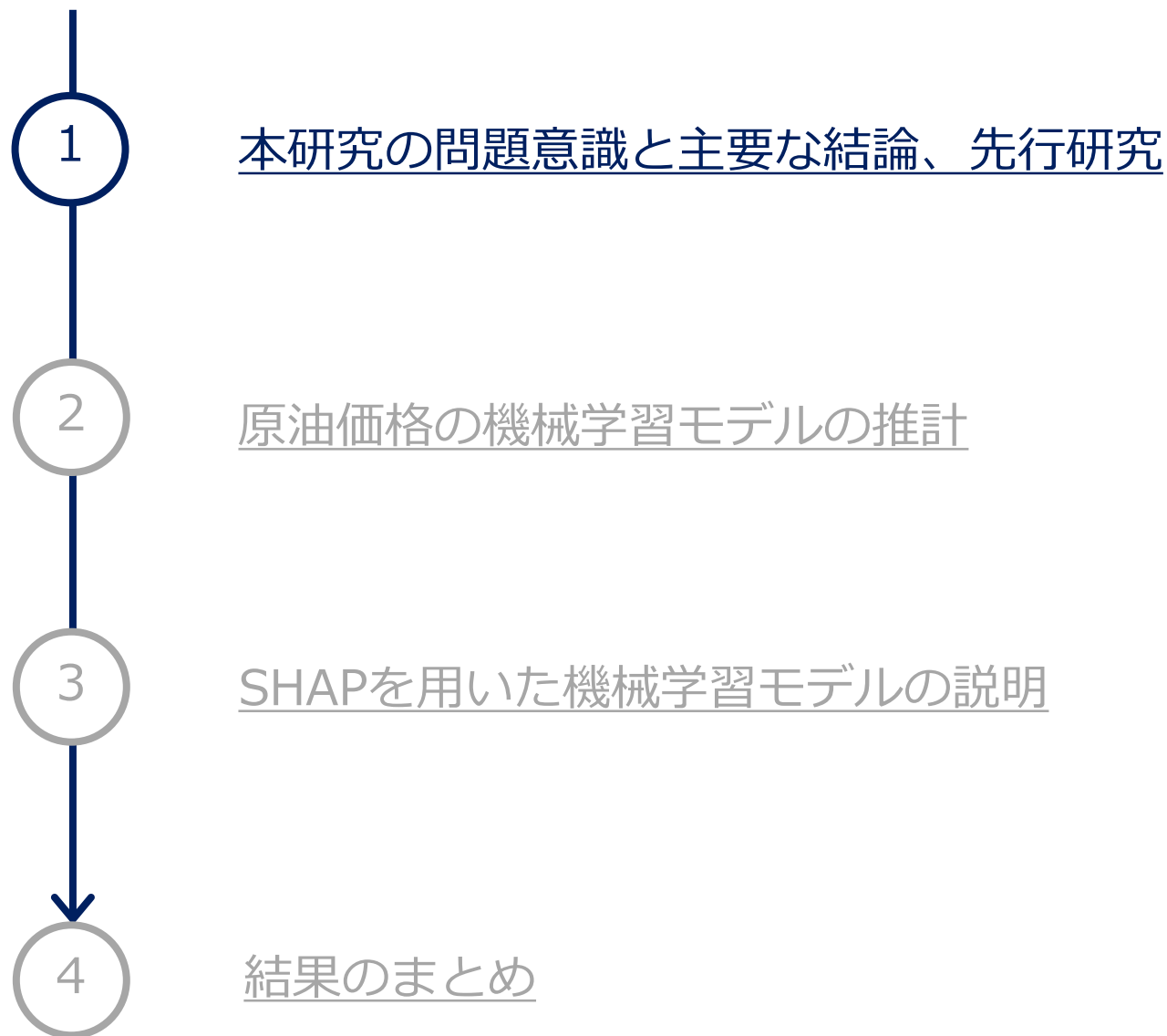


SHAPを用いた機械学習モデルの解釈 — 原油価格の変動要因分析を例に —

2022年11月11日

日本銀行金融研究所	金田 規靖
	木全 友則
	平木 一浩
住友生命保険	松栄 共紘

※本資料に示されている意見は、筆者ら個人に属し、日本銀行および住友生命保険の公式見解を示すものではない。また、ありうべき誤りはすべて筆者個人に属する。本資料はファイナンス・ワークショップ参加者限り。



本研究の問題意識と概要

- 近年、ファイナンスを含む幅広い分野において、AI・機械学習モデルを用いたデータ分析や研究が盛んになっている。
- 実務への利用にあたっては、**機械学習モデルの説明可能性**が課題として指摘。
 - 説明可能性とは、モデル予測値の要因や根拠を理解し、説明できる状態（大坪他[2021]）。
 - 予測精度やデータへのフィットが良くても、予測値がどの要因（説明変数）によりもたらされたかを理解できなければ、意思決定や対外説明に利用できないといった問題が生じる（総務省「AI開発ガイドライン」や「G20 AI Principles」などが指摘）。
- 機械学習の「ブラックボックス化」に対応するため、近年、**説明可能なAI（XAI: eXplainable AI）**への注目が高まっている。SHAP等のモデル説明手法の研究や、それらを用いたファイナンス・経済分析（Buckmann et al. [2022]等）が進展。
 - SHAPは「時点毎・説明変数毎に、モデル予測値を寄与度分解し、定量化・可視化することができるモデル説明手法」。
- 本研究では、「原油価格」を題材に、機械学習モデルの推計とモデル説明のための分析ワークフローを示したうえで、その有用性と実務上の留意点をまとめる。
 - 本研究では、アウトオブサンプル予測の検証ではなく、過去データの要因分析と情勢判断などへの活用を想定し、インサンプル予測精度の検証と説明可能性に焦点を絞って分析。

原油価格の予測や変動要因に関する先行研究

- ①原油価格の変動要因は、複合的かつ、局面毎に非線形関係を有する可能性
- ②原油価格の予測・要因分解は、構造VARを用いた分析が主流
- ③近年、機械学習モデルが原油価格予測で優れた精度を示している

	先行研究	手法	概要
①	Hamilton (2003)	線形モデル、 非線形モデル	原油価格の主要な変動要因は、供給や需要など複合的であり、非線形性を考慮できるモデルが有用であると指摘。
②	Kilian (2009)	構造VAR	実質原油価格、原油生産、原油需要の3変数構造VARを構築し、要因分解を行った。変動要因の把握は、経済への影響や政策対応を検討するために重要と言及。
	Fueki et al. (2016)	構造VAR	Killian (2009)を拡張し、将来の需給要因や金融要因を追加すると精度が向上。局面毎に主要な変動要因が異なることを示した。
③	Costa et al. (2021)	線形モデル、 機械学習モデル	原油価格の予測精度の比較を実施（線形モデルや機械学習モデルなど22種類。経済金融の大規模データセット）。短中期予測では、機械学習モデルの予測精度が高かった。重要な説明変数の特定などモデルの結果解釈を試みている。
	Calomiris et al. (2022)	線形モデル、 自然言語処理	ニュース記事から、自然言語処理によりエネルギー・ニュース指標を作成し、説明変数に使用したところ、アウトオブサンプル予測の精度向上が実現した。

→ ただし、原油価格の機械学習モデルの説明可能性に踏み込んだ研究は筆者の知る限り、知られていない。また、原油価格の変動要因分解（需要・供給・金融要因など）は中銀にとって関心の高いテーマ。

機械学習モデルの説明手法（SHAP）に関する先行研究

概念の導入 手法の提案

協力ゲームにおいて「Shapley値」概念が導入 Shapley (1953)

モデル説明に対するゲーム理論の応用可能性 Strumbelj and Kononenko (2010)

モデル説明手法「SHAP」の提案 Lundberg and Lee (2017)

ファイナンス・経済 分析への利用

SHAPを用いた金融経済分析 Bluwstein *et al.* (2021)

- SHAPを用いて、金融危機の予測に重要な金融経済変数（イールドカーブの傾きなど）を可視化し、複数の変数が相互関連し、複合的に金融危機リスクを高める可能性を示した。

分析ワークフロー の提示

説明可能性を意識した機械学習の分析ワークフローを提示 Buckmann *et al.* (2022)

- 米国失業率を対象に、モデルの予測精度の比較評価、SHAPによる予測値の要因分解・可視化、SHAP値の統計的検証（Shapley Regression）などを実施。ドメイン知識により変数を絞り込んだ場合、機械学習モデルの予測精度と説明可能性の最適なバランスが達成されることを示唆。

リサーチクエスションと主要な結果

■ Q1：機械学習は、伝統的手法と比較して、予測精度の高いモデルを推計できるか？

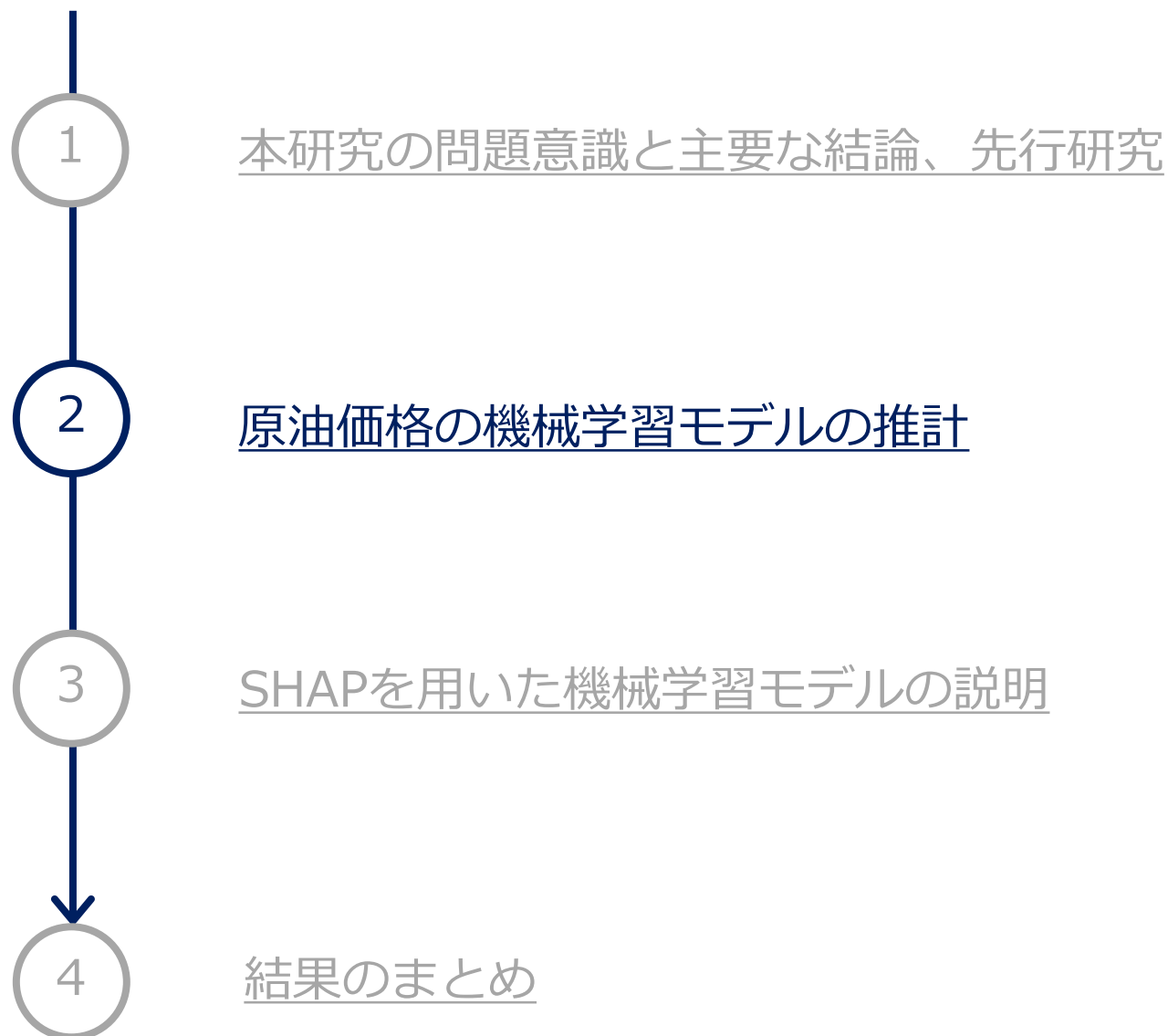
- ✓ A1：機械学習モデルは、構造VAR等と比較して、高いインサンプル予測精度を示した。多くの説明変数（特徴量）を用いたり、原油価格と説明変数の非線形関係を扱えるため。

■ Q2：機械学習モデル選択時に、説明可能性をいかに検証し、勘案すれば良いか？

- ✓ A2：SHAPを用いれば、機械学習モデルの可視化や統計的検証が可能となる。
機械学習モデルの選択は「予測精度」と「説明可能性」のトレードオフを踏まえ、双方の観点から、最適なモデルを選択することが重要。

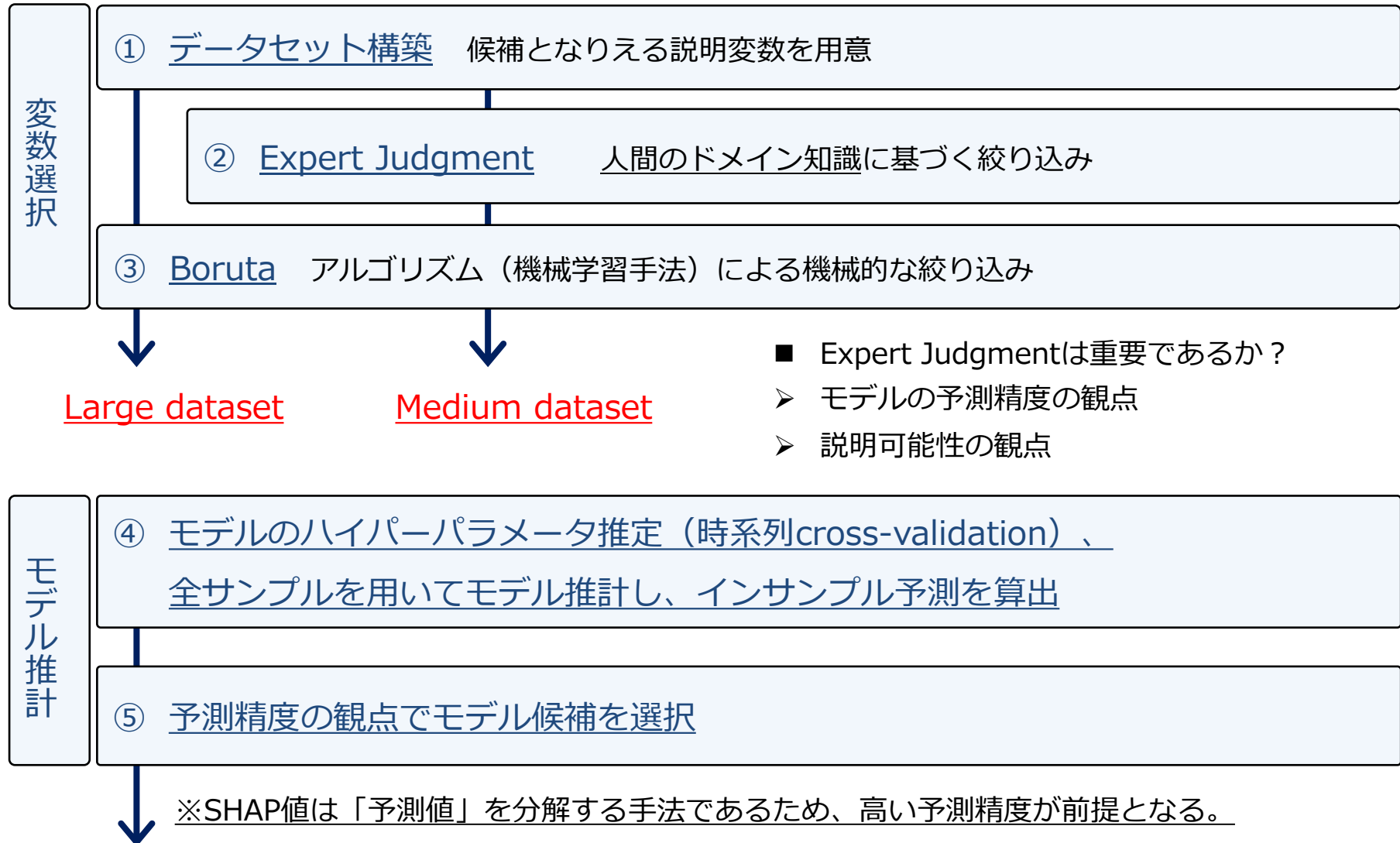
■ Q3：機械学習モデルは、原油価格の変動要因を分析できるか？

- ✓ A3：SHAPを用いて可視化された原油価格の変動要因をみると、原油需給およびマーケットが主因であり、金融政策（FRBのバランスシートサイズ）も大きく寄与していた。経済情勢を映じ、局面毎に各変動要因の寄与度や符号が変化していく結果が示された。こうした「データに語らせる分析」は解釈に留意が必要であるが、構造モデルを補完する有益なツールとなりうる。



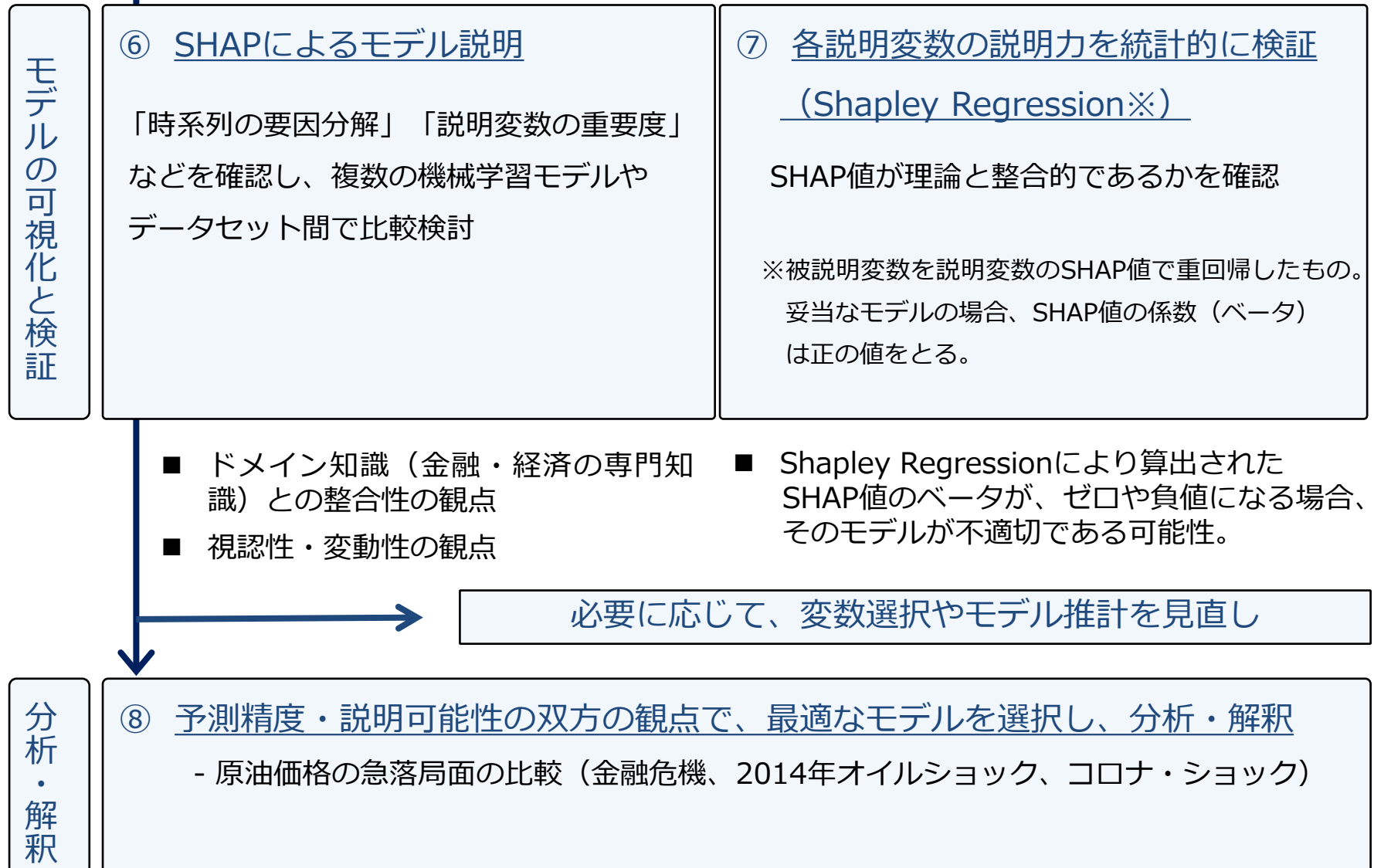
機械学習モデルの推計と説明のための分析ワークフロー

■ 機械学習モデルの推計：第2節



機械学習モデルの推計と説明のための分析ワークフロー

■ 機械学習モデルの説明：第3節



データセット

- 被説明変数（**実質ブレント原油価格**）：1 か月先の水準（1991/2月～2021/6月の月末値）
- 説明変数：3 か月前からの変化率、6 か月前からの変化率（ただし、金利などは変化幅）
 - **Large dataset**
 - ー 金融経済指標（Costa et al. [2021]）やエネルギー関連指標などのデータセット（319説明変数<638変数>）をBorutaにより絞り込んだ**75説明変数<97変数>**
 - **Medium dataset**
 - ー 上記データセットをExpert Judgmentに基づき変数選択し（38説明変数<52変数>）、そのうえでBorutaにより絞り込んだ**15説明変数<19変数>**

