

Design and Analysis of Commonly Used Comparative Horticultural Experiments

George C.J. Fernandez¹

College of Agriculture, Biotechnology and Natural Resources, University of Nevada–Reno, Reno, NV 89557

A comparative horticultural experiment is a procedure for collecting scientific data in a systematic way to maximize the chance of testing a research hypothesis correctly. It is very important that comparative horticultural experiments are well designed, correctly analyzed, and reported accurately to achieve the full potential of the research. Horticulturists are strongly encouraged to review standard experimental designs statistics textbooks (Kuehl, 1996; Littell et al., 2006), online statistical resources (Government of Alberta 2006; Schabenberger 2006; Sit, 1995), and several published articles on advanced statistical methods related to designed experiments (Chew, 1976; Jones, 1984; Petersen, 1977; Piepho et al., 2003; Saville, 1990; Waldo, 1976) before designing their comparative studies and analyzing their data. Also, it is wise always to consult with experimental station statisticians 1) before designing comparative studies, 2) during the analysis, and 3) when interpreting the results. It is not my goal to discuss in this paper an extensive description of theory (this is beyond the scope of this paper), flawed and proper examples, or a description of statistical methods applicable to all comparative horticultural experiments. The aim of this paper is to document the checklists and statistical guidelines applicable to commonly used horticultural experiments, which I presented during a colloquium at the 2004 American Society for Horticultural Science conference in Austin, TX. Download instructions are given in this paper to obtain a hard-copy of this PowerPoint presentation, which contains additional information on sample SAS codes for analyzing data from comparative horticultural experiments discussed in this paper. By reviewing these recommended statistical texts, online resources, and journal papers, and by following the guidelines presented herein, horticulturists may ensure that their research publications conform to the very highest scientific standards with respect to experimental design and statistical methods.

STATISTICAL ISSUES RELATED TO DESIGNING HORTICULTURAL EXPERIMENTS

Designed experiments are considered randomized, controlled comparative studies containing adequate numbers of true replications. A well-designed experiment will avoid bias and will be sufficiently powerful to

detect treatment effects likely to be of biological significance. Adequate number of replications is necessary to estimate the experimental error and to have enough power to detect true treatment effects. Proper randomization ensures unbiased estimation of treatment effects and experimental errors. Four types of designed experiments are common in horticultural research: 1) field experiments, 2) greenhouse experiments, 3) growth chambers studies, and 4) laboratory experiments. The characteristics of these different experiments and the important statistical issues related to statistical designs are described next.

Field experiments. Field studies are suitable for conducting yield trials, genotype screening, and comparing horticultural management practices, such as irrigation, nutrition, pest, disease, weed control, and so on. The statistical issues related to designing field experiments are noted in the following paragraphs.

Multiyear studies are important to investigate the interaction between the treatment and the year to remove the confounding effects of climatic variation. The year effects are usually treated as “random effect” in the subsequent analysis.

Multisite studies in a given year are important to evaluate the treatment effects across different sites and to investigate the interaction between the treatment and the site, and to remove the confounding effects of soil variation. If the field investigation is aimed at testing specific site effects and making inferences only about these specific sites, then the site effect should be treated as “fixed effect” in the subsequent analysis. The site effects should be treated as “random effect” when the selected sites represent a random selection of potential sites and the treatment comparison is aimed at making a nonsite-specific overall treatment comparison.

Individual homogeneous field plots with more than single plants are commonly used as experimental units in annual and perennial crop studies. Blocking across the field variability gradient or measuring additional continuous covariates from each experimental unit and using analysis of covariance (Fernandez, 1990) are highly recommended in reducing experimental errors in designed comparative experiments. In tree crop studies, individual trees are usually treated as the experimental units. In both cases, the treatments should be randomly assigned to the experimental units and each experimental unit should have an equal chance of receiving one of the treatments. All response measurements, such as yield, height, biomass, and other horticultural attributes should be

recorded for each experimental unit and used in the analysis. Recording response measurements from subsamples within each experimental unit adds additional source of subsample variation and provides a means to estimate both experimental and subsample error that complicate the data analysis of horticultural experiments with subsamples. Therefore, when there is no missing value in the data, multiple subsample measurements taken within an experimental unit (individual plots or trees) could be averaged to simplify the analysis. However, to estimate both the experimental error variation and the subsample error variation, and to use the proper error term to test the treatment effects, a mixed model analysis should be used.

Repeated response measurements are usually taken from the same experimental unit in tree and perennial crops in multiple years or from multiple harvests within the same year and when both are in consideration (e.g., with multiple year trial with many harvests each year). Because these repeated measurements are recorded within the same experimental units, these measurements are considered nonindependent and underestimate the experiment errors. Therefore, repeated-measures analysis of variance (ANOVA) (Fernandez 1991; Littell et al., 2006) should be used when analyzing the horticultural experiments involving repeated measurements to test the significance of the repeated-measures factor and any interaction effects associated with the repeated-measures factor.

