

ニューラルネットワークを導入した 土地分級評価精度の向上

小島尚人¹・大林成行²・山森英俊³

¹正会員 工博 東京理科大学講師 理工学部土木工学科 (〒278 千葉県野田市山崎 2641)

²正会員 工博 東京理科大学教授 理工学部土木工学科 (同上)

³学生員 東京理科大学理工学部土木工科大学院 (同上)

本研究は、著者らが開発した土地分級評価モデル(潜在因子モデル)の評価精度の向上を目的とし、従来までの数量化理論を中核とした分析機能に加えて、新たにニューラルネットワークを導入する方法を提案したものである。潜在因子モデルでは、トレーニングデータを基準として土地の性状が類似する箇所を判別・分析することから、どのような分析手法を用いても「理想的には分級評価結果は同一」にならなければならない。この基本的な考えのもとに、数量化理論とニューラルネットワークによるそれぞれの解が一致したトレーニングサンプルを再抽出する方法を提案した。この方法を用いれば判別精度が向上するとともに、数量化理論とニューラルネットワークの解が互いに歩み寄ってくるといった効果を示し、提案手法の実用性を示している。

Key Words : *land use capability classification, neural network, quantification method, training data, satellite multispectral data, geographical information*

1. はじめに

土地利用計画, 地域計画等に地理情報システム(GIS: Geographical Information System)をはじめ衛星リモートセンシングの技術が導入されるようになり, より付加価値のある情報を抽出しようとする研究開発が多くなってきている^{1,2)}。筆者らも「国土の調査・計画」と「衛星データの実利用化」といった問題に両面から取り組み, 地形, 表層, 土壌といった各種地理情報と衛星データを融合利用し, 土地の性状を分析する土地分級評価モデル, いわゆる潜在因子モデルを開発するとともに³⁾, 土地分級評価における衛星データそのものの有効性を示した⁴⁾。さらに, 評価主題の異なる複数の分級評価図を組合せ, 土地の性状を多角的な視点から分析できる土地利用構想図の作成アルゴリズムを提案し⁵⁾, 現地との照合を通して構想図の有用性を検証した。

しかし, ここまでの研究は, 潜在因子モデルの構築とその実用性・拡張性を立証することに研究の主眼が置かれたことから, 分級評価結果の良否や分級評価精度を向上させるといった問題については十分な議論を展開することができなかった。

潜在因子モデルは, 数量化理論(II類とIII類)を中核としてトレーニングデータを基準としたミニマックス2群判別手法から構成されるが⁶⁾, 質的データを数量化して判別分析を行う上で, このような線形モデルである必

然性はない。最近では非線形モデルに代表されるニューラルネットワークが各種の判別分析に有効であるといった研究成果が数多く発表されてきており⁶⁾, 本研究で対象とする土地分級評価においてもニューラルネットワークの判別能力について検討することは不可欠の課題となってきた。

ニューラルネットワークは, トレーニングデータを基準として複雑な入出力関係を高精度で推定することが可能であるが, その内部構造に対して意味解釈を加えることが困難であるため, ブラックボックス化されていることは古くから指摘されている。また, 数量化理論とニューラルネットワークから得られる判別解の違いも比較・議論されるようになってきており^{7,8)}, 適用分野別に未だ検討すべき課題は多いと言える。

以上の背景のもとに, 本研究では, 潜在因子モデルの分析機能の一つとして, 数量化理論に加えてニューラルネットワークを導入し, 分級評価精度を向上させる方法を提案することを目的とする。潜在因子モデルでは, トレーニングデータを基準として土地の性状が類似する箇所を抽出・分析することから, どのような分析手法を用いても理想的には分級評価結果は同一にならなければならない。この基本的な考えのもとに, 数量化理論とニューラルネットワークによる分級評価結果の違いを差画像を通して解釈するとともに, これらの解の違いを利用したトレーニングデータの再設定方法を提案する。この方法

を用いれば、トレーニングデータと類似する箇所の判別精度が向上するだけでなく、線形解と非線形解の違いが減少することも示し、提案手法の実用性を述べる。

2. 従来までの研究と本研究の特徴

(1) 分級評価精度を高める意義

筆者らが提案してきた潜在因子モデルでは「適・不適」といった2群判別結果によって曖昧性を排除し、具体性のある判断が下せる分級評価プロセスを有している。従来から利用されている代表的な分級評価手法と比較することによって、潜在因子モデルにおける2群判別の有用性と発展性については既に示している³⁾⁻⁵⁾。詳細は割愛するが、例えば畑地適性を評価する場合、土地の持つ潜在的な価値の評価の範囲内において

- ①現況畑地を将来にわたって保全すべき地域
- ②現況畑地を他の土地利用に転換してもよい地域
- ③現況畑地と同様な性質をもち、畑地適性のある地域
(新たに農地振興候補地として指摘していく地域)
- ④畑地適性のない地域

といった基本的な判断を支援する情報が必要となる。潜在因子モデルでは、トレーニングデータを基準とした2群判別結果からこれら4つの情報を色分けして表示し、評価者の意思決定を支援できる点が特徴となっている。

さらに潜在因子モデルでは、トレーニングデータの設定方法の違いによって現状型、規範型、仮定型といった3つの分級評価プロセスや³⁾、複数の分級評価図を組み合わせる土地利用構想図を作成する等の評価アルゴリズムを有していることから⁵⁾、評価主題別に作成される分級評価図の精度を高める方法を考えることは意義ある検討課題となる。この課題に対して、本研究ではニューラルネットワークの土地分級評価への導入方法を提案する。

(2) トレーニングデータの再設定方法の提案

ニューラルネットワークの特徴は、学習によってトレーニングデータの認識能力を高められることであるが、一旦設定されたトレーニングデータに対する学習を経て構成されるネットワーク構成が必ずしも良好な判別結果を得ることができるとは限らない。なぜならトレーニングデータ内には異常値を含むデータが存在する可能性もあるからである。この問題に対する最近の研究事例を見ると、衛星データを対象とした土地被覆分類において、最尤法とニューラルネットワークの解の違いを利用してトレーニングデータ内に混在する異常値を除去する方法も提案されているが⁶⁾、必ずしも画像全体にわたる分類精度の向上については顕著な効果が得られないことが報告されている。また、電力系統の事故区間判定への

判別分析において、数量化理論とニューラルネットワークの解の違いを比較した研究も見られるが⁷⁾、線形解と非線形解の整合性の問題については今後の課題とされている。

そこで、本研究では数量化理論とニューラルネットワークの解の違いを分析するとともに、この違いを利用したトレーニングデータの再設定方法を提案する。この方法を用いれば、トレーニングデータと類似する箇所の判別精度が向上するだけでなく、線形解と非線形解の違いが減少することを示す点が、本研究の特色となる。

(3) ニューラルネットワークと数量化II類併用の意義

分級評価精度は、トレーニングデータに対する判別精度と対象領域全体にわたる判別精度に分けて考えなければならない。前者はトレーニングデータ自身に判別された画素数の百分率(的中率³⁾)によって定量的に評価できるが、後者は分級評価図作成後に現地と照合できる箇所に対する判定結果で評価せざるを得ないことから、客観的に分級評価精度を検証する上でのずと限界がある。

潜在因子モデルでは、トレーニングデータに類似する箇所を抽出・分析することから、数量化理論とニューラルネットワークの解が互いに類似すれば、少なくとも評価対象領域全体にわたる分級評価精度は高まったものと推定できる。ニューラルネットワークと数量化II類を併用してトレーニングデータの再設定を行い、その有効性を検討する本研究の意義はこの点にある。