説明変数		データセット (絞り込み前)	Large (Boruta)	ドメイン知識 で絞り込み	Medium (Boruta)
原油供給	原油生産等	42	5	8	4
原油需要	原油消費	10	-	4	2
原油在庫	原油在庫等	20	-	4	-
原油先物	原油先物価格、スプレッド	16	4	8	2
コモディティ価格 マーケットデータ	商品価格、株価、金利等	88	9	22	9
金融政策	中銀バランスシートサイズ	8	2	2	2
経済指標	雇用、物価、生産等指標	454	77	4	-
合計		638	97	52	19

以降、Borutaによる絞り込み後の「**Large dataset (97変数)**」「**Medium dataset (19変数)**」で推計。

変数選択

■ Medium dataset (Borutaによる説明変数選択後) : 15説明変数<19変数>

原油供給	マーケットデータ
OPEC加盟13か国 原油生産 (3か月変化率)	米国債5年金利 (6か月変化率)
OPEC加盟13か国 原油生産 (6か月変化率)	ユーロ・ストックス50指数 (3か月変化率)
OECD加盟国 原油生産 (6か月変化率)	S&P500エネルギー株指数 (3か月変化率)
その他国 原油生産 (6か月変化率)	投機筋WTIネットポジション (3か月変化率)
原油需要	投機筋WTIネットポジション (6か月変化率)
OECD加盟国 原油消費 (3か月変化率)	フィラデルフィア連銀ADS Index (6か月変化率)
OECD加盟国 原油消費 (6か月変化率)	金融政策
原油先物価格	FEDバランスシート (3か月変化率)
原油1か月先先物価格 (3か月変化率)	FEDバランスシート (6か月変化率)
原油3年先先物価格 (3か月変化率)	
コモディティ価格	
S&P GS 商品指数 (エネルギー) (3か月変化率)	
CRB 商品先物指数 (3か月変化率)	
CRB 商品先物指数 (油脂) (3か月変化率)	

(出所) EIA、FRB、Bloomberg

※Large dataset (Borutaによる説明変数選択後) は、Appendix (スライド41) を参照。

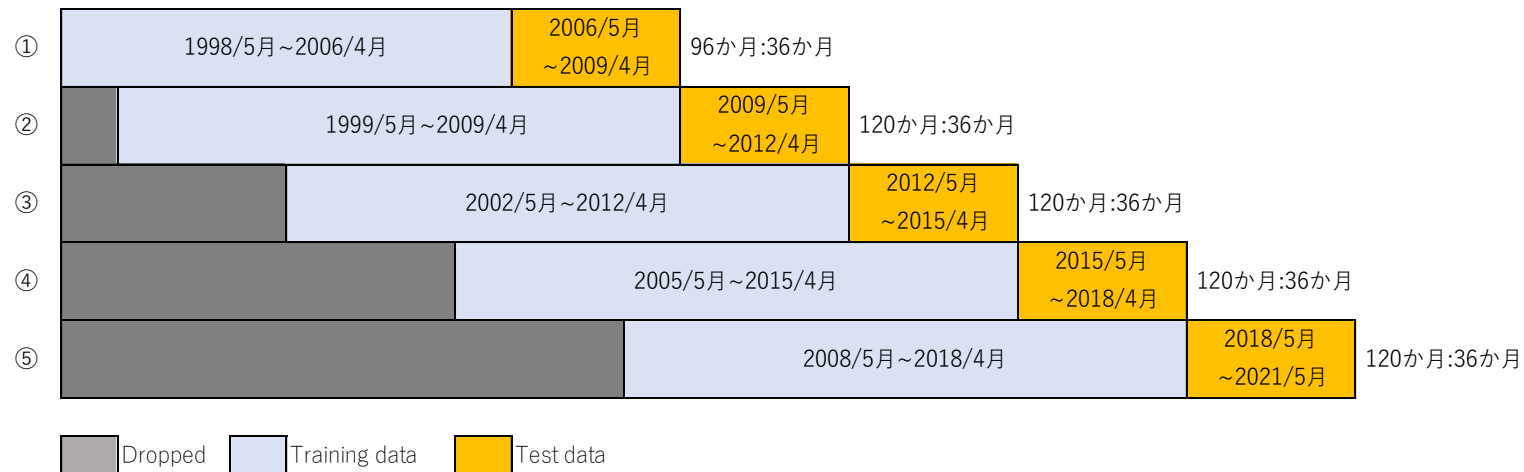
機械学習モデルの推計

■ 機械学習モデル

- ✓ Random Forest、Light GBM
- ✓ Ridge、LASSO、Elastic Net、Support Vector Machine
- ✓ Neural Network（3層モデル）

■ 時系列クロスバリデーション（ハイパーパラメーターを決定） <Bergmeir and Benítez [2012]>

（１）データを5分割し、それぞれをTraining Data（最大10年間）とTest Data（3年間）に振り分け。



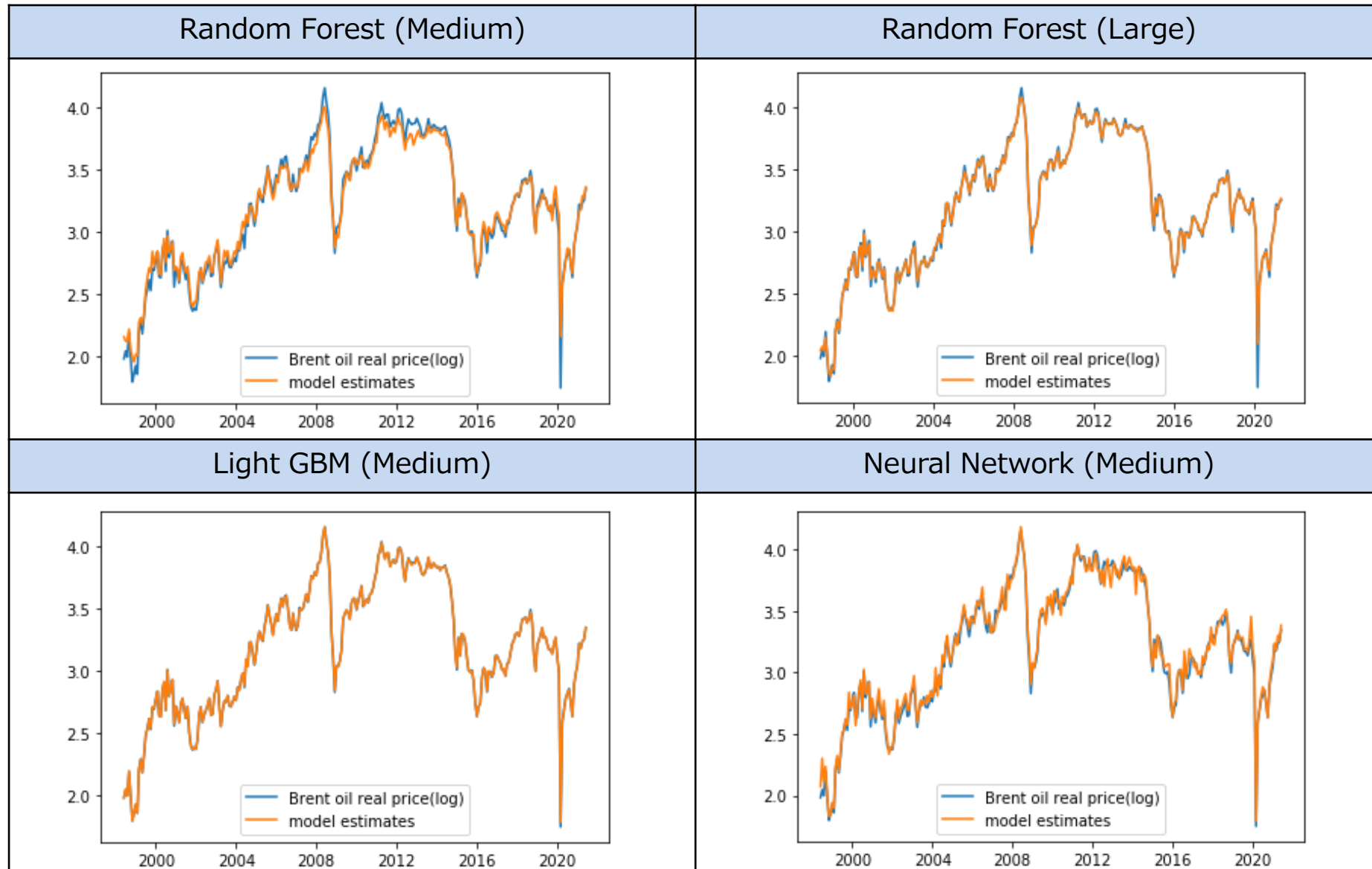
（２）各サンプルを用いて、予測誤差を最小化するモデルパラメーターを個別に推計。

（３）各サンプルの学習結果を踏まえ、全期間のインサンプル予測誤差を最小化するようなハイパーパラメーターを決定。

（４）推計したモデルに全期間のデータをフィットさせ、インサンプル予測値を算出。予測誤差によりモデル予測精度を評価。

機械学習モデルの推計結果（1か月先の原油価格）

■ 予測精度が高い機械学習モデルによる推計値と予測誤差



※原油価格の推計値は対数表示。

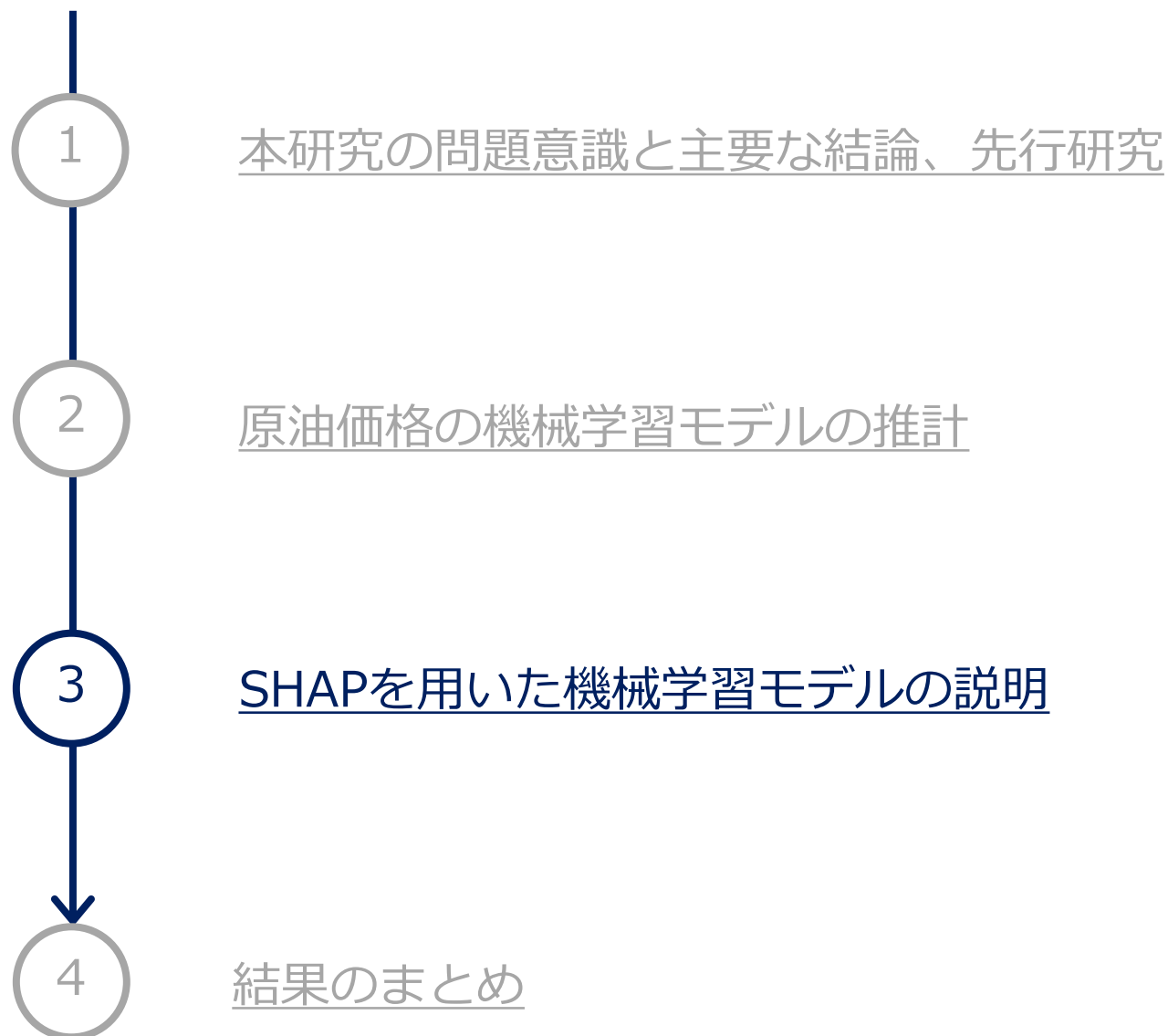
機械学習モデルの推計結果（1か月先の原油価格）

- ✓ 上段の伝統的モデルと比較して、下段の機械学習モデルの予測精度は優れている。
- ✓ 次節では、下段の4つの機械学習モデルを説明可能性の観点から検討する。
 - 予測精度の高いLGBM（Medium）やRF（Large）は説明可能性も高いといえるか？

	Random Walk	ARIMA	5変数VAR	Ridge	LASSO	Elastic Net	SVM
RMSE	0.105	0.081	0.122	0.134	0.132	0.133	0.137
MAE	0.075	0.061	0.101	0.086	0.086	0.086	0.083
R2 score	0.054	0.108	0.176	0.094	0.115	0.108	0.059

	Random Forest (Medium)	Random Forest (Large)	LGBM (Medium)	Neural Network (Medium)
RMSE	0.025	0.020	0.004	0.022
MAE	0.013	0.008	0.002	0.016
R2 score	0.792	0.864	0.996	0.825

※Random Forest（Large dataset）以外はMedium datasetを用いて推計。

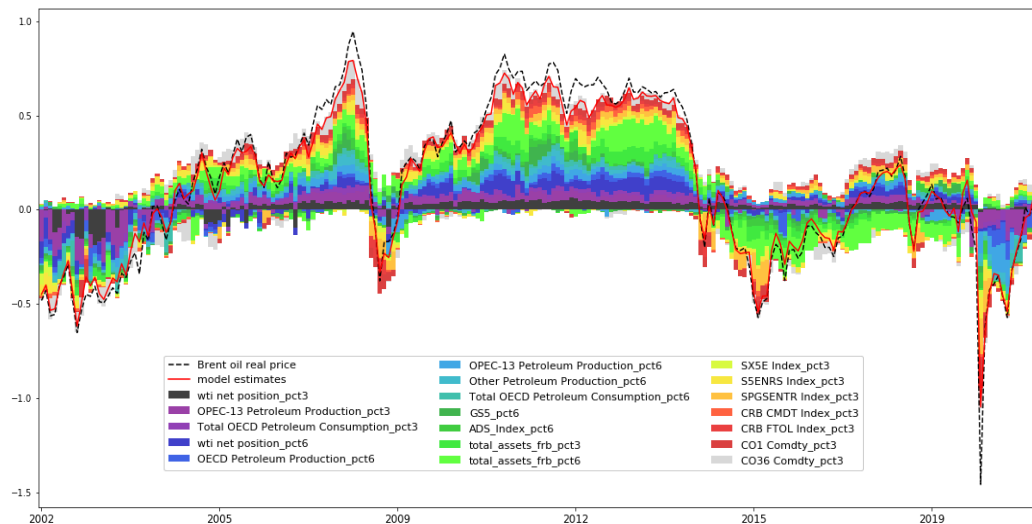


XAIを用いて機械学習モデルを説明するための視点

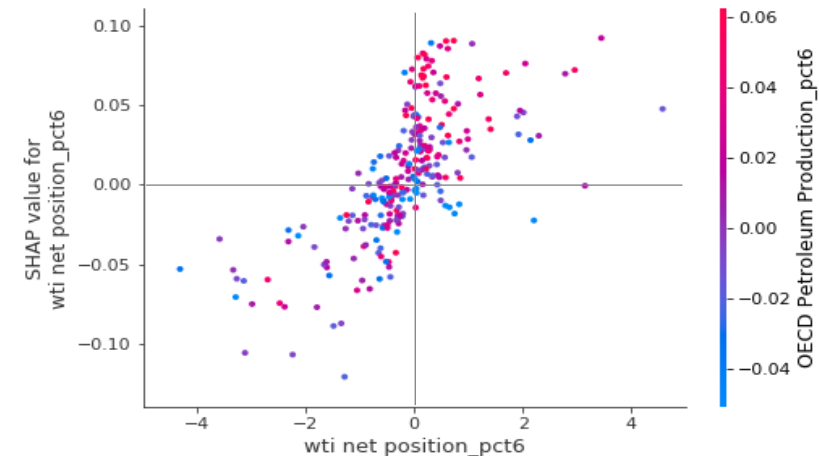
- 機械学習モデルは高い予測精度を達成できる。一方、モデル構造が複雑になるほど被説明変数と説明変数の関係性を理解することが難しくなる傾向がある。
 - ⇒ 一般的に、機械学習モデルが複雑になれば説明可能性は低下する(ブラックボックス化する)。
- XAIは機械学習モデルの予測根拠を可視化し、説明可能性を高めることができる。ただし、モデルの内部構造や計算過程を解析しているとは限らない点には留意。
- XAIを用いて、予測モデルの説明可能性を高めるにあたり、2つの視点が重要となる。
 - ① 予測モデル全体において、重要度の高い説明変数を把握する**大局的視点**
 - 機械学習モデルのパフォーマンス評価・運用改善
 - ② 各時点の予測値において、説明変数の寄与度（影響度）を把握する**局所的視点**
 - 機械学習モデルが想定通りに学習できているか、妥当な挙動をしているか確認
- 近年提案されたモデル説明手法である「SHAP」を用いれば、上記の2つの視点から機械学習モデルの説明可能性を高め、モデル間の比較検討が可能となる。

SHAPによる機械学習モデルの可視化や検証の例

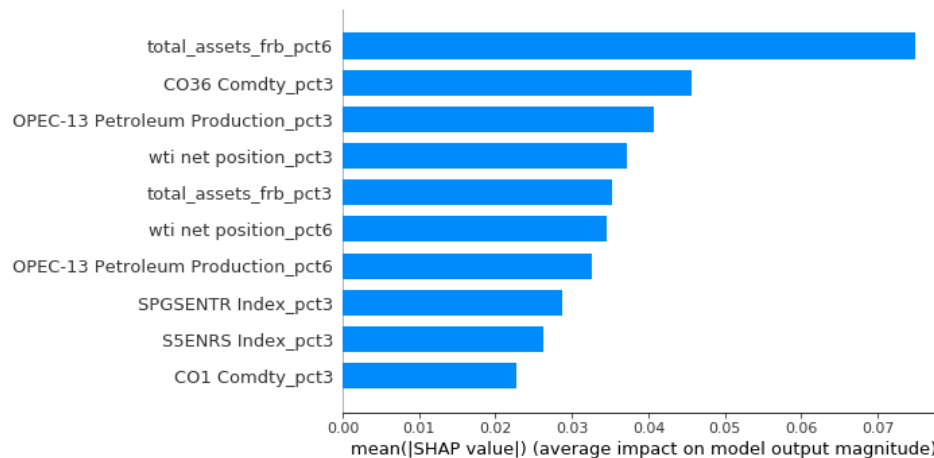
SHAP値の時系列推移



説明変数とSHAP値の関係



説明変数の重要度 (SHAP値の絶対値)



Shapley Regression

$$y = \phi_0^S + \sum_{k=1}^n \phi_k^S(\hat{f}, x) \hat{\beta}_k^S + \hat{\varepsilon}$$

y : 原油価格 $\phi_k^S(\hat{f}, x)$: 説明変数 k のSHAP値

OLSの枠組みで機械学習モデル・SHAP値の妥当性を統計的に検証できる。

機械学習モデルの説明手法：SHAP

- 「**SHAP**(**SH**apley **A**dditive ex**P**lanations)」とは、協力ゲーム理論におけるShapley値のコンセプトを機械学習モデル説明に応用し、予測値に対する説明変数の寄与度を計算する手法。
- 協力ゲーム理論におけるShapley値とは、複数プレイヤーの協力により得られた報酬を各プレイヤーに公正に分配する(貢献度を計測する)ための手段の一つ。

