RBGDR

LONGITUDINAL MIXED MODELS FOR THE GROSS VALUE OF AGRICULTURAL PRODUCTION: A SPATIAL AND REGIONAL ASSESSMENT IN THE STATE OF PARANÁ-BRAZIL

MODELOS MISTOS LONGITUDINAIS PARA O VALOR BRUTO DA PRODUÇÃO AGROPECUÁRIA: UMA APRECIAÇÃO ESPACIAL E REGIONAL NO ESTADO DO PARANÁ-BRASIL

RBGDR

REVISTA BRASILEIRA DE GESTÃO E DESENVOLVIMENTO REGIONAL
V.20, Nº1, Jan-Abr/2024 | Taubaté/SP - Brasil | ISSN 1809-239x

LONGITUDINAL MIXED MODELS FOR THE GROSS VALUE OF AGRICULTURAL PRODUCTION: A SPATIAL AND REGIONAL ASSESSMENT IN THE STATE OF PARANÁ-BRAZIL

MODELOS MISTOS LONGITUDINAIS PARA O VALOR BRUTO DA PRODUÇÃO AGROPECUÁRIA: UMA APRECIAÇÃO ESPACIAL E REGIONAL NO ESTADO DO PARANÁ-BRASIL

Elizabeth Giron Cima[1] | **Márcio Paulo de Oliveira**[2]
Weimar Freire da Rocha-Junior[3] | **Miguel Angel Uribe-Opazo**[4]

Received: 10/27/2023
Accepted: 02/01/2024

[1] PhD in Agricultural Engineering (UNIOESTE).
Professor at the State University of Western Paraná.
Cascavel - PR, Brazil.
Email: egcima74@gmail.com

[2] PhD in Agricultural Engineering (UNIOESTE).
Professor at the Federal Technological University of Paraná. Toledo - PR, Brazil.
Email: marciooliveira@utfpr.edu.br

[3] PhD in Production Engineering (UFSC).
Professor at State University of Western Paraná.
Toledo - PR, Brazil.
Email: weimar.junior@unioeste.br

[4] PhD in Statistics (USP).
Professor atState University of Western Paraná.
Cascavel - PR, Brazil.
Email: miguel.opazo@unioeste.br

ABSTRACT

Longitudinal mixed linear models are important because they analyze measurements at repeated frequencies in the average evolution over time. The objective of this this work is to evaluate spatially through longitudinal mixed models, the gross value of agricultural production, considering a historical series, having as explanatory variables the quantity of poultries produced per head, quantity of pigs produced per head, quantity of cattle produced per head and milk production in thousand liters produced by municipalities for the years 2018, 2019 and 2020 in the state of Paraná-Brazil. The results showed that there is a relation between the variable of interest the gross value of agricultural production and the explanatory variables quantities of poultries produced per head, amount of pigs and amount of cattle produced per head in the municipalities that make up the state of Paraná studied in this paper, and, it is possible to identify mesoregions with the variables that explain the gross value of agricultural production distinct for the model. In the analysis of variance with the models, the variable milk production was not significant to explain the gross value of agricultural production in each year, significant at the level of 5% of probabilities. The longitudinal mixed linear model provided necessary information on the relation associated with the weights that each agricultural variable provided in the gross value of agricultural production that was analyzed in the economic scenario of the state of Paraná-Brazil.

Keywords: Productive chain. Regional Development. Restricted Maximum Likelihood. Mesoregions.

RESUMO

Os modelos lineares mistos longitudinais são importantes por analisarem medições em frequências repetidas na evolução média no tempo. O objetivo deste trabalho é avaliar espacialmente, por meio do modelo linear misto longitudinal, o valor bruto da produção agropecuária, considerando uma série histórica, tendo como variáveis explicativas, quantidade de aves produzidas por cabeça, quantidade de suínos produzida por cabeça, quantidade de bovinos produzida per capita e produção de leite em mil litros, produzidos por municípios para os anos de 2018, 2019 e 2020, no estado do Paraná-Brasil. Os resultados mostraram que há um relacionamento entre a variável de interesse valor bruto da produção agropecuária, e as variáveis explicativas quantidades de aves produzidas por cabeça, quantidade de suínos e quantidade de bovino nos municípios do estado do Paraná estudados neste trabalho, sendo possível identificar mesorregiões com as variáveis que explicam o valor bruto da produção agropecuária distintos para o modelo. Na análise de variância com os modelos, a variável produção de leite não foi significativa para explicar o valor bruto da produção agropecuária em cada ano ao nível de 5% de probabilidade, já as demais variáveis foram significativas. O modelo linear misto longitudinal forneceu informações necessárias sobre o relacionamento associado aos pesos que cada variável agropecuária forneceu no valor bruto da produção agropecuária que foi analisado no cenário econômico do estado do Paraná-Brasil.

Palavras-chave: Cadeia produtiva. Desenvolvimento Regional. Máxima Verossimilhança Restrita. Mesorregiões.

INTRODUCTION

Technological advances which embrace studies that focus on agriculture and livestock are of necessary importance to explain different scenarios, within the scope of regional, economic and agro-industrial development**.** Longitudinal mixed linear models are advanced statistical techniques that seek alternative responses for contribute to the regional development of the area under study and other regional and national areas. The state of Paraná has its main economic activity in the agricultural and livestock sector, based on large-scale production that is integrated into the complex and the agro-industrial production chain, which jointly aim at sustainable food production in a resilient way, respecting the environment and biodiversity (Cima *et al.*, 2022)**.**

The food supply associated with the increase in the growth of the global human population requires the adoption of economic, social, political and environmental sustainability principles, thus the agricultural sector plays a very important role in globalized food security, this principle extends production and consumption models that are associated with the optimization of productive resources, humanitarian protocols and green economy concepts related to the natural protection of the environment (Kalinowska *et al.*, 2022).

Currently there are studies that focus on the agro-industrial production chain and regional, rural, local and sustainable development in Paraná, associated with initiatives, programs of agro-industrial development and governance structure like a contracts for instance (Kuhn; Rocha Jr; Staduto, 2005).

The differentiated view of the present work is the use of the statistical model called a longitudinal mixed linear model that is aimed at a more detailed investigation of possible events that may occur in different agricultural scenarios, considering a historical series of original data, it aims to verify several behaviors of subgroups that share common profiles with each other. Global agricultural and food production has increased significantly in the last century, stimulated by the globalized economic scenario and characterized by population growth, cultural aspects, technological changes and production practices. There is a relevant concern about ecologically correct and safe production that adopts principles of environmental sustainability (Hemathilake; Gunathilake, 2022).

The increasing supply of agricultural production in Paraná has presented the need to better understand its variability in relation to the value of agricultural production (VBP) that is expressed in reais, which represents an estimate of the income generation of rural areas, becoming a relevant variable for monitoring the sector performance as a whole (Alves, 2022).

Recently agricultural production has grown at significant annual rates. The locational indicators of economic growth reported lack clarity regarding the information they seek to explain, thus observing the need for greater knowledge in certain regions of the state of Paraná-Brazil (Alves *et al.*, 2020). There is an increasing number of public policies and government programs (Federal and State) that value space and state initiatives (regions), because it is in these places that economic growth and development are triggered (Lima, 2020). In this sense of analysis, the agro-industrial system deserves to be better evaluated because it represents a major factor in local and regional development. The state of Paraná-Brazil is considered promising in agricultural production and that incentives to the agro-industry system should be through mechanisms that aim to optimize productive resources.

In most computer programs, this type of data is analyzed using linear models of mixed effects, which considers the estimate of an average population trajectory, parameterized in terms of fixed effects and the variation of single trajectories (individual) around this average (Herle, *et al*., 2020). This fixed effect is captured by variations and covariances of random effects specific to the study object. Currently, the focus of modeling such data is more focused on a more accurate investigation, which aims to verify if there are several typical trajectories throughout the analysis process, leading to the characterization of subgroups of individuals that share a common profile among themselves (Herle, *et al*., 2020).

