

# Análise e quantificação do risco para a gestão eficiente do portfólio agrícola das seguradoras

Vitor Augusto Ozaki<sup>1</sup>  
Carlos Tadeu dos Santos Dias<sup>2</sup>

**Resumo:** A presença de eventos considerados catastróficos pode comprometer o balanço financeiro da carteira agrícola das seguradoras. Por exemplo, a estiagem pode afetar não somente um produtor, mas milhares de produtores em uma vasta extensão territorial. Para contornar esse problema, as seguradoras podem diversificar geograficamente suas operações. Nesse sentido este estudo aplica a análise de agrupamentos em um conjunto de dados de produtividade municipal de milho, no período de 1990 a 2005, para o estado do Paraná. Ao todo, 39 grupos foram formados conforme características similares de produtividade esperada e risco relativo. Para que as perdas sejam minimizadas, o estudo exclui alguns municípios de alto risco e sugere a classificação dos grupos, de acordo com o grau de risco relativo.

**Palavras-chaves:** seguro agrícola, análise de agrupamento, diversificação de risco.

**Abstract:** *In the agricultural insurance, the presence of adverse events can compromise the financial health of the agricultural insurance companies. Drought, for example, can affect not only one producer, but thousand of producers in a huge territorial extension. To skirt this problem, insurance companies diversify geographically their operations. This paper applies the cluster analysis in a corn yield data set for the regions of Paraná state, considering the period from 1990 to 2005. In total, 39 groups were formed according to similar characteristics of expected yield and relative risk. Based on one of insurability criteria, some high-risk regions were excluded, and groups were classified according to the degree of relative risk.*

**Key-words:** *crop insurance, cluster analysis, risk diversification.*

**Classificação JEL:** Q19.

<sup>1</sup> Professor Doutor do Departamento de Economia, Administração e Sociologia da Esalq/USP, coordenador do Geser (Grupo de Estudos em Seguros e Riscos). E-mail: vaozaki@esalq.usp.br

<sup>2</sup> Professor Associado do Departamento de Ciências Exatas, Área de Matemática e Estatística da Esalq/USP. E-mail: ctsdias@esalq.usp.br

## 1. Introdução

A agricultura é uma atividade econômica que apresenta elevado grau de risco, pois depende de condições que o homem não consegue controlar. Fenômenos climáticos adversos, pragas e doenças podem afetar negativamente a produção, causando graves prejuízos econômicos aos produtores. Por exemplo, no Rio Grande do Sul, na safra 2004/05, as perdas na soja, milho e trigo, devido à estiagem, resultaram em prejuízos da ordem de US\$ 1,5 bilhão (EMATER/RS, 2005)<sup>3</sup>.

Eventos dessa natureza são caracterizados como catastróficos, generalizados ou também denominados sistêmicos<sup>4</sup>. A ocorrência de tais eventos dificulta sobremaneira a continuidade das companhias seguradoras no ramo de atividade. Isso porque a seca apresenta elevada severidade e atinge não apenas uma propriedade rural, mas milhares de propriedades em uma grande extensão territorial.

Por esse motivo, diz-se que o risco é altamente correlacionado entre unidades seguradas. Esse fato viola um dos princípios básicos do mercado de seguros: as unidades expostas devem ser homogêneas e independentes (OZAKI, 2006b). Esta condição de segurabilidade diz que as unidades seguradas devem estar geograficamente dispersas.

Para ilustrar, considera-se que  $X$  seja uma variável aleatória, representando a indenização recebida por um grupo de  $n$  indivíduos independentes e idênticos. Assim,  $X_i$  ( $i = 1, 2, \dots, n$ ) será uma amostra de variáveis aleatórias independentes e identicamente distribuídas, tal que  $E(X_i) = \mu$  e  $\text{Var}(X_i) = \sigma^2$ .

Nesse contexto, denota-se o coeficiente de variação  $CV_X = (\sigma/\mu) \times 100$  como uma medida de risco associada à variável aleatória e relativo ao seu tamanho esperado. Definindo uma nova variável  $S$ , tal que  $S = \sum X_i$ , então  $E(S) = n\mu$  e  $V(S) = n\sigma^2$ . O coeficiente de variação da nova variável será dado por  $CV_S = \sigma/\mu (n)^{1/2} \times 100$ .

Dessa forma, quando  $n \rightarrow \infty$ ,  $CV_S \rightarrow 0$ , supondo-se que  $\sigma$  e  $\mu$  não se alterem à medida que  $n$  aumenta. Em outras palavras, quando os sinistros são independentes e identicamente distribuídos, o risco agregado se torna menor do que o risco individual. Este resultado é uma consequência da “Lei dos Grandes Números”.

No seguro agrícola, devido à existência do risco catastrófico, o risco agregado se torna, algumas vezes, muito maior que o risco individual. Nesse caso, sinistros generalizados podem resultar em enormes prejuízos à carteira agrícola de uma seguradora, tornando insustentável sua continuidade.

Uma possível solução a ser adotada pelas seguradoras é a atuação em diversas regiões distintas umas das outras em termos de características de

<sup>3</sup> Comunicação pessoal.

<sup>4</sup> Furacões, terremotos e erupções vulcânicas são outros exemplos de fenômenos catastróficos que atingem diversos países.

produtividade agrícola e estrutura de risco, a fim de pulverizar espacialmente o risco. Nesse sentido, este artigo estuda o comportamento da produtividade nos municípios do Paraná e agrupa aqueles com características semelhantes. Para isso, a seção 2 deste estudo explica, de modo geral, o funcionamento do seguro agrícola. A seção 3 detalha a metodologia estatística da análise de agrupamentos e a seção 4 aborda os dados utilizados no estudo. Na seção 5, os resultados são discutidos e as possíveis implicações para o seguro agrícola são abordadas na seção 6 e na subseção 6.1. As considerações finais estão na seção 7.