(例) プレイヤーA、B、Cの3人協力ゲームにおける報酬60万円の分配方法

- すべての組み合わせで各プレイヤーの限界貢献度を算出し、それらの平均値を各プレイヤーの報酬として定義

				限界貢献度 (万円)		
				Aさん	Bさん	Cさん
なし	0万円	+ Aさん	10万円	10	20	30
				(10 - 0)	(30 - 10)	(60 - 30)
なし	0万円	+ Aさん	10万円	10	38	12
				(10 - 0)	(60 - 22)	(22 - 10)
なし	0万円	+ Bさん	6万円	24	6	30
				(30 - 6)	(6 - 0)	(60 - 30)
なし	0万円	+ Bさん	6万円	44	6	10
				(60 - 16)	(6 - 0)	(16 - 6)
なし	0万円	+ Cさん	4万円	18	38	4
				(22 - 4)	(60 - 22)	(4 - 0)
なし	0万円	+ Cさん	4万円	44	12	4
				(60 - 16)	(16 - 4)	(4 - 0)
平均				25	20	15
				万円	万円	万円

(出所) DataRobot社 ブログ記事

<https://www.datarobot.com/jp/blog/explain-machine-learning-models-using-shap/>

SHAPの特長

- あらゆる機械学習モデルでSHAP値は計算可能 (モデルフリーな手法) で、モデル比較が容易。
- SHAP値は加法性を持つため、SHAP値を足し上げるとモデル予測値に一致する。
- この他、対称性、線形性、強単調性などの理論的に望ましい性質を持つ。

(参考) SHAP値の一般化式

・ Additive feature attribution method

M個の説明変数 $X = (X_1, \dots, X_M)^T$ からなる機械学習モデル $f(X)$ の各地点の予測値 $f(x)$ を要因分解する。

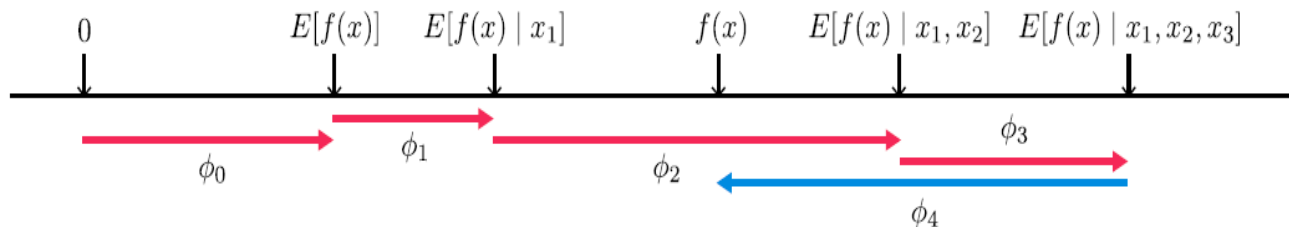
⇒ モデルの平均的な予測値 $E[f(X)]$ と各地点の予測値 $f(x)$ の差を各説明変数の限界貢献度 ϕ_j に分解する。

$$f(x) - E[f(X)] = \sum_{i=1}^M \phi_i \quad \xrightarrow{\phi_0 := E[f(X)]} \quad f(x) = \phi_0 + \sum_{i=1}^M \phi_i$$

限界効果を全ての組み合わせで平均 = Shapley Value

$$\phi_j = \sum_{S \subseteq \mathcal{M} \setminus \{j\}} \underbrace{\frac{|S|! (|\mathcal{M}| - |S| - 1)!}{|\mathcal{M}|!}}_{\text{組み合わせの出現しやすさ}} \underbrace{(v(S \cup \{j\}) - v(S))}_{\substack{\text{\mathcal{S}にプレイヤーjが参加したときの限界貢献度} \\ \text{\text{※ } v(\cdot) \text{ は報酬を表す関数}}}}$$

— 下図は、平均的な予測値 $E[f(X)]$ から各地点の予測値 $f(x)$ に至るまでのある1つの順序を示したもの（4変数の場合）。モデルが非線形であったり、説明変数同士が独立でない場合、説明変数 X を加える順序により、限界貢献度 ϕ_j は変化する。



⇒ そのため、全てのありうる順序で説明変数を加えた場合の限界貢献度を個別計算し、その平均値を各説明変数のSHAP値として算出している。

SHAPを用いた機械学習モデルの比較検証ポイント

■（大局的視点）モデル全体で重要な説明変数は何か。ドメイン知識と整合的なか？

（例）原油価格と関連性が低いと想定される説明変数（小国の経済指標など）が重要と判定されていないか。

■（局所的視点）個別予測値や局面別の要因分解はドメイン知識と整合的なか？

（例）不況期に需要要因が原油価格の押し上げに寄与するなど経済理論と矛盾した変動を示していないか。

■（局所的視点）個別予測値や局面別の要因分解の視認性や変動性に問題ないか？

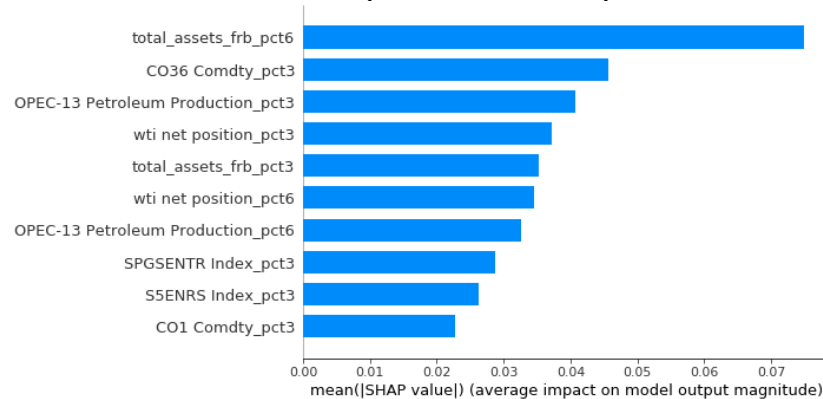
（例）SHAP値に振れが発生（ノイズに対する過剰適合）、符号の向きが頻繁に入れ替わる等がみられていないか。

■（統計的検証）Shapley Regressionによる統計的な検証結果は理論と整合的なか？

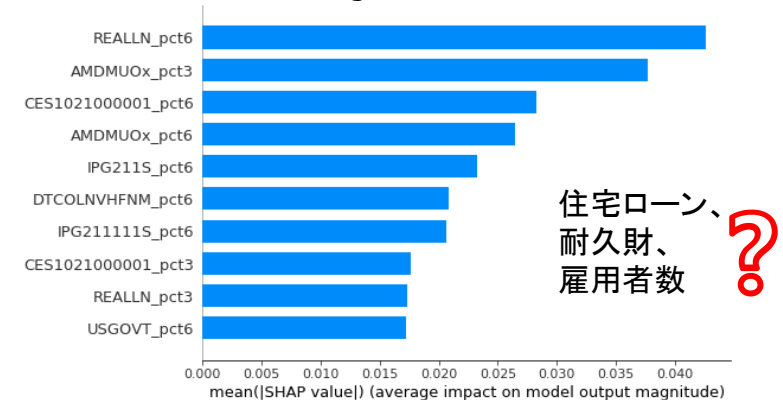
SHAPによる説明変数の重要度 (SHAP値の絶対値平均)

- RF・LGBM・NN (Medium dataset) は「FRBのバランスシート」「原油先物価格」「原油生産」が上位。
 - 右上のLarge datasetは「米国住宅ローン残高」「米国耐久財受注残高」「米国雇用者数（鉱業）」が上位。
- Expert Judgmentで変数選択していないRF (Large dataset) <右上> は、原油価格との関連度が低い説明変数でモデル推計されている可能性。

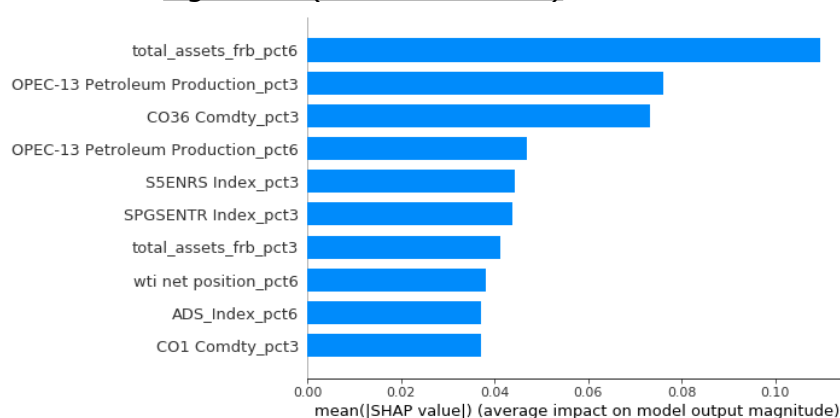
Random Forest (Medium dataset)



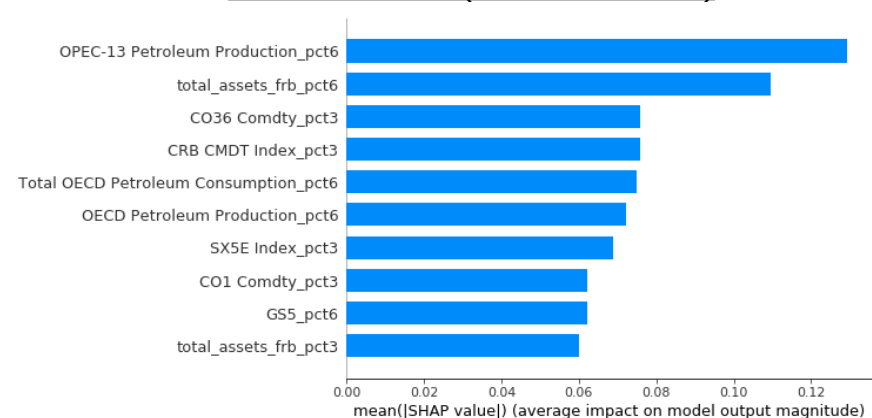
Random Forest (Large dataset)



LightGBM (Medium dataset)



Neural Network (Medium dataset)



SHAP値の統計的検証 (Shapley Regression) ①

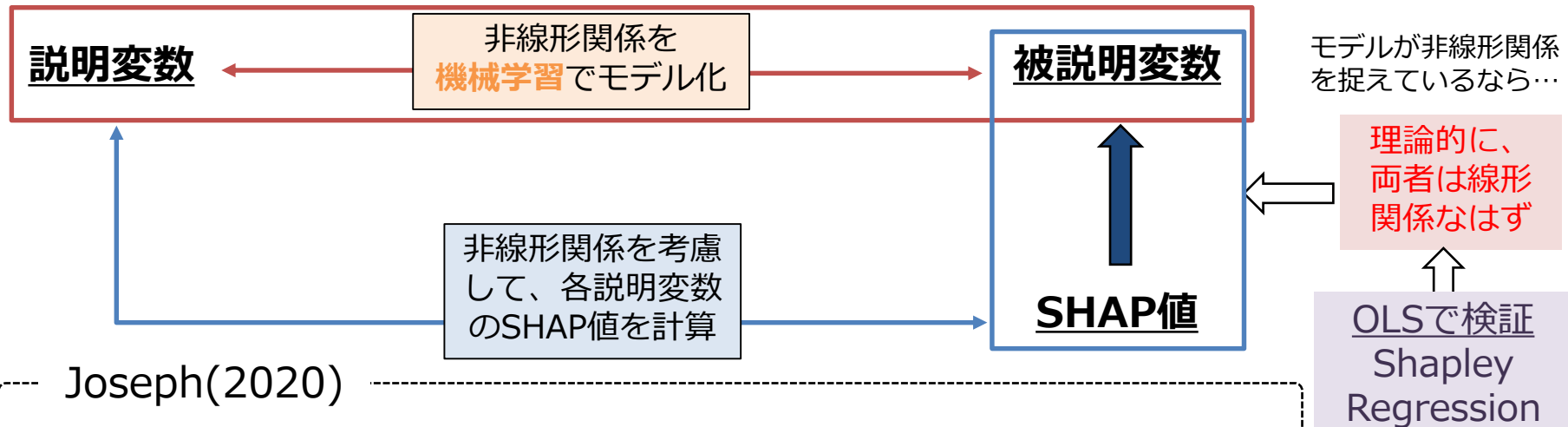
- Shapley Regression : 被説明変数 (原油価格) をSHAP値で重回帰して検証する手法。

$$y = \phi_0^S + \sum_{k=1}^n \phi_k^S(\hat{f}, \mathbf{x}) \hat{\beta}_k^S + \hat{\varepsilon}$$

y : 原油価格 (翌月末)

$\phi_k^S(\hat{f}, \mathbf{x})$: モデル $\hat{f}$ 、説明変数ベクトル $\mathbf{x}$ のもとでの説明変数 k のSHAP値

- 被説明変数と各説明変数の非線形関係をSHAPを通じて表現することで、通常のOLSの枠組みで機械学習モデルを統計的に検証するフレームワーク。



Joseph(2020)

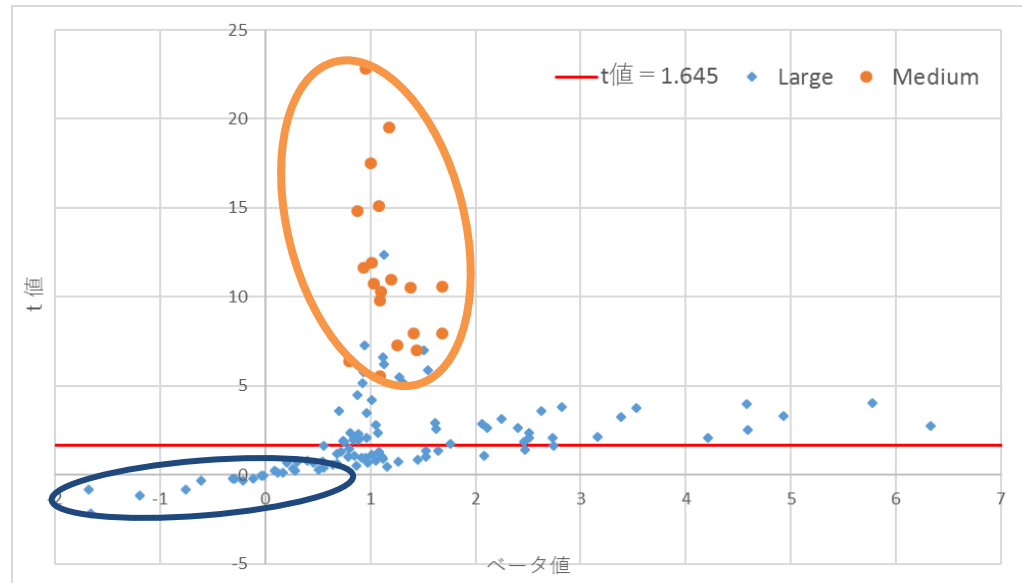
- ① 帰無命題 $H_0^k : \{\beta_k^S \leq 0\}$ を棄却
⇒説明変数 k は被説明変数に対して予測力を有する

- ② β の漸近的性質
 $\widehat{\beta}_k^S \rightarrow 1$ (k は説明力がある)
 $\widehat{\beta}_k^S \rightarrow 0$ (k はノイズ)

SHAP値の統計的検証 (Shapley Regression) ②

- Random Forest (Large dataset) と Random Forest (Medium dataset) について、それぞれ Shapley Regression を行い、係数 (β) をプロットした結果は下記のとおり。

Large dataset と Medium dataset のベータ値および t 値の関係



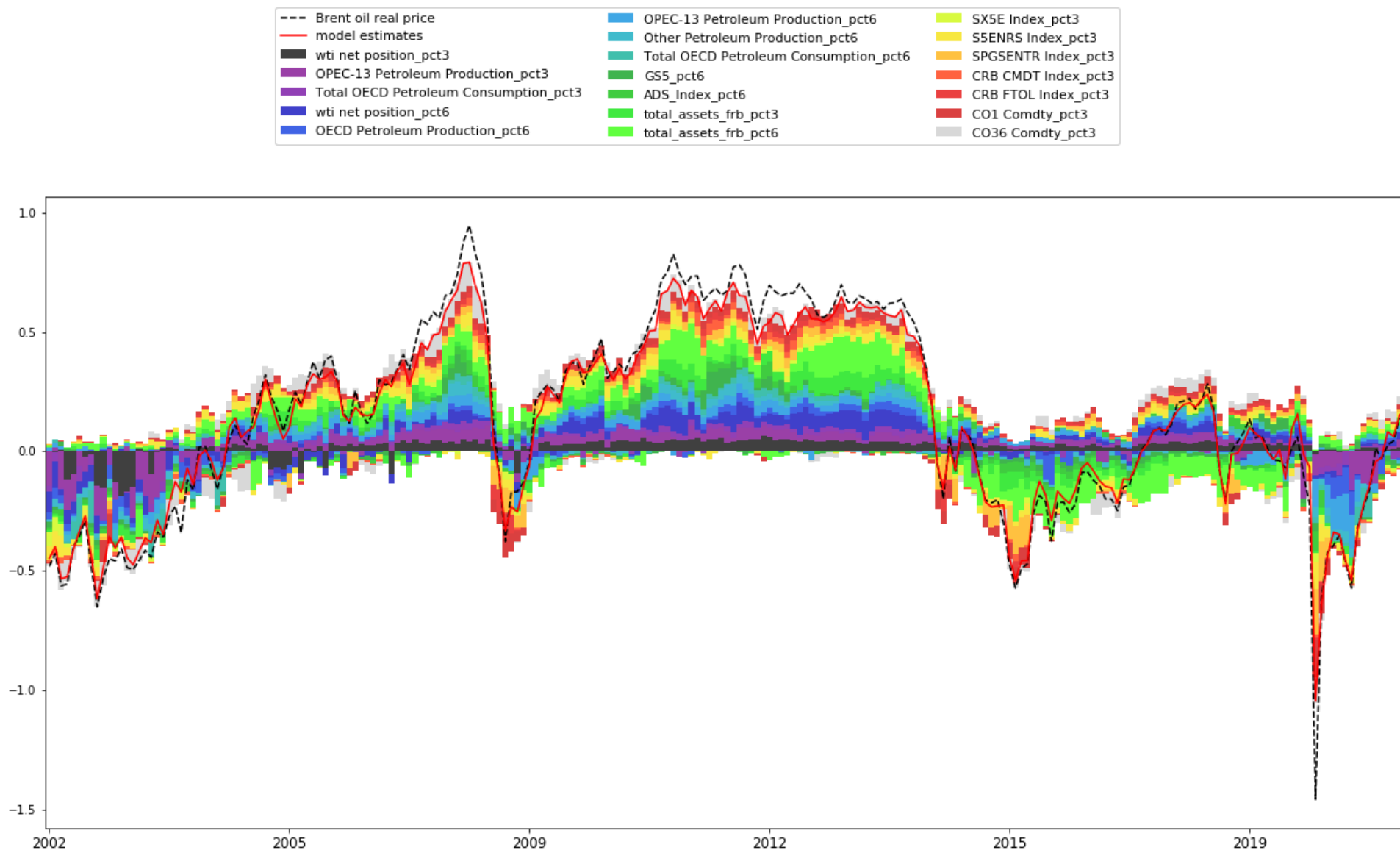
□ SHAP値の係数が有意に正かつゼロから十分に離れている

- Medium datasetで推計されたモデルは、「実際の原油価格の変動への説明力を有する」という仮定は統計的に頑健。
- 事前に変数選択していることもあり、Medium datasetモデルは妥当といえる。

