

Estimation and computational evaluation of the coefficient of intraspecific competition in edges in the context of linear models

Estimación y evaluación computacional del coeficiente de competencia intraespecífica en bordes en contexto de modelos lineales

Estimativa e avaliação computacional do coeficiente de competição intraespecífica em arestas no contexto de modelos lineares

Aquiles Enrique Darghan Contreras*  

Carlos Armando Rivera Moreno  

Nair Jose González Sotomayor  

Jose Luis Castellanos Coronel  

Rev. Fac. Agron. (LUZ). 2022, 39(1): e223918

ISSN 2477-9407

DOI: [https://doi.org/10.47280/RevFacAgron\(LUZ\).v39.n1.18](https://doi.org/10.47280/RevFacAgron(LUZ).v39.n1.18)

Department of Agronomy, Universidad Nacional de Colombia, Bogotá, Colombia. Postal code: 111321

Received: 24-09-2021

Accepted: 13-11-2021

Published: 20-02-2022

Crop Production

Associate editor: Dr. Jorge Vilchez-Perozo

Keywords:

Border effect
Competition coefficient
Reparameterization,
Monte Carlo Simulation

Abstract

In experimental trials, it is usually of interest to give special regard to the response of experimental units at the edges, since it is well known that the performance of these can be greater than that of the rest of the units due to having less competition from neighboring units. When treatments are available, it is possible that the differences in the mean crop response are attributable to the edge effect. Therefore, it is important to consider the edge or not in the modeling process. In this case, using the Kempton-Besag model and the reparameterization of the model, the intraspecific competition coefficient was estimated through least quadratic estimation that in this case was associated with the edge effect. Its distributional pattern was studied using Monte Carlo simulation. Simulated variance analyses were carried out to see the distributional effect of the F-statistic in the presence of the edge effect as a form of spatial dependence that was evaluated with the Moran index. The coefficient associated with the edge effect showed a clear normal distribution in all the considered edge scenarios. The sign of the coefficient and the confidence intervals generated made it possible to discriminate the presence/absence of edge effect. In addition, a method was proposed to allow a user to mitigate the fuzziness that may result from the point estimate of the coefficient. This procedure can be used in other neighborhood patterns and other design models of importance in agricultural research.

Resumen

En los ensayos experimentales suele ser de interés prestar especial atención a la respuesta de las unidades experimentales en los bordes, ya que es bien sabido que el rendimiento de estas puede ser superior al del resto de unidades por tener menor competencia de unidades vecinas. En un experimento, donde se tienen tratamientos, es posible que las diferencias en la respuesta media de las variables medidas sean atribuibles al efecto de borde. Por lo tanto, es importante considerar el borde o no en el proceso de modelado. En este caso, utilizando el modelo de Kempton-Besag y la reparametrización del modelo, se estimó el coeficiente de competencia intraespecífica mediante estimación mínima cuadrática que en este caso se asoció con el efecto de borde. Su patrón de distribución se estudió mediante simulación Monte Carlo. De cada análisis de varianza con datos simulados se extrajo el valor del estadístico F y se representó su comportamiento distribucional para relacionarlo con la ausencia o presencia del efecto del borde como una forma de dependencia espacial, la cual se confirmó desde el punto de vista estadístico con el índice de Moran. El coeficiente asociado con el efecto de borde mostró una clara distribución normal en todos los escenarios de borde considerados. El signo del coeficiente y los intervalos de confianza generados permitieron discriminar la presencia/ausencia de efecto de borde. Además, se propuso un método para permitir al usuario mitigar la falta de claridad que puede resultar de la estimación puntual del coeficiente. Este procedimiento se puede utilizar en otros patrones de vecindad y otros modelos de diseño de importancia en la investigación agrícola.

Palabras clave: efecto borde, coeficiente de competición, reparametrización, simulación Monte Carlo.