Recent data studies with spatial dependence were published to show the spatial behavior of possible profiles of agricultural production (gross value of agricultural production, poultries, pigs, cattle and milk), in this sense the longitudinal mixed linear models constitute an important technique to analyze this type of data set, since this family of models makes assumptions not only about the distribution of waste, but also about the distribution of random effects (Schielzeth *et al*., 2020). Longitudinal studies with dependent structure have become central in a wide variety of research areas, which have led different researchers to be interested in the better understanding of the logic that involves the *layout* of the longitudinal study, for it is the only *design* of study that allows to relate certain measures at the beginning of the study and to changes in the condition of individuals over time (Wu, 2010). Models for such longitudinal analysis were discussed by Wu (2010). A modeling tactic is based on time series, as this series is considered agricultural production in a certain period of time (Verbeke; Molenberghs, 2005; Molenaar, 2017; Asparouhov *et al*., 2018; Ariyo *et al*., 2020). This form of modeling is centered on longitudinal data analysis in which observations are performed considering the time series. Such analyzes are based on model dynamics in which the current results are regressed in previous results (Millar, 2018; Asparouhov; Muthén, 2019).

It should be noted that studies involving longitudinal analyzes in general present great potential for explanation of the phenomena analyzed (Verbeke; Molenberghs, 2005), report that this planning is a particular case of verification of repeated measures, because this is characterized by the observation of two or more measures in each individual, however what distinguishes the longitudinal structure within the repeated measures is the form of association between occasions and individuals. Given the importance of the effects that agricultural production causes to leverage regional economic development, strategic decision-making actions are necessary in order to promote well-being, agri-food supply and the strengthening of productive chains (Lima, 2020). Thus, the objective of this this work is to evaluate spatially through longitudinal mixed models, the gross value of agricultural production, considering a historical series, having as explanatory variables the quantity of poultries produced per head, quantity of pigs produced per head, quantity of cattle produced per head and milk production in thousand liters produced by municipalities for the years 2018, 2019 and 2020 in the state of Paraná-Brazil.

REGIONAL DEVELOPMENT AND THE AGRICULTURAL SECTOR

Regional development management is necessary because it promotes public policies in favor of the agricultural sector in municipalities and regions that have a vocation for this type of economic activity (Rizardi, 2022).

Efficient public and private management stands out for regional development, in addition to contributing to the agribusiness sector, it aims at humanized work, the well-being and quality of life of rural producers, suppliers and consumers, as well as all agents who they are part of the supply chain and agroindustrialization (Silva, 2015).

Guarnieri *et al*. (2022) reports that regional, economic and social development is associated with the number of industrial organizations and private companies that operate in the industrial and agro-industrial sector, which makes this sector attractive for public policies, aiming at the regional development of regions, favored by innovations technological.

Therefore, regional development promotes initiatives that generate economic competitiveness between regions, favoring the creation of jobs, improving the income and quality of life of populations and leveraging economies (Monteiro-Neto, 2023).

MATERIAL AND METHODS

The theoretical foundation and conceptual aspects of longitudinal linear mixed models are explored in several papers (Verbeke; Molenberghs, 1997; Werbeke; Molenbergers, 2005; Wu, 2010). They are very relevant models in diagnostics and statistical applications in the areas of economics, agriculture, among others, which favor the understanding of many events that occur in nature.

The longitudinal mixed linear model according to Pusponegoro *et al*. (2017) provides flexible analysis on longitudinal data that are correlated and also allows researchers the possibility of modeling the covariance structures that portray their random effects.

Spatial studies require the use of data obtained from georeferenced matrices. Figure 1 shows the location map of the study area with the 399 municipalities of the state of Paraná-Brazil associated with their ten mesoregions. Spatial data were used on the gross value of agricultural production (VBP) [R$], quantity of poultries (QTAVES) number of cattle (QTBOV), quantity of pigs (QTSUI) with units in quantity/head and milk production (PROLE) [in thousand liters] of the years 2018, 2019 and 2020, obtained from the *Instituto Paranaense de Desenvolvimento Econômico e Social* (IPARDES, 2022). The time span from 2018 to 2020 was worked on, available by municipality and presenting georeferenced characteristics for the state of Paraná, which comprises 399 municipalities.

Figure 1 | Study Area Location Map

Source: Adapted from IBGE (2023)

The response vector corresponding to the observations obtained under t different conditions (years) for j the i-th umpteenth $(j = 1, 2, \ldots, n_i)$ research units (municipalities) and i for the umpteenth subpopulations $(i = 1, 2, \ldots, g)$ obtained by considering spatial dependence through the univariate Geary site index (C_i). The longitudinal data structure is given by the vector $t \ x \ 1$ according to Equation 1:

$$y_{ij} = (y_{ij1}, y_{ij2}, \ldots, y_{ijt})^T, \quad (1)$$

where: t are the years used in the study; j are each municipalities in the Paraná state; n_i are the number of the municipalities in each subpopulation g, where g are the subpopulations groups from the univariate Geary site index.

The set of all observations corresponding to the $n = \sum_{i=1}^{g} n_i$ observations can be represented by Equation 2:

$$Y_I = vec\{[y_{11} \quad \cdots \quad y_{gng}]\}, \tag{2}$$

where: Y_I is the vector of the observations of the dependent variable VBP, where $vec(\cdot)$ is the vector obtained from the matrix columns stacking $(\cdot)$.

The spatial dependence structure obtained by the local Geary index, which varies from 0 to 2 according to Almeida (2012) being defined by Equation 3:

$$c_i = \sum_{u=1}^{n} w_{iu} (x_i - x_u)^2 \text{, i=1,... g,} \tag{3}$$

for *i*=1,...,*g,* being $(x_i - x_j)^2$ the square of the difference between the pairs of values of the variable under study (Almeida, 2012).

The longitudinal mixed linear model (Wu, 2010) is presented by Equation 4:

$$Y_I = X_i\beta + Z_i b_i + \varepsilon_i, \tag{4}$$

for $j = 1, 2 \ldots, n_i$, being $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$ a vector of unknown parameters of fixed order effects (*p* + 1) *x* 1; $b_i = (b_{i1}, \ldots, b_{iq})^T$, is a vector of random effects with spatial dependence, of order *q* x1, being $b_i \sim N_q(\mathbf{0}, D)$, with D the covariance matrix of random effects; X_i: is a matrix of the fixed effect, of order $n_i\ x\ (p+1)$, containing the covariables of the individual i; Z_i: it is an array, of order $n_i\ x\ q$ (Z_i it is often a submatrix of X_i); $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{ini})^T$, are random errors, of order $n_i\ x\ 1$, of intra-individual measurements, being $\varepsilon_i \sim N_{ni}(\mathbf{0}, R_i)$, with R_i the $n_i\ x\ n_i$ covariance matrix of individual measurements. The estimated value of the response variable ($\hat{Y}$) is obtained from the Equation 5 considering the fixed effects $X_i\beta$, and the random effects $Z_i b_i$. The parameters of the longitudinal mixed linear model are estimated using the restricted maximum likelihood function (REML) (Verbeke; Molenberghs, 2005). It is necessary to understand the assumptions of the model in order to understand its behavior throughout the analysis.

In the linear random effects model, the regression coefficients b_i are called random effects and the set of variances and covariances ***D.*** Each individual measurement differs from the total average of the individuals by a certain amount that is defined by random error. These errors are independent and are used to determine the variance components. Covariance is determined by the errors observed in estimating the intra-individual mean and presents the characteristic of dependence.

The random effects in the intercept represent the natural heterogeneity between individuals. Some of the data variability may also be caused by unmeasured differences at the level of covariates, between individuals and within individuals. When this happens, it is necessary to include another level of nesting in the model or, in other words, include random effects on another covariate (Fausto *et al*., 2008). The covariates included in vector $\boldsymbol{Z}$ are time-dependent and are also part of vector $\boldsymbol{X}$. Therefore, QTAVES, QTSUI, QTBOV and PROLE are covariates that depend on time and, therefore, must be included in models with random intercept and fixed effects for covariates.

