

UM MODELO DE DESAGREGAÇÃO DE ENCARGOS DE PRODUÇÃO VARIÁVEIS DA BASE RICA POR ACTIVIDADE

MARIA DO SOCORRO ROSÁRIO¹
Gabinete de Planeamento e Políticas

RUI MANUEL DE SOUSA FRAGOSO²
Universidade de Évora

MARIA LEONOR DA SILVA CARVALHO³
Universidade de Évora

RESUMO

A base de dados da RICA contém informação relativa ao total dos custos por tipo de factor não sendo, no entanto, desagregada por actividade. A obtenção dos custos variáveis por actividade é extremamente importante para a gestão da exploração e análise da política agrícola.

O objectivo deste artigo é estimar os coeficientes de afectação dos custos às actividades agrícolas, a partir de informação incompleta. Um modelo de desagregação dos custos de factores de produção variáveis por actividade com base na teoria da máxima entropia foi desenvolvido e aplicado à região Alentejo, usando os dados da base RICA de 2008.

A utilização da teoria da entropia mostrou ser uma opção adequada para a estimação daqueles coeficientes de afectação. Os estimadores de entropia apresentam de um modo geral bons resultados do ponto de vista estatístico e econométrico.

PALAVRAS-CHAVE

Máxima entropia; custos variáveis; modelo de desagregação; base RICA

¹ Gpp, MAMAOT, Rua Pe. António Vieira n° 1, 3°-1099-073 Lisboa, Portugal, socorro@gpp.pt

² Departamento de Gestão, CEFAGE, Universidade de Évora. Largo dos Colegiais, 2 –7000 Évora, Portugal, rfragoso@uevora.pt

³ Departamento de Economia, ICAAM, Universidade de Évora. Largo dos Colegiais, 2 –7000 Évora, Portugal, leonor@uevora.pt

1. Introdução

Na União Europeia há uma grande diversidade de estruturas e sistemas de produção e, apesar da crescente tendência para a especialização, a maioria das explorações agrícolas são multiproducto.

Para facilitar a análise e para haver comparabilidade das características estruturais e dos resultados económicos das explorações agrícolas dos diferentes Estados Membros, foi desenvolvida uma Rede de Informação de Contabilidades Agrícolas (RICA) com o objectivo de recolher anualmente toda a informação relativa à contabilidade da exploração agrícola, e a Tipologia para a Classificação Comunitária das Explorações Agrícolas (Tipologia) pertencentes à RICA.

Em Portugal, desde os anos 80 que a RICA tem desenvolvido um trabalho de recolha de informação a nível da exploração agrícola, numa amostra constante conforme o Regulamento n.º 79/65/CEE que institui a RICA e o Regulamento n.º 2143/81/CEE que a melhora. A informação disponível baseia-se nos dados de cada exploração agrícola numa visão contabilística, em que os custos totais por tipo de factor são agregados, não sendo detalhados por actividade.

Sendo a competitividade um objectivo amplamente divulgado e solicitado às empresas do sector agrícola, ela não pode ser alcançada apenas por via dos preços de comercialização de produto, pois as empresas compradoras querem cada vez mais pagar menos, independentemente da origem do produto. A forma de aumentar a competitividade nesse contexto consiste em melhorar a eficiência (gestão) na utilização dos factores de produção nas diversas actividades.

A análise económica das actividades agrícolas em resposta às mudanças de medidas de políticas e de mercado, sistemas de produção, tecnologias e competitividade também se baseiam na estrutura de custos de produção das actividades. Estas estruturas são inerentes à existência de uma empresa agrícola e esses custos podem ser variáveis, fixos e atribuídos, sendo os primeiros os custos específicos de uma actividade. A informação relativa aos encargos fixos e variáveis de produção por unidade de actividade é extremamente importante, não só ao nível da gestão da empresa agrícola mas também na análise de políticas. O agricultor, apesar de dispor da informação necessária, não consegue detalhá-la na forma contabilística mais amigável para a análise económica da empresa e a gestão das suas actividades, e que permita saber quais as actividades mais interessantes a desenvolver e as que podem ser alteradas. As explorações agrícolas que fazem parte da RICA têm essa informação agregada para a exploração por tipo de factores de produção e nunca separados por actividade, ou seja, as actividades têm a sua produção individualizada mas os custos dos factores de produção estão organizados de uma forma agregada para a totalidade da exploração. Desta forma não é possível conhecer a estrutura de custos e, por conseguinte, a rentabilidade por actividade.

O uso de técnicas baseadas em métodos econométricos ou na metodologia da máxima entropia pode ser uma solução válida para estimar os encargos com um custo relativamente reduzido. Para tentar responder ao problema da falta de informação relativa aos custos dos factores por actividade agrícola, este artigo tem como principal objectivo estimar os coeficientes de afectação desses custos às actividades agrícolas, a partir da base de dados da RICA. Para o efeito é desenvolvido um modelo de desagregação dos custos de factores de produção variáveis por actividade com base na teoria da máxima entropia. Este modelo é aplicado no contexto da região agrária do Alentejo, segundo os dados da RICA de 2008.

Este trabalho encontra-se organizado em 6 partes. No ponto 2 faz-se uma breve revisão dos principais métodos utilizados na desagregação de custos, dando especial atenção aos métodos baseados na teoria da Máxima Entropia; no ponto 3 descreve-se a metodologia e no ponto 4 os dados empíricos utilizados. No ponto 5 analisam-se os resultados e finalmente no ponto 6 sintetizam-se as principais conclusões.

2. Desagregação de Custos de Produção e Estimação de Coeficientes de Afectação

A falta de informação sobre os sistemas de produção é um problema determinante que pode ser atenuado com o recurso a métodos de desagregação de dados, nomeadamente dos custos da produção agrícola por actividades. Deste modo o problema da falta de informação transforma-se num problema

de desagregação de custos, formalizado para uma amostra de T explorações, que realizam K actividades agrícolas utilizando factores de produção que dão origem a I itens de custo:

$$X_i^t = \sum_{k=1}^K \alpha_{ik} Y_k^t + u_i^t \quad i=1,2,\dots,I \quad e \quad t=1,2,\dots,T \quad (1)$$

X_i^t = Custo do *input* i por exploração t;

α_{ik} = Coeficiente de afectação de custos por *input* i e por produto k;

Y_k^t = Receita total ou produto bruto por produto k e por exploração t;

u_i^t = Resíduo aleatório da distribuição que é específico de cada *input* i e cada exploração t.

Os coeficientes de afectação α_{ik} são os parâmetros desconhecidos da equação (1) e são definidos em termos dos seus valores por actividade agrícola ou produto e por item de custo ou input para a globalidade das empresas da amostra objecto de estudo.

Existem vários métodos para estimar os custos específicos das actividades das explorações agrícolas. Estas estimativas são baseadas nas relações que se estabelecem entre os coeficientes de afectação da produção e os respectivos factores, assumindo a hipótese de que os agentes económicos têm um comportamento de maximização do lucro e tendo em conta as devidas desagregações dos factores, estimativas de custos de produção e coeficientes de produtividade (Just et al., 1983; Lence & Miller, 1998a e 1998b).

Para se proceder à desagregação dos custos por actividade agrícola através da estimativa de parâmetros, pode recorrer-se às técnicas de regressão linear. No entanto, a fiabilidade dos parâmetros estimados depende do número de empresas ou de explorações agrícolas da amostra. Segundo Love (1999) e Zhang & Fan (2001), quando a informação provém de uma amostra pequena, as técnicas de regressão linear não podem ser utilizadas, pelo que deverão ser adoptados outros procedimentos metodológicos.

Por outro lado, Peeters & Surry (2002; 2005) referem que a imposição de restrições contabilísticas, como a que garante a igualdade do rendimento bruto total aos custos totais, dificultam ou inviabilizam mesmo a utilização de algumas técnicas de análise para estimar os coeficientes de desagregação dos custos. Estes autores também se referem às técnicas de regressão linear e à técnica de estimação bayesiana, como as metodologias clássicas mais usadas, apesar dos problemas que ambas comportam na sua aplicação.

Nas técnicas de regressão linear, os principais problemas prendem-se com a possibilidade de obtenção de estimativas negativas e com o facto de todas as equações serem interdependentes (Bewley, 1986). O problema da negatividade dos coeficientes pode ser resolvido com a utilização de técnicas de estimação bayesiana ou através de método dos mínimos quadrados com restrições (Moxey e Tiffin, 1994). No entanto, estes métodos, para além de não permitirem incorporar as restrições de saldo contabilístico, a sua aplicação empírica é muito complexa.

No seu estudo, Pires et al. (2010) concluiu que o método dos mínimos quadrados não era adequado para estimar parâmetros a partir de amostras pequenas, contrariamente ao método de máxima entropia (ME) que nessas condições permitiu melhores estimativas na desagregação de dados. Estas conclusões estão de acordo com os resultados apresentados por outros autores, como Golan & Judge (1996), Golan & Perloff (2002) e Kitamura & Stuzer (1997), que foram testemunhando ao longo dos anos a capacidade da ME para estimar a desagregação de dados económicos.

