

2.3 各種経済（景気）指標との関係

本項では、まず景気ウォッチャー調査のコメントから、同期の現状判断 DI および先行き判断 DI の原数値および各種経済指標を予測する分析を行う。すなわち、前項までの分析で景気指標との関係性があるとして抽出された有意単語 400 語の、特定年月における登場頻度を主たる説明変数とし、当該年月の現状判断 DI と先行き判断 DI の原数値または他の経済指標（一部指標は四半期系列）を目的変数とした回帰分析を行うことで、そのような予測を行う予測式が構成できないかの検討を行う。また、特に DI 値の予測においては、昨年度実施時に比べ、説明変数の変更などによって予測精度の改善が見込まれるかについての検証を行うことを主な目的とする。

本項では大きく分けて以下の分析群 A、B、C を実施している。

分析群 A：有意単語の発生頻度などによる現状判断 DI の予測・推計

分析群 B：有意単語の発生頻度などによる先行き判断 DI の予測・推計

分析群 C：有意単語の発生頻度などによる GDP など各種経済指標の予測・推計

また、分析群 A~C の実施においてはそれぞれ説明変数や目的変数を変化させ多数の分析を行っており、それぞれ 2.3 - 2 以降に後述する。

2.3 - 1 分析手順

分析群 A ~ C についてはいずれも、以下のステップにしたがっている。

説明変数、目的変数を定義する。

データセットを学習期間と予測期間に分割する。

学習データにおける「目的変数と説明変数の間の関係性モデル」を抽出する。

当該モデルを使用した予測の精度の評価と考察。

2.3.2 分析群A： 有意単語の発生頻度などによる現状判断 DI の予測

2.3.2.1 説明変数、目的変数の定義

分析群 A においては、表 2.3.1・表 2.3.2 に記載する A1～A8 の分析を行った。表 2.3.1 に A1～A8 それぞれの分析概要（除く説明変数）を、表 2.3.2 に各分析の説明変数を示す。

分析 A1：まず、昨年の分析と同様の説明変数・目的変数での分析を実施した。本分析は、今回分析における予測精度等の改善度合いを測る上でのベンチマークとなる分析である。

分析 A2：次に、目的変数を、DI 値から「前年との DI 値の差」に変更し、さらに説明変数において

- ・単語の出現数を出現頻度への変更と、
- ・シソーラスの上位語の出現頻度を導入し、
- ・さらに過去の DI 値に関する説明変数を大幅に省いて、再度分析した。これは、目的変数を前年差にすることでより単語の変化との関係が現れると考えたこと、および、予測精度向上のために出現数から出現頻度への変更や、シソーラスでの上位語の導入が有効であると考えられたからである。また、過学習¹が起こるという課題に対する解決策として、過去の DI 値に関する説明変数を省くことで説明変数数を削減した。

分析 A3：分析 A2 に対して、目的変数を DI 値に再変更して分析を実施した。これは分析 A2 の結果を解釈し、DI 値を目的変数にしたほうが良い結果になることが類推されたからである。

分析 A4：分析 A3 において、学習アルゴリズムをリッジ推定（L2 正則化付き重回帰分析）²に変更して分析を実施した。これは、分析 A3 の結果を解釈し、よりシンプルなモデルの方が良い結果になる可能性が考えられるとともに、多重共線性の影響を軽減するために妥当と考えたからである。

分析 A5：分析 A3 に対して、シソーラスの上位語のみに限定し、かつ前月との出現頻度差も説明変数に追加して分析を実施した。これは、分析 A4 までの分析によって前月との出現頻度差の追加が有効である可能性が考えられたが、単に説明変数を追加すると過学習が起こることが確実であったためである。そのため、変数削減を目的にシソーラスの上位語に限定した。

分析 A 6：分析 A 5 に対して、学習データから東日本大震災前後のデータを省いて分析を実施した。これは、分析 A 5 までにおいて、東日本大震災前後における個票内のことばが他の期間と比べ非常に異なっていたため、それを省くことで予測精度が向上する可能性があったからである。

分析 A 7：分析 A 3 に対して、説明変数から「前月の DI 値」など過去の DI 値に関する情報を全て省いて分析を実施した。これは、分析 A 3 3 に対して予測精度の向上はおそらく見込まれないが、個票内のことばと DI 値の関係という本調査の目的を考えると、直近の DI を含まないモデルの方が、よりことばと DI 値の関係が現れ有益な知見を得られると考えたからである。

分析 A 8：分析 A 7 に対して、シソーラスの上位語に限定した分析も実施した。

¹ モデルの複雑さに対し学習データサンプルが少なすぎる場合などに生じる現象。データサンプルの数が少ないと、学習期間にたまたま現れた説明変数の値を用いて目的変数を満たしてしまうケースが起こりやすくなる。この結果、予測期間においては精度が大きく低下する、すなわちモデルが十分な説明力を持っていなかったことが判明する現象。

² 通常の回帰分析の最小二乗法に、パラメータの二乗和に相当する正則化項を加えた式を最小にするように変更した重回帰分析。最適化対象が $(y - X\theta)^T (y - X\theta) + \alpha \theta^T \theta$ となる（ θ はパラメータ。回帰係数のことである）。 α は複雑度に関するパラメータであり、分析時に分析者が設定した。但しパラメータの推定量には偏りがある点に注意が必要である。

表 2 . 3 - 1 分析群 A における分析の概要

分析	条件設定	
A 1	学習期間 予測期間 目的変数 学習アルゴリズム	2011 年 03 月～2013 年 12 月（34 か月） 2010 年 1 月～2011 年 2 月は説明変数内に「1 年前の D I 値」を含むことにより対象外となった 2014 年 1 月～2014 年 12 月（12 か月） 現状判断 DI の原数値（全国および地域別） 異種混合学習
A 2	学習期間 予測期間 目的変数 学習アルゴリズム	A 1 と同じ A 1 と同じ 現状判断 DI の原数値の前年との差（全国および地域別） A 1 と同じ
A 3	学習期間 予測期間 目的変数 学習アルゴリズム	A 1 と同じ A 1 と同じ A 1 と同じ A 1 と同じ
A 4	学習期間 予測期間 目的変数 学習アルゴリズム	A 1 と同じ A 1 と同じ A 1 と同じ リッジ推定（L 2 正則化付き重回帰分析）
A 5	学習期間 予測期間 目的変数 学習アルゴリズム	A 1 と同じ A 1 と同じ A 1 と同じ A 1 と同じ
A 6	学習期間 予測期間 目的変数 学習アルゴリズム	2011 年 04 月～2011 年 2 月、2011 年 8 月～2013 年 12 月 2011 年 3 月～2011 年 7 月までを震災直後のため除外した A 1 と同じ A 1 と同じ A 1 と同じ
A 7	学習期間 予測期間 目的変数 学習アルゴリズム	2010 年 01 月～2011 年 2 月 A 1 と同じ A 1 と同じ A 1 と同じ

A 8	学習期間 予測期間 目的変数 学習アルゴリズム	A 7 と同じ A 1 と同じ A 1 と同じ A 1 と同じ
-----	----------------------------------	--