## 2. O seguro agrícola

Basicamente, a idéia do seguro é indenizar ao produtor toda vez que sua produtividade se situe em um patamar abaixo da produtividade garantida pela seguradora. O mecanismo de compensação  $I$  para cada propriedade rural  $i$  pode ser expressa da seguinte forma:

$$I_i = \max\{[(y_i^c - y_i) / y_i^c], 0\} \quad (1)$$

Em que:

$y_i^c$  é o nível de produtividade garantida da  $i$ -ésima propriedade rural; e,  
 $y_i$  é a produtividade observada da  $i$ -ésima propriedade rural.

A eq. (1) mostra que se a produtividade agrícola  $y_i$ , em determinado período, for menor que a produtividade garantida  $y_i^c$  (fixada a priori), o produtor é indenizado no montante igual a diferença<sup>5</sup>. Para ter direito a indenização, o segurado deve pagar um prêmio (preço do seguro) e seguir certas condições presentes na apólice de seguro.

Neste tipo de contrato o produtor tem a opção de escolher o nível de cobertura  $\alpha_i$  que desejar, tal que  $0 \leq \alpha_i \leq 1$ . Assim, a produtividade garantida é calculada de acordo com a equação:  $y_i^c = \alpha_i \mu_i$ , em que  $\mu_i$  é a produtividade esperada do produtor  $i$ .

Normalmente, sobre o montante a ser indenizado, é aplicado um percentual denominado franquia ou dedutibilidade. Essa é a parte do prejuízo sob responsabilidade do segurado. Este mecanismo reduz o problema do risco moral (fraude) (SPENCE e ZECKHAUSER, 1971; PAULY, 1974; NELSON e LOEHMAN, 1987; GOODWIN e SMITH, 1996; QUIGGIN et al., 1994; OZAKI, 2006a).

<sup>5</sup> Para se calcular o valor da responsabilidade por hectare (em termos monetários), deve-se multiplicar esta diferença por um preço determinado a priori.

### 3. Metodologia

#### 3.1. Análise de agrupamentos<sup>6</sup>

A análise de agrupamentos, como técnica multivariada, tem por objetivo reunir as observações de uma amostra em grupos, de tal forma que os elementos pertencentes a um mesmo grupo sejam semelhantes (homogêneos) entre si, e os elementos contidos em grupos diferentes sejam distintos (heterogêneos), tendo como base alguma medida de (di)similaridade sobre um conjunto de variáveis medidas ou avaliadas.

Em diversas áreas, a análise de agrupamentos tem sido aplicada com o intuito de se detectar padrões de semelhanças nos dados analisados; em economia e finanças, com a finalidade de se determinar o agrupamento de ativos relacionados à agroindústria, por exemplo: derivativos agropecuários e ações na Bolsa de Valores (KAMOGAWA et al., 2006); na análise de mercado, em que o objetivo é agrupar consumidores, de acordo com determinadas características (MACLACHLAN e JOHANSSON, 1981; PUNJ e STEWART, 1983).

No campo da ergonomia, a metodologia foi aplicada a um problema de seleção e avaliação de assentos para trabalho em dados de questionários, que foram, posteriormente, padronizados (LINDEN et al., 2002). Em estudos biogeográficos da fauna e flora na Bélgica, por meio da detecção de áreas homogêneas quanto a variáveis ecológicas, tais como: temperatura média, precipitação média, altitude, entre outros (DUFRENE e LEGENDRE, 1991). Na educação, analisando padrões de comportamento de estudantes, utilizando inventários de comportamento em salas de aula (SPEECE et al., 1985).

Em dados de mineração<sup>7</sup> o problema consiste em extrair informações implícitas, determinar padrões, associações e anomalias em grandes quantidades de dados. Assim, pelo menos uma das possibilidades em dados de mineração é a análise de agrupamentos (HAND, 1998). Aplicações em genética, em que os genes são alocados a determinados grupos conforme suas similaridades em padrões de expressões genéticas (EISEN et al., 1998; DE SMET, et al., 2002; TAVAZOIE et al., 1999).

Para se agrupar um conjunto de dados com  $n$  elementos amostrais e  $p$  variáveis aleatórias, deve-se, inicialmente, adotar alguma medida de similaridade ou dissimilaridade. Neste estudo, optou-se por uma medida de dissimilaridade  $D(X_l, X_m)$  conhecida como medida do quarteirão (também denominada "city-block" ou "Manhattan"), em que  $X_l$  e  $X_m$  são vetores  $p$ -variados das propriedades rurais  $l$  e  $m$ , respectivamente. Nesse caso, quanto

<sup>6</sup> Do termo em inglês "cluster analysis". Também denominada análise de classificação ou conglomerados.

<sup>7</sup> Do termo em inglês "Data mining".

menor for o seu valor, maior a similaridade entre os elementos comparados (MINGOTI, 2005). De forma geral, pode-se descrever essa medida como:

$$D(X_l, X_m) = \left\{ \sum_{i=1}^p |X_{il} - X_{im}|^v \right\}^{1/v} \quad (2)$$

em que  $p$  é o número de variáveis, e  $v = 1, 2, 3, \dots, \infty$ .

