BU. 674-M
Rev. May 1980

Intersite Transfer of Estimated Response Surfaces

CONSTANCE L. WOOD

Department of Statistics, University of Kentucky,
Lexington, Kentucky 40506, U. S. A.

FOSTER B. CADY

Biometrics Unit, Cornell University,
Ithaca, New York 14853, U. S. A.

Summary

Transferability of agrotechnology assumes the feasibility of extrapolating a response-input relationship, estimated from experimental sites, to other sites with similar conditions. One specific conjecture is that crop production technology is transferable across sites within a soil family classification. The general approach to evaluating the transfer conjecture incorporates into the data analysis the prediction of yields not used in the estimation of the transfer function. A transfer model, using a second order response surface and measured site variable information, is formulated and yields for each experimental site are predicted from a transfer function estimated from the other sites. The resulting transfer residuals are compared with the ordinary within-site residuals. Based on a sum of squares criterion, a prediction test statistic is developed and shown to have a distribution of a ratio of independent quadratic forms. The methodology of transfer residuals is applied to data from the Benchmark Soil Project, where the major objective is to assess the feasibility of transferring agrotechnology among sites having soil of the same taxonomic classification.

Key Words: Regression; Prediction; Extrapolation; Controlled and uncontrolled variables; Agrotechnology transfer.

1. Introduction

Agrotechnology transfer is the extrapolation of a response-input relationship, estimated from a series of experiments, to new but similar sites. A major objective of the Benchmark Soils Project, established by U.S.A.I.D. (Agency for International Development) in cooperation with the Universities of Hawaii and Puerto Rico, is to assess the feasibility of crop production technology transfer from one tropical site to another on the basis of similarity of soils as indicated by the soil family in the Soil Taxonomy Classification System (Soil Survey Staff 1975). The conjecture is that experimental results, specifically the response surface relating maize yield to applications of phosphorus and nitrogen, obtained from a set of sites can be applied to new sites on the same soil family.

The soil family was selected because the family classification integrates soil factors with long-term environmental factors that influence crop yield. However, because of natural and past management variability, soil properties are not constant within discrete soil families. Consequently, homogeneous response to applied fertilizer treatments usually will not be found in practice. Interpreting this to mean that agrotechnology is not transferable can be faulty. In particular, the individual site response surface may be affected by the specific biotic environment of the site. Only by measuring uncontrolled site variables which reflect differences in environments, and including them in the response surface, will the response to the applied variables be clearly focused.

Statistical analysis of the transfer conjecture involves

(i) identification and estimation of a response surface model which adequately relates maize yields from several experimental sites to both the applied fertilizer levels and measured site variables,

(ii) evaluation of the predictive ability of the resulting estimated response surface for sites within the same soil family, but not included in the estimation process.

Least squares estimation procedures for incorporating site variable information in the analysis of orthogonal response surface models have been discussed generally by Cochran and Cox (1957, Chapter 14), and specifically by Colwell (1967), and are used throughout. Here the focus is on quantitative evaluation of the predictive ability of the resulting response surface.

In order to assess the predictive ability, the actual transfer of agrotechnology to sites where experimentation has not been carried out needs to be simulated. Our approach is to predict the yields for one of the k experimental sites, say the ith site, using the response surface estimated from the other (k - 1) sites. Since the ith site was excluded in the estimation procedure, the resulting (n × 1) column vector of predicted yields for the ith site will be denoted by $\hat{\underset{\sim}{Y}}_{[-i]}$. This prediction procedure is then repeated for each of the k sites, i.e., the predicted yields at a particular site are based on a response surface estimated from the other (k - 1) sites. The resulting prediction error is reflected in the k vectors of transfer residuals $\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_{[-i]}$, where $\underset{\sim}{Y}_i$ is the (n × 1) vector of observed yields for the ith site.

A quantitative evaluation of the predictive ability of the estimated response surface can then be based on a comparison of the transfer residuals $\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_{[-i]}$, with the least squares within-site residuals, $\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_i$, calculated by fitting the p variate response surface individually to each of the k sites. The specific objectives here are (i) to develop the transfer residual methodology for evaluating prediction and (ii) to demonstrate the methodology with yield and site variable information from maize transfer experiments on

the thixotropic, isothermic soil family of Hydric Dystrandepts. The first step is the development of a statistic for evaluating the transfer residuals.

2. Prediction Using Site Variables

Our approach utilizes a sum of squares criterion to compare the magnitude of the transfer residuals, $\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_{[-i]}$, to the ordinary within-site residuals, $\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_i$. In particular, Cady (1974, Experimental strategy for transferring crop production information, Technical Report 502 in the Biometrics Unit Mimeo Series, Cornell University, Ithaca, New York) proposed the ratio of the pooled sum of squared transfer residuals to the pooled sum of squared within-site residuals; i.e.,

$$P = \sum_{i=1}^{k} \left(\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_{[-i]}\right)' \left(\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_{[-i]}\right) \Big/ \sum_{i=1}^{k} \left(\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_i\right)' \left(\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_i\right) \quad .$$

For two sites (P - 1) is a symmetrized version of Gardner's (1972) ratio bias statistic used for assessing the predictive ability of one sample for a second sample. In the more general case of k sites, P is the natural extension comparing the predictive ability of the ith site for itself with the predictive ability of the remaining (k - 1) sites.

The distribution of (P - 1) is considered for a prediction model describing an experimental situation for which

(a) the same equally replicated treatment design was used at each site generating independent, normally distributed yields with common unknown experimental error variance, σ^2,

(b) the same p variate response surface model; e.g., a quadratic polynomial in two treatment factors ($p = 5$), adequately fits each site,

(c) differences among the site means, which would not affect the economically optimal rates of the treatment factors, have been eliminated by subtracting the site mean from yield data within each site, and

(d) differences in the estimated model parameters can be explained by interactions between the p x-variables of the prediction equation and the measured site variables. All x-variables used in interactions are required to be orthogonal and centered at zero.

