


## Research Article

# Effect of Experimental Blocking on the Suppression of Spatial Dependence Potentially Attributable to Physicochemical Properties of Soils

Aquiles E. Darghan,<sup>1</sup> Giovanni Reyes ,<sup>2</sup> Carlos A. Rivera,<sup>1</sup> and Edwin F. Grisales<sup>1</sup>

<sup>1</sup>Facultad de Ciencias Agrarias, Universidad Nacional de Colombia, Bogotá 110111, Colombia

<sup>2</sup>Research Group SIPAF, Facultad de Ciencias Agrarias, Universidad Nacional de Colombia, Bogotá 110111, Colombia

Correspondence should be addressed to Giovanni Reyes; greyesm@unal.edu.co

Received 4 June 2021; Revised 17 July 2021; Accepted 3 August 2021; Published 6 September 2021

Academic Editor: Angelos Markopoulos

Copyright © 2021 Aquiles E. Darghan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the basic principles of experimental design is blocking, which is an important factor in the treatment of the systematic spatial variability that can be found in the edaphic properties where agricultural experiments are conducted. Blocking has a mitigating or suppressing effect on the spatial dependence in the residuals of a model, something desirable in standard linear modeling, specifically in design models. Some computer programs yield a  $p$  value associated with the blocking effect in the analysis of variance table that in many cases has been incorrectly used to discard it, and although it may improve some properties of the analysis, it may affect the independence assumption required in several models. Therefore, the present research recommends the use of the  $H$  statistic associated with the corrected blocking efficiency to show the role of blocking in modeling with the incorporation of an additional advantage rarely considered related to the suppression or mitigation of spatial dependence. With the use of the Moran index, the spatial dependence of the residuals was studied in a simple factorial design in a completely randomized and blocking field layout, which evidenced the advantages of blocking in the mitigation or suppression of the spatial dependence despite the apparently little or no importance it seems to show in the analysis of variance table.

## 1. Introduction

In agricultural experiments, control of experimental error is often important to improve research precision. In the field, the existence of systematic spatial variability attributable to the physical-chemical properties of the soil is common, for example, that associated with gradients due to slope, fertility, or humidity that are easily recognized in the field. The sources of this variability are very diverse, so in these situations, it requires an adequate experimental design and data analysis, where blocking and its analysis of variance continue to be the important techniques to consider this variability. The *European Journal of Agronomy* (2012) published an article related to the review of assumptions of analysis of variance (AOV) in the agronomic context, revealing a large number of inappropriate applications due to the

few times they review the assumptions of the technique, implying the revision of the assumption of independence [1].

There are several investigations that generate data that can show dependency structures, for example, those that involve designs in repeated measures, time series data, hierarchical or grouped studies, and the one that we consider in the current investigation that could be associated with the study of the physical or chemical properties of the soil, of which multiple studies have been accomplished showing clear structures of spatial dependence in variables such as apparent electrical conductivity, soil organic carbon content, texture, pH, moisture, apparent density, porosity, and mineral content of soils [2, 3]. This spatial variability caused Fisher [4] almost 100 years ago to address the issue of blocking in agricultural experimentation to minimize the

systematic error caused by the inherent variability of the soils or environments where the experiments were conducted.

Despite being a fairly well-known technique, some benefits in data analysis and their role in the assumption of independence of the residuals in the models are still unknown or cursorily considered. The great diversity of blocking modalities, such as complete, incomplete, generalized, fixed, random, solvable, and nested blocking, may not be adequately treated, thus rendering AOV tables not only incomplete but also erroneous. This could include, for example, when there is a generalized blocking where the interaction that could occur between blocks and factors is usually omitted [5–8].

The omission of the role of blocking as a factor for attenuation or suppression of spatial dependence can complicate data analysis. Gotway and Cressie [9] first showed a spatial AOV to consider dependence, pointing out that in the presence of spatial autocorrelation, the usual tests for the presence of treatment effects may no longer be valid. This is where blocking is important as it can eliminate spatial dependency on error terms. This assumption is part of the methods usually considered in the basic courses of experimental design or standard linear models [5, 10], specifically in the design models [11] since in advanced courses techniques to deal with this dependency are usually taught [12].

In this research, multiple AOV were made by Monte Carlo simulation for both a simple factorial design in a completely randomized arrangement (CRD) and a randomized complete block with one observation per cell (RCB). Spatial dependence was induced, and each plot was assumed as an experimental unit and as a response to any indicator of the yield of a crop. In addition, a factor with three levels was considered, and it was blocked at 10 levels for reasons of a hypothetical gradient of some soil property. The imposition of spatial dependency was done in order to note the effect of considering the blocking in the modeling process when comparing the models with and without blocking. The blocking efficiency was evaluated with a new strategy much more convenient to demonstrate the suppressive effect of spatial dependence using the blocking efficiency statistic, since the simulations evidenced a lower probability of suggesting the elimination of the blocking effect due to the misuse of the  $p$  value generated in the analysis of variance table; in addition, its measurement provides in most cases evidence of the role of blocking in the mitigation of spatial dependence, which was evaluated by the Moran index (MI) and the matrix of spatial weights associated with the distribution of the plots in the lot. The results revealed the great utility of blocking in minimizing the effect of spatial dependence or even in suppressing it, which makes this methodology a current and highly useful tool in teaching experimental design at the level of undergraduate and postgraduate courses where the statistical tools that are considered are still classical or elementary, such as the one dictated by following the texts of Mead et al. [13] and Montgomery [14] and where the usual design model taught at these stages requires the assumption of independence of residuals [15].