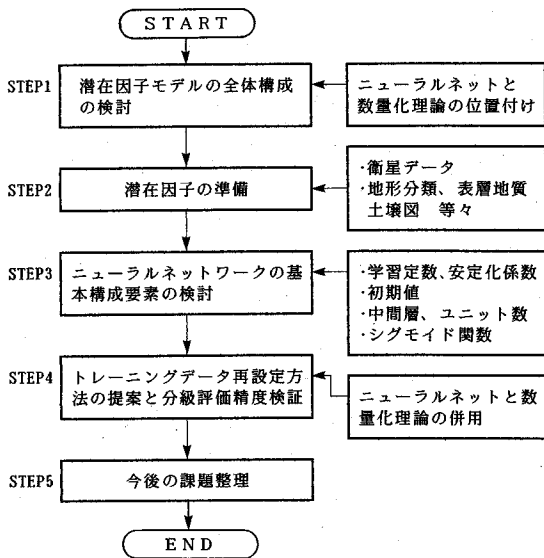
また、地理情報を利用した土地分級評価においてニューラルネットワークを導入した研究も見られるが⁸⁾、「トレーニングデータに対する再現性は良い」ものの評価図全体の精度やネットワーク構造の問題については言及されていない。衛星データと地理情報を融合利用し、土地分級評価を行う研究は、土地利用構想・計画、地域計画等での活用が期待されているが、手法別の分級評価結果の比較、さらには異なる手法を併用して分級評価精度を向上させようとする研究は、筆者らの知る限り見あたらない。

3. 研究の構成

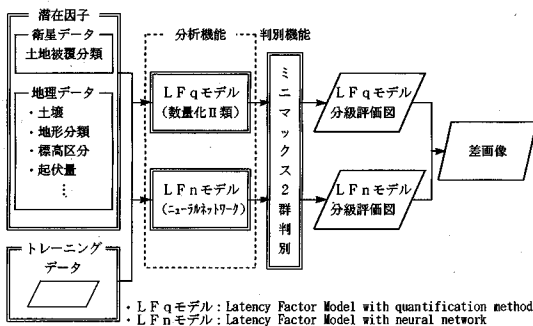
本研究の全体構成を図-1に示す。

(1) 潜在因子モデルの全体構成の検討 (STEP 1)

潜在因子モデルにおけるニューラルネットワークの位置付けを図-2に示す。従来まではトレーニングデータと評価対象データを外的基準として数量化II類を通して個体に数量を付与した後に、ミニマックス2群判別によって分級評価図を作成していた³⁾。本研究では、さらに個体に数量を付与する機能にニューラルネットワークを



図一 研究の全体構成



図二 ニューラルネットワークの位置付け

導入する。以降の説明を判り易くするために、数量化理論とニューラルネットワークをベースとした分級評価機能をそれぞれ LFq モデル (Latency Factor model with quantification method) および LFn モデル (Latency Factor model with neural network) と定義する。

(2) 潜在因子の準備 (STEP 2)

a) 分級評価項目の設定と対象領域

潜在因子モデルでは、分級評価項目を設定する際に、農村的要因と都市的要因の2つに区分するようになっている。この区分の考え方は、参考文献5)に詳述してあるので割愛するが、本研究では農村的要因として「畑地適性、水田適性、樹林地適性」を、都市的要因として「宅地適性」を評価主題として設定した。

対象領域は、著者らの大学の近くに位置し、分級評価図と現地との照合を繰り返し行えることを念頭に置き、千葉県流山市の近郊とした。地域地区指定の線引きや逆

表一 本研究で使用した潜在因子

土地分類基本調査成果図			衛星データ因子 (TMデータ) 1990年11月5日観測	
表層地質因子	地形分類因子	土壌因子		
埋立地堆積物 泥がち堆積物 砂がち堆積物 砂質堆積物 ローム1 ローム2	火山灰台地 谷底平野 自然堤防 切土地 盛土地 崖	八街F統 八街統 船木統 上砂F統 布佐統 平三統	クラス1 クラス2 クラス3 クラス4 クラス5 クラス6	
DTMから作成				
標高因子	起伏量因子	谷密度因子	斜面方位因子	傾斜区分因子
以上~未満 0~8 8~12 12~16 16~20 20~24 24~	以上~未満 (m) 0~2 2~4 4~6 6~8 8~10 10~	(個) 0 1 2 3 4	北 北東 東 南東 南 南西 西 北西 平地	以上~未満 (度) 0~3 3~6 6~9 9~12 12~

線引き等の問題に対しても分級評価図を支援的に利用していくことを念頭に置き、都市計画によって地域指定がなされている箇所と白地領域が含まれている箇所を選定した。地域地区指定がなされている箇所は市街化しており、白地領域は水田や畑地、樹林地等の農村的立地条件に依存する土地利用となっている。白地領域は、今後の土地利用計画を展開する上で注目される箇所であり、精度の高い分級評価図が支援情報として必要となる。

b) 衛星データを含めた素因データの準備

本研究を進める上で使用した潜在因子を表一に示す。土地分類基本調査の成果図面を数値化した地理データとDTM (Digital Terrain Model) からコンピュータ処理によって作成される数値データ (標高, 起伏量, 谷密度, 斜面方位, 傾斜区分の各因子) である。

c) 衛星データの選定

本研究では30m/ピクセルという高分解能を有し、広く利用されているLANDSAT/TMデータ (1990年11月5日) を使用した。潜在因子モデルでは、多次元で構成される画像濃度値の情報を集約化 (クラスタリング) した上で、1つのアイテムとして組み込むようになっている⁹⁾。

(3) ニューラルネットワークの基本構成要素の検討 (STEP 3)

ニューラルネットワークを導入したモデルを考える場合には中間層の数, 各層におけるユニットの数などの各種のパラメータが必要となるが, それらの決定には未だ一般的な方法論は見い出されていない^{6),10)}。本研究のように地理情報と衛星データを融合利用する場合において

も同様な問題に直面する。研究の前提条件や範囲を絞っておくことが、実用性のある土地分級評価支援システムを構築する上で重要となる。この検討は4章で述べる。

(4) トレーニングデータの再設定方法の提案(STEP 4)

潜在因子モデルの分級評価精度は、トレーニングデータの良否に左右される。初期設定したトレーニングデータを用いた場合には、LFnモデルとLFqモデルの分級評価図にかなりの違いが現れることを示し、トレーニングデータ内に混在する画素を除去することの必要性を述べる。具体的な対策として、LFnモデルとLFqモデルそれぞれの解が一致した画素を再抽出し、新たなトレーニングサンプルとする方法を提案する。このトレーニングデータを用いれば、2群判別精度が向上するだけでなく、LFnモデルとLFqモデルの分級評価結果が類似して行くことを示し、提案したトレーニングデータの再設定方法は、対象領域全域にわたって分級評価精度を向上させる上で実用性の高いことを述べる。このステップの検討内容は、第5章で述べる。

4. 土地分級評価へのニューラルネットワークの導入

(1) ニューラルネットワークの概要

本研究で使用するニューラルネットワークの構造を図一3に示した。この構成の内容と特徴については後述するが、ここでは以降の説明を判りやすくするために、3層構造のニューラルネットワークについて説明する。

j 番目の中間層ユニットへの入力層ユニットからの総入力 X_j は、次式で定義される。

$$X_j = \sum_i Y_i \cdot W_{ij} \quad (1)$$

ここで、 Y_i : j 番目の中間層ユニットとシナプス結合した i 番目の入力層ユニットからの出力値

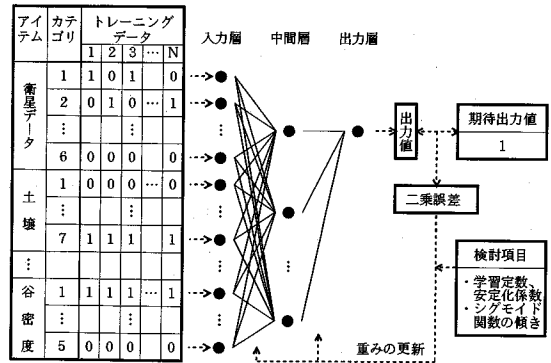
W_{ij} : シナプスウェイト (以下、重み)

