

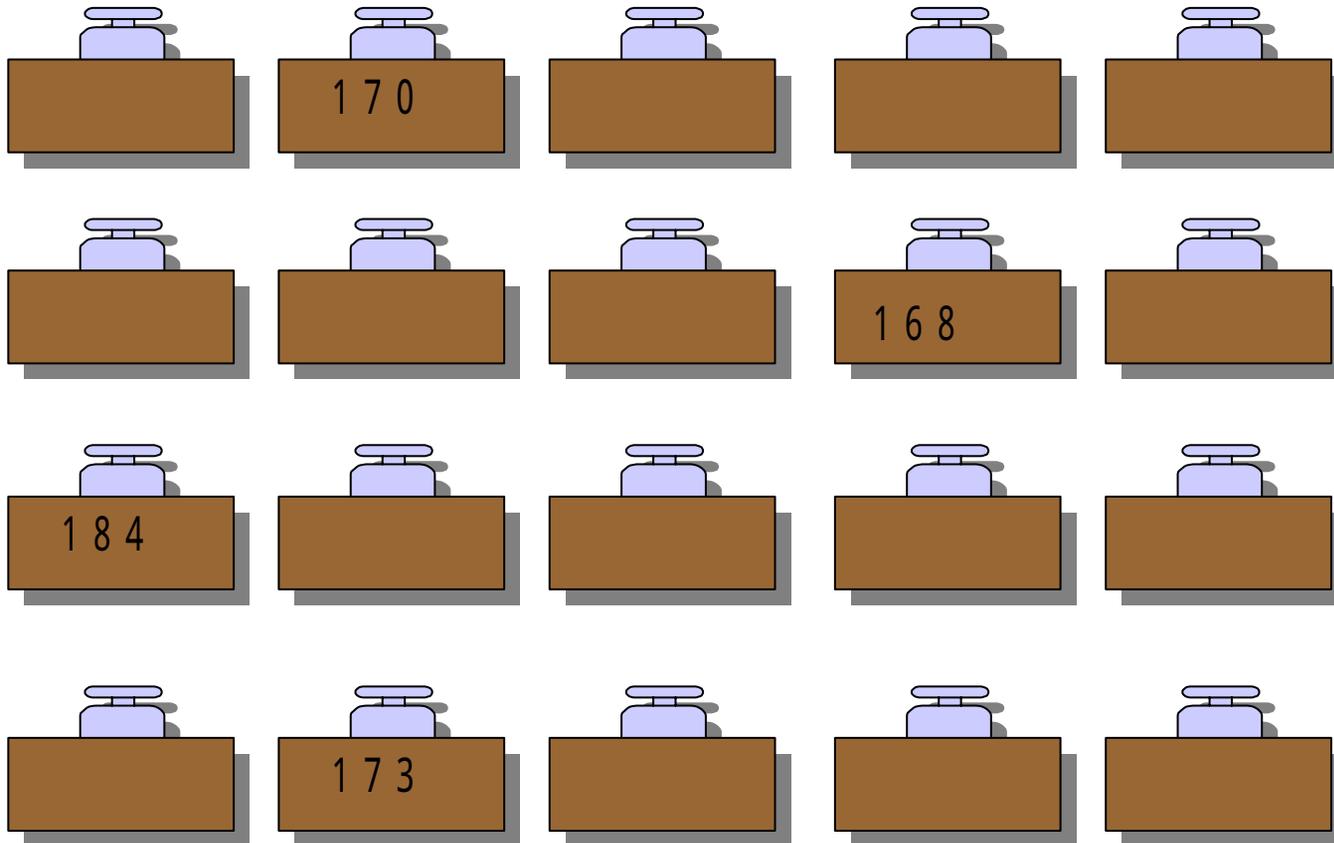
空間データの可視化ツール としてのS-PLUS

岡山大学
垂水共之

目次

- 空間データとは
- Latticeデータ
 - 地域メッシュナビゲータ
 - SAMAS
 - SVGを用いた町丁字等別小地域データ分析結果の可視化システム

データ(標本)



空間データとは

- どこで採った値
 - 位置情報
吉井川、旭川、
高梁川
- どんな値
 - 観測値



空間データの種類

- GeoStatisticalデータ (地理統計データ)
測定地点が決まっている
- Latticeデータ (格子データ)
隣接関係があるもの
- Spatial Point Patternデータ
(空間点パターン)
発生地点が主要な目的のもの

Latticeデータ

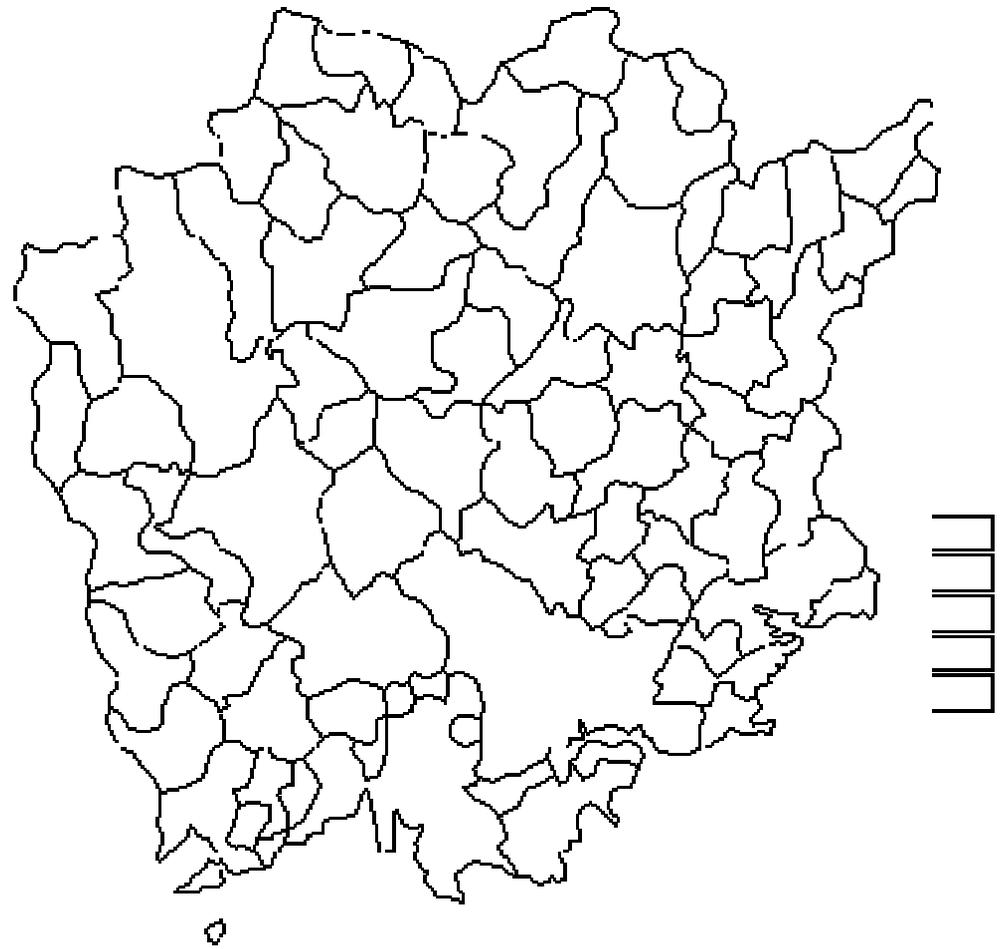
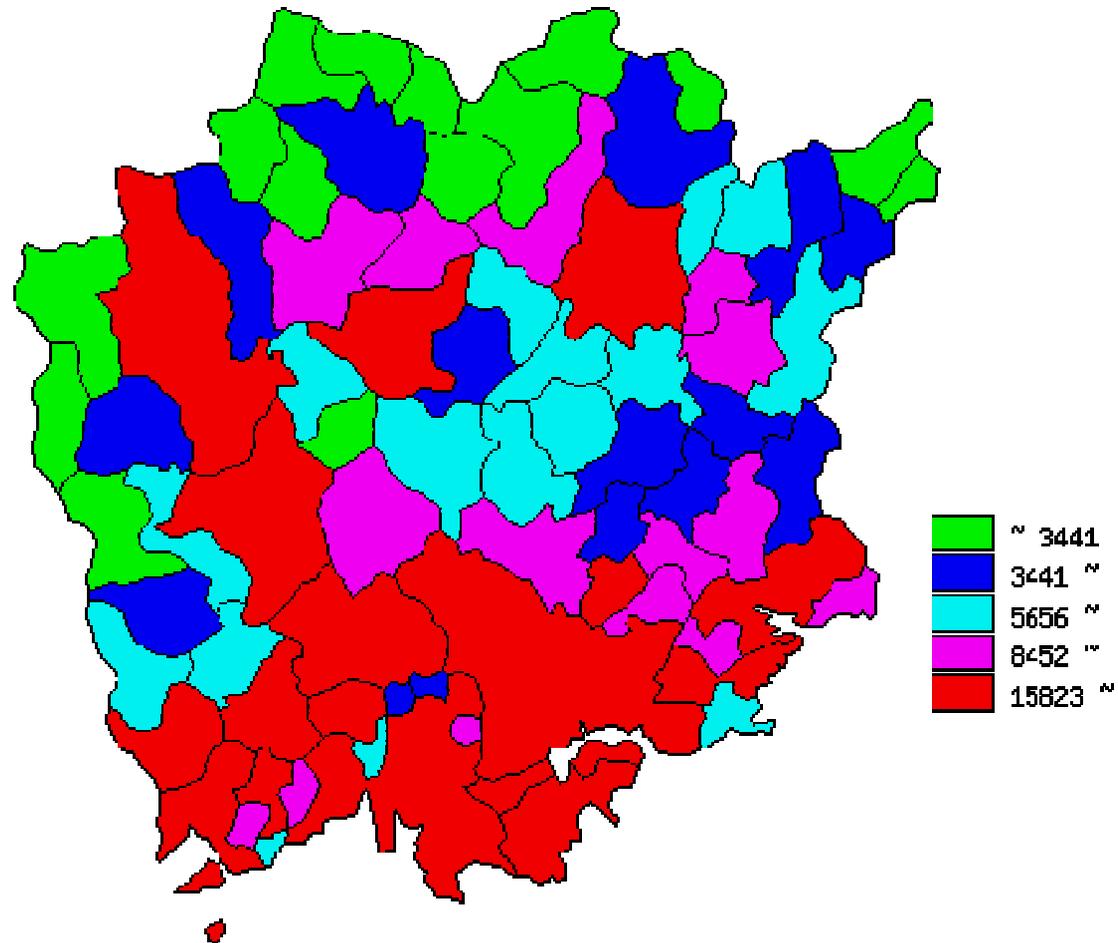