Schuielzeth *et al*. (2020) report that the assumptions of the generalized linear mixed model for longitudinal data are that the explanatory variables are linearly related to the response. Errors must assume constant variance. Errors must be independent and normally distributed. The residues must be plotted in relation to the explanatory variable that will indicate if the skewed model has been adjusted, also plot the residues against the adjusted values will indicate if there is variance of non-constant error, that is, if the variance increases with the mean where the waste spreads as the adjusted value increases. For the authors Schuielzeth *et al*. (2020) the restrictions of the model may be related to bimodal distributions where the presence of up and down shifted means may occur in or $\pm\ 1{,}5$ units, these changes may arise if there is a missing fixed effect influence on the model (Schuielzeth *et al*., 2020).

For this work, the objective of using the longitudinal mixed linear model was to describe the longitudinal behavior of the gross value of agricultural production and also to predict the profile of the quantity of poultry, pigs, cattle and milk produced.

In this work, the model of the gross value of the average agricultural production (Equation 5) was estimated, with the following parameters that consider the years 2018, 2019 and 2020.

$$\widehat{VBP}_{Full} = \hat{b} + \hat{\beta}_0\text{QTAVES} + \hat{\beta}_1\text{QTSUI} + \hat{\beta}_2\text{QTBOV} + \hat{\beta}_3\text{PROLE}, \qquad (5)$$

where, $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$: are the estimated parameters of fixed effects; *VBP*: gross value of agricultural production; *QTAVES*: quantity of poultries; Q*TBOV:* amount of cattle; *QTSUI*: quantity of pigs; and PROLE: milk production. A *stepwise* selection procedure was applied in the Equation model (5), to choose the explanatory variables in order to identify the most appropriate model to explain the variability in VBP. The data of the variables: gross value of agricultural production, quantity of poultries, quantity of cattle, quantity of pigs and milk production, were processed, following the transformation of Box-Cox (Box-Cox, 1964), as presented in Equation 6:

$$Z = \begin{cases} \ln(U), & se\ \lambda = 0, \\ \frac{U^{\lambda}-1}{\lambda}, & se\ \lambda \neq 0 \end{cases} \quad (6)$$

where, λ: is the parameter that defines the transformation of the data; U corresponds to the original data of the; Z: corresponds to processed data. In the models estimated for the random effect, variations in the interception of the models were considered, and the factor used for the estimation was in the mesoregions of the state of Paraná-Brazil. The parameters of the models were estimated with the use of the restricted maximum likelihood function, which allows to obtain estimates of variance of the model and also corrects possible biases that can occur when using the maximum likelihood. To obtain the longitudinal mixed models, the R (R development core team, 2023) software was used with the *lmer function* of the *lme4* package as presented by Verbeke and Molenberghs (2005).

RESULTS AND DISCUSSIONS

The results found in this work contribute to the regional development of the municipalities of Western Paraná, the study showed the behavior of the gross value of agricultural production throughout the analyzed period, which allowed this economic indicator of high relevance in the agribusiness sector the possibility of identification of important characteristics associated with poultry, pig and cattle production, which serves as privileged information for the assessment of managers and entities that represent the agro-industrial development of the area under study.

Table 1 presents the descriptive statistics of the variables: Gross value of agricultural production (VBP), quantity of poultries (QTAVES), quantity of pigs (QTSUI), quantity of cattle (QTBOV) and milk production (PROLE) for the years 2018, 2019 and 2020. It can be observed that the lowest standard deviations occurred in the QTSUI variable, followed by the PROLE variable for the years 2018, 2019 and 2020. The variables VBP18, VBP19 and VBP20 presented normality 5% of likelihood with the Kolmogorov–Smirnov test, disregarding the disparate points. This scenario of regional agricultural development is corroborated by the results of research by Krelling *et al*. (2023).

Table 1 | Descriptive analysis of the data of VBP, QTAVES, QTSUI, QTBOV and PROLE for the years 2018, 2019 and 2020.

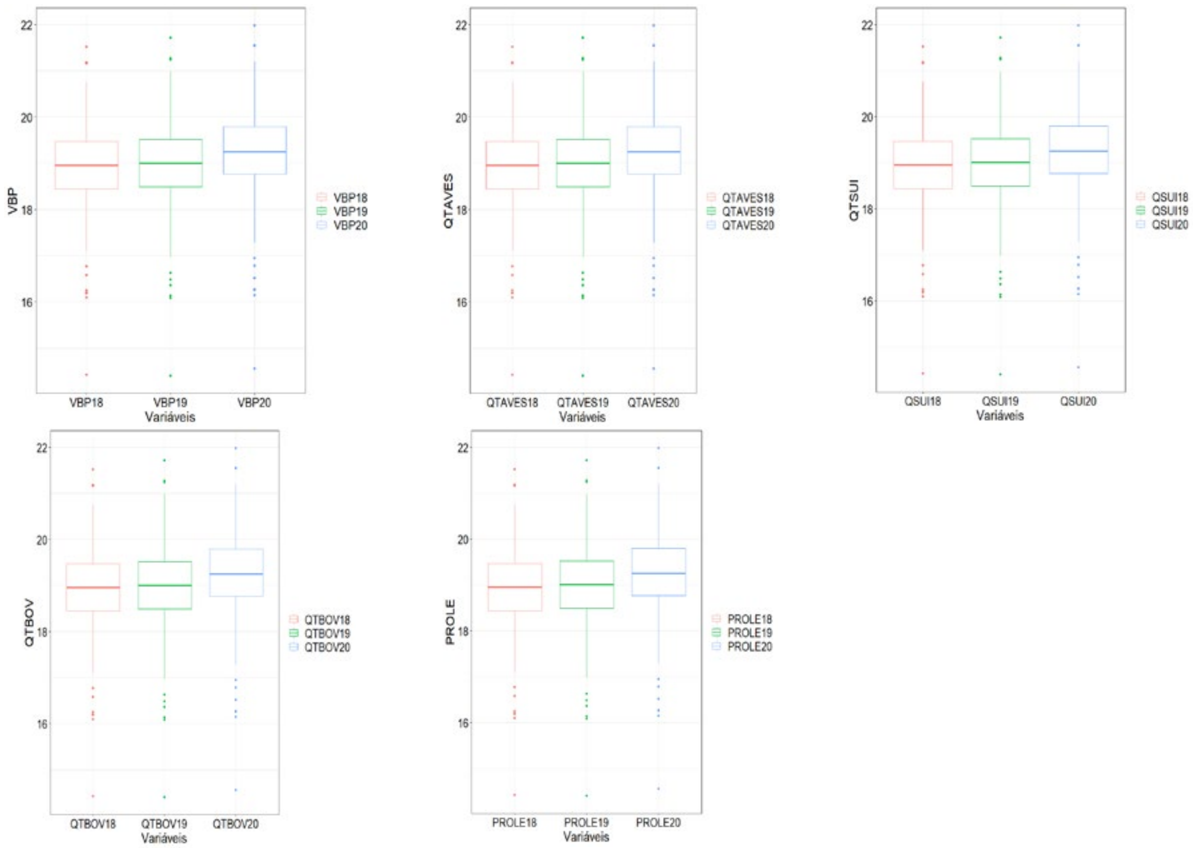
Variable	n	Min	Q1	Q2	Mean	Q3	Max	SD
VBP18*	399	14.426	18.440	18.949	18.916	19.461	21.518	0.818
QTAVES18	399	6.273	11.479	12.857	12.559	13.924	16.627	1.906
QTSUI18	399	2.944	6.875	7.772	8.025	8.961	13.973	1.700
QTBOV18	399	3.761	8.752	9.607	9.473	10.379	12.017	1.252
PROLE18	399	2.639	7.039	8.426	8.231	9.509	12.586	1.646
VBP19*	399	14.397	18.493	19.000	18.991	19.521	21.713	0.830
QTAVES19	399	6.295	11.434	12.853	12.568	13.934	16.647	1.898
QTSUI19	399	2.398	6.791	7.741	7.975	8.891	13.975	1.707
QTBOV19	399	3.091	8.700	9.577	9.430	10.358	11.967	1.265
PROLE19	399	2.565	7.004	8.423	8.202	9.411	12.543	1.669
VBP20*	399	14.559	18.769	19.254	19.255	19.795	21.973	0.847
QTAVES20	399	6.282	11.542	12.887	12.608	13.981	16.649	1.864
QTSUI20	399	1.386	6.685	7.673	7.911	8.912	13.976	1.786
QTBOV20	399	3.296	8.663	9.537	9.379	10.271	11.886	1.269
PROLE20	399	1.386	7.070	8.410	8.210	9.518	12.805	1.728

Note: N: number of municipalities, Min: minimum value, Q1: first quartile, Q2: second quartile, Med: Mean, Q3: third quartile, Max: maximum value, SD: standard deviation, * significance at 5% of likelihood level.

