Theoretical-empirical Article

An Application of Geographically Weighted Quantile Lasso to Weather Index Insurance Design

Uma Aplicação do Lasso Quantílico Geograficamente Ponderado ao Seguro de Índice Climático

Open Data Open Access

Daniel Lima Miquelluti*[1]
Vitor Augusto Ozaki[1]
David José Miquelluti[2]

■ **ABSTRACT**

Objective: this article studies the efficiency of a novel regression approach, the geographically weighted quantile lasso (GWQlasso), in the modeling of yield-index relationship for weather index insurance products. GWQlasso allows regression coefficients to vary spatially, while using the information from neighboring locations to derive robust estimates. The lasso component of the model facilitates the selection of relevant explanatory variables. **Methodology:** a weather index insurance (WII) product is developed based on one-month standardized precipitation index (SPI) derived from a daily precipitation dataset for 41 weather stations in the state of Paraná (Brazil) for the period from 1979 to 2015. Soybean yield data are also used for the 41 municipalities from 1980 to 2015. The effectiveness of the GWQlasso product is evaluated against a classic quantile regression approach and a traditional yield insurance product using the spectral risk measure (SRM) and the mean semi-deviation. **Results:** while GWQlasso proved as effective as quantile regression, it outperformed the yield insurance product. **Conclusion:** the GWQlasso is an alternative to the crop insurance market in Brazil and other locations with limited data.

Keywords: GWQlasso; index insurance; systemic risk; crop insurance.

■ **RESUMO**

Objetivo: este artigo estuda a eficiência de uma nova abordagem de regressão, o Lasso quantílico geograficamente ponderado (GWQlasso), na modelagem da relação índice-rendimento para produtos de seguro de índice climático. O GWQlasso permite que os coeficientes de regressão variem espacialmente, enquanto usa as informações de locais vizinhos para gerar estimativas robustas. O componente Lasso do modelo facilita a seleção de variáveis explicativas relevantes. **Metodologia:** um produto de seguro de índice climático (WII) é desenvolvido com base no *standardized precipitation index* (SPI) de um mês derivado de um conjunto de dados de precipitação diária para 41 estações meteorológicas no estado do Paraná (Brasil) para o período de 1979 a 2015. Os dados de produção de soja também são usados para os 41 municípios de 1980 a 2015. A eficácia do seguro modelado utilizando-se GWQlasso é avaliada em comparação com uma abordagem de regressão quantílica clássica e um produto de seguro de produtividade tradicional usando a medida de risco espectral (SRM) e o semidesvio médio. **Resultado:** embora o GWQlasso tenha se mostrado tão eficaz quanto a regressão quantílica, ele superou o produto de seguro de produtividade. **Conclusão:** o GWQlasso mostra-se como alternativa para o mercado de seguro agrícola no Brasil e em outros locais com limitação de dados.

Palavras-chave: GWQlasso; seguro paramétrico; risco sistêmico; seguro rural.

# of invited reviewers until the decision:

	1	2	3	4	5	6	7	8	9
1st round	x	x	x	x	x	x	🤝	🤝	x
2nd round	🤝	🤝							

1

INTRODUCTION

The unpredictability of climatic variations is the principal risk factor in soybean cultivation on the south of Brazil. Reports on indemnities paid by government risk management programs, the Program for the Guarantee of Agricultural and Livestock Activity (Proagro) and Rural Insurance Premium Subsidization Program (PSR) (Banco Central do Brasil [BACEN], 2018; Ministério da Agricultura, Pecuária e Abastecimento [MAPA], 2015), show that the occurrence of droughts are the main event of loss (85% of the insured sum), followed by excessive rain (7.6% of the insured sum) and hail (4.2% of the insured sum). In addition, losses due to strong wind, excessive temperature fluctuation, and flood are also mentioned.

Crop insurance is recognized as one of the most efficient mechanisms of income protection in agriculture, transferring risk from agriculture to other agents and economic sectors. Insurance tends to stimulate the increase of cultivated area and the use of technology, especially as it acts as an additional guarantee for access to credit (Goodwin, Vandeveer, & Deal, 2004). In this sense, it not only contributes to the achievement of lower interest rates (Cai, 2016) by the rural producer, since the reduction of agricultural risk translates into lower credit risk, but also contributes to the development of financial, insurance, and capital markets. As a result, it minimizes the pressure for subsidized credit and ex-post government financial bailout, reducing the recurring pressure for renegotiations of rural debts.

However, the degree of penetration of agricultural insurance, considering the size and relevance of Brazilian agribusiness, is still insignificant. One of the reasons for the restriction of the subsidized crop insurance program and the massification of rural insurance in the country is the limited availability of budgetary resources to fund the policies. In addition, these budgetary resources depend on congress approval, thus preventing the long-term planning of investments by the private sector, imposing costs on the beneficiaries and generating dissatisfaction of the target public (Ministério da Agricultura, Pecuária e Abastecimento [MAPA], 2017).

The Proagro risk management program also faces difficulties. According to Oñate, Ozaki and Bravo-Ureta (2016), there was no increase in welfare for participating farmers. Considering the fact that the pricing of Proagro does not take into account regional differences, only crop type and cultural management practices such as the use of irrigation (BACEN, 2018), we believe different approaches must be sought by the government.

A possible alternative to overcome these issues is parametric insurance, which has lower administrative and regulation costs when compared to traditional insurance. The absence of in situ claim adjustment and moral hazard monitoring greatly reduces the administrative costs of this type of insurance, permitting a subsidy free crop insurance (Jensen & Barrett, 2017). Another advantage of parametric insurance is the rapid payment of indemnities.

Parametric insurance in Brazil is quite limited, however, with only one insurer offering tailored weather index insurance products (Swiss Re) as of 2018. Past initiatives include a yield index product, commercialized by AgroBrasil (Carter, Janvry, Sadoulet, & Sarris, 2017) in the state of Rio Grande do Sul and a hypothetical yield index insurance for the municipality of Castro in the state of Paraná (Ozaki & Shirota, 2005).

Therefore, aiming to contribute for the expansion of parametric insurance in Brazil, we intend to assess if the Paraná state presents a suitable environment for this type of product, that is, an index insurance product provides better or at least equal protection than a traditional yield insurance product. This study specifically targets soybean in Paraná, the second largest soybean producer in Brazil with a total of 19,073,706 tons produced in 2017, being also the second in average yields (3,663 kg/ha in 2017). We develop a weather index product based on the standardized precipitation index (SPI) and analyze its hedging effectiveness against a common yield insurance.

We also extend the work of Conradt, Finger and Bokusheva (2015), who proposed the use of quantile regression to model the yield-index relationship, by applying the geographically weighted quantile least absolute shrinkage and selection operator (GWQlasso) (Wang, Xu, & Yan, 2018) framework. Our hypothesis is that the spatial component, captured by the latter, plays an important role in the determination of the yield-index relationship, resulting in a better risk protection than a product developed with quantile regression. In addition, this methodology is less data intensive, as it borrows information from neighboring locations. The effectiveness of our model is compared to the traditional yield insurance and the quantile regression approach by means of two risk measures, the spectral risk measure (SMR) and the mean-semideviation model.

This article is organized as follows: in the empirical framework section, we present in detail the different methodologies utilized throughout the article, and then in empirical application we give some context in our database and the proposed index insurance product for Paraná. Our findings and discussion are found in Results and Discussion and we finish with conclusions.

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

2

DEVELOPMENT OF INDEX INSURANCE

Parametric insurance first appeared in the pioneering written by Chakravarti (1920). After more than a decade studying the subject, the author developed an insurance product based on rainfall levels for Chitradurga in India. Indemnities were paid if total rainfall measures in the beginning of the agricultural year were 35% below normal. The payouts were divided in two periods, from January to July and from July to October, according to the production cycle. The author noted that the area should be as uniform as possible, in respect to rainfall, for the insurance to work properly. The premiums were calculated to be as close as possible to land tax value, with both premiums and indemnities depending on the land's quality. In order to keep the farmers enrolled in the insurance scheme, contracts would be ranging from five to ten years, so that each farmer would receive at least one indemnity and thus perceive the value of crop insurance (Mishra, 1995; Rao, 2011).

Halcrow (1949) devised a different form of index insurance, based in the area yields. The main idea was to develop an insurance product where indemnities would be due when the mean yield of a uniform area fell below a predefined level (which could be defined as a proportion of the expected mean yield). The size of the area could vary as long as the homogeneity of yields was maintained, and the insured farmer would select a percentage of the expected yield for the area.

The main advantage of this type of insurance over the traditional crop insurance products is the reduction of moral hazard (when the insured incur in risk increasing activities or stop taking risk-mitigating actions due to being covered by the insurance). Since the farmer could not significantly alter the area yield, risk-increasing measures are not economically viable. This would also lead to a reduction in deductibles and coverage levels limitation by the insurers (Miranda, 1991). This author also notes that adverse selection (the inability to correctly measure farmer risk leads insurers to price the insurance incorrectly and in consequence to a greater proportion of high-risk farmers in their portfolio, incurring in a market collapse), which is caused by information asymmetry, is reduced in area yield insurance as this information is available to the general public. Adding to the advantages of this type of insurance are the reduced administrative costs since an index-based insurance does not require individual assessment of yields, a major cost for traditional crop insurance.

Two years after the work published by Miranda (1991), a yield-based index insurance was developed by the United States Federal Crop Insurance Corporation (FCIC) in conjunction with Skees, Black, and Barnett (1997). The product named Group Risk Plan (GRP) was expanded in

1994 and reached 70% of market participation in 1997, considering the seven major crops and excluding forage. An additional feature of GRP was the possibility to scale the protection (the product of expected yield and expected price) up to 150%. This option was intended to increase protection since farm and county yields are not perfectly correlated. The difference between the county yields, the index, and the value of individual yields is called basis risk, a problem that is always present in index-based insurance. In this way, GRP was designed to reduce basis risk by using double exponential smoothing to forecast the central tendency of yields, scaling the protection and paying indemnities based on the percentage reduction of yields rather than the weight/volume reduction. Since yield data provided by the National Agricultural Statistics Service (NASS) are available only at the county level, it was not possible to change the area in order to increase homogeneity of yields.

