



Persistence effect determination of variability in forecasting of agricultural and road machinery national production

Tailon Martins^{1*}  Alisson Castro Barreto²  Daniel Arruda Coronel³  Luciane Flores Jacobi⁴ 
Valentina Wolff Lirio⁵  Adriano Mendonça Souza⁴ 

¹Programa de Pós-graduação em Engenharia de Produção, Universidade Federal de Santa Maria (UFSM), 97105-900, Santa Maria, RS, Brasil. E-mail: martins.tailon@gmail.com. *Corresponding author.

²Programa de Pós-graduação em Engenharia de Produção, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brasil.

³Departamento de Economia e Relações Internacionais, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brasil.

⁴Departamento de Estatística, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brasil.

⁵Graduação em Estatística, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brasil.

ABSTRACT: *The objective of this research was to forecast the Brazilian national production of agricultural and road machinery in the short term by BOX & JENKINS methodology and determine the persistence effect. Data were obtained at National Association of Automotive Vehicle Manufacturers (ANFAVEA) from January 1960 to October 2019, totaling 718 monthly observations. The Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Conditional Heteroscedasticity (ARCH) methodology were used. The ARIMA (2,1,1)-ARCH (2) model was fitted and persistence of 0.60 was determined, showing that the instability in the series will be for a long period of time.*

Key words: *time series, prediction, volatility, agricultural machinery, road machinery.*

Determinação do efeito de persistência da variabilidade na previsão da produção nacional de máquinas agrícolas e rodoviárias

RESUMO: *O objetivo desta pesquisa é prever a produção nacional de máquinas agrícolas e rodoviárias no Brasil, no curto prazo por meio da metodologia BOX & JENKINS e determinar o efeito de persistência na série. Os dados foram obtidos no site da Associação Nacional dos Fabricantes de Veículos Automotores (ANFAVEA) no período de janeiro de 1960 a outubro de 2019, totalizando 718 observações mensais. Os modelos Autoregressivos Integrados e de Médias Móveis (ARIMA) e de Heteroscedasticidade Condicional Autoregressiva (ARCH) foram utilizados para ajustar a média e a variabilidade da série. O modelo ARIMA(2,1,1) - ARCH(2) foi selecionado por meio das estatísticas de ajustes e a persistência determinada foi de 0,60 mostrando que a instabilidade na série é duradoura.*

Palavras-chave: *séries temporais, previsão, volatilidade, máquinas agrícolas, máquinas rodoviárias.*

INTRODUCTION

The product industry assumed a prominent position in terms of technological development, especially with regard to the agricultural machinery and implements segment (GONÇALVES et al., 2015).

Performing an advanced search of the Web of Science database using the search string TS = ((ARIMA OR forecasting OR forecast) AND (“road machinery” OR “agricultural machinery”)) it was reported 9 papers, all of them written in the English language. Of these 9 papers 3 are in automation control

systems, 3 in computer science, interdisciplinary applications, 2 are in multidisciplinary agriculture, 2 are in management, 2 are in mechanical engineering and 1 in computer science, artificial intelligence, 1 in electrical engineering, 1 in manufacturing engineering, 1 in computer science information systems, 1 in multidisciplinary engineering.

Brazil is considered one of the most productive country addressing agrobusiness, but no one related to forecast agricultural and road machineries. We consider relevant to estimate the demand for agricultural tractors using univariate

linear model.

AI(2015) analyzing economic development as a function of agricultural mechanization, used nonlinear relationship models to predict the demand based on GDP. In this study the joining of linear models to study the level and a nonlinear to study the volatility will be used, to determine the persistence and the forecast in a short term.

JU et al., (2013) used agriculture-related data to make predictions through neural network capability and fuzzy logic, exploiting the full potential power of agricultural machinery. Here models of the general class ARIMA and ARCH will be used together.