Resumo

Em ensaios experimentais, geralmente é de interesse dar atenção especial à resposta das unidades experimentais nas bordas, uma vez que é bem conhecido que o desempenho destas pode ser maior do que o do resto das unidades devido à menor competição de unidades vizinhas. Quando os tratamentos estão disponíveis, é possível que as diferenças na resposta média da cultura sejam atribuídas ao efeito de borda. Portanto, é importante considerar a borda ou não no processo de modelagem. Nesse caso, utilizando o modelo de Kempton-Besag e a reparametrização do modelo, o coeficiente de competição intraespecífico foi estimado por meio da estimativa do mínimo quadrático que, neste caso, foi associada ao efeito de borda. Seu padrão de distribuição foi estudado usando simulação de Monte Carlo. Análises de variância simuladas foram realizadas para verificar o efeito distributivo da estatística F na presença do efeito de borda como uma forma de dependência espacial que foi avaliada com o índice de Moran. O coeficiente associado ao efeito de borda apresentou uma distribuição normal clara em todos os cenários de borda considerados. O sinal do coeficiente e os intervalos de confiança gerados permitiram discriminar a presença / ausência do efeito de borda. Além disso, foi proposto um método para permitir ao usuário mitigar a imprecisão que pode resultar da estimativa pontual do coeficiente. Este procedimento pode ser usado em outros padrões de vizinhança e outros modelos de projeto importantes na pesquisa agrícola.

Palabras chave: efeito de borda, coeficiente de competição, reparameterização, Simulação de Monte Carlo

Introduction

It has been 100 years since the publications of Arny (1922), describing the edge effect in agricultural experiments and where proposing ways to avoid it. Even the term competition was used to associate it with the performance of off-edge row performance with the rows at the edge, recognizing since then that removing the rows from the edge could in certain situations be an unnecessary operation, something that recent studies endorse (Romani *et al.*, 1993; Kuemmel, 2003). However, others recommend the elimination of edges depending on the crop, with the purpose of making a better estimation of the yield (Gałęzewski *et al.*, 2013). According to Keddy (2001). The definition presents a challenge since it is an example of a generalized phenomenon that can occur in quite diverse conditions. Current research focuses on an autoregressive competition model (Ord, 1975) used by several authors to consider such an effect and for which various proposals are put forth (Connolly *et al.*, 1993). But this time, rather than estimating the parameters of the model associated with the effects considered, computational analysis was done on the least-squared estimator of the Kempton-Besag competition model. These generated scenarios associated with situations described for performance when there is no edge effect and when it is perceived not only at the most exposed edge but also at the two outermost edges of the batch, as well as in partial edge situations (Besag and Kempton, 1986; Shukla and Subrahmanyam, 1999). With the calculation of the Moran index, the spatial dependence of the residuals of the model was studied. Computationally the effect obviating the spatial dependence of the residuals on the analysis of variance for the two-way model was shown. Finally, with the simulated construction of the confidence interval for the competition estimator, the presence/ absence of competition could be inferred. This is something that many research usually describe in their trials (Phillips *et al.*, 2020).

Materials and methods

This section considers a two-way classification model including its reparametrization from the perpendicular projection operators approach using the estimator obtained through least-squared estimation of the coefficient of competition associated with the edge effect while considering a series of closed-edge scenarios as partial edges and off-edge. This suggests a greater response in units located in any edge modality.

Model Structure and estimation approach

The Besag-Kempton autoregressive model based on the argument that the random response of neighboring units affects or competes is written as:

$$Y = X\beta + \kappa WY + \varepsilon, \quad (1)$$

where Y is a random vector of length n and denotes the response of the n units, X is a known design matrix of dimension $n \times p$, β is the vector of unknown parameters of length p and consists of the effects of the model (treatments and blocks), κ is the unknown competition coefficient that was associated with the edge effect, W is the matrix of weights of dimension $n \times n$ that in this case were associated with the inverses of the Euclidean distances between the units with the assumption that the non zero weights of the neighbors on a particular unit yield a Markov matrix, and the effect of unit on itself is zero, yielding a square matrix of the Hollow type (In our case, the weights matrix was standardized). Finally, independence and identical normal

distribution with zero mean and variance $\sigma^2\mathbf{I}$ was assumed for the vector of errors, \mathbf{I} being an identity matrix of dimension $n \times n$. The design matrix was associated with a two-way model, to which lateral conditions were imposed. The model is written as follows:

$$y_{ij} = \mu + \tau_i + \delta_j + \varepsilon_{ij} \quad i = 1, 2, \dots, t; \quad j = 1, 2, \dots, b, \quad (2)$$

with τ_i as the effect of the i -th treatment and δ_j as the effect of the j -th block that in matrix form is written as $\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$. In model (2) there is a deficiency in the range of magnitude $p-q=2$, with $p=t+b+1$ and $q=t+b-1$. To cover this deficiency, the usual non-estimable functions in this design model were used, that is, $\sum_{i=1}^t \tau_i = 0$ and $\sum_{j=1}^b \delta_j = 0$. Under the hypotheses $H_0: \kappa=0$ of the model in (1) we have the linear model $\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ and solving the problem of the range of \mathbf{X} with the imposition of lateral conditions we proceeded to the estimation of the parameters. For model (1), the logarithm of the likelihood is given by.

$$l(\boldsymbol{\beta}, \sigma^2, \kappa) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) + n \ln |\mathbf{I} - \kappa \mathbf{W}| - \frac{\|(\mathbf{Y} - \kappa \mathbf{W}\mathbf{Y}) - \mathbf{X}\boldsymbol{\beta}\|^2}{2\sigma^2}, \quad (3)$$

with $|\mathbf{I} - \kappa \mathbf{W}|$ like the Jacobian transformation.

