IEESE International Journal of Science and Technology (IJSTE), Vol . 1, No. 4, December 2012,11-15
ISSN : 2252-5297

Model Comparison for Multienvironment Trials

Mohammad Masjkur

Department of Statistics, Faculty of Sciences-Bogor Agricultural University
Bogor, Indonesia
m_masjkur@yahoo.com

Abstract. In multienvironment trials, the local spatial tendency within trials and the residual heterogeneity between trials can be jointly modeled in the context of mixed linear models. The objective of this study was to know the suitability of spatial models within trials and the residual heterogeneity between trials for multienvironment fertilizer trials. The research used field trials data of rice fertilization in Karawang and Kebumen region. Fertilizer treatments consist of fourteen levels of fertilizer type with three replications. In Kebumen region the research was conducted over two growing seasons. The experiment used a randomized complete block design. The results show that for single experiment the best model for Karawang and Kebumen1 location was RCBCE power isotropic model, whilst for Kebumen2 location was RCBCE power anisotropic model. In combined experiments the best model was CE power anisotropic model with heterogeneous residual variance across locations. Modeling of the local spatial tendencies using an analysis of variance including heterogeneous residual variances increased the ability to identify differences among treatments.

Key-Words: *randomized complete block design, multienvironment trials, local spatial, residual heterogeneity.*

1. Introduction

Fertilizer recommendation usually based on field trials conducted in a wide range of environment (multienvironment trials). The objectives was to calibrate critical levels of nutrient for soil testing, optimum doses of fertilizer or the efficient nutrient management. The design usually used was randomized complete block design (RCBD). Field conditions such as fertility variation or elevation level of soil relatively homogen was subject to blocking. Then fertilizer treatments placed randomly within block.

Randomization of treatment within block actually was an effort to satisfy the independence assumption of error for model design and to overcome spatial correlations ([1]-[2]). However, much of research revealed the spatial correlation of treatments within block although had been used randomization. Therefore some researcher recommended significant examination of spatial correlation structure (*data-driven approach)* beside basic model of randomized complete block design *(model-driven approach*) to know better information of treatment effects ([2]-[5]).

Some spatial models potentially can be used to improve precision of experiment. The classical approach was nearest neighbour method ([6]-[9]). Another approach was geostatistical methods such as linear variance, first order autoregressive (AR1), gaussian, spherical and power models ([2]; [4]-[5]; [10]). The nearest neighbor methods indirectly take spatial dependencies between neighboring plots into account, whereas geostatistical methods model the covariance function directly. The suitability of spatial models varies from study to study. It depends on environment and crop condition, candidate models and also layout of randomized complete block design (([2]; [4]-[5]; [11]).

In multienvironment trials (MET), the local spatial tendency within trials and the residual heterogeneity between trials can be jointly modeled in the context of mixed linear models. By using a two-dimensional coordinate system at each trial, it is possible to define the plot location in a field, for example, from latitude and longitude of plot centers ([4]; [9]). For METs, the modeling of the between trial residual heterogeneity can greatly improve the combined analysis, both in models with block effects and with a spatial correlation structure.

The objective of this study was to know the suitability of spatial models within trials and the residual heterogeneity between trials for multienvironment fertilizer trials.

2. Linear Mixed Model

Linear mixed model takes the following form **:**

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{u} + \varepsilon \qquad (1)$$

where $\boldsymbol{y}$ is an N observation vector, $\boldsymbol{\beta}$ is a fixed effect vector, $\boldsymbol{u}$ is a random effect vector, $\boldsymbol{\varepsilon}$ is a residual vector, **X** and **Z** are a design matrix. The random effect u has normal distribution with mean 0 dan variance matrix **G**. The distribution of residual was normal with mean 0 dan variance matrix **R.**

Basic assumption of this model is $\boldsymbol{u}$ and $\boldsymbol{\varepsilon}$ were not correlated and the expected value was zero. If the variance of random effect var ($\boldsymbol{u}$) = $\sum_{\boldsymbol{u}}$ and $\mathbf{V_1}$= var ($\mathbf{Z}\boldsymbol{u}$) = $\mathbf{Z}\sum_{\boldsymbol{u}}\mathbf{Z}'$ and variance of error $\mathbf{V_2}$ = var (ε), then variance of $\mathbf{y}$ is var ($\mathbf{y}$) = $\mathbf{V_1} + \mathbf{V_2}$. Therefore any correlation of observation can be specified in $\mathbf{V_2}$ and/or $\mathbf{V_1}$ ([3]; [6]; [12]).

