

# 空間データ分析マシン「SDAM」ユーザガイド

1 . 開発の経緯.....	1
2 . 特色 .....	2
3 . 機能 .....	2
4 . 構成 .....	3
5 . ダウンロード・インストールの方法 .....	5
6 . 操作方法 .....	6
6 . 1 地図作成 .....	8
6.1.1 主題図.....	8
6.1.2 カルトグラム .....	9
6 . 2 検索.....	13
6 . 3 空間解析 .....	17
6.3.1 オーバーレイ .....	17
6.3.2 バッファー.....	19
6.3.3 ユニオン/ディゾルブ .....	21
6.3.4 TIN (不規則三角形網) .....	23
6.3.5 ポロノイ (ティーセン) .....	24
6.3.6 凸包 .....	25
6 . 4 地域分析 .....	26
6.4.1 基本統計量.....	26
6.4.2 多変量解析.....	26
6.4.3 探索的空間分析 (ESDA) .....	32
6.4.4 ポイント・パターン分析 .....	35
6.4.5 グローバル空間的自己相関.....	37
6.4.6 ローカル空間的自己相関 .....	43
6.4.7 Join 統計量.....	48
6.4.8 空間的相互作用分析 .....	49
6.4.9 GWR (地理加重回帰分析) .....	55
(1) GWR とは.....	55
(2) 入力画面 .....	55
7 . 評価.....	61

## 1 . 開発の経緯

1950年代末以降，計量地理学や地域科学の分野では，空間モデルや計量的手法が数多く開発され，実証研究に活用されるとともに，都市計画，地域政策に援用されてきた．これらのモデル・手法を操作可能にするにはコンピュータの助けが必要になるが，これまで統計パッケージ

にせよ GIS ソフトウェアにせよ、空間分析に対するユーザのニーズを十分に満たす機能を備えたものはきわめて少なかった。最近、GIS ベンダーが操作機能を充実させる動きをみせているが、洗練されたソフトウェアは一般に価格が高く、個人ユーザには手が出ないのが実状である。

本章執筆者の一人、村山は大学で空間情報科学の講義や GIS の実習を担当している。これまで実習ではラボの PC にインストールされた ArcView を用いてきたが、学生が自分のパソコンで場所を選ばず自由に空間分析を行える環境は構築できないものかと考え、この「空間データ分析マシン」(Spatial Data Analysis Machine, 以下 SDAM と記載)の構想を思い立った。本章のもう一人の執筆者、尾野の賛同を得て、このシステムの開発に着手したのは 2001 年秋のことであった。村山は操作機能の選別とシステムの設計を、尾野はプログラミングを担当した。拡張性を考え、GIS エンジンや統計解析用の言語はオープンソースを使うことにし、それ以外のアプリケーションは独自開発することに決めた。大学での GIS 教育を想定し、使いやすさを最優先にシステム構築を進めた。こうして 2002 年 10 月に試作版(統合型空間分析システムと呼称)が完成した(村山・尾野, 2003)。その後、逐次改良を加えるとともに、機能を追加してできあがったのがこの SDAM (2004 年 8 月リリース)である。

## 2. 特色

SDAM は 地図表示機能と統計解析ソフトウェアを融合させたシステムである。GUI によって、面倒なコマンド入力なしに、空間分析が行えるよう設計されている。周知のように、空間分析には多種多様な手法が存在するが、SDAM では最近の動向を踏まえ、とくに空間分析関連の授業に有用と考えられる機能を優先的に採用している。

## 3. 機能

SDAM は次の操作機能を有する。地図作成(コロプレス図, 円ドット図, カルトグラム), カーネル密度推定, 空間・属性検索, オーバーレイ, バッファー, ユニオン, TIN, ポロノイ, 凸包, 記述統計, 多変量解析(回帰分析, 因子分析, クラスタ分析), 探索的空間分析(ESDA), ポイント・パターン分析(方格法, 最近隣法, K 関数法), 空間的自己相関分析(グローバル・モラン統計量, グローバル・ギアリ統計量, モラン・プロット, ローカル・モラン統計量, ローカル G 統計量, ローカル G\* 統計量), 空間的相互作用分析(古典的重力モデル, 発生制約型モデル, 吸収制約型モデル, 二重制約型モデル), 地理加重回帰分析(GWR)。

SDAM は Windows98/Me/NT/2000/XP 上で動作する。本システムの構築にあたり、以下のアプリケーションを利用したが、これらはオープンソースのフリーウェアである。

### GIS 関連

GeoTools for Java(英国 CCG が開発した Java 用 GIS エンジン)

JTS(Java Topology Suite) (オーバーレイ解析モジュール)

### 空間分析関連

R 言語(統計解析用言語)

J(D)COM(システム本体と RSTAT サーバ間の通信ソフトウェア)

RSTAT サーバ(JCOM と R 言語間の通信, 独自開発)

VRML プラグイン(カーネル密度推定, 3D 表示)

R 言語は商用統計分析ソフトウェア S 言語のクローンであるが、S 言語に劣らない拡張パッケ

ージを数多く有している (ヴェナブルズ, 2001)。SDAM では, プログラム本体 (Java 言語で作成) と R 言語を統合し, J(D)COM, RSTAT サーバの利用により, GIS 機能と空間分析機能との Tight なカップリングを実現した。

#### 4 . 構成

図 1 は SDAM のシステム構成を示している。

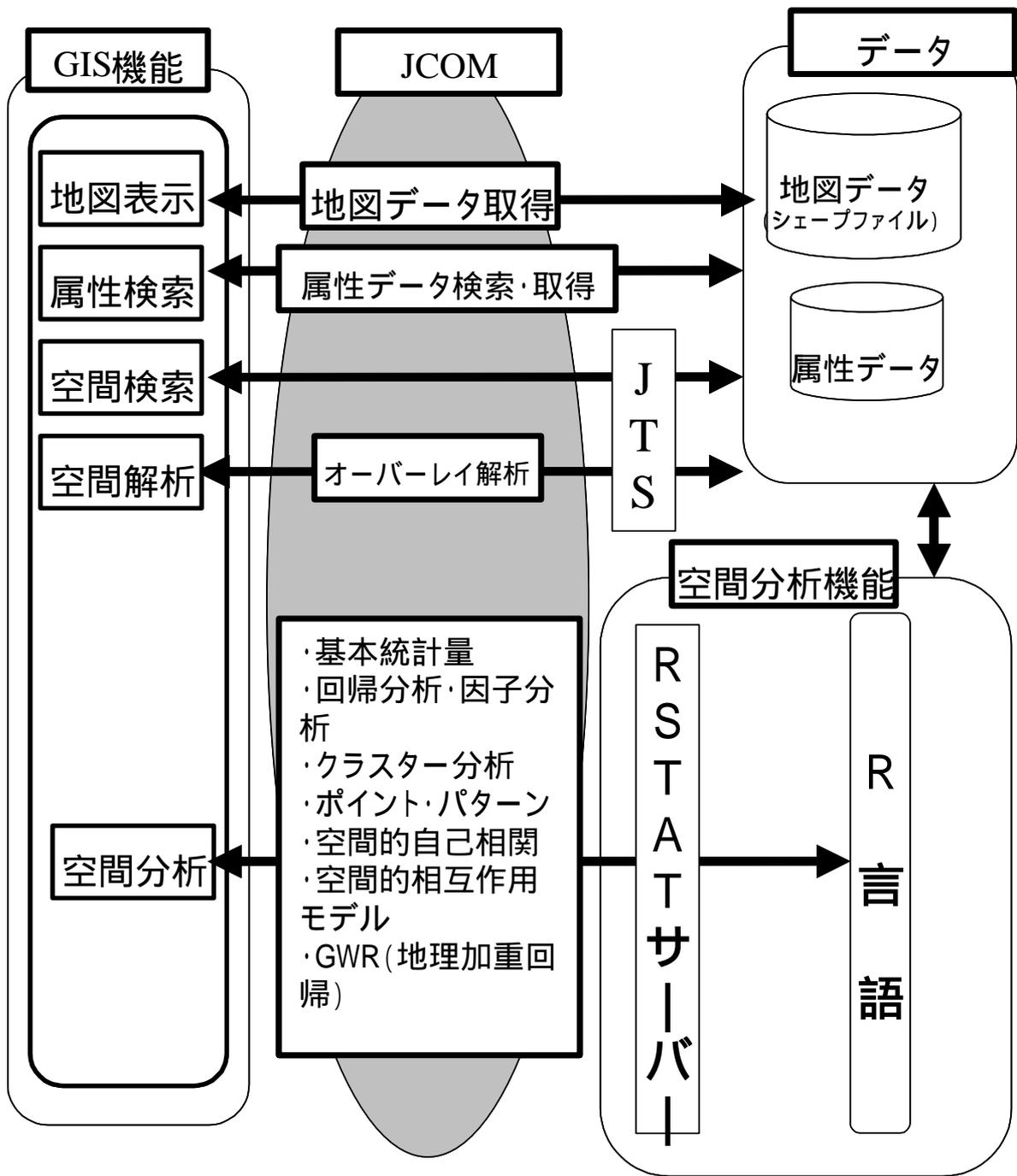


図1 システム構成図

SDAMでは、ESRI社のArcViewシェープファイルを処理できる。シェープファイルは、以下の3つのファイルより構成される。

- ・ SHPファイル(拡張子".shp")・・・図形情報
- ・ SHXファイル(拡張子".shx")・・・インデックス情報(SHPファイルのインデック

ス情報)

- ・ DBase ファイル ( 拡張子 ".dbf" ) . . . 属性情報 ( 地理行列 )

図形情報は、ポイント、ライン、ポリゴンからなる 2 次元図形として格納される。属性情報は DBase 形式で地理行列の形で管理されるので、Excel で読み込め、データ変更も容易である。なお、図形情報と属性情報とはリンクされ、DBase ファイル内にユニークな ( 他と重複しない ) 整数型の ID のフィールドが定義される。

## 5 . ダウンロード・インストールの方法

SDAM は無償であり、以下のサイトから自由にダウンロードできる<sup>1</sup>。

<http://land.geo.tsukuba.ac.jp/teacher/murayama/sdam/>

SDAM は、CPU が Pentium 以上 ( 推奨 )、メモリー 256MB ( 推奨 512MB ) 以上、HD20GB 以上のハードウェアで動作する。OS は Windows2000/XP。SDAM を動かすために必要なソフトとその入手先を表 1 に示す。

表 1 SDAM の動作に必要なソフト

ソフト名	入手先	使用バージョン
Java 2 SE	<a href="http://java.sun.com/j2se/1.4/ja/index.html">http://java.sun.com/j2se/1.4/ja/index.html</a>	1.4.2X 以上
R 言語	<a href="http://www.r-project.org/">http://www.r-project.org/</a>	1.9.1
R(D)COM	<a href="http://cran.r-project.org/contrib/extra/dcom">http://cran.r-project.org/contrib/extra/dcom</a>	1.3.5

R 言語は、図 2 のホームページから CRAN(The Comprehensive R Archive Network) ( 図中丸で囲んだ部分 ) にアクセスすればダウンロードできる。R(D)COM は、VisualBasic や VC++ などで作成されたプログラム、あるいは Excel の VBA より、R 言語の使用を可能にするアプリケーションである。

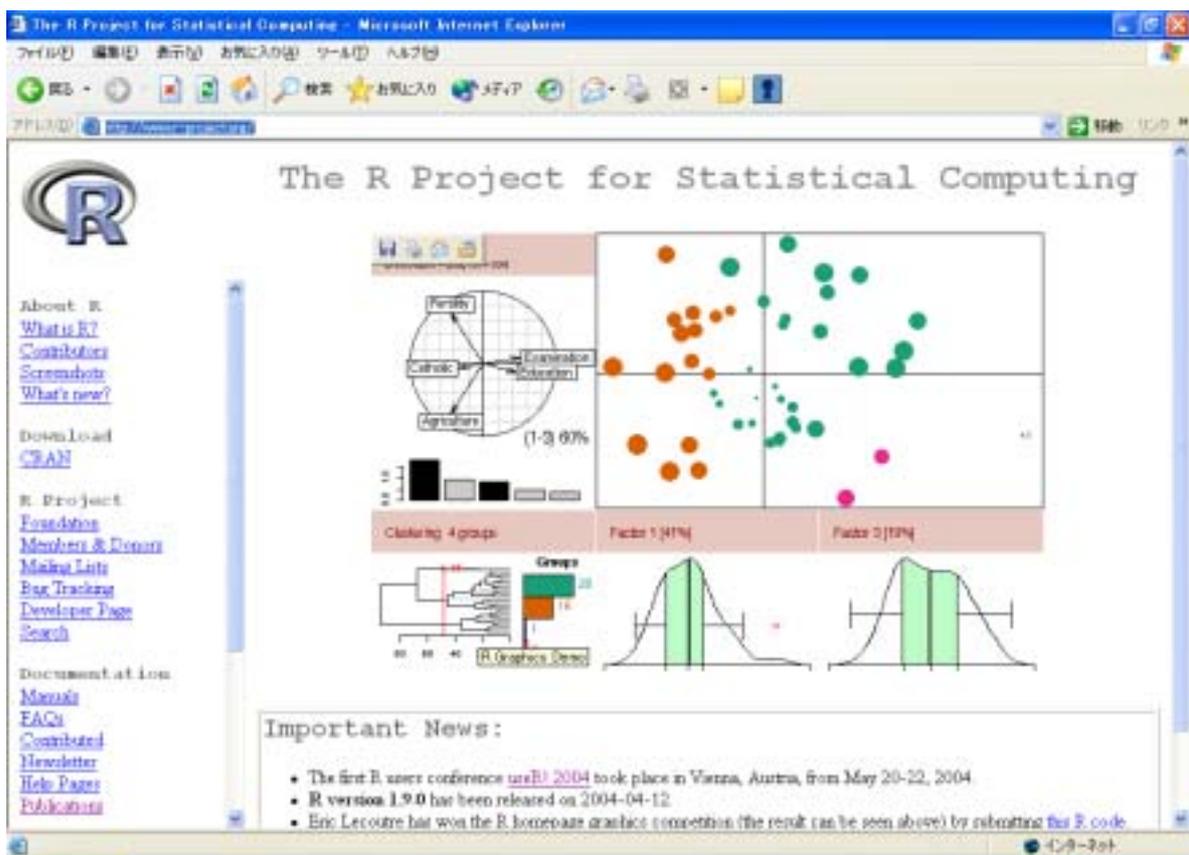


図2 R言語のホームページ

R言語本体には、すでに多変量解析用のパッケージなどが含まれているが、TINや空間的自己相関などを実現するにあたって、いくつかのR言語の追加パッケージが必要になる(表2)。SDAMでは、これらの追加パッケージは、インストールプログラム実行時に自動的にR言語にインストールされるよう設計されている<sup>2</sup>。以上がインストールされると、SDAMが動作する。

表2 SDAMに必要なR言語の追加パッケージ

パッケージ名	機能
fields	カーネル密度推定
issa	因子分析と各種ユータリティ
maptools	ArcViewシェープファイルの読み込み
spgwr	地理加重回帰(GWR)
splancs	最近隣法, K関数法
spdep	空間的自己相関, ウェイト・マトリックス作成
tripack	TIN・ボロノイ図・凸包の作成

## 6. 操作方法

Windowsのエクスプローラで、インストール先のsdam.jarをダブル・クリックすると、SDAMが起動する(図3)。

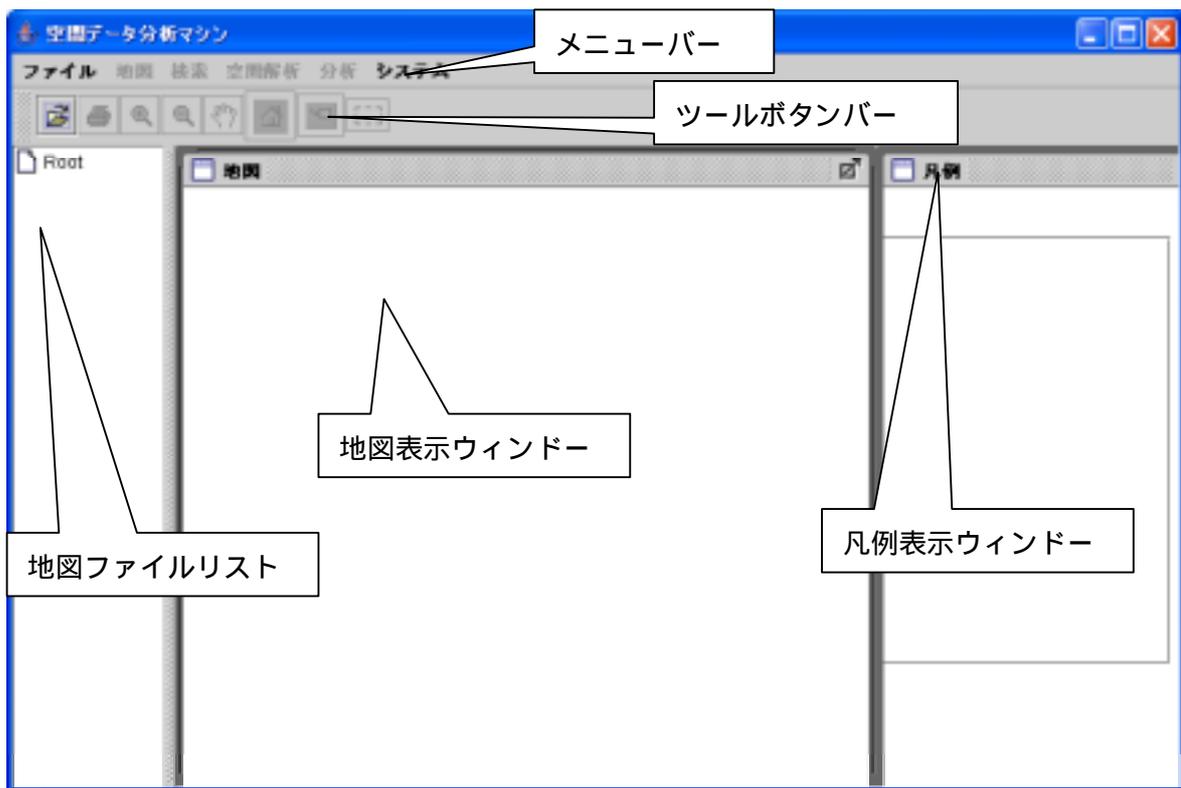


図3 SDAMの起動画面

「ファイル」メニューより、「開く」を選択するか、ツールボタンのを押すと、「地図ファイル選択」ウィンドウが開く。「地図ファイル選択」ウィンドウのを押すと、「シェープファイル選択」ウィンドウが表示される。表示したいシェープファイルを選び、ID 選択項目で、シェープファイルと属性データを結びつける ID 用のユニークな項目を選択する。すると、「地図表示」ウィンドウに図4の地図が表示される。「凡例」ウィンドウの下には縮尺が表示される。

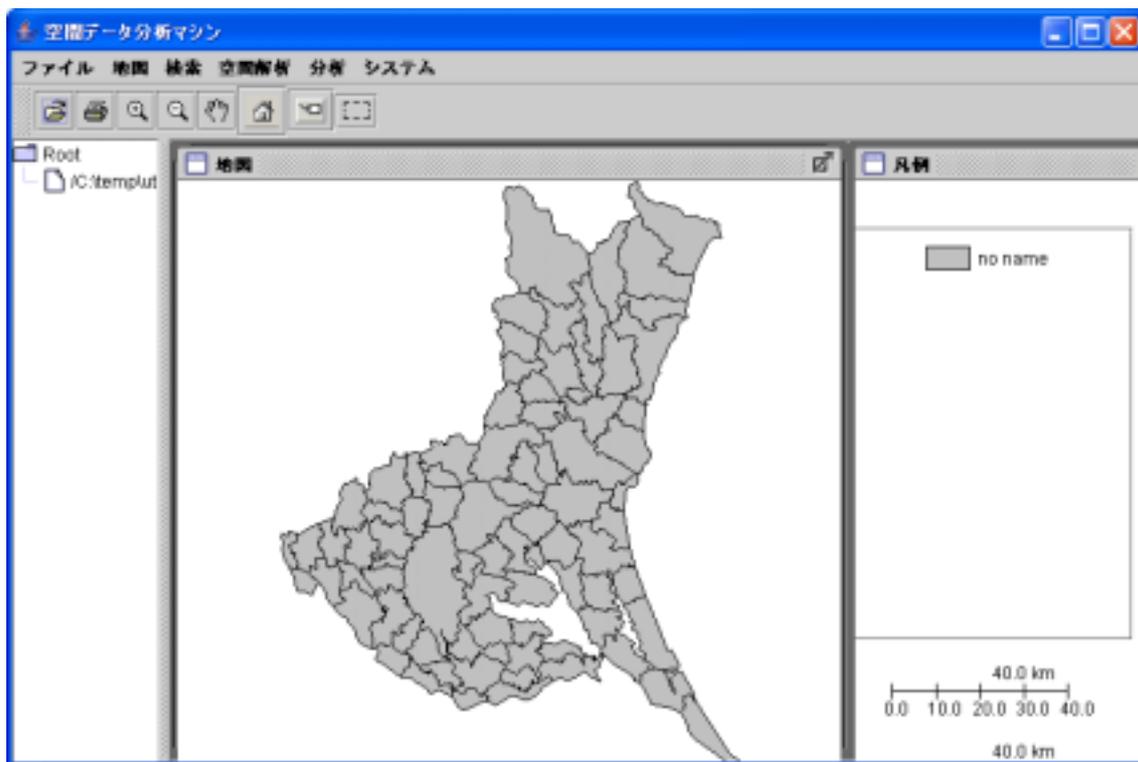


図4 地図の表示

## 6.1 地図作成

### 6.1.1 主題図

地図データの数値属性の値に基づいて、色の塗りわけが可能なコロプレス図を描くことができる(図5)。「地図選択」で表示したい地図を、「項目選択」で表示したい項目(属性)を、そして「表示タイプ」を選び、「階級数選択」で階級値を選択すると、コロプレス図が地図表示ウィンドーに描画され、これに対応した凡例が凡例表示ウィンドーに表示される。

階級区分の仕方は、「等間隔」、「等サイズ」、「自然階級分類」の3タイプが選択できる。「等間隔」は、各階級値の幅を同一に設定する表示法、「等サイズ」は1つの階級に収まる個数を同数に設定する表示法、そして「自然階級分類」は非階層型クラスター分析の一手法を用いて、各階級内の値と各階級の平均値との差の平方和を最小にする表示法である。図6は「自然階級分類」によるコロプレス図を作成した例である。