## 2. Materials and Methods

On the same data set, two linear models with a single factor were fitted, which differed by the incorporation of the effect of the blocks to the CRD model, thus generating the RCB model, i.e., the randomized block model with one observation per cell. In general, the matrix form of both models is written as

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \quad (1)$$

where  $\mathbf{X}$  is associated with the matrix of each design model,  $\mathbf{Y}$  represents the response variable,  $\beta$  represents the vector of parameters, and finally  $\varepsilon$  is associated with the vector of residuals, which for the case of standard linear modeling, it is assumed that  $\varepsilon \sim N_n(0, \sigma^2 I)$ , which is evidenced by the diagonal shape of the matrix  $\sigma^2 I$  and where  $0$  represents the vector of null means and  $I$  is an identity matrix of dimension  $n \times n$ . This expression suggests normality in distribution, homogeneity of variances, and independence of residuals, the latter assumption being the one that is questioned when spatial dependence is shown in this case, since it may also imply a different source of dependence. In addition, the vast majority of design models studied in experimental design courses require this independence assumption [8, 16–18]. However, in more advanced courses, it applies in the case of partially linear models, nonparametric models, and linear and nonlinear autoregressive models where independence of residuals is assumed and where blocking can be a systematic variance minimization strategy [19]; however, a measure analogous to the  $H$  statistic should be developed, as this was developed for the linear model.

For the management of spatial information, artificial coordinates were created associated with the distribution of the plots in both rows and columns. The center of the plot was used as a representative coordinate to create the distance matrix of all the neighbors and thus generate the spatial weight matrix ( $\mathbf{W}$ ) [20] and be able to evaluate the MI [21]. As the data was intentionally ordered in the direction of the gradient of the potential physicochemical property of interest to create the artificial spatial dependence, the MI was calculated for the model without blocks, and 1000 AOV generated by Monte Carlo simulation were run, the vectors were extracted as residuals per model, and the MI was applied, from which a majority of results with induced spatial dependence were obtained. Subsequently, the model with blocks was adjusted, and its residuals were extracted, thus generating four regions of interest, a first region labeled A, associated with those results that showed spatial dependence in the CRD model and that resulted in spatial dependence in the RCB model, a B region for those that showed spatial dependence in the CRD model and that resulted in spatial independence in the RCB model, a C region with those with independence in both models, and a D region with the results with spatial independence in the CRD model and that resulted with spatial dependence in the RCB model. A scatter plot was constructed to illustrate the four regions, and a confusion matrix was constructed to show the distribution of the analysis count in each region.

```

Function tests( r )
Start
pvSW = pvalue( ShapiroWilks_NormalityTest( r ) )
pvSF = pvalue( ShapiroFrancia_NormalityTest( r ) )
pvB = pvalue( Bartlett_HomoscedasticityTest( r ) )
pvFK = pvalue( FlignerKilleen_HomoscedasticityTest( r ) )
MI = MoranIndex( r, W )
pvMI = pvalue( MI )
Return [pvSW, pvSW, pvSF, pvAD, pvB, pvL, pvFk, MI, pvMI]
End

```

ALGORITHM 1: Pseudocode for calculating normality and homoscedasticity tests.

```

Var
Trt: vector [1 ... 3] #treatments
Block: vector [1... 10] #blocks
Nsim: #simulations

Start
Solution: matrix [1... Nsim, 1... 20]
For i =1 to Nsim
  Resp = vector [1... 30] ~ N ( $\mu = 300, \sigma = 20$ )
  Resp = PartialSpatialSorted (Resp)
  Mod1 = AnalysisOfVariance (Resp ~ Trt)
  Mod2 = AnalysisOfVariance (Resp ~ Bloq + Trt)
  H = BlockEfficiency (Mod2)
  Solution [i, 1... 9] = tests (Residuals (Mod1))
  Solution [i, 10... 18] = tests (Residuals (Mod2))
  Solution [i, 19] = H
  Solution [i, 20] = pvalue (block (Mod2))
End For
End

```

ALGORITHM 2: Pseudocode to form the structure of the adjusted models.

In order not only to work with synthetic data, several databases of recognized experiments available in the R *agridat* library [22] were used to show the mitigating or suppressing effect of dependency, specifically the bases: *rothamsted.brussels*, *mead.strawberry*, *federer.tobacco*, and *rothamsted.oats*, where RCB designs were run. In our case, we ran it with the absence and presence of blocking to see the effect of blocking on those cases and to explain the importance of incorporating other variables in the modeling process to finish correcting the effect of spatial dependence that with only blocking is not possible and manages to correct. It is important to note that the models adjusted in principle are models with blocks, to which the removal of the blocks is done to see their effect, and it is not that they were completely randomized models to which the blocking was done later.

As the analysis of spatial dependence was carried out on the residuals of each model, the dependency relationship on the residuals with the spatial dependence on the response variable was explored, thus generating a result related to this practice. To summarize, the information and scatter diagrams were constructed with the regions associated with the classifications according to the  $p$  values of each method and model.