表1 市町村別、男女別人口、世帯数、1世帯当たり人員、人口密度及び面積

市町村別	平成12年10月1日						
	人口総数	男	女	総世帯数	1世帯 当たり 人員	人口密度	面積
岡山県計	1,950,828	936,044	1,014,784	691,620	2.82	274.3	7,112.1
岡山市	626,642	302,339	324,303	244,010	2.57	1,220.9	513.3
倉敷市	430,291	207,626	222,665	152,510	2.82	1,439.5	298.9
津山市	90,156	43,082	47,074	32,435	2.78	485.4	185.7
玉野市	69,567	33,314	36,253	24,680	2.82	671.8	103.6
笠岡市	59,300	27,894	31,406	19,834	2.99	436.1	136.0
井原市	34,817	16,681	18,136	11,024	3.16	387.1	90.0
総社市	56,531	27,125	29,406	18,909	2.99	294.0	192.3
高梁市	25,374	12,160	13,214	9,983	2.54	110.8	229.0
新見市	24,576	11,643	12,933	8,494	2.89	69.8	352.0
備前市	28,683	13,695	14,988	9,815	2.92	214.6	133.7
御津郡	23,402	11,054	12,348	7,590	3.08	67.8	345.1
御津町	10,214	4,753	5,461	3,244	3.15	89.3	114.4
建部町	6,989	3,276	3,713	2,330	3.00	78.1	89.5

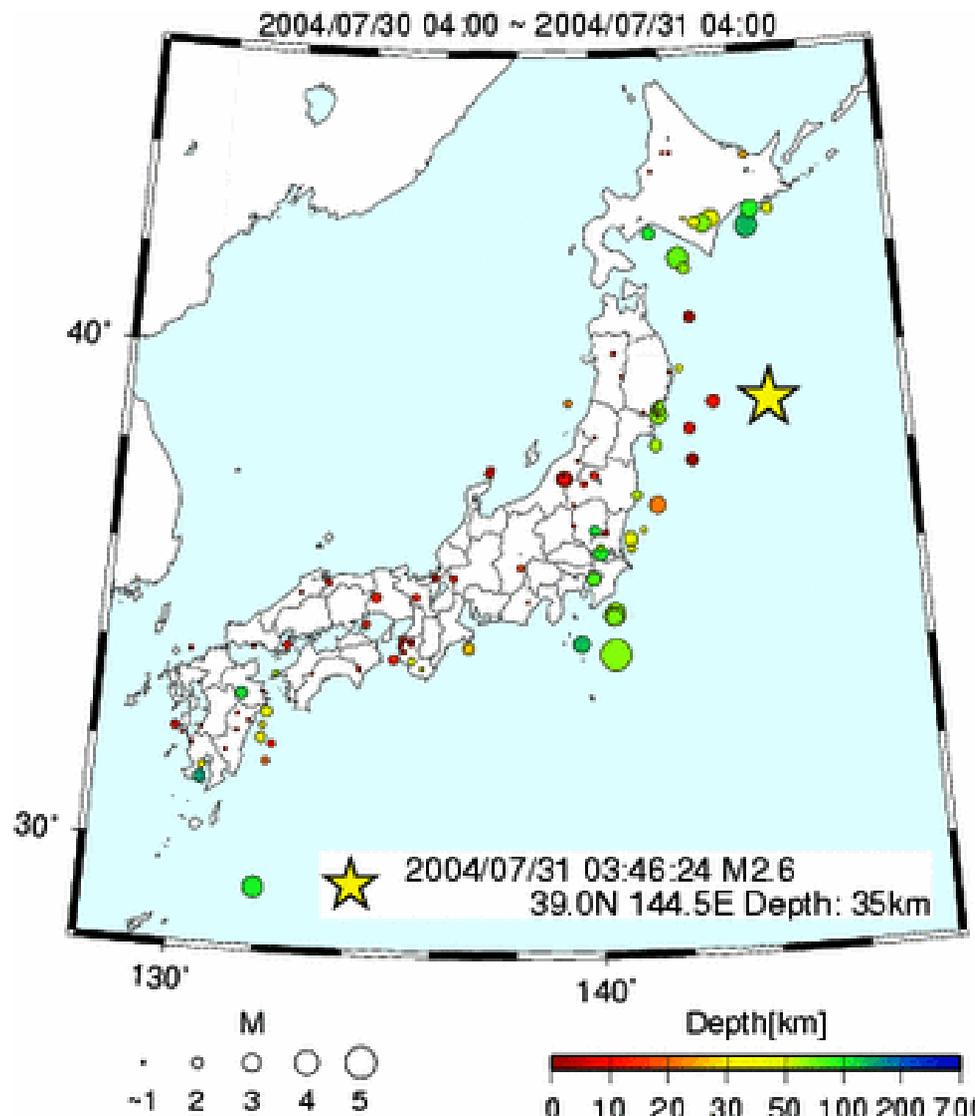
Latticeデータ

2000年国調人口



Spatial Point Pattern(空間点パターン)

2004/07/30 04:00~2004/07/31 04:00 震源データ



空間データ分析

- クリギング

- 空間現象を連続空間確率場でモデル化
- 空間内部の平均もしくはは特定の位置における確率場の値を観測されたデータから予測
- 通常型クリギング

- 空間領域 D の任意の位置 \mathbf{x} における確率場が2次定常

$$E[Z(\mathbf{x} + \mathbf{h})] = E[Z(\mathbf{x})]$$

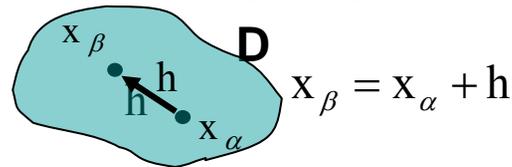
$$\text{cov}[Z(\mathbf{x} + \mathbf{h}), Z(\mathbf{x})] = C(\mathbf{h})$$

- 領域 D のある点 $Z(x_0)$ における値

$$Z^*(\mathbf{x}_0) = \sum_{i=1}^n \omega_{\alpha} Z(\mathbf{x}_{\alpha}) \quad \left(\sum_{\alpha=1}^n \omega_{\alpha} = 1 \right)$$

離間距離に対する非類似度

- 地理的空間における2点 x_α , x_β をベクトル $h = x_\alpha - x_\beta$ によって結びつけることができる。
- 非類似度 γ^* は、ベクトル によって記述される点対の離間距離と方向に依存する。



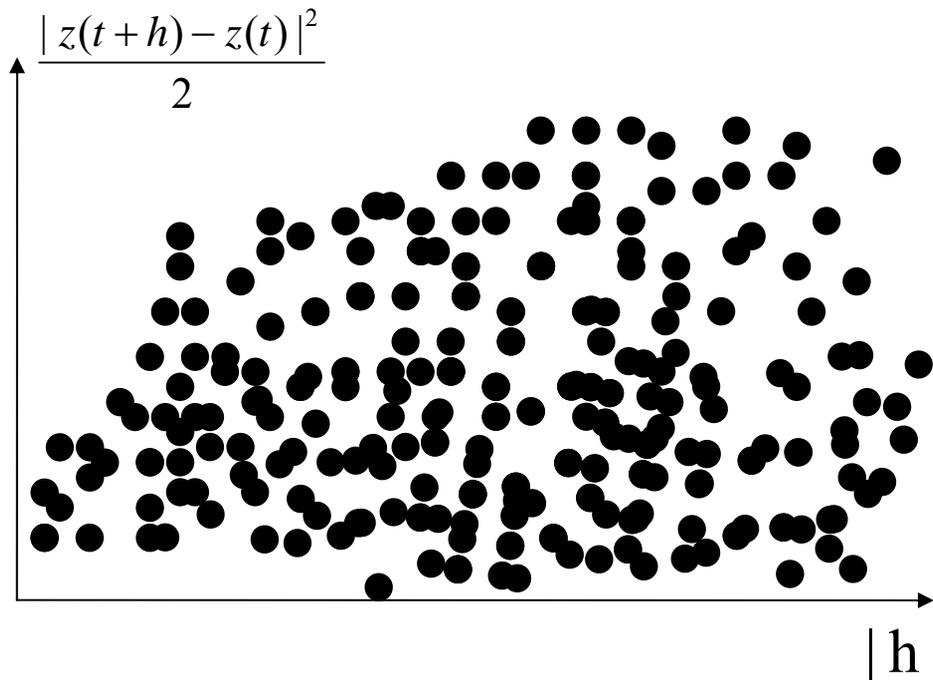
$$\gamma^*(h) = \frac{1}{2} (z(x_\alpha + h) - z(x_\alpha))^2$$

- 非類似度はベクトル h に関して対称。

$$\gamma^*(-h) = \gamma^*(+h)$$

バリオグラム雲(variogram cloud)