Na eq. (1)  $D(X_l, X_m)$  representa a métrica de Minkowski entre duas observações  $l$  e  $m$ . Quando  $v = 1$ ,  $D(X_l, X_m)$  se torna a distância quarteirão (“city-clock”) e quando  $v = 2$ ,  $D(X_l, X_m)$  é a distância euclidiana<sup>8</sup>. Pode-se notar que o parâmetro  $v$  altera o peso associado à diferença entre  $X_l$  e  $X_m$ . Por exemplo, considerando a distância euclidiana, a presença de valores discrepantes nos dados tem o efeito de aumentar consideravelmente a estimativa da distância.

Entretanto, tais medidas têm a desvantagem de não serem invariantes em escala. Em outras palavras, as distâncias entre observações podem se alterar em função da mudança na escala da variável medida. Em casos em que não é possível se obter dados medidos em escalas comparáveis, podem ser utilizadas distâncias estatísticas, tais como: distância euclidiana para dados padronizados e a distância de Mahalanobis (SHARMA, 1996).

Nesse estudo, as variáveis utilizadas possuem a mesma escala (quilogramas por hectare). Porém, constatou-se a presença de valores discrepantes<sup>9</sup> na maioria dos municípios analisados. Por esse motivo, optou-se pela distância quarteirão. Além disso, considerou-se a priori que o número de grupos  $K$  é conhecido. Esse número está baseado nas microrregiões pré-estabelecidas pelo Instituto Brasileiro de Geografia e Estatística (IBGE), que subdividiu o estado em 39 microrregiões diferentes<sup>10</sup>.

Tendo em vista que  $K$  é conhecido, o método não-hierárquico denominado  $K$ -médias será utilizado para organizar  $n$  elementos (propriedades), em  $p$  dimensões, em  $K$  grupos. Basicamente, o algoritmo aloca cada ponto ao grupo que tenha o centróide (média vetorial) mais próximo, para que a soma dos quadrados  $V$  dentro de cada grupo  $i$  seja minimizada (HARTIGAN e WONG, 1979; WONG e LANE, 1983):

$$V_i = \sum_{j=1}^k \sum_{j \in K_i} |X_j - \mu_i|^2 \quad (3)$$

em que  $j$  é o número de elementos em cada grupo  $K_i$ , tal que  $X_j \in K_i$ ; e,  $\mu_i$  é o centróide de cada grupo  $C_i$ .

<sup>8</sup> Johnson e Wichern (2002, p.670 e 671) apresentam outras medidas de dissimilaridade.

<sup>9</sup> Por meio da análise exploratória dos dados (“box plots”).

<sup>10</sup> No anexo estão as 39 microrregiões e seus respectivos municípios.

Formalmente, dado uma matriz  $X$  de  $n$  elementos em um espaço  $p$ -variado  $\mathfrak{R}^p$ , o processo de agrupamento consiste em particionar as linhas de  $X$  em  $K$  grupos, tal que a soma dos quadrados das distâncias dos elementos de  $X$  em relação ao centróide mais próximo seja minimizada. Assim, para um conjunto de elementos  $x$ , e um elemento  $s$ , ambos em  $\mathfrak{R}^p$ ,  $D(s,x) = \min_x ||xs||$ , em que  $|| \cdot ||$  é a norma e  $x \in X$ , então:

$$K\text{-Médias} = \min_x \sum_s \{D(s,x)\}^2 \quad (4)$$

De modo geral, o processo de agrupamento ocorre em três passos. Inicialmente, determina-se  $K$  centróides<sup>11</sup>. O segundo passo consiste na comparação de cada uma das observações com cada centróide inicial e sua subsequente alocação para o grupo cuja distância é menor. Recalcula-se o centróide para os novos grupos. No terceiro passo, repete-se o passo anterior até que não ocorram novas realocações (estabilização do grupo) ou que o valor do centróide não se altere no recálculo (JOHNSON e WICHERN, 2002).

## 4. Fonte de dados

O banco de dados em painel de produtividade agrícola municipal é proveniente do Instituto Brasileiro de Geografia e Estatística (IBGE). As variáveis temporais compreendem os anos de 1990 até 2005<sup>12</sup>. A cultura analisada é o milho. Como as séries temporais são relativamente curtas, optou-se por analisar os municípios com o máximo de informação disponível (apenas aqueles com 16 observações temporais). Assim, dos 399 municípios existentes no Paraná, 291 participaram deste estudo<sup>13</sup>.

<sup>11</sup> Ressalta-se que existem diferentes maneiras de se escolher as sementes iniciais de agrupamento, que podem influenciar o agrupamento final. Neste estudo as sementes iniciais foram baseadas nas  $K$  primeiras observações do banco de dados. Mingoti (2005) sugere a utilização desse procedimento quando as  $K$  primeiras observações são discrepantes entre si, como ocorre nos dados analisados na pesquisa.

<sup>12</sup> Estimativa preliminar para a safra 2004/05.

<sup>13</sup> Os *missing values* poderiam ser estimados por métodos de imputação. Porém, isso teria pouca relevância tendo em vista o objetivo central do estudo que é analisar o agrupamento dos municípios por meio da análise estatística multivariada.

## 5. Aplicação empírica

O método de agrupamentos foi aplicado aos dados a fim de se obterem grupos de municípios que apresentem padrões semelhantes no comportamento das séries de produtividade agrícola. A Figura 1 mostra os resultados dos agrupamentos.

De forma geral, pode-se observar duas características marcantes nos mapas da Figura 1: a formação de subgrupos dentro de um mesmo grupo, espalhados em diferentes regiões; e, os municípios pertencentes a um mesmo subgrupo são, na maioria das vezes, contíguos. Esse fato é especificamente marcante no grupo 6, 7 e 9 (Figura 1).