Source: Own calculations

Figure 2 shows the boxplot chart of the variables under study. The VBP dependent variables are originally obtained in reais, the variable PROLE in a unit of 1,000 liters and the other independent variables in unit of quantity/head. Symmetry is observed in the data distributions, and the presence of disparate points in the charts can also be observed as VBP for the years 2018, 2019 and 2020, and these points can be explained by the fact that we use in the research all the municipalities of the state of Paraná-Brazil. Some municipalities are considered the main producers of the state and others that have little representation in the production chain, which reflects in Figure 2 as the outliers below the lower limit and above the upper limit of the charts. Research such as Vieira *et al*. (2021) presents a similar discussion which is related with regional and agricultural development.

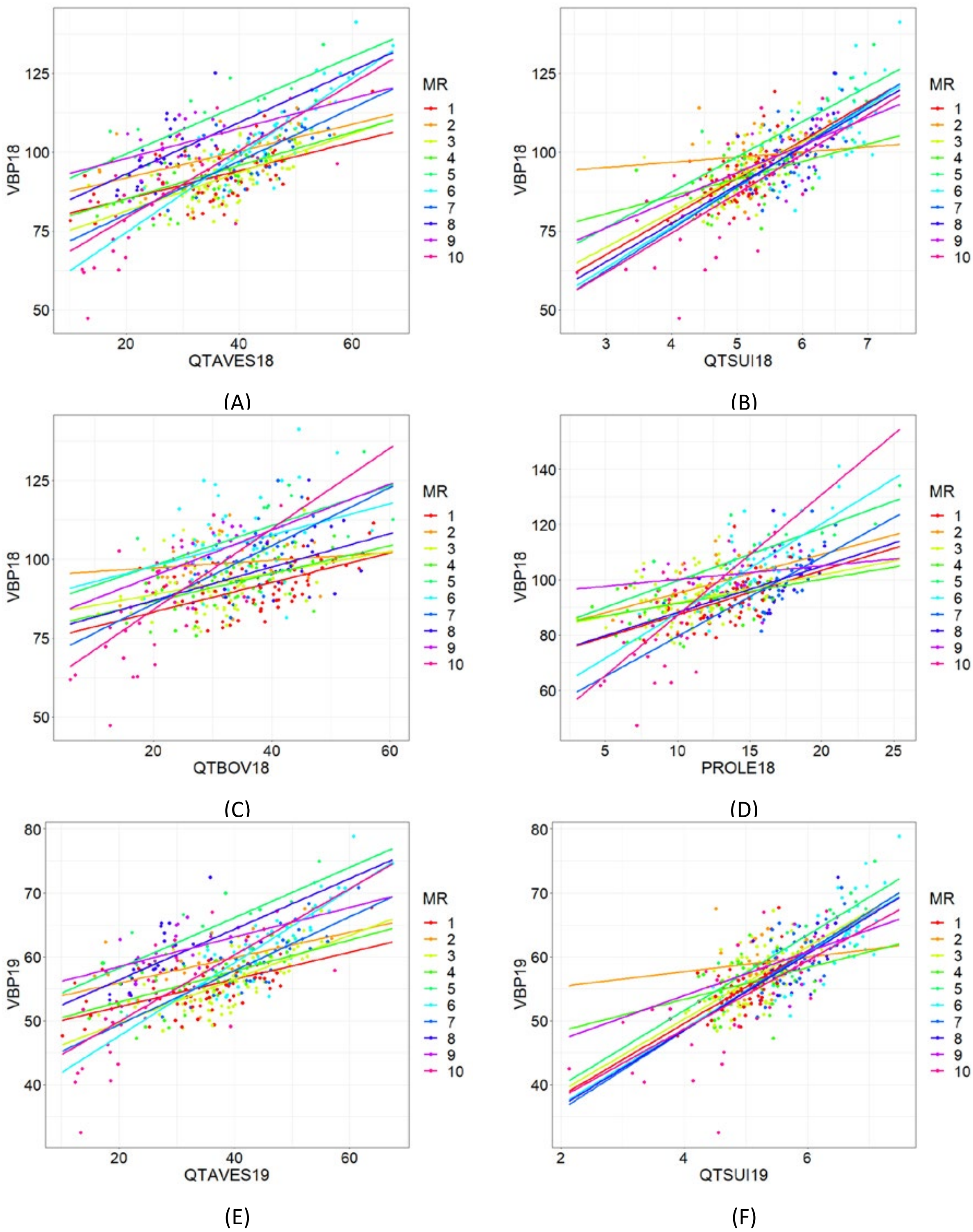
Figure 2 | Boxplot of the variables: Gross value of agricultural production (VBP), quantity of poultries (QTAVES), quantity of pigs (QTSUI), quantity of cattle (QTBOV) and milk production (PROLE) for the years 2018, 2019 and 2020.

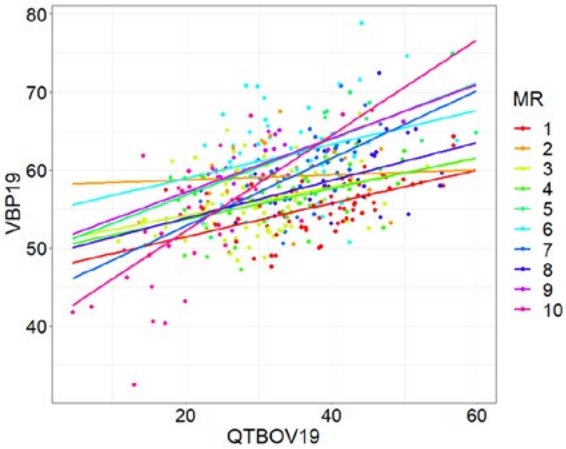


Source: Own research

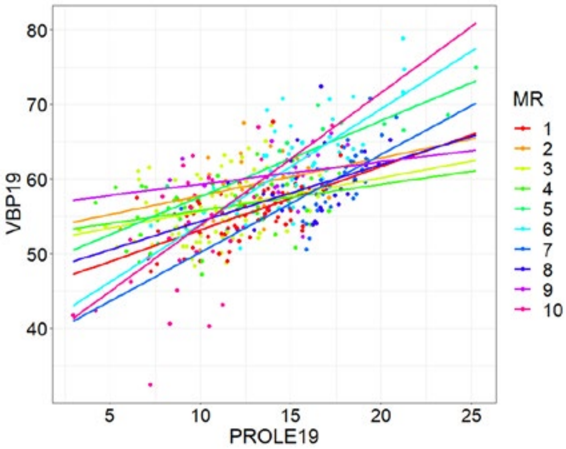
In figures (3A) to (3L) regression lines are showed considering the mesoregions (MR) of Paraná are denoted as 1: Paraná state Northwest , 2: Paraná state Western Center, 3: Paraná state North Central, 4: Paraná state North Pioneer, 5: Paraná state Eastern Center, 6: Paraná state West , 7: Paraná state South-West , 8: Paraná state South Center, 9: Paraná state South-East, 10: Metropolitan of Curitiba. The mesoregions of Paraná are used as a factor for the generation of longitudinal mixed models, so that Figures (3A) to (3L) indicate the relationship between each dependent variable and each independent variable before generating the mixed models. It can be observed in these Figures for the years 2018, 2019 and 2020 that the intercepts of the models are different in each mesoregion, which indicates the pertinence of using mesoregions as a factor to obtain the random effects of interception of the spatial longitudinal mixed models in this study.

