Ciência Rural, Santa Maria, v.54:04, e20220335, 2024
http://doi.org/10.1590/0103-8478cr20220335
ISSNe 1678-4596
AGRIBUSINESS

Multidimensionality of agricultural grain road freight price: a multiple linear regression model approach by variable selection

Alcília Mena José Sitoe Macarringue[1] **Andréa Leda Ramos de Oliveira**[1*] **Carlos Tadeu dos Santos Dias**[2] **Karina Braga Marsola**[1]

[1]Laboratório de Logística e Comercialização Agroindustrial (LOGICOM), Faculdade de Engenharia Agrícola (FEAGRI), Universidade Estadual de Campinas (UNICAMP), 13083-875, Campinas, SP, Brasil. E-mail: aleda@unicamp.br. *Corresponding author.
[2]Departamento de Ciências Exatas, Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ), Universidade de São Paulo (USP), Piracicaba, SP, Brasil.

ABSTRACT: The road system is the main mode used for the transportation of agricultural cargo, and in some cases, it is the only option for handling this type of product. This dependence means that the implementation of tools to support the management of logistical costs can reduce the financial impact with the transport felt by the economic agents operating in the soybean chain. This study contributed to a better understanding of the variables that make up the cost of road freight, generating a system of road freight prediction from a multiple linear regression model using the selection of variables Stepwise, Forward, and Backward elimination. This being said, this research intends to evaluate whether the behavior of soybean road freight is influenced by the variables that make up the productive, economic, and infrastructure dimensions in price formation. The regression models had an explanatory power of 87.20%. In the infrastructure dimension, the most impact variable in soybean road freight was the distance traveled; in the economic dimension, the variables of inflation and fuel price stood out; while in the productive dimension, the main contribution was the volume of production. A more assertive predictability of logistical costs and better understanding of the dynamics of freight price formation helps industry agents in planning and decision-making. Another contribution of this study is that it can be used as a practical tool for predicting soybean road freight on several transportation routes.
Key words: agroindustrial logistics, predictive models, transport routes, price behavior.

Multidimensionalidade do preço de frete rodoviário de grãos agrícolas: uma abordagem de um modelo de regressão linear múltipla por seleção de variáveis

RESUMO: O sistema rodoviário é o principal modal utilizado para o transporte de cargas agrícolas, e em alguns casos, é a única opção para movimentação desse tipo de produto. Esta dependência faz com que a implementação de ferramentas para apoiar a gestão dos custos logísticos possa reduzir o impacto financeiro com o transporte sentido pelos agentes econômicos que atuam na cadeia da soja. O presente trabalho visa contribuir para um melhor entendimento das variáveis que compõem o custo dos fretes rodoviários, gerando um sistema de previsão de fretes rodoviários a partir de um modelo de regressão linear múltipla usando a seleção de variáveis Stepwise, Forward e a eliminação Backward. Isto posto, a pesquisa tem por objetivo avaliar se o comportamento do frete rodoviário da soja é influenciado pelas variáveis que compõem as dimensões produtivas, econômicas e de infraestrutura na formação do preço. Os modelos de regressão tiveram um poder explicativo de 87,20%. Na dimensão infraestrutura o destaque foi a variável distância percorrida como a mais impactante no frete rodoviário da soja, na dimensão econômica as variáveis inflação e preço do combustível, enquanto que na dimensão produtiva a principal contribuição foi o volume de produção. A previsibilidade mais assertiva dos custos logísticos e melhor compreensão da dinâmica de formação do preço do frete auxilia os agentes do setor no planejamento e na tomada de decisão. Outra contribuição deste trabalho é que pode ser usado como uma ferramenta prática para a predição de fretes rodoviários da soja em diversas rotas de transporte.
Palavras-chave: logística agroindustrial, modelos preditivos, rotas de transporte, comportamento dos preços.

INTRODUCTION

In the past decade, Brazil has been achieving records in several agricultural chains, both in production and export, despite current infrastructure conditions. To maintain this good performance, continuous advances are required in new technologies that lead to productivity gains and/or value addition to the final product; however, a recurring aspect for sustaining the agricultural sector is overcoming obstacles related to distribution logistics (OECD/FAO, 2021).

Roads are the main means used for the transport of agricultural cargo, and often the only option for the displacement of this type of product, due to the reduced navigability of waterways and the little availability of railways that connect large distances and, at the same time, are located near the agricultural producing poles (OLIVEIRA et al., 2021). In addition to this predominance in the transport matrix, road freight is the reference of freight of other modes; that is, it is an indicator of the freight market. The operators of both rail and waterway modes associate their prices with the one practiced by the road mode (OLIVEIRA et al., 2013).

Received 06.09.22 Approved 26.06.23 Returned by the author 09.01.23
CR-2022-0335.R1
Editor: Leandro Souza da Silva

The pricing of agriculture freight can be influenced by a set of dimensions that bring together different variables (OLIVEIRA et al., 2022). In the productive dimension we have variables such as: harvest and off-season period, regions with high rates of production and productivity, planted area and harvested area. In the infrastructure dimension we have variables such as: road quality, road tolls, monitoring compliance, distance traveled, warehousing capacity, processing/crushing industry. Other relevant variables to predict soybean freight by road are related to the economic dimension, such as: inflation, fuel prices, exchange rate, volume of agricultural exports and fertilizers imports (CANGUSSU et al., 2013; COOPE et al., 2020; LIMA et al., 2016; MACHADO et al., 2019; MARTINS et al., 2004; PÉRA et al., 2018). The variables associated with the operational dimension also affect prices, in particular, the possibility of return cargo, loading and unloading time, specificity of the cargo transported, type of vehicle and delivery time (CANGUSSU et al., 2013; MACHADO et al., 2019; MARTINS et al., 2004; PÉRA et al., 2018).

To support the agribusiness sector, regarding planning and management strategies, it is necessary to expand the forecast of transport expenses, especially road ones. The use of forecast models has been used for different situations, such as in predicting the national production of agricultural and road machinery in Brazil (MARTINS et al., 2020), in the calculation of total costs of a transport network (JANIC, 2007), and for modeling and simulation of design and development processes (MARTH et al., 2021). Concerning the predictive models for road freight, previous studies do it more commonly by the use of different statistical techniques (COOPE et al., 2020; GÜLER, 2014), evaluating the behavior of variables such as road quality (TUNDE& ADENIYI, 2012), economic instability (MOSCHOVOU& GIANNOPOULOS, 2021), and hybrid models of price forecasting (MENHAJ& KAVOOSI-KALASHAMI, 2022).

This being said, this research evaluated whether the behavior of soybean road freight is influenced by the variables that make up the productive, economic, and infrastructure, dimensions in price formation. In addition, to predict the road freight price of agricultural grains by multiple regression models and compare their forecasting power. The hypothesis we want to verify is whether prediction techniques, when submitted to a diverse data set, are capable of generating more assertive models to predict the behavior of the freight price of agricultural grains.

This paper seeks to shed light on the price behavior of road freight for agricultural grains. The objective is to evaluate if the behavior of soybean road freight is influenced by the variables that make up the productive, economic and infrastructure dimensions in price formation. The intersection of variables in three different dimensions is not usual in the literature but is necessary to analyze the real effect on the price of freight, which is one of the contributions of this research. Furthermore, the analysis of grain transport from isolated flow data does not represent the impact on the structure of the road network (FRIED, et al. 2018) and this research advances in this direction.

