

Fundamentos Estatísticos da Análise-Diagnóstico de Sistemas Agrários:

uma interpretação baseada na Teoria
da Evidência de Dempster-Shafer

Benedito Silva Neto¹

Resumo

Neste texto procura-se demonstrar que a Teoria da Evidência de Dempster-Shafer permite fundamentar estatisticamente os procedimentos adotados na Análise-Diagnóstico de Sistemas Agrários. A partir de uma discussão das dificuldades para se constituir inferências estatísticas sobre o comportamento de sistemas complexos, seguida de uma apresentação da Análise-Diagnóstico de Sistemas Agrários, é realizada uma discussão da Teoria da Evidência de Dempster-Shafer enquanto uma fundamentação estatística dos procedimentos da Análise-Diagnóstico. Conclui-se que a Análise-Diagnóstico de Sistemas Agrários é compatível com a Teoria da Evidência de Dempster-Shafer e, portanto, com uma análise estatística rigorosa, especialmente quando se leva em conta a complexidade do desenvolvimento da agricultura.

Palavras chave: Complexidade. Sistemas agrários. Teoria da evidência.

¹ Professor do Departamento de Estudos Agrários e docente do Mestrado em Desenvolvimento da Unijui.
bsneto@unijui.edu.br

Abstract

In this work it's argued that the Dempster-Shafer Theory of Evidence permits a statistical foundation of the Analysis-Diagnosis of Agrarian Systems. From a discussion of the difficulties to make statistical inferences about behavior of complex systems, it's made a presentation of the Analysis-Diagnosis of Agrarian Systems, followed by a discussion of the Dempster-Shafer Theory of Evidence as statistical foundation of the Analysis-Diagnosis of Agrarian Systems. It is concluded that the Analysis-diagnosis of Agrarian Systems is compatible with Dempster-Shafer Theory of Evidence and, so, with a rigorous statistical analysis, especially when it takes in account the complexity of the agricultural development.

Keywords: Complexity. Agrarian systems. Theory of evidence.

A Análise-Diagnóstico de Sistemas Agrários (ADSA) consiste em um método de estudo da agricultura concebido originalmente para o estabelecimento de linhas estratégicas de desenvolvimento local relacionadas a este setor. Este método vem sendo desenvolvido há mais de 30 anos pela cátedra de Agricultura Comparada e Desenvolvimento, do Instituto das Indústrias dos Seres Vivos e do Ambiente – AgroParisTech (do qual faz parte o antigo Instituto Nacional Agrônômico de Paris-Grignon – INA-PG). Uma preocupação central no desenvolvimento deste método foi a sua aplicabilidade aos países do terceiro mundo, cujos complexos problemas agrícolas, normalmente associados a técnicas pouco usuais em relação aos padrões ocidentais e a grandes dificuldades de intervenção do poder público, tornam a elaboração de projetos de desenvolvimento uma tarefa extremamente difícil e complexa.

Embora a ADSA esteja fundamentada em uma sólida interpretação da evolução da agricultura, proposta por Mazoyer e Roudart (1997), sendo neste sentido discutida também por Dufumier (1996, 2004), a sua adequação para investigações científicas pode ser contestada, na medida em que seus procedimentos diferem substancialmente dos propostos pelos métodos usuais de pesquisa. Dentre estas diferenças uma das principais é o fato de a ADSA normalmente não incluir análises estatísticas.

Por outro lado, é possível que os procedimentos da ADSA se mostrem coerentes com os princípios estatísticos desde que estes sejam analisados a partir de teorias mais abrangentes do que a Teoria da Probabilidade, tal como a Teoria da Evidência de Dempster-Shafer (TE). Isto porque esta teoria permite que se trabalhe com representações holísticas do objeto de análise, além de reconhecer explicitamente a existência de ignorância, o que a torna particularmente interessante para o tratamento da incerteza relacionada aos processos de busca de conhecimento. Assim, a partir da Teoria da Evidência de Dempster-Shafer (TE), neste artigo procura-se demonstrar que os procedimentos adotados na ADSA são compatíveis com os princípios estatísticos que devem reger a análise

se de sistemas complexos, de forma inclusive muito mais rigorosa do que a permitida pela Teoria da Probabilidade. Para tanto o artigo foi organizado em quatro partes, além desta introdução e das conclusões. Na primeira parte é realizada uma breve discussão sobre as dificuldades da aplicação de inferências estatísticas no estudo de sistemas complexos; na segunda parte é feita uma discussão da TE, procurando-se mostrar sua adequação à análise de sistemas complexos e evolutivos. A terceira parte descreve sucintamente os princípios metodológicos e as principais etapas da ADSA. Na quarta parte é analisada a compatibilidade dos procedimentos adotados na ADSA com a TE.