The technical evolution present in this sector has generated supply of equipment that uses advanced technology, which contributes more and more to the field productivity (VIAN et al., 2013). One of the alternatives for subsidizing production control and planning systems is predicting using time series analysis (MARCHESAN & SOUZA, 2010).

The volatility of the world economy affects the predicted values, both in the short and long term. The BOX & JENKINS models are able to capture the series level behavior and useful to forecast it, and volatility models capture the variability on the series. Total agricultural and road machinery corresponds to wheeled tractors, crawler tractors, motorized cultivators, grain harvesters, cane harvesters and backhoes, in units.

The researches problems are identifying if the agricultural and road machinery production series can be represented in level by a linear model from an ARIMA family and its variability by an ARCH family model.

The aim of the research is to determine the best forecast values in a short-term to Brazilian agricultural and road production machinery. Assertive forecasts are made with the Autoregressive Integrated Moving Averages (ARIMA) models, which is an effective econometric approach for performing level prediction (ROCKENBACH et al., 2016). The ARIMA-ARCH models estimated jointly bring the novelty in this productive sector because the average is forecasted but also considering the volatility.

MATERIALS AND METHODS

National production of agricultural and road machinery was collected on the National Association of Automotive Vehicle Manufacturers website (<http://www.anfavea.com.br/estatisticas.html>), from January 1960 to October 2019, totaling 718 monthly observations. The variable considers

the wheeled tractors, crawler tractors, motorized cultivators, grain harvesters, cane harvesters and backhoes, in units.

The series stationarity was tested by the Augmented Dickey-Fuller (DICKEY & FULLER, 1979), Phillips-Perron (PHILLIPS & PERRON, 1988), Ng-Perron (NG & PERRON, 2001) and Elliott-Rothenberg-Stock (ELLIOT et al., 1996) unit root tests, the stationarity is condition to fit an ARIMA model.

The variable under analysis was fit by BOX & JENKINS methodology, based on interactive cycle: a) Identification, used to select the possible model based on autocorrelation function (FAC) and partial autocorrelation function (PACF); b) Estimation, the filter AR or MA define in the step before are estimated by Minimum Least Squares or Maximum Likelihood estimators; c) Verification, the model fit is analyzed if it presents white noise characteristics and adjustments statistics are used; d) Forecasting, the best model was used to forecast the future values, in this stage the evaluation statistics are used and the DIEBOLD – MARIANO test (1995) was applied to decide the best model to forecast (MORETIN & TOLOI, 2004; PEREIRA & REQUEIJO, 2008).

The general model proposed by BOX & JENKINS (1970), represented in equation 1.

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (1)$$

The model adequacy will be performed through residual analysis, which must present *white noise* characteristics, that is, zero mean, constant variance and non-autocorrelated, and present the lowest values for Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC), equations 2a

$$AIC_{(p,q)} = \ln \sigma_{p,q}^2 + \frac{2(p+q)}{n} \quad BIC_{(p,q)} = \ln \sigma_{p,q}^2 + (k+1) \frac{\ln(n)}{n} \quad (2a) \quad (2b)$$

The p and q values in the AIC and BIC criteria equation are used to penalize models with many parameters in order to find a model more parsimonious (SENNA & SOUZA, 2016).

The forecasts values will be verified evaluated by *U-Theil* statistics which provided if the model established give better forecasting than the mean series values. The Mean Squared Error – MSE and Root Means Squared Error – RMSE to evidence the better model to forecast, according to equations 3a, 3b, and 3c. The best forecasted model will be evaluated by DIEBOLD-MARIANO test (1995) give to us the best model. The null hypothesis of DM test is that all forecasted values

have the same accuracy,

$$UTheil = \frac{\sqrt{\sum_{t=1}^n (Z_t - \bar{Z}_t)^2}}{\sqrt{\sum_{t=1}^n (Z_t - Z_{t-1})^2}} \quad MSE = \frac{1}{n} \sum_{t=1}^n a_t^2 \quad (3a) \quad (3b)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Z_t - \bar{Z}_t)^2} \quad (3c)$$

