

PREDICTING BREEDING VALUES
with Applications in Forest Tree Improvement

FORESTRY SCIENCES

Volume 33

This is Journal Series No. R00051 of the Institute of Food and Agricultural Sciences of the University of Florida, Gainesville, FL, 32611.

For a list of titles in this series see final page of this volume.

Predicting Breeding Values

with Applications in Forest Tree Improvement

by

TIMOTHY L. WHITE

and

GARY R. HODGE

Department of Forestry, University of Florida, Gainesville, U.S.A.



Springer-Science+Business Media, B.V.

Library of Congress Cataloging in Publication Data

White, Timothy L., 1951-

Predicting breeding values with applications in forest tree
improvement / Timothy L. White, Gary R. Hodge.

p. cm. -- (Forestry sciences ; 33)

Includes bibliographical references.

ISBN 0-7923-0460-8 (U.S.)

1. Trees--Breeding. 2. Trees--Selection. 3. Forest genetics.
4. Breeding. 5. Quantitative genetics. I. Hodge, Gary R., 1958-
. II. Title. III. Series.

SD399.5.W47 1989

631.5'2--dc20

89-19903

ISBN 978-90-481-4055-8 ISBN 978-94-015-7833-2 (eBook)

DOI 10.1007/978-94-015-7833-2

printed on acid free paper

All Rights Reserved

© 1989 by Springer Science+Business Media Dordrecht

Originally published by Kluwer Academic Publishers in 1989.

Softcover reprint of the hardcover 1st edition 1989

No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical including photocopying, recording, or by any information storage and retrieval system, without written permission from the copyright owner.

To Mary, Dorothy and Suzanne

and Terry, Daniel and Megan

Contents

	Preface	ix
Chapter 1	Matrix Algebra	2
Chapter 2	Statistics	20
Chapter 3	Concepts of Progeny Test Analysis	48
Chapter 4	Theory of Best Linear Prediction (BLP)	62
Chapter 5	BLP with Half-sib Progeny Test Data	86
Chapter 6	BLP with Full-sib and Multiple Sources of Data	112
Chapter 7	BLP: Further Topics	136
Chapter 8	BLP: An Operational Example	172
Chapter 9	Selection Index Theory	208
Chapter 10	Selection Index Applications	232
Chapter 11	Best Linear Unbiased Prediction: Introduction	276
Chapter 12	Best Linear Unbiased Prediction: Applications	300
	Literature Cited	328
	Appendices	336
	Answers to Problems	346
	Index	364

Preface

In most breeding programs of plant and animal species, genetic data (such as data from field progeny tests) are used to rank parents and help choose candidates for selection. In general, all selection processes first rank the candidates using some function of the observed data and then choose as the selected portion those candidates with the largest (or smallest) values of that function. To make maximum progress from selection, it is necessary to use a function of the data that results in the candidates being ranked as closely as possible to the true (but always unknown) ranking.

Very often the observed data on various candidates are messy and unbalanced and this complicates the process of developing precise and accurate rankings. For example, for any given candidate, there may be data on that candidate and its siblings growing in several field tests of different ages. Also, there may be performance data on siblings, ancestors or other relatives from greenhouse, laboratory or other field tests. In addition, data on different candidates may differ drastically in terms of quality and quantity available and may come from varied relatives. Genetic improvement programs which make most effective use of these varied, messy, unbalanced and ancestral data will maximize progress from all stages of selection.

In this regard, there are two analytical techniques, best linear prediction (BLP) and best linear unbiased prediction (BLUP), which are quite well-suited to predicting genetic values from a wide variety of sources, ages, qualities and quantities of data. These techniques were largely developed by the late C.R. Henderson and his colleagues for predicting breeding values of dairy cows from massive amounts of extremely messy data. BLUP is now widely used in the United States and other countries in the dairy industry. These analytical techniques (BLP and BLUP) are equally well-suited to breeding programs of other plant and animal species that generate messy genetic data, but have not been widely used. While these techniques are well-documented in the literature, many of the papers require a fair amount of mathematical sophistication and often deal with the extraordinary computational problems encountered in dairy cow breeding programs.

The purpose of this book is to describe these techniques of predicting genetic values in a manner requiring a minimum of prerequisite background (one year of statistical methods and a course in quantitative genetics). While we have tried to be somewhat rigorous in the presentation, we try to stress concepts and the intuitively pleasing manner with which these techniques handle different types of data.

There are three analytical techniques discussed in this book: BLP, selection index and BLUP. The common thread of these techniques is that in the derivation of their analytical formulas, genetic values are considered random variables to be predicted rather than fixed effects (i.e. constants) to be estimated. While this seems subtle at first, it results in quite different analytical methods and formulas than often employed in breeding programs.

While these three methods are similar in some regards (in terms of treating genetic values as random variables and in terms of some characteristics in the way they handle data), BLP, selection index and BLUP differ from one another in some other critical underlying assumptions. Thus, the appropriate technique to use for any given problem depends on which set of assumptions of the three methods is most suitable in that situation. For this reason, we try to be quite explicit in describing the underlying assumptions, problems and issues involved in the application of all three techniques.

The book is intended for students and scientists in quantitative genetics and for breeders in genetic improvement programs. While the examples and applications given are necessarily limited to our experiences in forest genetics, we hope the book finds appeal in other disciplines. The concepts are wholly general.

The first three chapters develop some background material in matrix algebra (Chapter 1), statistics (Chapter 2) and progeny testing (Chapter 3) that is useful in the ensuing discussions of BLP (Chapters 4 - 8), selection index (Chapters 9 - 10), and BLUP (Chapters 11 - 12). Chapters 4, 9 and 11 develop the theory and properties of BLP, selection index and BLUP, respectively. These chapters are the most difficult mathematically; however, they can be skimmed lightly the first reading without disrupting continuity. These theory chapters are included for completeness and for later reference. The other chapters dealing with each technique rely heavily on numerical examples to 1) demonstrate how to apply the technique to real data sets and 2) develop intuitive concepts about how the technique handles data (e.g. how different qualities of data get different weights in predicting genetic values). In particular, Chapters 5 and 6 develop many concepts about BLP that also apply to the other two techniques. We view these as key chapters to read even if primary interest is on selection index or BLUP.

The treatment of these three techniques in this book (particularly BLUP) is certainly not exhaustive. Rather, we try to develop concepts and demonstrate applications to motivate and allow use of these techniques by other geneticists. Errors in the presentation inevitably exist and we would appreciate them being brought to our attention. Also, we are still learning about new applications and ramifications of these techniques and would appreciate discussing them with other scientists.

Many people have helped us in the preparation of this book and we are deeply grateful. Drs. Mike DeLorenzo and Charles Wilcox stimulated our original interest and were extremely helpful in our early applications of BLP to forestry problems. The 40 students in two shortcourses (the USDA Forest Service course in Corvallis, Oregon, and the Tree Breeding course in Canberra, Australia) made tremendous contributions to earlier drafts of the manuscript and helped in many other ways. Two reviewers, Dr. Dale van Vleck and Dr. Sharon Friedman, read the entire final draft. They provided simply excellent review comments for which we are grateful. We also thank other reviewers that read one to several chapters: Dr. Floyd Bridgwater, Mr. Dudley Huber, Dr. Steve McKeand, Ms. Sonia De Souza and Dr. Claire Williams. We also want to thank the 15 industrial and state agencies that are the members of the Cooperative Forest Genetics Research Program. Their support and patience made this possible. Finally, we want to thank Mr. Greg Powell for everything he does: managing and analyzing the data, keeping the computers running, picking up the slack and steering a steady course even in high winds.