Greenhouse or glass house experiments. Greenhouse experiments for treatment comparisons are usually conducted under managed environments using artificial growth media and controlled irrigation and atmospheric conditions to minimize soil and climatic variation. The commonly used experimental units are single or multiple plants grown in individual pots or trays. Although greenhouse experiments are conducted under managed or controlled environments, air temperature, humidity, solar radiation, and moisture variation may occur across the greenhouse as well as within the greenhouse benches. To control the experimental error resulting from the environmental variations, suitable blocking and the use of additional covariates must be used. Furthermore, if the experimental units are pots or planting trays, these could be randomly rotated within a bench during the experimental period to minimize any systematic variation and to reduce edge effect in the greenhouse. The statistical issues related to analyzing greenhouse experimental data are noted in the following paragraphs.

¹To whom reprint requests should be addressed; e-mail gcejf@unr.edu

Multiyear or multisite studies are uncommon because both climatic and soil variations are controlled here. Main emphasis is given here to compare the main effects and interaction among treatment combinations.

Individual homogeneous pots or trays with more than single plants are commonly used as experimental units in greenhouse studies. The treatments are randomly assigned to the experimental units (individual pots or trays with more than single plants), and each experimental unit should have an equal chance of receiving one of the treatment combinations. Recording response measurements from subsamples (different plant parts within a single plant or pot) within each experimental unit is also very common and adds additional source of subsample variation. Therefore, multiple subsample measurements taken within an experimental unit (individual pots) should be averaged when there are no missing values to simplify the analysis or a mixed model analysis should be used to separate the experimental error variation and the subsample error variation.

Repeated response measurements are also very common in greenhouse studies, and repeated-measures ANOVA should be used when analyzing the horticultural experiments involving repeated measurements.

Growth chamber experiments. Growth chamber studies are special controlled environment studies in which the researcher can study the temperature or light effects on plant growth by controlling specific environmental factors at a given level throughout the experimental study. The lack of or inadequate number of replications is the main problem in growth chamber studies. Because the treatment condition (fixed temperature or light intensity/duration) are randomly applied to the growth chamber and not to the pots or individual plant, the growth chamber unit becomes the experimental unit. The pots, Petri dishes, trays, and plants included within a growth chamber are considered subsamples and should not be treated as true replicates. To provide valid replications when the number of growth chambers are limited, the experiments should be repeated by using the same growth chamber with different randomly assigned treatment levels multiple times to collect statistically valid experimental data.

Laboratory experiments. Horticultural laboratory experiments can be classified into two major categories. The first is studies involving chemical or nutritional analysis of plant samples obtained from other field or greenhouse studies. These are extensions of the field or greenhouse studies, and therefore the experimental design and same sampling plans should be used in the statistical analysis of the laboratory investigations. The second category includes stand-alone experiments using horticultural materials (e.g., seeds, tissue culture propagules, cellular components such as proteins, fats, carbohydrates, hormones, postharvest plants, and plant products) handled under laboratory conditions.

Thus, the recommended statistical principles in conducting valid experiments such as randomization, adequate replication, and error control measures should be used in designing these laboratory experiments involving plant and their parts.

STATISTICAL ISSUES RELATED TO ANALYZING DATA FROM HORTICULTURAL EXPERIMENTS

The data collected from comparative experiments should be evaluated using the appropriate statistical analysis. The aim of the proper statistical analysis is to extract all valid information present in the data in such a way that it can be interpreted, taking into account the inherent variability and measurement errors associated with the experiment. The choice of statistical analysis will depend on the purpose of the study, the design of the experiment, and the nature of the response variable.

Designed experiments conducted by horticultural scientists often give rise to several random sources of variation. Relevant examples are split plot designs, multiyear and multisite yield trials, and repeated measurements taken on the same field plot. Unbalanced data are a common problem in plant research based on designed experiments. The analysis of repeated measurements needs to account for the correlation among observations on the same experimental unit. There are many different types of models that can be imposed on the correlation structure, including autoregressive and spatial models. These correlation structures can be embedded in a mixed model framework. Data arising from such experiments may be conveniently analyzed by mixed models.

A fixed effect is considered repeatable (i.e., the same levels of the factor could be repeated by another scientist), and scientists are only interested in making treatment comparisons about the specific levels used in that experiment. A random factor is an effect for which the levels of the factor are considered to be a random sample from an infinite set of possible levels. When the levels of a treatment cannot be exactly replicated by another researcher, then that treatment should be thought of as a random effect. However, even treatments that are repeatable may be considered random effects if they can be regarded as a random sample from some larger population of treatment levels. With random effects, scientists are only interested in estimating the population variances related to randomly selected factor levels. Also, inferences about the random effect are not concerns about treatment mean comparisons (as they are with fixed effects), but concerns about the estimation of the variance covariances of the random effects.

A split plot experiment requires two random error terms for main plots and subplots. Repeated measurements taken at different points in time or space are correlated, which may be accounted for by a mixed

model with an appropriate variance-covariance structure. Experiments replicated at several sites or in several years require a linear model with random site or random year effects, providing environments may be considered as a random factor and the objective is to compute treatment means across a wide range of environments.