Where a_t is the error defined by subtracting the values from the original series by the forecasts e is the value of the original series lagged by $t-1$ periods. The models showed values close to zero are the best suited, the forecast values will be done six steps ahead. To decide which forecast value is better the modified DIEBOLD-MARIANO (DM) test was used to decide the best alternative forecast. The DM is a more versatile test than any other alternative test to decide the best forecast, and used the critical values of Student's test t distribution with $(n-1)$ degrees of freedom. And, the modified DM brought improvement in cases where the sample sizes are small (HARVEY et al., 1997). The best ARIMA model fitted will have its residues tested by conditional heteroscedasticity test - LM-ARCH test (ENGLE, 1982), which will show if the quadratics residuals have autoregressive characteristics that can be represented by the Generalized Autoregressive Heteroscedastic (GARCH) model presented in equation 4

$$\sigma_t^2 = \alpha + \alpha_1 a_{t-1}^2 + \dots + \beta_1 \sigma_{t-1}^2 + \varepsilon_t \quad (4)$$

In the ARCH (p, q) model, p represents the parameters of the autoregressive part associated with β_1 and q represent the parameters of moving average, associate with α_1 , and α represent the mean process.

If quadratic residuals derived from linear

modeling present significant autocorrelation and the LM-ARCH test is statistically significant, so there is evidence of heteroscedasticity and ARCH model should be fit and de parameters estimated. The sum of the parameters α and β represent the volatility persistence, as closer to 1 the sum of $\alpha + \beta$ greater is the persistence, showing a long period of instability.

RESULTS AND DISCUSSION

Performing the ADF test applied to the series in level was considered stationary ($p=0.413$). To the KPSS test, the series in level was considered non-stationary with $LM_{calc} = 0,7525$ and LM_{tab} at 5% 0,4630, but in first differences $LM_{calc} = 0,0975$ and LM_{tab} at 5% 0,4630, the series was considered stationary. Performing the Phillips-Perron test applied to the series in level it was considered stationary with $p = 0.001$. To the Elliott-Rothenberg-Stock Point - Optimal test, considering the series in level stationary $P_{calc} = 6,95$ and $P_{stat} = 3,26$, the series is considered stationary. To the Ng-Perron test consider series in level with $MZa = -3,026$ and MZa at 5% -8,10 the series is non-stationary, in the first differences the series become stationary $MZa = -22,83$ and MZa at 5% -8,10. Due to controversial results of the unit root tests the series will be considered non-stationary to be fit.

The ARIMA competing models are displayed in table 1, the best fit was an ARIMA(2,1,1)-ARCH(2) model. To corroborate the presence of volatility the LM-ARCH test was carried out, which showed that the residuals are not homoscedastic ($p = 0.5257$).

The sum of parameters values corresponding to a_{t-1}^2 and a_{t-2}^2 from the ARCH model (2), using table 1 is $0.25 + 0.35 = 0.60$. The value 0.60

Table 1 - Competing models of the ARIMA and GARCH general class for Brazil's agricultural and road machinery production series.

| Model | Parameter | AIC | BIC | RMSE | MSE | U-Theil | -----DM----- | |
|---|--|-------|-------|--------|--------|---------|-------------------|-------------------|
| | | | | | | | F _{stat} | F _{prob} |
| SARIMA (1,1,0) (1,0,0) ₁₂ - GARCH(1,1) | $\phi_1 = -0,200$ $\Phi_{12} = 0,570$ $\alpha^2_{t-1} = 0,121$ $\sigma^2_{t-1} = 0,865$ C= 2135,1 | 15,15 | 15,18 | 628,29 | 535 | 0,081 | 84,01 | 0,002 |
| ARIMA (2,1,1)-ARCH(2) | $\phi_1 = 0,097$ $\phi_2 = 0,543$ $\theta_1 = -0,310$ $\alpha^2_{t-1} = 0,25$ $\alpha^2_{t-2} = 0,35$ C= 134476,2 | 15,30 | 15,34 | 523,36 | 410,64 | 0,066 | 1,33 | 0,384 |

was close to one indicating that the series was highly persistent, showing that the volatility of domestic production of agricultural and road machinery will affect the series level over long periods of time.

