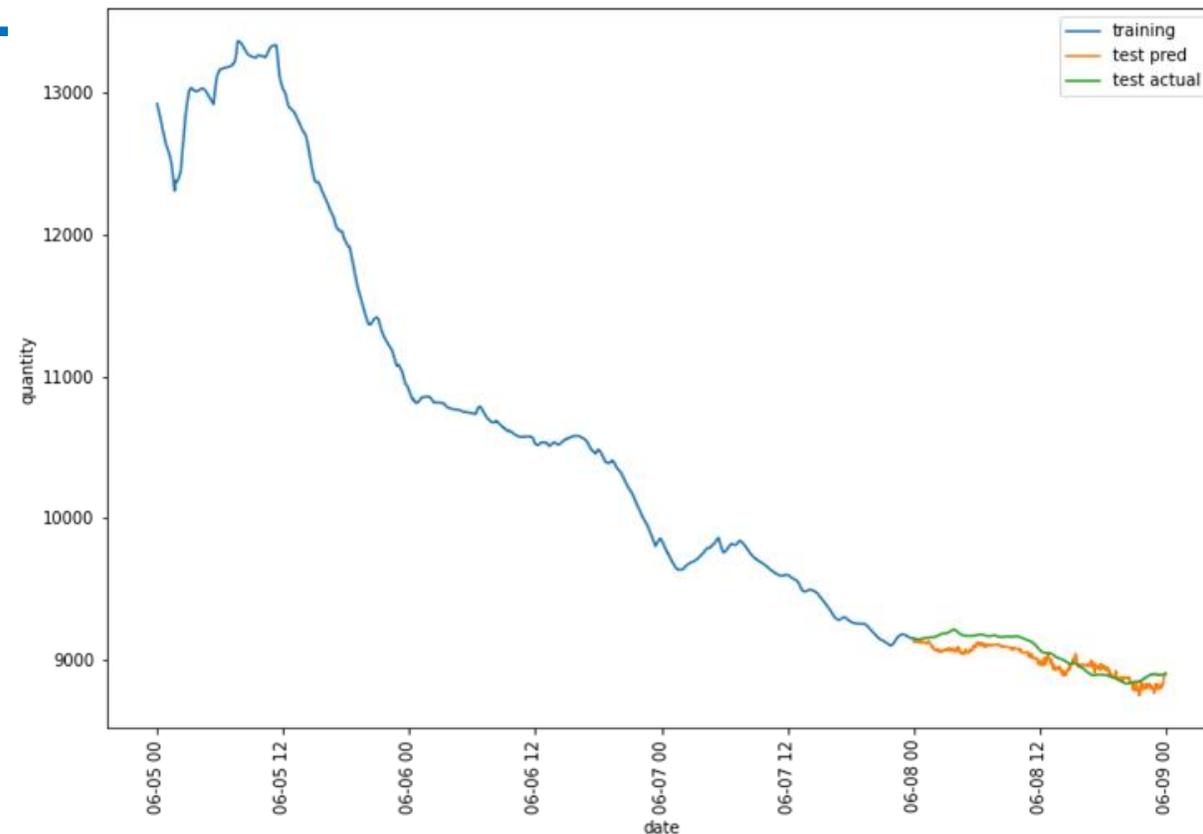
ESP坑井の流動圧カプロファイル

作成した予測モデル

項目	設定
使用するデータ期間	学習期間：2017年6月5日～6月7日 検証期間：2017年6月8日
目的変数	ガス産出レート（6時間移動平均）
説明変数	坑口圧力、坑口圧力差分値 P2、P8、P8_Tbg、P10、diff_P2-P7 T2、T8、T8_Tbg、T10 ※すべて6時間移動平均値
予測手法・評価指標	アルゴリズム：MLPR 評価指標：MAE ※ハイパーパラメータの探索にGridSearchCVを使用



坑内圧力差（P2-P7）、坑口圧力差分値、温度T2がガス産出レート予測モデルに影響がある説明変数



ガス産出レート（6時間移動平均）の学習・検証結果

※ 青：学習期間の実測値
緑：検証期間の実測値
橙：予測値

まとめ

AIアルゴリズムに関する知見、貯留層工学や産出テストのエンジニアリングサイドの知見を活用し、

- 砂層型メタンハイドレート貯留層評価におけるAI アルゴリズム適用可能性における課題を抽出・整理した。
- ガス産出レートを予測できるモデルを作成し、検証手順を取り纏めることができた。

本資料は経済産業省の委託により実施しているメタンハイドレート研究開発事業において得られた成果に基づいております。