The mixed procedure of SAS software version 9.1 of the SAS System for Windows (Littell et al., 2006), allows for flexible inclusion of a combination of continuous and categorical random effects as well as nested and hierarchical structures. This will provide restricted maximum likelihood (REML) estimates of variance and covariance components among model factors, and permits fitting both fixed and random model effects in mixed models analyses (Littell et al., 2006).

The statistical guidelines and checklists for analyzing data obtained from commonly used comparative designed experiments involving a continuous response variable are presented in Tables 1 through 4. The choice of the correct statistical analysis depends on the scale of the response variable, type of experimental design [complete randomized design (CRD), randomized complete-block design (RCBD), split plot, and so on], treatment design (single factor, factorial design, fixed vs. random factor, qualitative vs. quantitative factors, independent vs. correlated or repeated factor levels, presence of subsamples, and so on). Therefore, horticulturists are encouraged to review the checklists and statistical analysis guidelines presented in Tables 1 through 4 before analyzing their data and to select the correct analysis plans, including the appropriate mean comparison methods. This exercise will also help them to describe the experimental design used and statistical methods used effectively in the materials and methods section. For additional information regarding running these statistical analyses using SAS software and to see the sample SAS code, please visit my home page at www.cabnr.unr.edu/gf and download a copy of the PowerPoint presentation by clicking the running puppy clip art and following the online instructions.

CHECKING FOR VIOLATIONS OF ANOVA ASSUMPTIONS

Checking for any violations of statistical assumptions such as 1) nonindependence of error, 2) unequal error variation within treatment, and 3) nonnormal error distribution is an important component of statistical analysis. Proper randomization methods in experimental designs ensure that the residuals are independent. However, when randomization is not possible with repeated-measures treatment, experimental errors become correlated, resulting in the underestimation of experimental errors. Therefore, repeated-measures ANOVA using the mixed model methodology should be used to correct this problem. Constant or equal variance assumption can be assessed graphically by plotting the residuals

Table 1. Checklist and guidelines for analyzing data from designed experimental studies: step 1.

Treatment Design	
Single-factor experiments	Multiple-factor experiments
a) Experiment design: Complete randomized design (CRD) or randomized complete-block design (RCBD) _____	_____
b) Input name of the experimental unit (e.g., plot, tree, pot, tray) _____	_____
c) Input the response variable (e.g., yield, height, biomass) _____	_____
Input treatment name (e.g., variety, nitrogen) _____	Input factor name(s) (e.g., variety, nitrogen) _____
Factor type: Qualitative (QL) or Quantitative (QT) _____	Input factor types and number of levels: _____
Number of treatment levels (t) = _____	Factor A: QL or QT _____
Number of replications (r) in case of CRD or number of blocks in case of RCBD _____	Number of levels (m) _____
Total number of observations (n) _____	Factor B: QL or QT _____
_____	Number of levels (n) _____
_____	Factor C: QL or QT _____
_____	Number of levels (o) _____
_____	Total number of treatment combinations: _____
_____	($m \times n$) _____ or ($m \times n \times o$) _____
_____	Number of replications (r) in case of CRD or number of blocks in case of RCBD per treatment combinations _____
_____	Total number of observations (n) _____
Go to Table 2 for data analysis checklists and guidelines.	Go to Table 3 for data analysis checklists and guidelines.
_____	Go to Table 4 for data analysis checklists and guidelines.

CRD, complete randomized design; m, total number of levels; n, number of observations; QL, qualitative; QT, quantitative; r, replications; RCBD, randomized complete-block design, t, treatment levels.

against the predicted values (Fernandez, 1992). The ANOVA F test, treatment standard errors, and mean comparison results can be misleading when the equal variance assumption is not met. A natural log transformation [$Y_t = \text{Log}(Y + 0.1)$] is often appropriate when the residual distribution is significantly skewed (Kuehl 1996). When the response measurements are expressed as proportions or percentages, such as the proportion of seed germination, and many percentage values fall below 20% or above 80%, an arcsin transformation [$Y_t = \arcsin[\text{SQRT}(p)]$] is recommended where the proportion (p) can take any values between 0 and 1, without including 0 or 1 (Kuehl 1996). Response measurements involving small counts (insect count/plot, disease lesions, number of seed/pot) usually have a Poisson distribution, and a square root transformation [$Y_t = \text{SQRT}(Y + 0.5)$] is usually recommended (Kuehl, 1996) for count data. Mixed models have also been extended to allow for nonnormal data. The extension is known as a generalized linear mixed model (Littell et al., 2006; Schabenberger and Pierce, 2001), and it is highly relevant for problem data such as count data (e.g., weeds and insects) and percentages (e.g., disease incidence, weed coverage, and emergence rates) (Piepho, 1999). Moreover, nonlinear mixed models can accommodate intrinsically nonlinear regression models (e.g., plant growth models) (Schabenberger and Pierce, 2001). Categorical or qualitative responses need to be analyzed using categorical data analysis methods such as χ^2 test, logistic regression for binary response, and so forth. Spatial analysis is recommended for spatially correlated data by using special correlation structures within the mixed model framework (Schabenberger, 2004).