One use of the lateral conditions is the reparameterization of the model, where the Besag- Kempton incomplete-range model is transformed into a full-range model. This simplifies the estimation procedure and allows obtaining unique estimators (Chistensen, 2011). The method on model (1) is described below:

Let $\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\kappa\mathbf{W}\mathbf{Y}+\boldsymbol{\varepsilon}$ be the model and $\mathbf{L}\boldsymbol{\beta}=\mathbf{0}$ the matrix expression associated with the lateral conditions, where \mathbf{L} is a matrix of dimension $l \times p$, which is linked to the two-way model $l=p-q$. $\mathbf{L}\boldsymbol{\beta}$ is a set of non-estimable functions. If $\mathbf{L}\boldsymbol{\beta}=\mathbf{0}$, then $\boldsymbol{\beta}'\mathbf{L}'=\mathbf{0}$, so that $\boldsymbol{\beta}'$ belongs to the orthogonal complement onto column space of \mathbf{L}' ($\boldsymbol{\beta}' \in C(\mathbf{L}')$). To identify the reparametrized model it is necessary to select a matrix \mathbf{Z} (Supplemental Material) such that it has the same $C(\mathbf{L}')$; and, in this way, the vector $\boldsymbol{\beta}$ belongs to $C(\mathbf{Z})$, that is, $\boldsymbol{\beta}=\mathbf{Z}\boldsymbol{\gamma}$, for some vector $\boldsymbol{\gamma}$. By substituting $\boldsymbol{\beta}=\mathbf{Z}\boldsymbol{\gamma}$ in model (1) the reparameterized model is obtained:

$$\mathbf{Y}=\mathbf{X}\mathbf{Z}\boldsymbol{\gamma}+\kappa\mathbf{W}\mathbf{Y}+\boldsymbol{\varepsilon}, \quad (4)$$

and doing the reparameterized model of Besag-Kempton is finally written as:

$$\mathbf{Y}=\mathbf{P}\boldsymbol{\gamma}+\kappa\mathbf{W}\mathbf{Y}+\boldsymbol{\varepsilon}, \quad (5)$$

From model (5) we can obtain the logarithm function of the likelihood as well as the estimators for $\boldsymbol{\gamma}$ and σ^2 for a restricted model under H_0 . However, as the estimator for κ in this estimation mode did not present a closed solution (6), the least square estimation was used.

$$\mathbf{y}'\mathbf{W}'(\mathbf{I} - \hat{\kappa}\mathbf{W} - \mathbf{M})\mathbf{y} = \text{tr}\{(\mathbf{I} - \hat{\kappa}\mathbf{W})^{-1}\mathbf{W}'\mathbf{y}'(\mathbf{I} - \mathbf{M})\mathbf{y}\} \quad (6)$$

where $\mathbf{M}=\mathbf{P}(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'$ is the perpendicular projection matrix on $C(\mathbf{P})$ and tr represents the trace. In the case of the least-square estimation, for which no distributional assumption is required, we seek to minimize $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$, that is,

$$\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} = \sum_{i=1}^n \hat{\varepsilon}_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \mathbf{P}\hat{\boldsymbol{\gamma}} - \hat{\kappa}\mathbf{W}y_i)^2 \quad (7)$$

Note that the predicted value \hat{y}_i estimates $E(y_i)$ and not y_i , however, it is usually written as \hat{y}_i . To find the estimator for $\hat{\kappa}$ from (7) we differentiated with respect to κ and set it equal to zero and obtained the closed solution for $\hat{\kappa}$ that was used to detect the edge effect of agronomic trials and computationally studied its distributional pattern based on simulation Monte Carlo. The estimator turned out to be:

$$\hat{\kappa} = \mathbf{y}'(\mathbf{I} - \mathbf{M})\mathbf{W}\mathbf{y} / \mathbf{y}'\mathbf{W}(\mathbf{I} - \mathbf{M})\mathbf{W}\mathbf{y} \quad (8)$$