Golan et al. (1994) usou a ME e a mínima entropia cruzada (EC) para derivar uma matriz de contabilidade social (*social accounting matrix*). Mais tarde, Golan et al. (1997) e Golan et al. (2001) utilizaram o método da máxima entropia generalizada (MEG) para estimar os parâmetros de um modelo de regressão. Paris & Howitt (1998) recorreram a uma formulação de MEG para estimar os parâmetros das funções custo das explorações a serem incorporadas na função objectivo de um modelo de programação matemática positiva.

Leon et al. (1999) e Peeters & Surry (2002) mostraram que a qualidade de precisão dos coeficientes extraídos de amostras de explorações agrícolas com dados incompletos através dos métodos de

entropia é elevada. Estes autores, para além de concluírem que as características estatísticas demonstradas pelos estimadores da entropia são boas, verificaram também que os resultados não apresentam diferenças significativas, quando comparados com uma amostra de encargos observados desagregados por actividade.

Fraser (2000) utilizou a MEG para estimar a procura de carne no Reino Unido e Zhang & Fan (2001) recorreram à MEG para estimar as tecnologias de produção na agricultura chinesa e para proceder à desagregação dos factores de produção por província.

Garvey & Britz (2002) estimaram a desagregação de custos a partir da contabilidade da RICA também com a MEG e Campbell & Hill (2005 e 2006) mostraram como colocar restrições de informação num modelo de MEG utilizando diferentes vectores de suporte.

Howitt & Msangi (2006) recorreram à MEG para estimar funções de produção numa amostra pequena de 27 explorações no Rio Bravo, México. Rezek & Campbell (2006) apoiaram-se na MEG para estimar os custos de diversos efluentes /poluentes a partir de dados de painel.

Hansen & Surry (2006) derivaram a procura de quantidades de factores para diferentes produções na Alemanha utilizando o método de máxima entropia aplicado aos dados das Contas Económicas Regionais

Fragoso et al. (2008) utilizaram um modelo EC para proceder à desagregação das superfícies agro-florestais de acordo com zonas agro-ecológicas pré-definidas para o ecossistema montado no Alentejo.

Lips em 2009, partindo da informação da RICA, utilizou a ME para desagregar os custos das várias actividades pelas diversas áreas de culturas arvenses na Suíça.

Xavier et al. (2010) utilizou a ME para criar um modelo de desagregação dos dados da utilização da terra, dos efectivos pecuários e de produções. Papalia (2010) também usou a MEG para a estimação de um modelo de interacção espacial a partir de dados de painel.

3. Metodologia

Para proceder à desagregação dos custos de produção da exploração agrícola por actividade a partir da base de dados da RICA utilizou-se o método da ME, método que permite estimar parâmetros sem a imposição de restrições à distribuição da probabilidade e dos erros a partir de amostras de tamanho reduzido (Golan et al., 1996; Campbell & Hill, 2005, 2006). Este método tem também a vantagem de não requerer informação sobre o comportamento do agricultor e de poder integrar no modelo informação adicional que não consta da amostra.

Considerou-se um modelo de MEG, cujos resultados foram melhorados com ajuda de um *prior* de informação numa formulação de entropia cruzada generalizada (ECG). A aplicação desses modelos de ME permite estimar os coeficientes de afectação dos custos dos factores de produção da exploração agrícola pelas suas actividades através de um sistema de equações lineares. Este conjunto de equações não é mais do que as restrições de informação que condicionam as estimativas dos coeficientes de afectação. Os custos dos factores são assumidos como uma função linear do valor da produção das actividades.

A Máxima Entropia Generalizada

A metodologia da máxima entropia generalizada (MEG) tem sido muito utilizada para a resolução de problemas relacionados com a desagregação de factores. Por exemplo, Peeters & Surry (2002 e 2005) adoptaram este método para estimar os coeficientes de afectação dos custos dos factores de produção às actividades agrícolas a partir de uma amostra de explorações. Golan et al., (2001) e Peeters & Surry (2002, 2005) utilizaram esta metodologia tendo introduzido uma formulação Tobit nos seus modelos de ME. Este procedimento permite separar as situações em que a informação está disponível daquelas em que não existe. A sua aplicação é muito útil no caso da aplicação dos modelos de ME ao contexto das explorações agrícolas, na medida em que nem todas as explorações realizam todas as actividades agrícolas disponíveis, correspondendo, por conseguinte, nestes casos, a uma situação de ausência de informação, que é necessário corrigir. Deste modo, tendo em conta a formulação Tobit, as equações de informação na sua forma matricial e vectorial passam a ser as seguintes:

$$X = YA + u = \begin{bmatrix} x_1 > 0 \\ x_2 = 0 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} A + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (2)$$

onde x_1 , y_1 e u_1 dizem respeito aos respectivos valores não nulos da informação observada e x_2 , y_2 e u_2 representam a ausência de informação.

De acordo com Golan et al. (1996b) e Peeters & Surry (2002), o modelo estatístico linear pode ser alterado, reparametrizando A e u em função dos seus parâmetros aleatórios p e w_1 e w_2 e dos vectores de suporte z_1 , z_2 , v_1 e v_2 :

$$\begin{bmatrix} x_1 > 0 \\ x_2 = 0 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} z'p + \begin{bmatrix} v'_1 & w_1 \\ v'_2 & w_2 \end{bmatrix} \quad (3)$$

A seguir apresenta-se a formulação Tobit simplificada do modelo de MEG que foi utilizado para estimar os coeficientes de afectação dos custos dos factores de produção agrícola.

$$\begin{aligned} \text{Max}(p, w_1, w_2) = & \\ & - \sum_{i=1}^I \sum_{k=1}^K \sum_{m=1}^M p_{k,i}^m \ln(p_{k,i}^m) - \\ & \sum_{t=1}^T \sum_{i=1}^I \sum_{n=1}^N w_{t,i}^n \ln(w_{t,i}^n) - \sum_{t=1}^T \sum_{i=1}^I \sum_{n=1}^N w_{t,i}^n \ln(w_{t,i}^n) \end{aligned} \quad (4)$$

Sujeito a:

$$X1_i^t = \sum_{k=1}^K \sum_{m=1}^M p_{k,i}^m \cdot z_{k,i}^m \cdot y1_k^t + \sum_{n=1}^N w_{t,i}^n \cdot v_{t,i}^n, \quad \forall t1 \text{ e } \forall i \in \{1, \dots, T\} \quad (5)$$

$$X2_i^t = \sum_{k=1}^K \sum_{m=1}^M p_{k,i}^m \cdot z_{k,i}^m \cdot y2_k^t + \sum_{n=1}^N w_{t,i}^n \cdot v_{t,i}^n, \quad \forall t2 \text{ e } \forall i \in \{1, \dots, T\} \quad (6)$$

$$\sum_{m=1}^M p_{k,i}^m = 1 \quad \forall k \in \{1, 2, \dots, I\} \quad (7)$$

$$\sum_{n=1}^N w_{t,i}^n = 1, \quad \forall i \in \{1, 2, \dots, I\} \text{ e } \forall t \in \{1, 2, \dots, T_1\} \quad (8)$$

$$\sum_{n=1}^N w_{t,i}^n = 1, \quad \forall i \in \{1, 2, \dots, I\}, \forall t \in \{1, 2, \dots, T_2\} \text{ e } T_1 + T_2 = T \quad (9)$$

$$\sum_{i=1}^I \sum_{m=1}^M p_{k,i}^m \cdot z_{k,i}^m = 1, \quad \forall k \in \{1, 2, \dots, K\} \quad (10)$$

A equação (4) diz respeito à maximização da entropia conjunta da distribuição de probabilidades dos coeficientes de afectação dos custos dos factores de produção ($p_{k,i}^m$) e das distribuições de probabilidades dos erros ($w_{t,i}^n$ e $w_{t,i}^n$).

As equações (5) e (6) representam as restrições de consistência da informação, referindo-se a primeira às $t1$ explorações com valores positivos em $y1_k^t$ e a segunda às $t2$ explorações com valores nulos em $y2_k^t$. Essas equações compreendem a necessidade de consistência de informação com um mínimo de conteúdo, tratando as relações da equação linear do problema (1) como um sistema independente de equações, em que todos os factores de produção i são contabilizados em conjunto e em simultâneo. No entanto os valores positivos são separados dos valores nulos.

As equações (7)-(9) são restrições relativas às propriedades da probabilidade e criam no modelo a normalização dos valores de probabilidade de $p_{k,i}^m$ e $w_{t,i}^n$ relativamente às suas dimensões M e N , respectivamente. Garante-se deste modo que a soma das probabilidades estimadas é igual à unidade.

A equação (10) corresponde a uma condição adicional que estabelece que o somatório dos coeficientes de afectação dos custos dos factores de produção estimados para cada actividade k seja igual a 1.

O Modelo de Mínima Entropia Cruzada

Para tentar melhorar os resultados das estimativas dos coeficientes de afectação dos custos, utilizou-se um *prior* de informação exterior à amostra de explorações agrícolas objecto de estudo, recorrendo à formulação ECG. Neste caso, o *prior* de informação utilizado diz respeito à estrutura dos custos utilizada pelo Gabinete de Planeamento e Políticas (GPP) para realizar o apuramento das margens brutas padrão das actividades agrícolas para o ano de 2004.

