

1. はじめに
2. EBPM と ICT の有用性
3. データ間の因果性と改善値の導出
4. 提案手法
5. 数値実験並びに考察
6. おわりに

証拠に基づく政策立案のための オープンデータを利活用した WebGIS 可視化によるデータフュージョン

Data Fusion through Web-GIS Visualization
Using Open Data for Evidence-Based Policy Making

長瀬 永遠 (Towa Nagase)
t815048@st.pu-toyama.ac.jp

富山県立大学 工学部 電子・情報工学科 情報基盤工学講座

Teams, 14:05-14:20 Wednesday, February 16, 2022.

1.1 本研究の背景

近年、世界各国で証拠に基づく政策立案 (Evidence-Based Policy Making: EBPM) に注目が集まっている。日本においても例外ではなく、研究機関でも取り上げられている。政府を中心に推進されている EBPM であるが、地方自治体ではいまだにエビソードベースの意思決定が根強い。

エビデンスベース

政策によって改善したい対象を明確化したうえでデータを収集し意思決定。



エピソードベース

住民によって役場に持ち込まれた問題に 対して対面処理的に意思決定

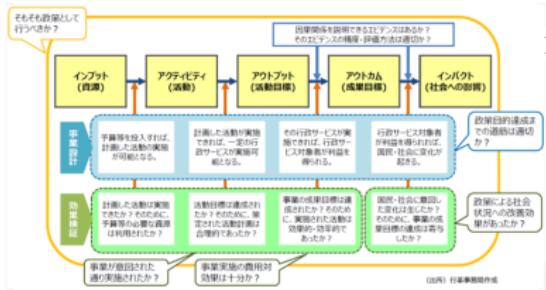


図1 EBPMのロジックモデル

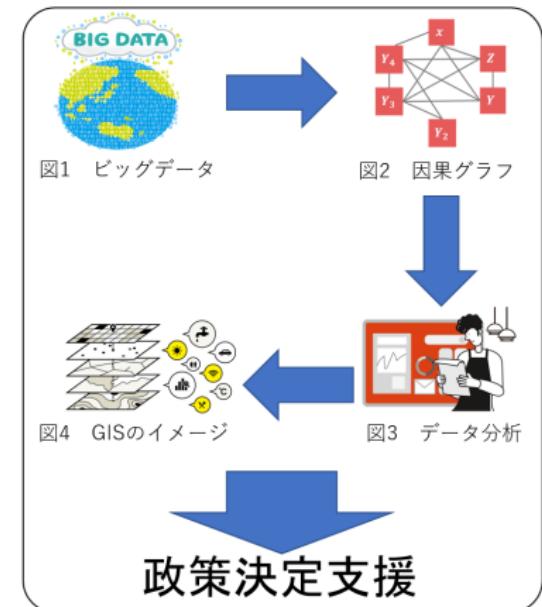
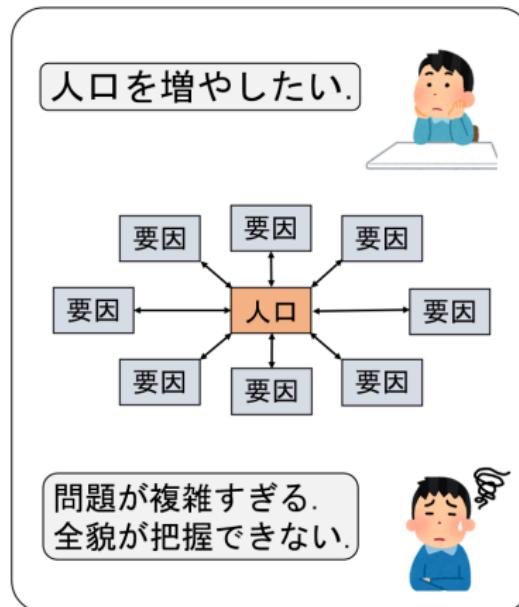
EBPMの利点

- ・データに基づくため予測を行うことができ、問題が顕在化する前に対策を打つことが可能。
 - ・その場限りのエピソードによるものではなく、明確な根拠があるため、住民の理解が得やすい。
 - ・政策による効果が事前に逆算できるため、状況に応じた微調整が可能。