「地図表示」メニューで、「円ドット図」メニュー項目を選択すると、円ドット図を作成することができる。

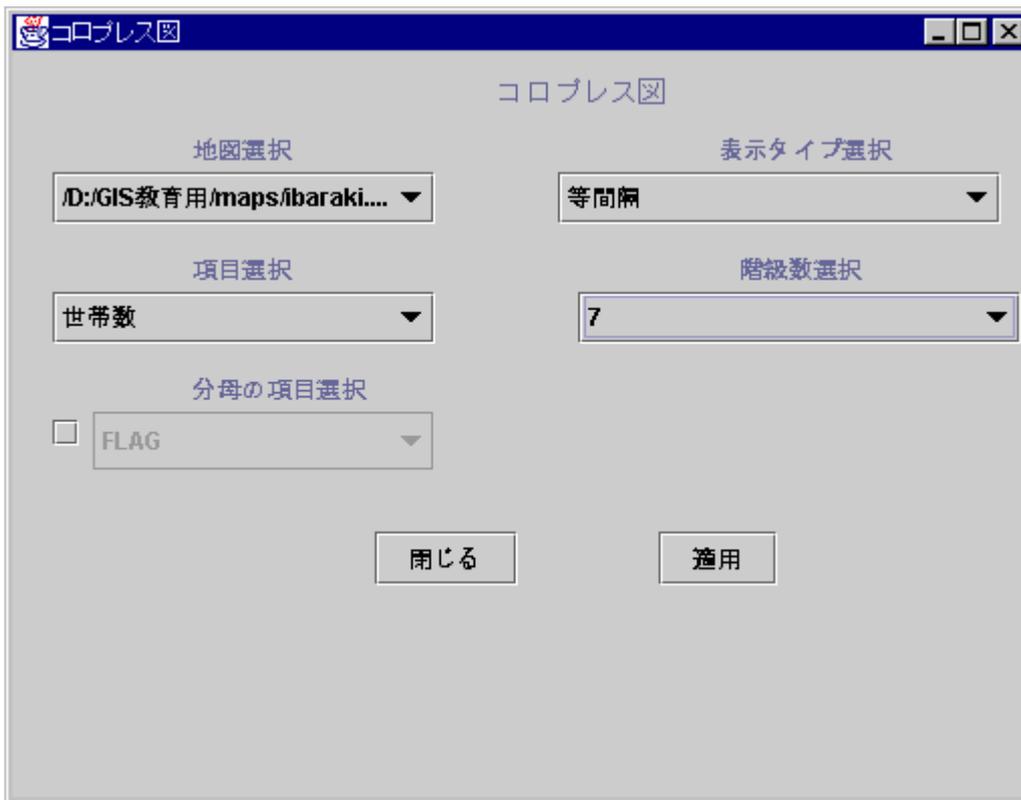


図5 「コロプレス図」のウィンドウ

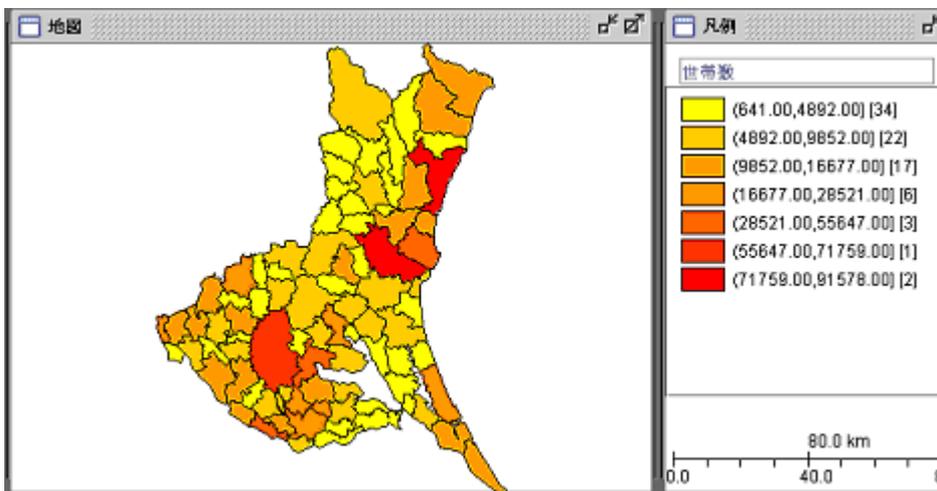


図6 コロプレス図「自然階級分類」の表示

### 6.1.2 カルトグラム

カルトグラム（変形地図）には多種多様な方法が存在するが(Dorling, 1996), SDAM では連続面積カルトグラムと非連続面積カルトグラムを描画できる。

「連続面積カルトグラム」を作成する場合は、図7のウィンドウで、地図と項目を選択し、最後に収束回数を入力して、「適用」ボタンを押す。表示までには多少の時間がかかる(図8)。「非連続面積カルトグラム」では図9のような図が描かれる。描画はほぼ瞬時である。



図7 「連続面積カルトグラム」のウィンドー

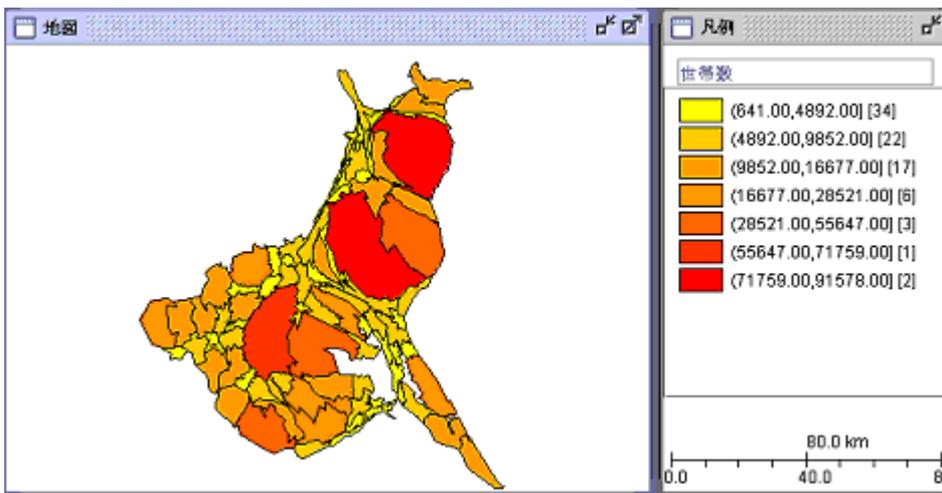


図8 連続面積カルトグラム

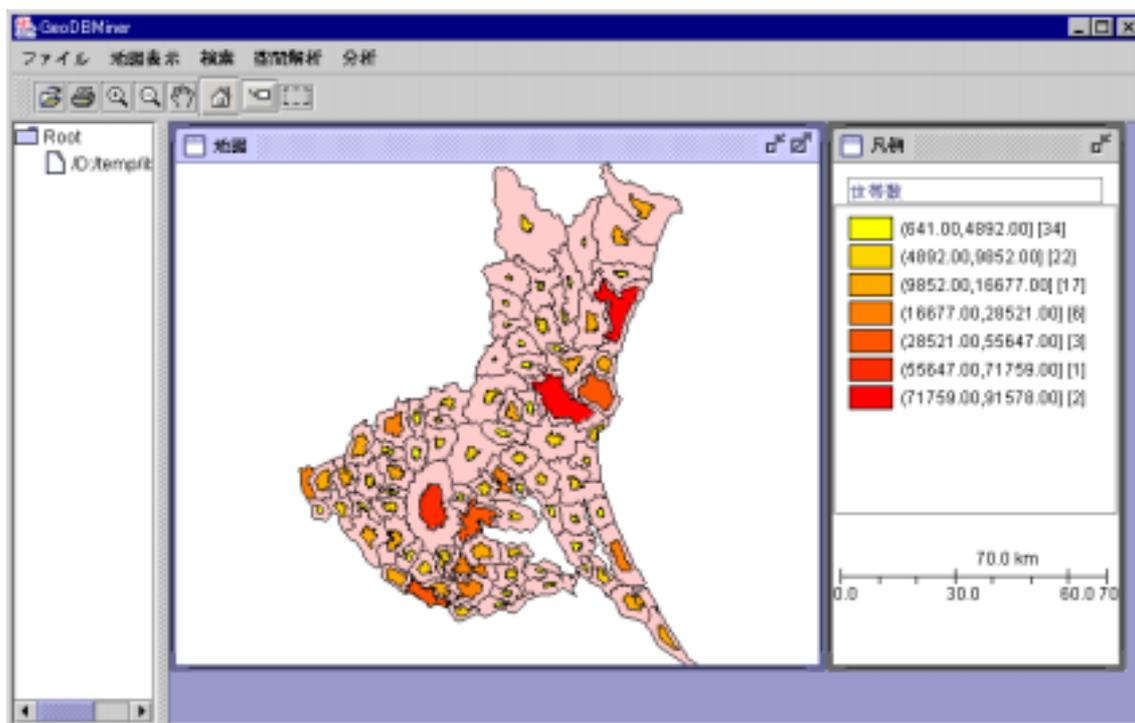


図9 非連続面積カルトグラム

### 6.1.3 カーネル密度推定 - サーフェス (平滑化) -

カーネル密度推定法は、中心から重みが距離逓減していく様を関数で示す。人口分布など離散量の表示に適する手法で、気候値（気温）などの連続量には向かない。本システムでは、属性のないポイントデータと離散量の属性をもつポイントデータを処理できる。

「カーネル密度推定法」サブメニューを選択すると、図10の「カーネル密度推定」ウィンドウが表示される。このウィンドウにおいて、「地図選択」でターゲットになる地図名を、「分布関数」で「ガウス分布」または「指数分布」を選択する。さらに属性の利用に関して、選択か未選択かを選ぶ。次にバンド幅を指定し、次に、グリッドのセルサイズ、グリッドファイル名などを指定する。3D表示を行う場合は、出力するVRMLのファイル名を指定する。図11と図12に実行結果の例を示す。



図 10 「カーネル密度推定」のウィンドウ

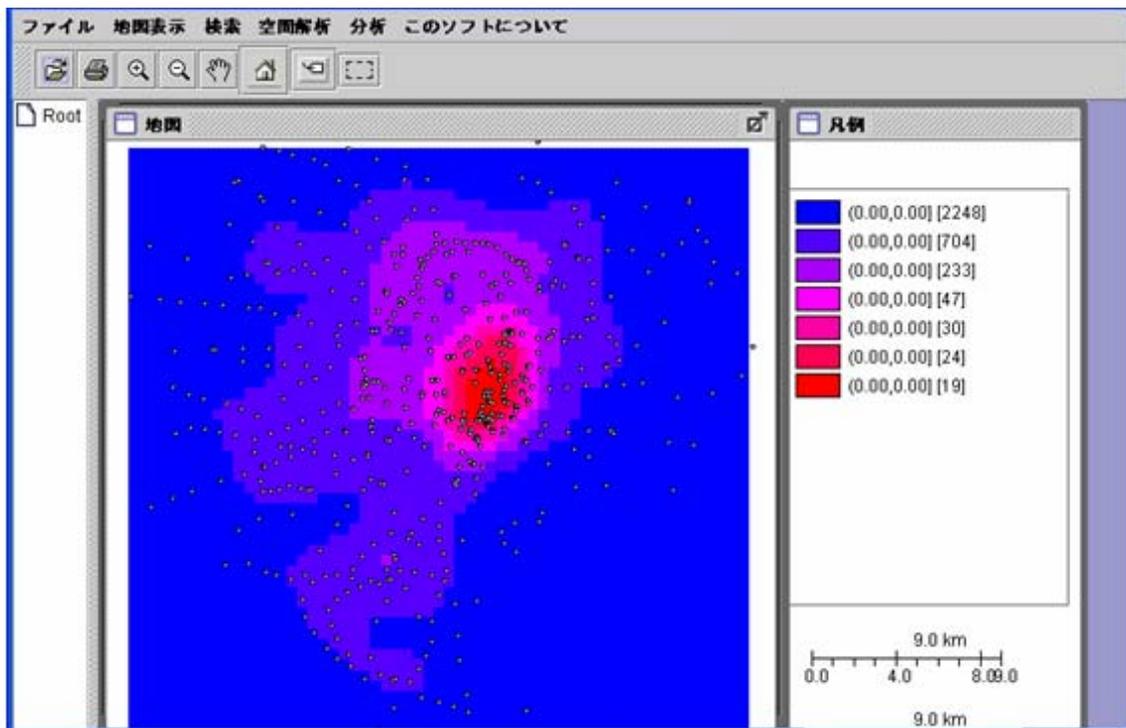


図 11 カーネル密度推定（ガウス分布，属性選択無し）

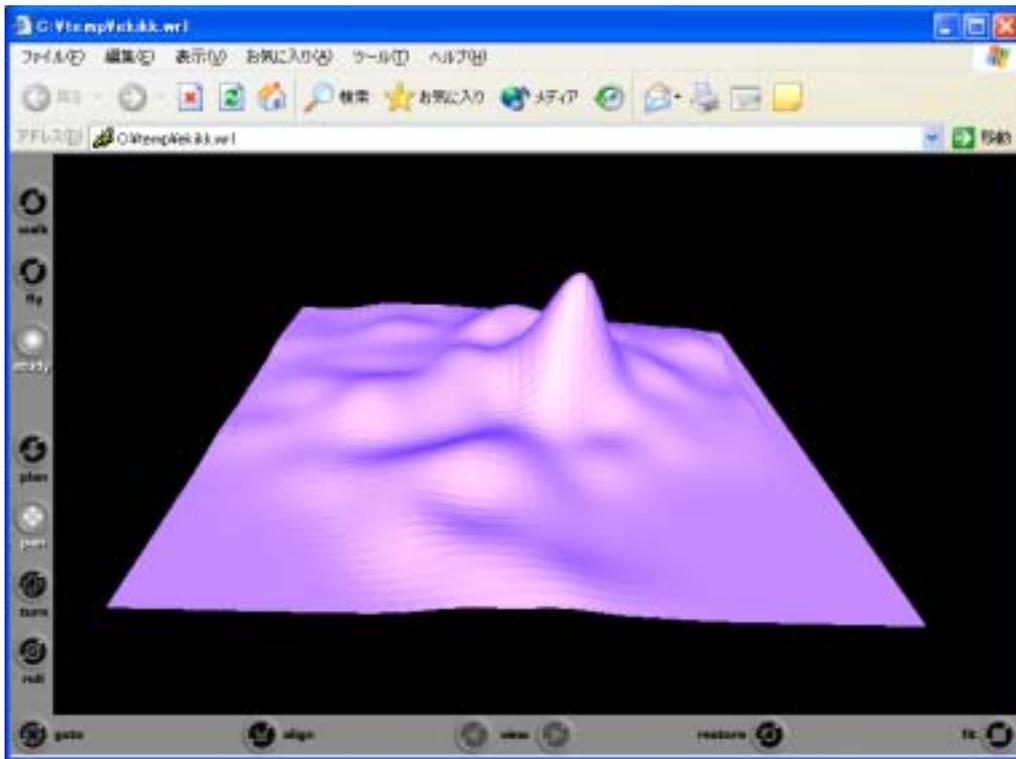


図 12 カーネル密度推定 ( 3D 図, VRML 出力 )

## 6.2 検索

SDAM は、空間検索と属性検索の機能を有する。空間検索では、表 3 のオペレーターがサポートされている。図 13 は equals の、そして図 14 は intersects の操作結果の例である。

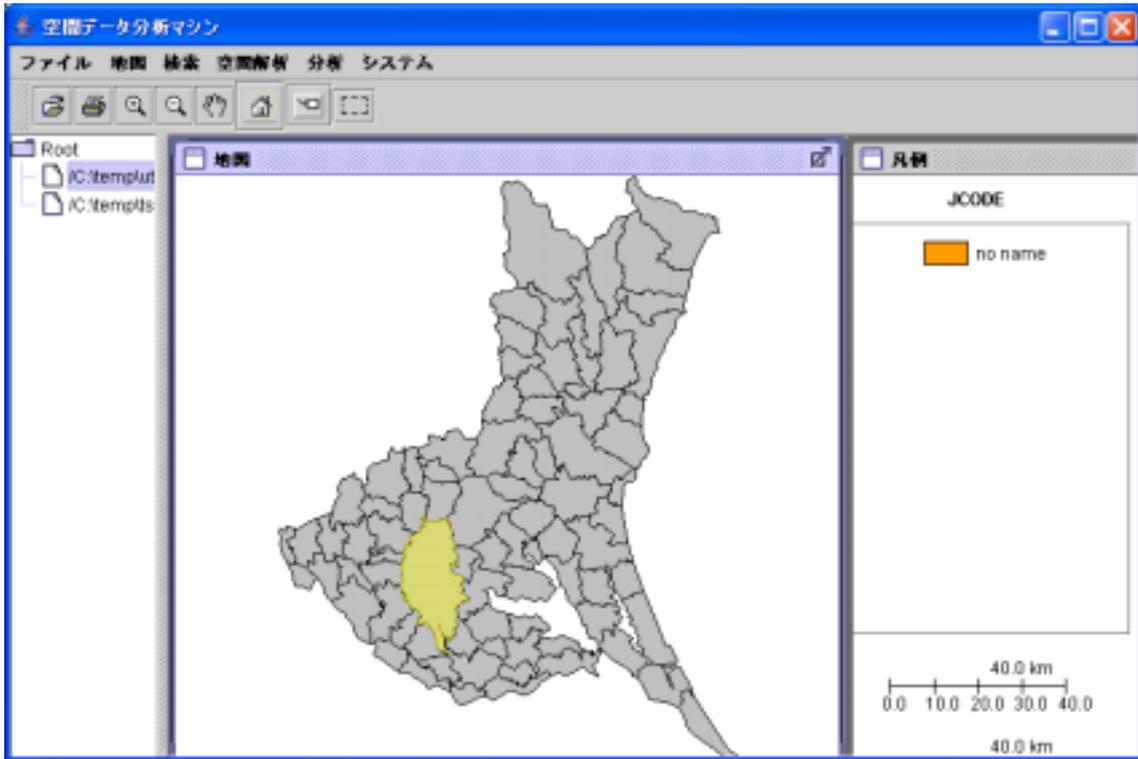


図 13 「equals」の例（ターゲット「茨城県」，サブジェクト「つくば市」）

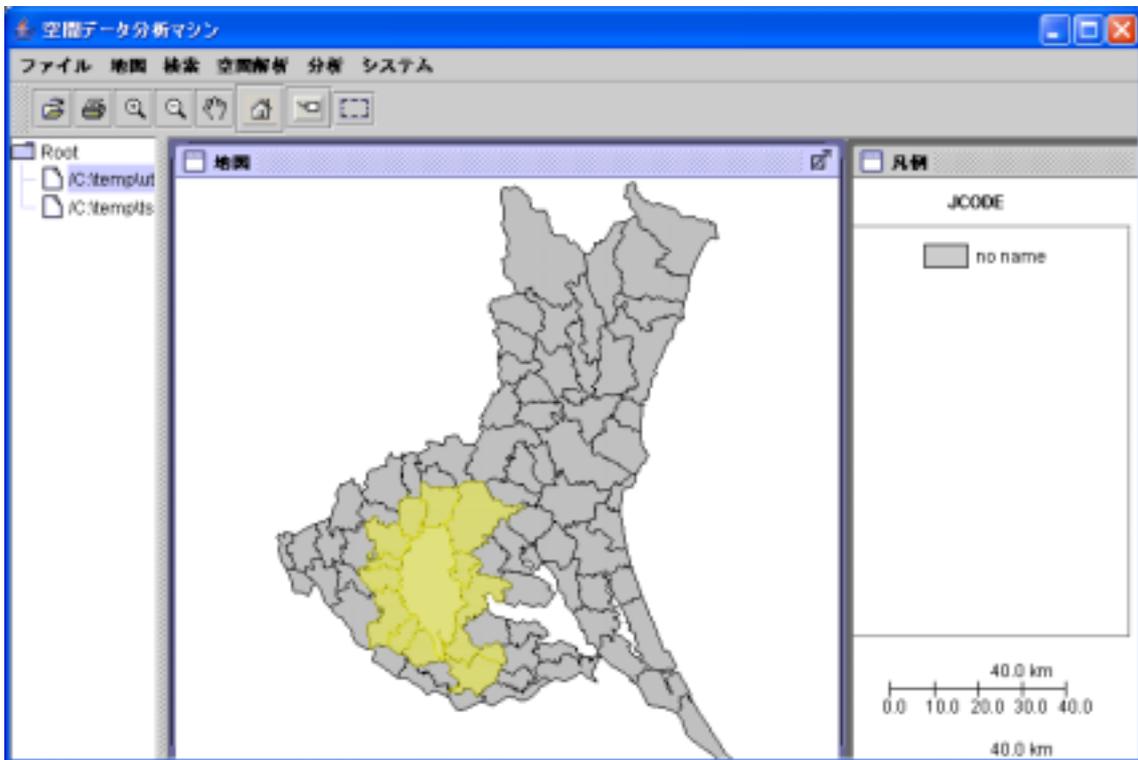


図 14 「intersects」の例（ターゲット「茨城県」，サブジェクト「つくば市」）

表3 空間検索でサポートされているオペレーター

オペレーター名	機能
contains	ターゲット・レイヤーのフィーチャーをサブジェクト・レイヤーのフィーチャーが含む場合、該当のターゲット・レイヤーのフィーチャーを選択
contained	ターゲット・レイヤーのフィーチャーをサブジェクト・レイヤーのフィーチャーが含む場合、該当のサブジェクト・レイヤーのフィーチャーを選択
cross	異なる図形タイプのターゲット・レイヤーのフィーチャーとサブジェクト・レイヤーのフィーチャーが交差して共通部分をもつ場合、該当のターゲット・レイヤーのフィーチャーを選択
disjoint	ターゲット・レイヤーのフィーチャーとサブジェクト・レイヤーのフィーチャーが交差も接触もしない共通部分をもたない場合、該当のターゲット・レイヤーのフィーチャーを選択
equals	ターゲット・レイヤーのフィーチャーとサブジェクト・レイヤーのフィーチャーが等しい場合、該当のターゲット・レイヤーのフィーチャーを選択
intersects	ターゲット・レイヤーのフィーチャーとサブジェクト・レイヤーのフィーチャーが交差して共通部分をもつ場合、該当のターゲット・レイヤーのフィーチャーを選択
overlaps	同じ図形タイプのターゲット・レイヤーのフィーチャーとサブジェクト・レイヤーのフィーチャーが境界部で接し、かつ、重なり合う場合、該当のターゲット・レイヤーのフィーチャーを選択
touches	ターゲット・レイヤーのフィーチャーとサブジェクト・レイヤーのフィーチャーの境界が接する場合、該当のターゲット・レイヤーのフィーチャーを選択
within	ターゲット・レイヤーのフィーチャーをサブジェクト・レイヤーのフィーチャーが含む場合、該当のターゲット・レイヤーのフィーチャーを選択

属性検索では、論理式に基づき、条件に合致する検索が行える。「条件検索」ウィンドーで、地図・項目・論理式を選択した後、値を入力し、「適用」ボタンを押す(図15)。すると、図16の地図(世帯数が2万以上)が描画され、設定された条件に合う統計単位が選択される。