2.1. Randomized Block Design with iid Errors (RCBiid)

In classical randomized block design, the error within block assumed iid errors **(**$\boldsymbol{\varepsilon}$) where $\mathbf{V_1} = \mathbf{Z}\sum_{\boldsymbol{u}}\mathbf{Z}'$ and $\mathbf{V_2} = \sigma^2 \boldsymbol{I_n}$; $\boldsymbol{I_n}$ is an identity matrix. Consequently, any spatial correlation of observations is reflected only in $\mathbf{V_1}$.

2.2. Randomized Block Design with Correlated Errors (RCBCE)

This design model was a nonclassical randomized block design with spatially correlated error where $\mathbf{V_1} = \mathbf{Z}\sum_{\boldsymbol{u}}\mathbf{Z}'$ and $\mathbf{V_2}= \sigma^2 \mathbf{W}$; **W** is an n x n spatial covariance matrix whose *ij*th element is defined as a function of distance (h_{ij}) between site i and j. In this approach, any spatial correlation of observations is reflected in both $\mathbf{V_1}$ and $\mathbf{V_2}$. Clearly the RCBiid model is a reduced form of the RCBCE model, because if no spatial correlation is present, **W** will reduce to $\boldsymbol{I_n}$.

2.3. Correlated Error Analysis (CE method)

The CE method assumes that $\mathbf{V_1} = 0$, which means the random block effects are not considered and are removed from the model in Eq. [1]. Any spatial correlation present is defined only in $\mathbf{V_2}$. Clearly the CE model is a reduced form of the RCBCE model. If the random effects of the RCBCE models are determined ineffective, that is, the elements of the matrix $\sum_{\boldsymbol{u}}$ are all zeros, RCBCE models become CE models. Additionally, all the considerations discussed above for the RCBCE method about specifying a function to model the spatial covariance also apply to CE implementation.

3. Data and Method

This research used field trials data of rice fertilization in Karawang and Kebumen region. Fertilizer treatments consist of fourteen levels of fertilizer type with three replications. In Kebumen region the research was conducted over two growing seasons. The experiment used a randomized complete block design. Yield data of each trial were analysed by spatial model as,

$$Y_{ij} = \mu + \tau_{k(ij)} + T_{ij} + \varepsilon_{ij}$$

where Y_{ij} is the rice yield (ku/ha) of the *j*th plot and the *i*th block or plot *ij*, the term $\mu + \tau_{k(ij)}$ represents the mean of *k*th fertilizer effect on plot *ij*, T_{ij} represents spatial variability effect on that plot, and ε_{ij} is a random error. The basic model is RCBiid variance analysis. In this case, the spatial effect T_{ij} were assumed to be constant for all plot within the same block, that is $T_{ij} = \beta_i$ (the *i*th block effect). We compared different spatial analyses (RCBCE isotropic and anisotropic power correlation) with the classical RCBiid model with random blocks at each location.

After comparing these models within locations, we conducted an across-locations analysis using the following MET models. The first two procedures were based on analysis of variance for an RCBD at each location:

$$Y_{ijk} = \mu + L_j + B(L)_{k(j)} + F_i + FL_{(ij)} + \varepsilon_{ijk} \quad (2)$$

where Y_{ijk} is the rice yield (ku/ha) of the *i*th fertilizer, in location j and the *k*th block; μ is the overall mean; L_j is the effect of location j with j=1, ..., s; $B(L)_{k(j)}$ is the random effect of block k in location j with k=1, ..., n; F_i is the effect of fertilizer i with i =1, .., f; $FL_{(ij)}$ is the effect of interaction of fertilizer i and location j; and ε_{ijk} is an error term associated with observation Y_{ijk} .

Except for ε_{ijk} and the block effects, all of the model factors were considered as fixed effects. The ε_{ijk} were assumed independent with a constant variance σ^2 in the first method, assuming that local spatial variation and heterogeneous residual variance between locations do not exist. The block effect variances were also assumed to be homogeneous (RB model). The second procedure denoted as an RBH model was also based on Eq. [2], but permitted heterogeneous residual variances across locations.

Besides we compared different spatial analyses (CE isotropic and anisotropic power correlation) with RB and RBH models. Model estimation was used restricted maximum likelihood (REML) procedure for linear mixed model with random block effects. The best model selection was based on Bayesian Information Criteria (BIC) ([3]-[4]; [7]; [9]).