Esse procedimento tem sido muito utilizado. Harris (2002) utilizou técnicas de mínima entropia cruzada para reconciliar dados referentes a diferentes fontes na estimativa da matriz contabilística social regional no México, onde as regiões rurais são diferenciadas por sistemas de produção agrícola com diversas tecnologias. Fragoso et al. (2008) mostrou que a utilização da mínima entropia cruzada pode ser um processo muito útil para a desagregação espacial de dados. Xavier et al. (2010) utilizou a mínima entropia cruzada para criar um modelo de desagregação dos dados da agricultura da região Algarve para um nível local zona de intervenção florestal.

Neste caso, o modelo de ECG tem a mesma estrutura de equações do que o modelo de MEG, ou seja, inclui as restrições de coerência de informação que permitem relacionar as variáveis dependentes com as variáveis independentes do problema e as restrições das propriedades das probabilidades. A única alteração que se verifica é ao nível da função objectivo que passa a ser a minimização da entropia cruzada entre distribuição dos coeficientes a estimar e a distribuição dos coeficientes que constituem o prior de informação, como se representa a seguir:

$$\text{Min}(\alpha, w_1, w_2) = \sum_{i=1}^I \sum_{k=1}^K \alpha_k^i \ln \left(\frac{\alpha_k^i}{q_k^i} \right) + \sum_{t=1}^T \sum_{i=1}^I \sum_{n=1}^N w_{1,t,i}^i \ln (w_{1,t,i}^i) + \sum_{t=1}^T \sum_{i=1}^I \sum_{n=1}^N w_{2,t,n}^i \ln (w_{2,t,n}^i) \quad (11)$$

Onde α é o coeficiente de afectação dos custos que se pretende estimar e q é o coeficiente conhecido do *prior* de informação utilizado.

Limites para os intervalos dos vectores de suporte

A escolha dos vectores de suporte, nomeadamente, dos limites do intervalo a que pertencem os seus valores e número de pontos suporte desses vectores é sempre uma tarefa difícil e delicada, sendo necessário ter em conta não só a amplitude do intervalo, como também a simetria da distribuição e por conseguinte da forma como se distribuem os seus valores no espaço.

Golan et al. (1996a) considera que se aumentar a largura do intervalo dos limites do vector suporte e o número de pontos suporte, pode-se aumentar a informação produzida dos dados estimados. No entanto, Paris & Howitt (1998) defendem que, apesar dos valores dos vectores suporte z poderem modificar as estimativas de uma forma significativa, não há uma regra para proceder à escolha de um vector certo para a distribuição.

No que se refere ao vector de suporte do erro, é muito comum o uso da regra dos três sigmas (3σ), que consiste em considerar o erro centrado na origem e admitir uma amplitude de 3 desvios padrão à direita e de 3 desvios padrão à esquerda (Pukelsheim, 1994). Exemplos da utilização deste procedimento para tratar o erro da distribuição são os trabalhos de Golan et al. (1996a), Peeters & Surry (2002), Chakir (2009), Martins et al. (2011), Pires et al. (2010) e Fragoso & Carvalho (2011).

Peeters & Surry (2002 e 2005) utilizaram, respectivamente, num estudo feito para o Canadá e num estudo feito para França, um vector suporte z definido no intervalo $[0,1]$, que nesses casos eram os únicos valores teoricamente aceitáveis, onze pontos de suporte ($M=11$) e a regra dos três sigmas na definição do erro com um vector de suporte v com três pontos de apoio ($N=3$).

Fragoso e carvalho (2011) definiram os limites do vector suporte z no intervalo $[0,1]$ e consideraram apenas 3 pontos de apoio ($M=3$). Para o erro consideraram a regra dos três sigmas na amplitude do intervalo do vector v . Fragoso et al. (2008) e Van Delden et al., (2006) também consideraram o mesmo procedimento, que está de acordo com (Golan et al., 1996). O processo de definição do erro seguido por esses autores também foi verificado por Chakir (2009) e Martins et l. (2011).

Neste estudo vai-se considerar para o vector de suporte das probabilidades estimadas $M=3$ pontos e $Z=\{0, 0,5, 1\}$ porque os resultados não pareceram ser sensíveis a outras alternativas de z . No que diz respeito à definição do erro adoptou-se a regra dos três sigmas e estabeleceram-se três pontos de apoio ($N=3$) para o vector v , i.e., $v = \{-\sigma, 0, \sigma\}$.

Indicadores de qualidade e precisão

Para avaliar a qualidade e a precisão das estimativas dos modelos de entropia apresentados na secção anterior, considerou-se o cálculo dos indicadores da entropia normalizada, do pseudo-R², dos ganhos

de informação obtidos no processo de desagregação de custos e do desvio absoluto médio. Esses indicadores têm sido utilizados por vários autores para aferir as propriedades estatísticas dos estimadores da entropia e para concluir da sua validade prática, por comparação com dados observados (Peeters & Surry, 2002; Howitt & Msangi, 2006; e Fragoso & Carvalho, 2011).

O indicador de entropia normalizada (S) é definido por Golan et al. (1996) como a proporção de incerteza que permanece na estimativa. Este indicador é muito útil para avaliar sobre o desempenho relativo das estimativas dos coeficientes da entropia no que diz respeito à geração de nova informação e à participação da informação inicial contida nos dados nessas estimativas.

No caso deste estudo, com base em Golan et al. (1996b) e em Peeters e Surry (2003), considerou-se o cálculo da entropia normalizada dos coeficientes de afectação dos custos dos factores de produção ($S(\hat{p})$) e o cálculo da entropia normalizada para a estimativa dos erros ($S(\hat{w})$):

$$S(\hat{p}) = \frac{-\hat{p}' \ln \hat{p}}{IK \ln(M)}$$

$$S(\hat{w}) = \frac{-\hat{w}' \ln \hat{w}}{IK \ln(N)}$$

sendo I o número de factores de produção ou itens de custo, K o número de actividades agrícolas, T o número de explorações e M e N o número de pontos de apoio considerados para os vectores suporte z e v.

O valor do indicador da entropia normalizada varia entre 0 e 1. O valor zero indica que não há incerteza na estimativa e que a informação dos dados está totalmente reflectida nas estimativas, não se tendo por conseguinte produzido qualquer tipo de informação nova no processo de desagregação dos dados. Pelo contrário, se o valor indicado é igual à unidade, a incerteza é máxima e a participação dos dados na informação contida nas estimativas é mínima, indicando que a produção de informação nova é máxima.

Para avaliar a capacidade de estimação do modelo utilizou-se como indicador o Pseudo R^2 , que foi também utilizado com o mesmo fim por Peeters & Surry (2002), Fragoso & Carvalho (2011), Fragoso et al. (2008), Chakir (2009), Martins et al. (2011) e Xavier et al. (2010).

O Pseudo R^2 é definido como o quadrado do coeficiente da correlação entre os valores estimados e os valores observados:

$$R_i^2 \equiv \frac{|\sum_{t=1}^T \hat{x}_t^i \hat{x}_t^i|^2}{|\sum_{t=1}^T \hat{x}_t^i \hat{x}_t^i|^2}, \quad \text{com } \hat{x}_t^i = \sum_{k=1}^K \hat{\alpha}_k^i y_{k,t}$$

Os valores do indicador Pseudo R^2 variam entre 0 e 1, significando no primeiro caso uma reduzida capacidade de estimação do modelo e no último a capacidade estimação é máxima.

Uma das medidas da qualidade da aderência das estimativas com a realidade e simultaneamente da quantidade de informação produzida no processo de desagregação é o indicador de ganhos de informação desagregada (GID). Este indicador foi estabelecido por Howitt & Reynaud (2003) sob a designação original de Disaggregation Informational Gain (DIG) e entre nós foi utilizado por Xavier et al., (2010) e por Fragoso et al., (2008). O cálculo deste indicador é dado por:

$$GID = 1 - \frac{\hat{E} C}{E C} = 1 - \frac{\sum_{k=1}^K \sum_{i=1}^I \hat{\alpha}_k^i \ln \left(\frac{\hat{\alpha}_k^i}{\alpha_k^i} \right)}{\sum_{k=1}^K \sum_{i=1}^I \alpha_k^i \ln \left(\frac{\alpha_k^i}{\alpha_k^i} \right)}$$

em que $\hat{E} C$ é a entropia cruzada estimada e $E C$ é a entropia cruzada observada. A $\hat{E} C$ mede a entropia entre a distribuição estimada e a distribuição agregada. Quando o valor da $\hat{E} C$ é zero, a distância é nula e não existe heterogeneidade nos dados. A $E C$ mede a entropia cruzada dos coeficientes da informação observada no nível desagregado e no nível agregado, dando a heterogeneidade de informação observada quando desagregada.

Segundo Howitt & Reynaud (2003), o GID deverá aumentar à medida que se ganha informação com a desagregação. O ganho é nulo quando a desagregação é uniforme e sem heterogeneidade. Quando o GID é 1, atinge o seu valor máximo, e indica que toda a informação foi recuperada.

No que diz respeito à validação e avaliação da utilidade prática do modelo, optou-se por realizar a comparação das estimativas com os dados observados, através do cálculo do desvio absoluto médio (DAM) por factor de produção ou item de custo (i) e por actividade agrícola (k):

$$DAM_k^i = \left| \frac{\alpha_k^i - \hat{\alpha}_k^i}{\alpha_k^i} \right| \times 100$$

com α_k^i representando os dados observados que são comparados com os resultados estimados.