図 15 「条件検索」のウィンドー

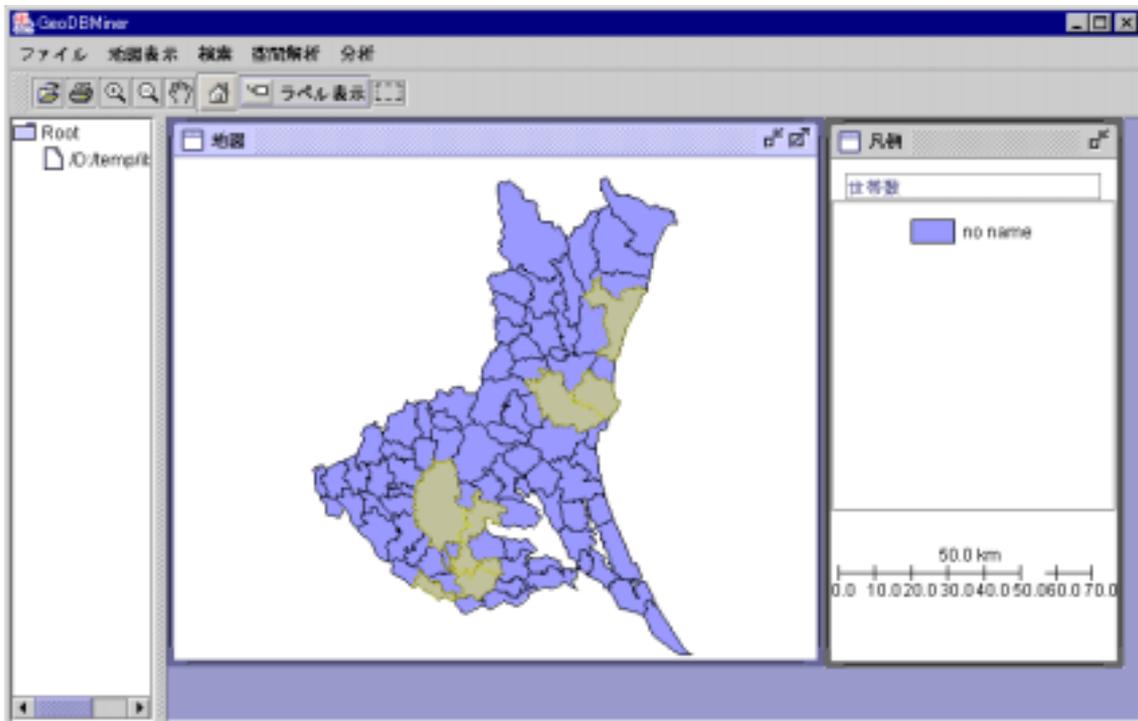


図 16 「条件検索」実行後の結果

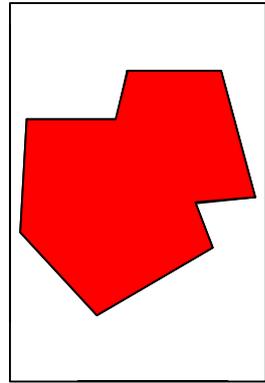
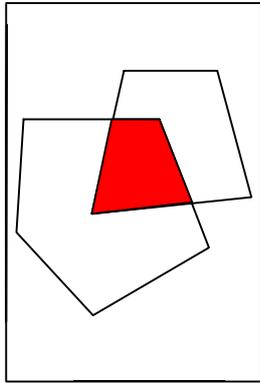
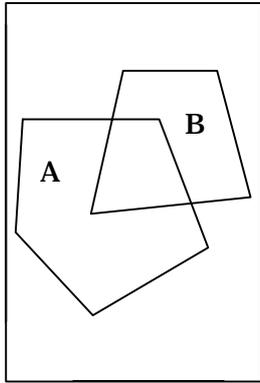
## 6.3 空間解析

### 6.3.1 オーバーレイ

SDAM は、表 4 のオーバーレイ解析を可能にする(図 17)。「空間解析」メニューより、「オーバーレイ解析」サブメニューを選択すると、「オーバーレイ解析」ウィンドーが表示される(図 18)。図 19 は「intersection」の実行例である。

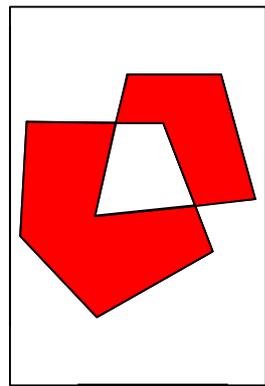
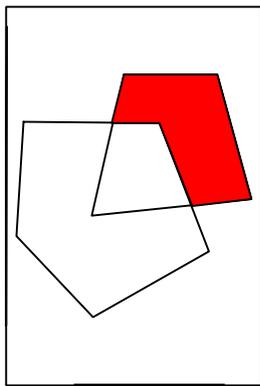
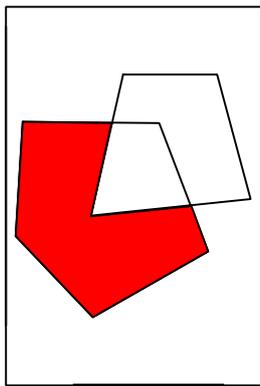
表 4 オーバーレイ解析のオペレーター

オペレーター名	説明
intersection	ターゲットレイヤーとサブジェクトレイヤーの共通部分
difference	ターゲットレイヤーのサブジェクトレイヤーと共通しない部分
symDifference	ターゲットレイヤーとサブジェクトレイヤーの共通しない部分
union	ターゲットレイヤーとサブジェクトレイヤーを結合



A. intersection (B)

A. union (B)



A. difference (B)

B. difference (A)

A. symDifference (B)

図 17 オーバーレイ解析の概念図 (Vivid Solutions, 2002)



図 18 「オーバーレイ解析」のウィンドウ

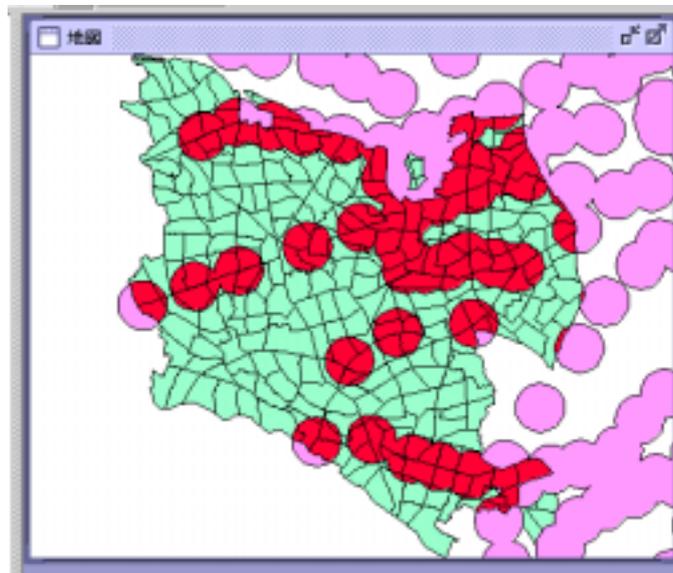


図 19 「intersection」の例（ターゲット駅の 500m のバッファー，サブジェクト世田谷区）

### 6.3.2 バッファー

これは、ポイント・ライン・ポリゴンからなる図形から等しい距離にある領域を確定する手法である。「空間解析」メニューで、「バッファー」メニュー項目を選択すると、「バッファー」ウィンドウが表示される。次に、地図名およびバッファー・サイズ（ポイントの場合はバッファ

の半径)を選択する。さらに、「保存ファイル」ボタンを押して、保存先のディレクトリとファイル名(.shpの拡張子はタイプしない。タイプした場合エラーになり、地図が表示されない)を入れて、最後に「実行」ボタンを押す。実行結果の例を図20(ポイントのバッファー)と図21(ラインのバッファー)に示す。

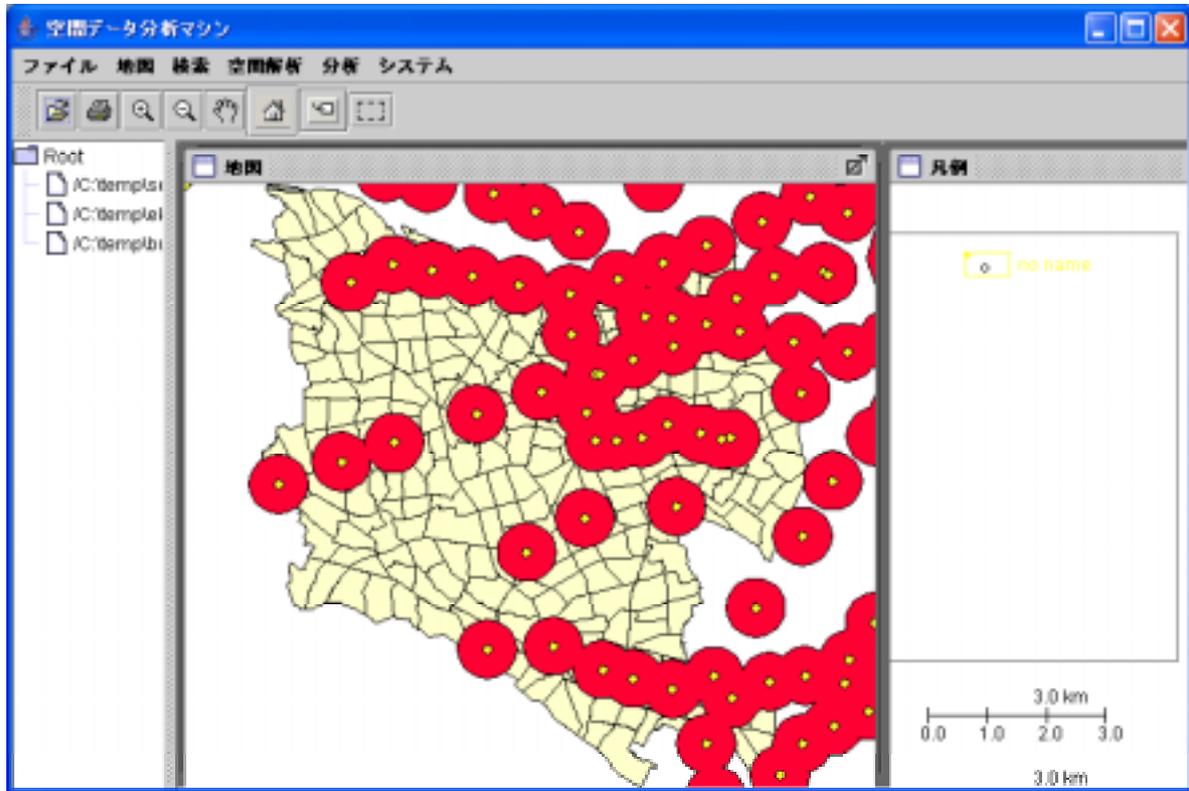


図20 ポイントのバッファーの実行結果

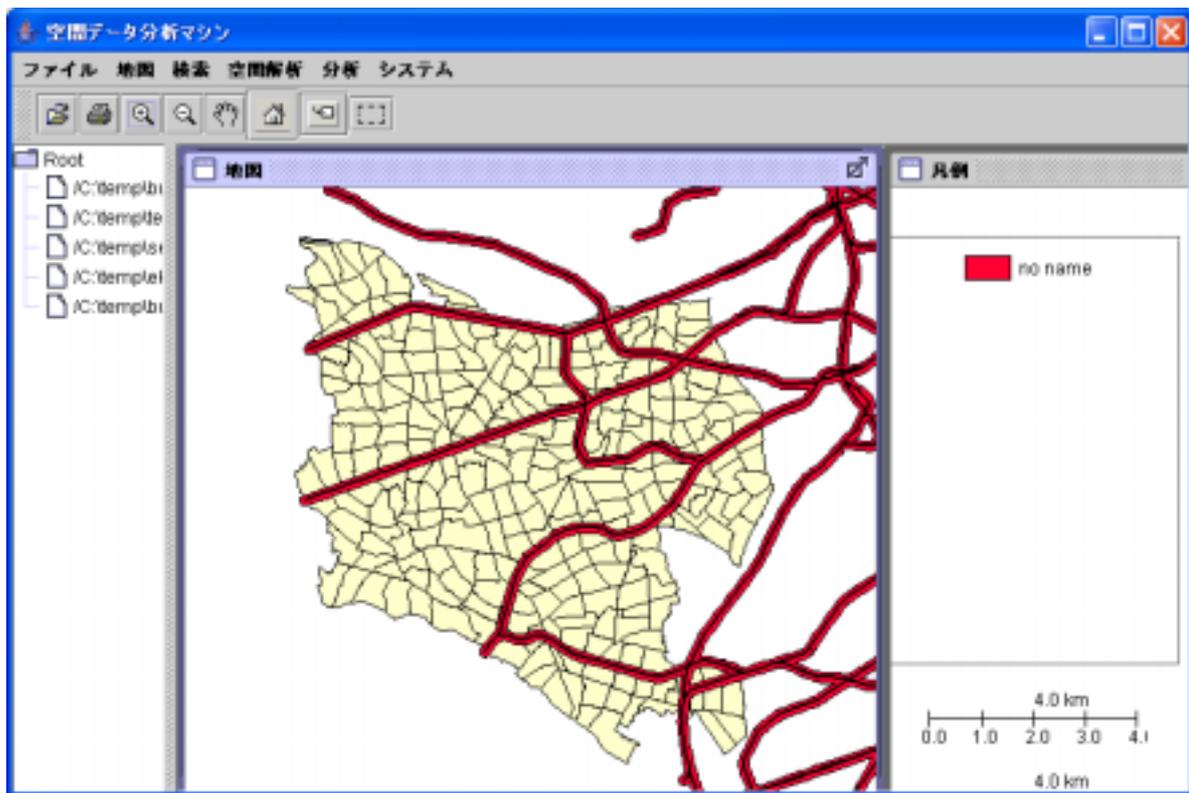


図 21 ラインのバッファの実行結果

### 6.3.3 ユニオン

「ユニオン」ウィンドーで、保存先のディレクトリとファイル名をタイプして、「実行」ボタンを押すと、表示されている地図の各図形（フィーチャー）がひとつの図形として結合（ユニオン）される（図 22, 23）。

さらに、新規に結合（ユニオン）されるレイヤーに、フィーチャーリストで属性を選択し、加工方法（合計・平均・最初のレコードの値）を指定して、挿入することもできる。



図 22 キーフィールド (CITY2) の指定

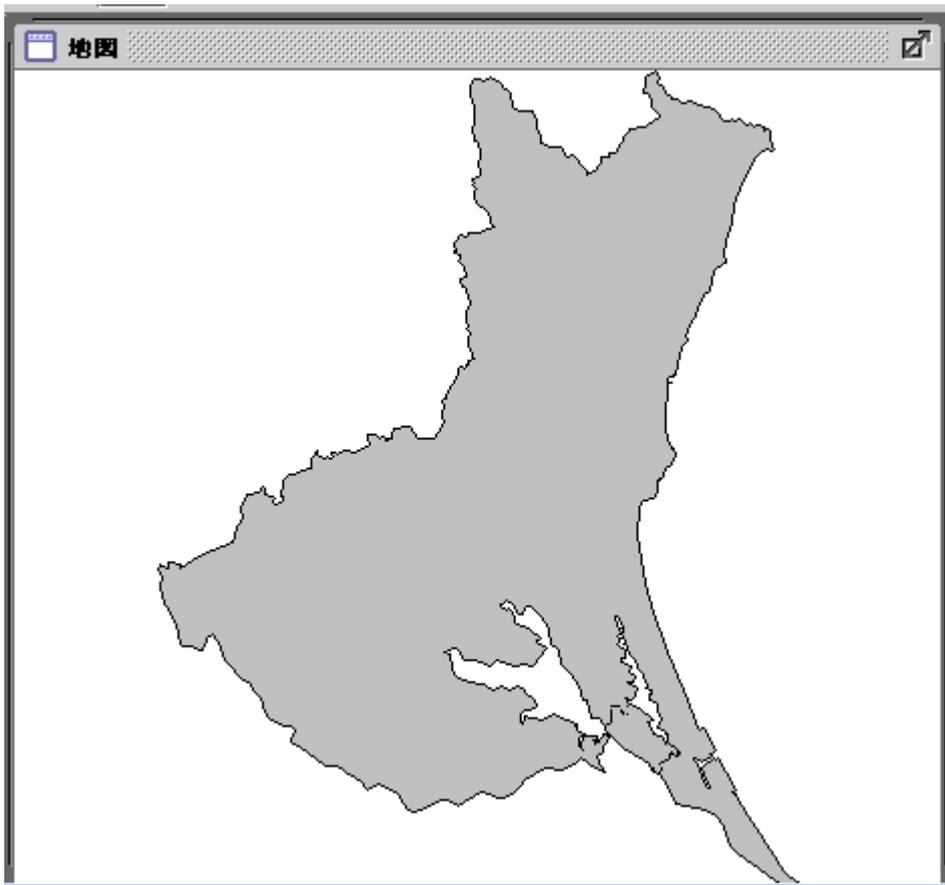


図 23 キーフィールド (CITY2) を指定したディゾルブの例 (市・町・村で分類)

#### 6.3.4 TIN (不規則三角形網)

TIN とは、互いに重なり合わない連続的な小三角形の集合で地表面の形態を近似するモデルである (地理情報システム学会・地理情報科学用語集, 2000)。SDAM では、ポイントまたはポリゴンの中心点をベースに TIN を生成できる (図 24 はポリゴンの例)。

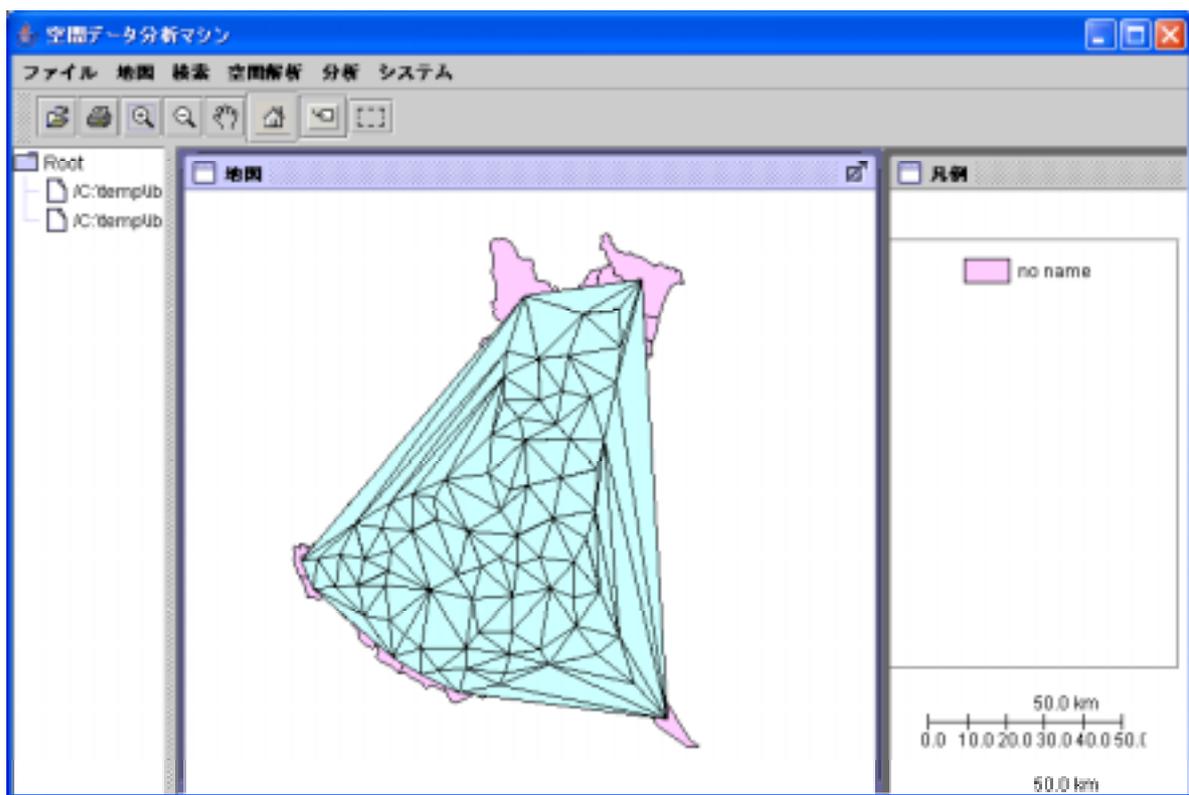


図 24 「TIN」の表示

### 6.3.5 ポロノイ (ティーセン)

平面上に分布する  $n$  点のうち、任意の点と他の点を結ぶ直線の垂直 2 等分線を順次引くことによって描かれた、点のまわりの多角形をティーセン図形といい、この図形を作成する手続きをポロノイ (ティーセン) 分割と呼ぶ。この手法は、勢力圏 (たとえば商圈や駅勢力圏) の理論的設定などに用いられる (図 25)。

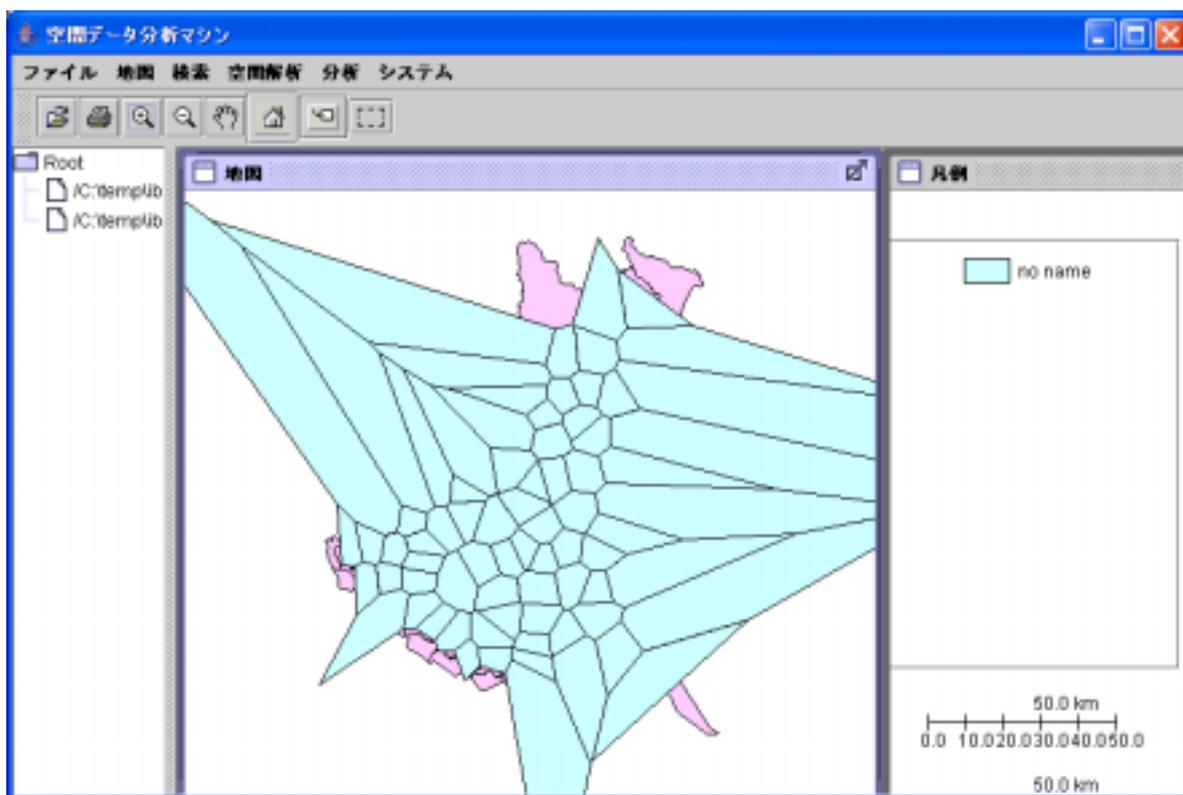


図 25 ポロノイ図の表示

### 6.3.6 凸包

「空間解析」メニューで、「凸包」メニュー項目を選択すると、「凸包」ウィンドーが表示される。地図名を選択し、「実行」ボタンを押すと、凸包が描画される（図 26）。

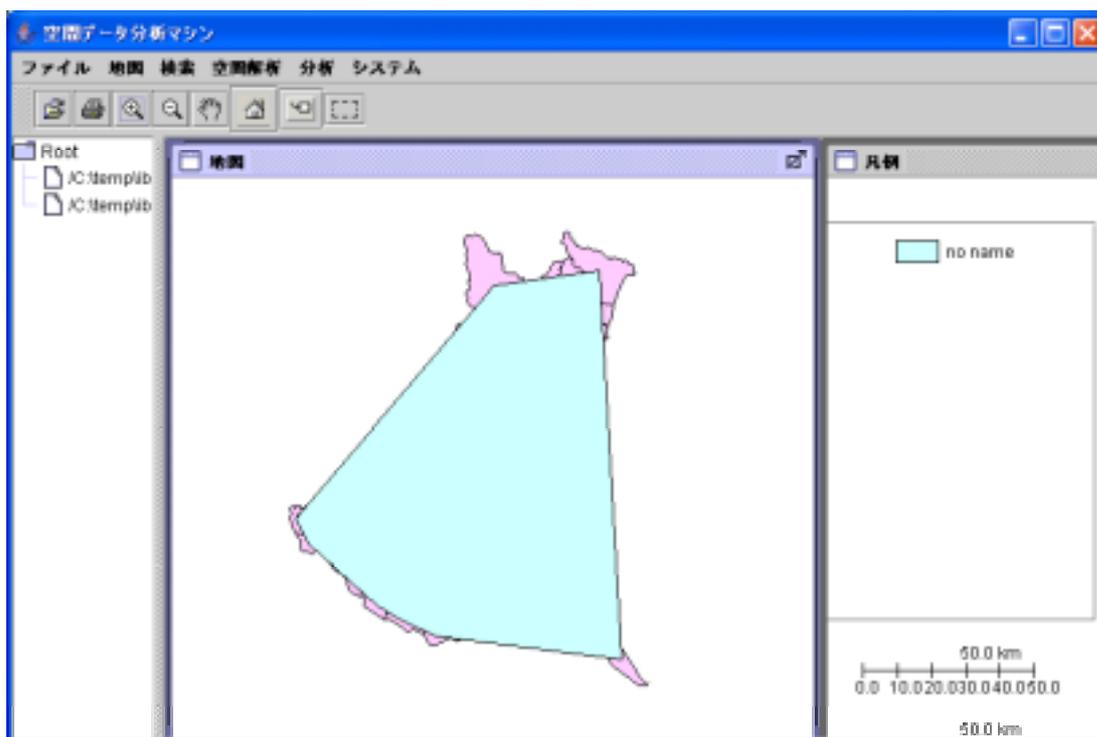


図 26 凸包の表示

## 6.4 地域分析

計量地理学の発達とともに、地域分析のための手法が数多く開発されてきた（村山，1990；村山・尾野，1993）。SDAM を利用すると、基本統計量の導出、多変量解析、ポイント・パターン分析、空間的自己相関分析、空間的相互作用分析、地理加重回帰分析などが行える。

