

## 森林立地研究の統計

西村 五月\*

### 1. はじめに

研究結果には、精度と再現性が要求される。精度を向上させるためには、サンプルの吟味が重要である。ここでは、森林立地研究のような複雑な条件下での資料抽出と再現性向上のための解析法について考えてみたい。

一般の試験報告では、資(試)料と実験方法について詳しく述べるが、それは如何なる条件下において得られた結果であるかの説明であり、同じ条件下においての再現性を期するものでもある。この場合資料の吟味は精度向上に貢献する。そして、資料と条件が同じであれば、何処で誰が試みても同じ結果が得られる筈である。しかし、実験室や限られた条件下でのポット試験、苗畑試験までは、同じ条件設定はできても、立地研究のような現地を対象とする分野には通用しない。

森林立地の研究では、まず二度と再び同じ条件がすべて揃う現場はないといっても過言ではない。現場は大筋では似ていても、何処かに大きく、或は小さい点で異なるものである。このような場合の再現性を期する手法はどうあるべきか？ 再現性が期せねば常に結果は千差万別で学問たり得ない。これまでの手法や結果に対して、疑問を抱く訳ではないが、後で述べるが種々の難点があった事実は否めない。つまり、得られた結果がその対象地域では、非常に高い精度に達したとしても、他の地区への適用という点に関しての配慮が不十分であったことである。

或地点、地域で得られた情報は点情報である。再現性のある手法によって得られた結果は、面情報化することができる。しかし、再現性に欠ければ、それを如何程積上げてても点で面情報になることはない。点情報を全面に並べて面化というならば全面調査しかない。それには時間と金と人が多く必要である。その上、これでは研究ではなくて調査である。

このような悩みの中で、試行錯誤を繰返して20余年、まだ解決した訳ではないが、問題点や留意事項も含めて筆者の研究方法を述べて見たい。

### 2. 資料の吟味

条件が十分に調整管理されているならば、サンプル(資料)の蒐集は容易である。例えば、苗木を同じ条件下で育てて、或は特定条件設定下で育てた場合など、その中から苗の大きさを区分して選び出すことができる。葉の大きさを測定するならば、特定条件を設けたり、全木から均等に

---

\*長崎県総合農林試験場

資料を抽出するように留意する。その結果、正規分布に近い形に分布する資料が得られる。まず、これでサンプリングの第一関門が通過できる。

広域を対象とし、複雑な条件下において森林調査を行うためのサンプリングとは、調査点の抽出と配列である。対象区域は県下全域、〇〇〇森林基本区、△△川流域などのことが多い。ここでこの対象地域が全域森林で、同一地形、同じ地質、同一樹種などというように条件が整っていれば問題はない。しかし、対象地域の中には、平地（農耕地）や市街地、村落などがあり、山あり谷あり、傾斜畑ありで、更に複数の地質が入り合い、複雑な条件の中に複雑な森林分布となっている。この中からの抽出である。

筆者は数年前にある森林調査を依頼されたが、調査点は既に指定されていた。抽出資料は、森林簿から広葉樹林を各林小班毎に齡階別に呼出して、調査に必要な点数を決めて、機械的に抽出したものである。林分位置を5万分の1地図と森林基本図に記入してある。地形図上の配列を見ると資料の偏在は少なく、全域にわたって割付けられているように見える。しかし、これを詳細に見ると、道路の少ない奥地でプロジェクトが少なく、集落周辺の里山に多く割振られている傾向が強い。更に森林基本図上で観察すると面積の狭い林小班が多く抽出されている。広い面積の小班が抽出される場合は稀である。奥地では小さな面積での所有者も少ないが、狭い面積での施業も少ないので、林小班が小さくなることは少ない。この抽出では、林小班の条件の中に面積が入っていないため、すべての林小班の面積は等しいとの仮定になってしまったのである。

筆者は、土壌とその母材および気象を要因とするヒノキの生産力研究<sup>1)</sup>に従事した。その際、気象条件は海から陸地内部に向かって変化する。リアス式海岸のため、海岸部から造林地のこともある。林地が陸地内部へ連続していれば問題なく気象条件の要因は機械的抽出で満たされる。しかし、その間には農地や集落、人工林でない山林などがあり、各種の母材があったり地形が変わったりして、調査点の配列はまゝならない。つまり、条件を整えると必要な場所に林地が存在しない場合に出会う。その結果、調査に欠落した部分が生ずる。これを無視して統計処理をすれば、資料数の偏りのため、精度は別としても再現性の点に不十分さが残る。これを修正するには、資料の欠落部分から生じた歪を裁く必要がある。