The GRP insurance was later expanded in 1999 to cover price variations and the index turned into a revenue index, named the Group Risk Income Protection (GRIP). The expected price was calculated individually for each crop and region. Both GRP and GRIP were replaced by the Area Risk Protection Insurance Policy (ARPI) in 2013. This new policy is formed of three insurance plans, Area Revenue Protection (ARP), Area Revenue Protection with Harvest Price Exclusion (ARPwHPE), and Area Yield Protection (AYP). The ARP and ARPwHPE are similar to the GRIP and the AYP is similar to the GRP, with the harvest price exclusion option meaning the amount protected will not rise if harvest prices rise (Schnitkey, 2014).

Weather-based index products were to be operated only from 2006 with the approval of flood insurance by the Peruvian government (Khalil, Kwon, Lall, Miranda, & Skees, 2007). Following that, several studies and pilots were launched, mostly in developing countries (Giné, Menand, Townsend, & Vickery, 2010; Leblois, Quirion, Alhassane, & Traoré, 2014; Maestro, Barnett, Coble, Garrido, & Bielza, 2016; Skees, Gober, Varangis, Lester, & Kalavakonda, 2001; Skees, Hartell, & Murphy, 2007).

EMPIRICAL FRAMEWORK

This section outlines the conceptual framework strategy used in the article. We present an overview of the methods used to model the yield-index relationship and to evaluate the proposed index insurance contract.

Geographically weighted quantile lasso

A natural extension to the geographically weighted regression (GWR) is the geographically weighted quantile regression (GWQR) model, which has the following form:

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022 | doi.org/10.1590/1982-7849rac2022200387.en | e-ISSN 1982-7849 | rac.anpad.org.br

3

$$Y_i = X_i^T \beta_\tau(u_i, v_i) + \epsilon_{\tau,i} \qquad (1)$$

where $\epsilon_{\tau,i}$ is the random error term, τ is the quantile of interest, Y_i and $X_i^T = [X_{i1}, \ldots, X_{ip}]$ are respectively the response variable Y and the explanatory variables $X_1, \ldots, X_p$ at the geographical location (u_i, v_i) $(i = 1, \ldots, n)$.

Consider $\rho_\tau(y) = y(\tau - I_{(y<0)})$ as the check loss function at quantile $\tau \in (0,1)$, with $I(\cdot)$ as the indicator function. For two locations (u_t, v_t) and (u_i, v_i), let $d_{it} = \| (u_i, v_i) - (u_t, v_t) \|$, where $\|\cdot\|$ is the Euclidean norm. According to Chen, Deng, Yang and Matthews (2012), the local linear GWQR estimates of the coefficients, and their partial derivatives, are the ones that minimize the local weighted quantile loss function:

$$\mathcal{L}_h(u_t, v_t) = \Sigma_{i=1}^n \rho_\tau \left\{ Y_i - X_i^T \begin{bmatrix} \beta_\tau(u_t, v_t) + \beta_\tau^{(u)}(u_t, v_t)(u_t - u_i) + \ldots \\ \ldots + \beta_\tau^{(v)}(u_t, v_t)(v_t - v_i) \end{bmatrix} \right\} K_h(d_{it}) \qquad (2)$$

With respect to $\beta_\tau(u_t, v_t)$, the partial derivatives of $\beta_\tau(u, v)$; $\beta_\tau^{(u)}(u_t, v_t)$ and $\beta_\tau^{(v)}(u_t, v_t)$ for a specified kernel function $K_h(\cdot) = K(\cdot/h)/h^2$ and bandwidth h. The latter is chosen via a cross validation procedure that is identical to its GWR counterpart, while the only difference is in the check loss function replacing the quadratic loss function.

Applying the aforementioned lasso method to the GWQR we have:

$$\mathcal{L}_{h,\lambda} = \sum_{t=1}^n \mathcal{L}_h(u_t, v_t) + \sum_{j=1}^p \begin{pmatrix} \lambda_{1j} \| [\beta_j(u_1, v_1), \ldots, \beta_j(u_n, v_n)]^T \| + \cdots \\ \ldots + \lambda_{2j} \| \begin{bmatrix} \beta^{(u)}(u_1, v_1), \ldots, \beta^{(u)}(u_n, v_n), \\ \beta^{(v)}(u_1, v_1), \ldots, \beta^{(v)}(u_n, v_n) \end{bmatrix}^T \| \end{pmatrix} \qquad (3)$$

where $\lambda_1 = (\lambda_{11}, \ldots, \lambda_{1p})^T \in \mathbb{R}^p$ and $\lambda_2 = (\lambda_{21}, \ldots, \lambda_{2p})^T \in \mathbb{R}^p$ are the tuning parameters. This combination of the GWQR technique and the lasso method is named by Wang, Xu and Yan (2018) as the geographically weighted quantile lasso (GWQlasso).

Both the local weighted quantile loss function and the penalty function in (3) are nondifferentiable at the origin, which results in the common derivative-based algorithm being unusable for obtaining the minimum point of $L_{h,\lambda}$. Therefore, a quadratic approximation (Hunter & Lange, 2000) is used to approximate the local weighted quantile loss function, while the local quadratic approximation (Fan & Li, 2001) is used to approximate the penalty function and establish the iterative algorithm of the GWQlasso.

Spectral risk measures

Traditional risk measures, such as the value at risk (VaR) and expected shortfall (ES), have some limitations. The two measures of risk do not explicitly consider the degree of risk aversion of the user of the method (Cotter & Dowd, 2010). It is implicit, when using VaR as a risk measure, that the agent has a negative risk aversion, whereas the choice of ES implies risk neutrality (Grootveld & Hallerbach, 2004). In the case of VaR, the negative risk aversion is explicit when it is verified that the agent does not weigh the losses that exceed the VaR. For ES, risk neutrality is illustrated by the fact that the agent weighs losses that exceed the VaR uniformly. Therefore, Acerbi (2002), Dowd, Cotter, and Sorwar (2008), and Cotter and Dowd (2010) argue that VaR and ES are not consistent risk measures when the agent using the technique has risk aversion.

To overcome this limitation, Acerbi (2002) proposed a measure of spectral risk that is consistent when applied to agents with risk aversion. Thus, consider the risk measure defined by:

$$M_\varphi = \int_0^1 q_p \, \varphi(p) dp \qquad (4)$$

where q_p is the quantile p of the distribution of losses, $\varphi(p)$ is a weight function defined in p, and p is a cumulative probability interval such that $p \in [0,1]$.

The measure of risk M_φ satisfies the conditions of coherence if and only if $\varphi(p)$ satisfies the following properties:

$\varphi(p) > 0$: the weights must always be non-negative.

$\int_0^1 \varphi(p) dp = 1$: the sum of the weights must be equal to the unit.

$\varphi'(p) \geq 0$: high losses are associated with weights greater than or equal to losses of smaller magnitude.

Now, one must select a suitable risk aversion function that satisfies the above properties. Here we use the exponential function of risk aversion:

$$\varphi(p) = \frac{k e^{-k(1-p)}}{1 - e^{-k}} \qquad (5)$$

where $k>0$ is the absolute risk aversion coefficient. This measure of spectral risk attributes greater weights to losses in the higher levels of cumulative probability distribution (the worst losses). In addition, for any infinitesimal variation in p, the weights vary more rapidly the more risk averse the agent is. The growth rate depends on the value of k, that is, the more risk averse the investor, the more the weights will grow.

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

4

Mean semi-deviation

The standard deviation considers both the below and above average values to be equally undesirable, and this may not be consistent with the objectives of the farmers, as the concern is generally about losses. Alternatively, we can use an indicator that considers only the dispersion of values on the left side of the distribution, that is, the semi-deviation, given by:

$$\sigma_{ssd} = \sqrt{\frac{\sum_{i=1}^{n}\left(min[0, w_i - \bar{w}_i]\right)^2}{n}} \tag{6}$$

where σ_{ssd} is the default semi-deviation of the wealth stream, w_i is the wealth, $\bar{w}_i$ is the critical point below which the farmer cares, and n is the number of observations. The value of $\bar{w}_i$ represents the minimum acceptable return, that is, the point at which the dispersion of the left distribution is measured.

The concept of semi-deviation is not new, and its applications in the area of finance have emerged with Markowitz (1959), who in his classic book notes that the choice between the two measures depends on the convenience, familiarity, and differences between the portfolios produced by different metrics, among other pertinent characteristics.

An important feature to be emphasized is that the numerical value of the standard deviation is at least equal to the semi-deviation. The immediate implication is that we cannot make a comparison between the standard deviation and the semi-deviation, even though the two have equal units.

Thus, similarly to the mean-variance utility function, the mean semi-deviation utility function is expressed by σ_{ssd} and:

$$U_{it} = \begin{cases} W_{it} - E(\boldsymbol{W}_i), & \text{if } W_{it} < E(\boldsymbol{W}_i) \\ 0, & else \end{cases} \tag{7}$$

where U_{it} is the farmer utility and E is the expectation operator. The exposure to adverse weather conditions relative to the semi-deviation is then measured by:

$$V_i = E(\boldsymbol{W}_i) - \frac{1}{2}k\sigma_{ssd} \tag{8}$$

where V_i is the revenue risk. A higher value of V_i is conditioned to a lower level of semi-deviation, thus indicating less exposure to weather risk.

For both risk measures, we chose a k value of 0.5 following Conradt et al. (2015) and Dowd et al. (2008).

EMPIRICAL APPLICATION

Data cleaning and yield detrending

We utilized the National Water Agency (ANA) daily precipitation data set, focusing only on municipalities, in the state of Paraná, with an operational weather station. The time series spans from January 20, 1979 to April 01, 2015 for 41 weather stations, one per municipality. We also used the series of annual soybean yields for these 41 municipalities, from 1980 to 2015, obtained from the National Institute of Geography and Statistics (IBGE).

Crop yields were detrended using the following equation (Duarte, Braga, Miquelluti, & Ozaki, 2018):

$$\tilde{y}_{t,i} = \hat{y}_{2015,i}\left(1 + \frac{\hat{e}_{t,i}}{\hat{y}_{t,i}}\right) \tag{9}$$

where $\tilde{y}_{t,i}, \hat{y}_{t,i}, \hat{y}_{2015,i}$, and $\hat{e}_{t,i}$ are, respectively, the corrected yield, the fitted yield, the fitted yield for 2015 and the residual for year t and municipality i.

Data pre-processing and clustering

In order to fill missing values we applied multiple imputation by chained equations (MICE) using the R software (Buuren & Groothuis-Oudshoorn, 2000) and then calculated the standardized precipitation index (SPI) with a three-month scale, thus capturing severe drought events during the crop season (McKee, Doesken, & Kleist, 1993).