Enfim, é importante salientar que a discussão dos fundamentos estatísticos da ADSA, principal foco deste artigo, tornou imprescindível a formalização matemática de alguns conceitos, especialmente nas duas primeiras partes. Tal formalização, no entanto, é acompanhada de descrições literárias ou da sua associação a noções intuitivas de forma que, mesmo que eventualmente algum leitor possa sentir certa dificuldade em compreender algumas passagens das duas primeiras partes, sua compreensão das partes posteriores não seja prejudicada.

Complexidade e Inferência Estatística

Talvez o melhor exemplo de sistema complexo, no sentido atribuído a este termo neste artigo, seja a própria biosfera terrestre (Gribbin, 2004, p. 219). Termodinamicamente a atmosfera terrestre possui uma composição altamente improvável, que só pode ser explicada pela presença de vida. Nas condições de pressão e temperatura vigentes, o gás oxigênio (O_2), que constitui 21% da atmosfera da Terra, é altamente reativo, e o gás carbônico, que é quimicamente pouco ativo, mantém-se em níveis relativamente baixos (0,03% da atmosfera aproximadamente). Isto só pode ser explicado pelo efeito da fotossíntese realizada por vegetais e algas, utilizando a energia solar. Ocorre que, se a composição

da atmosfera estivesse próxima do equilíbrio químico, no qual a concentração de gás oxigênio seria muito baixa e a de gás carbônico muito mais alta, a vida seria impossível (excetuando talvez algumas espécies de microorganismos). Uma alta concentração de gás carbônico tornaria a Terra tão quente (o preocupante efeito estufa) que impossibilitaria a sobrevivência de vegetais e animais. Já uma grande diminuição da presença do gás oxigênio impediria a utilização de energia pelos organismos por meio da respiração, também impossibilitando sua sobrevivência. Assim sendo, é a própria vida que é responsável por manter as condições ambientais que permitem a sua continuidade, o que caracteriza a biosfera como um sistema auto-sustentado (Gribbin, 2004, p. 219).

Nem sempre, porém, a biosfera terrestre foi assim. O comportamento apresentado pela biosfera atualmente é fruto de um longo processo evolutivo, no decorrer do qual várias espécies, sociedades, ecossistemas e biomas se sucederam, cada qual apresentando características específicas e imprevisíveis. Em outras palavras, a trajetória da biosfera no tempo é plena de “surpresas” e “novidades”, não apenas em relação aos seus componentes, mas também no que diz respeito às suas formas de organização.

A biosfera é um sistema altamente heterogêneo, apresentando uma grande diversidade de subsistemas, os quais são, também, sistemas complexos e evolutivos. Dentre estes sistemas destacam-se os ecossistemas e as sociedades.

Estatisticamente, sistemas² simples e sistemas complexos apresentam comportamentos muito distintos. Por exemplo, dado o número de combinações P de um sistema homogêneo com N componentes, dividido em dois compartimentos de mesmo tamanho, temos,

$$P(N_1 N_2) = \frac{N!}{N_1! N_2!} \quad (1)$$

² Referimo-nos aqui, como ao longo de todo este artigo, a sistemas dissipativos. Para uma discussão mais rigorosa ver Prigogine e Stengers (1986).

onde

P = número de combinações

N = número total de componentes

N_1 = número de componentes no compartimento 1

N_2 = número de componentes no compartimento 2

O estado mais provável de um sistema simples (isto é, não complexo) é aquele em que cada compartimento conteria o mesmo número de componentes. Este é o estado ao qual um sistema simples tenderia espontaneamente e após atingi-lo, nele permaneceria. Por isto o estado em que os componentes estivessem distribuídos uniformemente no sistema é chamado de estado de equilíbrio. É interessante observar que é no estado de equilíbrio que se dispõe de menos informação a respeito de um componente isolado de um sistema, ou seja, é o estado em que o sistema apresenta-se com um grau máximo de desordem. Por exemplo, se 90% dos componentes do sistema estivessem contidos no primeiro compartimento e 10% no segundo, caso estivessemos procurando localizar um componente do sistema a probabilidade de encontrá-lo no primeiro compartimento seria nove vezes maior do que no segundo. Neste caso disporíamos de muito mais informação sobre a localização dos componentes do que se estes estivessem igualmente distribuídos nos dois compartimentos. Assim, pode-se afirmar que qualquer estado mais ordenado do sistema corresponde a uma quebra da sua simetria. Um aumento da organização de um sistema corresponde assim à formação de padrões assimétricos espaciais (relativos à localização dos seus componentes) ou temporais (relativos à sucessão dos estados dos seus componentes ao longo do tempo).