Since grain production is unstable and dependent on weather factors that are sometimes difficult to predict, grain logistics chains must be sufficiently reliable and elastic (PITTMAN et al., 2020), i.e. reliable, flexible and agile. Therefore, it is necessary to develop methods capable of analyzing the predictions for road freight price of agricultural grains by multiple regression models and compare their forecasting power.

The hypothesis we want to verify is whether prediction techniques, when submitted to a diverse data set, are capable of generating more assertive models to predict the behavior of the freight price of agricultural grains.

A better predictability of logistical costs and a better understanding of the dynamics of freight price formation helps industry agents in planning and decision-making, and are important tools to anticipate fluctuations in the road transportation market.

MATERIALS AND METHODS

The data set used to develop the soybean road freight price forecast models is based on the road routes (origin/destination) per municipality (Table 1).

The data was collected from official sources within a 2012 to 2019 time frame which covered monthly and yearly information period, that resulted in $n = 19{,}400$ datasets. The freight price and the distance traveled by route were based on the Freight Information System (SIFRECA). The data related to production aspects had the Brazilian Institute of Geography and Statistics (IBGE) as a source. The harvest period, storage and soybean domestic price were based on data from the National Supply Company (CONAB). The soybean international price was from the Chicago Board of Trade (CBOT) and fuel prices from The Brazilian National Agency of Petroleum, Natural Gas and Biofuels (ANP). The inflation rates were obtained from the IBGE and the

Table 1 - Soybean database for principal component analysis.

Data	-------Geographic Outline-------	----------------Unit----------------	---------------Source---------------
Freight	Route	R$/t	SIFRECA (2021)
---------Productive Dimension---------			
Harvest Period	State	Harvest	CONAB (2021)
Planted Area	State	ha	IBGE (2021)
Harvested Area	State	ha	IBGE (2021)
Production	State	t	IBGE (2021)
Yield	State	kg/ha	IBGE (2021)
---------Infrastructure Dimension---------			
Distance	Route	km	SIFRECA (2021)
Warehousing	Origin State	thousand t	CONAB (2021)
Warehousing	Destination State	thousand t	CONAB (2021)
Crushing	Origin State	t/day	ABIOVE (2021)
Crushing	Destination State	t/day	ABIOVE (2021)
---------Economic Dimension---------			
Domestic Price	State	R$/bag (60kg)	CONAB (2021)
International Price	-	US$/t	CONAB (2021)
Average Diesel Price	Origin State	R$/l	ANP (2021)
Average Ethanol Price	Origin State	R$/l	ANP (2021)
Average Oil Price	Origin State	R$/l	ANP (2021)
Inflation rate IGPM	Brazil	%	FGV (2021)
Inflation rate IPCA	Brazil	%	IBGE (2021)
Exchange Rate	US Dollar	R$	Bacen (2021)
Soybean Export	Origin State	t	Comex Stat (2021)
Diesel Oil Import	Brazil	m^3	Comex Stat (2021)

Getulio Vargas Foundation (FGV) and the exchange rate from the Central Bank of Brazil (Bacen). Finally, soybean exports and diesel imports were sourced from the Brazilian foreign trade statistics system (Comex Stat) linked to the Ministry of Development, Industry and Foreign Trade (MDICS) of Brazil (Table 1).

For the Principal Component Analysis (PCA) a database was used which counted on 22 variables grouped into three dimensions. The variables were grouped into different dimensions because they present similarities in relation to productive, economic, and infrastructure aspects. Each dimension is composed of the independent variables that were tested in the regression model. For example, the productive dimension was composed of variables related to soybean production (harvest period, harvested area and production) that can interfere in the formation of the freight price (Figure 1).

In this research we used two techniques. Principal Component Analysis (PCA) and multiple linear regression. Principal component analysis was used to reduce the dimensionality of the data set and subsequent analysis in multiple linear regression (Table 2).

PCA is a technique used to group variables according to behavior in a given data set. Considering a data matrix $X_{n \times p}$ represented by p variables of n individuals from a π population (MANLY& ALBERTO, 2016).

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \quad (1)$$

Obtaining the principal components represented by the set of p variables $X_1, X_2, \ldots, X_p$ with averages $\mu_1, \mu_2, \ldots, \mu_p$ and variance $\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2$, since these variables are not independent and have the covariance between the i-th and k-th variable defined by σ_{ik}, for every $i \neq k = 1,2, \ldots p$. Thus, p variables can be represented in vector form by $X = [X_1, X_2, \ldots, X_p]'$ with the same vector of averages $\mu = [\mu_1, \mu_2, \ldots, \mu_p]'$ and with the same covariance matrix Σ.

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{1p}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{p1}^2 & \cdots & \sigma_{pp}^2 \end{bmatrix} \quad (2)$$

The pairs of eigenvalues and eigenvectors $(\lambda_1, e_1), (\lambda_2, e_2), \ldots, (\lambda_p, e_p)$, in which $\lambda_1 \geq \lambda_2, \ldots, \lambda_p$ is associated with $\sum$ and then the i-th principal component is defined by (HONGYU et al., 2015).

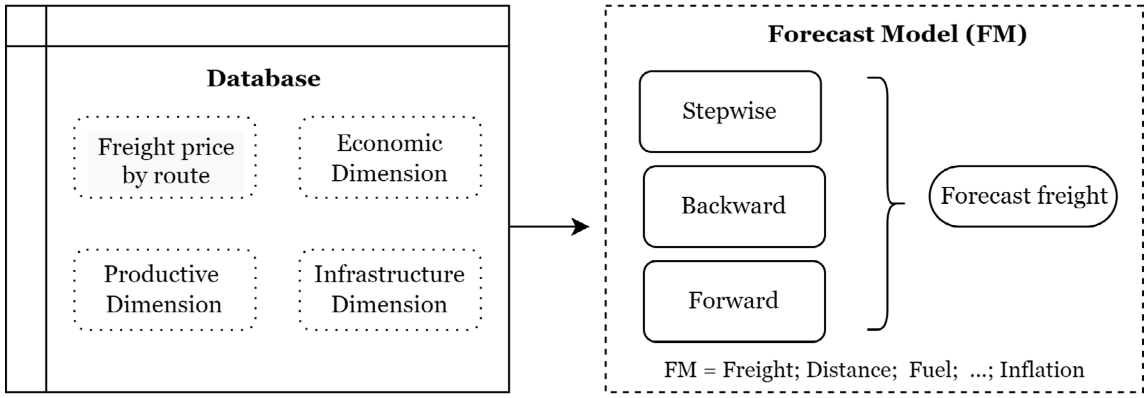

Figure 1 - Diagram of the soybean road freight forecast model.

$$Z_i = e_I^T X = e_{i1}X_1 + e_{i2}X_2 + ? + e_{ip}X_p \quad (3)$$

where $i = 1,2, \dots, p$.

As the variable Z_i is not obtained from the experiment, it is considered latent. Therefore, the total variability contained in the original variables is equal to the total variability contained in the PCs (JOHNSON & WICHERN, 2013). The contribution of each PC (Z_i) is expressed as a percentage, and the individual contribution of each PC can be calculated, for example, for the k-th PC.

$$C_k = \frac{Var\,(Z_i)}{\Sigma_{i=1}^{p} Var\,(Z_i)} \times 100 \quad (4)$$

After re-expressing the data in a new space (Figure 1), since the original data is rotated and translated, each axis Z_1 and Z_2 represents a principal component (Figure 2). The axes are selected according to their variance and, after rotation, the new generated variables, PCs, will be independent (JOHNSON & WICHERN, 2013).