Given that index insurance is based in the fact that risks are systemic, thus affecting a wide area, prior to the calculation of the product parameters a determination of the municipalities that present similar risk behavior must be carried. In order to spatially segment these areas, we chose the Ward's clustering method with an Euclidean distance matrix since it has already proved successful in defining homogenous precipitation regions in Brazil (Keller, Assad, & Lima, 2005). The optimal number of clusters was obtained through majority vote of 30 indices, an algorithm implemented in Charrad, Ghazzali, Boiteau and Niknafs (2014).

Weather index insurance

The state of Paraná is an important producer of soybean, being the second largest producer in Brazil. In spite of the evolution in crop technology and crop management, yields are highly susceptible to drought in some regions of the state, with as much of 50% of the final yields being dependent on water availability (Carmello & Sant'anna, 2016; Farias et al., 2001).

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

5

Our WII hypothetical product is based on the standardized precipitation index (SPI) rather than cumulative rainfall. We chose this approach as there is a weak correlation between monthly precipitation and yields. This is because water availability depends on variables other than rainfall, such as water storage capacity in the soil and evapotranspiration potential, which is greatly influenced by air temperature (Lier, 2014). The option for a rainfall-based index is also due to the better coverage of rainfall stations in Paraná.

The standardized precipitation index (SPI) is based on the probabilities of overcoming a certain accumulated precipitate volume. Rainfall values are summed over several scales, for example 3, 6, 12, or 24 months, depending on the interest or need of the analyst. For a given month, for example, October, the seven-month SPI (SPI-7) is obtained from the sum of the precipitations over the seven months preceding the reference month.

A probability function is fitted to the time series of monthly precipitation. In the original formulation and the one we utilize in this work, McKee, Doesken and Kleist (1993) used the gamma distribution. From the fitted probability distribution, each element of the fitted series is assigned a probability of non-overrun. Each of these probabilities of non-overrun is finally associated with the corresponding quantile of the standard normal distribution. The quantile value of the $N(0,1)$ associated with the probability calculated in the period of interest is the SPI value for the month.

One of the advantages of using SPI, according to McKee et al. (1993), is that SPI is only a function of probability. Thus, regardless of the probability distribution function to be used, the SPI can be properly calculated. Other advantages are that SPI is able to characterize both dry and rainy periods, as well as the fact that it is suitable for any hydrological variable. However, the use of this index also has limitations. Mishra and Singh (2010) argue that the main one is the need for long historical records for its consistent calculation, which is not always possible (Weschenfelder, Pickbrenner, & Pinto, 2011).

The relationship between SPI and soybean yields is then modeled using the GWQlasso framework, where the yields are the response variable and the SPIs are explanatory variables. We followed Conradt et al. (2015) and used a method based on the inverse function of the estimated regression to determine the triggers and exits of the contract. This approach permits a precise definition of the coverage level and does not require individual tinkering of the product parameters for each location, thus facilitating and streamlining product development. For our study we chose a coverage level of 100% of the expected yield.

In preliminary assessments we found that the one-month SPI has the highest correlation with soybean yields, thus we only present here the results for this index from October to March, the months that correspond to the planting and harvesting of soybean in most of Paraná. Planting and harvesting progress reports are available at the state level, with the months of October and March corresponding to more than 50% of the total crop area planted/harvested. These months are also assumed as planting and harvesting dates in Franchini, Balbinot, Nitsche, Debiasi and Lopes (2016).

Payout structure

The payout (indemnity) function will depend on the relation of the index and losses, being the most common: linear, accelerated, and step payout function. The base format of the payout structure that we utilized follows:

$$payoutrate = \begin{cases} 0, for\ x > trigger \\ g(x), for\ trigger \geq x > exit \\ 1, for\ x \leq exit \end{cases} \quad (10)$$

where *exit* is the lower threshold of the insurance, *trigger* is the point where the contract starts to pay out, and $g(x)$ is the payout function. The payout rate is then multiplied by the max payout to obtain the actual payout for the insured.

The $g(x)$ function is generally linear. However, it may be defined differently to better capture losses by the insureds. Functions that provide lower payments for minor events and higher ones for extreme events lead policyholders to have more confidence in the product (Carter, 2009). However, this tends to increase the frequency in payments, what may increase the premium and thus reduce the sum insured. In addition, as noted in Skees (2008) and Collier, Skees and Barnett (2009), savings and credit perform better in smoothing consumption when less severe events occur. Index insurance is best applied for severe and catastrophic events.

Premium estimation

The insurance premium is derived from the probability density function (PDF) of indemnities, or an approximation of this distribution. In our study, we use the historical burn analysis (HBA) method to approximate the PDF of indemnities. This method is based in the application of the proposed product structure to historical realizations of the index. It is thus possible to calculate the historical payouts and the premium, calculated as the mean of payouts.

HBA is the simplest method to estimate an insurance premium, it also does not require assumptions on the PDF

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

6

parameters, in contrast to other methods such as historical distribution analysis (HDA) and Monte Carlo-based methods (Hess, Skees, Stoppa, Barnett, & Nash, 2005). The HDA method consists in finding the probability distribution that fits best the historical index values, and obtaining the expectation of payouts analytically. In the case where finding a closed form for the expectation of the payouts is not possible, simulations may be used, resulting in the Monte Carlo method for premium estimation. We refrain from using these latter methods as our data is aggregated at the municipality level and thus it may misrepresent variability at the farm level.

In order to provide a representative data set for the premium estimation we use the first 30 years of data for this part of the analysis, with the remaining six years being used for the evaluation of the methods.

Product evaluation

For the product evaluation, we use the values of the final wealth realizations for a hypothetical farm with an area of 1 ha. The only assets present in such farm are the soybean yield and the proposed weather index insurance contract:

$$W_{it} = (1/60)vy_{it} + I_{it} - P_i,$$

where W_{it} is the final wealth, v is the price paid to the farmer for each 60 kg of soybean, Y_{it} is the corrected yield, I_{it} is the indemnity, and P_i is the premium, with i being the municipality and t the year. Considering we corrected yields, we utilized the 2015 average prices of soybean provided by the Department of Rural Economy (Deral) of the State Secretariat for Agriculture and Food Supply (Seab) in Paraná. Final wealth realizations are calculated for farmers without insurance, thus having only the first component of the right-hand side, and for farmers with the WII parameters estimated by the quantile regression and the GWQlasso.

In order to measure the efficiency of the proposed index insurance to mitigate the risk faced by farmers, we used two risk measures, namely the spectral risk measure (SRM) and the mean semi-deviation, coupled with a Bayesian bootstrap procedure. The latter is necessary given that we dispose of only six years of data for the evaluation. In this step, a cross-validation (CV) method would be ideal but the computational requirements of GWQlasso make the use of CV not feasible in our case.

The Bayesian bootstrap (Rubin, 1981) is very similar to its classical counterpart (Efron, 1979), differing only in how probabilities are attached to each data value. While in the classical bootstrap a $1/n$, being n the sample size, is attributed to all n observations, in the Bayesian bootstrap the probabilities are given by a posterior distribution centered in $1/n$ but varying across replications. The main difference is in the interpretation of the results as the Bayesian bootstrap is a simulation of the posterior distribution of the parameter being estimated, whereas the classical bootstrap simulates the sampling distribution of an estimator for the parameter of interest.

The relative risk reduction (RR) is structured to compare the risk exposure of farmers in three situations: the first one being a farmer with a WII insurance designed with the GWQlasso method against a farmer without insurance; the second situation is a farmer with WII insurance designed with GWQlasso versus a farmer with WII insurance designed with quantile regression; and a third situation for a WII insurance (designed with GWQlasso) versus a traditional yield insurance (YI) product — this type of insurance protects the farmer in the case where observed yields fall below the expected yield, requiring on site claim adjustment. Thus, the general formula for the RR is:

$$RR_{case\ 1/case\ 2} = \frac{RM(W_{case\ 1}) - RM(W_{case\ 2})}{RM(W_{case\ 2})},$$

where RM stands for the risk measurements previously described and W for the final wealth realizations.

In our evaluations, we consider 4,000 Bayesian bootstrap replications to provide better estimates of the relative risk reduction. The latter is also tested against a hypothesis of null relative risk reduction by means of a non-parametric Wilcoxon test.

RESULTS AND DISCUSSION

The optimal number of clusters from the precipitation data was two. These clusters managed to capture the different precipitation regimes identified by Keller, Assad and Lima (2005), with cluster 1 representing municipalities with higher total precipitation in the year aggregate but greater variability among years, and cluster 2 indicating municipalities with a lower total precipitation but with less variability. For the yield data, the optimal number of clusters was also two, with both clusters presenting a similar yield level from the beginning of the series through 1990 and from 2001 onward. However, in the period comprised between 1991 and 2000, cluster 1 has lower yields. Cluster 1 contains municipalities in regions prone to drought, and thus presents lower yields and higher variability. A thorough discussion of the clustering analysis results is found in Miquelluti (2019).

