1984

A study of weaning weights in Hereford cattle in the state of Rio Grande do Sul, Brazil

Luiz Alberto Fries

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Agriculture Commons, Animal Sciences Commons, and the Genetics Commons

Recommended Citation

https://lib.dr.iastate.edu/rtd/8163

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages √
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) ________ lacking when material received, and not available from school or author.
12. Page(s) ________ seem to be missing in numbering only as text follows.
13. Two pages numbered _________. Text follows.
14. Curling and wrinkled pages _____
15. Other ________________________________________________________
A study of weaning weights in Hereford cattle in the state of Rio Grande do Sul, Brazil

by

Luiz Alberto Fries

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of the Requirements for the Degree of DOCTOR OF PHILOSOPHY

Department: Animal Science
Major: Animal Breeding

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1984
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. BRIEF HISTORY OF BEEF CATTLE BREEDING IN RS</td>
<td>5</td>
</tr>
<tr>
<td>III. LITERATURE REVIEW OF SOME ENVIRONMENTAL EFFECTS ON WEANING WEIGHT</td>
<td>15</td>
</tr>
<tr>
<td>A. Date of Birth and Age at Weaning</td>
<td>15</td>
</tr>
<tr>
<td>1. Birth date</td>
<td>16</td>
</tr>
<tr>
<td>2. Environment from birth to weaning</td>
<td>16</td>
</tr>
<tr>
<td>3. Growth curve</td>
<td>16</td>
</tr>
<tr>
<td>B. Nursing Status of the Dam</td>
<td>24</td>
</tr>
<tr>
<td>IV. DATA DESCRIPTION (INCLUDING BIOLOGICAL MODEL)</td>
<td>27</td>
</tr>
<tr>
<td>A. Editing</td>
<td>28</td>
</tr>
<tr>
<td>B. What Was Not Edited and Why Not</td>
<td>30</td>
</tr>
<tr>
<td>C. Edited Data Description</td>
<td>32</td>
</tr>
<tr>
<td>D. Biological Model</td>
<td>45</td>
</tr>
<tr>
<td>V. METHODS</td>
<td>51</td>
</tr>
<tr>
<td>A. Parameter Estimation</td>
<td>52</td>
</tr>
<tr>
<td>1. Ordinary least squares (OLS)</td>
<td>54</td>
</tr>
<tr>
<td>a. Dummy variables</td>
<td>55</td>
</tr>
<tr>
<td>b. Ordinary polynomials</td>
<td>58</td>
</tr>
<tr>
<td>c. Grafted polynomials</td>
<td>60</td>
</tr>
<tr>
<td>d. Estimation of the jointing points of the grafted polynomials</td>
<td>76</td>
</tr>
<tr>
<td>2. Generalized least squares (GLS)</td>
<td>92</td>
</tr>
<tr>
<td>a. Heteroscedasticity of disturbances (WLS)</td>
<td>94</td>
</tr>
<tr>
<td>b. Correlated errors (mixed model) and prediction problems</td>
<td>96</td>
</tr>
</tbody>
</table>
3. Robust estimation
   a. M-estimators
   b. Equivariance, invariance, and scale estimation
   c. Resistance and robustness of efficiency
   d. Breakdown bound
   e. W-estimation
   f. How to obtain a perfect, absolute fit to your model
   g. Robust weighted least squares (ROBWLS)

B. Variance Component Estimation
   1. Definition of REML
   2. A few computation aspects

C. Solving Least Squares Problems
   1. Constructing the records of the smallest possible example
   2. Direct solution using NE and the inverse of X'X
   3. Direct solution using the sweet operator
   4. Direct solution using Givens planar rotations

D. Using Givens Rotations to Obtain GLS Estimates of the Effects in the Model and to Obtain REML Variance Component Estimates
   1. Absorption of CG effects in the data matrix
   2. Apply Givens rotations
   3. Data matrix augmentation
   4. Extracting information from T*

VI. ALGORITHMS USED, WITH CRITICISMS
   A. Two-Way Fixed Model (GIVFIX)
   B. Criticisms on GIVFIX Program
   C. Results from GIVFIX Program
   D. Obtaining Variance Components of a Mixed Two-Way Model Using Henderson's Iterative Algorithm (HIA)
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.</td>
<td>A Two-way Mixed Model Program (GIVMIXCV) to Obtain REML Estimates of Variance Components Using the Common Intercept Approach</td>
<td>176</td>
</tr>
<tr>
<td>F.</td>
<td>Programs to Solve a Two-way Mixed Model, With Covariates (Programs ABSSIRER and REML)</td>
<td>178</td>
</tr>
<tr>
<td>VII.</td>
<td>RESULTS AND DISCUSSION</td>
<td>185</td>
</tr>
<tr>
<td>A.</td>
<td>Comparison of Methods and Models</td>
<td>186</td>
</tr>
<tr>
<td>B.</td>
<td>Environmental Effects</td>
<td>196</td>
</tr>
<tr>
<td>1.</td>
<td>Dam nursing status (STAT1 and STAT2)</td>
<td>196</td>
</tr>
<tr>
<td>2.</td>
<td>Age of dam</td>
<td>198</td>
</tr>
<tr>
<td>3.</td>
<td>Seasonality within fall effects</td>
<td>200</td>
</tr>
<tr>
<td>4.</td>
<td>Seasonality within spring effects</td>
<td>202</td>
</tr>
<tr>
<td>5.</td>
<td>Non-linear age of calf effects</td>
<td>205</td>
</tr>
<tr>
<td>6.</td>
<td>Weaning date effects</td>
<td>215</td>
</tr>
<tr>
<td>C.</td>
<td>Heritability and Error and Sire Variance Estimates</td>
<td>224</td>
</tr>
<tr>
<td>D.</td>
<td>Reduced Model</td>
<td>235</td>
</tr>
<tr>
<td>VIII.</td>
<td>APPLICATION OF RESULTS AND POSSIBLE USE OF GIVENS ROTATIONS FOR WITHIN HERD AND ACROSS HERD EVALUATIONS</td>
<td>243</td>
</tr>
<tr>
<td>A.</td>
<td>Taking Environmental Effects Into Account</td>
<td>243</td>
</tr>
<tr>
<td>1.</td>
<td>Precorrection of records: restricted least squares</td>
<td>244</td>
</tr>
<tr>
<td>2.</td>
<td>Combining information from extraneous sources with the sample own information</td>
<td>247</td>
</tr>
<tr>
<td>B.</td>
<td>The Operation of Promebo as a GLS Problem</td>
<td>253</td>
</tr>
<tr>
<td>1.</td>
<td>Cholesky decomposition vs. Givens rotations</td>
<td>253</td>
</tr>
<tr>
<td>2.</td>
<td>Combining Cholesky and Givens</td>
<td>255</td>
</tr>
<tr>
<td>3.</td>
<td>Within herd evaluation</td>
<td>260</td>
</tr>
<tr>
<td>4.</td>
<td>Across herd evaluations</td>
<td>263</td>
</tr>
<tr>
<td>5.</td>
<td>Advantages of using Givens rotations</td>
<td>266</td>
</tr>
<tr>
<td>IX.</td>
<td>CONCLUSIONS</td>
<td>269</td>
</tr>
<tr>
<td>X.</td>
<td>THINGS THAT SHOULD HAVE BEEN DONE DIFFERENTLY</td>
<td>271</td>
</tr>
<tr>
<td>XI.</td>
<td>CITATIONS</td>
<td>273</td>
</tr>
</tbody>
</table>
XII. COMMITTEE MEMBERS

XIII. APPENDICES

Appendix A. Problem Set Number 7 of Harville (1983)

Appendix B. GIVFIX - a Two-way Fixed Model Program Using Givens Rotations

Appendix C. Partial Print-out of GIVFIX Execution with Records Given in Appendix A

Appendix D. Results from Using the Method of Successive Approximations for REML Estimation of a Variance Ratio (records from Appendix A)

Appendix E. GIVMIXCV - A Two-way Mixed Model Program to Obtain REML Estimates of Variance Components Using the Common Intercept Approach

Appendix F. Results From Using GIVMIXCV on Records Given in Appendix A

Appendix G. Results From Using GIVMIXCV, Modified to Achieve Faster Convergence, on Records Given in Appendix A

Appendix H. Listing of Program ABSSIRER

Appendix I. Listing of Program REML

Appendix J. Sire Solutions in the Full and in the Reduced Models
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Distribution of weaning weight records among herds, years, and seasons within years</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>Distribution of fall birth dates (in the Julian calendar) and corresponding weaning ages</td>
<td>36</td>
</tr>
<tr>
<td>3</td>
<td>Distribution of spring birth dates (in the Julian calendar) and corresponding weaning ages</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>Distribution of weaning weight records according to age (AOD) and nursing status (STATUS) of the dam</td>
<td>41</td>
</tr>
<tr>
<td>5</td>
<td>Means and dispersion measures on several variables in the data set</td>
<td>44</td>
</tr>
<tr>
<td>6</td>
<td>Non-linear least squares grid search for the knots ($\alpha_1$ and $\alpha_2$) of a cubic-quadratic-cubic system created to estimate age of calf effects on weaning weight</td>
<td>83</td>
</tr>
<tr>
<td>7</td>
<td>Investigation of multicollinearity among the explanatory variables as given by the tolerance and variance inflation factor (VIF) statistics</td>
<td>90</td>
</tr>
<tr>
<td>8</td>
<td>Coefficients of determination ($R^2$), mean squared errors (MSE), estimates of regression coefficients and their standard errors (in parentheses, below the estimate) for the seven models used in this study</td>
<td>187</td>
</tr>
<tr>
<td>9</td>
<td>Analysis of variance table for model (OLS - complete given by Equation (38))</td>
<td>192</td>
</tr>
<tr>
<td>10</td>
<td>Type I (sequential) sum of squares for the first four sources in the model (OLS-complete) given by Equation (38)</td>
<td>192</td>
</tr>
<tr>
<td>11</td>
<td>Type IV sum of squares for sires and regressors in the model (OLS-complete) given by Equation (38)</td>
<td>193</td>
</tr>
<tr>
<td>12</td>
<td>Type IV sum of squares for the regressors in the GLS model given in Equation (39)</td>
<td>194</td>
</tr>
</tbody>
</table>
Table 13. Iterative REML estimates of error and sire variances, and their ratio, using CIA (Schaeffer, 1983)  

Table 14. Comparison of complete and reduced models, using OLS (sires fixed) and GLS (sires random) estimators (variance component estimates for GLS are REML)  

Table 15. Number of operations needed to solve a within herd problem via Cholesky decomposition during a period of 10 years  

Table 16. Number of operations needed to solve a within herd problem via Givens rotations during a period of 10 years  

LIST OF FIGURES

Figure 1. Biological model hypothesized for weaning weights of Hereford calves in RS, being tested in the present study 47

Figure 2. Age of dam (within sex of calf) effects on weaning weight 199

Figure 3. Seasonality within fall effects on weaning weight 203

Figure 4. Seasonality within spring effects on weaning weight 205

Figure 5. Non-linear age of calf effects on real weaning weight (kg) 206

Figure 6. Non-linear age of calf effects on average daily gain (kg/day) 208

Figure 7. Non-linear age of calf effects on corrected weaning weight (kg) 210

Figure 8. Seasonality within fall (SF1) and non-linear age of calf (age) effects on weaning weight (kg) 211

Figure 9. Seasonality within fall (SF1) and non-linear age of calf (age) effects on weaning weight (kg) 212

Figure 10. Seasonality within spring (SS1) and non-linear age of calf (age) effects on weaning weight (kg) 213

Figure 11. Seasonality within spring (SS1) and non-linear age of calf (age) effects on weaning weight (kg) 214

Figure 12. Effect of weaning date on corrected weaning weight according to date of birth within fall 217

Figure 13. Effects of weaning date on corrected weaning weight of fall born calves according to age of calf 218

Figure 14. Effect of weaning date on corrected weaning weight according to date of birth within spring 221

Figure 15. Effect of weaning date on corrected weaning weight of spring born calves according to age of calf 222

Figure 16. Plot of sire solutions in the full and in the reduced models 240
I. INTRODUCTION

The purpose of this work is to identify important environmental effects, account for these effects, and estimate the relative importance of genetic differences for weaning weight in the beef cattle population participating in a record of performance (ROP) program in the state of Rio Grande do Sul (RS), Brazil.

The important environmental effects on weaning weight in the United States (US) were researched and understood by the late 1960s using large bodies of field and experimental station data. In RS, only small data sets from experiment stations were available, but after 1974 a ROP was initiated that followed the guidelines of the Beef Improvement Federation (BIF) of 1972. Data analyzed from the first two herds suggested that higher corrections for young dams would be necessary than those recommended by BIF. Rocha (1978) and Cardellino and Cardellino (1983) found very low heritability estimates (0.00 to 0.05). Pons et al. (1980) reported a repeatability of 0.05. Ferreira (1982), using a model with only the factors recommended by BIF in 1981, found a coefficient of determination \( R^2 \) of 0.29, while similar models used to describe US field data gave \( R^2 \) values of 0.50 to 0.60.

These results suggest that there is a low correlation in the present ROP data between the breeding value and phenotypic value for weaning weight. Of all the selection practiced among bulls based on performance records, weaning weight selection accounts for from 50-80% of the total cullings made. Difficulties with managing large groups of bulls on pasture and/or the high feed costs are responsible for such an emphasis.
This use of weaning weight as a selection criteria emphasizes the importance of researching methods to increase the heritability of these weaning weight records.

Not only is weaning weight (WW) crucial, but also other measured variables appear to be more in line with the literature. Rocha (1977) estimated heritability for post-weaning gain, on pasture, from 205 to 550 days as 0.67. Cardellino and Cardellino (1983) reported an estimate of heritability for weaning conformation score of 0.32.

The consistency of results by several researchers suggests that the problem of low estimates of heritability of WW is real. Models that better describe the biology and management constraints of the situation in RS appear necessary. Some of the possible causes of the problem are as follows:

1. The estimation methods and implicit assumptions were partially inadequate.
2. The analytical tools employed were limited.
3. Some data sets studied were small.
4. Records of questionable value were not checked.
5. The identification of parentage was not perfect.

The beef cattle production system in RS, when compared with an intensive system in a temperate or cold climate, can be expected to yield lower estimates of heritability. This hypothesis can not be tested in the present work. Probably causes for a lower real heritability are the following:

1. The environment is less attenuated by uniform husbandry practices and can be controlled only to a lesser degree,
resulting in less standardization.

2. Ecto- and endo-parasites reach higher infestations in regions where the ground does not freeze during the winter. Hoof-and-mouth disease also has a minor role through animals in a sub-clinical stage. Individual differences in susceptibility remain unexplained by the model and can affect estimates of genetic merit and inflate residual error.

3. The trait weaning weight may have a different or more broad connotation. In RS, the phenotype results not only from appetite, milk production and potential for growth, but also from genes regulating fitness and adaptation to the overall environment. These genes, with their additive and non-additive effects, may add more complexity to the part of the genome determining weight.

The work on this problem centers on investigating some of the possible causes by the following procedures:

1. Doing an exploratory analysis to study the data, using means, plots, tables, distributions, and identifying extreme values so that peculiarities of the data can be understood.

2. Developing an adequate model that more nearly describes the environmental influences.

3. Testing different assumptions about the distribution of the residuals and their influence on the estimates of the parameters.

4. Using statistical procedures that are better than ordinary
least squares, ordinary polynomials and, for variance components, method III of Henderson.
II. BRIEF HISTORY OF BEEF CATTLE BREEDING IN RS

In Chapter VIII of this work, some suggestions will be made with the aim of improving the efficiency of the present ROP program offered to the beef cattle industry of RS. It is difficult to propose or to critically analyze proposed changes in one of the parts of a (non-integrated) system without an overall perspective. This is more the case when field data are collected in one country and are analyzed and conclusions drawn in another one.

Even if one were interested in the subject, but could not rely on his Portuguese expertise, very little up-to-date information would be found. Joandet (1977), reviewing crossbreeding results from South America stated this situation as follows: "In Brazil, animal breeding research has been conducted for many years, but very little information is available from that country." Results of research are seldom reported outside the country. A few papers can be found in Animal Production and in World Animal Review. Every paper or thesis in Brazil must contain a summary in a foreign language. Most of these are in English, but very few articles in English deal with production systems or breeding methods being operated in Brazil.

Veiga (1955) described how beef cattle were produced in Central Brazil and the research done with Zebu breeds. Hill (1967) had a broader scope but his paper is more concerned with dairy cattle breeding in Central Brazil. Pearson de Vaccaro (1974) gave an account of dairy breeding in South America as a whole and reviewed some Brazilian results. Monteiro et al. (1981) described the bio-economic system of beef cattle
production in Central Brazil. No account of beef cattle breeding (in English) in Southern Brazil is known.

Portuguese colonizers first brought cattle to what is now the state of Bahia in 1531. Cattle from Portugal and from the Azores Islands came in several voyages to the early settlements.

Contrary to what happened in the rest of the country, the first cattle were introduced in RS from Argentina in the early 17th century. Indian missions were formed by priests of the Jesuit Order in the north of Argentina. Some were built across the Uruguay River. Pe. Cristovão de Mendonza, one of the early missionaries, crossed this river with large herds for the flourishing new communities. This mission work lasted for 60 years in that part of South America.

In that short period, agriculture was established and Indians became good herdsman and supreme horseman. Large shipments of wheat were sent regularly to the Headquarters of the Order in Buenos Aires. Coal, copper, and gold were found in RS. Bronze bells were made in the missions. Schools, hospitals, churches, and the arts were established. Soon the missionaries began to care more about the people than with the interests of the Order. The missions became too powerful within the colonies. When the missionaries were ordered to dismantle the communities and return, most sided with the people. The civilization was destroyed on both sides of the Uruguay River by the Spanish Army, since during that time the crowns of Portugal and Spain were united. Besides the lives lost, an advanced political system with equitable social laws and high economic output was also lost.

When settlers came from Central Brazil to RS, they found large herds
of Criollo cattle living in the feral state and grazing native grass. An extractive industry based on these herds was established. First, animals were slaughtered only to export their horns and hides to the center of the country. Later, salt works were established, harbors constructed, and slaves brought into the state to prepare the meat. It was cured with salt and sun, and exported to central and northern Brazil. This system prospered, forming the first economic cycle.

In 1822, Brazil declared its independence from Portugal. Emperor Pedro I had family ties with royal houses in Europe. He and his son promoted immigration to Brazil, since only the coast was really occupied.

In 1825, the first German immigrants arrived in the central east part of RS. Fifty years later, the Italians arrived and were located in the northeastern portion of RS. Together, they occupied the agricultural land in the northern half of the state and developed its industry on the eastern side. The telephone book of Porto Alegre, capital of RS, traces about 40% of the family names to a German ancestry.

The southern half of the state holds the majority of the state's 13 million head of cattle and 11 million head of sheep. The pastoral vocation of this region is assured by its soil. Basalt formations are underneath a thin layer of soil, which is covered by native grass. Only small areas can be cultivated. Rice, soybeans, wheat, corn, millet, peas, and cultivated pastures are produced in the southern half of RS also.

Until the late 1800s, husbandry practices were limited to hot iron-branding, castration, and working the herds on horseback once a year, when animals were chosen to be trailed to the coast. Frontiers were being
settled and disputed by armed struggles. RS' separatist movement fought for ten years with Federal forces. Only with the declaration of the Republic, in 1889, did this period of almost continuous wars cease, and then ranchers and peons returned to tend their land and animals.

By that time, ranchers began to visit livestock shows in Uruguay and Argentina. These countries had established herds of "improved" breeds from Great Britain. Ranchers began to use bulls of these "improved" breeds on their criollo cows.

In 1883, the Agronomy School of Pelotas was founded. Shows, animal husbandry practices, animals, and the heritage were imported in a package from England.

In 1906, Herd Book Collares was initiated by Eng. Agron. Leonardo Collares and they have been active ever since, keeping pedigree records of all European beef breeds in their closed herd books. Today, Herd Book Collares has the official name of Associação Nacional de Criadores (ANC), with 700 plus members and broader responsibilities.

Specific breed associations were soon formed and, with the help from the State Department of Agriculture (SAGRI), began to promote grading-up to these breeds and selection based on breed standards and show regulations. Later, absolute performance levels have been incorporated into the list of selection criteria.

Several A.I. cooperatives, working with refrigerated and frozen semen, were operating from the middle to late '50s. Several private and one state A.I. company have been established. Official records show that about 5% of the cows are serviced by A.I. Many D.V.M.s work in the field preparing pellets and straws for owners of bulls. There are no official
records on this private activity. Imported semen is getting a smaller share of the market, being used almost exclusively in seedstock herds. In 1972, RS exported 70,000 tons of frozen meat to foreign markets. Since then, the tonnage has been reduced greatly. The demand exists now for cooked, canned meat. In 1972, half of the foreign currency earned from meat exports went to pay for the imported semen (300,000 vials). In 1984, Brazil became the second net exporter of meat (mostly manufactured) in the world, with 550,000 tons contracted.

The role Zebu will play in meat production in RS is anyone's guess. One composite breed (5/8 Angus, 3/8 Nelore), the Ibagé, after 30 years of work from the Exp. Stat. Cinco Cruzes, Bagé, RS, was recognized by the government in 1979, had its Association formed in the same year, and membership is increasing steadily. Another composite, the Santa Clara Cattle (1/2 Nelore, 1/2 Hereford), has 15 years of work from a particular breeder, Rubem Vasconcellos, in Rosário county, RS.

In 1971, ANC began a program (CDP) to take official repeated (at 45 days intervals) weights of purebred animals from birth to 720 days. This program has its operational costs paid for by the breeders. Animals participating in National or in the State shows must have these records. Participation now is at a rate of 3,500 new calves per year, from 300 herds. Breeders can preselect which animals will be recorded and only since 1978 has the program been computerized, allowing for comparisons among animals managed alike to be offered in the reports. The idea of weight at standard ages was introduced to breeders in the process.

Also in 1971, SAGRI opened a network of Central Testing Stations, following the example of the State of São Paulo which has conducted such
tests since 1951. Participation in 1983 was around 250 weaned male calves, but initially the numbers were higher. Until 1983, these stations followed the same procedures and feeding levels as was done in São Paulo and in the US, giving them many of the same problems. A more realistic approach was adopted. This year (1984), animals are being raised on cultivated pastures in a large central experiment station, until long (18 mths) yearling weights are recorded. Costs to the breeders (feed + health care) will be much less. Bulls will not be overconditioned for the breeding season and have less hoof problems (related to high energy intakes). Around 50% of the animals are approved at the completion of the test. Superior animals are candidates to be progeny tested.

In 1973, SAGRI began to progeny test outstanding animals from the central testing stations. The first team of bulls consisted of 4 Hereford and 3 Charolais bulls. They went under the most severe scrutiny known in a progeny test, from cariotyping the red blood cells from sires to making histological studies on the testis and ovaries of the progenies. Not one of these bulls was recommended for use through A.I., since every one failed in at least one of the several criteria. This applied research effort brought several state institutions together. New teams of bulls are tested every year. New groups are becoming larger. Outstanding bulls that are recommended by groups of breeders, that have on farm own or progeny records but were not tested in a central station, are involved. No ties were made among teams of bulls in the first years. The model used for the analysis is a fixed one, relying on experimental design (with a target of equal numbers of progenies in all subcells) and
on experimental conduct (extensive supervision in the cooperating herds and a post-weaning phase all on the same experiment station, as examples). A description of the initial methodology and complete results of the first team of bulls are reported by Poli (1977).

In 1974, ANC offered to commercial and seedstock breeders a ROP program to help them with within-herd selections. This work was suggested and developed by Professor Joel J. Kemper, then working at the Graduate School of the Federal University of RS (UFRGS), under an agreement between the University of Wisconsin and UFRGS. In its first year, 343 calves had their WW recorded. In 1981, 8882 calves were recorded. There has been a steady slow growth of the program without any real explosion in participation. Besides WW, calves are weighed again, by the end of the fall season, at 550 days (spring born) or 365 days (fall born). Conformation scores and descriptive grades are also taken at the weigh days. Computing services are hired from a company. Reports for bull and heifer selections, within herd sire evaluation and dam production lists are delivered to the breeders within 1 day to 3 months after the weighings.

Currently, the ROP program collects records on about 50% of the purebred calves registered per year (circa 10,000) on ANC (European breeds). A large expansion in numbers tested is needed, especially in the commercial bull area, where an estimated 30,000 replacement bulls are needed per year to service 4 million beef brood cows.

This historical overview can help trace parallels among countries which were once colonies. The productivity level of beef industries in different countries should not be equated to just genetic level of the
beef population (misconception responsible for the century long importation of genetic material), natural resources available or management skills of the producers. These levels can be better understood if market prices, costs of production, and effective demand are also considered. These last factors explain why Brazil has an "average" (with a huge positively skewed curve) beef consumption of less than 17 kg/year/person, which has been declining since 1979 and still Brazil is a net exporter. Given the low buying power of the majority of the population, internal market prices are stabilized politically to accommodate the situation. With prices defined by decree, there is little incentive to change production systems. If this government intervention is a hindrance on productivity of the beef herd, one has to recognize that it allows for at least a low ingestion of beef meat by the majority of the Brazilian population. Economic studies have shown that the use of cultivated pastures or of concentrate feeding would only be economical when offered to animals which would die otherwise or to rapidly finish developed steers to offer fresh meat in the long interval between the slaughtering periods (fall). During this period, only the price of frozen meat is regulated by the government. At best, one can hope for a slow evolution of the ratio of minimum wage to market price of meat. Only in this economically constrained environment for production, can realistic breeding plans be developed.

With the exception of National Sire Evaluation Programs in the US, the record systems available to the breeders to make breeding decisions are comparable. There are differences in sizes of purebred European beef cattle populations, relative participation by breeders on the programs
and some lag time on the application of methods. But this lag time is getting shorter. In 1970, the lag was at least some 20 years.

Cardellino-Stercken (1980) raised the question if any genetic progress is taking place for production traits in the beef cattle population in RS. The question can be extended backwards one hundred years. Age at slaughter (50% at 3.5 years and 50% at 4.5 years) and carcass weight (220 kg) has remained the same over this period, if one does not take into consideration improvements made in veterinary hygiene. Fertility rate is at a 50% level. Population growth of a species measures its adaptation to an environment. The beef cattle population in RS is a static figure and only old cows are sent to the market. As a rule, no selection is practiced or is possible on the female side of commercial herds.

In the same review paper, Cardellino-Sterken (1980) enumerated and sharply described as limitants to the genetic progress the following:

1) lack of a clear definition of the objectives for a beef herd;
2) lack of objectivity and accuracy in the evaluation of production traits;
3) lack of fair comparisons among animals raised in different environmental conditions;
4) overvalue given to the pedigree as an indicator of breeding value;
5) too much emphasis given to some individual animals or to some absolute records;
6) overstressing shows and judging as improvement tools;
7) too much importation of genetic material;
8) too small a participation in the available programs;
9) underutilization of records and reports for selection decisions; and
10) lack of convergence of the objectives and criteria of seedstock and commercial producers.

Historical background, appraisal of a situation, and technical knowledge are to be integrated and be used (invoking the social responsibilities of its beholders) to propose changes.

What can be done immediately is to more closely integrate the several programs offered both within and among institutions. More help from extension agencies is badly needed. This must be followed by a reshaping and/or improvements of some programs. This can be done using the present human and physical resources.

Further improvements will depend on developing better trained personnel. A threshold critical mass exists to accomplish this within RS. Using RS records will not only yield needed answers to local problems but will also better train the personnel.
III. LITERATURE REVIEW OF SOME ENVIRONMENTAL EFFECTS ON WEANING WEIGHT

To fairly compare animals born in the same herd-year-season, of the same sex and receiving equal management (this set of factors and all their interactions defines contemporary group (CG)), BIF (1981) recommend adjusting WW for linear age of calf and to use additive correction factors for age of dam. Fries (1974) and Guitou (1983) have made extensive reviews of works which studied these effects, most of them from the US. These reviews will simply be referred to here. Efforts will concentrate on those effects that will be studied but have been less often investigated.

A. Date of Birth and Age at Weaning

Most ROP programs record at least the birth date, weaning date, and weight. Several relationships exist among these variables and others. The simple methodology recommended by BIF (1981) should be improved, at least for RS' data, given the results found by Ferreira (1982).

Growth from birth to weaning is a very complex process with a multitude of variables affecting it, not always in a linear or in an additive manner. Most papers present partial or individual facets of it only. Too few of these reports integrate some of the variables.

Before going to this peacemeal approach, an integrated overview is useful. A simplified verbalization of this process is as follows:
1. Birth date

Birth date has a direct effect on birth weight and post-partum milk yield. Within the year, birth date also determines how good an environment the calf will have in which to grow.

The environment of the cow influences birth weight and, since most of the weight gain of the fetus occurs in the last third of the gestation period, it is generally derived from this that the environment is especially important in this period. In RS, fall born calves are heavier at birth than early spring born ones.

Birth date also affects milk yield of the dam, by determining her body condition at the birth of the calf and the pasture quantity and quality available to maintain milk production.

2. Environment from birth to weaning

The calf's environment during his growth period from birth to weaning is by and large a consequence of his birth date. Availability of pasture in the first three months of life will affect the calf's weight through the milk production of his dam. After this approximate point (zebu calves and crosses are dependent for a much longer period), direct consumption begins to count more in the total nutrients ingested. Quality of the grass is paramount for direct utilization by the young calf.

3. Growth curve

Absolute weights of growing animals, when plotted against age, follow a sigmoid curve, not a linear one. A linear approximation can be
used only in small intervals of age. This underlying physiological fact is interacting with the previous factors and it is mostly confounded with them.

The only known work done putting some of the above factors together was by Silveira Júnior (1979), in another setting. Using growth records from the breed Ibagé, in RS, he utilized as a growth model the Brody's cubic root of the weight, to which periodic components were added in order to eliminate seasonal variation. After performing the harmonic analysis, it was observed that the periodic oscillation was due mainly to the influence of the semestral and annual waves. The objective of the work was to predict slaughter weight, on a given date, from earlier measurements.

All the papers to be reviewed have a common element missing; all have only one explanatory variable (date of birth or age at weaning) to describe the two events. They are really superimposed and in most cases (calves born in the same season and all weaned at the same date), collinearity precludes independent estimation of both effects. Thus, alternating names given by the authors does not guarantee that the named effect is really isolated from the other neither that it is the predominant one.

Scott et al. (1976) made a review and discussion of several research results which investigated environmental factors associated with summer and fall growth rates of cattle and sheep. In their Figure 1, they plot several results, from both hemispheres, showing liveweight gains of 4- to 9-month old cattle under grazing conditions. None of their plots show a pure linear relationship of liveweight with age or with period of the
year. Their understanding of seasonal effects can be depicted from: "As summer progresses, pasture species start to mature, with a subsequent decline in nutritive value. In addition, less nutritious summer growing species may become dominant as the season progresses and would lead to a lowering in overall digestibility of the herbage." These observations fit perfectly for RS.

Schaeffer and Wilton (1974) studied several environmental effects and interactions on preweaning average daily gain (ADG) of Angus and Hereford Canadian ROP records. Editing limits on weaning weights were 68-431 kg, and on age at weaning were 130-280 days. For Angus, the regressions of ADG on age at weaning were $b_1 = -0.0010$ kg/d and $b_2 = -0.000008$ kg/d$^2$; for Herefords they were $b_1 = -0.0006$ kg/d and $b_2 = -0.000006$ kg/d$^2$. They remarked: "Although the hypothesis tests were highly significant these regressions have little effect on ADG considering the range of age encountered. The inclusion of age at weaning and age at weaning squared as covariates in the model could be ignored in future analysis on preweaning ADG."

If one does some calculations with their coefficients some questions can be raised about their conclusions, on the grounds of what is biologically important. Consider BIF (1981) recommendation for adjusting WW linearly for age of calf to a standard age of 205 days, within the limits of ± 45 days. According to the results above, one should correct the ADG of a 160-day old Angus calf by subtracting 0.0288 kg/d (0.01485 kg/d) and the ADG of a 250-day old by adding 0.0612 kg/d (0.03915 kg/d). Values within parenthesis refer to corresponding Hereford figures. The numbers by themselves look small indeed. Considering a contrast between
two calves with the same genetic make-up for absolute growth rate and being weaned at those opposite age limits, without correcting measured ADG for non-linear age of calf (AOC) effects, would result in a bias of 8.10 kg (4.86 kg). Looking at the age limits on the data studied, the bias for a similar comparison would be 22.50 kg (13.50 kg).

As any of these errors can easily influence the decision to cull or keep a calf (and/or its dam) and the adjustments are so easily made, Schaeffer and Wilton's conclusion becomes surprising. What is at play are not just random errors but errors with a marked tendency, the very definition of bias, to overestimate genetic values of young calves and to underestimate older ones. The established practice of following AI with clean-up bulls enlarges the consequences of this bias, which is even worse in the case of using different genetic grouping for AI and natural service bulls.

Pherigo et al. (1969) studied the association between day of birth and corrected (205 days) weaning weight in beef cattle. They found a significant quadratic effect of day of birth on WW. But the coefficient for some years was positive while for others it was negative. The birth date span was from day 15 to day 165 (Julian dates). The extremes of the yearly parabolas were always around day 90. But for some years the parabola was upward and in others it was downward. Their study involved calves only from one herd and considered only calves weaned in a fixed month (October). This situation determines a complete negative association between the effects of day of birth and of AOC at weaning on WW. As they preadjusted the records linearly for AOC, their regression coefficient estimates contain deviations from the assumed linearity. In
such a situation, one can name the effect with whatever label one desires. Different years may deserve different labels also. Even so, the authors could explain the variability in the results by the registered rainfall. They concluded that calves born late in the spring appeared to have a greater reduction in their adjusted WW when drought conditions occurred in mid to late summer.

Cundiff et al. (1966) observed that age of dam accounted for 7% of the variation in WW. This same figure (7%) of the total variance was accounted for by month-of-birth. The range of effects went from -10.4 kg (October) to +13.6 kg (March). They found significant interactions between month-of-birth and management (creep vs non-creep), breed and type of pasture.

These two last reports are suggesting both the magnitude of the effects of seasonality and also the difficulty of establishing correction factors that can be safely used by the industry. This is especially true if herds are spread over large geographic areas. This may also be saying that it is impossible to use a single set of correction factors for date of birth (seasonality effects) across herds and years, like it is done with sex and age of dam.

In an extensive experiment conducted in Rhodesia, Richardson et al. (1979) when regressing calf body gain from birth to 150 days (age constant) on date of birth (1 Oct = day 1), found a linear coefficient of -0.097 kg/d, after correcting for the effects of birth weight and milk production. They concluded that the negative relationship between growth to 150 days and date of birth was probably a reflection of the quantity and quality of grass available to suckling calves from 90 days of age.
on, when grass forms a substantial part of their diet. These results are in agreement with earlier study in Rhodesia, conducted by Vorster (1964). This author found a progressive decrease in the weight of calves born from December to March. Since in every year (1939 to 1952) all calves (5,825) were weaned at the same time, irrespective of the date of birth, the above author wrote: "these differences in weight may obviously be attributed to variations in age and to the fact that calves born early have the full advantage of the growing season and would naturally tend to grow more rapidly than those born latter." In Vorster's Table 38, ADG is presented according to month of birth (December to March) as 0.342, 0.331, 0.320, and 0.298 kg/d, respectively. One has to consider the fact that this difference would be larger if non-linear AOC effects had been simultaneously estimated. Even though Vorster's results are biased, they reflect an underestimate of the month-of-birth effects that are of a magnitude big enough to preclude any attempt to drop them from a model in the name of parsimony.

Barlow et al. (1974), in Australia, found significant linear and quadratic regression coefficients of ADG on weaning age. They state that "little variation was removed by these regressions." The average age at weaning was 208 ± 31 days which resulted from concentrated breeding and calving seasons. This is the prevalent situation in field beef data and it is known that when the independent variable has such a distribution the regression sum-of-squares (SS) will be relatively small. Actually, this regression SS can be manipulated. As suggested in elementary statistics texts, if one wants to increase the coefficient of determination ($R^2$), it is enough just to enlarge the range of the
independent variable or to collect more observations at more extreme values. This is saying that regression SS are distribution dependent and are of lesser value to assess the practical importance of a regression coefficient than its value by itself or relative to its standard error.

The same above authors, Barlow et al. (1974), used a second model, fitting regressions within each "CG" (herd, year, and management). They broke down these nested regressors into a part common to all "CG" and an interaction part. Only 4 of the 27 interaction terms were significantly different from zero and the authors mistakenly concluded that "correction of ADG for age at weaning is not justified in most years." Their results show that ADG should be adjusted for age at weaning and that, within the limits of inference set by their data, very few "CG" would deviate enough from the common regressors as to require specific regressors.

Pabst et al. (1977) analyzed data collected on the Meat and Livestock Commission's pedigree recording scheme. In a model which explained from 36 to 49% of the variation (several breeds) in 200-day weights, they found significant month-of-birth effects. For most of the breeds with large numbers analyzed, these effects followed a seasonal pattern. Even so, the authors suggested correction factors for whole seasons, lumping together months with more than two standard errors of difference among their least squares constants.

As far as their review goes, it is evident that the problem is still unsolved, that seasonality and non-linear AOC effects exist, are important and remain entangled. As a result, no better options exist than to follow BIF (1981) recommendations.
It may be helpful to have a brief look at how the dairy industry has solved this problem. Besides other advantages, their situation is at least one component less complex. Instead of having just one indirect measurement (weaning weight of the calf) they take repeated measurements of milk production. Instead of having to look simultaneously to the effects of lactation curve, growth curve of the calf and seasonality, with only one observation (at weaning) per calf, the dairy industry needs to handle only lactation stage and month of calving and has a better qualified data base to do it.

Correction factors for region, age of cow, breed and month of calving to adjust milk, fat, solids-non-fat and protein yields (Norman et al., 1978) are in use by the Dairy Herd Improvement Association (DHIA). These consist of different set of multiplicative factors that adjust for all these factors simultaneously.

The basic work is the one from Miller et al. (1970). They used over 240,000 records from purebred Holstein cows from the Northeastern region of the US. Using "maximum likelihood" methods they estimated the joint effects of age of cow and month of calving on milk yields.

They concluded that "Monthly mature-equivalent factors could not be adequately grouped into a few seasons and that the most appropriate method of adjusting for both age and month of calving was by multiplicative factors which simultaneously adjust records to the expected yield, had all other conditions remained the same." Their final remarks were: "Use of the joint month-age adjustment should improve the accuracy of both within-herd and between-herd comparisons of individual records. However, of greater economic importance is the removal of
biases from AI sire proofs which should result from this adjustment.
These biases in AI proofs occur because the age distribution of daughter
records is greatly different from old bulls than from young bulls, and
because first AI proofs, on which decisions to keep or cull are often
made, frequently contain only records from very young daughters calving
in a single season." Extension of these remarks to National Sire
Evaluation Programs (NSEP) in beef cattle is automatic. A problem close
to this was reported by Pollak et al. (1977). They found that the
genetic group (based on sire's birth year) solutions were heavily
influenced by the use of breed-wide corrections factors for age of dam
and these correction factors were inadequate for the herd studied.

The work from Miller et al. (1970) was replicated in Canada. Mao et
al. (1974) suggested constants for joint adjustment for age and seasons,
given that the interactions were present. Their comments were:
"Seasonal change is continuous and grouping of months into a few seasons
is arbitrary and can not be done satisfactorily, since differences
between months in the same season are sometimes large." Their statement
is saying that they prefer smaller than larger classes, but they are
still grouping (all days of the month are the same) what is a pure
continuous function.

B. Nursing Status of the Dam

This variable can be defined as the nursing status of the dam the
year before the present calf is born and weaned. In suboptimal
environments where a cow can not maintain yearly calving intervals one
can expect that cows which had a longer "resting period" to rebuild their
body reserves will be able to maintain a higher milk yield during the following nursing period.

Freeden et al. (1981) studied the effects of gestation and lactation on growth patterns of cows under two environments in Canada. Females at Brandom (the "good" environment) gained weight during gestation and lost weight during nursing. Those at Manyberries (an "average" environment) invariably gained weight during nursing with substantial gestation weight loss in all 3 years of the experiment. "Nursing status of the female had large effects on weight changes, with barren years providing the opportunity for large compensatory gains."

These results raise important questions on what is universally taken to be true: that the lactation period is a more stressful phase than the gestation period. Maybe a beef cow can more successfully limit the proportion of the metabolizable nutrients to be directed to the lactation process than she can limit the fetus growth, when in a period of negative balance of energy or total nutrients.

Lôbo et al. (1983) studied several environmental effects on milk yield of Gyr cows in Brazil. They reported a curvilinear effect of length of previous dry period on milk yield, confirming the fact that there is an optimum resting period for the mammary gland, after which yield starts falling again.

Vorster (1964), in Rhodesia, found a difference of 15.9 kg between weaning weights of calves whose dams had alternate vs. consecutive calvings. These groups of cows had weights of 395 and 385 kg, respectively. This author concluded that, owing to the prolonged winter (dry season) period, cows are not able to build up sufficient body
reserves after weaning to stand the high strain of two successive pregnancies and lactations.

Thorpe et al. (1981) analyzed several native breeds and their crosses in Zambia. Average WW of calves from dams which were dry by mating was 168.2 kg; while the ones from lactating dams on the previous mating was 159.1 kg. More striking was the effect of dam status on fertility. Calving rate of dams dry at mating was 89.1%; contrasted with 39.8% for those dams that were lactating at mating and that had calved late in the season. As a comparison to these figures, several observations made in RS have shown that, if first calf dams are managed like mature cows, only 20% of them are bred in the immediate breeding season.
IV. DATA DESCRIPTION (INCLUDING BIOLOGICAL MODEL)

Hereford and Polled Hereford calves, with ROP records, born in fall and spring calvings from 1974 to 1979, raised by their dams on native and cultivated pasture, in conditions fairly representative of the commercial production in RS have been analyzed.

Only the Hereford (in Brazil, the polled is considered a variety, with its merits and large representation, but are kept in the same Herd Book) breed was analyzed since it comprises just above 50% of all records. Other 5 breeds (Angus, Charolais, Devon, Ibagé and Santa Gertrudis) have 5 to 10% of the records each. The balance is made up of several much smaller breeds and crosses. This proportion of Herefords is a good approximation to its overall contribution to the RS' beef herd. This proportion is even higher in the southern (border) region, which have larger ranches than in other parts of the state.

Other breeds need to be studied also but to do this would mean an excessive enlargement of the problem and the numbers would be too small to make useful inferences.

Only records from calves born before the spring of 1979 were included because at least 90% of these were collected and all processed by the same person. Every herd and owner, each with its special practices and particularities, is known. The extensive knowledge about the data is judged as a major input for this study. More recent records could be included, allowing the use of twice as many records, but the degree of understanding about them would be less.

Herds included in this study are all located in the southern part of
RS, from the sea coast to the borders of Uruguay and Argentina. All are near to or south of highway BR-290. There are too few herds in each of the regions to allow for the inclusion of region as factors in the study, although soil, elevation and climatic differences would warrant this.

Individual records from the ROP file can be segmented in four parts, as follows:

1) CG identification, with breeder, breed or cross, herd, year, season, sex, and management codes.

2) Birth records, with the same needed information supplied by a pedigree record, with the option for birth weight.

3) Weaning records, with weights of calf and cow, date, and conformation scores (USDA and Ankony systems).

4) Yearling records, with weight (365 or 550 days), date and conformation scores.

This information, with all other computed figures, as gains, corrected and adjusted weights, ratios, and indexes are packed into a 300 byte record.

A. Editing

The total number of records read was 40,129. Herefords comprised 21,066 of these. Deleting records from breeders with only one calf crop and birth records of dams born before 1969 cut the number to 16,004.

This file was sorted by dam within herd and by calf birth date within dam. Calving intervals were calculated and assigned to the last calf record from each interval.

From this file, the following editing criteria were used:
1) Only fall or spring calf crops from 1974 to 1979 were kept.

2) Records were deleted which were from 2 years old dams or younger and from 17 year old dams or older.

3) Complete management groups were deleted when codes were indicating calves that were nursed by a foster cow, or received some grain supplement (creep feeding) or were full-fed, or developed severe clinical signs of hoof-and-mouth disease, babesiosis, anaplasmosis, or were blind (pink-eye related).

4) Individual records from calves with a condition score of 1, from a scale of 1 to 5, indicating some unspecified severe health problem (cachexia) were deleted.

5) Records from castrated calves or from twins were deleted.

6) Records of calves from unknown parentage were deleted.

7) Records outside arbitrary confidence limits for several explanatory variables were printed out and examined on a case by case bases. A few were deleted representing some rare possible extreme values or some errors in the data base.

8) After the editings some further complete herds, which became too small or were not repeated across years, were deleted.

9) Records of calves from sires with less than 5 progeny were excluded to reduce the variance of the sire variance estimate and to reduce the size of the task.

The editings were not done in the straightforward way as it is related but more in a sequential analysis-edit-analysis steps. Only after the work was completed was it realized that some further edits should have taken place.
B. What Was Not Edited and Why Not

No one record was eliminated because of the magnitude of the WW itself.

Some authors suggest, for messy data, to delete all records farther than 2 standard deviations from the overall mean. This procedure has an impact on estimates of error variance. Estimated parameters and variance components will behave nicely but this action will limit inferences that can be made. Just how established this procedure is among animal breeders is not known and one can only hear or see a few scattered tips about it. Extreme individuals in one trait or animals which are at extremes of pleiotropic networks are eagerly sought. But if the practice of trimming large data sets has widespread use among animal breeders, these target populations are seldom studied and, hence, little known. This may cause realized genetic gains to fall short of their expectations.

Huber (1981) made his plea for robust estimation procedures in comparison with a two-step approach (that would be by itself more correct than what is described above), defined as cleaning the data by applying some rule for outlier rejection followed by classical estimation and testing procedures, as follows:

"1) It is rarely possible to separate the two steps clearly. For instance, in multiparameter regression problems outliers are difficult to recognize unless we have reliable, robust estimates for the parameters.

1Extension of this assertion for multifactorial problems seems logical."
2) Even if the original batch of observations consists of normal observations interspersed with some gross errors, the cleaned data will not be normal (there will be statistical errors of both kinds, false rejections and false retentions), and the situation is even worse when the original batch derives from a genuine nonnormal distribution, instead of from a gross-error framework. Therefore, the classical normal theory is not applicable to cleaned samples, and the actual performance of such a two-step procedure may be more difficult to work out than that of a straight robust procedure.

3) It is an empirical fact that the best rejection procedures do not quite reach the performance of the best robust procedures. The latter apparently are superior because they can make a smooth transition between full acceptance and full rejection of an observation."

Measurement errors on weights and dates, misidentification of ancestry, bad information relative to management codes and errors in transcription exist in beef cattle data sets. Besides having influence on estimated parameters they lead to wrong decisions in selection among animals. This problem has to be dealt with by animal breeders.

A maximum likelihood robust estimator, in the classification given by Huber (1981), is proposed in Chapter V, section A.3, was used and comparative results are given in Chapter VII.

---

Most probably this is the situation animal breeders envision when analyzing field data. The organization owning the records (e.g.: breed associations or A.I. companies) may "clean" the data by themselves so that the animal breeder is saved from making a decision on what to do with the observations that are not well behaved.
C. Edited Data Description

A total of 8,323 records were used in this study. They can be grouped by breeders (13), herds (18), registered (PO = "Pure by Origin," 8 herds, 2479 obs., 29.8% of total) or grade (PC = "Pure by Crossing," 10 herds, 5844 obs., 70.2% of total), year (1974 to 1979), season (fall, with 1429, or spring, with 6894 records), sex (48.3% males and 51.7% females), contemporary groups (214), or sires (130 in 3 disconnected sets).

Table 1 gives the distribution of records across herds, years and seasons nested in years. From the 198 possible subclass cells, 103 cells are empty, or 47.98% are filled. One of the objectives of the editing phase was to check for confounding of herds with year*season. There is no comparative measure for the sparseness in this data set but fewer cells are probably filled than in larger and longer time data bases. Deleting some herds, a few that discontinued participation after 1 year and almost all that joined the ROP program more recently, accomplished this level of subclass filling. Before the last two editing steps were performed, 9398 observations were present, from 265 CG and 232 sires. From the 61,480 possible subclasses only 1,041 were filled, or 1.69%. In a two-way model, the left hand side (LHS) of the normal equation (NE), a 498*498 matrix, would have only 1.44% of its subclasses filled.

After the final editings, 3.12% of the CG*sire subclasses were filled.

The name given to the season in which a calf crop is born follows no strict rules or limits in terms of Julian dates. "Productions" (as calf crops are called in RS) are managed like a spring one (weaned by fall and
Table 1. Distribution of weaning weight records among herds, years, and seasons within years

<table>
<thead>
<tr>
<th>Year</th>
<th>Season</th>
<th>CAA</th>
<th>CAB</th>
<th>EZ</th>
<th>FBA</th>
<th>FBB</th>
<th>IF</th>
<th>JAA</th>
<th>JAB</th>
<th>MAA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974</td>
<td>S</td>
<td>108</td>
<td>*</td>
<td>*</td>
<td>15</td>
<td>20</td>
<td>18</td>
<td>37</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>F</td>
<td>47</td>
<td>100</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>33</td>
<td>73</td>
<td>*</td>
<td>48</td>
<td>43</td>
<td>21</td>
<td>34</td>
<td>35</td>
<td>*</td>
</tr>
<tr>
<td>1976</td>
<td>F</td>
<td>52</td>
<td>110</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>15</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>50</td>
<td>*</td>
<td>*</td>
<td>35</td>
<td>100</td>
<td>26</td>
<td>22</td>
<td>52</td>
<td>6</td>
</tr>
<tr>
<td>1977</td>
<td>F</td>
<td>23</td>
<td>83</td>
<td>27</td>
<td>20</td>
<td>10</td>
<td>20</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>75</td>
<td>188</td>
<td>362</td>
<td>73</td>
<td>102</td>
<td>28</td>
<td>19</td>
<td>33</td>
<td>2</td>
</tr>
<tr>
<td>1978</td>
<td>F</td>
<td>68</td>
<td>74</td>
<td>58</td>
<td>*</td>
<td>10</td>
<td>11</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>88</td>
<td>167</td>
<td>405</td>
<td>53</td>
<td>112</td>
<td>68</td>
<td>48</td>
<td>77</td>
<td>13</td>
</tr>
<tr>
<td>1979</td>
<td>F</td>
<td>97</td>
<td>121</td>
<td>115</td>
<td>11</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>101</td>
<td>112</td>
<td>395</td>
<td>51</td>
<td>74</td>
<td>52</td>
<td>20</td>
<td>52</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>742</td>
<td>1028</td>
<td>1362</td>
<td>291</td>
<td>466</td>
<td>261</td>
<td>161</td>
<td>286</td>
<td>34</td>
</tr>
</tbody>
</table>

*aS = Spring and F = Fall Calving Season.*
<table>
<thead>
<tr>
<th>NPA</th>
<th>NPB</th>
<th>PA</th>
<th>PM</th>
<th>PlA</th>
<th>PlB</th>
<th>WP</th>
<th>Z9</th>
<th>JFS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>*</td>
<td>52</td>
<td>52</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>302</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>41</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>188</td>
</tr>
<tr>
<td>12</td>
<td>*</td>
<td>60</td>
<td>73</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>432</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>177</td>
</tr>
<tr>
<td>7</td>
<td>31</td>
<td>36</td>
<td>138</td>
<td>*</td>
<td>424</td>
<td>*</td>
<td>78</td>
<td>*</td>
<td>1005</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>40</td>
<td>*</td>
<td>129</td>
<td>*</td>
<td>352</td>
</tr>
<tr>
<td>23</td>
<td>26</td>
<td>59</td>
<td>99</td>
<td>18</td>
<td>18</td>
<td>480</td>
<td>*</td>
<td>34</td>
<td>1639</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>28</td>
<td>*</td>
<td>249</td>
</tr>
<tr>
<td>*</td>
<td>34</td>
<td>81</td>
<td>68</td>
<td>2</td>
<td>*</td>
<td>53</td>
<td>99</td>
<td>78</td>
<td>1446</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>15</td>
<td>25</td>
<td>*</td>
<td>*</td>
<td>79</td>
<td>*</td>
<td>463</td>
</tr>
<tr>
<td>*</td>
<td>144</td>
<td>10</td>
<td>26</td>
<td>523</td>
<td>497</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>2070</td>
</tr>
<tr>
<td>42</td>
<td>91</td>
<td>288</td>
<td>615</td>
<td>45</td>
<td>69</td>
<td>1520</td>
<td>596</td>
<td>496</td>
<td>8323</td>
</tr>
</tbody>
</table>
long yearling weights taken the next fall) or like a fall one (weaned by late
spring and yearling weights taken in the fall). With this arrangement,
at most two visits per year are needed to collect records from a herd.
Some herds have the production concentrated in early spring; others try
to get extra calves and have the birth season extended into the summer.
A couple of summer and winter productions were deleted.

Males in the data set were 4020, from 105 CG. The average size of a
male CG was 38.3 ± 43.1 calves. Females were 4303, from 109 CG. Average
size of a female CG was 39.5 ± 41.9. The numbers are in agreement with
what was empirically observed in some herds; a larger number of males
are kept in creep-feeding or are full fed. Also, some herds castrate
poor animals before weaning, especially fall born ones. This means that
male CG are smaller and, believing in the breeders' ability, have a
reduced genetic variability. This is not the rule but is more frequent
in registered (PO) herds. If this happens in other field data studies,
questions can be raised about real causes of significant interactions
involving sex of calf, reported in the literature. Biological
explanations for higher heritabilities found for females may also be in
jeopardy.