Although, for inferential purposes a multivariate normal distribution of the dataset is usually assumed, PCA as a descriptive tool needs no distributional assumptions (JOLLIFFE & CADIMA, 2016). It is based on the calculation of eigenvalues and high vectors and only requires that the matrix of variances, covariances and correlations be positive definite (MANLY & ALBERTO, 2016), these conditions are met by our data.

To obtain a forecast of freight costs, it was necessary to relate two or more independent variables (predictors) to predict a result or dependent variable (freight price). The correlation provides a quantitative way to measure the degree or strength of a relationship between two variables, and the regression analysis mathematically describes this relationship, such that the observed Y and predicted Y have the maximum possible correlation in the simple or multiple regression model. The regression analysis makes it possible to predict the value of a dependent variable based on the value of at least one independent variable (RENCHER& SCHAALJE, 2007).

The response y is often influenced by more than one predictor variable. A linear model that relates the response to several independent variables (MONTGOMERY et al., 2012)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \mathcal{E} \quad (5)$$

The parameters β_0, β_1, ..., β_k are called regression coefficients. As in (5), $\mathcal{E}$ provides random variation in y not explained by variables x. This random variation may be due in part to other variables that affect y, but are not known or not observed. The model in (6) is linear in the parameters β, but it is not necessarily linear in the variables x. Thus, models such as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_2 + \beta_4 sinx_2 + \mathcal{E} \quad (6)$$

For forecasting, we only need to assume that the same correlations that existed when the data were collected also remain in place when the forecasts are made. Showing that there is a significant relationship between y and the variables x in (6) does not necessarily prove that the relationship is causal. To establish causality, the researcher must choose the values of the variables x in the model and use randomization to avoid the effects of other variables KAIRA, 2017; RENCHER & SCHAALJE, 2007).

Correlation and causality are fundamental to multiple linear regression modes as they allow to identify patterns and relationships that can be used to predict the future behavior of variables. Correlation is needed to identify patterns

Table 2 - Soybean database for regression models.

Data	--------Geographic Outline-------	-----------------Unit----------------	---------------Source---------------
Freight	Route	R$/t	SIFRECA (2021)
--------Productive Dimension--------			
Harvest Period	State	Harvest	CONAB (2021)
Harvested Area	State	ha	IBGE (2021)
Production	State	t	IBGE (2021)
--------Infrastructure Dimension--------			
Distance	Route	km	SIFRECA (2021)
Warehousing	Origin State	thousand t	CONAB (2021)
Warehousing	Destination State	thousand t	CONAB (2021)
--------Economic Dimension --------			
Domestic Price	State	R$/bag (60kg)	CONAB (2021)
International Price	-	US$/t	CONAB (2021)
Average Diesel Price	Origin State	R$/l	ANP (2021)
Average Ethanol Price	Origin State	R$/l	ANP (2021)
Average Oil Price	Origin State	R$/l	ANP (2021)
Inflation rate IGPM	Brazil	%	FGV (2021)
Inflation rate IPCA	Brazil	%	IBGE (2021)
Exchange Rate	US Dollar	R$	Bacen (2021)
Soybean Export	Origin State	t	Comex Stat (2021)

between variables while causality is needed to determine the relationship between the variables with the dependent variable. Without correlation and causality multiple linear regression would be much more difficult to perform (KAIRA, 2017; RENCHER & SCHAALJE, 2007).

Stepwise regression is usually used for exploratory studies, with an emphasis on describing the relationships between variables. Evaluating all possible regressions can be costly computationally, and several methods have been developed to evaluate only a small number of subset regression models, adding or deleting variables one at a time. These methods are generally referred to as step-by-step procedures, and they can be classified into three broad categories: Forward selection, Backward elimination, and Stepwise regression, the latter being a combination of the Forward and Backward procedures (MONTGOMERY et al., 2012).

In Forward selection, the procedure starts with the assumption that there are no other variables in the model other than the intercept, then a procedure is done to find an optimal subset by inserting variables into the model one at a time. The first variable selected for entry into the equation is the one that has the highest simple correlation with the response variable y. Suppose this variable is x_1; this is the variable that will have the highest value of the F statistic to test the significance of the regression. This variable is entered if the F statistic exceeds a preselected F value, say F_{IN}. Then, the second variable chosen to enter the model is the one that has the highest correlation with y after adjusting for the effect of the first variable inserted (x_1) in y, and we refer to these correlations as partial. They are the simple correlations between the residue of regression.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 \tag{7}$$

and residual of the regressions of the other candidate variable in x_1, which is

$$\hat{x}_j = \hat{a}_{0j} + \hat{a}_{1j} x_1 \tag{8}$$

where $j = 2,3, \ldots, k$.

Suppose that in the second step, the variable with the highest partial correlation with y is x_2; this implies that the greatest partial F statistic is if this value of F exceeds F value,

$$F = \frac{SS_R(x_2/x_1)}{MS_E(x_1, x_2)} \tag{9}$$

If the F value is greater than the F_{IN} (F verified in Table F according to the number of variables, degrees of freedom), then x_2 is added to the model. In general, at each step the variable with the highest partial correlation with y (or equivalently the highest partial F statistic, given the other variables already in the model) is added if its partial F statistic exceeds the preselected entry level F. The procedure ends when the partial F

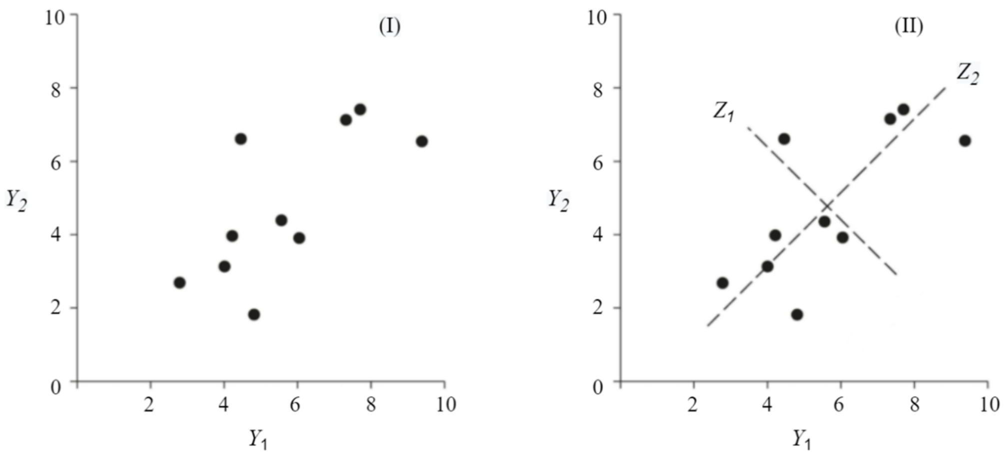

Figure 2 - Projection of a data set and the PCs.
(I): original data set. (II): new data set and axes after Principal Component Analysis (PCA).
Source: Adapted from JOHNSON & WICHERN (2013).

statistic in a specific step does not exceed *F* or when the last candidate variable is added to the model (MONTGOMERY et al., 2012).