### 6.4.1 基本統計量

総数、最小値、最大値、合計、平均値、標準偏差、分散の統計量が得られる。「基本統計」ウィンドーにおいて、「地図選択」と「項目選択」で基本統計量を導出したいデータを指定すると、右側のテーブルに結果が表示される。

### 6.4.2 多変量解析

#### (1) 回帰分析

単回帰と重回帰の分析が実現できる。図 27 において、「地図選択」で当該地図を選択し、従属変数を選択メニューより選ぶ。独立変数を「独立変数用項目選択」リストからマウスでクリックした後、「選択」ボタンを押すと、「選択独立変数名リスト」に、選択された項目が追加される。独立変数は、複数個を選択できる。選択された変数を取り消すには、「選択独立変数名リスト」内の選択項目をマウスで選択し、「取り消し」ボタンを押せばよい。従属変数および独立変数を選択した後、「適用」ボタンを押すと、計算が実行される。しばらくして、図 28 の「回帰分析の計算結果」ウィンドーが表示され、パラメータ、決定係数および検定の結果が得られる。このウィンドーで、上部にあるタブをマウスでクリックすると、図 29 の予測値と残差のテーブルが得られる。その下部にある「予測値地図表示」ボタンを押すと、予測値のコロプレス図が表示され

る。「残差地図表示」ボタンを押すと、残差のコロプレス図が得られる（図 30）。



図 27 「回帰分析」のウィンドウ



図 28 「回帰分析の計算結果」のウィンドウ

回帰分析の計算結果

計算結果 予測値・残差

実測値	予測値	残差
718.0	630.9304666978004	87.0695333021996
3546.0	4022.555914966457	-476.55591496645684
4138.0	4393.5315731165165	-255.53157311651648
2510.0	2729.2245892297947	-219.22458922979467
3413.0	3617.0598960197703	-204.05989601977035
731.0	694.769490088824	36.230509911176
7351.0	7524.008127550412	-173.0081275504117
5554.0	5705.068842560989	-151.06884256098874
7063.0	7157.406624706774	-94.40662470677398
5686.0	5579.400542811622	106.59945718837844
7299.0	6333.5285617215095	965.4714382784905
17154.0	14981.351822931458	2172.6481770685423
17728.0	22902.237760379547	-5174.237760379547
35455.0	28470.773906282106	6984.226093717894
941.0	891.4882584641263	49.51174153587374
1966.0	1735.8184530172562	230.18154698274384
3590.0	3343.970740402614	246.02925959738604
4656.0	4764.507231266575	-108.50723126657522
2104.0	2374.445127754866	-270.445127754866
272.0	-115.98610697717504	387.98610697717504
451.0	142.56193775647034	308.43806224352966
649.0	614.4978291952963	34.502170804703724
2654.0	2567.4990633058687	86.50093669413127
1143.0	1169.660891869827	-26.660891869827083

予測値地図表示      残差地図表示

図 29 予測値と残差のテーブル

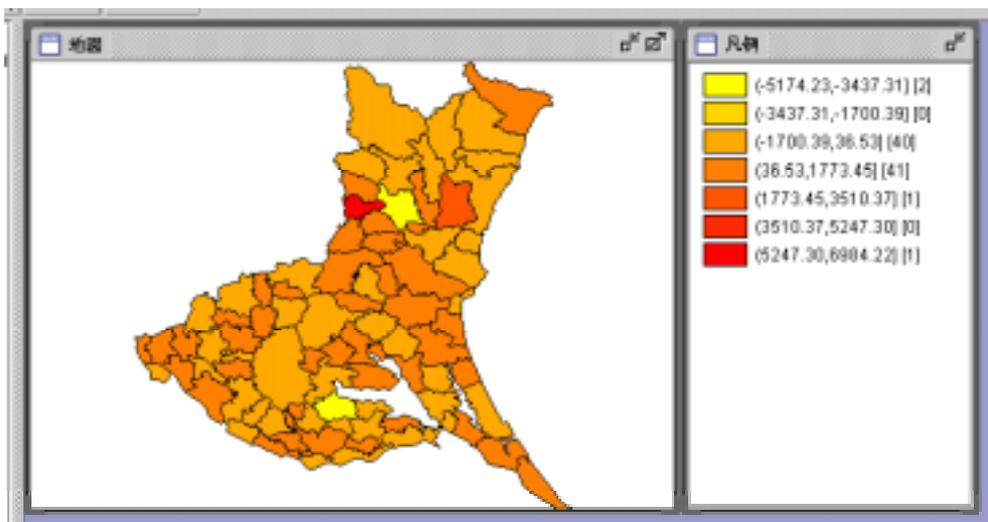


図 30 残差のコロプレス図

## (2) 因子分析

因子分析は、互いに相関のある変数を潜在的な少数個の因子に縮約する手法である。「因子分析」ウィンドーにおいて、左側の「変数選択」リストより、因子分析に用いる変数をマウスで選ぶ。「選択」ボタンを押し、右側にある「選択変数リスト」に入れる。「適用」ボタンを押すと、因子分析（主因子法・バリマックス回転）が実行される。計算が終了すると、「因子分析の計算結果」ウィンドーに、固有値および累積寄与率が表示される。このウィンドーにおいて、上にある「因子得点」タブをマウスでクリックすると、因子負荷量が得られる（第9因子まで）。ここで、「クラスター図表示」ボタンを押すと、クラスター分析（Ward法）の計算が始まる。結果が「クラスター分析」ウィンドーに示される。このウィンドーで、「デンドログラム表示」ウィンドーを押すと、図31のデンドログラムが現れる。デンドログラムは、マウスを使って、拡大・縮小が可能である。元の表示に戻すには、「再描画」ボタンを押す。

図31中のグループ化リストで、グループ数を指定する。デンドログラムがグループ化され（図32）、地図が表示される（図33）。グループ名には、自動的にアルファベットが割り当てられる。「クラスター分析」ウィンドーの「グループ名」ボタンを押すと、図34の「グループ名指定」ウィンドーが表示される。新しいグループ名を入力して、「クラス名変更」ボタンを押すと、グループ名が変更される。

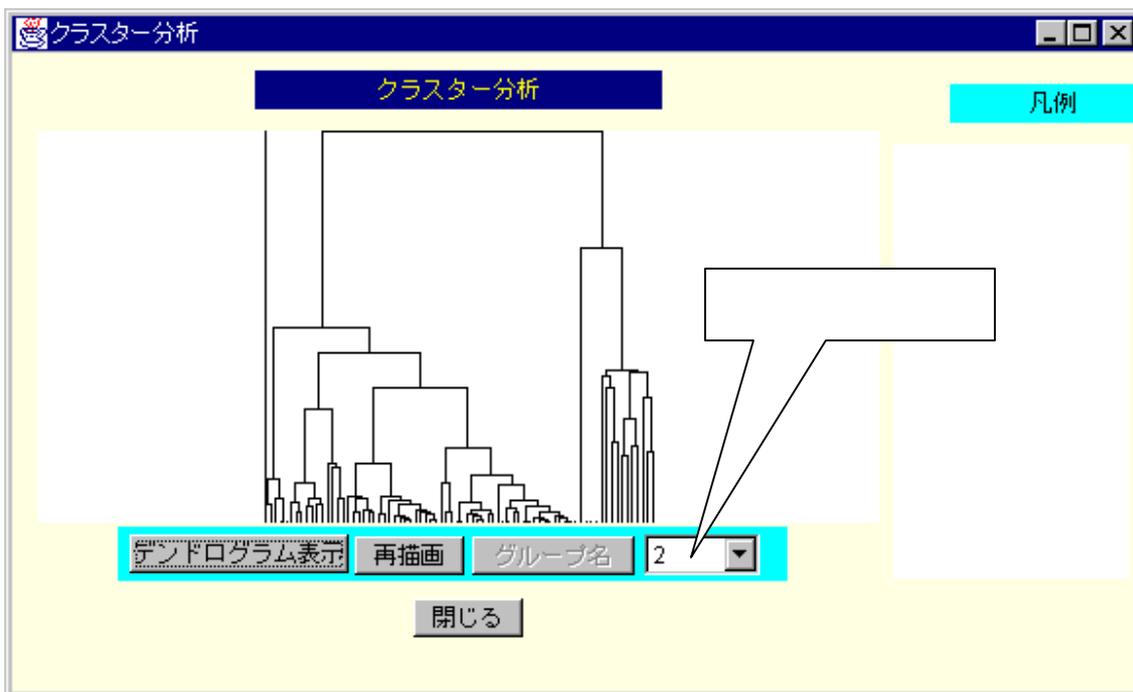


図 31 デンドログラムの表示

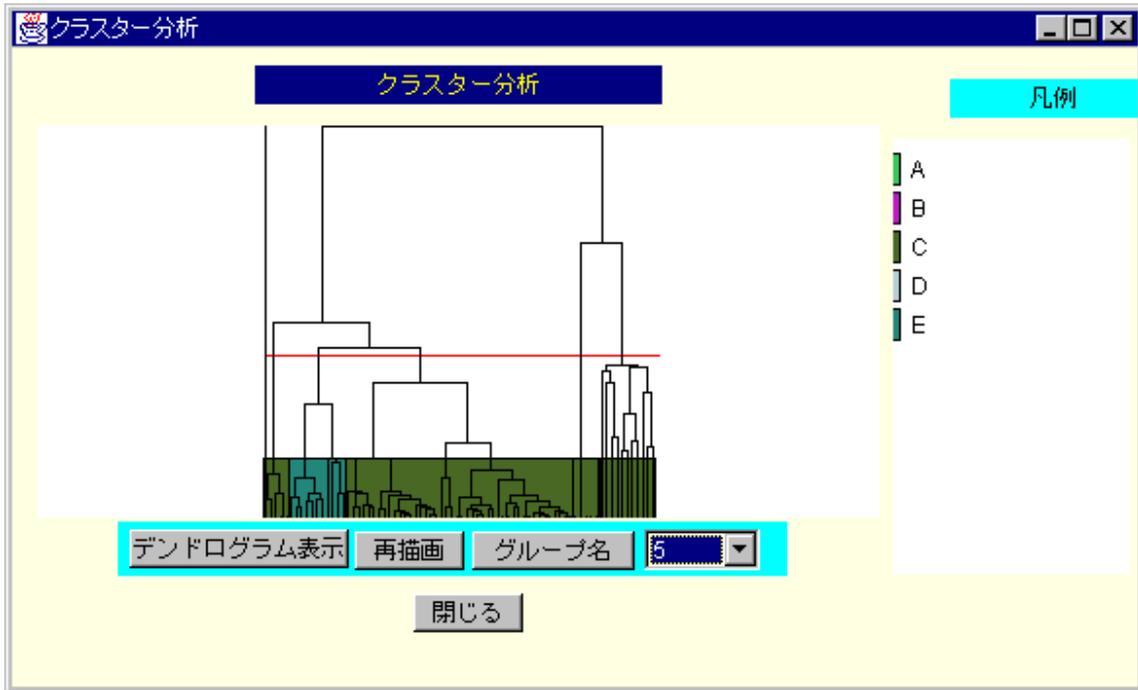


図 32 クラスタ分析によるグループ化の結果

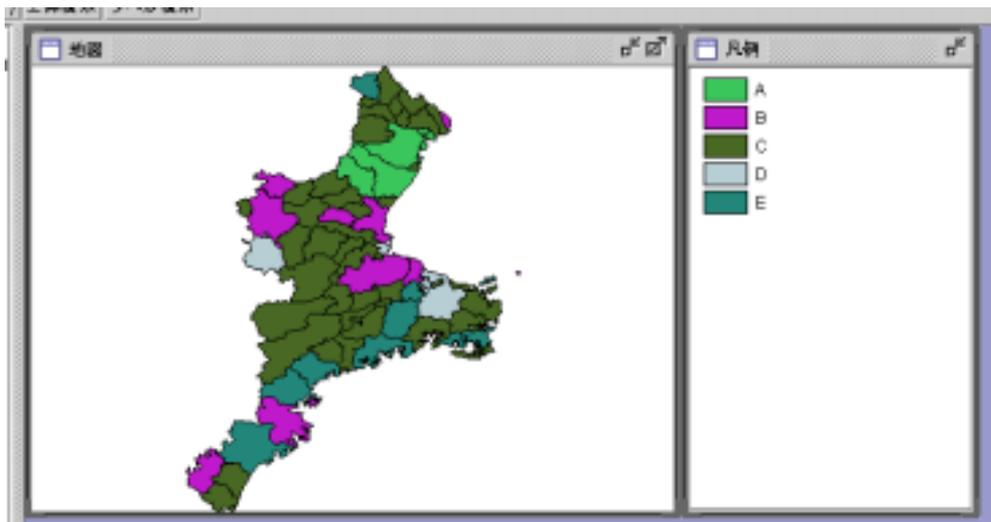


図 33 クラスタ分析によるグループ化の地図表示

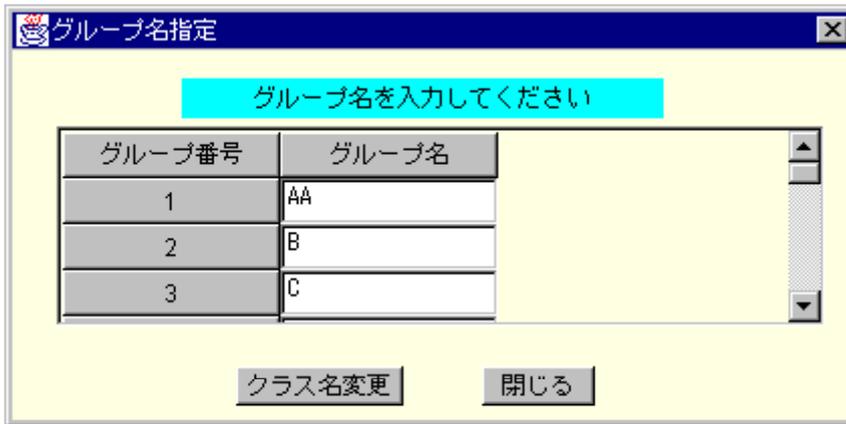


図 34 「グループ名指定」のウィンドー

### (3) K Means 非階層型クラスター分析

非階層型クラスター分析においては、予め、クラスター数を指定する必要がある。本システムは、代表的な K Means 法をサポートする。階層型クラスター分析に比べて、この方法は画像処理など大量データの分析に向いている。「非階層型クラスター分析」ウィンドーにおける左の「変数選択」リストより、分析する変数を選択し、グループ数（デフォルトのグループ数は 2）を定める。演算が終了すると、「地図表示」ウィンドーに、指定されたグループ数で分類された地図が表示される。

#### 6.4.3 探索的空間分析 (ESDA)

この機能はデータマイニングに用いられる。「散布図」ウィンドー（図 35）において、X 軸と Y 軸の変数を選んで、「実行」ボタンを押すと、散布図が表示される。他の地図の属性を表示したければ、「地図名選択リスト」より選択する。表示された散布図で、マウスを任意のポイントに重ねると、それと対応する地図の統計単位が内部を赤く塗りつぶされる。逆に地図側で、マウスが重なった統計単位は赤く塗りつぶされ、該当する散布図のポイントも赤くなる。マウスをドラッグして、任意の点を選択すると、それに対応する地図の統計単位が選択される（図 36，図 37）。