As the proposal with blocking was used as a spatial dependence attenuator or suppressor and since it is known there is no valid test for the effect of blocking, the proposal of the  $H$  statistic of Lentner et al. [23] was used to evaluate the efficiency of blocking, and it was compared with the  $p$  value that is automatically generated in the AOV tables of the RCB model. The objective of this measure was to provide a more convenient tool and of innovative use for researchers to rule out the usefulness of blocking in their modeling and thus avoid two inconvenient procedures, namely, to remove the blocking reason only because the output of the AOV showed an apparent  $p$  value erroneously considered a measure of the effect of blocking and second to provide an updated measure of blocking efficiency that more clearly demonstrated the suppressive effect of spatial dependence caused by blocking. Finally, tables of compliance with the assumptions of both normality in distribution with two tests (Shapiro-Wilk and Shapiro-Francia) and that of homoscedasticity with two tests (Barlett and Fligner-Killeen) were presented. As the interest was not interpretive of the effects of the factor or the efficiency of blocking, the emphasis was always kept on the assumptions, specifically that of independence. All simulation and statistical analyses were performed with RStudio software.

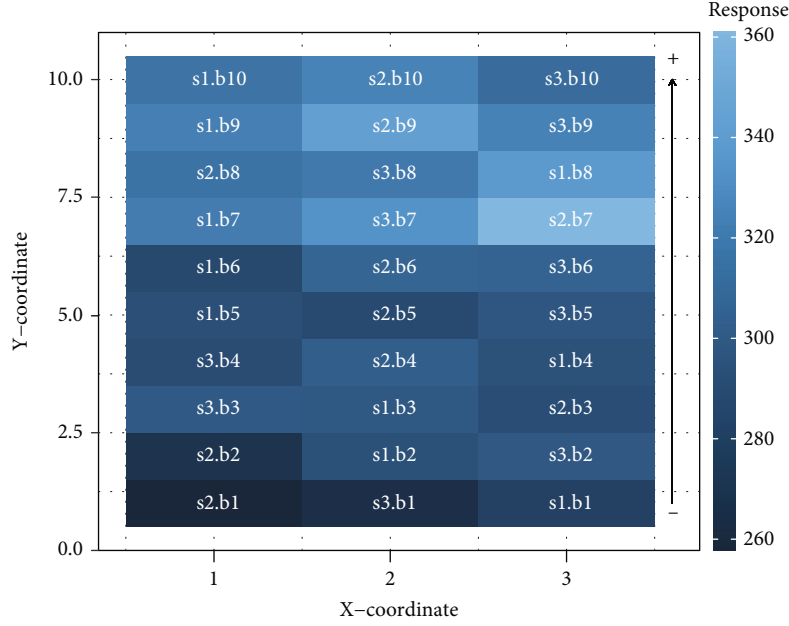


FIGURE 1: Arrangement of treatments and blocks in the gradient in the artificial coordinates and distribution of the response variable displayed in a heat diagram (a particular simulation).

### 2.1. Pseudocode for the Generation of Statistical Analysis.

Next, two algorithms were presented in pseudocode, the first (Algorithm 1) to calculate the tests of normality and homoscedasticity of residuals, as well as for the calculation of the MI, and it was associated with the  $p$  value. The second (Algorithm 2) formed the structure of the adjusted models to carry out the simulations. The distribution to be used and its parameters were set, the response was given a partial order to induce dependence, and the two models were fitted. Subsequently, Algorithm 1 was invoked to perform the tests on the extracted residuals.

## 3. Results and Discussion

The arrangements of the treatments and blocks were made in the perpendicular direction of the gradient of the soil property (indicated by the arrow in Figure 1). The gradient was indicated from a lower value (-) to a maximum value of the property (+). The equispaced levels were constructed; in this way, a linear gradient in the direction of the Y coordinate was assumed. The labels used the general expression  $s(k) \cdot b(m)$ , with  $k = 1, 2, 3$ ;  $m = 1, 2, \dots, 10$ , where  $k$  is the  $k$ th treatment and  $m$  is the  $m$ th level of blocking. The blue scale was associated with the magnitude of the simulated response, which changed in each simulation, maintaining the same probability distribution. To take advantage of the opportunities that arise when considering spatial dependence [24], data with induced spatial dependence were generated based on computational ordering (Algorithm 2).

The CRD design models were adjusted:  $y_{km} = \mu + \tau_k + \varepsilon_{km}$  with  $k = 1, 2, 3$ ;  $m = 1, 2, \dots, 10$ , assuming first the absence of blocking and the CRB model:  $y_{km} = \mu + \tau_k + \beta_m + \varepsilon_{km}$ , with  $k = 1, 2, 3$ ;  $m = 1, 2, \dots, 10$ , for the case with blocks. For both models, the independence of residuals was

	Spatial I	RCB Spatial D	$\Sigma$
Spatial I	14.3% 143	0.7% 7	15% 150
CRD Spatial D	15.4%	6.2% 62	85% 850
$\Sigma$	93.1% 931	6.9% 69	1000

FIGURE 2: Confusion matrix associated with the  $p$  values (MI) for CRD and RCB models ( $\alpha = 0.05$ ).

required as previously described. For the 1000 simulations, the usual value of the significance level at  $\alpha = 0.05$  was used as a threshold in the visualization to generate the four classification regions associated with the  $p$  values obtained from the MI calculated from  $\mathbf{W}$  (weight matrix), standardized by rows using the inverse of the Euclidean distance between the centers of each plot for the established artificial coordinates. It should be remembered that the response data with a predominance in spatial dependence were artificially created; in fact, 85.0% of the simulations presented spatial dependence (see margin with spatial dependence in the confusion matrix in Figure 2), and expressed otherwise,