4. Results and Discussion

4.1. Single Experiment

Based on BIC the best model for Karawang and Kebumen1 location was RCBE power isotropic model. This model has BIC value smaller than the RCBDiid and RCBE power anisotropic models. In Kebumen2 location the best model was RCBE power anisotropic model. The BIC value smaller than RCBDiid model and the RCBE power isotropic models (Table 1). This indicated that in Karawang and Kebumen location, the RCBDiid-model fit residuals were significantly spatially correlated, providing evidence of field heterogeneity within blocks that could make the RCBiid method less powerful than a method that incorporated the spatial correlation.

Table 1. Candidate covariance models and their related statistics and covariance parameters including randomized complete block with iid errors (RCBiid), randomized complete block with correlated errors (RCBCE) models for Karawang and Kebumen location (Italic type indicates the selected model).

Models	BIC‡	*p value*¶	SE#	Covariance parameter estimate		
				Range	Block	Residual
			Karawang			
RCBiid	148.4	0.6728	1.90	-	2.41	5.44
RCBE power isotropic model	132.2	0.1022	1.18	0.66	2.31	5.84
RCBE power anisotropic model	132.4	0.0519	1.11	0.32r 0.73c	2.56	4.39
			Kebumen1			
RCBiid	74.4	0.0150	0.62	-	0.31	0.58
RCBE power isotropic model	73.4	0.0001	0.62	-0.34	0.46	0.56
RCBE power anisotropic model	74.7	0.0009	0.62	0.39r 5.9e-18c	0.19	0.67
			Kebumen2			
RCBiid	77.7	0.0004	0.55	-	0.03	0.48
RCBE power isotropic model	77.6	<0.0001	0.55	-0.18	0.04	0.46
RCBE power anisotropic model	77.1	<0.0001	0.55	-0.37r 0.08c	0.06	0.46

‡ Bayesian Information Criterion
¶ *p* value of *F* test for treatment effect
# The average of standard errors of treatment comparisons.
r, c Range of power anisotropic model in row or column direction

In Karawang and Kebumen location trials, the errors were spatially correlated and the RCBiid method tended to give higher estimates of the standard errors of treatment comparisons than the other models (Table 1). Consequently, the RCBiid and the selected covariance model could have resulted in different conclusions regarding treatment significance. The selected covariance model for Karawang location resulted in a smaller significance level for treatment comparison, $p = 0.10$ compared with $p = 0.67$ for the RCBiid model. In Kebumen1 and Kebumen2 location the selected covariance model resulted smaller significance level for treatment comparison $p = 0.0001$ and $p =< 0.0001$ compared with $p = 0.0150$ and $p = 0.0004$ for the RCBiid model, respectively.

4.2. Combined Experiment

In the MET evaluated for all models, the treatments x location interaction was significant (p < 0.0001) (Table 2). This is probably because the treatments x environment interaction was random by nature. Also because plot spatial correlation models differ among locations. Therefore, the treatments were compared within each location.

When the block effect within the location was considered, it was not significant (p > 0.05). Tables 2 present the BIC values associated with each model. According to BIC, the Powa model was better than RB and Pow models. This indicate that anisotropic spatial correlation analysis was better than random block effect or isotropic spatial model. The RBH model was not superior to RB model. The differences in residual variances between locations were negligible (Tables 2). Also the differences between the *p value* for the hypothesis test of no treatment effect obtained within each location were very similar for the homogeneous and heterogeneous block design models (Table 3).

Table 2. Candidate covariance models and their related statistics and covariance parameters for combined experiment. (Italic type indicates the selected model)

Models	BIC‡	*p value*¶	SE#	Covariance parameter estimate		
				Range	Block	Residual
RB	169.0	<0.0001	0.43	-	0.07	0.23
RBH	172.2	<0.0001	0.43	-	0.07	0.22a 0.29b 0.19c
Pow	169.8	<0.0001	0.44	0.27	-	0.31
PowH	164.0	<0.0001	0.46	0.67a 0.19b -0.09c		0.42a 0.38b 0.20c
Powa	166.8	<0.0001	0.43	0.47r 0.16c	-	0.32
PowaH	153.5	<0.0001	0.43	0.76ra 0.64ca 0.65rb 0.01cb -0.30rc 0.21cc	-	0.41a 0.47b 0.21c

RB, random block effects; RBH, random block effect with heterogenous residual variances;
Pow, CE isotropic power model; PowH, CE isotropic power model with heterogenous residual variances;
Powa, CE anisotropic power model; PowaH, CE anisotropic power model with heterogenous residual variances.
‡ Bayesian Information Criterion
¶ *p* value of *F* test for interaction treatment x location effect
# The average of standard errors of treatment comparisons.
r, c Range of power anisotropic model in row or column direction
a, b, c Range or residual for Karawang, Kebumen1 and Kebumen2 location respectively.

