

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

長瀬 永遠

富山県立大学 情報基盤工学講座

January 28, 2022

# 1.1 本研究の背景

2/23

## 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ピソードベースの意思決定が根強い．

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

現状報告

## 1.2 本研究の目的

3/23

### 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

現状報告

## 2.1 ICT とデータの利活用による行政施策

4/23

### ICT を用いた EBPM 推進の例

EBPM の考え方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。

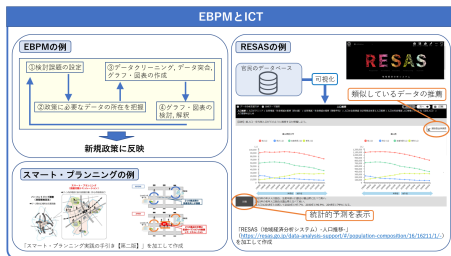


図 1: EBPM と ICT

## 2.2 GIS による地域的な政策決定支援

5/23

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。

### GIS を使用することの 3 つの利点

- 業務効率化によるコスト削減
- 最適な意思決定の促進
- コミュニケーション性の向上

### GIS を用いた政策決定の実例

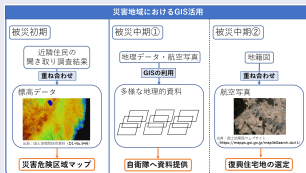


図 2: 福島県相馬市の例

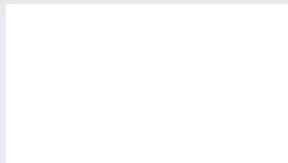


図 3: 奈良県橿原市の例

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

現状報告

## 2.3 WebGIS によるデータフュージョン

6/23

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。

### 複数分野のデータフュージョンによる新知見の発見

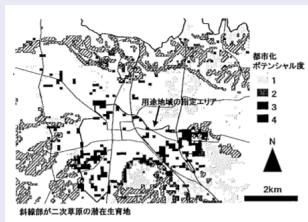


図 4: 生態学と都市計画分野の重ね合わせ

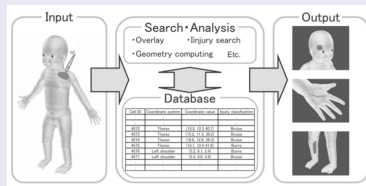


図 5: 身体情報に対する GIS の応用

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

現状報告

## 3.1.1 因果探索によるデータ間の関係性1

7/23

因果探索とは、観測データを用いて、その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。

### LiNGAM

因果探索における手法の一つ。

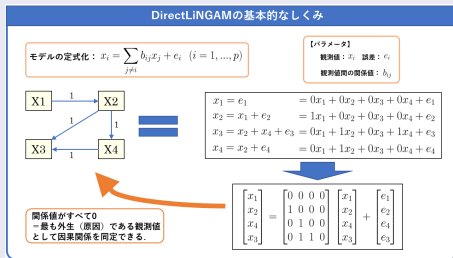


図 6: LiNGAM のアルゴリズム

## 3.1.2 因果探索によるデータ間の関係性2

8/23

### Direct-LiNGAM

- 観測変数群から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観測変数群全体における因果の始まりとなる変数を探す。
- その変数を観測変数群から除外し、残った変数のみで再度、観測変数群を形成する。

という2つの操作を観測変数群に属する変数が存在しなくなるまで繰り返すことによって元の観測変数群の因果グラフを同定

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

現状報告



## 3.2.1 DEA による効率値と入力・出力改善値の導出 1

9/23

### DEA

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データ指向のアプローチ。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）。DEA での分析の利点の一つとして、複数の入力・出力があるデータを扱うことができることが挙げられる。

### CCR モデル

<CCR モデルの主問題>

$$\text{Objective} \quad \frac{u^T y_o}{v^T x_o} = z \rightarrow \max \quad (1)$$

$$\text{Constrained} \quad -v^T X + u^T Y \leq 0 \quad (2)$$

$$u \geq 0 \quad (3)$$

$$v \geq 0 \quad (4)$$

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

現状報告

## 入力指向モデル・出力指向モデル（CCR の双対問題）

### < 入力指向モデル >

$$\text{Objective} \quad w = \theta \rightarrow \min \quad (5)$$

$$\text{Constrained} \quad Y\lambda \geq y_o \quad (6)$$

$$-X\lambda + x_o\theta \geq 0 \quad (7)$$

$$\lambda \geq 0 \quad (8)$$

### < 出力指向モデル >

$$\text{Objective} \quad w = \eta \rightarrow \max \quad (9)$$

$$\text{Constrained} \quad X\mu \leq x_o \quad (10)$$

$$-Y\mu + y_o\eta \leq 0 \quad (11)$$

$$\mu \geq 0 \quad (12)$$

## 3.2.3 DEA による効率値と入力・出力改善値の導出 3

11/23

### 参照集合

入力指向モデルおよび出力指向モデルを用いて効率値を求めた際に  $\lambda$  が値を持つ DMU の集合を参照集合と呼び、これは効率的な DMU のみを集めたものになる。参照集合内の DMU における入力・出力を用いて対象の DMU の入力・出力に対する改善案が算出可能。

### 入力・出力改善案

< 入力改善案 >

$$\hat{x}_i = \sum_{k=1}^K x_{ik} \lambda_k \quad (i = 1, 2, \dots, m) \quad (13)$$

