

研究資料

次代検定林データを用いたBLP法による精英樹の系統評価

久保田正裕⁽¹⁾・栗延 晋⁽¹⁾・宮浦富保⁽²⁾

Masahiro Kubota⁽¹⁾, Susumu Kurinobu⁽¹⁾ and Tomiyasu Miyaura⁽²⁾

Evaluation of plus trees with the data from clonal tests and progeny tests by Best Linear Prediction

要旨：次代検定林データを用いた精英樹の特性評価を行うため、BLP法による精英樹の系統評価システムを構築した。システムにおける計算の方法を述べるとともに、解析に使用するコンピュータプログラムの使用法を説明した。また、本システムを用いた次代検定林データへのBLP法の適用結果について、これまでに報告した内容の概略をまとめた。

(1) 林木育種センター

〒319-1301 茨城県日立市十王町伊師3809-1
Forest Tree Breeding Center
3809-1 Ishi, Juo, Hitachi, Ibaraki 319-1301 Japan

(2) 龍谷大学理工学部

〒520-2194 大津市瀬田大江町横谷1番5
Ryukoku University Faculty of Science and Technology
1-5 Yokotani, Seta Oe-cho, Otsu, Shiga 520-2194 Japan

1 はじめに

林木の次代検定では、設定した多数の検定林の全体を通じて成績の良い精英樹が遺伝的に優れていると判断し、その結果を、採種園の改良や第二世代育種集団の親の選抜に役立てている。このような次代検定林は、何年にもわたって多数の箇所に設定していることから、検定林ごとの供試系統や各系統の供試回数、調査年次は不揃いとなり、データの精度は不均一となることが避けられない。このようなアンバランスなデータに対しては、偏差値平均法や最小2乗法が用いられてきたが、データの精度も反映させて育種価を予測する方法としてBLP法 (Best Linear Prediction: 最適線型予測) があり、林木の次代検定データへの適用結果が報告された^{7, 8)}。著者らは、このように多量でアンバランスな次代検定林データから、精英樹の特性評価や選抜をより合理的に行うため、関東育種基本区の次代検定林データを材料として、BLP法による精英樹の系統評価システムの構築を行なった。本報告では、まず、システムにおける計算方法を述べ、次に、計算に使用するコンピュータプログラムの使用法を説明した。最後に、本システムを用いた次代検定林データへのBLP法の適用結果について、これまでに報告した内容の概略をまとめた。

2 精英樹の系統評価システムにおけるBLP値予測のためのデータ解析

2.1 対象とした次代検定林データ

このシステムでは、関東育種基本区内のスギ、ヒノキ、カラマツの一般次代検定林の5年ごとの定期調査結果のプロット平均値を用いた。スギは、さし木クローンおよび自然交配家系、ヒノキ、カラマツは、自然交配家系である。形質は、調査年次によって異なるが、樹高、胸高直径、幹曲がり、根元曲がり、生存率である。樹高は (m) 単位で、胸高直径は (cm) 単位で実測し、幹曲がりおよび根元曲がりは、目視により5段階 (1が最も曲がりが大きく、5は曲がりなし) の指数で評価した。生存率は、植栽した本数のうち調査時に生存していた本数の割合である。

2.2 BLP値の計算方法

林木の次代検定林データによるBLP値の計算方法は、WHITEらによって詳細に記されており⁸⁾、本システムも、概ねWHITEらの方法によって計算を行なった。BLP値は、対象形質の検定林平均値からの偏差、観測値と予測対象形質の育種価との共分散行列 (C)、観測値相互間の分散共分散行列 (V) から求めることができる。行列Cおよび行列Vの要素は、スギ等の樹種では利用可能な推定値がないため、次代検定林データから直接推定する必要がある。