この入力値 X_j を受けた後の中間層ユニット j からの出力値 $f(X_j)$ は、

$$f(X_j) = 1 / (1 + \exp(-X_j/T)) \quad (2)$$

で表される非線形関数を用いて定義される。なお、 T はシグモイド関数の形状を決めるパラメータ (以下、形状係数) である。

中間層と出力層の間も式 (1) および式 (2) によって結合される。出力層の値は期待出力データ d_i (教師データ) と比較され、次式で示す誤差 E を最小化するように最急降下法でネットワークの重み W_{ij} を更新していく。いわゆるバックプロパゲーション法である。



図一3 LFnモデルの構成

$$E = (1/2) \sum_i (d_i - Y_i)^2 \quad (3)$$

具体的には、学習信号を

$$\delta_i = (d_i - Y_i) f'(X_i) \quad (4)$$

として、次式にしたがって重みが更新されていく。

$$\Delta W_{ij}(n+1) = \eta \cdot \delta_i \cdot Y_j + \alpha \Delta W_{ij}(n) \quad (5)$$

ここで、 n : 重みの更新回数

η : 学習定数、 α : 安定化係数

以上のことから、ニューラルネットワークの基本構成要素として、 (η, α, T) が重要となることが判る。この点も含めてニューラルネットワークを土地分級評価へ導入する上での前提条件等について検討を進める。

(2) ニューラルネットワーク導入上の前提条件

ニューラルネットワークを導入した土地分級評価モデルの定式化に際して、以下の範囲を設定した。

a) 層数について

中間層の数を変えて解の精度を検討する事例が多いが、学習定数や安定化係数、ユニット数、シグモイド関数の形状を決定付ける係数との関係から検討すべき組合せ数が膨大となるだけでなく、判別精度を比較検証する上での信頼性も問題となる。そこで、本研究では実用性を重視し、最も基本的で単純な構造である「入力層、中間層、出力層」といった3層構造に範囲を絞った上で、その他のパラメータ (学習定数、安定化係数、シグモイド関数) について検討の組合せを考える。この3層構造のネットワークを用いても十分に分級評価精度を高められることを示し、提案手法の有効性を示す。

b) ユニット数について

表一1に示したように地理情報や衛星データは、カテゴリに分割され名目尺度として扱われることから、各カテゴリに対して1つのユニットを対応させる。1画素の構成データを入力するときに、反応するユニットに「1」を割当て、それ以外は「0」とする。

ユニット数を増やすと、コンピュータ処理上の効率が極端に低下することもあり、実用性の点で問題となる。本研究では提案手法の適用効果を示す上で、最も単純なユニット構造とする。つまり、図-3に示したようにアイテムの数（衛星データを含めた各種地理情報の数）と合わせる。このような単純なユニット構造でも、提案手法を用いて分級評価精度を向上できることを示す。

c) 出力ユニットについて

潜在因子モデルは、土地の適性の有無を2群に判別することから、ニューラルネットワークでの出力ユニット数を考えることが必要となる。トレーニングデータに類似する箇所と、そうでない箇所といったように出力ユニットを2つ用意することも考えられるが、土地の適性の有無が明確に区別でき、しかもそれぞれにおいて一様性の高いトレーニングデータを用意することは現実問題として困難である。この問題に対して数量化理論をベースとしたLFnモデルでは、例えば畑地適性評価の場合、現状が畑地として利用されている箇所をトレーニングデータとし、これに類似した箇所をミニマックス2群判別手法を用いて判別するようになっている（現状型分級評価³⁾）。LFnモデルにおいてもこのLFqモデルの構成と同様とし、図-3に示すようにニューラルネットワークの出力ユニットは一つとしている。

(3) LFnモデルの処理の流れ

LFnモデルの処理は、次の3つのステップから成る。

① STEP 1：トレーニングデータに対する学習

トレーニングデータを構成する入力データに対して出力信号が「1」になるように学習を行う。学習が完了するとトレーニングデータを構成する各画素には、実数値が付与されることになる。学習方法は、トレーニングデータを基準として広く一般に利用されているバックプロパゲーション法を採用した。

② STEP 2：評価対象データに対する出力値の計算

学習が完了した後に、評価対象データの入力値をニューラルネットワークに入力し、一画素毎に出力値を付与する。個体に付与される数量は実数値となる。

③ STEP 3：ミニマックス2群判別

トレーニングデータを基準としたミニマックス2群判別を行う。これはLFqモデルと同様であり³⁾、この判別を通して初めてLFnモデルとLFqモデルによる分級評価図の比較ができることになる。ミニマックス2群判別手法を介して、第2章で述べた4つの情報を有する分級評価図を作成する点において、本研究のニューラルネットワークを導入した判別の考え方は従来とは異なる。

(4) 潜在因子の決定

以降の検討を進める上で、まず表-1で示した潜在因

表-2 潜在因子間の相関係数（数量化III類）

	標高	土壌	表層	衛星	地形	斜面方位	谷密度	傾斜区分
標高	○	—	—	—	—	—	—	—
土壌	○	○	—	—	—	—	—	—
表層地質	○	○	○	—	—	—	—	—
衛星データ	○	○	○	○	—	—	—	—
地形	○	○	○	○	○	—	—	—
斜面方位	○	○	○	○	○	○	—	—
谷密度	○	○	○	○	○	○	○	—
傾斜区分	○	○	○	○	○	○	○	○
起伏量	○	○	○	○	○	○	○	○

表-3 潜在因子の選定結果

潜在因子	分級評価項目			
	農村的要因			都市的要因
	畑地	水田	樹林地	宅地
衛星データ	○	○	○	○
表層地質				
地形分類				
土壌	○	○	○	○
標高	○	○	○	○
起伏量	○	○	○	○
谷密度	○	○	○	○
斜面方位	○	○	○	○
傾斜区分		○		○

子のうち実際に分級評価に使用する因子を絞り込む。

LFqモデルでは、数量化III類によってあらかじめ設定した潜在因子間（表-1）の相関係数を算出するとともに、互いに相関の高い潜在因子のいずれか一方を除去した上で分級評価図を作成するようになっている³⁾。これは、数量化II類の信頼性を高める上で重要なプロセスに位置付けられる。このとき、相関の高い潜在因子の組を判断する際の「しきい値」が問題となる。筆者らの今までの検討^{3),4)}と統計処理上の一般的な判断として、本研究では、相関係数が0.7以上のものを削除の対象となる組とした。

一例として畑地適性分級評価について、数量化III類によって算出した潜在因子間の相関係数を表-2に示した。表中、相関係数が0.7以上のものを枠で囲って示してある。さらにこれらの組のうち、いずれか一方を除去する際の指標として数量化II類の偏相関係数を用いる。表-2では判り易くするために偏相関係数の高い潜在因子を上から順に記載してあるので、表層地質因子と地形因子および起伏量因子が除去されることが判る。最終的に残される潜在因子は「標高、土壌、衛星データ、斜面方位、谷密度、傾斜区分」の6種類となる。

他の分級評価項目についても同様な処理を通して潜在因子を選定するが、紙面の都合上、最終的に選定された潜在因子を表-3に整理した。ネットワークへの入力アイテムもこの表で示されたものとする。相関の高いデータをネットワークに入力すると解の信頼性が保証されなくなることもあり、数量化理論による分級評価結果と比較する上でも以上の処理は重要となる。