このような資料の欠落部分ができないように対象地域の広域化を避けて、似た条件下の特定の小地域を対象とする方法が一般に多い。この方法では、その区域内では高い精度を期待することができる。しかし、この結果は条件の異なる他地区へは移せない。結果の再現性が期待できないのである。類似の例は、気象値の推定で地形図を図幅単位で式化している場合がある。図幅によって式の係数などが若干異り、隣接図幅と併せて用いようとすると、図幅の接点に不整合が生じる。上に述べた小区域対象の場合も全く同じで、主要な作用を持つ因子の種類が区間で違うなどの原因で、この点情報を面化することができない。

### 3. データの検討

林地生産力に関する報告は多い。その中から少し古い<sup>1)</sup>が、比較的広く知られており、調査箇所数も多い事例で数量化による地位指数の推定<sup>2)</sup>を試みたものを取り上げたい。断っておくが、同書には問題点や今後解決を要する事柄が述べてあることは熟知しており、決して批判しようとするものではない。

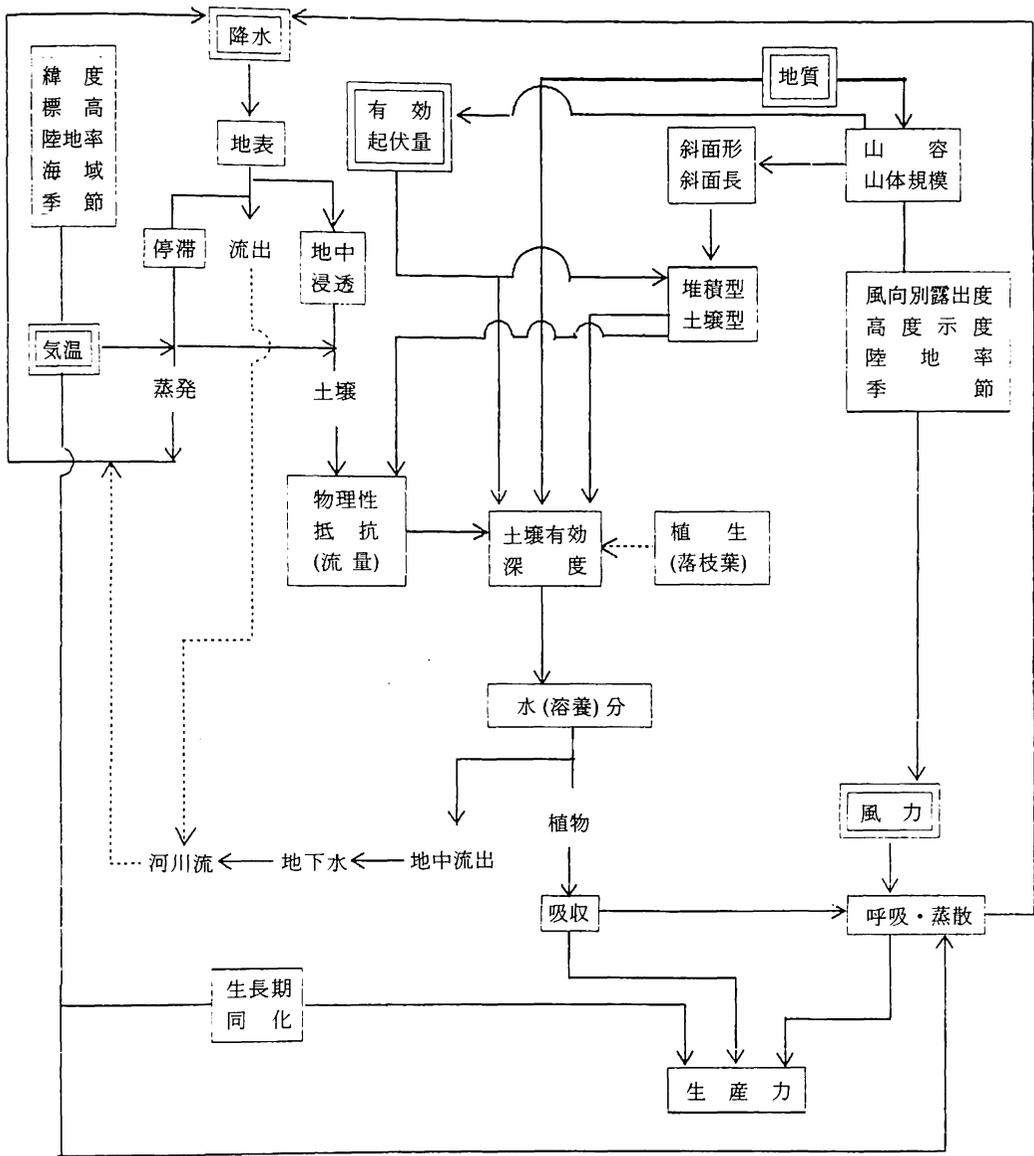
この書の方法では地域を単位にしてある。地域の範囲の定義については大きく意が注がれている。調査結果は各調査区共高い精度である。しかし、地位指数に大きな影響を与える因子はそれぞれ調査区で異なっている。この傾向は本書の例に止まらず、これまでの多く報告も同じで、“主要因は場所によって異なり……”の感がある。こうなると、調査結果の適用範囲は、その調査対象区域内に止まり、他地区への適用ができない。