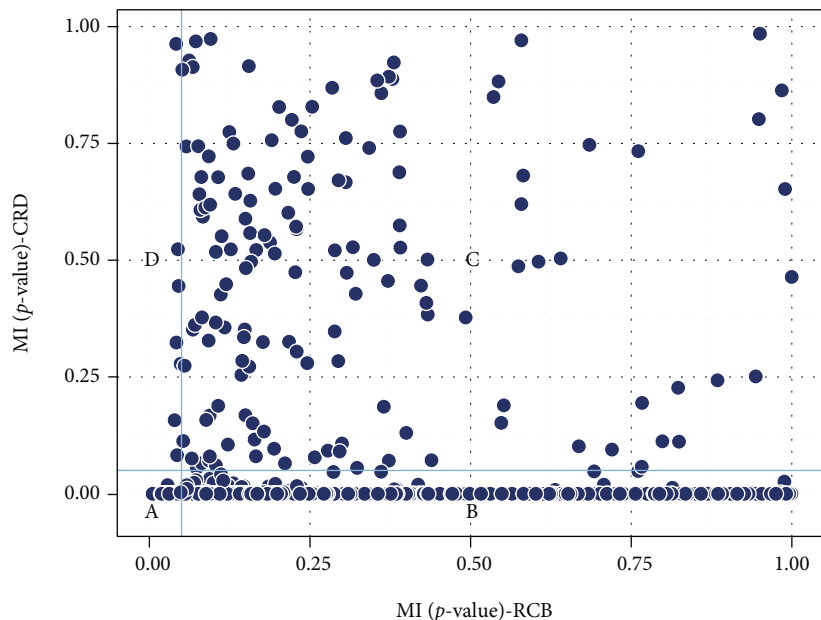


FIGURE 3: Distribution of the  $p$  values (MI) of the residuals obtained by simulation for the CRD and RBC design model.

evidence was found in the simulations against the independence hypothesis.

Figure 3 shows the distribution of the  $p$  values of MI for both models, where a large number of points, belonging to the dependency regions with the CRD model, moved to region B; that is, they became independent (spatially), showing the effect of the blocking in eliminating or reducing the spatial dependence in the modeling. Undoubtedly, this result is subject to the correct construction of the block [25] and therefore to the appropriate selection of the gradient; otherwise, it could happen that the results were invalidated and a redesign might be necessary, since some authors have already declared the inadequacy of blocking in these circumstances [26].

Figure 2 shows the counts and the percentage distribution of the rankings in the four regions in Figure 3. As mentioned before, region D was the most important in the current research, since it corresponds to 788 simulations of 1000 carried out (78.8%) with the effect of interest, namely, reduce or eliminate from a statistical point of view the spatial dependence, thus reaching the fulfillment of one of the important assumptions in the modeling with AOV in the basic experimental designs where independence is assumed in the residuals, which in this case is considered potentially attributable to the gradients of some soil properties that are considered spatial in nature, as shown in various studies [27, 28]. If instead of  $\alpha = 0.05$ , one such as  $\alpha = 0.01$  had been used in order not to exaggerate the evidence of spatial independence, the results would have been more forceful not to accept the advantages of blocking in mitigating or eliminating spatial dependence.

Currently, there are many tools available to perform an adequate blocking available for students of basic design courses, such as spectral information using different indices, generated from remote or proximal sensing, which with simple processing can easily define the structure and dimension

of the blocks so that the orientation of an important gradient in the modeling process can be simple [29].

MI has undoubtedly proven to be an appropriate measure to assess the spatial dependence of soil properties [30, 31]. Before a spatial inferential analysis, it has become routine to do an exploratory analysis of the variables involved in the modeling process; this usually involves calculating the MI but using the response variable as descriptive statistics prior to spatial modeling or a spatial interpolation process [32]. However, in our research, we found that the application of MI on a response variable in an exploratory way evidencing spatial dependence did not guarantee that this was precisely the result when this variable was incorporated into the modeling, because precisely the effect of the blocking can remove such dependency, something that was not observed before modeling where it had not been blocked.

Figure 4 shows the scatter diagrams of the  $p$  values of the MI of the residuals (Y-axis) and MI of the response (X-axis) for both models. It is remarkable how the point cloud is concentrated in region B in the RCB model since the dependency present in the CRD model had been corrected. Once blocked, most of the responses showed dependence, but this was not the case for the residuals; these were actually predominantly independent. This is the desired assumption in modeling that could confuse those with little expertise of the subject, who might believe that some modeling technique would be necessary for special spatial avoiding the possibility of adjusting to a simpler model, thus discarding a fairly controversial principle of modeling, known as the “principle of parsimony” or as “Ockham’s razor.” It is that if you have two valid proposals to model a situation, all things being equal, usually the simplest is the best, not to say that the simplest theories are always the best option. However, in the context of teaching experimental design, when the basis for understanding more complete theories