本システムにおける計算の流れを図-1に示す。

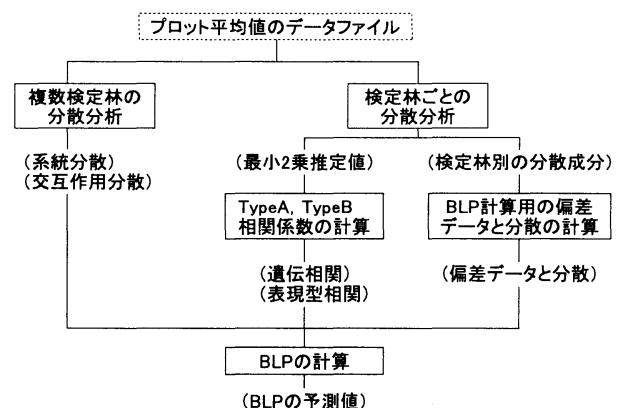


図-1 BLP計算の流れ

この計算の過程は、1) それぞれの検定林での形質ごとの分散分析、2) 各形質の分散共分散を推定するための複数検定林の分散分析、3) BLPを計算するための系統ごとの行列V, Cの作成、4) それぞれの系統のBLP値の計算の4つのステップに分けることができる。以下、そのステップごとに、計算の内容と留意すべき点について述べる。

2.2.1 検定林ごとの分散分析

検定林ごとの形質ごとに、以下の線形モデルに基づいて分散分析を行なう。

$$Y_{jk} = \mu + \beta_j + \gamma_k + \varepsilon_{jk}$$

ここで、 Y_{jk} はj番目のブロックのk番目の系統のプロット平均値、 μ は全体の平均値、 β_j は、j番目のブロックの効果、 γ_k はk番目の系統の効果、 ε_{jk} は誤差である。

分散分析の結果、系統平均値の表現型分散、系統分散を推定する。また、検定林では、プロットが欠測する場合もあるので、ベクトルyの要素として使用するために系統の最小2乗推定値を求める。系統分散が負の値に推定された検定林については、データの信頼性が低いものと判断して、そのデータをのぞく必要がある。

2.2.2 複数検定林の分散分析

予測対象の系統分散を推定するため、以下の線形モデルに基づいて分散分析を行う。

$$Y_{ijk} = \mu + \alpha_i + \alpha\beta_{ij} + \gamma_k + \alpha\gamma_{ik} + \varepsilon_{ijk}$$

ここで、 Y_{ijk} はi番目の検定林のj番目のブロックのk番目の系統のプロット平均値、 μ は全体の平均値、 α_i はi番目の検定林の効果、 $\alpha\beta_{ij}$ はi番目の検定林のj番目のブロックの効果、 γ_k はk番目の系統の効果、 $\alpha\gamma_{ik}$ はi番目の検定林とk番目の系統の交互作用、 ε_{ijk} は誤差である。

分散分析によって、系統分散および検定林と系統の交互作用分散を推定する。共通に供試される家系が、なるべく多くなるように検定林を選出して分析を行なう必要がある。

2.2.3 系統ごとの行列V, Cの作成

分散分析の結果、得られた最小2乗推定値と分散成分のファイルをもとにして、検定林別の最小2乗推定値相互間の相関係数を算出した。同一形質の検定林間の相関は、TypeB相関、同一検定林の異なる形質間の相関はTypeA相関⁸⁾と呼ばれる。行列Vの対角要素は、個々の試験地における系統平均値の表現型分散で、非対角要素は、TypeA相関、TypeB相関により得ることができる。また、行列Cの要素は、予測の対象となる形質間の系統値の共分散、系統相関および系統分散によって計算される。系統相関および系統分散は、複数検定林の分散分析から得られるパラメータである。

2.2.4 BLP値の計算

2.2.1, 2.2.2, 2.2.3で得られた行列C, V, 偏差データファイルを用いて、BLP値を以下の式にしたがって計算する⁸⁾。

$$g = C' \cdot V^{-1} \cdot y$$

ここで、gは予測値のベクトル、yは対象形質の検定林平均値からの偏差（観測値）ベクトル、Cは観測値と予

測対象形質の育種価（予測値）との共分散行列， V は観測値相互間の分散共分散行列である。また，予測値と育種価との相関係数を行列 C ， V ，系統分散により計算する。