4. Implementação empírica do modelo

Para tentar responder ao problema da falta de informação relativa aos custos dos factores por actividade agrícola, desenvolveu-se um modelo de desagregação dos custos de factores de produção variáveis por actividade com base na teoria da máxima entropia. Este modelo é aplicado no contexto da região agrária do Alentejo, segundo os dados da RICA de 2008.

Uma análise preliminar dos dados ajuda a organizá-los de acordo com as categorias de custo específico e de rendimento, seleccionar explorações, actividades e custos a estudar.

Do ficheiro de resultados da RICA consta um extenso número de actividades, para cada uma das regiões de Portugal, tendo-se elegido o Alentejo como região de estudo, pela dimensão física elevada das explorações agrícolas e por existir também alguma especialização das actividades produzidas.

TABELA 1
N.º de explorações por classes de OTE e MBT no Alentejo

OTE	MBPT (€)	%	nº explorações	%
Culturas Arvenses	2312814	22	40	16
Horticultura Intensiva	305837	3	10	4
Culturas Permanentes	1077999	10	36	15
Herbívoros	4519264	43	115	47
Granívoros	279879	3	2	1
Policultura	372920	4	7	3
Polipecuária	533736	5	9	4
Mistas Culturas e Pecuária	1145318	11	28	11
Total	10547768	100	247	100

Fonte: RICA, 2010

A Tabela 1 apresenta a distribuição de explorações por classes de OTE e Margem Bruta Total (MBT), na amostra RICA de 2008 para a região Alentejo. Da sua análise, verifica-se que as OTEs que agregam mais explorações agrícolas são a OTE dos Herbívoros (47%), a OTE das Culturas Arvenses (16%) e a OTE das Culturas Permanentes (15%), representando 78% do total de explorações da amostra para o Alentejo e contribuem com 75% para a MBPT da região.

Cada uma dessas OTE apresenta um número de explorações que permite tratar cada uma delas como uma amostra autónoma, o que permite obter resultados por OTE.

Foi extraída uma amostra de conveniência de explorações agrícolas da base RICA 2008, por OTE. No sentido de tornar essa amostra mais equilibrada em termos de culturas, áreas e produções, desenvolveram-se alguns critérios, que passam também por apostar principalmente nas culturas com alguma área produtiva (culturas que sejam comercialmente activas) e nas explorações com alguma dimensão (para o Eurostat as empresas têm de ter mais de 5 ha de superfície agrícola utilizada). Também se analisaram e eliminaram as explorações que apresentavam problemas de ordem técnica como: a) as explorações que se encontravam a fazer implantação de culturas permanentes (em período

de carência de rendimentos), b) as empresas que apresentavam coeficientes económicos negativos, como as margens brutas, c) as explorações que tinham tido problemas de sanidade animal e consequente vazio sanitário e d) as empresas com uma especificidade determinante no sistema de produção que as tornavam diferentes, como a produção de cavalos (equídeos), com grandes e cuidadosos regimes alimentares e os suínos e as aves com uma vertente industrial.

De acordo com a metodologia usada para calcular as margens brutas padrão, os itens de custo considerados para as actividades vegetais foram sementes e plantas, fertilizantes, fitofármacos, outros encargos com as culturas. Para as actividades animais esses itens foram rações compradas, alimentos forrageiros comprados e produzidos na exploração e assistência veterinária e outros encargos com a pecuária.

O valor da margem bruta é definido em euros por actividade (vegetal ou animal). A unidade usada para as culturas é o hectare e para os animais é cabeça normal (CN).

Para validar o modelo de desagregação de custos por actividade é necessário dispor dos coeficientes dos custos observados e desagregados por factor de produção e produto. Para esse efeito utilizou-se a estrutura de custos de referência das actividades das MBP de 2004.

Nas culturas forrageiras e pastagens, a área utilizada é a que consta da ficha de exploração, uma vez que esta área não consta do Fresult. Também as produções destas actividades, sendo auto-utilizadas não apresentam valor na base de dados, pelo que o valor da sua produção teve de ser estimado através do Valor de Produção Padrão (VPP) 2008.

As 29 empresas seleccionadas na OTE 1- Culturas Arvenses têm uma área média de 97 ha de superfície agrícola, com um mínimo de 31 ha e um máximo de 309 ha de área total, mostrando uma grande variação na sua dimensão física. As actividades que apresentam valores de receita mais elevados são as culturas horto-industriais, milho grão e milho forragem. As produções médias alcançadas pelas empresas (€ de receita/ha produzidos apresentam um intervalo relativamente grande em quase todas as actividades. As actividades que mais contribuem para receita global da OTE, são as horto-industriais (44%). Quando se analisam os encargos por hectare também se verifica que os gastos são muito diferentes, sendo os intervalos com maior amplitude, os apresentados pelos fitofármacos e pelos fertilizantes. A margem bruta também apresenta alguma variabilidade mas bem menor que os encargos.

As empresas seleccionadas na OTE 3- Culturas Permanentes apresentam uma área média de 33 ha, com mínimo de 6 ha e um máximo de 149 ha. As actividades que apresentam maior receita média por ha são a fruticultura seguida de viticultura. O peso da contribuição de cada actividade no rendimento da OTE 3 é de 48% por parte da viticultura e 40% da olivicultura. No que se refere à área, a olivicultura é dominante (67%), seguida da vinha (20%) e o trigo com 8%. De notar que a fruticultura ocupa apenas 2% da área, no entanto apresenta os melhores resultados médios por ha (6076€), seguida da viticultura (3447€). Quando se observam os encargos, verifica-se que as variações são grandes nos outros encargos e fitofármacos. A margem bruta apresenta a menor das variações mínimas.

As empresas orientadas para herbívoros com actividade agrícola (OTE 4) apresentam áreas diversas desde 60 ha a 413 ha, sendo a média de 163 ha. A nível das actividades, as maiores receitas médias por ha são provenientes de milho forrageiro, trigo e pastagens temporárias.

No entanto, a maior contribuição em termos de receita para a OTE é dada pela actividade de prados temporários com 33% do valor, seguido de outras forragens e milho forrageiro com 21% e prados permanentes com 16%, num total de 91%, mostrando claramente a prioridade da empresa em fornecer alimentação aos seus animais.

A maior área média está concentrada nas pastagens permanentes que condiciona cerca de 52% da área total envolvida e as outras culturas forrageiras apresentam um peso de 21%.

As empresas orientadas para herbívoros e com actividade vegetal também apresentam flutuações bastante elevadas nos encargos como sucede com os outros encargos e os fitofármacos, variações muito elevadas.

As empresas orientadas para herbívoros (animal) apresentam quantidades de cabeças normais diversas desde 6 CN a 269 CN, com uma média de 72 CN.

A nível das actividades, a maior receita média por CN é proveniente da actividade de bovinos de leite, sendo os bovinos de carne a actividade que mais contribui para as receitas da OTE, com cerca de 46%. São estes os que apresentam o maior efectivo médio (43CN), seguido dos ovinos com 25CN. Por outro lado, essas mesmas espécies contribuem para com 64% e 34% do peso do efectivo da OTE. Os encargos por CN apresentam variações muito grandes.

5. Resultados

Neste ponto apresentam-se os resultados dos modelos de Tobit-MEG e Tobit-ECG aplicados a quatro amostras de explorações agrícolas da RICA no Alentejo para o ano de 2008.

De acordo com Rosário (2012), o modelo Tobit-ECG foi utilizado com o objectivo de melhorar as estimativas do modelo Tobit-MEG, fazendo uso de priores de informação que não estão contidos nas amostras objecto de estudo. Neste caso utilizou-se como prior de informação, $q_{i,k}$, a estrutura de custos utilizada pelo GPP do Ministério da Agricultura em Portugal para calcular as margens brutas padrão do ano 2004 por actividade agrícola.

Deste modo, os resultados dos modelos de Tobit-MEG e Tobit-ECG são apresentados para cada uma das amostras de explorações agrícolas consideradas e são analisados, primeiro em termos dos indicadores de entropia normalizada e de capacidade de prescrição (pseudo- R^2), depois em termos dos ganhos de informação obtidos no processo de desagregação e por último em termos da sua validade prática através da comparação entre os resultados estimados e os dados observados para os coeficientes de afectação dos custos dos factores por actividade agrícola, através do cálculo do DAM.

Na Tabela 2 apresentam-se os resultados dos modelos MEG e ECG relativos aos indicadores da entropia normalizada e do pseudo- R^2 das explorações agrícolas por OTE.

Para a OTE das culturas arvenses, o modelo de MEG apresenta para os parâmetros estimados um valor de entropia normalizada $S(\hat{p})$ de 0,708. Este valor é relativamente próximo da unidade, o que denota uma percentagem razoável de criação de nova informação no processo de desagregação de custos. No modelo de ECG, como seria de esperar, o valor da entropia normalizada $S(\hat{p})$ é menor, i.e., mais perto de zero. No que diz respeito à componente do erro, os valores obtidos para entropia normalizada $S(\hat{w})$ são em ambos os casos muito próximos da unidade (0,981 e 0,991).