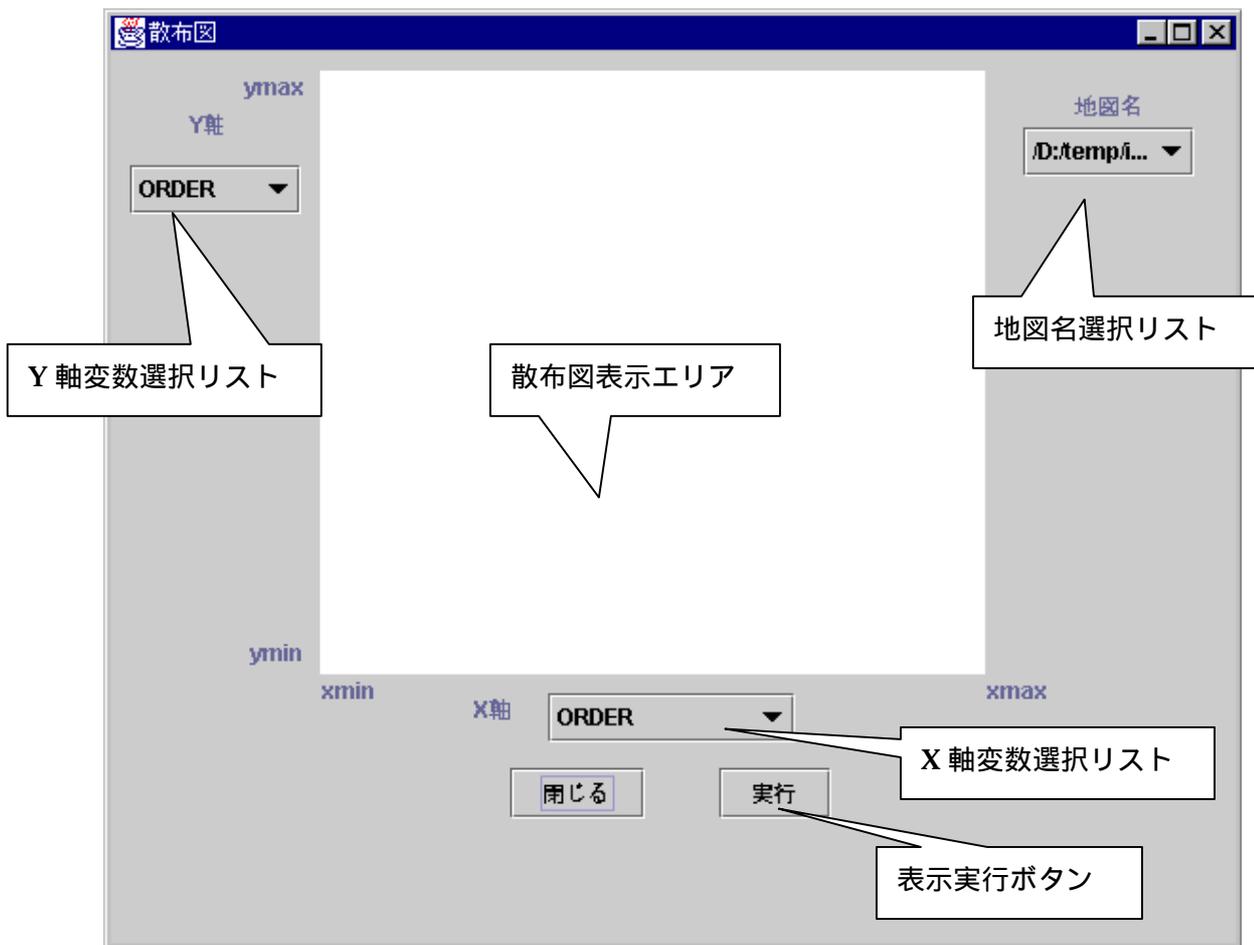


図 35 「散布図」のウィンドー

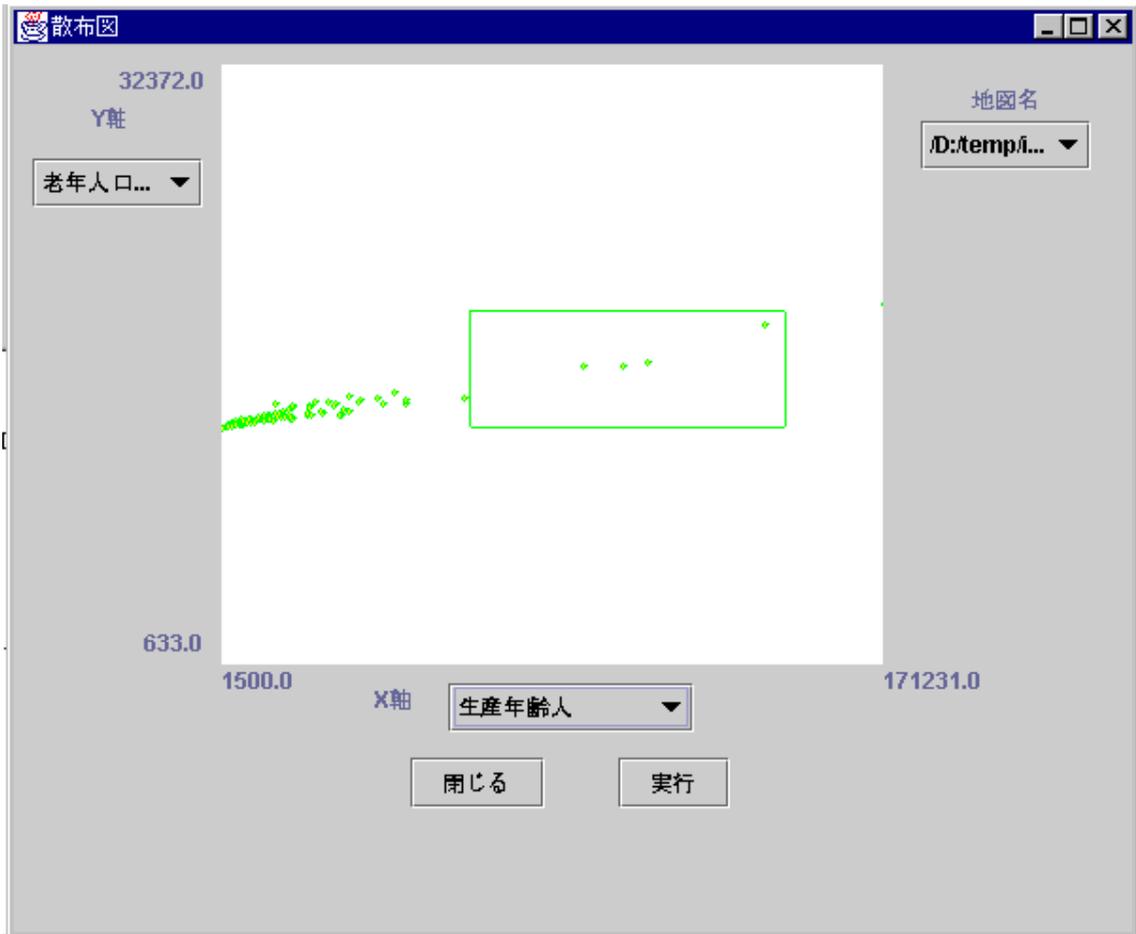


図 36 散布図の選択部分

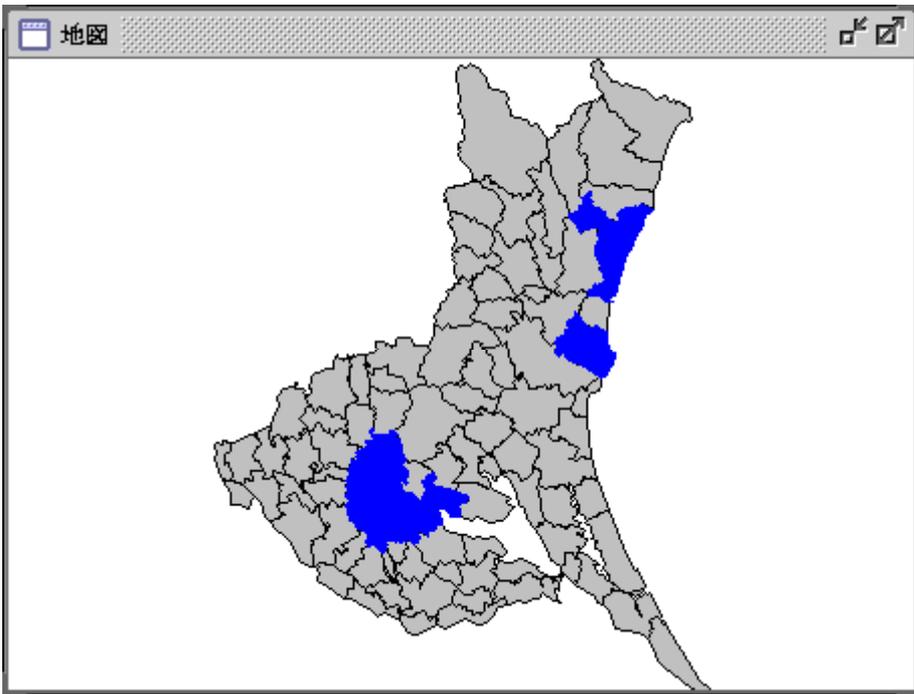


図 37 地図の散布図で選択された部分に対応する統計単位

#### 6.4.4 ポイント・パターン分析

ポイント・パターン分析とは、マクロレベルでは都市や集落，交通結節点など，ミクロレベルでは駅，公共施設，家屋などを点としてみなし，そのばらつき度合い（分布）を判定する手法である。SDAM では、「方格法」，「最近隣法」，「K 関数法」の 3 つの手法がサポートされ，ランダム，凝集，均等の 3 つの型が識別される。

##### (1) 方格法

点が分布する範囲に方格をかぶせ，各方格に含まれる点をカウントし，各方格の点の度数分布に最も適合する理論的な分布を与える関数を見いだす方法である。「分析」メニューから，「ポイント・パターン分析」を選択する。サブメニューより「方格法」メニュー項目を選択すると，「方格法」ウィンドーが表示される。「方格法」ウィンドーで，ターゲットの地図名，グリッド名，グリッド・セルサイズ，グリッド分割数，グリッド分割数を指定する。「適用」ボタンを押すと，「方格法」ウィンドーに実行結果が表示される（図 38）。

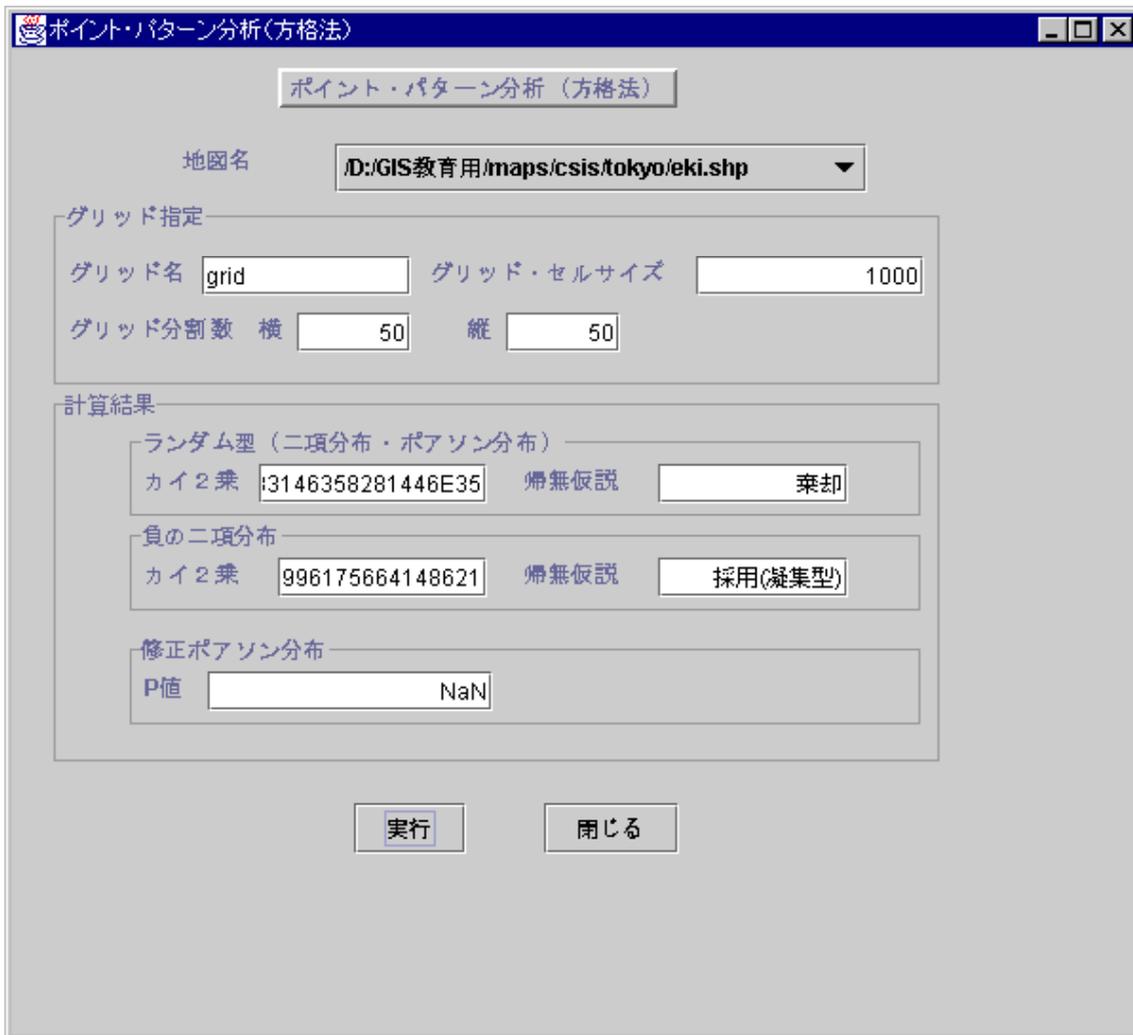


図 38 「方格法」のウィンドーと実行結果

## (2) 最近隣法

これは点間の距離を利用して、点の分布状態を把握する手法である。ターゲットの地図名を指定し、「適用」ボタンを押すと、「最近隣法」ウィンドーに実行結果が表示される。「最近隣法」ウィンドーで、「最近隣点表示」ボタンを押すと、最近隣点が表示される（図 39）。

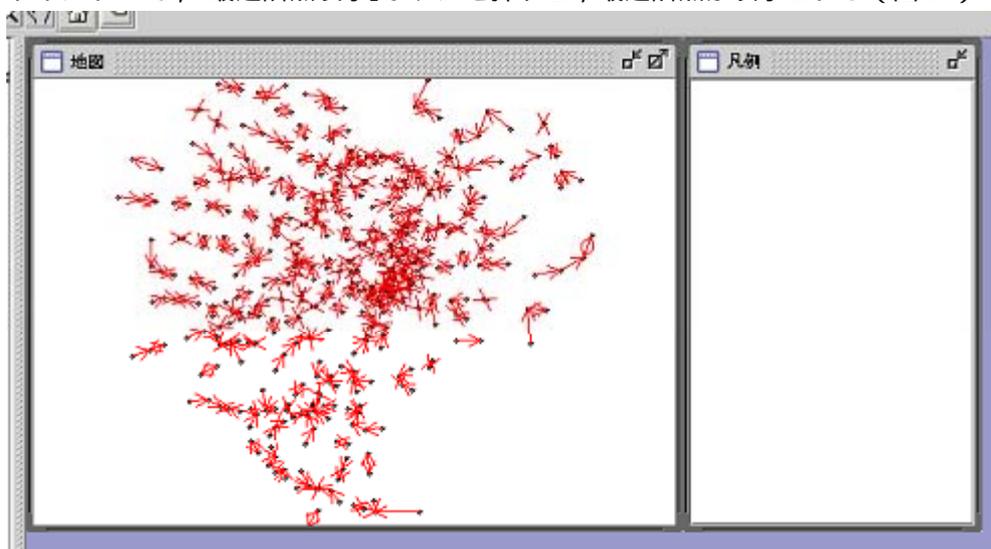


図 39 「最近隣点表示」

## (3) Ripley の K 関数法

「K 関数法」メニュー項目を選択すると、図 40 が表示される。ターゲットの地図名を選択し、「距離」および「増分」を指定する。実行後、「メッセージ」フィールドに「計算終了」というメッセージがでるまで待つ。「グラフ表示」ボタンを押すと、図 41 のグラフが表示される。

図 40 「K 関数法」のウィンドー

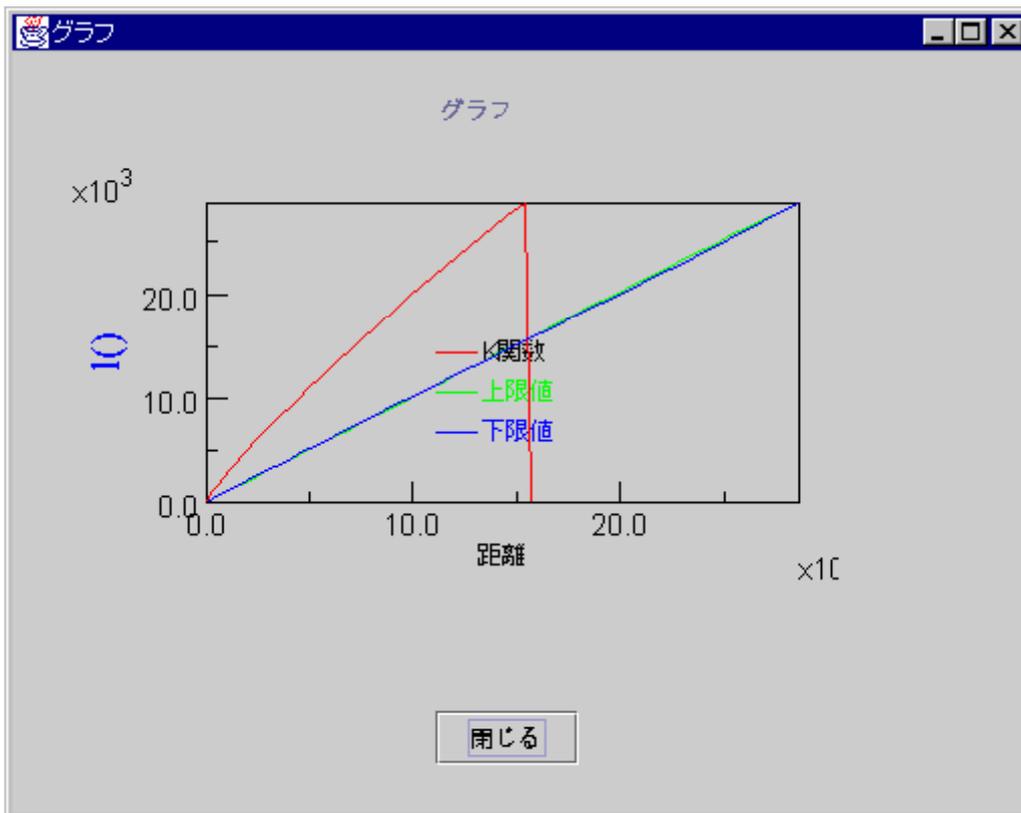


図 41 「K 関数法」のグラフ

#### 6.4.5 グローバル空間的自己相関

##### (1) グローバル・モラン統計量

グローバル・モラン統計量は空間的自己相関分析の中で、最もよく使用される統計量である（杉浦，1986）。

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{W \sum_{i=1}^n (x_i - \bar{x})^2} \quad \dots \text{式(1)}$$

ここで、 $n$  は、観測地点の総数、

$w_{ij}$  は、観測地点  $i$ 、 $j$  間の結合関係をあらわすウェイト値、

$x_i$  は、地点  $i$  の観測値、 $\bar{x}$  は観測値  $x$  の平均値。

$$W = \sum_{i=1}^n \sum_{j=1}^n w_{ij} \quad \dots \text{ウェイト値の合計 ( } i = j \text{ )}$$

Moran の検定は以下の式で行われる。

$$Z = \frac{I - E(I)}{\sqrt{\text{Var}(I)}}$$

ただし、 $E(I)$ は、 $I$ の期待値、 $\text{Var}(I)$ は $I$ の分散で、以下のように定義される。

$$E(I) = \frac{-1}{(n-1)}$$

$$\text{Var}(I) = E(I^2) - E(I)^2$$

有意性の仮定には正規性仮定とランダム化仮定の2つがある。正規性仮定の場合、期待値は、

$$E(I^2) = \frac{n^2 S_1 - n S_2 + 3W^2}{W^2(n^2 - 1)}$$

である。ここで、

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2 \quad i \neq j$$

$$S_2 = \sum_{i=1}^n \left( \sum_{j=1}^n w_{ij} + \sum_{j=1}^n w_{ji} \right)^2 \quad i \neq j$$

である。ランダム仮定では、以下のようになる。

$$E(I^2) = \frac{n((n^2 - 3n + 3)S_1 - nS_2 + 3W^2) - k((n^2 - n)S_1 - 2nS_2 + 6W^2)}{(n-1)(n-2)(n-3)W^2}$$

ここで、 $k$ は $x_i$ の尖度である<sup>3</sup>。

各空間パターンとグローバル・モラン統計量  $I$  との間には表5の関係がある (Lee & Wong, 2000)。

表5 空間パターンとグローバル・モラン統計量  $I$  との関係

空間パターン	グローバル・モラン統計量 $I$
凝集型 (空間クラスター)	$I > E(I)$
ランダム型	$I \sim E(I)$
均等型	$I < E(I)$

## (2) グローバル・ギアリ統計量

グローバル・ギアリ統計量  $c$  は式(2)で求められる。

$$c = \frac{(n-1) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{2W \sum_{i=1}^n (x_i - \bar{x})^2} \quad \dots \quad \text{式(2)}$$

検定方法はグローバル・モランと同様で、Z スコアを求める。期待値は正規仮定およびランダム仮定ともに 1 である。

$$E(c) = 1$$

正規仮定の場合、分散は以下ようになる。

$$\text{Var}(c) = \frac{[(2S_1 + S_2)(n-1) - 4W^2]}{2(n+1)W^2}$$

ランダム仮定の場合、分散は以下ようになる。

$$\begin{aligned} \text{Var}(c) &= \frac{(n-1)S_1[n^2 - 3n + 3 - (n-1)k]}{n(n-2)(n-3)W^2} \\ &\quad - \frac{(n-1)S_2[n^2 + 3n - 6 - (n^2 - n + 2)k]}{4n(n-2)(n-3)W^2} \\ &\quad + \frac{\{(n^2 - 3) - (n-1)^2 k\}}{n(n-2)(n-3)} \end{aligned}$$

ここで、

$$S_1 = \frac{\sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2}{2}$$

$$S_2 = \sum_{i=1}^n (w_i + w_i)^2$$

$$k = \frac{n \sum_{i=1}^n (x_i - \bar{x})^4}{\left( \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}$$

表 6 に各空間パターンとグローバル・ギアリ統計量 c との関係を示す (Lee & Wong, 2000)。

表 6 空間パターンとグローバル・ギアリ統計量 c との関係

空間パターン	グローバル・ギアリ統計量 c
凝集型 (空間クラスター)	$0 < c < 1$

ランダム型	$c = 1$
均等型	$1 < c$

### (3) モラン・プロット

モラン・プロットは、Anselin(1995)が考案した ESDA の概念にもとづく。モラン統計量にもとづき散布図を作成し、外れ値(outliers)を考察する手法である。式(1)のグローバル・モラン統計量  $I$  において、 $y$  を属性値  $x$  の偏差のベクトル、 $W$  を行で標準化したウェイトマトリックスとする。

$$I = \frac{n}{W} \cdot \frac{y'W_x y}{y'y}$$

これは  $y$  に対する  $W_y$  (空間ラグ) の単回帰式の斜辺をあらわす式である。これより以下の単回帰式が導かれる。

$$x = \alpha + I W_y \quad \dots \quad \text{式(3)}$$

ここで、 $\alpha$  は切片であり、 $I$  は斜辺をあらわす回帰係数(グローバル・モラン統計量  $I$ ) である。横方向に  $x$ 、縦方向に空間ラグ  $W_y$  を取って散布図を描画し、式(3)で得られる回帰直線を追加したのがモラン・プロットである。本システムでは、表示がみやすいように両軸を標準化している。

### (4) 実行方法

「空間的自己相関」メニューを選択すると、「空間的自己相関」ウィンドーが表示される。グローバル・モラン、グローバル・ギアリ、およびモラン・プロットを同時に計算できる。

「地図選択」で地図を、「項目選択」で項目を、「分類手法選択」で手法を選ぶ。「ウェイト・マトリックスタイプ選択」は、R 言語のパッケージ `spdep` が「クィーン型」に固定されているので、SDAM は「クィーン型」のみ対応とした。「ウェイト値計算法指定」では「バイナリ」か「距離の逆数」のいずれかを、「ウェイト値加工」では「バイナリ」、「標準化(マトリックスの行の値)」、「標準化(マトリックスのすべての値)」のうち、いずれかを選択する。シェープファイルより、ウェイト・マトリックスを求める場合には、演算に時間がかかることがある。計算終了後、「空間的自己相関」ウィンドーの右側に計算結果が表示される(図 42)。

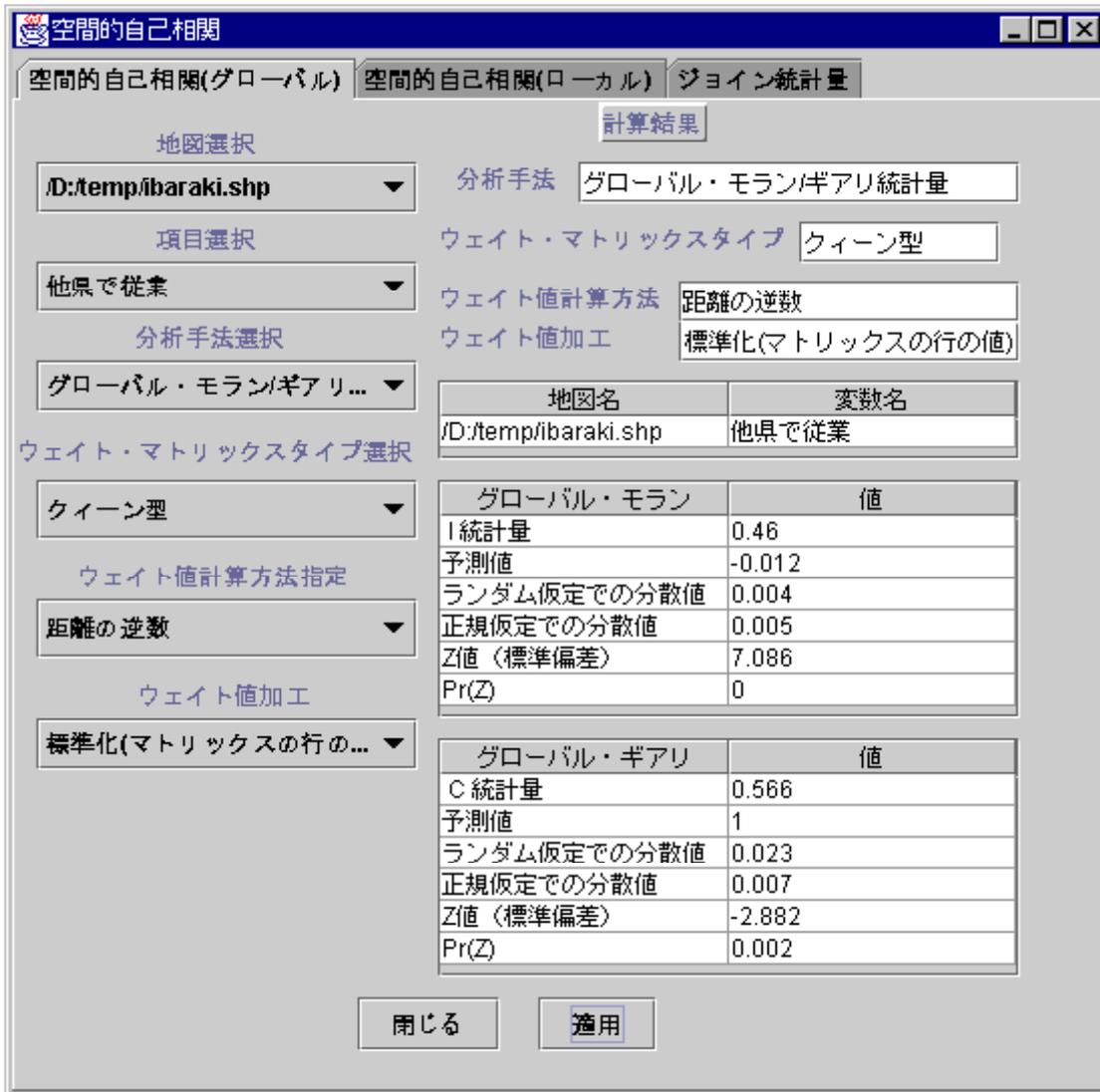


図 42 グローバル空間統計量の計算結果

図 42 により, 表 5 からみて I 統計量は期待値より大きく, 表 6 からみて C 統計量は 1 より小さいので, 当該事象は正の自己相関を示すことが理解できる. 検定については, Z 値は I および C 統計量とともに  $\alpha = 0.05$  (5%水準) の Z 値 (1.96) より大きいので, 帰無仮説は棄却されて有意である. したがって空間クラスターが存在することになる. モラン・プロットも同時に表示される. 図 43 が示すように, 散布図上で, 外れ値をマウスで選択できる. この操作は地図とリンクされており, 地図上で該当統計単位が選ばれる (図 44, 図 45).