- 標本対の空間的離間距離 h に対する非類似度 γ^* を plot したものの

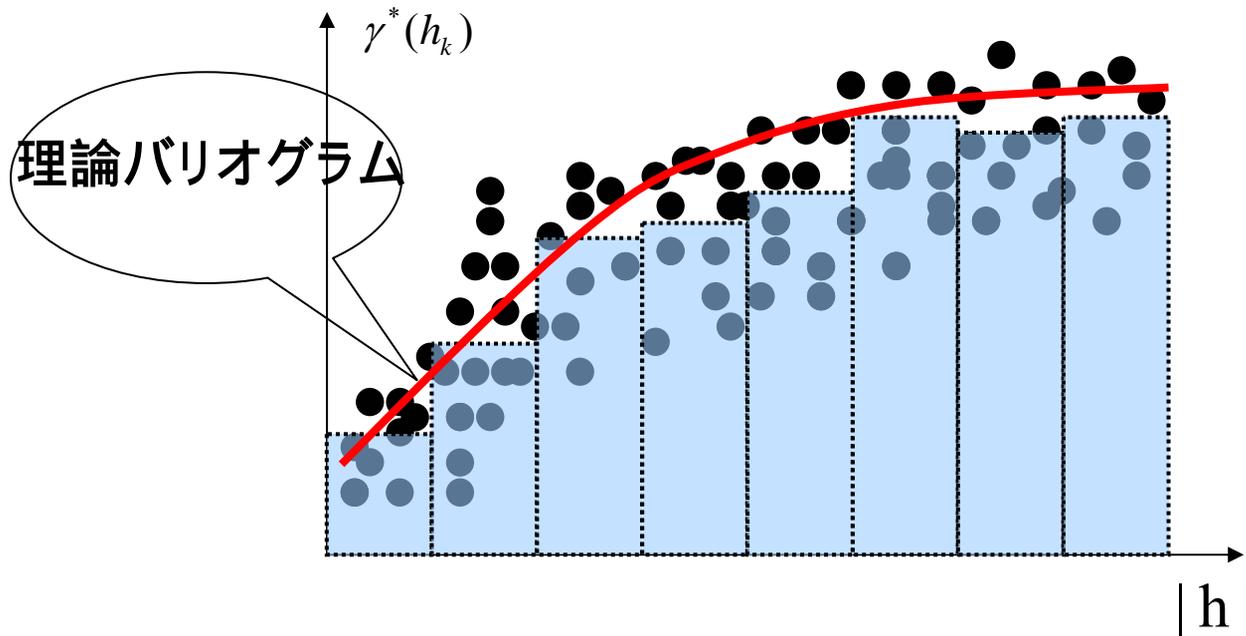


- 近接した標本間に類似した傾向があるほど、標本間の距離の増加に伴って非類似度が増加していく

標本バリオグラム(experimental variogram)

- ベクトルの階級 h_k に関する非類似度の平均値 $\gamma^*(h)$

$$\gamma^*(h_k) = \frac{1}{2n_c} \sum_{\alpha=1}^{n_c} (z(\mathbf{x}_\alpha + \mathbf{h}) - z(\mathbf{x}_\alpha))^2 \quad (\mathbf{h} \in h_k)$$



● 階級 h_k を設定することで指定された範囲内の長さ、与えられた許容角度内の方向をもつベクトルを抽出してグループ化することができる。

S-PLUSでのツール

- Spatialライブラリ
 - W.N.Venables and B.D.Ripley(1999)の第14章
- S+SpatialStats アドオンモジュール
- Environmental Stats for S-PLUS アドオンモジュール

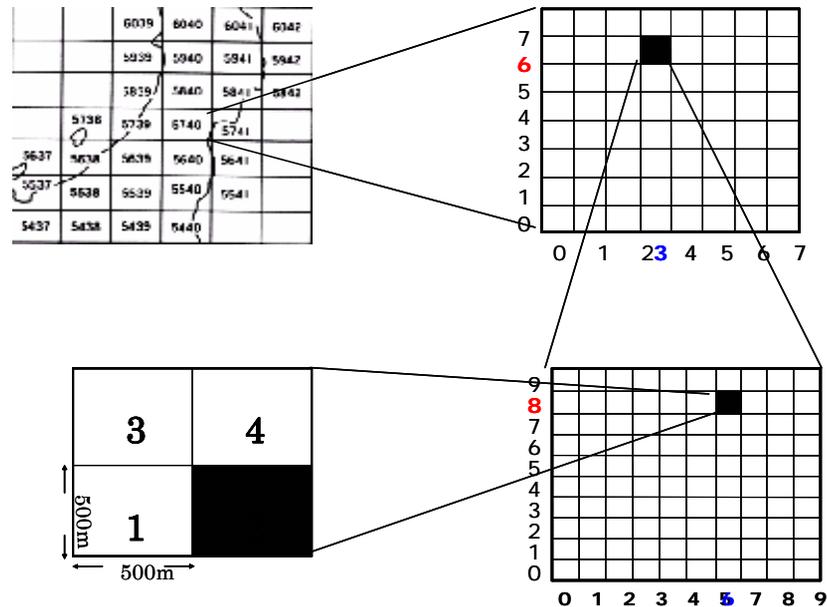
Latticeデータ

- 都道府県別データ
- 市区町村別データ
- 小地域データ
 - メッシュデータ ← 牛
 - 町丁字等別データ ← 亀川

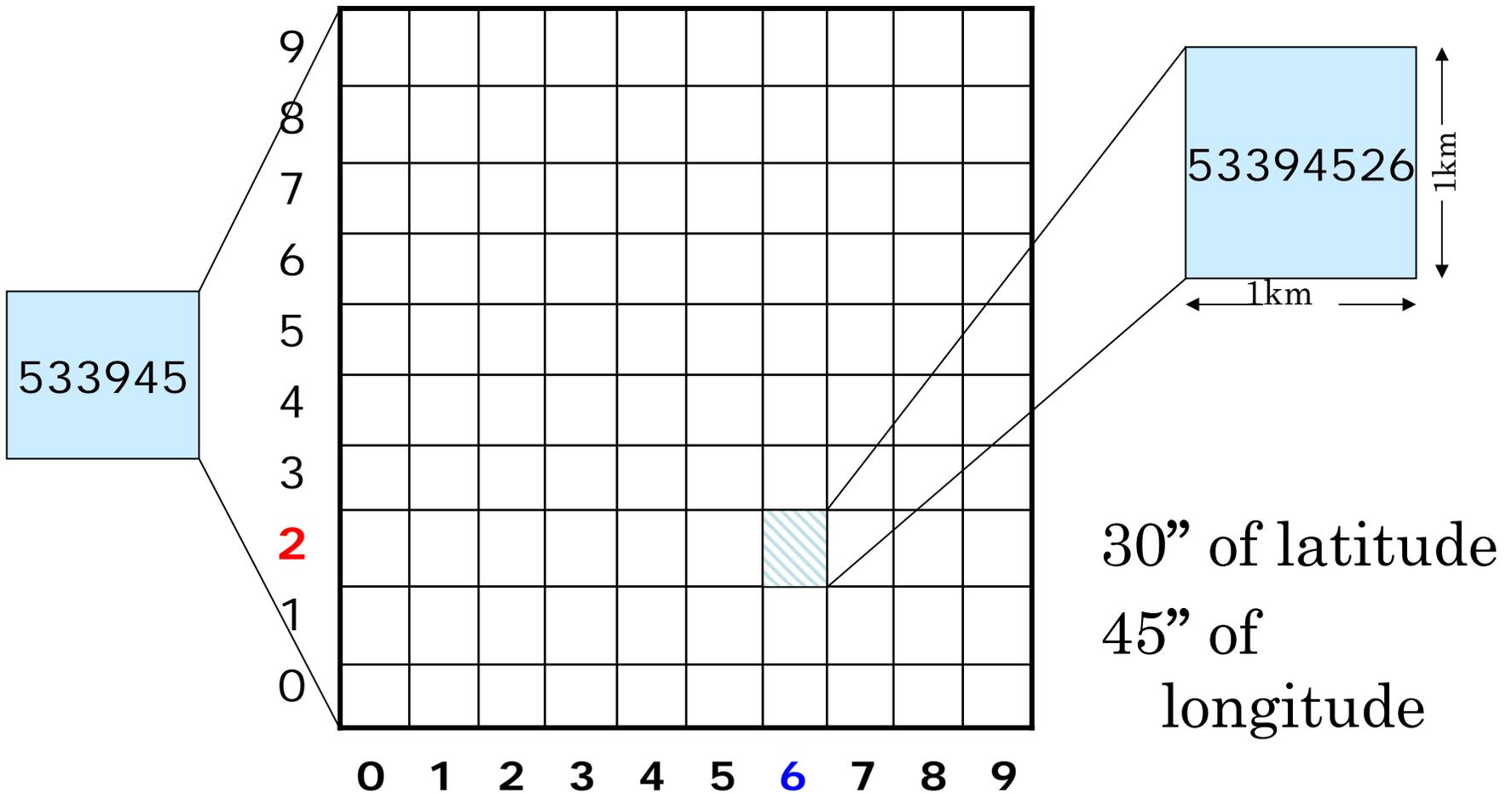
メッシュデータ

Area Mesh Statistics

- One of the small area statistics
- The whole area of Japan is divided into small square area like a mesh
- Based on latitudinal and longitudinal line.