Backward elimination attempts to find a good model working in the opposite direction, i.e., it does not start with any model that includes all K candidate variables. Next, the partial *F* statistic is calculated for each variable as if it were the last variable to enter the model. The smallest of these partial *F* statistics is compared with a preselected value, F_{out} (or *F* to remove), for example, and if the smallest partial *F* value is lower than *F* to remove, that variable is removed from the model. Now a regression model with k-1 variables is adjusted, partial *F* statistics for this new model are calculated, and the procedure is repeated. The Backward elimination algorithm ends when the smallest partial F value is not lower than the preselected cutoff value of *F* to remove (MONTGOMERY et al., 2012).

In a more simplified way, the flowchart (Figure 3) represents the Stepwise method, showing the steps along the construction of the model, when a variable enters or when a variable leaves the model.

For the linear regression model, the assumption of normality of the residuals, influential points and homogeneity of variances was evaluated and all were met. The need for transformation was verified using the optimal power method of Box-Cox, which suggested not transforming (DEAN & VOSS, 1999).

RESULTS AND DISCUSSION

In order to characterize the observations and verify the relationships between the variables used in the regression models, PCA was applied to the data set. The principal components were determinants to characterize the observations and verify the relationships between the variables used in the regression models. JOHNSON & WICHERN (2013) point out that, for applications in several areas of knowledge, the number of components used has accumulated 80% or more of the proportion of total variance.

In this principal component analysis, of the 22 variables considered, six main components were generated (Table 3). We verified that, from the first to the fifth component, the eigenvalue is above 1; from the Kaiser criterion (MANLY&ALBERTO, 2016); therefore, it is used to explain the dependent variable, freight price. Values lower than 1 are disregarded, which helps in the selection of components.

The score coefficient matrix of each component (Table 4) indicates which variables are included in each component. Based on this analysis, a nomenclature was defined for each component depending on the nature of these variables. Thus, it was possible to relate the dimension of greater interference in the formation of the price of soybean road freight, which involves from aspects related to agricultural production to the behavior of *commodity* prices in the international market. Therefore, we

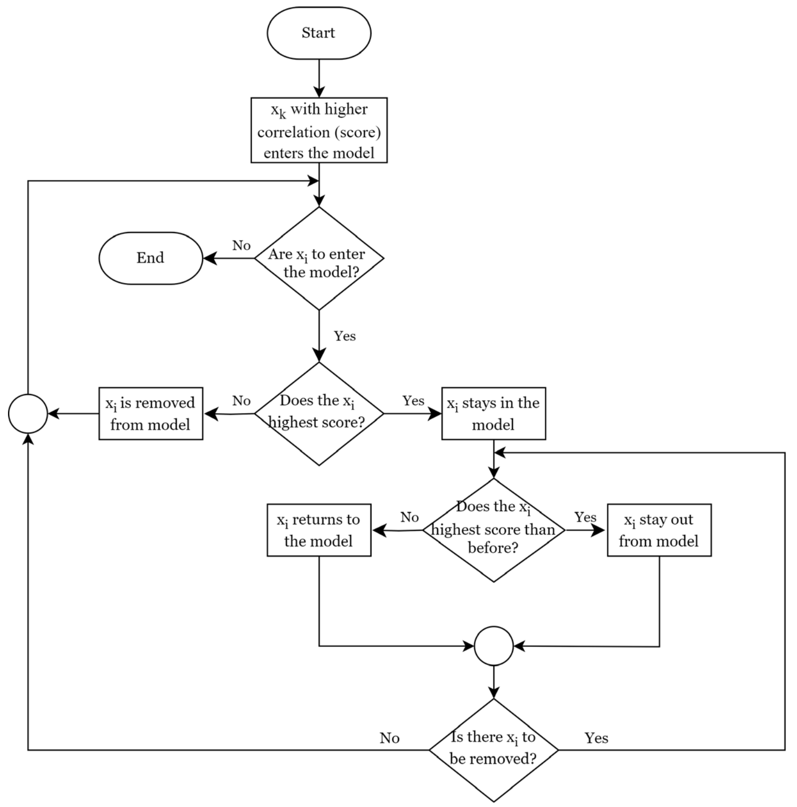

Figure 3 - Stepwise method flowchart.
Source: Adapted from Alves et al. (2013).

classify the components as: component 1 - production, component 2 - fuel, component 3 - processing, component 4 - inflation, component 5 - seasonality.

Component 1 - production, explained by the variables state production, planted area, and harvested area, has a highlighted importance in the behavior of the road freight price and positively or directly affect the formulation of freight price (MACHADO et al., 2019; PÉRA et al., 2018; CANGUSSU et al., 2013; MARTINS et al., 2004).

Component 2 - fuel, explained by the variables of average ethanol and diesel prices, has been widely studied as one of the main variables that influence the price of soybean road freight (MACHADO et al., 2019; PÉRA et al., 2018; CANGUSSU et al., 2013; MARTINS et al., 2004).

Component 3 - processing is explained by distance traveled, warehousing capacity at state destination and crushing capacity at state destination. MARTINS (2008) point out that the greater the industrial processing capacity, in this case soybean crushing, the greater the negotiation capacity among the agents in the sector, which directly influences the formulation of soybean road freight. DELAI et al. (2017) showed that the warehousing capacity of agricultural products allows a better marketing strategy and increases the profits of producers, who can expect moments when prices are on the rise. This dynamic of storing the product and taking advantage of times of higher prices and lower transport costs represents the windows of opportunity, which avoid the so-called “sales *rush*”, which causes a storage deficit the leads the producer to harvest and soon put the product on the market (OLIVEIRA, 2014).

The distance traveled has been the determining factor for the establishment of road freight prices (BEILOCK et al., 1986; SOARES& CAIXETA FILHO, 1997; ARAÚJO et al., 2014), and, in this analysis, it stood out as one of the variables that make up the third component.

Component 4 - inflation is explained by the monthly variation in inflation by IGPM and IPCA indices, corroborating with KUSSANO & BATALHA (2012), who stated that inflation is one of the factors

Table 3 - Total explained variation of soybean road freight price variables.

Component	------------Total------------	--------Difference--------	-------% of variation------	-------% cumulative-------
1	7.6022	3.3438	34.56	34.56
2	4.2584	1.6515	19.36	53.91
3	2.6069	0.4757	11.85	65.76
4	2.1312	0.9549	9.69	75.45
5	1.1763	0.3039	5.35	80.80
6	0.8724	0.0243	3.97	84.76
7	0.8481	0.1849	3.86	88.62
8	0.6633	0.1743	3.01	91.63
9	0.4890	0.0364	2.22	93.85
10	0.4526	0.1661	2.06	95.91
11	0.2865	0.1025	1.30	97.21
12	0.1839	0.0394	0.84	98.05
13	0.1446	0.0556	0.66	98.71
14	0.0890	0.0190	0.40	99.11
15	0.0700	0.0285	0.32	99.43
16	0.0415	0.0073	0.19	99.62
17	0.0343	0.0022	0.16	99.77
18	0.0320	0.0163	0.15	99.92
19	0.0157	0.0137	0.07	99.99
20	0.0020	0.0020	0.01	100.00
21	0.0000	0.0000	0.00	100.00
22	0.0000	-	0.00	100.00

that directly influences the formation of road freight price, in addition to being one of the main indices of road freight price correction.