■ SHAP値の係数がマイナスやゼロ近傍となっている

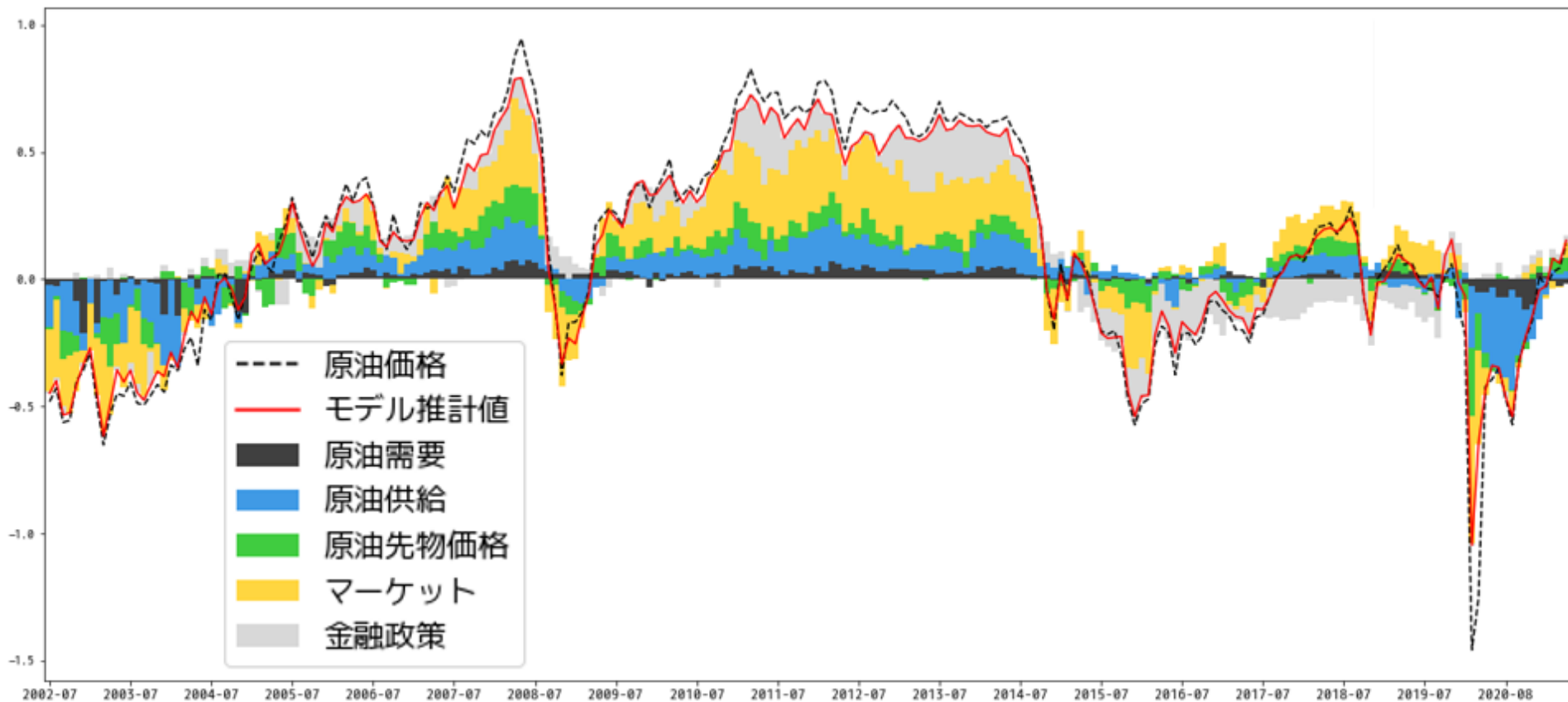
- Large datasetで推計されたモデルは、説明変数の多くが説明力を持たないとみられる。
- Large datasetモデルは妥当でない可能性（説明変数の見直しなどが必要）。

SHAP値の時系列推移（Random Forestモデル<Medium>）



※y軸：モデル推計値から期間平均値を差し引き正規化。原油価格にも同様に適用。

■ SHAP値の時系列推移（Random Forestモデル<Medium>）

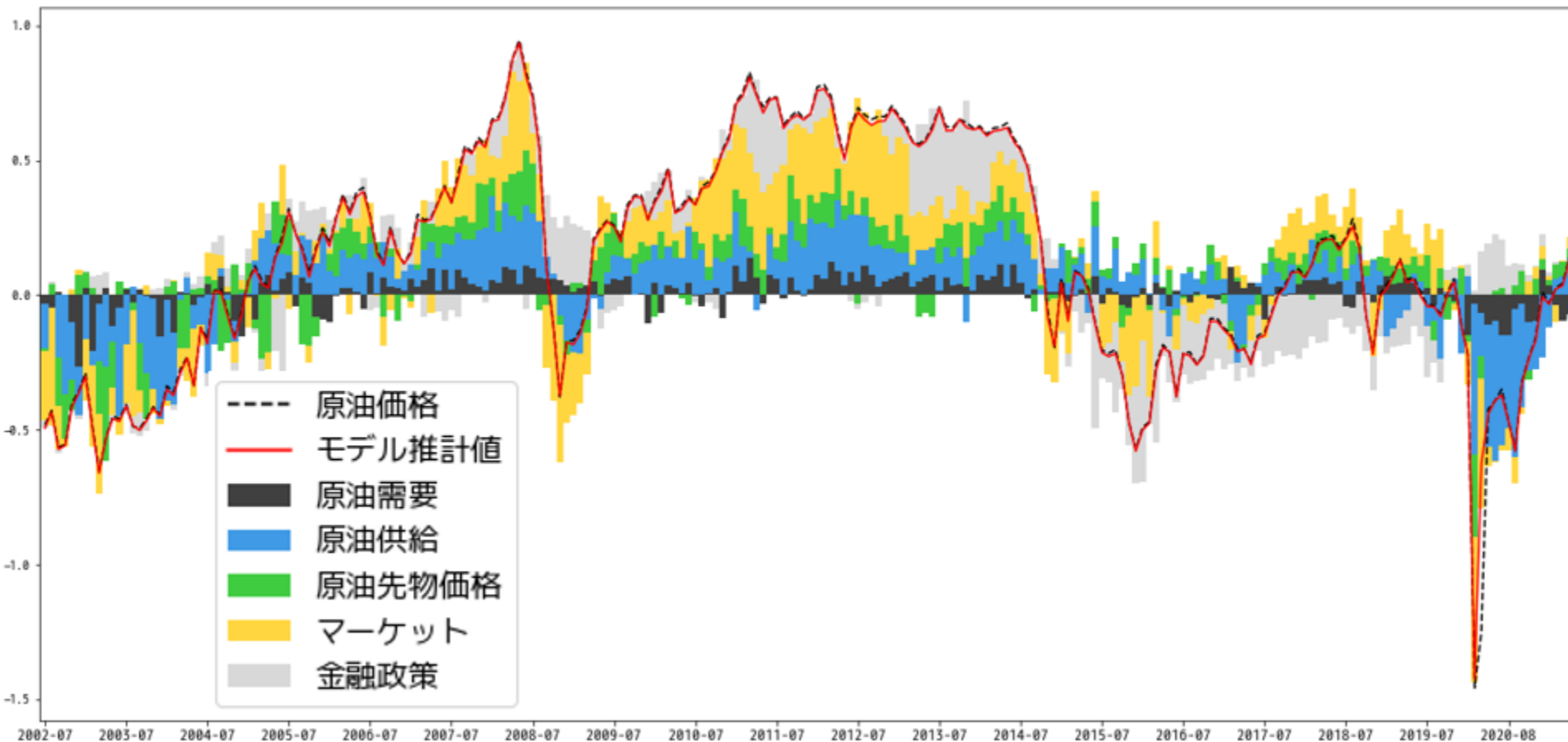


- SHAP値の加法性により、5 要因（需要、供給、先物、マーケット、金融政策）に集約表示可能。
 - ならだかに推移しており、局面毎に主要な変動要因の移り変わりが、無理なく視認できる。金融危機、コロナ・ショックを含め、金融経済動向と概ね整合的な姿となっている。
- 説明可能性があるモデル候補として残る。

※y軸：モデル推計値から期間平均値を差し引き正規化。原油価格にも同様に適用。

SHAPによる時系列の寄与度分解

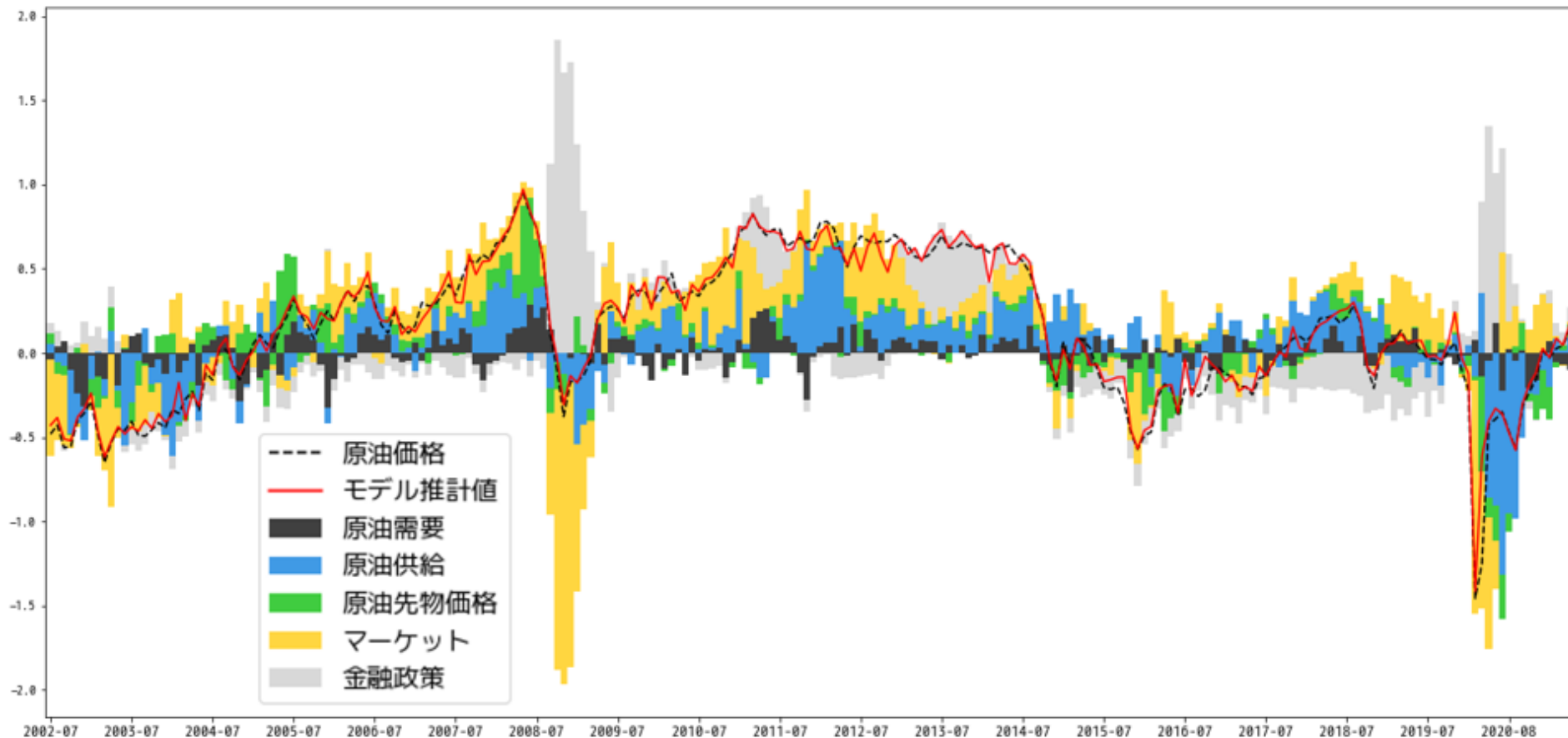
■ SHAP値の時系列推移（Light GBMモデル<Medium>）



- Random Forestより、データへのフィットが良い反面、SHAP値がボラタイルで、視認性がやや劣るが、概ね同様の傾向であり、細かな変動を柔軟にモデリングしている可能性がある。

→ 説明可能性があるモデル候補として残る。

■ SHAP値の時系列推移（Neural Networkモデル<Medium>）



- 金融危機、コロナ・ショック時は、市場とFRB要因が反対方向に激しく変動。
- OECDと他地域の需要動向が同時期に正反対となるなど、金融経済の動向と整合的でない局面が目立つ。

→ NNモデルは、経済的知見に基づいた解釈が難しく、視認性も低いため、採用できない。

※y軸：モデル推計値から期間平均値を差し引き正規化。原油価格にも同様に適用。

SHAPを基にした最適モデルの選択基準

- 視認性や経済的知見等の観点から検討し、分析目的に応じて、モデルの予測精度と説明可能性のトレードオフどちらを優先するかを判断し、最適なモデルを選択することが重要。

		RF (Medium)	RF (Large)	LGBM (Medium)	NN (Medium)
予測精度		○	○	◎	○
説明可能性	説明変数の重要度	○	×	○	○
	時系列の要因分解	◎	×	○	×
	Shapley Regression	○	×	○	○
	SHAP値の変動性	◎	×	○	○
(データ追加に伴う安定性)		◎	×	×	○

- ✓ 実務上、モデルの継続的な運用を行う上で、データ追加に伴うSHAP値の安定性は重要な視点。
⇒ データ追加時に過去の要因分解が大きく変化しないという点では RFが一番安定している。

➤ モデル更新時のSHAP値の変動幅

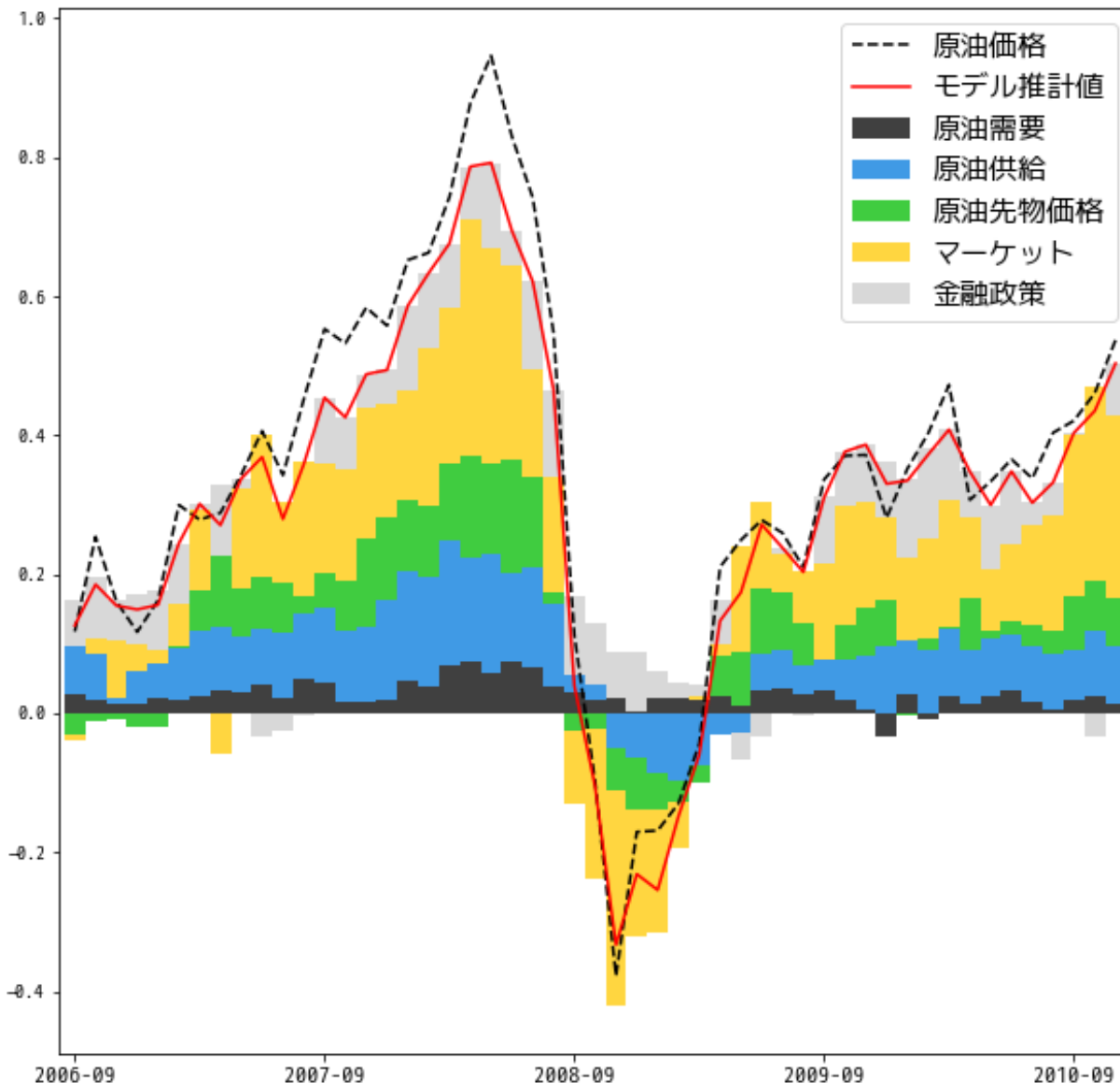
	原油供給	原油需要	原油先物価格	マーケットデータ	金融政策	平均
RF	0.015	0.041	0.038	0.058	0.032	0.037
LGBM	0.030	0.069	0.061	0.103	0.070	0.067
NN	0.018	0.029	0.029	0.133	0.022	0.046



※「2010/12月までのデータで学習したモデル」をベンチマークとし、翌四半期（2011/1-3月期）のデータを加え再学習した時のSHAP値の変動幅を計算するという操作を足もとまで繰り返す。表中ではSHAP値の変動幅の絶対値の平均値を記載。

局面比較詳細（Random Forestモデル<Medium>）①

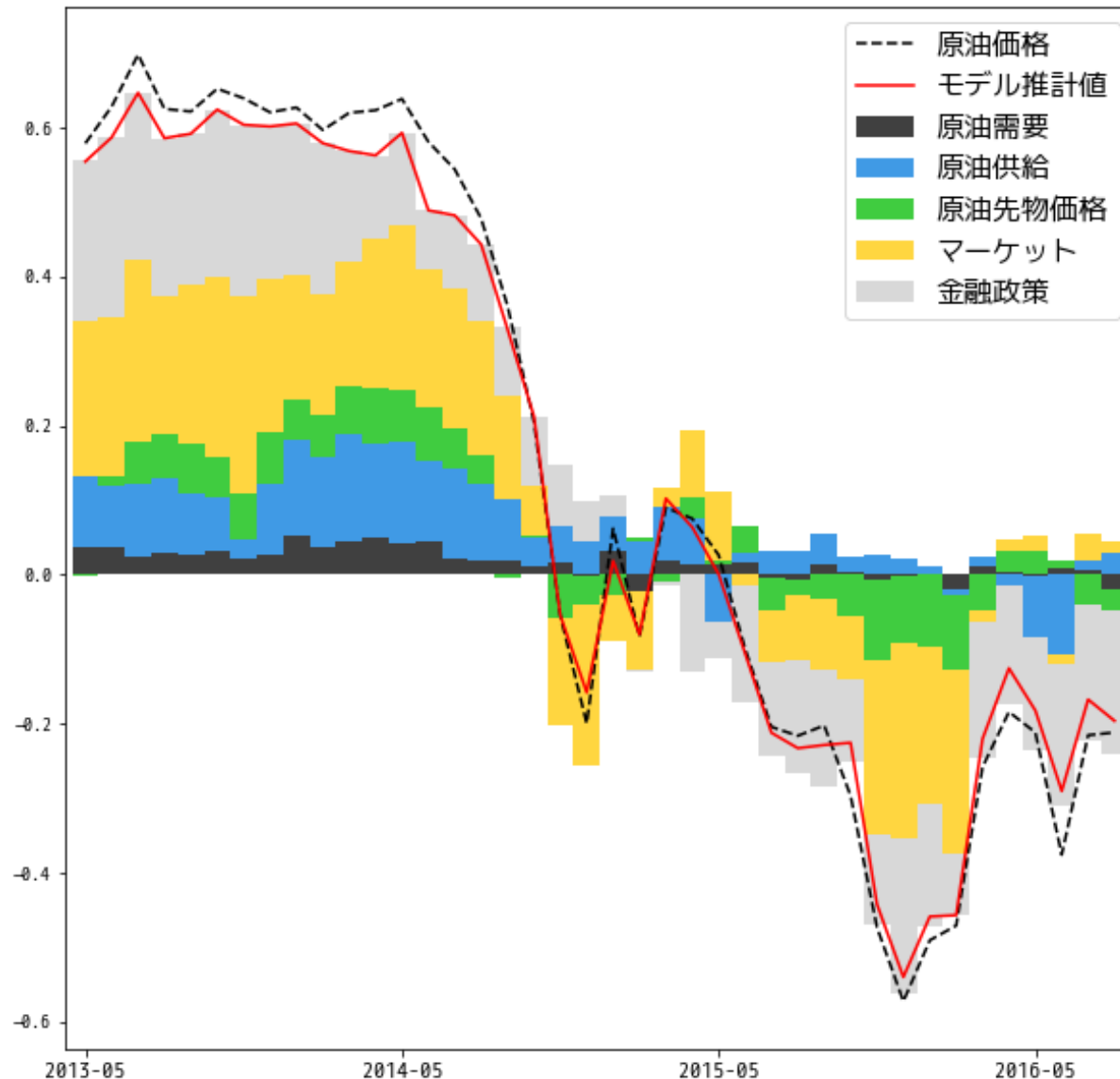
■ 金融危機時の動き



- 2008/6月までは、全ての要因が原油価格の押し上げに寄与。
- 2008/7月にマーケット要因のプラス寄与が急速に縮小。先物価格、供給要因もそれに続いた。
- 2008/9月にマーケットおよび先物価格がマイナスとなり、下押し寄与に転じた。ややラグをもって、供給要因も下押し寄与に転じた。
- 金融危機を受けたFRBの金融緩和はマーケットを下支えた。原油市況回復後も、金融緩和は価格上昇に寄与した。