Prigogine e Stengers (1986, p. 239), evocando a “lei dos grandes números”, salientam que quanto maior o número de componentes de um sistema, menor é a magnitude das suas flutuações em relação ao seu

estado de equilíbrio. No exemplo citado anteriormente, isto significa que quanto maior o número de componentes N do sistema, relativamente menores serão as flutuações em torno da repartição $N/2$, de maior probabilidade, em seu estado de equilíbrio. Considerando a média X das observações do estado do sistema como

$$X = \frac{1}{n} \sum_{i=1}^n X_i \quad (2)$$

onde

X_i = número de componentes em cada compartimento na observação “ i ”

n = número de observações

e a dispersão das observações em relação à média X como

$$(X - N/2)^2 = \frac{1}{n} \sum (X_i - N/2)^2 \quad (3)$$

o valor relativo da dispersão, descrito por

$$\frac{\sqrt{(X - N/2)^2}}{X} \quad (4)$$

é da ordem de

$$\frac{1}{\sqrt{N}} \quad (5)$$

Assim, na medida em que o número de componentes de um sistema aumenta, a sua dispersão relativa tende a diminuir, o que indica que, se um sistema for suficientemente grande, as flutuações do seu estado podem ser consideradas negligenciáveis. A lei dos grandes números sintetizada anteriormente se constitui em um dos fundamentos mais importantes das distribuições estatísticas mais utilizadas, como a de Gauss e de Poisson (Prigogine; Stengers, 1986, p. 240).

Nos sistemas complexos, caracterizados termodinamicamente como sistemas dissipativos que se mantêm longe do equilíbrio, entretanto, a lei dos grandes números não é respeitada, isto é, ela é

contrabalançada por outros processos (Prigogine; Stengers, 1986, p. 240). Os autores mostram que a lei dos grandes números é válida quando as flutuações amortecem ao longo do tempo ou se neutralizam mutuamente e os efeitos das alterações das condições de contorno (como a temperatura e a pressão) são linearmente proporcionais a sua magnitude. Isto só é observado nos sistemas em equilíbrio, ou que se encontram próximos dele. Os sistemas complexos, porém, apresentam relações não-lineares que provocam a amplificação de certas flutuações, as quais deixam assim de ser neutralizadas por flutuações no sentido contrário, provocando uma auto-organização que é específica (e muitas vezes única) do sistema, o qual passa a se afastar ainda mais do equilíbrio. É fundamentalmente sobre esta capacidade de criação de “ordem por flutuação”, apresentada pelos sistemas complexos (Prigogine; Stengers, 1986, p. 239), que estão baseados os processos evolutivos apresentados pela biosfera e seus subsistemas, incluindo-se entre estes as sociedades humanas.

Os processos evolutivos, caracterizados por mudanças, inovação, diversidade e surpresas, são, portanto, tipicamente abertos, o que contrasta com o “mundo fechado”, pressuposto pela Teoria da Probabilidade (Smets, 1990). Por esta razão, o caráter evolutivo dos sistemas complexos constitui-se em um sério obstáculo à realização de inferências estatísticas sobre o seu comportamento. Isto porque, na medida em que os processos evolutivos representam o surgimento de características novas do sistema, inclusive por meio de mudanças das relações entre seus componentes, a projeção do seu comportamento futuro como uma simples extensão do passado torna-se problemática. A aplicação de inferências ao estudo de sistemas complexos é ainda mais problemática quando a análise inclui as suas propriedades emergentes, isto é, propriedades que não podem ser atribuídas aos componentes do sistema tomados isoladamente, mas sim às relações que estes mantêm entre si. Por este motivo, a consideração de propriedades emergentes como objeto de estudo de determinados sistemas, especialmente os relacionados às sociedades humanas, só pode ser realizada sob condições não controladas (Harvey; Reed, 2004).