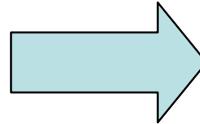


Third Area Mesh (Basic Area Mesh)



The Advantages of Area Mesh Statistics

- The same size and shape of partition
- The location and partition is fixed
- Partitioned by the longitude and latitude squarely

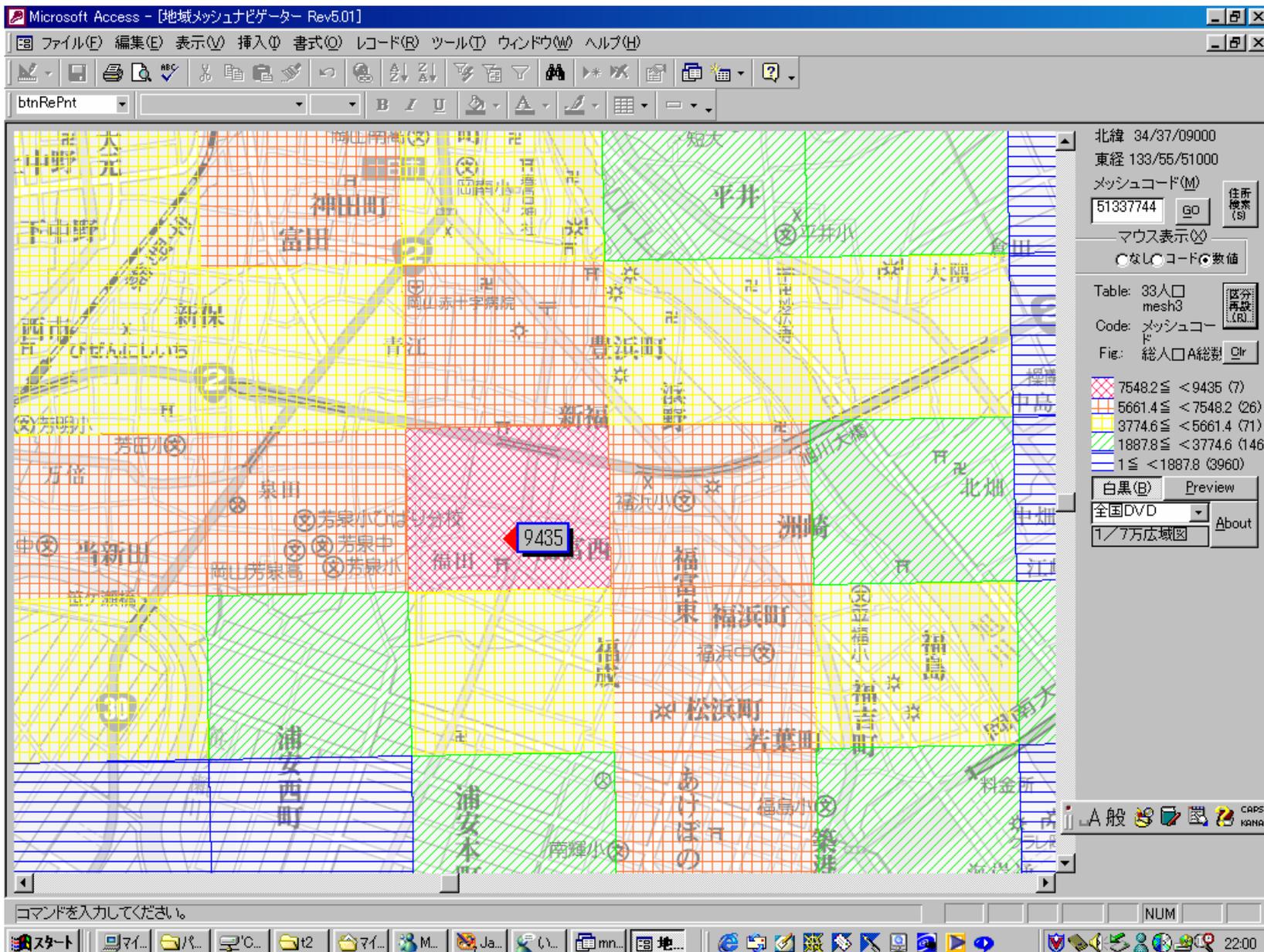


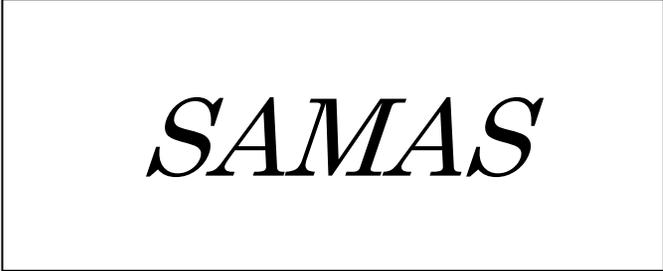
- Measure and compare inter-Area-Mesh characteristics
- Compare in time-series without being affected by the change of local districts
- Identify the location, analyze the data easily

地域メッシュナビゲーター

- メッシュデータの地図上への表示
- Microsoft Access
+
ProAtlas2000/2001/2002/
- <http://www.sinfonica.or.jp/>
- 榎田 直木 (1998)
[「地域メッシュナビゲーター」](#)

メッシュ統計(1995年国勢調査人口)





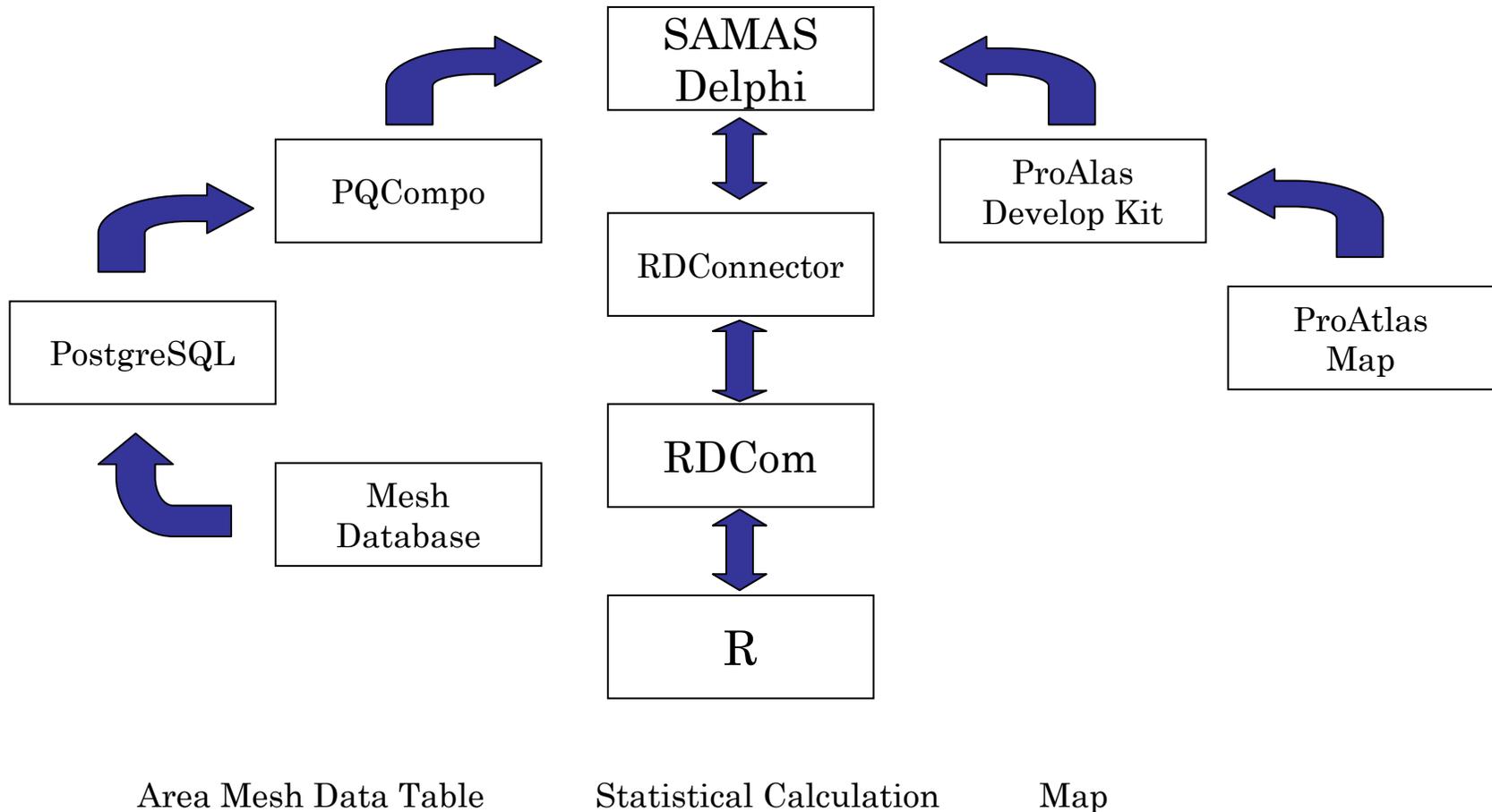
SAMAS

Statistical Area Mesh Analysis System

The Features of SAMAS

- Features of statistical analysis
 - Predict the geostatistical data
 - Calculate Echelon numbers, & etc.
- Features of displaying
 - Display the Data (Mosaic Plot) on the Map (Inherited from software Mesh Navigator)
 - Display the Result of Echelon Analysis (Echelon numbers of each mesh data)

System Figure of SAMAS



Demo of SAMAS

- Kriging the observed geostatistical data (Modified SO₂ data of Okayama Prefecture in 2003, 52 points) to mesh data (aim locations are meshcode selected from mesh database)



- Show the mesh data on the map

Variogram Form

Initial Values

Model's Number

Model	Sill	Range	Nugget
<input type="text" value="Spherical"/>	<input type="text" value="4.94117891"/>	<input type="text" value="0.122011"/>	<input type="text" value="0"/>