局面比較詳細（Random Forestモデル<Medium>）②

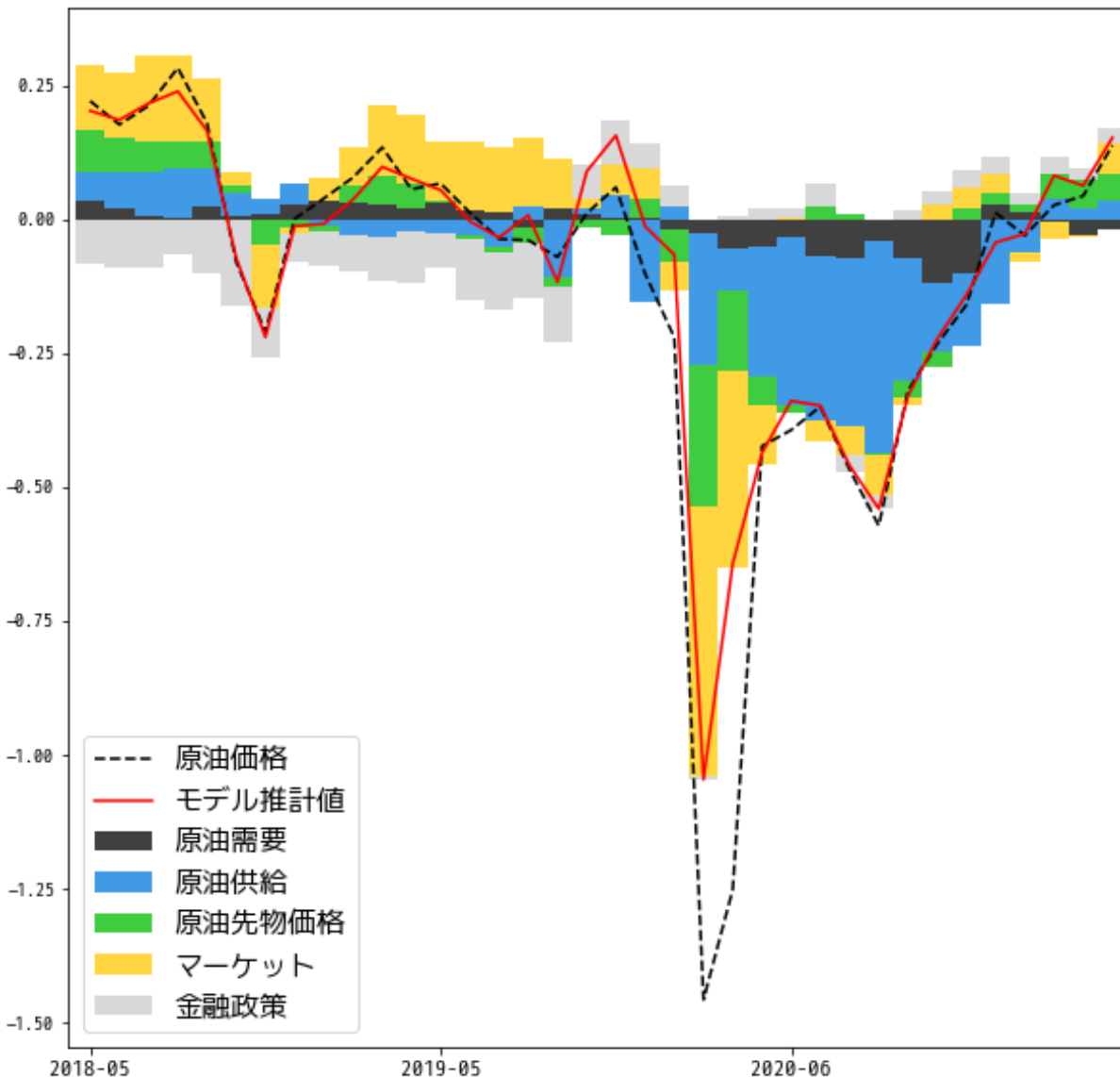
■ 2014年の原油ショック、チャイナ・ショック時の動き



- 2014/5月以降、金融政策を除く全ての要因のプラス寄与が縮小し始めた。
- 2014/10月にはマーケットと先物価格要因を主因にマイナスに転じた。
- 2015年半ばからは、米国の金融緩和縮小がマイナス寄与となりはじめ、原油市況を下押しした。
- 2015年後半のチャイナ・ショック時は、マーケットと先物価格要因が大きなマイナス寄与を示した。

局面比較詳細（Random Forestモデル<Medium>） ③

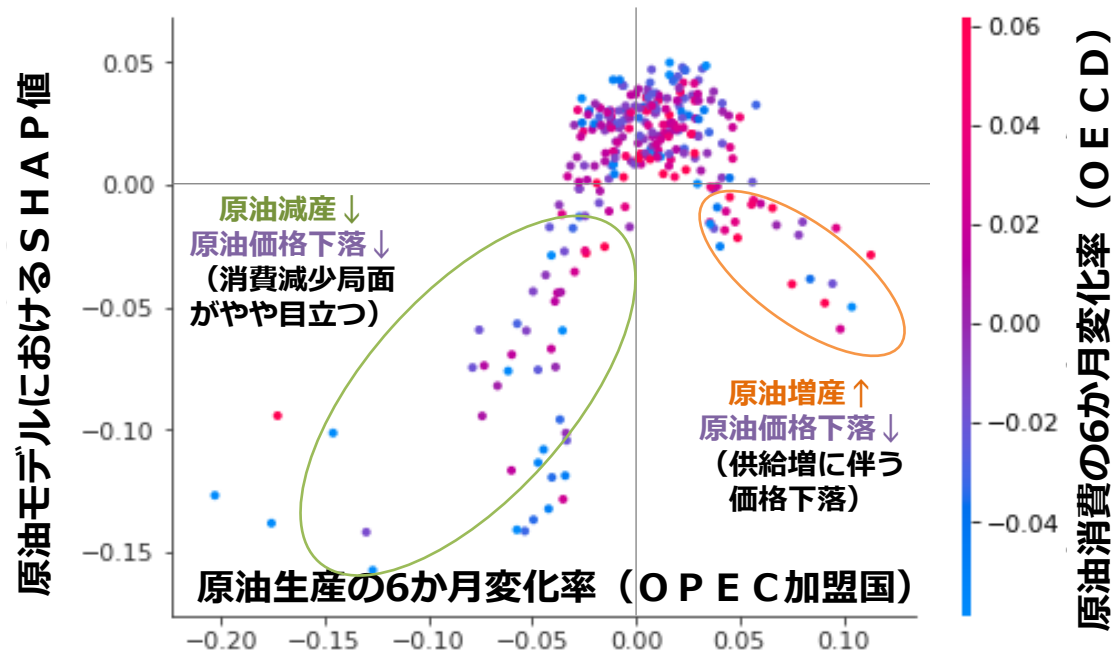
■ コロナ・ショック時の動き



- コロナ・ショック発生後、まずマーケットや先物価格要因が大幅なマイナス寄与を示した。
- 次に、実体経済の大幅な落ち込みで、需給バランスは急速悪化（供給、需要がマイナス寄与）。
- その後、マーケットなどのマイナス寄与は縮小したが、需要の弱さが継続する下で、供給過剰（供給要因による下押し）が1年程度続いた。
- その後、需給バランスは着実に改善し、足元は先物価格やマーケット要因により小幅ながら上昇に転じている。

SHAPによる散布図 (RF <Medium>)

■ 説明変数と原油モデル上のSHAP値（≒原油価格に対する影響度）の関係



- 原油生産と原油価格の間には非線形な関係が確認できる。
 - 原油生産が増加すると原油価格は下落（供給増に伴う価格下落）。
 - 原油消費の減少局面（青点）を中心に、原油生産が減少した局面でも原油価格は下落。
- ただし、SHAP値は個別の関係のみならず、その他の説明変数の影響が含まれうることを勘案して解釈することが重要。こうしたSHAPの散布図表示により、説明変数間の交絡作用（SHAP Interaction Values）を検証することは今後の課題。
 - 需要の増減に対する非対称な原油生産の調整メカニズムを示唆するなど、経済メカニズムを考えるうえで有益な情報が得られる可能性。

SHAPによる機械学習モデル解釈に関する論点

■ 機械学習モデルの予測精度と説明可能性のトレードオフ

モデルの予測精度の観点のみならず、説明可能性の観点からも検討し、分析目的に応じてどちらを相対的に優先するかを決定し、最適なモデルを選択することが重要。

■ ドメイン知識の重要性

ドメイン知識に基づいて、モデル推計時の変数選択（Expert Judgment）や、SHAPによる可視化の妥当性検証を行うことは機械学習モデルの解釈において重要。

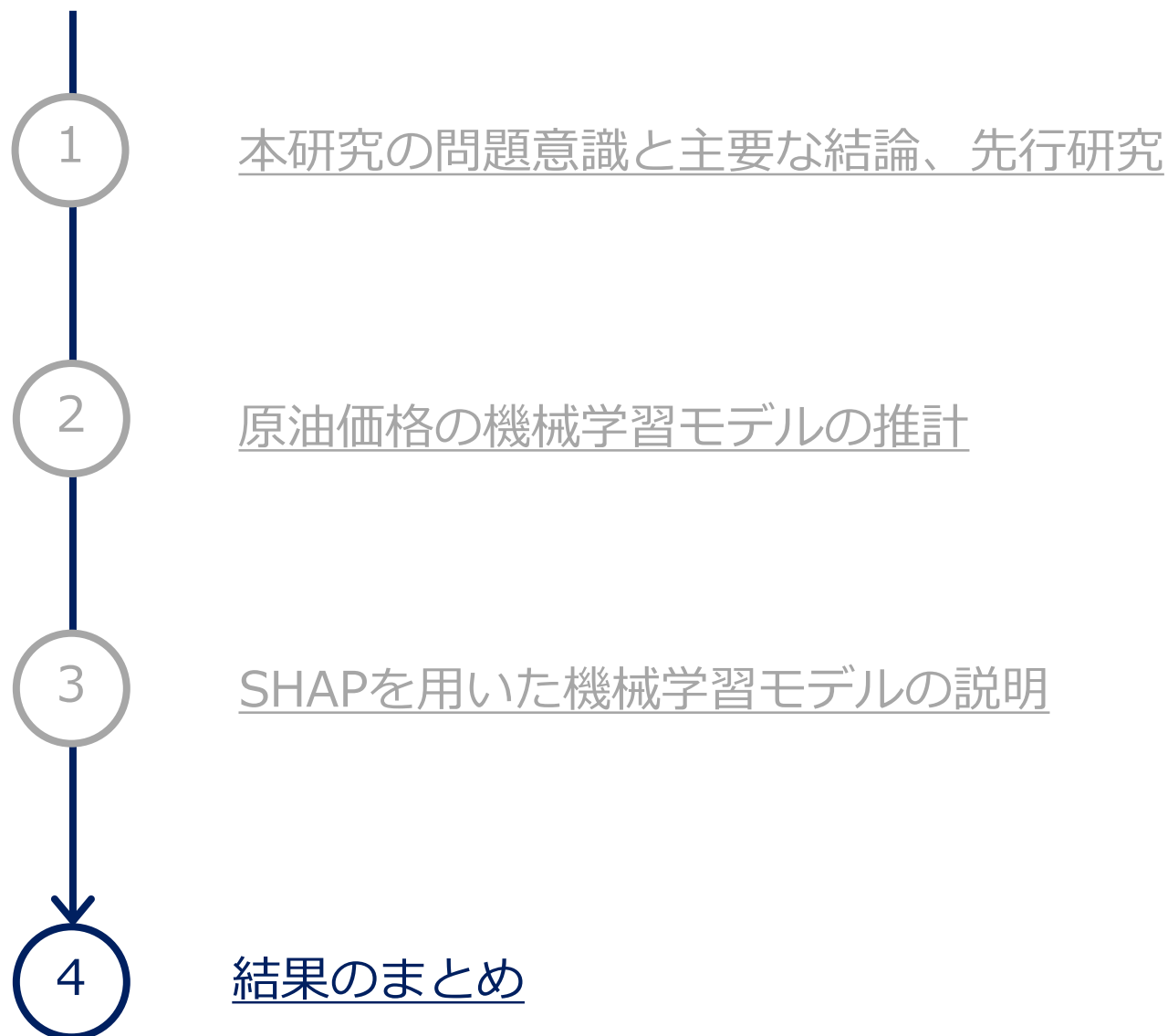
■ 被説明変数と説明変数の因果関係

SHAPは説明変数と予測値の「関係性」を示しているが「因果関係」を直接的に示していない点に留意する必要がある（森下 [2021]）。

【例 1】原油生産と原油価格（前スライド）⇒ 需要要因の影響も含めた「関係性」を捕捉。

【例 2】米国金融政策と原油価格⇒ アネクドタルな言説を裏付ける結果。もっとも、経済学的メカニズムからの検証が必要。

➤ SHAPによる機械学習モデルは「データに語らせている分析」であるため、メカニズムの解釈や因果関係の検証には、別途、因果推論等の手法を用いる必要である。ただし、構造推定のための予備的分析や変動要因の把握などで、機械学習モデルは有用なツールとなりえる。



まとめ（分かったこと）

- 機械学習モデルは、伝統的手法対比でインサンプル予測で高い予測精度を示した。
同時に、説明可能性の低下に直面する。
- SHAPを用いれば、機械学習モデルの説明や妥当性の検証が可能になる。
 - 機械学習モデルを推計時には、ドメイン知識に基づいて、説明変数絞り込みやSHAPによる可視化・検証を行うことが、モデルの説明可能性を高めることが示唆された。
 - 予測モデルにおける重要変数の特定、時系列推移の視認性・変動性の確認（大きな振れ、ノイズへの過剰適合が生じていないかなど）、統計的検証。
- 機械学習モデルの「予測精度」と「説明可能性」のトレードオフを踏まえ、分析目的に応じて、最適なモデルを選択することが重要。
 - 今回の例では、NNやRF(Large)は、予測精度は良いが、説明可能性が低く、実務での利用は難しい。RF(Medium)とLBGMは、定性的には同様の結果となっているが、モデルの変動性や安定性の観点を考慮して、RF(Medium)を選択。
- SHAPを用いた原油価格の変動要因分解は、原油需給・マーケット・金融政策要因が主因であるが、経済情勢を反映して、寄与度や符号の方向は大きく入れ替わる姿が示された。

まとめ（留意点、今後の検討事項）

- SHAPは、被説明変数と説明変数の相関を含めた関係性を示すものの、**因果関係を直接的に示していない**点には留意。被説明変数と説明変数の因果関係を示すためには、因果推論等による検証が必要。また、分析対象データの構造変化をどのように勘案していくかも論点。
- 本研究では、過去の原油価格変動の局面比較や足もとの情勢判断への活用を念頭に、SHAPを用いて、インサンプルの予測精度とモデルの説明可能性にフォーカスした。当ワークフローはアウトオブサンプル予測の最適なモデル選択と異なる可能性。
 - 原油価格の将来予測は難しいことが知られているが、将来予測に優れた機械学習モデルを推計・選択する際に、説明可能性の観点をどのように勘案するかは今後のリサーチ・テーマ。
- 機械学習モデルを継続運用する場合、XAI（モデル説明手法）を活用し、モデルの定期的な見直しや説明可能性の確保を図ることが、モデル・リスク管理やモデルの説明責任の観点で有効な対応策となりえる。
 - 概念ドリフト（モデルの被説明変数の統計的特性が時間経過とともに予期せず変化し、機械学習モデルの精度が悪化する可能性がある）への対処は、予測精度の向上と同様に重要（Lu et al. [2020], Baumeister, Korobilis and Lee [2020]）。

ご清聴ありがとうございました。

照会先

日本銀行金融研究所

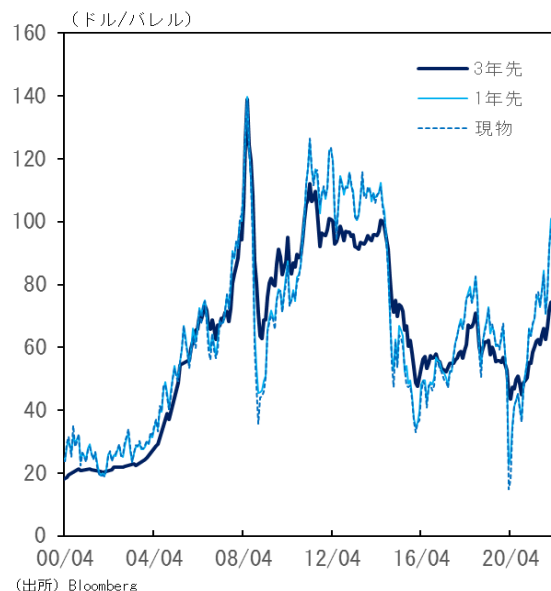
金田 規靖	noriyasu.kaneda@boj.or.jp
木全 友則	tomonori.kimata@boj.or.jp
平木 一浩	kazuhiro.hiraki@boj.or.jp

- 大坪直樹・中江俊博・深沢祐太・豊岡 祥・坂元哲平・佐藤 誠・五十嵐健太・市原大暉・堀内新吾、2021年、「XAI（説明可能なAI）—そのとき人口知能はどう考えたか？—」、リックテレコム
- 三浦翔・井實康幸・竹川正浩、「入出金情報を用いた信用リスク評価－機械学習による実証分析－」、日本銀行ワーキングペーパーシリーズNo.19-J-4、日本銀行、2019年
- 森下光之助、2021年、「機械学習を解釈する技術 予測力と説明力を両立する実践テクニック」、技術論評社
- Arrieta A. B., N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila and F. Herrera (2020). “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Information Fusion*, 58, pp.82-115.
- Baumeister C., Korobilis, D. and Lee, T. (2022). “Energy markets and global economic conditions,” *Review of Economics and Statistics*, 104 (4), pp. 828–844.
- Bergmeir Christoph, José M. Benítez (2012). “On the use of cross-validation for time series predictor evaluation,” *Information Sciences*, 191, pp.192-213.
- Bluwstein K., M. Buckmann, A. Joseph, M. Kang, S. Kapadia, and O. Simsek (2021). “Credit growth, the yield curve and financial crisis prediction: evidence from a machine learning approach,” Working Paper Series 2614, European Central Bank.
- Breiman Leo (2001). “Random forest,” *Machine Learning*, 45(1), pp.5-32.
- Buckmann M. and A. Joseph (2022). “An interpretable machine learning workflow with an application to economic forecasting,” Staff Working Paper No. 984, Bank of England.
- Calomiris Charles W. and Çakır Melek, Nida and Mamaysky, Harry (2022). “Predicting the Oil Market,” Federal Reserve Bank of Kansas City Working Paper, No. 20-20.
- Consoli S., D. R. Recupero and M. Saisana (2021). “Data Science for Economics and Finance. Methodologies and Applications,” Springer.
- Costa A. B. R., P. C. G. Ferreira, Wagner. P. Gaglianone, Osmani. T. C. Guillén, J. V. Issler and Y. Lin (2021). “Machine Learning and Oil Price Point and Density Forecasting,” *Energy Economics*, 102, 105494.
- Friedman Jerome H. (2001). “Greedy function approximation: A gradient boosting machine,” *Ann. Statist*, 29(5), pp.1189-1232.