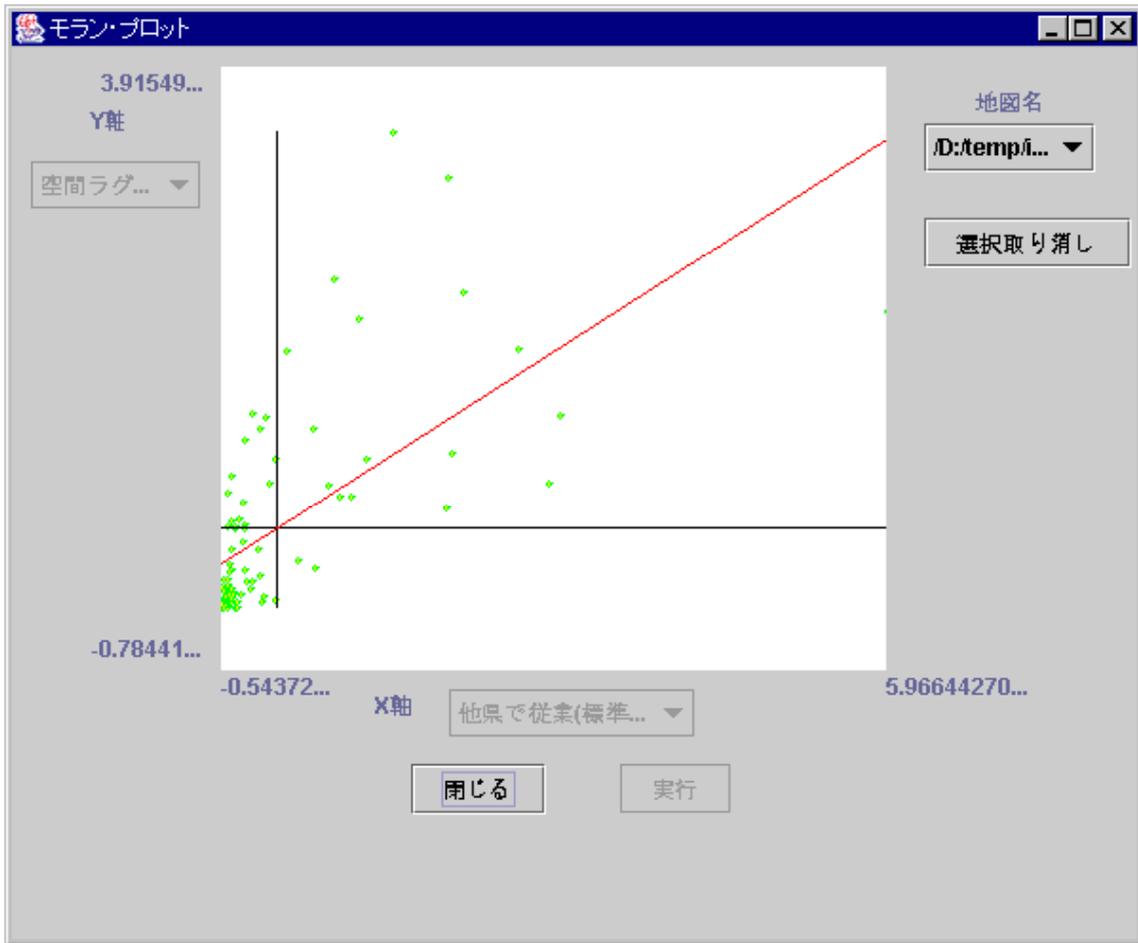


図 43 モラン・プロット

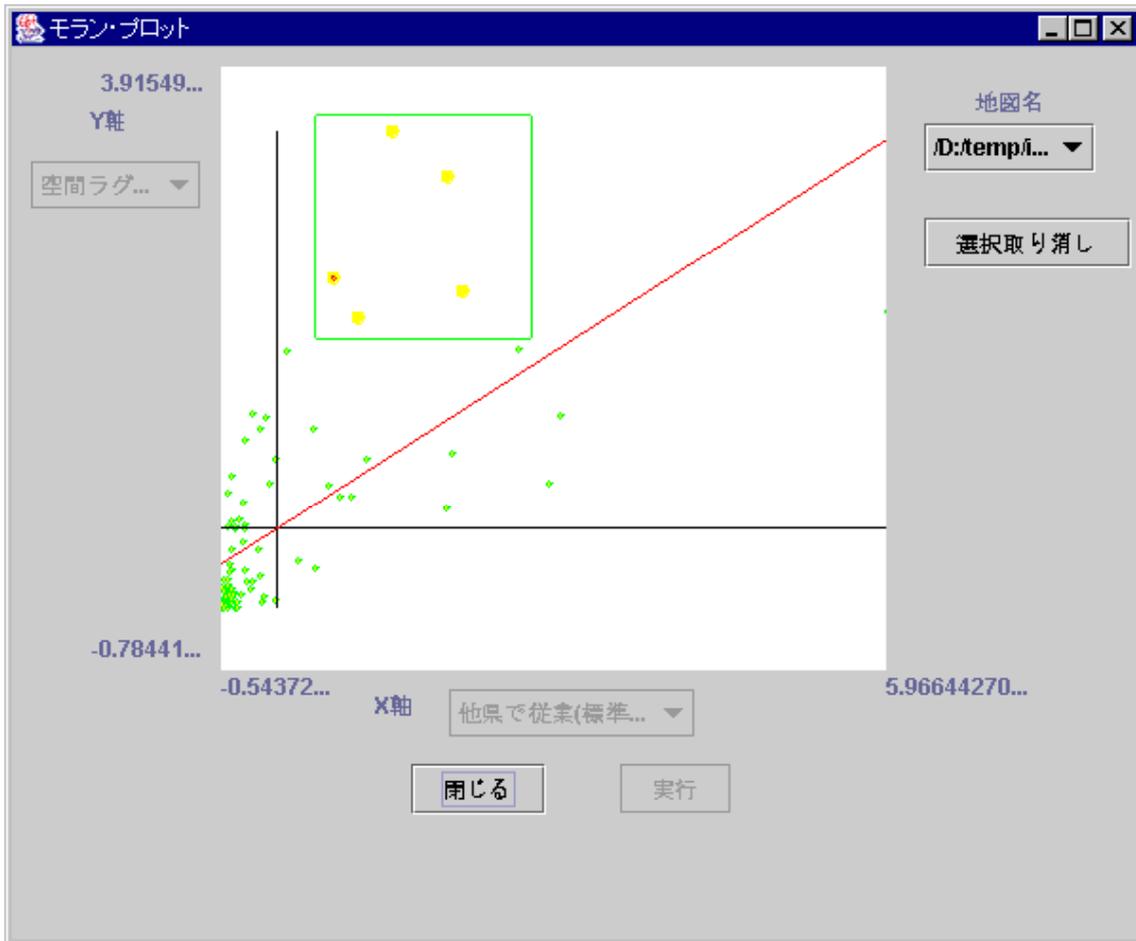


図 44 外れ値選択

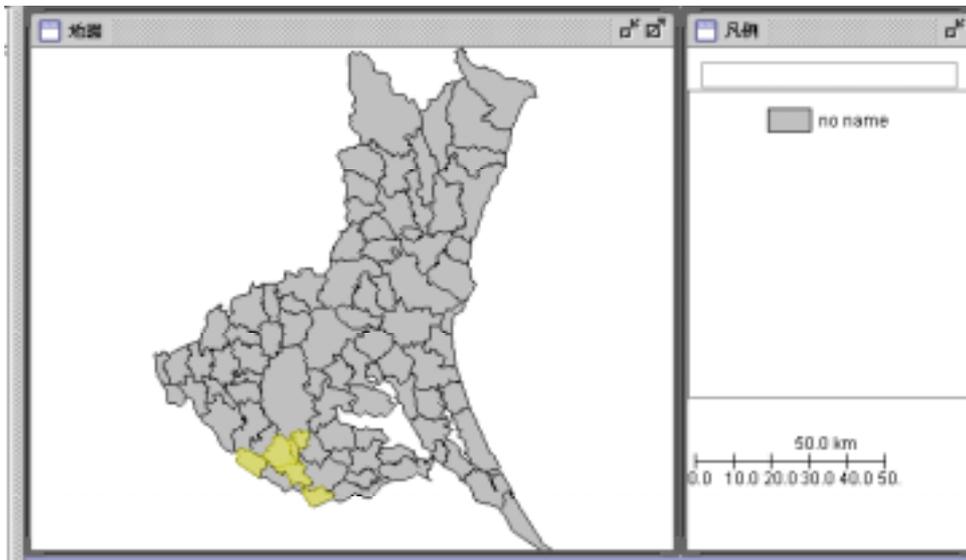


図 45 外れ値に該当する地図表示

#### 6.4.6 ローカル空間的自己相関

##### (1) ローカル・モラン統計量

ローカル・モラン統計量  $I_i$  は以下の式で示される (Anselin, 1995) .

$$I_i = \frac{(x_i - \bar{x}) \sum_{j=1}^n w_{ij} (x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \dots \quad \text{式(4)}$$

期待値は,

$$E_i = -\frac{w_i}{(n-1)}$$

分散は,

$$\text{Var}(I_i) = w_i^2 \text{Var}(I)$$

標準偏差は,

$$Z_i = \frac{I_i - E_i}{\sqrt{\text{Var}(I_i)}}$$

ここで,  $\text{Var}(I)$  は, ランダム仮定でのグローバル・モラン統計量の分散を示す.

「空間的自己相関(ローカル)」タブを押し, 「地図選択」で地図を, 「項目選択」で項目を, 「分析手法選択」で「ローカル・モラン統計量」を選ぶ. そして「ウェイト値計算方法指定」と「ウェイト値の加工」を指定する(図46). 計算が終了すると, 結果が得られる. ローカル・モラン統計量  $I_i$  は「I.i 地図表示」, 期待値  $E_i$  は「E.I.i 地図表示」, 分散  $\text{Var}(I_i)$  は「Var.I.i 地図表示」, そして

標準偏差  $Z_i$  は「Z.I.i 地図表示」の各ボタンを押すと, それぞれに該当する地図が描画される.

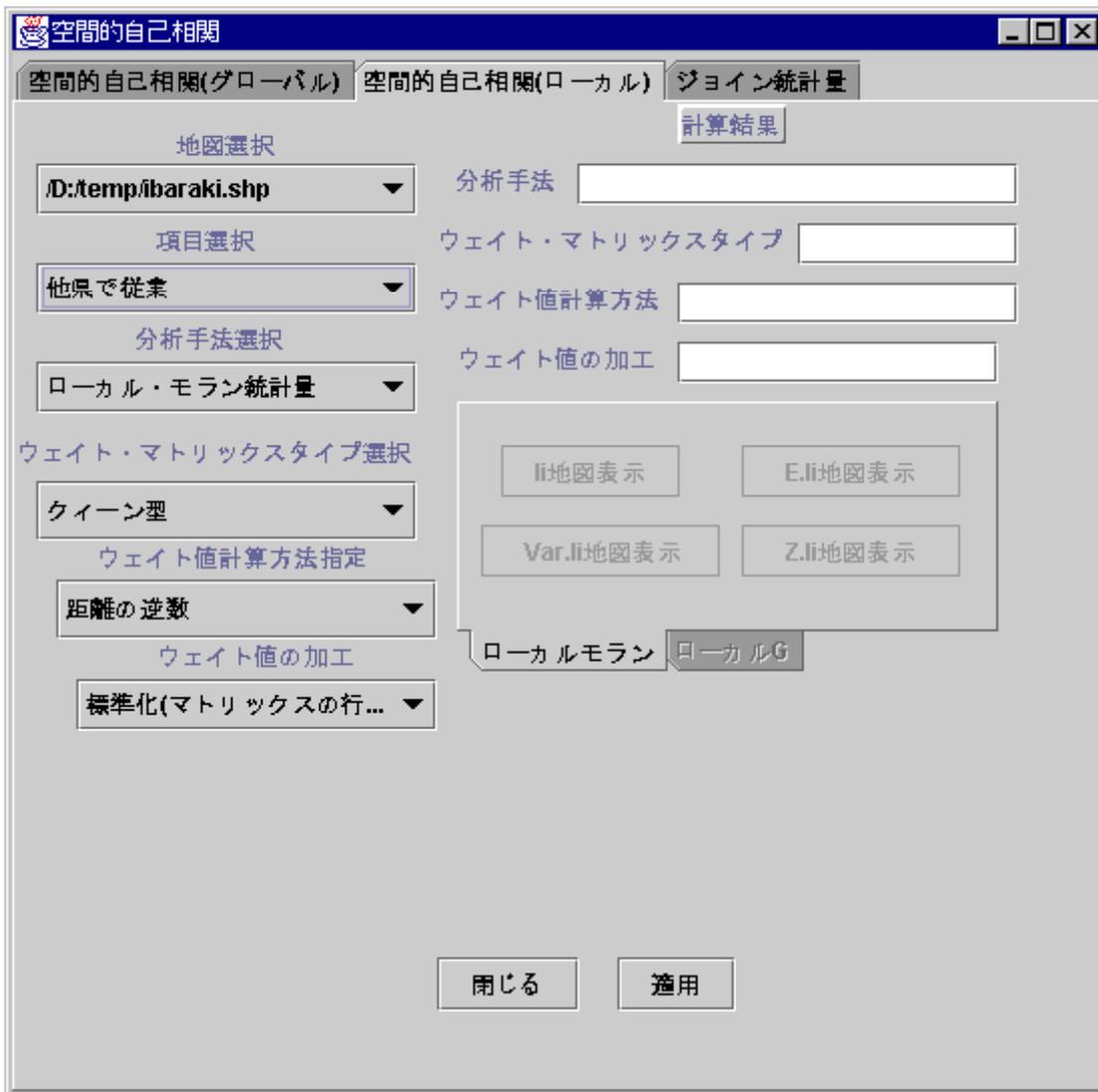


図 46 ローカル・モラン統計量

(2) ローカル G 統計量およびローカル G\* 統計量

これは、Getis & Ord(1992)による距離統計に基づいたバイナリのウェイトマトリックスのみを扱う  $G_i$  および  $G_i^*$  統計量について、バイナリ以外のウェイト・マトリックスを扱えるように拡張した統計量である(Ord & Getis, 1995)。ローカル G 統計量とローカル G\* 統計量とは、前者が  $i \neq j$  であるのに対して、後者は  $i = j$  を含む点が異なる。

$$G_i(d) = \frac{\sum_j w_{ij}(d)x_j - W_i \bar{x}(i)}{s(i) \sqrt{\left\{ \frac{[(n-1)S_{ii}] - W_i^2}{n-2} \right\}}} \quad i \neq j \quad \dots \quad \text{式(5)}$$

ローカル G\* 統計量の式は以下ようになる。

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^* \bar{x}}{s \sqrt{\left\{ \frac{[(nS_{1i}^*) - W_i^{*2}]}{n-1} \right\}}} \quad i = j \text{ を含む} . \quad \dots \quad \text{式(6)}$$

ここで、 $w_{ij}(d)$ は、観測地点  $i$  より距離  $d$  内にあるリンクの地点のウェイト・ベクトルである。

ローカル  $G$  統計量の場合、 $w_{ii}(d)$ は 0 になる。ローカル  $G^*$  統計量では、 $w_{ii}(d)$ は 0 ではない。

ウェイトの合計は以下のようになる。

$$W_i = \sum_j w_{ij}(d) \quad i \neq j$$

$$W_i^* = W_i + w_{ii} \quad i = j \text{ を含む} .$$

$$S_{1i} = \sum_j w_{ij}^2 \quad i \neq j$$

$$S_{1i}^* = \sum_j w_{ij}^2 \quad i = j \text{ を含む} .$$

各平均は

$$\bar{x}(i) = \frac{\sum_j x_j}{(n-1)} \quad i \neq j$$

$$\bar{x} = \frac{\sum_j x_j}{n} \quad i = j \text{ を含む} .$$

分散は以下のようになる。

$$s(i)^2 = \frac{\sum_j x_j^2}{(n-1)} - [\bar{x}(i)]^2 \quad i \neq j$$

$$s^{*2} = \frac{\sum_j x_j^2}{n} - \bar{x}^2 \quad i = j \text{ を含む} .$$

### (3) 実行方法

「空間的自己相関(ローカル)」タブを押すと図 47 が表示される。「地図選択」で地図を、「項目選択」で項目を、「分析手法選択」で「ローカル  $G$  統計量」を選択し、さらに距離などを入力する。計算が終了すると、図 48 の結果が得られる。上記と同様なやり方で、ローカル  $G^*$  統計量を計算すると、結果は図 49 のようになる。

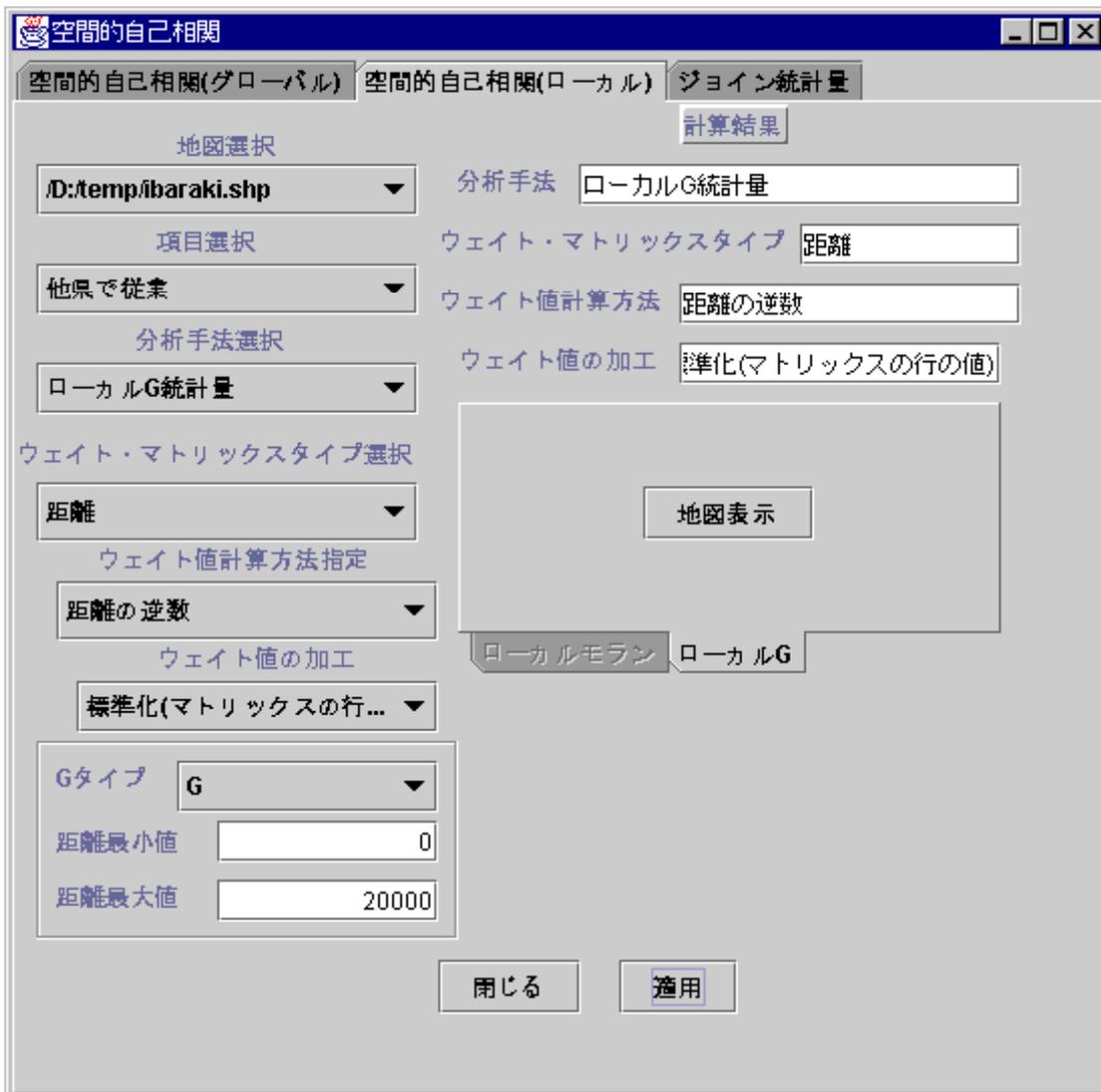


図 47 ローカル G 統計量

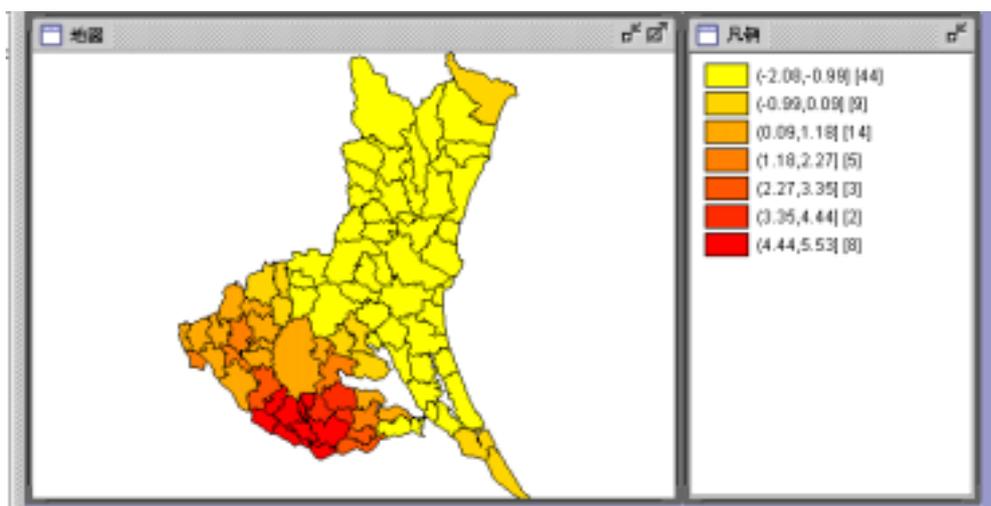


図 48 ローカル G 統計量の地図表示 (距離 20km)