Once the design matrix \mathbf{X} , the matrix \mathbf{Z} for the reparameterization of the model, and the neighborhood pattern \mathbf{W} were defined, the functions were created in the RStudio software to perform the Monte Carlo simulations of all the scenarios. This involved randomizing the treatments so that the response vector \mathbf{Y} associated with performance did not show the edge effect. Higher response measures were generated in the treatments that resulted in the complete edge, and a double row of higher response was generated in the two outer and complete edges. Unilateral, bilateral, and trilateral partial borders with edge effect were generated. For the blocking reason, the reader can abstract to some situation that can be linked with simplicity and that causes some randomization restriction to create the blocks, such as the origin of the seed.

Finally, whatever the estimation methodology is used in the estimation of parameters, for example, the lateral conditions, the current interest is related to the competition coefficient associated with the edge effect in this case, so there is no It is of interest to search for information on the edge effect with different proposals for estimating the competition coefficient, since it is unprecedented as an alternative to study the edge effect.

Scope of the methodology

The usefulness of the statistical-computational approach is subject to a previous exploration of the data set to detect the possible presence of the edge effect, since the simple obtaining of the $\hat{\kappa}$ statistic is not directly associated to the absence or presence of the edge effect but rather to the presence or absence of the competition effect such as the nature of the model used for the construction of the statistic.

Motivating dataset

Figure 1 illustrates the arrangement of the treatments and blocks in the study area. The label associated with the treatment, block, and repetition appears. The number above each label represents the response as it was recorded in the field and the number to the right represents the position in the response vector due to randomization. In the case of the simulated responses, they were generated by Monte Carlo method to obtain all the scenarios described above. For all the scenarios where the response was generated it was guaranteed that the treatment means were not different. Therefore, once the analysis of variance (aov) of the two-way model without competition or edge effect was run (standard linear model in 2), no differences in the treatment means were seen. Therefore, the distribution of the F-statistic of the 3000 simulations followed the characteristic F-distribution under the null hypothesis and thus could detect any departure from this distribution not because of the differences in means but because of the spatial dependence induced by the edge effect.

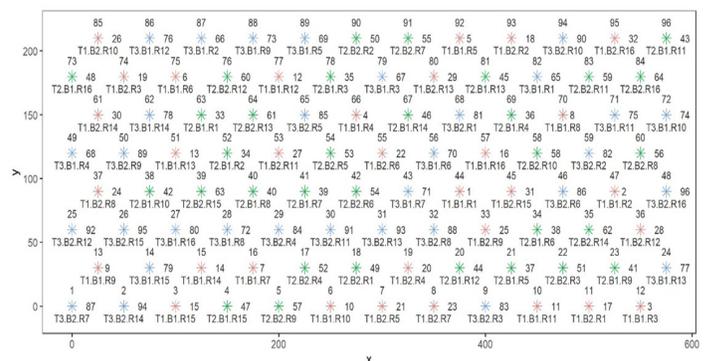


Figure 1. Distribution of blocks and treatments on the experimental units simulated.

The following figures show different patterns of the different scenarios developed that are associated with edge effect in agricultural experiments. The first group of patterns (Fig. 2) distributes the response (value causing 1) absent edge effect (left), 2) complete outer edge around the entire convex contour (center) and 3) two outer edges for the two successive convex contours (right).

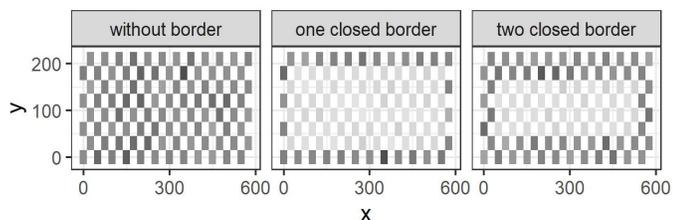


Figure 2. Distribution of responses in the three main border patterns.

Figure 3 shows all possible cases to illustrate the unilateral edge effect, for which a higher yield was associated with the more accentuated gray points.

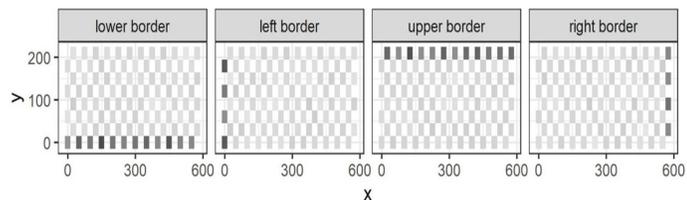


Figure 3. Distribution of responses on one-sided edges.