次いで取り上げられた因子の種類数が多いことに注目したい。因子数が多いことは見落としをなくし、計算上相関係数を高くする利点がある。しかし、因子数が多いと、各因子間で内部相関の高いものが多くなり、類似の作用を複数回にわたって評価するおそれもある。因子数の制限が目的ではない。調査研究の段階での因子数は多い方が良い。要は内部相関の高いものや、一次的作用と二次的作用などに整理することを提案したいのである。また、研究結果を現地に適用する上でも、因子数が多いことは種々難点となる。

地位指数に影響を与える主要因子が、調査対象区によって異なる場合、その理由について客観的、理論的に説明ができればよい。その説明が対象区の環境特性の強調に止まるのでは、客観的、理論的とはいえない。何故なら、それぞれの区でおのおの同じような理由があるからである。ここで、因子の作用を一次的、二次的などに分けて考えると、その整理に片手落ちがなければ、一次的作用の因子の数はかなり減ずると思われる。また、その結果、相関係数を高めるために因子数をふやすことはためらわざるを得なくなる。

筆者は因子の水準選択について仮説を立てた<sup>3)</sup>。その進め方は図に示したとおりである。実際はこの図の中には見落としもあり、未調査の部分もある。しかし、内部相関の存在を探索する意義は認めてもらえるだろう。方法的には、調査目的で若干異なることが予想される。或は、似た作用、内部相関などによってクラスター化をはかり、これをデンドログラム化する方法も考えられよう。

内部相関の例について述べよう。林地生産力の要因解析の報告の中に、有効な因子として標高がしばしば登場する。標高の持つ意味の中で、降水量と気温はその最たるものであろう。山地では標高が大きくなるにつれて、降水量が増大し、気温は低下傾向を辿る。標高が比較的小さい範囲では気温は林木成長を十分満たしているから降水量と共に成長が大きい傾向にあるが、標高が大きくなって気温が低下し、林木成長に不十分になると降水量の増大よりは気温の低下による成長阻害が見えてくる<sup>4)</sup>。したがって、標高はこの両者を区別しておく必要がある。このようにそ



林地生産力の構造模式

の因子の持つ意義・作用を十分に考慮した選択がないと客観的説明ができず、他の地点への適用で再現性を欠くことになる。

#### 4. 要因の作用度の推定

要因の整理が終ると、数量化による因子の作用の推定にかゝる。この中で当初に述べた資料の偏在と、現在の手法が重回帰による直線化した計算である点を問題としたい。資料が偏在したまま代表値を計算で求めると、歪が生じるのは当然である。また、自然界における多くの現象のすべてが直線的関係で表示できるとは考えられない。むしろ、どのような線型を示すかわからない場合が多く、その線型は因子の種類によって特性を示すものと考えた方が、客観的・理論的とも思われる。

