

平成 30 年 5 月 30 日現在

機関番号：13301

研究種目：挑戦的萌芽研究

研究期間：2016～2017

課題番号：16K12658

研究課題名(和文) 疎林はどのように疎林か。また何故。 - 乾燥地植林のための個木インベントリ -

研究課題名(英文) How sparse are sparse trees, and why?

研究代表者

川西 琢也 (Kawanishi, Takuya)

金沢大学・自然システム学系・准教授

研究者番号：80234087

交付決定額(研究期間全体)：(直接経費) 1,700,000円

研究成果の概要(和文)：西オーストラリア疎林の個木分布状況を解析するとともに、分布の幾何学形状の統計的解析手法について検討した。前者については、画像解析による個木の位置情報取得について、ほぼ目処が立つところまでできた。後者については、一般極値分布 (generalized extreme value distribution, GEV) の指数 (extreme value index, EVI) を求めるための手法を提案した。新手法はこれまでの手法と計算コストは変わらないが誤差は小さく、最尤法より良い近似となる。

研究成果の概要(英文)：We have been tried to develop methods for characterizing the sparse trees in arid region to facilitate the afforestation for carbon sequestration. During the analysis of sparse trees, we came up with a new method of statistical parameter estimation. In the analysis of sparse trees, we concern the generalized extreme value distribution (GEV). For some range of parameters, the maximum likelihood (ML) method does not work for estimation of GEV parameters. Among the alternatives of ML, the method of maximum spacings (MPS) are convenient because it can be applied wider range of parameters and relatively small error of estimation. Our method is a variant of MPS and it is in general less bias than MPS, and theoretical more proxy to ML.

研究分野：化学工学

キーワード：統計 環境 パラメータ推算

## 1. 研究開始当初の背景

申請者は西オーストラリア州Leonora近郊における植林プロジェクトに参加し研究を進めてきた。このプロジェクトでは、過放牧で裸地になったところに植林することに成功し、これが温暖化対策になりうることを明らかにした。次の段階として、植林可能な場所の探索が課題となっている。

## 2. 研究の目的

本研究では、衛星データから、乾燥地の植生パターンを解析することにより、植林適地を見つけ出す手法の開発を目指す。

## 3. 研究の方法

研究期間内に、Leonora近郊の1000 km × 1000 kmの地域において、衛星データから個木位置情報を算出しデータベース化、幾何学的指標、統計的指標を用いて疎林のキャラクター化および分類を試み、さらには、気象情報、地理情報などから疎林生成要因の解析を行う。研究期間内に以上について手法を確立し、できる限りの自動化を行う。以上に基づいて、全球レベルの解析のフェージビリティスタディを行う。

この研究期間では、結果的に、統計的指標の算出のためのパラメーター推算法の開発に注力することになった。これは、統計的手法の maximum product of spacings (MPS)法の改良を行ったことになる。

## 4. 研究成果

西オーストラリア疎林の個木分布状況を解析するとともに、分布の幾何学形状の統計的解析手法について検討した。前者については、画像解析による個木の位置情報取得について、ほぼ目処が立つところまでできた。後者については、一般極値分布 (generalized extreme value distribution, GEV) の指数 (extreme value index, EVI) を求めるための手法を考案し、

これを提案するに至った。

GEVのパラメーター推定においては、パラメーターの値によっては最尤法が使えないなど、定番の手法がない。諸手法の中ではMPSがほとんどの場合に収束し誤差も比較的小さいため、我々も種々の統計パラメーター推定をMPS法で行ってきた。MPSは、順序統計にもとづく経験分布関数のスペーシングの積を最小にするものであり、最尤法の密度関数の代わりに分布関数を使うものと解釈できる。我々は、このスペーシングの積について、両端の2つのスペーシングをその平方根で置き換えることにより、誤差(バイアス)を大幅に減らすことができることに気がついた。MPS法では(サンプル数+1)個のスペーシングの積を取るが、我々の手法では、実質上サンプル数に等しい数のスペーシングの積をとることになる。結果、MPS法よりも最尤法のより良い近似となる。新方法は、様々なメリットにもかかわらず、推算に要する計算コストはMPSと変わらない。この手法については論文一報を投稿中であり、新たにもう一報準備中である。課題全体としては、予定より遅れることになったが、新たな統計的手法を得たことは大きな成果である。