1.2 本研究の目的

政策決定における対象の複雑性という課題に対して、多種多様なデータから対象と因果関係のあるもののみを抜き出して分析し、GISを用いて結果を提示することで政策決定の支援を行う手法を提案する。

- はじめに
- EBPM と ICT の有用性
- データ間の因果性と改善値の導出
- 提案手法
- 数値実験並びに考察
- おわりに



2.1 ICT とデータの利活用による行政施策

4/15

- はじめに
- EBPM と ICT の有用性
- データ間の因果性と改善値の導出
- 提案手法
- 数値実験並びに考察
- おわりに

EBPM を実行するためには膨大なデータの収集・分析が必要になる。EBPM を広範囲に的確かつ効率的に取り入れるには ICT の有効活用がもとめられる。このことから、内閣府を中心に地方自治体における EBPM を支援するためのシステムが複数提供されている。

EBPM推進の向けたシステム

- ・内閣エビデンスシステム
研究機関の運営評価に特化。研究機関における「研究力」、「教育力」、「資金獲得力」などを見える化。
- ・地域経済分析システム
経済に関するデータを中心に幅広いデータを扱う分析が可能。
- ・V-RESAS
COVID-19による経済の変化を分析可能。
- ・内閣府経済社会総合研究所
内閣府を支えるシンクタンク。政策研究を担う人材育成も行う。



図6 地域経済分析システム



図5 内閣エビデンスシステム



図7 V-RESAS

2.2 GIS による地域的な政策決定支援

1. はじめに

2. EBPM と ICT
の有用性

3. データ間の因果
性と改善値の導出

4. 提案手法

5. 数値実験並びに
考察

6. おわりに

GIS とは、地理空間データを総合的に管理・加工し、地理的位置とデータを結び付けて視覚化できる技術。近年、多くの自治体で運用されている。

GISの特徴

- ・データの可視化
統計データ等を2Dや3D、アニメーションなど多様な表現方法で地図上にプロット。
- ・データ間の関係性把握
複数データを同一画面に表示し比較。
- ・データの統合と分析
位置情報をキーとして異なる特徴を持つデータを統合。
- ・データの作成と更新
データ更新の負担を軽減し鮮度を保つ。

GISの利点

- ・業務効率化による情報管理のコスト削減
電子化により作業時間や人員を縮小
- ・最適な意思決定の促進
同一画面上に複数データを表示
- ・コミュニケーション性の向上
視覚的な情報共有や議論が可能

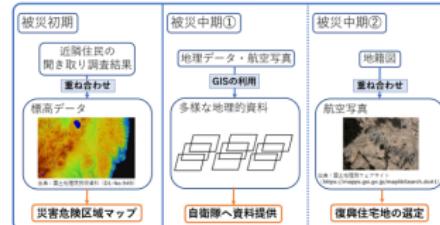


図8 被災下でのGIS



図9 地方自治体におけるGIS

2.3 Web-GIS によるデータフュージョン

6/15

- はじめに
- EBPM と ICT の有用性
- データ間の因果性と改善値の導出
- 提案手法
- 数値実験並びに考察
- おわりに

GIS は多様なデータに対して地理空間を介在することで同一プラットフォーム上に可視化できる。そのため、普段は別々に研究が行われている分野の研究結果を重ね合わせて新たな知見を得ることが可能。