Yield index modeling

We observe for the cluster representing the western and northern portions of Paraná that the December SPI presents the greatest impact on yields (Figures 1 and 2). Given that

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022 | doi.org/10.1590/1982-7849rac2022200387.en | e-ISSN 1982-7849 | rac.anpad.org.br

7

we assume, based on state reports (Secretaria da Agricultura e do Abastecimento, 2021), soybean planting dates are beginning on October, the crop would be in the reproductive stage in December, thus highly sensible to water shortage. Therefore, for the premium estimation in cluster 1 we select the December SPI as the index. For the central and eastern portions of the state, however, both December and February SPI are impacting yields, with the February SPI having a slightly higher impact and thus being the one selected as the index for cluster 2.
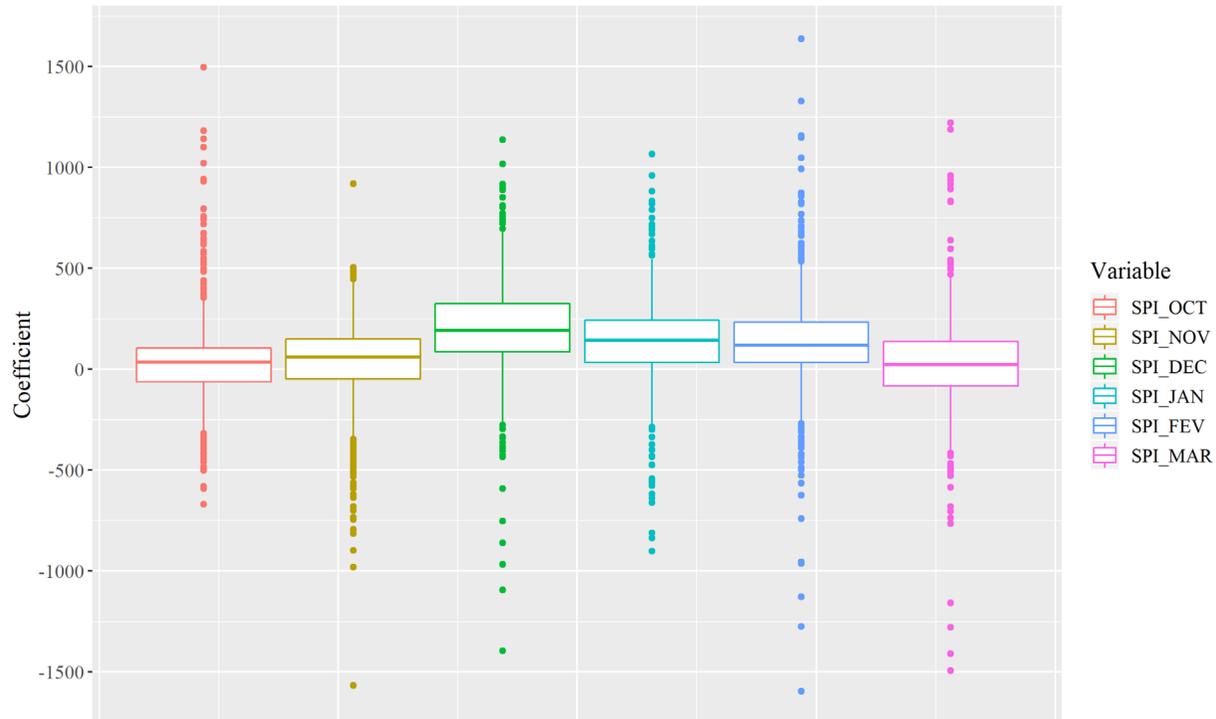

Figure 1. Boxplot of the GWQlasso coefficients for cluster 1.
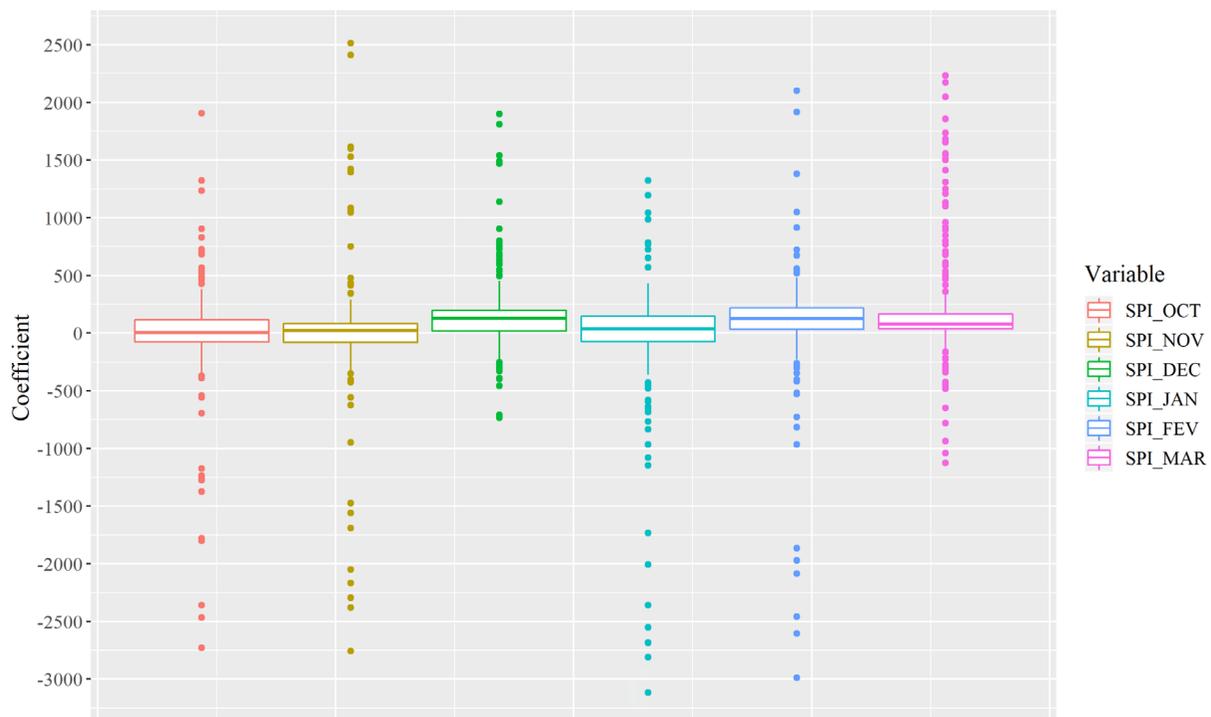

Figure 2. Boxplot of the GWQlasso coefficients for cluster 2.

Note that these coefficients are selected based on their boxplot as GWQlasso does not account for temporal structure of the data, thus assuming that all observations come from a unique point in time and resulting in 30 coefficients for each location. Albeit this represents a limitation to the modeling, the impact is reduced as we used time-detrended yields and are not interested in the temporal behavior of the yield/index series, just in their intrinsic relationship. In addition, the incorporation of the lasso method permits a better identification of relevant variables, as the ones with little importance to yields rapidly converge to zero.

The possibility to model yields at several locations simultaneously while considering the spatial structure of the relationship between yields and explanatory variables and also having a method to quickly dismiss unimportant variables presents a great opportunity to WII scalability. One of the major issues of WII is its low ability to grow at scale as the models developed for one location may prove completely obsolete as you move away from it; however, with GWQlasso one may inspect both the general significance of the explanatory variables and their coefficient for each location. This permits a faster screening of possible indices, along with their respective triggers when using the methodology detailed here and proposed in Conradt et al. (2015).

Another hindrance to the spread of WII is the absence of long series of yield data. The GWQlasso method is less affected by this issue, as it uses information from neighboring yields in the estimation process. This characteristic is especially important in developing countries, which in general do not have long series of yield and weather data.

The relationship between the December SPI and yields, for cluster 1, varies by more than five times when we compare municipalities on the west of Paraná to the ones in the center region of the state (Figure 3). This goes in line with the characteristics of these regions: soils to the west, mostly in the northwest, are sandy, and the climate is classified as Cfa, with higher temperatures in the summer, both unfavorable to soybean. These characteristics result in diminished yields and higher susceptibility to drought, which are translated in the coefficients for December SPI. A higher variability in yields, when compared to other regions in the state, coupled with susceptibility to drought is also observed by Franchini et al. (2016).

GWQLASSO β^{Dec}
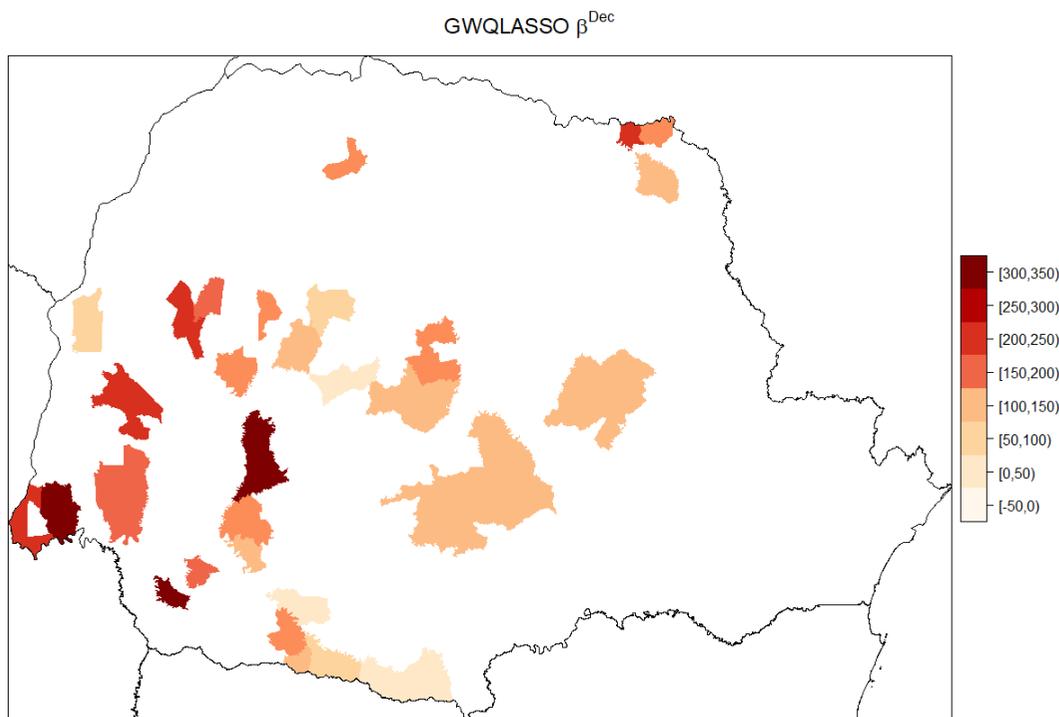
Figure 3. GWQlasso β^{Dec} coefficients spatial distribution for cluster 1.

The center and east of the state are classified as Cfb in Koeppen's system. This means that these municipalities have lower temperatures in the summer, what benefits soybean plants. In addition, the soil in these regions have more clay, what leads to a higher capacity to contain water and consequentially mitigates the effects of drought, resulting in a less pronounced relation to the index, for both clusters 1 and 2 (Figures 3 and 4).

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

9

Figure 4. GWQlasso β^{Feb} coefficients spatial distribution for cluster 2.

Weather index insurance premium and performance

WII is generally associated with lower premiums, mainly due to the lack of in situ crop inspection after a claim is filed. Here we compare a traditional yield insurance product with a 65% coverage to our proposed WII product with a 100% coverage. While this comparison may not be fair to our product, as by definition a higher coverage means a higher premium, we found the 65% coverage level to be the most common for soybean in Paraná. Coverage levels above 90% are rare in Brazil and suffer from two problems, the higher premiums and the inferior percentage in subsidization by the government, thus they are not attractive to the farmer.