Future values were projected from November to April 2020 as shown in figure 1. The model adequacy is displayed at table 1. Figure 2 represents original and expected variance serie values from November to April 2020. It was identified that the first volatility cluster occurred in May 1980, which compared to the previous month showed 87.02% increase in monthly production.

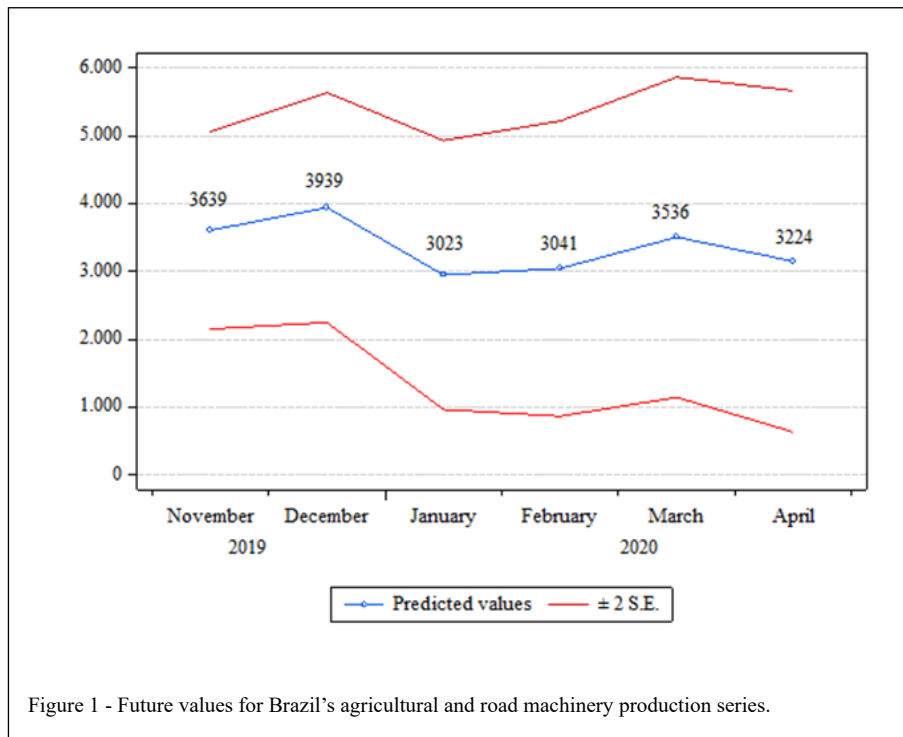
This volatility effect can be explained by the country's economic instability in relation to agricultural and livestock activities. The 60's were marked by industrialization, city growth, population increase and high purchasing power, but the situation was food shortage (EMBRAPA, 2017). To solve the problem, the government has instituted rural credit policies in the country to foster modernization and increase agricultural and livestock productivity (BIANCHINI, 2015).