Analysis of variance methods are normally robust against minor departures from the normally distributed residuals. The residual error term is assumed to be both homoscedastic and independent. Graphical analysis including frequency histograms and normal probability plots of the residuals can be used to assess normality visually. Evidence for the violation of normality assumptions can be assessed graphically with quantile-quantile plots in which a straight line indicates normality (Fernandez, 2002). Presence of significant outliers can be responsible for the normality assumption violation, and therefore before taking any remedial measures, the data should be thoroughly examined for the presence of significant outliers (Fernandez, 1992, 2002). Nonparametric statistical methods are recommended for data with severe nonnormality issues. Consult a statistician for the proper selection and use of any advanced statistical methods reported here.

SUMMARY

Horticulturists are encouraged to review experimental designs textbooks, online statistical resources, and published articles on

Table 2. (Continued) Checklists and guidelines for analyzing data from single-factor (QL or QT)-designed experiments in complete randomized design (CRD) or randomized complete-block design (RCBD): step 2.

Experimental design	Treatment and experiment designs		Recommended analysis		
	Treatment levels are independent (IND) or correlated (COR; repeated measures)	Presence of subsamples when $n > (t \times r)$ or no subsamples when $n \leq (t \times r)$	Presence of covariates (yes) or no covariates (no)	Analysis type	
		With covariates	<p>Fixed blocks: fixed two-way ANCOVA (assumption: no block \times treatment interaction) in RCBD, use SAS PROC GLM</p> <p>Random blocks: mixed two-way ANCOVA (no block \times treatment interaction) in RCBD, use SAS PROC MIXED with RANDOM statement, estimate block variance</p> <p>*Verify first that treatment \times covariate interaction is not significant</p>	<p>Mean computation, qualitative factor (QL): trend analysis, qualitative factor (QT) treatment: LSMEANS at the mean level of the covariate (SAS LSMEAN option), or use ESTIMATE statement</p> <p>QT treatment: Fit covariate-adjusted regression model and test for linear, quadratic effects using type ISS</p>	Mean comparison methods, QL factor; trend analysis for QT factor
	With subsamples	No covariates	<p>Mixed two-way ANOVA: In RCBD with two random errors (experimental and subsample errors), use SAS PROC MIXED with RANDOM statement, estimate block and block \times treatment error variance</p> <p>*Verify first that treatment \times covariate interaction is not significant</p>	<p>QL treatment: LSMEANS at the mean level of the covariate (SAS LSMEAN option), or use ESTIMATE statement</p> <p>QT treatment: Fit covariate adjusted regression model and test for linear, quadratic effects using type ISS</p>	Use block-adjusted LSMEANS with and without Tukey or Bonferroni adjustment, or use preplanned group comparisons using (CONTRAST or ESTIMATE statements
	With covariates		<p>Mixed two-way ANCOVA: In RCBD with two random errors (experimental and subsample errors), use SAS PROC MIXED with RANDOM statement, estimate covariate-adjusted block and block \times treatment error variance</p> <p>*Verify first that treatment \times covariate interaction is not significant</p>	<p>LSMEANS at the mean level of the covariate (SAS LSMEAN option) or use ESTIMATE statement</p>	

(Continued on next page)

Table 2. (Continued) Checklists and guidelines for analyzing data from single-factor (QL or QT)-designed experiments in complete randomized design (CRD) or randomized complete-block design (RCBD): step 2.

Treatment and experiment designs		Recommended analysis	
Experimental design	Treatment levels are independent (IND) or correlated (COR; repeated measures)	Presence of covariates (yes) or no covariates (no)	Mean computation, qualitative factor (QL); trend analysis, qualitative factor (QT)
COR	Presence of subsamples when $n > (t \times r)$ or no subsamples when $n \leq (t \times r)$	No covariates	Mean comparison methods, QL factor; trend analysis for QT factor
	No subsamples	No covariates	QL treatment: Use block-adjusted LSMEANS with and without Tukey or Bonferroni adjustment, or use preplanned group comparison (CONTRAST or ESTIMATE)
		Analysis type	QT treatment: Estimate linear, quadratic parameters and report the regression model
		Fixed blocks: fixed two-way repeated measures ANOVA (assumption: no block \times treatment interaction) in RCBD, use SAS PROC MIXED with REPEATED statement (select best covariance structure first)	QL treatment: LSMEAN option, or use ESTIMATE statement
		Random blocks: mixed two-way repeated measures ANOVA (no block \times treatment interaction) in RCBD, use SAS PROC MIXED with RANDOM (estimate block variance) and REPEATED statements (to select best covariance structure)	QT treatment: Fit regression model and test for linear, quadratic effects using type ISS
		With covariates	QL treatment: LSMEANS <i>at the mean level of the covariate</i> (SAS LSMEAN option), or use ESTIMATE statement
		Fixed blocks: fixed two-way repeated measures ANCOVA (assumption: no block \times treatment interaction) in RCBD, use SAS PROC MIXED with REPEATED (select best covariance structure)	QT treatment: Fit covariate adjusted regression model and test for linear, quadratic effects using type ISS
		Random blocks: mixed two-way repeated measures ANCOVA (no block \times treatment interaction) in RCBD, use SAS PROC MIXED with RANDOM (estimate block variance) and REPEATED statements (to select best covariance structure)	
		*Verify first that treatment \times covariate interaction is not significant	

(Continued on next page)

Table 3. Checklists and guidelines for analyzing data from multifixed-factor full-factorial experiments in CRD or RCBD: step 2.