For this prediction model, the statistic (P - 1) can be written as the ratio of two quadratic forms; i.e.,

$$(P-1) = \frac{k^2}{(k-1)^2} \cdot \frac{\underset{\sim}{Y}'\underset{\sim}{B}_1\underset{\sim}{Y}}{\underset{\sim}{Y}'\underset{\sim}{B}\underset{\sim}{Y}} ,$$

where $\underset{\sim}{B}_1$ and $\underset{\sim}{B}$ are $(kn \times kn)$ symmetric matrices with $\underset{\sim}{B}_1\underset{\sim}{B} = \underset{\sim}{0}$, $\underset{\sim}{B}\underset{\sim}{B} = \underset{\sim}{B}$ and $\underset{\sim}{Y}' = [\underset{\sim}{Y}_1' \vdots \underset{\sim}{Y}_2' \vdots \cdots \vdots \underset{\sim}{Y}_k']$. (See Appendix for details.) Since $\underset{\sim}{B}_1\underset{\sim}{B} = \underset{\sim}{0}$, the numerator and denominator are independent. Also, $\underset{\sim}{B}\underset{\sim}{B} = \underset{\sim}{B}$ implies that the denominator is distributed as $\sigma^2\chi^2[k(n-p-1)\text{d.f.}]$. The distribution of the numerator, unfortunately, cannot be so easily identified. However, the quadratic form, $\underset{\sim}{Y}'\underset{\sim}{B}_1\underset{\sim}{Y}$, can be represented as $\sigma^2\Sigma_{\ell=1}^{kn}\theta_\ell\chi^2_\ell$ (1 d.f.), where χ^2_ℓ (1 d.f.) are independent χ^2 random variables, each with 1 degree of freedom (d.f.), and θ_ℓ are the eigenvalues of $\underset{\sim}{B}_1$.

In general, the distribution of a linear combination of χ^2 variables cannot be simplified. However, in a particular problem, once P has been computed from the data and the eigenvalues of $\underset{\sim}{B}_1$ have been determined, the attained significance level of P can be accurately estimated by Monte Carlo methods.

One complication arises because the dimension of $\underset{\sim}{B}_1$ is extremely large for moderate k and n . This makes direct numerical computation of the eigenvalues not feasible. Due to the orthogonality and centering at zero of the x-variables used in interactions, this problem can be reduced to the computation of the eigenvalues of several matrices, each of dimension k, a problem

readily handled by available computer packages.

Specifically, if <u>two</u> orthogonal, centered, x-variables are used to form interactions with measured site variables, the eigenvalues of $\underset{\sim}{B}_1$ can be thought of as arising in three groups. The eigenvalues in the first are either zero or one, the number depending only on $(p-2)$ and the number of sites. Each of the remaining groups corresponds to one of the x-variables and the eigenvalues depend only on the site variables used as interactions with that x-variable. Two matrices, $\underset{\sim}{C}_1$ and $\underset{\sim}{C}_2$, with the required eigenvalues, are computed as follows:

(i) Construct $\underset{\sim}{T}_{(-i)1}$ as a $(k-1) \times m_1$ matrix of m_1 site variables used to form interactions with the first x-variable for all sites except the <u>i</u>th site. Each column of $\underset{\sim}{T}_{(-i)1}$ is centered at zero. The k excluded $(1 \times m_1)$ row vectors are denoted as $\underset{\sim}{T}_{i1}$. Similarly, construct $\underset{\sim}{T}_{(-i)2}$ and $\underset{\sim}{T}_{i2}$ for the second x-variable.

(ii) Form $\underset{\sim}{T}_1$ as a $(k \times k)$ matrix with diagonal elements equal to zero and the remaining elements in the <u>i</u>th row are given in order by the elements of the alias matrix, $\underset{\sim}{T}_{i1}\left(\underset{\sim}{T}'_{(-i)1}\underset{\sim}{T}_{(-i)1}\right)^{-1}\underset{\sim}{T}'_{(-i)1}$. Similarly form $\underset{\sim}{T}_2$.

(iii) Calculate $\underset{\sim}{C}_1$ as a $(k \times k)$ matrix by

$$\underset{\sim}{C}_1 = (\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k) + k^{-2}(k-1)^2[\underset{\sim}{T}'_1\underset{\sim}{T}_1 - k(k-1)^{-1}(\underset{\sim}{T}_1 + \underset{\sim}{T}'_1) + (k-1)^{-1}(\underset{\sim}{J}_k\underset{\sim}{T}_1 + \underset{\sim}{T}'_1\underset{\sim}{J}_k)]$$

where $\underset{\sim}{I}_k$ is a $(k \times k)$ identity matrix and $\underset{\sim}{J}_k$ is a $(k \times k)$ matrix of ones. Similarly, form $\underset{\sim}{C}_2$ from $\underset{\sim}{T}_2$.

If $\theta_1, \cdots, \theta_k$ denote the eigenvalues of $\underset{\sim}{C}_1$ and $\theta_{k+1}, \cdots, \theta_{2k}$ denote the eigenvalues of $\underset{\sim}{C}_2$, then the remaining eigenvalues of $\underset{\sim}{B}_1$ will either be zero or one. In particular, only $(k-1)(p-2)$ of the remaining eigenvalues will be equal to one; the rest will be zero. (See Appendix, equation (2).) The

term (p - 2) is the number of x-variables which are not used to form interactions with site variables.

Note that if site variables are not included in the response surface model and the above approach is followed using only x-variables, then (P - 1) is proportional to the usual F-statistic for testing equality of the k response surfaces.