Estimate of Variogram

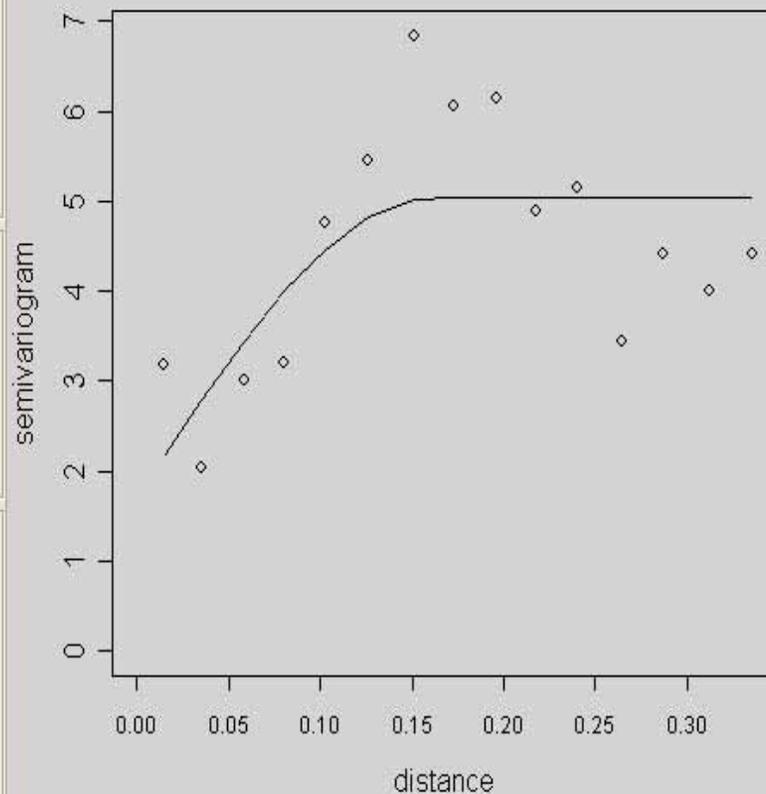
Model	Sill	Range	Nugget
<input type="text" value="Spherical"/>	<input type="text" value="3.32077281"/>	<input type="text" value="0.160644"/>	<input type="text" value="1.71173611"/>

RSS

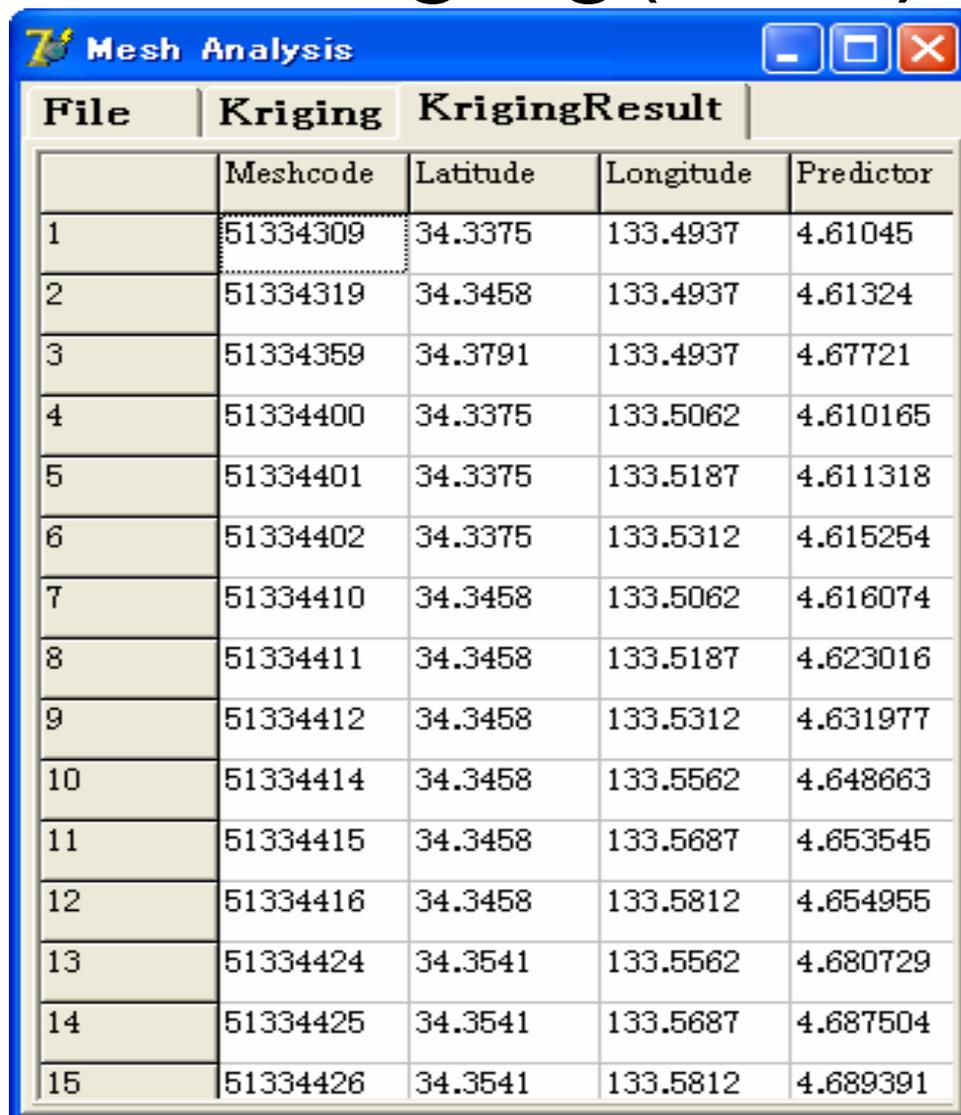
Using Variogram

Model	Sill	Range	Nugget
<input type="text" value="Spherical"/>	<input type="text" value="3.32077281"/>	<input type="text" value="0.160644"/>	<input type="text" value="1.71173611"/>