Pode-se notar que o grupo 6 é formado por dois subgrupos formados por municípios que apresentam fronteiras em comum e um município isolado no extremo norte do estado. O primeiro subgrupo situa-se no nordeste do estado (formado por três municípios), e o outro, no extremo leste (com quatro municípios). O grupo 7, por sua vez, subdivide-se em dois grupos: o primeiro deles localizado na parte central do estado (três municípios) e o segundo, na região oeste (quatro municípios). O grupo 9 é formado por três subgrupos e quatro municípios espalhados nas regiões central e sul do estado. Nesse grupo, o primeiro subgrupo (com três municípios) localiza-se na parte centro-norte do estado paranaense. O segundo, situa-se na parte leste (dois municípios) e o último subgrupo, na parte sul (quatro municípios).

Como era de se esperar, as produtividades em municípios próximos tendem a apresentar o mesmo comportamento, devido às características edafoclimáticas similares. Por isso, tais municípios formaram um mesmo grupo.

Analisando o grupo 13 com mais atenção, percebe-se que nesse grupo estão os municípios com os níveis de produtividade esperada mais elevados. São eles: Castro (6.448 kg/ha), Ponta Grossa (5.999 kg/ha), Marilândia do Sul (5.934 kg/ha), Tibagi (5.689 kg/ha), Catanduvas (5.494 kg/ha) e Céu Azul (5.185 kg/ha). Por outro lado, no grupo 6, encontram-se os municípios com os menores valores de produtividade esperada: Paranapoema (2.131 kg/ha), Pinhalão (1.982 kg/ha), Japira (1.978 kg/ha), Jaboti (1.939 kg/ha), Guaratuba (1.642 kg/ha), Morretes (1.642 kg/ha), Paranaguá (1.614 kg/ha) e Antonina (1.583 kg/ha).

A Tabela 1 mostra os grupos, com o número de seus respectivos municípios, em ordem decrescente de produtividade esperada.

Figura 1. Grupos formados pelo método de agrupamento K-médias, no Paraná.



Fonte: Resultados da pesquisa.



**Tabela 1.** Produtividade esperada (kg/ha) e coeficiente de variação médio (em %) dos 39 grupos formados e seus respectivos municípios.

Grupos	Municípios	Produtividade esperada	Coeficiente de variação médio
13	6	5791	24.7
22	3	4960	24.0
9	13	4795	30.4
37	5	4659	42.1
10	6	4423	23.5
29	7	4252	25.1
11	8	4144	32.1
31	4	4116	22.4
7	7	3960	21.6
17	12	3761	33.2
33	1	3749	31.8
19	8	3683	23.1
34	8	3592	29.7
8	3	3538	25.1
20	4	3483	25.4
2	3	3471	23.2
35	v5	3470	20.7
4	17	3404	37.1
32	1	3386	29.8
24	2	3347	29.1
28	13	3306	23.1
38	15	3098	26.2
5	11	3068	25.9
15	13	3035	33.4
36	8	2918	28.6
30	3	2895	25.7
3	12	2894	24.5
27	13	2843	22.1
39	9	2789	35.9
1	8	2775	28.1
21	7	2730	29.3
16	3	2715	29.0
23	5	2643	23.4
18	4	2557	29.7
25	13	2509	23.2
12	9	2459	33.1
26	12	2444	29.6
14	2	2152	27.1
6	8	1814	20.2
<b>total</b>	<b>291</b>	-	-

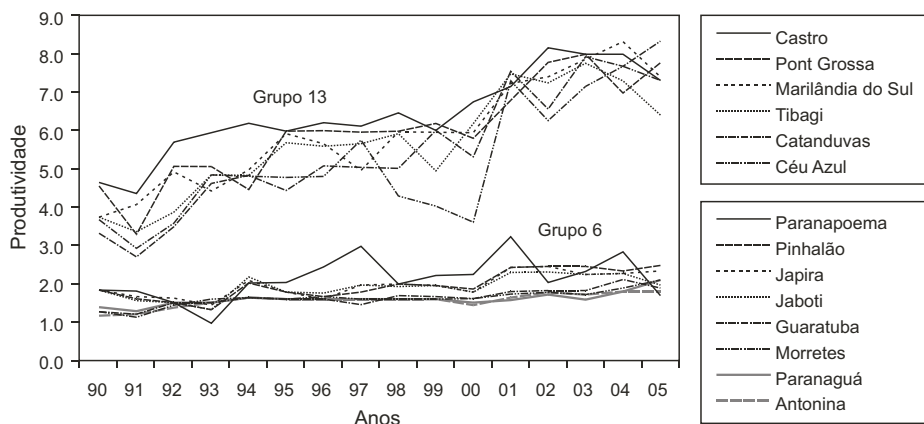
Fonte: Resultados da pesquisa.

Observa-se certa consistência nos grupos formados, em relação ao comportamento das séries de produtividade esperada. Os municípios foram alocados aos seus respectivos grupos em função da semelhança em suas produtividades esperadas.

Comparando o grupo 13 ao 6, percebe-se que existem dois grupos com comportamento bastante distintos (Figura 2). Nota-se de forma bem evidente uma tendência pronunciada da produtividade nos municípios do primeiro grupo 13, em função da incorporação de tecnologia nesse período, que resultou em ganhos consideráveis de produtividade. Por outro lado, embora os municípios do grupo 6 apresentem uma leve tendência de crescimento, suas taxas de crescimento são bem menores em relação às dos municípios do grupo 13.