TABELA 2
Resultados de Entropia normalizada e pseudo- R^2 para os modelos de MEG e ECG das explorações agrícolas por OTE

	OTE CULTURAS ARVENSES		OTE CULTURAS PERMANENTES		OTE HERBÍVOROS-VEGETAL		OTE HERBÍVOROS - ANIMAL	
	MEG	ECG	MEG	ECG	MEG	ECG	MEG	ECG
Entropia Normalizada	0,708	0,307	0,450	0,253	0,489	0,198	0,755	0,403
Entropia Normalizada do erro	0,981	0,991	0,974	0,981	0,975	0,980	0,992	0,993
Pseudo R^2								
Sementes e Plantas	0,867	0,928	0,636	0,643	0,501	0,479	-	-
Fertilizantes	0,822	0,864	0,845	0,934	0,590	0,582	-	-
Pesticidas	0,753	0,803	0,893	0,920	0,229	0,357	-	-
Outros Custos	0,773	0,800	0,293	0,337	0,231	0,239	0,654	0,654
Margem Bruta	0,916	0,968	0,995	0,997	0,987	0,983	0,895	0,920
Concentrados	-	-	-	-	-	-	0,940	0,931
FORAGENS	-	-	-	-	-	-	0,801	0,784

Os valores obtidos para o pseudo- R^2 , como referido anteriormente, mostra a capacidade explicativa do modelo. No MEG, os custos com sementes e plantas e com fertilizantes apresentam valores de 0,867 e

0,822, enquanto os gastos com os pesticidas e outros custos registam valores de 0,753 e 0,773. Todos esses valores, assim como o valor obtido para a margem bruta (0,916) estão perto da unidade, o que evidencia a existência de um forte poder explicativo que permite relacionar a distribuição dos custos dos factores de produção com a distribuição do produto bruto nas actividades das explorações agrícolas da OTE das culturas arvenses. No modelo da ECG, esse poder explicativo é ainda maior, uma vez que os valores do pseudo- R^2 são ainda superiores.

No que diz respeito às culturas permanentes, o modelo MEG apresenta uma entropia normalizada de 0,450, com uma criação de informação mais contida e para o vector de suporte do erro, $S(\hat{p})$, o valor de 0,974, mais perto da unidade.

Nos valores obtidos pelo pseudo- R^2 , os custos com os pesticidas e os fertilizantes apresentam os valores mais próximos da unidade, com 0,893 e 0,845 respectivamente, enquanto os encargos com as sementes e plantas são 0,636 e outros custos apresentam valores de 0,293. Aquele grupo está bem mais perto da unidade. O item que apresenta valores mais próximos da unidade é a Margem Bruta (0,977).

No modelo da ECG, a entropia normalizada de $S(\hat{p})$ é mais pequena (0,253), portanto mais perto do zero, o que mostra que a incerteza de informação é bem pequena. A entropia normalizada do erro ruído também está muito perto da unidade (0,981).

Os valores do pseudo- R^2 dos custos de fertilizantes e pesticidas no modelo ECG passam para 0,934 e 0,920, aproximando-se mais da unidade. No caso dos custos com as sementes e plantas (0,643), os valores do pseudo- R^2 crescem muito pouco, apenas 0,007, e nos outros custos passam para 0,337. A margem bruta está com 0,997, praticamente na unidade.

O grupo Herbívoro na sua componente de custos agrícolas para actividades vegetais, apresentam para o modelo MEG uma entropia normalizada para o vector $S(\hat{p})$ de 0,489 e para o erro, $S(\hat{w})$ o de 0,975. Nos valores obtidos para o pseudo- R^2 , e no mesmo modelo de MEG, os custos com as sementes e plantas e fertilizantes estão com 0,501 e 0,590, enquanto os pesticidas e outros custos apresentam valores de 0,229 e 0,231. A Margem Bruta atinge o valor de 0,987 que, quando comparado com a diversidade das outras variáveis, é bem melhor.

No modelo de ECG, a entropia normalizada dos coeficientes de probabilidade estimados $S(\hat{p})$ é mais pequena (0,198), mais perto do zero e a entropia normalizada e do ruído está muito perto da unidade (0,980). Nos valores do pseudo- R^2 , os custos com os fertilizantes e as sementes e plantas passam para 0,582 e 0,479, mostrando uma pequena diminuição. Os outros custos crescem muito pouco, apenas 0,008, os pesticidas atingem 0,337 (crescem 0,218) e a margem bruta fica em 0,983.

O grupo Herbívoro na sua componente de custos direccionados para animais, apresenta no modelo de MEG uma entropia normalizada para as probabilidades estimadas $S(\hat{p})$ de 0,755 e para o erro, $S(\hat{w})$, o valor de 0,992. Os valores obtidos para o pseudo- R^2 , no MEG, para os custos com concentrados e com as forragens são 0,940 e 0,801, enquanto os outros custos apresentam valores de 0,654. A Margem Bruta atinge valores de 0,895, relativamente abaixo do valor atingido pelo concentrado.

Quando se passa para o modelo de ECG, há uma diminuição dos valores na entropia normalizada $S(\hat{p})$ e o valor da entropia normalizada do erro $S(\hat{w})$ está ainda mais perto da unidade, mantendo-se as posições relativas das diversas actividades. Já a margem bruta apresenta um valor superior, 0,920.

Nos ganhos de informação desagregada total com a utilização dos modelos (Tabela3), para as culturas arvenses, o indicador é de 0,871 para o MEG, que pode ser considerado de bom, mas ainda é melhor no ECG, com 0,999.

O GID parcial evidencia na generalidade das actividades envolvidas, os ganhos que ocorrem parcialmente e como influenciam o GID.

No modelo de MEG, são visíveis 2 grupos com valores menos elevados mas mesmo assim bons: as proteaginosas, a olivicultura e a pastagem permanente com valores entre 0,82 a 0,84, e depois o trigo e as oleaginosas com 0,87.

TABELA 3
Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE
culturas arvenses

	MEG	ECG
Ganhos de informação desagregada total	0,871	0,999
Ganhos de informação desagregada parcial		
Trigo	0,87	1,00
Milho	0,99	0,99
Arroz	0,98	0,99
O. cereais	0,98	1,00
Proteaginosas	0,82	1,00
Horto industriais	0,99	0,99
Oleaginosas	0,87	1,00
Olivicultura	0,83	1,00
Pastagens Permanentes	0,84	1,00
Pastagens Temporárias	0,90	1,00
Milho forrageiro	0,91	1,00
Outras forragens	0,92	1,00

No modelo de ECG, os ganhos de informação são quase totais e os que apresentam no MEG um valor menor acabam elevando esse valor quando se muda de modelo. O milho, o arroz e as horto-industriais ficam com o valor de 0,99.

No caso do modelo de ECG, os ganhos de informação são superiores aos obtidos para o modelo de MEG e são quase totais, uma vez que o valor do indicador é na maior parte dos casos igual à unidade.

TABELA 4
Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE
culturas permanentes

	MEG	ECG
Ganhos de informação desagregada total	0,977	0,994
Ganhos de informação desagregada parcial		
Trigo	0,98	0,98
Outros cereais	0,88	0,90
Fruticultura	0,98	0,98
Olivicultura	0,99	1,00
Viticultura	0,99	0,99

Para as culturas permanentes, os ganhos de informação total (Tabela 4), são bons, o indicador é de 0,977, mas ainda é melhor no modelo de ECG com 0,994. Quando se analisa os GID parciais, apenas os outros cereais tem um comportamento algo diferente.

Nos ganhos de informação desagregada total (Tabela 5), para os herbívoros na componente vegetal, o indicador é de 0,924 mas é superado pelo valor dado pelo modelo de ECG (0,992). As pastagens pobres e a olivicultura com ganhos relativamente menores no modelo de MEG, de 0,82 e 0,90, chegam a 1,0 no modelo de ECG. As restantes actividades ou estão com 1,0 ou nas redondezas de 0,98 e 0,99.

TABELA 5
Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE herbívoros componente vegetal

	MEG	ECG
Ganhos de informação desagregada total	0,924	0,992
Ganhos de informação desagregada parcial		
Trigo	0,97	0,99
O. cereais	0,97	0,98
Olivicultura	0,90	1,00
Pastagens Permanentes	0,95	0,98
Pastagens Pobres	0,82	1,00
Pastagens Temporária	1,00	0,99
Milho forragem	0,99	0,99
Outras forragens	0,99	1,00

Ao analisar os herbívoros na sua componente animal e os ganhos de informação desagregada total (Tabela 6), com o modelo de ECG, os valores são muito superiores quando comparados com os do modelo MEG. Nos ganhos de informação parcial os valores relativos aos ovinos e caprinos são, no modelo de MEG, 0,45 e 0,78, respectivamente, e atingem 1,00 no modelo de ECG.

TABELA 6
Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE herbívoros componente animal

	MEG	ECG
Ganhos de informação desagregada total	0,814	0,995
Ganhos de informação desagregada parcial		
Bov. carne	0,96	0,99
Bov. leite	0,96	0,99
Ovinos	0,75	1,00
Caprinos	0,78	1,00

Nas Tabelas seguintes apresentam-se os resultados das estimativas dos coeficientes de afectação dos custos dos factores de produção às actividades agrícolas obtidos com os modelos de MEG e de ECG, e procede-se à sua comparação com os dados observados através do cálculo do Desvio Absoluto Médio (DAM).