表 2 . 3 - 2 分析群 A における各分析の説明変数

項目 / 分析	A 1	A 2	A 3	A 4	A 5	A 6	A 7	A 8
月		○	○	○	○	○	○	○
エリアコード	○	○	○	○	○	○	○	○
地域ごとの有意表現を含む コメント数/ 出現割合	○ 数	○ 割合	○ 割合	○ 割合	○ 割合	○ 割合	○ 割合	○ 割合
同上・うちシソーラスの上位語に限定					○			○
地域ごとの有意表現を含む 出現割合の 1 か月前との差					○	○		
1 ヶ月前の DI	○	○	○	○	○	○		
2 ヶ月前の DI	○	○	○	○	○	○		
3 ヶ月前の DI	○							
4 ヶ月前の DI	○							
5 ヶ月前の DI	○							
6 ヶ月前の DI	○							
7 ヶ月前の DI	○							
8 ヶ月前の DI	○							
9 ヶ月前の DI	○							
1 0 ヶ月前の DI	○							
1 1 ヶ月前の DI	○							
1 2 ヶ月前の DI	○							
1 ヶ月前と 2 か月前の DI 差分	○	○	○	○	○	○		
2 ヶ月前と 3 か月前の DI 差分	○							
3 ヶ月前と 4 か月前の DI 差分	○							
4 ヶ月前と 5 か月前の DI 差分	○							

5ヶ月前と6か月前のDI差分	○							
6ヶ月前と7か月前のDI差分	○							
7ヶ月前と8か月前のDI差分	○							
8ヶ月前と9か月前のDI差分	○							
9ヶ月前と10か月前のDI差分	○							
10ヶ月前と11か月前のDI差分	○							
11ヶ月前と12か月前のDI差分	○							
12ヶ月前と13か月前のDI差分	○							
1ヶ月前から3か月前のDI平均	○	○	○	○	○	○		
2ヶ月前から4か月前のDI平均	○							
3ヶ月前から5か月前のDI平均	○							
4ヶ月前から6か月前のDI平均	○							
5ヶ月前から7か月前のDI平均	○							
6ヶ月前から8か月前のDI平均	○							
7ヶ月前から9か月前のDI平均	○							
8ヶ月前から10か月前のDI平均	○							
9ヶ月前から11か月前のDI平均	○							
10ヶ月前から12か月前のDI平均	○							
11ヶ月前から13か月前のDI平均	○							
12ヶ月前から14か月前のDI平均	○							

2.3-2-2 予測精度の評価結果

表2.3-3に各分析における精度の一覧を示す。なお精度一覧は平均誤差率を記載している。平均誤差率は、各予測値から実績値を引いた誤差の絶対値の平均値を、実績値の平均値で割ったものである。また、表中の赤字は複数の分析の中で評価期間の誤差率が最小であったもの、青字は2番目に小さかったものである。また表内の「全体」は、全国および各地域の誤差の平均値である。

なお、機械学習の精度評価において学習期間の精度の大小はそのモデルの優劣を評価する上では不適切であるため、比較する際には評価期間の精度を比較するのが適切である。また、学習期間の精度と評価期間の精度が大幅に異なる場合は、過学習がより強く発生していると考えられる。

表2.3-3 分析群Aにおける分析結果の精度一覧（平均誤差率を記載）

	A1:ベース分析		A2:前年差+シソーラス+割合化		A3:シソーラス+割合化		A4:シソーラス+割合化 #L2付き重回帰		A5:割合化+シソーラス限定+前月差分		A6: A5 + 震災除外		A7:テキストのみ (単語+シソーラス)		A8:テキストのみ (シソーラス限定)	
	学習	評価	学習	評価	学習	評価	学習	評価	学習	評価	学習	評価	学習	評価	学習	評価
全体	4.25%	9.29%	4.16%	16.75%	2.79%	5.52%	3.19%	5.06%	2.95%	6.14%	2.45%	5.58%	2.57%	5.81%	3.45%	5.99%
全国	2.21%	24.32%	2.30%	16.13%	1.43%	3.71%	1.58%	3.75%	1.24%	4.47%	1.53%	4.00%	1.42%	2.74%	1.79%	4.32%
北海道	4.79%	8.95%	4.37%	18.64%	2.97%	4.85%	3.09%	5.00%	2.71%	6.14%	2.30%	5.93%	2.77%	6.77%	3.50%	8.00%
東北	4.02%	7.77%	4.05%	20.35%	3.69%	5.04%	3.90%	4.76%	3.50%	7.02%	2.59%	5.56%	3.03%	4.50%	4.60%	4.99%
北関東	5.05%	8.17%	4.25%	21.61%	2.73%	8.31%	3.16%	7.65%	3.24%	6.73%	1.89%	8.90%	2.29%	8.73%	3.12%	10.27%
南関東	2.61%	7.82%	4.33%	17.55%	2.57%	4.98%	3.18%	3.84%	2.73%	6.41%	2.17%	4.17%	2.48%	5.23%	2.71%	4.77%
東海	3.67%	8.73%	5.32%	16.33%	2.61%	5.29%	2.73%	4.91%	2.90%	5.60%	2.21%	6.04%	2.94%	4.58%	2.59%	5.85%
北陸	5.32%	7.30%	4.82%	11.98%	2.67%	5.48%	3.05%	6.04%	3.15%	6.04%	2.80%	5.59%	2.32%	8.02%	3.58%	6.51%
近畿	3.83%	6.28%	3.84%	13.50%	2.52%	3.19%	3.04%	3.10%	2.58%	3.50%	2.58%	3.91%	2.70%	4.77%	3.22%	3.58%
中国	4.24%	7.83%	3.68%	19.36%	3.57%	6.15%	4.07%	6.14%	3.29%	6.54%	3.01%	5.89%	3.08%	5.90%	3.49%	5.09%
四国	4.30%	8.28%	4.75%	17.46%	3.22%	5.56%	3.67%	4.99%	3.47%	7.26%	2.48%	5.49%	3.08%	4.40%	4.09%	5.49%
九州	4.52%	7.73%	3.49%	13.53%	2.84%	5.55%	2.97%	5.52%	2.83%	6.28%	2.15%	6.06%	2.31%	5.63%	3.55%	6.22%
沖縄	6.26%	8.39%	4.72%	15.36%	2.68%	8.07%	3.85%	5.20%	3.74%	7.71%	3.53%	5.67%	2.43%	8.34%	5.03%	7.03%

2.3.2.3 結果の考察

表 2.3.3 の結果を元に考察する。

まず、分析 A1～A4 の結果を比較する。図 2.3.1 に、各分析における全国の DI 値の評価期間の予測値を実際の値（実績値と呼ぶ）と比較した結果を示す。なお、各地域の評価期間の予測値は資料編に全て記載されている。

分析 A2 において、分析 A1 と比較して評価期間の誤差が悪化した。これは、目的変数を前年差にしたことにより、東日本大震災の翌年 3 月（2012 年 3 月）が最高値（前年と比較して最も DI 値が増加した）となったためであると考えられる。なぜなら 2012 年 3 月の DI 値は前後と比較して特別高い期間でないにも関わらず目的変数が非常に大きいため、この期間に特別数が多いことばに関する説明変数を重視しすぎ、本来あるべきモデルとの乖離が起き（2012 年の 3 月付近への過学習が起こったと言える）結果として評価期間の予測精度が悪化したと考えられる。景気ウォッチャー調査の DI 値やデータ系列が現在月単位であることを考えると、目的変数を前年差にする分析は今後も同様の問題が発生すると考えられるため本テーマには適さない手法であることがわかった。

分析 A3、A4 において、精度が分析 A1 と比較して精度が良いことが分かった。特に、全分析結果を比較しても A4 が最高精度であった。これは、A4 が最も単純なモデルを出力しやすい手法であり、現在の学習データ数からは複雑なモデルではより過学習が起きやすいことが原因であると考えられる。しかし、リッジ推定はパラメータの推定量に偏りがあることから、推定値の期待値（平均値）が真のパラメータ値とならないことや、また今後データの数が増加した際には、より複雑なモデルの方が高精度を発揮する可能性があるため、本件は引き続き検討が必要であると思われる。