Entretanto, um olhar mais atento revela que o grupo 13 apresenta variabilidade média bem maior do que o grupo 6. De fato, o desvio padrão médio do primeiro grupo é quase quatro vezes maior do que o desvio padrão do segundo grupo, sugerindo que o risco nesse grupo é comparativamente mais elevado.

**Figura 2.** Produtividades esperadas (1000 kg/ha) para os municípios dos grupos 13 e 6, no período 1990 a 2005.



Fonte: Resultados da pesquisa.

## 6. Implicações para o seguro agrícola

Diferentemente dos outros ramos, o seguro agrícola apresenta uma característica peculiar que a torna uma modalidade eminentemente arriscada: a presença do risco de eventos generalizados ou catastróficos (MIRANDA e GLAUBER, 1997; OZAKI e SHIROTA, 2005). Esse risco é caracterizado por fenômenos adversos que afetam uma grande quantidade de segurados, em certo período de tempo.

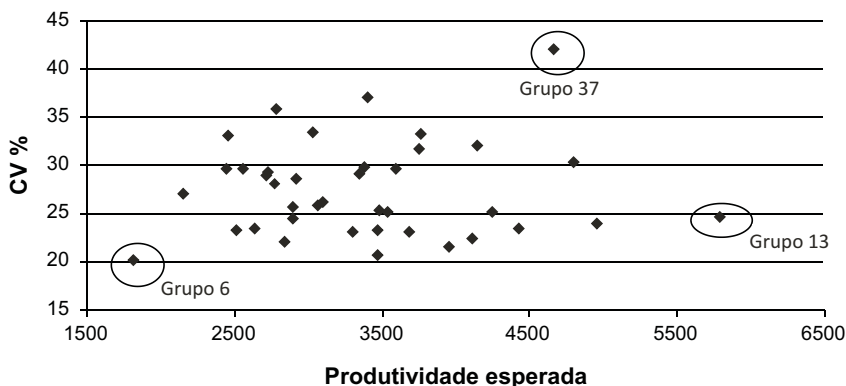
A estiagem, por exemplo, pode afetar não somente um produtor, mas milhares de produtores em uma vasta extensão territorial. Nessa situação, o pagamento das indenizações para todos os segurados pode comprometer o balanço financeiro da carteira agrícola das seguradoras, podendo levar ao encerramento de suas atividades, como ocorreu com a seguradora Porto Seguro, que encerrou a carteira agrícola em 1998, e mais recentemente, com a seguradora Minas Brasil (OZAKI, 2006c), que operou até a safra 2005/06.

Para contornar esse problema, as seguradoras necessitam classificar os municípios, de acordo com o grau de risco, e diversificar as operações em diferentes regiões geográficas, a fim de pulverizar o risco retido. Essa estratégia permite que os sinistros ocorridos em determinadas regiões possam ser compensados por resultados positivos em outras localidades.

Os municípios de alto risco não serão cobertos pelo seguro, em bases puramente privadas<sup>14</sup>. Aos municípios restantes, aplica-se o esquema de diversificação. A métrica utilizada para quantificar o risco será o coeficiente de variação, que reflete a magnitude relativa do risco (Figura 3).

A Figura 3 mostra as produtividades esperadas e os respectivos riscos associados a cada grupo. Nota-se que o grupo 37<sup>15</sup> possui o maior risco agregado, embora apresente alta produtividade esperada. Por outro lado, o grupo 6 possui o menor risco e também a menor produtividade média. O grupo 13 possui risco intermediário, próximo a 25%, mas com a maior produtividade esperada.

**Figura 3.** Produtividades esperadas (kg/ha) de todos os grupos (1 a 39) e seus respectivos coeficientes de variação (CV), em percentuais.



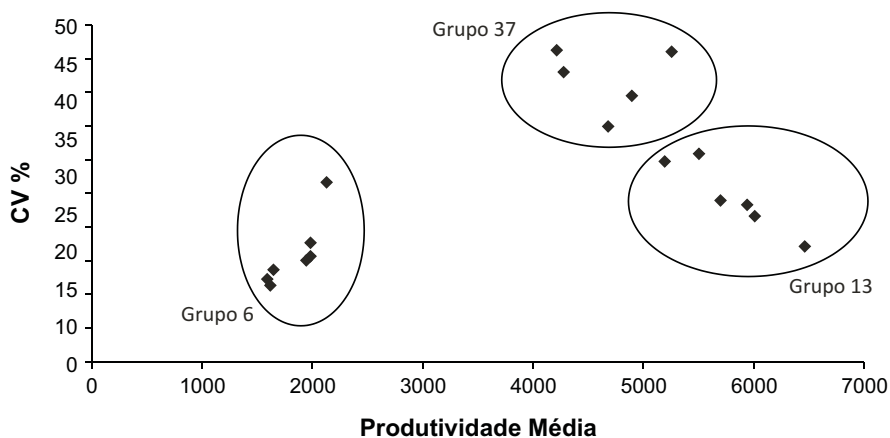
Fonte: Resultados da pesquisa.

<sup>14</sup> Para esses municípios, é necessário que haja subvenção governamental, para que as seguradoras tenham incentivo em aceitar o elevado risco.

<sup>15</sup> Composto pelos seguintes municípios: Campo Bonito, Pato Branco, Vitorino, Paula Freitas e São Mateus do Sul.

A Figura 4 mostra as produtividades esperadas e os coeficientes de variação dos municípios dos grupos 37 e 6. Nota-se a formação de dois grupos: um de baixo risco (6) e outro de alto risco (37) e um grupo de risco intermediário (13).

**Figura 4.** Produtividades médias (kg/ha) e riscos associados (CV) aos municípios dos grupos 37 e 6.



Fonte: Resultados da pesquisa.