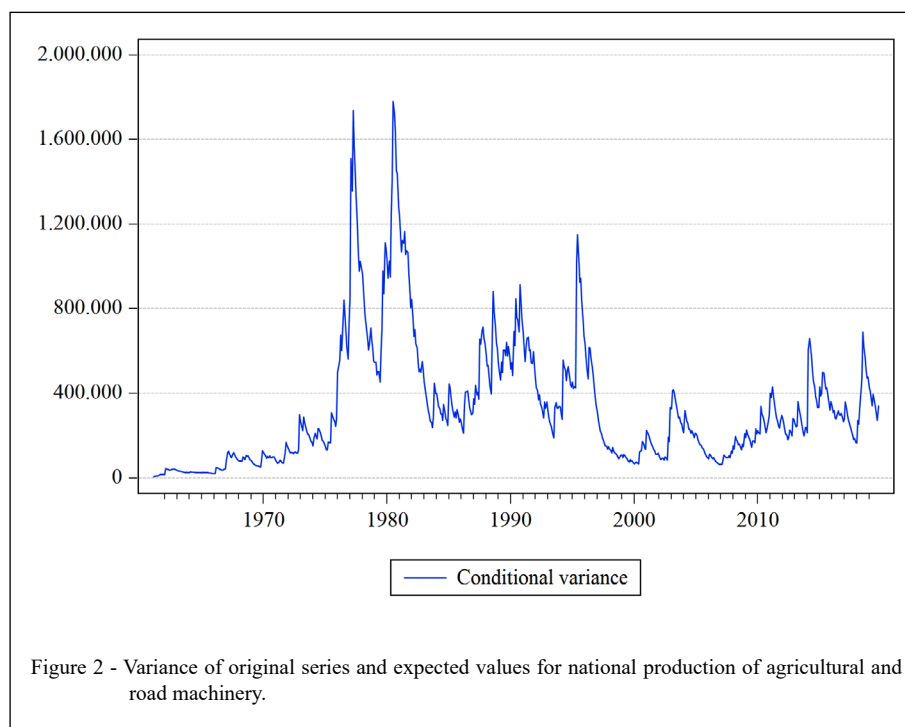
The volatility cluster formed in May 1980 can be explained by production to meet the demand for agricultural machinery intended to subsidize family farming, especially in soybean farming. It was also observed that production declined significantly

and may be justified by the end of the harvest. Another factor that contributed to the reduction in production was that the credit model managed for the 1970s, called Miracle Economic. The end of the decade the sector starts to enter in a crisis, due to changes in international policies that affected the Brazilian economy. The credit line was equal to the other financing lines, and the resources were scarce and expensive, not being adjusted to the agricultural price, causing delinquency in rural credit (BIANCHINI, 2015).

The economic crisis impacted the reduction of sales, resulting in bankruptcy, closure of activities and denationalization of a large number of companies that participated in this industrial segment (AMATO NETO, 1985), which justifies the reduction of production in the subsequent period called Lost Decade.

The second highest volatility peak occurred on February 2017, when production increased by 97.59% compared to January. In this same period the country was in the process of modernization, agribusiness production chains impacted agricultural activities, which were of substantial participation in GDP. In 2016 agribusiness accounted for 46% of the export's values, and in total generating 23% of GDP.





In 2017, agribusiness employed 4.12 million people, services 5.67 million and the agribusiness input segment 227.9 thousand (BRASIL, 2017).

This volatility cluster can be justified by the increase of exports, besides supplying the national demand, the international market was favorable, justified by the low dollar price. The high production of agricultural and road machinery occurs mainly in the first months of the year, so that the machinery works due to the soybean crop, which occurred between February and May.

The third largest volatility cluster occurred in September 1995, with 41.99% increase in total production of agricultural and road machinery, which may be justified by the emergence of the National Family Farming Strengthening Program (PRONAF) and macroeconomic stabilization (EMBRAPA, 2017; BIANCHINI, 2015).

In the first half of the 1990s problems with rural credit were persistent. In this post-crisis scenario of the agricultural model, and the end of the military regime, farmers' organizations rearticulated themselves in the National Confederation of Agriculture (CNA), and the Brazilian Agribusiness Association (ABAG). The mobilization of family farmers in 1995 allowed the consolidation of

PRONAF, establishing a differentiated credit line for the productive restructuring of family farming. The PRONAF's rural credit was created by Central Bank Resolution 2191 of August 24, 1995 (BIANCHINI, 2015), and its first year of operation provided a small number of financing for agricultural production, reflecting in an increasing production in the same year (SCHNEIDER, 2004).