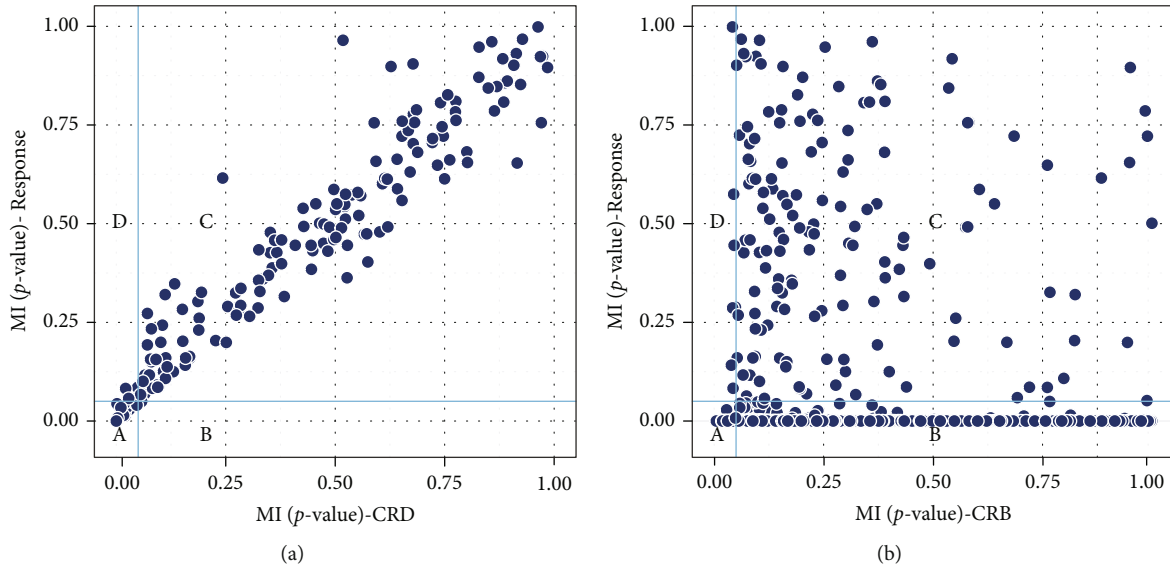


FIGURE 4: Distribution of the  $p$  values of the Moran index of the residuals of the CRD (a) and CRB (b) models against the  $p$  values of the same index but using the response variable.

and methods is not part of the course syllabus, the simplest solution may undoubtedly be the most convenient, for its ease in interpretation and explanation [33].

Once the positive effect of blocking in the modeling process was evidenced, this result was used to discuss another aspect of interest associated with the  $p$  value that is usually obtained automatically in statistical computer programs such as RStudio, where after AOV with blocking appears a  $p$  value associated with this effect. This value has generated important interpretation controversies, since the user thinks that there is some statistical hypothesis associated with it so that this default value can be obtained. The truth is that there are no valid block effect tests for the RCB model [34]. So why does the  $p$  value appear in the table? The answer is simple; the output is associated with a two-way model for fixed or random effects with a test for its effect. In this way, for the  $p$  value for a factor (one way) and a  $p$  value (for another way), the interpretation of the  $p$  value is given to the factor and not to the block. This is simply a restriction of randomization [35].

As the  $p$  values in our case are uninterpretable, the blocking efficiency correction proposed by Lentner et al. [23] was used to evaluate the blocking efficiency not only in the suppression of spatial dependence but also to incorporate updated statistical proposals and thus avoid epistemic errors, statistical slipups, analytical aberrations, or mathematical miscues by interpreting the default results in the computer programs [36].

The automatically generated  $p$  value of the AOV that is run in RStudio was extracted with the function AOV (analysis of variance) for the blocks, the  $H$  statistic was calculated, and a scatter diagram was constructed. Clearly, the vast majority of the  $H$  statistics were greater than unity (955 of 1000 simulations) (Figure 5), highlighting the result of 120, since in these cases the effect of the

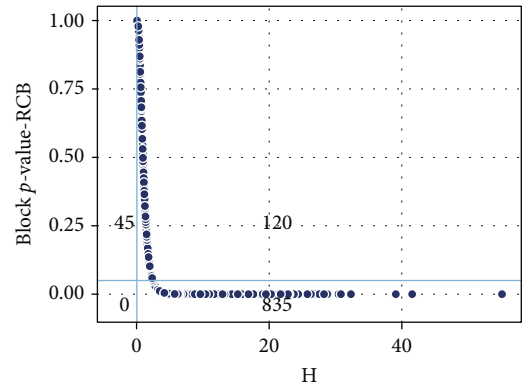


FIGURE 5: Distribution of the  $p$  values of the AOV and the  $H$  statistic of the blocking efficiency.

blocks could have been eliminated by improperly using their  $p$  value, when the  $H$  statistic indicates blocking efficiency. In this way, blocking continues to provide advantages in the modeling and design of experiments [37]. Of the 165 simulations that generated a  $p$  value greater than 5% where the propensity to eliminate the effect of blocks could have been generated, 72.7% of these simulations showed blocking efficiency, which may make the researcher not inclined to eliminate the effect, and thus, the role of blocking in spatial dependence suppression may be maintained.