Component 5 - seasonality is explained by the variables soybean price in the national market and harvest and off-season period. These variables have been widely studied as one of the main factors that influence the price of soybean road freight (MACHADO et al., 2019; PÉRA et al., 2018; CANGUSSU et al., 2013; MARTINS et al., 2004). The correlation between these variables happens because, with a large volume of product in the market in harvest periods, commodity prices fall; however, the freight price rises. In the off-season, the price of the commodity has a downward trend and road freight tends to rise (OLIVEIRA, 2006; KAVUSSANOS& ALIZADEH-M, 2001).

The components 1, 5 are part of the productive dimension. The component 3 is part of the infrastructure dimension. While components 2, 4 are included in the economic dimension. In this way the PCA results confirmed the multimensinality in the formation of the price of road freight. After the PCA, the variables were re-sized and the components generated, which also reaffirms the analyses grouped into three different dimensions. From these dimensions, the variables that were selected to make up the regression model were: freight, distance, production, harvested area, diesel price, ethanol price, origin warehousing, destination warehousing, soybean price in the national market, soybean price in the international market, export volume, IGPM and IPCA inflation indexes.

Multiple linear regression is widely used to understand which set of variables explain a dependent variable; in this case, three types of methods were performed as mentioned: Stepwise selection, Backward elimination, and Forward selection, to verify which model would better explain soybean road freight. The results are similar and there was no significant statistical difference between the methods. In table 5 we present the statics resulting from the Stepwise and the regression had an 87.20% explanatory power for soybean road freight.

In the infrastructure dimension, the distance traveled variable was the most important for predicting road freight. This variable has been widely established as the main one (COOPE et al., 2020; LIMA et al., 2016; MACHADO et al., 2019; MARTINS et al., 2004; SOARES & CAIXETA

Table 4 - Component score coefficient matrix.

Characteristic	-------C1-------	-------C2-------	-------C3-------	-------C4-------	-------C5-------
Distance	0.1562	0.0469	0.4256	−0.0942	0.1291
Production	0.3436	0.1161	0.0459	−0.0105	−0.0289
Yield	0.1534	−0.1318	0.1025	−0.0648	−0.3567
Planted area	0.3370	0.1533	0.0232	0.0028	0.0334
Harvested area	0.3371	0.1532	0.0238	0.0026	0.0313
Average diesel price	0.2435	−0.3279	−0.0018	−0.0293	−0.0166
Average ethanol price	0.1067	−0.3839	−0.1047	−0.0460	−0.0687
Average oil price	0.2007	−0.3770	−0.0557	−0.0436	−0.0523
Warehousing capacity at origin	0.3082	0.1927	−0.1150	0.0287	0.0476
Warehousing capacity at destination	0.1314	0.1343	−0.4543	0.1231	−0.1959
Domestic price	0.0677	−0.2219	−0.2792	0.0656	0.5285
International price	−0.1957	0.3320	0.0152	−0.0290	0.1438
Crushing capacity at origin	0.2788	0.2307	−0.1145	0.0558	0.0501
Crushing capacity at destination	0.1043	0.1646	−0.4512	0.1407	−0.2071
Inflation rate IGPM	−0.0098	−0.0314	0.1290	0.6535	0.1353
Inflation rate IPCA	−0.0098	−0.0314	0.1290	0.6535	0.1353
Exchange rate	0.2120	−0.3599	−0.0844	0.0307	0.0874
Diesel oil import	−0.1190	0.1998	0.0153	−0.0972	0.0546
Soybean export (annual)	0.2482	0.1471	0.1054	−0.0219	0.0487
Soybean export (monthly)	0.3031	0.1834	0.0616	−0.0103	0.0370
Harvest period	0.0067	−0.0264	0.2340	0.2566	−0.6243
Soybean freight	0.1981	−0.0241	0.4078	−0.0833	0.1367

Note: Variables that best explain the main components are highlighted.

FILHO, 1997) and the relevance of the distance traveled in the model is about 87.19% in the forecast of the road freight model; this was verified in both Stepwise regression and Forward selection. In this analysis, it was verified that, in all regressions, the infrastructure – distance dimension had greater relevance in the forecast of road freight.

The infrastructure dimension has been pointed out as one of the most important; however, regarding the specificity of each route, it may undergo some type of change in the formation of prices. Authors such as ARAUJO et al. (2014); CAIXETA FILHO (1997); CANGUSSU et al. (2013); MARTINS (2008); MARTINS et al. (2005); and PÉRA et al. (2018) have stated that the road freight market in Brazil and, specifically, that of agricultural cargo, does not undergo any type of mediation by the government concerning transport activities, which implies that prices are formed based on the free negotiation between supply and demand for the transport service.

The economic dimension also has a great impact on the formation of freight prices, mainly due to the variation in fuel prices and inflation (COOPE et al., 2020). And one can note that diesel and ethanol fuel prices are the ones with the highest coefficient in the regression equations. In this analysis, this dimension was the most important for the prediction of soybean road freight.

The productive dimension had an effect on the prediction of road freight. The seasonality of the harvest has effects on the fluctuation of the freight value, for example, at times of harvest when there is greater demand for transport services, prices tend to rise and, at the time of the off-season, prices tend to fall (CANGUSSU et al., 2013; COOPE et al., 2020; LIMA et al., 2016; MACHADO et al., 2019; MARTINS et al., 2004; PÉRA et al., 2018; MENHAJ& KAVOOSI-KALASHAMI, 2022).

The freight price can also differ according to the route. For instance, a region with a greater demand for transport has more expensive freight. There is evidence of movements occurring in quite diverse intensity; regions with better storage structure, as well as regions well served by more than one mode, might put less pressure on the market in the harvest periods, causing lower seasonality in the prices practiced in the freight market (DELAI et al., 2017).

In this analysis variables grouped in three dimensions were used and the regressions were

Table 5 - Stepwise regression statistics.

Characteristic	Estimated Parameter	--Standard Error--	Type II Sum of Squares	-------F Value------	-------Pr > F-------
Intercept	−58.23939	3.17659	190140	336.13	<.0001
Distance	0.09554	0.00037916	35912224	63486.5	<.0001
Production	0.00000150	1.282537E-7	77033	136.18	<.0001
Harvested Area	−0.00000659	4.849062E-7	104342	184.46	<.0001
Diesel Price	39.28108	0.92817	1013140	1791.05	<.0001
Ethanol Price	-14.52713	0.62307	307502	543.61	<.0001
Origin Storage	0.00090885	0.00005538	152330	269.29	<.0001
Destination Storage	−0.00056720	0.00001935	486259	859.62	<.0001
National Price	0.08434	0.02749	5325.26919	9.41	0.0022
Chicago Price	0.02364	0.00354	25266	44.67	<.0001
Export	−2.12684E-7	−7.11984E-8	5047.67264	8.92	0.0028
IGPM Inflation	2.07549	0.32577	22960	40.59	<.0001
IPCA Inflation	1.37251	0.69395	2212.78990	3.91	0.0480

Note: n = 19,400; significance level a = 0,05.

explanatory to soybean road freight by 87.2%, a percentage that is very explanatory in this type of analysis and for the sample size used.

The three models have the same behavior. In the analysis of root mean squared error (RMSE), the best models were Stepwise and Backward, which do not include the gasoline variable in the model. However, this difference is not significant (Table 6). Therefore, all models can be used for freight forecasting.