都市計画と生態学

都市計画と生態学のそれぞれで作成されたマップを重ね合わせることで異分野の知見を融合。



図10 二次草原潜在生育地マップ



図11 都市化ボテンシャルマップ

二次草原が失われる可能性のある土地



図12 重ね合わせマップ

類似な特徴を持つ土地を分析

身体情報に対する応用

GISにおける位置情報をキーにして情報を重ね合わせるという特徴を人体モデルに応用

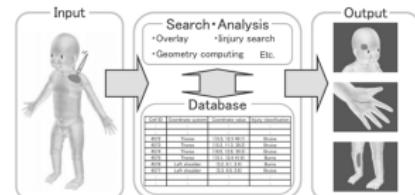


図13 身体地図情報システムの構成

オーバーレイ機能

レイヤ同士を重ね合わせることで重畳部分のみの新たなレイヤを作成。部位別の受傷頻度等を算出。

受傷面積算出機能

傷害として塗りつぶされた範囲を自動算出し、体表の何%に当たるかを計算。熱傷の重症度の測定等に使用。

ICD-10コード変換機能

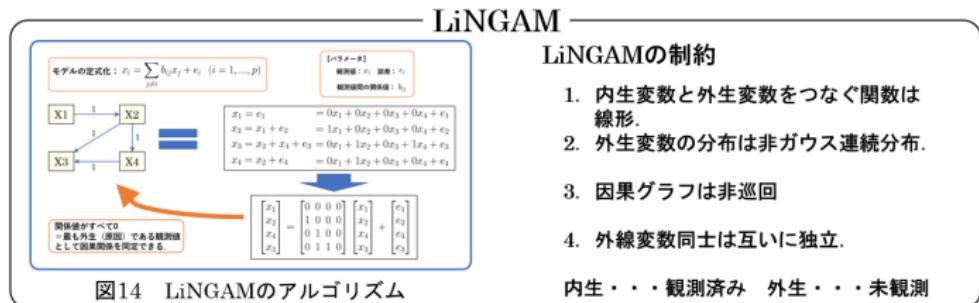
傷害情報と受傷位置それぞれの情報を持つレイヤを重ね合わせ、傷害情報と受傷位置をICD-10コードに自動変換。

3.1 因果探索によるデータ間の関係性

7/15

- はじめに
- EBPM と ICT の有用性
- データ間の因果性と改善値の導出
- 提案手法
- 数値実験並びに考察
- おわりに

因果探索とは、観測データを用いて、データ群の因果グラフ（複数の観測データにおいて、それぞれの値がお互いに及ぼしあっている影響の度合いを構造的に示したもの）を導出するための教師なし学習。



Direct-LiNGAM

Direct-LiNGAMのアプローチ
・・・回帰分析を用いる手法

- 内生変数群から2変数を取り出しそれらの変数間に成り立つ因果関係を同定することを繰り返して因果グラフの始点を探索。
- その変数を内生変数群から除外し、残った変数のみで内生変数群を再形成。

$$\begin{cases} x_1 = e_1 \\ x_2 = b_{21}x_1 + e_2 \end{cases} \quad \left\{ 1 - \frac{b_{21} \text{cov}(x_1, x_2)}{\text{var}(x_2)} \right\} e_1 - \frac{b_{21} \text{var}(x_1)}{\text{var}(x_2)} e_2$$

ダルモア・スキットピッチの定理

2つの確率変数 y_1, y_2 が互いに独立な確率変数 $s_i (i = 1, \dots, q)$ を用いて下記のように表されるとき、 y_1, y_2 が独立なら、 $\alpha_j, \beta_j \neq 0$ となるような変数 s_i はガウス分布に従う。

$$y_1 = \sum_{i=1}^q \alpha_i s_i \quad y_2 = \sum_{i=1}^q \beta_i s_i$$

3.2 DEA による効率値と入力・出力改善値の導出

8/15

- はじめに
- EBPM と ICT の有用性
- データ間の因果性と改善値の導出
- 提案手法
- 数値実験並びに考察
- おわりに

DEA とは、ある分野における組織の集合において、対象の組織の業績を評価するために生み出されたノンパラメトリックなアプローチ。組織とは、その活動においていくつかの種類の入力（投入）をいくつかの出力（産出）に変換することに携わる生産体（Decision Making Unit: DMU）。

<CCR モデルの主問題>

$$\begin{aligned} & \text{maximize} \quad \frac{u^T y_o}{v^T x_o} = z \\ & \text{subject to} \quad -v^T X + u^T Y \leq 0 \\ & \quad u \geq 0 \quad v \geq 0 \end{aligned}$$