3. Example

The Benchmark Soils Project is described by Silva and Beinroth (1978, Research on agrotechnology transfer in the tropics based on the soil family (Progress Report 1, Benchmark Soils Project), Technical Report, Department of Agronomy and Soil Science, University of Hawaii, Honolulu, Hawaii). As indicated in Section 1, a major objective of the project is to assess the feasibility of transferring agrotechnology in the tropics on the basis of soil taxonomic units, thereby reducing the amount of site specific experimentation. Specifically, the conjecture that an estimated response-input relationship can be transferred within the same soil family needs to be evaluated. This example uses grain yield (kg/ha) data from five maize experiments on the Hydric Dystrandept soil family; two sites (PUC-K and BUR-B) are in the Philippines, two in Hawaii (KUC-C and KUK-D) and one in Indonesia (LPH-E). The same 13-point treatment design with three replications was used at each site, a partial 5 X 5 factorial with applied phosphorus and applied nitrogen as the controlled variables. An estimated second order response surface model in the two factors adequately fits the treatment means.

Given here are the numerical details of calculating the P statistic under a prediction model which introduces site variable information in the transfer function as interactions between the site variables and the linear effects of applied phosphorus (P) and applied nitrogen (N). Table 1 gives

the basic site variable information, (EXTN = extractable soil nitrogen, MINT = average daily minimum temperature for eight weeks around tasseling and EXTP = extractable soil phosphorus), the within-site residual sums of squares (SS) based on fitting the second order response surface to each site, and the transfer SS based on the prediction equation estimated from the other sites. Four interactions are included in the prediction equation, applied phosphorus with EXTP and EXTN and applied nitrogen with MINT and EXTN.

[Insert Table 1 here]

In particular, the data matrix of x-variables corresponding to the second order response surface in P and N individually fitted to each site is given by:

$$\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{X}^{*} \\ \vdots \\ \underset{\sim}{X}^{*} \\ \vdots \\ \underset{\sim}{X}^{*} \end{bmatrix}, \quad \text{where} \quad \underset{\sim}{X}^{*} = \begin{bmatrix} P & N & P^2 & N^2 & NP \\ -.85 & -.85 & .340 & .340 & .723 \\ -.85 & 0 & .340 & -.383 & 0 \\ -.85 & .85 & .340 & .340 & -.723 \\ -.40 & -.40 & -.223 & -.223 & .160 \\ -.40 & .40 & -.223 & -.223 & -.160 \\ 0 & -.85 & -.383 & .340 & 0 \\ 0 & 0 & -.383 & -.383 & 0 \\ 0 & .85 & -.383 & .340 & 0 \\ .40 & -.40 & -.223 & -.223 & -.160 \\ .40 & .40 & -.223 & -.233 & .160 \\ .85 & -.85 & .340 & .340 & -.723 \\ .85 & 0 & .340 & -.383 & 0 \\ .85 & .85 & .340 & .340 & .723 \end{bmatrix}.$$

A column of ones is not included because the yields have been centered at zero and, due to the three replications, the $\underset{\sim}{X}^{*}$ is repeated three times. Note that coded levels of the treatments are used and the P^2 and N^2 columns have been centered at zero by subtracting the original column mean.

TABLE 1

Site Variable Data, Residual Sums of Squares

and Transfer Sums of Squares

Site	EXTN (ppm)	MINT (°C)	EXTP (ppm)	Residual SS ($x10^3$)	Transfer SS ($x10^3$)
PUC-K	79	23.00	10	5,869	14,700
BUR-B	29	21.50	5	25,055	36,584
KUK-C	46	18.83	74	13,602	18,695
KUK-D	29	17.90	62	25,599	32,792
LPH-E	119	16.76	23	17,880	23,660
				88,005	126,431

Next, since interactions with the first two columns of $\underset{\sim}{X}$, say $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$, respectively, are used in the prediction equation, the data matrices for the first site, PUK-K, for calculating $\hat{\underset{\sim}{Y}}_{[-1]}$ are

$$\underset{\sim}{X}_1 = [\underset{\sim}{X} \vdots \overset{P \times EXTP}{(10)\underset{\sim}{x}_1} \vdots \overset{P \times EXTN}{(79)\underset{\sim}{x}_1} \vdots \overset{N \times MINT}{(23)\underset{\sim}{x}_2} \vdots \overset{N \times EXTN}{(79)\underset{\sim}{x}_2}]$$

and

$$\underset{\sim}{X}_{(-1)} = \begin{matrix} \\ \underset{\sim}{X} \\ \underset{\sim}{X} \\ \underset{\sim}{X} \\ \underset{\sim}{X} \end{matrix} \begin{matrix} \overset{P \times EXTP}{} \\ \\ \\ \\ \end{matrix} \begin{matrix} \overset{P \times EXTN}{} \\ \\ \\ \\ \end{matrix} \begin{matrix} \overset{N \times MINT}{} \\ \\ \\ \\ \end{matrix} \begin{matrix} \overset{N \times EXTN}{} \\ \\ \\ \\ \end{matrix}$$

$$\begin{bmatrix} (5)\underset{\sim}{x}_1 & (29)\underset{\sim}{x}_1 & (21.5)\underset{\sim}{x}_2 & (29)\underset{\sim}{x}_2 \\ (74)\underset{\sim}{x}_1 & (46)\underset{\sim}{x}_1 & (18.83)\underset{\sim}{x}_2 & (46)\underset{\sim}{x}_2 \\ (62)\underset{\sim}{x}_1 & (29)\underset{\sim}{x}_1 & (17.90)\underset{\sim}{x}_2 & (29)\underset{\sim}{x}_2 \\ (23)\underset{\sim}{x}_1 & (119)\underset{\sim}{x}_1 & (16.76)\underset{\sim}{x}_2 & (119)\underset{\sim}{x}_2 \end{bmatrix} .$$

Then $\hat{\underset{\sim}{Y}}_{[-1]} = \underset{\sim}{X}_1\left(\underset{\sim}{X}'_{(-1)}\underset{\sim}{X}_{(-1)}\right)^{-1}\underset{\sim}{X}'_{(-1)}\underset{\sim}{Y}_{(-1)}$, where $\underset{\sim}{Y}_{(-1)}$ is the $(k-1)n \times 1$ vector of yields for all sites except the first, gives the transfer residuals for the first site, $\underset{\sim}{Y}_1 - \hat{\underset{\sim}{Y}}_{[-1]}$. The transfer residuals for the other sites are computed similarly.