上述の条件を満たすには、要因の数値的变化に伴う従属変量との間の曲線変化に、円滑な規則性を保たせながら、次の要因との交互作用を含ませた曲線が得られねばならない。これに適した方法として図相関解析がある。主として気象関係で利用されるが、林学にも例がある<sup>1,3,4,5,6,7)</sup>。この方法では一番初めに、因子の中で成可く大きな影響力を持つ因子を取上げるが、客観的な数値表示がなされていることが必要である。最初の因子 $X_1$ と $Y$ の関係は著しく分散しているが、これを次の因子 $X_2$ で標示し、その分散境界に線を入れる。この図から $X_1$ と $X_2$ の値を用いて推定値 $\hat{Y}_{12}$ を求める。更に $Y$ と $\hat{Y}_{12}$ の座標と $X_3$ の座標を用いて $X_3$ 因子というように次々と総合的な解析を進めてゆく。

$X_1$ と $X_2$ の $Y$ に対する値は、目視的に $X_2$ の標識区分の境界によって求めるが、 $X_2$ の各ランク別に得られた曲線は整然とした関係を示す訳ではない。交絡や逆転なども見られる。そこで、これを全体の形と各ランク間の関係などの理論性を考えて、交絡や前後関係、曲線の形状などを規則的に補正する。以下の各々については目視的に最小二乗的に線を入れて行く。ここで最小二乗法による計算で線を求めてはいけぬ。この時の線は必ず原点を通らねばならない。 $X_2$ 以下の因子は必ずしも数値化されている必要はなく、質的なものでもよい。こうして、作用の大きな因子から、順次役割りの小さい因子へと解析を進める。あたかも篩のメッシュの粗なものから順次細かい方へ選別して行くようなものである。

こうして得られた図は、最終的に各点が $45^\circ$ の傾斜の線の周辺に集ってくる。これらの各解析図から総合的な位置づけで、各因子の単独的、又は相互的な関係を推知することができる。したがって、未知の地点の林地の生産力の推定が可能となる。しかし、この方法は曲線を目視的に求めるため、資料が少ないと苦慮することになる。資料数は各ランクや因子数などにより異なるだろうし、理論的根拠もないが、筆者は300点以上を目標にしている。実際、解析していても100点以上はほしいように思われる。

また、各ランク別に線を求める際に、現場で出現頻度の高い値を示す部分の資料が多くなっており、上又は下などの両端部分の資料が少なくなる。これは自然界の成立ちの関係上避け難い一

面ではある。しかし、この両端部にも線を設けなければならないが、この時は全体の分布の形と理論的關係を考慮して少ない資料を利用して線を決定する。また、独自の線を設ける必要がない時は隣接のランクに繰入れる。この作業は経験を積みれば決して難かしくはない。この意味でも資料数が多いことがのぞましいが、成可く意識した上下両端部の資料がほしい。

## 5. 実際の方法

ここで筆者の調査手順を紹介したい。資料の入手は調査目的の確立に従って、当該林分をリストアップし、5万分の1地図上に記入する。林分数が多くて作業が混乱するようであれば、希望する点数よりはるかに多い数を抽出する。抽出地の分布を地図上で把握しておく。これは資料抽出の円滑の他に、結果の考察の上でも参考になる。いわゆる“山を翔ける”という全体を掌握する作業である。

全体像を得た上でサンプリングに入る。資料抽出には機械的方法がある。一定間隔又は地理的に均等に機械的に調査点を配置する。この方法は一定区域内、一流域内などでは比較的容易に配置できるが、対象区域が広がると前述したように配列は乱れてくる。今一つは主観的配列である。調査の目的や主旨によって主観の条件が異なる。的確な条件設定と鋭い洞察力があれば、大きな効果が期待できるが、往々にして要因の見落としなどがあって、十分な成果を挙げ得ない欠点がある。しかし、多くの人々はこの主観的方法を応用しているものと思われる。

筆者はこの両者を折衷的に併用することが多い。例えば、九州地方のヒノキの地位指数と樹形について検討<sup>9)</sup>したが、その時は次のようにした。九州各県の調査資料1870点の中から300点を目標に抽出した。したがって、全量の16%である。まず県別にその県の資料総数の中から1/6を、資料番号によって機械的に抽出する。地図上に記されている調査位置と抽出番号を照合する。全域に偏りなく資料が配分されていればサンプリングは終了する。若し、抽出された番号の分布が地図上で偏っていればその修正をはかる。資料が集中している部分があれば疎な配列の地区へ移す。集中している部分はないが、疎な部分や飛地があれば、その部分には改めて追加する。このような操作で資料を得ている。