Treatment and experiment designs		Recommended analysis	
Experimental design	Treatment levels are independent (IND) or correlated (COR); repeated measures)	Subsamples or no subsamples when $n > (t \times r)$ or 2) no subsamples when $n \leq (t \times r)$	Mean computation, all factors are qualitative (QL), LSMEAN computation or at least one factor; quantitative (QT), trend analysis by levels of the qualitative factor
	Complete randomized design (CRD)	No subsamples	1) Two and three factors are QL Two or three QL factor interaction <i>significant</i> : Use LSMEAN comparisons with or without the <i>P</i> value adjustment for the significant interaction terms. Significant $A \times B$ interaction: Compare between <i>A</i> levels at a given level of <i>B</i> and vice versa Significant $A \times B \times C$ interaction: Compare between <i>A</i> levels at a given level of <i>B</i> and vice versa
Complete randomized design (CRD)	Presence of covariates (yes) or no covariates (no)	No covariates, all experimental units are assumed to be homogeneous	QL \times QL or (QL \times QL \times QL): Estimate LSMEANS using SAS LSMEAN option, or use SAS ESTIMATE statement. One QT factor: Fit regression model and test for significant linear, quadratic main effects and the interactions between QL factor levels and the linear and quadratic effects using type 1 sequential SS QL \times QT factor interaction significant: Fit regression models with <i>different slopes</i> for the various levels of QL factor
	With covariates, experimental units are assumed to be <i>not</i> homogeneous	Two factors: Fixed-effects two-way ANCOVA in CRD, use SAS PROC GLM Three factors: Fixed-effects three-way ANOVA in CRD, use SAS PROC GLM *Verify first that factors \times covariate interactions are not significant Two factors: Mixed-effects two-way ANOVA with subsamples in CRD, use SAS PROC MIXED, RANDOM statement	2) One QT factor Factor <i>A</i> is QL and factor <i>B</i> is QT and $A \times B$ interaction is <i>significant</i> : Fit separate regression models for factor <i>B</i> and response for <i>each level of factor A</i> ; estimate linear, quadratic parameters and report the regression model Treatment means: LSD, Tukey, or Bonferroni comparisons for <i>A</i> means at more than one specified level of factor <i>B</i>
Complete randomized design (CRD)	With covariates, experimental units are assumed to be <i>not</i> homogeneous -	With covariates, all experimental units are assumed to be homogeneous	QL \times QT factor interaction <i>not</i> significant: Fit regression models with <i>same slopes</i> for the various levels of QL factor In the presence of covariate, estimate LSMEANS at the mean level of the covariate Two or more QT factors: Perform response surface regression using SAS PROC RSREG
	With covariates, experimental units are assumed to be <i>not</i> homogeneous -	Two factors: Mixed-effects two-way ANCOVA with subsamples in CRD, use SAS PROC MIXED, RANDOM statement	

(Continued on next page)

Table 3. (Continued) Checklists and guidelines for analyzing data from multifactorial experiments in CRD or RCBD: step 2.

Treatment and experiment designs		Recommended analysis	
Treatment levels are independent (IND) or correlated (COR); repeated measures)	Subsamples or no subsamples when $n > (t \times r)$ or 2) no subsamples when $n \leq (t \times r)$	Presence of covariates (yes) or no covariates (no)	Analysis type
Experimental design			
Randomized Complete Block Design (RCBD)	No subsamples	No covariates	Fixed blocks: Fixed two- or three-way ANOVA in RCBD, use SAS PROC GLM Random blocks: Mixed two-way ANOVA in RCBD, use SAS PROC MIXED with RANDOM statement, estimate block variance
	With covariates		Fixed blocks: Fixed two- or three-way ANCOVA in RCBD, use SAS PROC GLM Random blocks: Mixed two- or three-way ANCOVA in RCBD, use SAS PROC MIXED with RANDOM statement, estimate block variance *Verify first that factors \times covariate interactions are not significant Mixed two- or three-way ANOVA in RCBD with two random errors (experimental and subsample errors)
	With subsamples	No covariates	Fixed blocks: Use SAS PROC MIXED with RANDOM statement to estimate experimental error block \times (A \times B) or block (A \times B \times C) experimental error variance Random blocks: Additionally estimate the random block effects using PROC MIXED RANDOM statement
			Mean computation, all factors are quantitative (QL), LSMEAN computation or at least one factor; quantitative (QT), trend analysis by levels of the qualitative factor (QL \times QL) or (QL \times QL \times QL): Estimate LSMEANS using SAS LSMEAN option, or use SAS ESTIMATE statement One QT factor: Fit regression model and test for significant linear, quadratic main effects and the interactions between QL factor levels and the linear and quadratic effects using type 1 sequential SS QL \times QT factor: interaction significant: Fit regression models with <i>different slopes</i> for the various levels of QL factor QL \times QT factor: interaction <i>not</i> significant: Fit regression models with <i>same slopes</i> for the various levels of QL factor. *In the presence of covariate, estimate LSMEANS at the <i>mean level of the covariate</i> Two or more factors quantitative: Perform response surface regression using SAS PROC RSREG.