CONCLUSION

The ARIMA(2,1,1)-ARCH(2) model fit to predict production and showed that the national production of agricultural and road machinery had linear and nonlinear components. The model fitted show that only an ARIMA model would not be able to capture all the movements in the series. The benefit to estimate linear and nonlinear models, go ahead for just made a simultaneous estimation of parameters, but provided the series behavior volatility along period.

The volatility captured in the national production of agricultural and road machinery was caused by the variation in credit development policies. It is suggested for future analysis the incorporation of exogenous macroeconomic variables, in order

to contribute to the prediction's assertiveness. To capture the long memory effects fractional models represented by an ARIMA with fractional differentiations (ARFIMA) jointly with volatility models, autoregressive conditional heteroscedasticity ARCH models are encouraged to be applied fitting an ARFIMA – GARCH models, probably providing better candidate models than other conditional heteroscedastic models for volatility.

ACKNOWLEDGMENTS

We are thankful to the anonymous reviewers for their research suggestions, to the Laboratory of Analysis and Statistical Modeling (LAME) at Universidade Federal de Santa Maria (UFSM). This paper was carried out with the financial support of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Financing Code 001.

DECLARATION OF CONFLICT OF INTERESTS

The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

AUTHORS' CONTRIBUTIONS

The authors contributed equally to the manuscript.