From Table 1, we see that the prediction statistic is

$$P = \frac{\text{Transfer SS}}{\text{Residual SS}} = \frac{126,431,000}{88,005,000} = 1.44 \quad .$$

In other words, a 44% increase in unexplained variability when predicting the ith site from the remaining sites is observed using the model with five design variables (quadratic polynomial) and four interactions with the site variables.

The next step is to assess whether this 44% increase is to be expected, or is so large as to contradict the ability to transfer results from one experiment to another. From Section 2, we have that

$$\frac{(k-1)^2}{k^2}(P-1) = 0.64\,(P-1) \sim \frac{\sigma^2 \sum_{\ell=1}^{22} \theta_\ell \chi^2_\ell (1 \text{ d.f.})}{\sigma^2 \chi^2 (165 \text{ d.f.})}$$

where $22 = k + k + (k-1)(p-2)$ and $165 = k(n-p-1)$ and θ_ℓ are the eigenvalues of $\underset{\sim}{B}_1$.

Following the construction method outlined in Section 2, $\underset{\sim}{C}_1$ is computed from the site variables EXTP and EXTN. While the original values of the site variables are to be used in $\underset{\sim}{X}_{(-i)}$, the values in $\underset{\sim}{T}_{(-i)j}$ are the deviations from the mean of the (k - 1) sites involved. (See Appendix, equation (1).) In particular, $\underset{\sim}{C}_1$ and its eigenvalues are:

$$\underset{\sim}{C}_1 = \begin{bmatrix} 3.870 & -2.037 & 0.067 & 0.262 & -2.163 \\ -2.037 & 1.240 & 0.459 & -0.707 & 1.045 \\ 0.067 & 0.459 & 1.457 & -1.670 & -0.313 \\ 0.262 & -0.707 & -1.670 & 1.945 & 0.170 \\ -2.163 & 1.045 & -0.313 & 0.170 & 1.261 \end{bmatrix} \quad \text{and} \quad \begin{aligned} \theta_1 &= 6.208 \\ \theta_2 &= 3.564 \\ \theta_3 &= 0.000 \\ \theta_4 &= 0.000 \\ \theta_5 &= 0.000 \end{aligned} \quad .$$

Similarly, $\underset{\sim}{C}_2$ is computed from minimum temperature (MINT) and extractable nitrogen (EXTN). This yields the eigenvalues

$$\theta_6 = 11.705, \quad \theta_7 = 2.666, \quad \theta_8 = 0.000, \quad \theta_9 = 0.000, \quad \text{and} \quad \theta_{10} = 0.000 \quad .$$

Combining these facts we see that

$$0.64\,(P-1) \sim \sigma^2 \sum_{\ell=1}^{22} \theta_\ell \chi^2_\ell \text{ (1 d.f.)} / \sigma^2 \chi^2 \text{ (165 d.f.)} \quad .$$

We need to compare the observed value of $[(k-1)^2/k^2](P-1) = 0.280$ with the quantiles of the distribution of $\Sigma_\ell \theta_\ell \chi^2_\ell$ (1 d.f.)/χ^2 (165 d.f.) . As stated earlier, the distribution of such a linear combination of χ^2_ℓ (1 d.f.) as found in the numerator cannot be simplified, while the denominator is an independent χ^2 (165 d.f.) variable.

Even though no tables exist for the distribution of $[(k-1)^2/k^2](P-1)$, the attained significance level may be readily estimated by Monte Carlo simulation. Using the fact that a standard normal variable squared is χ^2 (1 d.f.) and that the numerator and denominator are independent, many variables with the above distribution may be computed and the proportion which falls above the computed value of 0.280 recorded. This will give an accurate estimate of the attained significance level. In this example, we may make a further simplification. Since the number of d.f. of the denominator is so large, the residual mean square is very close to σ^2, the unknown experimental error, with high probability. Rewriting

$$[(k-1)^2/k^2](P-1) \doteq \frac{\sigma^2 \sum_{\ell=1}^{22} \theta_\ell \chi^2_\ell (1 \text{ d.f.})}{165\sigma^2} \quad \text{or} \quad 165[(k-1)^2/k^2](P-1) \doteq \sum_{\ell=1}^{22} \theta_\ell \chi^2_\ell (1 \text{ d.f.}) \ .$$

This implies that we need only compare $165(0.280) = 46.2$ with the quantiles of $\Sigma_\ell \theta_\ell \chi^2_\ell$ (1 d.f.) .

Ten thousand random variables with the distribution given above were generated. In particular, at each iteration, 22 standard normal random variables, say N_ℓ, $\ell = 1, \cdots, 22$, were generated using GGUSN from the IMSL Library (Houston, Texas, U.S.A.). Then each variable was formed as the linear combination of χ^2_ℓ (1 d.f.) (N_ℓ^2) variables given above. The attained significance level is 0.236 .

Constant experimental error variances across sites have been assumed. The residual sums of squares in Table 1 make this assumption dubious. However, if the error variances are heterogeneous, $\sigma_1^2, \cdots, \sigma_k^2$, but known, the above analysis remains valid with minor modifications of the eigenvalues. In particular, if

$$\underset{\sim}{D} = \text{diag}\left[\left(\sigma_1^2 \Big/ \sum_{i=1}^{k}\sigma_i^2\right), \cdots, \left(\sigma_k^2 \Big/ \sum_{i=1}^{k}\sigma_i^2\right)\right] ,$$

then $\theta_1, \cdots, \theta_k$ are the eigenvalues of $\underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{C}_1\underset{\sim}{D}^{\frac{1}{2}}$, and $\theta_{k+1}, \cdots, \theta_{2k}$ are the eigenvalues of $\underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{C}_2\underset{\sim}{D}^{\frac{1}{2}}$. Unfortunately, the last (k-1)(p-2) eigenvalues are no longer ones but are the eigenvalues of $\underset{\sim}{D}^{\frac{1}{2}}(\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k)\underset{\sim}{D}^{\frac{1}{2}}$, each with multiplicity (p-2) .