表-4 パラメータの検討

ケース	T	η	α	最大の中率(%)
1	0.5	0.1	0.8	65.3
2	1.0	0.1	0.9	67.0
3	1.5	0.1	0.9	66.3
4	2.0	0.1	0.9	65.8

注) T: シグモイド関数の形状係数
 η : 学習定数、 α : 安定化係数

(5) ニューラルネットワークのパラメータの決定

a) 検討ケース設定の考え方

ニューラルネットワークの基本構成(層数, ユニット数)と使用する因子が決まった後に, さらに検討すべきパラメータとして前述した「シグモイド関数の形状係数 T, 学習定数 η , 安定化係数 α 」が挙げられる。

シグモイド関数の形状係数は小さくなるほど(関数形状の傾きが大)学習速度が速くなるが, 学習精度が低下するといった報告もある¹⁰⁾。 η を大きくすると学習速度が速くなるが, 出力ユニットからの出力値と教師信号との誤差が振動する可能性があり, ネットワークにおいて重み係数が一定の値に収束しない。逆に小さく設定すると重み係数は一定の値に収束するが, 学習速度が遅くなるといった問題がある。 α によって振動誤差を抑制することができるが, η, α の値の組合せについては試行錯誤的に決定されている。また, 衛星データの土地被覆分類において, η の値と分類精度について検討した例もあるが¹¹⁾, α との関係までには言及されていない。各種の適用事例を見ても(T, η, α)の3つのパラメータの組合せについて検討した例は意外にも見当たらない。これは, 層数やユニット数との組合せを考えると膨大な数にのぼり, 試行検討と解の比較分析におのずと限界があるためである。

そこで, 本研究では層数とユニット数を絞った上で, η, α, T の値について検討することとした。このような検討の流れは, 土地分級評価のみならず, 適用分野別にニューラルネットワークの構成を考えていく上で見逃されている点と考えている。

b) パラメータの決定

Tの値は一般に適用されている「0.5, 1.0, 1.5, 2.0」の4ケースを考える。それぞれのケースに対して η, α の値を0.1~0.9まで変化させ, トレーニングデータを基準としたミニマックス2群判別精度を表す的中率の値を計算する。検討ケースの総数は,

$$(T: 4 \text{ ケース}) \times (\eta: 9 \text{ ケース}) \times (\alpha: 9 \text{ ケース}) = 324 \text{ ケース}$$

となる。処理結果を表-4にまとめた。

ここでは紙面の都合上, Tの4ケースに対して「最大の中率」を示した η および α の値のみを抽出してある。

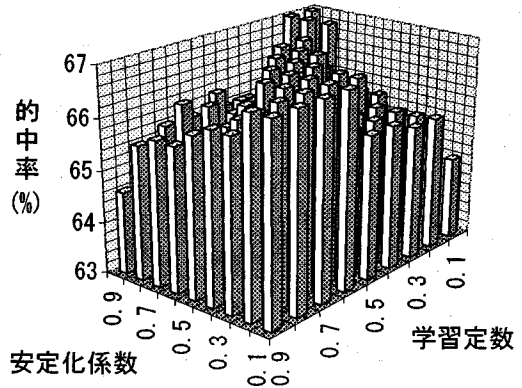


図-4 的中率と(学習定数 η , 安定化係数 α)の関係
 注) シグモイド関数の形状係数 T=1.0の場合であり, 表-4のケース2に対応する。

さらに, 表中の最大の中率のうち, 最も高い値を示したケース2の全ての組合せ事象についてグラフ化したものを図-4に示す。

この検討から, (T, η, α)=(1.0, 0.1, 0.9)で分級評価図を作成すれば最も高い2群判別精度が得られることが判る。的中率が最も高い(T, η, α)の組合せを見つけた際には, コンピュータ内で自動処理され, その後分級評価図の作成処理へ移行できるようにシステム設計および操作面での配慮もなされている。

以上のように, ニューラルネットワークのパラメータを決定した後, トレーニングデータの再設定処理について論述を進める。

5. トレーニングデータの再設定方法の提案

(1) 処理の流れ

本研究で提案するトレーニングデータの再設定処理の流れは, 図-5に示す7つのステップから成る。

① STEP 1: トレーニングデータの初期設定

詳細な現地調査のもとに, 候補となるトレーニング領域を選定する。本研究では, 現状の土地利用を基準としてトレーニング領域を設定した。例えば畑地適性評価であれば, 現状が畑地である箇所となる。

② STEP 2 および STEP 3: トレーニングデータの判別

LFnモデルおよびLFqモデルそれぞれを通して数量が付与された画素に対してミニマックス2群判別を行い, トレーニングデータにおいて「適地」として判別された画素を抽出する。

③ STEP 4: トレーニングデータの再設定

LFnモデルとLFqモデルともに「適地」として判定されたトレーニングサンプルのみを抽出し, 新たなトレ

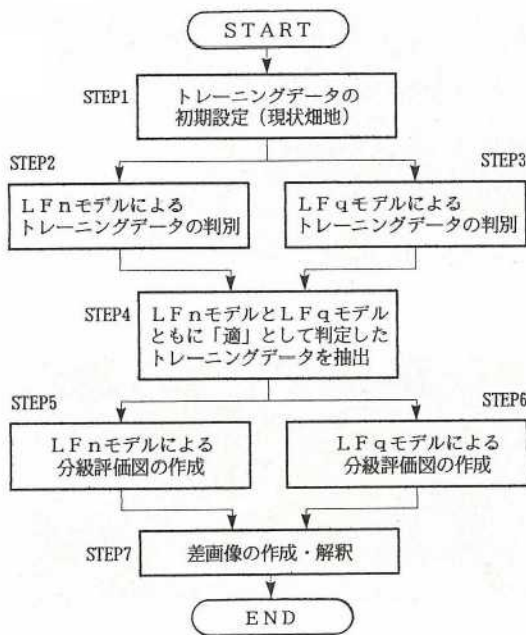


図-5 トレーニングデータの再設定処理の流れ

ニングデータとして再設定する。

④ STEP 5 および STEP 6 : 分級評価図の作成

再設定されたトレーニングデータを用いて、LFn モデルおよび LFq モデルそれぞれによって分級評価図を作成する。

⑤ STEP 7 : 差画像の作成・解釈

STEP 5 と STEP 6 で得られた分級評価図の違いをコンピュータ処理を通して色分けして表現した差画像を作成し、さらに解釈表を作成する。以上までの処理は、いわゆるトレーニングデータを代表する土地の性状と異なる領域(画素)を除去するものであり、本研究ではこれを「トレーニングデータの再設定あるいは洗練化処理」と呼ぶ。この適用効果の有無について検討を進める。

(2) 検討ケースの設定

トレーニングデータの再設定方法の適用効果を検証するために、次の3つの検討ケースを設定した。

ケースA: 初期設定したトレーニングデータを用いて分級評価図を作成する場合

ケースB: LFnモデル, LFqモデルそれぞれ個別にトレーニングデータを再設定し, 分級評価図を作成する場合(図-5のSTEP2とSTEP3)

ケースC: LFnモデルとLFqモデルで一致したトレーニングデータを再抽出し, 分級評価図を作成する場合(図-5のSTEP4)



写真-1(a) ケースA: LFqモデル

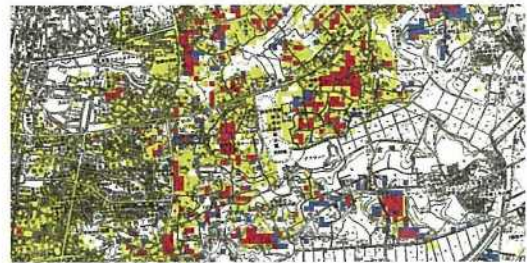


写真-1(b) ケースA: LFnモデル

表-5 分級評価図(写真-1の凡例)