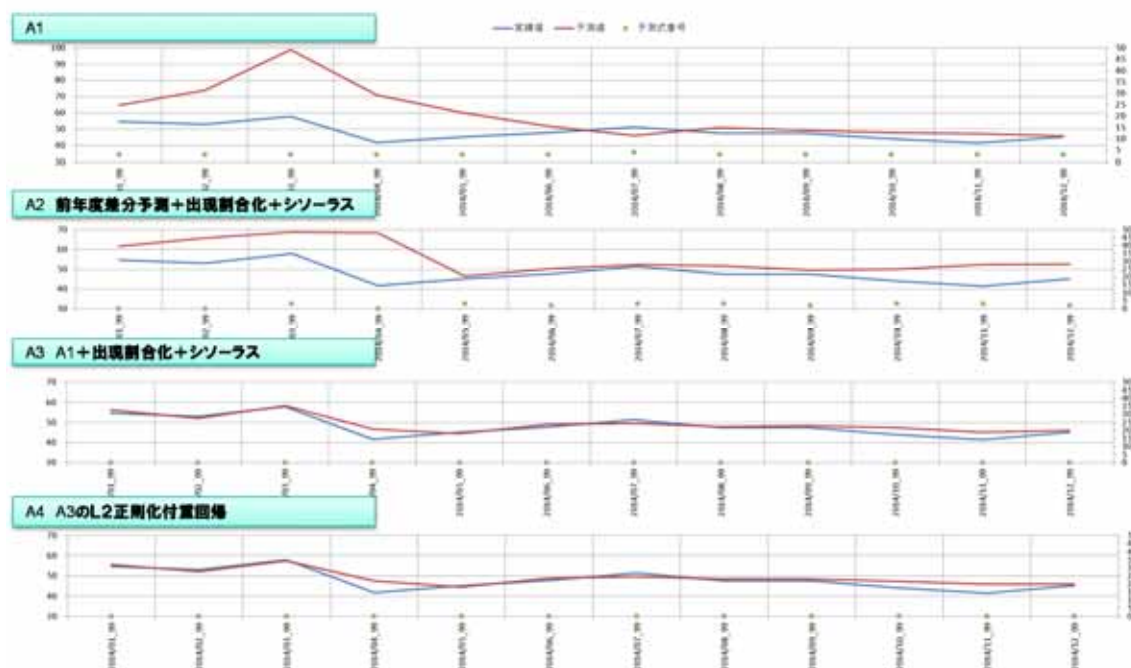
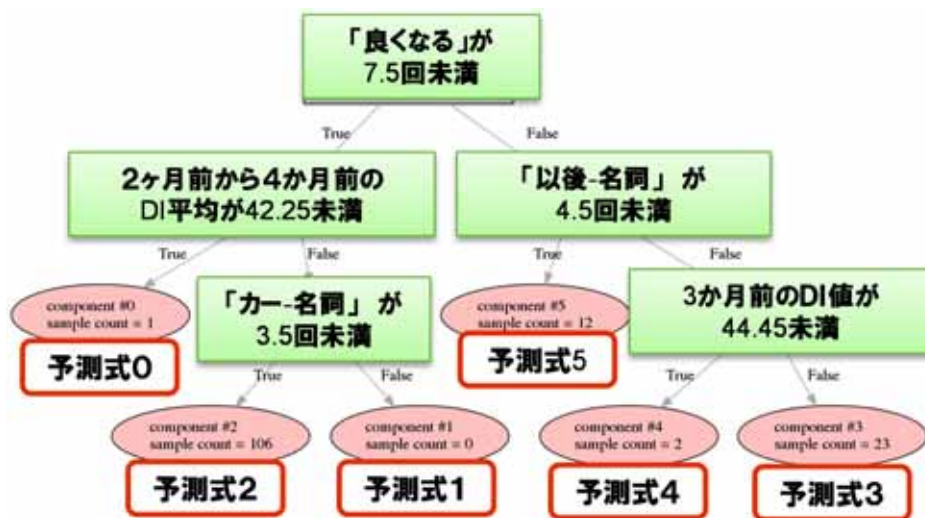


図2.3-1 分析A1～A4における全国のDI値の評価期間の予測値・実績値

図2.3-2に分析A1において学習器が出力した予測モデルを、図2.3-3と図2.3-4と図2.3-5に分析A2、A3、A4の予測モデルをそれぞれ示す。なお、予測モデルの上段は異種混合学習によって分割された分割条件であり、分割されたデータごとに別の予測式が割り当てられている。例えば図2.3-3の上段は、例えば「カー」という名詞が、ある地域のある月で3.5回未満だった場合に予測式2番となるなど、各地域・月におけることばの出現回数やDI値などによって予測式を選択する条件を示している。そして下段の棒グラフはそれぞれの予測モデルにおける説明変数の係数を示している。棒グラフにおいて正に大きい係数は、その変数が大きいときに目的変数(DI値)が大きくなることを示しており、負に大きい係数は、その変数が小さいときに目的変数(DI値)が大きくなることを示している。また、モデル内に「@」付きで記載されたことばは、シソーラスの上位語である。

モデルの比較を行うと、分析A3、A4において、シソーラスの上位語に対して、正または負に大きな係数が付与されていることがわかる。これは、シソーラスの導入によって景気と関係がある語がまとめられたことを示唆しており、上位語としたことによる効果であると考えられる。分析A3、A4の精度がよいことから、シソーラスの導入は予測精度の向上に寄与していると思われる。一方で、分析A3、A4においてもシソーラスで括られていない単独の単語に大きな係数があることがある(「キャンセル」など)。これは、今後そのようなことばを上位語に含むようにシソーラスの見直しをすることにより、さらに良いモデル・精度を導出できる可能性があることを示唆している。