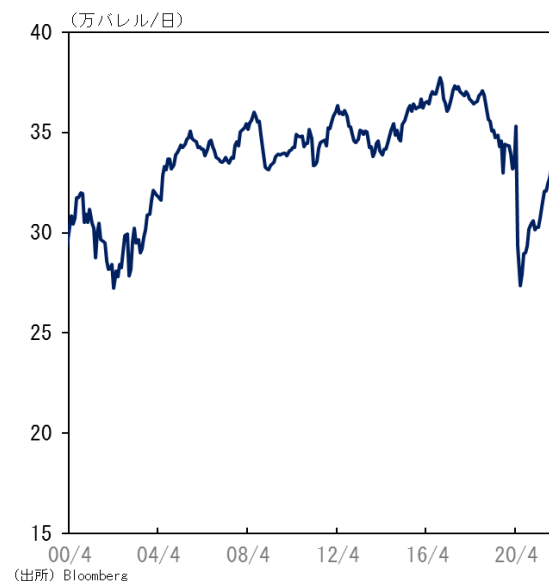
- Fueki, T., H. Higashi, N. Higashio, J. Nakajima, S. Ohyama, and Y. Tamanyu (2016). “Identifying Oil Price Shocks and Their Consequences,” Bank of Japan Working Paper Series 16-E-17, Bank of Japan.
- Goldstein Alex, A., Kapelner, J., Bleich, and E., Pitkin (2015). “Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation,” *Journal of Computational and Graphical Statistics*, 24(1): pp.44-65.
- Hamilton James D. (2003). “What is an oil shock?,” *Journal of Econometrics*, 113(2), pp.363-398
- Joseph A. (2020). “Parametric Inference with Universal Function Approximators,” Staff Working Paper No.784, Bank of England.
- Ke G., Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, Tie-Yan Liu (2017). “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” 31st Conference on Neural Information Processing Systems (NIPS 2017).
- Kilian Lutz (2009). “Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market,” *American Economic Review*, 99(3), pp.1053-1069.
- Kursu M. B. and W. R. Rudnicki (2010). “Feature Selection with the Boruta Package,” *Journal of Statistical Software*, 36(11), pp.1-13.
- Lu Jie, Anjin Liu, Fan Dong, Feng Gu, Joao Gama and Guangquan Zhang (2020). “Learning under Concept Drift: A Review,” arXiv:2004.05785.
- Lundberg, Scott and Su-In Lee (2017). “A unified approach to interpreting model predictions,” *Advances in Neural Information Processing Systems*, pp. 4765-4774
- Lundberg, Scott, Gabriel Erion and S.- I. Lee (2018). “Consistent individualized feature attribution for tree ensembles,” arXiv:1802.03888.
- Shapley L. S. (1953). “A Value for N-Person Games,” *Contributions to the Theory of Games*, 2, pp.307-318.
- Strumbelj E. and I. Kononenko (2010). “An Efficient Explanation of Individual Classifications Using Game Theory,” *Journal of Machine Learning Research*, 11, pp.1–18.
- Molnar Christoph (2021). “Interpretable Machine Learning: A Guide for Making Black Box Models Explainable,” website, URL: <https://hacarus.github.io/interpretable-ml-book-ja/index.html>

Appendix

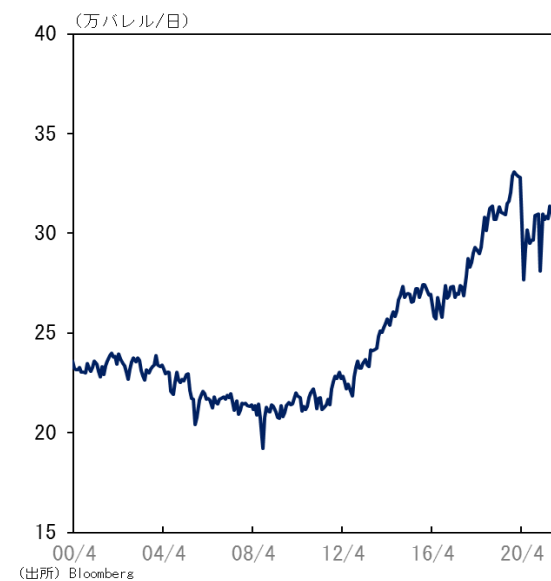
図表1：原油価格



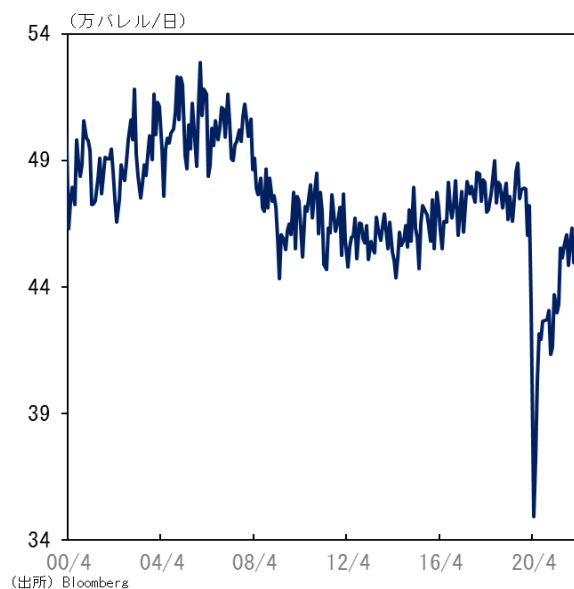
図表2：OPEC原油生産量



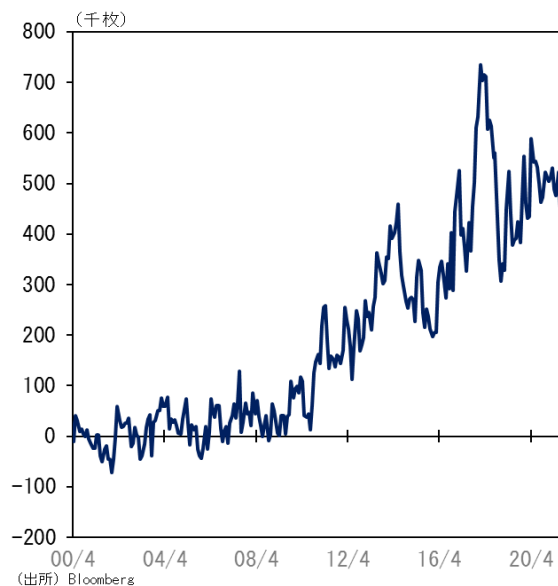
図表3：OECD原油生産量



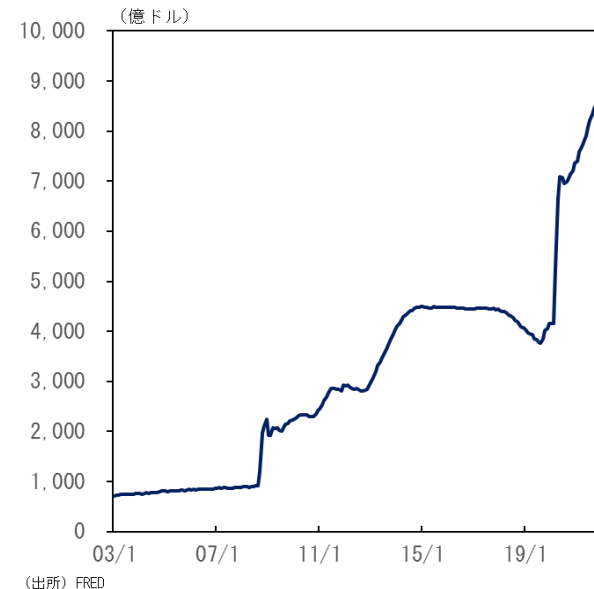
図表4：OECD原油消費量



図表5：IMM原油ネットポジション



図表6：FEDのバランスシート



機械学習モデルの推計（変数選択結果＜Large＞）

■ Large dataset（Borutaによる説明変数選択後）：75説明変数＜97変数＞

原油供給	在庫（米国）	マネー（米国）	景気動向指数等
米国原油生産（6か月変化率）	耐久消費財の未履行数（3か月変化率）	マネタリーベース（6か月変化率）	中国（3か月変化率）
世界原油生産（6か月変化率）	耐久消費財の未履行数（6か月変化率）	商業および産業ローン（3か月変化率）	フランス（3か月変化率）
米国原油・天然ガスの稼働数（3か月変化率）	消費財新規注文数（6か月変化率）	商業および産業ローン（6か月変化率）	ギリシャ（3か月変化率）
米国原油・天然ガスの稼働数（6か月変化率）	事業用在庫（6か月変化率）	消費者信用比率（6か月変化率）	アイルランド（3か月変化率）
地熱発電発電量（3か月変化率）	実質個人消費支出（6か月変化率）	カーローンおよびカーリース（3か月変化率）	韓国（3か月変化率）
原油先物価格	労働（米国）	カーローンおよびカーリース（6か月変化率）	ロシア（3か月変化率）
原油2年先物価格（3か月変化率）	平均賃金（製品生産）（6か月変化率）	消費者ローンおよびリース全体（3か月変化率）	スウェーデン（3か月変化率）
原油3年先物価格（3か月変化率）	雇用者数（鉱業）（3か月変化率）	消費者ローンおよびリース全体（6か月変化率）	トルコ（3か月変化率）
原油スプレッド（12か月-2か月）（6か月変化率）	雇用者数（鉱業）（6か月変化率）	銀行貸付（3か月変化率）	イギリス（3か月変化率）
原油スプレッド（3年-1年）（6か月変化率）	平均賃金（製造）（6か月変化率）	銀行貸付（6か月変化率）	オーストラリア（6か月変化率）
コモディティ価格	雇用者数（製造）（6か月変化率）	M1マネースtock（6か月変化率）	チリ（6か月変化率）
ロイター商品指数（6か月変化率）	雇用者数（非耐久消費財）（3か月変化率）	実質M2マネースtock（3か月変化率）	ギリシャ（6か月変化率）
マーケットデータ	雇用者数（非耐久消費財）（6か月変化率）	実質M2マネースtock（6か月変化率）	アイルランド（6か月変化率）
S&P PER指数	雇用者数（サービス）（6か月変化率）	M2マネースtock（3か月変化率）	韓国（6か月変化率）
AA3か月物CPレート（6か月変化率）	27週以上被雇用者数（6か月変化率）	M2マネースtock（6か月変化率）	ロシア（6か月変化率）
米国3か月金利（6か月変化率）	雇用者数（建設）（3か月変化率）	信用貸付全体（3か月変化率）	スペイン（6か月変化率）
米国6か月金利（6か月変化率）	雇用者数（建設）（6か月変化率）	信用貸付全体（6か月変化率）	スウェーデン（6か月変化率）
米国1年金利（6か月変化率）	雇用者数（金融）（3か月変化率）	住宅ローン（3か月変化率）	トルコ（6か月変化率）
Moody's Aaa社債金利（6か月変化率）	雇用者数（金融）（6か月変化率）	住宅ローン（6か月変化率）	スロヴァキア鉱工業生産（6か月変化率）
Moody's Baa社債金利（6か月変化率）	雇用者数（政府）（3か月変化率）		
米ドル／ポンド為替レート（6か月変化率）	雇用者数（政府）（6か月変化率）		
金融政策	雇用者数（卸売）（6か月変化率）		
ECBバランスシート（3か月変化率）	住宅（米国）		
日米欧バランスシート（6か月変化率）	住宅建築許可数・西部地区（6か月変化率）		
鉱工業生産（米国）	消費（米国）		
化学製品生産（6か月変化率）	実質個人所得（6か月変化率）		
エネルギー生産（6か月変化率）	物価指数（米国）		
石油・天然ガス抽出（6か月変化率）	CPI(アパレル)（6か月変化率）		
米油・ガス抽出（6か月変化率）	CPI(医薬品)（6か月変化率）		
製薬（3か月変化率）	CPI(サービス)（6か月変化率）		
製薬（6か月変化率）	個人消費（サービス除く）（6か月変化率）		
非耐久化学製品（6か月変化率）	PPI(金属製品)（6か月変化率）		
航空宇宙その他輸送（6か月変化率）	PPI(中間財)（6か月変化率）		
天然ガス抽出（6か月変化率）	PPI(原油製品)（6か月変化率）		

アルゴリズムによる説明変数の絞り込み

■ Borutaとは「与えられた説明変数の中から重要度の低いものを削除するアルゴリズム」。

①各説明変数の時系列順序をランダムに入れ替えたダミー変数を作成し、オリジナルの説明変数と組み合わせ、全ての説明変数を用いてRandom Forestモデルを学習

②各モデルにおける各説明変数の重要度（説明力）を計測し、ダミー変数より重要度が高い説明変数を記録

③上記の手順を一定数繰り返し、ダミー変数より重要度の高いオリジナルの説明変数を重要変数として選択

⇒ 重要と判定された説明変数はBorutaにより絞り込まれた説明変数として、機械学習モデルの推計に使用。

■ Borutaによる説明変数選択のイメージ

①

真の説明変数			ダミー		
真①	真②	真③	偽①	偽②	偽③
○	○	○	△	×	○
△	△	△	×	△	×
×	×	×	○	○	△

... + ...

⇒これら全ての説明変数を用いてRandom Forestモデルを学習

②

Random Forestモデルにおける説明変数の重要度

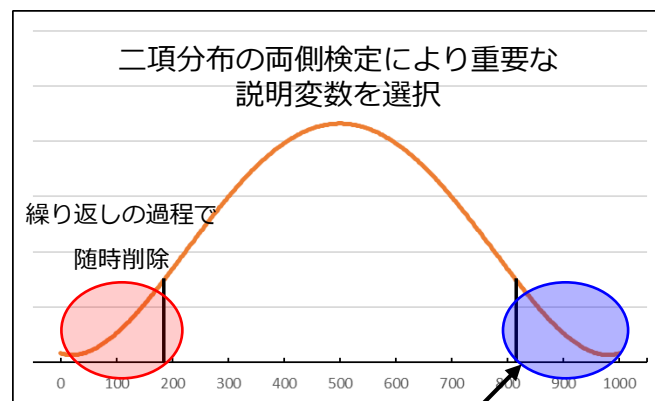
真②	偽③	偽②	真①	偽①	真③	...
0.3	0.25	0.2	0.15	0.08	0.02	...

ダミー説明変数より重要度が高いものを記録

③

繰り返した結果

	真①	真②	真③
ダミーに 勝った回数	500	800	700



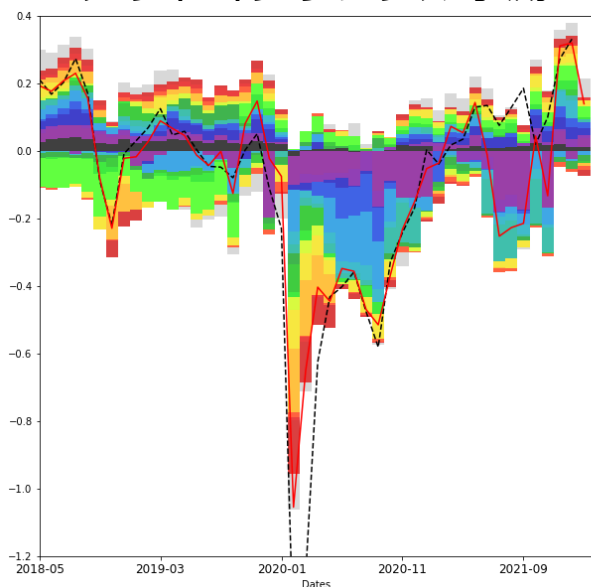
(参考) その他の説明変数選択の手法

手法	概要	例
Filter Methods	説明変数の評価指標を基に選択する方法。	- 分散、相関係数 - Mutural Infomation
Wrapper Methods	機械学習モデルを使用して説明変数の組み合わせを評価する方法。Filter Methodsでは不明であった変数間の関係性を捉えられる。	- Forward Feature Selection - Backward feature Elimination - Exhaustive Feature Search
Embedded Methods	機械学習モデルの学習結果で変数選択を行う方法。	- LASSO - Tree ModelのFeature Importance
その他	説明変数から新たな説明変数を作り出す方法。	- 主成分分析 - T-SNE

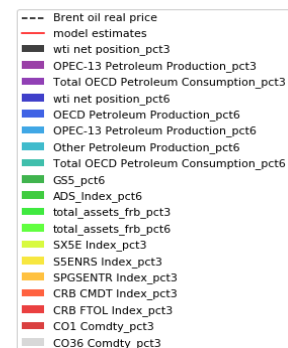
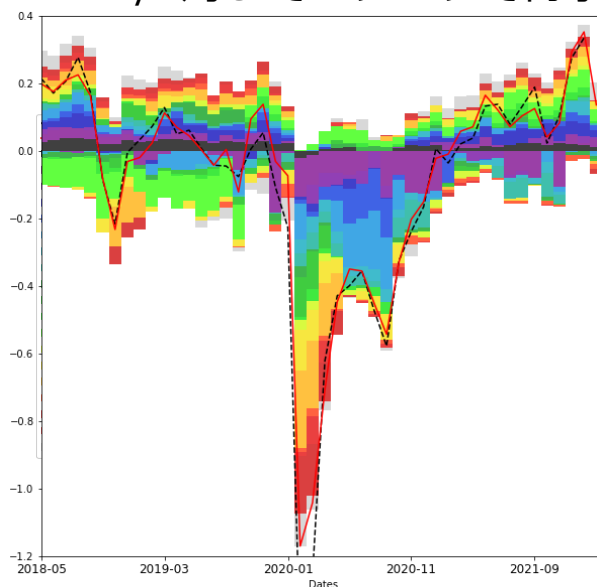
機械学習モデルの説明手法に関する先行研究（詳細）

先行研究	手法	備考
Strumbelj and Kononenko (2010)	モデル解釈へのゲーム理論の応用可能性	協力ゲーム理論に基づき、プレイヤー（説明変数）のゲームに対する貢献度の計算を援用すれば、モデル予測値の重要度とみなせることを数学的に証明。全ての機械学習モデルに適用可能なことも示した。
Lundberg and Lee (2017)	SHAPの提案	協力ゲーム理論のShapley値（平均限界貢献度）を機械学習に応用したSHAPを提案。説明変数の重要度（予測への影響力）を加法的に表現。
Bluwstein et al. (2021)	SHAPの適用	機械学習で金融危機の予測モデルを推計。SHAPを用いて、 <u>金融危機の予測に有用な金融経済変数を可視化し、複数の変数が相互関連し、複合的に金融危機リスクを高めうる可能性を示した。</u>
Buckmann et al. (2022)	SHAPの適用と統計的検証	経済的解釈を意識した機械学習モデルの推計、検証のワークフローを提案。米国失業率を対象に、モデルの予測精度の比較評価、SHAPによる予測値の要因分解の可視化、SHAP値の検証（Shapley Regression）などを実施。 <u>ドメイン知識により変数を絞り込んだ場合、機械学習モデルの予測精度と説明可能性がバランスよく達成するモデルが推計できると主張。</u>
森下 (2021)	機械学習モデルの説明手法の留意点	機械学習モデルの説明手法（PFI,PD,ICE,SHAP）は「 <u>説明変数と予測値の関係</u> 」を示しており、「 <u>説明変数と目的変数の因果関係</u> 」ではない <u>点に留意</u> 。別途、因果推論による検証が必要。
大坪他 (2021)	機械学習モデルの説明可能性について考察、説明手法の留意点	機械学習モデルの予測精度と説明可能性のトレードオフに言及したうえで、機械学習の実務利用のうえで、モデル予測値の解釈を通じた納得感の醸成が重要であると指摘。分析目的や用途に応じたXAIの活用方法（「大局説明」、「局所説明」）を考察し、各手法のコードや活用例を紹介。

■ アウトオブサンプル予測

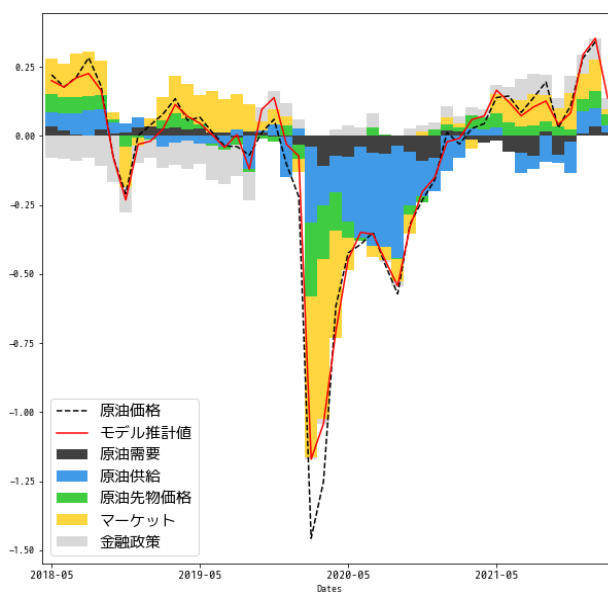
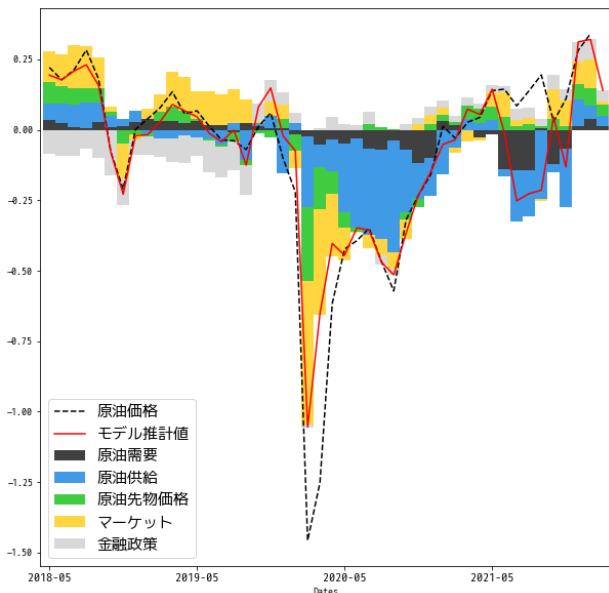