< 出力改善案 >

$$\hat{y}_j = \sum_{k=1}^K y_{jk} \mu_k \quad (j = 1, 2, \dots, n) \quad (14)$$

## 3.3 Folium を用いた WebGIS の開発

12/23

### Web-GIS と Folium

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。

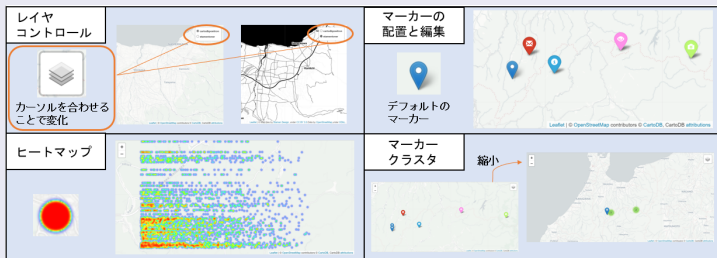


図 7: Folium による Web-GIS 実装例

## 4.1 統計データのデータベース作成と因果探索によるデータの選定

13/23

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

現状報告

## 4.2.1 選定されたデータに基づく DEA 分析 1

14/23

### DEA による分析

選定された入力・出力データを用いて DEA による評価値，参照集合，入力・出力の改善案を算出．

< 評価値の算出式 >

$$\text{Objective} \quad \theta \rightarrow \min \quad (15)$$

$$\text{Constrained} \quad \sum_{d=1} y_{id} \lambda_d \geq y_{io} \quad (i = 1, 2, \dots, m) \quad (16)$$

$$- \sum_{d=1} x_{id} \lambda_d + x_{io} \theta \geq 0 \quad (i = 1, 2, \dots, m) \quad (17)$$

$$\lambda \geq 0 \quad (18)$$

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

現状報告

## 4.2.2 選定されたデータに基づく DEA 分析 2

15/23

### DEA のシステムにおける流れ

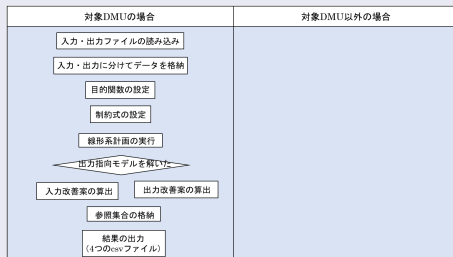


図 8: DEA 部分のフロー

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

現状報告

## 4.3 WebGIS を用いたデータフュージョンのシステム開発

16/23

### EBPM-GIS

以上の結果を Web-GIS 上に表示, 地理情報データとの重ね合わせることでデータフュージョンを行う EBPM-GIS を実装. DEA による評価値はマーカー, 改善案および参照集合はポップアップ, 地理情報はマーカークラスタによって表示.

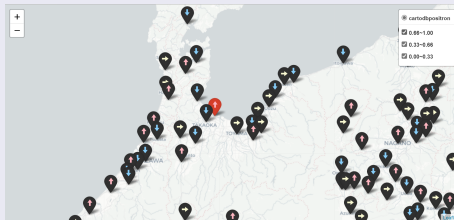


図 9: EBPM-GIS

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

現状報告



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

現状報告

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

現状報告



## 因果探索

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

現状報告

就業者数	人口	事業所数	地価_住宅	地価_商業	乱数_1	乱数_2	乱数_3
209403	418686	6154	41924	96000	1280.693	3222.395	133870
85832	172125	2354	32246	74666	925.9376	4843.729	162275.7
21940	42935	505	36000	65700	1489.198	7234.628	194195.5
23738	47992	318	33025	49900	723.5999	6674.72	93591.08
17065	32755	172	26233	40000	917.6694	6798.834	178289.7
20811	40991	327	26500	52750	1342.32	8397.706	110471.9
25712	49000	451	31700	42633	1830.67	341.4045	151687.2
15812	30399	269	22166	50600	1365.636	4760.153	177582.3
27029	51327	572	24116	41200	517.4427	2769.082	105788
46605	92308	1194	29228	44450	862.4601	1425.979	121991.3
1586	2982	10	19500	26200	756.4645	1022.445	197765.4
10496	20930	151	18900	30200	1286.429	6111.826	104155.9
13700	26317	100	17700	34500	1135.203	8172.202	67857.17
13099	25335	123	21450	35300	1727.469	8260.787	48688.03

15	
1	1
事業所数	地価_商業
float	float
6154	96000
2354	74666
505	65700
318	49900
172	40000
327	52750
451	42633
269	50600
572	41200
1194	44450
10	26200
151	30200

図 11: LiNGAM インプット

図 12: LiNGAM アウトプット

## DEA

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

現状報告

	7			
	2	2		
医師数	平均入院数	一日平均床	医療収益	
float	float	float	float	
17	15.5	226	2.3	
58	29.1	661	5.6	
72	15.8	1695	9.79	
19	20.4	514	2.68	
11	19.2	543	2.21	
54	13.1	1447	11.05	
8	24.9	390	1.82	

図 13: DEA インプット

効率値

city_code	effic_value
1	0.644373
2	0.467123
3	0.971215
4	0.694685
5	1
6	1
7	1

改善案

imp_item	imp_value
医師数	13.19901
平均入院数	14.17157
一日平均床	739.9039
医療収益	3.857865

参考市区町村  
(入力指向)

city_code	weight_in
5	0.627373
6	0.112768
7	0.026054

参考市区町村  
(出力指向)

city_code	weight_out
5	0.903104
6	0.16233
7	0.037505

図 14: DEA アウトプット

## Folium

実際に見せます

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

現状報告

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

現状報告

## 今後の課題

- GIS の施設位置レイヤの作成・数値実験（今週末まで）
- 概要・本論の完成（来週月曜朝まで）
- 発表資料・デモ動画の作成