(注) 各色の表示部分の長さが、予測式の係数の大きさを示す(以下の図も同様)。

図2.3-2 分析A1の予測モデル

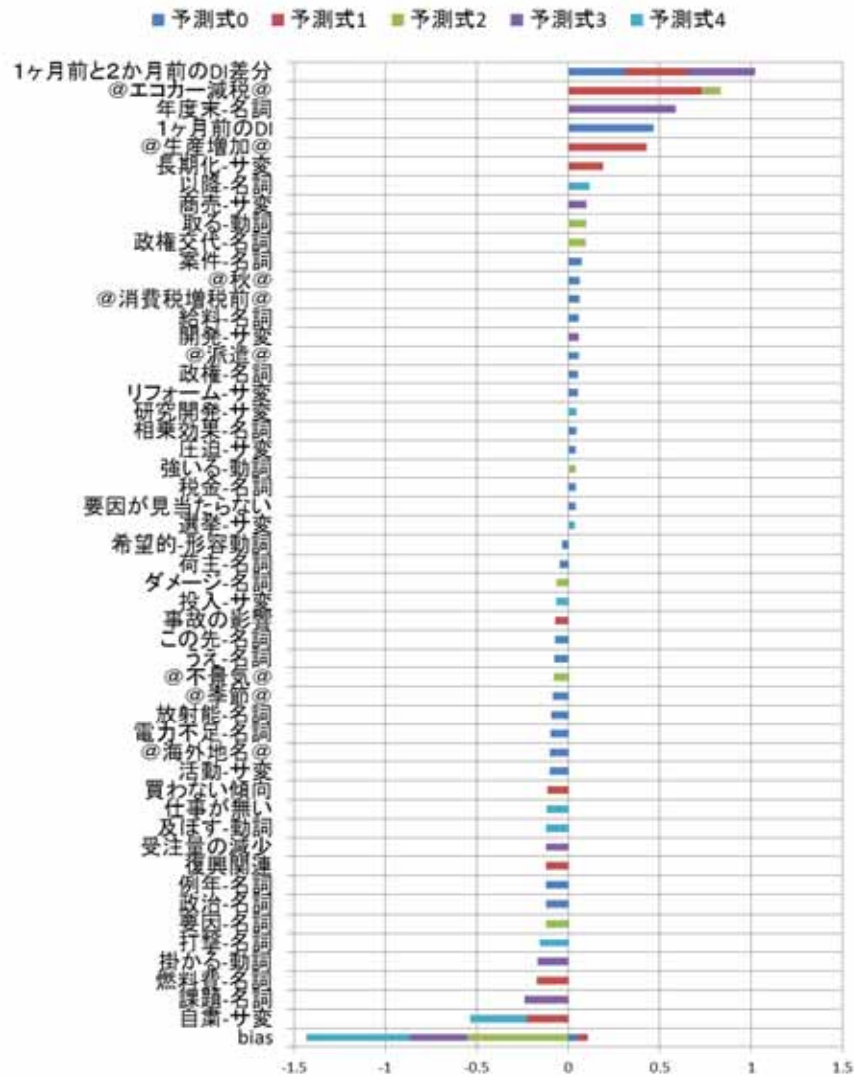
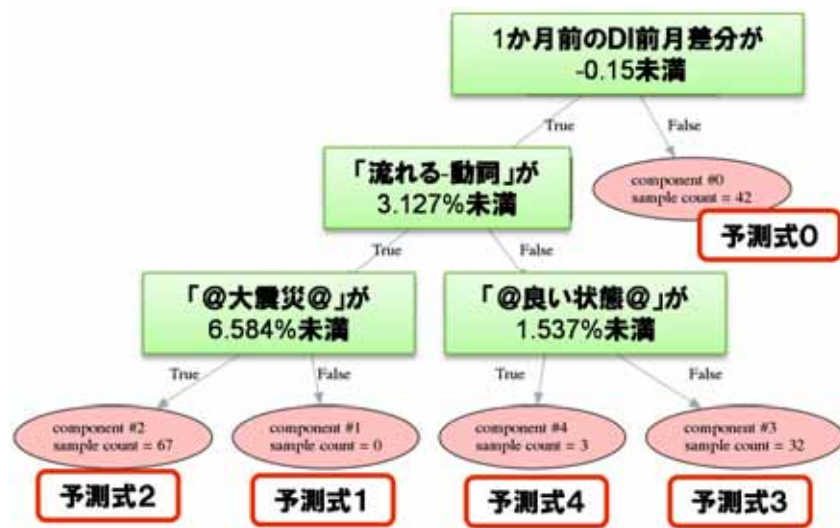