REFERENCES

- AI, H. Study on prediction of the total power of agricultural machinery based on fuzzy BP network. **Lecture Notes in Electrical Engineering**, p.551–558, 2015.
- AMATO NETO, J. A Indústria de máquinas agrícolas no Brasil - Origens e evolução. **Revista de Administração de Empresas**, v.25, n.3, 1985. Available from: <<http://dx.doi.org/10.1590/S0034-75901985000300005>>. Accessed: Jul. 15, 2019. doi: 10.1590/S0034-75901985000300005.
- ASSOCIAÇÃO NACIONAL DOS FABRICANTES DE VEÍCULOS AUTOMOTORES – ANFAVEA. Available from: <<http://www.anfavea.com.br/estatisticas.html>>. Accessed: Mar. 28, 2019.
- BIANCHINI, V. **Vinte anos do PRONAF, 1995-2015: avanços e desafios**. 2015. Available from: <http://www.mda.gov.br/sitemda/sites/sitemda/files/ceazinepdf/PRONAF_20_ANOS_VALTER_BIANCHINI.pdf>. Accessed: Mar. 28, 2019.
- BOX, G.; JENKINS, G. **Time series analysis: forecasting and control**. Holden-Day, San Francisco, 1970.
- BRASIL, Ministério da Agricultura, Pecuária e Abastecimento. **Agropecuária puxa o PIB de 2017**. 2017. Available from: <<http://www.agricultura.gov.br/noticias/agropecuaria-puxa-o- PIB-de-2017>>. Accessed: Jun. 17, 2019.
- DICKEY, D. A; FULLER, W. A. Distributions of the estimators for autoregressive time series with a unit root, **Journal of American Statistical Association**, v.74, n.366, p.427-481, 1979. Available from: <https://www.jstor.org/stable/2286348?seq=1#metadata_info_tab_contents>. Accessed: Jun. 17, 2019. doi: 10.2307/2286348.
- DIEBOLD, F. X; MARIANO, R. S. Comparing predictive accuracy. **Journal of Business and Economic Statistics**, v.13, p.253-63, 1995. Available from: <<https://www.tandfonline.com/doi/abs/10.1198/073500102753410444>>. Accessed: Jun. 17, 2019. doi: 10.1198/073500102753410444.
- ELLIOTT, G.; et al. Efficient tests for an autoregressive unit root, **Econometrica**, 64, 813-836, 1996. Available from: <<https://www.nber.org/papers/t0130>>. Accessed: Jun. 22, 2019. doi: 10.3386/t0130.
- EMBRAPA. **Trajatória da agricultura brasileira**. 2017. Available from: <<https://www.embrapa.br/visao/trajetoria-da-agricultura-brasileira>>. Accessed: Jun. 17, 2019.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. **Econometrica**, p.987-1007, 1982. Available from: <https://www.jstor.org/stable/1912773?seq=1#metadata_info_tab_contents>. Accessed: Jun. 17, 2019. doi: 10.2307/1912773.
- GONÇALVES, E.; et al. Padrões de inovação da indústria de bens de capital agrícola no Brasil. **Pesquisa & Debate**, v.26, n.1, v.47, p.216-235, 2015. Available from: <[jstor.org/stable/1912773?seq=1#metadata_info_tab_contents](https://www.jstor.org/stable/1912773?seq=1#metadata_info_tab_contents)>. Accessed: Jun. 19, 2019.
- HARVEY, D. et al. Testing the equality of prediction mean squared errors. **International Journal of Forecasting**, v.13, n.2, p.281-291, 1997. Available from: <<https://www.sciencedirect.com/science/article/abs/pii/S0169207096007194>>. Accessed: Jun. 17, 2019. doi: 10.1016/s0169-2070(96)00719-4.
- JU, J. Y. et al. Forecasting and Analysis the Demand of Agricultural Mechanization for Economic Development. **Advanced Materials Research**, v.694-697, p.3512–3515, 2013. Available from: <<https://www.scientific.net/AMR.694-697.3512>>. Accessed: Jun. 17, 2019. doi: 10.4028/www.scientific.net/amr.694-697.3512.
- MARCHESAN, A; SOUZA, A. M. Previsão do preço dos principais grãos produzidos no Rio Grande do Sul. **Ciência Rural**, v.40, n.11, p.2368-2374, 2010. Available from: <http://www.scielo.br/scielo.php?pid=S0103-84782010001100019&script=sci_abstract&tlng=pt>. Accessed: Jun. 22, 2019. doi: 10.1590/s0103-84782010001100019.
- MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. São Paulo: Edgard Blücher, 2004.
- PEREIRA, Z. L.; REQUEIJO, J. G. **Qualidade: planejamento e controle estatístico de processos**. Lisboa: Tipografia Lousanense, Ltda, 2008.
- PHILLIPS, P. C. B.; PERRON, P. Testing for a unit root in time series regression, **Biometrika**, v.75, n.2, p.335–346, 1988. Available from: <<https://academic.oup.com/biomet/article>>

abstract/75/2/335/292919>. Accessed: Jun. 22, 2019. doi: 10.2307/2336182.

ROCKENBACH, F. L. et al. Economic feasibility of biogas production in swine farms using time series analysis. **Ciência Rural**, v.46, n.7, p.1295-1300, 2016. Available from: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-4782016000701295&lng=en&tlng=en>. Accessed: Jul. 15, 2019. doi: 10.1590/0103-8478cr20141848.

SCHNEIDER, S. et al. Histórico, caracterização e dinâmica recente do Pronaf – Programa Nacional de Fortalecimento da Agricultura

Familiar. **Políticas Públicas e Participação Social no Brasil Rural**. p.21-50, 2004.

SENNA, V. SOUZA, A. M. Assessment of the relationship of government spending on social assistance programs with Brazilian macroeconomic variables. **Physica A: Statistical Mechanics and its Applications**, v.462, p.21-30, 2016. Available from: <<https://ideas.repec.org/a/eee/phsmap/v462y2016icp21-30.html>>. Accessed: Jul. 25, 2019. doi: 10.1016/j.physa.2016.05.022.

VIAN, C. E. F. et al. Origens, evolução e tendências da indústria de máquinas agrícolas. **Revista de Economia e Sociologia Rural**, v.51,