Figure 4 incorporates the successive or separate bilateral border structures with six possible cases. It would be expected that if the statistics $\hat{\kappa}$ detects the edge effect, these structures show a clearer effect than those of the previous figure.

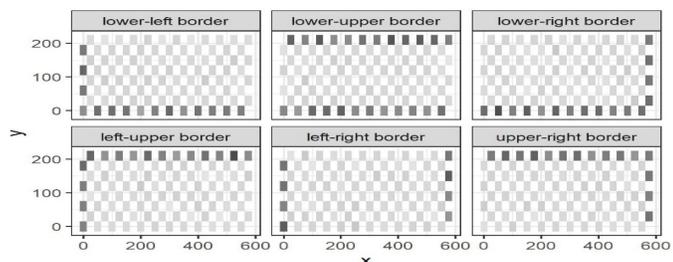


Figure 4. Distribution of responses on the bilateral border with the highest yields

Finally, the trilateral border patterns that are the most similar structures to the usual simple border in agricultural trials, thus, a behavior in the $\hat{\kappa}$ statistic is expected to be much more like that obtained in the presence of a simple edge (Fig. 5).

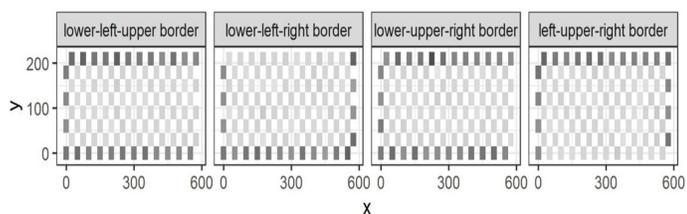


Figure 5. Distribution of responses at the trilateral borders.

Mathematical classification and construction of simulated density

Let Ψ the space with points $\Psi_i = A = \bigcup_{j=1, i=j}^2 \Psi_j$ associated with the combination of treatments and blocks in an experimental design ($\Psi = \Psi_1 \cup \Psi_2 = \Psi_1 \cup \Psi_2$) and let $S_i (i=1,2)$ the i -th subset of Ψ containing the points identifying the outer single edge or outer double edge in a two-dimensional spatial array. The indicator function S_i denoted by $I_{S_i}(\cdot)$, is the function with domain Ψ and codomain equal to the set consisting of two real numbers 0 and 1 defined by $I_{S_i}(\Psi_i) = 1$ if $\Psi_i \in S_i$ or $I_{S_i}(\Psi_i) = 0$ if $\Psi_i \notin S_i$; $I_{S_i}(\cdot)$ clearly indicates S_i .

With the previous description, the two situations are mathematically formalized and will be illustrated when a user wants to know if the data of his experiment show the edge effect. The selection of S_i depends on the distribution of the units in the field and as in our case it was a staggered planting with a rectangular lot with a defined number of rows and columns, the points with the minimum common x-coordinate and maximum common x-coordinate as well as the minimum and maximum common y-coordinate were identified. For the single edge effect case $I_{S_1}(\Psi_i) = 1$ if $\Psi_i \in S_1$ indicates the combination of blocks and treatments on this, the outermost edge. For the case of two outer edges, for the purposes of the algorithm (see supplementary material) all elements of S_1 in Ψ , i.e., we obtain $\Psi - S_1 = \Psi^* (\Psi \setminus S_1)$ and again identify the points with the minimum common x-coordinate and maximum common x-coordinate as well as the minimum and maximum common y-coordinate but in Ψ^* , being Ψ_2 the selected points on this inner edge. To count the two outer edges we do $\Psi_1 \cup \Psi_2$, in this way, $I_{S_2}(\Psi_2) = 1$ if $\Psi_2 \in S_2$.

Once the single and double external edges have been indicated, we proceed to propose a computational method that the user can use to have clarity of the edge effect. It should be remembered that an experimental test only generates a measure of the $\hat{\kappa}$ coefficient that we have associated with the edge effect, so it would not be simple to clearly discriminate the presence of the edge effect. To avoid this problem, we use RStudio's `simulate()` function (R Core Team, 2021) to generate a simulated density for $\hat{\kappa}$ (found by the researcher) using RStudio's `lm()` function for the formula: response $\sim S_1$ and do the desired simulations to construct the density of the outermost edge. The element vector in containing zeros and ones of length equal to the evaluated response is simulated over a user-defined number of simulations, in our case 500 simulations were used and then we plotted the density, and we can notice the pattern that certainly tells us if it is contained to zero or not to decide if we have the effect associated with the outer edge. If the formula response $\sim S_1 + S_2$ is used and the linear model is fitted, we now have the possibility to obtain the density for the outermost edges and again the edge effect considered will be clear (supplementary material).