Table 2 presents the distribution of fall birth dates simultaneously
with weaning ages. The table shows a strong association between the
variables given the fact that almost all calves from a production are
weaned on the same day. Few productions are too large to have all calves
weighted and classified on a single day. The association is not perfect
and allows for simultaneous estimation of both effects on WW. Some
researchers express the justifiable fear that when all the calves from a
Table 2. Distribution of fall birth dates (in the Julian calendar) and corresponding weaning ages

<table>
<thead>
<tr>
<th>Birth Date</th>
<th>160</th>
<th>170</th>
<th>180</th>
<th>190</th>
<th>200</th>
<th>210</th>
<th>220</th>
<th>230</th>
<th>240</th>
<th>250</th>
<th>260</th>
<th>270</th>
<th>280</th>
<th>290</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
</tr>
<tr>
<td>64</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>2</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>*</td>
<td>5</td>
<td>*</td>
<td>19</td>
</tr>
<tr>
<td>74</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>3</td>
<td>3</td>
<td>32</td>
<td>6</td>
<td>13</td>
<td>3</td>
<td>15</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>84</td>
<td>1</td>
<td>*</td>
<td>*</td>
<td>8</td>
<td>13</td>
<td>63</td>
<td>48</td>
<td>35</td>
<td>10</td>
<td>51</td>
<td>14</td>
<td>2</td>
<td>*</td>
<td>245</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>*</td>
<td>*</td>
<td>2</td>
<td>10</td>
<td>36</td>
<td>111</td>
<td>95</td>
<td>45</td>
<td>52</td>
<td>9</td>
<td>1</td>
<td>*</td>
<td>*</td>
<td>361</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>*</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>12</td>
<td>123</td>
<td>118</td>
<td>24</td>
<td>32</td>
<td>17</td>
<td>4</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>334</td>
</tr>
<tr>
<td>114</td>
<td>*</td>
<td>1</td>
<td>3</td>
<td>17</td>
<td>87</td>
<td>51</td>
<td>12</td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>199</td>
</tr>
<tr>
<td>124</td>
<td>*</td>
<td>1</td>
<td>3</td>
<td>33</td>
<td>26</td>
<td>3</td>
<td>16</td>
<td>4</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>86</td>
</tr>
<tr>
<td>134</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>34</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>59</td>
</tr>
<tr>
<td>144</td>
<td>1</td>
<td>3</td>
<td>14</td>
<td>1</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>20</td>
</tr>
<tr>
<td>154</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>3</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>7</td>
<td>31</td>
<td>92</td>
<td>144</td>
<td>236</td>
<td>329</td>
<td>221</td>
<td>136</td>
<td>98</td>
<td>68</td>
<td>30</td>
<td>8</td>
<td>3</td>
<td>1406</td>
</tr>
</tbody>
</table>

\[a\] Values are centers of classes from 155-164, 165-174, ..., 285-295.

\[b\] Birth dates in Julian dates. Values are centers of classes from 50-59, 60-69, ..., 150-160.
CG are weaned on the same day, the absorption of CG will wipe out one of these effects and set one variable to zero. This may be the case if one of the variables included were weaning date. This could get confounded with CG. But what the absorption of CG does to the continuous variables (like date of birth and weaning age) is just to center them around their CG means, leaving the association unchanged.

Average birth date for fall born calves is day 100.7 (April, 10th).

Table 3 presents the distribution of spring birth dates simultaneously with weaning ages. The same kind of association (calves born early in the season are older at weaning) is shown as for fall born calves. The upper left corner of the table indicates the presence in the data of 45 animals born late in the winter that were managed almost as a fall calf crop, being weaned in early summer. It is not understood if their occurrence is detrimental, innocuous, or beneficial to the analysis. As their number is small, they were kept with the hope that they would help to break strong associations. It is worth noting that the spring calving season spreads out for a longer period, across and within herds, with a corresponding larger range in weaning ages.

Average birth date for spring born calves is day 275.9 (Oct., 3rd). The distribution of birth dates suggests a mild bimodal shape. This could be explained empirically by having heifer and dry (on previous mating season) cows calving earlier in the season and nursing (on previous mating season) cows calving later, generally. This is not a claim since it was not numerically studied.

It is important to notice from Tables 2 and 3 that if weaning age limits at 160–250 days were followed, 159 fall calves (11.31% of fall
Table 3. Distribution of spring birth dates (in the Julian calendar) and corresponding weaning ages

<table>
<thead>
<tr>
<th>Birth Date</th>
<th>120</th>
<th>130</th>
<th>140</th>
<th>150</th>
<th>160</th>
<th>170</th>
<th>180</th>
<th>190</th>
<th>200</th>
<th>210</th>
<th>220</th>
<th>230</th>
<th>240</th>
<th>250</th>
<th>260</th>
<th>270</th>
<th>280</th>
<th>290</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>174</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>2</td>
<td>18</td>
<td>*</td>
<td>47</td>
</tr>
<tr>
<td>184</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>2</td>
<td>5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>*</td>
<td>16</td>
<td>*</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>194</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>7</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>14</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>204</td>
<td>*</td>
<td>*</td>
<td>2</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>16</td>
<td>12</td>
<td>23</td>
<td>16</td>
<td>37</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>214</td>
<td>*</td>
<td>2</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>*</td>
<td>16</td>
<td>12</td>
<td>23</td>
<td>16</td>
<td>37</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>224</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>27</td>
<td>14</td>
<td>23</td>
<td>64</td>
<td>91</td>
<td>47</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>234</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>11</td>
<td>18</td>
<td>29</td>
<td>54</td>
<td>107</td>
<td>75</td>
<td>19</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>244</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>4</td>
<td>23</td>
<td>76</td>
<td>89</td>
<td>100</td>
<td>142</td>
<td>82</td>
<td>36</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>254</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>5</td>
<td>24</td>
<td>36</td>
<td>84</td>
<td>172</td>
<td>213</td>
<td>81</td>
<td>25</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>264</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>37</td>
<td>32</td>
<td>90</td>
<td>155</td>
<td>241</td>
<td>126</td>
<td>42</td>
<td>12</td>
<td>14</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>274</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>1</td>
<td>26</td>
<td>35</td>
<td>89</td>
<td>176</td>
<td>202</td>
<td>94</td>
<td>42</td>
<td>24</td>
<td>15</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>714</td>
</tr>
<tr>
<td>284</td>
<td>*</td>
<td>*</td>
<td>1</td>
<td>2</td>
<td>24</td>
<td>52</td>
<td>132</td>
<td>171</td>
<td>205</td>
<td>71</td>
<td>26</td>
<td>26</td>
<td>19</td>
<td>5</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>737</td>
</tr>
<tr>
<td>294</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>10</td>
<td>94</td>
<td>159</td>
<td>212</td>
<td>165</td>
<td>98</td>
<td>35</td>
<td>19</td>
<td>21</td>
<td>5</td>
<td>*</td>
<td>13</td>
<td>2</td>
<td></td>
<td></td>
<td>833</td>
</tr>
<tr>
<td></td>
<td>304</td>
<td>314</td>
<td>324</td>
<td>334</td>
<td>344</td>
<td>354</td>
<td>362</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_birth</td>
<td>*</td>
<td>*</td>
<td>15</td>
<td>74</td>
<td>138</td>
<td>182</td>
<td>118</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_dates</td>
<td>*</td>
<td>8</td>
<td>72</td>
<td>79</td>
<td>120</td>
<td>79</td>
<td>63</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>4</td>
<td>23</td>
<td>41</td>
<td>57</td>
<td>40</td>
<td>43</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_centers</td>
<td>*</td>
<td>14</td>
<td>17</td>
<td>33</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_from</td>
<td>2</td>
<td>11</td>
<td>11</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_to</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_Total</td>
<td>19</td>
<td>34</td>
<td>357</td>
<td>420</td>
<td>594</td>
<td>420</td>
<td>357</td>
<td>291</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

^Values are centers of classes from 115-124, 125-134, ..., 285-295.

^Birth dates in Julian dates. Values are centers of classes from 170-179, 180-189, ..., 350-359, 360-365.
total) and 1,548 spring calves (22.38%) wouldn't be analyzed. Also, breeders wouldn't have information on which to base their selections. The alternative, to wean a production at different dates, implies more visits, more travelling expenses and smaller CG.

Table 4 gives the distribution of dams by age and by nursing status (STATUS). AOD at birth of the calf less than 3 years were eliminated from the analysis since they were too few and also because it indicates that they were not raised on pasture (native or cultivated). For similar reasons AOD=3 for STATUS higher than 1 and AOD=4 for STATUS=3 were also eliminated.

The categorization of dams by STATUS was done as follows:

1) STATUS 1: includes first-calf heifers, dams with only one calf recorded, few dams with a calving interval (CI) of 273 to 290 days and dams with a CI larger than 900 days. This is a group of dams with "no previous information."

2) STATUS 2: dams with a calving interval of 290 to 450 days. This is the "consecutive calving" group.

3) STATUS 3: dams with a calving interval of 451-900 days. This is the "alternate calving" group.

A categorical partition of dams depending both on CI and on whether the last calf nursed his dam until weaning or not was tried. Partition by CI had 5 classes: only one calf; CI around 1 year; 1.5 years; 2 years; and above 2 years. Together with the other dichotomy (BY SIDE or not) 9 classes were formed. A preliminary analysis did not show an understandable pattern. Differences among effects of CI at 1.5 years and longer were non-significant. Some of the 9 classes had very small
Table 4. Distribution of weaning weight records according to age (AOD) and nursing status (STATUS) of the dam

<table>
<thead>
<tr>
<th>Age of Dam</th>
<th>Status *</th>
<th>Status 2</th>
<th>Status 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>818</td>
<td>*</td>
<td>*</td>
<td>818</td>
</tr>
<tr>
<td>4</td>
<td>1414</td>
<td>123</td>
<td>*</td>
<td>1537</td>
</tr>
<tr>
<td>5</td>
<td>845</td>
<td>220</td>
<td>181</td>
<td>1246</td>
</tr>
<tr>
<td>6</td>
<td>624</td>
<td>265</td>
<td>384</td>
<td>1273</td>
</tr>
<tr>
<td>7</td>
<td>499</td>
<td>295</td>
<td>290</td>
<td>1084</td>
</tr>
<tr>
<td>8</td>
<td>396</td>
<td>201</td>
<td>234</td>
<td>831</td>
</tr>
<tr>
<td>9</td>
<td>295</td>
<td>204</td>
<td>177</td>
<td>676</td>
</tr>
<tr>
<td>10</td>
<td>157</td>
<td>139</td>
<td>127</td>
<td>423</td>
</tr>
<tr>
<td>11</td>
<td>86</td>
<td>54</td>
<td>74</td>
<td>214</td>
</tr>
<tr>
<td>12</td>
<td>56</td>
<td>30</td>
<td>50</td>
<td>136</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>8</td>
<td>19</td>
<td>42</td>
</tr>
<tr>
<td>14</td>
<td>8</td>
<td>4</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>*</td>
<td>2</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>5218</td>
<td>1546</td>
<td>1559</td>
<td>8323</td>
</tr>
</tbody>
</table>

*Status: 1 = no previous information.
2 = repeat calving.
3 = alternate calving.
numbers. A very high percentage of the dams (72%) had only one usable WW record, so that including dams in the model wouldn't be rewarding. Given these problems and the low potential of the records to furnish answers to more ambitious questions, the more limited categorization as presented above was used. This approach, the categorization being results-oriented, is criticizable and a more elaborate study, with a less limited data set, should be made.

The distribution of dams among the AOD classes is not truly representative of commercial productions in RS for the following reasons:

1) Only if cows have been raised on cultivated pasture during their first winter after weaning can they be bred at 24 to 27 months of age. The rule, on native pasture, is for Hereford heifers to reach 280-300 kg only when they are 36-39 months old. The proportion of 3 year-old dams in this study is higher than in the overall industry.

2) The decrease in numbers for ages older than 4 years does not reflect yearly mortality rates among mature cows (2-3%) neither the selection practiced (null in a practical sense), but more the fact that several breeders of PC herds join the program, in their first year, only with heifers or AI serviced dams.

3) Commercial Hereford cows are kept barren the last year of their lives, with teeth still good enough to fatten on native grass. This happens when cows are 8 to 10 years old. Only PO cows (and in a few PC herds) are cows kept "until they can still produce a viable calf."

With the exception of the 2\textsuperscript{nd} reason above, the distribution would be representative of seedstock herds.

Average number of calves in the CGs was 38.89. Two CGs had 257
calves; 4 had more than 200 calves and 18 CGs had above 100 calves. On the other extreme, 1 CG had 2 calves only. No direct selection was made to keep only large CGs. Some of the herds included are considered small herds in RS. The average size of the CGs is comparable or higher than in the US, when using as reference the Angus and Hereford data sets used for National Sire Evaluation. This comparison reflects the fact that in RS: 1) mixed farming is more the exception in RS, where ranches depend mostly on their beef and sheep (wool types) operations; and 2) to generate enough income for the owner's family, comparable to an urban professional, farm size must be at least 1,000-2,000 acres, if no grain crops are produced.

Table 5 presents several other variables with their means and dispersion measures.

Only records from sires having more than 4 calves were kept. This reduced number of sires by almost half their original total. One hundred-thirty sires remained, belonging to disconnected sets. Average half-sib family size was 64.0. Twenty-two sires had more than 100 progeny (101 to 636), with a total of 5040 calves and an average of 229.1 progeny/sire. These were all AI sires, from imported or national semen; some were from imported sires and owned by the breeders; and most are Polled Herefords. Other 107 sires, mostly used in natural service, had a total of 2402 progeny, with an average of 22.5 calves/sire.

The last so-called "sire" had 941 "progeny." Under this class, some 50 sires are represented. All of them are selected sons of an AI sire, also heavily used (270 progenies in 10 herds), Jones Rollo 0628, from American Breeders Service. Grandsire-grandson families can also be used
Table 5. Means and dispersion measures on several variables in the data set

<table>
<thead>
<tr>
<th>Variable (unit)</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum Value</th>
<th>Maximum Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dam birth year</td>
<td>8323</td>
<td>1971.1</td>
<td>2.9</td>
<td>1958</td>
<td>1976</td>
</tr>
<tr>
<td>Age of dam (year)</td>
<td>8323</td>
<td>6.3</td>
<td>2.4</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Dam weight(^a) (kg)</td>
<td>7294</td>
<td>389.7</td>
<td>58.2</td>
<td>212</td>
<td>733</td>
</tr>
<tr>
<td>Cow efficiency(^b)(%)</td>
<td>7294</td>
<td>36.3</td>
<td>6.5</td>
<td>15.9</td>
<td>63.2</td>
</tr>
<tr>
<td>Calving interval (day)</td>
<td>3492</td>
<td>529.3</td>
<td>179.3</td>
<td>273</td>
<td>1536</td>
</tr>
<tr>
<td>Fall birth date (Julian)</td>
<td>1406</td>
<td>100.7</td>
<td>38.3</td>
<td>51</td>
<td>155</td>
</tr>
<tr>
<td>Spring birth date (Julian)</td>
<td>6917</td>
<td>275.9</td>
<td>108.1</td>
<td>170</td>
<td>364</td>
</tr>
<tr>
<td>Weaning age (day)</td>
<td>8323</td>
<td>213.5</td>
<td>33.1</td>
<td>115</td>
<td>295</td>
</tr>
<tr>
<td>Real weaning weight(^c) (kg)</td>
<td>8323</td>
<td>143.4</td>
<td>29.9</td>
<td>60</td>
<td>270</td>
</tr>
<tr>
<td>Average daily gain (g/day)</td>
<td>8323</td>
<td>519.1</td>
<td>127.0</td>
<td>161</td>
<td>1025</td>
</tr>
</tbody>
</table>

\(^a\)Taken at weaning date of her calf.

\(^b\)Ratio of weaning weight of the calf, corrected to 205 days, not adjusted for AOD, divided by dam weight, and multiplied by 100.

\(^c\)After a complete fasting period of 12-14 hours.
to estimate genetic and environmental variances. This group of grand progeny is large, in two herds, helps to tie the data set, and comprises about 40% of all progeny from STATUS 2 dams. Sire editings were done last and deleting this group was considered too big a loss, even if it is a non-orthodox group.

When the ROP program began, only calves from identified sires and dams were accepted. Large PC herds, weaning 1000 or more calves, are unable to AI all cows or to provide paddocks for groups of 20-25 cows. Cow reports are regarded as a very useful feature of the program but it requires complete information on all calves of a dam. In 1977, calves from multiple sires began to be accepted and breeders know that there is a trade-off of information. Most breeders use teams of bulls grouped by their sire, to keep inbreeding in check. Some use the best two-year olds on their own herds, as multiple sires, and then sell them in next spring's auction, when they are 3-years old. The adaptation of the ROP program to these peculiarities of large PC herds, allowed for larger participation of complete PC herds, especially for those with a fast genetic turnover, and for obtaining more complete cow reports.

D. Biological Model

Several identifiable environmental factors influencing WW have been discussed in this and the previous chapter. In this section, a more concise treatment to these factors, especially to their interrelationships will be given. More emphasis is placed on clarity of what is biologically important than with statistical rigor. How these factors translate to regression conglomerates and necessary assumptions to
accompany a mathematical model will be presented in Chapter V.

Figure 1 summarizes a personal view of the factors which have an influence on WW of beef cattle in RS; at least those perceived and contained in the data set studied. Such a scheme allows for the study and quantification of the explicit relationship, being more a "useful fiction" (this is how Weisberg (1980) defines models) than a simplified representation of the real world.

Since this is an observational study, no claim can be made about finding cause-and-effect pathways. No predictions or inferences can be drawn outside the data set itself, but until experiments indicate that the findings are wrong, the results of such a study remain as a working hypothesis.

Contemporary groups (CG) is the big classificatory factor that partitions the records. Their effects, and from the interactions among their factors, are so unpredictable that precorrection is not feasible. If the range of AOC is not large, CG is the single most important factor in accounting for variation in WW. If mixed models are not used, comparisons among calves, for selection purposes, can only be made within CG. This puts a serious restrain on the range of comparisons a breeder is allowed to make.

Sire evaluation is a valuable fringe benefit from ROP programs. Its inclusion in the model is mandated for the purpose of estimating heritability. If sires form disconnected sets, each set should be analyzed separately, if under a fixed model framework. If the assumption that all sires came from a common population is tenable, the complete data set can be analyzed under a mixed model.
Figure 1. Biological model hypothesized for weaning weights of Hereford calves in RS, being tested in the present study. (Broken arrows indicate that these variables were not used in the present study)
Sire by CG (or any of its components) interaction is not indicated in Figure 1 neither is this interaction used in any model. Within herd reports from the ROP program have often shown changes of rank of sires across years and seasons. It is expected that the inclusion of cow status, seasonality and non-linear age of calf effects will reduce these changes of rank. Pure management problems exist, since breeders select calves for the various feeding regimes and cows to be bred to each sire. The reduction of CG sizes (more sampling errors), the confounding of genetic merit with different environmental levels and the inconsistencies across years in dam selections are not burden-free decisions the breeder takes. Inclusion of this interaction in the model would result in higher estimates of heritability. It would not be easy to explain, even if there are reasons to expect it to be real, given the unbalanceness of the records.

The above paragraph is not trying to play down the importance of genetic-environmental interactions. The massive imports of germplasm and the environmental distance between exporting centers and RS makes this a very important question. What is recognized above is the impotence to deal effectively with the problem. Looking at genotype stability, as described by Falconer (1981) and Bulmer (1980), possibly using response surface methods, appears to be a very promising way to look at this problem and to give understandable information to breeders.

Climatic effects influence weaning weights of calves in several ways, some directly affecting his metabolism or indirectly through his environment. Some of these effects were reviewed in Chapter III, but interactions exist also, as follows:
1) AI sires are followed by clean-up sires, as a rule. Hence, for the correct estimation of seasonal effects and of breeding values of sires, their simultaneous estimation is needed.

2) A high proportion of AI serviced dams are STATUS 2 or heifers, and they will have most of their calves born early in the season.

Age-of-dam (AOD) effects are estimated within sex of calf because in a preliminary analysis the AOD by sex interaction was significant.

Weaning weight (WW) is the measured response for all the effects presented in Figure 1. WW is a composite trait, being used both to estimate the genetic potential for growth of the calf and the maternal ability of his dam.

In dairy cattle cow evaluations, the response is milk production, with its own set of explanatory variables. In beef cattle, milk production is substituted by its proxy, WW, and the only explanatory variable remaining is AOD (parity). All environmental effects on milk production are broken down in two parts:

1) what happened before the onset of the first lactation and will affect lifetime production is designated as "permanent effects"; and

2) all changes occurring between last and actual weaning are put into "temporary effects" and they inflate residual error.

The result of this procedure is that producing ability (genetic + permanent environmental effects) is estimated for dams from the series of WWS. True transmitting abilities (genetic) for WW can not be estimated unbiasedly, since known fixed effects were not accounted for.

Ideally, the structural system represented in Figure 1 could be solved using simultaneous-equation methods. At least a pair of equations
should be written: one for the weaning weight of the calf and other for
the milk production of the dam. This group of equations has correlated
errors and correlated random effects, and could be solved using Zellner
(1962) methods or its extension, multiple trait methods, as described by
Henderson and Quaas (1976).

As milk production is not measured, the whole equation is removed.
Weaning weight records are adjusted for AOD and the records are
considered a trait of the calf. Including cow status and AOD in the
equation for the calf is a small step in the direction of removing these
biases. An animal model, including dam in the equation is a large step,
recently examined by Wilson (1984b), after being proposed to the industry
for a long period. Adding to this animal model AOD and STATUS within
herd and dam CG (given by birth year and season, and management group
until first breeding) would improve the estimators.

In this study, given the few repeated records per cow and almost
complete lack of information on how each cow was raised, only AOD and
STATUS are included in the model.
V. METHODS

A submodel containing the environmental variables described in Figure 1 the last chapter was used in conjunction with 7 different sets of assumptions. The models will be described later and are cited here as follows:

1. **WITHOUT SIRES** (incomplete models)
   a) Ordinary least squares (OLS)
   b) Weighted least squares (WLS)
   c) Robust Weighted least squares (ROBWLS)

2. **WITH SIRES FIXED**
   a) OLS
   b) WLS
   c) ROBWLS

3. **WITH SIRES RANDOM**
   a) Generalized least squares (GLS)

Variance components (sire and error) were estimated using the Restricted Maximum Likelihood Method (REML) and convergence of the iterative solutions were obtained using the Common Intercept Approach (CIA), as in Schaeffer (1983).

The data were analyzed after being reduced to the Normal Equations (NE) form, using SAS (1982a,b) package or row operations were done directly on the data matrix, using Givens (1954) rotations. This same process was used in the iterative process (REML-CIA) to find converged values for the variance component estimates.
A. Parameter Estimation

In the mathematical world, variables (y and x) relate to each other in an exact (within machine accuracy bounds) form as in the slope (b) form:

\[ y = f(x) = bx; \]

or in intercept-slope form:

\[ y = f(x) = a + bx. \]

In the natural or physical world, this perfect relationship is hardly ever found. Hence, only approximations \((\hat{y})\) are possible, like \(y \approx f(x) \approx a + bx\).

Under this form, persons are free to choose the parameters \((a,b)\). This is not a comfortable situation since one has to decide how good the different sets of parameters proposed are. A heuristic idea is that "good" parameters will give a "good approximation." Using some proposed parameters, one can calculate the values of the expected response or fitted values \((\hat{y})\) under the chosen parameters. Contrasting this expected \((\hat{y})\) versus actual \((y)\) response gives a measure of how good the approximation is for each of the data points \((y,x)\).

One may want \((y - \hat{y})\) to be small and to describe the relationship really well for certain ranges or domains of the variables. Others may be interested in having an equally good approximation over all the values of \(y\) and \(x\). If this is the case, one can just sum all the observed residuals \((\hat{r} = y - \hat{y})\). Since different sets of parameters can yield the same sum value due to cancellations of positive and negative values of \(\hat{r}\),
a better approach is to sum the absolute value of the residuals. This is called the $l_1$ norm, being a linear ($l$) form (a simple sum) of the absolute value of the residuals raised to the power 1. This norm has all possible appeals but it is not easily tractable in mathematical formulations.

The dominant kind of norm is the $l_2$. The measure of how good the approximations are is given by the sum of the squares of the residuals. The leverage of each residual over this measure of approximation and, in the last instance, over the set of parameters to be chosen, is the value of each residual squared. This is the origin of the lack of resistance of the $l_2$-norm fitting to large residuals. A single one of these residuals, if it is large enough and being squared, can force the parameters to approximate the function less well over most of its domain so that the squared value of this influential residual can be reduced.

A major plus for $l_2$-norm fitting is that it is much easier to develop mathematical formulations and statistical properties than the $l_1$-norm. Working under the $l_2$-norm allows one to use this norm itself as the criteria to be minimized in obtaining numerical solutions for general models. Also, the $l_2$-norm relates directly to second moments (variances and covariances) and to the decomposition of total variance, so dear to animal breeders.

In the framework of the $l_1$-norm, a measure of variability exist, the Median Absolute Deviation (MAD), that is more resistant to gross-errors or outliers, and has a higher Asymptotic Relative Efficiency (ARE) when compared with the square root of the variance and when the distribution of the observations is not normal.
Other norms are being investigated, especially relating to the range of $l_1$ to $l_2$ (residuals are raised to powers between 1 and 2).

This discussion intended to bring up the point that the $l_2$-norm or the principle of least squares is not the only way to investigate relationships among variables although it is the most common; and also that it may not always be the best to work with, and this fact gives rise to adaptations to the $l_2$-norm, like some robust estimators.

1. **Ordinary least squares (OLS)**

Under this method, estimates of the parameters are chosen such that the sum-of-squares of the residuals ($l_2$-norm) is minimized. The method is a pure mathematical formulation, does not depend on any assumption concerning the residuals and will yield least squares estimates of the parameters even if the model is completely inappropriate to describe a real situation.

Development of the method, for simple regression or in general matrix form, can be found in several texts, using some calculus. Guttman (1982) presents these, a geometric interpretation for linear models, and a simple algebraic development of the least squares method, not requiring any calculus.

The explanatory variables are always assumed to be known without error. Adding the assumption that the residuals are independent from each other and come from the same population with mean zero and a common variance, give least squares estimates (linear functions of the response variable) that have the smallest possible variance. This is an informal translation of the Gauss-Markov theorem.
Adding the assumption that the residuals are normally distributed, the estimates will also be normally distributed, normal-theory tests of significance can be performed, and the least squares estimators are the same as the maximum likelihood ones.

**a. Dummy variables** The least squares procedure is not restricted to continuous explanatory variables but can also handle classificatory ones. Dummy variables (with values of 1 or 0, normally) are used to indicate in which group a given observation belongs. Sets of dummy variables, possibly for each set of classes or levels in a factor, can be contained and manipulated more easily in the form of incidence or data matrices.

Classificatory factors used in a model create rank dependencies in the column space of the data matrix or in the normal equations. No unique solutions are possible and constraints on the solutions or generalized inverses have to be used. Or, the model can be reparameterized to full rank.

This approach was taken in the present study with respect to estimates of STATUS effects. The reparameterization used is equivalent to applying the restriction that the sum of STATUS effects equal zero.

Consider

\[ y_{ij} = \mu + \alpha_i + e_{ij} \text{ where } i = 1, 2, 3 \text{ and } j = 1, \ldots, n_i \]

and adjoin
\[ \frac{3}{2} \sum_{i=1}^{\infty} a_i = 0 = a_1 + a_2 + a_3, \text{ or} \]

\[ a_3 = -a_1 - a_2. \]

Or in matrix notation (with \( n_i = 1 \)):

\[
y = x\alpha + e = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} + e.
\]

We now want a matrix \( C \) such that \( CA = y \)

\[
C\alpha = C \begin{bmatrix} \mu \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = y = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ 0 \end{bmatrix}.
\]

Such a \( C \) is

\[
C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}
\]

and its inverse is

\[
C^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{bmatrix}
\]

such that \( C^{-1}C = I \). We can use a special form of \( XI = X \):

\[
y = x\alpha + e = x_1\alpha + e = xC^{-1}\alpha + e = xC^{-1}y + e = Wy + e
\]
and $W = XC^{-1} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 \end{bmatrix}$.

Since we are imposing the condition that $\gamma_3 = 0 = \alpha_1 + \alpha_2 + \alpha_3$, the last column in $W$ can be deleted.

Dummy variables $ST1$ and $ST2$ were created, under this reparameterization, and used in the analysis according to the following rule:

If STATUS = 1 then ST1 = 1 and ST2 = 0.
If STATUS = 2 then ST1 = 0 and ST2 = 1.
If STATUS = 3 then ST1 = -1 and ST2 = -1.

$ST1$ and $ST2$ correspond to the 2nd and 3rd columns of $W$ above. The parameters they are set to estimate are $\alpha_1$ and $\alpha_2$, corresponding to the restriction $\alpha_1 + \alpha_2 + \alpha_3 = 0$.

Since $\alpha_3 = - \alpha_1 - \alpha_2$, and $\alpha_1$ and $\alpha_2$ are estimable, then $\alpha_3 = - \alpha_1 - \alpha_2$.

Notice that contrasts among parameters are always estimable if observations exist. This means that contrasts are independent of the set of solutions obtained under a given restriction, by a given generalized inverse or by a specific reparameterization.

Possible contrasts, and their estimable function, are as follows:

STATUS 1 - STATUS 3 = $\alpha_1 - \alpha_3 = \alpha_1 - \alpha_1 - \alpha_3 = \alpha_1 - \alpha_1 - \alpha_2 = \alpha_1 + \alpha_2$
STATUS 2 - STATUS 3 = $\alpha_2 - \alpha_3 = \alpha_1 - \alpha_3 = \alpha_2 - \alpha_1 - \alpha_2 = \alpha_1 + \alpha_2$
STATUS 1 - STATUS 2 = $\alpha_1 - \alpha_2 = \alpha_1 - \alpha_1 - \alpha_2$
Dummy variables were created for the estimation of contemporary group (CG) effects. CG contain main effects of herd, year, season, feed code (management) and sex, and all possible interactions. As later used, CG may contain a constant $\mu$. A numerical code for CG was created using the following rule:

$$CG = \text{SEX} \times (1 \text{ or } 2) + 10 \times \text{FEED CODE} \times (1 \text{ to } 9) + 100 \times \text{SEASON} \times (1 \text{ for Fall and } 2 \text{ for Spring}) + 1000 \times \text{YEAR} \times (4 \text{ to } 9, \text{ corresponding to } '74 \text{ to } '79) + 10000 \times \text{HERD} \times (1 \text{ to } 23).$$

Another place where dummy variables could play a role is in the estimation of seasonal effects or month-of-birth effects on WW, as most of the work done in this area have used this classificatory approach.

In this study, the approach to estimate seasonality effects was to use grafted polynomials or spline functions. The dummy variable approach was not used because corrected records from animals born in the middle of the month would probably contain very little of the effects; but comparisons made between animals when one is born at the end of a given month and the other is born in the beginning of the subsequent month will contain a bigger error than if "corrections" were not made.

b. Ordinary polynomials In the regression framework, polynomials in the explanatory variables have been heavily used to approximate the response function, when this is believed to be continuous and curvilinear. Theoretically, any such function can be approximated to
any degree of accuracy by a polynomial in the explanatory variable. In the limit, a p-th degree polynomial should perfectly describe the relationship between two variables, based on a sample of n observations, given p ≤ n. Of course, considerations about computational costs, the principle of parsimony, and the need to biologically explain any found relationship, work to keep p to a minimum or to a compromise level.

Ordinary polynomials can give rise to several problems for the data analyst, some may be easy to remedy, others may go undetected, as follows:

1) For positive real numbers, the correlation between x and x^2 is above .90 (higher with small numbers). This creates problems of multicollinearity and increases sampling variances of the estimators. If several higher power terms are added, the problem may become intractable. Worst, due to rounding error a solution to the problem can be found, but any solution is possible.

2) As Rice (1969) demonstrated, "ordinary polynomials are inadequate in many situations. This is particularly the case when one approximates functions which arise from the physical world rather than from the mathematical world. Functions which express physical relationships are frequently of a disjoint or dissociated nature. That is to say that their behavior in one region may be totally unrelated to their behavior in another region. Polynomials, as with most other mathematical functions, have just the opposite property. Namely their behavior in any small region determines their behavior everywhere." A
potential problem area in this work is with estimating age-of-dam (AOD) effects with a polynomial with AOD to the first and second powers. One can expect big differences among young age effects, but after maturity one expects a less steep decreasing function. As there is a bulk of dams in younger ages and even with the "balancing" power of the linear term, a very symmetric parabola-like curve could be estimated.

3) The measure of how good the approximation given by the polynomial normally used is the reduction in the total sum of squares of all observations due to this regressions. The level of approximation for the entire curve will be constant only if the explanatory variable is equally distributed over the entire range. If the data points are concentrated in a given section, the curve will be very well determined in that portion; sections with few observations will have shapes almost completely determined by what the relation in the cluster of data points is.

The plan in this study was to estimate AOD, seasonality (within fall and spring) and age of calf (AOC) effects, each by a continuous function. The problems described above can have a negative influence on the solutions, either when considering all regressors together or when considering the distributions of each of the explanatory variables.

c. **Grafted polynomials** Related terms to grafted polynomials are: piecewise polynomials, partition regression, spline functions, piecewise regression, segmented polynomials, switching regressions,
segmented curves, and other combinations.

Some applied texts which present the subject are Bliss (1970), Seber (1977), and Draper and Smith (1981). Several applications for agriculture and economics along with theoretical development are given by Fuller (1969). The treatment given in this section is based largely on Fuller (1977, 1982).

Rice (1969) defines spline functions as "piecewise polynomials of degree n that are connected together (at points called knots) so as to have n-1 continuous derivatives." This same author states that "spline functions are the most successful approximating functions for practical applications so far discovered," and comparing them with ordinary polynomials, "splines do not suffer from those handicaps" (defined in last sub-section) "since they are defined piecewise and yet, they represent nice, smooth curves in the physical world."

Poirier (1973) citing I. J. Schoenberg: "Polynomials are wonderful even after they are cut into pieces, but the cutting must be done with care. One way of doing the cutting leads to so-called spline-functions."

Fuller (1969) recommends that "in a practical situation, the choice of functional form will rest upon theoretical considerations, ease of estimation and acceptance by the data." In the same paper, Fuller listed some desirable properties of an approximating function:

1) to be continuous everywhere;
2) to possess continuous first derivatives and, if the degree of the segmented polynomial is p, to have continuous (p-1)-th derivatives;
3) to be easy to estimate (linear in the parameters); and
4) to permit easy computation of some optimality criteria.

The major obstacle for the use of grafted polynomials is that the knots or joining points are required to be known. Prior knowledge on the subject matter, use of plots and means for gross estimation or use of a non-linear model for the simultaneous estimation of the knots and the regression coefficients can be employed in this task. Before going to the description of how these points were chosen, these joining points are assumed to be known and the regression systems are presented, for AOD, seasonality within fall, and seasonality within spring.

The shape of the curve of AOD effects on WW is well known, being given by a linear plus a quadratic component. Say that it is suspected that on each side of the maximum point (6 years) a different system (linear plus quadratic terms) should be fit. Keeping the equations as simple and informal as possible, what is wanted is as follows:

The equation can be written as

\[ y \leq a_1 + b_{11}x + b_{12}x^2 \quad \text{for} \quad x \leq 6; \quad \text{and} \quad (1) \]
\[ y \geq a_2 + b_{21}x + b_{22}x^2 \quad \text{for} \quad x \geq 6. \quad (2) \]

There are 6 parameters to be estimated, but the following restrictions have to be imposed:

1) a continuous mean function:

\[ a_1 + b_{11}x + b_{12}x^2 = a_2 + b_{21}x + b_{22}x^2 \quad \text{at} \quad x = 6; \quad (3) \]
2) a continuous first derivative:

\[ b_{11} + 2b_{12}x = b_{21} + 2b_{22}x \quad \text{at } x = 6. \quad (4) \]

After the two restrictions, four parameters are to be estimated and \( a_1, b_{11}, b_{12} \) and \( c = b_{22} - b_{12} \) are chosen. Using the relationships given in the restrictions, the 6 original parameters and the 2 original equations have to be rewritten to contain only the 4 chosen parameters.

Substituting \( x = 6 \) in the 1\(^{\text{st}}\) restriction (Eq. 3) and solving for \( a_2 \):

\[ a_2 = a_1 + b_{11}6 + b_{12}6^2 - b_{21}6 - b_{22}6^2. \quad (5) \]

Substituting \( x = 6 \) in the 2\(^{\text{nd}}\) restriction (Eq. 4) and solving for \( b_{21} \):

\[ b_{21} = b_{11} + b_{12}6^2 - b_{22}6^2. \quad (6) \]

Substituting in (5) \( b_{21} \) for its value in (6)

\[ a_2 = a_1 + b_{11}6 + b_{12}6^2 - b_{22}6^2 \]
\[ - b_{11}6 - b_{12}6^2 + b_{22}6^2 \]
\[ a_2 = a_1 - b_{12}6^2 + b_{22}6^2. \quad (7) \]

In Eq. (2), values of \( a_2 \) as in (7) and of \( b_{21} \) as in (6) are substituted:
Rewriting Eq. (1):

\[ y = a_1 + b_{11}x + b_{12}x^2 \text{ ; for } x \leq 6. \]  

(1)

Eq. (1) and (9) substitute the original system (Eq. (1) and (2)) and Eq. (9) is similar in parameters to Eq. (1) but for the addition of the term \((b_{22} - b_{12})(x - 6)^2\).

This suggests a way to put all data together by creating a third explanatory variable.

Say that:

\[ \text{AOD}_1 = x = \text{AOD in years ; } \]

\[ \text{AOD}_2 = x^2 = \text{AOD}^2 ; \text{ then} \]

\[ \text{AOD}_3 = (x - 6)^2 = (\text{AOD} - 6)^2, \text{ if AOD > 6 } \]

\[ = 0, \text{ otherwise.} \]

What the regression coefficients will estimate are:

\[ b_{11}, \text{ the linear effect of AOD, over the whole range of AOD;} \]
65

\( b_{12} \), the quadratic effect of AOD, over the whole range of AOD;

\( (b_{22} - b_{12}) \), the difference in the quadratic effects of AOD,

between dams which are older than 6 years

from the ones which are 6 years old or younger.

Because of the significant interaction found between AOD and sex, in
a preliminary analysis, when AOD effects were being estimated in a clas-
sificatory mode, regression coefficients for AOD effects were estimated
within each sex. The rule to create the explanatory variables, to yield
6 regression coefficients, was:

If sex = 1 (male), then:

\[
\begin{bmatrix}
AOD11 \\
AOD21 \\
AOD31
\end{bmatrix} = \begin{bmatrix}
AOD1 \\
AOD2 \\
AOD3
\end{bmatrix} ; \text{ else: } \begin{bmatrix}
AOD11 \\
AOD21 \\
AOD31
\end{bmatrix} = 0 .
\]

If sex = 2 (female), then:

\[
\begin{bmatrix}
AOD12 \\
AOD22 \\
AOD32
\end{bmatrix} = \begin{bmatrix}
AOD1 \\
AOD2 \\
AOD3
\end{bmatrix} ; \text{ else: } \begin{bmatrix}
AOD12 \\
AOD22 \\
AOD32
\end{bmatrix} = 0 .
\]

Another approach that can be taken, given by Fuller (1982), which is
more elegant and general, using matrices and reparameterization ideas, is
as follows:

Given that the knot, point A along the x-values, is known, it is
desired to fit:
for \( x \leq A \): \( y = b_1 + b_2x + b_3x^2 \);  

for \( x \geq A \): \( y = b_4 + b_5x + b_6x^2 \); and

putting this together:

\[
\begin{bmatrix}
1 & x & x^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & x & x^2 \\
0 & 0 & 0 & 1 & x & x^2 \\
0 & 0 & 0 & 1 & x & x^2 \\
\end{bmatrix}
\begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
b_4 \\
b_5 \\
b_6 \\
\end{bmatrix}
= \begin{bmatrix}
x \\
x \\
x \\
x \\
\end{bmatrix}
\]

So far, nothing is forcing the two "pieces" of the polynomial to be strictly continuous at \( A \). Restrictions on the mean function and on the first derivative are added, at \( A \). These restrictions are interpreted as a Lagrange objective function, and equated to zero, so that later corresponding columns can be deleted from the transformed data matrix.

The restrictions are:

\[
b_1 + b_2A + b_3A^3 = b_4 + b_5A + b_6A^2 \quad \text{, or}
\]

\[
b_1 + b_2A + b_3A^3 - b_4 - b_5A - b_6A^2 = 0
\]

(12)
\[ 0 + b_2 + 2Ab_3 = 0 + b_5 + 2Ab_6, \] or
\[ 0 + b_2 + 2Ab_3 - 0 - b_5 - 2Ab_6 = 0. \] (13)

With these 2 restrictions, 4 parameters are to be estimated. The reparameterization is as follows:

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 \\
1 & A & A^2 - 1 & -A & -A^2 \\
0 & 1 & 2A & 0 & -1 & -2A
\end{bmatrix}
\begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
b_4 \\
b_5 \\
b_6
\end{bmatrix}
= 
\begin{bmatrix}
\gamma_1 \\
\gamma_2 \\
\gamma_3 \\
\gamma_4 \\
\gamma_5 \\
\gamma_6
\end{bmatrix}
= 
\begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
b_4 - b_3 \\
0 \\
0
\end{bmatrix}.
\]

Notice that \( \gamma \) above, in its "b" terms, corresponds to (9) and (1).

Again, in
\[ y = \chi \theta = \chi \theta \] with \( Z = \chi \theta \), where
\[
Z = \begin{bmatrix}
1 & x & x^2 & 0 & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
1 & x & x^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & x & x^2 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & 1 & x & x^2
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & A^2 - 1 & A \\
0 & 1 & 0 & -2A & 0 & -1 \\
0 & 0 & 1 & 1 & 0 & 0
\end{bmatrix}
\]
Since the two restrictions we are imposing ($y_5$ and $y_6$) are both equal to zero, corresponding columns in $Z$ (5-th and 6-th) can be deleted. Columns 2, 3, and 4 in $Z$ correspond exactly to the system given in (10), using a long-hand algebraic derivation.

Seasonality effects for fall born calves, or the effect of date of birth for animals born between Julian days 50 to 160, were also estimated using a grafted polynomial.

The display of residuals showed that a quadratic-quadratic curve would be a good approximation. The joint point used was Julian day 85 and, without repeating a demonstration since the system of equations is exactly the same as for AOD, the rule to create the explanatory variables is as follows:

If BIRTHDATE (Julian) is not between days 50 and 160, then:

$$
\begin{bmatrix}
1 & x & x^2 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & x & x^2 & (x-A)^2 & -1 & (A-x) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & x & x^2 & (x-A)^2 & -1 & (A-x) \\
\end{bmatrix} = 0.
$$

If BIRTHDATE (Julian) is between days 50 and 160, then:

$$
\begin{bmatrix}
SF1 \\
SF2 \\
SF3
\end{bmatrix} = 0.
$$

(14)
$\text{SF}_1 = \text{BIRTHDATE (Julian)}$;

$\text{SF}_2 = (\text{SF}_1)^2$;

$\text{SF}_3 = (\text{SF}_1 - 85)^2$, if $\text{SF}_1 \geq 85$;

$= 0$, otherwise.

What the regression coefficients, corresponding to these explanatory variables, will estimate are:

- $b_{\text{SF}_1}$, the linear effect of date of birth, within fall season;
- $b_{\text{SF}_2}$, the quadratic effect of date of birth, within fall season;
- $b_{\text{SF}_3}$, the difference in the quadratic effects of date of birth, between animals born before or after Julian date 85 (March 26).

Notice, in Table 2, that only 25% of fall born calves were born before March 26.

For animals born in the "spring," Julian days 170 to 365, the data indicated that a more complex system would be needed. The system is a cubic-quadratic-cubic. In its simplest form, the system can be described as follows:

$$y_{1i} = a_1 + b_1 x_i + c_1 x_i^2 + d_1 x_i^3 + e_{1i} \quad \text{if } x_i \leq 245; \quad (15)$$

$$y_{2i} = a_2 + b_2 x_i + c_2 x_i^2 + e_{2i} \quad \text{if } 245 \leq x_i \leq 315; \quad (16)$$

$$y_{3i} = a_3 + b_3 x_i + c_3 x_i^2 + d_3 x_i^3 + e_{3i} \quad \text{if } x_i \geq 315; \quad (17)$$

This system involves 11 parameters and 3 sets of restrictions are
imposed:

(1) continuous mean functions

\[ a_1 + b_1 x_1 + c_1 x_1^2 + d_1 x_1^3 = a_2 + b_2 x_1 + c_2 x_1^2 \text{ at } x_1 = 245; \quad (18) \]
\[ a_2 + b_2 x_1 + c_2 x_1^2 = a_3 + b_3 x_1 + c_3 x_1^2 + d_3 x_1^3 \text{ at } x_1 = 315; \quad (19) \]

2) continuous first derivatives

\[ b_1 + 2c_1 x_1 + 3d_1 x_1^2 = b_2 + 2c_2 x_1 \text{ at } x_1 = 245; \quad (20) \]
\[ b_2 + 2c_2 x_1 = b_3 + 2c_3 x_1 + 3d_3 x_1^2 \text{ at } x_1 = 315; \quad (21) \]

3) continuous second derivatives

\[ 2c_1 + 6d_1 x_1 = 2c_2 \text{ at } x_1 = 245; \quad (22) \]
\[ 2c_2 = 2c_3 + 6d_3 x_1 \text{ at } x_1 = 315. \quad (23) \]

These restrictions would allow the third derivatives to "jump" at the knots, satisfying one of the conditions for cubic splines given by Poirier (1973).

Using these 6 restrictions, we can solve the system for 5 unknown parameters: \( a_2, b_2, c_2, d_1, \) and \( d_3 \). The strategy to be followed is to obtain from the restrictions values for \( a_1, b_1, \) and \( c_1 \) in terms of the 5 parameters and substitute them in Eq. (15) and for \( a_3, b_3, \) and \( c_3 \) in terms of the 5 parameters and substitute them in Eq. (17).

From Eq. (22):

\[ c_1 = c_2 - 3(245)d_1. \quad (24) \]
From Eq. (20):

\[ b_1 = b_2 + 2(245) c_2 - 2(245) c_1 - 3(245)^2 d_1 \]

and, substituting \( c_1 \) from Eq. (24):

\[ b_1 = b_2 + 2(245) c_2 - 2(245)[c_2 - 3(245) d_1] - 3(245)^2 d_1 \]
\[ b_1 = b_2 + 2(245) c_2 + 6(245)^2 d_1 + \]
\[ - 2(245) c_2 - 3(245)^2 d_1 \]
\[ b_1 = b_2 + 0 + 3(245)^2 d_1 \] \quad (25)

From Eq. (18):

\[ a_1 = a_2 + (245)b_2 + (245)^2 c_2 - (245)b_1 - (245)^2 c_1 - (245)^3 d_1 \]

and, substituting \( c_1 \) from Eq. (24) and \( b_1 \) from Eq. (25):

\[ a_1 = a_2 + (245)b_2 + (245)^2 c_2 - (245)^3 d_1 + \]
\[ - 245 \left[ b_2 + 3(245)^2 d_1 \right] + \]
\[ - (245)^2 \left[ c_2 - 3(245) d_1 \right] \]
\[ a_1 = a_2 + (245)b_2 + (245)^2 c_2 - (245)^3 d_1 + \]
\[ - (245)b_2 + 3(245)^3 d_1 + \]
\[ - (245)^2 c_2 - 3(245)^3 d_1 \]
\[ a_1 = a_2 + 0 + 0 - (245)^3 d_1 \] \quad (26)

Substituting values of \( a_1 \) from Eq. (26), \( b_1 \) from Eq. (25), and \( c_1 \) from Eq. (24) in Eq. (15):
\[ y_{11} = \left[ a_2 - (245)^3 d_1 \right] + \left[ b_2 + 3(245)^2 d_1 \right] x_i + \\
\left[ c_2 - 3(245)d_1 \right] x_i^2 + d_1 x_i^3 + e_{11} + \\
= a_2 + b_2 x_i + c_2 x_i^2 + d_1 x_i^3 - 3(245)^3 d_1 + e_{11} + \\
y_{11} = a_2 + b_2 x_i + c_2 x_i^2 + d_1 (x_i - 245)^3 + e_{11}; \text{ if } x_i \leq 245. \quad (27) \]

Notice that Eq. (27) and (16) differ only in the coefficient for \( d_1 \).

Focus now is turned to Eq. (17) and its parameters. From Eq. (23):

\[ c_3 = c_2 - 3(315)d_3. \quad (28) \]

From Eq. (21):

\[ b_3 = b_2 + 2(315)c_2 - 2(315)c_3 - 3(315)^2 d_3 \]

and, substituting \( c_3 \) from Eq. (28)

\[ b_3 = b_2 + 2(315)c_2 - 3(315)^2 d_3 + \\
-2(315) \left[ c_2 - 3(315)d_3 \right] \\
= b_2 + 2(315)c_2 - 3(315)^2 d_3 + \\
-2(315)c_2 + 6(315)^2 d_3 \\
\]

\[ b_3 = b_2 + 0 + 3(315)^2 d_3. \quad (29) \]

From Eq. (19)

\[ a_3 = a_2 + (315)b_2 + (315)^2 c_2 - (315)b_3 - (315)^2 c_3 - (315)^3 d_3 \]
and, substituting $c_3$ from Eq. (28) and $b_3$ from Eq. (29)

$$a_3 = a_2 + (315)b_2 + (315)^2c_2 - (315)^3d_3 +$$
$$-315[b_2 + 3(315)^2d_3] +$$
$$-315^2[c_2 - 3(315)d_3]$$

$$a_3 = a_2 + (315)b_2 + (315)^2c_2 - (315)^3d_3 +$$
$$-(315)b_2 - 3(315)^3d_3 +$$
$$-(315)^2c_2 + 3(315)^3d_3$$

$$a_3 = a_2 + 0 + 0 - (315)^3d_3 . \quad (30)$$

Now, substituting values of $a_3$ from Eq. (30), $b_3$ from Eq. (29), and $c_3$ from Eq. (28), in Eq. (17):

$$y_{3i} = [a_2 - (315)^3d_3] + [b_2 + 3(315)^2d_3]x_i +$$
$$[c_2 - 3(315)d_3]x_i^2 + d_3x_i^3 + e_{3i}$$

$$= a_2 + b_2x_i + c_2x_i^2 +$$
$$d_3x_i^3 - d_3(3)(315)x_i^2 + d_3(3)(315)^2x_i - d_3(315)^3 + e_{3i}$$

$$= a_2 + b_2x_i + c_2x_i^2 +$$
$$d_3[x_i^3 - 3(315)x_i^2 + 3(315)^2x_i - (315)^3] + e_{3i}$$

$$y_{3i} = a_2 + b_2x_i + c_2x_i^2 + d_3(x_i - 315)^3 + e_{3i} \text{ if } x_i \geq 315 . \quad (31)$$

At this point, the system given in Eq. (15), (16), and (17), for the estimation of seasonality within spring effects can be rewritten, substituting Eq. (15) by (27) and Eq. (17) by (31), as follows:

$$y_{1i} = a_2 + b_2x_i + c_2x_i^2 + d_1(x - 245)^3 + 0 + e_{1i} ; \text{ if } x_i \leq 245 ; \quad (27)$$
If the assumptions that \( V(e_1, e_1') = V(e_2, e_2') = V(e_3, e_3') \), where \( V \) indicates variance-covariance matrix, are expected to hold true, then the above system can be rewritten in one single equation, as follows:

\[
y = a_2 + b_2 x_1 + c_2 x_1^2 + d_3 (x_1 - 315)^3 + e_{2i} + e_{3i} + e_{4i}; \quad \text{if } 245 \leq x_1 \leq 315;
\]

\[
y = a_2 + b_2 x_1 + c_2 x_1^2 + d_3 (x_1 - 345)^3 + e_{3i}; \quad \text{if } x_1 \geq 315;
\]

A complete algebraic development to find these transformed explanatory variables, to be used with the grafted polynomials, was presented instead of just using the very useful rules, in a general form, given by Fuller (1969). It is the opinion of the author that those rules are clear only for the experienced (with spline functions) user and that relying only on them here would be a disservice for the one planning to
use grafted polynomials in a related problem and it would reduce chances of errors being checked.

For another environmental source of variation, age of calf (AOC), the use of segmented polynomials was tried. A simpler form, a polynomial with linear and quadratic terms, appeared to explain well enough the relationship between WW and AOC. A cubic-quadratic-cubic system didn't prove to be significantly better than the ordinary polynomial in the first two powers of AOC. Explanatory variables to estimate AOC effects were created according to the following rule:

If AOC is between 115 and 295, then:

\[
\begin{align*}
CA_1 &= AOC \\
CA_2 &= (AOC)^2
\end{align*}
\]  \hspace{1cm} (33)

So far in this section V.A.1., all environmental variables used in this study have been presented. These variables or factors have also been shown in Figure 1, at the end of Chapter IV. Together, they build up the first model cited at the beginning of this chapter: an incomplete model (without sires), with very unrealistic assumptions, an OLS model. Errors are assumed to be independent (while there are big half-sib families), and homogeneous.

From Figure 1, leaving SIRES:SET out, the environmental factors can be written in terms of an approximate equation as follows:

\[
WW \approx CG + STATUS + AOD:SEX + BIRTHDATE:SEASON + AOC
\]
These factors can be broken down in their components, according to
the coding used in Section V.A.1.a, and in Eqs. (11), (14), (32), and (33).

\[ \begin{align*}
W &= CG + \begin{bmatrix}
ST1 \\
ST2
\end{bmatrix} + \begin{bmatrix}
AOD11 \\
AOD12 \\
AOD21 \\
AOD22 \\
AOD31 \\
AOD32
\end{bmatrix} + \begin{bmatrix}
SF1 \\
SF2 \\
SF3
\end{bmatrix} + \begin{bmatrix}
SS1 \\
SS2 \\
SS3 \\
SS4
\end{bmatrix} + \begin{bmatrix}
CA1 \\
CA2
\end{bmatrix} \cdot (34)
\end{align*} \]

A problem which has been deferred until now is the one of finding
the knot values, and this is the subject of the next section. At the end
of that section a formal presentation of the model is given.

d. Estimation of the jointing points of the grafted polynomials

Grafted polynomials have several advantages for their use as
approximating functions, as was discussed in the last section. The main
disadvantage is that the jointing points have to be known. If they are
not known and it is desired to estimate these points from the data,
simultaneously with the regression coefficients, in a multiple regression
framework, the resulting model will be non-linear in the unknown
parameters.