図2 . 3 - 3 分析A2の予測モデル

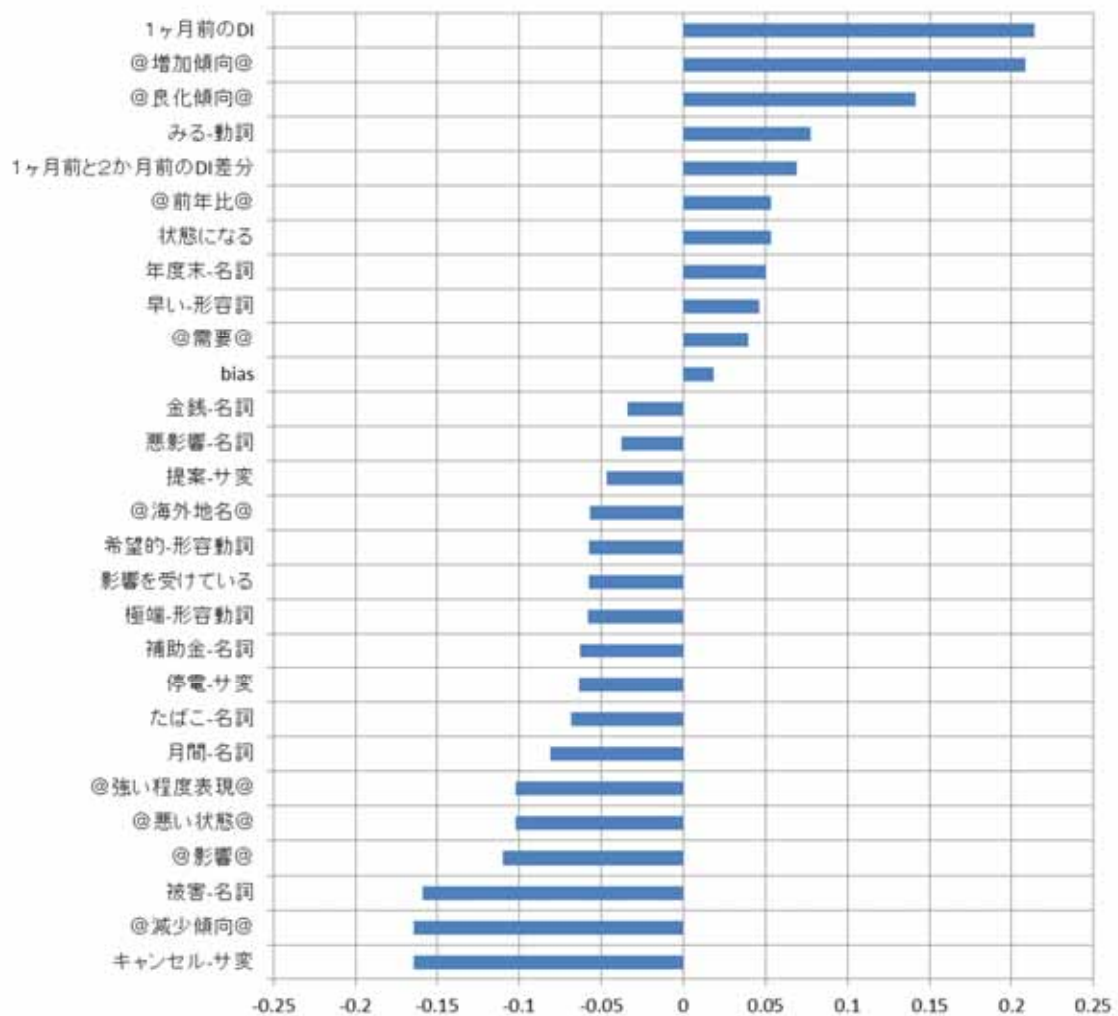


図 2 . 3 - 5 分析 A 4 の予測モデル

次に、分析 A5～A8の結果を比較する。図2.3-6に、各分析における全国のDI値の評価期間の予測値を実際の値（実績値と呼ぶ）と比較した結果を示す。

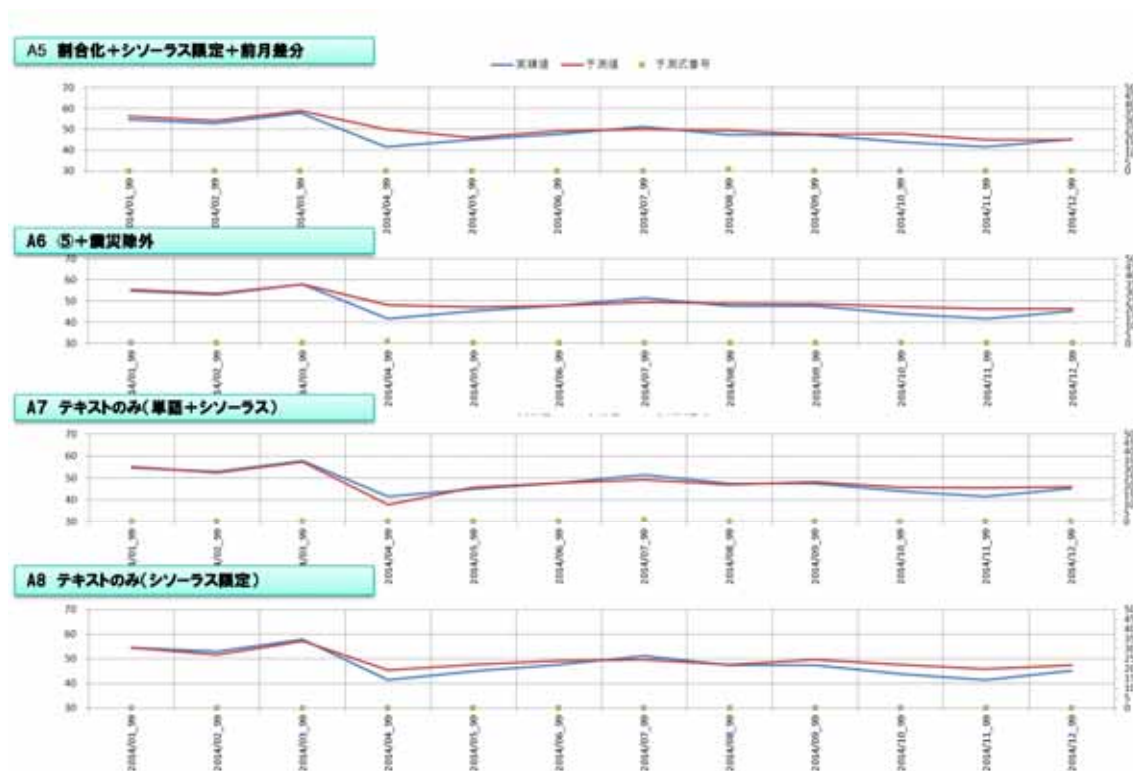


図2.3-6 分析A5～A8における全国のDI値の評価期間の予測値・実績値

まず、表2.3-3にあるように分析A5からA8までの結果の精度は地域ごとにやや差がある。また、分析A5およびA6は図2.3-6にあるように2014年4月や11月の予測精度が少し悪い。これは、学習データに前月とのことば頻度の差を追加したこと（分析A5、A6）や震災データを除外したこと（分析A6）により、説明変数数が増えた、及び学習サンプル数が減ったことにより過学習が増加したことが原因であると考えられる。

また、分析A7、A8の予測精度はA3、A4よりやや劣るが、大差ないことがわかった。これは、過去のDI値の指標を使わず、ことば情報だけでも十分な予測ができる可能性を示唆している。