3 精英樹系統評価システムの使用法

3.1 プログラムの開発経過

計算に使用したプログラムは，NEC社製のパーソナルコンピュータで動作するN88Basicで栗延らが開発したものを，現在のパーソナルコンピュータのオペレーティングシステムの主流であるMicrosoft社のWindows環境に宮浦らが移植したものである。プログラムの移植には，富士通社のF-basic（Ver. 6.3）を使用した。また，汎用性が高いと考えられる分散分析プログラムは，Microsoft社VisualBasic⁶⁾ およびExcelのVBAにも移植された。

3.2 システムの動作環境およびプログラムの構成

システムを使用するために必要なものは，オペレーティングシステムとしてMicrosoft社のWindowsが動作しているパーソナルコンピュータおよびF-Basic（富士通社）である。

システムを構成するプログラムファイルを表-1に示す。ステップごとに，それぞれのプログラムを実行してデータを計算し，「結果ファイル」を得る。

3.3 プログラムの使用法

これらのプログラムは，テキスト（ASCII）形式でファイルに保存しておいたデータを読み込み，計算結果をテキスト形式でファイルに出力するものである。以下，それぞれのプログラムの使用法を示す。プログラムファイルは，アイコンをダブルクリックするとプログラムが起動し，利用可能となる。それぞれのプログラム間のファイルの受け渡しを示すため，データファイルと「結果ファイル」の例を表-1に示す。

3.3.1 検定林ごとの分散分析

使用するプログラムは「PLMSGLO2」である。プログラムを起動すると，ウィンドウが開くので，計算に必要な情報を順に入力する（図-2）。“File-name to read”，データファイル名を入力する。“Number of observations”には，データファイルのレコード数を，“Number of tests”は，検定林数を，“Number of replications”は，反復数を，“Number of families”は，系統数を，“Number of traits”は，形質数を入力する。これらは，プログラム内で配列変数等の領域を確保するために用いるものであり，正確な値がわからない場合は，実際のデータよりも大きめの数値を入力すればよい。分散分析は，調査年次ごとに行なうので，“Age for analysis”で年次を入力する。“File name to output LSE”，“File name to output ANV”は，それぞれ，最小2乗推定値，分散成分を出力するためのファイル名を入力する。ファイルの内容がわかるように，それぞれ，拡張子は，lse，anvとしておく。“File name for work-space”は，作業スペースを確保するため，“work”など任意の名称を入力する。

解析に用いるデータファイルの例を図-3に示す。1から5桁目は検定林コード，6から10桁目はブロック，11から15桁目は精英樹コード，16から17桁目は選抜された育種区コード，18から20桁目は調査年次となっている。

表-1 評価システムを構成するプログラムファイルとデータファイル, 結果ファイルの例

プログラムファイル	データファイル	結果ファイル
検定林ごとの分散分析		
PLMSG102	sugix.dat	sugi 5 .anv sugi 5 .lse sugi10.anv sugi10.lse sugi15.anv sugi15.lse sugi20.anv sugi20.lse sugi20.lse
エディター等で結合する		
	sugi 5 .anv sugi10.anv sugi15.anv sugi20.anv sugi 5 .lse sugi10.lse	sugiall.anv sugiall.lse
1 系統を 1 レコードにまとめる		
PLMDTANV	sugiall.anv	sugittl.tab
PLMARA	sugiall.lse	sugittl.lse
複数検定林の分散分析		
LSQABS21	sugi.dat	sugi.res
系統ごとの行列V, Cの作成		
CORPOOL 3	sugittl.tab sugittl.lse	sugitpa.crl
CORPOOL 2	sugittl.tab sugittl.lse	sugittl.crl
CALCORL	sugitpa.crl	sugittl.prm sugittl.crl
偏差分散ファイルの作成		
PLMDTVCL	sugiall.anv sugittl.lse	sugittl.vct
BLPの計算		
BLPCLD	sugittl.prm sugittl.vct	sugi.blp