Our results show that the index insurance may vary from half to three times the price of the common yield insurance. There is a tendency of pricier index insurance, compared to the yield insurance, as we move to the western portion of Paraná (Figure 5), what is expected as this region is more susceptible to drought.

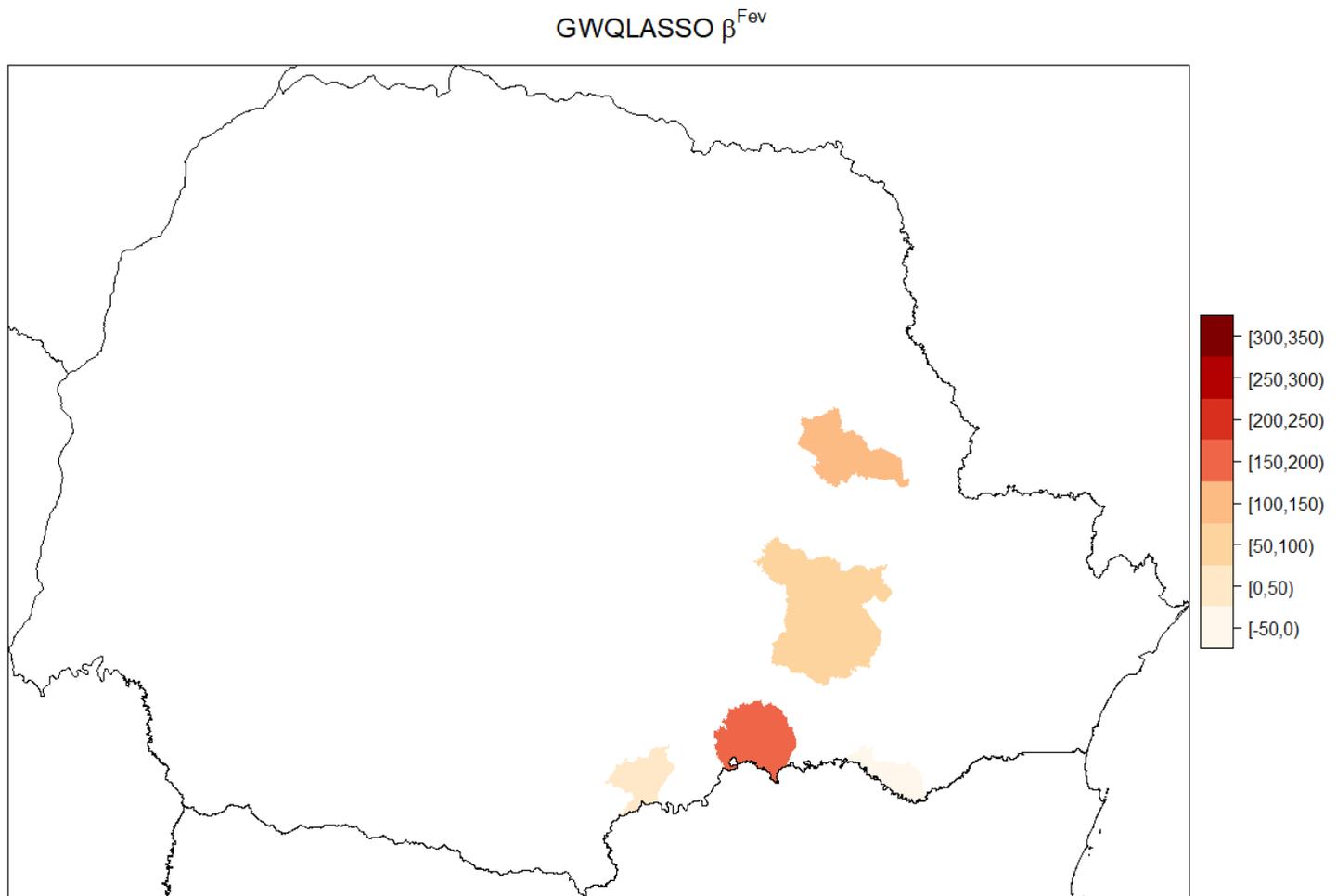In our design, we do not consider gains from scale and the spatial diversification of risk by the insurer. This would lead to a lesser difference between our product and the commercial product depicted here. Even so, the WII results in a net gain for the producer, as for both clusters and risk measurements it performs better than the yield insurance. The results from both risk measures indicate that both GWQlasso and quantile regression provide similar risk reduction, with both being more effective than a yield insurance product with a 65% coverage level (Tables 1 to 4). Borrowing from results in Conradt et al. (2015), we can also derive that GWQlasso is superior to ordinary least squares.

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022 | doi.org/10.1590/1982-7849rac2022200387.en | e-ISSN 1982-7849 | rac.anpad.org.br

10

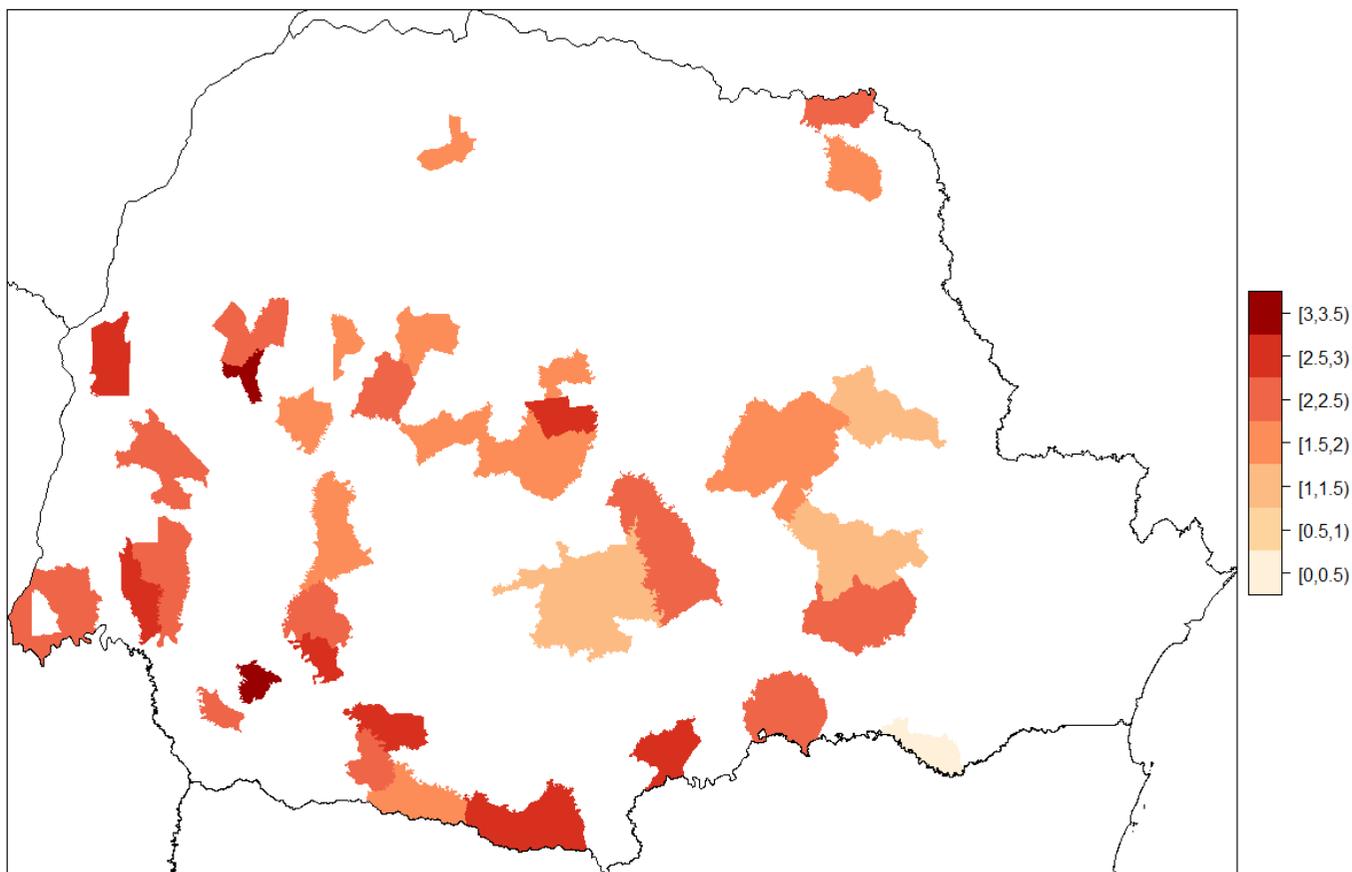**WII and Standard Yield Insurance premium ratios**

Figure 5. Ratio of GWQlasso calculated premiums to standard yield insurance premiums.

Table 1. Relative risk reduction per municipality in cluster 1 according to the spectral risk measure (SRM).