On the other hand, by fitting models with spatial correlation and heterogeneous residual variance between environments, the BIC values were reduced in the isotropic and anisotropic analysis (Tables 2). In the PowH and PowaH model, the difference between the residual variances was important with a ratio between the highest and the lowest residual variances greater than two (Tables 2). This suggested that the models for a heteroscedastic residual variance are more appropriate than their homogeneous residual variance version. The important *p value* change was found in Karawang location (Table 3).

Table 3. P-value of treatment effect for six models for each location.

Location	Models					
	RB	RBH	Pow	PowH	Powa	PowaH
Karawang	0.6527	0.6139	0.8854	0.4579	0.7889	0.1247
Kebumen1	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Kebumen2	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0005

RB, random block effects; RBH, random block effect with heterogenous residual variances; Pow, CE isotropic power model; PowH, CE isotropic power model with heterogenous residual variances; Powa, CE anisotropic power model; PowaH, CE anisotropic power model with heterogenous residual variances.

5. Conclusion

The models with block effect (RCBiid) was not the best model to fit the fertilizer data. For single or combined experiment modeling the local spatial tendencies using an analysis of variance (RCBE or CE) was better than random block effect (RCBiid). For single experiment the best spatial model to fit the fertilizer data was differ among locations. In combined experiment, an analysis of variance including heterogeneous residual variances was more apropriate than their homogeneous residual variance versions. Modeling of the local spatial tendencies using an analysis of variance including heterogeneous residual variances increased the ability to identify differences among treatments.

6. References

[1] van Es, H. M., C. P. Gomes, M. Sellman, and C. L. van Es. 2007. *Spatially-Balanced Complete Block Designs for field experiments*. Geoderma 140:346-352.

[2] Christel Richter and Bärbel Kroschewski. 2012. *Geostatistical Models in Agricultural Field Experiments: Investigations Based on Uniformity Trials*. Agron. J. 104:91–105.

[3] Hong, N., J. G. White, M. L. Gumpertz, R. Weisz. 2005. *Spatial Analysis of Precision Agriculture Treatments in Randomized Complete Blocks: Guidelines for Covariance Model Selection*. Agron. J. 97:1082–1096.

[4] Casanoves, F., R. Macchiavelli, and M. Balzarini. 2005. *Error Variation in Multienvironment Peanut Trials: Within-Trial Spatial Correlation and Between-Trial Heterogeneity*. Crop Sci. 45:1927–1933.

[5] Bettina U. Müller, Kathrin Kleinknecht, Jens Möhring, and Hans-Peter Piepho. 2010. *Comparison of Spatial Models for Sugar Beet and Barley Trials*. Crop Sci. 50:794–802.

[6] Lambert, D. M., J. L. DeBoer, and R. Bongiovanni. 2003. *Spatial Regression Models for Yield Monitor Data: A Case Study from Argentina.* Paper prepared for presentation at American Agricultural Economics Association Annual Meeting, Montreal, Canada.

[7] Rong-Cai Yang, Terrance Z. Ye, Stanford F. Blade, and Manjula Bandara. 2004. *Efficiency of Spatial Analyses of Field Pea Variety Trials.* Crop Sci. 44:49–55.

[8] Bullock, D. S. and J. L. DeBoer. 2006. *Using Spatial Analysis to Study the Values of Variable Rate Technology and Information.* Invited paper prepared for presentation at the International Association of Agricultural Economists Conference, Gold Coast, Australia, August 12-18, 2006.

[9] Casanoves, F., R. Macchiavelli, and M. Balzarini. 2007. *Models for multi-environment yield trials with fixed and random block effects and homogeneous and heterogeneous residual variances*. J. Agric. Univ. P.R. 91(3-4):117-131.

[10] Heike Knörzer, Bettina U. Müller, Buqing Guo, Simone Graeff-Hönninger, Hans-Peter Piepho, Pu Wang, and Wilhelm Claupein. 2010. *Extension and Evaluation of Intercropping Field Trials Using Spatial Models*. Agron. J. 102:1023–1031.

[11] Fagroud, M. and M. Van Meirvenne. 2002. *Accounting for Soil Spatial Autocorrelation in the Design of Experimental Trials.* Soil Sci. Soc. Am. J. 66:1134–1142.

[12] Searle, S. R. 1987. *Linear Models for Unbalanced Data*. John Wiley and Sons, Inc. New York.