(Continued on next page)

Table 3. (Continued) Checklists and guidelines for analyzing data from multifixed-factor full-factorial experiments in CRD or RCBD: step 2.

Treatment and experiment designs		Recommended analysis	
Treatment levels are independent (IND) or correlated (COR); repeated measures	Subsamples or no subsamples when $n > (t \times r)$ or 2) no subsamples when $n \leq (t \times r)$	Presence of covariates (yes) or no covariates (no)	Analysis type
Experimental design	With covariates	With covariates	Mixed two- or three-way ANCOVA in RCBD with two random errors (experimental and subsample errors), use SAS PROC MIXED with RANDOM statement to estimate block \times (A \times B) or block (A \times B \times C), the experimental error variance
		No covariates	Random blocks: Estimate the random block variance using PROC MIXED RANDOM statement
COR	No subsamples	No covariates	Fixed blocks: Fixed two or three-way repeated measures ANOVA in RCBD, use SAS PROC MIXED with REPEATED statement
		With covariates	Random blocks: Mixed two- or three-way repeated measures ANOVA in RCBD, use SAS PROC MIXED with RANDOM and REPEATED statements (model the repeated measures correlation by selecting the best covariance structure using the smallest AICC)
			One QT factor: Fit regression model and test for significant linear, quadratic main effects and the interactions between QL factor levels and the linear and quadratic effects using type 1 sequential tests in PROC MIXED
			QL \times QT factor interaction significant: Fit regression models with <i>different slopes</i> for the various levels of QL factor
			QL \times QT factor interaction <i>not</i> significant: Fit regression models with <i>same slopes</i> for the various levels of QL factor.
			*In the presence of covariate, estimate LSMEANS at the <i>mean level of the covariate</i>
			Two or more QT factors: Perform response surface regression using SAS PROC MIXED
			*Verify first that factors \times covariate interactions are not significant
			(QL \times QL) or (QL \times QL \times QL): Estimate LSMEANS using SAS LSMEAN option, or use SAS ESTIMATE statement
			Mean computation, all factors are qualitative (QL), LSMEAN computation or at least one factor; quantitative (QT), trend analysis by levels of the qualitative factor
			Mean comparison methods: QL factors or trend analysis; for QT factors, fitting linear, quadratic regression models

(Continued on next page)

Table 3. (Continued) Checklists and guidelines for analyzing data from multifixed-factor full-factorial experiments in CRD or RCBD: step 2.

Treatment and experiment designs		Recommended analysis	
Experimental design	Treatment levels are independent (IND) or correlated (COR; repeated measures)	Presence of covariates (yes) or no covariates (no)	Analysis type
	Subsamples or no subsamples: 1) subsamples when $n > (t \times r)$ or 2) no subsamples when $n \leq (t \times r)$	No covariates	Fixed or random blocks: mixed two- or three-way repeated-measures ANOVA in RCBD with two random errors (experimental and subsample errors)
	With subsamples		Random blocks: Additionally, estimate the random block variance using the PROC MIXED RANDOM statement
			Use SAS PROC MIXED with RANDOM statement to estimate block \times (A \times B) or block (A \times B \times C), the experimental error variance
			*Model the repeated measures correlation using PROC MIXED REPEATED statement by selecting the best covariance structure based on the smallest AICC
		With covariates	Fixed or random blocks: mixed two- or three-way repeated-measures ANCOVA in RCBD and with two random effects (experimental and subsample errors)
			Random blocks: Additionally, estimate the random block variance using PROC MIXED RANDOM statement
			Use SAS PROC MIXED with RANDOM statement to estimate block \times (A \times B) or block (A \times B \times C), the experimental error variance
			*Model the repeated measures correlation using PROC MIXED REPEATED statement by selecting the best covariance structure based on the smallest AICC
			*Verify first that factors \times covariate interactions are not significant
			Mean computation, all factors are qualitative (QL), LSMEAN computation or at least one factor; quantitative (QT), trend analysis by levels of the qualitative factor
			Mean comparison methods: QL factors or trend analysis; for QT factors, fitting linear, quadratic regression models

AICC, Akaike Information Criterion; ANCOVA, analysis of covariance; ANOVA, analysis of variance; COR, correlated; CRD, complete randomized design; IND, independent; n, total number of observations; QL, qualitative; QT, quantitative; r, replications; RCBD, randomized complete-block design; t, treatment levels.

Table 4. Checklists and guidelines for analyzing data from split plot or split-split plot experiments in CRD or RCBD: step 2.