トレーニングデータ (現状畑地)	赤色	現地において畑地のピクセルであるもののうち、畑地として判別されたピクセル。
	青色	現地において畑地のピクセルのうち「畑地適性・無」として判別されたピクセル。 (本来ならば赤色として判別されるべきピクセル)
評価対象データ	黄色	評価対象データのうち「畑地適性・有」として判別されたピクセル。

ケースCがトレーニングデータを再設定する提案手法である。また、ケースBはこの提案手法の有用性(数量化理論とニューラルネットワークの併用効果)を裏付けるために設定したものである。いずれのケースにおいてもそれぞれ LFn モデルと LFq モデルを通して2種類の分級評価図が作成され、これらを相互に比較する。

(3) 分級評価図上に現れる違い

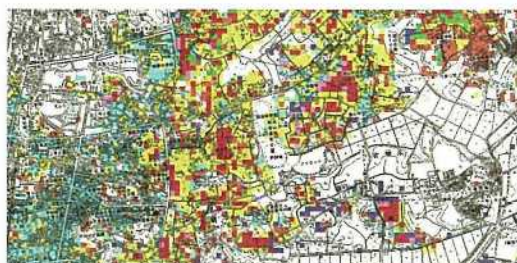
a) ケースAについて

(写真-1(a), 写真-1(b), 写真-2(a))

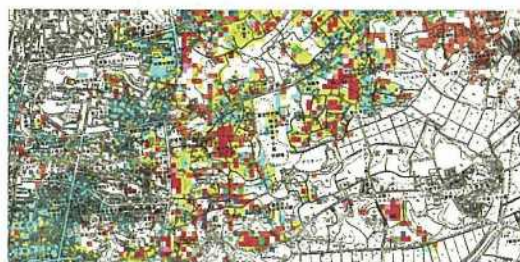
ケースAの分級評価図を写真-1(a)と写真-1(b)に示す。色の意味は表-5のとおりである。これはトレーニングデータを基準としたミニマックス2群判別結果から導かれる情報の区分であり³⁾、土地利用計画策定段階で最も基本となる情報を表現したものとして潜在因子モデルの特徴の一つである。例えば、赤色の部分は将来にわたって畑地として保全すべき地域、青色の部分は他の土地利用への転換可能地域、黄色の部分は今後の農地振

表一六 差画像の作成とその解釈

	分級評価図		差画像の解釈	差画像の色付け写真-2 (a)
	LFqモデル写真-1 (a)	LFnモデル写真-1 (b)		
トレーニングデータ	適	適	畑地保全地域	赤
		不適	LFqモデルのみ「畑地適性・有」と判別した地域。「畑地保全地域」か「他用途への転換候補地」かの判断は現地との照合が必要な地域。	緑
	不適	適	LFnモデルのみ「畑地適性・有」と判別した地域。「畑地保全地域」か「他用途への転換候補地」かの判断は現地との照合が必要な地域。	ピンク
		不適	他用途への転換候補地	青
評価対象データ	適	適	農業振興候補地	黄
		不適	LFqモデルのみ「畑地適性・有」と判別した地域。「農業振興候補地」か「畑地適性のない地域」かの判断は、現地との照合が必要な地域。	茶
	不適	適	LFnモデルのみ「畑地適性・有」と判別した地域。「農業振興候補地」か「畑地適性のない地域」かの判断は、現地との照合が必要な地域。	薄青
		不適	畑地適性のない地域	白



写真一2 (a) ケース A の差画像



写真一2 (b) ケース B の差画像



写真一2 (c) ケース C の差画像

興地域として計画するといった基礎評価を加えることができる。ここで、評価適地以外の白色の部分に加えて分級評価図は4通りの解釈ができることから、写真一1 (a)と(b)の違いを抽出する場合には、表一六のような組合せが考えられる。本研究ではこのようにLFnモデルとLFqモデルによる分級評価図の違いを抽出した画像を差画像と呼ぶこととする。表一六からも判るように差画像では8通りの解釈に拡張される。この解釈にしたがって作成した差画像を写真一2 (a)に示す。

この差画像と表一六の解釈表を用いて多角的な視点から評価が可能となることは既に参考文献4)に詳述してあるので割愛するが、ここで注目すべき点は、LFqモデルとLFnモデルの解が違う箇所、つまり「緑色、ピンク色、茶色、薄青色」で表現されている画素である。特に薄青色の画素が目立つが、これはLFnモデルによって「畑地適性有」として判別された画素が多いためである。初期設定トレーニングデータを用いた場合に、LFqモデルとLFnモデルにおいて、これ程多くの違いが現れることは、分級評価を行う上で問題となる。この違いを出来る限り少なくすることが本研究の目標となる。

b) ケース B について (写真一2 (b))

ケース B の差画像を写真一2 (b)に示す。LFnモデルとLFqモデルによって個別にトレーニングデータを再設定したとしても、分級評価図はケース A よりもかえって解の違いが多くなっている。この点は次項で述べる「適合率」によって定量的にも現れており、個別にトレーニングデータを再設定する効果はないと言える。

c) ケース C について (写真一2 (c))

ケース C の差画像を写真一2 (c)に示す。ケース A およびケース B の差画像と比較すると、ケース C の差画像は、LFnモデルとLFqモデルの分級評価結果が類似していることは明らかである。LFqモデルとLFnモデルを併用してトレーニングデータを再設定する本研究の提案手法の適用効果は極めて高いと言える。

さらに、提案手法(ケース C)の有用性について、次の2つの視点から定量的に分析する。

- ① ミニマックス2群判別を行う際の「個体の分布状態」と2群判別精度を表す「的中率」
- ② LFnモデルとLFqモデルによる分級評価結果がどの程度似通っているのかを表す「適合率」