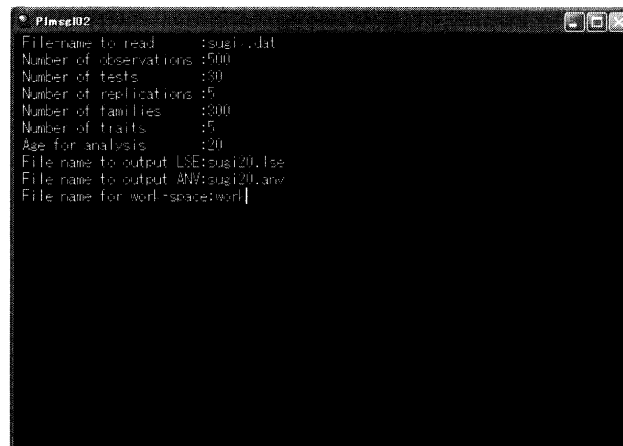


図-2 分散分析プログラムPlmsg102を起動してパラメータを入力した例

データは、26桁目から始まり、5桁がひとつの調査形質で、樹高、胸高直径、幹曲がり、根元曲がり、生存率となっている。21から25桁目、どの形質のデータがあるかを示している。この例では、樹高、胸高直径、生存率のデータがあることがわかる。

335	2	1544	9	2012	513.2017.30	66.7
335	2	1545	9	2012	511.4214.50	80.0
335	2	1548	9	2012	59.83 12.33	40.0
335	2	1553	9	2012	512.1516.38	86.7
335	2	1577	9	2012	511.6416.09	73.3
335	3	1530	9	2012	511.8516.08	86.7
335	3	1531	9	2012	511.5015.00	53.3
335	3	1532	9	2012	512.2516.25	53.3

-----+-----1-----+-----2-----+-----3-----+-----4-----+-----5----- 桁数

図-3 検定林ごとの分散分析のためのデータファイルの例

得られた「結果ファイル」は、調査年次ごとのファイルであるので、エディター等を使用し、一つのファイルに統合する。年次の区別ができるように、あらかじめ各レコードの最初の2カラムに年次を入力しておく。統合したファイルは、プログラム「PLMARA」、「PLMDTANV」により、1系統の結果が1レコードとなるように形を整える。これらのプログラムの使用法も、「PLMSGL02」とほぼ同様である。

3.3.2 複数検定林の分散分析

使用するプログラムは、「LSQABS21」である。使用法は、先のプログラム「PLMSGL02」とほぼ同じである。プログラムを起動し、計算に必要な情報を順に入力していく。“File-name to read”, “File-name to output”には、データファイル名および出力ファイル名を入力する。以下、先のプログラムと同様に、データファイルのレコード数、検定林数、最大の反復数、系統数、形質数を入力する。また、作業スペースの名前を入力する。

解析に用いるデータファイルの例を図-4に示す。1から10桁目は精英樹コード、11から15桁目は検定林コード、15から20桁目はブロック、21から22桁目は選抜された育種区コード、23から25桁目は調査年次となっている。データは、26桁目から始まり、5桁がひとつの調査形質で、樹高、胸高直径、幹曲がり、根元曲がり、生存率の順に並んでいる。この例では、樹高、胸高直径、幹曲がり、根元曲がりのデータがあることがわかる。

これらの分散分析は、調査年次ごとに行なうので、「結果ファイル」に年次を入れて保存するとわかりやすい。

2932	281	3 8 20	913.77	4.88	4
2942	281	3 8 20	9.8	15	3.7 3.8
2942	281	2 8 20	13.117.21	3.89	4.28
2942	281	1 8 2010.	9215.65	4.4	3.65
2961	281	3 8 20	8.3313.33	2.66	4.33
2961	281	1 8 2011.	5616.25	3.87	3.75
2961	281	2 8 20	11.515.83	3.91	4.75
3035	281	3 8 20	9.0411.72	3.81	3.63