この改良法(便宜上RAMPS法とよぶ)のポイントは、MPS法がn個のデータから作られたn+1個のスペーシングにより計算されるのに対し、実質上n個のスペーシングを用いることである。実質スペーシングをn個にすることにより、RAMPSのスコア関数は、最尤法のスコア関数の近似として、MPSよりも格段に誤差が小さくなる。また、MPSでは、n+1個のスペーシングを用いることで、スコア関数の極限分布に対する中心極限定理の適用が、最尤法よりも不自然となっていたが、RAMPSでは、n個の、近似的に独立なスコアが得られることにより、中心極限定理の適用がよ

り自然に行える。

理論的には、最尤法でパラメーターが推定できる場合にかんしては、RAMPS のほうが常に MPS よりも良い近似である。これを、最尤法のスコア関数と、MPS、あるいは RAMPS 法のスコア関数とを比較することにより、理論的に示した。

漸近的には、最尤法、MPS, RAMPS 全てが同じ値に収束するが、収束速度は、MPS の収束速度が最尤法より遅いのに対し、RAMPS は最尤法と同等である。例えば、ロジスティック分布の場合、バイアスの収束速度は MPS 法では  $O(\log n/n)$  であるのに対し、最尤法、RAMPS では  $O(1/n)$  である。

サンプル数の小さい場合においては、RAMPS は、一様分布の場合を除いて、MPS、最尤法よりも小さいバイアスを示す。ただし、標準誤差、平均二乗誤差については、分布によって優劣が異なる。

サンプル数が大きいときには、MPS と RAMPS に違いはない。しかし、サンプル数 200 程度でも、RAMPS によりバイアスが著しく改善される場合もある。

原理的には、MPS が使える場合は RAMPS を使うことができる。では、常に RAMPS を使うべきだろうか。一様分布では明らかに MPS が優れているほか、指数分布では MPS のほうが小さい標準誤差を示しているため、必ずしも RAMPS が常に MPS より優れているわけではない。しかし、それらの欠点はマイナーであり、常に RAMPS を選択する、ということのデメリットは常に MPS を選択するデメリットより小さいといえよう。

さらに、多くの計算例で、RAMPS が最尤法よりもよい結果を出していることは注目に値する。この点については、理論的に解析中であるが、大まかに言って、分布の端点の情報  $F(-)$ 、 $F(+)$  における情報

をどれだけ取り込めるか、という点に違いがあると考えている。

現在のところ、最尤法と MPS、RAMPS のパフォーマンスの差の理由は明らかではないが、数値計算で用いられる近似解法の理論が役立つのではないかと考えている。最尤法、MPS、RAMPS は全て Kullback-Leibler divergence を最小にする方法であるが、その最小化問題を重み付残差法で表現したときの離散化法が異なる 3 つのバリエーションと考えることができる。このときに、MPS、RAMPS はガラーキン法を用いた離散化に対応しており、また、境界条件の違いにより MPS、RAMPS が導かれる。

分布が定めれば、境界条件により、適用すべき数式が MPS になるか RAMPS になるかが定まるが、これと、パフォーマンスに関しては、これまで計算した例においてはよく一致しており、この方向での解析が有効であることを示唆している。

乾燥地植林のパターン解析においては、極値統計のパラメーター推定を行うが、これに関してシミュレーションした結果は、多くの場合において、RAMPS は MPS に比べて小さいバイアス、ほぼ同等の標準誤差を与えることが分かった。

MPS では、EVI が -0.5 の近くでバイアスが 0 になるが、この値が -0.5 を離れるにしたがって急速にバイアスが大きくなり、パラメーター推定の観点から、特に MPS を使う理由はない。

また、最尤法を用いることができない、EVI が -1.0 より小さい場合に、概して小サンプルでの最小二乗誤差は RAMPS のほうが MPS より優れている。

乾燥地のデータでは、個木の分布から算出した 2 次データを用いるため、データ数が少ない場合にも対応できるパラメーター推算法が望ましく、この意味でも RAMPS

を得たことは意義深い。

#### 5．主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 0 件)

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

〔その他〕

ホームページ等

<http://t-kawanishi.w3.kanazawa-u.ac.jp>

#### 6．研究組織

(1)研究代表者

川西琢也 (金沢大学・自然システム学系・准教授)

研究者番号：80234087