Figure 3 | Linear response profile in each mesoregion (MR) of the state of Paraná.

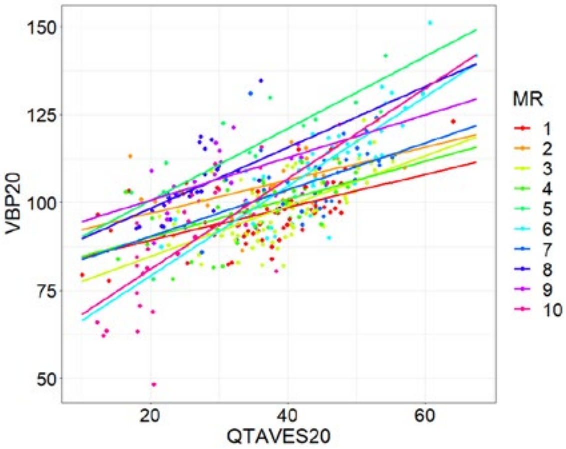




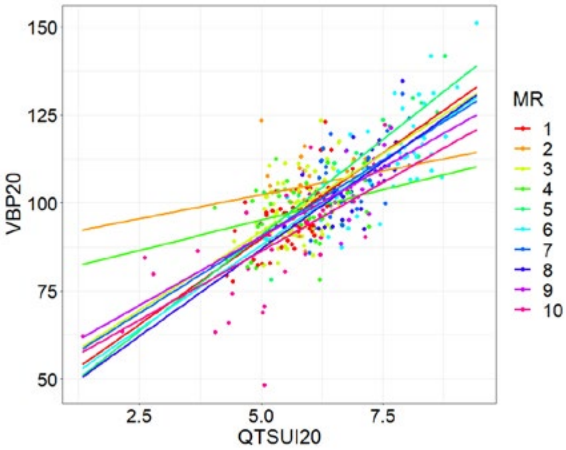
(G)



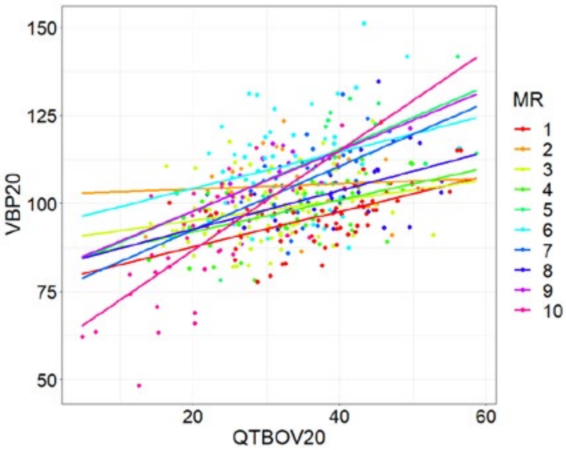
(H)



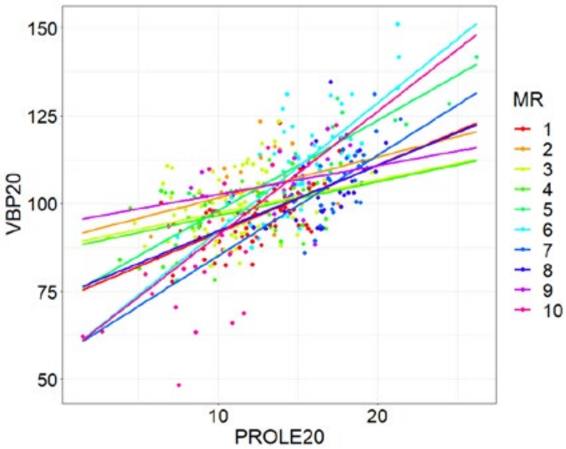
(I)



(J)



(K)

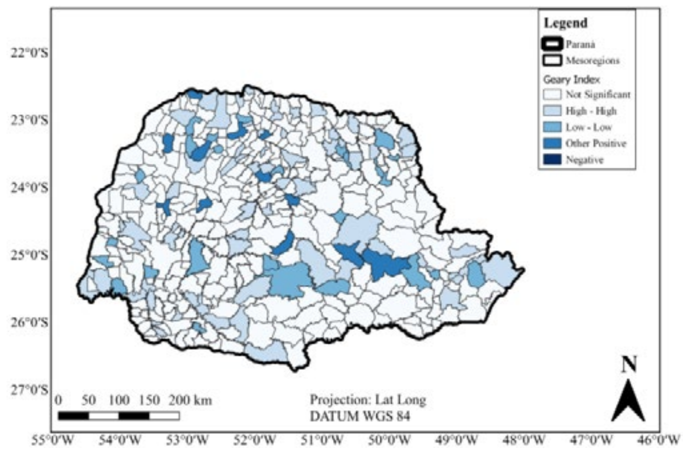


(L)

Source: Own research

The response profiles in the ten mesoregions (MR) of the state of Paraná for the variables under study are presented in Figures 3A to 3L. The relational behavior between the independent variables and the dependent variable is called the response profile (Kim, *et al*., 2003; Narvand *et al*., 2013). The analysis of the response profiles allows to identify a linear relationship between the dependent variable VBP and the independent variables QTAVES, QTSUI, QTBOV and PROLE in the mesoregions. The profiles indicate different interception parameters for the models in the mesoregions, thus an average parameter can be used (Bates *et al*., 2015). Although, in some mesoregions, differences in model angulation parameters are noticed, it is evident in Figures 3A to 3L that the intercept is the parameter with more distinction in the profiles. The Geary index, calculated from the Equation 3 is shown in Figure 4. The formed groupings are used to capture the random effects of the variables in the longitudinal mixed model under study.

Figure 4 | Geary index for the municipalities of the state of Paraná.



Source: Own research

Figure 4 shows that a grouping of five classes of the univariate local Geary index (C_i) was performed. The high-high category represents municipalities that have high value in the VBP and are surrounded by municipalities that also have values, which can be considered high in the VBP variable, being observed in municipalities such as Toledo, Tibagi, Castro and others, are part of the West and Center-West mesoregion. In the case of low-low class are municipalities that have low values of the variable of interest surrounded by neighbors that also present low values, and in this situation we observe municipalities such as Guarapuava, Irati and Campo Largo, belonging to the Central-South, Southeast and Metropolitan mesoregions of Curitiba respectively. In the other positive class, they are municipalities that have high values surrounded with municipalities that correspond to a positive correlation, as in the municipality of Pitanga, which is part of the Center-South mesoregion.

In the economic development context sustainable, the spatial behavior observed suggests an expressive spatial variability in all municipalities of the state of Paraná, and it is related to the fact that many municipalities associated with their mesoregions, have high agricultural production and others do not follow this productive dynamic. Thus, the variability identified in the univariate Geary local index is consistent, since it adequately reflects the behavior observed in the agricultural activity of the state of Paraná-Brazil. In the research by Kalinowska *et al*. (2022) spatial aspects of agricultural development associated with sustainability are discussed and which also corroborate the results obtained in this work.