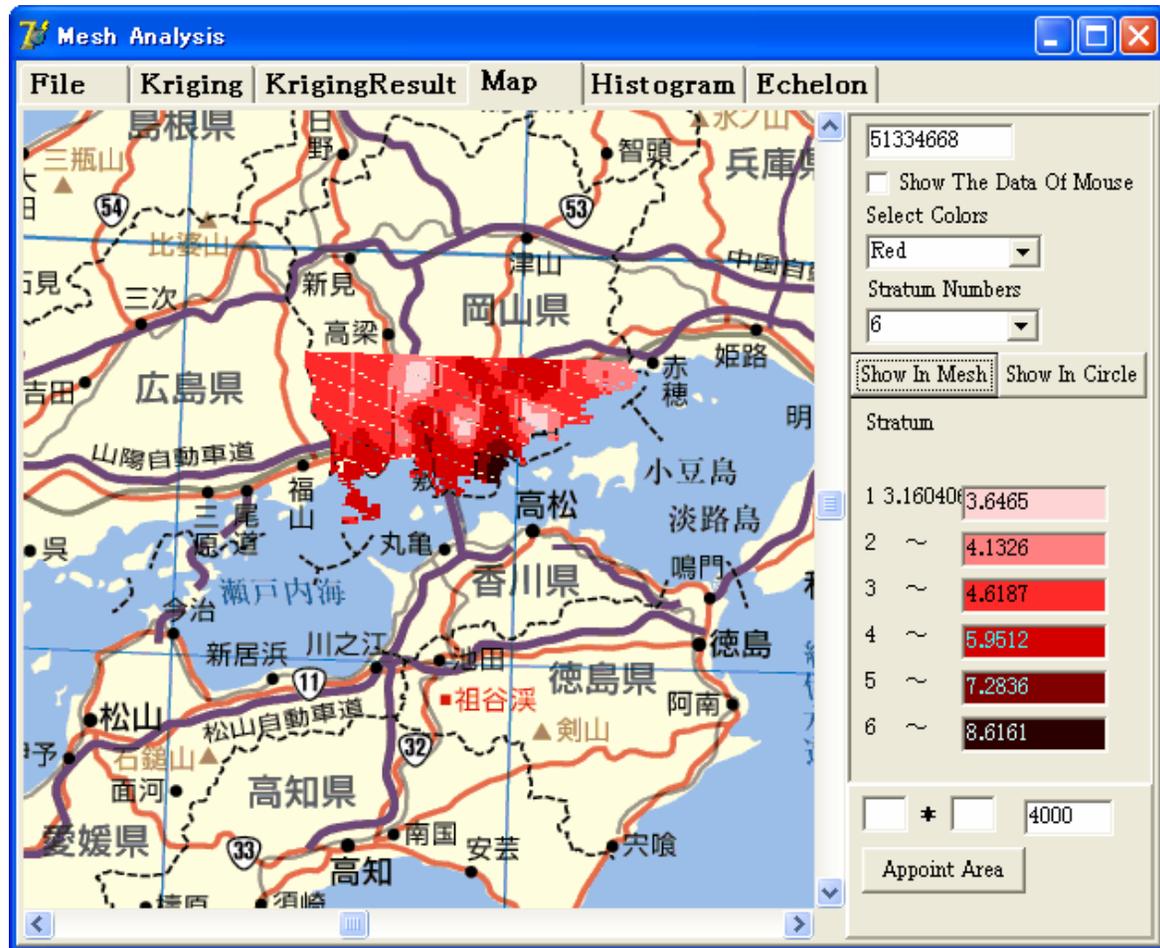
RSS

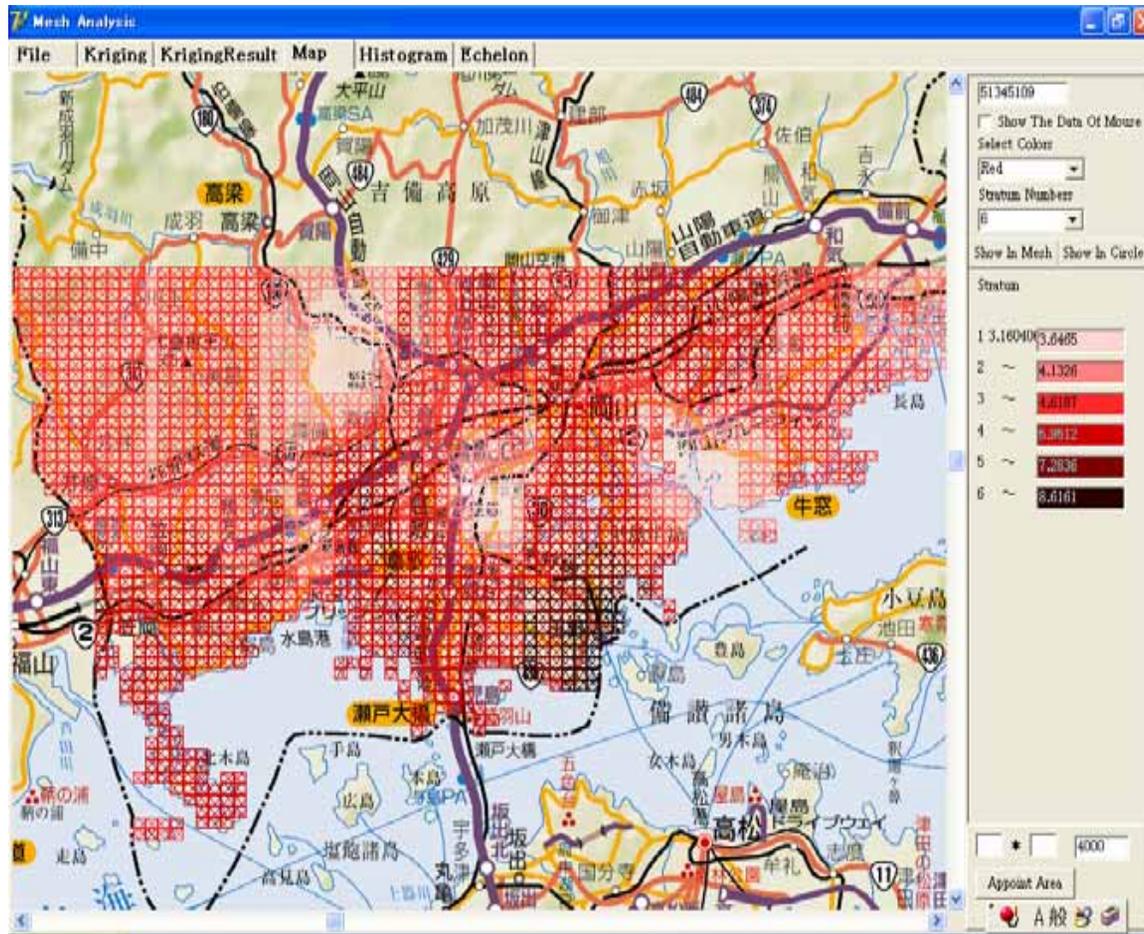


Kriging(予測)



File	Kriging	KrigingResult		
	Meshcode	Latitude	Longitude	Predictor
1	51334309	34.3375	133.4937	4.61045
2	51334319	34.3458	133.4937	4.61324
3	51334359	34.3791	133.4937	4.67721
4	51334400	34.3375	133.5062	4.610165
5	51334401	34.3375	133.5187	4.611318
6	51334402	34.3375	133.5312	4.615254
7	51334410	34.3458	133.5062	4.616074
8	51334411	34.3458	133.5187	4.623016
9	51334412	34.3458	133.5312	4.631977
10	51334414	34.3458	133.5562	4.648663
11	51334415	34.3458	133.5687	4.653545
12	51334416	34.3458	133.5812	4.654955
13	51334424	34.3541	133.5562	4.680729
14	51334425	34.3541	133.5687	4.687504
15	51334426	34.3541	133.5812	4.689391





Kriging & Variogram

- **Kriging**

$$\hat{Z}(x) = m + \sum_{i=1}^n w_i (Z(x_i) - m)$$

x : predict location, x_i : observed data location, w_i : weight

Under $E\{\hat{Z}(x) - Z(x)\} = 0$

Minimize $\sigma_E^2 = \text{Var}(\hat{Z}(x) - Z(x))$

- **Second-order Stationarity & Covariogram**

$$E\{Z(s)\} = \mu$$

$$\text{Cov}\{Z(s_1), Z(s_2)\} = C(s_1 - s_2)$$

- **Intrinsic Stationarity & Variogram**

$$E\{Z(s_1) - Z(s_2)\} = 0$$

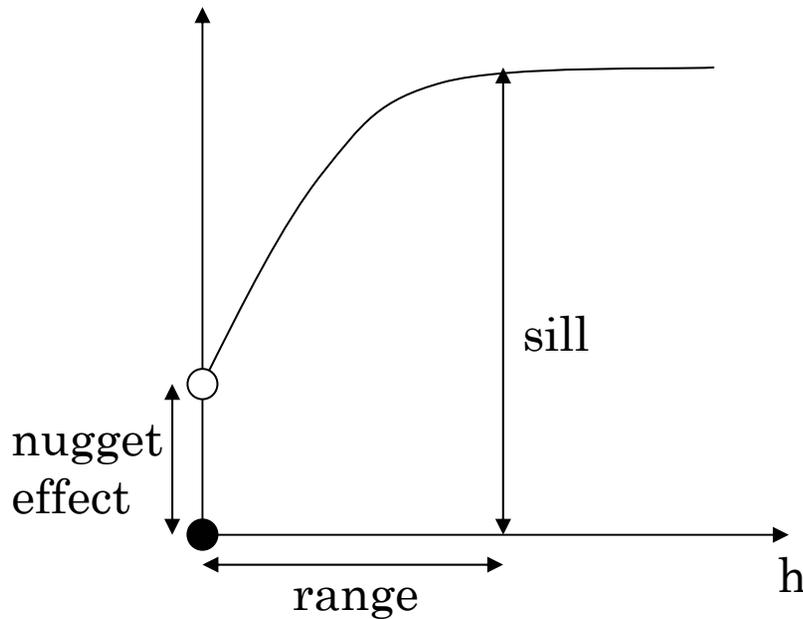
$$\text{Var}\{Z(s_1) - Z(s_2)\} = 2\gamma(s_1 - s_2)$$

γ is called semivariogram and often used in geostatistics.

- ◆ **Relation between Semivariogram and Covariogram**

$$\gamma(h) = C(0) - C(h) \quad C(0) = \text{Var}\{Z(s)\}$$

Typical Semivariogram



- Parameter

- sill: The supremum of (semivariogram).

- range: Minimum distance where reaches it's sill.

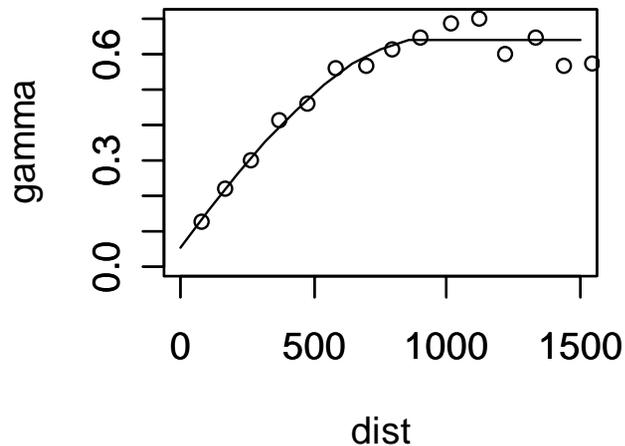
- nugget effect: White noise.

- Spatial dependence

- sill/nugget

- range

Estimate from Empirical Semivariogram



- Empirical Semivariogram
- Theory Semivariogram

nonlinear least squares method (nls)

find θ , to minimize

$$S(\theta) = \sum_{k=1}^K (\hat{\gamma}(h_k) - \gamma(h_k; \theta))^2 \quad (1)$$

where $\theta = (\theta_1, \theta_2, \theta_3)'$

$$\begin{cases} \theta_1 > 0 : \textit{sill} \\ \theta_2 > 0 : \textit{range} \\ \theta_3 \geq 0 : \textit{nugget} \end{cases}$$

let $\phi = (\phi_1, \phi_2, \phi_3)'$,

$$\theta_i = \phi_i^2, \quad i = 1, 2, 3$$

then (1) become

$$S(\phi) = \sum_{k=1}^K (\hat{\gamma}(h_k) - \gamma(h_k; \theta(\phi)))^2 \quad (2)$$

nls for Semivariogram Estimate

$$S(\phi) = \sum_{k=1}^K (\hat{\gamma}(h_k) - \gamma(h_k; \theta(\phi)))^2 \quad (2)$$

have minimize at $\phi = \hat{\phi}$, then

$$\left. \frac{\partial S}{\partial \phi_j} \right|_{\phi = \hat{\phi}} = 0 \quad (j = 1, 2, 3) \quad (3)$$