To give greater validity to the simulations and not to create a confusing effect by suggesting the advantages of blocking in minimizing the effects of spatial dependence, when the results could be due more to the nonfulfillment of other assumptions, it is well known that the departure from normality of residuals and homoscedasticity of treatment variances can lead to nonfulfillment of other assumptions such as independence, for which two tests considered with high

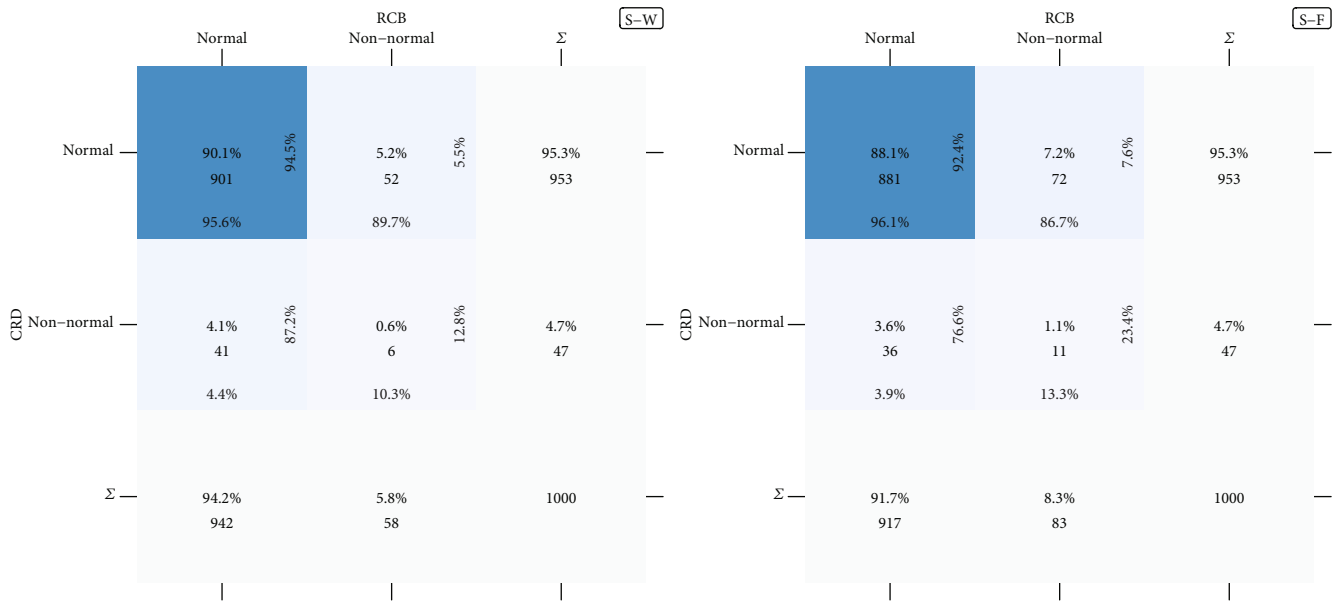


FIGURE 6: Confusion matrices associated with the  $p$  values from the AOV on CRD and RCB models for the S-W and S-F normality tests.

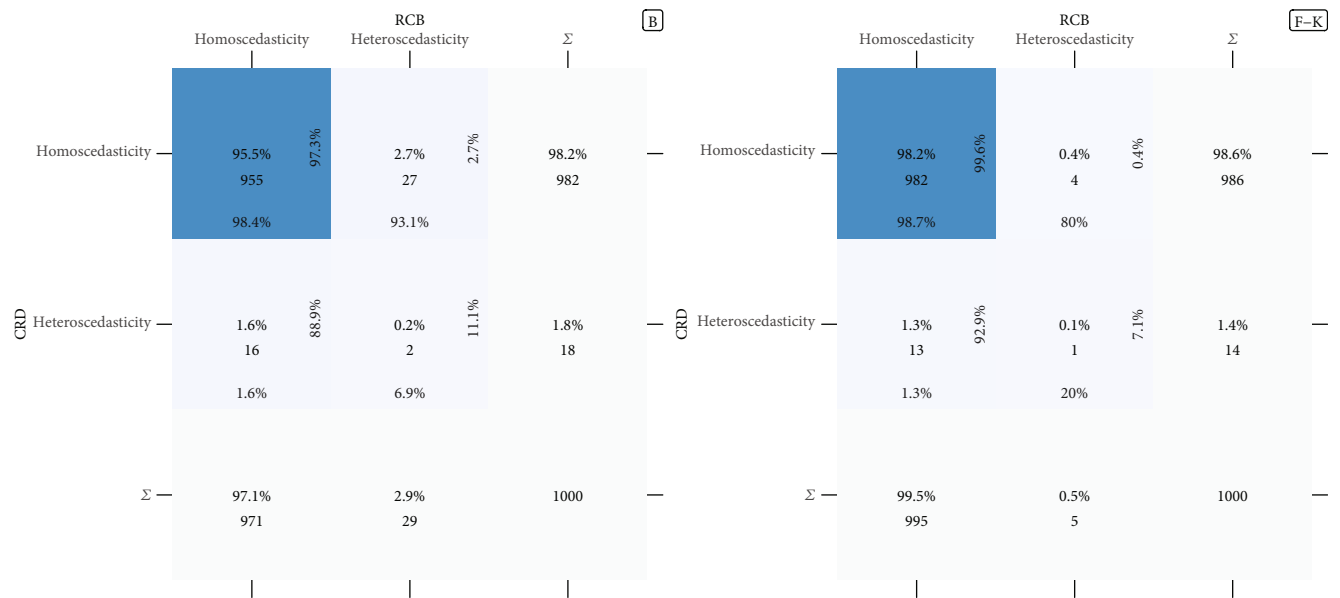


FIGURE 7: Confusion matrices associated with the  $p$  AOV values on CRD and RCB models for the tests of homogeneity of variances of B and F-K.

power were applied to detect normality of residuals, Shapiro-Wilk (S-W) and Shapiro-Francia (S-F), as well as the homogeneity of variances by power Bartlett (B) and that of Fligner-Killeen (F-K) for being robust to deviations from normality [38–40]. The results were quite satisfactory both in the normality tests (Figure 6) and in the homogeneity of variances (Figure 7).