Results and discussion

When the usual aov is used in linear modeling the technique uses spatial dependence to modify the usual inferential procedures (Gotway and Cressie 1990). In fact, the distributional assumption based on the F-curve may not be valid. By maintaining the null hypothesis of equality of means of the treatments, it is possible by simulation to see the effect in the F-distribution. Figure 6-left shows the distribution that is expected for two degrees of freedom associated with the treatments and 92 degrees of freedom for the error in the two-way model. Figures 6-center and 6-right show the

departure from this distribution when there is a clear edge effect generating spatial dependence. Some might confuse the result of these two final figures if they do not realize that the samples did not provide statistical evidence against the null hypothesis of equality of means in the simulations generated. The figure also shows the results of the Kolmogorov-Smirnov test for the adjustment for the distribution and its D statistic. Only in the case of no edge (Fig. 6-left) the distribution is precisely the F.

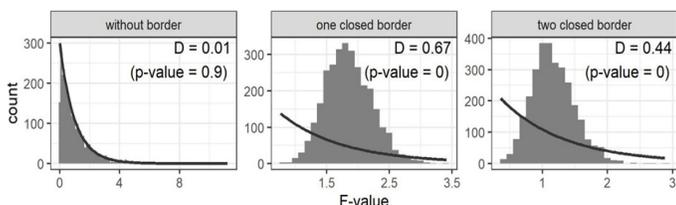


Figure 6. Distribution of the F statistic generated in the 3000 Monte Carlo simulations.

In Figure 7 the distribution of the p values of the Moran index of the cases associated with figure 2 (no border, single border, and double border) is illustrated. As can be seen in 7 (left), most of the simulations did not provide evidence against the hypothesis of spatial independence of the residuals of the analysis of variance model, with which the results of the technique are applicable if the assumptions are met. In Figures 7-center and 7-right most simulations show spatial dependence, especially in the case with two edges. The vertical lines illustratively demarcate the cut for $\alpha=5\%$ and the simulation count is also printed where the hypothesis of spatial independence is rejected. Undoubtedly, the presence of an edge is associated with a modality of spatial dependence, so results of an aov to the standard linear model may not be valid, and hence the importance of taking some remedial measure in the presence of an edge effect. This is done in the study of the speed of the root front of corn and soybeans (Ordóñez *et al.*, 2018). Although an edge effect could not always be present in our research as was the case of pasture and gramineous, some authors have concluded that both the vegetation and soil properties are influenced by the proximity to the edge (Ruwanza, 2018).

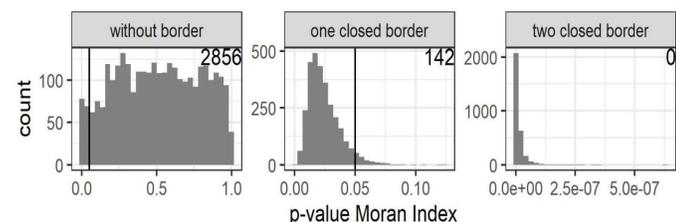


Figure 7. Distribution of the p value of the Moran index generated in the 3000 simulations.

From (5), the coefficient associated with the edge effect in model (1) was obtained by the least-squared estimation for all scenarios, showing a clear normal distribution by simulation, that was corroborated with the Shapiro-Wilks Normality test (p-values within each image in figure 8). With the mean for $\hat{\kappa}$ in the distribution of the simulations, a 95% confidence interval was constructed for κ , specifically for the cases without edge effect, single and double edge effect, and it is here where one of the most important results. Clearly the first interval (Fig. 8-left) contains zero, with which we assert the absence of edge effect, while Figures 8-center and 8-right

not only do not contain zero, but also obtain negative lower and upper limits, that is, the negative coefficient associated with the edge effect shows the effect of interest.

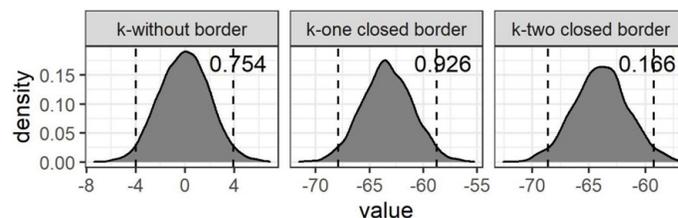


Figure 8. Distribution of the coefficient associated with the competition model or edges.