-----+-----1-----+-----2-----+-----3-----+-----4-----+-----5----- 桁数

図-4 複数検定林の分散分析のためのデータファイルの例

3.3.3 系統ごとの行列V, Cの作成

使用するプログラムは、相関係数の計算では、プログラム「CORPOOL2」, 「CORPOOL3」および「CALCORL」, 偏差データと分散計算値の算出では、プログラム「PLMDTVCL」である。これらのプログラムの使用法は、先のプログラム「PLMSGLO2」とほぼ同じである。プログラムを起動し、入力、出力データファイル名、形質数、レコード数等のデータ、ワークスペースのファイル名を入力する。

3.3.4 系統のBLP値の計算

使用するプログラムは、「LSQABS21」である。このプログラムの使用法は、先のプログラム「PLMSGLO2」とほぼ同じである。プログラムを起動し、入力、出力データファイル名、形質数、レコード数等のデータ、ワークスペースのファイル名を入力する。

4 精英樹評価システムを用いたBLP法の次代検定林データへの適用

本システムを用いて、関東育種基本区のスギ、ヒノキ、カラマツの次代検定林データにBLP法を適用した。それぞれの結果の詳細については、それぞれの報告を参照されたい。ここでは、それぞれの結果の概要を述べる。

スギさし木クローン次代検定林の5, 10, 15年次の樹高, 胸高直径, 幹曲がり, 根元曲がり, 生存率を使用し, 15年次の樹高, 幹曲がり, 生存率を予測した。北関東育種区では, 45検定林の結果から, 190クローンの育種価を予測することができた。BLP予測値とクローンの育種価の相関は, 樹高でもっとも高く, 生存率, 幹曲がりの順に低くなった。クローン別の供試回数とBLP予測値の関係は, 供試回数の少ないクローンほど分布の中心に集中する傾向がみられた^{3,4)}

スギ自然交配家系の次代検定林の5, 10, 15, 20年次の樹高, 胸高直径, 幹曲がり, 根元曲がりおよび生存率のデータを使用し, 20年次の樹高, 胸高直径および生存率, 10年次の幹曲がり, 根元曲がりを予測した。BLP予測値は312家系で得ることができた。従来の最小2乗法で20年次評価値を得られたのは, 163家系であった。樹高, 胸高直径, いずれもBLP予測値は, 供試回数が少ない家系の値は平均近くに多く分布する傾向が見られた。また, 予測した育種価を用いた樹高による選抜は, 胸高直径では樹高と同程度の改良が期待され, 幹曲がり, 根元曲がりおよび生存率も低下することはないことを確認した²⁾。

ヒノキ自然家系に次代検定林の10, 15年次の樹高, 胸高直径, 幹曲がり, 根元曲がり, 生存率のデータを使用し, 10年次の樹高, 幹曲がり, 生存率を予測した。BLP予測値の頻度分布は, 供試回数の少ない家系のBLP予測値が分布の中心に集中する傾向にあった。BLP値と真の育種価の相関係数は, 相関係数の高い家系は, BLP予測値の変動が大きく, 相関係数の低い家系では, BLP予測値は, 平均値の近くに集中していた。また, 供試回数の多い家系については, 相関係数が高くなる傾向にあった⁵⁾。

カラマツ自然交配家系の次代検定林の10, 15, 20年次の樹高, 胸高直径, 幹曲がり, 根元曲がりのデータを使用し, 20年次の樹高, 胸高直径, 根元曲がりを予測した。いずれの形質でも年次間での順位の変動は小さかった。樹高, 胸高直径, 根元曲がりいずれもBLP予測値は, 供試回数が2回以下と少ない家系の値は平均近くに多く分布する傾向が見られた。また, 樹高の上位家系を選抜した場合の20年次における胸高直径および根元曲がりの選抜効果を推定した¹⁾。