O critério de escolha dos municípios de alto risco a serem excluídos será a comparação entre o risco individual e o risco agregado. Dessa forma, os municípios que apresentarem valores de risco (coeficiente de variação) maiores em relação ao risco agregado, representado pelo coeficiente de variação agregado dos 291 municípios analisados (37,6%), serão excluídos das carteiras das seguradoras. A idéia geral do critério consiste no fato de que o risco agregado deve ser menor que cada risco isolado, para que o princípio da diversificação seja eficientemente aplicada.

Desagregando os grupos e analisando isoladamente o risco de cada município, nota-se que 23 municípios apresentaram risco mais elevado do que o risco agregado. São eles: Agudos do Sul (pertencente ao grupo 39), Antônio Olinto (4), Balsa Nova (4), Campo do Tenente (4), Cândido de Abreu (12), Carlópolis (15), Contenda (4), Corumbataí do Sul (4), Guaraci (36), Irati (15), Jaguariaíva (39), Mariópolis (9), Pato Branco (37), Paula Freitas (37), Paulo Frontin (17), Pitanga (15), Rebouças (17), Rio Negro (4), Roncador (4), São Mateus do Sul (37), Sengés (4), Verê (4) e Doutor Ulysses (37). A nova configuração dos grupos pode ser observada na Tabela 2:

**Tabela 2.** Produtividade média (kg/ha) e coeficiente de variação (em %) dos grupos modificados.

Grupos	Número de Municípios	Produtividade média	Coefficiente de Variação
4	8	3349	31.6
9	12	4763	29.6
12	8	2456	32.3
15	10	3002	31.1
17	10	3785	32.3
36	7	2886	26.5
37	1	4675	35.0
39	7	2762	32.3

Fonte: Resultados da pesquisa.

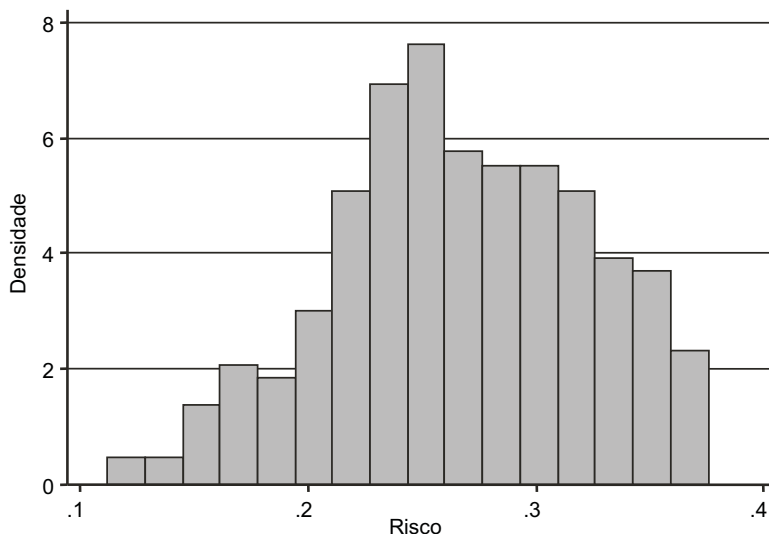
Para melhor visualizar os grupos de acordo com o seu respectivo risco associado, o Quadro 1 mostra a classificação dos grupos de acordo com a faixa de risco em que se encontram. Nesse sentido, os grupos foram subdivididos em 8 classes, sendo que a classe 1 representa os grupos de baixo risco – coeficiente de variação entre 20 e 22%, e a classe 8, aqueles de maior risco relativo – coeficiente de variação maiores do que 34%.

**Quadro 1.** Estratificação do risco e classificação dos grupos.

Classes	Faixas de risco (%)	Grupos								
		6	35	7						
1	20-22	6	35	7						
2	22-24	27	31	19	28	25	2	23	10	22
3	24-26	3	13	29	8	20	30	5		
4	26-28	38	36	14						
5	28-30	1	16	24	21	9	26	34	18	32
6	30-32	15	4	33						
7	32-34	11	17	12	39					
8	> 34	37								

Fonte: Resultados da pesquisa.

O Quadro 1 mostra que quase 50% dos grupos se encontram nas três primeiras classes de risco, consideradas de baixo risco. Pouco mais de 31% são considerados de risco mediano, e aproximadamente 20%, de alto risco. Com relação aos municípios, a Figura 5 mostra o histograma representativo do risco relativo de cada município.

**Figura 5.** Histograma do risco relativo dos respectivos municípios.

Fonte: Resultados da pesquisa.

Pelo histograma nota-se que há uma freqüência maior dos municípios com risco intermediário (ao redor de 0.25) e também uma leve assimetria negativa, sugerindo que exista uma freqüência maior de municípios com alto risco relativo.

### 6.1. Precificação e níveis de cobertura

As Tabelas 1 e 2 fornecem as produtividades esperadas e os riscos relativos de todos os grupos e seus respectivos municípios. Com essas informações, as seguradoras podem estrategicamente definir os locais para ofertar seus produtos. Levando em conta a estratificação apresentada no Quadro 1, o seguro poderá se restringir apenas às regiões menos arriscadas, por exemplo, municípios do grupo 6, 7 e 35, ou a todos os grupos, independentemente do risco.

Em ambos os casos, a taxa de prêmio irá variar conforme o nível de risco de cada unidade segurada. O ideal seria precificar cada propriedade rural separadamente. Entretanto, como dados estatísticos de produtividade por propriedades rurais são praticamente inexistentes, os contratos são precificados, utilizando-se dados de produtividade municipal.