No caso do modelo de MEG das culturas arvenses, cujos resultados são apresentados no Tabela 7, verifica-se uma coerência generalizada dos valores estimados. No caso dos itens de custo relativos às sementes e plantas, o arroz, as proteaginosas e outras forragens apresentam valores de desvios elevados. Para os fertilizantes, têm-se o arroz, as horto-industriais, a olivicultura e ainda as oleaginosas com um valor bem elevado. Os pesticidas têm desvios elevados nos outros cereais e olivicultura. Os outros custos têm desvios consideráveis nas diversas actividades, talvez devido a ser um grupo de custos muito diversificado e ter também adesões diferentes. As margens brutas apresentam desvios entre 17% nas horto-industriais e 77% nas proteaginosas. Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (44%), nos fertilizantes (45%), nos pesticidas (67%), nos outros custos (25%) e margens brutas (33%).

A análise descritiva dos DAM no modelo de MEG para as culturas arvenses, mostra que existem medianas que se comportam de uma forma aceitável, uma variação muito elevada entre os máximos e os mínimos originando uma média mais elevada, consequentemente um desvio padrão e coeficiente de variação também elevados.

TABELA 7
Coefficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE culturas arvenses

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,177	60,9	0,206	17,6	0,154	18,9	0,185	0,0	0,279	36,6
Milho	0,162	32,5	0,236	24,2	0,139	39,0	0,184	67,3	0,278	20,6
Arroz	0,190	216,7	0,210	133,3	0,237	5,2	0,131	0,8	0,232	51,7
O. cereais	0,153	17,7	0,225	2,2	0,143	615,0	0,180	100,0	0,299	44,6
Proteaginosas	0,198	120,0	0,200	0,0	0,195	0,0	0,201	0,0	0,206	77,4
Horto indust.	0,078	13,3	0,152	204,0	0,086	4,4	0,064	220,0	0,621	17,2
Oleaginosas	0,164	2,5	0,231	670,0	0,175	0,0	0,154	413,3	0,275	64,7
Olivicultura	0,196	0,0	0,204	308,0	0,196	553,3	0,193	0,0	0,212	76,7
P.Permanente	0,200	0,0	0,200	81,8	0,190	0,0	0,198	120,0	0,213	73,4
P.Temporária	0,127	0,0	0,199	5,2	0,119	0,0	0,154	3,8	0,401	36,3
M. forrageiro	0,124	55,0	0,211	75,8	0,130	0,0	0,135	237,5	0,399	47,5
Out. forragens	0,196	292,0	0,202	34,7	0,195	0,0	0,197	0,0	0,210	72,7

Na observação dos desvios no ECG para as culturas arvenses (Tabela 8), percebe-se um melhor relacionamento dos coeficientes de afectação e os desvios com os valores reais, estão relativamente próximos. As situações mais elevadas são os fertilizantes e outros custos nas horto-industriais, devido a esta actividade ser um agregado de culturas, os pesticidas no milho (que agrega o milho grão, quer de sequeiro quer de regadio, híbrido e regional), outros cereais e outras forragens, que também são grandes agregados de culturas com comportamentos diferentes em produção.

TABELA 8
Coefficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE culturas arvenses

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,113	28,0	0,23	4,4	0,163	37,0	0,000	0,0	0,494	3,8
Milho	0,166	42,5	0,255	72,1	0,149	203,0	0,180	63,6	0,249	17,9
Arroz	0,082	11,9	0,116	58,3	0,267	93,1	0,074	43,1	0,461	2,6
O. cereais	0,100	36,4	0,238	32,1	0,029	180,1	0,140	55,6	0,494	2,8
Proteaginosas	0,091	28,2	0,000	0,0	0,000	0,0	0,000	0,0	0,909	4,1
Horto indust.	0,082	31,8	0,106	137,4	0,083	58,1	0,053	165,0	0,676	7,9
Oleaginosas	0,125	39,9	0,038	66,7	0,000	0,0	0,033	10,0	0,803	13,2
Olivicultura	0,000	0,0	0,051	6,3	0,031	61,6	0,000	0,0	0,918	0,6
P. Permanente	0,000	0,0	0,104	19,8	0,000	0,0	0,084	6,7	0,811	11,8
P.Temporária	0,000	0,0	0,121	24,6	0,000	0,0	0,099	38,1	0,780	45,5
M. forrageiro	0,046	56,8	0,089	9,9	0,000	0,0	0,029	27,5	0,836	17,2
Out. forragens	0,049	22,0	0,149	15,3	0,01	119,8	0,020	0,0	0,772	2,8

Fonte: Resultados do Modelo

Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (78%), nos fertilizantes (64%), nos pesticidas (14%), nos outros custos (50%) e margens brutas (92%).

A análise descritiva do DAM das culturas arvenses em ECG apresenta medianas aceitáveis, com desvios reduzidos; a variação de máximos e mínimas são bem menores, os desvios da média estão relativamente perto do valor da mediana e conseqüentemente os desvios padrão e coeficientes de variação são bem menores.

No caso das explorações da OTE Culturas Permanentes, quando se observa os coeficientes de afectação e o DAM (Tabela 9) no modelo de MEG, apenas os pesticidas nos outros cereais apresentam um valor muito elevado. Por outro lado, as margens brutas também apresentam desvios elevados para o trigo e outros cereais. Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nos fertilizantes (60%), nos pesticidas (50%) e margens brutas (20%).

TABELA 9
Coefficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE culturas permanentes

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,143	58,7	0,301	1,1	0,241	19,6	0,017	0,0	0,297	154,7
Out. cereais	0,062	74,6	0,232	45,0	0,181	1370,1	0,047	90,6	0,478	478,4
Fruticultura	0,001	0,0	0,118	15,8	0,050	0,0	0,001	99,6	0,831	97,1
Olivicultura	0,003	0,0	0,114	32,5	0,070	0,0	0,007	93,7	0,806	28,0
Viticultura	0,010	0,0	0,037	77,7	0,083	0,0	0,016	59,3	0,854	58,6

No caso das culturas permanentes, a análise descritiva do DAM no modelo MEG, revela que a mediana e a média têm valores algo diferentes, sendo nos outros custos a mediana superior à média. Os desvios máximos são de um modo geral inferiores aos que se obtiveram na OTE Culturas Arvenses, com excepção dos pesticidas que também apresentam desvios e coeficientes de variação elevados.

Ao se mudar para o modelo de ECG das culturas permanentes (Tabela 10), os coeficientes de afectação e o DAM dos fertilizantes na olivicultura têm um desvio considerável e de novo os pesticidas nos outros cereais apresentam um valor muito elevado seguido do trigo. Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (50%), nos fertilizantes (20%), nos outros custos (100%) e margens brutas (80%).

TABELA 10
Coefficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE culturas permanentes

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,128	18,5	0,065	73,0	0,311	161,4	0,000	0,0	0,496	4,2
Out. cereais	0,040	74,6	0,131	27,3	0,126	1117,2	0,026	1,0	0,677	40,9
Fruticultura	0,000	0,0	0,101	85,2	0,000	0,0	0,000	2,0	0,898	6,0
Olivicultura	0,001	0,0	0,121	152,3	0,054	0,0	0,007	3,0	0,817	11,5
Viticultura	0,008	0,0	0,028	46,8	0,064	0,0	0,016	4,0	0,884	0,1

A análise descritiva para as culturas permanentes no modelo ECG mostra que a mediana é muito aceitável, excepção para os fertilizantes e os pesticidas que apresentam máximos e médias muito grandes e têm desvios padrão e coeficientes de variação elevados.

Nos coeficientes de afectação e desvios do DAM nos herbívoros na sua componente vegetal (Tabela 11), no modelo de MEG, os fertilizantes apresentam desvios grandes e distribuídos pelas diversas actividades, assim como os outros custos nas actividades de outros cereais, pastagens permanentes e milho forragem e os pesticidas, nos outros cereais e olivicultura

TABELA 11
Coefficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE herbívoros-componente vegetal

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,103	58,5	0,333	311,1	0,095	32,6	0,029	0,0	0,439	38,5
O. cereais	0,103	37,3	0,307	338,6	0,062	342,9	0,022	266,7	0,505	39,5
Olivicultura	0,142	0,0	0,293	2563,6	0,151	843,8	0,076	0,0	0,339	65,2
P. Permanente	0,070	0,0	0,222	753,8	0,046	0,0	0,012	140,0	0,651	32,8
Pastagens Pobres	0,203	0,0	0,209	0,0	0,200	0,0	0,176	0,0	0,211	78,9
P. Temporária	0,002	0,0	0,047	24,2	0,003	0,0	0,001	94,1	0,947	2,2
Mil. Forragem	0,022	42,1	0,070	141,4	0,018	0,0	0,009	350,0	0,880	5,5
Out. forragens	0,055	161,9	0,144	300,0	0,014	250,0	0,017	1600,0	0,771	17,7

Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (25%), nos fertilizantes (14%), nos pesticidas (25%) e margens brutas (75%).

Na análise descritiva do DAM no modelo de MEG para as culturas da OTE herbívoros, os valores da mediana flutuam dentro de limites aceitáveis a pouco aceitáveis, donde para este grupo de actividades os coeficientes de variação são muito elevados.

No caso do modelo de ECG (Tabela 12), na OTE herbívoros na sua componente vegetal, os valores estão dentro de limites aceitáveis. Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (25%), nos fertilizantes (57%), nos pesticidas (50%), nos outros custos (14%) e margens brutas (88%).