(4) ミニマックス2群判別の的中率

潜在因子モデルでは、トレーニングデータを基準とした2群判別にミニマックス2群判別手法を用いている。

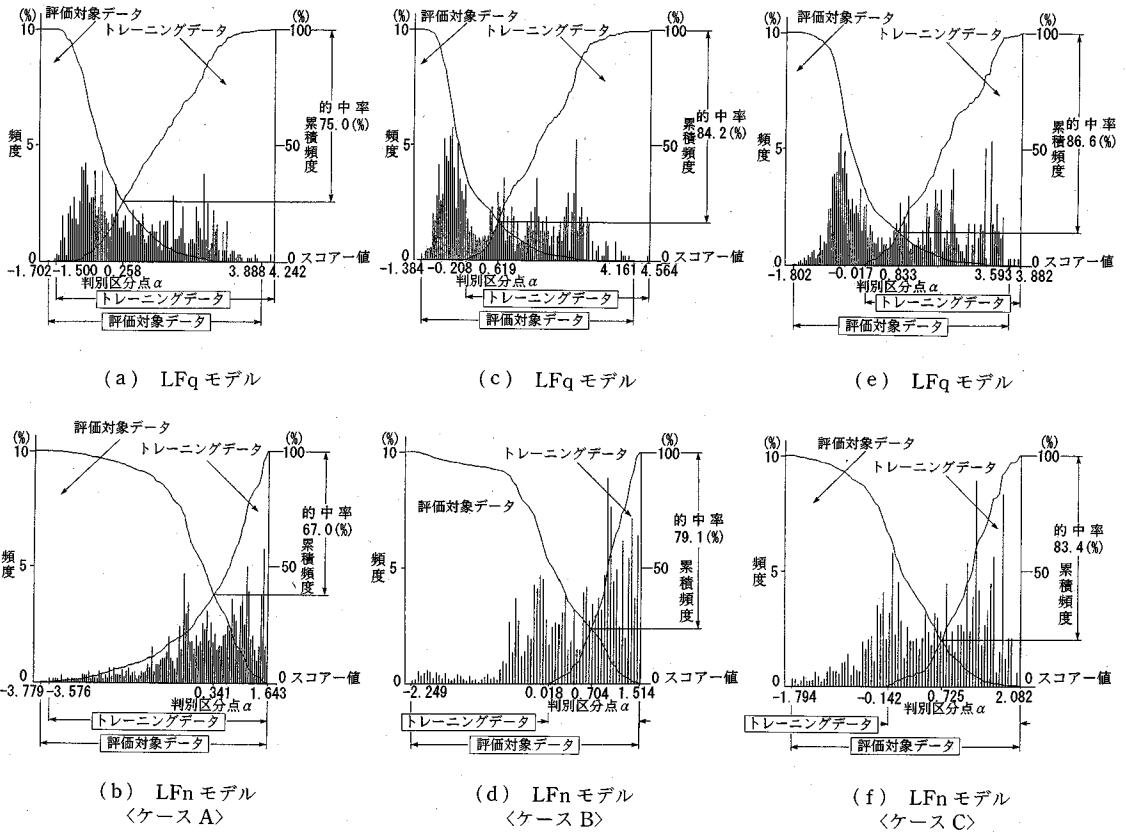


図-6 トレーニングデータと評価対象データの分布状態 (ミニマックス 2 群判別)

ケース A~ケース C までのミニマックス 2 群判別結果を図-6 に示した。この図に示されるように負極側と正極側からそれぞれトレーニングデータと評価対象データの累積頻度曲線を描く。つまり、トレーニングデータを A グループ、評価対象データを B グループとすると、

$$F_A(X) = 1 - \int_{-\infty}^X f_A(x) dx \quad \left(\int_{-\infty}^{+\infty} f_A(x) dx = 1 \right) \quad (6)$$

$$F_B(X) = \int_{-\infty}^X f_B(x) dx \quad \left(\int_{-\infty}^{+\infty} f_B(x) dx = 1 \right) \quad (7)$$

(但し、 $f_i(x)$: $i=A, B$ グループのスコア値 x の度数) と表現でき、2 つの累積頻度分布 $F_A(X)$, $F_B(X)$ の交点が求める判別区分点 α となる。すなわち、

A グループとして判定される領域: $X \geq \alpha$

B グループとして判定される領域: $X < \alpha$

であり、このときの判別の精度を表す的中率 P は、

$$P = 1 - F_i(\alpha) \quad (i=A, B) \quad (8)$$

として求められる。潜在因子モデルにおけるミニマックス 2 群判別に関わるその他の詳細は、参考文献 3) に詳述してあるので、紙面の都合上割愛する。

a) ケース A について (図-6 (a) と (b))

個体の分布状態は LFq モデルではマイナス側に、LFn

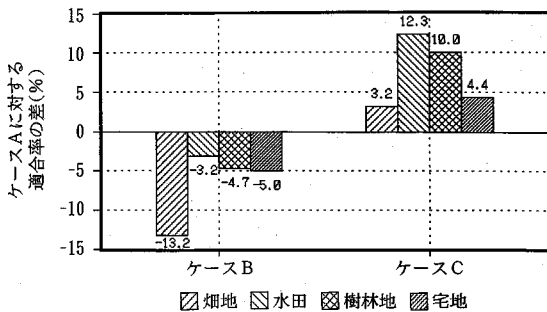
モデルでは、プラス側に分布する傾向にある。的中率は、LFq モデルの方が 8 % も高い。このことが、写真-1 (a) と (b) の違い、すなわち写真-2 (a) の差画像となって現れている。

b) ケース B について (図-6 (c) と (d))

的中率は、ケース A に比べて高くなっているが、ケース A にも増して個体の分布状態は、LFq モデルではマイナス側に、LFn モデルではプラス側に偏る傾向が一層強くなっている。つまり、トレーニングデータの再設定を LFn モデルと LFq モデル個別に行くと個体の分布状態がかけ離れ、ケース A よりもさらに分級評価図の違いが多くなることは避けられないと言える。

c) ケース C について (図-6 (e) と (f))

ケース A とケース B に比べて、個体の分布状態は互いに中央部に歩み寄ってきている。LFn モデルと LFq モデルを併用してトレーニングデータを再抽出する効果が認められる。この効果が写真-2 (c) に現れ、LFn モデルと LFq モデルによる分級評価図の違いも少なくなってきたと言える。この違いを定量的に把握するために、適合率を用いてさらに検証を進める。



図一七 ケース A に対する適合率の差

(5) 適合率

LFn モデルと LFq モデルによって作成される分級評価図の違いを定量的に評価する指標として、本研究では式 (9) の適合率を定義する。

$$\text{適合率 (\%)} = \frac{P(\text{LFn} \cap \text{LFq})}{P(\text{LFn} \cup \text{LFq})} \quad (9)$$

ここに、

$P(\text{LFn} \cap \text{LFq})$: 評価対象データのうち、LFq モデルと LFn モデルともに適地と判別された画素の積集合 (写真一 2 の黄色の画素)

$P(\text{LFn} \cup \text{LFq})$: 評価対象データのうち、LFn モデルと LFq モデルそれぞれにおいて適地と判別された画素の和集合 (写真一 2 の黄色、茶色、薄青色の画素)

ここでは判り易くするために、ケース A に対するケース B とケース C の「適合率の差」を縦軸にとってグラフ化したものを図一七に示した。ケース B はすべての分級評価項目についてケース A よりも適合率が低下していることが判る。このことは、LFq モデルと LFn モデルそれぞれ個別にトレーニングデータを再設定すると「的中率は向上するが、互いの分級評価図はかえってかけ離れた結果になる」ことを意味する。この影響が写真一 2 (b) に現れたことになる。つまり、数量化理論とニューラルネットワークそれぞれの研究領域において判別精度を何らかの方法で向上できたとしても、互いの解が違ってくることは無視できない問題となる。

一方、ケース C はすべての分級評価項目においてケース A よりも適合率が向上している。本研究で提案したトレーニングデータの再設定方法の適用効果は高い。トレーニングデータに対する的中率 (2 群判別精度) が高まることに加えて、各々の分級評価図の結果が歩み寄ってくるのが、評価図全体の精度が高まったと判断する上での新たな必要条件として重要な評価要素となり得る。

畑地適性評価の場合、現状が畑地となっている箇所をトレーニングデータとして選定するが、必ずしも土地の

性状が同じ条件の箇所であるとはかぎらない。平地な箇所のところもあれば、段々畑のように傾斜地に立地している箇所もある。これらを融合してトレーニングデータとして設定すると、判別精度が低下する。このような場合に、トレーニングデータの精度 (一様性) を高めることが必要となり、本研究で提案したトレーニングデータの再設定方法の適用意義がある。

6. まとめ

本研究は、著者らが開発してきた数量化理論を中核とした土地分級評価モデル (潜在因子モデル) に、新たにニューラルネットワークを導入し、分級評価精度を向上させる方法を提案したものである。

ニューラルネットワークと数量化理論を中核とした分析機能をそれぞれ、LFn モデルおよび LFq モデルと定義した上で、分級評価精度について検討を進めた。得られた成果は次の 4 点にまとめられる。

①ニューラルネットワークの基本構造を検討する場合には、層数とユニット数を設定した上で、シグモイド関数の形状係数 T の値別に、学習定数 η 、安定化係数 α の組み合わせについて検討するといった手順が重要であることを示した。システム利用者は、的中率を指標として最も高い 2 群判別精度を示す (T, η, α) の値を決定できる。この処理はコンピュータを通して自動決定でき、効率的に精度の高い分級評価図を得ることができる。