いずれにしても現地調査を行う場合は、選び出された林分が調査に適するか否かの検討が必要である。その判断基準は調査目的により異なる。次いで、対象林分内を一巡し林分内の状況を把握する。目的が生産力調査であれば、林分内で成長の最高と最低を示す部分を選んでプロットを設ける。したがって広い林分でも成長差がなければプロット数は増えない。賦存調査では、その林分の代表的、標準的な場所を調査点とする。

XとYの關係を知るのが目的である調査で、標準的、平均的資料に依存した解析をおこなうと、類似の数値に資料が集中する。この場合、XとYの關係の傾斜度は資料の偏りによって左右される。したがって、計算上相関があることは確認できても、その程度は研究事例で異なることになる。的確にその傾きを知るには、成可く両極端の資料を補充することである。敢えて両極端を漁っ

ても、中間的値の資料は多すぎる位である。このようにして得られた資料を前項で述べた図相関法で解析している。

結果の的確性や再現性の向上を図る上で、もっとも肝要なのは因子の選択・配列と思われる。これは上述のように、因子の整理の基準、因子の重要順の配列が鍵である。実際にはかなり経験的、主観的とも考えられよう。しかし、因子の整理、順位配列に理論性が認められれば、必ずしもそうばかりはいえない。数量化方式の検討でも因子の選択はすべて研究者自身の手によることでは同じである。とすれば、因子の選択・配列が総合的であり、客観性・理論性があることの方が合理的であると思われる。

また、相関値については、見かけ上の相関の場合もあり(因子の選択・配列が原因する)、相関値よりは誤差(標準誤差や誤差率)の方が精度としては重要である。したがって、数量化方式での解析でも因子の選択・配列に客観性と理論性をもたせればかなりの改善が期待される。ただし、現時点では曲線的関係を直線で表示することから生ずる歪を是正することはできない。

## 6. おわりに

座標軸による図相関解析の紹介と、因子の選択・配列の客観性・理論性を訴えることに終始した。この方法は、より具体的な説明の文献<sup>5,6,7)</sup>などがある。パソコン導入がない頃は、大変な作業であったが、現在では $X_m$ から $X_n$ への $\hat{Y}_{mn}$ の求めなど極めて容易である。そのために $X$ の順序なども試行錯誤が繰返せるようになった。

そして、もう一つ付加しておきたいことは、やゝもすれば分析結果の説明や解釈を、数値にのみ求めようとするが、それよりは結果に至るまでの過程において、資料の獲得は正しいか、分析手法は目的に適合しているか、資料はその分析に備える内容を有するかなどの検討がほしいことである。更にその上で、分析者の洞察力と思考が重要性を持つ。これが満たされてこそ、統計の結果は活かされると考えている。

## 参考文献

- 1) 西村五月：西海地方の林地生産力に関する環境解析。長崎総農林試研報(林) 17：pp.100, 1986
- 2) 渡辺定元・田中正則・若月 勇：地位指数調査の実際 一数量化による地位指数の推定法について。pp.233 日本林業調査会, 東京, 1966
- 3) 安藤愛次：中部山地の林地生産力に関する研究。pp.195, 農林出版, 東京, 1962
- 4) 竹下敬司：山地の地形形成とその林業的意義。福岡林試時報 17：pp.109, 1964
- 5) 竹下敬司・福島敏彦・萩原幸弘・斉藤 巧：林地生産力(林分材積)に関する立地解析とその分布推定。福岡林試時報 18：41-76, 1966
- 6) 竹下敬司：北九州市門司・小倉区における山地崩壊の予知とその立地的解析。pp.85, 福岡県, 1971
- 7) 福島敏彦・高木潤治・竹下敬司・田形正義：ヒノキ林地生産力(樹高)の立地解析について。福岡林試時報 23：1-34, 1974

8) 西村五月：ヒノキ林の生産解析 XIX. 長崎総農林試研報(林) 14：1-12, 1983