Treatment and experiment designs		Recommended analysis	
Split plot or split-split plot design	Subsamples or no subsamples: 1) subsamples when $n > (t \times r)$ or 2) no subsamples when $n \leq (t \times r)$	Mean computation, all factors qualitative (QL) or at least one factor quantitative (QT): trend analysis by levels of the qualitative factor	Mean comparison methods, all factors QL or trend analysis (for at least one factor is QT: fitting linear, quadratic regression models)
	Main plot design	Presence of covariates (yes) or no covariates (no)	
Split plot	CRD	Analysis type	1) Both main plot and split plot factors are QL
	No covariates, all experimental units are assumed to be homogeneous	Mixed effects split plot ANOVA in CRD, use SAS PROC MIXED with RANDOM statement to separate main plot error from split plot error	Interaction <i>significant</i> : Use LSMEAN comparisons with or without the P value adjustment for the significant interaction terms or use PROC MIXED ESTIMATE or CONTRASTS to estimate
	With covariates, main plot or split plot experimental units are assumed to be <i>not</i> homogeneous	Mixed-effects split plot ANCOVA in CRD, use SAS PROC MIXED with RANDOM statement to separate main plot error from split plot error	Significant $A \times B$ interaction: Compare between main plot (A) levels at a given level of split plot (B) and vice versa
		*Verify first that main plot or the split plot factor \times covariate interactions are not significant	Nonsignificant $A \times B$ interaction, but main effect significant: Use LSMEAN comparisons with or without the P value adjustment for the significant main plot and/or split plot levels, or use group comparison using preplanned CONTRAST or ESTIMATES
	With subsamples	Mixed effects split plot ANOVA with subsamples in CRD, use SAS PROC MIXED with separate main plot error, split plot error, and subsample errors	2) Either main plot or split plot factor is QT
	No covariate, all main plot experimental units are assumed to be homogeneous	Mixed effects split plot ANOVA with subsamples in CRD, use SAS PROC MIXED with separate main plot error, split plot error, and subsample errors	QL \times QT factor interaction is <i>significant</i> : Fit separate regression models for QT factor and response for <i>each level of QL factor</i> ; estimate linear, quadratic parameters and report the regression models for each QL levels
	With covariates, main plot or split plot experimental units are assumed to be <i>not</i> homogeneous	Mixed-effects split plot ANCOVA with subsamples in CRD, use SAS PROC MIXED with separate main plot error, split plot error, and subsample errors	Treatment means: LSMEAN comparison with or without Tukey or Bonferroni adjustments for QL factor means <i>at more than one specified levels of QT factor</i>
		*Verify first that main plot or the split plot factor \times covariate interactions are not significant	
		*In the presence of covariate, estimate LSMEANS <i>at the mean level of the covariate</i>	

(Continued on next page)

Table 4. (Continued) Checklists and guidelines for analyzing data from split plot or split-split plot experiments in CRD or RCBD: step 2.

Treatment and experiment designs		Recommended analysis	
Split plot or split-split plot design	Main plot design	Subsamples or no subsamples: 1) $n > (t \times r)$ or 2) no subsamples when $n \leq (t \times r)$	Presence of covariates (yes) or no covariates (no)
RCBD	RCBD	No subsamples	No covariate, main plot or split plot experimental units are assumed to be homogeneous
		With subsamples	With covariates, main plot or split plot experimental units are assumed to be <i>not</i> homogeneous
			Analysis type
			Mixed-effects split plot ANOVA in RCBD with random or fixed block effects, use SAS PROC MIXED with RANDOM statement to separate main plot error from split plot error and random block effects
			Mixed-effects split plot ANOVA with subsamples in RCBD with random or fixed block effects, use SAS PROC MIXED with RANDOM statement to separate main plot error from split plot error and random block effects
			Verify first that main plot or the split plot factor \times covariate interactions are not significant
			Mixed-effects split plot ANOVA with subsamples in RCBD with random or fixed block effects, use SAS PROC MIXED with RANDOM statement to separate main plot and split plot errors from subsample errors, and random block effects
			Mixed-effects split plot ANOVA with subsamples in RCBD with random or fixed block effects, use SAS PROC MIXED with RANDOM statement to separate main plot and split plot errors from subsample errors, and random block effects
			*Verify first that main plot or the split plot factor \times covariate interactions are not significant
			Mean computation, all factors qualitative (QL) or at least one factor quantitative (QT): trend analysis by levels of the qualitative factor
			Mean comparison methods, all factors QL or trend analysis (for at least one factor is QT: fitting linear, quadratic regression models
			Both main plot and split plot factors are QT: Perform response surface regression using SAS PROC MIXED to handle main plot and split plot error structures
			QL \times QT factor interaction is <i>not significant</i> : Fit regression models with the same slopes for QT factor and response for <i>each level of QL factor</i> ; estimate linear, quadratic parameters and report the regression models for each QL levels
			Treatment means: LSMEAN comparison with or without Tukey or Bonferroni adjustments for QL factor means <i>at the mean level of QT factor</i>
			3) Both main plot and split plot factors are QT: Fit mixed-model response surface regression and study the response surfaces

(Continued on next page)

Table 4. (Continued) Checklists and guidelines for analyzing data from split plot or split-split plot experiments in CRD or RCBD: step 2.