For example, in the system to estimate seasonality within spring
given by Eq. (27), (16), and (31), terms like: \( d_1(x_1 - 245)^3 \) and
\( d_3(x_1 - 315)^3 \) appear. If those points are to be estimated simultaneously
with the coefficients, then these terms should be written like:
\( d_1(x_1 - \alpha_1)^3 \) and \( d_3(x_1 - \alpha_3)^3 \). These pairs of unknowns, \((d_1, \alpha_1)\) and
\((d_3, \alpha_3)\) are in a non-linear form and preclude the use of standard
multiple regression estimation routines.
This problem has created a major concern among statisticians and a short and summarized review is presented, in what follows, of the methods suggested to obtain a solution.

For the estimation of the joint points of a grafted polynomial, Quandt (1958), Robinson (1964), and Hinkley (1971) have proposed maximum likelihood (ML) techniques; Hudson (1966) proposed a steepest descent of the sum of all segments' residual SS method; Bellman and Roth (1966) used segmented differential approximation; McGee and Carleton (1970), Guthery (1974), and Ertel and Fowlkes (1976) used different clustering methods; and Halpern (1973) used a Bayesian approach, assuming that the location of all possible knots were known, and assigning prior probabilities to each point within a subset of each possible location.

Gallant and Fuller (1973) proposed a generalized version of Hartley's modified Gauss-Newton method for solving the (non-linear) problem of finding the estimates of the coefficients on each section of the curve and of the joint points that would minimize residual SS. If the curve is in the mode quadratic-quadratic, the solutions are guaranteed to converge. To find the initial values for the iterative procedure, the authors above recommend plotting the data and obtain the initial values for the joint points by inspection. Using these joint points, the other initial values for the parameters are obtained by standard multiple linear regression methods.

This procedure of Gallant and Fuller (1973) can be used directly in PROC NLIN of SAS (1982b). It requires the statement of the model, initial values and a vector of partial derivatives of the model with respect of each of the parameters.
Lerman (1980) suggested a grid search in the transition points combined with the Marquardt technique for obtaining estimates in segmented regression models. This technique is highly recommended if the parameters are correlated. The author presents three examples of the method, one of them relating voluntary feed intake in ruminants to energy concentration in the diet in a curve with three segments. This method can also be handled by PROC NLIN of SAS (1982b).

An alternative to the problem of unknown join points is given by Poirier (1973), who has integrated ideas of piecewise regression and of spline theory and has pondered that nowhere does there occur an instantaneous structural change in the relation among variables and that the change can be viewed as occurring in a smooth fashion. Poirier (1973) states that, in a cubic spline regression model: "the parameters being estimated are actual ordinate values (not slopes) of the true regression functions for particular abscissa values, namely the knots. In simple regression only the constant term has this interpretation."

This side issue is equivalent to what is espoused by some in linear model theory that the parameters being estimated are $\gamma$ and not $\beta$.

The first step in finding the initial values for the knots was "to inspect" the data, following Gallant and Fuller (1973), to be later used in PROC NLIN.

Weaning weight ratios (WWR) are also recorded on the tape containing ROP data. These ratios are obtained after correcting the WW records according to BIF (1972) recommendations, dividing them by the CG mean, and multiplying this ratio by 100. If BIF (1972) recommendations are valid for RS data (no suggestion like that is made in the publication),
then plots of WW ratios against all environmental variables should show no pattern of dependency. Plots were not even tried, given the number of observations. Means of WW ratios were calculated along the range of all studied continuous environmental variables and they should deviate very little from 100 if the corrections were efficient.

This method proved to be a very sensible and economical way to inspect the data.

Average WWR according to age of dam showed that very young AOD were underadjusted and that 6 year-old dams had highest WWR. Given the large number of dams at classes 3, 4, 5, and 6 year-old, this point (6 years) was chosen as a knot, to allow for an independent estimation of the curvature for older ages from the curvature at younger ages, according to the ponderings of Rice (1969). If this segmentation of the curve is not needed, or the chosen knot is not a correct one, then the regression coefficients for AOD3 (or for AOD31 and AOD32) will be non-significant.

Average WWR also gave some indications on how to form STATUS and BYSIDE groups. A model containing this and all other environmental variables were analyzed using PROC GLM of SAS (1982b) and showed non-significant differences among several of the STATUS * BYSIDE classes. The less ambitious partition, as given in V.A.l.a, was used in all further models.

Average WWR for fall born calves, divided in groups in which each contained all calves born within a period of ten days (DECA), were calculated, including Julian days from 1 to 180. Small representation resulted in trimming records from fall calves born before day 50 or after day 180. The averages of WWR were smaller than 100, linearly decreasing
from day 85 to day 50. After day 85, average WWR were around 100 or a little above. The system given by Eq. (14) was used. Average WWR were consistent across years (75 to 79).

DECA groups were also formed for spring born calves and average WWR calculated for these groups. Original birth days included from Julian day 143 to 386 (January of following year). Too low frequencies at the tails resulted in the trimming to within days 170 to 365. A very strong linear trend was apparent, with early born calves having WWR averaging higher than 100 and late ones less than 100. Near the middle of the distribution the situation was less clear, with the averages bouncing around 100. Within years the situation was not so clear but several years had shorter distributions and some had too small numbers. The degree of agreement across years suggests that it would probably be beneficial to estimate these spring curves within each year, if enough numbers were available. A cubic-quadratic-cubic spline was tried, cutting the curve in three nearly equal intervals and based on points where average WWR were 100 with averages below or above this mark on either side of the chosen point. According to Poirier (1973), a third degree piecewise polynomial, having smooth changes from one part of the curve to another, is less dependent on the exact knowledge of the location of the knots as, for example, if the fitted system was a linear-linear one. These "initial values," as recommended by Gallant and Fuller (1973) were used in all further models and have proved to be good approximations, as will be presented in Chapter VII.

For Age of Calf a similar approach was taken. Calves with ages within a period of ten days (DECA) formed different groups and their
average WWR was calculated. Younger calves had higher than 100 averages and older ones below 100 averages. Linear and quadratic components were evident, but higher degree or segmentations were not.

A cubic-quadratic-cubic system for AOC was supposed to exist, with the range of AOC (115 to 295 days) divided in 3 equal periods. That means, knots were assumed to exist at AOC days 175 and at 235. Adding these two cubic terms to the "model" presented at end of the last section (V.A.1.c), and making assumptions pertinent to OLS, PROC GLM from SAS (1982b), using the option to absorb CG, was used to give initial estimates of the parameters. Estimates for the two cubic components for AOC were non-significant. What was needed was to look for other possible knots, using a non-linear model.

To reduce computation cost, records were corrected in two ways: 1) using all the regression coefficients from the preceding step; or, 2) with the exception of the AOC coefficients. Averages for CG of the corrected records the first way were calculated and records corrected in the second way were transformed into corrected deviation records (CDR) from CG means. At this stage, these CDR still contained information about AOC effects but none from other environmental effects.

These CDR were used to investigate AOC effects, estimating jointly the knots. PROC NLIN from SAS (1982b) was used, and the example (Segmented Model: Example 4, pp. 31-33) given by the above source was instrumental.

A grid search for the knots together with Marquardt method were used. The knot for young AOC was looked for between ages 135 to 175 in 10-day intervals and similarly for old AOC, between ages 235 to 275.
Table 6 presents the Residual SS for each pair of these points. What is clear from the table is that as the knot $\alpha_1$ moves from day 175 to 135, the Residual SS was reduced. The same happens when $\alpha_2$ moves from day 235 to 275. Note that the smaller Residual SS is given by points (135, 275). The procedure is pushing the knots to the extreme values of AOC and it looks like that, if the range of the grid search were large enough, the procedure would push the knots past the admissible range.

This result is even more surprising when considering the fact that as the knots get closer to the edges of the range, fewer animals or observations are used in computing SSs. If the residual SS is getting smaller, then the cubic components of the system are introducing more error than they are accounting for.

The initial values for the regression coefficients of WW on the AOC system, obtained from PROC GLM, were as follows:

\[
\begin{align*}
    b_1 &= 0.58160528 \text{ Kg/d} ; \\
    b_2 &= -0.00049161 \text{ Kg/d}^2 ; \\
    b_3 &= -0.000077167 \text{ Kg/d}^3 ; \text{ and} \\
    b_4 &= 0.0000071959 \text{ Kg/d}^3 .
\end{align*}
\]

These coefficients corresponded to the following explanatory variables:

\[
\begin{align*}
    CA_1 &= \text{AOC} ; \\
    CA_2 &= \text{AOC}^2 ;
\end{align*}
\]
Table 6. Non-linear least squares grid search for the knots ($a_1$ and $a_2$) of a cubic-quadratic-cubic system created to estimate age of calf effects on weaning weight

<table>
<thead>
<tr>
<th>$a_1$</th>
<th>$a_2$</th>
<th>Residual SS$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>135</td>
<td>235</td>
<td>990,733</td>
</tr>
<tr>
<td>135</td>
<td>245</td>
<td>935,022</td>
</tr>
<tr>
<td>135</td>
<td>255</td>
<td>908,944</td>
</tr>
<tr>
<td>135</td>
<td>265</td>
<td>898,613</td>
</tr>
<tr>
<td>135</td>
<td>275</td>
<td>895,360</td>
</tr>
<tr>
<td>145</td>
<td>235</td>
<td>996,428</td>
</tr>
<tr>
<td>145</td>
<td>245</td>
<td>940,716</td>
</tr>
<tr>
<td>145</td>
<td>255</td>
<td>914,639</td>
</tr>
<tr>
<td>145</td>
<td>265</td>
<td>904,307</td>
</tr>
<tr>
<td>145</td>
<td>275</td>
<td>901,054</td>
</tr>
<tr>
<td>155</td>
<td>235</td>
<td>1,014,431</td>
</tr>
<tr>
<td>155</td>
<td>245</td>
<td>958,720</td>
</tr>
<tr>
<td>155</td>
<td>255</td>
<td>932,642</td>
</tr>
<tr>
<td>155</td>
<td>265</td>
<td>922,311</td>
</tr>
<tr>
<td>155</td>
<td>275</td>
<td>919,058</td>
</tr>
<tr>
<td>165</td>
<td>235</td>
<td>1,064,644</td>
</tr>
<tr>
<td>165</td>
<td>245</td>
<td>1,008,932</td>
</tr>
<tr>
<td>165</td>
<td>255</td>
<td>982,855</td>
</tr>
<tr>
<td>165</td>
<td>265</td>
<td>972,523</td>
</tr>
<tr>
<td>165</td>
<td>275</td>
<td>969,270</td>
</tr>
<tr>
<td>175</td>
<td>235</td>
<td>1,193,275</td>
</tr>
<tr>
<td>175</td>
<td>245</td>
<td>1,137,564</td>
</tr>
<tr>
<td>175</td>
<td>255</td>
<td>1,111,486</td>
</tr>
<tr>
<td>175</td>
<td>265</td>
<td>1,101,155</td>
</tr>
<tr>
<td>175</td>
<td>275</td>
<td>1,097,902</td>
</tr>
</tbody>
</table>

$^a$True Residual SS can be obtained by adding above figures to 3,000,000.
\[
CA_3 = (AOC - 175)^3, \text{ if } AOC \leq 175, \text{ and } \\
= 0, \text{ otherwise; }
\]
\[
CA_4 = (AOC - 235)^3, \text{ if } AOC \geq 235, \text{ and } \\
= 0, \text{ otherwise. }
\]

After the grid search found initial values for the knots (\(\alpha_1 = 135; \alpha_2 = 275\)), the procedure began to iterate solutions: \(b_3\) and \(b_4\) converged very fast to zero; \(b_1\) and \(b_2\) had almost no change, until the iterations stopped. The derivative with respect to \(\alpha_2\) involves the term \(b_4\) that multiplies all other elements of the equation. This derivative was deemed to be zero and the procedure failed to converge.

Clearly, the procedure is saying that \(b_3\) and \(b_4\) are zero, the knots do not exist and that the proposed system for AOC is over-parameterized. A simpler system, only with its linear and quadratic components was used for all further models.

The logical next step would be to repeat the above process, estimation of the knots, hopefully with more success, for each of the systems estimating the effects of AOD, seasonality within fall, and seasonality within spring; each of these as spline functions. These procedures would refine what OLS was estimating with the assumption that the knots were known and fine-tune the exact position of the knots.

But at this point three facts were considered, as follows:

1) What the gross approximation method (inspection of averages of WWR) was unable to detect or find a pattern, proved to be non-existent with a more powerful technique.
2) Ordinary least squares analyses were showing that the basic shapes of the spline functions were known already, since all the regression coefficients were significantly different from zero.

3) The computing costs of using a non-linear model, even with the process of narrowing the parameters to be estimated to the spline function of each of the variables at a time, were too high.

Based on these facts, and accepting the reasonings of Poirier (1973), the knots were assumed to be known at this point. This is a very strong assumption that should be checked in the future.

A formal definition of the first mathematical equation used in this study can now be given. This equation completely ignores sires and all other relationships among calves. Together with assumed different distributional properties of the random errors, this same equation gives rise to the first three models (OLS, WLS, and ROBWLS) mentioned at the beginning of this chapter.

The mathematical equation used for the weaning weights is:

\[
\begin{align*}
\hat{Y}_{ijklmn} & = \mu + \delta_{ijklmn} + b_1 Y_{ijklmn} + b_2 V_{ijklmn} + Y_1 b_3 C_{ijklmn} \\
& + Y_2 b_4 C_{ijklmn} + Y_1 b_5 C_{ijklmn} + Y_2 b_6 C_{ijklmn} + Y_1 b_7 C_{ijklmn} \\
& + Y_2 b_8 C_{ijklmn} + \Omega b_9 O_{ijklmn} + \Omega b_{10} O_{ijklmn} + \Omega b_{11} O_{ijklmn} \\
& + \pi b_{12} P_{ijklmn} + \pi b_{13} P_{ijklmn} + \pi b_{14} P_{ijklmn} \\
& + \pi b_{15} P_{ijklmn} + b_{16} \ell_{ijklmn} + b_{17} \ell_{ijklmn} + u_{ijklmn} 
\end{align*}
\]

(35)

Convention: "n-th" observation means "ijklmn-th" observation.
Where:

\( y_{ijklmn} \) is the weaning weight of the \( n \)-th individual of the \( ijklm \)-th sex, belonging to the \( ijk \)-th management group and born in the \( ijk \)-th herd-year-season;

\( \mu \) is a constant term;

\( g_{ijklm} \) is the \( ijklm \)-th contemporary group effect, containing main effects of the \( m \)-th sex, \( l \)-th management group, \( k \)-th season, \( j \)-th year, \( i \)-th herd and all their interactions;

\( b_1 \) to \( b_{17} \) are regression coefficients of weaning weight on each of the concomitant variables;

\( v_{ijklmn} \) is a dummy variable representing nursing status of the dam of the "n-th" individual. All possible values for \( v_1 \) are: 1 if dam belongs to the "no information" category; 0 if dam had "consecutive calvings"; -1 if dam had "alternative calvings";

\( v_{2ijklmn} \) is the 2\textsuperscript{nd} dummy variable representing nursing status of the dam of the "n-th" individual. All possible values for \( v_2 \) are: 0 if dam belongs to the "no information" category; 1 if dam had "consecutive calvings"; -1 if dam had "alternative calvings." This parameterization in terms of \( v_1 \) and \( v_2 \) corresponds to the constraint that the sum of the effects of the three nursing status levels equals zero.
\(Y_j\) is a Kronecker delta operator, with values:

1, if the "n-th" observation is from a male;
0, if the "n-th" observation is from a female.

\(Y_2\) is a Kronecker delta operator, with values:

0, if the "n-th" observation is from a male;
1, if the "n-th" observation is from a female.

\(Y_j\) and \(Y_2\) allow for the estimation of AOD effects within sex.

\(c_{lijklmn} = \text{AOD of the "n-th" observation.}\)

\(c_{2lijklmn} = \text{AOD}^2 \text{ of the "n-th" observation.}\)

\(c_{3lijklmn} = (\text{AOD} - 6)^2 \text{ of the "n-th" observation, if } \text{AOD} \geq 6;\)

\(= 0, \text{ otherwise.}\)

\(\Omega\) is a Kronecker delta operator which allows for the estimation of seasonal effects within fall, with values:

1, if the "n-th" calf is born between Julian days 50 and 160;
0, otherwise.

\(\omega_{lijklmn} = \text{Julian BIRTHDATE of the "n-th" calf.}\)

\(\omega_{2lijklmn} = (\omega_{lijklmn})^2.\)

\(\omega_{3ijklmn} = (\text{BIRTHDATE} - 85)^2,\)

if \(\text{Julian BIRTHDATE} \geq 85, \text{ for the "n-th" calf;\}

\(= 0, \text{ otherwise.}\)

\(\Upsilon\) is a Kronecker delta operator which allows for the estimation of
seasonal effects within spring, with values:

1, if the "n-th" calf is born between Julian days 170 and 365;
0, otherwise.

\[ P_{ijklmn} = \delta_{ijklmn} \]

\[ P_{ijklmn} = \delta_{ijklmn} \]

\[ P_{ijklmn} = (BIRTHDATE - 245)^3, \] if Julian BIRTHDATE of the
"n-th" calf \( \leq 245; \)
\[ = 0 , \] otherwise.

\[ P_{ijklmn} = (BIRTHDATE - 315)^3, \] if Julian BIRTHDATE of the
"n-th" calf \( \geq 315; \)
\[ = 0 , \] otherwise.

\[ t_{ijklmn} \] is the age (AOC) of the "n-th" calf at weaning date.

\[ t_{ijklmn}^2 = (t_{ijklmn})^2 \]

\[ u_{ijklmn} \] is the unexplained part, in the equation, of the
observations, the weaning weights. As a constant term was
included in the equation, these residual terms \( u_{ijklmn} \) will
sum to zero over all observations. They will also sum to zero
within each of the contemporary groups. It is assumed that if
expectations are taken over any other environmental variable
the result will be zero.

The model represented in Eq. (35) is a formal translation of
the one given in Eq. (34). Eq. (35) has a column rank deficiency of 1 (assuming no perfect collinearity among covariables), since the $g_{ijklm}$-columns will sum to 1.

This rank situation is true when multicollinearity problems are not important. When using large numbers of regressors, like in this study, most of them with low values, this problem is highly probable, even if subject matter knowledge alone does not predict such problems. One statistic relevant to look at this problem is the one given by the TOLERANCE option in PROC GLM from SAS, as described by Freund and Littel (1981). Tolerance equals $1 - R^2_{i.others}$, where $R^2_{i.others}$ is the coefficient of determination of each of the (i-th) independent variables as related to other independent variables. When $R^2$ is calculated for the variable i-th with respect to previously entered variables, the result is a Type I Tolerance; when with respect to all other variables, the result is a Type IV Tolerance.

Table 7 presents these tolerance levels for the explanatory variables. Type I Tolerances are useful to indicate how much the tolerance is reduced by adding higher powers of a polynomial, like when moving from AOD11 to AOD21, or from AOD12 to AOD22, or from SF1 to SF2, or from SS1 to SS2, or from CA1 to CA2. Reductions in tolerance levels for adding the square of the previous variable was of a magnitude 5 to 10. The use of segmented polynomials instead of arbitrary higher powers prevent the problem from getting unsolvable.

Type IV Tolerances are more useful to locate variables with serious multicollinearity problems. Three groups of variables with tolerances smaller than 0.01 can be identified, as follows:
Table 7. Investigation of multicollinearity among the explanatory variables as given by the tolerance and variance inflation factor (VIF) statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type I Tolerance</th>
<th>Type IV Tolerance</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST1</td>
<td>0.66713</td>
<td>0.41985</td>
<td>2.4</td>
</tr>
<tr>
<td>ST2</td>
<td>0.60599</td>
<td>0.53507</td>
<td>1.9</td>
</tr>
<tr>
<td>AOD11</td>
<td>0.16149</td>
<td>0.00062</td>
<td>1605.9</td>
</tr>
<tr>
<td>AOD12</td>
<td>0.16779</td>
<td>0.00061</td>
<td>1627.1</td>
</tr>
<tr>
<td>AOD21</td>
<td>0.01905</td>
<td>0.00080</td>
<td>1243.5</td>
</tr>
<tr>
<td>AOD22</td>
<td>0.01939</td>
<td>0.00079</td>
<td>1258.0</td>
</tr>
<tr>
<td>AOD31</td>
<td>0.01239</td>
<td>0.01236</td>
<td>80.9</td>
</tr>
<tr>
<td>AOD32</td>
<td>0.01235</td>
<td>0.01233</td>
<td>80.1</td>
</tr>
<tr>
<td>SF1</td>
<td>0.02747</td>
<td>0.00022</td>
<td>4492.4</td>
</tr>
<tr>
<td>SF2</td>
<td>0.00400</td>
<td>0.00041</td>
<td>2412.5</td>
</tr>
<tr>
<td>SF3</td>
<td>0.04038</td>
<td>0.02875</td>
<td>34.8</td>
</tr>
<tr>
<td>SS1</td>
<td>0.05132</td>
<td>0.00058</td>
<td>1730.4</td>
</tr>
<tr>
<td>SS2</td>
<td>0.00457</td>
<td>0.00094</td>
<td>1063.7</td>
</tr>
<tr>
<td>SS3</td>
<td>0.41837</td>
<td>0.38822</td>
<td>2.6</td>
</tr>
<tr>
<td>SS4</td>
<td>0.68420</td>
<td>0.60375</td>
<td>1.7</td>
</tr>
<tr>
<td>CA1</td>
<td>0.03041</td>
<td>0.00301</td>
<td>332.7</td>
</tr>
<tr>
<td>CA2</td>
<td>0.00434</td>
<td>0.00434</td>
<td>230.3</td>
</tr>
</tbody>
</table>

1) Among the AOD variables, given by the automatic correlation existing between small numbers and their squares. The relationship between numbers from 3 to 16 and their squares (a parabola) is very well approximated by the straight line in this region, and a sample correlation coefficient measures exactly this, the linear relationship among pairs of variables.

2) Besides being correlated within themselves, the groups of variables SF and CA are also correlated, as shown in Table 2.

3) This is also true for groups SS and CA, as shown in Table 3. Recall that the distribution in Table 3 is more sparse than in Table 2 and, hence, tolerance levels for SS are twice less
severe than for SF variables.

As stated by Freund and Littel (1981): "Extreme multicollinearity can cause severe rounding errors in the matrix inversion computation and result in incorrect answers. Since all SAS procedures perform computations in double precision (16-17 decimal digits), this type of error is not likely using SAS." The default value that SAS procedures use to deem the matrix as singular is a tolerance value of $1 \times 10^{-8}$, according to SAS (1982b). These considerations mean that the results obtained in this study are free from rounding errors, for any meaningful purpose.

Another severe consequence of multicollinearity is the increase in the variance of the parameters estimated, artificially rendering significance tests suspicious, since they will fail to reject null hypothesis with a probability much higher than expected. As will be discussed, this study is very much affected by this multicollinearity consequence. It may be worth raising the possibility that the same problem existed in the works reviewed in Chapter II (to a lesser extent since the numbers of regressors was always smaller in those studies than in the present one) and that this influenced the conclusions of those authors.

Snee (1973) coined the term variance inflation factor (VIF), according to Freund and Littel (1981). VIF is given by the reciprocal of the Type IV Tolerance statistic. The variance of each partial regression coefficient is inflated by this factor; that is, the variance is larger by that factor than it would be if all independent variables were
states that VIF provides information on how much is the variance of each
regression coefficient increased by comparison with the variance obtained
from a model (assuming the same error variance) that has each variable,
by turns, as the only explanatory variate.

VIF for each variable in the model are presented in Table 7. Large
consequences on variances of the regression coefficients can be expected.
Rounding errors are not expected, since they would become severe only if
VIF values were in the millions, if calculations are done with double
precision, according to Freund and Littel (1981). Maindonald (1984) is
more cautious, stating that: "a VIF as large as, say, 1000 calls for
scrutiny."

2. **Generalized least squares (GLS)**

This section is based on Chapter VII of Johnston (1972). Complete
results and proofs can be found from that source. Eq. (35) from the
last section describes the mathematical relation between the
environmental variables (treated as fixed effects) and the response, WW,
investigated in this study. The non-explained part of that equation was
designated \( u \), a non-observable random element. A property of the
equation is given by \( I'u = 0 \).

If to that equation the simplistic assumptions that:

1) \( E(u) = 0 \); and

2) \( E(u,u') = \sigma^2 u' n \); are added, the analysis can proceed to yield ordinary least squares
estimators (OLS).

The second assumption above is compactly stating the belief that the random variable $u$ has elements uncorrelated with each other and come from a unique population with mean zero and variance $\sigma_u^2$, with $\sigma_u^2$ probably unknown.

An assumption that is much less restrictive and that has more harmony with the data analyzed is the following:

$$E(u, u') = \sigma^2 V,$$

where $V$ is a known symmetric positive definite matrix of order $n$ (the number of observations) and $\sigma^2$ is probably unknown.

If the model is written as

$$y = \mathbf{X} \beta + u,$$

then $E(u, u') = \sigma^2 I$ is an OLS problem, while $E(u, u') = \sigma^2 V$ can be solved using GLS (Aitken) equations. Further,

$$\hat{\beta}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$$

with $V(\hat{\beta}_{OLS}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$; and

$$\hat{\beta}_{GLS} = (\mathbf{X}'V^{-1}\mathbf{X})^{-1}\mathbf{X}'V^{-1}y$$

with $V(\hat{\beta}_{GLS}) = \sigma^2 (\mathbf{X}'V^{-1}\mathbf{X})^{-1}$.

The two assumptions produce different results and are not to be chosen freely. If the data require the use of GLS but instead OLS is used, then $\hat{\beta}_{OLS}$ is still unbiased, but the minimum variance property is destroyed and $V(\hat{\beta}_{OLS}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'V\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$, in this situation.

If residuals are calculated as:

$$\hat{\varepsilon}_{OLS} = y - \mathbf{X}\hat{\beta}_{OLS}$$

and
\[ \hat{e}_{GLS} = y - X\hat{b}_{GLS}; \]

then in OLS what is being minimized is \((\hat{e}_{OLS}^\prime \hat{e}_{OLS})\), while in GLS the objective is to minimize \((\hat{e}_{GLS}^\prime V^{-1}\hat{e}_{GLS})\). In the simple case that \(V\) is a diagonal matrix, then the last expression above says that what is being minimized under GLS is the sum of squares of the observations, each weighted by the reciprocal of its variance.

**a. Heteroscedasticity of disturbances (WLS)**

Herds from different owners, with different management practices or different quality levels in their data collection practices; animals born in different years and seasons, with different seasonal constraints or enhancements; animals of different sexes; and animals managed in a more controlled environment or being raised in large numbers, with their dams, on native pastures, are all reasons to discredit the assumption that the variance within each CG is the same.

When computing ROP reports, based on BIF (1982) recommendations, large differences among standard deviations within CG were noticed. This sample statistic gives a measure of the quality of the information from each CG. Heuristically, one would like to allow observations coming from a group with smaller "unexplained errors" to have higher leverage on the estimation of parameters, breeding values and variance components.

This approach was taken, in different situations, by Eriksson et al. (1978), using different error variances according to breed of sire, and by Berger and Freeman (1978), using different error variances according to parity of cows, estimated from another data set.

When \(V\) is diagonal, with at least 2 different values for its
elements, the simpler form of a GLS problem occurs. This special case is also called a Weighted Least Squares (WLS) problem. More precisely, the $i$-th diagonal element of $V$ is $\sigma^2/\lambda_i$, with $\lambda_i \neq \lambda_i'$ for at least one $i$ and $i'$; $\sigma^2$ is generally unknown but $\lambda_i$ values are known a priori, for all $i$; and all off-diagonals are null.

In animal breeding work, it is a common practice to obtain second moments of the random elements in the model from the same data set subsequently used to obtain estimates of the parameters and "predictions" of the random elements. This will cause an "over-fit," that is, the results obtained behave better than if the true components of (co-)variance were used. The consequence of this approach, under strict statistical theory judgment, is that no property (other than unbiasedness) can be claimed for the solutions thus obtained. But sometimes (when no previous knowledge exist), this is the only alternative left, so that the analysis can proceed. No recommendation exists, so far as the author is aware, to reduce this problem in the form of, as an example, adjusting degrees of freedom for the estimation of MSE or reinflating MSE or the variance of the estimates and predictions.

Within CG error variances, after adjustment for all the 17 regressors expressed in Eq. (35), were calculated and used directly to form $V$ or to transform the data matrix as given in Johnston (1972). This action shows the extreme naivete of the author of the present study.

Unrealistic low estimates of error variances, found in some small CG, gave to the observations from these CG an unfounded high leverage (15 times higher than if the pooled variance had been used, in the most extreme case).
WLS was used with the purpose of reducing the inference of observations from "bad" CG, in a smooth fashion. This happened, but the opposite effect from the "too good" CG almost completely wiped out the expected advantages of the method, as will be shown in Chapter VII.

The arithmetic mean of all (214) within CG error variance estimates was 411.4 kg², with a standard deviation of 125.5 kg². Extreme low estimates were 26.3, 72.9, 81.7, 94.2, 95.8, and 98.3 kg² for groups of size 5, 8, 6, 10, 25, and 11, respectively. Extreme high estimates were 814.6, 844.6, 929.7, 1027.4, 1089.7, 1240.4, and 1737.6 kg² for groups of size 4, 6, 6, 9, 4, 4, and 2, respectively. Recalling that the average size of CG was 38.89 calves, the above numbers are suggesting that these extreme estimates are due to sampling errors, at least in part.

These points only became apparent three months after the analysis phase of this study was terminated. In Chapter X some alternatives to circumvent these problems will be suggested.

b. Correlated errors (mixed model) and prediction problems

Eq. (37) represents Eq. (35) in matrix form as:

\[ y = X\beta + u \]

where: \( E(u) = 0 \)

\[ V(y) = V(u) = E(u,u') = V, \text{ an } n \times n \text{ matrix.} \]

If \( V = I_0^2 \), OLS yield best linear unbiased estimates (BLUE) of \( \beta \); if \( V \neq I_0^2 \) but diagonal, WLS will yield BLUE of \( \beta \); otherwise, only the use of Aitken equations will yield BLUE of \( \beta \), which requires \( V^{-1} \).

In this study, calves are considered to be paternal half sibs or
unrelated. Sires are considered unrelated and non-inbred.

If we define:

- $Z$ as the incidence matrix relating calves and their sires;
- $\sigma_a^2$ as the additive genetic variance of $WW$;
- $\sigma_s^2 = \frac{1}{4} \sigma_a^2$ as the among sire variance;
- $\sigma_e^2$ as the variance among progeny of the same sire ($\sigma_e^2 = \frac{3}{4} \sigma_a^2 + \sigma_s^2$);
- $\sigma_c^2$ as the assumed pure environmental variance among $WW$; and
- $R\sigma_e^2$ as the var-cov matrix among observations, after accounting for
  the effect of common sires;

then $V = ZZ\sigma_s^2 + R\sigma_e^2$.

With the further assumption that $R$ is diagonal and if observations
are ordered by sire, then $V$ is a block diagonal matrix, with diagonal
elements equal to $\sigma_y^2 = \sigma_s^2 + \sigma_e^2$ and off-diagonals equal to $\sigma_s^2 = \frac{1}{4} \sigma_a^2$ if
calves are half sibs or equal to zero otherwise.

Assuming $R = \mathbf{I}\sigma_e^2$, Henderson (1950) proposed the use of the mixed
model equations (MME). Considering full rank matrices, Henderson et al.
(1959) proved that the use of MME yield the same solutions as the Aitken
equations for the fixed effects and, hence, the same BLUE properties can
be attached to these estimators.

Selection is the single most important source to obtain long-term
genetic response in beef cattle. Ability to identify superior animals is
a primary tool for animal breeders to achieve their desiderata.

Economists have to take decisions on a day-by-day basis, using all
available information they have, trying to predict future developments in
a world they do not control but is subject to sudden changes, in order to maximize profit (generally). Econometric and animal breeding theory and practice have many common problems which may be solved using the same statistical methods. More channels of communication between these fields should have a synergetic effect on science.

Goldberger (1962) gave the truly seminal result for prediction in a GLS setting. Quotations that follow are from the Johnston (1972) review of Goldberger (1962) paper.

"Prediction problems require special treatment in a GLS model," as in Eq. (37). "The problem is to predict a single value of the regressand \( y^o \) (the predictand is the production of a future progeny of a sire) given the row vector of prediction regressors \( x^o \). We can write

\[
y^o = x^o \beta + u^o
\]

where \( u^o \) is the true but unknown value of the prediction disturbance \("(u^o \)

is the sire's predicted difference). "We assume:

\[
\begin{align*}
E(u^o) &= 0 \\
E(u^2_o) &= \sigma^2_o \\
E(u^o, u) &= E \begin{bmatrix} u_1, u_0 \\ u_2, u_0 \\ \vdots \\ u_n, u_0 \end{bmatrix} = w
\end{align*}
\]

Where \( w \) is the \( n \times 1 \) vector of covariances of the prediction disturbance with the vector of sample disturbances."

In animal breeding terminology, if \( G \) is a \( q \times q \) matrix representing
the var-cov matrix of sires then $ZG$ is the covariance between observations $y$ and sires' true breeding values. The column from $GZ'$ corresponding to the sire of interest $u_o$, should be interpreted as $w$ above.

"Define a linear predictor $p = c'y$ where $c$ is a vector of $n$ constants. If $p$ is to be a best linear unbiased predictor we must then choose $c$ to minimize the prediction variance

$$
\sigma_p^2 = E[(p - y_o)^2]
$$

subject to $E(p - y_o) = 0$" (unbiasedness).

From $p = c'y$, $y_o = x_o \beta + u_o$ and $y = X\hat{\beta} + u$, "we have

$$
p - y_o = c'y - x_o \beta - u_o
$$

$$
= c'X\hat{\beta} + c'u - x_o \beta - u_o
$$

$$
p - y_o = (c'X - x_o)\beta + c'u - u_o.
$$

The condition for an unbiased predictor thus require that $c$ must satisfy

$$
c'X - x_o = 0 \text{ or } c'X = x_o.
$$

The prediction error is then

$$
p - y_o = c'u - u_o
$$

and the prediction variance is

$$
\sigma_p^2 = E[(p - y_o)^2]
$$
\[
\sigma_p^2 = E[(p - y_o)(p - y_o)'] \text{ since } (p - y_o) \text{ is a scalar} \\
= E[(c'u - u_o)(c'u - u_o)'] \\
= E[c'u'u'c + u_o^2 - 2c'u'w]
\]

\[
\sigma_p^2 = c'Vc + \sigma_o^2 - 2c'w .
\]

To minimize \( \sigma_p^2 \) subject to unbiasedness, we minimize

\[
\phi = c'Vc - 2c'w - 2(c'X - x_o)\lambda
\]

Where \( \lambda \) is a vector of Lagrange multipliers. Differentiating with respect to \( c \) and \( \lambda \), and setting the results equal to zero vectors gives the equations

\[
2Vc - 2w - 2X = 0 \\
-2X'c + 2x_o' = 0
\]

or

\[
\begin{bmatrix} V & X \\ X' & 0 \end{bmatrix} \begin{bmatrix} c \\ -\lambda \end{bmatrix} = \begin{bmatrix} w \\ x_o' \end{bmatrix}
\]

or

\[
\begin{bmatrix} \hat{c} \\ -\hat{\lambda} \end{bmatrix} = \begin{bmatrix} V & X \\ X' & 0 \end{bmatrix}^{-1} \begin{bmatrix} w \\ x_o' \end{bmatrix}
\]

Applying the rule for the inverse of a partitioned matrix then gives

\[
\hat{c} = V^{-1}[I - X(X'V^{-1}X)^{-1}X'V^{-1}]w + V^{-1}X(X'V^{-1}X)^{-1}x_o
\]

so that the best linear unbiased predictor is
\[ \hat{p} = c'y \]
\[ \hat{p} = x_o'[(X'V^{-1}X)^{-1}X'V^{-1}y + w'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}y] \]

and, since \[ \hat{b}_{GLS} = (X'V^{-1}X)^{-1}X'V^{-1}y \],

\[ \hat{p} = x_o \hat{b}_{GLS} + w'V^{-1}(y - X\hat{b}_{GLS}). \]

Further
\[ \hat{p} = x_o \hat{b}_{GLS} + u_o \]

with \[ u_o = w'V^{-1}(y - X\hat{b}_{GLS}) \], or
\[ u_o = (GZ')_o V^{-1}(y - X\hat{b}_{GLS}) \].

Henderson (1963) obtained, after Goldberger (1962), exactly the same result. Henderson (1984) states that "Henderson (1963) and Goldberger (1962) proved that BLUP of \( u \) is \( \hat{u} = GZ'V^{-1}(y - X\hat{b}_{GLS}) \)," but no claim of independence of results or otherwise is made.

Henderson (1963) also proved that the use of MME would yield a GLS estimator of \( u \), with BLUE or BLUP properties. Henderson (1973, 1977) developed methods and properties for GLS prediction of future records, along with results of Goldberger (1962).

GLS estimators were obtained in this study. Sires were included in the model (35), as random variables, with the following assumptions:

\[ E(s) = 0 \], independently from which set the sire was connected;
\[ E(s^2) = 1.0 \sigma^2_s \], that means a common sire variance for all sets and no relationships among sires.
\[ R = V(e) = 1.0 \sigma^2_e \], that means a common variance and that the only source of covariance (phenotypic) among calves was from common sires.
R # I. was intended to be used in this study, but given the poor results obtained from the WLS-fixed model and the problems discussed in the last subsection, this was not attempted.

According to Henderson Jr. and Henderson (1979), use of the MME for a model containing covariates is a simple extension and solutions have BLUE properties.

3. Robust estimation WW and residuals (RESWW) from fitting model (35) under OLS assumptions were analyzed using PROC UNIVARIATE from SAS (1982a). Sample statistics of interest from these two variables are as follows:

<table>
<thead>
<tr>
<th>WW(kg)</th>
<th>RESWW(kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKEWNESS</td>
<td>0.3129</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>0.0243</td>
</tr>
<tr>
<td>D : NORMAL</td>
<td>0.0366</td>
</tr>
<tr>
<td>PROB &gt; D</td>
<td>&lt; 0.01</td>
</tr>
</tbody>
</table>

EXTREMES

- LOWEST
  60  -60.6
  61  -60.4
  62  -60.4
  62  -59.8
  63  -59.5

- HIGHEST
  252  61.3
  255  61.4
  260  61.5
  269  61.6
  270  62.1

The author had no previous experience with these higher moment (3rd
and 4th) statistics and no comparative values for ROP data could be found.

PROC UNIVARIATE presents also bar charts, boxplots, and normal probability plots. Mostly based on these, and also on the definitions of the above statistics, some statements can be made, as follows:

1) The probabilities associated with the values of the Kolmogorov D statistics are expressing a high level of disbelief that the samples came from a normal distribution.

2) WW are positively skewed, maybe because of a failure by the part of the breeders to report that some animals were temporarily under a special treatment.

3) The variables have too heavy tails, as expressed by the kurtosis statistics, especially RESWW. Heavy tails diagnoses too many extreme observations, that can be too influential on parameter estimation (and even more on variance components estimation).

It was assumed that the non-normality was due to the presence of outliers and a method to reduce the leverage of these observations was devised and used. This will be presented after a short review on Robust statistics and their uses.

Laboratory measurements of fat and protein contend in milk delivered from farmers to 6 different Norwegian dairies were used to compare 57 robust estimators of location by Spjøtvoll and Aastveit (1983). Using Monte Carlo techniques they looked at the increase in precision of those estimators relative to the sample mean. They found that the mean had a
strikingly poor performance; the average increase of efficiency by using the better robust estimators is more than 50 percent (for some dairies it was more than 200 percent). Not all robust estimators did equally well, but there is a host to choose from. They noticed that the efficiency of the robust estimators increases with sample size; and that this effect is most pronounced for the populations where the advantage of using robust estimators is greatest.

This latter conclusion of Spjøtvoll and Aastveit (1983) challenges animal breeders' traditional posture, where there exists a complete reliance on asymptotic properties of the sample mean and, by extension, on $R = \frac{\text{OLS}}{\text{Median}}$; or on the fact that as sample sizes get bigger, the OLS will be unbiased and most efficient. In a non-robust setting, if a sire has a few progeny, one or a few extreme observations, inflated by gross measurement or intentional error, have the power to rank the sire far from the true placement or from future ranks, when more information is available. The same is true for the estimates of fixed effects.

Bickel (1976) presented a remarkably clear paper, a philosophical review and heated discussion on robustness against gross errors. Bickel sees the history of robust estimation as a continuous factional strife, Bayesians vs. non-Bayesians, subjective vs. objective probability theorists, believers in the likelihood principle against disbelievers and so on. He makes two quotations to exemplify the extremes of the spectrum:

"The hallmark of good science is that it uses models and theory but never believes them." M. Wilk (cf. Bickel, 1976)
And on rejection of outliers:

"We believe that only through a firm adherence to this rule (never rejecting an observation for internal reasons in calculating the mean) have we been able to remove arbitrariness from our results."

Bessel (cf. Bickel, 1976)

Bickel (1976) views robustness ideas as attempts to bridge the gap between these points of view. He points out that robust fitting is an iterative method and presents a form of robust estimator that is equivalent to a least squares estimator. Bickel states that most robust estimators can be interpreted in this way and his arguments (not literally) are as follows:

1) Obtain least squares estimates from the model \( y = X + e \)
   \((e \sim N(0, \sigma^2_e)\) assumed) for \( \beta, \hat{y} \), and \( \hat{e} \).

2) Express the confidence held for each \( \hat{e}_i \) with the function \( \psi(\hat{e}_i) \), preserving the sum-to-zero property of \( \hat{e} \). \( \psi \) must be a continuous (not necessarily everywhere) function of \( \hat{e} \).

3) Construct pseudo-observations \( y^* = \hat{y} + \psi(\hat{e}_i) \).

4) Obtain OLS estimators from \( y^* = X \beta + e \) \((e \sim N(0, \sigma^2_e)\) more tenable now) and \( \hat{\beta}^* \) will be the robust estimator according to the \( \psi(\hat{e}_i) \) used. The process may begin a new cycle here.

Askin and Montgomery (1980) combined biased and robust regression techniques. Biased estimators that may reduce variance of the estimates and MSE more than they induce squared bias considered were:
1) ridge estimators, proposed by (Hoerl and Kennard, 1970, as cited by Askin and Montgomery, 1980),
2) variations of the principal components estimator, discussed by (Marquardt, 1960, as cited by Askin and Montgomery, 1980), and
3) the uniform shrinkage estimator (Stein, 1960, as cited by Askin and Montgomery, 1980).

To deal with non-normality of the errors, especially if errors are drawn from a distribution with heavier tails than the normal or are either location or scale contaminated, Askin and Montgomery (1980) considered M-estimators, originally proposed by Huber (1964).

Ridge regression has a strong resemblance, in the form of the equations, to the MME of Henderson (1950) as noted by, e.g., Schaeffer (1983). It may help to broaden perspectives to consider that if the sires considered are the only ones of interest, are not related and have about equal information, then the use of MME will produce biased estimates of sire effects (in comparison with estimates from a fixed model).

Askin and Montgomery (1980) elaborated more on the use of the ridge estimator

\[ \hat{\beta} = (X'X + kI)^{-1}X'y, \quad k > 0, \]

where \( \hat{\beta} \) is a p-vector, and noted, after Marquardt (1970), that to perform ridge regression on a standard regression computer package, all that is needed is to augment the original data with the p points \( (X, y) = (k^{1/2}I_p, 0) \). The same can be said for the MME, using PROC GLM of SAS.

To obtain robust M-estimates Askin and Montgomery (1980) suggest using a diagonal matrix $W$ whose $w_{ii}$ element is a function of the $i$-th residual, in such a way that the function of the residuals being minimized is less influenced by large residuals than in least squares. Reestimating residuals at each step corresponds to the method called Iteratively Reweighted Least Squares (IRLS). Augmented observations always have a weight of 1.0. The proposed combined estimator (biased-robust), in examples tested by the authors produced superior results when compared to either robust biased or estimators alone. The form of these estimators is given by

$$\beta^o = [X'WX + X'_\text{Aug}X_{\text{Aug}}]^{-1}X'Wy$$

Partitioning $X$ and $X_{\text{Aug}}$ into fixed and random parts of a mixed model, and interpreting $W$ by $R$ gives a framework to obtain robust-GLS estimates using Henderson's MME on general regression computer packages.

Robust estimation in the mixed model was studied by Fellner (1982). This author later extended robustness (down-weighting of moderate outliers) to variance component estimates, via Restricted Maximum Likelihood (Fellner, 1983).

General texts covering the subject of robust statistics include: Andrews et al. (1972), where an extensive list of robust estimators are explained and their small- and large-sample properties are examined; Huber (1977 and 1981) is a highly regarded source among statisticians but the author of this study was unable to grasp much of it at all; and Hoaglin et al. (1983) is a highly readable and understandable source.
Colin Goodall who wrote Chapter 11 in Hoaglin et al. (1983), presented an outline of the theory on M-estimators of location, on which what follows is based, until the presentation of the weighing function for the observations in this study.

Three classes—L-estimators (linear combinations of the ordered observations), R-estimators (derived from rank tests) and M-estimators (the class that offers the greatest advantages in performance, flexibility and convenience)—have played a prominent role in recent research on robust estimation.

a. M-estimators  M-estimators minimize functions of the deviations of the observations from the estimate that are more general than the sum of squared deviations or the sum of absolute deviations. In this way, the class of M-estimators includes the mean and the median as special cases. Instead of squaring the deviation of each observation $x_i$ from the estimate $t$, a function $\rho(x;t)$ is applied and the objective function is formed by summing over the sample:

$$\sum_{i=1}^{n} \rho(x_i;t).$$

Often $\rho(x;t)$ depends on $x$ and $t$ only through $x-t$, so that it is possible to write $\rho(x-t)$. The nature of $\rho$ determines the properties of the M-estimator. Viewed in another way, M-estimators generalize the idea of the maximum likelihood estimator of the location parameter in a specified distribution. Thus, it is reasonable to expect that a suitably chosen M-estimator will have good robustness of efficiency in large samples. In
fact, the original theoretical development was motivated by achieving robustness in a neighborhood of the Gaussian distribution.

**DEFINITION:** The M-estimate $T_n(x_1, x_2, \ldots, x_n)$ for the function $\rho$ and the sample $x_1, x_2, \ldots, x_n$ is the value of $t$ that minimizes the objective function

$$
\sum_{i=1}^{n} \rho(x_i; t).
$$

When the derivative of $\rho$ with respect to $t$ is known, a function which (except for a multiplicative constant) can be denoted by $\psi$, it may be more convenient to calculate $T_n$ by finding the value of $t$ that satisfies

$$
\sum_{i=1}^{n} \psi(x_i; t) = 0.
$$

[$\rho$ is expected to be continuous, but not necessarily everywhere; the derivative may fail to exist at a finite number of points.]

The most familiar M-estimate is the sample mean, the least square estimator of location, when $\rho$ is the square of the residual:

$$
\rho(x; t) = (x-t)^2 \quad \text{and} \quad Q = \sum_{i=1}^{n} (x_i-t)^2
$$

is minimized (the objective function). Differentiating $Q$ with respect to $t$ leads to
From there,

\[ t = \frac{\sum_{i=1}^{n} x_i}{n} \]

solves the equation and the sample mean is the M-estimator when the objective function is the square of the residual.

A second example is the sample median (when \( n \) is an even number any answer between the two middle observations is acceptable). The objective function is the absolute value of the residual,

\[ \rho(x; t) = |x - t| \]

and the corresponding \( \psi \)-function is

\[ \psi(x; t) = \text{sgn}(x - t) \]

where

\[ \text{sgn}(x - t) = \begin{cases} +1 & \text{if } (s - t) > 0 \\ 0 & \text{if } (s - t) = 0 \\ -1 & \text{if } (s - t) < 0 \end{cases} \]

The expression

\[ \sum_{i=1}^{n} \psi(x_i; t) = \sum_{i=1}^{n} \text{sgn}(x_i - t) \]

counts each observation above \( t \) as +1 and each observation below \( t \) as -1, so that when \( t = \text{median} \) the sum will be zero. The function \( \psi \) that yields the median is a two-step function, with a jump at \( t \).

Through the choice of objective functions \( (\rho) \), M-estimation
generalizes least squares estimations, but \( \psi \) functions are more convenient to work with than \( \rho \) since they solve for the estimators.

b. Equivariance, invariance, and scale estimation

It is desirable for M-estimators to have some properties known, like the ways that estimators respond to systematic changes in the sample.

When the whole sample is shifted by an amount \( a \), any reasonable estimator of location should follow by the same amount. Specifically, what is wanted is

\[
T_n(x_1 + a, ..., x_n + a) = T_n(x_1, ..., x_n) + a,
\]

so that such an estimator can be said to be location-equivariant (Huber, 1981). Note that in hypothesis testing, \( a \) may be the quantity of primary interest.

Often the whole sample may be multiplied by a non-zero constant \( b \) and then shifted by \( a \). Again, the estimator of location should follow this simultaneous change in location and scale:

\[
T_n(bx_1 + a, ..., bx_n + a) = bT_n(x_1, ..., x_n) + a.
\]

Such estimators are said to be location-and-scale-equivariant. Another way to write this property, possibly more suggestively, is:

\[
T_n(x_1, ..., x_n) = BT_n((x_1-A)/B, ..., (x_n-A)/B) + A
\]

Most M-estimators achieve this property only by incorporating some measure of the scale of the sample, here generally denoted by \( S_n \). It is reasonable to require that shifting the whole sample leaves the value of
unchanged ($S_n$ is location-invariant, and that $S_n$ follow changes in scaling):

$$S_n(bx_1, \ldots, bx_n) = |b| S_n(x_1, \ldots, x_n);$$

that is

$S_n$ is scale-equivariant.

In general, an M-estimator of location must take account of the scale of the sample in order to be location-and-scale-equivariant. Two exceptions are the mean and the median. Other M-estimators of location depend on the scale of the arguments of $\rho$ and $\psi$. Therefore, an auxiliary estimator of scale $S_n(s_1, \ldots, x_n)$ has to be chosen and used, together with a constant $c$, to rescale the residuals. Observations are centered and rescaled by $u_i = (x_i - t)/cS_n$. $T_n$ is a value of $t$ that minimizes

$$\sum_{i=1}^{n} \rho(u_i)$$

or satisfies

$$\sum_{i=1}^{n} \psi(u_i) = 0 .$$

If $S_n$ is scale-equivariant then $T_n$ is location-and-scale-equivariant.

In general, an M-estimator of location involves a fixed measure of scale $S_n$. The constant $c$ in the definition of $u_i$ matches the basic scale of $\psi$ or $\rho$ to the scale of the data, measured in units of $S_n$. An inherent part of the definition of the M-estimator, $c$ is known as the tuning constant because it can be chosen to fine-tune the estimator so that it
has a specified asymptotic efficiency at a chosen distribution (usually the Gaussian). Commonly, the \( \psi \)-function is curved or is composed of several linear pieces, and some \( \psi \)-functions are identically zero outside a given interval. Thus, an M-estimator of location is actually a family of estimators, all involving the same scale measure or auxiliary scale estimator; the value of the tuning constant determines the individual member of the family. Most often used auxiliary scale estimators are:

The Median Absolute Deviation, \( \text{MAD} = \text{med}_i \{ |x_i - \text{med}_j \{ x_j \}| \} \); and the fourth-spread (\( d_k \), the difference between the lower fourth and upper fourth) or its close relative, the inter-quartile range.

Maximum-likelihood techniques can be used to find simultaneous estimates of location and scale. A family of probability density functions is given by \( f(x; \theta) \), with \( \theta \) and \( \sigma \) being, respectively, the location and scale parameters and \( f' = \sigma \psi \). Differentiating the log-likelihood with respect to \( \theta \) and \( \sigma \), simultaneous equations are found that, ultimately, will yield (after incorporation of the tuning constant \( c \)):

\[
\sum_{i=1}^{n} \psi \left( \frac{x_i - T_n}{c S_n} \right) = 0 \quad \text{and} \quad \sum_{i=1}^{n} \chi \left( \frac{x_i - T_n}{c S_n} \right) = 0.
\]

\( \psi \) and \( \chi \) are chosen to give \( T_n \) and \( S_n \) the desired properties. Usually, \( \psi \) is an odd function and \( \chi \) is an even function as in maximum-likelihood estimation when \( f \) is symmetric.
c. Resistance and robustness of efficiency  
An estimator is resistant if it is affected to only a limited extent either by a small number of gross errors or by any number of small rounding and grouping errors. An estimator is resistant to gross errors if a small subset of the sample cannot have a disproportionate effect on the estimate (a single observation can dominate the mean of a sample). An estimator has robustness of efficiency over a range of distributions if its variance (or, for biased estimators, its mean squared error) is close to the minimum for each distribution.

The extreme observations associated with sampling from a heavy-tailed distribution suggest the inconsistency of the data with an underlying Gaussian distribution. Sampling from a light-tailed distribution (like after trimming extreme observations) produces no comparable indication of inconsistency.

d. Breakdown bound  
As estimator is resistant if it is altered to only a limited extent by a small proportion of outliers. The estimator breaks down if the proportion becomes too large.

DEFINITION: The breakdown bound (or point) of an estimator is the largest possible fraction of the observations for which there is a bound on the change in the estimate when that fraction of the sample is altered without restriction.

An estimator is resistant only if its breakdown bound is greater than 0. As a single observation is made larger, the mean increases without bound. Therefore, the breakdown bound of the mean is 0. The
breakdown bound of the median is the largest possible for a location estimator that treats observations on each side of the estimate symmetrically. It is \((1/2 - 1/n)\) of even \(n\), and \((1/2 - 1/2n)\) for odd \(n\). The breakdown bound of any estimator that treats observations equivariantly cannot exceed 0.5. There do exist estimators with breakdown bounds greater than 0.5.