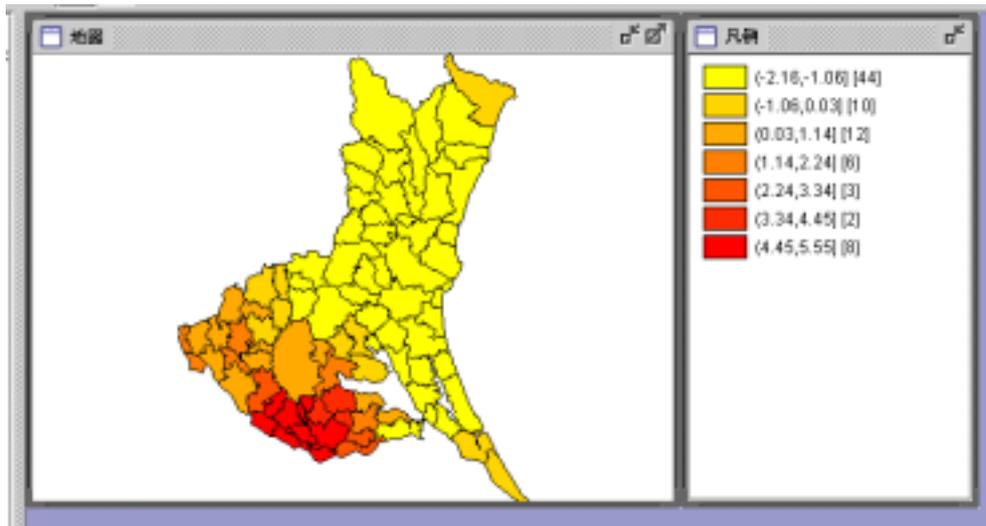


図 49 ローカル  $G^*$  統計量の地図表示 (距離 20km)

#### 6.4.7 Join 統計量

Join 統計量とは、空間的に隣接する一組のポリゴンやグリッドのセルの間でのクラスタリングまたは分離の度合いを計量的に測る方法である (Lee & Wong, 2000; Fotheringham *et al.*, 2000)。この統計量は名目尺度のみを対象とする。以下に 2 値 (黒と白 (存在する場合としない場合)) の例を示す。ここで、ポリゴン  $i$  の色が黒い場合が 1、色が白い場合は 0 とする。

BB 結合 (Joint):

$$O_{BB} = \frac{\sum_i \sum_j w_{ij} x_i x_j}{2}$$

WW 結合 (Joint):

$$O_{WW} = \frac{\sum_i \sum_j w_{ij} (1-x_i)(1-x_j)}{2}$$

BW 結合 (Joint):

$$O_{BW} = \frac{\sum_i \sum_j w_{ij} (x_i - x_j)^2}{2}$$

ここで、 $w_{ij}$  は観測点  $i$  と  $j$  間の 1 または 0 のウェイト値、 $x_i$  は観測点  $i$  の、 $x_j$  は観測点  $j$  のそれぞれの属性値を示す。図 50 は計算結果である。

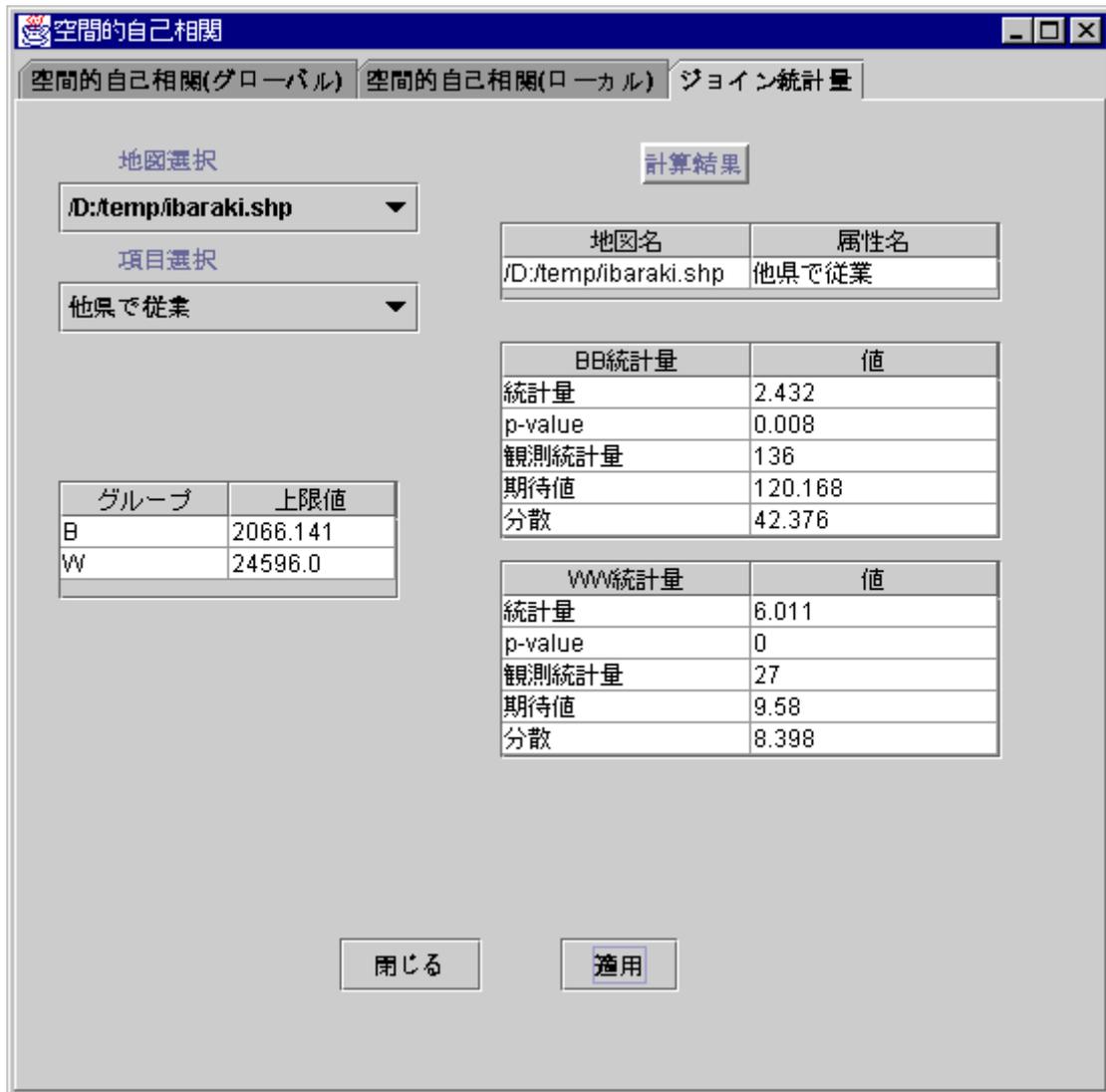


図 50 Join 統計量の計算結果

#### 6.4.8 空間的相互作用分析

##### (1) 空間的相互作用モデルとは

空間的相互作用モデルでは、地表における財・旅客・人口移動者・資金・情報・アイデアといったさまざまな流れを、発地区と着地区の規模、両地区間距離の関数で示す(石川,1988)。空間的相互作用モデルは以下のように定式化される。

$$T_{ij} = f(V_i, W_j, S_{ij}) \quad \dots \text{式(7)}$$

ここで、 $T_{ij}$  は地区  $i$  ,  $j$  間の相互作用、 $V_i$  は地区  $i$  の放出性(規模)、 $W_j$  は地区  $j$  の吸引力(規模)、 $S_{ij}$  は分離性(距離)である。

空間的相互作用分析は、古典的重力モデルを経て、1970年代に開発されたエントロピー最大

化モデルにより新たな段階に入り、地理学のみならず、多くの社会科学分野で利用されてきた。その後、1990年代に入り、複雑系の科学であるニューラルネット (Openshaw, 1993: 中谷, 1995: Openshaw and Openshaw, 1997) や遺伝的アルゴリズム (GA) (Diplock and Openshaw, 1996: Openshaw and Openshaw, 1997) を援用した空間的相互作用モデルが構築されるとともに、地理加重回帰 (GWR) と結びつけた研究も台頭している (Nakaya, 2003)。

### (2) 使用データ

SDAM で使用する入力データは属性データ (シェーブ・ファイルの DBase ファイル) と OD データ (DBase ファイル) の 2 つである。表 7 に OD データの入力方法を示す (3 地区からなる地域の例)。

表 7 OD データの入力例

ORID (発地 ID)	DESID (着地 ID)	TIJ (移動量)	DIJ (距離)
1	1	300	5
1	2	20	25
1	3	30	50
2	1	10	15
2	2	50	5
2	3	20	20
3	1	10	15
3	2	25	10
3	3	200	5

### (3) 古典的重力モデル

古典的重力モデルは以下のように定式化される<sup>4</sup>。

$$T_{ij} = k \frac{P_i^\alpha P_j^\beta}{D_{ij}^\gamma} \quad \dots \text{式(8)}$$

ここで、 $T_{ij}$  は発地  $i$ 、着地  $j$  間の流動量、 $P_i$  と  $P_j$  は発地  $i$  および着地  $j$  の (人口) 規模、 $D_{ij}$  は発地  $i$  と着地  $j$  間の距離であり、 $k$  は定数、 $\alpha$ 、 $\beta$ 、 $\gamma$  はパラメータである。 $\gamma$  は距離パラメータと呼ばれる。

古典的重力モデルでは、式の両辺の変数を対数変換し、通常重回帰分析でパラメータを求める。モデルの適合度は、以下の調整済み逸脱度で定義される。

$$d = 2 \left[ \left\{ \sum_i \sum_j T_{ij} \ln \left( \frac{T_{ij}}{\hat{T}_{ij}} \right) \right\} - \left\{ \sum_i \sum_j (T_{ij} - \hat{T}_{ij}) \right\} \right] \quad \dots \text{式(9)}$$

逸脱度は値が 0 のとき、モデルの適合度は完全で、値が大きくなるにつれ、適合度は低くなる。この逸脱度の式は、後述のエントロピー最大化モデルでも採用されている。

地図データを選択後、メニューバーの「分析」より、サブメニュー「空間的相互作用モデル」を選択すると、図 51 が表示される。

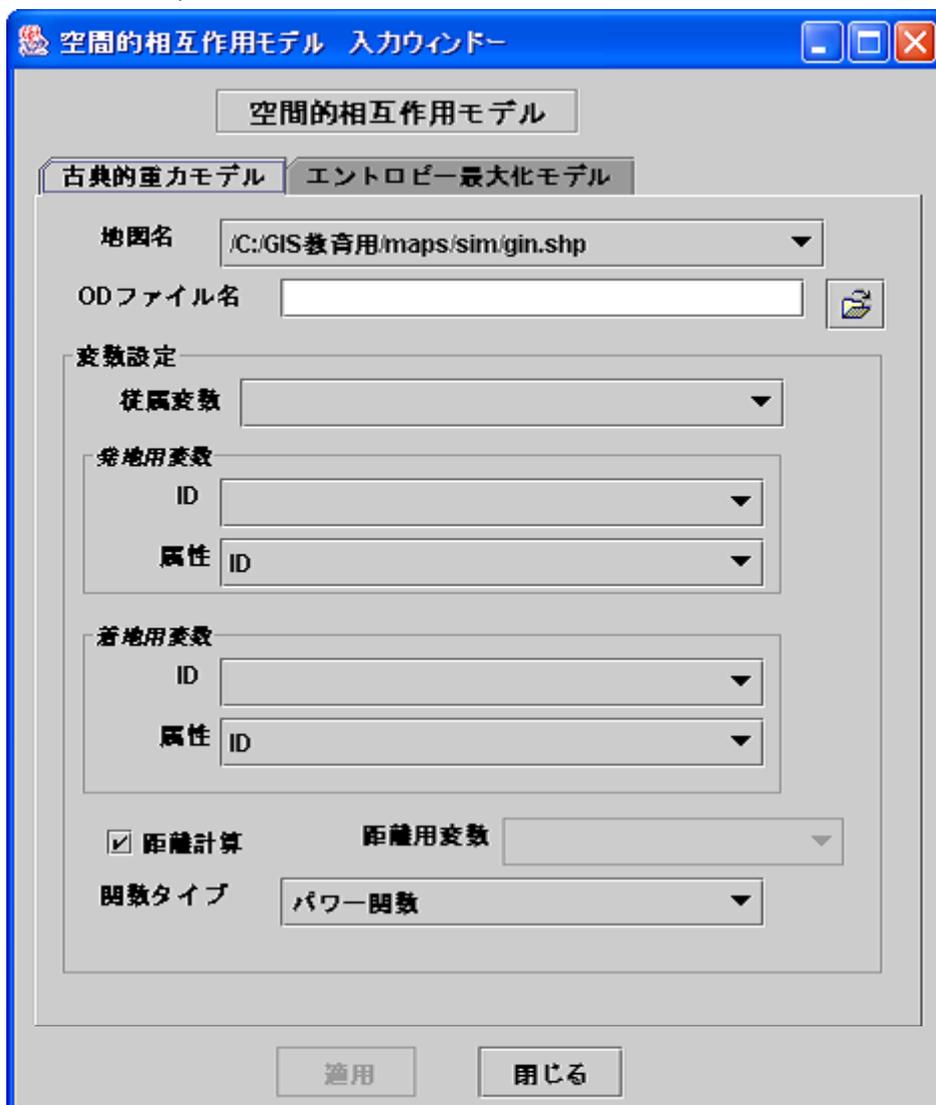


図 51 「空間的相互作用モデル」の初期画面

距離は、「距離計算」のチェックを入れると、地図データのポイントおよびポリゴンのシェープファイルより計算され、チェックを外すと、OD ファイル内の「距離用変数」の選択が有効になる。「関数タイプ」については、「パワー関数」と「指数関数」のいずれかを選択できる。図 52 が実行結果で、切片 (intercept) と各変数のパラメータと調整済み逸脱度が得られる。「予測値」タブを押すと、図 53 の予測値と残差が得られる。「Excel 出力」ボタンを押すと計算結果が出力される (図 54)。



図 52 計算結果：パラメータおよび逸脱度



図 53 計算結果：予測値・残差

1	2	3	4	5	6	7	8
1	空間的相互作用モデル(古典的重力モデル)						
2							
3	変数名						
4	(Intercept)	POPDの対数	POPDの対数	DIJの対数			
5	-4.565	0.601	1.452	-3.833			
6							
7	修正済みの逸脱度		252.056				
8							
9	発地ID	着地ID	実測値	予測値	残差		
10	1	1	106	35.964	72.066		
11	1	2	2	5.935	-3.935		
12	1	3	9	5.648	3.352		
13	2	1	2	9.958	-6.958		
14	2	2	46	13.387	32.613		
15	2	3	18	20.009	-2.009		
16	3	1	2	1.417	0.583		
17	3	2	3	3.334	-0.334		
18	3	3	715	1,012.54	-297.535		

図 54 計算結果の Excel への出力

#### (4) エントロピー最大化モデル

エントロピー最大化モデルのパラメータ推定は最尤法で行う。R 言語や商用ソフトの SPSS や SAS など多くの統計分析パッケージでサポートされている GLM (Generalized Linear Model, 一般化線形(型)モデル)を使えばパラメータが求められる(矢野, 1991)。GLM とは、変数間の線形回帰を一般化して、t 検定、分散分析、共分散分析、回帰分析、重回帰分析、ロジスティック回帰や正準相関分析など多くの分析手法を共通に扱えるようにした数学モデルのことをいう(中澤, 2003)。一般化線形モデルとよく似た用語に一般線形モデル(General Linear Model)があるが、このモデルは正規分布に従う 1 つの連続的従属変数と複数の同じく連続的独立変数からなる重回帰分析のことをさす。これに対し、一般化線形モデルでは、従属変数は連続量および正規分布を示さなくてもよく、カテゴリカルデータを含み、独立変数は線形式に限らない(ドブソン, 1993)。一般線形モデルは一般化線形モデルのサブセットといえる。

##### (a) 発生制約型モデル

・制約式

$$O_i = \sum_j T_{ij} = \sum_j \hat{T}_{ij} \quad \dots \quad \text{式(10)}$$

・発生制約型モデル式

$$\hat{T}_{ij} = A_i O_i w_j^\alpha d_{ij}^\beta \quad \dots \quad \text{式(11)}$$

ここで、 $T_{ij}$  は地区  $i, j$  間の流動量の観測値、 $\hat{T}_{ij}$  は流動量の予測値、 $A_i$  は発地  $i$  の均衡因子、 $O_i$  は発地  $i$  の発生量、 $w_j$  は着地  $j$  の吸引力（魅力度）、 $d_{ij}$  は地区  $i, j$  間の距離、 $\alpha$  と  $\beta$  はパラメータ（ $\beta$  は距離パラメータ）である。均衡因子  $A_i$  の式は以下のようになる。

$$A_i = \frac{1}{\sum_j w_j^\alpha d_{ij}^\beta} \quad \dots \quad \text{式(12)}$$

(b) 吸収制約型モデル

・制約式

$$D_j = \sum_i T_{ij} = \sum_i \hat{T}_{ij} \quad \dots \quad \text{式(13)}$$

・吸収制約型モデル式

$$\hat{T}_{ij} = v_i^\lambda B_j D_j d_{ij}^\beta \quad \dots \quad \text{式(14)}$$

ここで、 $T_{ij}$  は地区  $i, j$  間の流動量の観測値、 $\hat{T}_{ij}$  は流動量の予測値、 $v_i$  は発地  $i$  の放出性、 $B_j$  は着地  $j$  の均衡因子、 $D_j$  は着地  $j$  の吸収量、 $d_{ij}$  は地区  $i, j$  間の距離、 $\lambda$  と  $\beta$  はパラメータ（ $\beta$  は距離パラメータ）である。均衡因子  $B_j$  の式は以下のようになる。

$$B_j = \frac{1}{\sum_i v_i^\lambda d_{ij}^\beta} \quad \dots \quad \text{式(15)}$$

(c) 二重制約型モデル（発生・吸収型制約モデル）

・制約式

$$O_i = \sum_j T_{ij} = \sum_j \hat{T}_{ij} \quad \dots \quad \text{式(16)}$$

$$D_j = \sum_i T_{ij} = \sum_i \hat{T}_{ij} \quad \dots \quad \text{式(17)}$$

・二重制約型モデル式

$$\hat{T}_{ij} = A_i O_i B_j D_j d_{ij}^\beta \quad \dots \quad \text{式(18)}$$

ここで、 $T_{ij}$  は地区  $i, j$  間の流動量の観測値、 $\hat{T}_{ij}$  は流動量の予測値、 $A_i$  は発地  $i$  の均衡因子、 $O_i$  は発地  $i$  の発生量、 $B_j$  は着地  $j$  の均衡因子、 $D_j$  は着地  $j$  の吸収量、 $d_{ij}$  は地区  $i, j$  間

の距離， $\alpha$  と  $\beta$  はパラメータ ( $\alpha$  は距離パラメータ) である．均衡因子  $A_i$  および  $B_j$  の式は以下のようになる．

$$A_i = \frac{1}{\sum_j w_j^\alpha d_{ij}^\beta} \quad \dots \quad \text{式(19)}$$

$$B_j = \frac{1}{\sum_i v_i^\lambda d_{ij}^\beta} \quad \dots \quad \text{式(20)}$$

#### 6.4.9 GWR (地理加重回帰分析)

##### (1) GWR とは

GWR (Geographically Weighted Regression) とは，異なった位置で観測された変数間の関係は地域全体では一定でないという前提のもとで，空間的に変動する関係を統計的に明らかにする手法である (Fotheringham, Brunsdon & Charlton, 2000)．

周知のように，一般的な回帰式は以下のように定式化される．

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \quad \dots \quad \text{式(21)}$$

ここで， $y$  は従属変数， $x$  は独立変数， $\beta$  はパラメータ， $i$  は地点， $k$  は独立変数の番号をあらわす．

これに対し，GWR はローカル・パラメータを推定するよう式(21)を拡張したもので，以下のように定式化される．

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad \dots \quad \text{式(22)}$$

ここで， $(u_i, v_i)$  は地点  $i$  の座標， $\beta_k(u_i, v_i)$  は地点  $i$  における連続関数  $(u, v)$  の値である．GWR では地点ごとにウェイトつき最小二乗法によりパラメータを求めていく．パラメータのベクトルを  $\beta$ ，独立変数の観測値のマトリックスを  $X$ ，従属変数の観測値のベクトルを  $y$ ，地点  $i$  のウェイトマトリックスを  $W(u_i, v_i)$  とすると，地点  $i$  のパラメータは以下の式で求められる．

$$\beta(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y$$

GWR では，任意の観測地点の従属変数は遠くの地点よりも近くの地点により多くの影響を与えるという前提を有しているため，ウェイトマトリックスは地点  $i$  からの距離逓減を考慮に入れた空間カーネルによって定義される．ウェイトマトリックスの値は対角要素以外は 0 になる．

##### (2) 入力画面

従属および独立変数，バンド幅および固定型を選択した入力画面の例を図 55 に示す．

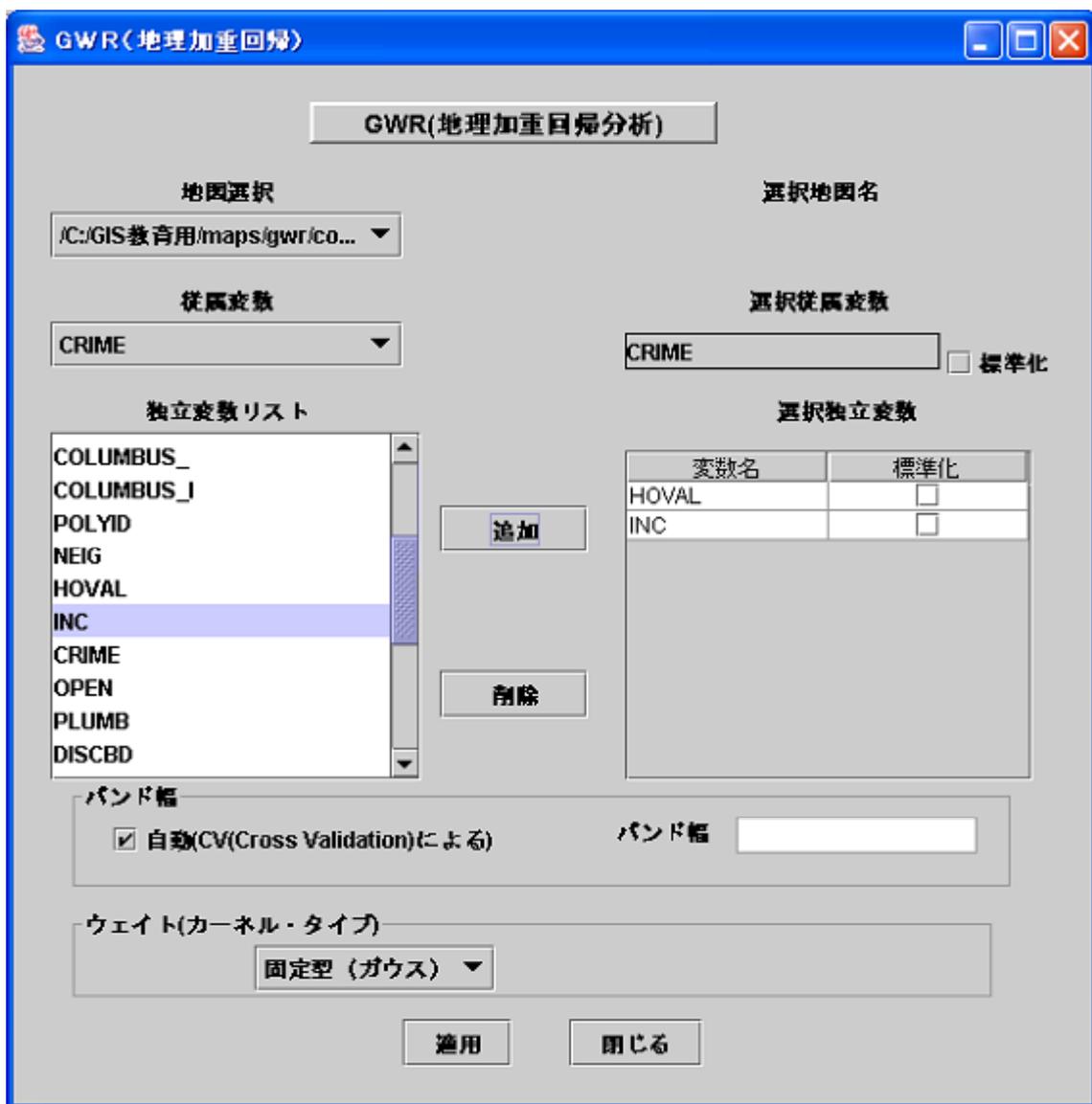


図 55 GWR の入力画面

入力変数を標準化する場合は次のようにする。従属変数では図 55 の右側にある「選択従属変数」右横にチェックを入れる。独立変数では「選択独立変数」のテーブルにおける標準化する変数の行の「標準化」カラムにチェックを入れる。

SDAM ではバンド幅を相互検証法のクロスバリデーション (Cross Validation) によって推定するか、ユーザがバンド幅を指定するか、いずれかの方法がとれる。ウェイト (カーネル・タイプ) については固定型空間カーネルのみを処理し、ガウスと距離の二乗の 2 タイプをサポートする。

### (3) 計算結果

図 56 の「パラメータ」タブを押すと、決定係数と地点ごとの各変数のパラメータが得られる。Excel へ出力する場合には、図の右上の「Excel 出力」ボタンを押せばよい。パラメータの地図表示については後述する。「検定」タブを押すと、検定情報が得られる。



図 56 パラメータ

SDAM では , 「 Shapiro-Wilk 正規性検定 」 ( デフォルト表示 ) , 「 Brunson-Fotheringham-Charlton ANOVA テーブル 」 ( 図 57 ) , 「 Leung-Mei-Zhang F\_1 テスト および F\_2 テスト 」 ( 図 58 ) が行える .