TABELA 12
Coefficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE herbívoros-componente vegetal

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,068	56,7	0,265	10,2	0,058	51,3	0,004	47,3	0,605	27,1
O. cereais	0,082	47,9	0,226	25,5	0,009	13,1	0,014	91,9	0,669	39,3
Olivicultura	0,000	0,0	0,050	4,3	0,015	21,8	0,006	37,5	0,930	0,7
P. Permanente	0,028	0,0	0,146	68,1	0,019	0,0	0,010	94,7	0,796	9,7
Pastagens Pobres	0,000	0,0	0,000	0,0	0,000	0,0	0,000	0,0	1,000	0,0
P. Temporária	0,000	0,0	0,022	86,3	0,000	0,0	0,000	99,9	0,978	82,4
Mil. Forragem	0,022	79,4	0,051	48,4	0,017	0,0	0,009	88,9	0,901	26,3
O. Forragens	0,049	22,0	0,088	31,9	0,001	87,3	0,017	67,8	0,845	12,6

A análise descritiva do DAM para o modelo de ECG nos herbívoros na sua componente vegetal, mostra valores aceitáveis da mediana. Como os máximos não tem desvios muito elevados, a média está relativamente perto da mediana, os desvios padrão são relativamente baixos. No entanto, os coeficientes de variação são relativamente mais elevados.

Ao se analisar os coeficientes e os desvios no modelo de MEG, para a componente animal da OTE Herbívoros, os bovinos apresentam desvios acentuados em todos os tipos de custos e os caprinos apenas nas forragens (Tabela 13). Os desvios dos diversos custos com valores aceitáveis abaixo de 40% encontram-se apenas no item relativo a outros custos (25%).

TABELA 13
Coefficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE herbívoros-componente animal

	Concentrados		Forragens		Outros custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM
Bov. de carne	0,332	149,6	0,062	264,7	0,204	264,3	0,402	49,4
Bov. de leite	0,411	152,1	0,056	460,0	0,138	160,4	0,395	49,0
Ovinos	0,328	0,0	0,054	68,8	0,214	22,3	0,403	49,2
Caprinos	0,274	0,0	0,176	309,3	0,281	72,4	0,270	66,0

Fonte: Resultados do Modelo

A análise descritiva em MEG para a OTE Herbívoros na sua componente animal, salienta medianas e médias muito próximas e os desvios padrão algo elevados assim como os coeficientes de variação relativamente elevados, com excepção da margem bruta.

Quando se passa para o modelo de ECG, os valores encontrados estão mais normalizados, só os outros custos continuam elevados para os bovinos (Tabela 14). Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nos custos relativos a concentrados (50%), às forragens (25%), aos outros custos (50%) e às margens brutas (100%).

TABELA 14

Coefficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE herbívoros-componente animal

	Concentrados		Forragens		Outros custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM
Bov. de carne	0,254	16,0	0,046	55,9	0,124	220,9	0,577	9,6
Bov. de leite	0,403	48,9	0,052	17,7	0,116	211,5	0,429	31,8
Ovinos	0,000	0,0	0,034	83,0	0,163	27,1	0,802	19,4
Caprinos	0,000	0,0	0,146	42,5	0,156	39,3	0,698	10,1

A análise descritiva do DAM no modelo de ECG, mostra que também neste caso a média e a mediana estão muito próximas e o máximo atinge um valor elevado nos outros custos, e consequentemente se traduz num aumento do desvio padrão. O coeficiente de variação apresenta também uma grande oscilação de valores.

De um modo geral, os valores obtidos com a intervenção de indicadores de qualidade, precisão e estimação têm comportamentos muito diferentes, apesar de, em alguns casos serem muito semelhantes aos que são apresentados em diversos trabalhos como Fragoso & Carvalho (2011), Peeters & Surry (2002), nomeadamente, para as entropias normalizadas e para o pseudo-R².

Também os valores apresentados pelos GID são relativamente semelhantes aos apresentados por Fragoso & Carvalho (2011), Xavier et al. (2010) e Howitt & Reynaud (2003).

Os valores referentes aos coeficientes de afectação e aos DAM também são relativamente semelhantes aos apresentados por Fragoso & Carvalho (2011).

As principais dificuldades sentidas foram tal como referem Peeters & Surry (2002), problemas com as categorias heterogéneas como as outras oleaginosas e outros cereais onde se encontram grupos de culturas. Por outro lado, as variações ou melhor a heterogeneidade entre as empresas e, nomeadamente, a variabilidade entre os custos de factores também levaram a que existisse alguma disparidade entre os valores estimados e os valores reais observados, como aconteceu com Fragoso & Carvalho (2011).

Não obstante, os mesmos autores avançam com alguns valores para analisar o DAM na calibração de modelos, considerando que o modelo se encontra bem validado para valores inferiores a 15%. No entanto, um DAM superior a 15% pode ser considerado razoável, mas mostra que o modelo precisa de ser melhorado. Howitt & Msangi (2006) consideram aceitáveis valores de DAM entre 1.5 % e 40.9%.

6. Conclusões

A falta de informação desagregada relativamente a factores de produção é um problema generalizado na área das ciências agrárias. Para se ter esse detalhe recorre-se à utilização de inquéritos directos e pessoais junto do agricultor, que são dispendiosos e demorados.

O uso de técnicas e métodos alternativos para estimar os coeficientes de afectação dos custos às actividades agrícolas, a partir de informação incompleta, pode ser uma experiência válida para estimar os custos variáveis de produção unitários da exploração a um custo significativamente mais baixo.

A base de dados da RICA é uma fonte de informação muito útil para estudar diversos aspectos da política agrícola. A informação sobre as explorações constantes das bases de dados da RICA respeita ao total dos custos por tipo factores, não sendo desagregada por actividade. Para além disso, apesar da crescente viragem para a especialização, a maioria das explorações agrícolas são ainda multi-produto.

A obtenção dos custos variáveis unitários de produção por unidade de actividade é extremamente importante quer para a gestão da exploração quer na perspectiva da análise da política agrícola. Daí a necessidade do propósito deste estudo em desenvolver um modelo que permitisse a obter a partir da base de dados RICA, os encargos de produção desagregados por actividade. Para o efeito, utilizou-se uma metodologia baseada na teoria da Máxima Entropia, desenvolvendo-se dois modelos Tobit, um de Máxima Entropia Generalizada e outro de Mínima Entropia Cruzada, para estimar os coeficientes de afectação tendo em vista a desagregação do custo dos factores de produção ao nível da exploração

agrícola. Entre as principais vantagens da opção por este método está a forma como é possível ultrapassar dificuldades relacionadas com a informação incompleta, e ainda a sua flexibilidade e facilidade de manuseamento e também a existência de valores nulos na amostra.

Os dados utilizados foram os da base da RICA 2008 para região Alentejo, para explorações agrícolas com Orientação Técnico-Económica em culturas arvenses, culturas permanentes, herbívoros componente vegetal e herbívoros componente animal.

A qualidade do modelo para estimar os coeficientes de afectação dos custos de produção por actividade agrícola foi avaliada tendo em conta os critérios da entropia normalizada, do pseudo R², o ganho de informação desagregada e o desvio absoluto médio.

Os resultados apresentados mostram que se conseguiu desagregar a informação a partir dos modelos de máxima entropia generalizada e da entropia cruzada generalizada com a estimação de coeficientes de afectação por actividade, sendo este último modelo o que apresenta valores mais interessantes.

A aplicação de critérios de qualidade de precisão e de previsão mostram que os modelos aplicados dão boas indicações no sentido da desagregação dos factores e que se encontram num limiar muito perto de outros autores que se basearam nas mesmas questões do sector agrícola, com os mesmos métodos e modelos e com a utilização dos mesmos critérios. Os estimadores de entropia apresentam assim, de um modo , valores que se podem considerar bons tendo em conta as suas propriedades estatísticas e econométricas.

A comparação dos resultados dos modelos de Máxima Entropia Generalizada e Mínima Entropia Cruzada mostra de uma forma acentuada, a utilização da segunda variante do modelo conduz a uma melhoria dos resultados apresentados. Na origem dessa melhoria de resultados está o recurso a valores observados a priori e retirados das margens brutas padrão do Gabinete de Planeamento e Políticas do Ministério da Agricultura, o que indicia, tal como de um modo geral na literatura, que o modelo de Máxima Entropia Generalizada poderá dar melhores resultados se for utilizada informação exterior à amostra.

No entanto, há que salientar que nem todos os resultados podem ser considerados bons, nomeadamente com a análise feita com o DAP, há alguns valores menos interessantes que precisam ser melhorados. Mencionando Hazell & Norton (1986), a melhoria do modelo é obtida com testes de produção, testes de preço e informação da estrutura de modo a recolher sempre mais informação, incorporar as relevantes e dinamizar novas situações pois a criação de um modelo é um processo onde há sempre hipótese de introduzir novos elementos.

A melhoria do modelo pode requerer mais algum trabalho, mais algum esforço para solucionar os desvios percebidos: amostras mais alargadas, amostras com dimensão física semelhante, não utilização de critérios de OTE, não regionalização, com cruzamento de outra informação como painéis de opinião, líderes de opinião e valores provenientes de consulta bibliográfica.