■ 22/2月までのデータで再学習



・コロナ・ショック後の回復局面は、過去にみられない動きであり、予測（黒点線）と実績（赤線）は乖離が目立つ（左図）。

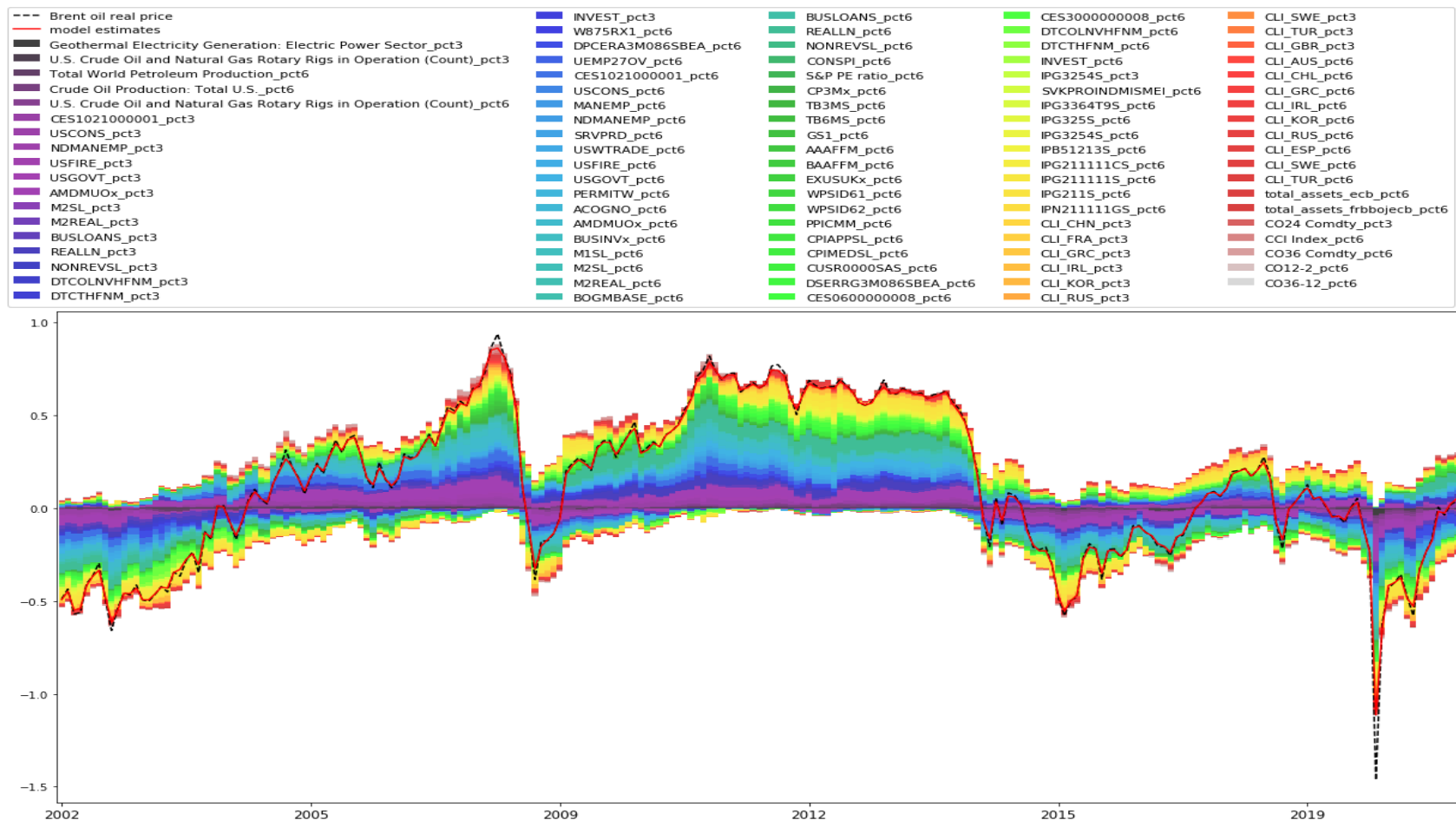
・22/2月までのデータで再学習したところ、高いデータへのフィットを示した（右上下図）。21年後半の原油価格は、マーケットや先物価格要因を主因にプラスに転じ、供給要因などのマイナス寄与は縮小している姿（右下図）。なお、地政学リスクの高まり（ウクライナ危機）による原油価格上昇は、マーケット・原油先物・供給要因により説明されているとみられる。



※y軸：モデル推計値から期間平均値を差し引き正規化。原油価格にも同様に適用。

SHAPによる時系列の寄与度分解（集約表示前）

SHAP値の時系列推移（Random Forest <Large>）



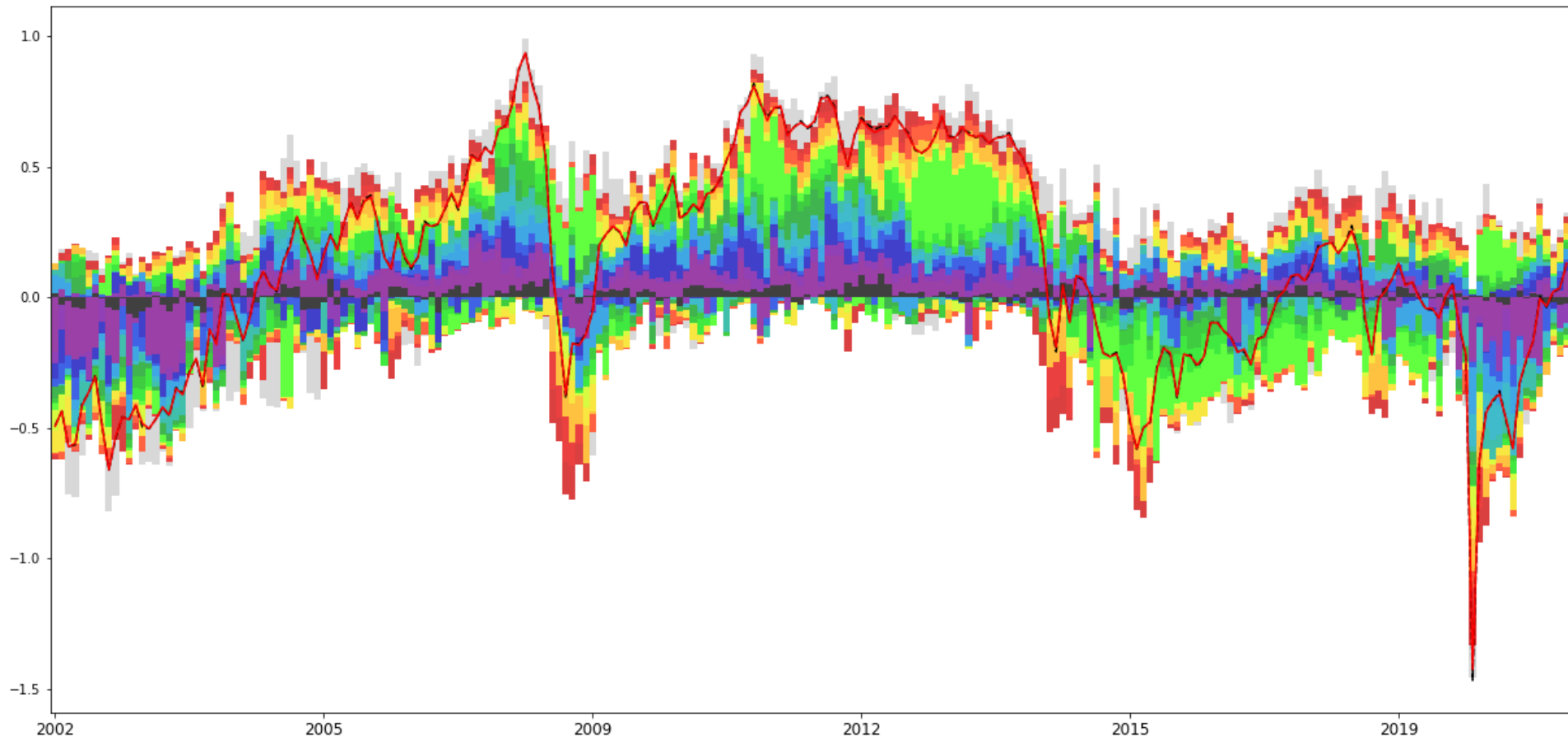
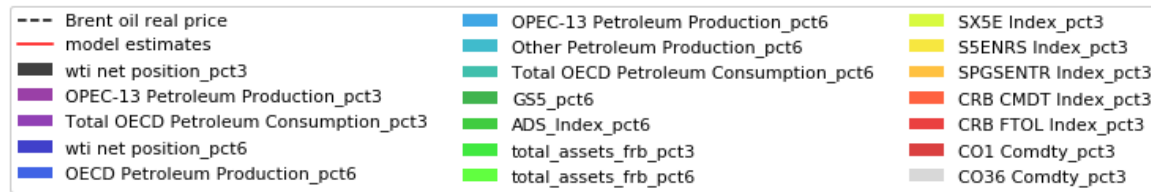
- 原油価格との関連が比較的低い要因（失業率、貸出、新興国の経済指標など）が多く出現。使用変数が多岐にわたるため、カテゴリー集約は不能。

→ Large dataset は、どの要因が主因であるかが解釈困難であり、採用することが難しい。

※y軸：モデル推計値から期間平均値を差し引き正規化。原油価格にも同様に適用。

SHAPによる時系列の寄与度分解（集約表示前）

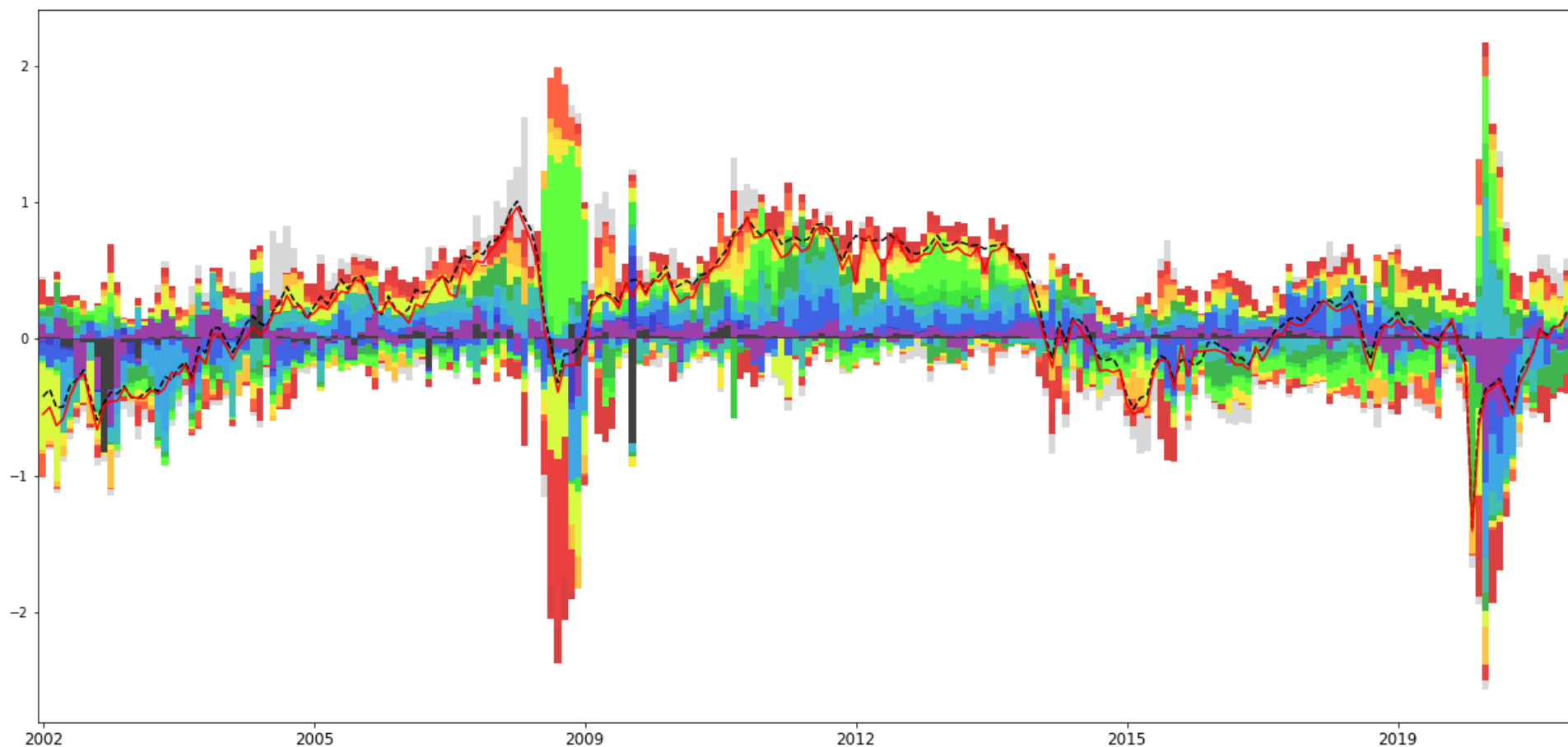
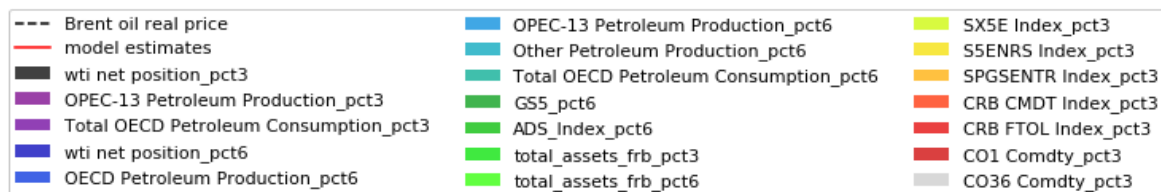
■ SHAP値の時系列推移（Light GBMモデル）



※y軸：モデル推計値から期間平均値を差し引き正規化。原油価格にも同様に適用。

SHAPによる時系列の寄与度分解（集約表示前）

■ SHAP値の時系列推移（Neural Networkモデル）



※y軸：モデル推計値から期間平均値を差し引き正規化。原油価格にも同様に適用。

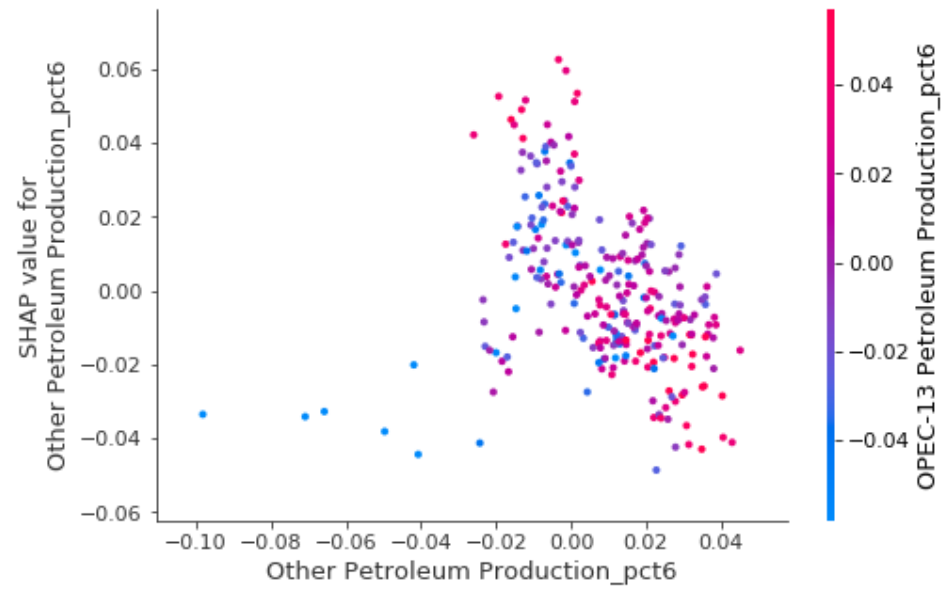
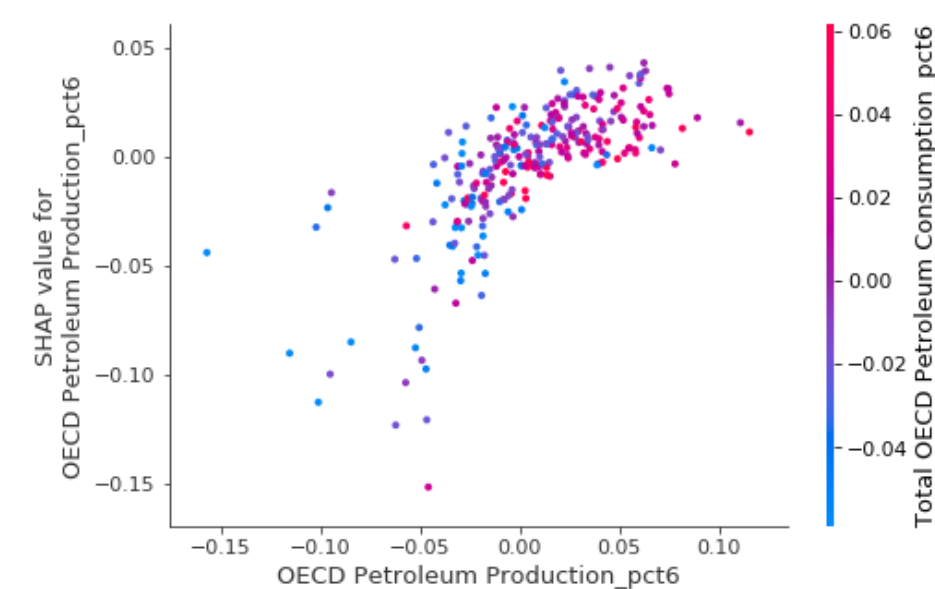
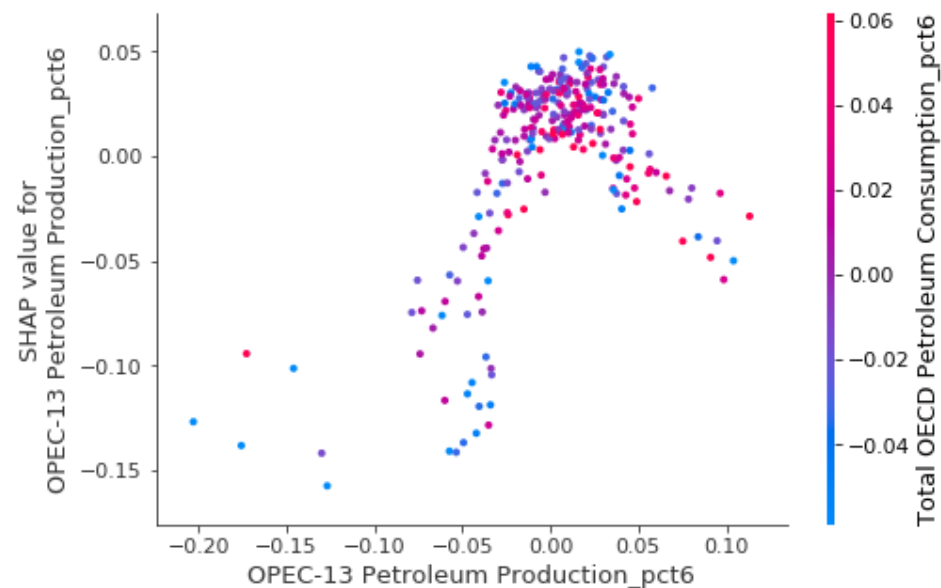
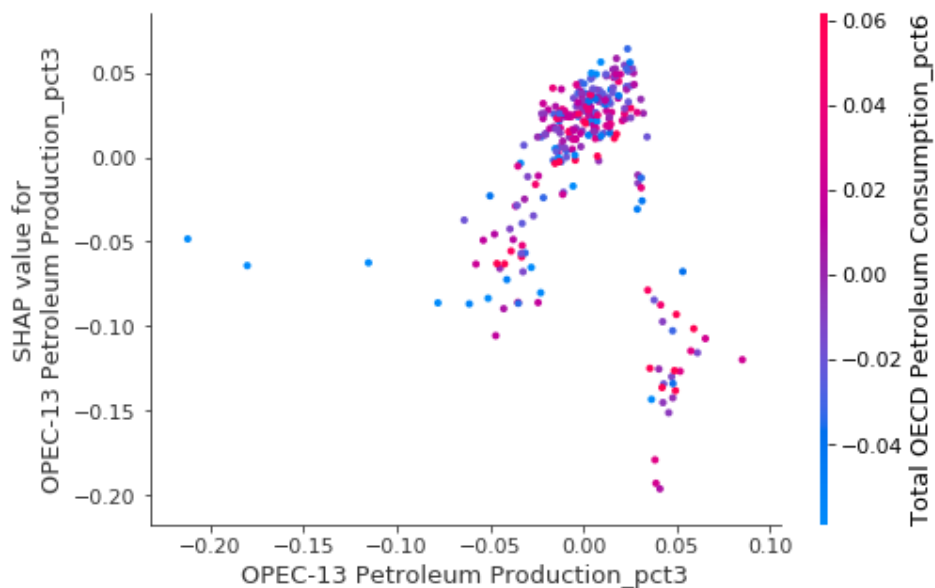
- ドメイン知識や視認性の観点に加え、SHAP値の変動性（標準偏差）といった定量指標がモデルの説明可能性の判断基準となりえる。
- ✓ 3モデルを比較すると、Random ForestモデルのSHAP値の変動性が一番小さい（振れが小さい）ため、視認性の観点では優れている。