対象の効率性 = 対象の仮想出力
対象の仮想入力
を可能な限り最大化する重みを定める。

他の DMU に對しても、効率値は 1 を超えないように制約。

重みが負の値をとらないように制約。

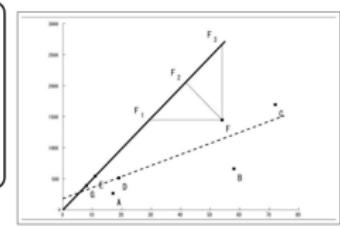


図15 1入力・1出力のDMU群

双対問題

<入力指向モデル>

$$\begin{aligned} & \text{minimize} \quad w = \theta \\ & \text{subject to} \quad Y\lambda \geq y_o \\ & \quad -X\lambda + x_o\theta \geq 0 \\ & \quad \lambda \geq 0 \end{aligned}$$

<出力指向モデル>

$$\begin{aligned} & \text{maximize} \quad w = \eta \\ & \text{subject to} \quad X\mu \leq x_o \\ & \quad -Y\mu + y_o\eta \leq 0 \\ & \quad \mu \geq 0 \end{aligned}$$

<入力改善案>

$$\hat{x}_i = \sum_{k=1}^K x_{ik} \lambda_k \quad i = 1, 2, \dots, m$$

<出力改善案>

$$\hat{y}_j = \sum_{k=1}^K y_{jk} \mu_k \quad j = 1, 2, \dots, n$$

それぞれの DMU に対して各入力をどれだけ減少、各出力をどれだけ増加させれば評価値が 1 になるかが算出できる。また、その際に参考とした DMU もわかる。

3.3 Folium を用いた Web-GIS の開発

9/15

1. はじめに
2. EBPM と ICT の有用性
3. データ間の因果性と改善値の導出
4. 提案手法
5. 数値実験並びに考察
6. おわりに

Web-GIS とは、html 形式で記述され、World Wide Web 上でアプリケーションとして動作する GIS の総称。また、Web-GIS における html は直接手書きで作成するほかプログラミング言語で自動的に生成も可能。Web-GIS を記述する Python のモジュールを Folium と呼ぶ。

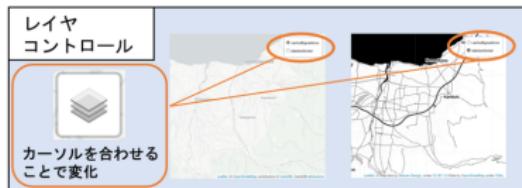


図16 レイヤコントロールの例

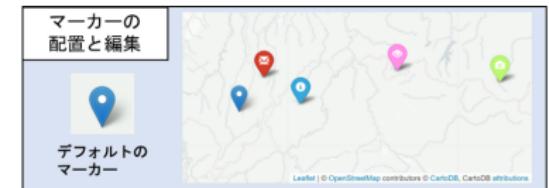


図17 各種マーカーの例

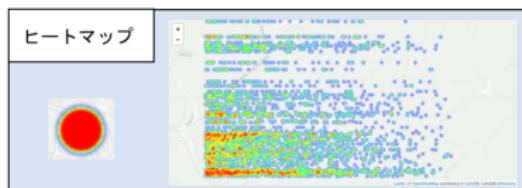


図18 ヒートマップの例



図19 マーカークラスタの例

4.1 データベース作成と因果探索によるデータの選定

10/15

1. はじめに
2. EBPM と ICT の有用性
3. データ間の因果性と改善値の導出
4. 提案手法
5. 数値実験並びに考察
6. おわりに

データベース内には地理情報を持たない数値データ、地理情報を持つ数値データ、施設等の位置データがある。データは RESAS API および国土交通省のウェブサイトを用いて収集した。それらを因果探索で絞りこみ、DEA の入力・出力に振り分ける。