By Newton method, the k th approximation of ϕ is :

$$\sum_{k=1}^3 \frac{\partial^2 S}{\partial \phi_j \partial \phi_k} \bullet \Delta \phi_k = - \frac{\partial S}{\partial \phi_j} \quad (j = 1, 2, 3) \quad (4)$$

$$\phi^{(k+1)} = \phi^{(k)} + \Delta \phi \quad (5)$$

町丁字等別小地域データ

亀川

研究目的

- 空間統計解析

- 空間データを用いて解析

- 位置情報と属性値で構成されるデータ

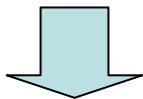


町丁字等別である領域(エリア)で表される小地域データ

- 解析法

- クリギング

- 空間上の未観測地点での属性値の推定法



- 一つの地点での属性値を未知とした推定値を求め、
推定誤差や推定誤差分散を求める

- 可視化

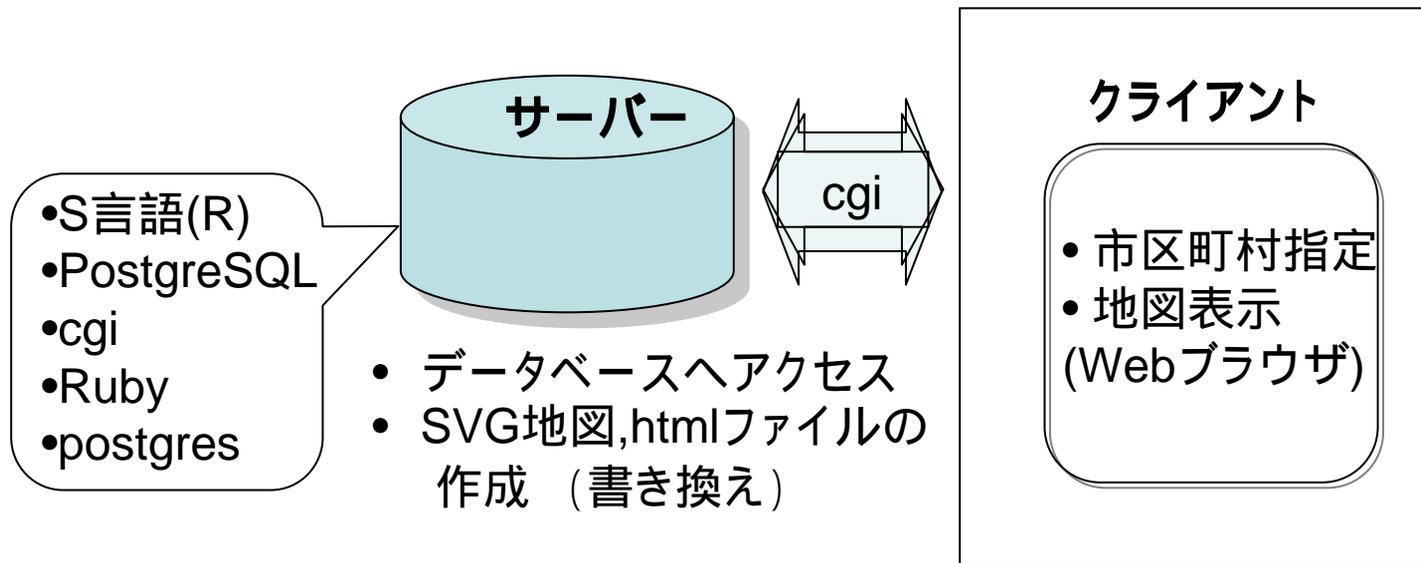
- 視覚的に分析結果を評価

可視化ツール

- 小地域データの分析と結果をWebブラウザ上で提供
 - 町丁字等別データに対応する位置情報が必要
 - ↳ 数値地図の可視化
 - » SVG (Scalable Vector Graphics)



可視化ツール



- 小地域の地図の表示
- 小地域における国勢調査集計データの表示
- データでの地図の塗りわけ
- 空間分析の実行と結果の表示、塗りわけ

实例

- 岡山県

- http://face.f7.ems.okayama-u.ac.jp/~ykameka/cgi-bin/ruby/db/dbfile/index_02.html

- 指定してください。
- 岡山市
 - 倉敷市
 - 津山市
 - 玉野市
 - 笠岡市
 - 井原市
 - 総社市
 - 高梁市
 - 新見市
 - 備前市
 - 御津郡
 - ト御津町
 - ト建部町
 - ト加茂川町
 - 赤磐郡
 - ト瀬戸町
 - ト山陽町
 - ト赤坂町
 - ト熊山町

地図表示

リセット

地図を見るには
SVGViewerが必要
です。



←市区町村を選択してください。

- 指定してください。
- 岡山市
 - 倉敷市
 - 津山市
 - 玉野市
 - 笠岡市
 - 井原市
 - 総社市
 - 高梁市
 - 新見市
 - 備前市
 - 御津郡
 - 御津町
 - 建部町
 - 加茂川町
 - 赤磐郡
 - 瀬戸町
 - 山陽町
 - 赤坂町
 - 熊山町

地図表示
リセット

地図を見るには
SVG Viewerが必要
です。



SVG MAP

Information

← 地域をクリックして下さい。

小地域名:
*

小地域コード:
*

データで塗り合わせる

選択してください

クリギングの結果で塗り合わせる

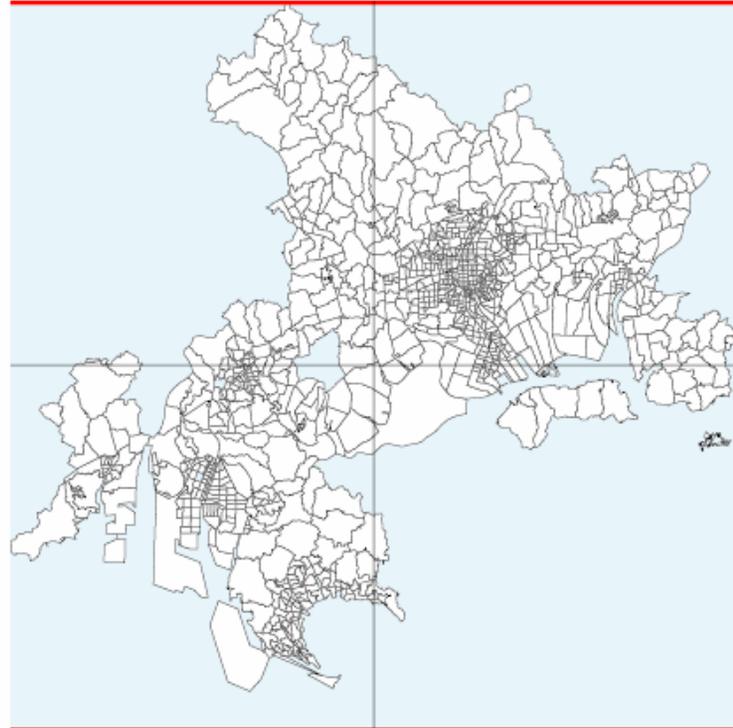
選択してください

- 指定してください。
- 岡山市
 - 倉敷市
 - 津山市
 - 玉野市
 - 笠岡市
 - 井原市
 - 総社市
 - 高梁市
 - 新見市
 - 備前市
 - 御津郡
 - ト御津町
 - ト建部町
 - ト加茂川町
 - 赤磐郡
 - ト瀬戸町
 - ト山陽町
 - ト赤坂町
 - ト熊山町

地図表示
リセット

地図を見るには
SVG Viewerが必要
です。



SVG MAP	Information
 <div data-bbox="466 1135 1199 1320">  </div>	<p>← 地域をクリックして下さい。</p> <p>小地域名: *</p> <p>小地域コード: *</p> <p>地域情報</p> <hr/> <p>データで塗りわける</p> <p>選択してください</p> <p>送信</p> <hr/> <p>クリギングの結果で塗りわける</p> <p>選択してください</p> <p>送信</p>

- 指定してください。
- 岡山市
 - 倉敷市
 - 津山市
 - 玉野市
 - 笠岡市
 - 井原市
 - 総社市
 - 高梁市
 - 新見市
 - 備前市
 - 御津郡
 - ト御津町
 - ト建部町
 - ト加茂川町
 - 赤磐郡
 - ト瀬戸町
 - ト山陽町
 - ト赤坂町
 - ト熊山町

地図表示

リセット

地図を見るには
SVG Viewerが必要
です。



SVG MAP



Information

← 地域をクリックして下さい。

小地域名:

津島中1丁目

小地域コード:

33201268001

地域情報

データで塗りわかる

選択してください

送信

クリギングの結果で塗りわかる

選択してください

送信

小地域名:津島中1丁目,小地域コード:33201268001

人口,世帯数	人口	男	女	世帯数	老人化率
	817	433	384	307	0.00367197

15歳以上年齢,配偶関係	分類項目	総数	男(配偶関係不詳を含む)	未婚	有配偶	死別	離別	女(配偶関係不詳を含む)	未婚	有配偶	死別	離別	1人の一般世帯(単独世帯)
	総数	577	322	47	269	2	4	255	36	214	2	2	90

65歳以上年齢,配偶関係	分類項目	総数	男(配偶関係不詳を含む)	未婚	有配偶	死別	離別	女(配偶関係不詳を含む)	未婚	有配偶	死別	離別	1人の一般世帯(単独世帯)
	65歳以上	3	0	0	0	0	0	3	0	1	1	1	0

世帯の種類,世帯人員	世帯数	世帯人員が1人	2人	3人	4人	5人	6人	7人	8人	9人	10人以上	世帯人員	1世帯当たり人員	施設等の世帯数	世帯人員	間借り・下宿などの単身者	会社などの独身寮の単身者
		307	90	57	48	94	15	3	0	0	0	0	817	2.66	0	0	1

世帯の家	分類	総数	親族世帯	核家族世帯	夫婦のみ	夫婦と子供	男親と子供	女親と子供	その他の親世帯	夫婦と両親	夫婦とひとり親	夫婦、子供と両親	夫婦、子供とひとり親	夫婦と他の親族(親、子を含まない)	夫婦、子供と他の親族(親を含まない)	夫婦、親と他の親族(子供を含まない)	夫婦、子供、親の親族	兄弟姉妹のみ	他に類されない親世帯	非親族世帯	単独世帯	3代世帯
	一般世		307	217	217	53	156	2	6	0	0	0	0	0	0	0	0	0	0	0	0	90



指定してください。

- 岡山市
- 倉敷市
- 津山市
- 玉野市
- 笠岡市
- 井原市
- 総社市
- 高梁市
- 新見市
- 備前市
- 御津郡
 - 御津町
 - 建部町
 - 加茂川町
- 赤磐郡
 - 瀬戸町
 - 山陽町
 - 赤坂町
 - 熊山町

地図表示

リセット

地図を見るには
SVGViewerが必要です。



SVG MAP

- 選択してください
- 人口
 - 男
 - 女
 - 世帯数
 - 老人化率
 - 15歳以上年齢, 配偶関係
 - ト総数
 - ト男(配偶関係不詳を含む)
 - ト未婚
 - ト有配偶
 - ト死別
 - ト離別
 - ト女(配偶関係不詳を含む)
 - ト未婚
 - ト有配偶
 - ト死別
 - ト離別
 - 1人の一般世帯(単独世帯)
 - 65歳以上年齢, 配偶関係
 - ト総数
 - ト男(配偶関係不詳を含む)
 - ト未婚
 - ト有配偶
 - ト死別
 - ト離別
 - ト女(配偶関係不詳を含む)
 - ト未婚
 - ト有配偶
 - ト死別