For this example, we estimate $\underset{\sim}{D}$ = diag(0.067,0.285,0.155,0.291,0.203) . More especially, the mean square residual for each site with 33 d.f. provides a sufficiently close estimate of the unknown error variance. Then the eigenvalues are:

$\theta_1 = 4.737$	$\theta_6 = 13.496$	$\theta_{11} = \theta_{12} = \theta_{13} = 1.439$
$\theta_2 = 3.554$	$\theta_7 = 2.699$	$\theta_{14} = \theta_{15} = \theta_{16} = 1.208$
$\theta_3 = 0.0$	$\theta_8 = 0.0$	$\theta_{17} = \theta_{18} = \theta_{19} = 0.881$
$\theta_4 = 0.0$	$\theta_9 = 0.0$	$\theta_{20} = \theta_{21} = \theta_{22} = 0.472$
$\theta_5 = 0.0$	$\theta_{10} = 0.0$	

A Monte Carlo simulation yielded a significance level of 0.240 for P - 1 .

Acknowledgments

The authors acknowledge their appreciation to the editors and referees for valuable suggestions, and to Dr. James A. Silva of the University of Hawaii and Dr. T. S. Gill of U.S.A.I.D. for their support and encouragement. We also acknowledge the use of the data from the Benchmark Soils Project, a U.S.A.I.D. Project with the Department of Agronomy and Soil Science, University of Hawaii.

References

Cochran, W. G. and Cox, G. M. (1957). Experimental Designs, 2nd edition. John Wiley and Sons, New York.

Colwell, J. D. (1967). Calibration and assessment of soil tests for estimating fertilizer requirements. Australian Journal of Soil Research 5, 275-293.

Gardner, M. J. (1972). On using an estimated regression line in a second sample. Biometrika 59, 263-274.

Searle, S. R. (1971). Linear Models. John Wiley and Sons, New York.

Soil Survey Staff, Soil Conservation Service, U.S. Department of Agriculture. (1975). Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys. Agriculture Handbook 436, U.S. Government Printing Office, Washington, D. C.

Appendix

Distribution of (P - 1)

Determining the distribution of P - 1 begins by expressing the combined vector of transfer residuals as a linear combination of kn independent normal errors $\underset{\sim}{\epsilon}$, each with common variance $\sigma^2 > 0$; i.e., $\underset{\sim\sim}{R}\underset{\sim}{\epsilon}$. Then

$$P - 1 = \frac{\underset{\sim}{\epsilon}'(\underset{\sim\sim}{R}'\underset{\sim\sim}{R} - \underset{\sim\sim}{B})\underset{\sim}{\epsilon}}{\underset{\sim}{\epsilon}'\underset{\sim\sim}{B}\underset{\sim}{\epsilon}}$$

where $\underset{\sim}{\epsilon}'\underset{\sim\sim}{B}\underset{\sim}{\epsilon}$ is the pooled sum of squared residuals divided by σ^2 . Next the numerators and denominators are shown to be independent. Lastly the eigenvalues of the quadratic form in the numerator, say $\theta_1, \cdots, \theta_q$ with $q = q(n,k,p,m_1,m_2)$, are determined.

First we develop the data matrices used for estimating the response surface required at each step. Let $\underset{\sim}{X}$ denote the $(n \times p)$ matrix of x-variables with each column centered at zero. Further, let $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$ be two orthogonal columns of $\underset{\sim}{X}$ which will be used to form interactions with site variables. Then, using the notation developed in Section 2, the true response function at the ith site is given by $\underset{\sim}{X}_i\underset{\sim}{\beta}$, where

$$\underset{\sim}{X}_i = \left[\underset{\sim}{X} \vdots \underset{\sim}{T}_{i1} \otimes \underset{\sim}{x}_1 \vdots \underset{\sim}{T}_{i2} \otimes \underset{\sim}{x}_2\right] , \qquad (1)$$

where $\otimes$ denotes the right Kronecker product. Note that without loss of generality, centered site variables may be used to form interactions. This follows from the fact that $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$ are centered at zero. Then the predicted values for the ith site based on the remaining $(k-1)$ sites is $\hat{\underset{\sim}{Y}}_{[-i]} = \underset{\sim}{X}_i \underset{\sim}{b}_{[-i]}$, where $\underset{\sim}{b}_{[-i]}$ is estimated only from the remaining $(k-1)$ sites.

In order to compute $\underset{\sim}{b}_{[-i]}$, we require the data matrix for all sites except the ith, say $\underset{\sim}{X}_{(-i)}$. In particular,

$$\underset{\sim}{X}_{(-i)} = \left[\underset{\sim}{1}_{k-1} \otimes \underset{\sim}{X} \vdots \underset{\sim}{T}_{(-i)1} \otimes \underset{\sim}{x}_1 \vdots \underset{\sim}{T}_{(-i)2} \otimes \underset{\sim}{x}_2\right] , \qquad i = 1, \cdots, k .$$

In Section 2, the yields were adjusted for the individual site means. In terms of the true response function, the adjusted yields are given by

$$\underset{\sim}{Y}_i = \underset{\sim}{X}_i\underset{\sim}{\beta} + (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)\underset{\sim}{\epsilon}_i , \qquad i = 1, \cdots, k$$

and, similarly, the combined $(k-1)n \times 1$ vector of adjusted yields excluding the ith site is

$$\underset{\sim}{Y}_{(-i)} = \underset{\sim}{X}_{(-i)}\underset{\sim}{\beta} + \left[\underset{\sim}{I}_{(k-1)} \otimes (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)\right]\underset{\sim}{\epsilon}_{(-i)} , \qquad i = 1, \cdots, k ,$$

where $\underset{\sim}{\epsilon}_{(-i)}$ is analogously defined.