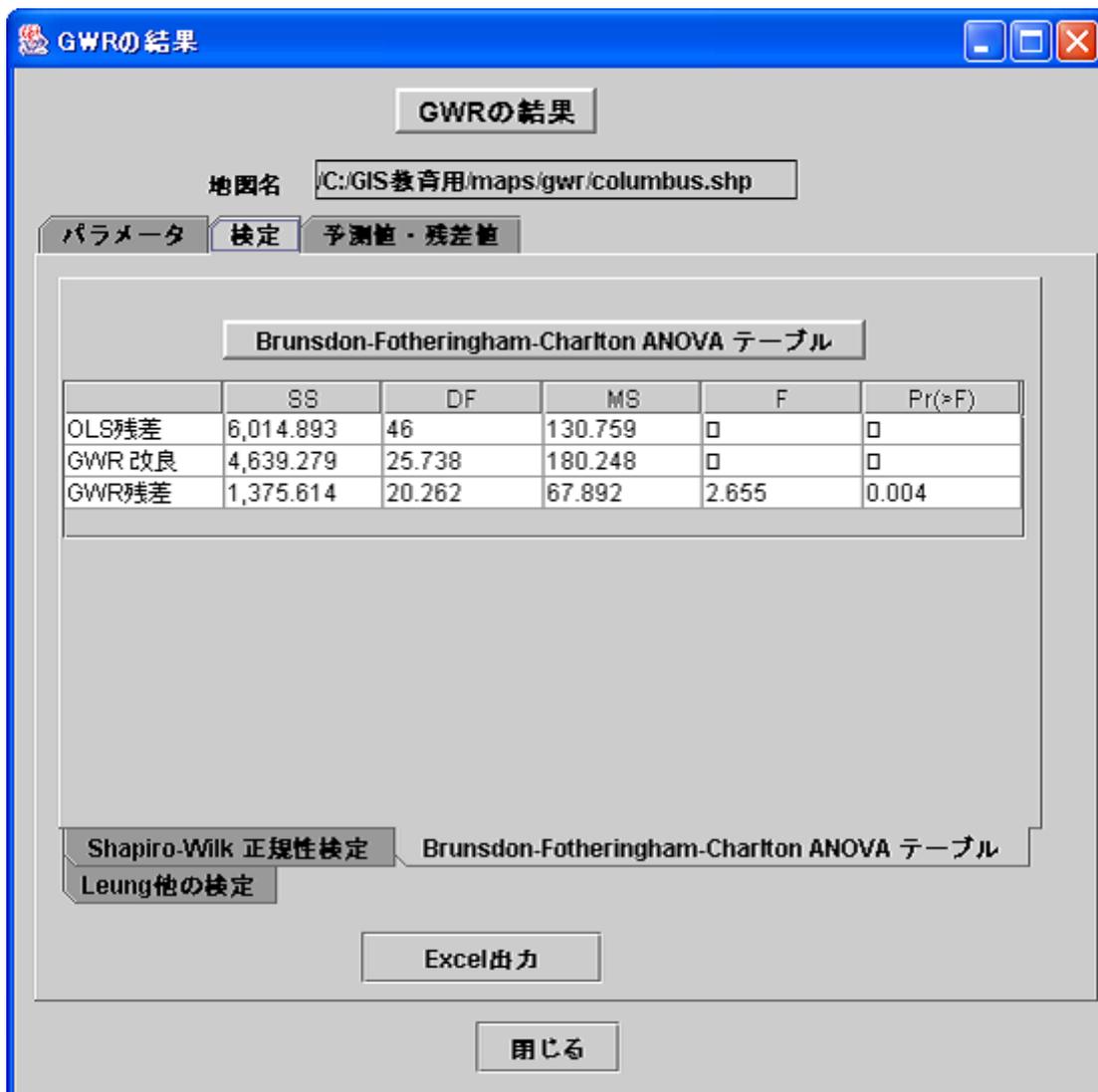


図 57 Brunsdon-Fotheringham-Charlton ANOVA テーブル

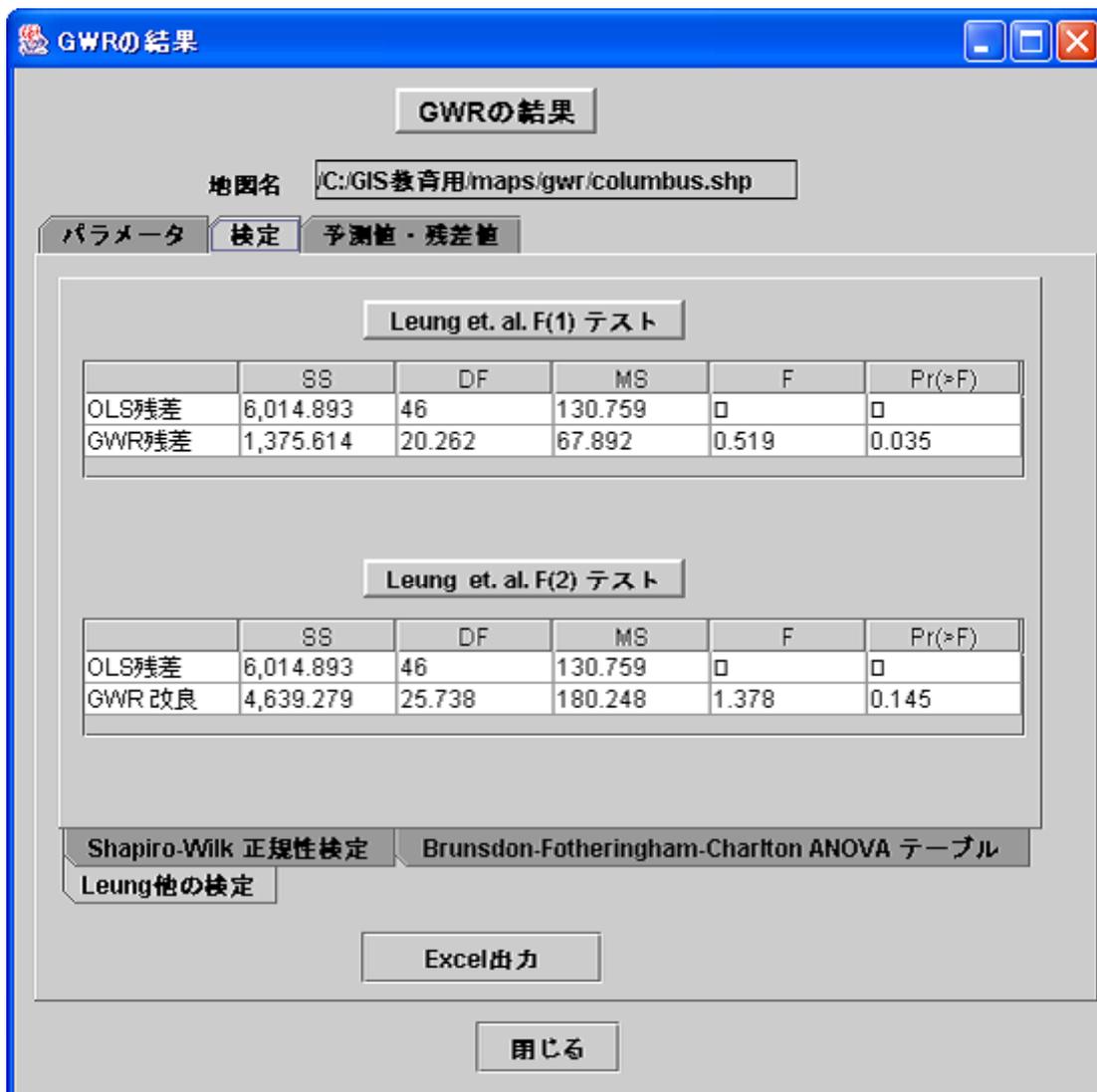


図 58 Leung-Mei-Zhang F<sub>1</sub> テストおよび F<sub>2</sub> テスト

ウィンドウ上部の「予測値・残差値」タブを選択すると、予測値と残差値を求めることができる（図 59）。



図 59 予測値・残差値

(4) 地図表示

SDAM では、GWR の計算結果のうち、地点ごとのパラメータ、予測値および残差値の地図表示をサポートしている。地図表示は「パラメータ」および「予測値・残差値」の各パネルの下にある図 60 の「地図表示」パネルで操作をおこなう。

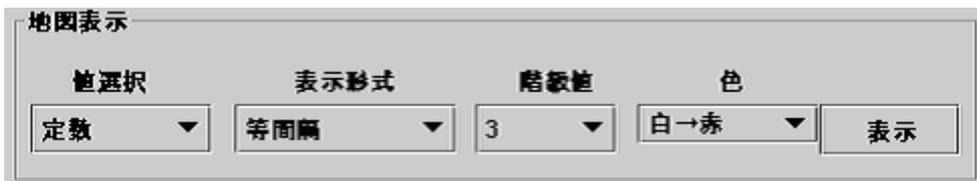


図 60 地図表示パネル

「値選択」で表示したい定数・変数名、「表示形式」、「階級値」を指定し、「色」を選択後、「表示」ボタンを押すと、パラメータ値、予測値、残差値の地図が表示される。図 61 はパラメータ値、図 62 は残差値に関する地図の出力例である。

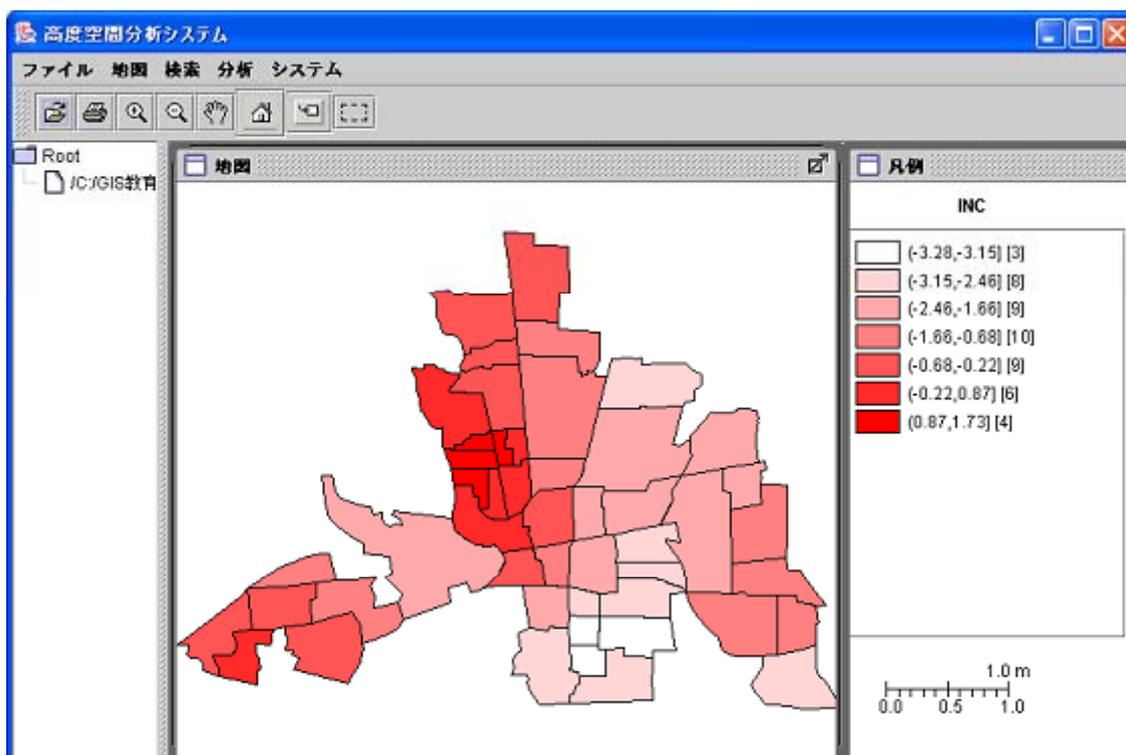


図 61 パラメータ値

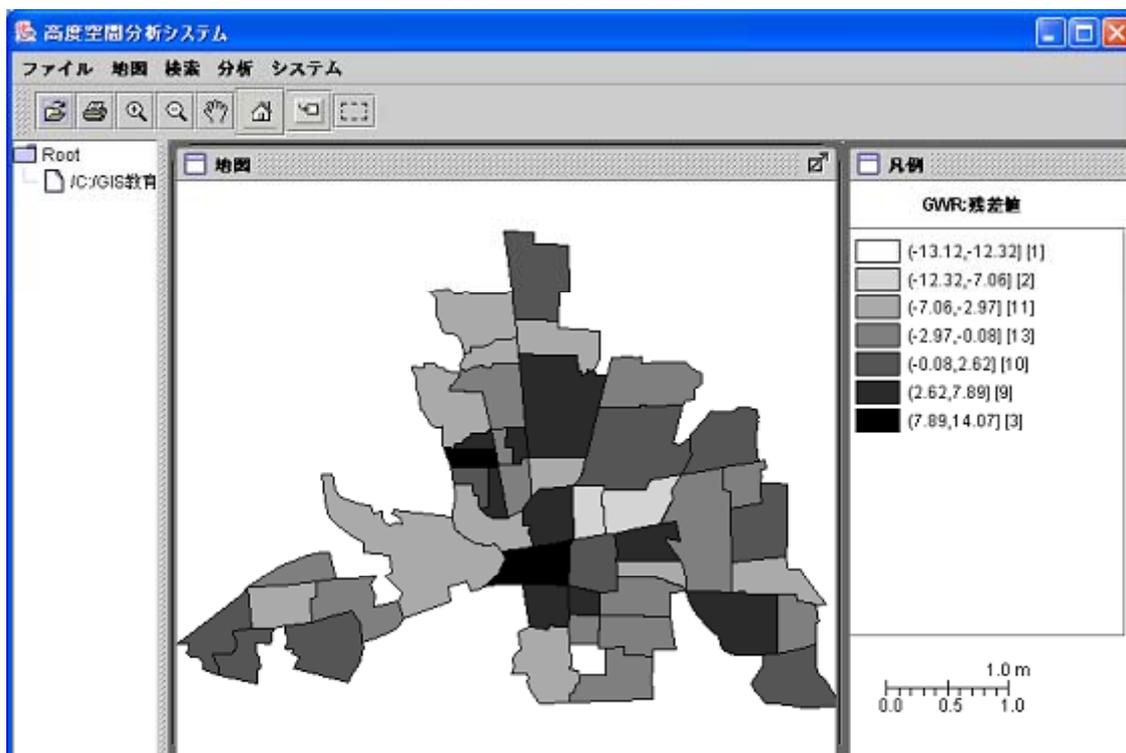


図 62 残差値

## 7. 評価

SDAM は R 言語を主体として構築された空間分析のフリーウェアである。SDAM のセールスポイ

ントとして、拡張性に優れている点を強調しておきたい。オープンソースである R 言語では、ソースコードが自由に閲覧できることもあり、新しいプログラムや修正プログラムが随時公開されている。プログラムは世界中のユーザによって評価され、絶えず修正・改良が行われている。GIS 関連では、空間統計学者の Roger Bivand 教授（ノルウェー）によって、R 言語を用いた空間分析のサイト(R Spatial projects)<sup>5</sup>が運営されており、ここでは洗練された空間分析プログラムが集積している(Bivand & Gebhardt, 2000)。このライブラリーを活用しながら、漸次、空間解析機能を充実させ、アドオンモジュール方式により SDAM のバージョンアップを図っていくことが可能である。さしあたって、R によるアプリケーションが公開されている等値線(コンター)生成、遺伝的アルゴリズム、マルチエージェント、セルオートマトン、自己組織化マップ、フラクタル、マルチレベルモデリング、GAM<sup>6</sup>などの手法を SDAM で操作可能にすることが望まれる。

## 8. おわりに

わが国でも、R 言語と GIS を結びつけた研究が進展しつつある。R 言語の情報交換を目的とするサイトに RjpWiki<sup>7</sup>があるが、このなかで R 言語による空間分析の詳細な情報が得られる。たとえば、「R で GIS」では、R 言語で実現できる GIS の各手法が説明されている。「Rmap を使った地図表示」は、R 言語で地図表示をおこなうパッケージである Rmap(多数の地図フォーマット、地図投影変換をサポート)について解説している。「GIS で利用可能な R パッケージ内データ」では、地図データと属性データに分けて所在情報が一覧表の形で示されている。「R と PostGIS」では、PostGIS を使って GIS の空間検索、空間分析を行う方法を説示している。

SDAM は GIS あるいは空間分析の教育実習用に開発されたソフトウェアであり、行政支援やビジネス利用を想定したシステムではない。それゆえ、データベース機能や可視化機能は劣り、また、システムにはバグやエラーが潜んでいるかもしれない。ユーザには不具合に対する情報の提供、忌憚のない意見・アドバイスをお願いしたい。

## 参考文献

- 石川義孝(1988)『空間的相互作用モデル』, 地人書房。
- ヴェナブルズ, W.N.・リプリー, B.D. 著, 伊藤 幹夫・大津 泰介・戸瀬 信之・中東 雅樹訳(2001)『S PLUS による統計解析』, 日本シュプリンガー・フェアラーク。
- 岡田昌史編(2004)『The R Book』, 九天社。
- 杉浦芳夫(1986)統計学的方法の基礎, 野上道男・杉浦芳夫『数理地理学演習』, 古今書院, 75-83。
- ドブソン, A. J. 著, 田中豊・森川敏彦・栗原考次訳(1993)『統計モデル 回帰モデルから一般化線形回帰まで』, 共立出版。
- 村山祐司(1990)『地域分析 地域の見方・読み方・調べ方』, 古今書院。
- 村山祐司・尾野久二編(1993)『地域分析のための地理情報システム - Arc/INFO を利用して -』, 文部省重点領域研究「近代化と環境変化」技術資料。
- 村山祐司・尾野久二(2003)オープンソースを利用した統合型空間分析システムの開発, 『筑波大学人文地理学研究』, 27, 71-105。
- 中澤 港(2003)『R による統計解析の基礎』, ピアソンエデュケーション。
- 中谷友樹(1995)PR 型パーセプトロンによる空間的相互作用モデリング - 日本の人口移動を事

- 例として,「人文地理」,47, 521-540.
- 矢野桂司(1991)一般線形モデルによる空間的相互作用モデルの統合,「地理学評論」, 64,367-387.
- 矢野桂司(2003)空間的相互作用モデル,杉浦芳夫編『地理空間分析』,朝倉書店,84-106.
- 渡辺利夫(1994)『使いながら学ぶS言語』,オーム社.
- Anselin, L. (1995) Local indicators of spatial association LISA. *Geographical Analysis*, 27, 93-115.
- Bivand, R. S. and Gebhardt, A. (2000) Implementing functions for spatial statistical analysis using the R language. *Journal of Geographical Systems*, 2, 307-317.
- Diplock, G. and Openshaw, S. (1996): Using simple genetic algorithms to calibrate spatial interaction models. *Geographical Analysis*, 28, 262-279.
- Dorling, D.(1996) Area cartograms: their use and creation. *Concepts and Techniques in Modern Geography (CATMOG)*, 59.
- Fotheringham, A. S.(2002) *Geographically Weighted Regression - The Analysis of Spatially Varying Relationships*, London: John Wiley & Sons.
- Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2000) *Quantitative Geography Perspectives on Spatial Data Analysis -*, London: SAGE Publications.
- Getis, A. and Ord, J.K. (1992): The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24, 189-199.
- Lee, J. and Wong, W. S. (2000) *Statistical Analysis with ArcView GIS*, New York: John Wiley & Sons.
- Nakaya, T. (2003) Local spatial interaction modelling based on the geographically weighted regression approach. In Boots, B., Okabe, A. and Thomas, R., eds. *Modelling Geographical Systems: Statistical and Computational Applications*, Dordrecht: Kluwer Academic Publishers, 45-68.
- Openshaw, S. (1993) Modelling spatial interaction using a neural net. In Fischer, M. and Nijkamp, P., eds. *Geographical Information Systems, Spatial Modelling, and Policy Evaluation*, New York: Springer-Verlag, 17-34.
- Openshaw, S. and Openshaw, C. (1997) *Artificial Intelligence in Geography*, Chichester: John Wiley & Sons.
- Ord, J.K. and Getis, A. (1995) Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27, 286-296.
- Vivid Solutions (2002): JTS technical specifications (<http://www.vividsolutions.com/jts/jtshome.htm>).

<sup>1</sup> SDAMはフリーウェアである。著作権は村山祐司・尾野久二がもつが,使用は自由に行って差し支えない。ただし,本ソフトを使用することでデータの消失等の損害が生じても,著作権者は責任を負いかねる。使用者各自の責任で使用願いたい。なおPCの環境によっては,SDAM本体もしくは機能の一部が正常に作動しないことがあり得ることをことわっておく。

<sup>2</sup> 表2の issa パッケージに含まれる因子分析プログラムは慶應義塾大学渡辺利夫教授によって開発され

---

た .

<sup>3</sup> 以上の記述は杉浦 (1986) にもとづく .

<sup>4</sup> 古典的重力モデルおよびエントロピー最大化法に関する数式の説明は矢野(1991, 2003)を参考にした .

<sup>5</sup> <http://sal.agecon.uiuc.edu/csiss/Rgeo/>

<sup>6</sup> Geographical Analysis Machine

<sup>7</sup> <http://www.okada.jp.org/RWiki/>