Municipality	$RR_{GWQLASSO}$			$RR_{GWQLASSO/QR}$			$RR_{GWQLASSO/YI}$		
	2.5%	Median	97.5%	2.5%	Median	97.5%	2.5%	Median	97.5%
Alto Piquiri	-0.1269	-0.0194	0.1459	-0.1127	-0.0056*	0.1117	-0.0691	0.0478*	0.2450
Ampére	-0.1291	-0.0150	0.1711	-0.1113	-0.0012	0.1221	-0.0816	0.0401*	0.2398
Andirá	-0.1317	0.0841	0.4253	-0.1722	0.0191*	0.2521	-0.0758	0.1470*	0.4714
Cambará	-0.1907	0.0076	0.2491	-0.1984	0.0025*	0.2538	-0.1523	0.0732*	0.3356
Campo Mourão	-0.1007	-0.0090	0.0953	-0.0927	0.0041*	0.1055	-0.0546	0.0398*	0.1621
Céu Azul	-0.1610	0.0004	0.2551	-0.1595	-0.001*	0.2099	-0.1101	0.0543*	0.2525
Clevelândia	-0.1156	-0.0077	0.1239	-0.1149	-0.0089*	0.1219	-0.0639	0.0424*	0.1830
Coronel Vivida	-0.2051	-0.0012	0.2519	-0.1915	0.0027	0.2465	-0.1703	0.0452*	0.2845
Formosa do Oeste	-0.1223	0.0513	0.3600	-0.1938	-0.0255*	0.1583	-0.0884	0.0939*	0.3761
Foz do Iguaçu	-0.2061	-0.0061	0.2812	-0.1968	-0.0007	0.2641	-0.1477	0.0468*	0.2916
Guaraniaçu	-0.0838	0.0086	0.1902	-0.1242	-0.0096*	0.1387	-0.0047	0.1035*	0.3146
Guarapuava	-0.0654	-0.0146	0.0306	-0.0536	-0.0015*	0.0517	-0.0159	0.0365*	0.0877
Ivaiporã	-0.0754	0.0101	0.0962	-0.0729	0.006*	0.0817	-0.0289	0.0664*	0.1580
Janiópolis	-0.0819	0.0148	0.1739	-0.1032	-0.0142*	0.0665	-0.0403	0.0678*	0.2515
Mamborê	-0.0609	0.0161	0.1068	-0.0766	-0.0017*	0.0708	-0.0072	0.0737*	0.1718
Manoel Ribas	-0.0960	-0.0051	0.0966	-0.0940	-0.0049*	0.0866	-0.0569	0.0423*	0.1548
Mariluz	-0.1125	0.0021	0.1730	-0.1149	0.0011*	0.1650	-0.0496	0.0722*	0.2705
Mariópolis	-0.1355	-0.0196	0.1762	-0.1117	0.001	0.1146	-0.0928	0.0255*	0.2021
Matelândia	-0.2177	-0.0355	0.2408	-0.1952	-0.0178*	0.2228	-0.1713	0.0107*	0.2166
Nova Esperança	-0.1745	-0.0164	0.1832	-0.1541	0.0047*	0.1982	-0.1191	0.0520*	0.2792
Palmas	-0.0514	0.0006	0.0545	-0.0511	0.0003	0.0556	-0.0098	0.0432*	0.1019
Pato Branco	-0.1411	-0.0264	0.1552	-0.1354	-0.0198*	0.1392	-0.0949	0.0344*	0.2290
Pitanga	-0.1261	-0.0224	0.0809	-0.1127	-0.008*	0.0985	-0.0768	0.0360*	0.1489
Prudentópolis	-0.0671	-0.0100	0.0406	-0.0643	-0.0025	0.0627	-0.0207	0.0386*	0.0914
Quedas do Iguaçu	-0.1423	-0.0341	0.1200	-0.0852	0.0212*	0.1112	-0.0926	0.0204*	0.1876
Roncador	-0.0955	-0.0026	0.1026	-0.0945	-0.0007	0.1072	-0.0466	0.0548*	0.1758
Salto do Lontra	-0.1587	-0.0316	0.1312	-0.1331	-0.0128*	0.1211	-0.1310	0.0035*	0.1887
Santo Antônio da Platina	-0.1100	0.0181	0.1963	-0.1099	0.0089*	0.1605	-0.0549	0.0811*	0.2803
São Jorge d'Oeste	-0.1958	-0.0146	0.2451	-0.1800	-0.0077*	0.2292	-0.1669	0.0148*	0.2518
São Miguel do Iguaçu	-0.1811	0.0151	0.3174	-0.2162	-0.0129*	0.2007	-0.1401	0.0603*	0.3187
Terra Roxa	-0.2323	0.0040	0.3032	-0.2307	-0.0034	0.2747	-0.1869	0.0432*	0.2877
Tibagi	-0.0546	0.0040	0.0777	-0.0510	-0.0017*	0.0481	-0.0094	0.0529*	0.1341
Toledo	-0.2095	-0.0251	0.2577	-0.1832	0.0057*	0.2260	-0.1671	0.0121*	0.2199
Ubiratã	-0.1043	-0.0065	0.0989	-0.1068	-0.0033*	0.1055	-0.0677	0.0425*	0.1584

Note. * significant at the 5% significance level. Source: Authors.

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

12

Table 2. Relative risk reduction per municipality in cluster 1 according to the mean semi-deviation.

Municipality	$RR_{GWQLASSO}$			$RR_{GWQLASSO/QR}$			$RR_{GWQLASSO/YI}$		
	2.5%	Median	97.5%	2.5%	Median	97.5%	2.5%	Median	97.5%
Alto Piquiri	-0.1264	-0.0063	0.1975	-0.1179	-0.0075*	0.1260	-0.0695	0.0611*	0.2836
Ampére	-0.1329	0.0027	0.2175	-0.1182	-0.0010	0.1383	-0.0803	0.0626*	0.2942
Andirá	-0.1192	0.1149	0.5029	-0.1770	0.0145*	0.2711	-0.0740	0.1663*	0.5406
Cambará	-0.2145	0.0042	0.2565	-0.2207	0.0039*	0.2835	-0.1647	0.0705*	0.3415
Campo Mourão	-0.1131	-0.0083	0.1159	-0.1062	0.0035*	0.1195	-0.0691	0.0454*	0.1772
Céu Azul	-0.1823	0.0135	0.3513	-0.1821	0.0034	0.2479	-0.1391	0.0501*	0.2829
Clevelândia	-0.1209	-0.0042	0.1430	-0.1177	-0.0021	0.1336	-0.0727	0.0500*	0.2040
Coronel Vivida	-0.2210	-0.0029	0.3084	-0.2217	-0.0002	0.2879	-0.1888	0.0433*	0.3327
Formosa do Oeste	-0.1360	0.0647	0.4381	-0.2105	-0.0324*	0.1685	-0.0956	0.1005*	0.4181
Foz do Iguaçu	-0.2108	0.0094	0.3584	-0.2213	-0.0010*	0.3078	-0.1740	0.0479*	0.3105
Guaraniaçu	-0.0795	0.0211	0.2260	-0.1127	-0.0073*	0.1362	0.0077	0.1257*	0.3703
Guarapuava	-0.0719	-0.0177	0.0360	-0.0585	-0.0012*	0.0587	-0.0217	0.0341*	0.0897
Ivaiporã	-0.0740	0.0154	0.1006	-0.0688	0.0062*	0.0830	-0.0230	0.0713*	0.1632
Janiópolis	-0.0900	0.0280	0.2251	-0.1225	-0.0184*	0.0775	-0.0409	0.0883*	0.3023
Mamborê	-0.0645	0.0197	0.1175	-0.0765	-0.0024	0.0782	-0.0064	0.0777*	0.1889
Manoel Ribas	-0.1054	-0.0015	0.1108	-0.1014	-0.0023	0.1090	-0.0597	0.0513*	0.1739
Mariluz	-0.1200	0.0115	0.2177	-0.1216	0.0068*	0.1904	-0.0566	0.0846*	0.3290
Mariópolis	-0.1422	-0.0030	0.2378	-0.1351	-0.0043*	0.1226	-0.1032	0.0394*	0.2540
Matelândia	-0.2383	-0.0270	0.2879	-0.2206	-0.0153*	0.2695	-0.2096	0.0042*	0.2616
Nova Esperança	-0.1800	-0.0154	0.2114	-0.1637	0.0031*	0.2192	-0.1216	0.0619*	0.3204
Palmas	-0.0509	0.0004	0.0562	-0.0529	0.0009	0.0563	-0.0098	0.0432*	0.1041
Pato Branco	-0.1490	-0.0101	0.1972	-0.1407	-0.0062*	0.1673	-0.0984	0.0503*	0.2741
Pitanga	-0.1350	-0.0257	0.0858	-0.1160	-0.0076*	0.1105	-0.0837	0.0336*	0.1576
Prudentópolis	-0.0706	-0.0164	0.0356	-0.0613	0.0001*	0.0714	-0.0274	0.0324*	0.0884
Quedas do Iguaçu	-0.1467	-0.0209	0.1694	-0.1007	0.0139*	0.1168	-0.1034	0.0335*	0.2417
Roncador	-0.1031	0.0001	0.1166	-0.1056	-0.0033	0.1189	-0.0498	0.0580*	0.1901
Salto do Lontra	-0.1636	-0.0222	0.1632	-0.1433	-0.0165*	0.1332	-0.1265	0.0203*	0.2314
Santo Antônio da Platina	-0.1060	0.0345	0.2447	-0.1200	0.0155*	0.1951	-0.0464	0.1032*	0.3303
São Jorge d'Oeste	-0.2115	-0.0019	0.3096	-0.2074	-0.0022*	0.2820	-0.1863	0.0238*	0.3015
São Miguel do Iguaçu	-0.1936	0.0278	0.4372	-0.2257	-0.0153*	0.2169	-0.1591	0.0621*	0.3582
Terra Roxa	-0.2673	0.0036	0.4015	-0.2827	-0.0060	0.3495	-0.2494	0.0257*	0.3068
Tibagi	-0.0538	0.0110	0.0876	-0.0578	-0.0040*	0.0493	-0.0107	0.0594*	0.1430
Toledo	-0.2288	-0.0173	0.3539	-0.2166	-0.0005	0.2440	-0.2068	0.0039	0.2347
Ubiratã	-0.1190	-0.0087	0.1096	-0.1136	-0.0044*	0.1130	-0.0768	0.0434*	0.1789

Note. * significant at the 5% significance level. Source: Authors.

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

13

Table 3. Relative risk reduction per municipality in cluster 2 according to the spectral risk measure (SRM).

Municipality	$RR_{GWQLASSO}$			$RR_{GWQLASSO/QR}$			$RR_{GWQLASSO/YI}$		
	2.5%	Median	97.5%	2.5%	Median	97.5%	2.5%	Median	97.5%
Palmeira	-0.0863	0.0083	0.0971	-0.0900	0.0039*	0.1028	-0.0470	0.0544*	0.1531
Piraí do Sul	-0.0359	-0.0019	0.0475	-0.0289	-0.0033*	0.0248	0.0055	0.0425*	0.0964
Ponta Grossa	-0.0427	-0.0058	0.0285	-0.0329	0.0032*	0.0444	0.0022	0.0411*	0.0782
Porto Amazonas	-0.1065	-0.0043	0.0976	-0.1010	0.0011*	0.1135	-0.0623	0.0428*	0.1542
Rio Negro	-0.0640	-0.0006	0.0623	-0.0598	0.0002	0.0665	-0.0096	0.0570*	0.1286
São Mateus do Sul	-0.0684	-0.0331	0.0048	-0.0318	-0.0045*	0.0240	-0.0160	0.0209*	0.0609
União da Vitória	-0.0705	-0.0017	0.0723	-0.0579	0.0056*	0.0690	-0.0279	0.0411*	0.1218

Note. * significant at the 5% significance level. Source: Authors.