3つのいずれの樹種においても、系統ごとのデータの信頼度や供試回数が育種価の予測値に反映されていることが確認された。このことは、予測値の大きい系統、あるいは、逆に小さな家系を選抜する場合、予測値の信頼度の低い家系を選抜する危険性が少なくなることを意味する。系統を評価した特性表は、採種園の改良等のために上位または下位の系統を選抜するために利用されることが多い。このような場合に、予測値の信頼度の低い系統を選抜する危険性を低くすることができ、BLP法が有効である点の一つである。また、BLP予測値は育種価の予測値であることから、選抜した家系のBLP予測値の平均からの差にもとづいて選抜の効果を検討することが可能であった。

Summary : A set of computer programs for evaluation of clones and half-sib progenies of plus-tree by Best Linear Prediction (BLP) were transferred from N88Basic to F-basic under MS-Windows. We explain calculation of BLP briefly, and describe the procedure for using the programs. The results of application of BLP method to progeny test data with these programs is also reviewed.

引用文献

- 1) 久保田正裕・宮浦富保：カラマツ精英樹家系の次代検定林データを用いたBLP法による系統評価，日林関東支論51, 97~98 (2000)
- 2) 久保田正裕：スギ精英樹家系の次代検定林データを用いたBLP法による系統評価，平成15年度林木育種センター年報, 60~61 (2004)
- 3) KURINOBU, S.・MIYAURA, T. : An application of a best linear prediction to clonal tests of Sugi (*Cryptomeria japonica*) in the northern Kanto region., J. Jpn. For. Soc. 75, 350-355 (1993)
- 4) 栗延晋・宮浦富保・久保田権：関東育種基本区のスギさしき検定データに対するBLP法の適用について，日林論104, 441-442 (1993)
- 5) 宮浦富保・栗延晋・久保田権：関東育種基本区のヒノキ次代検定データについてのBLP法による系統評価，日林論115, 319-320 (1994)
- 6) 宮浦富保：検定林データの分散分析プログラム，林育研報15, 251~258 (1998)
- 7) WHITE, T.L. and HODGE, G.R. : Best linear prediction of breeding values in a forest tree improvement program. Theor. Appl. Genet. 76 : 719~727 (1988)
- 8) WHITE, T.L. and HODGE, G.R. : Predicting breeding values with applications in forest tree improvement., Kluwer Academic Pub, Dordrecht, Netherlands, 367pp (1989)

「林木育種センター研究報告 No.22」正誤表

ページ	行	誤	正
目次	12	渡辺 敦史	渡邊 敦史
35	3	戸田忠雄 ⁽¹⁾	戸田忠雄 ⁽²⁾
35	5	Yoshitake Fujisawa and Tadao Toda	Yoshitake Fujisawa ⁽¹⁾ and Tadao Toda ⁽²⁾
35	欄外	(未記入)	(2)日中協力林木育種科学技術センター計画 (安徽省松材線虫抵抗性育種センター担当) <u>The Japan-China Cooperation Science and Technology Center</u> <u>forest Tree Improvement Project</u>
51	3	渡辺敦史	渡邊敦史
60	下から6	渡辺敦史	渡邊敦史
214	下から5	河野楯蔵	河野楯蔵
巻末「論文審査者」	2	宇都宮大学農学部 附属演習林	宇都宮大学農学部附属演習林助教授
同上	6	北海道立林業試験場林業経営部	北海道立林業試験場 林業経営部 主任研究員
同上	6	Makoto Kuromaru	Dr.Makoto Kuromaru
同上	9～10	東北大学大学院農学研究科附属 複合生態フィールド教育研究センター	東北大学大学院農学研究科 附属複合生態フィールド教育研究センター 生物共生科学研究室助教授
同上	13	(未記入)	日本森林技術協会 北海道事務所 主任研究員
同上	13	Tsugio Ohshima	Mr.Tsugio Ohshima