**e. W-estimation**  
An alternative form of M-estimation is called W-estimation. The M-estimate \(T_n\) is defined by

\[
\sum_{i=1}^{n} \psi \left( \frac{x_i - T_n}{cS_n} \right) = 0
\]

When \(w\) is defined accordingly to \(uw(u) = \psi(u)\) and substituted for \(\psi(u)\) above, the result is

\[
\sum_{i=1}^{n} \left[ \frac{x_i - T_n}{cS_n} \right] w \left[ \frac{x_i - T_n}{cS_n} \right] = 0 , \text{ or, rearranging,}
\]

\[
T_n = \left[ \frac{\sum_{i=1}^{n} x_i w \frac{x_i - T_n}{cS_n}}{\sum_{i=1}^{n} w \frac{x_i - T_n}{cS_n}} \right]
\]

\(T_n\) is thus a weighted mean of the \(x_i\), defined iteratively by the equation above, said to be a W-estimate based on the weight function \(w\). If \(\psi\) is an odd function, then \(w\) is even: \(w(u) = w(-u)\).

A trivial example is the sample mean, without any auxiliary scale, and the \(\psi\)-function is simple \(\psi(u_i) = u_i\) with \(u_i = x_i - T_n\). The weight
function is identically equal to 1 (so that $\psi(u_i) = u_i = u_i w(u_i)$), and the equation above becomes

$$T_n = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} w(u_i)}.$$ 

To obtain a solution, the formula above has to be iterated to produce a numerical solution. If $T_n(k)$ is the estimate at the $k$-th iteration, and

$$u_i^{(k)} = \frac{x_i - T_n^{(k)}}{cS_n},$$

then

$$T_n^{(k+1)} = \frac{\sum_{i=1}^{n} x_i w(u_i^{(k)})}{\sum_{i=1}^{n} w(u_i^{(k)})},$$

and this is an example of a general procedure known as IRLS estimation, with weights $w_i^{(k)} = w(u_i^{(k)})$. WLS estimate, $T_w$, based on fixed weights $w_i$ is the choice of $t$ that minimizes

$$\sum_{i=1}^{n} \frac{x_i - t}{w_i}^2,$$

namely $T_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i}$. 

IRLS is a further development. Each weight depends on the residual at the preceding iteration. Starting with a sensible initial estimate (for W-estimation of location the sample median is a good choice), iterations are performed until the sequence of estimates has converged to within the desired accuracy. The weights at each iteration correspond to the contribution of each observation to the estimate.

The OLS estimator is the ML estimator at the Gaussian distribution. Through the connection between density and objective functions, M-estimation may be regarded as a generalization of ML estimation. This is the origin of the term "M-estimator." Another generalization of least squares estimation is to IRLS—and thus to W-estimation. It is illuminating to see that the two generalizations, M- and W-estimation, are very closely related. The theory of M-estimation provides insight into IRLS-estimation.

The author finds it necessary to state again the material covered in subsections a) until here is comprised of excerpts (not always literal ones) from Goodall in Hoaglin et al. (1983) that should be consulted for a much more complete coverage. This large transcription of material was deemed as necessary by the author because, although the reader probably is worried about the lack of resistance of the estimates he/she obtains on his/her analysis, formal courses in robust statistics are a rare commodity and readable material even more so. Another point to be made with this last sections is that the method employed in this study is not the product of imagination of the author but it is embodied in statistical science. Maybe the message behind Huber's (1981) text is that robust statistics is not an inconsequent amateur's hobby.
f. How to obtain a perfect, absolute fit to your model

serious treatment under such a heading is possible except in dealing with a deterministic model. This joke-subsection is useful to show the extreme power of W-estimators (ROBWLS) and the dullness of the author or, at least, to save some computer expenses for someone else.

To get rid completely of the leverage of residuals (this really happened) why not do the following: 1st) fit an OLS model and obtain $\sigma_{e_i}^2$ for each CG; 2nd) fit a WLS and obtain the residuals $\hat{r}_{WLS}$; 3rd) fit a WLS again, using as weights: $[\hat{r}_{WLS}^2]^{-1}$.

The results were (everybody can cure his/her LOF syndrome):

1) MSE = 0 ;
2) F- and t-tests = how many nines as the format allows to be printed;
3) Variance of (Perfect fit) estimators = 0.0; and
4) $R^2 = 1.0$

g. Robust weighted least squares (ROBWLS)

A W-estimator was designed to be used with PROC GLM of SAS (1982b). It should take care both of the non-homogeneity of variances among CG, and also down-weight extreme observations, without removing (weight = 0) any observation, even if a probable outlier. The $\rho$-function employed (or the weights) express the belief that the underlying distribution of WW is Gaussian, before being contaminated by gross errors and some selection across management groups within HYS and Sex. At the same time, as $\rho$ (a symmetric step-function around the initial location estimate) is uniform between $-\hat{\omega}_e$ and $+\hat{\omega}_e$, the ROBWLS is not excessively dependent on the initial
estimates (OLS). To reduce computation costs, ROBWLS can be defined as a two-steps estimator.

To form the weights, the following steps were taken:

1) Calculate \( \hat{r}_{ij} \), the OLS \( i,j \)-th residuals, and \( \hat{\sigma}^2_{e_i} \), the \( i \)-th within CG OLS error variance estimates.

2) Compute \( S_{ij} = \text{INTEGER VALUE} \cdot \frac{|\hat{r}_{ij}|}{\hat{\sigma}^2_{e_i}} \).

3) Form a diagonal matrix \( W \) of weights with \( [(S_{ij} + 1) \cdot \hat{\sigma}^2_{e_i}]^{-2} \) or use

\[ [(S_{ij} + 1) \cdot \hat{\sigma}^2_{e_i}]^{-1} \]

to transform the data matrix.

The objective function being minimized is \( Q = \sum_{ij} \hat{r}_{ij} \cdot w_{ij} \). Clearly, \( \rho = |w_{ij} \hat{r}_{ij}| \) and a wording of the relative leverage \( (q_{ij}) \) of each observation is as follows:

if \( -1 \hat{\sigma}_{e_i} < \hat{r}_{ij} < +1 \hat{\sigma}_{e_i} \) then \( q_{ij} = \frac{\hat{r}_{ij}^2}{\hat{\sigma}^2_{e_i}} \);

if \( 1 \hat{\sigma}_{e_i} < |\hat{r}_{ij}| < 2 \hat{\sigma}_{e_i} \) then \( q_{ij} = \frac{\hat{r}_{ij}^2}{4 \hat{\sigma}^2_{e_i}} \);

if \( 2 \hat{\sigma}_{e_i} < |\hat{r}_{ij}| < 3 \hat{\sigma}_{e_i} \) then \( q_{ij} = \frac{\hat{r}_{ij}^2}{9 \hat{\sigma}^2_{e_i}} \);
if $3\sigma_{e_i} < |\hat{r}_{ij}| < 4\sigma_{e_i}$ then $q_{ij} = \frac{r_{ij}^2}{16\sigma_{e_i}^2}$.

No observation in the data set studied had $|\hat{r}_{ij}| > 3\sigma_{e_i}$.

Results from this method of estimation are presented in Chapter VII. ROBWLS was not used in combination with mixed models due to the problems with the estimates discussed in Section V.A.2.a.

The magnitude of the weights used is subject of a degree of subjectivity, while the form seems a good compromise. Only simulation studies could help in determining a "best" set of weights for ROP data, but that would be a complete new study. Beforehand, it can be said that ROBWLS performed much better that OLS or WLS.

B. Variance Component Estimation

Searle (1978), in the introduction of his summary paper, described not without a good dose of flair, the situation animal breeders are immersed in as follow:

Variance components estimation was, for several decades, the poor step-child of analysis of variance, but in recent years the subject has generated quite widespread interest. Until 1967, methods of estimation were based on equating sums of squares to their expected values. For balance data, this method involves sums of squares associated with traditional analysis of variance and useful minimum variance properties were derived since the late 1950's. For unbalanced data, Henderson's three methods of estimation, based on the same principle of equating sums of squares to their expected values, were used exclusively since 1953. Succeeding years saw expansion and explanation of these methods together with exploration of their properties. No new development happened until 1967 when maximum likelihood (ML) procedures, based on normality assumptions, were
described. Since then there has been a whole host of new methods, not only ML, but REML, MINQUE, I-MINQUE and MIVQUE - an doubtless some other alphabetic horrors also. In addition there are peripheral topics tangential to computing techniques - such as Henderson's MME and Pukelsheim's Dispersion-mean model. As foundation for all this there is a large corpus of matrix algebra, there are numerous notation that look sufficiently alike to add the traditional amount of confusion and, hanging like a thunder cloud over everything, are numerical and computing problems involved with very large data sets, sparse matrices, and the solving of non-linear equations subject of non-linear (non-negativity) constraints.

Given this overall panorama, as given by Searle (1978), the author does not feel too bad for being still under a state of shock and recognizes this area as his weakest point. Nevertheless, the problem has to be faced and estimates were needed.

Considering what is expressed in Chapter I, heritability estimates significantly different from zero were needed if the preweaning phase of the ROP program is to be justified. Considering the above; the high costs to obtain the field records and to expand the program; and that the data set is not too big or, better, that the number of random effects (sires) is not too large, it was decided that the "best available method" should be used.

Two qualities of the estimators were sought:
1) the estimates should be positive; and
2) the estimates should be minimum variance.

Harville (1983) presented a complete derivation of REML, properties of REML estimates, and a complete example (a mixed model) on how to obtain these estimates.

On the derivation of REML estimators, it is necessary to limit the
parameter space to positive values so that solution can be obtained from
equations (derivatives of the log-likelihood function). The property of
non-negativity of the estimates is generally attached to REML (and ML)
estimators. In two examples used by the author, when other estimators
gave negative estimates, the ML or REML methods (both iterative) failed
to converge. It is a temerity to make inferences based on these two
examples, but the author is in doubt about this non-negativity property.
For all practical purposes, a negative estimate and a failure in the
convergence of the estimates are the same result: no useful result.

ML estimators are asymptotic efficient and Searle (1979) shows that
REML, I-MINQUE and MIVQUE under normality are the same (if estimates are
positive). Thus, the two desired criteria above were reasonably well
fulfilled by REML.

Besides those criteria, a determining factor was the fact that the
author had gained some experience with REML. To illustrate the point, in
a recent seminar about embryo transfer, when asked about which is the
best methodology to follow for heat synchronization, multiple ovulation,
etc., the speaker boldly said: "The best method is the one you know how
to use. If the method is reasonable, after some practice you begin to
improve it and interact with it when under different circumstances. I
believe almost each professional has his own method and this is, truly,
the best method for him." Probably much of the discussion about variance
component methods could be saved if the above point was recognized.

Harville (1977) gives complete derivations of ML and REML estima-
tors, compares them, develops numerical procedures, originally proposes
an approximative REML (Henderson New Method, Schaeffer, 1983), and
relates ML and REML to other methods. Searle (1979) gives extensive
details of the attendant algebra and reviews much of the subject.
Schaeffer (1983) gives examples and discuss in more detail computational
aspects.

To obtain REML estimates for error and sire variance components, in
accordance with Harville (1983), the following computing formulas were
used:

\[ \hat{\sigma}^2_e = \frac{\text{Res SS}}{n - r(X)} \]

\[ \hat{\sigma}^2_s = \frac{\hat{s}'\hat{s}}{q - k*\text{tr}(C_{22})} \]

where:

\text{Res SS} = y'y - \hat{s}'X'y - \hat{s}'Z'y ;

\( n = \text{total number of observations;} \)

\( r(X) = \text{number of linearly independent fixed effects or} \)

\( = \text{number of CG + number of regressors;} \)

\( \hat{s}, \hat{s} = \text{GLS solutions, with BLUE properties, obtained with} \)

\( \text{the converged value of } k; \)

\( k = \frac{\hat{\sigma}^2_e}{\hat{\sigma}^2_s} ; \)

\( q = \text{number of sires; and} \)

\( \text{tr}(C_{22}) = \text{trace of the part of the complete inverse of the} \)

\( \text{coefficient matrix (LHS) of the MME corresponding to} \)

\( \text{the random element (sires).} \)

Some features of these formulas are as follows:
1) the denominator of $\hat{\sigma}_e^2$ (unbiased, while ML estimator is biased) is not adjusted for loss of degrees of freedom because $\hat{s}$ are being estimated, although Res SS is the residual after sires (and all fixed effects). In an ANOVA method $q$ would be subtracted from Res df. One reasoning about it can be that to estimate $\hat{\sigma}_e^2$ we used $n$ observations and prior information (the initial estimates) on $q$ sires; all together, $(n + q)$ pieces of information were used and $n - r(x) = (n + q) - r(X) - q$. When using an augmented set ($q$ points with $(X,Z,y) = (0, k^{1/2} * I_q', o)$, as suggested by Harville (1983) and Askin and Montgomery (1980), to obtain "ridge regressors" with a standard statistical computing package, the obtained MSE will be exactly the same as the above REML's $\hat{\sigma}_e^2$. When using, say, PROC GLM from SAS (1982b) to obtain GLS estimators and a 1-iteration REML variance component estimates, the reasoning above is absolutely true.

2) The equation for $\hat{\sigma}_s^2$ has an appeal on its own, being close to what the estimator of an univariate random variable is. In the no information case, $\hat{s} = o = \hat{s}'\hat{s}'$, but then, $\text{tr}(C_{22}) = q*k^{-1}$ and the denominator is also $\phi = q - k*q*k^{-1}$. As sires have more evenly distributed progenies, $\text{tr}(C_{22}) \rightarrow 0$ as $n \rightarrow \infty$ and the denominator of $\hat{\sigma}_s^2 \rightarrow q$. Another point is that as $k$ increases ($h^2 \rightarrow o$), $\hat{s}'\hat{s}'$ tends to zero (more "regressed" for incomplete $\hat{h}^2$), but then the denominator also tends to zero. Note also that $\hat{\sigma}_s^2$ cannot ever be negative.

3) In most methods were sum of squares are equated to their
expectations, all expectations contain the term $\sigma^2_e$. When these expectations are solved for the $j$-th component, an automatic negative correlation is created between $\hat{\sigma}^2_j$ and $\hat{\sigma}^2_e$ (the value being near $-1/df_j$). So that as $\hat{\sigma}^2_e$ increases, so does the chance of obtaining negative $\hat{\sigma}^2_j$. This negative correlation also exists among REML estimates, as the information matrix tells us. Always when decomposing a total phenotypic variance, or any fixed quantity, the covariance between the several components will be negative. This numerical property is quite bothersome in animal breeding problems. (As an example, when based on their calves WW records, the partition of cow effects on direct (genetic) and maternal effects will always yield a negative estimate of the covariance between these effects, even if the biological expectation would be a positive covariance). REML estimates contain, embodied in their formulas, this negative association also, but to a lesser degree than in other methods. Bigger $|\hat{s}|$ values make the numerator of $\hat{\sigma}^2_e$ smaller; bigger $\hat{\sigma}^2_e$ means higher $k$ and lower $\hat{s}\hat{s}$, reducing $\hat{\sigma}^2_s$. This entangling of the effects of each estimate on the estimates of other factors and/or variance components is possibly what allows for the simultaneous solution of the whole system of unknowns (fixed and random effects and interactions and variance components), with optimum properties as described by Hartley and Rao (1967); solutions are simultaneous if considered across rounds of iteration, until convergence.
1. **Definition of REML**

Assuming normality of the random effects, ML estimates are obtained from maximizing the logarithm of the likelihood of \( y \) (the response variable); in REML only that portion of the log-likelihood which is invariant to \( E(y) \) is maximized (Thompson, 1962) and Patterson and Thompson, 1971).

Harville (1977) described REML as the estimates obtained from the maximization of the log likelihood of \( A_1 y \), a set of \( n - p \) linearly independent error contrasts, where

\[
p = \text{rank}(X) = \text{number of linearly independent fixed effects and regressions},
\]

and an example of \( A_1 \) is \( A_1 = I - X(X'X)^{-1}X' \).

\( A_1 \) must be such that \( A_1 X = 0 \).

2. **A few computation aspects**

A simple method to obtain REML estimates—apparently first suggested by Henderson (1973), and hence here called Henderson's Iterative Algorithm (HIA)—is to use the MME, using for initial \( \hat{\sigma}^2 = \hat{s}^2 \) an easily obtained estimate or prior knowledge; and compute new \( \hat{\sigma}^2 \) and \( \hat{s}^2 \) estimates. If initial and new ratios are different then repeat the process until values converge. Number of needed iterations will depend on initial values. If the number of iterations is not limited, then final estimates will not depend on initial ones. The problem with HIA is that (REML is a slow converging process) when one decides to stop iterating based on magnitude of last change in solution one can never be sure how far off the final estimate is, since the process may go for hundreds of iterations until changes reach machine-error level. Working with HIA
in the previous mentioned example given by Harville (1983), the author found that an excess of 200 iterations were needed to arrive at the machine-error level. The example was a balanced incomplete block (BIB) design and the initial estimates were from an ANOVA method (Type IV SS from PROC GLM from SAS (1982b)). This method (HIA) was thence ruled out.

Harville (1977) describes several efficient algorithms, including the Newton-Raphson's. This algorithm has been described by Jennrich and Sampson (1976, 1977, and 1978) and implemented into the BMDP package for ML and REML estimation in general mixed models. The author never used this package and possibly would not have a feel for the outcomes, but it looks like a nice alternative to be tested by someone. Further, the author wanted to test the planar rotations of Givens in the iterative algorithm.

The same BIB example given by Harville (1983) was also tested using the Common Intercept Approach (CIA), as described by Schaeffer (1983). This approach allows for the establishment of bounds on the estimated values, i.e., at each couple of iterations a converged value is found and it is possible to assert a narrow bound for the final (if iterations were performed until machine-error level) estimate. Further, it works with variance ratios (\( \alpha \)), facilitating the use of mixed model equations or the corresponding mixed model data matrix. Schaeffer (1983) does not give references for the method neither claims authorship. The method had been used by Martinez (1982) in a dairy cattle breeding problem with very good results. Schaeffer (1983) clearly describes the method as follows:

This technique was motivated from plots of variance component estimates after each iteration. The curves appeared to asymptote toward the
"converged" value. By taking the changes from round \( k \) to \( k + 1 \) from different sets of a priori parameter values, then the "converged" value could be estimated as the point of intersection of the slopes of these changes.

Step 1: Select two sets of a priori values such that one set "underestimates" the expected or anticipated "converged" values and the other set "overestimates" the anticipated "converged" values.

We are essentially attempting to "bracket" around the "converged" values.

Step 2: Perform one iteration with each of the two sets of a priori values.

Let \( \alpha_1 \) and \( \alpha_2 \) be the two sets of a priori values for some parameters, and let \( \hat{\alpha}_1 \) and \( \hat{\alpha}_2 \) be the new estimates after one iteration.

Step 3: Estimate the "converged" value.

Let \( \Delta_1 = \hat{\alpha}_1 - \alpha_1 \) and \( \Delta_2 = \hat{\alpha}_2 - \alpha_2 \), then

\[
\alpha_c = \left( \frac{\Delta_2}{\Delta_1} \right)^{-1} \left( \alpha_1 \Delta_2 - \alpha_2 \Delta_1 \right) = \text{"converged" value of } \alpha.
\]

The method works even if \( \Delta_1 < 0 \) or \( \Delta_2 > 0 \).

New sets of original values can be given by \( \alpha_1^{(2)} = \alpha_c - \Delta_1 \) and \( \alpha_2^{(2)} = \alpha_c - \Delta_2 \), or, if \( |\Delta_1| \leq |\Delta_2| \), \( \alpha_1^{(2)} = \alpha_c - \Delta_1/2 \). Notice that if values are converging, then initial values for last or next (if \( |\Delta_1| \leq |\Delta_2| \)) iteration gives the bounds for the final estimate. In this way, a iteration gives the bounds for the final estimate. In this way, a final \( \alpha_c \) or \( \hat{\alpha}^2 \) can be found with any degree of accuracy, and one does not need to ask the terrible question: "What would be my estimates if I iterated for more than 5, 10, or more rounds?"
C. Solving Least Squares Problems

Animal breeders working with large files of field records need to develop, to be proficient in their jobs, numerical and computational (besides statistical) skills almost to the level of mastery. No Mathematics or Computer Science departments, in any university, will design courses for such a small clientele. Animal breeding sections could provide those, or maintain indication for sequential reading and worked examples, or promote round tables among students to increase their technological overlapping or adopt an learn-as-you-go (doing mistakes) posture.

Several text books exist that deal with the subject at any theoretical depth. Not that many texts are directed to a practitioner. Seber (1977) describes in a 40-page chapter several computational techniques for fixed models, with properties of each method and an overall comparison of methods. Applied and readable, Seber (1977) is a very good initial reference. Rice (1981) and Chambers (1977) cover the material in much more detail, keeping on the applied side. Lawson and Hanson (1974) should be a required reading for animal breeders; their text (which also lends its title to this subsection heading) presents several ways to solve animal breeding problems with algorithms and Fortran programs and subroutines. Adey and Brebbia (1983) is like an abridged version of Lawson and Hanson (1974), giving also some matrix operation Fortran subroutines. Dongarra et al. (1979) describe the problems, the approach and the theory behind a extensive collection of efficient Fortran subroutines (also given) for solving systems of linear equations and least squares problems. Maindonald (1984) is a text in the
same category as the one from Lawson and Hanson (1974), but it is of more help as a self-teaching guide given the intelligent use of completely worked algebraic examples. Maindonald (1984) gives a set of programs and subroutines in BASIC, designed for microcomputer users. To these users, Shoup (1983) and Miller (1981) are very useful sources.

This study relied heavily on the use of Givens (1954) planar rotations to obtain estimates of the effects and of variance components. Seber (1977) raised one's curiosity for this method and gives its basic advantages but only states the computational formulas. To understand how the formulas apply, an example is developed in what follows, comparing results with other methods.

1. Constructing the records of the smallest possible example

The example uses three observations showing the linear effect of AOC on WW.

What is given is a birth weight (BW) = \( b_0 = 731 \) lb, ADG of 21 lb/day, an average AOC of 205 days, and an average WW of \( 4831 \) lb.

The observations are as follows:

<table>
<thead>
<tr>
<th>Obs</th>
<th>x(_0) = Intercept</th>
<th>x(_1) = AOC</th>
<th>( \hat{y} - BW )</th>
<th>( \hat{y} )</th>
<th>e</th>
<th>y = WW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>185</td>
<td>370</td>
<td>443</td>
<td>+26</td>
<td>469</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>205</td>
<td>410</td>
<td>483</td>
<td>-52</td>
<td>431</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>225</td>
<td>450</td>
<td>523</td>
<td>+26</td>
<td>549</td>
</tr>
</tbody>
</table>

where

\[ 1'e = 0 \]

\[ e'e = 4056 \]

\[ 1'y = 1449 = x_0'y \]
1'\mathbf{x}_1 = 615
\mathbf{x}_1\mathbf{x}_1 = 126875
\mathbf{x}_1'\mathbf{y} = 298645
1'\mathbf{x}_0 = 3 = \mathbf{x}_0'\mathbf{x}_0

\mathbf{A} = [\mathbf{X} \ \mathbf{y}] = \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \mathbf{y} \end{bmatrix} = \begin{bmatrix} 1 & 185 & 469 \\ 1 & 205 & 431 \\ 1 & 225 & 549 \end{bmatrix}

2. **Direct solution using NE and the inverse of \mathbf{X}'\mathbf{X}**

\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y} = \begin{bmatrix} 3 & 615 \\ 615 & 126875 \end{bmatrix} \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \end{bmatrix} = \begin{bmatrix} 1449 \\ 298645 \end{bmatrix}

\mathbf{b} = \frac{1}{2400} \begin{bmatrix} 126875 & -615 \\ -615 & 3 \end{bmatrix} \begin{bmatrix} 1449 \\ 298645 \end{bmatrix}

\hat{\mathbf{b}} = \begin{bmatrix} 73.000 \\ 2.000 \end{bmatrix}

\mathbf{e}'\mathbf{e} = \mathbf{y}'\mathbf{y} - \hat{\mathbf{b}}'\mathbf{X}'\mathbf{y} = 707123 - 703067 = 4056

3. **Direct solution using the sweet operator**

This matrix decomposition algorithm is an hybrid from the Sequential Cholesky Decomposition Algorithm (SCDA) as described by Maindonald (1984) (which contains concepts of outer products and their sums as described by Schaeffer (1976) and Searle (1982)), and the SWEEP operator as described by Goodnight (1979). This operator was so named by Guitou (1984).

This algorithm reduces \mathbf{A}'\mathbf{A} to an upper triangular system \mathbf{T}, such
that $T'T = A'A$ and $T = \begin{bmatrix} U & z \\ \circ & r \end{bmatrix}$ where $U$ is upper triangular, $z$ is a column vector and $r$ is a scalar.

Let $T' = [t_0, t_1, t_2]$. The first step involves finding $t_0$ such that $A'A - t_0 t_0'$ has zeros in the first row and column.

$$A'A = \begin{bmatrix} X' \\ y' \end{bmatrix} \begin{bmatrix} X & y \end{bmatrix} = \begin{bmatrix} x'x & x'y \\ y'x & y'y \end{bmatrix} = \begin{bmatrix} 3 & 615 & 1449 \\ 615 & 126875 & 298645 \\ 1449 & 298645 & 707123 \end{bmatrix}$$

$$t_0 t_0' = \begin{bmatrix} t_{00} \\ t_{01} \\ t_{02} \end{bmatrix} \begin{bmatrix} t_{00} & t_{01} & t_{02} \end{bmatrix} = \begin{bmatrix} t_{00}^2 & t_{00} t_{01} & t_{00} t_{02} \\ t_{01} t_{00} & t_{01}^2 & t_{01} t_{02} \\ t_{02} t_{00} & t_{02} t_{01} & t_{02}^2 \end{bmatrix}$$

$$t_0 t_0' = \begin{bmatrix} \sqrt{3} \\ 205\sqrt{3} \\ 483\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 205\sqrt{3} & 483\sqrt{3} \end{bmatrix} = \begin{bmatrix} 3 & 615 & 1449 \\ 615 & 126075 & 297045 \\ 1449 & 297045 & 699867 \end{bmatrix}$$

$$A'A - t_0 t_0' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 800 & 1600 \\ 0 & 1600 & 7256 \end{bmatrix}$$

$$t_1 t_1' = \begin{bmatrix} 0 \\ \sqrt{800} \\ 2\sqrt{800} \end{bmatrix} \begin{bmatrix} 0 & \sqrt{800} & 2\sqrt{800} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 800 & 1600 \\ 0 & 1600 & 3200 \end{bmatrix}$$
Obviously $t_{22} = \sqrt{4056}$ such that $A'A = T'T = t_0 t_0' + t_1 t_1' + t_2 t_2'$ and

$$T = \begin{bmatrix} U & z \\ 0' & r \end{bmatrix} = \begin{bmatrix} t_0' \\ t_1' \\ t_2' \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 205\sqrt{3} & 483\sqrt{3} \\ 0 & \sqrt{800} & 2\sqrt{800} \\ 0 & 0 & \sqrt{4056} \end{bmatrix}.$$

Notice that $U'U = X'X$ and hence $U$ is the Cholesky decomposition of $X'X$. Describing the Cholesky decomposition as a sum of outer products makes very clear the word "decomposition," to follow the algebraic definition of the method, and to postulate alternative algorithms.

Notice that solutions are obtained directly from $T$, by backsubstitution, without need to solve another triangular system as in the original Cholesky.

$$b_1 = z_2/u_{22} = t_{12}/t_{11} = 2\sqrt{800}/\sqrt{800} = 2.$$

$$b_0 = (z_1 - b_1 u_{12})/u_{11} = (t_{02} - b_1 t_{01})/t_{00}$$

$$= (483\sqrt{3} - 410\sqrt{3})/\sqrt{3} = 73.$$

In common with the SWEEP operator, notice that the SWEET operator promptly gives:

$$SSE = \hat{\mathbf{e}}' \hat{\mathbf{e}} = r'r = r^2 = t_{22}' t_{22} = t_{22}^2 = 4056$$
\[8'X'y = z'z = SS_{\text{model}} = (483\sqrt{3})^2 + (2\sqrt{800})^2 = 699867 + 3200 = 703067\]

Notice that, using Searle (1971) reduction notation:

\[R(b_0) = t_{02}^2 = z_1^2 = (483\sqrt{3})^2, \quad \text{and}\]
\[R(b_1/b_0) = t_{12}^2 = z_2^2 = (2\sqrt{800})^2.\]

4. Direct solution using Givens planar rotations

The objective of Givens rotations is to transform any real rectangular matrix into an upper triangular system. Given a data matrix \(A\), with dimensions \(n \times p\), if \(n > p\), least square solutions are obtained from the upper triangular system by backsubstitution.

What follows is one of the most clear accounts of the Givens transformations as presented by Seber (1977).

Consider a matrix \(G\), with dimensions \(n \times n\), being an identity matrix with the exception of rows and columns \(j\) and \(k\), where

\[
G(jk) = \begin{bmatrix}
cos\theta & \sin\theta \\
-sin\theta & cos\theta
\end{bmatrix}
\]

for any \(j\)-th and \(k\)-th rows and columns.

\(G(j,k)\) consists of planar rotations such that when \(G\) is applied to \(A\), it simply rotates each of the two dimensional column vectors formed by the \(j\)-th and \(k\)-th rows through an angle \(\theta\), leaving the remaining rows unchanged. In particular, we can choose \(\theta\) so that two rows

row \(j\): 0, ..., 0, \(r_i\), \(r_{i+1}\), ..., \(r_m\), ..., \(r_p\)
row \(k\): 0, ..., 0, \(l_i\), \(l_{i+1}\), ..., \(l_m\), ..., \(l_p\)
become (orthogonal at column i-th, since $l'_i = 0$)

\[
\begin{align*}
\text{row j': } & 0, \ldots, 0, r'_i, r'_{i+1}, \ldots, r'_m, \ldots, r'_p \\
\text{row k': } & 0, \ldots, 0, 0, l'_{i+1}, \ldots, l'_m, \ldots, l'_p
\end{align*}
\]

where

\[
\begin{align*}
   r'_i &= r_i^2 + l_i^2 \\
   c &= \cos\theta = \frac{r_i}{r'_i} \\
   s &= \sin\theta = \frac{l_i}{r'_i} = \sqrt{1 - c^2} \\
   r'_m &= cr_m + sl_m \\
   l'_m &= -sr_m + cl_m
\end{align*}
\]


When working with classificatory models, $X$ is a sparse incidence matrix and computations can be greatly reduced by exploiting the presence of zeros. When $r_i = 0$ and $l_i > 0, \theta = \pi/2$, and $G$ simply interchanges the rows, with a change in sign in row $k'$, namely,

\[
\begin{align*}
\text{row j': } & 0, \ldots, 0, l_i, l_{i+1}, \ldots, l_p \\
\text{row k': } & 0, \ldots, 0, 0, -r_{i+1}, \ldots, -r_p
\end{align*}
\]
The sign of row j' is also changed if \( l_1 < 0 \).

If \( r_1 = l_1 = 0 \), then \( \theta = 0 \) and there is no change.

Elements of X or A below the diagonal can be annihilated by a series of Givens transformations. If \( t \) elements of A are to be zeroed, then the series of transformations can be represented by a matrix \( Q \), \( n \times n \), where

\[
Q' = G_t G_{t-1} \ldots G_2 G_1 \quad \text{and} \quad Q'Q = QQ' = I.
\]

Hence, \( Q' \) is an orthogonal matrix; it is independent of the order of the G's (\( Q' \) is always the same, independently of the order that the rotations are performed although the total number of rotations depends on the specific order the rotations occur); and when postmultiplies by A gives

\[
Q'A = \begin{bmatrix}
U & z \\
0' & r
\end{bmatrix}
\]

where \( U \), \( z \) and \( r \) are as before.

This reductions can be done in two ways (Wilkinson, 1965, pages 239-240):

1. “Transform the first row successively with the second, third, fourth, etc., row so as to annihilate the last \( n-1 \) elements of the first column; transform the second row successively with the third, fourth, etc., row so as to annihilate the last \( n-2 \) elements in the second column; in general, transform the \( j \)-th row successively with \((j + 1)\)-th, \((j + 2)\)-th, \ldots, \( n \)-th row so as to annihilate the last \( n - j \) elements in the \( j \)-th column.” [This way is suggesting a mode to perform a like absorption of a big classificatory factor. Say that in a model similar to the one in this study we have CG, and observations are sorted by contemporary groups. Then we know that for each CG (sequentially read and operated over) we only need to worry about one column for this factor. The first row for each CG will be exactly the same as the
corresponding row in the NE but divided by the square root of the number of observations in that CG. This way deserves further studies since it may greatly reduce number of operations].

2. "At the k-th step (k=1, 2, ..., n − 1) transform the (k + l)-th row with each of the rows above, beginning with the first, so that the first (k + l) rows of X or A are reduced to upper triangular form." This is the Sequential Accumulation process, as described by Lawson and Hanson (1974), that was used in this study.

The first method requires the storing of all of A in the memory whereas the second method can be applied to each row of A as it is read in.

Seber (1977) discusses two distinct advantages of Givens transformations: First, A can be processed one row at a time (the second method above) and second, zeros already present in A (even more if the structure of the data is known or modified by sorting) are readily exploited to reduce arithmetic cost.

The importance of these advantages is threefold (Gentleman, 1973):

1. In regression and particularly factorial designs, A may be too large to keep in high-speed storage, and therefore must be generated or fetched as required. Since there are generally many more rows than columns, processing by rows minimizes the high-speed storage and/or transfers to backing store (I/O operations) required. Also, it is often more natural to generate or fetch A by rows. For example, in regression problems each row corresponds to one observation in the model and we may wish to update A by adding more observations (rows)
((Chambers, 1971) used Givens rotations to update systems based on the NE and presents ways to add and delete observations); in factorial designs it is usually more convenient to generate rows than columns (rows of the incidence matrix can be generated as each observation is read and the rotations be applied, even without knowing the total number of levels in each factor; this actually reduces the number of operations needed for the rotations since the creation of columns according to the rows forces $A$ to acquire a banded pattern).

2. For classificatory models, $A$ frequently has a large number of zeros. Using Givens transformations and merely exploiting zeros in obvious ways (even without tailoring the computation to the particular matrix structure), Gentleman (1973) reports that reductions in computing cost of up to 70% have been measured for the analysis of some conventional unbalanced factorial designs.

3. Even after a model has been analyzed and $A$ or $A'^t$ have been reduced to an upper triangular system, it is often necessary to include further observations (as a side issue, new columns which enlarge the original model can be added directly to $U$ by using the Householder transformations) or impose constraints or use extraneous information as described by Goldberger (1964) in a NE setting or "augment" observations to obtain GLS estimators as proposed by Harville (1983). Irrespective of how the reduction was done (or if reductions were operated over the data matrix or over its summary form, the NE), the new matrix
can be retriangularized by Givens transformations, taking advantage of the structure already present.

Another advantage of methods which work directly on the data matrix (Householder and Givens) is discussed in depth and with a practical sense by Lawson and Hanson (1974). These authors show that methods based on forming and solving the normal equations typically require only about half as many operations as the Householder algorithm (and much less so when compared with the Givens rotations), but these operations must be performed using precision $k^2$ to satisfactorily encompass the same class of problems as can be treated using data matrix methods with precision $k$. Lawson and Hanson (1974) devote several chapters to precision related problems and define $k$ as the "relative precision of the normalized floating point arithmetic." Simply stated, the number $k$ is the smallest positive number (of decimal digits) needed to keep rounding errors within bounds (without propagation) for a given number and type of operations. Animal breeders working with pure classificatory models do not need to fear rounding errors when some precautions are taken (like deviating $y$ from some approximate average) and checks are included in the program, even when working with millions of records. But if the model contain regressors (like $CA2 = AOC^2$) or spline functions (which do not allow for deviating variables from approximate averages - what rules out the package developed by Harvey (1977)), this problem may become serious if number of observations is in the thousands. Even being much more precise than NE methods, the Gl subroutine of Lawson and Hanson (1974) has built-in operations to further reduce chances of propagating errors. Even the
small example used in this section, with only three observations, can
give an idea on how this problem may become crucial by looking at the
magnitude of numbers in the initial and intermediary steps of the three
methods developed.

Of course there are not only advantages to the Givens method. Com-
putational costs are higher than when using any other method based on the
NE. This explains why the method is still scarcely known and rarely used
even after more than 30 years of its formalization by Givens (1954). For
common least squares problems where the observations will furnish infor-
mation only once about a fixed set of parameters or tests, Givens rota-
tions should not be used. Marginal areas (away from the bulk of statisti-
cal computations) where the method can compete will be presented in the
next section and in chapter VIII. Number of calculation are far bigger
in Givens than in, say Cholesky method. Lawson and Hanson (1974) com-
pared operation counts for various least squares computational methods.
The comparisons are based on asymptotic number of operations, where an
operation is a multiplication or division plus an addition. To form the
normal equations $np^2/2$ operations are needed (considering a pure regres-
sion model) but for classificatory models only counters are needed and
the number of additions is only a fraction of the above. The Cholesky
solution of the normal equations require $p^3/6$ operations. The Givens
method applied on the data matrix requires $2np^2$ operations with the
addition of $np$ square roots. In the most favorable case, where $p = n$,
Givens method requires 12 times more operations than the Cholesky method,
even without counting the number of square roots. The prognosis is that
the above ratio will get even larger since many more people are working
Gentleman (1973) presented a modification of Givens method which avoids the computation of square roots and takes only three quarters as many multiplications, and latter this author gave an algorithm and program perform these modifications (Gentleman, 1974). Gentleman's method is especially suited for the WLS, but the author still cannot clearly understand how and why the method works, even after the description given by Seber (1977) and the availability of the program (Gentleman, 1974). Another point considered was the finding of George and Heath (1980) that "in practice, the modified scheme yields a far smaller reduction in actual running time than the reduction in operation count would imply, particularly since rescaling is often necessary to ensure stability."

Other sources from where some concepts about Givens rotations were obtained are Scott et al. (1982), Heath (1982), Bjorck (1976), and Gill and Murray (1976).

Given the basic theory, advantages and problems, the same original problem from this section will be solved using Givens rotations, in long-hand algebra, to enhance understanding on how the method works.

What will be done is to transform \( A \) into an upper triangular matrix by a series of Givens' planar rotations, or

\[
Q'A = Q'[X \ y] = \begin{bmatrix} U & z \\ 0' & r \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & z_1 \\ 0 & u_{22} & z_2 \\ 0 & 0 & r \end{bmatrix}
\]

The upper triangular system is given on the right hand side of what
follows and it is a null matrix at the beginning.

a. Read 1-st record:

\[ r = [ 0 \ 0 \ 0 ] \]

\[ l = [ 1 \ 185 \ 463 ] \]

Since \( r = 0 \), interchange \( r \) and \( l \).

\[
\begin{bmatrix}
1 & 185 & 469 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

b. Read 2-nd record:

\[ r = [ 1 \ 185 \ 469 ] = [ r_1 \ r_2 \ r_3 ] \]

\[ l = [ 1 \ 185 \ 463 ] [ 1 \ 1 \ 1 ] \]

b.a. annihilate \( l_1 \):

\[
r_1' = \sqrt{r_1^2 + l_1^2} = \sqrt{2} \\
c = \frac{r_1}{r_1'} = \frac{1}{\sqrt{2}} \\
s = \frac{l_1}{r_1'} = \frac{1}{\sqrt{2}} = 1 - c^2 \\
l_1' = 0 \\
[ r_2' \ r_3' ] = [ c \ s ] \begin{bmatrix} r_2 & r_3 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = [ c r_2 + s l_2 \ c r_3 + s l_3 ]
\[
\begin{bmatrix}
  r_2' & r_3'
\end{bmatrix}
= \begin{bmatrix}
  185 + 205 & 469 + 432
\end{bmatrix}
= \begin{bmatrix}
  390 & 900
\end{bmatrix}
\]

\[
\begin{bmatrix}
r_1' & r_2' & r_3'
\end{bmatrix}
= \begin{bmatrix}
  \sqrt{2} & 390 & 900
\end{bmatrix}
\begin{bmatrix}
  0 & 0 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
l_1' & l_3'
\end{bmatrix} = \begin{bmatrix}
  -s & c
\end{bmatrix} \begin{bmatrix}
  r_2 & r_3
\end{bmatrix}
= \begin{bmatrix}
  -sr_2 + cl_2 & -sr_3 + cl_3
\end{bmatrix}
\]

\[
\begin{bmatrix}
l_1' & l_3'
\end{bmatrix} = \begin{bmatrix}
  -185 + 205 & -469 + 432
\end{bmatrix}
= \begin{bmatrix}
  20 & -38
\end{bmatrix}
\]

\[
l_1' = l_1' \begin{bmatrix}
l_1 & l_2 & l_3
\end{bmatrix}
= \begin{bmatrix}
  0 & 20 & -38
\end{bmatrix}
\]

b.b. Annihilate \( l_2' \):

Since the second row of the triangular system is null, no operation is needed and \( l_1' \) becomes the second row in that system. After the interchange of \( l_1' \) and \( r \), \( l = 0 \) and so the operations with the second observation are ready.

\[
\begin{bmatrix}
  \sqrt{2} & 390 & 900
\end{bmatrix}
\begin{bmatrix}
  0 & 20 & -38
\end{bmatrix}
\begin{bmatrix}
  0 & 0 & 0
\end{bmatrix}
\]

c. Read 3rd record:

\[
l = \begin{bmatrix}
  1 & 225 & 549
\end{bmatrix}
\]
c.a. Since $l_1 \neq 0$, read first row from the triangular system and
annihilate $l_1$:

$$r = \begin{bmatrix} \sqrt{2} & 390 \sqrt{2} & 900 \sqrt{2} \end{bmatrix}$$

$$l = \begin{bmatrix} 1 & 225 & 549 \end{bmatrix}$$

$$r'_1 = \sqrt{r_1^2 + l_1^2} = \sqrt{2^2 + 1^2} = \sqrt{3}$$

$$c = \frac{r_1}{r'_1} = \frac{\sqrt{2}}{\sqrt{3}}$$

$$s = \frac{l_1}{r'_1} = \sqrt{1 - c^2} = \frac{1}{\sqrt{3}}$$

$$l'_1 = 0$$

$$\begin{bmatrix} r'_2 & r'_3 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \cdot \frac{390}{\sqrt{3}} + 225 \sqrt{2} \cdot \frac{900}{\sqrt{3}} + 549 \sqrt{2} \cdot \frac{549}{\sqrt{3}} \end{bmatrix}$$

$$r'_1 = \begin{bmatrix} \sqrt{3} \cdot \frac{615}{\sqrt{3}} + 1449 \sqrt{3} \frac{1449}{\sqrt{3}} \end{bmatrix}$$

$$\begin{bmatrix} \sqrt{3} \cdot \frac{615}{\sqrt{2}} - 38 \sqrt{2} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 20 \sqrt{2} & -38 \sqrt{2} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1_2 & l_3 \end{bmatrix} = \begin{bmatrix} -390 \sqrt{2} \cdot \frac{225}{\sqrt{3}} - 900 \sqrt{2} \cdot \frac{549}{\sqrt{3}} \end{bmatrix}$$
\[ 1' = \begin{bmatrix} 0 & 60 \\ \sqrt{6} & 198 \end{bmatrix} \]

c.b. Since \( l_2' \neq 0 \), read second row from the triangular system and annihilate \( l_2' \):

\[ r = \begin{bmatrix} 0 & 20 \\ \sqrt{2} & -38 \end{bmatrix} \]

\[ l' = 1 = \begin{bmatrix} 0 & 60 \\ \sqrt{6} & 198 \end{bmatrix} \]

\[ r_2' = \sqrt{r_2^2 + l_2^2} = \sqrt{\frac{20^2}{2} + \frac{60^2}{6}} = 20\sqrt{2} \]

\[ c = \frac{r_2}{r_2'} = \frac{20/\sqrt{2}}{20\sqrt{2}} = \frac{20}{40} = \frac{1}{2} \]

\[ s = \sqrt{1 - c^2} = \sqrt{3} = \frac{l_2}{r_2'} = \frac{60/\sqrt{6}}{20\sqrt{2}} = \frac{3\sqrt{3}}{2} \]

\[ l_2' = 0 \]

\[ r_3' = \begin{bmatrix} c & s \\ r_3 \\ l_3 \end{bmatrix} = \begin{bmatrix} -38 + \frac{\sqrt{3} \cdot 198}{2\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -19 + 99 \end{bmatrix} = \begin{bmatrix} 80 \end{bmatrix} = 40\sqrt{2} \]

\[ l_3' = \begin{bmatrix} -s & c \\ r_3 \\ l_3 \end{bmatrix} = \begin{bmatrix} \frac{(-\sqrt{3})(-38)}{2\sqrt{2}} + \frac{198}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} 156 \end{bmatrix} = 26\sqrt{6} \]
\[ r' = \begin{bmatrix} 0 & 20\sqrt{2} & 40\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 615 & 1449 \\ \sqrt{3} & 0 & \sqrt{3} \\ 0 & \sqrt{800} & 2\sqrt{800} \end{bmatrix} \]

\[ 1' = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \sqrt{4056} \end{bmatrix} \]

c.c. Since \( l_j \neq 0 \), it has to be annihilated but, since the third row of the upper triangular system is null, a simple interchange completes the rotations.

The final system then is:

\[
Q'A = \begin{bmatrix} u_{11} & u_{12} & z_1 \\ 0 & u_{22} & z_2 \\ 0 & 0 & r \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 615 & 1449 \\ \sqrt{3} & 0 & \sqrt{3} \\ 0 & \sqrt{800} & 2\sqrt{800} \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 205\sqrt{3} & 483\sqrt{3} \\ \sqrt{3} & 0 & \sqrt{800} \\ 0 & 0 & \sqrt{4056} \end{bmatrix}
\]

Looking at the results obtained with the sweet operator in previous subsections, it is evident that \( T = Q'A \). Hence, all the properties of the systems, solutions, interpretations and sums of squares are the same as before.

Differences among methods are very clear in this example:

1) Number of operations with Givens method is much larger;
2) Magnitude of numbers are never as small or as large as in the "NE" methods.
3) All operations with the Givens method were done involving only one row of data (1) and one row from \( T = Q'A (r) \); nowhere more than one row from \( T \) was in the "high speed memory."
D. Using Givens Rotations to Obtain GLS Estimates of the Effects in the
Model and to Obtain REML Variance Component Estimates

In this study, GLS estimates of fixed and random effects and REML estimates of sire and error variances were obtained by the following sequence of steps:

1. Absorb CG effects into the regressors and sire effects, directly on the data matrix, as records are read in.
2. Apply Givens rotations to the centered records as if the model was fixed, forming the upper triangular matrix $T$.
3. Augment data matrix with $[0 \ P \ 0]$, where $PP = G^{-1}$ and $G = \text{Var}(s,s') * \sigma_e^2$, and rotate augmented records against corresponding rows of $T$, giving $T^*$.
4. Backsolve, obtaining $1$-REML $\hat{\sigma}_e^2$ and GLS estimates of fixed and random effects.
5. Find the inverse of $T^*$ and, using sire GLS estimates and information from $G^{-1}$, calculate REML $\hat{\sigma}_s^2$ and a new $G$, $G^*$ say. If $G^* - G$ is larger than desired bound, use CIA approach and return to step 3, iterating until $G^* - G$ is as close to zero as desired.

Credits for each one of the above steps are due to different authors. Instrumental sources were, for each of the steps, the following:

1. Searle (1982);
2. Seber (1977) and Lawson and Hanson (1974);
Harville (1983); Lawson and Hanson (1974) and Harville (1983); and Schaeffer (1983) and Harville (1983).

Most of the above steps have been described in some detail in previous sections on this chapter. In this section, some further treatment is given, unifying them in the form they were used in this study.

1. Absorption of CG effects in the data matrix

When using the NE approach, animal breeders obtain a great reduction in computing costs by absorbing CG before obtaining solutions for other effects of interest. This and other techniques and "tricks" were and are being learned, making the NE approach even more economical in comparison with a data matrix approach.

Based on ideas from Searle (1982) the same technique was used on the data matrix. Later, the author found out that this is what is called as "Model A" by Harville (1984a). Now, with some hindsight, the author is not absolutely sure that the absorption of CG resulted in large benefits, neither that absorption is the best technique to handle the CG. This topic will be discussed in Chapter X.

If to the Eq. (35) sire effects are added, and the resulting equation is expressed in matrix form:

\[ y = Xb + Zs + Wr + e \]  (38)

where:

- \( b \) are the CG effects, added to the constant,
- \( s \) are the sire effects (treated as fixed, so far),
X, Z are incidence matrices,
W is the concomitant variables matrix,
\( r \) are the 17 regression coefficients,
y is the vector of WW, and
e is the vector of residuals with \( e \sim \text{IID}(0, \sigma_e^2) \).

If \( b \) has a large number of elements, they can be absorbed, when using the NE approach as follows:

\[
\begin{bmatrix}
X'X & X'Z & X'W \\
Z'X & Z'Z & Z'W \\
W'X & W'Z & W'W
\end{bmatrix}
\begin{bmatrix}
b^o \\
s^o \\
r^o
\end{bmatrix} =
\begin{bmatrix}
x'y \\
z'y \\
w'y
\end{bmatrix}
\tag{1}
\]

Solving for \( b^o \) in Eq. (1) of NE1:

\[
b^o = (X'X)^{-1}(x'y - X'Zs^o - X'Wr^o)
\]

Substituting \( b^o \) in (2) of NE1 for \( b^o \):

\[
Z'zs^o - Z'x(x'x)^{-1}x'zs^o + Z'wr^o - Z'x(x'x)^{-1}x'wr^o =
\]

\[
z'y - Z'x(x'x)^{-1}x'y
\]

and letting \( M = I - x(x'x)^{-1}x' \):

\[
Z'Mzs^o + Z'Mwr^o = Z'My
\tag{4}
\]

Substituting \( b^o \) in (3) of NE1 for \( b^o \):

\[
w'zs^o - w'x(x'x)^{-1}x'zs^o + w'wr^o - w'x(x'x)^{-1}x'wr^o =
\]

\[
w'My - W'x(x'x)^{-1}x'y
\]

\[
w'Mzs^o + w'Mwr^o = W'My
\tag{5}
\]
Putting (4) and (5) together in NE2:

\[
\begin{bmatrix}
Z'MZ & Z'MW \\
W'MZ & W'MW
\end{bmatrix}
\begin{bmatrix}
s^0 \\
r^0
\end{bmatrix} =
\begin{bmatrix}
Z'My \\
W'My
\end{bmatrix}
\]

Solutions \(s^0\) and \(r^0\) are the same in NE1 and NE2; size of NE2 is much smaller than of NE1; and \(Z'MZ\) is not a diagonal matrix as \(Z'Z\) was.

\(M\) is a symmetric idempotent (SIP) matrix since \(M = M' = MM = M'M = M'M'\). \(M\) is a "centering matrix" described by Searle (1982), of dimensions \(n \times n\).

When operating on columns of \(X, W,\) or \(y\), the effect of \(M\) is to deviate each variable from its CG mean.

If we define:

\[
\begin{align*}
\tilde{Z} &= MZ, \\
\tilde{W} &= MW, \\
\tilde{y} &= My, \text{ and} \\
\tilde{e} &= Me,
\end{align*}
\]

and write model (38) as:

\[
\tilde{y} = \tilde{Z}s + \tilde{W}r + \tilde{e} \quad ("\text{Model A" of Harville (1984)}) \quad (39),
\]

then we can write NE3 as:

\[
\begin{bmatrix}
\tilde{Z}'Z & \tilde{Z}'W \\
\tilde{W}'Z & \tilde{W}'W
\end{bmatrix}
\begin{bmatrix}
s^0 \\
r^0
\end{bmatrix} =
\begin{bmatrix}
\tilde{Z}'\tilde{y} \\
\tilde{W}'\tilde{y}
\end{bmatrix}
\]

or, by definition:

\[
\begin{bmatrix}
Z'M'MZ & Z'M'MW \\
W'M'MZ & W'M'MW
\end{bmatrix}
\begin{bmatrix}
s^0 \\
r^0
\end{bmatrix} =
\begin{bmatrix}
Z'M'My \\
W'M'My
\end{bmatrix}
\]

but, since \(M\) is SIP, \(NE4 = NE2\), the solutions for \(s\) and \(r\) obtained in
model (39) are the same as obtained in model (38).

Further, as demonstrated in the last section, solutions from \( T = Q'A \), the upper triangular system resulting from Givens rotation on the data matrix, are the same least squares solutions from the NE.

The conclusion is that what is accomplished with the absorption process in the NE mode can also be done using givens rotations on a "Model A" data matrix.

The algorithm to perform this absorption will be given in the next chapter.

A question that can be raised (as the author did) is that if the same problem which happened when using Harvey's LSMMML package (Harvey, 1977) for a model with grafted polynomials might not happen again when absorbing CG.

The package, to reduce rounding errors or chances of machine under- or overflow, requires information on the (approximate) means of all responses and covariates, and subtracts that mean from the respective variable's value for each observation. The records are "recoded" before the analysis begins. In most cases this technique is recommended, but not if the model contains grafted polynomials.

As an example, SS3 was formed by setting its value to \((\text{BIRTHDATE} - 245)^3\) if \(\text{BIRTHDATE}\) was greater than 170 and less than 245, and to 0 otherwise. The mean for all SS3 is \(-4849.8\). If this value is subtracted from SS3, this will leave the respective regression coefficient without any meaning. Weird results were found in this analysis when using the LSMMML package and the manual (Harvey, 1977) does not give any alert signal for this problem. Before beginning the analysis, this procedure
causes a change in the model. At first sight, this technique can be interpreted also as a "centering" or "absorption of the μ-equation" valid procedure. But the Z matrix is not being centered around an overall mean. A word of caution to users should be spelled out or an option given to not recode the variables.