Naturalmente, as taxas de prêmio para os municípios de baixo risco serão menores que as taxas para os municípios de alto risco (OZAKI et al., 2006). Além das taxas de prêmio, os níveis de cobertura também devem variar, de acordo com o risco, ou seja, coberturas menores aos municípios de alto risco e, vice-versa.

## 7. Conclusão

Por meio da análise de agrupamento, os municípios do estado do Paraná foram agrupados conforme suas similaridades na variável de produtividade agrícola de milho, entre os anos de 1990 a 2005. Embora a metodologia tenha sido aplicada aos municípios do Paraná, pode-se ampliar a análise para todos os estados do Brasil. Os resultados mostram que os grupos formados apresentam relativa semelhança entre os municípios do grupo na tendência e variabilidade das séries de produtividade.

Neste estudo, os dois primeiros momentos da distribuição de probabilidade – a esperança e a variância – têm grande importância, na medida em que formam a métrica (coeficiente de variação) utilizada para analisar o risco relativo da produtividade. Nesse caso, supõe-se que a distribuição de probabilidade seja simétrica. Para os casos em que as distribuições sejam consideradas assimétricas, momentos superiores deverão ser considerados na análise.

Para o seguro agrícola, a quantificação municipalizada do risco é fundamental para que as seguradoras possam trabalhar com risco diversificado e precificar adequadamente os contratos. Uma vez que os grupos estejam formados, excluem-se os municípios com risco maior que o risco agregado.

A metodologia utilizada no trabalho possibilita criar grupos de municípios que tenham perfis de produtividade esperada e risco semelhantes, com o objetivo de detectar os municípios de alto risco da carteira das seguradoras e, conseqüentemente, minimizar os efeitos dos sinistros e os prejuízos da seguradora.

Atualmente, as políticas de aceitação de risco (subscrição) variam de seguradora para seguradora. Entretanto, o seguro só é ofertado nos municípios em que existam informações do zoneamento agrícola<sup>16</sup>, do Ministério da Agricultura, Pecuária e Abastecimento. As seguradoras podem atuar em municípios de alto risco, porém, devem diversificar, atuando também em municípios de baixo risco.

Se optar pelos municípios de alto risco, a taxa de prêmio será proporcionalmente mais elevada. Nesse caso, poderá haver baixa adesão por parte dos produtores rurais. Para evitar esse problema, o governo federal criou o programa de subvenção ao prêmio do seguro rural, que reduz consideravelmente o prêmio do seguro.

Por exemplo, a subvenção para o milho se encontra atualmente em 60% do prêmio. Caso o produtor tenha um prêmio estimado em R\$ 100,00 deverá pagar R\$ 40,00 e o governo federal arcará com o restante, R\$ 60,00. Portanto, a

<sup>16</sup> Representa um importante mecanismo de gestão de risco e auxilia o mercado de seguro agrícola. Basicamente, identifica, para cada município, a melhor época de plantio das culturas nos diferentes tipos de solo e ciclos dos cultivares.

seguradora tem a opção de operar nas regiões de alto risco, cobrando um prêmio proporcional ao nível de risco, no entanto, a um custo relativamente menor ao segurado.

Outra implicação importante ao se detectar os municípios de alto risco pela metodologia proposta é a possibilidade de direcionar os recursos dos programas estaduais de subvenção para esses municípios. Ou seja, da mesma forma que o governo federal, os governos estaduais também estão criando seus próprios programas de subvenção ao prêmio, em que o percentual de subvenção poderá ser acumulado com o percentual de subvenção federal, reduzindo ainda mais o prêmio do seguro, principalmente nas regiões de risco mais elevado.

Desta forma, compensado o excesso de risco com subvenções, o seguro poderá avançar em áreas agrícolas importantes, mas que não atraem as seguradoras devido ao histórico elevado de perdas.

## 8. Referências Bibliográficas

De SMET, F.; MATHYS, J.; MARCHAL, K.; THIJS, G.; De MOOR, B.; MOREAU, Y. Adaptive quality-based clustering of gene expression profiles. **Bioinformatics**, v.18, n.5, p.735-746, 2002.

DUFRENE, M.; LEGENDRE, P. Geografic structure and potential ecological factors in Belgium. **Journal of Biogeography**, v.18, n.3, p.257-266, May 1991.

EISEN, M.B; SPELLMAN, P.T.; BROWN, P.O.; BOTSTEIN, D. Cluster analysis and display of genome-wide expression patterns. **Proceeding of the National Academy of Sciences**, v.95, p.14863-14868, Dec. 1998.

HAND, D.J. Data mining: statistics or more? **The American Statistician**, v.52, n.2, p.112-118, May 1998.

HARTIGAN, J.A.; WONG, M.A. Algorithm AS 136: a Kmeans clustering algorithm. **Applied Statistics**, v.28, n.1, p.100-108, 1979.

JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. New Jersey: Prentice Hall, 2002. 767p.

GOODWIN, B.K.; SMITH, V.H. Crop insurance, moral hazard and agricultural chemical use. **American Journal of Agricultural Economics**, v.78, n.2, p.428-438, May 1996.

KAMOGAWA, L.F.; OZAKI, V.A.; CRUZ JÚNIOR, J.C.; FONSECA, R.M. Uso da análise de componentes principais para a criação de clusters como mecanismo de diversificação de carteira de ativos do setor agroindustrial. In: **Simpósio Nacional de Probabilidade e Estatística (SINAPE)**. Caxambu, 2006. **Anais**. Caxambu: Associação Brasileira de Estatística, julho de 2006.