The mean absolute percentage error (MAPE) in all three models is 21%, which indicates that the model is well explanatory for freight forecasting (MAPE ranges from 0 to 100%, a lower MAPE indicates that the model is more accurate) (Table 6).

Considering the R^2_{ajusted} we see that on average 87% of the freight is explained by the independent variables. In absolute values the Forward model (87.2%) achieved the best value.

Figure 4 show the freight estimated using Stepwise Regression, Forward Selection and Backward Elimination and the real freight. It can be seen that this variation is mainly due to the distance traveled. In general, in the three models the estimated freight compared to the real freight and the estimated freight for distances up to 600 km had the same behavior.

For distances greater than 1,200 km the estimated freight was lower in all three models, and in this range the freights are more dispersed. This may indicate that the use of partitioned models for two distance ranges could generate better results for longer distances. The construction of a database for routes up to 600 km and another for longer distances could generate more efficient models for distances greater than 600 km.

In the soybean road freight predictive model there was no significant difference among the three models, all were highly explanatory in predicting freight price, with an R^2_{ajusted} of over 87%. HRYNKIEWICZ et al., (2019) did a study on structure activity prediction of inhibitory dipeptides using linear regression, Backward and Forward regression, and there were no significant differences between the methods and both were efficient in making the prediction in this chemometric approach.

CONCLUSION

The financial impact of transport felt by economic agents operating in agricultural chains could be reduced with the implementation of a freight management system. Soybean production continues to grow in Brazil, but with the challenge of maintaining competitiveness in the face of increasing production volumes that will continue to travel long distances to the main exporting ports. Knowing the variables capable of influencing freight price is necessary to support the decision-making regarding logistics, and, with the application of PCA, it was possible to identify seven main components capable of expressing the price of soybean road freight.

Table 6 - Statistical results.

Model	----------RMSE----------	----------MAPE----------	---------$R^2_{adjusted}$----------	--------------C_p-----------
Stepwise Regression	565.66709	21.10251	0.8719	11.5643
Backward Elimination	565.66709	21.06967	0.8719	11.5643
Forward Selection	565.68287	21.09712	0.8720	13.1050

Note: n = 19,400, significance level α = 0.05; RMSE: root mean squared error; MAPE: mean absolute percentage error.

The PCs were named according to the nature of the variables with the greatest contribution, being: 1. production, 2. fuel, 3. processing, 4. inflation, and 5. seasonality. It is worth mentioning that the distance traveled is one of the variables that has been directly linked to the price of soybean road freight. The application of PCA also showed the importance of considering other variables that had relevance in the components.

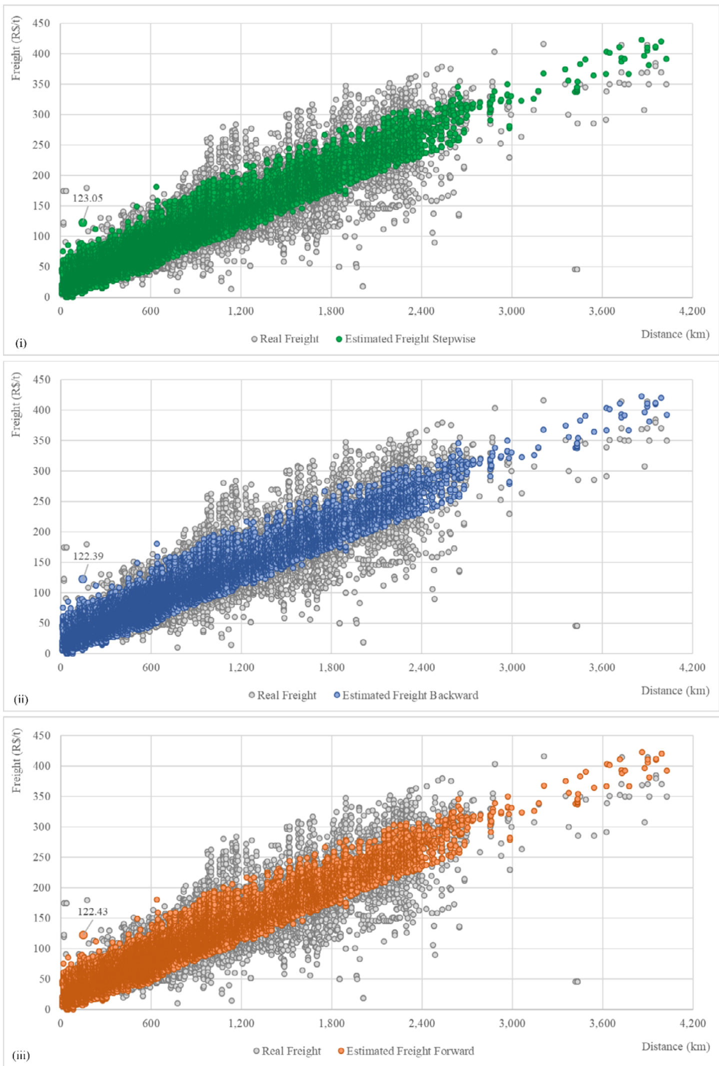

Figure 4 - Real and estimated freight by forecasting methods.

The three dimensions were relevant in the prediction of soybean road freight with the use of Stepwise multiple regression, in the following variables: distance traveled, production, harvested area, diesel price, ethanol price, warehousing capacity at origin, warehousing capacity at destination, domestic price, international price (Chicago/CBOT), soybean exports, IPCA inflation and IGPM inflation. According to the three regression equations obtained, the price of diesel and ethanol were the main factors for determining soybean road freight, and the three dimensions are responsible for about 87.2% of the explanatory power of the model, a very expressive percentage for this type of analysis.

After comparing the three models obtained via Stepwise, it was found that the three follow the same trend for price behavior and distance traveled. Prices have one behavior for distances up to 600 km and another for higher distances.

One of the contributions of our research is provide information for policy makers and agribusiness managers to more assertive predictability of logistical costs and better understanding of the dynamics of freight price formation helps industry agents in planning and decision-making. Another contribution of this study is that it can be used as a practical tool for predicting soybean road freight on several transportation routes.

ACNOWLEDGEMENTS

This research was financed in part by grant #2018/19571-1, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

DECLARATION OF CONFLICT OF INTEREST

We have no conflict of interest to declare.

AUTHORS' CONTRIBUTIONS

All authors contributed equally for the conception and writing of the manuscript. All authors critically revised the manuscript and approved of the final version.

REFERENCES

ABIOVE. Associação Brasileira das Indústrias de Óleos Vegetais. **Estatísticas**. Available from: <https://abiove.org.br/estatisticas>. Accessed: Jun. 20, 2021.

ALVES, M. F. et al. Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas. **Proceeding Series of the Brazilian Society of Applied and Computational Mathematics**, v.1, n.1, 2013. Available from: <https://proceedings.sbmac.org.br/sbmac/article/view/144/0>. Accessed: Sept. 10, 2021. doi: 10.5540/03.2013.001.01.0144.

ANP. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. **Série histórica do levantamento de preços**. 2021 Available from: <https://www.gov.br/anp/pt-br/assuntos/precos-e-defesa-da-concorrencia/precos/precos-revenda-e-de-distribuicao-combustiveis/serie-historica-do-levantamento-de-precos>. Accessed: Jun. 20, 2021.