② LFn モデルと LFq モデルそれぞれから作成される分級評価図を比較し、有意な違いが現れることは避けられないことを示した。このことから、トレーニングデータ内に混在する画素の除去の必要性について強調するとともに、ニューラルネットワークと数量化理論を併用したトレーニングデータの再設定方法を提案した。

③この提案方法を用いれば、トレーニングデータの一様性が向上するだけでなく、LFn モデルと LFq モデルによる分級評価図の違いが減少することが示された。換言すればトレーニングデータを基準とした対象領域全域にわたる分級評価精度の向上が示されたことになる。

④ LFn モデルと LFq モデルの解は、完全には一致しない。本研究では、この解の違いを評価図上に表現した差画像 (写真一 2) とその解釈表 (表一 6) を作成した。これらを活用すれば、現地調査時の支援情報となるだけでなく、トレーニングデータと土地の性状が異なる箇所の分析が多角的な視点から展開できる。

今後の課題は、次の 2 点にまとめられる。

①多段階処理の導入: 写真一 2 (c) から判るように、LFq モデルと LFn モデルの解は完全には一致しない。解が異なる画素については、さらに LFq モデルあるいは LFn モデルによって判別するといった多段階処理の導

入について検討することも興味深い課題となる。

②ネットワークの構成：本研究では、ネットワークを3層構造とした上で、シグモイド関数の形状係数に対する学習定数および安定化係数の組合せを変えて検討した。ネットワークの構成要素を決定する上でこのような検討の手順が重要であると考えている。この研究の範囲内において分級評価精度を高められ、当初の研究目標を達成できたが、中間層の層数やユニット数の違いに伴う分級評価精度への影響については今後の課題として残されている。これらの諸条件について検討することにより、今以上に高い分級評価精度が得られることも期待できる。

衛星データや地理情報を融合利用しようとする研究分野では、限られた狭い範囲での手法論、精度論的な研究が進められる傾向にあると言っても過言ではない。利用者にとって、いったいどの手法を採用したらよいのか、得られる解の精度をどのように評価したらよいのかといった根本的な問題も指摘されている。数量化理論とニューラルネットワークの解の違いを検討し、これら2つの手法を併用して分級評価精度を高めようとする本研究の着想と意義はまさにこの点にある。

衛星マルチスペクトルデータと地理情報の融合利用技術、特に空間データの分析モデリング機能の研究開発は、現在世界的にも注目されている研究課題の一つとなっている。本研究がこの方面の研究にも何等かの形で寄与できれば幸いである。

参考文献

- 1) 中村英夫, 川口有一郎, 清水英範, 齋藤林, 柴崎亮介: 地理情報システムを用いたシステム分析的都市計画, 土木学会論文集, No. 476/IV-21, pp.67-76, 1993年10月。
- 2) Lucas, L.F. and Marijke, N.J.: Integrating topographic data with remote sensing for land cover classification, *Photo Eng. and Remote Sensing*, Vol. 57, No. 6, pp.689-697, 1990.
- 3) 小島尚人, 大林成行: 衛星マルチスペクトルデータを適用した分級評価モデルの開発, 土木学会論文集, No. 427/VI-14, pp.65-74, 1991年3月。
- 4) 小島尚人, 大林成行: 土地分級評価モデルへの衛星マルチスペクトルデータ適用の有効性について, 土木学会論文集, No. 453/VI-17, pp.87-96, 1992年9月。
- 5) 小島尚人, 大林成行: 衛星マルチスペクトルデータを適用した土地利用構想図の作成に関する研究, 土木学会論文集, No. 504/VI-25, pp.91-100, 1994年12月。
- 6) 安田登, 白木渡, 松島学, 堤知明: ニューラルネットワークに基づいたコンクリート構造物点検技術者の思考過程の評価, 土木学会論文集, No. 427/VI-14, pp.65-74, 1991年3月。
- 7) 堀口晃, 大山力: 事故区間判定への判別分析理論の適用に関する一検討, 電力技術研究会資料, 電気学会, Vol. PE-91, No. 46, pp.75-84, 1991年。
- 8) Kanellopoulos, I. and Wilkinson, G.G.: Integration of neural network and statistical image classification for land cover mapping, *IGARSS' 93*, Vol. II, pp.511-513, 1993.8.
- 9) 山本由紀代, 築城幹典, 佐々木寛幸, 須山哲夫: ニューラルネットワークを利用した草地開発適地評価モデルの開発, 日本写真測量学会年次学術講演論文集, pp.89-92, 1993年5月。
- 10) 島崎敏一, 安田誠一: ニューラルネットワークによる交通手段選択モデル, 土木学会論文集, No. 494/IV-24, pp.79-86, 1994年7月。
- 11) 張立堅, 星仰: バックプロパゲーション学習法についての学習率の設定, 日本写真測量学会年次学術講演集, pp.11-14, 1994年5月。

(1995.8.28 受付)

IMPROVEMENT OF THE ACCURACY ON THE LAND USE CAPABILITY CLASSIFICATION APPLYING NEURAL NETWORK

Hirohito KOJIMA, Shigeyuki OBAYASHI and Hidetoshi YAMAMORI

The objective of this study is to improve the accuracy of the land-use-capability-classification model which is called the "LF (Latency Factor) model" developed by ourselves. In this LF model, we use not only geographical information but also satellite multispectral data. As for the two kernel functions in the LF model, "LFn model" and "LFq model" applying the Quantification method type II and the Neural Network respectively have been newly defined. As the similar area for the characteristic of the training data was discriminated, the land use capability classification map made by using LFn model and LFq model respectively must be ideally corresponded. Based on this basic concept, the purification procedure for the training data was proposed, which is extracted the corresponded pixels of discriminated results with both LFn model and LFq model. Through this procedure, the improvement of 15% or so on the land use capability classification accuracy could be achieved. Furthermore, it was confirmed that the differences of the land use capability classification maps with both of LFn model and LFq model was remarkably decreased.

非線形構造解析プログラム

MSC/ABAQUS

Version 5.4

MSC/ABAQUSは、広範囲な構造問題のシミュレーションを可能にする汎用有限要素法解析プログラムです。

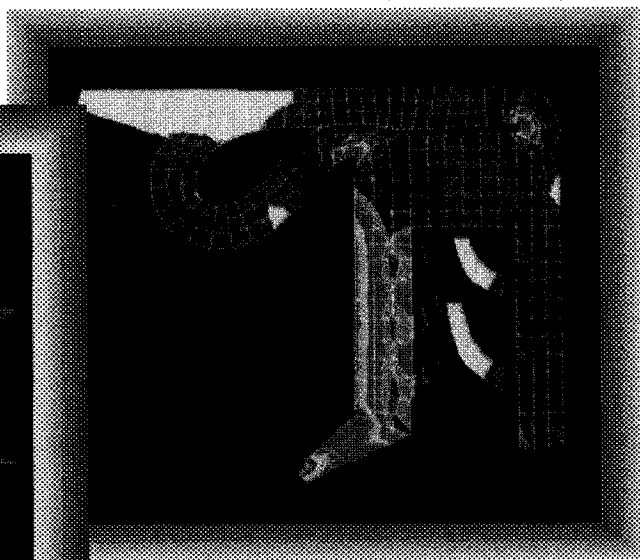
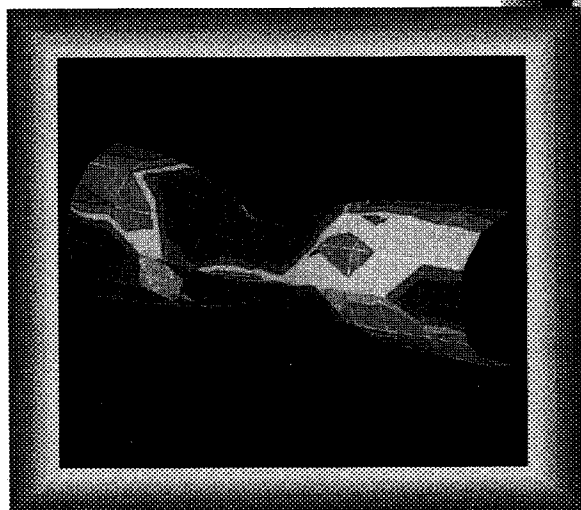
非線形解析に特化した長年の開発経緯を持ち、接触、材料破壊等の強烈的な非線形性を含む問題まで対応が可能です。モデル化のために豊富な要素、材料、解析機能ライブラリが用意され、