➤ 機械学習モデルのSHAP値の標準偏差

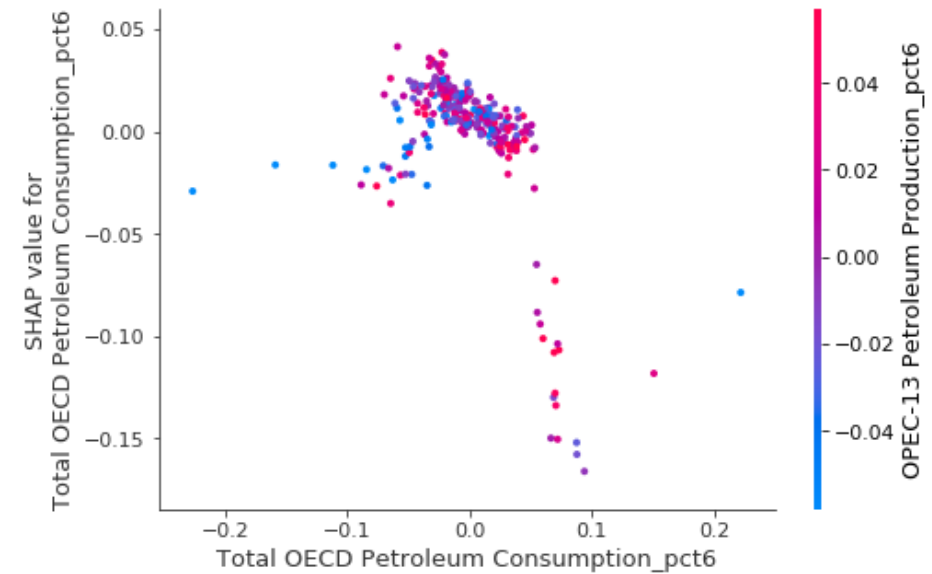
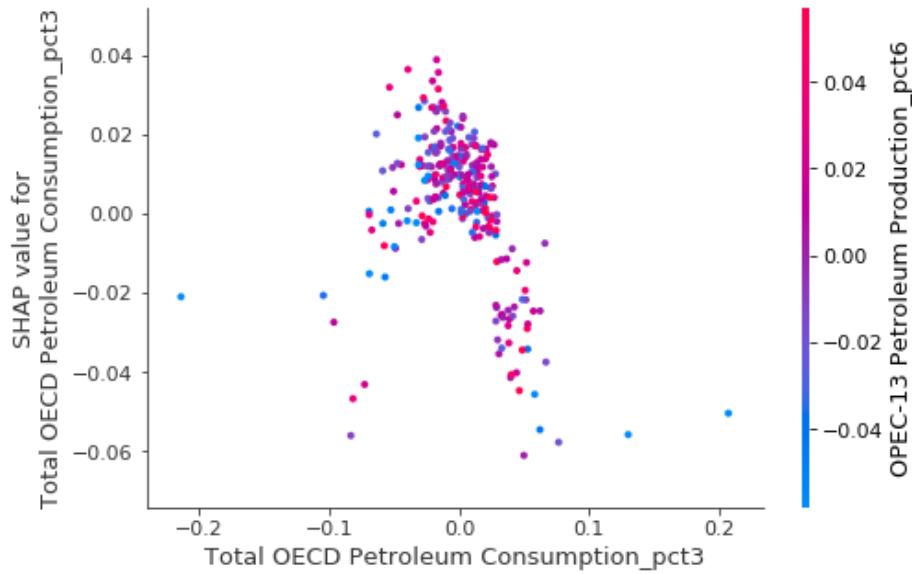
	原油供給	原油需要	原油先物価格	マーケットデータ	金融政策	モデル推計値
RF	10.5%	4.0%	6.4%	15.5%	10.3%	36.1%
LGBM	16.3%	6.2%	9.0%	17.6%	15.2%	40.2%
NN	24.4%	10.6%	12.4%	23.6%	30.2%	39.1%

被説明
変数
40.7%

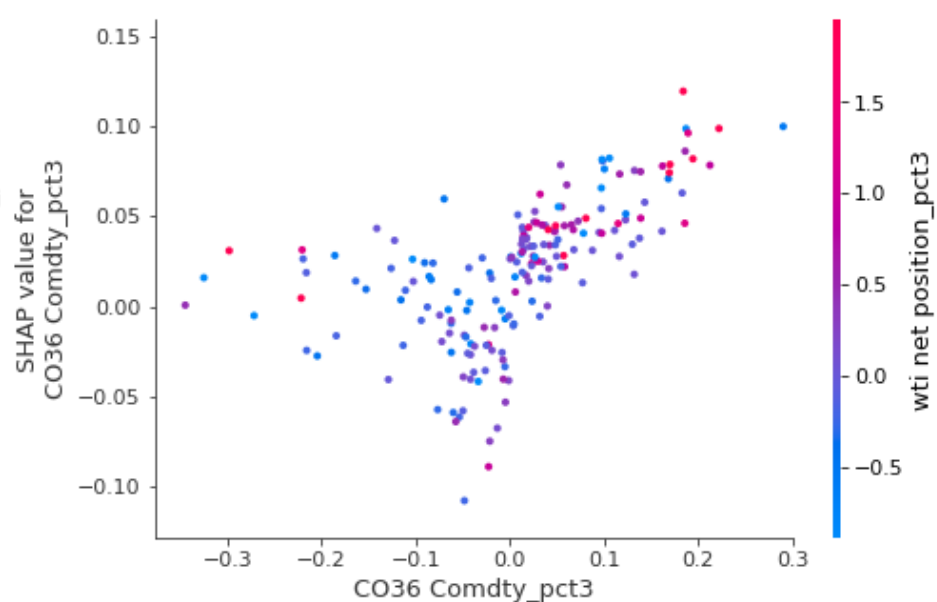
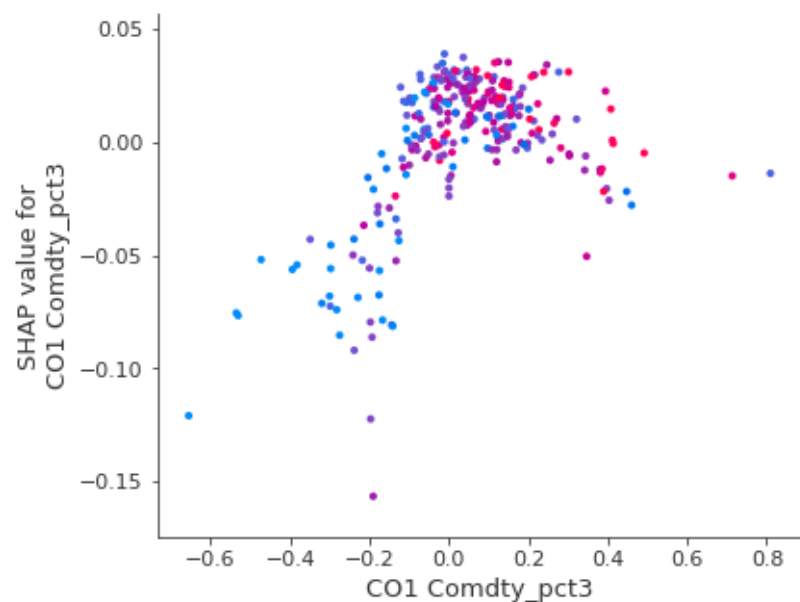
詳細分析 (原油供給)



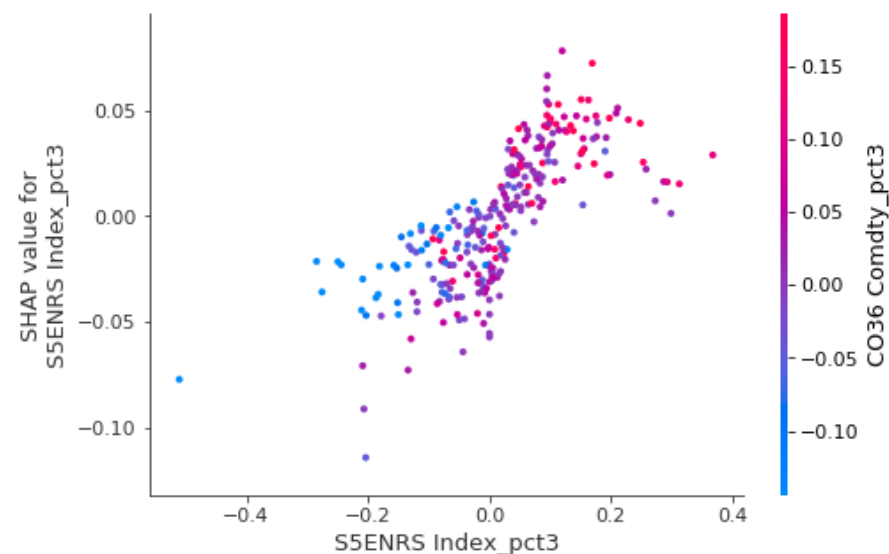
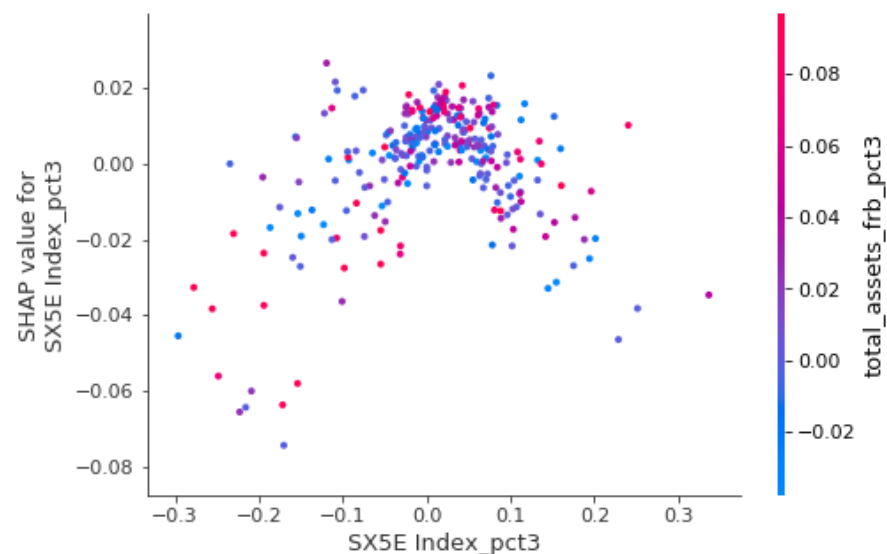
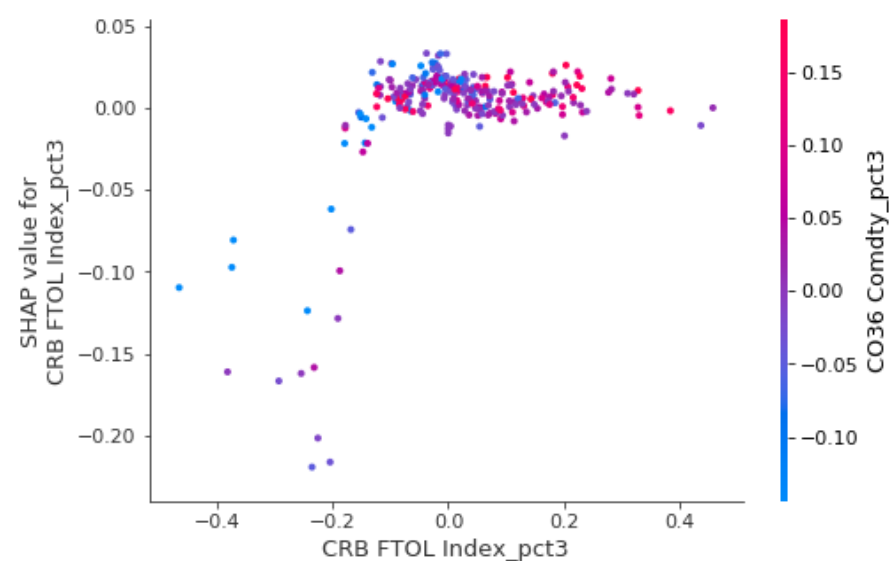
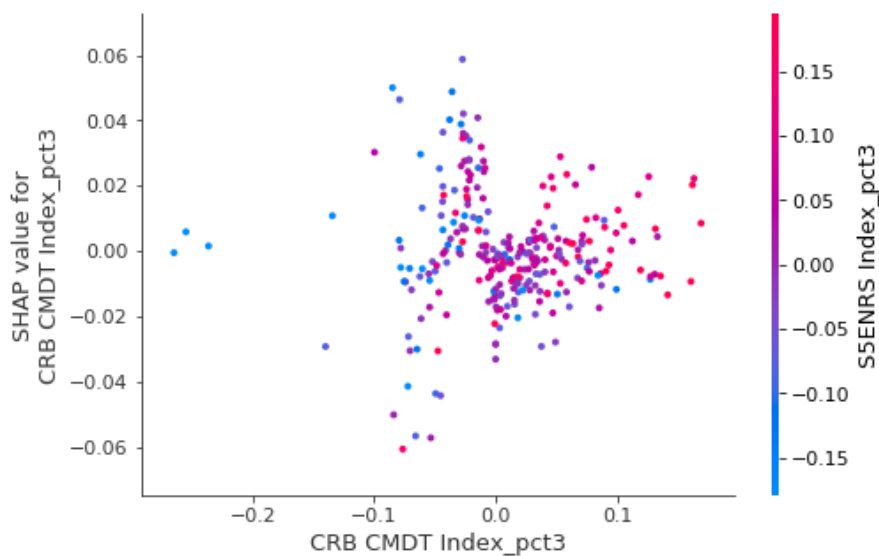
詳細分析 (原油需要)



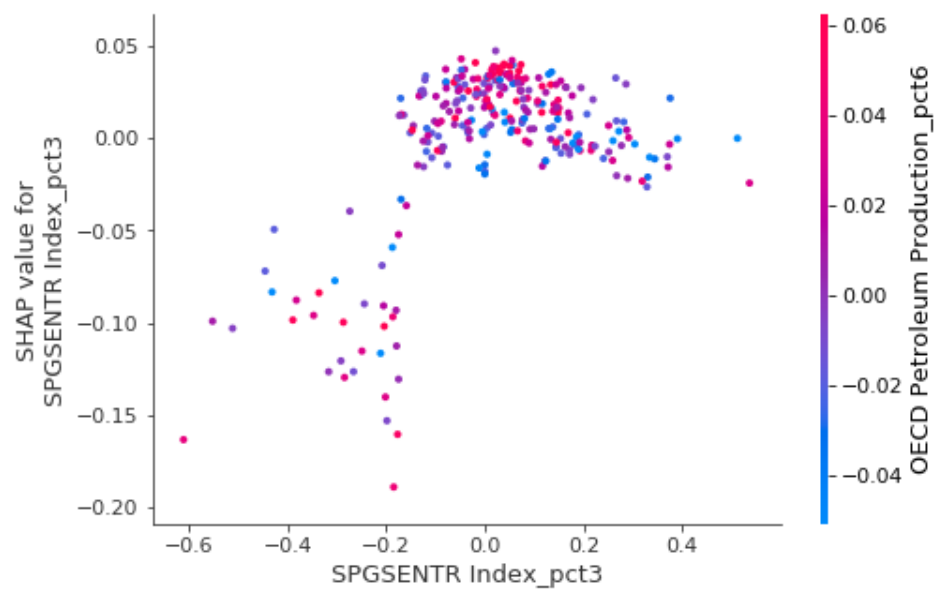
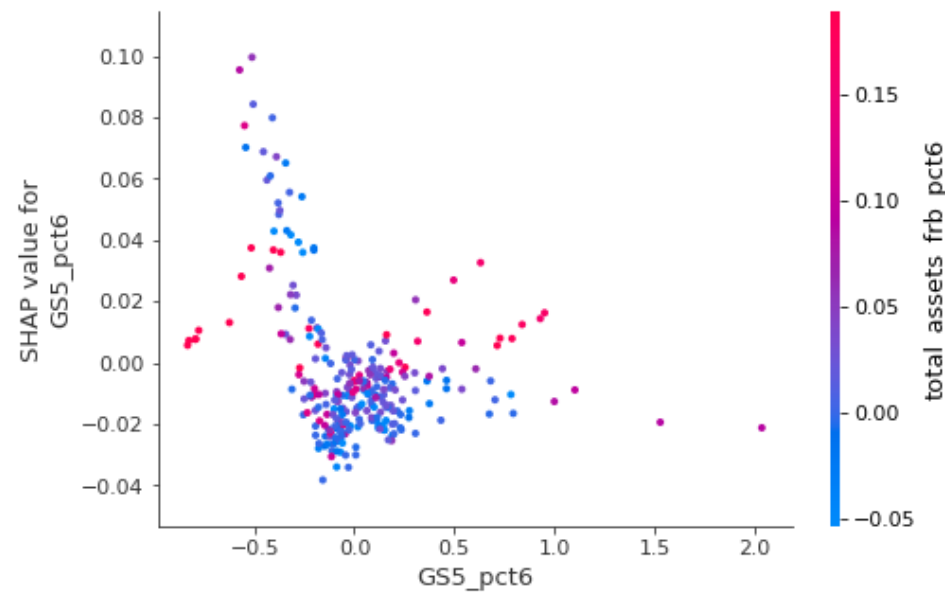
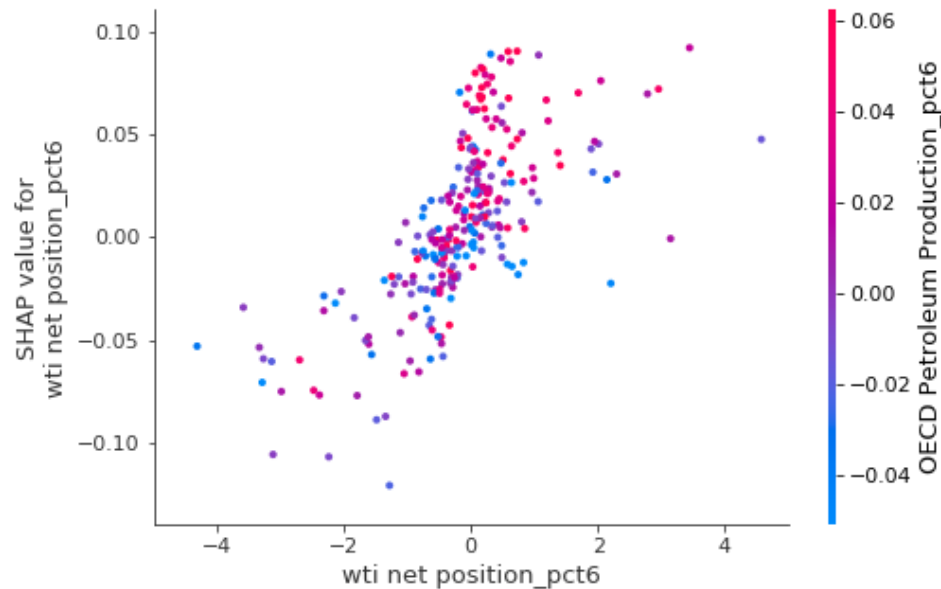
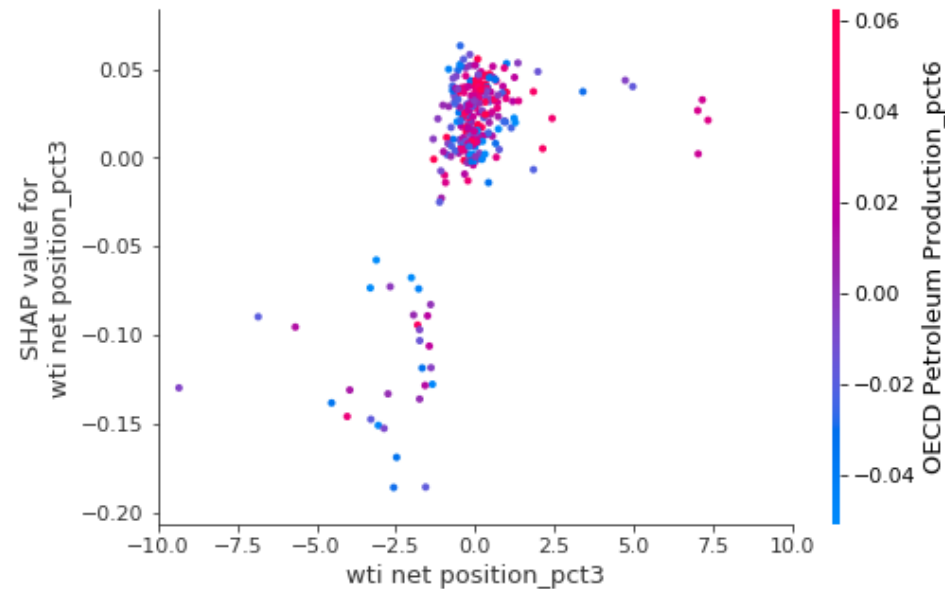
詳細分析 (原油先物価格)



詳細分析 (コモディティ価格)



詳細分析 (マーケットデータ)



詳細分析 (金融政策)