Tendo em vista que a abordagem de sistemas agrários constitui-se em uma interpretação da agricultura que destaca seu caráter complexo e evolutivo (Silva Neto, 2004, 2005, 2007, 2008), como a ADSA pode ser avaliada à luz dos princípios básicos da estatística? Algumas contribuições para responder a esta pergunta podem ser obtidas a partir da TE.

A Teoria da Evidência

A tolerância à incerteza, com sua diminuição controlada e progressiva, pode ser uma estratégia interessante para lidar com a complexidade. Segundo Klir (1994), ao lidarmos com um problema, a manutenção de alguma incerteza pode reduzir substancialmente a complexidade computacional do procedimento, permitindo, ao mesmo tempo, que a credibilidade da solução seja melhorada. De acordo com este autor, a TE, pelas suas características, pode ser uma ferramenta importante para o desenvolvimento de procedimentos baseados nessa estratégia.

A TE estuda as relações existentes entre um conjunto de proposições consideradas verdadeiras, denominadas “evidências”, e um conjunto de hipóteses falsificáveis, denominado “quadro de discernimento” (Shafer, 1976, p. 36; Saffiotti, 1994). A TE pode ser descrita sucintamente por meio de três funções básicas: a função de “alocação de massa³” m , a função de credibilidade Bel e a função de plausibilidade Pl (Klir, 1994). A função de alocação de massa define um mapa do conjunto de potência 2^X aplicado ao intervalo entre 0 e 1. O valor de uma função de alocação de massa para um dado conjunto A (ou seja, $m(A)$) expressa toda evidência disponível e relevante que suporta que um elemento do quadro de discernimento X pertence ao conjunto A . O valor de $m(A)$ refere-se ape-

³Tradução do inglês *mass assignment*, expressão que alguns autores consideram mais apropriada do que a expressão “atribuição básica de probabilidade” (do inglês *basic probability assignment* comumente adotado), devido ao fato de designar um conceito diferente do conceito bayesiano de probabilidade (Kruse; Klawonn, 1994).

nas ao conjunto A como um todo e não fornece informação sobre qualquer subconjunto de A . Uma evidência adicional sobre um subconjunto de A deve ser representada por outra função de alocação de massa, isto é, se $B \subset A$, $m(B)$ seria a função de alocação de massa do subconjunto B . Cada conjunto $A \in 2^X$ para o qual $m(A) \neq 0$ é chamado um elemento focal. Considerando F como o conjunto de elementos focais de m , o par (F, m) é denominado corpo de evidência.

Assim, formalmente a função de alocação de massa m é definida como

$$m: 2^X \rightarrow [0,1] \quad (6)$$

$$\sum_{A \subset X} m(A) = 1 \quad (7)$$

De acordo com Smets (1990), a definição anterior da função de alocação de massa não inclui $m(\emptyset) = 0$, o que significa que ela permite que se leve em conta explicitamente a possibilidade da hipótese verdadeira não ser prevista, ou seja, estar fora do quadro de discernimento corrente, decorrência do pressuposto de um “mundo aberto” adotado por esse autor.

A partir da função de alocação de massa, pode-se definir um limite inferior e um limite superior de probabilidade. Este intervalo é limitado por duas medidas contínuas não-aditivas chamadas de credibilidade, correspondente ao limite inferior, e plausibilidade, correspondente ao limite superior. Formalmente as funções de credibilidade $Bel(A)$ e de plausibilidade $Pl(A)$ podem ser descritas como

$$Bel(A) = \sum_{B|B \subseteq A} m(B) \quad (8)$$

$$Pl(A) = \sum_{B|B \cap A \neq \emptyset} m(B) \quad (9)$$

Assim, a função de credibilidade $Bel(A)$ indica a que ponto as informações fornecidas por uma fonte sustentam A . A função de plausibilidade $Pl(A)$ indica a que ponto as informações fornecidas por uma fonte não contradizem A . Estas funções podem ser derivadas uma da outra. Por exemplo:

$$Pl(A) = 1 - Bel(\bar{A}) \tag{10}$$

onde \bar{A} é o complemento clássico de A , ou seja,

$$Bel(\bar{A}) = \sum_{B|B \subseteq \bar{A}} m(B) = \sum_{B|B \cap A = \emptyset} m(B) \tag{11}$$

$$Pl(A) = \sum_{B|B \cap A \neq \emptyset} m(B) = 1 - \sum_{B|B \cap A = \emptyset} m(B) \tag{12}$$