Treatment and experiment designs		Recommended analysis	
Split plot or split-split plot design	Main plot design	Subsamples or no subsamples: 1) $n > (t \times r)$ or 2) no subsamples when $n \leq (t \times r)$	Analysis type
	RCBD	No covariate, main plot, split plot, or split-split plot experimental units are assumed to be homogeneous	Mixed-effects split-split plot ANOVA in RCBD with random or fixed block effects, use SAS PROC MIXED with RANDOM statement to estimate main plot, split plot and split-split plot errors and random block effects
		With covariates, main plot, split plot, or split-split plot experimental units are assumed to be <i>not</i> homogeneous	Mixed-effects split-split plot ANOVA in RCBD with random or fixed block effects, use SAS PROC MIXED with RANDOM statement to estimate main plot, split plot, and split-split plot errors and random block effects
			Treatment means: LSMEAN comparison with or without Tukey or Bonferroni adjustments for QL factor means at the mean level of QT factor
			3) Two factors are QT
			Fit mixed model response surface regression and study the response surfaces at each level of the QL factor
			*Verify first that main plot, split plot, or the split-split plot factor \times covariate interactions are not significant
			Mixed-effects split-split plot ANOVA with subsamples in RCBD with random or fixed block effects, use SAS PROC MIXED with RANDOM statement to separate main plot, split plot, and split-split plot error and random block effects from subsample errors
			Mixed-effects split-split plot ANOVA with subsamples in RCBD with random or fixed block effects, use SAS PROC MIXED with RANDOM statement to separate main plot, split plot, and split-split plot error and random block effects from subsample errors

ANCOVA, analysis of covariance; ANOVA, analysis of variance; COR, correlated; CRD, complete randomized design; IND, independent; n, total number of observations; QL, qualitative; QT, quantitative; r, replications; RCBD, randomized complete-block design; t, treatment levels.

advanced statistical methods related to designing experiments, and to consult with experiment station statisticians before designing comparative studies and analyzing horticultural data. Furthermore, checklists and statistical guidelines applicable to analyzing commonly used horticultural experiments conducted using CRD, RCBD, split plot, and split-split plot designs are presented in four tables. Download instructions are given to obtain a hardcopy of the PowerPoint presentation, which contains additional information on sample SAS codes for analyzing data from comparative horticultural experiments discussed in this paper. By reviewing these recommendations and following these guidelines, one may ensure that horticultural research publications conform to the very highest scientific standards with respect to experimental design and statistical methods.

Literature Cited

- Chew, V. 1976. Comparing treatment means: A compendium. *HortScience* 11:348–357.
- Fernandez, G.C.J. 1990. Evaluation of moving mean and border row mean covariance analysis for error control in yield trials. *J. Amer. Soc. Hort. Sci.* 115:241–244.
- Fernandez, G.C.J. 1991. repeated measure analysis of line-source sprinklers experiments. *Hort-Science* 26:339–342.
- Fernandez, G.C.J. 1992. Residual analysis and data transformations: Important tools in statistical analysis. *HortScience* 27:297–300.
- Fernandez, G.C.J. 2002. Data mining using SAS applications. CRC/Chapman Hall, Boca Raton, FL.
- Government of Alberta. 2006. Understanding design and analysis of research experiments. 8 May 2007. <[www1.agric.gov.ab.ca/\\$department/deptdocs.nsf/all/webdoc3033](http://www1.agric.gov.ab.ca/$department/deptdocs.nsf/all/webdoc3033)>.
- Jones, D. 1984. Use, misuse, and role of multiple-comparison procedures in ecological and agricultural entomology. *Environ. Entomol.* 13:635–649.
- Kuehl, R.O. 1996. Design of experiments: Statistical principles of research design and analysis. 2nd ed. Duxbury Press, Pacific Grove, CA.
- Littell, R.C., G.A. Milliken, W.W. Stroup, R.D. Wolfinger, and O. Schabenberger. 2006. SAS for mixed models 2nd ed. SAS Institute, Cary, NC.
- Petersen, R.G. 1977. Use and misuse of multiple comparison procedures. *Agron. J.* 69:205–208.
- Piepho, H.P. 1999. Analyzing disease incidence data from designed experiments by generalized linear mixed models. *Plant Pathol.* 48:668–674.
- Piepho, H.P., A. Buchse, and K. Emrich. 2003. A hitchhiker's guide to mixed models for randomized experiments. *J. Agron. Crop Sci.* 189:310–322.
- Saville, D.J. 1990. Multiple comparison procedures: The practical solution. *Amer. Stat.* 44:174–180.
- Schabenberger, O. 2004. Statistical methods for spatial data analysis. Texts in statistical science series. CRC Press, Boca Raton, FL.
- Schabenberger, O. 2006. SAS code for some experimental designs. 8 May 2007. <http://www.ats.ucla.edu/stat/SAS/library/SASexpDes_os.htm>.
- Schabenberger, O. and F.J. Pierce. 2001. Contemporary statistical models for the plant and soil Sciences CRC Press, Boca Raton, FL.
- Sit, V. 1995. Analyzing ANOVA designs. Biometrics Information Handbook No 5. Ministry of forest research program, Victoria British Columbia. Working paper. 8 May 2007. <<http://www.stat.sfu.ca/~cschwarz/Stat-650/MOF/WP07.anova.pdf>>.
- Waldo, D.R. 1976. An evaluation of multiple comparison procedures. *J. Anim. Sci.* 42:539–544.