LINDEN, J.C.S.; WERNER, L.; RIBEIRO, J.L.D. Aplicação de ferramentas estatísticas multivariadas para a análise da percepção sobre assentos de trabalho. In XII CONGRESSO LATINO-AMERICANO DE ERGONOMIA, 2002, Recife-PE. **Anais**. 2002.

MACLACHLAN, D.L.; JOHANSSON, J.K. Market segmentation with multivariate aid. **Journal of Marketing**, v.45, p.74-84, 1981.

MINGOTI, S.A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Minas Gerais: Editora UFMG, 2005. 295p.

MIRANDA, M. J.; GLAUBER, J. W. Systemic risk, reinsurance, and the failure of crop insurance markets. **American Journal of Agricultural Economics**, v.79, n.1, p. 206-215, fev 1997.

NELSON, C.H.; LOEHMAN, E.T. Further toward a theory of agricultural insurance. **American Journal of Agricultural Economics**, v.69, n.3, p.523-531, Aug. 1987.

OZAKI, V.A. O mercado de seguros à luz da teoria dos mercados contingentes. **Working paper**, Dept. de Economia, Administração e Sociologia ESALQ/USP, 29p., 2006a.

OZAKI, V.A. O papel do seguro na gestão do risco agrícola e os empecilhos para o seu desenvolvimento. **Working paper**, Dept. de Economia, Administração e Sociologia ESALQ/USP, 20p., 2006b.

OZAKI, V.A. O seguro rural estadual e as novas iniciativas privadas. **Agricultura em São Paulo**, v.53, n.1, p.91-106, jan/jun. 2006c.

OZAKI, V.A.; GOODWIN; B. K.; SHIROTA, R. Parametric and nonparametric statistical modeling of crop yield: implications for pricing crop insurance contracts. **Applied Economics**, v.40, n.7-9, p. 1151-1164, abril-maio 2008.

OZAKI, V.A.; SHIROTA, R. 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**, v.43, n.3, p.485-503, 2005.

PAULY, M.V. Overinsurance and public provision of insurance: the roles of moral hazard and adverse selection. **Quarterly Journal of Economics**, v.88, n.1, p.44-62, Feb. 1974.

PUNJ, G.; STEWART, D.W. Cluster analysis in marketing research: review and suggestions for application, **Journal of Marketing Research**, v.20, p.134-148, May 1983.

QUIGGIN, J.; KARAGIANNIS, G.; STANTON, J. Crop insurance and crop production: an empirical study of moral hazard and adverse selection. In: HUETH, D.L.; FURTAN, W.H. **Economics of agricultural crop insurance: theory and evidence**. Boston: Kluwer Academic Publishers, 1994. 380p.

SHARMA, S. **Applied multivariate techniques**. New York: John Wiley & Sons Inc., 1996. 493p.

SPEECE, D.L.; MCKINNEY, J.D.; APPELBAUM, M.I. Classification and validation of behavioral subtypes of learning-disabled children. **Journal of Educational Psychology**, v.77, n.1, p.67-77, 1985.

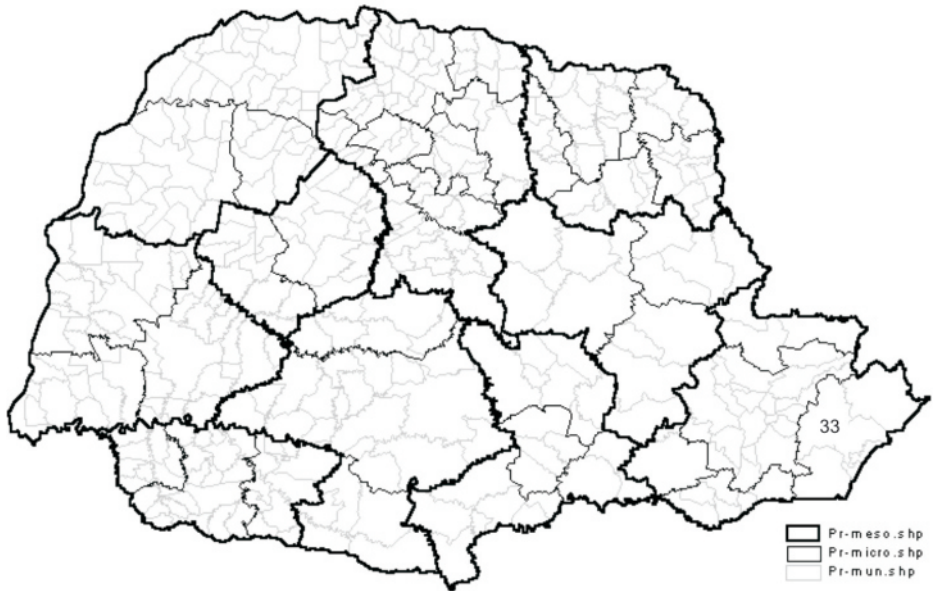
SPENCE, M.; ZECKHAUSER, R. Insurance, information, and individual action. **American Economic Review**, v.61, n.2, p.380-387, May 1971.

TAVAZOIE, S.; HUGHES, J.D.; CAMPBELL, M.J.; CHO, R.J.; CHURH, G.M. Systematic determination of genetic network architecture. **Nature Genetics**, v.22, p.281-285, 1999.

WONG, A.M.; LANE, T. A Kth nearest neighbour clustering procedure. **Journal of the Royal Statistical Society (B)**, v.54, n.3, 362-368, 1983.

## Anexo

### Paraná: Mesorregiões, Microregiões e Municípios - 1999



Fonte: IBGE (2006).