The absorption of CG effects (or even of the μ-equation, if it were the case) keeps the model unchanged and it is an artifact to obtain the same solutions by an "easier" way.

When comparing and checking these two procedures new (to the author) concepts were envisaged. Consider two sires, A and B, which have 10 and 20 calves, respectively, in a given CG. Say that CG were absorbed and the WW "deviated" record for a son of sire A is 18 Kg. The row of Z = MZ corresponding to this calf is [0.65 - 0.65]. If the model being used is a simple two way (fixed) one, then it is clear that this record is contributing +12 Kg and -12 Kg to the sum of the effects of sires A and B, respectively. Maybe because of the definition of incidence matrices, the author had the idea that a calf from sire A only "contained information" about sire A and none about sire B, even more in a fixed model with sires independent (unrelated). But really, the information contained in each observation is halved: half to furnish information about that observation's level and half partitioned among the other levels of the factor in proportion to the total number of observations on each of the other levels. Working with a 3 or more sires example allows for this to be seen. The proportion of the "deviated" record that will be credited to the observations level is the complement to one of the ratio of the number of observations from that level to the total number of observa-
tions. The same proportion is debited to the sum of the effects of the other levels. This is related to the characteristic of the least squares method of estimating the effects of each of the levels as if they all had the same number of observations (equal information assumption).

Maybe some misconceptions could be corrected if the idea that no effect is absolute, but it is always relative to the other levels of the factor, were emphasized more. All the technical jargon around the idea of "estimability" is not helping to reduce the problem.

Rows and columns of $\mathbf{Z} = \mathbf{MZ}$ sum to zero, the same property shared by $\mathbf{Z}'\mathbf{MZ}$ in the NE after absorption of CG. In the data matrix, this property illustrates clearly how the information contained on each observation is weighted by the number of calves, partitioned among sires and used to estimate sire effects. In the NE, this property is said to be "an adjustment for the competition among sires," but it is difficult to envisage much further than a set of simultaneous equations. In any other summarizing processes, some information is lost when the data matrix is condensed into its NE form.

Basic ideas of statistics like the above are too simplistic to appear in any statistics text but they would be helpful to a "field" animal breeder working to become a more "scientific" animal breeder.

What follows is a small example of the absorption process of the CG effects in the data matrix. The example is helpful to justify the statements made above. Conclusions drawn from a CG with 30 WW records, two sires, A and B, with 10 and 20 progeny each are the same as from a CG with 3 WW records with progeny sizes 1 and 2, respectively.
Data matrix A: \[A = (X \ Z \ y)\]

<table>
<thead>
<tr>
<th>OBS</th>
<th>CGI</th>
<th>CG2</th>
<th>SIRE A</th>
<th>SIRE B</th>
<th>SIRE C</th>
<th>WW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>188</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>166</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>156</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>189</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>183</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>164</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>161</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>178</td>
</tr>
</tbody>
</table>

Given these observations, the normal equations for the two way model, without interaction, is:

\[
\begin{bmatrix}
X'X & X'Z \\
Z'X & Z'Z
\end{bmatrix}
\begin{bmatrix}
h \\
s
\end{bmatrix}
= 
\begin{bmatrix}
3 & 0 & 1 & 2 & 0 \\
0 & 5 & 2 & 0 & 3 \\
1 & 2 & 3 & 0 & 0 \\
2 & 0 & 0 & 2 & 0 \\
0 & 3 & 0 & 0 & 3
\end{bmatrix}
\begin{bmatrix}
h_1 \\
h_2 \\
s_A \\
s_B \\
s_C
\end{bmatrix}
= 
\begin{bmatrix}
510 \\
875 \\
560 \\
322 \\
503
\end{bmatrix}
= 
\begin{bmatrix}
X'y \\
Z'y
\end{bmatrix}.
\]

Absorption of \(h\) \((h_i = \mu + c_g_i)\) yields:

\[
Z'MZs = Z'Zs = \begin{bmatrix}
28/15 & -2/3 & -6/5 \\
-2/3 & 2/3 & 0 \\
-6/5 & 0 & 6/5
\end{bmatrix}
\begin{bmatrix}
s_A \\
s_B \\
s_C
\end{bmatrix}
= 
\begin{bmatrix}
40 \\
-18 \\
-22
\end{bmatrix}
= Z'My = 
\begin{bmatrix}
Z'M'My = Z'y
\end{bmatrix}.
\]

Working with observations from the CGI:

<table>
<thead>
<tr>
<th>OBS</th>
<th>CGI</th>
<th>SIRE A</th>
<th>SIRE B</th>
<th>WW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>188</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>166</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>156</td>
</tr>
</tbody>
</table>

\[\begin{bmatrix}
3 \\
1
\end{bmatrix}
= \Sigma = h_1 NE
\]

\[\begin{bmatrix}
1/3 \\
2/3
\end{bmatrix}
= \text{"CGI mean vector."}
\]

Now, deviating observations from the CGI mean vector:
and \((X)_1\) is null; \(\bar{Z}_1\) rows and columns sum to zero; \(\bar{Y}_1\) sum to zero; and OBS 1 is contributing \(2/3(18)\) and \(-2/3(18)\) to the sum of sires A and B effects, respectively. If a function is estimable \((s_A - s_B)\) it has to be a linear function of the observations; \([\bar{Z}_1\bar{Y}_1]\) is showing that \((s_A - s_B)\) is estimable since it contains information in this contrast.

At this step, the first 3 observations \([\bar{Z}_1^y \bar{y}_1],[\bar{Z}_2^y \bar{y}_2],\) and \([\bar{Z}_3^y \bar{y}_3]\) can be transformed by Givens rotations.

Working with observations from CG2:

\[
\begin{array}{cccc}
\text{OBS} & \text{CG2} & \text{SIRE A} & \text{SIRE C} \\
4 & 1 & 1 & 0 \\
5 & 1 & 1 & 0 \\
6 & 1 & 0 & 1 \\
7 & 1 & 0 & 1 \\
8 & 1 & 0 & 1 \\
\end{array}
\]

\[
\begin{bmatrix}
5 \\
1
\end{bmatrix}
\begin{bmatrix}
2/3 \\
3/5
\end{bmatrix} = \begin{bmatrix}
175 \\
6*2 "2
\end{bmatrix} = \text{CG2 mean vector.}
\]

Deviating these observations from the CG2 mean vector:

\[
\begin{array}{cccc}
\text{OBS} & \text{CG2} & \text{SIRE A} & \text{SIRE C} \\
4 & 0 & 3/5 & -3/5 \\
5 & 0 & 3/5 & -3/5 \\
6 & 0 & -2/5 & 2/5 \\
7 & 0 & -2/5 & 2/5 \\
8 & 0 & -2/5 & 2/5 \\
\end{array}
\]
and again, $\bar{X}_2$ is null; $\bar{Z}_2$ rows and columns sum to zero; and $\bar{y}_2$ sum to zero. Notice that it is not needed to perform any operation on null segments of $X$ and $Z$ since they remain unchanged: this represents a major reduction in number of operations.

The data matrix, after the absorption of the CG effects is:

$$
\bar{A} = [\bar{X} \bar{Z} \bar{y}] = \begin{bmatrix}
0 & 0 & 2/3 & -2/3 & 0 & 18 \\
0 & 0 & -1/3 & 1/3 & 0 & -4 \\
0 & 0 & -1/3 & 1/3 & 0 & -14 \\
0 & 0 & 3/5 & 0 & -3/5 & 14 \\
0 & 0 & 3/5 & 0 & -3/5 & 8 \\
0 & 0 & -2/5 & 0 & 2/5 & -11 \\
0 & 0 & -2/5 & 0 & 2/5 & -14 \\
0 & 0 & -2/5 & 0 & 2/5 & 3
\end{bmatrix}
$$

Since $\bar{X} = 0$, these columns can be completely left out. Saving $\Sigma_1$ and $\Sigma_2$ allows for backsolving to obtain $\hat{h}_1$ and $\hat{h}_2$, after $\hat{s}$.

Performing the multiplications $\bar{Z}'\bar{Z}$ and $\bar{Z}'\bar{y}$ one obtains $Z'MZ$ and $Z'My$, the same results as if operating with the NE. If one takes the time to contemplate and stare at $\bar{A}$ one can gain some insight on what is the absorption process and what kind of information each observation holds.

2. **Apply Givens rotations**

As each record is deviated from its CG mean vector, immediately apply Givens rotations, according to the scheme given in the example in section C of this chapter.

The matrix $T$ has dimensions 148 * 148 (130 sires + 17 regressors +1 response) and requires 21,904 storage positions if full stored or 11,026 if half stored in a vector (vech of Searle, 1982). If CG were not absorbed, $T$ would required 131,044 positions for full storage or 65,703
for half storage as a vector (214 CG).

To perform these rotations, $3.65 \times 10^8$ operations ($2np^2$) are estimated to be needed, while if the Cholesky decomposition was used, only $5.40 \times 10^4$ operations ($p^3/6$) would be needed. The actual number of operations was different from the estimated one but, based on estimate numbers, they are saying that Givens rotations require 675 more operations than if Cholesky decomposition was used. The number for Cholesky does not consider the number of operations the form the NE; the numbers for neither method consider operations to absorb the CG effects.

There are three reasons why Cholesky decomposition was not used. They are as follows:

1) The author wanted to gain more experience with the Givens rotations, after having invested some two months to understand how it works and its properties.

2) The author did not have the time to invest in learning another method.

3) The only person in our group, in early 1984, with practical knowledge on Cholesky decomposition was Doyle Wilson and in visiting with Wilson (1984a), we were not sure that the same properties of T would apply to Q'A or vice versa.

If the LHS of the NE are bordered by the RHS (RHS is adjoined as the last column; RHS' is adjoined as the last row; and the bottom left position is filled with $y'y$ or $y'My$ — in the case of CG absorbed) and the resulting system is decomposed by the Cholesky (none of its variants) method, the resulting T is the same as Q'A. The method described above
is one of the forms of the SWEEP operator described by Goodnight (1979).

If one has such a problem, for given n and p, and the set of records is going to be used only once, there is no doubt in choosing Cholesky method with respect to computation time. Givens rotations are superior only in certain circumstances and several methods can be combined to obtain circumstances and several methods can be combined to obtain maximum overall performance.

3. Data matrix augmentation

The mixed model equations (MME) given by Henderson (1950, 1973) are:

\[
\begin{bmatrix}
\tilde{W}^T \tilde{W} & \tilde{W}^T \tilde{Z} \\
\tilde{Z}^T \tilde{W} & \tilde{Z}^T \tilde{Z} + G^{-1}
\end{bmatrix}
\begin{bmatrix}
\hat{r} \\
\hat{s}
\end{bmatrix}
= 
\begin{bmatrix}
\tilde{W}^T \tilde{y} \\
\tilde{Z}^T \tilde{y}
\end{bmatrix}
\]

where \( G^{-1} = A^{-1} * k = A^{-1} * \frac{\sigma_e^2}{\sigma_s^2} \).

A is the numerator relationship matrix (NRM). The ij-th off-diagonal element of A is the numerator of Wright's coefficient of relationship between animals i-th and j-th. The i-th diagonal element of A is equal to 1 + \( f_i \) where \( f_i \) is Wright's coefficient of inbreeding for the i-th animal. Henderson (1976) described methods for computing \( A^{-1} \) directly from a list of sires and dams and the diagonal elements of L. L is such that \( LL' = A \) and \( A^{-1} = (L')^{-1}L^{-1} \). Several improvements in the method have been made, beginning with Quaas (1976). The knowledge of L is what allows for the rapid computation of \( A^{-1} \), or, that if \( A^{-1} \) can be computed efficiently for large matrices, L or \( L^{-1} \) would also be, possibly even more efficient.

In this analysis, \( G^{-1} \) is assumed to be equal to \( I * k \). But the
arguments above allows one to say that what follows can also be applied, with the necessary modifications, when \( G^{-1} = A^{-1} \times k \).

Given \( G^{-1} \), a \( P \) matrix can always be found such that \( G^{-1} = P'P \). In the general case \( P = L^{-1} \times k \), and in this study, \( P = I \times k \) (sires are assumed unrelated, which is a gross oversimplification of the real situation).

What follows is due to Harville (1983) on how to use PROC GLM of SAS (1982b) to obtain GLS estimators (MM solutions). No CLASSES statement can be used when using PKOC GLM with this purpose.

Augment the data matrix with \( q \) pseudo-observations, where \( q \) is the number of sires and the dimensions of \( P \), and form the NE for the total data matrix.

\[
\begin{bmatrix}
\tilde{W} & \tilde{Z} \\
0 & P
\end{bmatrix}
\begin{bmatrix}
\tilde{r} \\
\tilde{s}
\end{bmatrix}
= 
\begin{bmatrix}
\tilde{y} \\
0
\end{bmatrix}
\]

(41)

\[
\begin{bmatrix}
\tilde{W}' & 0 \\
\tilde{Z}' & P'
\end{bmatrix}
\begin{bmatrix}
\tilde{W} & \tilde{Z} \\
0 & P
\end{bmatrix}
\begin{bmatrix}
\tilde{r} \\
\tilde{s}
\end{bmatrix}
= 
\begin{bmatrix}
\tilde{W}' & 0 \\
\tilde{Z}' & P'
\end{bmatrix}
\begin{bmatrix}
\tilde{y} \\
0
\end{bmatrix}
\]

(42)

By definition, \( G^{-1} = P'P \), therefore GLS estimators of \( r \) and \( s \) can be obtained by augmenting the data matrix with \([0, P, 0]\), and the MME will be formed when using a standard routine of forming NE from complete regression model matrices (Eq. (40) \( \equiv \) Eq. (42)).

As was seen before, solutions from \( Q'A = T \) are the same (from Givens rotations or Cholesky decomposition of the NE). Or, applying Givens
rotations to Eq. (41) yield the same results as solving Eq. (40) or Eq. (42).

Another property of Givens rotations is that the resulting upper triangular matrix is the same, independent of the order on which rows are processed (this property is that allows for the sequential accumulation of information, without the need to reprocess all previous records). This property can be used to our advantage.

The "real" observations (corresponding to the fixed part of the model) can be transformed (either by Givens or Cholesky - via NE - methods) into $T$. For each assumed $G^{-1}$, the pseudo-observations are rotated (this part corresponds to the shrinking of sire estimates phase) over $T$, which is computed only once and stored, yielding corresponding $T^*$'s.

So far, the following steps have been described:

1) absorb CG, transforming the data matrix according to Eq. (38) into one corresponding to Eq. (39) (Model A of Harville (1984a));

2) as the observations from each CG are being transformed by the projection matrix $M$, apply Givens rotations to them, forming $T$ (Cholesky could have been used with more efficiency), corresponding to the fixed model.

3) rotate the pseudo observations to obtain $T^*$.

4. **Extracting information from $T^*$**

The backsolving of $T^*$ will yield GLS estimators of $r$ and $s$ (from Eq. (41)) as shown above. The square of the bottom-right element of
This is equal to

\[ SSE = y'y - \hat{\beta}'W'y - s'Z'y = (y' - \hat{\beta}'W - s'Z')M'y \]

\[ = (y'y - \hat{\beta}'W'y - s'Z'y)M'y \]

Since \( M'y = (I - X(X'X)^{-1}X')y = y - X(X'X)^{-1}X'y \), then \( M'y = y - X\hat{\beta} \),

where \( \hat{\beta} = (X'X)^{-1}X'y = \text{OLS solution from the one-way model } y = X\beta + d \).

Substituting value of \( M'y \) in SSE:

\[ SSE = y'(y - X\hat{\beta}) - \hat{\beta}'W'(y - X\hat{\beta}) - s'Z'(y - X\hat{\beta}) \]

\[ = y'y - \hat{\beta}'W'y - s'Z'y - (y'X\hat{\beta} - \hat{\beta}'W'X\hat{\beta} - s'Z'X\hat{\beta}) \]

\[ SSE = y'y - \hat{\beta}'W'y - s'Z'y - K. \]

Since all elements in the equation are scalars, we can write \( K \) as:

\[ K = \hat{\beta}'X'y - \hat{\beta}'X'W\hat{\beta} - \hat{\beta}'X'Z\hat{\beta} \]

\[ = \hat{\beta}'(X'y - X'W\hat{\beta} - X'Z\hat{\beta}) \]

\[ K = y'X(X'X)^{-1}(X'y - X'W\hat{\beta} - X'Z\hat{\beta}). \]

In subsection 1 of this section, we obtained, after Eq. (38),

\[ \hat{\beta}^o = (X'X)^{-1}(X'y - X'W\hat{\beta}^o - X'Z\hat{\beta}^o), \] the OLS estimator of \( \beta \).

If \( G^{-1} \) were added to \( Z'Z \) in the NE corresponding to Eq. (38), we would obtain as a GLS estimator of \( \beta \):

\[ \hat{\beta} = (X'X)^{-1}(X'y - X'W\hat{\beta} - X'Z\hat{\beta}). \]
Substituting for $$\mathbf{b}$$ in $$\mathbf{K}$$:

$$
\mathbf{K} = \mathbf{y}'\mathbf{X}\hat{\mathbf{b}} \quad \text{or, since } \mathbf{K} \text{ is a scalar;}
\quad \mathbf{K} = \hat{\mathbf{b}}'\mathbf{X}'\mathbf{y}.
$$

Replacing $$\mathbf{K}$$ in $$\text{SSE}$$:

$$
\text{SSE} = \mathbf{y}'\mathbf{y} - \hat{\mathbf{b}}'\mathbf{X}'\mathbf{y} - \hat{\mathbf{r}}'\mathbf{W}'\mathbf{y} - \hat{\mathbf{s}}'\mathbf{Z}'\mathbf{y},
$$

which is the $$\text{SSE}$$ from the full mixed model, and equal to the original:

$$
\text{SSE} = \mathbf{y}'\mathbf{y} - \hat{\mathbf{r}}'\mathbf{W}\hat{\mathbf{y}} - \hat{\mathbf{s}}'\mathbf{Z}\hat{\mathbf{y}}, \text{ Q.E.D.}
$$

Notice that the difference between $$\mathbf{K}$$ and $$\mathbf{b}'\mathbf{X}'\mathbf{y}$$:

$$
\mathbf{K} - \mathbf{b}'\mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{W}\hat{\mathbf{r}} + \mathbf{X}'\mathbf{Z}\hat{\mathbf{s}}
$$

is what is assumed to be zero when working with "deviated records" or with WWR (not considering preadjustments for AOC and AOD).

If $$|t^*_{(148,148)}|^2$$, the square of the bottom-right element of $$\mathbf{T}^*$$, is equal to $$\text{SSE}$$ from the complete mixed model, then

$$
1 - \text{REML } \sigma^2_e = \frac{|t^*_{(148,148)}|^2}{8323 - (214 + 17)}
$$

$$
= \frac{\text{SSE}}{n - \text{(number of CG + number of regressors)}}
$$

$$
1 - \text{REML } \sigma^2_e = \frac{\mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y} - \mathbf{r}'\mathbf{W}'\mathbf{y} - \mathbf{s}'\mathbf{Z}'\mathbf{y}}{n - \text{column rank (X W)}}
$$

Where $$1 - \text{REML } \sigma^2_e$$ is the 1st iteration REML estimate of $$\sigma^2_e$$, which is dependent on the initial $$k$$ used but it is unbiased if $$k$$ comes from prior
knowledge completely independent from the data being analyzed.

To obtain REML $\hat{\sigma}_s^2$, the following formula was used:

$$\text{REML } \hat{\sigma}_s^2 = \frac{\hat{s}'\hat{s}}{q-k*\text{trace}(C_{22})}$$

where:

$\hat{s}$ are the GLS estimates,

$q = \text{number of sires} = 130$,

$k = \text{initial } \frac{\sigma_e^2}{\sigma_s^2} \text{ used, and}$

$\text{trace}(C_{22}) = \text{sum of the diagonal elements of the part of the}$

$\text{inverse of the MME corresponding to the random effect}$

$\text{being estimated.}$

Accepting the claim, without proof here, that the absorption of CG does not interfere with any of the results, and using:

$$\text{MME} = (T^*)'T^*$$

$$\text{(MME)}^{-1} = (T^*)^{-1}(T'^*)^{-1}$$

where $T^*$ is easily inverted, trace of $C_{22}$ can be computed by adding corresponding elements from the diagonal of

$$(T^*)^{-1}(T'^*)^{-1},$$

as was done in this study, or, more efficiently, as suggested by Maindonald (1984), just by adding up the squares of the elements from the rows corresponding to sires from $(T^*)^{-1}$.

With REML $\hat{\sigma}_e^2$ and $\hat{\sigma}_s^2$ a new $k = \frac{\hat{\sigma}_e^2}{\hat{\sigma}_s^2}$ is computed. If old and new $k$s
do not agree, then choose initial "low" and "high" values of k, and compute converged value of k using Schaeffer (1983) CIA.

Different values of k imply different $P = I \sqrt{k}$ and different pseudo-observations being rotated against the same $T$ which is kept in magnetic storage and computed only once.
VI. ALGORITHMS USED, WITH CRITICISMS

In this chapter, algorithms to solve some LS problems are presented while program sources and print-outs are given on the Appendices. Very few algorithms are original and directly copied ones will only be referenced here. The intent is to force consultation with the original source, mainly Lawson and Hanson (1974), whose reading should be very helpful to anyone interested in using Householder or Givens transformations or solving general least squares problems. The programs go from a simple two-way fixed model to the final mixed model used in this study, building up in complexity. All are based in triangularizing the data matrix with the Givens rotations.

All programs are written in Fortran WATFIV, as described by Moore and Makela (1981), capitalizing on the characteristic of this language for having the best diagnostic messages from the whole family of Fortran IV languages. Once a program was tested and ready, it should be converted and compiled as a Fortran H program. Fortran H is an optimizing compiler that has a faster execution time than any of the other Fortrans. Fortran G by itself has an execution time faster than WATFIV by a factor of 6 to 10, according to Lieberknecht (1984). What can be deduced is that the author wanted to have correct programs and to spend as little time as possible in programming. Also, once the program was correct, it would run only one time. The programs reflect little concern about the efficiencies, being more a straightforward solution; but a correct one.

One can raise some concerns about Fortran being used at all, since
more "efficient" languages exist; efficiency here meaning shorter execution time. Recently the author has developed some ideas (which can easily be labeled as biased, since they developed after a de facto situation), that may indicate that Fortran may be the most overall efficient language; overall efficiency now meaning total scientific output of an animal breeding research institution. Some reasons for this assertion are as follows:

1) If animal breeders have had any computing practice, it is most likely that the language used was Fortran. Hence, they can begin to work on a subject matter problem immediately; or have a reduced time to develop further programming skills; or, given a fixed amount of training time, be more efficient programmers if the language is Fortran, since they can capitalize on previous exposure to it.

2) If the institution is involved in long term projects or there exists a high rotation of personnel, and if systems, programs, subroutines and functions are written in Fortran, then new people coming into the projects can easily master the programs and soon begin to build on past experience, improving programs, adapting them to new situations or adding parts to the system.

3) There exists a large effort in programming consubstantiated into subroutine packages, like IMSL, LINPACK, BLAS and SPARSPACK, all written in Fortran, all in the public domain or at nominal charges, all of them useful (or at least potentially so) to animal breeders. Of course, once compiled, these subroutines can be called by programs written in other languages, but Fortran is still a requisite to understand what the subroutines are doing.
4) Research institutions are not characterized by doing large amounts of routine processing work, where execution processing time and fast printers are imperative.

5) A gray area exists when large tasks are involved. Then a decision needs to be made in choosing between how efficiently the job is to be performed and how broad a base the institution wants of people understanding the programs, improving them and insuring continuity of progress.

6) The availability of low priced and powerful microcomputers renders the axiom of "lower execution time" more and more irrelevant.

7) The power of Fortran to crank numbers efficiently and the fact the Fortran is the most worldwide language are basic features that guarantee its use on a large number of installed equipment and for years to come.

8) If one begins to follow the maxima of efficiency in a strict sensu (like the host of Algol 68 derivatives, like the C and ADA) soon we will all be "talking" a different language.

Maybe in the future, more "efficient" languages get widespread enough (like higher versions of BASIC) to replace Fortran by attaining a higher "overall efficiency."

The first two programs presented in this chapter are to analyze a two-way model, first as a fixed model (treatments and blocks fixed), and second as a mixed (blocks random) model.

Appendix A presents the data, as given by Harville (1983) as his problem set number 7. This is possibly one of the smallest examples and results can be checked using any standard statistical package.
A. Two-Way Fixed Model (GIVFIX)

In Appendix B, the listing of a source program to solve a two-way least squares problem is given. Data analyzed are given in Appendix A.

Dimensions of the upper triangular matrix are \( N \times N \) where \( N = \) number of treatments + number of blocks + number of responses = 4 + 6 + 1 = 11; since all elements below the diagonal are zero, only \( N \times (N+1)/2 = 66 \) positions are needed. The vector \( U \) stacks these elements, row by row.

Major steps of the program are given next, using the line number as references.

24-29 First and last positions of each row from the triangular matrix (rows of decreasing sizes since the first element of each row is the diagonal element) in the column vector \( U \) are determined.

Begin

For row = 1 until row = 11 do begin

First position (row) = \((N-row/2) \times (row-1) + row\)

Last position (row) = \((N-(row+1)/2) \times row + row\)

end

End

It is unnecessary to calculate the last position of a row since this is more easily given by

Last position (row) = First position (row+1) - 1.

The above formulas were being sought and tried to be reinvented by the author for a whole day of much frustration. Wilson (1984a) graciously gave them after a visit. Some time later, the author found a Fortran function, published by Schaeffer (1976), to find the corresponding position in a column vector which half-stores a full
symmetric matrix, from any position (row and column) of the full matrix. This publication, Schaeffer (1976), which explains and lists efficient programs in Fortran for sire evaluation and cows' repeated records analysis, should be part of all animal breeders' libraries but, unexplicably, is very seldom referenced. Actually, the author only got to know the existence and received a copy of Schaeffer (1976) by extended kindness of Pollack (1983).

30–35 First observation is read and transformed into the incidence vector R.

39–128 Apply Givens rotations to all other observations.

Begin

For record = 2 until record = last do

Begin

Read record.

If record is from a new TRT then.

Begin

Save vector R into U

Zero out R

Transform new record into incidence vector R

End

Else, if record is from the same TRT as previous record then

Begin

Transform new record into incidence vector S.

Rotate vectors R and S making them orthogonal at column corresponding to pre-
sent TRT, using subroutines G1 and G2. Keep R.

Find the first position in S that is not zero, which will be one of the BLK column; read that corresponding row from U and call it SU; rotate S against SU, using G1 and G2; save transformed SU into U. Go to next position in S which is not zero and repeat process until S is completely annihilated.

End

End

Save in U last R, which is the row corresponding to last TRT.

End

On the above part of the program, especially when S is rotated against SU, the basic structure comes from Lawson and Hanson (1974), algorithm SEQHT (27.10), on pages 210 and 275.

Apply the constraint that the sum of BLK effects is zero. Since this is a no-intercept model, the rank deficiency is one.

Begin

Create a vector S, containing a constraint on the solutions for BLK effects, with positions 5 to 10 equal to 1.0 and all others equal to 0.0. Rotate S against successive rows of U.
For each column of $S$ to be annihilated, read corresponding row of $U$ as a vector $R$. Rotations need to include last column of $S$, since $G_2$, when transforming the two vectors will put non-zero element in position 11.

End

Here, algorithm 27.10 of Lawson and Hanson is more closely followed. Calculate the complete inverse from the upper triangular, postmultiply it by its transpose, and obtain the inverse of the LHS of the NE.

If $p$ = number of parameters, then to obtain the inverse of the upper triangular requires $\frac{p^3}{6} + o(p^2)$ operations (same as a Cholesky decomposition). The postmultiplication of $U^{-1}$ by its transpose also requires $\frac{p^3}{6} + o(p^2)$ operations. In total, $\frac{p^3}{3} + o(p^2)$ operations are required to compute the elements of the inverse of the LHS.

All calculations are done with a column vector $(c)$, instead of a two-way matrix, and, with this particularity aside, algorithm COV (12.12) of Lawson and Hanson (1974, page 69), was used exactly (Chapter 12: The covariance matrix of the solution parameters).

If the model used contained a constant element then another constraint would be needed. If both constraints were that the sum of the respective effects are zero, then the resulting generalized inverse would be the Moore-Penrose inverse since the solutions are minimum norm. This result was demonstrated by the author and Kempthorne (1983) gave the proof. This generalized inverse is somehow unique. If the program is checked by using some package, the "inverses" may not agree, depending on
which restriction or generalized inverse was used to obtain the solution. Variances of estimable functions were checked with PROC GLM of SAS (1982b) and they agreed.

202-211 Solutions are obtained by backsolution of the upper triangular T.

This part is based, with modifications, on PROGl given by Lawson and Hanson (1974, page 279).

Solutions are saved on the column vector C, which also contains the unscaled "inverse," on positions which correspond to the last column of a half stored matrix.

The trace of the "inverse" is also computed at this point.

212-230 Degrees of freedom for error, mean squared error and printing statements.

236-258 Subroutine Gl, transformed to double precision, but for the rest is exactly as given by Lawson and Hanson (1974).

This subroutine is a reformulation of what was presented in Chapter V, Section C, designed to avoid unnecessary underflow or overflow.

Algorithm Gl (10.25) is presented in page 58 and the source on page 309 of Lawson and Hanson (1974).

264-270 Subroutine G2, also a double precision version, given by Lawson and Hanson (1974, page 59) (Algorithm 10.26) and page 309.

272-283 Records from Appendix A to be read.
B. Criticisms on GIVFIX Program

1. There is no need to read the first observation disjointly from
   the other observations. At the beginning, the first row of $U$ will be
   zero and the simple application of $G_1$ and $G_2$ will interchange record read
   and row of $U$.

   The awkward algorithm resulted in three vectors of length $p+1$ being
   necessary, where two would be sufficient.

2. The read statement governed by a do-loop is not efficient,
   according to Berger (1984). Consultation with personnel from specific
   computing centers should indicate the most efficient way of inputting
   records.

3. All operations are done like $U$ was in a virtual memory. This
   was intended to be so, allowing one to write an algorithm which would
   have only one row of $U$ in active memory at any given time and saving in
   input-output operations. But the complete column vector $U$ is present in
   active memory.

C. Results from GIVFIX Program

In Appendix C, the part of the print-out containing the solutions is
given.

The vector $C$, with 66 positions, is the vech of $LHS^{-1}$, and on the
last position of each row (or column) the solution to that row is kept.
That is $C$ has $(p+1)(p+2)/2$ positions. Position 66 could be used to save
$MSE$.

Solutions for $TRT_i$ correspond to the $(\mu+t_i)$ in an intercept model.
Contrasts among $TRT$ or among $BLK$ are unique. Solutions for $BLK$ are under
the constraint that $\sum b_j = 0$. Note the order of the solutions: the
design vectors are formed as records are read. Hence, the order that the
solutions are presented are dependent on the order that the TRT and BLK
appear on the records, keeping in mind that records from the same TRT
come in clusters.

Estimates and variances of estimates of estimable functions were
checked with the ones form PROC GLM of SAS (1982b).

D. Obtaining Variance Components of a Mixed Two-Way Model Using
Henderson's Iterative Algorithm (HIA)

The same example as given in Appendix A was also used to estimate
variance components when considering BLK as a random effect as suggested
by Harville (1983).

Using PROC VARCOMP or PROC GLM (declaring BLK as RANDOM) of SAS
(1982b) first estimates were obtained as: $\hat{\sigma}_e^2 = 139.3, \hat{\sigma}_b^2 = 9.416$, and $k$
$= \frac{\hat{\sigma}_e^2}{\hat{\sigma}_b^2} = 14.796467$.

The square root of $k$ was used to form $P = I_b^* k$ and pseudo-
observations were added to the data matrix as described in section D of
Chapter V. Using PROC GLM on this augmented data set gave first-
iteration-REML $\hat{\sigma}_e^2 = 139.30471, \hat{\sigma}_b^2 = 9.4422202$ and a new $k = 14.753385$.

If the solutions are in the parameter space ($\hat{\sigma}_1^2 > 0$) then REML
solutions are guaranteed to converge, but the rate of convergence is
slow. Schaeffer (1983) compared some computing algorithms for REML and
found that the equation, as given by Harville (1983) and used in this
study, to estimate, say, block variance as $\hat{\sigma}_b^2 = b^* (q-k*\text{tr}(C))^{-1}$ only
needed 8 rounds of iteration to arrive at a same $k$ as after 25 rounds
when using the alternative \( \hat{\sigma}_b^2 = \left( \hat{b}^* \cdot \mathbf{b}^* \mathbf{b}^T + \text{tr}(\mathbf{C}^2\hat{\sigma}_e^2) \right) \mathbf{q}^{-1} \).

Some authors suggest iterating until change will not affect the estimate of heritability within the second decimal place. This would be correct if the method had a fast convergence or one were sure that a few more iterations, with steep decreasing changes, would bring the converged value.

In Appendix D, results are shown for HIA. The final value for the ratio \( k \) was 13.79100182916 before being considered dependent on machine error.

Several conclusions can be taken from Appendix D, as follows:

1) To arrive at a final value (± 10^-12) some 400 iterations were performed, not all of those were needed, but this number says roughly how slow is the convergence and how inefficient is HIA.

2) If one were to adopt the criteria of stopping iterations whenever changes in solutions went under the threshold of 0.02, the final ratios would be 13.362 or 14.22, depending on whether one began with lower or higher approximations. If these were animal breeding data, resulting estimated heritabilities would be 0.2785 and 0.2628, respectively. Hence, that criteria fails to accomplish the desired level of precision on the \( \hat{h}^2 \).

3) Before arriving at the machine error level, HIA produce solutions that begin to diverge. Hence, besides being much less efficient than the CIA, this method is less stable or precise. The Common Intercept Approach, as will be seen, allows for greater precision before the series of solutions begin to diverge.
So poor results for HIA (even when initial estimate was an Anova estimator from a balanced design) allows one to suggest that even a grid search, at 0.1 intervals, would be more economical and much more precise than HIA.

E. A Two-way Mixed Model Program (GIVMIXCV) to Obtain REML Estimates of Variance Components Using the Common Intercept Approach

Expected results for HIA were so bad that it was discarded and another method sought. Schaeffer (1983) presents an alternative, CIA, that have a practical appeal and was implemented in the GIVMIXCV program, listed in Appendix E.

Until the transformation of the data matrix into the upper triangular $T$, the algorithm and program is exactly the same as GIVFIX, except the declaration of some extra variables. Also, it is not needed to impose restrictions since with BLK random, the data matrix is full rank.

The upper triangular matrix corresponding to the fixed model is saved in the vector $U$ (vech). All further operations are processed on $U_2$, that on the beginning of each iteration is copied from $U$.

Initial values for $k$ are set at 5.0 (low) and 20 (high). If TRT were CG and BLK were sires, these variance ratios would mean that one would expect $h^2$ for the response measured to be within the 0.19 and 0.67 interval. It is interesting to note that the CIA approach finds a converged value even if the initial ratios used do not contain the final solution.

The algorithm is as follows:
Begin
Set initial values (ALFA) at 5.00 (low), ILH=1; and 20.00 (high), ILH=2.

For STAGE=1 until STAGE=15 do
Begin
For ILH=1 (low) until ILH=2 (high) do
Begin
If STAGE=1 then RATIO=ALFA(ILH);
else; if STAGE > 1 then
RATIO=(converged value from previous stage) - 
(difference between new and old ratio estimates).

Copy U to U2.
For each row of augmented observations (S) and as each row is formed, apply Givens rotations against corresponding rows (R) of U2.
Find the inverse.
Calculate the trace of the part of the inverse corresponding to BLK.
Backsolve, calculate new $\sigma^2_e$, $\sigma^2_b$ and k(NRATIO) and the difference DELTA=NRATIO-RATIO.
End
Calculate converged value from the two pairs of low and high estimates, for each stage.
End
End
In Appendix F, the results of using this CIA iteratively is presented.

In only 5 stages (two inverses needed per stage), a converged value was reached. Comparing this result with the hundreds of iterations in Appendix D is a pleasant experience.

Only five cycles of iterations were needed and one can see how the converged value gets trapped inside closer and closer intervals.

Contrary to what happened with the successive approximations method the series of estimates never diverges, even if estimates are within machine error level.

For any practical purpose, three cycles would be more than enough.

A little better results were found by using $\Delta^{(1)}/2$ instead of $\Delta^{(1)}$ on the formula

$$k^{(i+1)} = \frac{a^{(1)}_{c} - \Delta^{(1)}}{c}$$

to find new ratios to use in the next stage. Results are presented in Appendix G. This modification was not tried on the real data from this study, since it was only thought of after the analysis was completed, but it may be useful in some other applications.

F. Programs to Solve a Two-way Mixed Model, With Covariates

(Programs ABSSIRER and REML)

After these exercises with a small example, the real problem of this study was attacked, also by pieces. The exercises helped the author to refresh his scant Fortran knowledge; to understand and to see how Givens rotations work; to gain some practice with REML estimation; and to decide in favor of the CIA.
A subset from the total number of observations, comprising 155 of the 8323 records, 4 of the 214 CGs and 5 of the 130 sires were used to test all programs before a final run.

First, a program containing only CG and the 17 regressions was used and results checked against PROC GLM of SAS (1982b). At this stage, it was learned how to extract information from the upper triangular $T$ to perform $t$-tests and sequential $F$-tests. Second, a program to handle the inclusion of sires (fixed) was developed. Third, a program to handle the absorption of CG on the data matrix before applying the rotations was developed. Fourth, the modification to handle sires as random effects and to obtain REML variance components estimates was implemented. At each of these steps, weighted (by within CG error variance) and robust (using also the step function of each residual) estimation was also programmed, simply by multiplying each observation (or transformed vector) by the corresponding weight. Each of the above programs, with the sample data set, had results carefully checked by comparison with results from PROC GLM; all fixed models with the complete data set were also checked with PROC GLM of SAS (1982b). The great number of problems and errors detected by this double checking have proved inequivocably to the author that a dangerous quality for a programmer is self-confidence. It may serve the best economical interests of an institution to keep two programmer experts, at the same level, which adopt the tactic of complete distrust on each other's programs and results. Algorithms are discussed in this chapter and programs given in the appendices also with the hope that errors and better alternatives will be pointed out.

Two main differences between previous programs and the ones used to
analyze the data exist: first, the total algorithm was sectioned in two parts, the cutting point being the obtaining of T; second, all operations were performed on a two dimensional matrix, even if it contained only zeros below the diagonal, to reduce number of operations and considering that the amount of memory is not at premium at the ISU Computation Center.

The first program, ABSSIRER, reads records sorted by CG, absorbs CG, applies rotations at end of each CG, and prints and saves the upper triangular, called U in the program, and the normal equations corresponding to the CG, called REALNE.

In Appendix H, the listing of the program ABSSIRER is given. Declared values for the variables NT0UCG=4, NTOURO=5 and ILAST=155 indicate that this is the source for running the test data set. The program used for the complete data set is exactly the same after changing the above values and corresponding matrix dimensions.

Some features of the program are as follows:

1. As records are read for each CG, the design vector for sires (LESIRE) is formed and adjoined to the covariates and WW, giving the matrix ANREAD.

As each row of ANREAD is formed, its contents are accumulated on REALNE. At the end of a CG, a row of REALNE contains the corresponding NE for that CG, with the "diagonal elements" (NANIM) of the CG*CG part of the LHS occupying the first column of REALNE. To complete the absorption of CG, each record (row of ANREAD) is deviated from its CG average (REALNE/NANIM). This is what the projection matrix
$M = I - X(X'X)^{-1}X'$ does when applied to the data matrix.

2. At the end of a CG, ANREAD contains the "deviated" records and design matrix, and to each of its rows the Givens rotations are applied. The fact that the design matrix for sires is formed as records are read (sire from the first record read will occupy the first column of the design matrix) is used to reduce the number of operations. NTUCG is a vector which keeps track of how many different sires were found until a given CG. All further columns corresponding to sires are zero. Therefore, no rotations are applied on these columns. Other "tricks" given by Seber (1977) and presented in Chapter V, Section C, are also used to reduce the number of unnecessary calls to subroutines G1 and G2.

3. If CG were not absorbed, U would be a 362*362 matrix and 2.18*10$^9$ operations estimated to be needed to complete the formation of U. If the elimination of CG was done without any cost, the number of estimated operations (2*8323*148$^2$) would be 3.65*10$^8$, about 6 times less. The number of statements executed by the actual run of the program for the complete data set, as given in the print-out, was 2.75*10$^6$, including the printings of U and REALNE. The execution time was 3,537.8 seconds, or only 778 statements/sec. When running the program on the test data, 234,012 statements were executed on 8.17 seconds, or about 28,600 statements/sec. PROC GLM for the full data set, also absorbing CG, took only 169 seconds, or about 21 times less; and it is necessary to consider that PROC GLM
performs several other tasks.

Recall that if $p=362$ and $n=8323$, then Givens is expected to require 675 times more operations than Cholesky. This ratio was brought down considerably, but at least two main flaws remain in the program ABSIRE:

1. Using WATFIV increases execution time by 6 to 10 times in comparison with FORTRAN G (Lieberknecht, 1984).

2. The reading form used is inefficient (Berger, 1984).

The most important conclusion is that if one has a specific problem to solve, Cholesky decomposition should be used. If the normal equations are arranged as suggested by Goodnight (1979) and the triangular matrix stored, then Givens rotations have a role to play in updating $T$ when new observations become available.

Program REML, given in Appendix I uses very much the same algorithms as program GIVMIXCV. Major steps of the program are:

1. Read $U$, calculated by ABSSIRE, as UDISK.

2. Use initial "low" and "high" estimates of the variance ratio $k$ and, for each one of these, create pseudo-observations using $k^{1/2}I_{130}$.

3. Rotate these observations against UDISK.

4. Backsolve and print solutions.

5. Obtain the inverse of the LHS.

6. Obtain REML $\hat{\sigma}_e^2$ and $\hat{\sigma}_g^2$ and a new $k$.

7. Obtain a converged $k$ and repeat the whole process once more.

8. At end, print the inverse.

The program performs a complete iteration for each of the two "low"
and the two "high" estimates. For the complete four rounds, REML took 384.4 sec of execution time; an average of 96.1 seconds per round.

This program shows the potential of Givens rotations for animal breeding applications.

The total number of statements executed, for the four rounds, was 2,320,341. This number could be greatly reduced by:

1) only printing solutions, and calculating standard errors of the estimates, t-tests, type I SS and sequential F-tests, with the final converged value.

2) only calculating (if desired) the complete inverse in the last round. Lawson and Hanson (1974) estimate that \( p^3/6 \) operations are needed to obtain \( U^{-1} \) from \( U \), and another \( p^3/6 \) to postmultiply \( U^{-1} \) by its transpose. As suggested by Maindonald (1984), if one is only interested in the diagonal elements of the inverse or on its trace, then summing up the squares of each row elements would give the desired diagonal elements. The average row would require \( (p+1)/2 \) multiplications and as many additions. For all rows, this would mean \( (p^2+p)/2 \) operations, or, for \( p=147 \), about 49 times less operations than the multiplication of \( U^{-1} \) by \( (U^{-1})' \).

The source program given in Appendix I eliminates some rows and columns of UDISK to obtain REML variance components and GLS estimates under a reduced model. This subject will be dealt with on Chapter VII, Section D. If lines 58 to 106 are deleted and dimensions of \( U \), \( C \), \( SST_1 \), \( SOL \) and \( RNAME \) kept in agreement with the ones from UDISK, then the
program is exactly the same as used for the complete (all regressions) model.
VII. RESULTS AND DISCUSSION

Results obtained in this study are considered as an intermediate product and figures obtained should not be used in any important decision making. In some cases, where no other information exists, results presented here can be used as a gross indications of the real effects. Such a warning or disclaimer is deemed necessary for the following reasons:

1. Total number of observations, and especially the number of observations in some subclasses, are not large enough to ensure high reliability.

2. Some of the estimation methods used, especially WLS and WTROB, were not used correctly or at their potentiality.

3. Some other factors should be included in the models, like dams and interactions among regressions. One such term that now seems obvious is the interaction between age of calf and seasonality effects.

4. The models used are the result of the predetermined views of the author. These preliminary results should be discussed with "geneticists" and "environmentalists" in RS before beginning a new cycle of analysis. At least at this point, some comfort can be taken from the thoughts of the late Jacob Bronowski as cited by Johnston (1984):

   Science is a very human form of knowledge. We are always at the brink of the known, we always feel forward for what is to be hoped. Every judgment in science stands on the edge of error, and is personal. The world is not a fixed, solid array of
objects, out there, for it cannot be fully separated from our perception of it. It shifts under our gaze, it interacts with us, and the knowledge that it yields has to be interpreted by us. There is no way of exchanging information that does not demand an act of judgment.

The author does not know how to pretend that there exists a "pasteurized science" and warns that the following results are very much subject to personal interpretation. Further, being away from RS precluded the chance of enriching discussions with colleagues about some of the effects and possible "causes."

This chapter has been divided in four parts: comparison of methods and models; environmental effect estimates; variance component estimates; and results from using a reduced model. Suggestions for a future study are presented in Chapter X.

A. Comparison of Methods and Models

Methods of estimation, differing in the assumptions made about $\sigma_e^2$, used in this study were: ordinary least squares (OLS), weighted least squares (WLS) and robust estimation (WTROB). All models included CG and the same 17 regressions, but preliminary models did not include sires (called 1-modes, for Incomplete), others considered sires fixed, while the last one considered sires random (GLS) assuming that $V(y) = ZZ' \sigma_s^2 + \sigma_e^2$.

All together seven models have been used and the major results are presented in Table 8.

As assumptions about $\sigma_e^2$ change, the models give a better fit, as measured by $R^2$ (WTROB > WLS > OLS) or by the variance of the estimates.

With the exception of the GLS model, all other results are from PROC
Table 8. Coefficients of determination ($R^2$), mean squared errors (MSE), estimates of regression coefficients and their standard errors (in parentheses, below the estimate) for the seven models used in this study.

<table>
<thead>
<tr>
<th>Without Sires</th>
<th>I-OLS</th>
<th>I-WLS</th>
<th>I-WTROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.551</td>
<td>0.574</td>
<td>0.668</td>
</tr>
<tr>
<td>MSE</td>
<td>412.0</td>
<td>1.001</td>
<td>0.562</td>
</tr>
<tr>
<td>STAT1</td>
<td>-0.580* (0.436)</td>
<td>-0.412* (0.421)</td>
<td>-0.602* (0.346)</td>
</tr>
<tr>
<td>STAT2</td>
<td>-0.692* (0.498)</td>
<td>-0.884* (0.493)</td>
<td>-0.658* (0.404)</td>
</tr>
<tr>
<td>AOD11</td>
<td>20.530 (2.503)</td>
<td>22.172 (2.477)</td>
<td>21.686 (2.051)</td>
</tr>
<tr>
<td>AOD12</td>
<td>10.280 (2.530)</td>
<td>12.294 (2.378)</td>
<td>12.175 (1.962)</td>
</tr>
<tr>
<td>AOD21</td>
<td>-1.648 (0.236)</td>
<td>-1.805 (0.233)</td>
<td>-1.750 (0.193)</td>
</tr>
<tr>
<td>AOD22</td>
<td>-0.707 (0.239)</td>
<td>-0.892 (0.224)</td>
<td>-0.882 (0.185)</td>
</tr>
<tr>
<td>AOD31</td>
<td>1.271 (0.292)</td>
<td>1.481 (0.289)</td>
<td>1.386 (0.241)</td>
</tr>
<tr>
<td>AOD32</td>
<td>0.261* (0.296)</td>
<td>0.486* (0.278)</td>
<td>0.473 (0.230)</td>
</tr>
</tbody>
</table>

*Indicates that the probability of the parameter estimated being equal to zero is larger than 0.05. All other estimates are statistically (p<0.05) different from zero.
<table>
<thead>
<tr>
<th></th>
<th>Sires Fixed</th>
<th>Sires Random</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>WLS</td>
</tr>
<tr>
<td>0.566</td>
<td>0.588</td>
<td>0.676</td>
</tr>
<tr>
<td>404.4</td>
<td>0.984</td>
<td>0.558</td>
</tr>
<tr>
<td>-0.667*</td>
<td>-0.583*</td>
<td>-0.712</td>
</tr>
<tr>
<td>(.448)</td>
<td>(.434)</td>
<td>(0.358)</td>
</tr>
<tr>
<td>-0.492*</td>
<td>-0.616*</td>
<td>-0.408*</td>
</tr>
<tr>
<td>(.522)</td>
<td>(.516)</td>
<td>(0.426)</td>
</tr>
<tr>
<td>(2.566)</td>
<td>(2.529)</td>
<td>(2.109)</td>
</tr>
<tr>
<td>10.606</td>
<td>12.321</td>
<td>12.311</td>
</tr>
<tr>
<td>(2.576)</td>
<td>(2.440)</td>
<td>(2.024)</td>
</tr>
<tr>
<td>-1.644</td>
<td>-1.739</td>
<td>-1.704</td>
</tr>
<tr>
<td>(.241)</td>
<td>(.237)</td>
<td>(.198)</td>
</tr>
<tr>
<td>-0.738</td>
<td>-0.895</td>
<td>-0.894</td>
</tr>
<tr>
<td>(.242)</td>
<td>(.229)</td>
<td>(.190)</td>
</tr>
<tr>
<td>1.231</td>
<td>1.378</td>
<td>1.319</td>
</tr>
<tr>
<td>(.297)</td>
<td>(.293)</td>
<td>(.246)</td>
</tr>
<tr>
<td>0.295*</td>
<td>0.488*</td>
<td>0.487</td>
</tr>
<tr>
<td>(.298)</td>
<td>(.282)</td>
<td>(.235)</td>
</tr>
</tbody>
</table>
Table 8. (Continued)

<table>
<thead>
<tr>
<th></th>
<th>I-OLS</th>
<th>I-WLS</th>
<th>I-WTROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF1</td>
<td>0.827</td>
<td>0.863</td>
<td>0.812</td>
</tr>
<tr>
<td></td>
<td>(.389)</td>
<td>(.365)</td>
<td>(.306)</td>
</tr>
<tr>
<td>SF2</td>
<td>-0.00480*</td>
<td>-0.00481*</td>
<td>-0.00435</td>
</tr>
<tr>
<td></td>
<td>(.00265)</td>
<td>(.00245)</td>
<td>(.00205)</td>
</tr>
<tr>
<td>SF3</td>
<td>0.00530*</td>
<td>0.00452*</td>
<td>0.00365*</td>
</tr>
<tr>
<td></td>
<td>(.00383)</td>
<td>(.00338)</td>
<td>(.00279)</td>
</tr>
<tr>
<td>SS1</td>
<td>0.193</td>
<td>0.218</td>
<td>0.239</td>
</tr>
<tr>
<td></td>
<td>(.086)</td>
<td>(.083)</td>
<td>(.067)</td>
</tr>
<tr>
<td>SS2</td>
<td>-0.000626</td>
<td>-0.000668</td>
<td>-0.000712</td>
</tr>
<tr>
<td></td>
<td>(.000217)</td>
<td>(.000207)</td>
<td>(.000171)</td>
</tr>
<tr>
<td>SS3</td>
<td>-0.0000305</td>
<td>-0.0000300</td>
<td>-0.0000273</td>
</tr>
<tr>
<td></td>
<td>(.0000112)</td>
<td>(.0000107)</td>
<td>(.0000088)</td>
</tr>
<tr>
<td>SS4</td>
<td>0.00122</td>
<td>0.0000953</td>
<td>0.000109</td>
</tr>
<tr>
<td></td>
<td>(.000047)</td>
<td>(.0000448)</td>
<td>(.000037)</td>
</tr>
<tr>
<td>CA1</td>
<td>0.584</td>
<td>0.544</td>
<td>0.528</td>
</tr>
<tr>
<td></td>
<td>(.123)</td>
<td>(.116)</td>
<td>(.095)</td>
</tr>
<tr>
<td>CA2</td>
<td>-0.000524</td>
<td>-0.000440*</td>
<td>-0.000378</td>
</tr>
<tr>
<td></td>
<td>(.000239)</td>
<td>(.000227)</td>
<td>(.000186)</td>
</tr>
<tr>
<td></td>
<td>Sires Fixed</td>
<td>Sires Random</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OLS</td>
<td>WLS</td>
<td>WTROB</td>
</tr>
<tr>
<td>1.008</td>
<td>0.924</td>
<td>0.877</td>
<td>0.966</td>
</tr>
<tr>
<td>(.395)</td>
<td>(.372)</td>
<td>(.313)</td>
<td>(.390)</td>
</tr>
<tr>
<td>-0.00610</td>
<td>-0.00519</td>
<td>-0.00486</td>
<td>-0.00578</td>
</tr>
<tr>
<td>(.00270)</td>
<td>(.00251)</td>
<td>(.00210)</td>
<td>(.00266)</td>
</tr>
<tr>
<td>0.00681*</td>
<td>0.00465*</td>
<td>0.00413*</td>
<td>0.00638*</td>
</tr>
<tr>
<td>(.00395)</td>
<td>(.00349)</td>
<td>(.00289)</td>
<td>(.00387)</td>
</tr>
<tr>
<td>0.232</td>
<td>0.235</td>
<td>0.250</td>
<td>0.220</td>
</tr>
<tr>
<td>(.088)</td>
<td>(.086)</td>
<td>(.070)</td>
<td>(.086)</td>
</tr>
<tr>
<td>-0.000754</td>
<td>-0.000743</td>
<td>-0.000769</td>
<td>-0.000703</td>
</tr>
<tr>
<td>(.000222)</td>
<td>(.000213)</td>
<td>(.000176)</td>
<td>(.000218)</td>
</tr>
<tr>
<td>-0.0000267</td>
<td>-0.0000257</td>
<td>-0.0000232</td>
<td>-0.0000301</td>
</tr>
<tr>
<td>(.0000114)</td>
<td>(.0000109)</td>
<td>(.0000070)</td>
<td>(.0000113)</td>
</tr>
<tr>
<td>0.000159</td>
<td>0.000128</td>
<td>0.000122</td>
<td>0.000129</td>
</tr>
<tr>
<td>(.000049)</td>
<td>(.000047)</td>
<td>(.000039)</td>
<td>(.000048)</td>
</tr>
<tr>
<td>0.541</td>
<td>0.509</td>
<td>0.506</td>
<td>0.568</td>
</tr>
<tr>
<td>(.128)</td>
<td>(.121)</td>
<td>(.100)</td>
<td>(.124)</td>
</tr>
<tr>
<td>-0.000515</td>
<td>-0.000431*</td>
<td>-0.000388</td>
<td>-0.000537</td>
</tr>
<tr>
<td>(.000250)</td>
<td>(.000240)</td>
<td>(.000198)</td>
<td>(.000243)</td>
</tr>
</tbody>
</table>

GLS results are from the last iteration, when estimating $\sigma^2_e$ and $\sigma^2_g$. $R^2$ for this model was calculated as:

$$R^2 = \frac{\text{CSS}_{\text{model}}}{\text{CSS}_{\text{total}}} = \frac{\text{CSS}_{\text{total}} - \text{SS}_{\text{error}}}{\text{CSS}_{\text{total}}} = 1 - \frac{\text{SS}_{\text{error}}}{\text{CSS}_{\text{total}}} ,$$

where $\text{SS}_{\text{error}} = [ n - r(W X)] * \sigma^2_e ,$

$$= (8323 - 231) * 405.3509 ,$$

$$= 8092 * 405.3509 ,$$

and therefore,

$$R^2 = 0.558 .$$

When comparing MSEs it is clear that GLS estimation is not minimizing MSE but $V(\hat{y} - \hat{y})$, where $\hat{y} = X\hat{b} + Z\hat{s} + \hat{w}$. Usage of the term "mean squared error of prediction" is only adding unnecessary confusion.