Table 4. Relative risk reduction per municipality in cluster 2 according to the mean semi-deviation.

Municipality	$RR_{GWQLASSO}$			$RR_{GWQLASSO/QR}$			$RR_{GWQLASSO/YI}$		
	2.5%	Median	97.5%	2.5%	Median	97.5%	2.5%	Median	97.5%
Palmeira	-0.0993	0.0032	0.0995	-0.0994	0.0014	0.1127	-0.0543	0.0502*	0.1535
Piraí do Sul	-0.0330	0.0061	0.0579	-0.0309	-0.0034*	0.0268	0.0111	0.0513*	0.1083
Ponta Grossa	-0.0435	-0.0056	0.0304	-0.0338	0.0060*	0.0504	-0.0003	0.0422*	0.0815
Porto Amazonas	-0.1155	-0.0058	0.1047	-0.1101	0.0013*	0.1335	-0.0773	0.0422*	0.1638
Rio Negro	-0.0657	-0.0009	0.0679	-0.0607	0.0007	0.0704	-0.0114	0.0576*	0.1330
São Mateus do Sul	-0.0620	-0.0289	0.0096	-0.0310	-0.0045*	0.0229	-0.0112	0.0261*	0.0685
União da Vitória	-0.0757	0.0013	0.0812	-0.0664	-0.0012	0.0696	-0.0323	0.0444*	0.1291

Note. * significant at the 5% significance level. Source: Authors.

As for the public policy implications, WII has proved to be a superior alternative to basic crop insurance products as a yield insurance with a 65% coverage level. Considering that Oñate et al. (2016) showed that Proagro, a risk management tool similar to a crop credit insurance, has not increased farmers welfare and is not priced according to regional characteristics, we favor the expansion of government operated or funded parametric insurance products. A WII product could be implemented as a microinsurance policy to small farmers or as a macroinsurance directly to the government. The latter would also be further advantageous, as the efficiency of WII grows with scale (Miranda & Farrin, 2012). Several products of this type have been successfully implemented in developing countries such as the Comité de Ayuda a Desastres y Emergencias Nacionales (CADENA) program in Mexico (Janvry, Ritchie, & Sadoulet, 2016) and the Pradhan Mantri Fasal Bima Yojana (PMFBY), an index insurance scheme in India (Rathore, 2017).

CONCLUSION

Despite the efforts by the central government, crop insurance is yet to take off in Brazil. Inconsistent budget, information asymmetry, and moral hazard are some of the issues that crippled the program and continue to impede its expansion. In this sense, parametric insurance may present an alternative to the local insurance market. Thus, aiming to foster the growth of parametric insurance in Brazil and contribute to the development of this type of insurance throughout the globe, we designed a WII product by using a novel approach to model the yield-index relationship, the GWQlasso. This methodology compounds the flexible modeling and robustness of quantile regression with the spatial component of geographically weighted regression and variable selection prowess of the lasso method.

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

14

We tested our assumptions using a crop insurance application in Paraná, Brazil. The long series of 36 years of precipitation and soybean yield data is split in design and evaluation sets, with the latter having only six years of data and thus requiring the use of Bayesian bootstrap to improve the reliability of results. To measure the ability of WII to reduce risk, when compared to yield insurance and between yield modeling approaches, we used two different risk measures, the spectral risk measure and the mean semi-deviation.

Regarding the performance of WII in Paraná, our findings indicate that index insurance is superior to a 65% coverage yield insurance in 41 municipalities of the state, despite being up to three times more expensive than this product. However, the GWQlasso approach proved as effective as the regular quantile regression. The latter may seem as a discouragement to the use of a more complex model; nevertheless, some of the characteristics of GWQlasso (less data intensive and simpler conjoint variable selection) argue in its favor.

Future studies are needed to confirm the viability of WII in other regions and crops throughout the country. In addition, regardless of our efforts to mitigate the effect of the level of aggregation in the crop yield data and lack of precise planting dates, these may lead to a loss of accuracy in product design that is unacceptable in a commercial environment. Therefore, tighter cooperation between risk bearers, insurance researchers, and product developers is needed.

REFERENCES

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, *26*(7), 1505-1518. https://doi.org/10.1016/s0378-4266(02)00281-9

Banco Central do Brasil (2018). *Programa de garantia da atividade agropecuária PROAGRO: Relatório Circunstanciado 2015 a 2018.* Retrieved from https://www.bcb.gov.br/content/estabilidadefinanceira/proagro_docs/relatorios_proagro/Relatorio-circunstanciado-2015_2018.pdf

Buuren, S. V., & Groothuis-Oudshoorn, C. G. M. (2000). *Multivariate imputation by chained equations: MICE V1.0 - User manual* [TNO Report PG/VGZ/00.038]. Leiden: TNO. Retrieved from https://stefvanbuuren.name/publication/2000-01-01_vanbuuren2000/

Cai, J. (2016). The impact of insurance provision on household production and financial decisions. *American Economic Journal: Economic Policy*, *8*(2), 44-88. https://doi.org/10.1257/pol.20130371

Carmello, V., & Sant'anna, J. L., Neto (2016). Rainfall variability and soybean yield in Paraná state, southern Brazil. *International Journal of Environmental & Agriculture Research, 2*(1), 86-97. Retrieved from https://ijoear.com/Paper-January-2016/IJOEAR-JAN-2016-16.pdf

Carter, M. (2009). Intelligent design of index-insurance for smallholder farmers and pastoralists [Working paper]. *IFPRI Innovations in Insuring the Poor, Focus 17-06.* Washington, D.C.: IFPRI. Retrieved from https://cgspace.cgiar.org/handle/10568/894

Carter, M., Janvry, A., Sadoulet, E., & Sarris, A. (2017). Index insurance for developing country agriculture: A reassessment. *Annual Review of Resource Economics*, *9*, 421-438. https://doi.org/10.1146/annurev-resource-100516-053352

Chakravarti, J. S. (1920). Agricultural insurance a practical scheme suited to Indian conditions. Bangalore: Government Press.

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software, 61*(6). https://doi.org/10.18637/jss.v061.i06

Chen, V. Y.-J., Deng, W.-S., Yang, T.-C., & Matthews, S. A. (2012). Geographically weighted quantile regression (GWQR): An application to US mortality data. *Geographical Analysis*, *44*(2), 134-150. https://doi.org/10.1111/j.1538-4632.2012.00841.x

Collier, B., Skees, J., & Barnett, B. (2009). Weather index insurance and climate change: Opportunities and challenges in lower income countries. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *34*(3), 401-424. https://doi.org/10.1057/gpp.2009.11

Conradt, S., Finger, R., & Bokusheva, R. (2015). Tailored to the extremes: Quantile regression for index-based insurance contract design. *Agricultural Economics*, *46*(4), 537-547. https://doi.org/10.1111/agec.12180

Cotter, J., & Dowd, K. (2010). Estimating financial risk measures for futures positions: A nonparametric approach. *Journal of Futures Markets*, *30*(7), 689-703. https://doi.org/10.1002/fut.20437

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022 | doi.org/10.1590/1982-7849rac2022200387.en | e-ISSN 1982-7849 | rac.anpad.org.br

15

Dowd, K., Cotter, J., & Sorwar, G. (2008). Spectral risk measures: Properties and limitations. *Journal of Financial Services Research*, *34*(1), 61-75. https://doi.org/10.1007/s10693-008-0035-6

Duarte, G. V., Braga, A., Miquelluti, D. L., & Ozaki, V. A. (2018). Modeling of soybean yield using symmetric, asymmetric and bimodal distributions: Implications for crop insurance. *Journal of Applied Statistics*, *45*(11), 1920-1937. https://doi.org/10.1080/02664763.2017.1406902

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics, 7*(1), 1-26. https://doi.org/10.1214/aos/1176344552

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, *96*(456), 1348-1360. https://doi.org/10.1198/016214501753382273

Farias, J. R. B., Assad, E. D., Almeida, I. R., Evangelista, B. A., Lazzarotto, C., Neumaier, N., & Nepomuceno, A. L. (2001). Caracterização de risco de déficit hídrico nas regiões produtoras de soja no Brasil. *Revista Brasileira de Agrometeorologia*, *9*(3), 415-421. Retrieved from http://www.cnpt.embrapa.br/pesquisa/agromet/pdf/revista/cap4.pdf

Franchini, J. C., Balbinot, A. A., Junior, Nitsche, P. R., Debiasi, H., & Lopes, I. O. N. (2016). Variabilidade espacial e temporal da produção de soja no Paraná e definição de ambientes de produção. *Embrapa Soja-Documentos (INFOTECA-E)*. Retrieved from https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1052786/variabilidade-espacial-e-temporal-da-producao-de-soja-no-parana-e-definicao-de-ambientes-de-producao

Giné, X., Menand, L., Townsend, R., & Vickery, J. (2010). Microinsurance: A case study of the Indian rainfall index insurance market [Policy Research Working Papers]. *The World Bank*. https://doi.org/10.1596/1813-9450-5459

Goodwin, B. K., Vandeveer, M. L., & Deal, J. L. (2004). An empirical analysis of acreage effects of participation in the federal crop insurance program. *American Journal of Agricultural Economics*, *86*(4), 1058-1077. https://doi.org/10.1111/j.0002-9092.2004.00653.x

Grootveld, H., & Hallerbach, W. G. (2004). Upgrading value-at-risk from diagnostic metric to decision variables: A wise thing to do? In G. P. Szego (Ed.), *Risk Measures for the 21st Century* (pp. 33–50). Chichester: Wiley.