It is reasonable to think that even with many simulations for the first three cases it is possible to detect partial edges effect with the same statistic. The cases of unilateral edge effect (Fig. 9) were run, finding again the detection of unilateral partial edge but with a mean for a normal distribution closer to zero on the left, especially at the edges of shorter length (lower and upper). So, a greater distance from zero on the left shows a more defined edge effect.

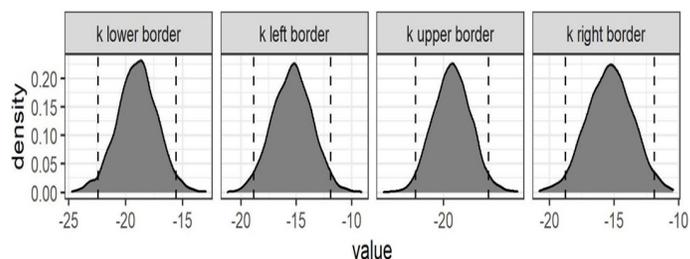


Figure 9. Distribution of the coefficient associated with the one-sided edges model.

In the cases of more defined but partial edges effect such as bilateral ones, the distribution of the statistic is maintained but its means are shifted further to the left. This defines the partial edge more clearly, especially when the edges are those with the greatest perimeter (Fig. 10). This result, which is becoming systematic, is validated with the trilateral cases of Figure 11. When the edges involve the widest perimeters, the coefficient associated with the most negative edge is obtained.

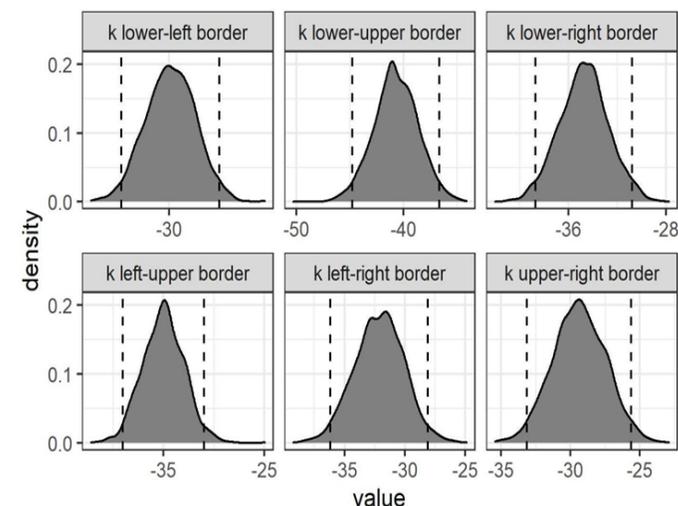


Figure 10. Distribution of the coefficient associated with the bilateral edges model.

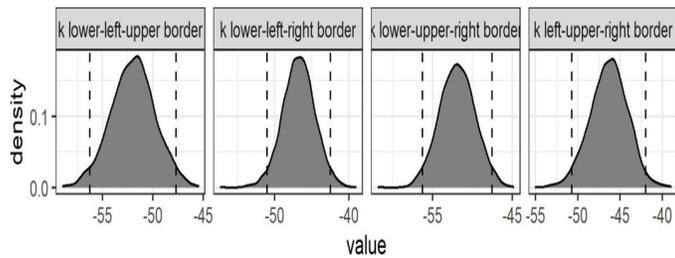


Figure 11. Distribution of the coefficient associated with the trilateral edges effect.

Application example

Using the same design structure and creating a single external edge condition in principle the researcher can obtain the $\hat{\kappa}$ estimator. As this value is not easy to use to discriminate the edge effect, we adjusted the models $\text{lm}(\text{response} \sim S_1)$ and $\text{lm}(\text{response} \sim S_1 + S_2)$ and using `simulate()` the simulations were generated. The plots of the respective densities and the simultaneous display of the estimated $\hat{\kappa}$ value allow us to conclude the presence of a single edge and not a double edge, thus, the computational strategy allows for edge detection and simplifies the analytical operations as described in (Paoella, 2019) to find the distribution of the quotient of quadratic forms associated with the proposed estimator.