表1 地理情報を持たない数値データ

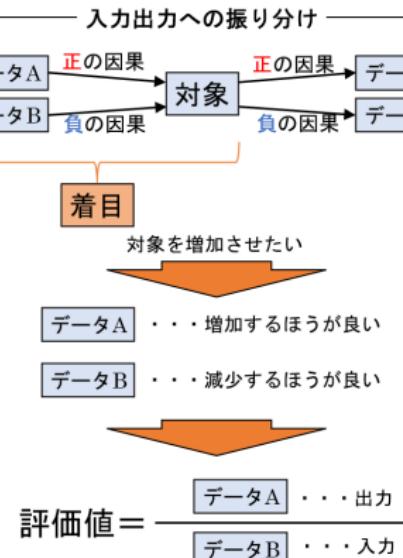
データ項目	単位	データ項目	単位
耕作放棄地率	%	経営耕地面積	1戸/経営体
農業生産額	千万円	労働生産性	なし
企業数	社	従業員数	人
歳出決算額 [総務費]	%	農地平均取引価格	円/m ²
歳出決算額 [民生費]	%	商業用地平均取引価格	円/m ²
歳出決算額 [生産費]	%	住宅用地平均取引価格	円/m ²
歳出決算額 [農林水産業費]	%	林地平均取引価格	円/m ²
歳出決算額 [商工費]	%	マンション等平均取引価格	円/m ²
歳出決算額 [土木費]	%	1人あたりの地方税	千円
歳出決算額 [警察費・消防費]	%	製造品出荷額	万円
歳出決算額 [教育費]	%	事業所数	事業所
歳出決算額 [公債費]	%	総人口	人
歳出決算額 [労務費]	%	老年人口	%
歳出決算額 [その他 (雑費)]	%	生産年齢	%
農業就業人口平均年齢	歳	年少人口	%
農業経営者平均年齢	歳	年間商品販売額	百万円
林作業請負収入	万円	海面漁獲物等販売額	万円
林産物販売金額	万円	付加価値額	万円
一人当たりの法人住民税	千円	1人あたりの固定資産税	千円

表2 地理情報を持つ数値データ

データ項目	単位
施設位置 [空港]	経度・緯度
施設位置 [工業団地]	経度・緯度
施設位置 [都市公園]	経度・緯度
施設位置 [道の駅]	経度・緯度
施設位置 [学校]	経度・緯度

表3 位置データ

データ項目	単位
施設数 [空港]	箇所
施設数 [工業団地]	箇所
施設数 [都市公園]	箇所
施設数 [道の駅]	箇所
施設数 [学校]	箇所



4.2 選定されたデータに基づく DEA 分析

11/15

- はじめに
- EBPM と ICT の有用性
- データ間の因果性と改善値の導出
- 提案手法
- 数値実験並びに考察
- おわりに

因果探索によって振り分けられた入力・出力を用いて DEA を行うことで評価値、入力・出力改善案、参照集合に属する市区町村とそれらにかかるウェイトを算出する。種類ごとの計 4 つの csv ファイルで結果を出力する。

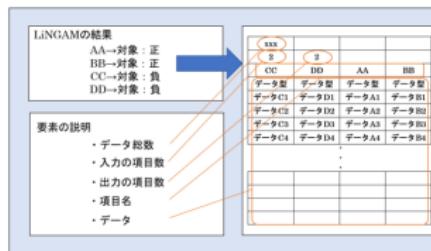


図20 DEAの入力・出力ファイル

都道府県	市	区	町	村	都道府県	市	区	町	村	都道府県	市	区	町	村
北海道	35	0	130	20	石川県	11	0	8	0	岡山県	14	4	10	2
青森県	10	0	22	8	福井県	9	0	8	0	広島県	13	8	0	9
岩手県	14	0	15	4	山梨県	13	0	8	6	山口県	13	0	0	6
宮城県	13	5	18	1	長野県	19	0	23	35	徳島県	8	0	15	1
秋田県	13	0	9	3	岐阜県	21	0	19	2	香川県	8	0	9	0
山形県	13	0	19	3	静岡県	21	10	12	0	愛媛県	11	0	9	0
福島県	13	0	31	15	愛知県	37	16	14	2	高知県	11	0	17	6
茨城県	32	0	10	2	三重県	14	0	15	0	福岡県	27	14	29	2
栃木県	14	0	11	0	滋賀県	13	0	6	0	福井県	10	0	10	0
群馬県	12	0	15	8	京都府	14	11	10	1	長崎県	13	0	8	0
埼玉県	39	10	22	1	大阪府	31	31	9	1	熊本県	13	5	23	8
千葉県	36	6	16	1	兵庫県	28	9	12	0	大分県	14	0	3	1
東京都	23	23	5	8	奈良県	12	0	15	1	宮崎県	9	0	14	3
神奈川県	16	28	13	1	和歌山县	9	0	20	1	鹿児島県	19	0	20	4
新潟県	19	8	6	4	鳥取県	4	0	14	1	沖縄県	11	0	11	19
富山県	10	0	4	1	島根県	8	0	10	1	合計	731	188	277	203