Halcrow, H. G. (1949). Actuarial structures for crop insurance. *American Journal of Agricultural Economics*, *31*(3), 418-443. Retrieved from https://www.jstor.org/stable/1232330

Hess, U., Skees, J. R., Stoppa, A., Barnett, B. J., & Nash, J. (2005). *Managing agricultural production risk: Innovations in developing countries*. Washington, DC: World Bank. Retrieved from https://openknowledge.worldbank.org/handle/10986/8797

Hunter, D. R., & Lange, K. (2000). Quantile regression via an MM algorithm. *Journal of Computational and Graphical Statistics*, *9*(1), 60-77. https://doi.org/10.1080/10618600.2000.10474866

Janvry, A. D., Ritchie, E. R., & Sadoulet, E. (2016). Weather index insurance and shock coping: Evidence from Mexico's CADENA Program [Policy Research Working Papers]. *The World Bank*. Retrieved from https://openknowledge.worldbank.org/handle/10986/24632

Jensen, N., & Barrett, C. (2017). Agricultural index insurance for development. *Applied Economic Perspectives and Policy*, *39*(2), 199-219. https://doi.org/10.1093/aepp/ppw022

Keller, T., Filho, Assad, E. D., & Lima, P. R. S. R. (2005). Regiões pluviometricamente homogêneas no Brasil. *Pesquisa Agropecuária Brasileira*, *40*(4), 311-322. https://doi.org/10.1590/S0100-204X2005000400001

Khalil, A. F., Kwon, H.-H., Lall, U., Miranda, M. J., & Skees, J. (2007). El niño–southern oscillation–based index insurance for floods: Statistical risk analyses and application to Peru. *Water Resources Research*, *43*(10). https://doi.org/10.1029/2006wr005281

Leblois, A., Quirion, P., Alhassane, A., & Traoré, S. (2014). Weather index drought insurance: An ex ante evaluation for millet growers in Niger. *Environmental and Resource Economics*, *57*(4), 527-551. https://doi.org/10.1007/s10640-013-9641-3

Lier, Q. D. J. V. (2014). Water availability to plants. In W. G. Teixeira, M. B. Ceddia, M. V. Ottoni, G. K. Donnagema (Eds.), *Application of Soil Physics in Environmental Analyses: Measuring, modelling and data integration* (pp. 435-452). Cham: Springer. https://doi.org/10.1007/978-3-319-06013-2_18

Maestro, T., Barnett, B. J., Coble, K. H., Garrido, A., & Bielza, M. (2016). Drought index insurance for the Central Valley Project in California. *Applied Economic Perspectives and Policy*, *38*(3), 521-545. https://doi.org/10.1093/aepp/ppw013

Markowitz, H. M. (1959). Portfolio selection efficient diversification of investments. *Monograph (Yale University. Cowles Foundation for Research in Economics)*, 16. *New York: John Wiley and Sons*

McKee, T. B., Doesken, N. J., & Kleist, J. (1993, January). The relationship of drought frequency and duration to time scales (pp. 179-183). *Proceedings of the Conference on Applied Climatology*, Anaheim, California, USA, 8. Retrieved from https://www.droughtmanagement.info/literature/AMS_Relationship_Drought_Frequency_Duration_Time_Scales_1993.pdf

Ministério da Agricultura, Pecuária e Abastecimento (2015). Relatório das indenizações pagas entre 2006 a 2015. *Programa de subvenção ao prêmio do seguro rural – PSR: protegendo o produtor rural*. Retrieved from https://www.gov.br/agricultura/pt-br/assuntos/riscos-seguro/seguro-rural/documentos-seguro-rural/RelatriodeSinistralidade_PSR_2006_2015.pdf

Ministério da Agricultura, Pecuária e Abastecimento (2017). Relatório geral 2017. *Programa de Subvenção ao Prêmio do Seguro Rural (PSR)*. Retrieved from https://www.gov.br/agricultura/pt-br/assuntos/riscos-seguro/seguro-rural/documentos-seguro-rural/RelatorioGeralPSR2017.pdf

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

16

Miquelluti, D. L. (2019). *Weather index insurance design: A novel approach for crop insurance in Brazil* (Doctoral dissertation). Universidade de São Paulo, Piracicaba, SP, Brazil. https://doi.org/10.11606/t.11.2019.tde-02082019-100224

Miranda, M. J. (1991). Area-yield crop insurance reconsidered. *American Journal of Agricultural Economics*, *73*(2), 233-242. https://doi.org/10.2307/1242708

Miranda, M. J., & Farrin, K. (2012). Index insurance for developing countries. *Applied Economic Perspectives and Policy*, *34*(3), 391-427. https://doi.org/10.1093/aepp/pps031

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, *391*(1-2), 202-216. https://doi.org/10.1016/j.jhydrol.2010.07.012

Mishra, P. K. (1995). Is rainfall insurance a new idea? Pioneering work revisited. *Economic and Political Weekly*, *30*(25), A84-A88. Retrieved from https://www.jstor.org/stable/4402912

Oñate, C. A., Ozaki, V. A., & Bravo-Ureta, B. (2016, July). Impact evaluation of the Brazilian crop insurance public program "Proagro Mais". *Proceedings of the Agricultural and Applied Economics Association – AAEA 2016 Annual Meeting*, Boston, Massachusetts. https://doi.org/10.22004/ag.econ.236096

Ozaki, V. A., & Shirota, R. (2005). Um estudo da viabilidade de um programa de seguro agrícola baseado em um índice de produtividade regional em Castro (PR). *Revista de Economia e Sociologia Rural*, *43*(3), 485-503. https://doi.org/10.1590/s0103-20032005000300005

Rao, K. N. (2011). Weather index insurance: Is it the right model for providing insurance to crops? *ASCI Journal of Management*, *41*(1), 86-101. Retrieved from https://drive.google.com/drive/folders/18Z4YazKMNL0jkgkdKCPx2QuFfPpXL5yy

Rathore, V. (2017). The performance of PMFBY and other crop insurance models in India. *International Journal of Advanced Research and Development*, *2*(5), 602-607. Retrieved from http://www.advancedjournal.com/archives/2017/vol2/issue5/2-5-222

Rubin, D. B. (1981). The bayesian bootstrap. *The Annals of Statistics*, *9*(1), 130-134. https://doi.org/10.1214/aos/1176345338

Schnitkey, G. (2014). Area risk protection insurance policy: Comparison to group plans. *Farmdoc Daily, 4*(2). Retrieved from https://farmdocdaily.illinois.edu/2014/01/area-risk-protection-insurance-policy.html

Secretaria da Agricultura e do Abastecimento (2021). Boletins conjunturais. *Governo do Estado do Paraná*. Retrieved from http://www.agricultura.pr.gov.br/Pagina/Boletins-Conjunturais

Skees, J. R., Black, J. R., & Barnett, B. J. (1997). Designing and rating an area yield crop insurance contract. *American Journal of Agricultural Economics*, *79*(2), 430-438. https://doi.org/10.2307/1244141

Skees, J., Gober, S., Varangis, P., Lester, R., & Kalavakonda, V. (2001). *Developing rainfall-based index insurance in Morocco* [Policy Research Working Papers]. *The World Bank*. Retrieved from http://hdl.handle.net/10986/19674

Skees, J. R., Hartell, J., & Murphy, A. G. (2007). Using index-based risk transfer products to facilitate micro lending in Peru and Vietnam. *American Journal of Agricultural Economics*, *89*(5), 1255-1261. https://doi.org/10.1111/j.1467-8276.2007.01093.x

Skees, J. R. (2008). Challenges for use of index-based weather insurance in lower income countries. *Agricultural Finance Review*, *68*(1), 197-217. https://doi.org/10.1108/00214660880001226

Wang, W., Xu, S., & Yan, T. (2018). Structure identification and model selection in geographically weighted quantile regression models. *Spatial Statistics*, *26*, 21-37. https://doi.org/10.1016/j.spasta.2018.05.003

Weschenfelder, A. B., Pickbrenner, K., & Pinto, E. J. D. A. (2011). Análise da frequência de ocorrência e a classificação das precipitações diárias máximas anuais na região da Laguna dos Patos (sub-bacia 87). *Anais do Simpósio Brasileiro de Recursos Hídricos*, Maceió, AL, Brazil, 19.

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

17

Authorship

Daniel Lima Miquelluti*

Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz", Departamento de Economia, Administração e Sociologia

Av. Pádua Dias, n. 11, Cx. Postal 9, 13418-900, Piracicaba, São Paulo, Brazil

E-mail address: danielmiq@usp.br

https://orcid.org/0000-0002-6335-3618

Vitor Augusto Ozaki

Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz", Departamento de Economia, Administração e Sociologia

Av. Pádua Dias, n. 11, Cx. Postal 9, 13418-900, Piracicaba, São Paulo, Brazil

E-mail address: vitorozaki@usp.br

https://orcid.org/0000-0002-0781-4475

David José Miquelluti

Universidade do Estado de Santa Catarina, Centro de Ciências Agroveterinárias, Departamento de Solos e Recursos Naturais

Av. Luiz de Camões, n. 2090, Conta Dinheiro, 88520-000, Lages, SC, Brazil

E-mail address: david.miquelluti@udesc.br

https://orcid.org/0000-0002-7369-6163

* Corresponding Author

Conflict of Interests

The authors have stated that there is no conflict of interest.

Funding

Plagiarism Check

The RAC maintains the practice of submitting all documents approved for publication to the plagiarism check, using specific tools, e.g.: iThenticate.

Authors' Contributions

1st author: conceptualization (lead); data curation (lead); formal analysis (lead); funding acquisition (lead); investigation (lead); methodology (lead); software (lead); validation (lead); writing-original draft (lead); writing-review & editing (equal).

2nd author: conceptualization (supporting); data curation (supporting); formal analysis (supporting); investigation (supporting); methodology (supporting); project administration (lead); supervision (lead); validation (supporting); writing-review & editing (equal).

3rd author: data curation (supporting); formal analysis (supporting); methodology (supporting); software (supporting); supervision (supporting); writing-review & editing (equal).

Copyrights

Peer Review Method

This content was evaluated using the double-blind peer review process. The disclosure of the reviewers' information on the first page, as well as the Peer Review Report, is made only after concluding the evaluation process, and with the voluntary consent of the respective reviewers and authors.

Data Availability

The authors claim that all data and materials have been made publicly available through the Harvard Dataverse platform and can be accessed at:

Miquelluti, Daniel Lima; Ozaki, Vitor Augusto; Miquelluti, David José, 2021, "Replication Data for: An application of geographically weighted quantile LASSO to weather index insurance design published by RAC - Revista de Administração Contemporânea", Harvard Dataverse, V1. https://doi.org/10.7910/DVN/UEZMJT

RAC encourages data sharing but, in compliance with ethical principles, it does not demand the disclosure of any means of identifying research subjects, preserving the privacy of research subjects. The practice of open data is to enable the reproducibility of results, and to ensure the unrestricted transparency of the results of the published research, without requiring the identity of research subjects.

Revista de Administração Contemporânea, v. 26, n. 3, e-200387, 2022| doi.org/10.1590/1982-7849rac2022200387.en| e-ISSN 1982-7849 | rac.anpad.org.br

18