Since the adjusted yields are not independent, $\underset{\sim}{b}_{[-i]}$ is computed using the methods of generalized least squares (Searle 1971, Section 5.8). Noting that the covariance matrix of the adjusted yields is $[\underset{\sim}{I}_{(k-1)} \otimes (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)]$, which is idempotent and $[\underset{\sim}{I}_{(k-1)} \otimes (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)]\underset{\sim}{X}_{(-i)} = \underset{\sim}{X}_{(-i)}$,

$$\underset{\sim}{b}_{[-i]} = \left(\underset{\sim}{X}_{(-i)}'\underset{\sim}{X}_{(-i)}\right)^{-1}\underset{\sim}{X}_{(-i)}'\left[\underset{\sim}{I}_{(k-1)} \otimes (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)\right]\left\{\underset{\sim}{X}_{(-i)}\underset{\sim}{\beta} + \left[\underset{\sim}{I}_{(k-1)} \otimes (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)\right]\underset{\sim}{\epsilon}_{(-i)}\right\}$$

$$= \underset{\sim}{\beta} + \left(\underset{\sim}{X}_{(-i)}'\underset{\sim}{X}_{(-i)}\right)^{-1}\underset{\sim}{X}_{(-i)}'\underset{\sim}{\epsilon}_{(-i)}, \qquad i = 1, \ldots, k \quad .$$

Therefore the $(n \times 1)$ vector of transfer residuals is given by

$$\underset{\sim}{Y}_i - \hat{\underset{\sim}{Y}}_{[-i]} = (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)\underset{\sim}{\epsilon}_i - \underset{\sim}{X}_i\left(\underset{\sim}{X}_{(-i)}'\underset{\sim}{X}_{(-i)}\right)^{-1}\underset{\sim}{X}_{(-i)}'\left[\underset{\sim}{I}_{(k-1)} \otimes (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)\right]\underset{\sim}{\epsilon}_{(-i)} \quad .$$

Next we must determine $\underset{\sim}{R}$. First consider $\underset{\sim}{Y}_1 - \hat{\underset{\sim}{Y}}_{[-1]}$. For simplicity, let $\underset{\sim}{e}_i = (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)\underset{\sim}{\epsilon}_i$, $\underset{\sim}{e}_{(-i)} = [\underset{\sim}{I}_{k-1} \otimes (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)]\underset{\sim}{\epsilon}_{(-i)}$ and $\underset{\sim}{e} = [\underset{\sim}{I}_k \otimes (\underset{\sim}{I}_n - n^{-1}\underset{\sim}{J}_n)]\underset{\sim}{\epsilon}$. Then

$$\hat{\underset{\sim}{Y}}_{[-1]} = \underset{\sim}{X}_i\underset{\sim}{\beta} + \left\{\left[(k-1)^{-1}\underset{\sim}{1}_{k-1}' \otimes \underset{\sim}{P}_X\right] + \left[\left(\underset{\sim}{T}_{11}(\underset{\sim}{T}_{(-1)1}'\underset{\sim}{T}_{(-1)1})\underset{\sim}{T}_{(-1)1}'\right) \otimes \underset{\sim}{P}_1\right]\right.$$

$$\left. + \left[\left(\underset{\sim}{T}_{12}(\underset{\sim}{T}_{(-1)2}'\underset{\sim}{T}_{(-1)2})^{-1}\underset{\sim}{T}_{(-1)2}'\right) \otimes \underset{\sim}{P}_2\right]\right\}\underset{\sim}{e}_{(-1)} \quad ,$$

where $\underset{\sim}{I}_k$ is a $(k \times 1)$ vector of ones, $\underset{\sim}{P}_X = \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'$, and $\underset{\sim}{P}_j = \underset{\sim}{x}_j(\underset{\sim}{x}_j'\underset{\sim}{x}_j)^{-1}\underset{\sim}{x}_j'$, $j = 1, 2$. Note that $\underset{\sim}{T}_{1j}(\underset{\sim}{T}_{(-1)j}'\underset{\sim}{T}_{(-1)j})^{-1}\underset{\sim}{T}_{(-1)j}'$, $j = 1, 2$ are $[1 \times (k-1)]$ row vectors. Augmenting each by a zero in the <u>first</u> position, we can define the $(1 \times k)$ row vectors:

$$\underset{\sim}{t}_{11} = \left[0 \vdots \underset{\sim}{T}_{11}(\underset{\sim}{T}_{(-1)1}'\underset{\sim}{T}_{(-1)1})^{-1}\underset{\sim}{T}_{(-1)1}'\right] \text{ and } \underset{\sim}{t}_{12} = \left[0 \vdots \underset{\sim}{T}_{12}(\underset{\sim}{T}_{(-1)2}'\underset{\sim}{T}_{(-1)2})^{-1}\underset{\sim}{T}_{(-1)2}'\right] \quad .$$

Then

$$\hat{\underset{\sim}{Y}}_{[-1]} = \underset{\sim}{X}_i\underset{\sim}{\beta} + \left[(k-1)^{-1}\underset{\sim}{1}_{k-1}' \otimes \underset{\sim}{P}_X\right]\underset{\sim}{e}_{(-1)} + \left[\underset{\sim}{t}_{11} \otimes \underset{\sim}{P}_1 + \underset{\sim}{t}_{12} \otimes \underset{\sim}{P}_2\right]\underset{\sim}{e} \quad ,$$

and

$$\underset{\sim}{Y}_1 - \hat{\underset{\sim}{Y}}_{[-1]} = \left\{ \left[\underset{\sim}{I}_n \vdots -(k-1)^{-1} \underset{\sim}{1}_{k-1}' \otimes \underset{\sim}{P}_X \right] - \left[\underset{\sim}{t}_{11} \otimes \underset{\sim}{P}_1 + \underset{\sim}{t}_{12} \otimes \underset{\sim}{P}_2 \right] \right\} \underset{\sim}{\epsilon} \quad .$$