ARAÚJO, M. D. P. S. et al. Custos e fretes praticados no transporte rodoviário de cargas: uma análise comparativa entre autônomos e empresas. **Journal of Transport Literature**, v.8, n.4, p.187-226, 2014. Available from: <https://www.scielo.br/j/jtl/a/R3BfL94BRgMV4TxCMsytHxr/abstract/?lang=pt>. Accessed: Sept. 10, 2021. doi: 10.1590/2238-1031.jtl.v8n4a8.

BACEN. Banco Central do Brasil. **Cotações e Boletins**. Available from: <https://www.bcb.gov.br/estabilidadefinanceira/historicocotacoes>. Accessed: Jun. 20, 2021.

BEILOCK, R. et al. Freight charge variations in truck transport markets: price discrimination or competitive pricing? **American Journal of Agricultural Economics**, v.68, n.2, p.226-236, 1986. Available from: <https://www.jstor.org/stable/1241424>. Accessed: Sept. 15, 2021. doi: 10.2307/1241424.

BRASIL. Ministério do Desenvolvimento, Indústria, Comércio e Serviços (MDICS). ComexStat. Estatística do Comércio Exterior Brasileira. **Exportações e Importações**. Available from: <http://comexstat.mdic.gov.br/pt/home>. Accessed: Jun. 20, 2021.

CANGUSSU, S. V. et al. Determinantes do preço do frete rodoviário de soja em grãos em Mato Grosso. **Revista de Economia da UEG**, v.9, n.1, p.78-94, 2013. Available from: <https://www.revista.ueg.br/index.php/economia/article/view/1425>. Accessed: Aug. 09, 2021.

CONAB. Companhia Nacional de Abastecimento. **Informações Agropecuárias**. Available from: <https://www.conab.gov.br/info-agro>. Accessed: Jun. 20, 2021.

COOPE, L. M. et al. Precificação de fretes rodoviários via modelos paramétricos. **Revista Produção Online**, v.20, n.4, p.1214-1237, 2020. Available from: <https://www.producaoonline.org.br/rpo/article/view/4021#:~:text=As%20vari%C3%A1veis%20preditoras%20que%20comp%C3%B5em,rotina%20dos%20analistas%20da%20empresa>. Accessed: Oct. 05, 2021. doi: 10.14488/1676-1901.v20i4.4021.

DEAN, A. M.; VOSS, D. T. **Design and analysis of experiments**. New York: Springer, 1999.

DELAI, A. P. D. et al. Armazenagem e ganhos logísticos: uma análise comparativa para comercialização da soja em mato Grosso do Sul. **Revista em Agronegócio e Meio Ambiente**, v.10, n.2, p.395-414, 2017. Available from: <https://periodicos.unicesumar.edu.br/index.php/rama/article/view/4579>. Accessed: Aug. 05, 2021. doi: 10.17765/2176-9168.2017v10n2p395-414.

FGV. Fundação Getulio Vargas. **Índice Geral de Preços**. Available from: <https://portalibre.fgv.br/igp>. Accessed: Jun. 20, 2021.

FRIED, T. et al. Evolving supply chains and local freight flows: a geographic information system analysis of Minnesota cereal grain movement. **Transportation Research Record**, v.2672, n.9, p.1-11, 2018. Available from: <https://journals.sagepub.com/doi/10.1177/0361198118759952>. Accessed: Mar. 06, 2023. doi: 10.1177/0361198118759952.

GÜLER, H. An empirical modelling framework for forecasting freight transportation. **Transport**, v.29, n.2, p.185-194, 2014. Available from: <https://journals.vilniustech.lt/index.php/Transport/article/view/1896>. Accessed: Nov. 10, 2022. doi: 10.3846/16484142.2014.930927.

HONGYU, K. et al. Análise de Componentes Principais: resumo teórico, aplicação e interpretação. **Engineering and Science**, v.1, n.5, p.83-90, 2015. Available from: <https://periodicoscientificos.ufmt.br/ojs/index.php/eng/article/view/3398>. Accessed: Dec. 15, 2020. doi: 10.18607/ES201653398.

HRYNKIEWICZ, M. et al. Structure-Activity Prediction of ACE Inhibitory/Bitter Dipeptides-A Chemometric Approach Based on Stepwise Regression. **Molecus**, v.24, n.5, p.1-13, 2019. Available from: <https://www.mdpi.com/1420-3049/24/5/950>. Accessed: Jan. 05, 2022. doi: 10.3390/molecules24050950.

IBGE. Instituto Brasileiro de Geografia e Estatística. **Banco de Tabelas Estatísticas**. Available from: <https://sidra.ibge.gov.br/home/ipp/brasil>. Accessed: Jun. 20, 2021.

JANIC, M. Modelling the full cost of an intermodal and road freight transport network. **Transportation Research Part D**, v.12, p.33-44, 2007. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S1361920906000794>. Accessed: Jan. 15, 2021. doi: 10.1016/j.trd.2006.10.004.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 6 ed. New York: Pearson, 2013.

JOLLIFFE, I. T.; CADIMA, J. **Principal Component Analysis: a beginner´s guide to PCA**. New York: Springer, 2016.

KAIRA, A. Decoding the Bland-Altman Plt: Basic Review. **Journal of the Practice of Cardiovascular Sciences**, v.3, n.1, p.36-8, 2017. Available from: <https://www.j-pcs.org/article.asp?issn=2395-5414;year=2017;volume=3;issue=1;spage=36;epage=38;aulast=Kalra>. Accessed: Oct. 05, 2021. doi: 10.4103/jpcs.jpcs_11_17.

KAVUSSANOS, M.G.; ALIZADEH, A. H. Efficient pricing of ships in the dry bulk sector of the shipping industry, **Maritime Policy & Management**, v.29, n.3, p.303-330, 2002. Available from: <https://www.tandfonline.com/doi/abs/10.1080/03088830210132588>. Accessed: Jan. 15, 2021. doi: 10.1080/03088830210132588.

KUSSANO, M. R.; BATALHA, M. O. Custos logísticos agroindustriais: avaliação do escoamento da soja em grão do Mato Grosso para o mercado externo. **Gestão & Produção**, v.19, n.3, p.619-632, 2012. Available from: <https://www.scielo.br/j/gp/a/fsxNvjK6dF9wBmYLY97dPdF/?lang=pt>. Accessed: Dec. 11, 2020. doi: 10.1590/S0104-530X2012000300013.

LIMA, L. M. et al. Fertilizer freight rate disparity in Brazil: a regional approach. **International Food and Agribusiness Management Review**, v.19, n.4, p.109-128, 2016. Available from: <https://www.wageningenacademic.com/doi/10.22434/IFAMR2015.0109>. Accessed: Feb. 12, 2021. doi: 10.22434/IFAMR2015.0109.

MACHADO, R. S. et al. Logística da BR-163 nas exportações de soja da Cooperlucas, Mato Grosso. **Revista Fatec Zona Sul**, v.5, n.4, p.1-12, 2019. Available from: <http://www.revistarefas.com.br/index.php/RevFATECZS/article/view/308 >. Accessed: May, 20, 2021.

MANLY, B. F. J.; ALBERTO, J. A. N. **Multivariate statistical methods: a primer**. 4 ed. Boca Raton: Chapman and Hall/CRC Press, 2016.