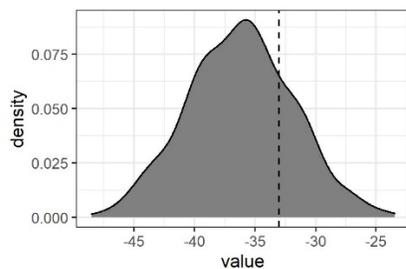


Figure 12. Simulated density plot for application example the line indicate the experimental kappa.

Conclusions

Current research uses the Besag-Kempton model and the least-quadratic estimation of parameters. From this model was possible to obtain the estimator for the competition coefficient that was associated with the edge effect. The results made it possible to visualize the effect on the distribution of the F statistic of the analysis of variance in the presence of spatial dependence evaluated using the Moran index, evident in the case of the synthetically generated edges effect.

Once a series of Monte Carlo simulations had been made for a normally distributed response, the distributional pattern of $\hat{\kappa}$ statistic associated with the edge effect was studied. In all cases, the normal distribution was evident κ in the scenarios of single, double, no border, partial borders, and the contrasting pattern of opposite borders effect. We observed that in the absence of the edge effect, this interval contained zero, while in the case of the presence of partial unilateral, bilateral, and trilateral borders.

The use of the sets of points associated with the edges of interest and the linear modeling allowed the construction of a simulated density for κ , which eliminated the ambiguity that could arise when making a point estimate for κ , so that a user of the method can easily judge the presence or absence of edge effect.

Supplemental material

Z and L matrices are available. Repository on Github of R code: The code for all the simulations in: CarlosRivera1212/BorderEffect ([github.com](https://github.com/CarlosRivera1212/BorderEffect))(edge effect coefficients)

Cited literature

- Army, A. C. (1922). Border Effect and Ways of avoiding it. *Journal of the American Society of Agronomy*, 14(7), 266-278.
- Besag, J. E. & Kempton, R. A. (1986). Statistical Analysis of field experiments using neighbouring plots. *Biometrics*, 42(2), 231-251.
- Kuemmel, Bernd (2003). Theoretical investigation of the effects of field margin and hedges on crop yields. *Agriculture, Ecosystems & Environment*, 95(1), 387-392.
- Connolly, T., Currie, I. D., Bradshaw, J. E., McNicol, J. W. (1993). Inter-plot competition in yield trials of potatoes (*Solanum tuberosum* L.) with single-drill plots. *Ann. Appl. Biol.*, 123(2), 367-377.
- Christensen, R. (2011). Plane Answers to Complex Questions. *The Theory of Linear Models. Fourth edition*. Springer, New York, Inc.
- Galęzewski, L., Piekarczyk, M., Jaskulska, I., Wasilewski, P. (2013). Border effects in the growth of chosen cultivated plant species. *Acta Scientiarum Polonorum. Agricultura*, 12(3), 3-12.
- Gotway, C. A., Cressie, N. A. C. (1990). A Spatial Analysis of Variance Applied to Soil-Water Infiltration. *Water Resources Research*, Vol. 26(11), 2695-2703.
- Keddy, P.A. (2001). Competition. *Population and Community Biology Series. 2nd Edition*. Springer Science+Business Media Dordrecht. USA.
- Ord, J. K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70, 120-126.
- Ordóñez, R. A., Castellanos, M. J., Halfeld, J. L., Halmers, M. J., Licht, M. A., Liebman, M., Dietzel, R., Fera, R. M., Ibbal, J., Puntel, L. A., Córdova, S. C., Togliatti, K., Wright, E. E., Archontoulis, S. V. (2018). Maize and soybean root front velocity and maximum depth in Iowa, USA. *Field Crops Research*, 215, 122-131.
- Paoella, M. S. (2019). Linear Models and Time Series-Analysis. Regression, ANOVA, ARMA and GARCH. John Wiley & Sons. USA.
- Phillips, X. A., Kndel, Y. R., Licht, M. A., Mueller, D. S. (2020). Estimating Soybean Radiation Use Efficiency Using a UAV in Iowa. *Agronomy*, 10(12).
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Romani, M., Borghi, B., Alberici, R., Delogu, G., Hesselbach, J., Salamini, F. (1993). Intergenotypic competition and border effect in bread wheat and barley. *Euphytica* 69(1-2):19-31
- Ruwanza, S. (2018). The Edge Effect on Plant Diversity and Soil Properties in Abandoned Fields Targeted for Ecological Restoration. *Sustainability*, 11(1), 140
- Shukla, G. K. & Subrahmanyam, G. S. V. (1999). A Note on a test and Confidence Interval for Competition and Overlap Effects. *Biometrics*, 55(1), pp. 273-276.