平易なインプットスタイルと自動増分解析機能により、

非線形解析シミュレーションが容易に行えます。

MSC/ABAQUSは、MSC/PATRANとの併用により

トータルなCAE環境をお約束します。



- コンクリート、岩盤等のダメージモデルを含む材料非線形大変形問題に適用（材料実験データ入力可能）
- 自動増分解析機能、修正Riks法による、オートマティックな非線形シミュレーション



日本エムエスシー株式会社 営業部

東京本社：〒107 東京都港区赤坂5丁目2-39 円通寺ガレリウスビル

TEL.03-3505-0269

FAX.03-3505-0914

関西支社：〒604 京都市中京区烏丸通六角下ル七観音町626 烏丸小泉ビル

TEL.075-255-3881

FAX.075-255-3515

Mr. SOIL Ver. 3.5 for Windows

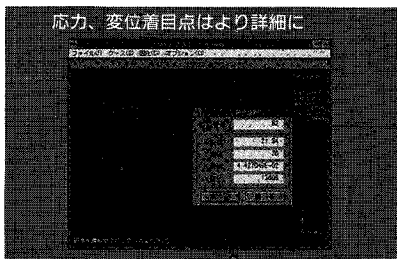
It's NEW!

業界すべてのエンジニアへ
優れた機能を抜群のコストパフォーマンスで。

簡易内臓CADによるメッシュエリア作成



応力、変位着目点はより詳細に



Step 1

オートメッシュで簡単に
メッシュデータ作成

Real Easy! Real Speedy!

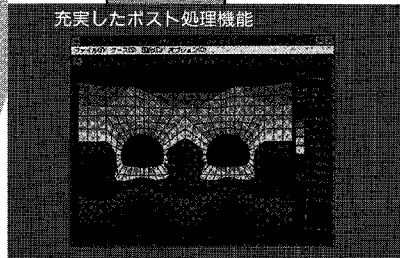
Mr. SOILは地盤の応力、
変形特性、土木構造物との
相互作用を総合的に解析可能
な強力なFEM解析ツールです。

販売実績：400本超
(国内外含：'95.3現在)

Step 4

さらに詳細検討へ

充実したポスト処理機能



Step 2

高速、高精度ソルバーへ

3000点規模の解析も可能



Step 3

各種ポスト処理へ

●動作環境

CPU：80386以上 (推奨 80486以上)
Windows 3.1
メモリ840Kbyte+7Mbyte以上
ハードディスク10Mbyte以上

●価格 ¥980,000- (税別)

旧バージョンをご使用中のお客様は、
バージョンアップ価格を、また、2本目以
降マルチライセンス価格も設定しておりま
すのでお問い合わせ下さい。

*Windowsは米国マイクロソフト社の商標です。

■詳細資料または、ご検討のためのDemo Systemをご希望のお客様は、
下記の問い合わせ先までご連絡下さい。

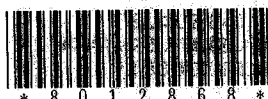
土木CAE業界をリードするCRC土壌ソフトウェアラインアップ

地下水解析：3次元地下水解析...SEEPAGE-3D
広域地下水変動解析...UNISSF(V-2)、PC/UNISSF
地質解析：3次元地質解析...GEORAMA
地盤解析：2、3次元地盤解析...Mr. SOIL
連成解析：土・水の連成(逆)解析...UNICOUP
岩盤・地盤解析：個別要素法...UDECC、3DECC
*これらのソフトを使用しての受託解析業務も行ってあります。

●開発/発売元 **CRC** 株式会社CRC総合研究所

日本技術開発総資料室

営業部/担当：村中
太郎 4-1-3
FAX：06-241-4136
エンジニアリング部
2-7-5
19 FAX：03-5634-7337



* 8 0 1 2 8 6 8 *

土と水の連成逆解析プログラム

未来設計企業
CRC

UNICOUP

応力解析と浸透解析がドッキングした!

軟弱地盤の解析に!

海洋開発・埋立

盛土・掘削

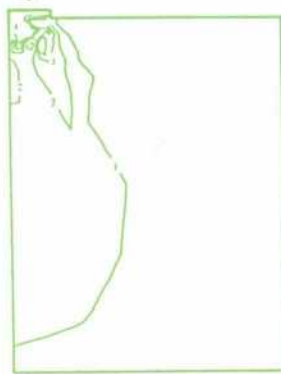
出力項目

- 各節点での変位、各要素での応力
- 各節点での全水頭・圧力水頭 他
- 豊富な図化処理
変位図、変位ベクトル図、応力ベクトル図、応力コンター図、安全率コンター図、水頭コンター図、圧力水頭コンター図

プログラムの特長

- 応力と地下水の流れをカップルさせた問題が解析可能です。(圧密含む)
- 地下水の流れは飽和・不飽和域を対象としています。
- 多段掘削・盛土や降雨等が扱えます。
- 柔や連結要素も扱え実用的です。
- 経時観測記録(変位・水位)があれば、非線形最小二乗法に基づき変形係数や透水性係数が逆解析できます。(順解析、逆解析がスイッチにて選択可能です。)
- 弾性・非線形弾性・弾塑性・弾粘塑性を示す地盤が扱えます。
非線形弾性(電中研式、タンカン・チャンの双曲線モデル)
弾塑性(ドラッカー・ブラガー、モール・クーロン、カムクレイモデル、ハードニング、ソフトニング)
弾粘塑性(関口・大田モデル)

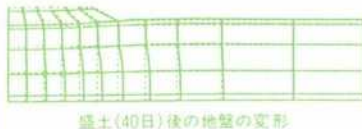
(荷重)



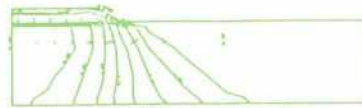
応力増分コンター (J・e・V)
(10日後)



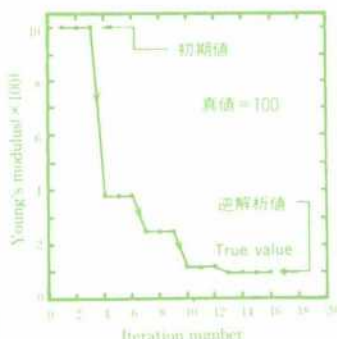
変位ベクトル図 (40日後)



盛土(40日)後の地盤の変形



盛土(40日)後の地下水の流れと水頭
コンターおよび自由水面



ヤング率と繰り返し回数
の関係
逆解析によるパラメータの推定

この製品は、情報処理振興事業協会の委託を受けて開発したものです。
通商産業省 特別認可法人

IPA 情報処理振興事業協会

株式会社 **CRC** 総合研究所 西日本事業部

〒105 東京都港区芝公園三丁目1番38号
TEL. (03) 3437-2301

問合せ先

〒541 大阪市中央区久太郎町4丁目1-3
(06) 241-4121 営業担当: 岩崎
(03) 3665-9741 本社窓口: 菅原