図2.3-7、2.3-8、2.3-9、2.3-10に、分析A5～A8の予測モデルを示す。

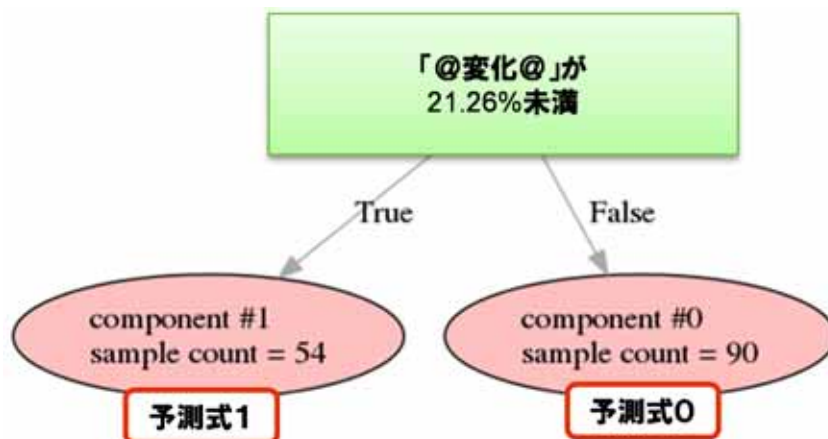


図2 . 3 - 7 分析 A5 の予測モデル



図2 . 3 - 8 分析 A6 の予測モデル

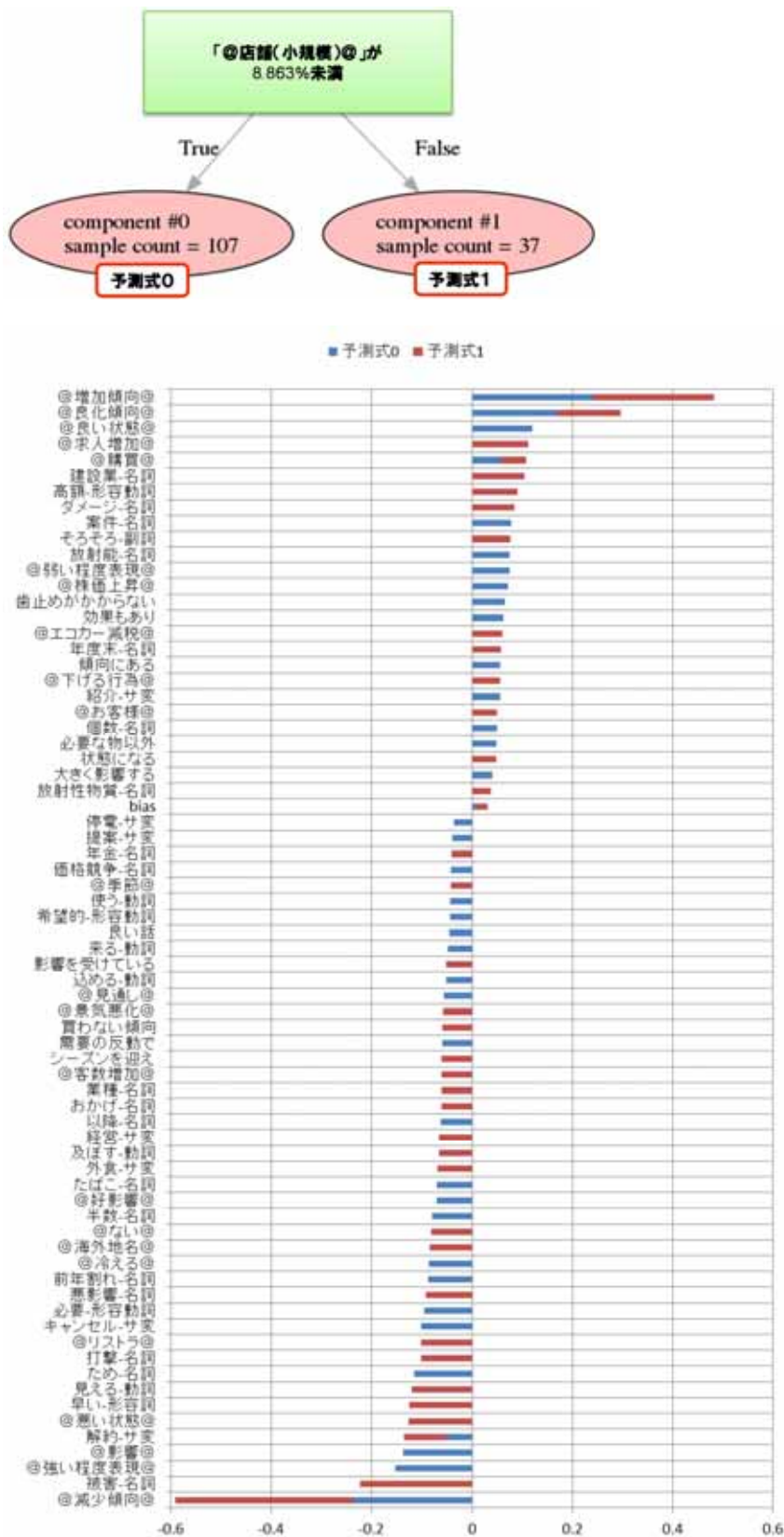


図2.3-9 分析A7の予測モデル

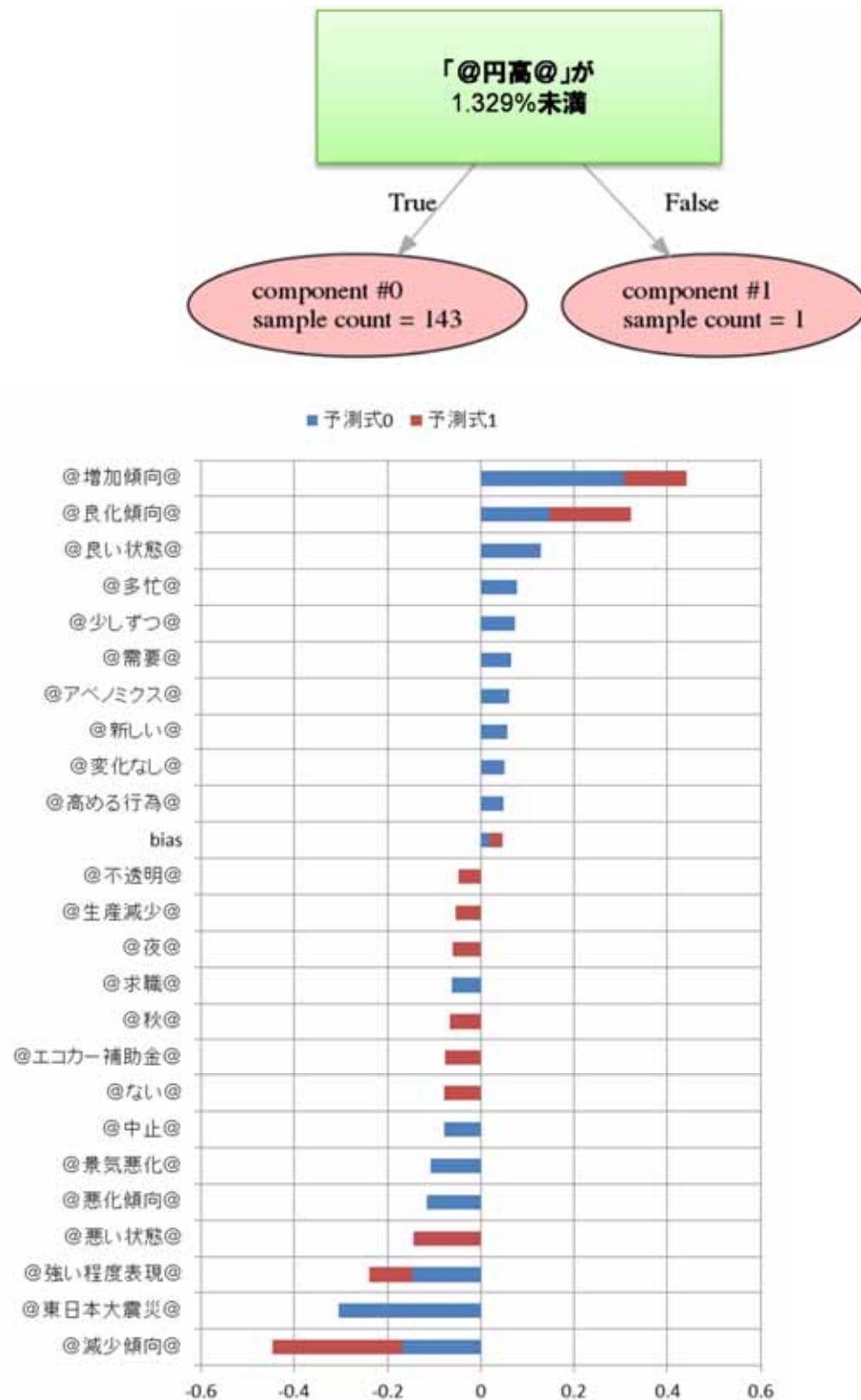


図 2 . 3 - 1 0 分析 A 8 の予測モデル

2 . 3 - 3 分析群 B： 有意単語の発生頻度などによる先行き判断 DI の予測

2 . 3 - 3 - 1 説明変数、目的変数の定義

分析群 B においては、分析群 A で実施した分析 A 1 ～ A 8 の全ての分析を、目的変数を先行き判断 DI 値に変更することで実施した。本項ではこれを分析 B 1 ～ B 8 と呼ぶ。分析 B1～B8 の概要は表 2 . 3 - 1 を先行き判断 DI 値に変更したものと同一であり、説明変数についても表 2 . 3 - 2 と同一である。

2 . 3 - 3 - 2 予測精度の評価結果

2 . 3 - 2 - 2 項と同様に、表 2 . 3 - 4 に各分析における精度の一覧を示す。

表 2 . 3 - 4 分析群 B における分析結果の精度一覧（平均誤差率を記載）

	B1: ベース分析		B2: 前年差 + シソーラス + 割合化		B3: シソーラス + 割合化		B4: ソーラス + 割合化 #L2 付き重回帰		B5: 割合化 + シソーラス限定 + 前月差分		B6: B5 + 震災除外		B7: テキストのみ(単語 + シソーラス)		B8: テキストのみ(シソーラス限定)	
	学習	評価	学習	評価	学習	評価	学習	評価	学習	評価	学習	評価	学習	評価	学習	評価
全体	3.73%	10.40%	3.57%	19.59%	2.41%	7.37%	2.93%	8.05%	2.54%	8.05%	2.35%	8.55%	2.59%	6.98%	3.41%	8.62%
全国	2.15%	12.94%	2.53%	19.89%	1.20%	5.05%	1.08%	6.49%	1.14%	6.41%	1.12%	7.18%	1.29%	5.71%	1.55%	7.69%
北海道	4.47%	12.29%	2.64%	20.95%	1.96%	10.92%	3.14%	9.15%	2.62%	10.23%	2.34%	9.76%	2.31%	7.25%	2.95%	7.20%
東北	3.52%	12.72%	3.67%	18.34%	2.76%	7.86%	3.10%	9.22%	2.60%	8.73%	2.12%	10.92%	2.58%	7.21%	4.11%	10.03%
北関東	3.46%	10.09%	3.84%	21.03%	2.48%	8.05%	3.34%	8.50%	3.07%	8.14%	2.64%	9.43%	2.89%	6.99%	3.65%	9.81%
南関東	3.74%	9.81%	3.92%	19.61%	2.84%	5.56%	2.60%	6.52%	2.14%	5.51%	2.08%	7.55%	2.61%	5.00%	3.04%	7.16%
東海	3.31%	9.41%	4.30%	22.78%	2.35%	7.40%	2.88%	9.08%	2.50%	8.38%	2.50%	6.89%	3.20%	7.75%	3.64%	8.99%
北陸	3.85%	11.08%	3.84%	19.60%	2.96%	8.61%	3.18%	6.99%	2.46%	8.25%	2.79%	8.74%	2.73%	6.79%	3.59%	7.72%
近畿	3.49%	9.06%	3.27%	22.52%	1.97%	6.54%	2.59%	6.79%	2.38%	6.73%	2.38%	6.95%	2.51%	6.20%	3.02%	7.29%
中国	3.66%	9.64%	3.78%	18.23%	2.71%	6.08%	3.16%	9.11%	2.91%	8.68%	2.34%	7.62%	2.79%	9.05%	4.07%	10.03%
四国	4.39%	11.09%	3.83%	20.43%	2.84%	7.89%	3.01%	8.06%	3.04%	10.09%	2.74%	10.62%	2.62%	8.31%	3.40%	8.81%
九州	3.40%	9.62%	3.53%	17.94%	2.46%	5.99%	3.15%	6.60%	2.82%	6.32%	2.69%	8.51%	2.41%	6.70%	3.13%	8.32%
沖縄	5.14%	7.35%	3.78%	14.39%	2.44%	8.59%	3.85%	10.06%	2.73%	9.21%	2.47%	8.69%	3.08%	6.89%	4.75%	10.33%