Some of the regression coefficients are not statistically ($p < 0.03$) different from zero. Attention is called to the contents of Table 7 and the effects of the VIF (Variance Inflation Factor) on the standard errors of these estimates.

Tables 9 and 10 present analysis of variance tables and Type I (sequential) sum of squares of some sources of variation. In Table 10, if SS for STAT1 and STAT2 are summed then the SS for the original classes (3) is obtained. Table 11 presents the Type IV SS for all the regressors and this is the only type of SS that should be used when looking at regressors. Tables 9, 10, and 11 were obtained by PROC GLM of SAS.
Table 9. Analysis of variance table for model (OLS - complete given by Equation (38))

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean square</th>
<th>F Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>356</td>
<td>4,204,771.502</td>
<td>11,811.156</td>
<td>29.21</td>
</tr>
<tr>
<td>Error</td>
<td>7966</td>
<td>3,221,444.803</td>
<td>404.399</td>
<td></td>
</tr>
<tr>
<td>Corr. Total</td>
<td>8322</td>
<td>7,426,216.304</td>
<td>$R^2 = 0.5662$</td>
<td></td>
</tr>
</tbody>
</table>

Table 10. Type I (sequential) sum of squares for the first four sources in the model (OLS-complete) given by Equation (38)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Type I SS</th>
<th>F Value</th>
<th>P(F*Value &gt; F Table)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG</td>
<td>213</td>
<td>2,712,298.560</td>
<td>31.49</td>
<td>0.0001</td>
</tr>
<tr>
<td>SIREs</td>
<td>126</td>
<td>333,757.514</td>
<td>6.55</td>
<td>0.0001</td>
</tr>
<tr>
<td>STAT1</td>
<td>1</td>
<td>17,905.061</td>
<td>44.28</td>
<td>0.0001</td>
</tr>
<tr>
<td>STAT2</td>
<td>1</td>
<td>26,301.969</td>
<td>65.04</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Table 11. Type IV\textsuperscript{a} sum of squares for sires and regressors in the model (OLS-complete) given by Equation (38)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Type VI SS</th>
<th>F Value</th>
<th>P(F*Value &gt; F Table)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIRES</td>
<td>126</td>
<td>112,166.682</td>
<td>2.20</td>
<td>0.0001</td>
</tr>
<tr>
<td>STAT1</td>
<td>1</td>
<td>898.016</td>
<td>2.22</td>
<td>0.1362</td>
</tr>
<tr>
<td>STAT2</td>
<td>1</td>
<td>359.380</td>
<td>0.89</td>
<td>0.3459</td>
</tr>
<tr>
<td>AOD11</td>
<td>1</td>
<td>26,069.793</td>
<td>64.47</td>
<td>0.0001</td>
</tr>
<tr>
<td>AOD12</td>
<td>1</td>
<td>6,855.344</td>
<td>16.95</td>
<td>0.0001</td>
</tr>
<tr>
<td>AOD21</td>
<td>1</td>
<td>18,796.665</td>
<td>46.48</td>
<td>0.0001</td>
</tr>
<tr>
<td>AOD22</td>
<td>1</td>
<td>3,754.446</td>
<td>9.28</td>
<td>0.0023</td>
</tr>
<tr>
<td>AOD31</td>
<td>1</td>
<td>6,941.590</td>
<td>17.17</td>
<td>0.0001</td>
</tr>
<tr>
<td>AOD32</td>
<td>1</td>
<td>393.961</td>
<td>0.97</td>
<td>0.3237</td>
</tr>
<tr>
<td>SF1</td>
<td>1</td>
<td>2,629.603</td>
<td>6.50</td>
<td>0.0108</td>
</tr>
<tr>
<td>SF2</td>
<td>1</td>
<td>2,067.630</td>
<td>5.11</td>
<td>0.0238</td>
</tr>
<tr>
<td>SF3</td>
<td>1</td>
<td>1,200.755</td>
<td>2.97</td>
<td>0.0849</td>
</tr>
<tr>
<td>SS1</td>
<td>1</td>
<td>2,810.912</td>
<td>6.95</td>
<td>0.0084</td>
</tr>
<tr>
<td>SS2</td>
<td>1</td>
<td>4,649.133</td>
<td>11.50</td>
<td>0.0007</td>
</tr>
<tr>
<td>SS3</td>
<td>1</td>
<td>2,219.938</td>
<td>5.49</td>
<td>0.0192</td>
</tr>
<tr>
<td>SS4</td>
<td>1</td>
<td>4,165.089</td>
<td>10.30</td>
<td>0.0013</td>
</tr>
<tr>
<td>CA1</td>
<td>1</td>
<td>7,261.165</td>
<td>17.96</td>
<td>0.0001</td>
</tr>
<tr>
<td>CA2</td>
<td>1</td>
<td>1,717.949</td>
<td>4.25</td>
<td>0.0393</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Type IV, in colloquial language, are the SS calculated for each factor as if that factor was included last in the model.
Table 12. Type IV sum of squares for the regressors in the GLS model given in Equation (39)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Type VI SS</th>
<th>F Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT1</td>
<td>1</td>
<td>875.939</td>
<td>2.16</td>
</tr>
<tr>
<td>STAT2</td>
<td>1</td>
<td>468.557</td>
<td>1.16</td>
</tr>
<tr>
<td>AOD11</td>
<td>1</td>
<td>27,255.795</td>
<td>67.24</td>
</tr>
<tr>
<td>AOD12</td>
<td>1</td>
<td>7,255.760</td>
<td>17.90</td>
</tr>
<tr>
<td>AOD21</td>
<td>1</td>
<td>19,762.362</td>
<td>48.75</td>
</tr>
<tr>
<td>AOD22</td>
<td>1</td>
<td>3,982.584</td>
<td>9.83</td>
</tr>
<tr>
<td>AOD31</td>
<td>1</td>
<td>7,585.053</td>
<td>18.71</td>
</tr>
<tr>
<td>AOD32</td>
<td>1</td>
<td>440.131</td>
<td>1.09</td>
</tr>
<tr>
<td>SF1</td>
<td>1</td>
<td>2,486.283</td>
<td>6.13</td>
</tr>
<tr>
<td>SF2</td>
<td>1</td>
<td>1,913.230</td>
<td>4.72</td>
</tr>
<tr>
<td>SF3</td>
<td>1</td>
<td>1,101.064</td>
<td>2.72</td>
</tr>
<tr>
<td>SS1</td>
<td>1</td>
<td>2,645.395</td>
<td>6.53</td>
</tr>
<tr>
<td>SS2</td>
<td>1</td>
<td>4,203.185</td>
<td>10.37</td>
</tr>
<tr>
<td>SS3</td>
<td>1</td>
<td>2,881.225</td>
<td>7.11</td>
</tr>
<tr>
<td>SS4</td>
<td>1</td>
<td>2,982.529</td>
<td>7.36</td>
</tr>
<tr>
<td>CA1</td>
<td>1</td>
<td>8,447.306</td>
<td>20.84</td>
</tr>
<tr>
<td>CA2</td>
<td>1</td>
<td>1,986.837</td>
<td>4.90</td>
</tr>
</tbody>
</table>
Table 12 presents Type IV SS for the regressors when using the GLS model, from the last iteration of the REML program (Appendix I). Small changes in Type IV SS in the OLS and GLS model are associated with changes in the estimated regression coefficients given in Table 8.

Some problems with the WLS and WTROB models are as follows:

1) Within CG error variance estimates $\sigma^2_{e_i}$ were used blindly even when the number of observations was too small. This caused observations from some CG (with small $\sigma^2_{e_i}$) to become too influential. Some "good" observations from some CG had very small weights because a few observations in the same CG inflated considerably the common $\sigma^2_{e_i}$.  

2) Residuals used to calculate the individual weights for the observation were calculated only once, from the OLS incomplete (without sires) model. Robust procedures are, always, iterative procedures. Residuals to be used on the WTROB should have been calculated from the complete OLS model. Also, to be more effective, a robust estimate of dispersion (like the MAD) should have been used in WTROB.

These two problems, fruits of the inexperience of the author, reduced the improvement these methods could bring about, especially in the complete models. Estimates of variance components were very sensitive to the above misuses. Estimates of $h^2$, obtained from Type IV sums of squares of PROC GLM of SAS (1982b) were 0.1212, 0.1020, and 0.0536, for OLS, WLS, and WTROB, respectively.
These results precluded the author from using GLS estimation with \( R \neq I \sigma_e^2 \).

**B. Environmental Effects**

1. **Dam nursing status (STAT1 and STAT2)**

In Chapter V, Subsection A.1.a, the reparameterization of the three classes into the two regressors, corresponding to the restriction that the sum of the three effects be equal to zero, was described.

The most interesting contrast (estimable function) is STATUS 2 - STATUS 3, or the difference between a repeat calver and an alternate calver. This is estimated by:

\[
\hat{\alpha}_1 + 2\hat{\alpha}_2 = \text{STAT1} + 2 \times \text{STAT2} = \\
= (-0.647882) + 2(-0.547698) \\
= -1.743278 \text{ Kg}.
\]

The variance of this contrast is given by:

\[
V(\text{STAT1} + 2\text{STAT2}) = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 0.00047920 & -0.00027096 \\ 0.00027096 & 0.00064020 \end{bmatrix} \begin{bmatrix} 1 \\ \text{Sym} \end{bmatrix} \sigma_e^2 \\
= (0.00195616) 405.35090256 \text{ Kg}^2
\]

\[
V(\text{STAT1} + 2\text{STAT2}) = 0.7929312 \text{ Kg}^2;
\]

and its standard error is 0.8904668 Kg.

Since \( t = \frac{-1.743278}{0.8904668} = 1.9577 \) and \( |t| = t_{0.05,8092} = 1.96 \), all one can say is that, with a bordering statistical significance, cows which calf in alternate years wean calves heavier than cows which calve in
It is surprising that the difference is so small. The difference in size and body condition among calves of these two groups suggest a much larger influence from previous lactating status of their dams.

The difference found by Vorster (1964) was of -15.9 Kg, but the way the results are presented gives the impression that this author analyzed each of the effects separately in simple models; therefore, the difference may contain some confounded effects.

Cardellino-Stercken (1983), using partly the same data as in this study, in a model containing cow (random) effects, found differences of -5 Kg, -25 Kg, and -13 Kg, for each of three different herds.

If the results obtained in this study were a reflection of the true situation, then probably calving rates would not be so low in RS. The possibility that the differences are larger in less favorable environment deserves to be investigated.

Maybe the fact that, on the average, earlier born calves are from cows which were dry the year before is helping to suggest a higher expectation for this contrast.

The only action that can be suggested at this point is to stimulate breeders to make more complete enrollments of their calf crops (reduce the percentage of STATUS 1 — "no information" — cows) and to include this factor in regular animal evaluations, since it will make due corrections in favor of cows with consecutive calving. This procedure will also decrease bias in sire evaluations since there is a tendency to AI mostly heifers and dry cows.
2. **Age of dam**

Dairy and beef cows have an optimum age for milk production, with young and old cows producing less milk than mature cows.

Lumping together effects smaller than 1.8 Kg. (41b.) one can consider cows mature when they are 6 - 9 years old, for dams of male calves; and from 6 - 10 years of age, for dams of female calves. For these two groups, peak production was attained at 7 and 8 year-olds, respectively.

Using the regression coefficients presented in Table 8, plots of age of dam effects, for male and female calves, are presented in Figure 2. The shapes of the curves resemble each other very much. For young dams of males, the AOD effects are more pronounced than for young dams of females.

AOD3 (nested in sex) measures the change in curvature for "post-maturity" dams from "pre-maturity" dams. The curvature is less steep for older ages than for younger ones, especially for dams of males. The quadratic effects, for all ages, is larger in dams of males than of females. If the curvature is smaller for the whole range of AOD in cows with female calves it is clear that there exists smaller chances for a change in the regression system to be statistically detected. This may be the reason why the regression coefficient for AOD32 was not significantly different from zero in most models. But, since there is no big harm in keeping this component in the system, its maintenance is suggested, especially if a more precise method of estimation, like W-estimation, is used. Using knots at years 7 (for dams of males) and 8
Figure 2. Age of dam (within sex of calf) effects on weaning weight.
(for dams of females) should also be tried.

Since cows, as a random effect, were not included in the models, there is always the possibility that the estimates are influenced by selection of dams. Older dams included in this study are all pedigree (closed herdbook) cows, selected more by their pedigree lines than for milk production.

Effects of dam age, as estimated in this study, are within their expectations, in agreement with the major part of the literature, and should be used in future genetic evaluations.

The ROP program in RS is using the multiplicative factors, as recommended by BIF (1972), according to AOD (in parentheses), as follows: $1.10(3)$, $1.05(4)$, $1.00(5-10)$, and $1.05(11-16)$. If the additive effects as found in this study are transformed to multiplicative ones, the following correction factors would be found: $1.11(3)$, $1.05(4)$, $1.01(5)$, $0.99(6)$, $0.99(7)$, $0.99(8)$, $1.00(9)$, $1.01(10)$, $1.03(11)$, $1.05(12)$, $1.08(13)$, $1.12(14)$, $1.17(15)$, and $1.23(16)$. The two sets of values agree remarkably well and the use of BIF's (1972) correction factors have been useful in reducing error variance.

3. **Seasonality within fall effects**

Using the GLS estimates of the regression coefficients of WW on SF1, SF2, and SF3, WW predicted values were plotted in Figure 3, keeping all other variables constant.

Seasonality effects within fall, isolated from non-linear age of calf effects, are not large by themselves. Calves born more toward the end of the fall season are slightly favored while calves born during late
Figure 3. Seasonality within fall effects on weaning weight
summer are lighter at weaning. There are justifications for this to happen, as follows:

1. Calves born during summer are expected to be lighter at birth than fall born ones due to the heat stress their dams suffer. As a mechanism to dissipate body heat, cows (animals in general) have their blood flow directed more to the surface (peripheral vasodilatation); thus, the placenta is less irrigated and the fetus receives less nutrients. Newborn calves can more easily get a navel infection or diarrhea if they are born in summer than if during late fall.

2. Mid- to late-fall born calves begin to consume forage at an age (about 5 months) when the fresh spring grass is sprouting. Summer born ones can only find dry, low quality, roughage at that same age.

Notice from Table 8 that the regression coefficient for SF3 is not statistically different from zero. The regression coefficient estimates the change in curvature (quadratic) after day 85 (Julian). In a future study, with more observations, it is recommended to better estimate this knot. The non-significance may also be just the product of the VIF (Variance Inflation Factor).

4. **Seasonality within spring effects**

Using the estimates given in Table 8 for the regression coefficients of WW on SS1, SS2, SS3, and SS4, the predicted values for WW were plotted in Figure 4, keeping all other variables constant.

Looking only at the linear trend, fall born calves do better if they
Figure 4. Seasonality within spring effects on weaning weight
are born late in the season while the opposite is the case for spring-born ones.

Breeders in RS say that the best calves are the ones born early in the spring. Several ROP reports for large herds gave higher WWR for calves born around October and therefore the author became suspicious that the saying was reflecting more age-of-calf than pure seasonality effects. Figure 4 is strongly backing the position and knowledge of the breeders.

Calves born at the end of winter have the minimum amount of health problems. Their dams can match the peak of lactation with the most exuberant and nutritious grass of the whole year.

Calves born during summer are lighter at birth, suffer more health hazards, and their dams may have their peak of lactation greatly reduced because it will coincide with the short summer drought (mid January to mid February in some years).

Calves born in fall or in spring are managed differently. Fall born calves are weaned in November-December, while spring born ones are weaned in April-May. There are only 10 days between the end of the fall birth season and the beginning of the spring one, but since animals born at these extremes are managed in separate groups, with different practices, it is not expected that the two curves (from Figure 3 and Figure 4) will give the idea of continuity in the "empty" periods. No constraint of continuity at the ends of the curves was imposed; letting SS3 and SS4 completely free makes the occurrence of this continuity even less likely; but even so, if one looks at the sequence SS-SF one gets an idea of continuity. An allowance has to be given to the change in management
between the two seasons, resulting in a "jump" between the curves.

As will be shown in sequence, spring effects had opposite effects to the non-linear age of calf regressions, in general terms. Calves born early in spring are favored by the environmental conditions associated with the period of time they are born, but are not favored by their older age (the rule is a fixed weaning date) at weaning. The reverse of this situation (late born = disfavored, but young age = favored) is absolutely true also, for spring born calves. The two effects almost "balance" each other out.

For fall born calves, the two effects (although fall effects are much smaller) are synergistic: early born = weaned old implies double disfavoring; late born = weaned young implies double favoring.

Maybe these sets of effects could be estimated in the present study because calves born in two different seasons were analyzed together. It only spring or only fall born calves were studied (studied disjointly), maybe some null or erratic effects (due to collinearity) would also be found in this study, in consonance with most of the reviewed papers.

5. Non-linear age of calf effects

Linear and quadratic GLS estimates, from Table 8, of regression coefficients of WW on CA1 (age of calf) and CA2 (age of calf squared) were used to produce the predicted WW, holding all other effects constant. These are plotted in Figure 5. Perpendicular to the abscissa are lines indicating the recommended correction age (205 days) and the limits to consider growth as linear on age, by BIF (1972, 1981). The straight line is the best fitting slope for the predicted values.
Figure 5. Non-linear age of calf effects on real weaning weight (kg)
Looking at Figure 5, it is possible to understand why so many researchers have completely dismissed any deviation from absolute linearity as not being important.

Figure 5 by itself is misleading because the intercept (using average values from Table 5 and regression coefficients from Table 8) is 46.61 Kg, while the average birth weight used are 32 Kg (females) or 35 Kg (males). Figure 5 is saying that if two weights from a calf were available, one before and one after 205 days, a linear interpolation would give very good results.

The estimated linear regression coefficient was 0.568 Kg/d. The real average daily gain (ADG) mean was 0.5191 Kg/d, with an average birth weight of 33.1 Kg and an average age at weaning of 213.5 days. Using the generated data to plot Figure 5 (uniformly distributed between days 155 and 295, with a mean of 205 days) and a birth weight of 33 Kg, ADG's were calculated according to BIF (1972, 1981). The ADG mean for this generate data was 0.531 Kg/d. This is saying that ADG is a poor estimator of the rate of growth around 205 days of age.

All predicted weights in Figure 5 were generated using the same regression coefficients. Thus, all have the same "genetic merit." ADGs for these calves are plotted in Figure 6 against their ages. If the method recommended by BIF (1972, 1981) were efficient, all ADGs should be the same; or independent of age of calf (AOC); or should form a line parallel to the AOC-axis. In Figure 6, the ADG mean is represented by such a line, at ADG=0.531 Kg/day.

If one uses the calculated ADG to compute a "corrected" WW, according to BIF (1972, 1981), then, again, "corrected" WW should be
Figure 6. Non-linear age of calf effects on average daily gain (kg/day)
independent of AOC. "Corrected" WW are plotted against AOC in Figure 7.

Considering the limits and procedure suggested by BIF (1972, 1981)
then a 160 day-old calf will have his WW over-adjusted by 9.1 Kg; a 250
day-old calf will have his WW under-adjusted by 7.7 Kg. Biases are
larger for younger than for older calves. In KS, given the long breeding
seasons, calves span more than 90 days of age. It is not feasible to
have two weaning dates. Of course the proportion of very young or old
animals is small, but these animals do occur and this situation requires
proper corrections.

After seeing the effects, one-by-one, it is time to look at these
effects together. Figures 8 and 9 show the addition of non-linear age of
calf (the same "corrected" WW from Figure 7) to seasonality within fall
effects. In Figures 8 and 9 the synergism between the two effects can be
seen.

In Table 2, the strong association between AOC and SF1 is shown.
This implies that large sections of the surface in Figures 8 and 9 do not
occur in practice. Using the cardinal points, with N (north) in the top
of the table, the absolute mass of records is gathered closely around the
line going from S (south) to N (S is at the bottom of the table).

Figures 10 and 11 present the addition of non-linear age of calf to
seasonality within spring effects. Again, the concentration of real data
points is around the line S to N.

The two sets of effects, for spring born calves, greatly counteract
each other. Intermediate values on the surface of figures 10 and 11 are
along the S-N line. This line, in Figures 8 and 9, holds the extreme
values on that surface.
Figure 7. Non-linear age of calf effects on corrected weaning weight (kg)
Figure 8. Seasonality within fall (SFJ) and non-linear age of calf (age) effects on weaning weight (kg)
Figure 9. Seasonality within fall (SF1) and non-linear age of calf (age) effects on weaning weight (kg). (Surface is more tilted over age-axis than in Figure 8)
Figure 10. Seasonality within fall (SF1) and non-linear age of calf (age) effects on weaning weight (kg)
Figure 11. Seasonality within spring (SS1) and non-linear age of calf (age) effects on weaning weight (kg). (Part of the bottom surface shown)
The addition of some interaction terms (SF*AOC and SS*AUC) to the model this would possibly allow for the estimation of more smooth surfaces.

6. Weaning date effects

Figures 8 and 9, for fall born calves, and Figures 10 and 11, for spring born ones, present overall pictures of the situation but are too general to comment on and to help in the decision taking.

Lines from these surfaces were extracted by considering fixed weaning dates. Adding a fourth variable to the system brings clarity to the effects of the other two explanatory variables in the response variable, WW "corrected" linearly for AOC according to BIF (1972, 1981).

Breeders decide when calves will be born; nature impose its effects associated with the chosen birth dates; the biological clock of the calves determines the shape of their (potential) growths curves; and breeders decide on how many buffering elements they will place (given economical feasibility to do so) between the animals and their environment and when they will wean the animals and record the WWs. This is the real picture, all these elements are needed, and comments based on incomplete models are only partial comments. As was discussed in Chapter IV, it is not possible to include in the model the birth date of each calf, his age at weaning and the actual weaning date since there is one absolute linear dependency amongst these three variables. Weaning date has effects on average adjusted WW of a calf crop, on postweaning gain (at least on a pasture based system of production), and, most important of all, on birth rates of the next two calf crops. Weaning date is a seldom
studied variable that deserves more attention. Breeders of different regions and even counties in RS, by trial-and-error, arrived at "best" periods, but more light is needed to be shed on this problem.

Figures 12 and 13 show the joint effects of weaning date, age of calf and date of birth on "corrected" WW of fall born calves. Figure 12 can be considered as different lines from Figure 9, each line perpendicular to the AGE-axis; Figure 13 as lines perpendicular to the SFI-axis.

Figure 12 shows in a dramatic way that the way the present ROP program in RS is handling the corrections of WW is a very poor one. Although birth date within fall effects are not large they add on to the age of calf effects.

If the corrections were working perfectly, all lines should be parallel to the abscissa. Late born and weaned early calves are being under-corrected up to 30 Kg; early born and weaned late calves are being overcorrected up to 16 Kg. These are large errors.

The earlier the weaning date the higher the corrected WW and also the greater the errors. The majority of fall calves are born in RS in April and May (Julian days 90 to 150) and are weaned from mid-November to mid-December. Use of AI is heavy for the fall calf crop; this is one of the reasons for the shorter breeding and birth seasons in most of the ranches.

The ROP program offers to the breeders a sire evaluation (within herd, within calf crop) based on selection index theory, every time a weaning or yearling report is produced. At the onset of the program, a breeder (with large calf crops — herds CAA and CAB in Table 1) began to notice changes in ranking of sires from one calf crop to the next.
Figure 12. Effect of weaning date on corrected weaning weight according to date of birth within fall
Figure 13. Effect of weaning date on corrected weaning weight of fall-born calves according to age of calf.
Several causes were suggested: different groups (only AI vs AI plus natural service) of sires used in the winter and summer mating seasons; different groups of dams (only STATUS 3 vs STATUS 2 and 3); selection of dams for each sire; sampling errors and the ubiquitous (credit goes to Professor Oscar Kempthorne; Kempthorne (1983)) genetic by environmental interaction. Those are all speculations and each may be playing a role. But Figure 12 shows that if sire progenies are born with 10 to 20 days of difference in their mean progeny age this will have a large effect on the sire estimates for "corrected" WW.

If no adjustment is made for date of birth and non-linear age of calf effects and a sire * CG interaction term is included in the model, then a big, non-genetic, source is there to make this "interaction" sizable. As a mere speculative point here, this situation could explain in part findings of Bertrand (1983), who estimated sire * CG variances of the same magnitude as sire variances.

When plotting real WW by date of birth within fall, the curve shows that there is no increase in real WW if the calf is born in any day between days 50 to 80. If a calf is born in late February or during March, his real WW will not be higher than if he is born at the beginning of April. This is another way the joint effects of seasonality within fall and non-linear age of calf are showing up.

Figure 13 reinforces the statement that the effects are more pronounced the earlier the calves are weaned and the younger they are.

Figures 14 and 15 show the joint effects of weaning date, non-linear age of calf and date of birth on WW "corrected" for linear age of calf, of spring born calves.
Comparing Figures 14 and 12, it is possible to conclude that the situation for spring born calves is not as bad as for fall born ones. The overall trend is the same, though. Late born, young weaned calves have higher "corrected" WWs.

A perpendicular line to the abscissa, at birth date (Julian) 315 allows for consideration of the weaning date effects. Points on this line show WW of calves born on the same date, all weights "corrected" to 205 days show a range in the weights of up to 34 Kg. This differences can be credited to the non-linear age of calf effects.

Figure 14 is consistent with earlier findings. The ROP program calculates the average WWR for groups of the same sex and management code for each calf crop in each herd. Most of those averages were indicating that October and November should be the months for calves to be born, in order to maximized "corrected" WW. That statement is correct; what is wrong is the form the WW are "corrected." The month of December normally have too few observations in each CG to give any patterned results.

Animals born in the period of 50 days, between mid-October to the end of November, if all weaned late (June, 14) have WW properly corrected by BIF (1972, 1981) adjustment recommendations. A good number of calves are born in this period although the highest concentration is during September. If one works with such a data base (animals born mostly in October and November — or, maybe in a larger corresponding period in the US); if one considers that for spring born calves the seasonality effects and the non-linear age of calf effects are antagonistic to each other (like in this 50-day period, they cancel each other); if one judges the adequacy of a model by the reductions of sum-of-squares and not by the
Figure 14. Effect of weaning date on corrected weaning weight according to date of birth within spring
Figure 15. Effect of weaning date on corrected weaning weight of spring born calves according to age of calf.
importance of the regression curves; and, if on recalls that in the early '50s the amount of calculations were a big part of the chores in running a ROP program, then the simple recommendations of BIF (1972, 1981) make sense. For the bulk of data, the simple model works remarkably well (probably, in the U.S., seasonality effects have a smaller effect on calves performance given the supplementary feeding levels practiced), but since breed associations own their own computers now, computing time is not a factor anymore in the compromise between efficiency and parsimony of competing models.

In Figure 14, curves get weird for very late born and weaned calves. It doesn't seem realistic. Probably the addition of more records in a future analysis and some interaction terms among regressors may help to smooth things up.

Figure 15 is corroborating with all the above arguments. It also shows that, for calves weaned on (DATE 7) June 14, calves with ages between 195 and 245 days have their weights correctly "corrected"; the same is true for (DATA 6) May 30, ages 180 to 230 days; and the same is true for (DATE 5) May 15, ages 160 to 220 days.

At this point, it is important to make a remark that spring and fall records should be handled differently. Spring seasonality effects are balanced out by the non-linear age effects. Fall seasonality effects are much smaller than spring ones but are in the same direction as non-linear age effects and large biases result from the lack of proper adjustment.
C. Heritability and Error and Sire Variance Estimates

PROC GLM of SAS (1982b) produced sires' Type IV SS (Table 11) and, using the option RANDOM, the $E(MS_{sires}) = \sigma_e^2 + 38.4450962 \sigma_s^2$, which, solved, produced $\hat{\sigma}_s^2 = 12.636526$. Thus, $h^2 = 0.1212032$ and $k = \hat{\sigma}_e^2 / \hat{\sigma}_s^2 = 32.002412$.

This value of $k = 32.0$ was used to obtain the first "low" and "high" prior values (22.0 and 42.0) to use in program REML (Appendix I). Table 13 presents the results from the 6 stages the CIA was used. For each stage, 2 inverses of the LHS were needed. The following observations can be made from results presented in Table 13:

1. From the first stage (first converged value), a heritability estimate of 0.157 was obtained. This was obtained from ratios which correspond to $h^2$ of 0.174 and 0.093.

2. From the second stage, $h^2$ was 0.155 and from the third stage $h^2$ was 0.155002. At the sixth stage, the final estimate of $h^2$ obtained was 0.1550005. This shows how fast a final value is obtained when using the CIA. For all practical purposes, the second stage would have given a very good approximate value already.

3. REML estimates of $\sigma_e^2$ are very robust to the use of incorrect initial estimates of the variances ratio.

4. The product prior ratio by corresponding trace of the inverse is also robust.

5. REML estimates of $\sigma_s^2$ and $\hat{e}'\hat{e}$ (the numerator of $\hat{\sigma}_e^2$) are the elements which are bouncing the most.
The program REML automatically calculates the new "low" and "high" estimates as $\alpha_c - \Delta$ for each 2 stages. That means that at stages 3 and 5 the "priors" were educated guesses, based on previous rounds. These educated guesses allowed for significant savings in the number of stages. As an example, at the third stage, the automatic calculated values would be 23.22 and 27.63, which are more spread than the guessed 24.5 and 24.9.

If instead of $\alpha_c - \Delta$, the new ratios to be used in the second stage as "priors" were calculated as

$$\frac{\alpha_c - \Delta}{\alpha_c - \text{OLD}},$$

were OLD are the priors used in that stage, then the new priors to be used would be 24.09 and 24.93, by far tighter around the final converged value than the ones calculated by the program. This formula is thus recommended since it can save a complete stage, but its use is not recommended after the first stage since $(\alpha_c - \text{OLD})$ can be smaller than 1.0. After stage one, $\alpha_c - \Delta/2$ can be used, as is shown in Appendix G.

The CIA was completely satisfactory since it converges very fast and it gives bounds to the converged values at each stage. Only half or less of the number of stages (6) would be needed if the only objective of the study was to obtain an estimate; later rounds allowed the author to gain some more insight.

The estimate of $h^2$, 0.15500058703, is saying that at least some of the causes listed in the introduction, hypothesized to explain previous estimates of $h^2$ for WW in RS around 0.0, may be true. The $R^2$s obtained, even for I-OLS, are more in line with reports from the U.S. The estimate of $h^2$ found in this study is similar to the ones reported by Bertrand (1983).
for Polled Hereford field data but lower than the one reported by Wilson (1984) of 0.25 for Angus field data, from the top herds of the breed.

In Table 11, the F-value for sires is associated with a probability smaller than 0.0001 that the true unknown value of \( \sigma_s^2 \) is zero. This indicates that \( h^2 \) is also not zero.

Animal breeders have traditionally expressed \( h^2 \) values with associated approximate standard errors and confidence intervals (CI). An approximation frequently used (when \( h^2 \) is based on paternal half-sib families) is

\[
V(h^2) = \frac{32h^2}{T}
\]

where \( T \) is the total number of progeny. For this study, this approximate value is

\[
V(h^2) = \frac{32(0.1550005)}{8323} = \frac{4.960016}{8323} = 0.00059594
\]

\[
S_{h^2} = [V(h^2)]^{0.5} = 0.0244119
\]

The value of \( S_{h^2} \) is sometimes used to establish CIs using values from a t-table, even if \( t \) is a symmetric distribution while that distribution of \( h^2 \) has to be asymmetric since the parameter space of \( h^2 \) is limited to the interval (0,1). Using this gross approach, a 95% CI would be

\[
0.1071531 \leq h^2 \leq 0.2028478
\]

while a 99% CI would be

\[
0.0921154 \leq h^2 \leq 0.2178855
\]

Harville and Fenech (1984) developed procedures to construct an
exact confidence interval for a variance ratio or for heritability. Also procedures for approximate CIs are developed and results are illustrated with an animal breeding example. One of these approximate methods (G1, as defined by the above authors) was used in this study and it is said to "provide a relatively good approximation ... for small values of \( \gamma \)," where \( \gamma = \frac{\sigma^2}{s^2} \frac{e^2}{e} \). In this study, \( \gamma = 24.806359^{-1} = 0.0403122 \). In Figure 1 of Harville and Fenech (1984), the pivotal quantities for constructing exact and approximate confidence intervals for the variance ratio \( \gamma \) are plotted, and for values of \( \gamma \) smaller than 0.1 the curves for \( G_1 \) and the exact method are superimposed. In the whole region of interest for animal breeders \( (0 \leq h^2 \leq 1.0 \text{ or } 0 \leq \gamma \leq 1/3) \), these two curves are very close; the closer as \( \gamma \) is nearer to zero.

Define (in Harville and Fenech, 1984, notation):

\[
r = \text{rank}(XWZ) - \text{rank}(XW) = (214 + 17 + 130 - 4) - (214 + 17) = 126;
\]

\[
f = n - \text{rank}(XWZ) = 8323 - 357 = 7966;
\]

\[
\gamma = \frac{\sigma^2}{s^2} \frac{e^2}{e} = \kappa^{-1} = 0.0403122;
\]

\[
K = (1/r) \text{tr}(C), \text{ where } C = Z'(I - P_{XW})Z,
\]

\[
P_{XW} = [XW] \begin{bmatrix} X'X & X'W \\ W'X & W'W \end{bmatrix}^{-1} \begin{bmatrix} X' \\ W' \end{bmatrix},
\]

and \( K = 38.9918795 \), given by PROC GLM of SAS (1982b) from
EMS(Sires) associated with the Sires' Type IV SS. From this same output, \( F = 2.20 = \text{EMS(sire)}/\text{EMS (Residual)} \).

The approximate 100(1 - \( \alpha \))% CI for \( \gamma \) is

\[
\frac{F - F^*_\alpha_1}{\alpha_1} \leq \frac{F - F^*_1 - \alpha_2}{\alpha_2} \leq \frac{F - F^*_1 - \alpha_2}{\alpha_2}, \quad \text{where}
\]

\( F^* \) are table values from the F distribution and \( \alpha = \alpha_1 + \alpha_2 \), with \( \alpha_1 = \alpha_2 = \alpha/2 \) possible. \( F^* \) used should be with \((r,f) = (126, 7966)\) degrees of freedom but values from the \((120, \infty)\) were used.

The approximate 95% CI for \( \gamma \), using

\[
F^*_{0.025}(120, \infty) = 1.27, \quad \text{and}
\]

\[
F^*_{0.975}(120, \infty) = \frac{1}{F^*_{0.025}(\infty, 120)} = \frac{1}{1.31} = 0.7633587, \quad \text{is:}
\]

\[
\frac{2.20 - 1.27}{38.991879(1.27)} \leq \frac{2.20 - 0.7633587}{38.991879(0.7633587)}
\]

\[
\gamma_L = 0.0187804 \leq \gamma \leq 0.0482664 = \gamma_H;
\]

hence,

\[
P[k^*_L = 20.718319 \leq \frac{\sigma^2}{\sigma^2} \leq 53.246975 = k^*_H] \approx 0.95 ;
\]

and,

\[
\text{using } h^2 = \frac{4\gamma}{1 + \gamma^*},
\]

\[
P \left[ h^2_{L} = 0.0737368 \leq h^2 \leq 0.1841768 = h^2_{H} \right] = 0.95 .
\]
The approximate 99% for \( \gamma \), using

\[
F^*_{.005(120,\infty)} = 1.36 \quad \text{and}
\]

\[
F^*_{.995(120,\infty)} = \frac{1}{F^*_{.005(\infty,120)}} = \frac{1}{1.43} = 0.6993007 \quad \text{is:}
\]

\[
\frac{2.20 - 1.36}{39.991879(1.36)} \leq \gamma \leq \frac{2.20 - 0.6993007}{38.991879(0.6993007)}
\]

\[\gamma = 0.015443 \leq \gamma \leq 0.0550371 = \gamma^*_H ;\]

hence,

\[
P \left[ k_L^* = 18.169562 \leq \frac{\sigma^2}{\sigma^2} \leq 64.748756 = k_H^* \right] = 0.99 ;
\]

and,

\[
P \left[ h_L^2 = 0.0608376 \leq h^2 \leq 0.2086641 = h_H^2 \right] = 0.99 .
\]

Notice that using the REML estimate, 0.1550, the situation \((\hat{h}^2 - \hat{h}_L^2) > (\hat{h}_H^2 - \hat{h}^2)\) is occurring for both intervals while one, intuitively, is expecting the reverse if \(h^2\) is more close to 0.0 than to 1.0. For the "fitting constants" estimate, 0.1212, the intervals are asymmetric, as expected.

Since \(\hat{\sigma}^2_s\) is significantly bigger than zero and neither CI for \(h^2\) includes zero it is reasonable to conclude that there exits identifiable genetic variation in WW of Hereford cattle in RS. The estimated value for \(h^2\) is 0.155 and there are reasons to believe that the true value is above this estimate. The arguments for this are as follows:
From Table 13, it can be concluded that the single factor which is affecting the iterated variance ratios the most is \( \hat{s}'\hat{s} \). The extrapolation that the single most important factor on the estimates of \( h^2 \) is also \( \hat{s}'\hat{s} \) is plausible. As was discussed in Chapter IV, there are 3 disconnected sets of sires in the data used in this study. In Appendix J, these sets are as follows:

a. Set 1: sires 1 to 32;
b. Set 2: sires 33 to 80 and 82 to 130; and
c. Set 3: sire 81.

Each of these sets have the property that \( \hat{1}'\hat{s} = 0.0 \). If the three sets were connected, then \( \hat{s}'\hat{s} \) would be certainly bigger than in the present situation. In the worst situation, \( \hat{s}'\hat{s} \) would be equal in either situation if and only if the property \( \hat{1}'\hat{s} = 0.0 \) within each set remained true even with the sets connected. Just as an example, the probability is small that if sire 81 were connected to either set 1 or 2 his \( \hat{s} \) would still be zero. The probability that each of the sires from set 1, when used within set 2 would retain the same \( \hat{s} \) as in his original is small. The compound probability that this would happen for all sires from set 1 is all but improbable.

Schaeffer (1975) studied the effects of disconnectedness on MINQUE variance components. He simulated records for 150 calves from 10 sires in 12 CG forming 4 disconnected sets of data. His model 4 is a two-way mixed model, without interaction. In this model, as he deleted more and more
Table 13. Iterative REML estimates of error and sire variances, and their ratio, using CIA (Schaeffer, 1983)

<table>
<thead>
<tr>
<th>Stage</th>
<th>LHC&lt;sup&gt;a&lt;/sup&gt;</th>
<th>OLD&lt;sup&gt;b&lt;/sup&gt;</th>
<th>NEW&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Δ = (NEW - OLD)</th>
<th>REML $\hat{\sigma}^2_e$</th>
<th>$\hat{\sigma}^2_s$</th>
<th>tr($C_{22}$)&lt;sup&gt;d&lt;/sup&gt;</th>
<th>REML $\hat{\sigma}^2_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L</td>
<td>22.0</td>
<td>23.06</td>
<td>1.06</td>
<td>405.1</td>
<td>917.0</td>
<td>3.54</td>
<td>17.6</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>42.0</td>
<td>34.61</td>
<td>-7.39</td>
<td>406.2</td>
<td>452.0</td>
<td>2.18</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>24.51</td>
<td>24.51</td>
<td>0.519</td>
<td>405.21</td>
<td>858.0</td>
<td>3.38</td>
<td>16.9</td>
</tr>
<tr>
<td>2</td>
<td>L</td>
<td>23.45</td>
<td>23.97</td>
<td>0.519</td>
<td>405.21</td>
<td>858.0</td>
<td>3.38</td>
<td>16.9</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>31.90</td>
<td>29.01</td>
<td>-2.887</td>
<td>405.97</td>
<td>617.0</td>
<td>2.69</td>
<td>14.0</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>24.74</td>
<td>24.74</td>
<td>-0.036</td>
<td>405.36</td>
<td>806.0</td>
<td>3.23</td>
<td>16.30</td>
</tr>
<tr>
<td>3</td>
<td>L</td>
<td>24.5</td>
<td>24.62</td>
<td>0.118</td>
<td>405.32</td>
<td>820.0</td>
<td>3.27</td>
<td>16.46</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>24.9</td>
<td>24.86</td>
<td>-0.036</td>
<td>405.36</td>
<td>806.0</td>
<td>3.23</td>
<td>16.30</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>24.8061</td>
<td>24.8061</td>
<td>0.000</td>
<td>405.36</td>
<td>806.0</td>
<td>3.23</td>
<td>16.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>24.69</td>
<td>24.73</td>
<td>0.0458</td>
<td>405.339</td>
<td>813.</td>
<td>3.2550</td>
<td>16.388</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>24.84</td>
<td>24.83</td>
<td>-0.0139</td>
<td>405.354</td>
<td>808.</td>
<td>3.2403</td>
<td>16.326</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>24.806323</td>
<td>(0.155)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>24.8062</td>
<td>24.80626</td>
<td>0.0000616</td>
<td>405.35089</td>
<td>809.429</td>
<td>3.24375</td>
<td>16.34067</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>24.8065</td>
<td>24.80645</td>
<td>-0.0000545</td>
<td>405.35092</td>
<td>809.418</td>
<td>3.24373</td>
<td>16.34055</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>24.80635912976</td>
<td>(0.1550005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>24.8063</td>
<td>24.8063210</td>
<td>0.0000238</td>
<td>405.350902</td>
<td>809.4257</td>
<td>3.24375</td>
<td>16.34063</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>24.8064</td>
<td>24.8063931</td>
<td>-0.0000219</td>
<td>405.350914</td>
<td>809.4214</td>
<td>3.24058</td>
<td>16.340638</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>24.80635912992</td>
<td>(0.1550005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a*Low prior, high prior and converged value of the variance ratio.

*b*OLD is ratio used to obtain the solutions at that stage; NEW is the ratio obtained from those solutions (h within parentheses).

*c*is the sum of squares of sire solutions obtained using OLD ratio.

*d*Trace of diagonal elements corresponding to sires of the complete inverse of the coefficient matrix LHS.
disconnected groups, $\hat{\sigma}_s^2$ went down while $\hat{\sigma}_e^2$ went up. This is somehow in conflict with what is stated above, but Schaeffer (1975) himself is not sure if this happening because the population size and number of degrees of freedom is too small, or because of some of the estimated values of the sires deleted or some other reason.

In another reference related to connectedness, Fernando et al. (1983) state that in a two way mixed model (α's are fixed CG effects, β's are random sire effects), contrasts $\alpha_i - \alpha_i'$ are estimable and BLUP estimators of the β's "are uniquely defined and obtainable ..."; "when a sire evaluation model contains fixed CG and random sires, the latter can be compared regardless of 'connectedness'." No doubt about the statistical properties; the assumption may be workable for variance component estimation although its effect on the estimates or on the variance of the estimates is far from clear; but no breed association would be willing to make public reports comparing sires across sets. Individual breeders will not accept and use estimates based on the assumption that his herd has the same average genetic value as any other herd in the breed.

Estimates of sire variance are not affected by disconnectedness if the method of estimation is REML; the method can not yield correct estimates if the model is wrong. The assumptions are an integral part of the model. The assumption that all sets of sires are equal may not be true. If there are genetic differences among the sets, this portion
of the total genetic variance existing in the population is not accounted for. The failure is from the model or from the size of the data set, but not from the method (REML). In light of this, the above argument should be understood.

2) Sires were assumed unrelated, while some close relationships (father-son) among sires analyzed are known to exist and many others are bound to exist. Use of $A^{-1}$ has the biggest effect on sires with small number of progeny. These sires, when $A^{-1}$ is used, will have their estimated predicted differences less "regressed" toward zero and thus will contribute to a bigger $\hat{s}\hat{s}$ than if $A^{-1}$ was not used (numerator would be calculated as $\hat{s}'A^{-1}\hat{s}$, according to Harville, 1984b).

3) "Sire" G028 represent a group of sons of 0628. Sire 0628 has 270 progeny WW records and 941 grand-progeny WW records. These 941 calves represent 11.31% of the total number of records. The effect of this situation on decreasing $h^2$ is two-fold, as follows:

a) because "G028" represents a reasonable large number of sires (a guess would be some 30 to 50, each used during one year), and the $V(\bar{s}) = \frac{1}{n}V(\bar{s}_1)$ or $ns^{-2} < \Sigma_1 (s_1^2)$, this reduces $\hat{s}\hat{s}$;

b) if one is estimating $h^2$ from paternal grandsire-grandson families, the numerator of $h^2$ is given by 16 times the between grandsires variance and not by 4 times that amount, which is done with between sires variance.
Only one factor can be perceived by the author to be influencing sire variance upward: selective matings are practiced in a few herds in an assortative manner. If the model included random cow effects this could be corrected but the net effect on \( \hat{h}^2 \) is not clear. The inclusion of dams may reduce \( \hat{\sigma}_e^2 \) substantially, increasing \( \hat{h}^2 \) after all.

D. Reduced Model

Figure 1, in Chapter IV, presents the biological model hypothesized in this study and the mathematical equation is given in (39). In this section, the question if this model is "better", or describes better the field WW records from RS, than a more parsimonious model as the one used by Ferreira (1982) and recommended by BIF (1972, 1981) is discussed and tentatively answered.

From equation (39), the segments \( X_b \) and \( Z_s \) remained the same, but, from \( W_r \) (the 17 regressors), only the linear and quadratic regressions on age of dam (\( AOD11, AOD12, AOD21, \) and \( AOD22 \)) and the linear regressions on age of calf (\( CAl \)) was retained.

This was done by reading the same \( T \) upper triangular matrix (the expensive part of the computations), corresponding to the full model, from disk, deleting all columns from \( T \) which were not wanted in the reduced model (equivalent to setting those regression coefficients to zero), and retriangularing the last six rows of \( T \) (after deleting the 12 columns) This can be seen in Appendix I, program REML, lines 58 to 106.

Table 14 compares the reduced and complete models. \( R^2 \) values of the reduced model remained very close to \( R^2 \) values of the complete model, either under OLS or GLS. This is saying that even after dropping 12 of
the 17 regressors the two models are equally good in explaining the variance of WWs. The main reason for this likelihood between the values of $R^2$ under the two models is that linear age of calf, in the reduced model, is picking up variation associated with seasonality effects in the complete model. The Type IV SS for CA1 in the complete model is 1,717.95 while in the reduced model it is an astonishing 1,016,214.03 (591 times bigger than in the complete model).

In the OLS models, $\sigma^2_e$ are very similar, while $\sigma^2_e$, from REML, is much larger for the reduced model. Sires variance estimates are also similar for OLS, while for REML $\sigma^2_s$, the estimate for the reduced model is lower than for the complete model. This may be saying that in the reduced OLS model (sires fixed) sire estimates are picking up (getting biased) variation associated with some regressors in the complete model.

In Appendix J, sire solutions (UF and UR) for the complete (UF) and for the reduced (UR) model are listed, along with standard errors of the estimates (EF and ER) and differences in solutions. UF has a larger range (15.0413 vs 13.7326 Kg) and standard deviation (2.505 vs 2.377 Kg) than UR. This is saying that the complete model allowed for a better discrimination among sires' breeding values.

Pearson sample correlation coefficient between UF and UR is 0.94165 while the rank correlation coefficient is 0.94827. If $1/EF$ are used as weights then the Pearson $r$= 0.93454. The values look big and are significant, but they are indicating that only 0.87 to 0.90 of the genetic progress attainable with the full model can be accomplished with the reduced model. Looking at the ratio of heritabilities, $h^2_F/h^2_R = 1.13475$ is also saying that the gain thru selection, by using the com-
Table 14. Comparison of complete and reduced models, using OLS (sires fixed) and GLS (sires random) estimators (variance component estimates for GLS are REML)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>GLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Complete</td>
<td>Reduced</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.566</td>
<td>0.562</td>
</tr>
<tr>
<td>( \hat{\sigma}^2_e )</td>
<td>404.4</td>
<td>407.95</td>
</tr>
<tr>
<td>( \hat{\sigma}^2_s )</td>
<td>12.64</td>
<td>12.39</td>
</tr>
<tr>
<td>( \hat{h}^2 )</td>
<td>0.1212</td>
<td>0.1179</td>
</tr>
<tr>
<td>( k=\hat{\sigma}^2_e/\hat{\sigma}^2_s )</td>
<td>32.00</td>
<td>32.93</td>
</tr>
<tr>
<td>Standard Deviation of EPDs</td>
<td>2.5049</td>
<td>2.3771</td>
</tr>
<tr>
<td>Range of EPDs</td>
<td>15.0413</td>
<td>13.733</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.5370</td>
<td>-0.3406</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>1.3994</td>
<td>1.0187</td>
</tr>
<tr>
<td>Kolmogoroff's D Test for Normality</td>
<td>0.0812</td>
<td>0.0973</td>
</tr>
<tr>
<td>Probability of obtaining a bigger D under Ho:Distribution is Normal</td>
<td>0.034</td>
<td>&lt; 0.01</td>
</tr>
</tbody>
</table>

\(^{a}\text{Expected Progeny Difference.}\)
piece model, can be incremented by 13.475%. Dempfe and Hagger (1983) did theoretical investigation and found that by going from the Contemporary Comparison to the best BLUP procedure the correlation between true and estimated breeding value would be increased by around 2 to 5%. This is saying that the 13.475% increase in expected response to selection is a big gain at very little cost in extra computations.

The model should be improved further and no point is made here that the model should be used as it is.

To obtain the initial values to begin the REML iterations, a different approach was used for the reduced model. PROC GLM of SAS (1982b) produced a value for \( k = 32.93 \) which was very close to the initial value obtained for the complete model. Thus, initial values for the reduced model were based on the final value (24.8064) for the complete model. Initial values were set at 23.8 and 27.8. The converged values for each stage were as follows:

\[
\begin{align*}
1 & : 28.26 \\
2 & : 28.252 \\
3 & : 28.25086 \\
4 & : 28.250873 \\
5 & : 28.250875 \\
\end{align*}
\]

On the process of going from the full to the reduced model, a simple way to obtain Type IV SS was thought of. For each desired variable(s), delete corresponding column(s) and retriangularize with Givens rotations. The position in the last column and row is the \( (SS_{\text{residual}})^{0.5} \); after the retriangularization is completed, the difference in \( SS_{\text{residual}} \) from the
complete and reduced model is the Type IV SS for the deleted source. Figure 16 shows the plot of sire solution in the Full (UF) and in the reduced model. This figure allows two possible interpretations, as follows:

1) One can be sloppy and say that, since the solutions are clustered around the correlation line (r=0.94) and what matters is the progress in the overall population, both models are equally good and the simpler model is to be preferred.

2) One can put the rancher boots on and be more worried with the correct classification of each one of his few sires. If expectations raised in previous evaluations do not get fulfilled in the results obtained or if ranking of sires keep changing, the breeder will probably lose confidence on the whole program and reports.

In Figure 16, four data points are signaled as "cases" and each is described in the following:

CASE 1: Based on the reduced model, this sire (observation 89 in Appendix J) is the second worst in EPD (-6.32 Kg) while in the complete model (EPD = -3.88 Kg) his ranking would be the ninth from bottom to top. The difference \(|UF-UR|\) is 2.45 Kg, about a full standard deviation of UF. This sire belongs to Set 2.

CASE 2: This sire belongs to Set 1, is the observation 28 in Appendix J, has the highest EPD under the complete model (+6.0583 Kg) while under the reduced model his EPD is 4.9128 Kg. This difference in
Figure 16. Plot of sire solutions in the full and in the reduced models.
EPD's of 1.1455 Kg was enough to put him in third place in ranking under the reduced model. This error could have large consequences if one were looking for the best sire to use extensively by AI. Part of the problem may be due to the disconnectedness showing that the observations made by Fernando et al. (1983), in practice, can not be fully accepted.

CASE 3: This is observation 112 in Appendix J. The sire, under the reduced model, had an EPD of +5.1416 Kg and was the second ranking while under the complete model his ranking would be the fifteenth, with an EPD of +2.7969 Kg and the \[|\text{UF-UK}| = 2.3447\], almost a complete standard deviation using the reduced model.

CASE 4: Observation 118 in Appendix J corresponds to the lowest ranking sire (EPD = -8.9830 Kg) and corresponds to 2.24 standard deviations of UF. His EPD under the reduced model is -3.3684 Kg and the corresponding ranking, from bottom to top, is the tenth. This sire was heavily used (look in column for EF of ER) and, under the present ROP program (reduced model), his large negative influence on beef production went undetected. When examining the complete variance-covariance matrix this sire has large (in most cases the largest) negative covariances with most of the regressors. Part of this is due to his large number of progeny, but it is also indicating that matings were done with special groups of cows or his progeny were born in periods such that the reduced model was not able to make proper corrections and thus, his UR contains environmental effects (is biased).