MARTH. S. et al. Stepwise modelling method for post necking characterisation of anisotropic sheet metal. **Modelling and Simulation in Materials Science and Engineering**, v.29, 2021. Available from: <https://iopscience.iop.org/article/10.1088/1361-651X/ac2797>. Accessed: Jan. 10, 2022. doi: 10.1088/1361-651X/ac2797.

MARTINS, R. S. Estudo da formação do frete rodoviário e potencial de conflitos em negociações em cadeias do agronegócio brasileiro. **Organizações Rurais & Agroindustriais**, v.10, n.1, p.75-89, 2008. Available from: <http://revista.dae.ufla.br/index.php/ora/article/view/93>. Accessed: Feb. 12, 2021.

MARTINS, T. et al. Persistence effect determination of variability in forecasting of agricultural and road machinery national production. **Ciência Rural**, v.50, n.6, e20190631, 2020. Available from: <https://www.scielo.br/j/cr/a/rhPqbvLrYtJFmmzpLcFZgJJ/?lang=en>. Accessed: Aug. 10, 2021. doi: 10.1590/0103-8478cr20190631.

MARTINS, R. S. et al. Sazonalidade nos fretes e preferências dos embarcadores. **Revista de Economia e Administração**, v.4, n.1, p.68-96, 2005. Available from: <http://www.spell.org.br/documentos/ver/25724/sazonalidade-nos-fretes-e-preferencias-dos-embarcadores-no-mercado-de-transporte-de-graneis-agricolas/i/pt-br>. Accessed: Feb. 12, 2021. doi: 10.11132/rea.2002.86.

MARTINS, R. S. et al. Formação de preços e sazonalidade no mercado de fretes rodoviários para produtos do agronegócio no estado do Paraná. **Revista Paranaense de Desenvolvimento**, v.106, p.113–136, 2004. Available from: <https://www.researchgate.net/publication/277848170_Formacao_de_precos_e_sazonalidade_no_mercado_de_fretes_rodoviarios_para_produtos_do_agronegocio_no_estado_do_Parana>. Accessed: Feb. 12, 2021. doi: 10.22434/IFAMR2015.0109.

MENHAJ, M., KAVOOSI-KALASHAMI, M. Developing a hybrid forecasting system for agricultural commodity prices (case study: Thailand rice free on board price). **Ciência Rural**, v.52, n.8, e20201128, 2022. Available from: <https://www.scielo.br/j/cr/a/mcdxrp7wjNdgBTZGbNZTJsh/abstract/?lang=en>. Accessed: Nov. 10, 2022. doi: 10.1590/0103-8478cr20201128.

MONTGOMERY, D. C. et al. **Introduction to Linear Regression Analysis**, 5 ed. Nova York: Wiley, 2012.

MOSCHOVOU, T. P., GIANNOPOULOS, A. G. Road freight transportation in a period of economic instability: A panel data study in four EU Mediterranean countries. **Research in Transportation Business & Management**, v.41, 100622, 2021. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S2210539521000055>. Accessed: Feb. 10, 2022. doi: 10.1016/j.rtbm.2021.100622.

OECD/FAO. Organisation for Economic Co-operation and Development/Food and Agriculture Organization. **Agricultural Outlook 2021-2030**, Paris: OECD Publishing, 2021. Available from: <https://www.oecd-ilibrary.org/agriculture-and-food/oecd-fao-agricultural-outlook-2021-2030_19428846-en>. Accessed: Jan. 12, 2022. doi: 10.1787/19428846-en.

OLIVEIRA, A. L. R. et al. Performance evaluation of agricultural commodity logistics from a sustainability perspective. **Case Studies on Transport Policy**, v.10, n.1, p.674-685, 2022. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S2213624X22000335?via%3Dihub>. Accessed: Oct. 10, 2022. doi: 10.1016/j.cstp.2022.01.029.

OLIVEIRA, A. L. R. et al. Logistical transportation routes optimization for Brazilian soybean: an application of the origin-destination matrix. **Ciência Rural**, v.51, n.2, e20190786, 2021. Available from: <https://www.scielo.br/j/cr/a/w4jKvsp5ZmcjFdS8ystPJDc/?lang=en>. Accessed: Dec. 12, 2021. doi: 10.1590/0103-8478cr20190786.

OLIVEIRA, A. L. R. Perfil da logística de transporte de soja no Brasil. **Informações Econômicas**, São Paulo, v.36, n.1, p.17-25, 2006. Available from: <https://docplayer.com.br/20028335-Perfil-da-logistica-de-transporte-de-soja-no-brasil-1.html>. Accessed: Feb. 12, 2021.

OLIVEIRA, A. L. R. A logística do agronegócio: para além do apagão logístico. In: BUAINAIN, A. M. (et al.). (org.). **O mundo rural no Brasil do século 21: a formação de um novo padrão agrário e agrícola**. 1ed. Brasília: Embrapa, 2014, p.337-370. Available from: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/994073/o-mundo-rural-no-brasil-do-seculo-21-a-formacao-de-um-novo-padrao-agrario-e-agricola>. Accessed: Jun. 20, 2021.

OLIVEIRA, A. L. R. et al. Estimativa do custo rodoviário da soja: uma análise da rota Sorriso-Santos. **Revista de Economia e Agronegócio**, v.11, n.2, p.255-276, 2013. Available from: <https://periodicos.ufv.br/rea/article/view/7545>. Accessed: Jun. 12, 2021. doi: 10.25070/rea.v11i2.221.

PÉRA, T. G. et al. Análise dos impactos da Medida Provisória nº 832 de 2018 (Política de Preços Mínimos do Transporte Rodoviário de Carga) na logística do agronegócio brasileiro. **Logística do Agronegócio: Desafios e Oportunidades**, v.3, 2018, 37p. Available from: <https://esalqlog.esalq.usp.br/categoria/serie-logistica-do- agronegocio>. Accessed: Feb. 1, 2021.

PITTMAN R. et al. The effectiveness of EC policies to move freight from road to rail: Evidence from CEE grain markets. **Research in Transportation Business & Management**, v.37, 100482, 2020. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S2210539519303633?via%3Dihub>. Accessed: Mar. 06, 2023. doi: 10.1016/j.rtbm.2020.100482.

RENCHER, A. C.; SCHAALJE, G. B. **Linear Models in Statistics**. 2 ed. Nova York: Wiley, 2007.

SIFRECA. Sistema de Informações de Frete. **Fretes rodoviários de soja**. Available from: <https://sifreca.esalq.usp.br/>. Accessed: Jun. 20, 2021.

SOARES, M. G.; CAIXETA FILHO, J. V. Caracterização do mercado de fretes rodoviários para produtos agrícolas. **Gestão & Produção**, v.4, n.2, p.186-204, 1997. Available from: <https://www.scielo.br/j/gp/a/jx9jxbMFz46pc3xYq4JBW9y/?lang=pt>. Accessed: Feb. 12, 2021. doi: 10.1590/S0104-530X1997000200007.

TUNDE, A. M.; ADENIYI, E. E. Impact of Road Transport on Agricultural Development: A Nigerian Example. **Ethiopian Journal of Environmental Studies and Management**, v.5, n.3, p.232-238, 2012. Available from: <https://www.ajol.info/index.php/ejesm/article/view/77955>. Accessed: Aug. 20, 2021. doi: 10.4314/ejesm.v5i3.3.