図21 各都道府県におけるDMUの内訳

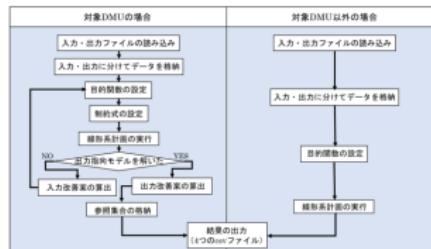


図22 DEA部分のフロー

評価値	入力・出力改善案	参照集合・重み（入力）	参照集合・重み（出力）
city_code,effic_value	imp_item,imp_value	city_code,weight_in	city_code,weight_out
1100	0.55679	AAA, 0	24303, 0.13164
1202	0.70558	BBB, 13.244	33663, 0.3689
1203	0.83029	CCC, 0.86798	34203, 0.02617
	.	DDD, 5.68598	34203, 0.02617
	.	EEE, 3.69084	39424, 0.05159
	.	FFF, 26.6893	44462, 0.69311
47381	0.30172	GGG, 6.20473	
47382	0.35638	HHH, 42026.3	

図23 DEA部分のアウトプット

4.3 Web-GIS を用いたデータフュージョンのシステム開発

12/15

1. はじめに
2. EBPM と ICT
の有用性
3. データ間の因果
性と改善値の導出
4. 提案手法
5. 数値実験並びに
考察
6. おわりに

動画

5.1 数値実験の概要

富山県射水市における少子高齢化問題について、その解決に向けた政策を行うというモデルケースに提案手法を適用することによって本研究の有効性を検証・考察する。

1. はじめに

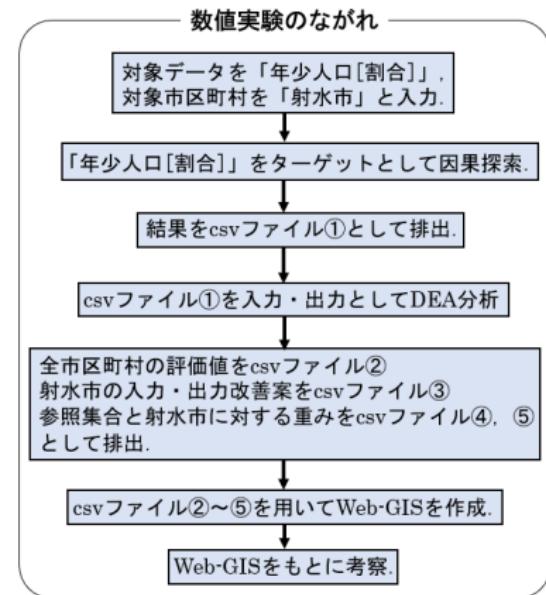
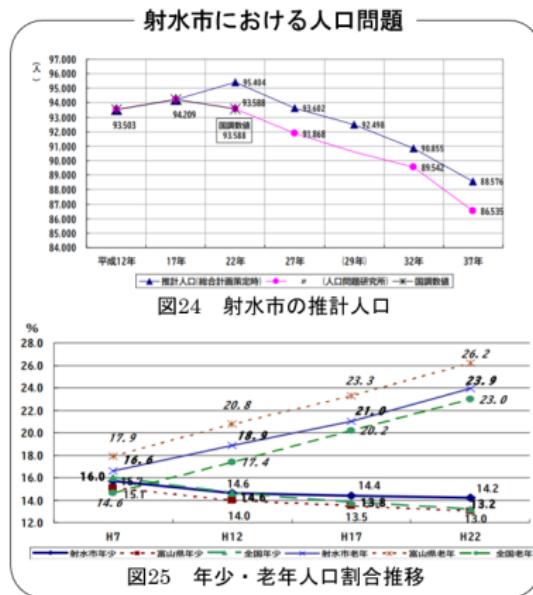
2. EBPM と ICT
の有用性