To validate the simulations performed, we used the databases of the *agridat* library of R [22] (Table 1), which has available data of a series of real experiments developed in the last decades. This is one of the purposes of this library and is precisely the one that has been given in this research, to be able to contrast the facts with real data with what others could do about the same data set. From these data-

TABLE 1:  $p$  values associated with MI in the CRD and RCB models in the *agridat* bases of R.

Database	MI-CRD ( $p$ value)	MI-RCB ( $p$ value)
Rothamsted.brussels	$5.67e-05$	0.47
Mead.strawberry	$2.00e-16$	0.47
Rothamsted.oats	$2.00e-16$	0.48
Federer.tobacco	$6.85e-13$	$1.09e-5$

bases, some designs with blocks were selected with the respective spatial coordinates for which the respective weight matrix was calculated and the AOV of the RCB and CRD models were run, from which the residual vectors were

extracted and the test associated with Moran's index was used to test the hypothesis of independence of the residuals. In all cases, the  $p$  value of the MI was obtained to study its absolute change. In all selected cases, spatial dependence decreased, and in some examples, the data showed evidence against spatial dependence. Undoubtedly, in some of the cases tested, the spatial dependence did not disappear; however, the analysis was carried out using different strategies [41].

The data from real and simulated experiments show the advantages of blocking in the suppression or mitigation of spatial dependence, and above all, the use of the  $H$  statistic is more appropriate to maintain the effect of the blocks within the model when its value is greater than unity even though the  $p$  value of blocking may give the false belief of the necessity to exclude it from the model.

#### 4. Conclusions

The benefits of blocking must be weighed against any consideration, including the reduction of degrees of freedom for error, since, as demonstrated by the simulations, it had a positive effect on mitigating and even eliminating the effect of spatial dependence that might be present in some response variables associated with some edaphic property. Undoubtedly, although blocking apparently did not seem desirable from the point of view of the  $p$  value obtained in the default analysis of variance table, this practice should be avoided, and instead, the efficiency of blocking with the  $H$  statistic should be used since the misused  $p$  value of blocking presented approximately a 3 : 1 ratio of highlighting the importance of leaving the effect of blocks within the model.

Appropriate blocking by considering some edaphic property or by spatial zoning can make standard linear modeling under the assumption of independence of the residuals of a design model meeting the assumptions for using analysis of variance as a statistical technique to compare the effects of treatments in the presence of blocks.

The strategy presented is appropriate for reviewing another important assumption in modeling using analysis of variance, since a result that has demonstrated efficient blocking may indirectly fit the data to meet the assumption of independence in the residuals and thus may be making the technique valid.

#### Data Availability

The data used come from two sources, namely, Monte Carlo simulations for which the respective pseudocodes are presented and the agridat library of R in the case of real data, and their names and extensions are given in the body of the document.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

The authors are very thankful to Sinha Surendra Prasad, professor at the Universidad de los Andes in Venezuela. The funding came from the employer Universidad Nacional de Colombia.

#### References

- [1] M. Acutis, B. Scaglia, and R. Confalonieri, "Perfunctory analysis of variance in agronomy, and its consequences in experimental results interpretation," *European Journal of Agronomy*, vol. 43, pp. 129–135, 2012.
- [2] K. K. Mann, A. W. Schumann, T. A. Obreza, M. Teplitski, W. G. Harris, and J. B. Sartain, "Spatial variability of soil chemical and biological properties in Florida citrus production," *Soil Science Society of America Journal*, vol. 75, no. 5, pp. 1863–1873, 2011.
- [3] M. A. Wani, N. Shaista, and Z. M. Wani, "Spatial variability of some chemical and physical soil properties in Bandipora district of Lesser Himalayas," *Journal of the Indian Society of Remote Sensing*, vol. 45, no. 4, pp. 611–620, 2017.
- [4] R. A. Fisher, "The arrangement of field experiments," *Journal of the Ministry of Agriculture*, vol. 33, pp. 503–513, 1926.
- [5] J. P. Queen, G. P. Quinn, and M. J. Keough, *Experimental Design and Data Analysis for Biologists*, Cambridge university press, 2002.
- [6] D. Raghavarao and L. V. Padgett, *Block designs: analysis, combinatorics and applications*, vol. 17 of Series on Applied Mathematics, World Scientific, 2005.
- [7] A. Dean, M. Morris, J. Stufken, and D. Bingham, *Handbook of Design and Analysis of Experiments*, vol. 7, Chapman & Hall CRC Press, 2015.
- [8] J. Lawson, *Design and Analysis of Experiments with R*, vol. 115, Chapman & Hall CRC Press, 2014.
- [9] C. A. Gotway and N. A. C. Cressie, "A spatial analysis of variance applied to soil-water infiltration," *Water Resources Research*, vol. 26, no. 11, pp. 2695–2703, 1990.
- [10] M. D. Casler, "Fundamentals of experimental design: guidelines for designing successful experiments," *Agronomy Journal*, vol. 107, no. 2, pp. 692–705, 2013.
- [11] F. A. Graybill, *Theory and Application of the Linear Model*, Duxbury Press, 1976.
- [12] M. A. Islam, R. I. Chowdhury, M. A. Islam, and R. I. Chowdhury, "Covariate-dependent Markov models," in *Analysis of Repeated Measures Data*, pp. 51–66, Springer, Singapore, 2017.
- [13] R. Mead, S. G. Gilmour, and A. Mead, *Statistical Principles for the Design of Experiments: Applications to Real Experiments*, vol. 36, Cambridge University Press, 2012.
- [14] D. C. Montgomery, *Introduction to Statistical Quality Control*, Wiley Global Education, 2012.
- [15] R. Christensen, *Plane Answers to Complex Questions: The Theory of Linear Models*, Springer International Publishing, 2020.
- [16] A. R. Hoshmand, *Design of Experiments for Agriculture and the Natural Sciences Second Edition*, Chapman & Hall CRC Press, 2006.
- [17] T. P. Ryan and J. P. Morgan, "Modern experimental design," *Journal of Statistical Theory and Practice*, vol. 1, no. 3–4, pp. 501–506, 2007.