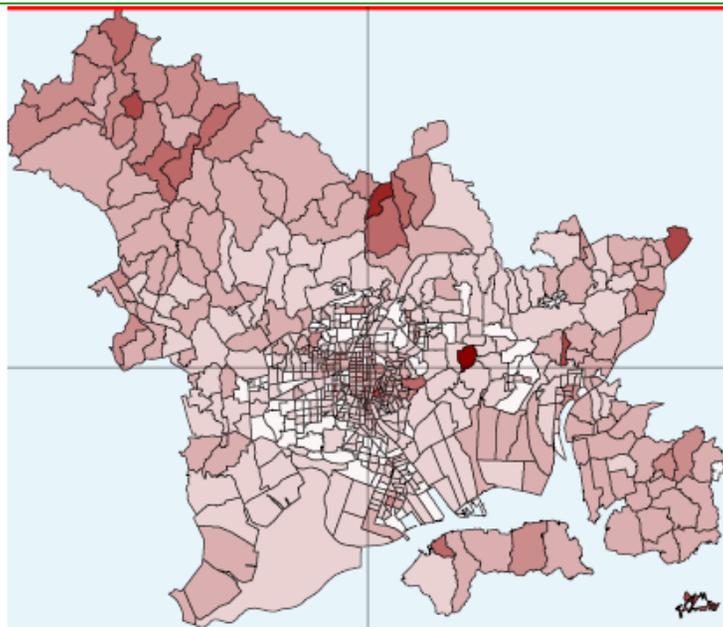
送信

クリギングの結果で塗りわかる

選択してください

送信

老人化率



Information

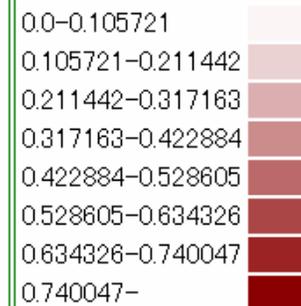
← 地域をクリックするとデータが表示されます

小地域名:

今谷

データ:

0.740047



- 指定してください。
- 岡山市
 - 倉敷市
 - 津山市
 - 玉野市
 - 笠岡市
 - 井原市
 - 総社市
 - 高梁市
 - 新見市
 - 備前市
 - 御津郡
 - ト御津町
 - ト建部町
 - ト加茂川町
 - 赤磐郡
 - ト瀬戸町
 - ト山陽町
 - ト赤坂町
 - ト熊山町

地図表示
リセット

地図を見るには
SVG Viewerが必要
です。



SVG MAP

Information

← 地域をクリックして下さい。

小地域名:
*

小地域コード:
*

データで塗りわける

選択してください

クリギングの結果で塗りわける

選択してください

- 選択してください
- 選択してください
- 老人化率の予測値
- 老人化率の予測誤差

予測値

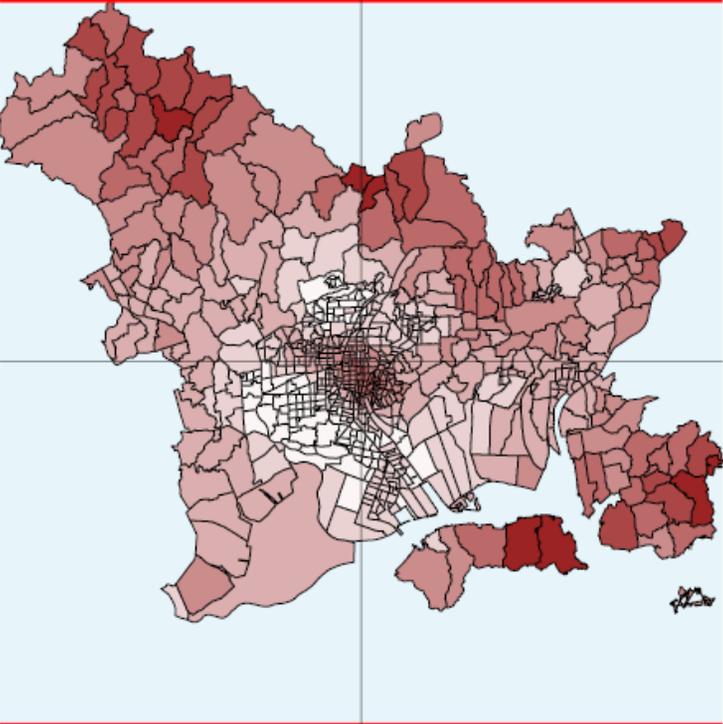
町字別SVG地図 - Microsoft Internet Explorer

ファイル(E) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)

戻る 検索 お気に入り

アドレス(D) http://face.f7.ems.okayama-u.ac.jp/~ykameka/cgi-bin/ruby/db/dbfile/krigergbmap2.cgi 移動

老人化率の予測値



Information

← 地域をクリックするとデータが表示されます

小地域名:

今谷

データ:

0.1928840212

0.0906695662-
0.1391651111
0.1391651111-
0.1876606559
0.1876606559-
0.2361562008
0.2361562008-
0.2846517456
0.2846517456-
0.3331472905
0.3331472905-
0.3816428353
0.3816428353-
0.4301383802
0.4301383802-1

+

-

←

↑

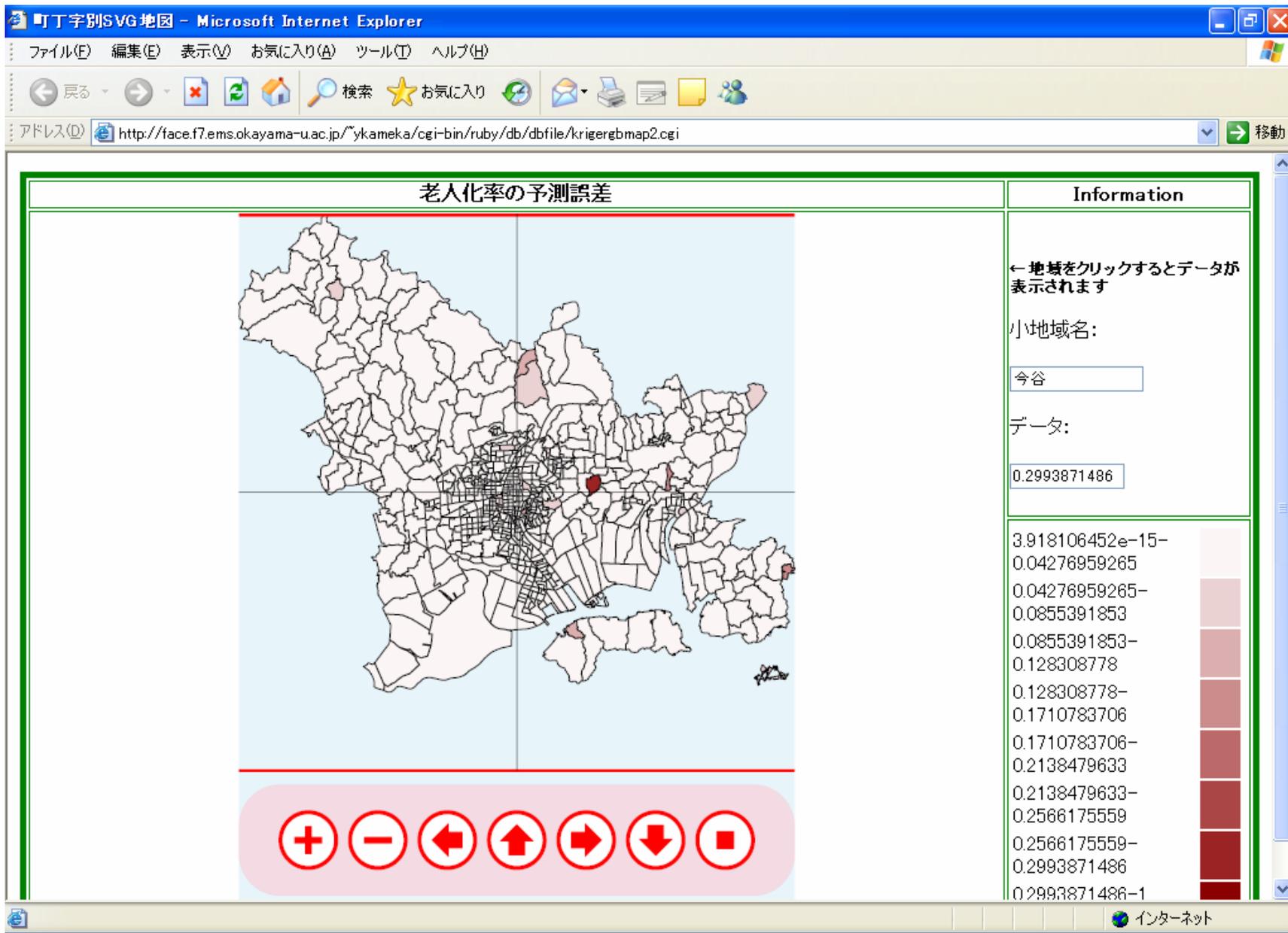
→

↓

■

インターネット

老人化率の予測誤差の地図



本システムの機能

- サーバにアクセスしデータを表示するWeb上ツールを作成した
- クリギングによる予測誤差を求めた
- 分析結果の可視化により視覚的にデータを眺めることができ、距離以外の要因でのホットスポットの検出を行うことができた

最後に

- S-PLUS / Rを統計エンジンとして利用している2つのシステムの紹介
- 各種数値地図の提供・配布
→利用が容易に
- 高精細地図の提供
(GoogleMap, GoogleEarth)
→APIの提供
→S-PLUSでの利用へ