Table 2 presents the estimated parameters for the complete longitudinal mixed model that considers the years 2018, 2019 and 2020. In the estimation of the parameters, mesoregions were assigned as a random effect factor, which allows for each mesoregion an intercept for the model, and the interception parameter of the obtained model is determined by the mean of mesoregion parameters. The model estimated is specified in Equation 7.

$$\widehat{VBP}_{Full} = 14.4136 + 0.1404\ QTAVES + 0.2336\ QTSUI + 0.0937\ QTBOV + 0.0155\ PROLE \qquad (7)$$

To select the significant explanatory variables of the Equation model (7), the *stepwise* process was performed using the *step* function of the *lmer* Test package of the R (R Development Core Team, 2023) software. The *stepwise* process data are presented in Table 2. It is observed in this table that the explanatory variables QTAVES, QTSUI and QTBOV were significant at the level of 5% likelihood. In the *stepwise* process the variable PROLE was not significant, being eliminated from the final model.

Table 2 | Stepwise selection process statistics

Variable	E	Gl	SQ	RSS	AIC	F-value	P-value
PROLE	1	1	0.229	338.06	-1505.4	0.8072	0.3691^{NS}
QTAVES	0	1	66.435	404.5	-1292.7	234.4456	0.0000^{*}
QTSUI	0	1	147.154	485.21	-1074.9	519.2986	0.0000^{*}
QTBOV	0	1	16.3	354.36	-1451.1	57.5203	0.0000^{*}

Note: E: Variable deleted, 1 for deleted and 0 for variable maintained in the model; Gal: degree of freedom; SQ: Sum of squares: RSS: Residual Sum Square; AIC: Akaike Information Criterion; F-value: F-statistic value; P-value: significance level; NS: not significant at 5% significance level; * significance at 5% likelihood level.

Source: Own calculations

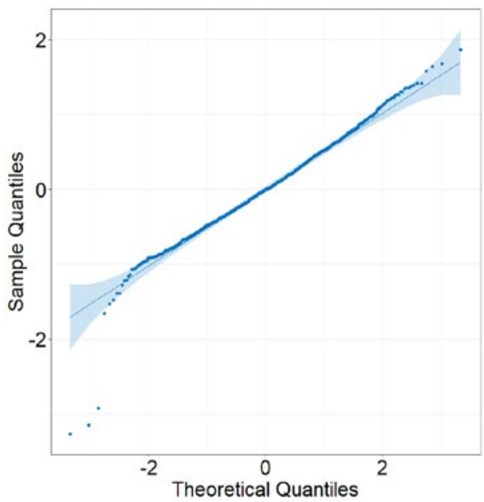
Thus, the final model estimated using the stepwise criterion is presented in Equation 8:

$$\bar{V}B\bar{P}_{stepwise} = 14.382 + 0.1404\ \text{QTAVES} + 0.2409\ \text{QTSUI} + 0.1046\ \text{QTBOV} \qquad (8)$$

The Equation model (6) has $R^2 = 0{,}733$ and $R^2_{Adjusted} = 0{,}732$. In Figure 5, for the estimated model without the variable PROLE, the theoretical quantile versus sample quantile chart is presented for the residues and bands with 95% confidence, considering the data of the years 2018, 2019 and 2020. It is possible to observe that there is an indication of normality for the residues of this model, so that it satisfies the assumption of normality.

The presence of residue values outside the confidence bands both below and above can be attributed to the fact that the study was conducted with the variables in the 399 municipalities of the entire state of Paraná-Brazil, and considering that there are in this set municipalities with high agricultural activity, that reflect in high values of VBP. These municipalities are represented in Figure 5 by the points with higher values outside the upper limits of the confidence bands. Municipalities with small agricultural activity are represented by points below the lower limits of the confidence bands. Thus, through the diagnosis used in the residues, there is no evidence to rule out the assumption of normality of the residues in the models obtained.

Figure 5 | Q-qqplot chart for the residues of the longitudinal mixed linear model with the variable PROLE eliminated for the years 2018, 2019 and 2020.



Source: Own calculations

In the model, it was possible to identify that QTAVES, QTBOV and QTSUI were important in predicting the gross value of agricultural production, since they were significant in the linear mixed longitudinal model selected to explain VBP in the years under study. Thus, the methodology used was adequate to identify spatial patterns and trends and comes against the need to explore the relations between agricultural activities and their impacts in the economic area.

In this study it was possible to identify a linear relationship of VBP with QTAVES, QTSUI and QTBOV with the use of mixed models. The evidence of the economic importance of these variables can be observed in the scientific literature that addresses the production in a joint way with the agricultural assets (Steinmets *et al*., 2021).

In this study, the economic importance of the number of bovine animals was evidenced, since the variable QTBOV presented statistical significance to explain the VBP. These results corroborate with several studies on cattle production systems and their impact on the estimation of economic values (Gathura *et al*., 2020; Alejandro *et al*., 2020). Similarly, the variable number of pigs presented statistical significance to explain the gross value of agricultural production (BVC), and although the pigs breeding may imply a certain cost of production and environmental cost, it is a relevant agricultural asset in the Brazilian productive system (Ali *et al*., 2018; Lima *et al*., 2021; Rocha *et al*., 2022).

The variable QTAVES in this study was also significant to explain the VBP being a result that agrees with Da Silva *et al*. (2020) on the importance of this variable in the economic context. The relevance of pig and poultry production can also be observed through systems that guide its integration and management (Dal Moro *et al*., 2021).

Milk production, the PROLE variable was not significant to explain VBP in Paraná for the years under study. This fact may be related to the family development and punctual way of dairy production in the regions that make up the state of Paraná-Brazil (Dorregão *et al*., 2020; Eurich *et al*., 2022), although family agriculture is relevant, with more than 50% of the gross domestic product (GDP) of production in the pig, poultry and milk chains (Eurich *et al*., 2022), the aggregate value associated with milk tends not to reach the market levels of the other assets evaluated in this study, thus, it is not significant to explain the VBP.

Agricultural activity is an important economic component for the generation of employment and income at regional, national and world level and contributes to the development of the State of the Countries (Esteves *et al*., 2016; Amaral *et al*., 2020; Cima *et al*., 2021). The mixed models used in this work allowed to measure the contribution of agricultural variables in the VBP of the state of Paraná, being possible to guide actions to increase these variables in order to strengthen the VBP, as well as contribute to public and private initiatives that make up the economic system.

CONCLUSION

The results presented are of regional and global economic representativeness, representing the profile of the longitudinal spatial behavior of agricultural variables. Thus, they highlight the importance of food security, as well as obtaining it in a sustainable way, thus contributing to social and environmental well-being.

The analysis with the longitudinal mixed model allowed to evaluate the behavior of the response variable, and it was the gross value of agricultural production (VBP) over time and allowed to identify important characteristics that allowed to follow its evolution over the period analyzed.

The variables QTAVES, QTSUI and QTBOV were significant to explain the variability in the VBP variable in the years 2018, 2019 and 2020. The values estimates for VBP can be performed with longitudinal mixed model with independent variables such as QTAVES, QTSUI and QTBOV, thus becoming a differential of possible analysis diagnoses in the agricultural sector, favoring the rural manager in his or her decision making.

Thus, it seeks to indicate recommendations to public policy managers, agribusiness entrepreneurs and local, regional, national and world development agents with a view to increasing the gross value of agricultural production, through actions in their respective sectors that propel the different segments that make up the economic system as a whole.

The variable PROLE was not significant to explain VBP in the years 2018, 2019 and 2020, and the longitudinal mixed linear model without the milk production variable was selected by the selection criteria used, as the best model to explain VBP in the years studied.

The random effect of the estimated longitudinal mixed models was obtained considering the variation of intercepts in the mesoregions, since the linear models suggest distinct intercepts between the mesoregions.

In the results found in the present work, it was evident the importance and contribution of the longitudinal mixed linear model for the analysis in the agricultural sector, thus contributing to scientific research, the economic sector and the agro-industrial sector.

The study showed that the results found do not limit the subject and that new scientific work is necessary to understand the universe that involves special analysis of agricultural data. Therefore, research in this field benefits regional, economic and social development, contributing to the economic understanding of municipalities, promoting practical and applied knowledge in the scientific community and in the agribusiness sector.

In view of this study, new works are suggested that focus on the spatial analysis of agricultural data considering other variables of interest such as agricultural GDP using quantitative methods of spatial econometrics. The results of these researches are relevant in regional, national and international development scenarios.