Next define $\underset{\sim}{t}_{i1}$ and $\underset{\sim}{t}_{i2}$, $i = 2, \cdots, k$, analogously to $\underset{\sim}{t}_{11}$ and $\underset{\sim}{t}_{12}$, i.e., $\underset{\sim}{t}_{21}$ is the $(1 \times k)$ row vector formed from $\underset{\sim}{T}_{21}(\underset{\sim}{T}_{(-2)1}'\underset{\sim}{T}_{(-2)1})^{-1}\underset{\sim}{T}_{(-2)1}'$ with a zero element inserted as the second element, etc. Then

$$\left[\left(\underset{\sim}{Y}_1 - \hat{\underset{\sim}{Y}}_{[-1]} \right)' \vdots \left(\underset{\sim}{Y}_2 - \hat{\underset{\sim}{Y}}_{[-2]} \right)' \vdots \cdots \vdots \left(\underset{\sim}{Y}_k - \hat{\underset{\sim}{Y}}_{[-k]} \right)' \right]$$
$$= \left[\underset{\sim}{A} - (\underset{\sim}{T}_1 \otimes \underset{\sim}{P}_1) + (\underset{\sim}{T}_2 \otimes \underset{\sim}{P}_2) \right] \left[\underset{\sim}{I}_k \otimes \left(\underset{\sim}{I}_n - n^{-1} \underset{\sim}{J}_n \right) \right] \underset{\sim}{\epsilon} \quad ,$$

where $\underset{\sim}{A} = \{ \underset{\sim}{I}_k \otimes [\underset{\sim}{I}_n + (k-1)^{-1} \underset{\sim}{P}_X] \} - [(k-1)^{-1} (\underset{\sim}{J}_k \otimes \underset{\sim}{P}_X)]$,

$$\underset{\sim}{T}_1 = [\underset{\sim}{t}_{11}' \vdots \underset{\sim}{t}_{21}' \vdots \cdots \vdots \underset{\sim}{t}_{k1}']' \quad \text{and} \quad \underset{\sim}{T}_2 = [\underset{\sim}{t}_{12}' \vdots \underset{\sim}{t}_{22}' \vdots \cdots \vdots \underset{\sim}{t}_{k2}']' \quad .$$

With the above representation and $\underset{\sim}{P}_X \underset{\sim}{1}_n = \underset{\sim}{0}$, $\underset{\sim}{T}_j \underset{\sim}{J}_k = \underset{\sim}{0}$, $\underset{\sim}{P}_X \underset{\sim}{P}_j = \underset{\sim}{P}_j$, $j = 1, 2$, and

$$\underset{\sim}{B} = \left[\underset{\sim}{I}_k \otimes (\underset{\sim}{I}_n - n^{-1} \underset{\sim}{J}_n) \right] \left[\underset{\sim}{I}_k \otimes (\underset{\sim}{I}_k \underset{\sim}{P}_X) \right] \left[\underset{\sim}{I}_k \otimes (\underset{\sim}{I}_n - n^{-1} \underset{\sim}{J}_n) \right] \quad ,$$

it can be shown that

$$\underset{\sim}{\epsilon}' (\underset{\sim}{R}' \underset{\sim}{R} - \underset{\sim}{B}) \underset{\sim}{\epsilon} = k^2 (k-1)^{-2} \underset{\sim}{\epsilon}' \left\{ \left[(\underset{\sim}{I}_k - k^{-1} \underset{\sim}{J}_k) \otimes \underset{\sim}{P}_X \right] + (\underset{\sim}{D}_1 \otimes \underset{\sim}{P}_1) + (\underset{\sim}{D}_2 \otimes \underset{\sim}{P}_2) \right\} \underset{\sim}{\epsilon} \quad ,$$

where

$$\underset{\sim}{D}_j = k^{-2} (k-1)^2 \left[\underset{\sim}{T}_i' \underset{\sim}{T}_i - k(k-1)^{-1} (\underset{\sim}{T}_i + \underset{\sim}{T}_i') + (k-1)^{-1} (\underset{\sim}{J}_k \underset{\sim}{T}_i + \underset{\sim}{T}_i' \underset{\sim}{J}_k) \right] , \quad j = 1, 2 \quad .$$

It now easily follows that $\underset{\sim}{B}_1 \underset{\sim}{B} = \underset{\sim}{0}$; i.e., the numerator and denominator of $k^{-2}(k-1)^2(P-1)$ are independent.

Finally we need to find the eigenvalues of

$$\underset{\sim}{B}_1 = \left[(\underset{\sim}{I}_k - k^{-1} \underset{\sim}{J}_k) \otimes \underset{\sim}{P}_X \right] + (\underset{\sim}{D}_1 \otimes \underset{\sim}{P}_1) + (\underset{\sim}{D}_2 \otimes \underset{\sim}{P}_2) \quad .$$

Define

$$\underset{\sim}{U}_0 = [\|\underset{\sim}{x}_1\|^{-2}\underset{\sim}{x}_1 \vdots \|\underset{\sim}{x}_2\|^{-2}\underset{\sim}{x}_2] \ ,$$

where $\|\underset{\sim}{x}_j\|^2 = \Sigma_{\ell=1}^{n} x_{j\ell}^2$. Let $\underset{\sim}{U}_1$ be any orthonormal matrix of rank p such that $\underset{\sim}{P}_X[\underset{\sim}{U}_0 \vdots \underset{\sim}{U}_1] = [\underset{\sim}{U}_0 \vdots \underset{\sim}{U}_1]$. Finally $\underset{\sim}{U}$ will be any orthonormal matrix of rank n with $\underset{\sim}{U} = [\underset{\sim}{U}_0 \vdots \underset{\sim}{U}_1 \vdots \underset{\sim}{U}_2]$ for some $\underset{\sim}{U}_2$. Then the eigenvalues of $\underset{\sim}{B}_1$ are the same as the eigenvalues of $(\underset{\sim}{I}_k \otimes \underset{\sim}{U})'\underset{\sim}{B}_1(\underset{\sim}{I}_k \otimes \underset{\sim}{U})$. By construction, $\underset{\sim}{U}'\underset{\sim}{P}_X\underset{\sim}{U}$ $= \text{diag}(\underset{\sim}{I}_p, \underset{\sim}{0}_{n-p})$, $\underset{\sim}{U}'\underset{\sim}{P}_1\underset{\sim}{U} = \text{diag}(1, \underset{\sim}{0}_{n-1})$, and $\underset{\sim}{U}'\underset{\sim}{P}_2\underset{\sim}{U} = \text{diag}(0,1,\underset{\sim}{0}_{n-2})$. Therefore,