2.3.3.3 結果の考察

表2.3.4の結果を元に考察する。

分析群 A と結果と異なり、精度が良い分析結果が B7 および B3 となった。しかし、分析 A7 や A3 も精度が低くはなく、全体的な傾向は分析群 A と分析群 B は類似していると思われる。

図2.3.11と図2.3.12に、各分析における全国の DI 値の評価期間の予測値を実際の値（実績値と呼ぶ）と比較した結果を示す。なお、各地域の評価期間の予測値は資料編に全て記載されている。

全体的に、2014 年 3 月の予測精度が良くない。これは消費税増税前に「購買」など景気に対して正に影響すると学習されたことばが多数存在したからであると考えられ、本件の改善は今後の課題とする。一方で、分析 B7 は、2014 年 3 月を除くと概ね傾向を示せていると考えられ、過去の DI 値情報がなくことば情報だけでも高い精度で DI 値を予測できる可能性があることが分かった。

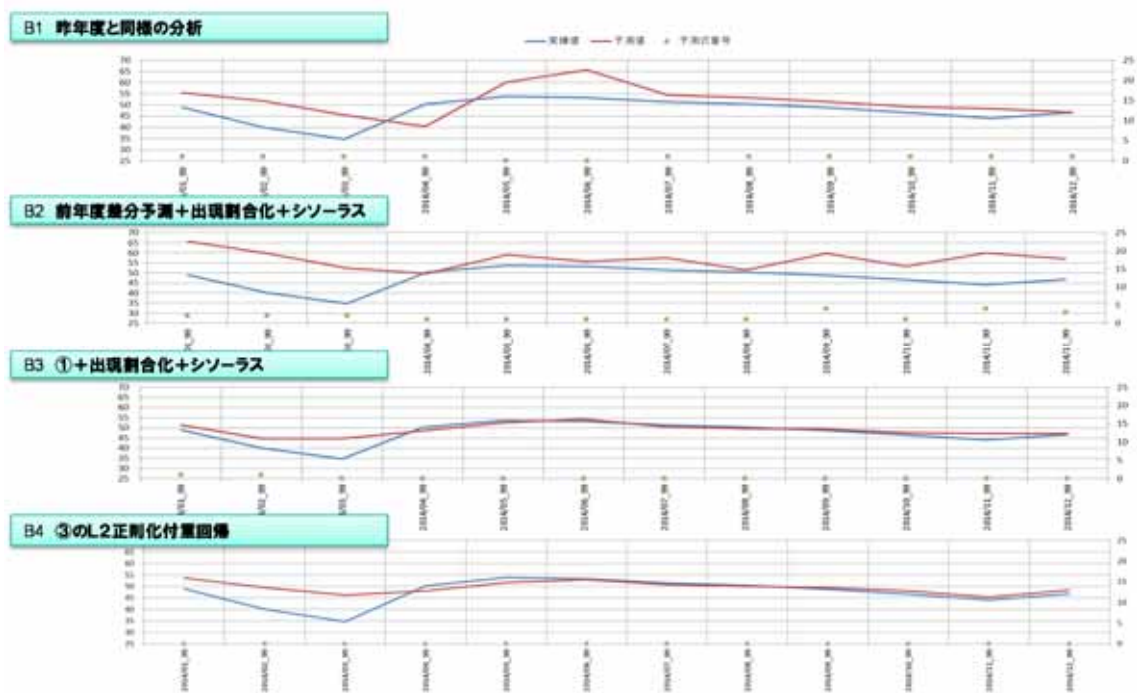


図2.3.11 分析 B1～B4 における全国の DI 値の評価期間の予測値・実績値

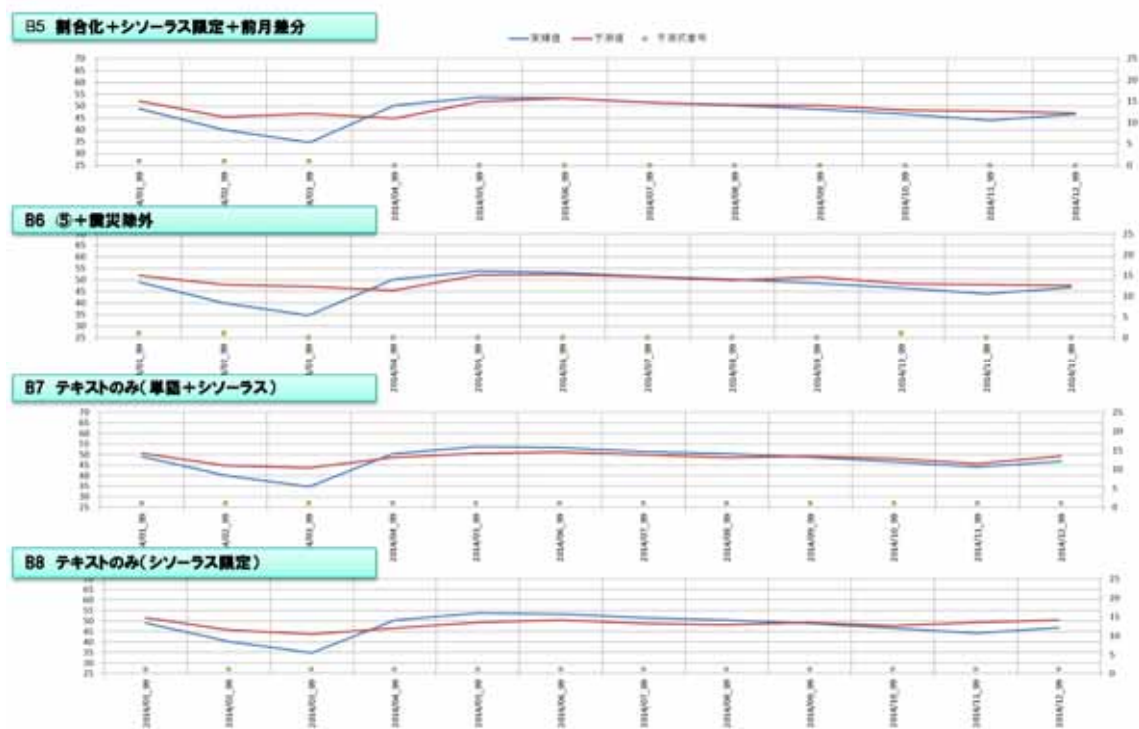


図 2 . 3 - 1 2 分析 B 5 ～ B 8 における全国の DI 値の評価期間の予測値と実績値

2 . 3 - 4 分析群 C：有意単語の発生頻度などによる GDP など各種経済指標の予測

2 . 3 - 4 - 1 説明変数、目的変数の定義

分析群 C では、分析群 A,B と同様の分析を、目的変数を別の経済指標にして実施した。

表 2 . 3 - 6 に分析群 C の各分析の概要を、表 2 . 3 - 7 に分析群 C の各分析における説明変数を示す。

表 2 . 3 - 6 に示したように各目的変数において 2 つずつ分析番号があるのは、説明変数において、景気ウォッチャー個票内の各有意表現の出現割合を現状に関するコメントで計算した場合と、先行きに関するコメントで計算した場合の 2 通りを実施したからであり、番号が小さい方を現状のコメントを用いた分析、番号が大きい方を先行きのコメントを用いた分析とする。