In this sense, the research showed that the magnitude of the information obtained can be explored in decision-making by presenting precise diagnoses of the possible scenarios that can be analyzed, providing the diversity of economic analysis in the comprehensive universe that is agribusiness.

ACKNOWLEDGMENT

The authors would like to thank the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Financing Code 001 for financial support, the National Council for Scientific and Technological Development (CNPq), the Graduate Program in Regional Development and Agribusiness (PGDRA) at the Western Paraná State University (UNIOESTE), and the of Spatial Statistics Laboratory (LEE-LISA).

REFERENCES

ALMEIDA, E. **Econometria Espacial Aplicada**. São Paulo: Alínea, 2012.

ALI, B. M.; DE MEY, Y.; BASTIAANSEN, J. W. M.; OUDE-LANSINK, A. G. J. Effects of incorporating environmental cost and risk aversion on economic values of pig breeding goal traits. **Journal of Animal Breeding and Genetics**, v.135, n.3, p.194–207, 2018. DOI: https://doi.org/10.1111/jbg.12331.

ALVES, L. R. Localização e reestruturação da base de exportação das regiões imediatas do estado do Paraná-Brasil entre 2010 e 2020, **Revista Informe GEPEC**. v. 26, n.3, p. 416-438, 2022.

ALVES, L. R.; ANDREICA, A. B.; ALCANTARA, I. R. Using industry 4.0 tools for increasing competitiveness of Clusters in Western Paraná. **Brazilian Journal of Development**, v. 6, n. 8, p. 56648-56672, 2020. DOI: 10.34117/bjdv6n8-179.

AMARAL, J. A.; RODRIGUES, K. C. T.; STADUTO, J. A. R. Emprego formal na agropecuária do Estado do Paraná: Uma análise por sexo. **Economia & Região**, v. 8, n. 1, p. 163-182, 2020. DOI: https://doi.org/10.5433/2317-627X.2020v8n1p163.

AMAYA, A.; GARRICK, D.; MARTÍNEZ, R.; CERÓN-MUÑOZ, M. Economic values for index improvement of dual-purpose Simmental cattle. **Livestock Science**, v. 240, p.104-224, 2020. DOI: https://doi.org/10.1016/j.livsci.2020.104224.

ARIYO, O.; QUINTERO,A.; MUÑOZ, J.; VERBEKE,G.; LESAFFRE, E. Bayesian model selection in linear mixed models for longitudinal data. **Journal of Applied Statistics**, v.47, n. 5, p. 890-913, 2020. DOI: 10.1080/02664763.2019.1657814. ISSN: 0266-4763.

ASPAROUHOV, T.; HAMAKER, E. L.; MUTHÉN, B. Dynamic structural equation models, **Structural Equation Modeling**, v.25, n. 3, p.359–388, 2018. DOI: 10.1080/10705511.2017.1406803. ISSN: 1532-8007.

ASPAROUHOV, T.; MUTHÉN, B. Latent variable centering of predictors and mediators in multilevel and time-series models. **Structural Equation Modeling**, v.26, n. 1, p.119–142, 2019. DOI: https://doi.org/10.1080/10705511.2018.1511375.

BANCO MUNDIAL. **A Economia nos Tempos de COVID-19**. Relatório Semestral sobre a América Latina e Caribe, p.1-66, 2020.

BATES, D.; MACHLER, M.; BENJAMIN, M.; BOLCKER, S.; WALKER, C. Fitting Linear Mixed-Effects Models using lme4. **Journal of Statistical Software**, v.67, n. 1, p. 1-51, 2015. DOI: 10.18637/jss.v067.i01.

BOX, G.; COX, D. An analysis of transformations. **Journal of the Royal Society,** v. 26, n.2, p. 211-252. Pennsylvania State University, United States. ISSN: 1467-9868, 1964.

CHAN, J.; GRANT, A. On the observed-data deviance information criterion for volatility modeling. **Journal of Financial Econometrics**, v. 14, n.4, p. 772–802, 2016. ISSN: 1479-8417.

CIMA, E. G.; ROCHA-JUNIOR, W. F.; DALPOSSO, G. H.; URIBE-OPAZO, M. A.; BECKER, W. R. Forecasting Grain Production and Static Capacity of Warehouses Using the Natural Neighbor and Multiquadric Equations. **AGRIS on-line Papers in Economics and Informatics**, v. 13, n. 3, p. 3-14, 2021. ISSN 1804-1930. DOI: 10.7160/aol.2021.130301.

CIMA, E. G; ROCHA-JUNIOR, W. F.; URIBE-OPAZO, M. A.; BOMBACINE, M. R. O novo traçado da Ferroeste: uma perspectiva para o desenvolvimento do Oeste do Paraná-Brasil, **Revista Brasileira de Gestão e Desenvolvimento Regional**. v.18, n.3, p.371-383, 2022.

DAL MORO, L.; MACULAN, L. S.; NECKEL, A.; VARGAS MORES, G.; PIVOTO, D.; BODAH, E. T.; OLIVEIRA, M. L. S. Geotechnologies applied to the analysis of buildings involved in the production of poultry and swine to the integrated food safety system and environment. **Journal of Environmental Chemical Engineering**, v. 9, n. 6, p. 64-75, 2021. DOI: https://doi.org/10.1016/j.jece.2021.106475.

DORREGÃO, V. V.; SALVARO, G. I. J.; ESTEVAM, D. de O. Contribuições da atividade leiteira para o desenvolvimento rural e para a reprodução da agricultura familiar em um município do sul catarinense. **Interações**, v.20, n.3, p.973–985, 2019. DOI: https://

doi.org/10.20435/inter.v20i3.1902.

ESTEVES, E. G. Z.; MICHELLON, E., ALVES, A. F.; SESSO-FILHO, U. A.; AMARAL, A. E. Impactos locais e inter-regionais no emprego e na renda da agropecuária no Estado do Paraná. **Economia & Região**, v. *3,* n.2, p.109-129,2016. DOI: https://doi.org/10.5433/2317-627X.2015v3n2p109.

EURICH, J.; WEIRICH NETO,; P. H.; ROCHA, C. H. Dairy production in a family farming colony in Palmeira, Paraná, Brazil. **Revista Ceres**, v.63, n.4, p. 454-460, 2022. DOI: https://doi.org/10.6084/m9.figshare.19928984.

GATHURA, D. M.; MUASYA, T. K.; KAHI, A. K. Meta-analysis of genetic parameters for traits of economic importance for beef cattle in the tropics. **Livestock Science**, v. 242, p.104-306, 2020. DOI: https://doi.org/10.1016/j.livsci.2020.104306.

GAYLE, V.; LAMBERT, P. What is Quantitative Longitudinal Data Analysis? **Bloomsbury Publishing**, London, p. 144, 2018. ISBN: 9781472515391.

GUARNIERI, M.C.G.; THOMÉ, K.M. **Agronegócios: perspectivas**. Brasilia. Editora Universidade de Brasília, 2020. 397 p., 2020.

HEMATHILAKE, D.M.K.S; GUNATHILAKE, D.M.C.C. Chapter 31 - Agricultural productivity and food supply to meet increased demands. **Future Foods**. Global Trends, Opportunities, and Sustainability Challenges, p. 539-553, 2022. DOI: https://doi.org/10.1016/B978-0-323-91001-9.00016-5.

HERLE, M.; MICALI, N.; ABDULKADIR, M.; LOOS, R.; BRYANT-WAUGH, R.; HUBEL, C.; BULIK, C.M.; Stavola, B. Identifying typical trajectories in longitudinal data: modelling strategies and interpretations. European. **Journal of Epidemiology**, v. 35, n. 3, p. 205–222, 2020. DOI: https://doi.org/10.1007/s10654-020-00615-6.

HOFFMAN, L. **Longitudinal Analysis: Modeling within-Person Fluctuation and Change, Routledge**. New York. p. 654, 2015. DOI: https://doi.org/10.4324/9781315744094. ISBN 9780415876025.