$$\begin{aligned}(\underset{\sim}{I}_k \otimes \underset{\sim}{U})'\underset{\sim}{B}_1(\underset{\sim}{I}_k \otimes \underset{\sim}{U}) &= (\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k) \otimes \text{diag}(0,0,\underset{\sim}{I}_{p-2},\underset{\sim}{0}_{n-p}) \\ &+ \left[(\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k) + \underset{\sim}{D}_1\right] \otimes \text{diag}(1,\underset{\sim}{0}_{n-1}) \\ &+ \left[(\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k) + \underset{\sim}{D}_2\right] \otimes \text{diag}(0,1,\underset{\sim}{0}_{n-2}) \ .\end{aligned} \tag{2}$$

Since the above three matrices are orthogonal, the required eigenvalues of $\underset{\sim}{B}_1$, $\theta_1, \cdots, \theta_q$, are the $(k-1)(p-2)$ eigenvalues of the first matrix which are identically equal to one, the k eigenvalues of $\underset{\sim}{C}_1 = (\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k) + \underset{\sim}{D}_1$ and the k eigenvalues of $\underset{\sim}{C}_2 = (\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k) + \underset{\sim}{D}_2$. This is the result stated at the end of Section 2.

Finally, the effect of unequal experimental error variances across sites must be investigated. Let σ_i^2, $i = 1, \cdots, k$, denote the variances of the yields of the ith site; i.e., $E[\underset{\sim\sim}{\epsilon\epsilon}'] = \text{diag}(\sigma_i^2 \underset{\sim}{I}_n) = \underset{\sim}{D} \otimes \underset{\sim}{I}_n$, where $\underset{\sim}{D} = \text{diag}(\sigma_1^2, \cdots, \sigma_k^2)$. It follows that

$$\frac{(k-1)^2}{k^2} \cdot (P-1) = \frac{\left[(\underset{\sim}{D}^{-\frac{1}{2}} \otimes \underset{\sim}{I}_n)\underset{\sim}{\epsilon}\right]'\left[(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n)\underset{\sim}{B}_1(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n)\right]\left[(\underset{\sim}{D}^{-\frac{1}{2}} \otimes \underset{\sim}{I}_n)\underset{\sim}{\epsilon}\right]}{\left[(\underset{\sim}{D}^{-\frac{1}{2}} \otimes \underset{\sim}{I}_n)\ \underset{\sim}{\epsilon}\right]'\left[(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n)\underset{\sim}{B}(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n)\right]\left[(\underset{\sim}{D}^{-\frac{1}{2}} \otimes \underset{\sim}{I}_n)\underset{\sim}{\epsilon}\right]}$$

where $(\underset{\sim}{D}^{-\frac{1}{2}} \otimes \underset{\sim}{I}_n)\underset{\sim}{\epsilon}$ is a vector of independent standard normal random variables. Note that $\underset{\sim}{B}(\underset{\sim}{D} \otimes \underset{\sim}{I}_n) = \underset{\sim}{B}(\underset{\sim}{D} \otimes \underset{\sim}{I}_n)\underset{\sim}{B}$, $(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n)\underset{\sim\sim}{B}\underset{\sim}{D}\underset{\sim}{B}_1(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n) = \underset{\sim}{0}$, and the numerator and denominator remain independent. Similarly, $(\underset{\sim}{I}_k \otimes \underset{\sim}{U})$ and $(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n)$ also commute and the eigenvalues of the numerator are the eigenvalues of

$$
\begin{aligned}
&(\underset{\sim}{I}_k \otimes \underset{\sim}{U})(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n)\underset{\sim}{B}_1(\underset{\sim}{D}^{\frac{1}{2}} \otimes \underset{\sim}{I}_n)(\underset{\sim}{I}_k \otimes \underset{\sim}{U}) \\
&\quad = \underset{\sim}{D}^{\frac{1}{2}}(\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k)\underset{\sim}{D}^{\frac{1}{2}} \otimes \operatorname{diag}\left(0, 0, \underset{\sim}{I}_{(p-2)}, \underset{\sim}{0}_{n-p}\right) \\
&\qquad + \underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{C}_1\underset{\sim}{D}^{\frac{1}{2}} \otimes \operatorname{diag}(1, \underset{\sim}{0}_{n-1}) + \underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{C}_2\underset{\sim}{D}^{\frac{1}{2}} \otimes \operatorname{diag}(0, 1, \underset{\sim}{0}_{n-2}) \quad .
\end{aligned}
$$

Thus the eigenvalues of the numerator consist of the eigenvalues of $\underset{\sim}{D}^{\frac{1}{2}}(\underset{\sim}{I}_k - k^{-1}\underset{\sim}{J}_k)\underset{\sim}{D}^{\frac{1}{2}}$ with multiplicity $(p-2)$ and the eigenvalues of both $\underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{C}_1\underset{\sim}{D}^{\frac{1}{2}}$ and $\underset{\sim}{D}^{\frac{1}{2}}\underset{\sim}{C}_2\underset{\sim}{D}^{\frac{1}{2}}$, which may be computed as in the equal variance case.

Also, for small n and k the eigenvalues of the denominator may be similarly calculated. In practice, however, the degrees of freedom of sum of squared residuals for each site will be large enough to assume that the resulting consistent within-site estimated error variance sufficiently approximates σ_i^2 and these eigenvalues will not be required. In particular,

$$
k(n-p)\left[\frac{(k-1)^2}{k^2}\ (P-1)\right] \xrightarrow{p} \frac{\underset{\sim}{\epsilon}'\underset{\sim}{B}_1\underset{\sim}{\epsilon}}{\Sigma_i \sigma_i^2}
$$

and the problem reduces to computing the eigenvalues of the numerator.