Each one of the above cases is enough to justify the choice for the complete model. If one divides all costs to run the ROP programs (by the
breeders, the Federal government, the breed association, and the university) by the number of calves participating, then the extra computing costs are probably the smallest investment that can be made to obtain an increase in the results of 13%. The actual structure to run the ROP program in RS is actually being underutilized since a breeder could do all calculations (under the reduced model) using a hand calculator.

Looking at |UF-UR| (ABSDIFF in Appendix J), the average is 0.5219 Kg with a standard deviation of 0.66 Kg. Forty percent of ABSDIFF are above 0.4506 Kg (about 1 lb.); 10% are above 0.9935 Kg (about 2 lb) and 5.4% above 1.5077 Kg (3.33 lb.). This is saying much more about the adequacy of a model than the correlations (Pearson, Spearman, Kendall $\tau$, or whatever, weighted or not).

The order given to sires by program ABSSIRER (and the "observation" number in Appendix J) comes from the order the records are read in. Since records are sorted by herd-year-season (also management and sex), in ascending order, the last sires in the Appendix J are sires that were used last on the last herd. Set 1 represents all sires from herd CA and the ordering of "observations" 1 to 32 are associated with the date the first calf (included in this study) from a sire was born. Just by looking at UF values for these 32 sires, it is very easy to observe that there is a positive trend. This indicates that genetic progress is occurring, at least in some herds, and that the situation is not as bad as described by Cardellino-Stercken (1980).
VIII. APPLICATION OF RESULTS AND POSSIBLE USE OF GIVENS ROTATIONS FOR WITHIN HERD AND ACROSS HERD EVALUATIONS

This chapter presents a discussion on why "correction factors" are not being suggested to adjust for the environmental effects and a proposal for the future operation of PROMEBO, the ROP program run in RS by the agreement ANC-UFPel.

These topics would have been the objective of study by the author during post-doctorate training at ISU, but instead, they will be studied upon the author's return to UFPel. They are presented here with the spirit of collecting helpful criticisms from the committee members. Some of the ideas may be helpful in other settings. Before any further analysis, this study should be redone, as will be presented in Chapter X. Thus, this chapter is more general, in dealing with variables included in the models.

The basic model in this chapter is:

\[ y = X\beta + Z_s + W_r + e \]  

(43)

where: \( \beta \) are CG effects;

s are random effects (possibly sires and dams);

r are other environmental effects;

other elements are as in (38); and

\( V(y) = \sigma^2_v \).

A. Taking Environmental Effects Into Account

This whole section is based on Goldberger (1964, Chapter 5. Extension of Linear Regression; Section 6. Use of Extraneous Information)
where complete development and proofs can be found.

1. Precorrection of records: restricted least squares

Ideally, correction factors, as used by animal breeders, should be calculated for each breed and herd, and possibly each CG. Most of the groupings are too small to obtain reliable estimates of environmental effects. Animal breeders then use correction factors obtained from much larger samples, assuming that the large sample used to obtain the correction factors and the CG or herds to which these factors are applied are samples of the same population. Econometricians say that they are using extraneous, unbiased estimates.

Suppose then that we have \( r^* \) such that \( r^* \) is an unbiased estimator of \( r \):

\[
  r^* = r + d
\]

with \( E(d) = 0 \)

and \( E(dd') = V_r^* \). Then, applying the restriction \( r^* = Ir \), we regress

\[
y - Wr^* \text{ on } [X Z] \]

to estimate \([\beta', s']\)' as

\[
  \begin{bmatrix}
    \beta^* \\
    s^*
  \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\
                             Z'Z & Z'Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X' \\
                                                                             Z' \end{bmatrix} \begin{bmatrix} y - Wr^* \\
                                                                                     Zr \end{bmatrix}.
\]

(44)

Substituting equation (43) in equation (44):

\[
  \begin{bmatrix}
    \beta^* \\
    s^*
  \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\
                             Z'Z & Z'Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X' \\
                                                                             Z' \end{bmatrix} [X\beta + Zs + Wr + e - Wr^*],
\]

where \( Wr^* = Wr + Wd \), and thus
The big assumption made to obtain this result is that $E[e, d'] = 0$. If no restriction were applied, then

$$V_2 = V_s = \sigma^2 \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix}^{-1} + \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'W \\ W'X \\ W'Z \\ Z'X \\ Z'Z + G^{-1} \end{bmatrix} \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'W \\ W'X \\ W'Z \\ Z'X \\ Z'Z + G^{-1} \end{bmatrix}.$$

(46)
where \( V_r = E[ (f - r)(f - r)' ] = \)
\[
g^2 \left\{ W'W - [W'X W'Z] \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix}^{-1} [X'W]^{-1} \begin{bmatrix} X'W \\ Z'W \end{bmatrix} \right\}. \tag{47}
\]

\( V_r \) is the covariance matrix of unrestricted GLS estimator of \( r \).

Subtracting (45) from (46) gives a comparison of the restricted and unrestricted GLS estimators of \( \begin{bmatrix} \beta \\ s \end{bmatrix} \) as

\[
V_2 - V_2^* = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'W \\ Z'W \end{bmatrix} [V_r - V_r^*] [W'X W'Z] \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'W \\ Z'W \end{bmatrix}
\]

= \( B [V_r - V_r^*] B' \)

where \( B = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'W \\ Z'W \end{bmatrix} \) and it is the coefficient matrix

of the "auxiliary regressions" of \( W \) on \( [X Z] \). If we let \( B_j \) denote the
\( j \)-th column of \( B' \), we see that the variance of the unrestricted estimator of the \( j \)-th element of \( \begin{bmatrix} \beta \\ s \end{bmatrix} \) will exceed the variance of its restricted
estimator by \( B_j [V_r - V_r^*] B_j \). There will be a gain in efficiency in the estimator of the \( j \)-th element of \( \begin{bmatrix} \beta \\ s \end{bmatrix} \) when this expression, \( B_j [V_r - V_r^*] B_j \),

is positive. This is seen to depend on whether the variances of the extraneous estimator of \( r (f^*) \) are small relative to those of its unrestricted GLS estimator \( f \). If \( [V_r - V_r^*] \) is non-negative definite, then \( V_2^* \) is smaller than \( V_2 \). Thus, it is advantageous to use extraneous
information (precorrect records) when their variance is relatively small. If one is analyzing large data sets (like in current sire evaluations), larger than those from where the correction factors were estimated, then \( \text{tr}[V_r - V_r^*] \) is positive and there is a smaller efficiency than if \( r \) were estimated directly or simultaneously with \( \hat{\beta} \) and \( \hat{a} \).

In RS, where records from one year for given breeds are greater in numbers or similar to previous accumulated records, this option of precorrecting records is not recommended. Even for breeds with much larger participation, this option may not be the best and the procedure developed in the next subsection may give better results.

2. **Combining information from extraneous sources with the sample own information**

The obvious deficiency of restricted least squares is that it fails to utilize the sample information to improve the estimates of the parameters that are assumed to be known.

Optimum use of the available information from RS' data is sought for the following reasons or cases:

1. Some breeds or crosses have too small numbers to ensure that the estimates have some minimal properties. Using overall information combined with that breed's records will result in estimates which would make sense and tilted in the direction of the particularities of that breed or sample. The more information a breed would have or accumulate the more differentiated could the final estimates be for each breed.
2. There are reasons to believe that some regions in the state of RS have climates different enough to make it necessary to obtain and use separate sets of "correction factors" for each of the regions. Most of them have too few herds, though. If one would fit a model containing interaction terms (fixed) of the regressors with regions, the resulting estimates could be completely determined by only a few observations if that region contained few herds or records. By combining information from all records and from a region's records one could obtain estimates which would be close to the ones obtained from a model which had regressors nested in regions; but the final estimates are not allowed to vary wildly if the region has few herds and some influential observations.

3. The same as in point 2 can be said about the "interaction" of regressors with years, especially regressors for seasonality effects.

4. When there exists an AI proof (the result of a progeny test or of a national sire evaluation) from a sire (or several) used in a herd then two courses of action can be taken: a restriction on sire estimates (equal to the outside information) can be imposed (in this way results from all herds are comparable since they are constrained to a common base); or outside information can be combined with within herd information to obtain better estimates of sires' EPDs. The restriction approach allows for immediate comparisons to be made across herds (say among dams and yearling bulls) but several untenable assumptions have to be
invoked. The approach which combines information across and within herds does not allow for instantaneous, absolute comparisons to be made. Only with time, as more and more ties link herds do these comparisons become more reliable. The transition between the two approaches can be exemplified by a herd which has only one progeny from a heavily used AI sire, and none of the other sires are used outside the herd. The results in either approach would be very similar but only if one is operating with the "combining information" approach can one see that comparisons across herds would be extremely unreliable. There is much discussion among animal breeders about these comparisons across herds after sire proofs are used in within herd evaluations. Probably the two groups of animal breeders are right, only that one group is thinking about using these proofs as "restrictions" but do not say so to the other group, which is thinking about "combining information" from outside and within herd. Probably an econometrician could solve the dispute.

A way to combine information follows. Again, these are results given by Goldberger (1964, pages 259-261). The equations will be made even more general, since they can be used for different purposes.

The problem to be solved can be written as

\[ y = X\beta + e \quad , \quad e \sim (0, \sigma^2 I); \]
\[ b_1^* = RB + d \quad , \quad d \sim (0, \sigma^2 W), \quad \text{with} \ E[u, d'] = 0 \quad \text{and} \]
\[ R = [I \ 0]. \]

The extraneous and sample (within herd, breed, or region) informa-
\[
\begin{bmatrix}
y \\
\hat{b}_1^* \\
\end{bmatrix}
= 
\begin{bmatrix}
X^T \\
R \\
\end{bmatrix}
\begin{bmatrix}
y \\
\hat{y} \\
\end{bmatrix}
+ 
\begin{bmatrix}
e \\
d \\
\end{bmatrix}
\]

The BLUE of \( \hat{\beta} \) in the sense of having minimum variance in the class of all unbiased estimators which are linear in \( y \) and \( \hat{b}_1^* \) is the GLS estimator:

\[
b^{**} = \left( X' R' \right) [\sigma^2 0]^{-1} [X]^{-1} [X' R'] [\sigma^2 0]^{-1} [y] \\
\]

\[
b^{**} = \left[ 1/\sigma^2 X'X + 1/\sigma^*_e R'W^{-1}R \right]^{-1} \left[ 1/\sigma^2 X'y + 1/\sigma^*_e R'W^{-1}b^*_1 \right].
\] (48)

Obviously, \( W \) is the inverse of the coefficient matrix used to obtain \( \hat{b}_1^* \). If the sample is large (\( \det|X'X| \) is large) and \( \hat{b}_1^* \) was not well estimated (\( \det|W| \) is large) then the sample information becomes more influential on \( b^{**} \). If the sample is small and \( \hat{b}_1^* \) was well-estimated, then \( b^{**} \) will be closer to \( \hat{b}_1^* \). In any case, the equation for \( b^{**} \) allows for optimum weighing of both sources of information.

Based on equation (43) and corresponding NE, in the case that we had an extraneous estimate of the last element (\( k^{th} \)) of the vector \( r, r_{k^*} \), the last row of the NE from where (48) was obtained can be written as

\[
\begin{bmatrix}
W'X_k \\
W'Z_k \\
W'W_{k-1} \\
\end{bmatrix}
\begin{bmatrix}
(b^{**})_1 \\
(b^{**})_2 \\
(b^{**})_3 \\
\end{bmatrix}
= 
\begin{bmatrix}
X'y \\
Z'y \\
W_{k-1}'y \\
W_{k}'y + c' r_{k^*} \\
\end{bmatrix}
\]

where: \( W'X_k \) and \( W'Z_k \) are the last rows of \( W'X \) and \( W'Z \).
$W'W_{k-1}$ is the last row of the first $k-1$ columns of $W'W$;

$w_{k}$ is the $k$-th, $k$-th element of $W'W$;

$r_{k}^*$ is the extraneous estimate of $r_{k}$; and

$c^k$ is the reciprocal of $[1/\sigma^2(\text{Var}(r_{k}^*))]$.

Thus, the only difference from the classical NE (or the MME in this example) is in the bottom right element of the LHS and in the bottom element of RHS.

This method requires knowledge of $\sigma^2$, $\sigma_*^2$, and $W$, up to a factor of proportionality. An approximate use of the method is to employ unbiased estimates of these variances and covariances.

In the case that some information exists, outside a given herd, about the random effect $s$, a $q \times 1$ vector, in the form of $Rs^*$, where $R$ is a diagonal matrix $q \times q$ with a 1 on the diagonal corresponding to the elements of $s$ from which there is information on $s^*$ and a 0, otherwise, the MME corresponding to (43) that uses the extraneous information is given by

$$
\begin{bmatrix}
X'X & X'Z & X'W \\
Z'X & Z'Z + G^{-1} + R'C^{-1} \lambda & Z'W \\
W'X & W'Z & W'W
\end{bmatrix}
\begin{bmatrix}
\beta^{**} \\
\lambda^{**} \\
s^{**}
\end{bmatrix}
= 
\begin{bmatrix}
X'y \\
Z'y + R'C^{-1}\lambda s^* \\
W'y
\end{bmatrix}
$$

where: $V(s^*, s^*) = C \sigma_*^2$;

and $\lambda = \sigma^2/\sigma_*^2$.

When working with the data matrix, $[X \ Z \ W \ y]$ was reduced to $T$, pseudo-observations in the form $[0 \ P \ 0 \ 0]$ were added and rotated against $T$ to produce $T^*$. If, corresponding to $[X \ Z \ W \ y]$ the matrix $[0 \ A \ 0 \ f]$ is rotated against $T^*$ to produce $T^{**}$, where
then the data matrix approach is also able to handle this optimum use of across and within herd information. Using $T^{**}$, new $A$ and $f$ can be formed to be used in the next herd being processed. $A$ and $f$ contain, at each instant, the most accurate and complete information about $s$.

If the situation is such that, during the weaning season (the same can be said about the yearling season), each day a different herd has its calves weaned and has its records processed then each herd will have back the reports to make selection decisions based on all information available up to that date. That is much better than using last year's across herd information or to wait until all records from all herds are submitted to receive the exact reports back.

Another advantage of this method is that the problem is not solved de novo for each calf crop, beginning by the construction of the NE using all previous records. This problem, that analysis based on the MME have, can preclude the widespread use of GLS estimates for within beef herd animal evaluations. One can say that all animal breeding evaluations based on extensive data collection schemes are "sequential accumulation" problems (as defined and solved by Lawson and Hanson (1974)), and there is a big waste of efforts when the problem is constrained to the NE or MME. One of the last proposals to reduce this inefficiency was made by Martinez and Rothschild (1983), still on the MME mode, using transition matrices in a recursive estimation procedure.

Thus, the method of combining within and across herd evaluations, with partial use of Givens rotations, and keeping an up-to-the-second
matrix of solutions (maybe in the form \([A f]\) above), should be studied by animal breeders which hold consulting positions with breed associations since it can replace, with advantages, present national sire evaluation programs and, more importantly, enhance the launching of within herd GLS estimation programs.

B. The Operation of Prombo as a GLS Problem

1. **Cholesky decomposition vs. Givens rotations**

   The Cholesky method, as more commonly presented, when applied to the NE requires that two triangular systems be solved. The SWEEP operator, as presented by Goodnight (1979), requires the solution of only one triangular system, and this can be done with the Cholesky decomposition. This form of Cholesky decomposition is what is being considered in this section.

   Lawson and Hanson (1974, Table 19.1) give the operation counts for various least squares computational methods. The asymptotic number of operations where an operation is a multiplication or division plus an addition is as follows:

   - Form the NE : \(np^2/2\);
   - Cholesky solution of the NE : \(p^3/6\); and
   - Solution of a data matrix by Givens rotations : \(2np^2 + np \text{ SQRTs}\).

   For a given problem whose records will be used only once, there is no doubt that Cholesky decomposition is the method of choice. What will be investigated next is a simulation of what could be the number of operations to solve a within herd program during 10 years.
The assumed parameters of the problem are as follows:

1. The herd has a calf crop of 100 calves weaned/year ($n^c$);
2. The cow herd has 120 cows and 20 cows are replaced/year ($n^c$);
3. The constant number of sires is 4 and one new sire is replaced every year ($n^s$); and
4. There are 3 CG per year ($n^h$).

The number of parameters $p = n^c + n^s + n^h = 107$ (in the first year) has been kept as low as possible since this is favoring the Cholesky method (number of operations is $O(p^3)$ while Givens method is $O(p^2)$).

Since rows corresponding to animals which do not produce progeny anymore are never operated on anymore in the Givens rotations, a fixed $p = 107$ and $n = 100$ are considered every year. A fixed number of observations (only last calf crop) is considered because once observations are transformed to $T$, they are never used again (when using Givens rotations).

Table 15 presents the calculated number of operations needed to solve the problem via the Cholesky method every year and the accumulation for the years. The number of operations needed to form the NE is for a pure regression model. Animal breeders deal mostly with classificatory models or models which contain only a few regressors. Very efficient programs exist to form the NE by counting techniques, although the author has never seen such a program. Schaeffer (1976) presents some ideas of sums of outer products, which operates with a record at a time, that are useful to regression problems. PROC SUMMARY is a very efficient procedure from SAS (1982a) that creates a file containing all non-zero elements of the LHS and RHS, completely indexed. Data are not needed to be sorted by any of the factors and the execution time of this procedure, as
experienced by the author, is much smaller than the DATA step (which was only reading the records) of SAS.

Again, numbers under column NE of Table 15 are for pure regression problems. This number of operations is not taken into account when calculating the accumulated, over all years, number of operations needed to solve the problem via Cholesky decomposition.

As the author sees the problem, when working with the NE or the MME, the major difficulty with the process is that the problem keeps getting bigger, both in terms of numbers of operations and with active memory requirements.

When comparing Tables 15 and 16, for the 1st year, the major advantage also becomes clear: Cholesky is much more efficient to solve a problem than Givens rotations, when that information is going to be used only once.

Table 16 shows that when there is a sequential accumulation of information, year after year, Givens rotations will become more efficient than Cholesky decomposition.

2. Combining Cholesky and Givens

The ROP program in RS has several herds with thousands of records collected. It is recommended that all those records be converted to an upper triangular T by Cholesky decomposition. Later updates would be done by Givens rotations. This would be the way to operate also over the "pseudo-observations" to obtain T* that will yield GLS estimators.

In animal breeding problems, n ≥ p generally, and it may be advantageous to combine the best traits of each method into a hybrid
Table 15. Number of operations needed to solve a within herd problem via Cholesky decomposition during a period of 10 years

<table>
<thead>
<tr>
<th>Year</th>
<th>nc</th>
<th>ns</th>
<th>nh</th>
<th>pa</th>
<th>nb</th>
<th>NEc</th>
<th>Choleskyd</th>
<th>Total/Yeare</th>
<th>Accumulatedf</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120</td>
<td>4</td>
<td>3</td>
<td>127</td>
<td>100</td>
<td>806,450</td>
<td>341,397</td>
<td>1,147,847</td>
<td>341,397</td>
</tr>
<tr>
<td>2</td>
<td>140</td>
<td>5</td>
<td>6</td>
<td>151</td>
<td>200</td>
<td>2,280,100</td>
<td>573,825</td>
<td>2,853,925</td>
<td>915,222</td>
</tr>
<tr>
<td>3</td>
<td>160</td>
<td>6</td>
<td>9</td>
<td>175</td>
<td>300</td>
<td>4,593,750</td>
<td>893,229</td>
<td>5,486,979</td>
<td>1,808,451</td>
</tr>
<tr>
<td>4</td>
<td>180</td>
<td>7</td>
<td>12</td>
<td>199</td>
<td>400</td>
<td>7,920,200</td>
<td>1,313,433</td>
<td>9,233,633</td>
<td>3,121,884</td>
</tr>
<tr>
<td>5</td>
<td>200</td>
<td>8</td>
<td>15</td>
<td>223</td>
<td>500</td>
<td>12,432,250</td>
<td>1,848,261</td>
<td>14,280,511</td>
<td>4,970,145</td>
</tr>
<tr>
<td>6</td>
<td>220</td>
<td>9</td>
<td>18</td>
<td>247</td>
<td>600</td>
<td>18,302,700</td>
<td>2,511,537</td>
<td>20,814,237</td>
<td>7,481,682</td>
</tr>
<tr>
<td>7</td>
<td>240</td>
<td>10</td>
<td>21</td>
<td>271</td>
<td>700</td>
<td>25,704,350</td>
<td>3,317,085</td>
<td>29,021,435</td>
<td>10,798,767</td>
</tr>
<tr>
<td>8</td>
<td>260</td>
<td>11</td>
<td>24</td>
<td>295</td>
<td>800</td>
<td>34,810,000</td>
<td>4,278,729</td>
<td>39,088,729</td>
<td>15,077,496</td>
</tr>
<tr>
<td>9</td>
<td>280</td>
<td>12</td>
<td>27</td>
<td>319</td>
<td>900</td>
<td>45,792,450</td>
<td>5,410,293</td>
<td>51,202,743</td>
<td>20,487,789</td>
</tr>
<tr>
<td>10</td>
<td>300</td>
<td>13</td>
<td>30</td>
<td>343</td>
<td>1000</td>
<td>58,824,500</td>
<td>6,725,601</td>
<td>65,550,101</td>
<td>27,213,390</td>
</tr>
</tbody>
</table>

a \( p = n_c + n_s + n_h \) = number of parameters.

b \( n \) = number of observations computed each year.

cNumber of operations needed to form the \( \text{NE}(n^2/2) \).

dNumber of operations needed to perform the Cholesky decomposition \( (p^3/6) \).

eTotal/Year = NE + Cholesky.

fOnly Cholesky operations are accumulated. NE are assumed to be formed at no expense.
Table 16. Number of operations needed to solve a within herd problem via Givens rotations during a period of 10 years. [Number of square roots needed are not computed as a trade-off for not computing the operations needed to form the NE in the Cholesky decomposition method]

<table>
<thead>
<tr>
<th>Year</th>
<th>$n_c$</th>
<th>$n_s$</th>
<th>$n_h$</th>
<th>$p^a$</th>
<th>$n^b$</th>
<th>Givens$^c$</th>
<th>Accumulated$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>2,289,800</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>4,579,600</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>6,869,400</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>9,159,200</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>11,449,000</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>13,738,800</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>16,028,600</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>18,318,400</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>20,608,200</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>4</td>
<td>3</td>
<td>107</td>
<td>100</td>
<td>2,289,800</td>
<td>22,898,000</td>
</tr>
</tbody>
</table>

$^a p = n_c + n_s + n_h = 107 = $ number of parameters on each year.

$^b n = $ number of observations computed each year.

$^c$ Number of operations calculated as $2np^2 = 2(100)(107)^2$. An additional np = 100(107) = 10,700 square roots are computed each year.

$^d$ Accumulated number of operations until a given year.
procedure.

It was discussed in Chapter V that the order in which the rotations were applied did not have any effect on the final result, \( T \). Say that each year the resulting \( T \) is \( T^1, T^2, \ldots, T^Y \). Before rotating records from the second year over \( T^1 \) to obtain \( T^2 \), these records can be rotated against themselves, forming \( T_2 \) as an intermediate step. Then, \( T_2 \) can be rotated against \( T^1 \) to obtain \( T^2 \). Since the order of rotations is not important, the resulting \( T^2 \) will be the same. What is envisaged is to form \( T_2 \) by Cholesky decomposition and then rotate \( T_2 \) against \( T^1 \) by Givens rotations. For the problem used in Tables 15 and 16, each one of \( T_1, T_2, \ldots, T_{10} \) would require the same 341,397 operations as in year 1 of Table 15. Note that \( T_1 = T^1 \) and hence only \( T^2, T^3, \ldots, T^{10} \) would be needed.

\( T_2 \) is an upper triangular with only \( p(p+1)/2 \) non-zero elements. Further, rows of \( T_2 \) that do not have correspondence (new cows and sires; all new CG) with rows of \( T^1 \) will be transferred directly, without any rotation being needed. Without considering this, \( T_2 \) has \( p \) rows and the average number of columns to the right of the diagonal element is \( p(p+1)/2 \). The number of operations to reduce a \( n \times p \) data matrix into an upper triangular form by Givens rotations is \( 2np^2 \). To operate \( T_2 \) over \( T^1 \), this number will be \( 2(p)((p+1)/2)^2 = p^3/2 + p^2 + 1 = 419,798 \) operations. Thus, for the 2\(^{nd} \) to the 10\(^{th} \) year, the number of operations would be 341,397 + 419,798 = 761,195. For the 10 years of the program, 7,192,152 would be needed. This is only 26.42% of the accumulated for the Cholesky decomposition method and 31.41% of the accumulated number of operations for the Givens rotations method. In the first year, the number of operations would be exactly the same as the Cholesky method, obviously. On the
second year, the hybrid method would require 187,370 more operations than
the Cholesky, but from the third year onward, it would be much more
efficient than either method. This method combines both qualities of the
"parent" methods (especially fixed amount of computations and memory
requirements per year, from the Givens rotations) and it is superior to
either one, in any "trait." Looks like there is really some hybrid vigor
in this method.

An iterative method that is used to solve large animal breeding
problems is the Gauss-Seidel or its modification, the successive over-
relaxations (SOR) method. Schaeffer (1976) presents a very good discus-
sion of this method, with a small example, properties and a efficient
Fortran program to be used in national sire evaluations. The number of
needed multiplications is $k^2 p$, where $k$ is the number of rounds needed to
obtain convergence of sire solutions and $p$ is the number of unknowns. In
National Sire Evaluations done at ISU, values of $k$ needed are between 10
and 15. Values more near to 15 are needed when all solutions are
required to change less than 0.005 before stopping the iterations. Using
values of $p$ given in table 15 and a $k = 10$, the number of operations
needed to solve the same within herd problem as before, from year 1 to 10
are, respectively: 161290, 228010, 306250, 396010, 497290, 610090,
734410, 870250, 1017610, and 1176490. After the first year, a fixed
number of operations (761,195) is needed to solve the same problem using
the combined Cholesky and Givens method. Thus, at the 8th year the
combined method is more efficient than the best indirect method, and from
there onwards this difference would be, asymptotically, of order $(p^2 -
p)$. 
3. **Within herd evaluation**

To handle the environmental effects in the new PROMEBO, what is planned is to use what was discussed in Subsection 2, Section A of this chapter. That is, combining information from an overall analysis (or breed-wide) and from each of the participating herds.

In practical terms this method requires more computations than the traditional precorrection of records approach, but it has three main advantages, as follows:

1. It allows for the estimation of environmental effects specifically for each herd, ensuring that the results are always reasonable by adequate weighing of extraneous and within each herd information. It is guessed that the effects which may be most variable (corresponding somehow to interaction terms) are age and nursing status of the dam (dependent on how "good" the environment is for each herd) and seasonality effects (changes according to regions and years would be picked up).

2. Since the set of estimates are always changing, one never has to write them out as a definitive set, which, given the human nature, always make changes and improvements more difficult to occur.

3. The estimates obtained for each herd can be used very well for management decisions. Plots like the ones used in the figures of this study should be given to the breeders, each year.

For an established herd, with an about fixed number of dams and sires and size of the calf crops, the method of Givens rotations requires an about fixed (across time) amount of memory and effort from the computing
equipment. Memory requirements are only for the program and two vectors at a time, with the length of the largest possible number of sires and dams producing a calf crop. Active memory requirements are easily supplied by modern microcomputers. As data accumulates, peripherals in the form of hard disks with backup tapes will be needed. Hard disks for microcomputers can be found with over one hundred megabytes. Laser bubble memories will be soon in the market with capacity to hold information measured on the gigabytes. The author has information (not guaranteed to be up-to-date) about hardware that is being produced in Brazil for microcomputers, peripherals, and hard disks with 20 megabytes (without backup in tape). It appears that information within herds will accumulated at a slower pace than the Brazilian microcomputer industry (the Brazilian government adopt a policy of reserving the internal market for micro- and minicomputers for companies with 100% of Brazilian capital). If a hard disk can not hold anymore the entire T matrix or for any other reason old information is not wanted anymore, rows corresponding to old CG, or sires, or dams can be deleted without any change in the remaining rows (each row contains the effect of a factor). The backsubstitution process begins from bottom to top. Hence, even if old (the upper) rows are deleted, the solution for the parameters of interest remain exactly the same and the solution is simultaneous to those effects which have been deleted. This can not be done if one is working with the NE or the MME. To obtain parts of the complete inverse, one can begin from any row of T until the last minus one row, find the $T^{-1}$ of this segment, postmultiply it by its transpose and the product will be the desired portion of the complete inverse, even if rows have been deleted
from T or upper ones are just not used in the computations.

To obtain these same results when using the NE or the MME, unwanted rows should be absorbed into the other ones (there is nothing wrong with absorbing some sire equations into the remaining ones) and not just deleted, as has been suggested by some authors. This would be a gross error and very easy to prevent.

Instead of just deleting rows from T, the rows can be kept in storage somewhere else, possibly on a tape. T is then partitioned into an active section and a temporarily inactive section (old). On the rows of the old section, there are coefficients corresponding to parameters in the active section. If and when solutions for the old parameters are wanted, the solutions from the active part can be used in backsubstitution in the old rows and the solutions computed. Thus, no information is lost and there is no sacrifice in accuracy of the solutions (all) when one decides to "delete" old rows of T to reduce storage requirements.

Before sectioning T, one may desire to impose a restriction that a given sire or dam or a linear combination of them equals zero (or any other value). The only thing needed to be done is to form $R\tilde{\mathbf{p}} = \mathbf{v}$, as described in Sections A.1 and A.2 of this chapter and to rotate $R\tilde{\mathbf{p}} = \mathbf{v}$ against T. When T is sectioned, the solutions for the "old" parameters will remain unchanged always (so that one knows what the solutions are) and the solutions for the active section will be conditional to the restrictions imposed, even after the old rows are deleted. The process can be repeated at different time intervals, always when one keeps the sets of restrictions consistent among themselves.
4. Across herd evaluations

The present rules which govern PROMEBO states that the breeder is the only owner of the information generated about his animals (that is the reason why Appendix J does not contain identification of the sires) and that no comparisons would be made among animals of different herds. There is a loss of valuable information resulting from these rules but that was demanded by most breeders at the beginning of the program. On the other side, if the information from each herd is not publicized and the breeder is the only one who decides which information will be made public (the ANC checks that a minimum set of figures are used and that values reported are correct) one has the guarantee that the records will be made available by the breeder in their most accurate possible form.

An open competitive program can increase the apparent rate of genetic gain but, as human history shows, it will also increase the use of false information and all possible tricks to put one herd ahead of their competition. In such a system, even when there are indications that the results are too good to be achievable, "prominent" breeders are respected, obtain commercial success, their seedstock gets extensively used in the breed, and they are used as a model by their peers. Until his manager quits the job or the breeder becomes less cautious. The whole program suffers when this happens.

The genetic progress occurring in a breed or in the beef population of a state is nothing more than the (weighted by use) summation of the genetic gains being made at each of the beef producing units, the individually managed herds. At this level is where gains are being made and they will be sound only if based on exact information.
"The backbone of the beef industry is the honesty of the breeders" is a nice sentence that is enough to make most people comfortable. "The opportunity is what engenders the thief" is more crude, but more near the truth and should prevail among personnel in breed associations truly responsible for the improvement programs.

When more and more outstanding bulls are being made it creates a sense that the breed is moving, it gives a subject matter for breeders to talk about with optimism, it makes everybody happy, even insiders who do not believe that all that is being, or can be, achieved. Use of genetic groups can also add to the euphoria. And it can really help in international competition. After all, the industry of a country may be best served, when measured by profit, under such a system that produces lots of results, even if not all that is true.

RS will probably never compete in the international genepool market, given its label of being technologically backward. Thus, one can plan more in terms of obtaining true results.

The real task for PROMEBO is to make clear to the breeders that they will maximize genetic gain only by being perseverant in the selection of their animals under RS' specific environmental conditions and for a specific market. Following international modisms only added to diversification and to the folklore; massive imports of genetic material have been mostly disruptive ("steers" in RS are being slaughtered at the same age and weight as the Criollo oxen were at the turn of the last century); and elite seedstock breeders whose only work is to use fashionable imported semen are just following the easy path to keep the status quo, with profit.
The author is not convinced that making across herd evaluations available will be more beneficial than upgrading PROMEBO from a ratios base to a GLS estimation base, within herds. This is a subject to be debated with the ANC, their breeders and technicians.

The immediate approach that may be decided is to allow some comparisons across herds of breeders who have common commercial interests and are co-owners of the same bulls. That may produce some gain without risks of giving an incentive to dishonesty. Probably some breeders will choose to remain isolated and, even if a few, their decision should be respected.

A very conservative approach that can be used is the one discussed in Section A.1 of this chapter. This is by using information about some sires as restrictions. In this way, one can link some herds together but the information from those herds does not influence the estimates of the sires used across herds. Under this form, the breeder can obtain a ranking of his sires against the rest of the breed, but that information would remain particular and not official.

Another way to ensure confidence in the results is to use extraneous information as in section A.2 about old sires only which are not producing in the herd any more. This does not give a chance to the breeder to plan matings or to give preferential treatment to the progeny of sires which the breeder has commercial interest in. Those old sires will have different categories of descendants in the herd and it would require a high level of expertise to plan matings and the rest to really manipulate any result, if that is possible at all. After several years of participation, when the breeders (and those breeders) become more
involved with their mutual benefit from such a system, more and more recent across herd information can be used to better tie the herds together and the resulting information be given more widespread use.

Information generated by SARGS, in their Progeny Test and Central Testing Stations, should also be used to improve accuracies of within herd estimates and increase ties across herds.

This informal proposed scheme of implementation of technologies may seem too cautious. To the author it seems preferable to ensure the continuity of PROMEBO, its expansion and use for selection of the generated information. Beef cattle genetic programs are never impact programs, at least in terms of results. Perseverance in the objectives and the education of the breeders are seen as better qualities or options than to take bold actions that may produce a little higher results (maybe only apparent ones), but that introduce the risk of destroying what was built in the last ten years.

5. Advantages of using Givens rotations

The use of Givens rotations to solve animal breeding problems has been dealt with sparsely in the last four chapters. The main advantages of using Givens rotations are listed in the section. Complete discussion of these topics have been given or referenced previously. The advantages are as follows:

1. Computations are done one row of data at a time; memory requirements are for one row of data and one row of T[2*(p+1) positions] plus the program; memory requirements are kept at minimum by increasing I/O operations. Thus, Givens rotations
are a very nice alternative technology for microcomputer owners and to less developed countries (LDC). The statement that "Mixed model technology requires a sophisticated computation center" is not true anymore.

2. Givens rotations increase the useful life of computers to solve animal breeding problems. For example, within herd mixed models, if based on the NE, can outgrow the static capacity of any computation center, given enough time for the records to accumulate.

3. Sparseness of the incidence (data) matrices are readily exploited to reduce number of computations.

4. Accuracy of results is increased. Something is lost when NE are formed, especially if the model used contain covariates. One has to work with precision $k^2$ ($k =$ number of decimal places) with the NE to obtain the same level of accuracy as in methods which work with the data matrix using a precision $k$.

5. Updating (adding new records) is natural to this process. New and old information are processed exactly the same way: one row of data at a time. Records can be deleted with the same ease: just add again the record to be deleted previously multiplied by $-1$. This is a big advantage for breed associations. ANC processes records for each herd-year-season with a frequency of near two. Working with ratios this is not a big problem but for within herd GLS evaluations based on the NE this is a big waste.

Animal breeding data bases have a continuous data stream; data accumulates sequentially. Animal evaluations are never
final; today's evaluation will be more accurate tomorrow.

6. The same updating idea is used to obtain REML estimates of variance components and sire solutions with BLUE properties.

7. Weighted and robust least squares can also be used. There is a possibility to reduce the operation counts and to eliminate the need to take square roots.

8. No precorrection of records is needed. Information from the herd can be used to obtain proper estimation of the environmental effects. After T is computed, add extraneous estimates for the regression coefficients, weighted by the inverse of their standard errors. This would ensure against weird regression coefficients being estimated in small herds.

9. Extraneous information about sires and MGS can be used for each herd to tie them together. No national sire evaluation "per se" is needed since at each moment we have evaluations using all available information and we are weighing them properly.

10. "Old" rows of T can be thrown out (to save storage), but other solutions are still adjusted for the effects these rows represent.

11. Some of the tricks used with the NE can also be done with the data matrix, like the absorption of CG. But a lot remains to be learned about this method, since it is not used as much as methods based on the NE.

12. It is possible to combine Givens and Cholesky methods to solve animal breeding problems in such a way as to obtain better results than using either one alone.
IX. CONCLUSIONS

This work began with very ambitious goals. Very few of them could be completely met. Weaning weight was the only response variable studied. Conclusions are as follows:

1. The estimate of heritability obtained at $0.155 (h^2 = 0$ with $p < 0.01$) is assuring that there exists identifiable genetic variation in the weaning weights of the Hereford cattle population of RS. This estimate of heritability was obtained from the REML covariance among paternal half sibs and REML error variance estimates. Several indications exist that the assumptions made are causing the value of 0.155 to be an underestimate.

2. A not-so-parsimonious model was investigated and this model appears to better approximate the environmental effects on range beef calves suckling on their dams, in the state of RS, than the previous model used by PROMEBO. The more complete model can better discriminate sires' expected progeny differences (EPD) and allows for higher responses to selection (at least 13% higher) than the reduced model being used, which may also produce biased sires' EPD.

3. A robust estimation procedure was tested, showed a potential to be helpful in animal genetic evaluations, but, given the use of a bad within CG error dispersion measure, failed to yield conclusive results. Suggestions are given for further study and possible improvement of the technique.

4. A combination of Givens rotations (Givens, 1954) and common intercept approach (CIA of Schaeffer, 1983) was used in a algorithm
(claimed to be efficient) to perform the iterations necessary to obtain REML estimates of variance components.

5. Givens rotations applied to the data matrix were used to obtain GLS estimates. A procedure to combine Cholesky and Givens methods is suggested, to solve animal breeding problems characterized by the sequential accumulation of information, which is expected to be more efficient in a medium range of time and much more efficient in the long range than methods based in solving the MME of Henderson (1950).

6. The result which is the most gratifying to the author is that a 30 year old technology can be used to solve today's animal breeding problems by an animal breeding worker which is living in a very simplistic computer environment. This can be useful in democraticizing the technology among nations, or to reduce the gap between developed and backward countries, or to increase the technological independence of the so-called third world countries. Only then a more pure, academic, free of commercial interests, sharing environment can be established in the so-called North-South and South-South relations to better serve the major objective of allowing each person to receive the necessary intake of daily animal protein.
X. THINGS THAT SHOULD HAVE BEEN DONE DIFFERENTLY

Several ideas from different authors were put together to solve a specific problem. Some of these ideas were learned in course work done at ISU (mostly from Fuller (1982) and Harville (1983, 1984a). Other ideas were obtained from the literature (mostly from Schaeffer (1983); Goldberger (1964) and Johnston (1972, 1984); Seber (1977), Lawson and Hanson (1974), and Maínndonald (1984); and Huber (1981) and Hoaglin et al. (1983)). The interpretation and application of the ideas obtained from the readings were done with complete independence by the author.

The naivete and lack of experience of the author, the reduced size and quality (for a few of the explanatory variables) of the data set used, and the, say, excessive independence of the author when conducting the analysis are the major causes for the shrinkage of the results from this study.

Before any application (although results obtained show that the procedures used in this study are better than the ones currently being used by PROMEBO), this work should be completely redone with the following changes:

1) Include in the data set the Hereford calves born from 1979 to 1983. Being close to the technicians which collected the records from 1980 to 1984 will ensure the same level of knowledge about the data as in the present study. Some 25,000 records should be available after all editings.

2) Include some more explanatory variables in the model to be tested, especially interaction terms between age and seasonality
regressors and, if the data allows, adjoin to the nursing status classes the condition of whether the cow weaned her previous calf or not. This should be done after discussing the present model with peers, ANC technicians and breeders participating in PROMEBO.

3) Improve the efficiency of the programs by using better input/output statements; testing Gentleman (1974) WLS approach for Givens rotations; by optimally combining Cholesky and Givens methods; and, before the last iteration when computing REML estimates of variance components, since only the trace of part of the complete inverse of the coefficient matrix is needed, compute the diagonal elements of the desired part of the inverse just by adding the squares of the elements of each of the corresponding rows of \((T^*)^{-1}\).

4) Add cows as random effects in the model and use complete relationships among sires.

5) Try to improve robust estimation techniques, in a GLS setting, possibly by using the Median Absolute Deviation (MAD) or some weighted estimate of within CG variance. Most importantly, use residuals from the complete model and establish a minimum (possibly the pooled value across CG) value for the dispersion estimate within CG, to prevent excessive leverage from observations of some CG.

6) If paternal grand-sires are identified and sires are not for some observations, and those observations are kept, form sire incidence vectors with values of 0.5 and 0.0.
XI. CITATIONS


Miller, A. R. 1981. BASIC Programs for Scientists and Engineers. SYBEX Inc., Berkeley, California.


Pollack, E. J. 1983. Set of notes given to participants of the Animal Breeding Summer Course at Texas A & M University, College Station, TX. Cornell University, Ithaca, New York.


XII. COMMITTEE MEMBERS


Freeman, Albert E., Professor of Animal Science; Charles F. Curtiss Distinguished Professor in Agriculture. B.S., 1952, M.S., 1954, West Virginia; Ph.D., 1957, Cornell.


Willham, Richard L., Professor of Animal Science; Charles F. Curtiss Distinguished Professor in Agriculture. B.S., 1954, Oklahoma State; M.S., 1955, Ph.D., 1960, Iowa State. (Major Professor).

The author deeply appreciates corrections, criticisms, and suggestions by the committee members.
Appendix A. Problem Set Number 7 of Harville (1983)

The following data set consists of 12 observations, classified by "treatment" and by "block"; and the pattern of "filled subclasses" is that associated with an experimental design known as a balanced incomplete block (BIB) design. Values in the cells are the measured responses (1 or 0 observations per cell).

<table>
<thead>
<tr>
<th>BLOCKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRT</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

Observations are read by the program ordered by TRT and records are listed at end of source program.

Blocks (BLK) are first considered as fixed and after as random effects, when variance components are estimated.
Appendix B. **GIVFIX - a Two-way Fixed Model Program Using Givens Rotations**

```
1. //LAFRIES JOB U3242,FRIES
2. //STEP1 EXEC WATFIV
3. //GO.SYSIN DD *
4. SJOB U3242FRIES
5. C
6. C USING GIVENS ROTATIONS TO SOLVE L.S. PROBLEMS.
7. C
8. C
9. C
10. INTEGER LAST/12/,TOTALT/4/,TOTALB/6/,BOOLE,FPOS,LPOS,N,M
11. INTEGER POS(11,2),NB/1/,NT/1/,INICIO,FIM,DIFF
12. REAL*8 X/0.D0/,Y/0.D0/,TRACE/0.D0/,MSEF,ERRO/0.D0/,SEF
13. REAL*8 R(11)/11*0.D0/,S(11)/11*0.D0/,SU(11)/11*0.D0/
14. REAL*8 C(66)/66*0.D0/,U(66)/66*0.D0/,COS,SIN,SE(10)
15. REAL Y,FF,LL
16. CHARACTER*2 T,B,TRT(4),BLK(6)
17. C
18. C
19. C
20. C
21. C
22. NP=TOTALT+TOTALB
23. N=TOTALT+TOTALB+1
24. DO 10 I=1,N
25. FF=(N-I/2.)*(1-1.)*1
26. LL=(N-(I+1.)/2.)*1+1
27. POS(I,1)=FF
28. POS(I,2)=LL
29. CONTINUE
30. READ(5,800)T,B,Y
31. TRT(NT)=T
32. BLK(NB)=B
33. R(NT)=1.D0
34. R(TOTALT+NB)=1.D0
35. R(N)=Y
36. C FIRST OBS IS READ AND TRANSLATED TO R(1).
37. C
38. C
39. DO 200 I=2,LAST
40. READ(5,800)T,B,Y
41. IF(T.NE.TRAT(NT)) THEN
42. INICIO=POS(NNT,1)
43. FIM=POS(NNT,2)
44. DO 30 J=INICIO,FIM
45. U(J)=R(J-INICIO+NT)
46. CONTINUE
47. R(J-INICIO+NT)=0.D0
48. NT=NT+1
49. TRT(NT)=T
50. R(NT)=1.D0
51. BOOLE=0
52. DO 40 J=1,NB
53. IF(B.EQ.BLK(J)) THEN
54. BOOLE=1
55. R(TOTALT+J)=1.D0
56. CONTINUE
57. IF(BOOLE.EQ.0) THEN
58. NB=NB+1
59. BLK(NB)=B
60. R(TOTALT+NB)=1.D0
```
END IF

R(N)=Y
ELSE
C END IF, IF NO CHANGE IN TRT OR 'HYS'.
C
S(NT)=1.000
BOOLE=0
DO 60 J=1,NB
  IF(B.EQ.BLK(J)) THEN
    BOOLE=1
    S(TOTAL+J)=1.000
END IF
CONTINUE
60
IF(BOOLE.EQ.0) THEN
  NB=NB+1
  BLK(NB)=B
  S(TOTAL+NB)=1.000
END IF
S(N)=Y
C
C NOW THAT YOU HAVE THE VECTORS R AND S, ELIMINATE THE
C COLUMN IN S CORRESPONDING TO THE ACTUAL TRT, NT.
C
CALL G1(R(NT),S(NT),C0S,SIN,R(NT))
S(NT)=0.000
K=TOTAL+1
DO 80 J=K,N
  CALL G2(C0S,SIN,R(J),S(J))
80
R AND S ARE TRANSFORMED.
C
S NOW CONTAINS A '0' FOR 'ALL' TRTS' COLS. NOW THE PROBLEM IS
C TO ELIMINATE EACH OF ITS ELEMENTS CORRESPONDING TO THE SECOND
C FACTOR, BLK, WHICH IS NOT ORDERED AS TRT WAS.
C
DO 100 J=K,N
  IF(S(J).EQ.0.000) GO TO 100
  DO 85 L=J,N
    SU(L)=U(POS(J,1)+L-J)
    CALL G1(SU(J),S(J),C0S,SIN,SU(J))
    S(J)=0.000
    JJ=J+1
    DO 90 L=JJ,N
      IF(L.LE.N)THEN
        CALL G2(C0S,SIN,SU(L),S(L))
      END IF
90
    CONTINUE
95
U(POS(J,1)+L-J)=SU(L)
DO 96 L=K,N
SU(L)=0.000
DO 100 J=K,N
100 CONTINUE
150 CONTINUE
S(J)=0.000
C
SAVE IN U LAST R (CORRESPONDING TO LAST TRT.).

INICIO=POS(TOTALT,1)
FIM=POS(TOTALT,2)
DO 300 I=INICIO,FIM
300 U(I)=R(TOTALT+I-INICIO)

APPLYING CONSTRAINTS ON THE SOLUTIONS.
NOTE THAT RANK DEFICIENCY IS ONE:
THIS IS A NO-INTERCEPT MODEL.
SUMMATION OF BLOCK EFFECTS=0. 

k=TOTALT+1
DO 505 I=1,TOTALT
505 S(I)=0.000
DO 510 I=K,NP
510 S(I)=1.000
S(N)=0.000
DO 600 I=1,N
IF(S(I).EQ.0.000) GO TO 600
DO 520 J=I,N
520 R(J)=U(POS(1,1)+J-I)
CALL G1(R(I),S(I),C0S,S1N,R(I))
S(I)=0.000
JJ=I+1
DO 530 J=JJ,N
IF(J.LE.N) THEN
530 CALL G2(C0S,SIN,R(J),S(J))
END IF
530 CONTINUE
DO 540 I=1,N
540 U(POS(1,1)+J-I)=R(J)
550 R(J)=0.000
600 CONTINUE

CALCULATING THE COMPLETE INVERSE FROM THE UPPER TRIANGULAR.

INICIO=POS(1,1)
FIM=POS(1,2)-1
DO 654 J=INICIO,FIM
654 C(J)=U(J)
C(INICIO)=1/C(INICIO)
NNP=NP-1
DO 670 J=1,NNP
K=I+1
DO 670 J=K,NNP
X=0.000
654 CONTINUE
DO 655 J=1,NP
655 INICIO=POS(1,1)
FIM=POS(1,2)-1
654 C(J)=U(J)
655 C(INICIO)=1/C(INICIO)
NNP=NP-1
DO 670 J=1,NNP
K=I+1
DO 670 J=K,NNP
X=0.000
M=J-1
DO 660 L=1,M
X=X+C(POS(1,1)+L-1)*C(POS(J,1)+L-J)
C(POS(1,1)+J-1)=-X*C(POS(J,1))
END

VECTOR C CONTAINS THE INVERSE OF U. NOW, TO GET THE INVERSE
OF LHS IT IS NECESSARY TO POSMULTIPLY C BY ITS TRANSPOSE.

GETTING SOLUTIONS BY BACKSOLVING THE UPPER TRIANGULAR SYSTEM.