- [18] N. M. Madhyastha, S. Ravi, and A. S. Praveena, *A First Course in Linear Models and Design of Experiments*, Springer, Singapore, 2020.
- [19] K. G. Russell, *Design of Experiments for Generalized Linear Models*, Chapman & Hall CRC Press, 2018.
- [20] G. Arbia, *A Primer for Spatial Econometrics with Applications in R*, Springer Nature, 2014.
- [21] H. Chen, D. L. Reuss, D. L. S. Hung, and V. Sick, "A practical guide for using proper orthogonal decomposition in engine research," *International Journal of Engine Research*, vol. 14, no. 4, pp. 307–319, 2013.
- [22] K. Wright, "agridat: agricultural datasets," 2021, R package version, 1.18.
- [23] M. Lentner, J. C. Arnold, and K. Hinkelmann, "The efficiency of blocking: how to use  $ms(blocks)/ms(error)$  correctly," *The American Statistician*, vol. 43, no. 2, pp. 106–108, 1989.
- [24] D. K. Cassel, O. Wendroth, and D. R. Nielsen, "Assessing spatial variability in an agricultural experiment station field: opportunities arising from spatial dependence," *Agronomy Journal*, vol. 92, no. 4, pp. 706–714, 2000.
- [25] R. D. Landes, K. M. Eskridge, P. S. Baenziger, and D. B. Marx, "Are Spatial Models Needed with Adequately Blocked Field Trials?," in *Conference on Applied Statistics in Agriculture*, Department of statistics Kansas State University, USA, 2001.
- [26] A. Venter, M. F. Smith, D. J. Beukes, A. S. Claassens, and M. van Meirvenne, "Spatial variation of soil and plant properties and its effects on the statistical design of a field experiment," *South African Journal of Plant and Soil*, vol. 26, no. 4, pp. 231–236, 2009.
- [27] C. C. Tsui, C. C. Tsai, and Z. S. Chen, "Soil organic carbon stocks in relation to elevation gradients in volcanic ash soils of Taiwan," *Geoderma*, vol. 209-210, pp. 119–127, 2013.
- [28] H. Rezaei, A. A. Jafarzadeh, A. Alijanpour, F. Shahbazi, and K. V. Kamran, "Effect of slope position on soil properties and types along an elevation gradient of Arasbaran forest, Iran," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 5, no. 6, pp. 449–456, 2015.
- [29] L. Polidori and M. el Hage, "Digital elevation model quality assessment methods: a critical review," *Remote Sensing*, vol. 12, no. 21, pp. 3522–3536, 2020.
- [30] C. Zhang, L. Luo, W. Xu, and V. Ledwith, "Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland," *Science of the Total Environment*, vol. 398, no. 1–3, pp. 212–221, 2008.
- [31] Q. Liu, W. J. Xie, and J. B. Xia, "Using semivariogram and Moran's I techniques to evaluate spatial distribution of soil micronutrients," *Communications in Soil Science and Plant Analysis*, vol. 44, no. 7, pp. 1182–1192, 2013.
- [32] R. Awal, M. Safeeq, F. Abbas et al., "Soil physical properties spatial variability under long-term no-tillage corn," *Agronomy*, vol. 9, no. 11, p. 750, 2019.
- [33] N. Lazar, "Ockham's razor," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 2, pp. 243–246, 2010.
- [34] G. W. Oehlert, "Comparing models: the analysis of variance," in *A First Course in Design and Analysis of Experiments*, pp. 44–52, WH Freeman and Co., New York, NY, 2000.
- [35] B. Jones, *Avoiding Data Pitfalls: How to Steer Clear of Common Blunders When Working with Data and Presenting Analysis and Visualizations*, John Wiley & Sons, 2019.
- [36] M. S. McIntosh, "Can analysis of variance be more significant?," *Agronomy Journal*, vol. 107, no. 2, pp. 706–717, 2015.
- [37] M. Krzywinski and N. Altman, "Analysis of variance and blocking," *Nature Methods*, vol. 11, no. 7, pp. 699–700, 2014.
- [38] H. R. Lindman, *Analysis of Variance in Experimental Design*, Springer Science & Business Media, 2012.
- [39] J. Gross and U. Ligges, "nortest: tests for normality," 2015, R package version 1.0-4.
- [40] A. G. Bunn, "A dendrochronology program library in R (dplR)," *Dendrochronologia*, vol. 26, no. 2, pp. 115–124, 2008.
- [41] M. O. Grondona and N. Cressie, "Using spatial considerations in the analysis of experiments," *Technometrics*, vol. 33, no. 4, pp. 381–392, 1991.