表 2 - 3 - 6 分析群 C における分析の概要

分析	条件設定	
C1	学習期間	2010 年 04 月～2013 年 12 月（45 か月） 2010 年 1 月～2010 年 3 月は説明変数内に「1ヶ月前から3ヶ月前の目的変数値平均」を含むことにより対象外となった
C2	予測期間	2014 年 1 月～2014 年 9 月（9 か月）
	目的変数	地域別消費総合指数
	学習アルゴリズム	異種混合学習
C3	学習期間	C1 と同じ
C4	予測期間	C1 と同じ
	目的変数	地域別民間住宅総合指数
	学習アルゴリズム	C1 と同じ
C5	学習期間	C1 と同じ
C6	予測期間	C1 と同じ
	目的変数	地域別民間設備投資総合指数
	学習アルゴリズム	C1 と同じ
C7	学習期間	C1 と同じ
C8	予測期間	C1 と同じ
	目的変数	地域別鉱工業生産指数
	学習アルゴリズム	C1 と同じ
C9	学習期間	C1 と同じ
C10	予測期間	C1 と同じ
	目的変数	公共投資総合指数
	学習アルゴリズム	C1 と同じ
C11 ～ C18	学習期間	2010 年 1-3 月期～2013 年 10-12 月期
	予測期間	2014 年 1-3 月期～2014 年 7-9 月期
	目的変数	実質 GDP(分析 C11、C12)・実質 GDP[財貨・サービス・輸出]（分析 C13、C14）・実質 GDP[財貨・サービス・輸入]（分析 C15、C16）・実質 GDP[民間在庫品増加]（分析 C17、C18）
	学習アルゴリズム	C1 と同じ
C19 ～ C26	学習期間	C11 と同じ
	予測期間	C11 と同じ
	目的変数	名目 GDP(分析 C19、C20)・名目 GDP[財貨・サービス・輸出]（分析 C21、C22）・名目 GDP[財貨・サービス・輸入]（分析 C23、C24）・名目 GDP[民間在庫品増加]（分析 C25、C26）
	学習アルゴリズム	C1 と同じ

表 2 . 3 - 7 分析群 C の各分析における説明変数

月
エリアコード
景気ウォッチャー個票内の各有意表現の出現割合（現状/先行き）
1 ヶ月前(または 1 期前)の目的変数値
2 ヶ月前（または 2 期前）の目的変数値
1 ヶ月前と 2 か月前（または 1 期前と 2 期前）の目的変数値差分
1 ヶ月前から 3 か月前（または 1 期前と 3 期前）の目的変数値平均

2 . 3 - 4 - 2 予測精度の評価結果

表 2 . 3 - 8 に各分析の結果の精度を示す。分析群 A、B と同様に、評価期間の誤差率の平均値を記載する。ただし、分析群 C においては、各分析の目的変数が異なることから、誤差率を比較することは適切ではない（目的変数の平均と分散によって、予測結果に対する誤差率の傾向が異なるからである）。

そこで、各分析結果の予測・実績のグラフを目視で確認し、定性的な評価を実施した。評価に際しては、値の上下を予測できているか、予測値の平均値が実績値の平均値と一致しているか、最大値と最小値の時期が一致しているかなどを判断基準とし、○、△、× の 3 種類の評価を実施した。評価結果は表 2 . 3 - 8 に併せて示す。

表 2 . 3 - 8 分析群 C の各分析の精度及び定性評価結果

分析 番号	目的変数	評価期間誤差率		定性 評価	定性評価コメント
		現状コメント利 用時	先行きコメ ント利用時		
C1C2	地域別消費総合指数	3.05%	3.39%	×	消費税増税前後の変化が全く予測できていない
C3C4	地域別民間住宅総合指数	3.89%	4.42%	○	概ね傾向を捉えられている
C5C6	地域別民間設備投資総合指数	4.82%	5.19%	×	予測値が実績値以上に上下し外れている
C7C8	地域別鉱工業生産指数	2.56%	3.47%		目的変数自体の変化量が小さく評価が難しい
C9C10	公共投資総合指数	3.87%	5.84%	○	全体的な傾向に追従できている
C11 ~ C18	各種実質 GDP	25.55%	27.20%	×	四半期単位でありデータ数が少なく過学習しており2014年の傾向を捉えられていない
C19 ~ C26	各種名目 GDP	23.54%	30.06%	×	四半期単位でありデータ数が少なく過学習しており2014年の傾向を捉えられていない

2 . 3 - 4 - 3 結果の考察

分析群 C においては、多数の分析を実施したが、その多くは過学習によって評価期間の予測精度が良くない結果となった。これは景気ウォッチャーと各種経済指標の意味する対象の違いも理由ではあるが、学習サンプル数の少なさが大きな理由であると考えられる。特に GDP は四半期単位のデータであることから、サンプル数が極端に少なく、過学習している。この問題は、GDP の各種構成項目を統合して学習することである程度改善する可能性があり、今後の課題としたい。

一方、地域別民間住宅総合指数や公共投資総合指数など全体的な傾向を予測できている指標もあり、これらの指標に関して、今後分析群 A、B で行ったような詳細な分析を実施することでより推定精度が高い予測モデルを作ることができると考えられる。

2.3-5 2.3全体のまとめ

今回、分析群 A、B で予測精度改善のための各種方式を検討した。その結果以下のことがわかった。

1. シソーラスの導入により、上位概念が予測モデルに取り入れられ、安定した予測モデルになり予測精度が改善する。一方で、上位概念・下位概念の構造の見直しによりさらに安定したモデルへ改善する可能性がある。今後、学習結果のモデルに下位概念が強く採用されている時に、概念構造の見直しを行うと良い。
2. 説明変数から過去の DI 値に関する情報を省いても予測精度はほぼ落ちない。個票内のことば情報だけでも DI 値の推定は可能である。これは、今後、twitter 情報、Web 内の他の情報を用いた DI 値の推定を行った際に再度確認が必要であるが、ことば情報の可能性を示したと言える。
3. 目的変数を前年差にすることは現状のデータ数では精度悪化を招く可能性が高く良くない。本精度悪化の対策として、説明変数のことば情報も差分を取った後の変数を採用することが考えられるが、完全に防ぐことは困難であると考えられるため、本施策は今後採用しないほうが良い。
4. 消費税率引上げ前の先行き DI 値等予測精度に課題が残る部分もあった。この時期の個票内のことばをさらに調べ、政策・施策の変化に対して安定的に予測できるモデルを出すためのシソーラスなどを検討すると良い。
5. モデルを解釈すると「@強い程度表現」などの、心情や度合を示す副詞や形容詞が、大きい係数として採用されていることが多かった。またその度合は先行き指標の方が強い。これは、景気をことばで表現するとき、「かなり」「とても」「非常に」などのことばを経済・景気・産業に関することばと組み合わせて用いていることを示している。今後、どのような形容詞・副詞がより多く記述されているのかについての分析を行うことで、さらに精緻な予測モデル構築のための知見を得ることができると考えられる。
6. 現状 DI 値の推定における北関東地方など、地方を比較したときに精度が低い地域があった。これは地域ごとにことばの傾向に差がある可能性を示しているが、詳細な調査は今後の課題としたい。

また、分析群 C において DI 値以外の目的変数の予測・推計を実施した。その結果以下のことが分かった。

- 7 . GDP 等四半期単位の指標は過学習が極めて起きやすく予測精度が悪い。
今後期間を大幅に増やすことや、GDP の各種構成項目を加えて分析するなどデータ数増加の施策が必要である。
- 8 . 地域別民間住宅総合指数や公共投資総合指数など全体的な傾向を予測できている指標がある。この結果について、これらの指標の予測精度を高める施策をさらに実施するとともに、他の指標との違いを調べ、どのような指標が景気ウォッチャーのことは情報との関係が深いのかについて検証すると良い。