Ao analisar-se detalhadamente as características das explorações verifica-se que há alguns pontos que eventualmente poderão ser melhorados. Na Orientação Técnico-Económica Culturas Arvenses, deveriam estar apenas culturas arvenses e não horto-industriais, culturas permanentes e culturas forrageiras anuais e permanentes, que são muito diferentes em área, em tecnologia/sistema de produção e rendimento e por conseguinte também em termos dos custos de produção. As áreas das explorações estão inseridas num intervalo muito grande, entre 31 e 309 ha. As receitas deste grupo têm uma amplitude muito grande, sendo as mais elevadas nas explorações de horto-industriais e as mais baixas nas culturas e prados permanentes. O produto bruto mostra as diferenças acentuadas que existem na mesma actividade, sobressaindo o grau de intensificação da agricultura e consequentemente o uso maior ou menor de factores de produção como sementes e plantas, fertilizantes, pesticidas e outros encargos, com reflexo directo na margem bruta. Nas outras duas classes de Orientação Técnico-Económica consideradas para o estudo, permanecem também as mesmas dificuldades que foram referidas para as Culturas Arvenses.

Com estas recomendações, considera-se que num futuro próximo, a aplicação das novas formas de abordagem podem fomentar o desenvolvimento de modelos com maior poder de previsão. Por outro lado, a metodologia da entropia apresentou-se como um instrumento de trabalho flexível e robusto para ajudar a compensar falhas de informação, como é o caso dos custos dos factores de produção desagregados por actividade agrícola, não sendo também a sua implementação difícil.

Referências bibliográficas

- Campbell, R., and Carter Hill, R. (2005). "A Monte Carlo study of the effect of design characteristics on the inequality restricted maximum entropy estimator", *Review of Applied Economics*, 1, pp.53-84.
- Campbell, R., and Carter Hill, R. (2006). "Imposing parameters inequality restrictions using the principle of maximum entropy", *Journal of Statistical Computation and Simulation*, 76, pp. 985-1000.
- Chakir, R. (2009). "Spatial downscaling of agricultural land use data: an economic approach using cross entropy", *Land Economics*, 85(2), pp. 238-251.
- Fragoso, R.; Martins, M.B. e Lucas, M.R. (2008). "Disaggregated soil allocation data using a Minimum Cross Entropy model", *WSEAS Transactions on Environment and Development*, 9 (4), pp. 756-766.
- Fragoso, R. e Carvalho, M.L. (2011). "Estimation of Farm Cost Allocation Coefficients at Farm Level Using an Entropy Approach", *CEFAGE-UE Working-Papers*, 2011/21.
- Fraser, I. (2000). "An application of maximum entropy estimation the demand for meat in the United Kingdom", *Applied Economics*, 32, 4pp. 5-59.
- Garvey, E.e Britz, W. (2002). "Estimation of Input Allocation from EU Farm Accounting Data using Generalized Maximum Entropy", *CAPRI.Working Paper 02-01*, U.Ireland & Bonn.
- Golan, A.; Judge, G. e Robinson, S. (1994). "Recovering information from incomplete or parcial multisectorial economic data", *Review of Economics and Statistics*, 76, pp. 541-551.
- Golan, A.; Judge, G. e Miller D. (1996a). *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, New York, John Wiley Editions.
- Golan, A.; Judge, G. e Perloff M. (1996b). "A Maximum Entropy Approach to Recovering Information From Multinomial Response Data". *Journal of the American Statistical Association*, Vol.91, n.434, pp. 841-853. Disponível em <http://www.jstor.org/stable/2291679>
- Golan, A. and Judge, M. (1996). "A maximum entropy approach to empirical likelihood estimation and inference", Working paper, University of California, Berkeley.
- Golan, A., Karp, S., Perloff, M. (1997). "Estimation and inference with censored and ordered multinomial response data", *Journal of Econometrics*, 73, pp. 23-52.
- Golan, A.; Perloff, M. e Shen, Z. (2001). "Estimating a demand system with the non-negativity constraints: Mexican meat demand". *Review of Economics and Statistics*, LXXXIII, pp. 541-551.
- Golan, A. and Perloff, J. (2002). 'Comparison of maximum entropy and higher-order entropy estimators', *Journal of Econometrics*, 107(1-2), pp. 195-211.
- Hansen, H. e Surry, Y. (2006). "Estimating the cost allocation for German agriculture: an application of the maximum entropy methodology", Conference paper, *46th Annual Conference of German Association of Agricultural Economists*, October 4-6.
- Harris, R. (2002). "Estimation of a Regionalized Mexican Social Accounting Matrix:using Entropy Techniques to Reconcile Disparate Data Sources", *TMD discussion paper n°97* WDC. USA.
- Hazell, P. e Norton, R. (1986). *Mathematical Programming for Economic Analysis in Agriculture*, NY, Mac Millan P.C. USA.
- Howitt, R. e Reynaud, A. (2003). "Spatial Disaggregation of Agricultural Production Data using Maximum Entropy", *European Review of Agriculture Economics*, Vol. 30, No. 3, pp.359-387.
- Howitt, R. e Msangi, S. (2006). "Estimating Disaggregate Production Functions: An Application to Northern Mexico", Draft paper.
- Just, R.; Zilberman, D. and Hochman, E. (1983). "Estimation of Multicrop Production Functions", *American Journal of Agricultural Economics*, 65 (November), pp.770-780.
- Kitamura, Y. e Stutzer, M. (1997). "An information-theoretic alternative to generalized method of moment estimation", *Econometrica*, vol. 66(4), pp. 861-874.

- Lence, H.L, and Miller, D. (1998a) "Estimation of Multi-Output Production Functions with Incomplete Data: A Generalized Cross Entropy Approach", *European Review of Agricultural Economics*, 25(December), pp. 188-209.
- Lence, H.L, and Miller, D. (1998b). "Recovering Output-Specific Inputs from Aggregated Input Data: A Generalized Cross Entropy Approach", *American Journal of Agricultural Economics*, 80(November), pp. 852-867.
- Leon, Y., Peeters, L., Quinqu, M. and Suury, Y. (1999). "The use of maximum entropy to estimate input-output coefficients from regional accounting data", *Journal of Agricultural Economics*, 50, pp. 425-439.
- Lips, M. (2009). "Full product costs on base of farm accountancy data by means of maximum entropy", Paper presented to the *International Association of Agricultural Economists Conference*, Beijing, China. August, 2009.
- Love, H. A. (1999). "Conflicts Between Theory and Practice in Production Economics", *Amer. J. Agr.Econ.* , 81, pp. 696-702.
- Martins, M.B.; Fragoso, R. e Xavier, A. (2011). "Spatial disaggregation of Agricultural Data: A maximum Entropy Approach", *JP Journal of Biostatistics*, Vol. 5,1, pp. 1-16.
- Moxey, A e Tiffin, R. (1994). "Estimating linear production coefficients from farm business survey data: A note", *Journal of Agricultural Economics*, 45, pp. 381-385.
- Papalia, R. (2010). "Generalized Maximum entropy Estimation of Spatial panel data interaction models", *WSEAS Transactions on Business and Economics*, 14, Vol 9, pp. 322-331.
- Paris, Q. e Howitt, R.E. (1998). "An analysis of ill-posed production problems using maximum entropy", *American Journal of Agricultural Economics*, n.80, pp. 124-138.
- Peeters, L. e Surry, Y. (2002). "Farm cost allocation based on the maximum entropy methodology- The case of Saskatchewan crops farms-Agriculture and Agri-Food Canada Strategic Policy Branch". Publication 2121/E, Project 02-011-tp, Ottawa, Ontario.
- Peeters, L. e Surry, Y. (2005). "Estimation d'un modèle à parameters variables par la méthode d'entropie croisée généralisée et application à la répartition des couts de production en agriculture", In : *Actes des Journées de Méthodologie Statistique*.
- Pires, C.; Dionísio, A. e Coelho, L. (2010). "GME versus OLS which is the best to estimate utility functions", *CEFAGE-UE Working-Papers*, 2010/02.
- Pukelsheim, F. (1994). The Three Sigms Rule. *American Statistician*, 48, pp. 88-91.
- Rosário, M.S. (2012). "Um modelo de desagregação de encargos de produção variáveis da base RICA por actividade", Dissertação de mestrado, Universidade de Évora, Évora.
- Rezek, J. e Campbell, R. (2006). "Cost estimates for multiple pollutants: A maximum entropy approach", *Energy Economics*, 29, pp. 503-519. Disponível em <http://www.sciencedirect.com>
- Van Delden, H. e Luja, P. (2006). "Integration of multi-scale dynamic spatial models for land use change analysis and assessment of land degradation and socio-economic processes", In : *Conference on Soil protection strategy - needs and approaches for policy support*, Poland.
- Xavier, A.; Martins, M.B. e Fragoso, R. (2010). "Combined disaggregation of agricultural land uses, livestock numbers and crops' production: an entropy approach", *Advances in Mathematical and Computacional methods*, pp 192-198.
- Zhang, X. e Fan, S. (2001). "Estimation Crop-Especific Production Technologies in Chinese Agriculture: a Generalized Maximum Entropy Approach", *American Journal of Agricultural Economics*, Vol.83, n.2, pp. 378-388. Disponível em <http://www.jstor.org/stable/1244680>