DO 700 I=1,10
TRACE=TRACE+C(POS(1,1))
END

DIFF=(LAST-1)-(TOTAL-1)-(TOTALS-1)
MSEF=(U(POS(N,1))^2)/DFF
SEF=DSQRT(MSEF)
PRINT,'TRACE=',TRACE
PRINT,'DFF=',OFF
PRINT,'MSEF=',MSEF
PRINT,'SEF=',SEF
PRINT,'UNSCALED C=',C
DO 750 I=1,10
C(POS(I,1))=C(POS(I,1))*MSEF
SE(1)=DSQRT(C(POS(I,1))
PRINT,('SOLUTION FIXED MODEL --- S.E. ESTIMATE')
END

SUBROUTINE G1(A,B,COS,SIN,SIG)
DOUBLE PRECISION A,B,COS,SIN,SIG,XR,YR
ZERO=0.0
ONE=1.0
IF (DABS(A).LE.DABS(B)) GO TO 10

241. XR = B / A
242. YR = DSQRT(ONE + XR**2)
243. COS = DSIGN(ONE / YR, A)
244. SIN = COS * XR
245. SIG = DABS(A) * YR
246. RETURN
247. 10 IF (B) 20, 30, 20
248. 20 XR = A / B
249. YR = DSQRT(ONE + XR**2)
250. SIN = DSIGN(ONE / YR, B)
251. COS = SIN**XR
252. SIG = DABS(B) * YR
253. RETURN
254. 30 SIG = ZERO
255. COS = ZERO
256. SIN = ONE
257. RETURN
258. END
259. C
260. C
261. C
262. C
263. C
264. SUBROUTINE G2(COS, SIN, X, Y)
265. DOUBLE PRECISION COS, SIN, X, Y, XR
266. XR = COS**X + SIN**Y
267. Y = -SIN**X + COS**Y
268. X = XR
269. RETURN
270. END
271. //
Appendix C. Partial Print-out of GIVFIX Execution with Records

Given in Appendix A
TRACE = 0.53611111111112D 01
DF= 3
MSEF= 0.1393333333333332D 03
SEF= 0.1180393413975051D 02

UNSCALED C= 0.4861111111111111D 00
-0.1527777777777778D 00
0.9722222222222222D 00
0.9722222222222222D 00
-0.1527777777777778D 00
0.9722222222222222D 00
-0.1527777777777778D 00
0.9722222222222222D 00
-0.1527777777777778D 00

ORDER OF SOLUTIONS FOR TRTS. IS
1 2 3 4
ORDER OF SOLUTIONS FOR BLKS. IS 1 3 5 4 6 2

STATEMENTS EXECUTED= 3993
CORE USAGE OBJECT CODE= 9808 BYTES, ARRAY AREA= 1508 BYTES.
DIAGNOSTICS NUMBER OF ERRORS= 0, NUMBER OF WARNINGS=
COMPILE TIME= 0.18 SEC, EXECUTION TIME= 0.13 SEC, 21.12.53
CSSTOP
Appendix D. Results from Using the Method of Successive Approximations for REML Estimation of a Variance Ratio (records from Appendix A)

<table>
<thead>
<tr>
<th>Old Ratio</th>
<th>New Ratio</th>
<th>Rounds</th>
<th>Change&lt;sup&gt;a&lt;/sup&gt;(- or +)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.0</td>
<td>14.25</td>
<td>21</td>
<td>-</td>
</tr>
<tr>
<td>14.796467</td>
<td>14.753385</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>14.0</td>
<td>13.991</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>13.9</td>
<td>13.895</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>13.8</td>
<td>13.7997</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>13.7901</td>
<td>13.7910096</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>13.791002</td>
<td>13.791001847</td>
<td>52</td>
<td>-</td>
</tr>
<tr>
<td>13.7910018292</td>
<td>13.791001829179</td>
<td>35</td>
<td>-, +(20)</td>
</tr>
<tr>
<td>13.79100182916</td>
<td>13.7910018291602</td>
<td>15</td>
<td>+, -(1)</td>
</tr>
<tr>
<td>13.791001829</td>
<td>13.79100182914</td>
<td>52</td>
<td>+</td>
</tr>
<tr>
<td>13.791001822</td>
<td>13.79100182809</td>
<td>49</td>
<td>+</td>
</tr>
<tr>
<td>13.79100166</td>
<td>13.79100181</td>
<td>52</td>
<td>+</td>
</tr>
<tr>
<td>13.791</td>
<td>13.79100164</td>
<td>52</td>
<td>+</td>
</tr>
<tr>
<td>13.7909</td>
<td>13.7909048</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.790561</td>
<td>13.7905799</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.7902</td>
<td>13.790234</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.7901</td>
<td>13.790139</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.79006</td>
<td>13.7901</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.79</td>
<td>13.79006</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.755</td>
<td>13.75661</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.752</td>
<td>13.7537</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.75</td>
<td>13.7518</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.7</td>
<td>13.7039</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.5</td>
<td>13.6799</td>
<td>21</td>
<td>+</td>
</tr>
<tr>
<td>13.5</td>
<td>13.489</td>
<td>21</td>
<td>+</td>
</tr>
<tr>
<td>13.343</td>
<td>13.362</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>13.0</td>
<td>13.512</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>5.0</td>
<td>10.39</td>
<td>21</td>
<td>+</td>
</tr>
</tbody>
</table>

<sup>a</sup>Change indicates sign of (NEW RATIO - OLD RATIO). Values in parentheses indicate last iteration, whose value is given on NEW RATIO, before solutions began to diverge, probably due to reaching machine error level.
Appendix E. GIVMLXCV - a Two-way Mixed Model Program to Obtain REML Estimates of Variance Components Using the Common Intercept Approach

1. \( \text{//LAFRIES Job U3242, FRIES} \)
2. \( \text{//STEP1 EXEC WATFIV} \)
3. \( \text{GO.SYSIN DD *} \)
4. \( \text{SJ0B U3242FRIES} \)
5. \( \text{C} \)
6. \( \text{C USING GIVENS ROTATIONS TO SOLVE L. S. PROBLEMS.} \)
7. \( \text{C} \)
8. \( \text{C} \)
9. \( \text{C MIXED MODEL} \)
10. \( \text{C} \)
11. INTEGER LAST/12/, TOTALT/4/, TOTALB/6/, BOOLE, LPOS, LPOS, N, M
12. INTEGER POS(11,2), NB/1/, NT/1/, INICIO, FIM, DFF
13. REAL*8 X/O.DO0/, V/O.DO0/, TRACE/O.DO0/, MSEF, ERROR/O.DO0/, SEF
14. REAL*8 R(11)/11*O.DO0/, S(11)/11*O.DO0/, SU(11)/11*O.DO0/
15. REAL*8 C(66)/66*O.DO0/, U(66)/66*O.DO0/, COS, SIN, SE(10)
16. REAL*8 RATIO, NRatio, SRatio, Diff
17. REAL*8 TRACB, REMLE, REMLB, BPB, DENOM
18. REAL*8 DIFF2/99.D00/, U2(66)/66*O.DO0/
19. REAL*8 ALFA(2), ALFA(15), DELTA(2)
20. REAL Y, FF, LL
21. CHARACTER*2 T, B, TRT(U), BLK(6)
22. CHARACTER*2 T, B, TRT(U), BLK(6)
23. CHARACTER*2 T, B, TRT(U), BLK(6)
24. CHARACTER*2 T, B, TRT(U), BLK(6)
25. CHARACTER*2 T, B, TRT(U), BLK(6)
26. CHARACTER*2 T, B, TRT(U), BLK(6)
27. CHARACTER*2 T, B, TRT(U), BLK(6)
28. CHARACTER*2 T, B, TRT(U), BLK(6)
29. CHARACTER*2 T, B, TRT(U), BLK(6)
30. CHARACTER*2 T, B, TRT(U), BLK(6)
31. CHARACTER*2 T, B, TRT(U), BLK(6)
32. CHARACTER*2 T, B, TRT(U), BLK(6)
33. CHARACTER*2 T, B, TRT(U), BLK(6)
34. CHARACTER*2 T, B, TRT(U), BLK(6)
35. CHARACTER*2 T, B, TRT(U), BLK(6)
36. CHARACTER*2 T, B, TRT(U), BLK(6)
37. CHARACTER*2 T, B, TRT(U), BLK(6)
38. CHARACTER*2 T, B, TRT(U), BLK(6)
39. CHARACTER*2 T, B, TRT(U), BLK(6)
40. CHARACTER*2 T, B, TRT(U), BLK(6)
41. CHARACTER*2 T, B, TRT(U), BLK(6)
42. CHARACTER*2 T, B, TRT(U), BLK(6)
43. CHARACTER*2 T, B, TRT(U), BLK(6)
44. CHARACTER*2 T, B, TRT(U), BLK(6)
45. CHARACTER*2 T, B, TRT(U), BLK(6)
46. CHARACTER*2 T, B, TRT(U), BLK(6)
47. CHARACTER*2 T, B, TRT(U), BLK(6)
48. CHARACTER*2 T, B, TRT(U), BLK(6)
49. CHARACTER*2 T, B, TRT(U), BLK(6)
50. CHARACTER*2 T, B, TRT(U), BLK(6)
51. CHARACTER*2 T, B, TRT(U), BLK(6)
52. CHARACTER*2 T, B, TRT(U), BLK(6)
53. CHARACTER*2 T, B, TRT(U), BLK(6)
54. CHARACTER*2 T, B, TRT(U), BLK(6)
55. CHARACTER*2 T, B, TRT(U), BLK(6)
56. CHARACTER*2 T, B, TRT(U), BLK(6)
57. CHARACTER*2 T, B, TRT(U), BLK(6)
58. CHARACTER*2 T, B, TRT(U), BLK(6)
59. CHARACTER*2 T, B, TRT(U), BLK(6)
60. CHARACTER*2 T, B, TRT(U), BLK(6)
R(TOTAL+J) = 1.000

END IF

40 CONTINUE

IF (BOOLE.EQ.0) THEN
NB = NB + 1
BLK(NB) = B
R(TOTAL+NB) = 1.000
END IF

R(N) = Y

ELSE

C ELSE, IF NO CHANGE IN TRT OR 'HYS'.

C

S(NT) = 1.000
BOOLE = 0
DO 60 J = 1, NB
IF (B.EQ.BLK(J)) THEN
BOOLE = 1
S(TOTAL+J) = 1.000
END IF

60 CONTINUE

END IF

60 CONTINUE

IF (BOOLE.EQ.0) THEN
NB = NB + 1
BLK(NB) = B
S(TOTAL+NB) = 1.000
END IF

S(N) = Y

NOW THAT YOU HAVE THE VECTORS R AND S, ELIMINATE THE
COLUMN IN S CORRESPONDING TO THE ACTUAL TRT, NT.

CALL G1(R(NT), S(NT), COS, SIN, R(NT))
S(NT) = 0.000

K = TOTAL + 1
DO 80 J = K, N
CALL G2(COS, SIN, R(J), S(J))
80

NOW THAT YOU HAVE THE VECTORS R AND S, ELIMINATE THE
COLUMN IN S CORRESPONDING TO THE ACTUAL TRT, NT.

CALL G1(R(NT), S(NT), COS, SIN, R(NT))
S(NT) = 0.000

K = TOTAL + 1
DO 80 J = K, N
CALL G2(COS, SIN, R(J), S(J))
80

R AND S ARE TRANSFORMED.

S NOW CONTAINS A '0' FOR 'ALL' TRTS' COLS. NOW THE PROBLEM IS
TO ELIMINATE EACH OF ITS ELEMENTS CORRESPONDING TO THE SECOND
FACTOR, BLK, WHICH IS NOT ORDERED AS TRT WAS.

DO 100 J = K, N
IF (S(J).EQ.0.000) GO TO 100
DO 95 L = J, N
SU(L) = U(POS(J, 1) + L - J)
CALL G1(SU(J), S(J), COS, SIN, SU(J))
S(J) = 0.000
J = J + 1
DO 90 L = JJ, N
IF (L.LE.N) THEN
CALL G2(COS, SIN, SU(L), S(L))
90 CONTINUE
END IF
DO 95 L = J, N
95 U(POS(J, 1) + L - J) = SU(L)
DO 96 L = K, N
SU(L)=0.000

CONTINUE

DO 150 J=1,N

S(J)=0.000

END IF

200 CONTINUE

SAVE IN U LAST R (CORRESPONDING TO LAST TRT.).

INICIO=POS(TOTALT,1)

FIM=POS(TOTALT,2)

DO 300 I=INICIO,FIM

U(I)=R(TOTALT+I-INICIO)

ADDED SQRT(RATIO) TO THE DIAGONAL PART CORRESPONDING TO THE RANDOM ELEMENTS. TO THE FIXED AND RHS ELEMENTS, ZEROS ARE ADDED.

USING THE COMMON INTERCEPT APPROACH AS IN SCHAEFFER'S NOTES.

THE INTERCEPT IS ESTIMATED AT FIFTEEN STAGES, SINCE THE CLOSER THE FIRST ESTIMATES ARE FROM THE TRUE VALUE, SO WILL BE THE CONVERGED VALUE. AND LESS ITERATIONS WILL BE NEEDED. NOTE THAT THE RATE OF CONVERGENCE IS NOT FAST IN REML.

LLL=POS(N,2)

ALFA(1)=5.0

ALFA(2)=20.0

DO 780 ISTAGE=1,5

DO 770 ILH=1,2

IF(ISTAGE.EQ.1) THEN

RATIO=ALFA(ILH)

ELSE

RATIO=ALFAC(ISTAGE-1)-DELTA(ILH)

ALFA(ILH)=RATIO

END IF

SRATIO=DSQRT(RATIO)

DO 503 I=1,LLL

DO 650 II=1,TOTALB

DO 510 J=1,N

S(J)=0.000

S(TOTAL+II)=SRATIO

K=TOTAL+1

DO 600 I=K,N

IF(S(I).EQ.0.000) GO TO 600
DO 520 J=1,N
R(J)=U2(POS(1,1)+J-1)
CALL G1(R(I),S(I),COS,SIN,R(I))
S(I)=0.000
JJ=I+1
DO 530 J=JJ,N
IF(J.LE.N) THEN
CALL G2(COS,SIN,R(J),S(J))
ENDIF
530 CONTINUE
DO 540 J=1,N
U2(POS(1,1)+J-1)=R(J)
540 CONTINUE
DO 550 J=1,N
R(J)=0.000
550 CONTINUE
DO 560 CONTINUE
DO 570 1=1,NNP
INICIO=POS(1,1)
FIN=POS(1,2)-1
DO 545 J=INICIO,FIN
C(J)=U2(J)
545 CONTINUE
C(INICIO)=1/C(INICIO)
NP=NP-1
DO 605 I=1,NP
INICIO=POS(I,1)
FIN=POS(I,2)-1
DO 545 J=INICIO,FIN
C(J)=U2(J)
545 CONTINUE
C(INICIO)=1/C(INICIO)
DO 655 1=1,NP
X=0.000
M=J-1
DO 660 L=1,M
X=X+C(POS(L,1)+J-L)*C(POS(J,1)+L-J)
660 CONTINUE
C(POS(M)+J-L)=X*C(POS(J,1))
655 CONTINUE
DO 675 1=1,NP
DO 675 J=1,NP
X=0.000
DO 672 L=J,NP
X=X+C(POS(L,1)+J-L-1)*C(POS(J,1)+L-J)
672 CONTINUE
C(POS(J,1)+L-J)=X
675 CONTINUE
TRACEB=0.000
K=TOTALT+1
DO 677 I=K,NP
677 CONTINUE
TRACEB=TRACEB+C(POS(I,1))
DO 700 I = 1, 10
  K = 11 - 1
  C(POS(K, 2)) = U2(POS(K, 2)) / U2(POS(K, 1))
  IF (K .GE. 2) THEN
    KK = K - 1
    DO 680 J = 1, KK
      U2(POS(J, 2)) = U2(POS(J, 2)) - C(POS(K, 2)) * U2(POS(J, 2) - 1)
    ENDIF
  ENDIF
  700 CONTINUE
  DFMN = LAST - TOTAL
  REMLE = U2(POS(N, 1))**2 / DFMN
  IF (ILH .EQ. 1) THEN
    PRINT, 'LOWER ESTIMATES-LOW-LOW-LOW-LOW-LOW-LOW'
  ELSE
    PRINT, 'HIGHER ESTIMATES-HIGH-HIGH-HIGH-HIGH-HIGH'
  ENDIF
  PRINT, 'DFMM=', DFMN
  PRINT, 'REMLE=', REMLE
  BPB = 0.0
  260 K = TOTALT + 1
  261 DO 760 L = K, NP
  262 BPB = BPB + (C(POS(L, 2))**2)
  263 760 BPB = BPB + (C(POS(L, 2))**2)
  264 DENOM = TOTALB - (RATIO * TRACEB)
  265 IF (DENOM .LT. 0.0) PRINT, 'DENOMINATOR IN REMLB IS NEGATIVE'
  266 REMLB = BPB / DENOM
  267 NRATIO = REMLE / REMLB
  268 DIFB = NRATIO / RATIO
  269 IF (DIFB .LT. 0.0) PRINT, 'DIFF BET RATIOS IS NEGATIVE'
  270 PRINT, 'BPB=', BPB
  271 PRINT, 'DENOM=', DENOM
  272 PRINT, 'RATIO=', RATIO
  273 PRINT, 'TRACEB=', TRACEB
  274 PRINT, 'REMEL=', REMLB
  275 PRINT, 'NRATIO=', NRATIO
  276 PRINT, 'DIFF BET RATIOS=', DIFB
  277 CONTINUE
  278 ALFAC(ISTAGE) = ALFA(1) * DELTA(2) - ALFA(2) * DELTA(1)
  279 ALFAC(1STAGE) = ALFAC(ISTAGE) / (DELTA(2) - DELTA(1))
  280 PRINT, 'CONVERGED VALUE AT STAGE=', ISTAGE, 'IS=', ALFAC(ISTAGE)
  281 CONTINUE
  282 800 FORMAT (A2, A2, F2.0)
  283 STOP
  284 END

SUBROUTINE G1(A,B,COS,SIN,SIG)
DOUBLE PRECISION A,B,COS,SIN,SIG,XR,YR
ZERO = 0.0
ONE = 1.0
IF (DABS(A).LE.0.0) GO TO 10
XR = B/A
YR = DSQRT(ONE + XR**2)
COS = DSIGN(ONE, XR, A)
SIN = COS * XR
SIG = DABS(A) * YR
RETURN
STOP
END
10 IF (B) 20, 30, 20
20 XR = A/B
30 XR = DSQR(T(ONE + XR**2))
30. SIN = DSQR(N(ONE / XR, B))
30. COS = SIN * XR
30. SIG = DABS(B) * XR
30. RETURN
30. SIG = ZERO
30. COS = ZERO
30. SIN = ONE
30. RETURN
30. END
31. C
31. C
31. C
31. C
31. C
31. SUBROUTINE G2(COS, SIN, X, Y)
32. DOUBLE PRECISION COS, SIN, X, Y, XR
32. XR = COS * X + SIN * Y
32. Y = SIN * X + COS * Y
32. X = XR
32. RETURN
32. END
33. SENTRY
33. 1 1 7
33. 1 3 5
33. 1 5 4 9
33. 2 4 3 2
33. 2 6 1 4
33. 3 2 1 5
33. 3 3 1 1
33. 3 6 2 2
33. 4 2 9
33. 4 4 1 5
33. 4 9 8
33. //
Appendix F. Results From Using GIVmixcv on Records Given in Appendix A

358. STAGE= 1
359. LOWER ESTIMATES-LOW-LOW-LOW-LOW-LOW-LOW
360. DFMM= 8.0000000
361. REMLE= 0.12940162907268150 03
362. BPB= 0.2810635611585349D 02
363. DENOM= 0.12030075187969920 01
364. RATIO= 0.5000000000000000D 01
365. TRACEB= 0.9593984962406016D 00
366. REMLB= 0.2336340852103222D 02
367. NEW RATIO= 0.5367437245226344D 01
368. DIFF BET RATIOS= 0.3674372452263444D 00
369. STAGE= 1
370. HIGHER ESTIMATES-HIGH-HIGH-HIGH-HIGH-HIGH
371. DFMM= 8.0000000
372. REMLE= 0.1415729166666663D 03
373. BPB= 0.2649739583333328D 01
374. DENOM= 0.3693181818181814D 00
375. RATIO= 0.2000000000000000D 02
376. TRACEB= 0.2815340909090909D 00
377. REMLB= 0.7174679487179481D 01
378. NEW RATIO= 0.1973229841411656D 02
379. DIFF BET RATIOS= -0.2677015858834366D 00
380. CONVERGED VALUE AT STAGE= 1 IS: 0.1367772274097112D 02
381. STAGE= 2
382. LOWER ESTIMATES-LOW-LOW-LOW-LOW-LOW-LOW
383. DFMM= 8.0000000
384. REMLE= 0.1415729166666663D 03
385. BPB= 0.5546969658495431D 01
386. DENOM= 0.534187964869727D 00
387. RATIO= 0.1331025957494780D 02
388. TRACEB= 0.410623975096057D 00
389. REMLB= 0.1037944315836480D 02
390. NEW RATIO= 0.1333087990840741D 02
391. DIFF BET RATIOS= 0.205944126626290-01
392. STAGE= 2
393. HIGHER ESTIMATES-HIGH-HIGH-HIGH-HIGH-HIGH
394. DFMM= 8.0000000
395. REMLE= 0.138789573645128D 03
396. BPB= 0.510455977108746D 01
397. DENOM= 0.5126570510778860D 00
398. RATIO= 0.1394542432685456D 02
399. TRACEB= 0.398689830690840D 00
400. REMLB= 0.9957065315266158D 01
401. NEW RATIO= 0.1393880317611447D 02
402. DIFF BET RATIOS= -0.6621150709794854D-02
403. CONVERGED VALUE AT STAGE= 2 IS: 0.1379090428753462D 02
404. STAGE= 3
405. LOWER ESTIMATES-LOW-LOW-LOW-LOW-LOW-LOW
406. DFMM= 8.0000000
407. REMLE= 0.138789573645128D 03
408. BPB= 0.5221096146856112D 01
409. DENOM= 0.5184779418128550D 00
410. RATIO= 0.1377030987487200D 02
411. TRACEB= 0.3980681702611045D 00
412. REMLB= 0.1007004486077062D 02
413. NEW RATIO= 0.1377119688260360D 02
414. DIFF BET RATIOS= 0.8870077316014413D-03
415. STAGE= 3
416. HIGHER ESTIMATES-HIGH-HIGH-HIGH-HIGH-HIGH
417. DFMM= 8.0000000
REMLE = 0.1386943007251424D 03
BPB = 0.5202724108184604D 01
DENOM = 0.5175646213733100D 00
RATIO = 0.1379752543824442D 02
TRACEB = 0.3973491770800061D 00
REMLB = 0.100523179006992D 02
NEW RAT10 = 0.1379724577908992D 02
01TF BET RAT10S = -0.2796591544997895D-03

CONVERGED VALUE AT STAGE = 3 IS: 0.1379100165537834D 02
CONVERGED VALUE AT STAGE = 4 IS: 0.1379100182891640D 02
CONVERGED VALUE AT STAGE = 5 IS: 0.1379100182916415D 02
CONVERGED VALUE AT STAGE = 6 IS: 0.1379100182916503D 02
CONVERGED VALUE AT STAGE = 7 IS: 0.1379100182916517D 02
CONVERGED VALUE AT STAGE = 8 IS: 0.1379100182916486D 02
CONVERGED VALUE AT STAGE = 9 IS: 0.1379100182916493D 02
CONVERGED VALUE AT STAGE = 10 IS: 0.1379100182916497D 02
CONVERGED VALUE AT STAGE = 11 IS: 0.1379100182916494D 02
CONVERGED VALUE AT STAGE = 12 IS: 0.1379100182916496D 02
CONVERGED VALUE AT STAGE = 13 IS: 0.1379100182916496D 02
CONVERGED VALUE AT STAGE = 14 IS: 0.1379100182916380D 02

STATEMENTS EXECUTED = 79383
CORE USAGE OBJECT CODE = 11320 BYTES, ARRAY AREA = 2188 BYTES.
DIAGNOSTICS NUMBER OF ERRORS = 0, NUMBER OF WARNINGS = 0.
COMPILE TIME = 0.22 SEC, EXECUTION TIME = 2.34 SEC.
Appendix G. Results From Using GIVMIXCV, Modified to Achieve Faster Convergence, on Records Given in Appendix A

<table>
<thead>
<tr>
<th>Stage</th>
<th>Lower Estimates</th>
<th>Higher Estimates</th>
<th>Diff Between Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LOW-LOW-LOW-LOW</td>
<td>HIGH-HIGH-HIGH-HIGH</td>
<td>0.2677015858834366</td>
</tr>
<tr>
<td>2</td>
<td>LOW-LOW-LOW-LOW</td>
<td>HIGH-HIGH-HIGH-HIGH</td>
<td>0.2677015858834366</td>
</tr>
<tr>
<td>3</td>
<td>LOW-LOW-LOW-LOW</td>
<td>HIGH-HIGH-HIGH-HIGH</td>
<td>0.2677015858834366</td>
</tr>
</tbody>
</table>

Converged Value at Stage = 1 is: 0.136772274097112D 02

Converged Value at Stage = 2 is: 0.136772274097112D 02

Converged Value at Stage = 3 is: 0.136772274097112D 02
418. REMLE= 0.1386903381611475D 03
419. BPB= 0.5206827210516934D 01
420. DENOM= 0.5177687371269682D 00
421. RATIO= 0.1379143677910785D 02
422. TRACEB= 0.3975098567101326D 00
423. REMLB= 0.100562964210719D 02
424. NEW RAT= 0.137914621921786D 02
425. DIFF BET RATIOS= 0.137914621921786D 02
426. CONVERGED VALUE AT STAGE= 3 IS: 0.1379100182561533D 02
427. CONVERGED VALUE AT STAGE= 4 IS: 0.1379100182916353D 02
428. CONVERGED VALUE AT STAGE= 5 IS: 0.1379100182916428D 02
429. CONVERGED VALUE AT STAGE= 6 IS: 0.1379100182916429D 02
430. CONVERGED VALUE AT STAGE= 7 IS: 0.1379100182916430D 02
431. CONVERGED VALUE AT STAGE= 8 IS: 0.1379100182916430D 02
432. CONVERGED VALUE AT STAGE= 9 IS: 0.1379100182916430D 02
433. CONVERGED VALUE AT STAGE= 10 IS: 0.1379100182916430D 02
434. CONVERGED VALUE AT STAGE= 11 IS: 0.1379100182916430D 02
435. CONVERGED VALUE AT STAGE= 12 IS: 0.1379100182916430D 02
436. CONVERGED VALUE AT STAGE= 13 IS: 0.1379100182916430D 02
437. CONVERGED VALUE AT STAGE= 14 IS: 0.1379100182916430D 02
438. STAGE= 15
439. LOWER ESTIMATES= LOW-LOW-LOW-LOW-LOW-LOW
440. DFMM= 8.0000000
441. REMLE= 0.1386900563663069D 03
442. BPB= 0.5207119061320898D 01
443. DENOM= 0.5177683252677358D 00
444. RATIO= 0.1379100182916423D 02
445. TRACEB= 0.3975212834595700D 00
446. REMLB= 0.1005656137852254D 02
447. NEW RAT= 0.1379100182916424D 02
448. DIFF BET RATIOS= 0.7105427357601002D-14
449. STAGE= 15
450. HIGHER ESTIMATES= HIGH-HIGH-HIGH-HIGH-HIGH-HIGH
451. DFMM= 8.0000000
452. REMLE= 0.1386900563663069D 03
453. BPB= 0.5207119061320898D 01
454. DENOM= 0.5177683252677358D 00
455. RATIO= 0.1379100182916423D 02
456. TRACEB= 0.3975212834595700D 00
457. REMLB= 0.1005656137852254D 02
458. NEW RAT= 0.1379100182916425D 02
459. DIFF BET RATIOS= 0.1865174681370263D-13
460. CONVERGED VALUE AT STAGE= 15 IS: 0.1379100182916425D 02
461. STATEMENTS EXECUTED= 79383
462. CORE USAGE= OBJECT CODE= 11336 BYTES, ARRAY AREA= 2188 BYTES,
463. CORE USAGE= OBJECT CODE= 11336 BYTES, ARRAY AREA= 2188 BYTES,
464. DIAGNOSTICS= NUMBER OF ERRORS= 0, NUMBER OF WARNINGS= 0,
465. COMPILE TIME= 0.22 SEC, EXECUTION TIME= 2.35 SEC,
466. CSSTOP
Appendix H. Listing of Program ABSSIRER

1. //LAFRIES JOB U3242,FRIES
2. //STO EXEC PGM=IEFBR14
3. //KILL DD DSN=L.U3242.UTEST,UNIT=3330,VOL=SER=AGS208,
4. // DISP=(OLD,DELETE)
5. //KILL2 DD DSN=L.U3242.NETEST,UNIT=3330,VOL=SER=AGS208,
6. // DISP=(OLD,DELETE)
7. //STEP1 EXEC WATFIV,REGION.G0=512K,TIME.G0=3
8. //GO.FT10F001 DD UNIT=3330,DISP=(OLD,KEEP),VOL=SER=AGS208,
9. // DSN=L.U3242.GIVTEST,DCB=(RECFM=FB,LRECL=515,BLKSIZE=12995)
10. //GO.FT11F001 DD DSN=L.U3242.UTEST,UNIT=3330,VOL=SER=AGS208,
11. // DISP=(NEW,KEEP),DCB=(RECFM=VB,S,LRECL=4232,BLKSIZE=4236),
12. // SPACE=(TRK,(1,1),RLSE)
13. //GO.FT12F001 DD DSN=L.U3242.NETEST,UNIT=3330,VOL=SER=AGS208,
14. // DISP=(NEW,KEEP),DCB=(RECFM=VS,LRECL=800,BLKSIZE=804),
15. // SPACE=(TRK,(1,1),RLSE)
16. //GO,SYSIN DD *
17. $JOB U3242,FRIES,TIME=180,PAGES=40
18. C
19. C
20. C
21. C
22. C
23. C
24. C
25. C
26. C
27. C
28. C
29. C
30. C
31. C
32. C
33. C
34. C R(1 TO 214) CONTAINS INCIDENCE VALUES FOR CG = HYSFS
35. C R(215) CONTAINS REGRESSOR VARIABLE (1,0,-1) FOR STAT1
36. C R(216) (0,1,-1) FOR STAT2
37. C R(217) LINEAR FOR AOD,WITHIN MALES.
38. C R(218) QUAD FOR AOD,WITHIN FEMALE.
39. C R(220) " MALES.
40. C R(221) " IF AOD>6,WITHIN MALES
41. C R(222) " ,WITHIN FEMALE.
42. C R(223) LINEAR FOR DATE OF BIRTH EFFECT
43. C R(224) QUAD FOR DATE OF BIRTH EFFECT
44. C R(225) " FOR DATE OF BIRTH EFFECT
45. C R(226) " IF BIRTHDATE>85,WITHIN FALL.
46. C R(227) " FOR DATE OF BIRTH EFFECT
47. C R(228) " WITHIN SPRING.
48. C R(229) " FOR DATE OF BIRTH EFFECT
49. C R(230) " FOR AOD,WITHIN FEMALES.
50. C R(231) " FOR AOD,WITHIN FEMALES.
51. C R(232) CONTAINS THE RESPONSE VARIABLE,WEANING WEIGHT.
USING A FIXED MODEL. SOLUTIONS WILL BE CHECKED AGAINST SAS.
IF (IBOOLE.EQ.0) THEN
    NSIRES = NSIRES + 1
    SIRE(I(NSIRES)) = SIRE(I)
    LESTRE(NANIM, NSIRES) = 1.000
END IF

DO 25 J = 1, NTOURO
    25 ANREAD(NANIM, J+1) = LESTRE(NANIM, J)
DO 30 J = 1, 18
    30 ANREAD(NANIM, J+NRBEG) = DBLE(LEREST(J))
GO TO 10
ELSE
    NANI = NANI - 1
    CONTINUE
REALNE(INCRCG, 1) = NANI
DO 50 I = 1, NANI
    50 REALNE(INCRCG, J) = REALNE(INCRCG, J) + ANREAD(I, J)
DO 70 I = 1, NANI
    70 ANRFAD(I, J) = ANREAD(I, J) - REALNE(INCRCG, J) / DFLOAT(NANI)
NANI = NANI + NANI + 1
    REALNE(INCRCG, NCOLNE) = DBLE(CG)
C
APPLY ROTATIONS AT END OF CG.
C
NTOUCC(INCRCG) = NSIRES
    IF (IBOOLE.EQ.0) NTOUCC(INCRCG) = NSIRES + 1
    K1 = NTOUCC(INCRCG)
    DO 110 I = 1, NANI
    DO 100 J = 2, NCOLLE
        IF (J.GT.K1.AND.J.LT.NRBEG) GO TO 100
        IF (ANREAD(I, J).EQ.0.000) GO TO 100
        ISSIGN = 1
        IF (ANREAD(I, J).LT.0.000) ISSIGN = -1
        DO 80 JJ = J, NCOLLE
            SWAP = -1 * U(J-1, JJ-1)
            U(J-1, JJ-1) = ISSIGN * ANREAD(I, JJ)
            80 ANREAD(I, JJ) = SWAP
        ELSE
            CALL G1(U(J-1, J-1), ANREAD(I, J), COS, SIN, U(J-1, J-1))
            IF (J.EQ.NCOLLE) GO TO 100
            JJ = J + 1
            DO 90 K = JJ, NCOLLE
            CALL G2(COS, SIN, U(J-1, K-1), ANREAD(I, K))
            90 CONTINUE
        END IF
    100 CONTINUE
    110 CONTINUE
END IF
IF (NANI.LT.ILAST) THEN
    DO 115 IS = 1, NTOURO
    115 ANREAD(1, IS) = LESTRE(NANIM+1, IS)
    NANIM = 1
    ULTCG = CG
    INCRG = INCRG + 1
    DO 120 J = 1, 18
    120 ANREAD(1, J+NRBEG) = DBLE(LEREST(J))
181. GO TO 10
182. ENDIF
183. DO 130 I=1,NTOUR
184. 130 RNAME(I)=SIREID(I)
185. C Printing of U or saving it.
186. C
187. C
188. PRINT 999
189. PRINT,'THE UPPER TRIANGULAR U'!
190. LUC=NCOLU/8+1
191. LMU=MOD(NCOLU,8)
192. DO 140 I=1,NCOLU
193. PRINT 901,1
194. DO 140 K=1,LU
195. JBE=K*8+1
196. JEND=K*8+LMU
197. 140 PRINT 902,(U(J),J=JBE,JEND)
198. C
199. C
200. WRITE U TO L.U3242.UTEST IN AGS208
201. C
202. C
203. WRITE (11) U
204. C
205. C
206. WRITE NORMAL EQUATIONS CORRESPONDING TO CG
207. C
208. C
209. WRITE (11) REALNE
210. C
211. C
212. PRINT 999
213. PRINT,'NORMAL EQUATIONS, CORRESPONDING TO THE CG:
214. C
215. C
216. C
217. C
218. C
219. C
220. C
221. C
222. C
223. C
224. C
225. C
226. C
227. C
228. C
229. C
230. DO 200 K=1,LNE
231. JBE=K*8+1
232. JEND=K*8+LMNE
233. 200 PRINT 910,(RNAME(J),J=1,7)
234. C
235. C
236. C
237. C
238. C
239. C
240. C
WRITE(FMT,899) IREP
899 FORMAT(('///,T2,'"CG"',11,'(9X,A6),9X,'"RHS"',/'))
PRINT FMT,(RNAME(J),J=JNBEG,JNLAST)
END IF
DO 200 I=1,INCRCG
ICGC=IFIX(SNGL(REALNE(I,NCOLNE)))
PRINT 912,ICGC,(REALNE(I,J),J=JBEG,JEND)
901 FORMAT('ROW=',I3)
902 FORMAT(T1,F6.0,A6,24X,F12.7,5(F2.0),5(F3.0),F5.0,F4.0,F3.0)
STOP
END

SUBROUTINE G1(A,B,COS,SIN,SIG)
DOUBLE PRECISION A,B,COS,SIN,SIG,XR,YR
ZERO=0.0
ONE=1.0
IF (DABS(A).LE.DABS(B)) GO TO 10
XR=B/A
YR=DSQRT(ONE+XR**2)
COS=DSIGN(ONE/YR,A)
SIN=COS*XR
SIG=DABS(A)*YR
RETURN
10 IF (B) 20,30,20
XR=A/B
YR=DSQRT(ONE+XR**2)
SIN=DSIGN(ONE/YR,B)
COS=SIN*XR
SIG=DABS(B)*YR
RETURN
30 SIG=ZERO
COS=ZERO
SIN=ONE
RETURN
END

SUBROUTINE G2(COS,SIN,X,Y)
DOUBLE PRECISION COS,SIN,X,Y,XR
XR=COS*X+SIGN(Y)
Y=SIN*X+COS*Y
X=XR
RETURN
END

ENTRY
//
Appendix I. Listing of Program REML

1. //LAFRIES JOB U3028,FRIES
2. /*KEY BB
3. //STEP1 EXEC WATFIV,REGION.GO=1280K,TIME.GO=10
4. //GO.FT11FOO1 DD DSN=L.U3242.US1RE,UNIT=3330,VOL=SER=AGS208,
5. // DISP=(OLD,KEEP),DCB=(RECFM=VBS,LRECL=10952,BLKSIZE=10956)
6. //GO.SYSIN DD *
7. $JOB U3028FRIES,TIME=600,PAGES=300
8. C
9. C
10. USING GIVENS ROTATIONS TO SOLVE G. L. S. PROBLEMS.
11. C
12. C
13. C MODEL IS A TWO-WAY CLASSIFICATION WITH COVARIATES(17).
14. C
15. C
16. C
17. C
18. C
19. C
20. C
21. C
22. C
23. C
24. C READ IN U, CALCULATED BY ABSSIRE.
25. C
26. C ADD SRATIO TO DIAGONAL AND ROTATE.
27. C
28. C ITERATE NEW RATIOS.
29. C
30. C
31. C
32. C
33. C
34. C
35. C
36. C
37. C BACKSOLVE AND PRINT SOLUTIONS.
38. C
39. C
40. C
41. C
42. C
43. C
44. C
45. C
46. C
47. C
48. C
49. C
50. C
51. C
52. C
53. C
54. C
55. C
56. C
57. C
58. C
59. C
60. C REDUCED MODEL

REAL*8 U(136,136),C(135,135)/18225*0.000/,X
REAL*8 SST1(135),SE,SOL(135),COS,SIN
REAL*8 UD1SK(148,148),ALFA(2),DELTA(2),ALFAC(10),S(136)
REAL*8 RATIO,NRATIO,REMLE,REMLS,BPB,DENOM,TRACES
CHARACTER RNAME*6(135)/130*/'A0Dn','A0D12','A0D21','A0D22','CA1'/

ILAST=8323
NCG=214
NSIRES=130
NREG=5
NP=NSIRES+NREG
NCOLU=NP+1
READ(11) UD1SK
DO 500 1=1,148
UDISK(1,131)=UDISK(1,133)
UDISK(1,132)=UDISK(1,134)
UDISK(1,133)=UDISK(1,135)
UDISK(1,134)=UDISK(1,136)
UDISK(1,135)=UDISK(1,146)
UDISK(1,136)=UDISK(1,148)
DO 500 CONTINUE

IT IS NEEDED TO TRIANGULARIZE IT TO 136 BY 136.

DO 600 1=132,148
DO 510 J=131,136
S(J)=UDISK(1,J)
DO 520 J=131,136
IF(S(J).NE.0.D00) THEN
CALL G1(UDISK(J,J),S(J),COS,SIN,UDISK(J,J))
S(J)=0.D00
IF(J.EQ.136) GO TO 520
K=J+1
DO 515 L=K,136
CALL G2(COS,SIN,UDISK(J,L),S(L))
END IF
DO 520 CONTINUE
DO 600 CONTINUE

ADDING SQRT(RATIO) TO THE DIAGONAL PART CORRESPONDING TO THE RANDOM ELEMENTS. TO THE FIXED AND RHS ELEMENTS, ZEROS ARE ADDED.
USING THE COMMON INTERCEPT APPROACH, TWICE, AS IN SCHAEFFER'S NOTES.

THE INTERCEPT IS ESTIMATED AT TWO STAGES, SINCE THE CLOSER THE
FIRST ESTIMATES ARE FROM THE TRUE VALUE, SO WILL BE THE CONVERGENCE. AND LESS ITERATIONS WILL BE NEEDED. NOTE THAT THE
RATE OF CONVERGENCE IS NOT FAST IN REML.

PRINT 999
ALFA(1)=23.8D00
ALFA(2)=27.8D00
DO 333 I=1,2

IF(I.EQ.1) THEN
  RATIO=ALFA(I)
ELSE
  RATIO=ALFA(I-1)-DELTA(I)
  ALFA(I)=RATIO
END IF
DO 5 I=1,NCOLU
DO 5 J=1,NCOLU
U(I,J)=UDISK(I,J)
DO 30 IS=1,NSIRES
DO 10 J=1,NCOLU
10 S(J)=0.D00
S(IS)=DSQRT(RATIO)
DO 30 J=IS,NCOLU
IF(S(J).EQ.0.D00) GO TO 30
CALL G1(U(J,J),S(J),COS,SIN,U(J,J))
JK=J+1
DO 20 K=JK,NCOLU
    IF(K.LE.NCOLU) CALL G2(COS,SIN,U(J,K),S(K))
20 CONTINUE
30 CONTINUE

CALCULATING THE INVERSE FROM THE UPPER TRIANGULAR.

DO 150 I=1,NP
11=1+1
C(I,I)=1.0D0/U(I,I)
DO 150 J=I,NP
150 CONTINUE

150 CONTINUE

NP=NP-1
DO 180 I=1,NP
11=1+1
180 CONTINUE

DO 180 J=11,NP
X=0.000
JK=J-1
DO 170 K=1,JK
X=X+C(I,K)*C(K,J)
170 C(I,J)=X
180 C(J,I)=X
190 C CONTAINS THE INVERSE OF U, NOW.
180 C NEXT IT'S NEEDED TO POSTMULTIPLY C BY ITS TRANSPOSE.
190 C
191 DO 200 I=1,NP
192 SST(I)=U(I,NCOLU)**2
193 DO 200 J=1,NP
194 X=0.000
195 DO 190 K=J,NP
196 X=X+C(I,K)*C(J,K)
197 C(I,J)=X
198 200 C(J,I)=X
199 TRACES=0.000
200 DO 210 I=1,NSIRES
201 210 TRACES=TRACES+C(I,I)
202 IDFEMM=LAST-NC=NCREG
203 REMLE=U(NCOLU,NCOLU)**2/IDFEMM
204 C
205 C
206 C BACKSOLVING
207 DO 300 I=1,NP
208 K=NCOLU-I
209 SOL(K)=U(K,NCOLU)/U(K,K)
210 IF(K.EQ.1) GO TO 300
211 KK=K-1
212 DO 250 J=1, KK
213 U(J,NCOLU)=U(J,NCOLU)-SOL(K)*U(J,K)
214 250 CONTINUE
215 CONTINUE
216 PRINT,'STAGE=',ISTAGE
217 ELSE
218 IF(ILH.EQ.1) THEN
219 PRINT,'LOWER ESTIMATES=LOW-LOW-LOW-LOW-LOW-LOW-LOW-LOW'
220 ELSE
221 PRINT,'HIGHER ESTIMATES=HIGH-HIGH-HIGH-HIGH-HIGH-HIGH-HIGH'
222 ENDIF
223 PRINT 950,1STAGE,REMLE
224 BPB=0.000
225 DO 220 I=1,NSIRES
226 220 BPB=BPB+SOL(I)**2
227 DENOM=NSIRES*(RATIO)**2
228 IF(DENOM.LE.0.000) PRINT,'DENOMINATOR IN REMLS IS NON-POSITIVE'
229 REMLS=BPB/DENOM
230 NRATIO=REMLES
231 230 DELTA(ILH)=NRATIO-RATIO
232 PRINT,'BPB=',BPB
233 PRINT,'TRACE SIRES=',TRACES
234 PRINT,'REMLES=',REMLES
235 PRINT,'OLD RATIO=',RATIO,'NEW RATIO=',NRATIO,'DIFF BET RATIOS'
236 322 CONTINUE
237 323 ALFAC(I STAGE)=ALFA(I)*DELTA(2)-ALFA(2)*DELTA(I)
239 ALFAC(I STAGE)=ALFAC(I STAGE)/(DELTA(2)-DELTA(1))
240 PRINT 960,ISTAGE,ALFAC(ISTAGE)
333 CONTINUE

PRINT 999
PRINT,'SOLUTION - MIXED MODEL:'
PRINT,'---------------------'
PRINT 910
DO 350 i=1,NP
SE=DSQRT(C(1,1)*REMLE)
TTEST=SOL(I)/SE
FTEST1=SST1(I)/REMLE
350 PRINT 920,RNAME(I),SOL(I),SE,TTEST,SST1(I).FTEST1
PRINTING OF XPX-INVERSE.
PRINT 999
PRINT,'UNSCALED XPX-INVERSE:'
LC=NP/9+1
LMC=MOD(NP,9)
IF(LMC.EQ.0) LC=LC-1
DO 230 K=1,LC
JBEG=(K-1)*9+1
JEND=K*9
IF(K.EQ.LMC) JEND=(K-1)*9+LMC
PRINT 930,(RNAME(J),J=JBEG,JEND)
230 PRINT 940,RNAME(I),(C(1,J),J=JBEG,JEND)
910 FORMAT;' SOURCE',5X,'SOLUTION',4X,'S.E.ESTIMATE',1X,
920 FORMAT,' T FOR HO: B=0',6X,'TYPE 1 SS',5X,'SEQUENTIAL F',//,
930 FORMAT,' ',19X,A6,8(7X,A6),//,
940 FORMAT(' ',4X,A6,5X,9(F13.8))
950 FORMAT(' ',12,'REMIE=',F22.16)
960 FORMAT(*DF=',12,'TH. STAGE IS:',F22.16,///)
999 FORMAT('1')
STOP
END

SUBROUTINE G1(A,B,COS,SIN,SIG)
DOUBLE PRECISION A,B,COS,SIN,SIG,XR,YR
ZER0=0.0
ONE=1.0
IF (DABS(A).LE.DABS(B)) GO TO 10
XR=A/B
YR=DSQRT(ONE+XR**2)
COS=DSIGN(ONE/YR,A)
SIN=COS*XR
SIG=DABS(A)*YR
RETURN
10 IF (B) 20,30,20
20 XR=A/B
30 YR=DSQRT(ONE+XR**2)
301.       SIN=DSIGN(ONE/YR,B)
302.       COS=SIN*XR
303.       SIG=DABS(B)*YR
304.       RETURN
305.       SIG=ZERO
306.       COS=ZERO
307.       SIN=ONE
308.       RETURN
309.       END
310.       C
311.       C
312.       C
313.       C
314.       C
315.       SUBROUTINE G2(COS,SIN,X,Y)
316.       DOUBLE PRECISION COS,SIN,X,Y,XR
317.       XR=COS*X+SIN*Y
318.       Y=-SIN*X+COS*Y
319.       X=XR
320.       RETURN
321.       END
322.       SENTRY
323.       //
### Appendix J. Sire Solutions in the Full and in the Reduced Models

**SYMBOLS ARE AS FOLLOWS:**

- $UF$=SIRE GLS ESTIMATES ON THE FULL MODEL.
- $EF$=STANDARD ERROR ($UF$).
- $UR$=SIRE GLS ESTIMATES ON THE REDUCED MODEL.
- $ER$=STANDARD ERROR ($UR$).
- $ABSDIFF$=ABSOLUTE VALUE $(UF-UR)$.
- $REALDIFF$=$(UF-UR)$.

<table>
<thead>
<tr>
<th>OBS</th>
<th>$UF$</th>
<th>$EF$</th>
<th>$UR$</th>
<th>$ER$</th>
<th>$ABSDIFF$</th>
<th>$REALDIFF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.6337</td>
<td>2.41520</td>
<td>-1.2855</td>
<td>2.40174</td>
<td>0.61175</td>
<td>0.6118</td>
</tr>
<tr>
<td>2</td>
<td>-1.1237</td>
<td>3.66330</td>
<td>-0.8327</td>
<td>3.61189</td>
<td>0.29101</td>
<td>-0.2910</td>
</tr>
<tr>
<td>3</td>
<td>0.9394</td>
<td>2.37642</td>
<td>0.5577</td>
<td>2.38274</td>
<td>0.37771</td>
<td>0.3777</td>
</tr>
<tr>
<td>4</td>
<td>-0.0447</td>
<td>2.72940</td>
<td>0.1863</td>
<td>2.72123</td>
<td>0.23099</td>
<td>-0.2310</td>
</tr>
<tr>
<td>5</td>
<td>2.4915</td>
<td>3.19630</td>
<td>3.0153</td>
<td>3.18186</td>
<td>0.52381</td>
<td>-0.5238</td>
</tr>
<tr>
<td>6</td>
<td>0.6598</td>
<td>3.45030</td>
<td>-0.0388</td>
<td>3.41493</td>
<td>0.69862</td>
<td>0.6986</td>
</tr>
<tr>
<td>7</td>
<td>-0.3122</td>
<td>3.29880</td>
<td>-0.5084</td>
<td>3.28142</td>
<td>0.19625</td>
<td>0.1962</td>
</tr>
<tr>
<td>8</td>
<td>-0.3187</td>
<td>3.44710</td>
<td>0.5402</td>
<td>3.41308</td>
<td>0.85910</td>
<td>-0.8599</td>
</tr>
<tr>
<td>9</td>
<td>0.6425</td>
<td>3.59190</td>
<td>0.8497</td>
<td>3.54577</td>
<td>0.20720</td>
<td>-0.2072</td>
</tr>
<tr>
<td>10</td>
<td>-4.7760</td>
<td>3.50740</td>
<td>-4.3230</td>
<td>3.47354</td>
<td>0.45299</td>
<td>0.4530</td>
</tr>
<tr>
<td>11</td>
<td>-0.3339</td>
<td>3.34180</td>
<td>-0.0738</td>
<td>3.32069</td>
<td>0.26009</td>
<td>-0.2601</td>
</tr>
<tr>
<td>12</td>
<td>-0.1500</td>
<td>3.70460</td>
<td>0.0798</td>
<td>3.64766</td>
<td>0.22978</td>
<td>-0.2298</td>
</tr>
<tr>
<td>13</td>
<td>-1.7844</td>
<td>2.50230</td>
<td>-1.1010</td>
<td>2.52768</td>
<td>0.68343</td>
<td>-0.6834</td>
</tr>
<tr>
<td>14</td>
<td>-1.7554</td>
<td>1.68910</td>
<td>-1.4871</td>
<td>1.70726</td>
<td>0.26831</td>
<td>-0.2683</td>
</tr>
<tr>
<td>15</td>
<td>-4.6527</td>
<td>1.72290</td>
<td>-4.7061</td>
<td>1.74299</td>
<td>0.05343</td>
<td>0.0534</td>
</tr>
<tr>
<td>16</td>
<td>-0.3807</td>
<td>3.12920</td>
<td>0.2220</td>
<td>3.12559</td>
<td>0.60271</td>
<td>-0.6027</td>
</tr>
<tr>
<td>17</td>
<td>1.1819</td>
<td>3.35360</td>
<td>0.6877</td>
<td>3.33356</td>
<td>0.49420</td>
<td>0.4942</td>
</tr>
<tr>
<td>18</td>
<td>2.0390</td>
<td>3.31560</td>
<td>2.0850</td>
<td>3.29755</td>
<td>0.04602</td>
<td>-0.0460</td>
</tr>
<tr>
<td>19</td>
<td>-4.3162</td>
<td>1.83100</td>
<td>-3.2194</td>
<td>1.85505</td>
<td>1.09838</td>
<td>-1.0968</td>
</tr>
<tr>
<td>20</td>
<td>1.5512</td>
<td>3.09080</td>
<td>1.8917</td>
<td>3.05832</td>
<td>0.34048</td>
<td>-0.3405</td>
</tr>
<tr>
<td>21</td>
<td>-3.6508</td>
<td>2.05070</td>
<td>-2.9380</td>
<td>2.07608</td>
<td>0.71275</td>
<td>-0.7128</td>
</tr>
<tr>
<td>22</td>
<td>1.7831</td>
<td>2.69740</td>
<td>0.2754</td>
<td>2.71509</td>
<td>1.50769</td>
<td>1.5077</td>
</tr>
<tr>
<td>23</td>
<td>2.3573</td>
<td>2.66960</td>
<td>2.1131</td>
<td>2.69107</td>
<td>0.24187</td>
<td>0.2418</td>
</tr>
<tr>
<td>24</td>
<td>-2.2603</td>
<td>3.70230</td>
<td>-2.3908</td>
<td>3.64745</td>
<td>0.13050</td>
<td>0.1305</td>
</tr>
<tr>
<td>25</td>
<td>-0.1586</td>
<td>2.60640</td>
<td>-0.3521</td>
<td>2.62824</td>
<td>0.19354</td>
<td>0.1935</td>
</tr>
<tr>
<td>26</td>
<td>2.7162</td>
<td>3.56200</td>
<td>0.5944</td>
<td>3.52005</td>
<td>2.12181</td>
<td>2.1218</td>
</tr>
<tr>
<td>27</td>
<td>4.6848</td>
<td>3.70960</td>
<td>4.4687</td>
<td>3.65371</td>
<td>0.17613</td>
<td>0.1761</td>
</tr>
<tr>
<td>28</td>
<td>6.0583</td>
<td>3.35470</td>
<td>4.9128</td>
<td>3.33409</td>
<td>1.14547</td>
<td>1.1455</td>
</tr>
<tr>
<td>29</td>
<td>1.2485</td>
<td>3.37420</td>
<td>1.6037</td>
<td>3.35057</td>
<td>0.35517</td>
<td>-0.3552</td>
</tr>
<tr>
<td>30</td>
<td>0.1610</td>
<td>3.56480</td>
<td>-0.4255</td>
<td>3.52418</td>
<td>0.58648</td>
<td>0.5865</td>
</tr>
<tr>
<td>31</td>
<td>-1.2018</td>
<td>3.68850</td>
<td>-0.7866</td>
<td>3.63308</td>
<td>0.41521</td>
<td>-0.4152</td>
</tr>
<tr>
<td>32</td>
<td>-0.6168</td>
<td>3.25120</td>
<td>0.3455</td>
<td>3.23946</td>
<td>0.96226</td>
<td>-0.9623</td>
</tr>
</tbody>
</table>

33  -0.8841  1.67340  -2.6541  1.68278  1.76997  1.7700  
34  3.1337  1.59940  1.6442  1.60685  1.48950  1.4895
|   | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 | 101 | 102 | 103 | 104 | 105 | 106 | 107 | 108 | 109 | 110 | 111 | 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 | 120 | 121 | 122 | 123 | 124 | 125 | 126 | 127 | 128 | 129 | 130 |
|   | 1.5172 | 2.90470 | 1.2752 | 2.90804 | 0.24200 | 0.2420 | 1.2202 | 3.73870 | 1.2008 | 3.67884 | 0.01944 | 0.0194 | -0.1280 | 3.73999 | -0.5831 | 3.67914 | 0.45517 | 0.4552 | 0.2530 | 3.45270 | -0.1278 | 3.42222 | 0.38080 | 0.3808 | -1.1234 | 3.00650 | -0.6873 | 3.09472 | 0.43611 | -0.4361 | -3.5912 | 3.15620 | -4.2116 | 3.12742 | 0.62038 | 0.6204 | -1.4279 | 3.26840 | -0.5698 | 3.21603 | 0.85812 | -0.8581 | 5.0192 | 3.61080 | 4.7814 | 3.56160 | 0.23783 | 0.2376 | 4.3077 | 2.69820 | 3.9146 | 2.67341 | 0.79311 | 0.7931 | -0.8779 | 3.69890 | -0.8451 | 3.64350 | 0.03285 | -0.0329 | 0.4077 | 2.75800 | 0.5363 | 2.76670 | 0.12861 | -0.1286 | 1.1671 | 3.74210 | 1.0276 | 3.68164 | 0.13950 | 0.1395 | -3.3199 | 3.33770 | -2.8892 | 3.31618 | 0.43066 | -0.4307 | 3.7651 | 3.30320 | 3.7071 | 3.28468 | 0.05800 | 0.0580 | -0.3034 | 2.68040 | 0.0732 | 2.69029 | 0.37659 | -0.3766 | 0.8608 | 3.62490 | 0.6509 | 3.57802 | 0.20990 | 0.2099 | 3.3886 | 2.56650 | 3.8931 | 2.57401 | 0.20449 | -0.5045 | 1.1789 | 2.47890 | 1.2701 | 2.48914 | 0.09120 | -0.0912 | 2.7969 | 2.37927 | 5.1416 | 2.39333 | 2.34472 | -2.3447 | 0.3165 | 3.7275 | -0.6662 | 3.66852 | 0.98273 | 0.9827 | -3.1632 | 3.7724 | -2.1845 | 3.70964 | 0.97867 | -0.9787 | -0.5504 | 3.8349 | -0.8703 | 3.76432 | 0.31991 | 0.3199 | 0.4674 | 3.7946 | 0.5678 | 3.72849 | 0.12043 | -0.1204 | -0.5434 | 3.7228 | -0.6134 | 3.66477 | 0.06999 | 0.0700 | -8.9830 | 3.2590 | -3.3684 | 2.35377 | 5.61462 | -5.6146 | -1.7380 | 3.6564 | -2.1886 | 3.60596 | 0.45064 | 0.4506 | -1.9111 | 3.5095 | -1.7162 | 3.47530 | 0.19493 | -0.1949 | -0.6567 | 3.1957 | 1.8892 | 3.18003 | 2.54593 | -2.5459 | 0.8488 | 3.0799 | 0.3368 | 3.07911 | 0.51197 | 0.5120 | -2.4203 | 2.6600 | -3.0192 | 2.67514 | 0.59890 | 0.5989 | -0.1170 | 3.2027 | -0.1971 | 3.19321 | 0.61408 | 0.6141 | 1.0350 | 2.5471 | 1.0758 | 2.55643 | 0.04082 | -0.0408 | 0.8037 | 2.4732 | 1.2232 | 2.48456 | 0.41954 | -0.4195 | -5.2322 | 3.0819 | -5.0559 | 3.07429 | 0.17626 | -0.1763 | 2.0377 | 2.84908 | 1.6481 | 2.84887 | 0.40936 | 0.4093 | -2.0806 | 3.3239 | -1.9894 | 3.32265 | 0.09123 | -0.0912 | -0.8312 | 3.2695 | -0.6181 | 3.25357 | 0.21308 | -0.2131 |