3. データ間の因果
性と改善値の導出

4. 提案手法

5. 数値実験並びに
考察

6. おわりに



5.2 数値実験の結果と考察

システムの各部分における結果を示す。

- はじめに
- EBPM と ICT の有用性
- データ間の因果性と改善値の導出
- 提案手法
- 数値実験並びに考察
- おわりに

表5 「年少人口[割合]」に対するDirect-LiNGAMの結果

データ項目	バス係数	データ項目	バス係数
施設数[空港]	0.350	民生費	0.027
農林水産業費	0.023	土木費	0.008
警察・消防費	-0.031	教育費	0.036
住宅用土地平均取引価格	-1.454×10^{-6}	マンション等平均取引価格	-2.236^{-6}
老年人口	-0.057	生産年齢人口	0.217

表6 DEAにおける入力・出力

入力	出力
警察費・消防費	施設数[空港]
住宅用土地平均取引価格	民生費
マンション等平均取引価格	農林水産業費
老年人口	土木費
	教育費
	生産年齢人口

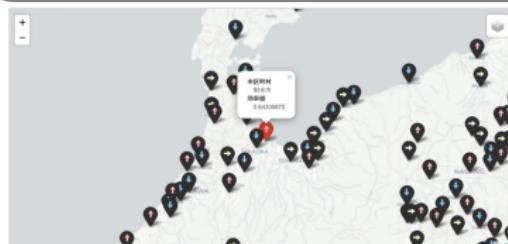


図25 EBPM-GISの結果

表7 元データと改善案の比較

	元データ	改善案
生産年齢人口	57.5	117.4422884327
民生費	31.67	64.6851420806
農林水産業費	1.87	14.934071761810001
土木費	12.25	40.47816951629995
警察費・消防費	2.88	1.41005485674
教育費	12.5	5.530992139156
住宅用土地平均取引価格	19507	5681.56383792
マンション等平均取引価格	0	0.0
老年人口	28.6	14.0026281516
施設数[空港]	0	0.73932497

表8 参照集合に属する市区町村とウェイト

参照市区町村	入力指向モデル		出力指向モデル	
	ウェイト	参照市区町村	ウェイト	参照市区町村
越前町	0.036873746	越前町	0.075313657	
忍野村	0.36197527	忍野村	0.73932497	
東近江市	0.61557443	東近江市	1.2572946	
三好郡東みよし町	0.44934957	三好郡東みよし町	0.91778469	

1. はじめに
2. EBPM と ICT の有用性
3. データ間の因果性と改善値の導出
4. 提案手法
5. 数値実験並びに考察
6. おわりに

今後の課題

本研究では、市区町村単位の自治体におけるEBPMの支援を目的として、政策に関する意思決定のためのデータ分析および結果の提示を行う手法を提案した。

- データベース内の情報量の充実
 - • • 扱える問題の汎用性と分析結果の確からしさ両方の向上につながる。
- 因果探索における手法の洗練
 - • • 因果グラフにおけるパス係数をどの程度まで有意とするか検討。
Direct-LiNGAMに限らず、より効果的な因果探索手法を模索。
- DEAにおけるモデルの深化
 - • • 本研究のテーマによりフィットしたモデルを模索・提案。

今後の展望

本研究では、政策における意思決定を用いて本研究を提案したが、我々が生活する現代社会では要因が複雑に絡み合った問題に遭遇する機会が多くあるため、分析の際に全容が確定していないデータ群に対して幅広く用いることができるよう提案手法の汎用性を広げることも考えられる。

Temporary page!

\LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it. If you rerun the document (without altering it) this surplus page will go away, because \LaTeX now knows how many pages to expect for this document.