IPARDES, Instituto Paranaense de Desenvolvimento Econômico e Social. (2022). **BDEweb - Base de dados do estado**. Available: https://www.ipardes.pr.gov.br/Pagina/Base-de-Dados-do-Estado-BDEweb. Accessed: 04 de abr de 2023.

KALINOWSKA, B.; BÓRAWSKI, P.; BEŁDYCKA-BÓRAWSKA, A.; KLEPACKI, B.; PERKOWSKA, A.; ROKICKI, T. Sustainable Development of Agriculture in Member States of the European Union. **Sustainability**, v.14, 4184, 2022. DOI: https://doi.org/10.3390/su14074184.

KIM, K.; MAHMOUD, M. A.; WOODALL, W. H. On the Monitoring of Linear Profiles. **Journal of Quality Technology**, v. 35, n. 3, pp.317–328, 2003. DOI: https://doi.org/10.1080/00224065.2003.11980225.

KRELLING , A.P.; ANTUNES, C.V.; BROADHURST, M.K. Investigating variability among fisheries litter accumulation on beaches in Paraná, Brazil. **Marine Pollution Bulletin**, v.187, 114-607, 2023, DOI: https://doi.org/10.1016/j.marpolbul.2023.114607.

KUHN, S. L., ROCHA-JUNIOR, W. F., STADUTO, J. A. R. Contratos no agronegócio: o caso de Cascavel, PR. **Revista Informe GEPEC**, v10, n.2, jul./dez., p. 133-151, 2005.

LI, L.; QIU, S.; ZHANG, B.; Feng, C.X. Approximating cross-validatory predictive evaluation in Bayesian latent variable models with integrated IS and WAIC. **Statistics and Computing**,v. 26, p.881–897, 2016. DOI: doi.org/10.1007/s11222-015-9577-2.

LIMA, J. F. Valor Adicionado Fiscal no Estado do Paraná: concentração e reestruturação regional. **Revista do Desenvolvimento Regional**, v.17, n.2, p.100-112, 2020. DOI: 10.26767/colóquio.v17i2.1664.

LIMA, M. S. M.; ERYARSOY, E.; DELEN, D. Predicting and Explaining Pig Iron Production on Charcoal Blast Furnaces: A Machine Learning Approach. **Informs Journal on Applied Analytics**, v.51, n.3, p. 213–235, 2021. DOI: https://doi.org/10.1287/inte.2020.1058.

MILLAR, R.B. Conditional vs marginal estimation of the predictive loss of hierarchical models using WAIC and cross-validation. **Statistics and Computing**, v. 28, p. 375–385, 2018. DOI: 10.1007/s11222-017-9736-8.

MONTEIRO NETO, A.; COLOMBO. L.A.; ROCHA NETO, J.M. (2023). **Desenvolvimento regional no Brasil: políticas, estratégias e perspectivas.** Rio de Janeiro: IPEA, 2023.

NARVAND, A.; SOLEIMANI, P.; RAISSI, S. Phase II monitoring of auto-correlated linear profiles using linear mixed model. **Journal of Industrial Engineering International**, v. 9, n.1, p.12, 2013. DOI: https://doi.org/10.1186/2251-712X-9-12.

PRAVIA, M. I.; RAVAGNOLO, O.; URIOSTE, J. I.; GARRICK, D. J. Identification of breeding objectives using a bioeconomic model for a beef cattle production system in Uruguay. **Livestock Science**, v. 160, p. 21–28, 2014. DOI: https://doi.org/10.1016/j.livsci.2013.12.006.

QUINTERO, A.; LESAFFRE, E. Multilevel covariance regression with correlated random effects in the mean and variance structure. **Journal Biométrica**, v.59, n. 5, p. 1047–1066, 2017. DOI: 10.1002/bimj.201600193.

QUINTERO, A.; LESAFFRE, E. Comparing hierarchical models via the marginalized deviance information criterion. **Statistics in Medicine,** v.37, n. 430, p. 2440–2454, 2018. DOI: 10.1002/sim.7649.

R DEVELOPMENT CORE TEAM, *R*. **A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, ISBN 3-90005107-0. Disponível em: http://www.R-project.org. Acesso em: Jul, 2022.

RIES, A. C. R.; COSTA-SILVA, V.; SANTOS, C. F.; BLOCHTEIN, B.; THYSSEN, P. J. Factors Affecting the Composition and Succession of Beetles in Exposed Pig Carcasses in Southern Brazil, **Journal of Medical Entomology**, v. 58, n.1, p. 104–113. 2021. DOI: https://doi.org/10.1093/jme/tjaa175.

RIZARDI, B. M. (2022). **Caminhos da inovação no setor público**. Brasília: Enap, 392 p. 2022.

ROCHA, G. C.; DUARTE, M. E.; KIM, S. W. Advances, Implications, and Limitations of Low-Crude-Protein Diets in Pig Production. **Animals (Basel)**, v.12, n. 24, p.1-17, 2022. DOI: https://doi.org/10.3390/ani12243478.

SÄFKEN, B.; RÜGAMER, D.; KNEIB, T.; GREVEN, S. Conditional model selection in mixed-effects models with cAIC4. **Journal of Statistical Software**, v.99, n.8, p. 1-30, 2018. DOI: https://doi.org/10.18637/jss.v099.i08.

SCHIELZETH, H. NIEL, J.; NAKAGAWA, S.; WESTNEAT, D.F.; ALLEGUE, H.; TEPLITSKY, C.; RÉALE, D.; DOCHTERMANN, N.A.; GARAMSZEGI, L.Z.; ARAYA-AJOY, Y.G. Robustness of linear mixed-effects models to violations of distributional assumptions. **Methods in Ecology and Evolution**, V.11, n.9, p. 1141- 1152, 2020.DOI: 10.1111/2041-210X.13434.

SILVA, E. P. **Sistema agroindustrial de origem vegetal: livro didático** . Palhoça: Unisul-Virtual. 2015.

SILVA, S. N.; BRUM-JÚNIOR, B. D. S.; VALENTE, B. S.; MANZKE, N. E.; LOPES, D. C. N.; RUTZ, F.; XAVIER, E. G. Composição corporal de frangos de corte fêmeas com 21 dias alimentadas com níveis de quirera de arroz na dieta. **Science And Animal Health**, v.7, n. 2, p.147-156, 2020. DOI: https://doi.org/10.15210/sah.v7i2.14462.

STEINMETZ, L.; VEYSSET, P.; BENOIT, M.; DUMONT, B. Ecological network analysis to link interactions between system components and performances in multispecies livestock farms. **Agronomy for Sustainable Development**, v. 41, n. 3, p. 1–16, 2021. DOI: https://doi.org/10.1007/s13593-021-00696-x.

VERBEKE, G.; MOLENBERGHS, G. **Linear mixed models in practice: a SAS-oriented approach**. In Springer Series in Statistics, p.319, 1997. ISBN 978-0-387-98222-9

VERBEKE, G.; MOLENBERGHS, G. **Linear Mixed Models For Longitudinal Data**. In Springer Series in Statistics, p.579. DOI: https://doi.org/10.1007/978-1-4419-0300-6, 2005. ISBN 978-0-387-95027-3.

VIEIRA, F. M. C.; MILITÃO, E. E.; JÚNIOR, A. W.; HERBUT, P.; MAYER, L. R. R .;VISMARA, E. S.; VIEIRA, A. M. C.; PARIS, W.; HERMES, C.; PAŚMIONKA, I. Impact of Shading of a Silvopastoral System on Microclimate and Botanical Composition of Polyphytic Pasture: A Preliminary Study in the Subtropics. **Journal of Ecological Engineering**, v. 22, n.10, p. 215–224. 2021. DOI: https://doi.org/10.12911/22998993/142206.

WU, L. **Mixed Effects, Models for Complex Data**. Vancouver: Chapman & Hall. 2010. ISBN 9780367384913.