Este exemplo mostra que quanto mais a plausibilidade de uma hipótese aumenta, menor se torna a credibilidade da hipótese contrária. Um aspecto interessante da TE é que é possível definir a ignorância a respeito do valor preciso de uma probabilidade, no sentido clássico, pela diferença entre a plausibilidade e a credibilidade, ou seja (Coutto et al, 2007):

$$I(A) = Pl(A) - Bel(A) \tag{13}$$

Funções de alocação de massa originadas de fontes independentes de evidência podem ser combinadas pela regra de Dempster, que pode ser descrita como (Zhang, 1994):

$$m(B) = \frac{\sum_{B_1 \cap B_2 = B} m_1(B_1)m_2(B_2)}{\sum_{B_1 \cap B_2 \neq \emptyset} m_1(B_1)m_2(B_2)} \tag{14}$$

A relação entre a TE e outras teorias estatísticas

As Teorias da Possibilidade e da Probabilidade podem ser interpretadas como casos particulares da TE.

Quando todos os elementos focais são consonantes, podendo assim ser ordenados por inclusão, as funções de plausibilidade e de credibilidade tornam-se, respectivamente, funções de possibilidade e de necessidade, o que torna a Teoria da Possibilidade um caso particular da TE (Klir, 1994).

Já a Teoria da Probabilidade se constitui em um caso particular da TE quando

$$Bel(A) = Pl(A) \quad (15)$$

o que implica em $I(A) = 0$.

Neste caso, ao analisarmos uma hipótese, esta deve ser considerada como um evento singular, à qual é atribuída certa probabilidade de ser verdadeira. Além disso, a somatória das probabilidades deve ser igual à unidade, o que implica que todos os eventos que podem ocorrer no sistema já estão previstos (pressuposto de “mundo fechado”).

Interpretação conceitual da TE

A maior abrangência da Teoria da Evidência permite que ela seja interpretada a partir de um quadro conceitual mais amplo do que o da Teoria da Probabilidade (Fioretti, 2001; Shafer, 1976, p. 274; Smets, 1990). Isto porque, segundo a TE, o conjunto de hipóteses a serem testadas não se constitui de uma agregação de elementos singulares. Ao contrário, segundo esses autores, um quadro de discernimento pode corresponder a uma visão global de um objeto de análise, a qual não necessariamente

precisa ser compreendida como a combinação de componentes mais fundamentais. Como salienta Shafer (1976, p. 120), isto permite que a TE traduza o fato de, na prática, a construção do conhecimento sempre se iniciar por um quadro relativamente grosseiro de discernimento, que é então sucessivamente refinado de acordo com as necessidades e as evidências disponíveis. É por esta razão que a TE pode ser considerada mais bem adaptada como base para investigações científicas, pois trata de quadros de discernimento em que se reconhece explicitamente a não abrangência de todo o conhecimento possível, permitindo assim que o seu refinamento seja incorporado na própria lógica da obtenção do conhecimento. A partir desta posição, Fioretti (2001, 2004) afirma que a TE possibilita justificar a existência de hipóteses imprevistas que surgem ao longo de um processo de busca de conhecimento. Assim, conclui Fioretti (2001), enquanto a Teoria da Probabilidade expressa a incerteza em termos de chances que governam a verdade, a TE expressa a incerteza em termos do conhecimento provido por uma mensagem cuja verdade é desconhecida. No primeiro caso, o conjunto de probabilidades indica onde a verdade pode estar; no segundo caso, o quadro de discernimento indica como um indivíduo pode interpretar uma mensagem, aproximando-se da verdade por meio da eliminação de contradições.

A Análise-Diagnóstico de Sistemas Agrários

O reconhecimento de que a realidade da agricultura é complexa e heterogênea confere uma nova perspectiva aos estudos que visam à definição de projetos de desenvolvimento. Com efeito, é crescente a tomada de consciência, entre pesquisadores e extensionistas, de que muitos dos problemas encontrados resultam de uma análise reducionista das condições em que ocorre a produção agropecuária, a qual é responsável pela inadequação das alternativas propostas aos agricultores (Dufumier, 1996; Michailoff, 1987). A ADSA procura contornar estes problemas por meio de uma abordagem sistêmica em vários níveis, cada qual relacionado a um conceito específico, os quais são apresentados sucintamente a seguir.

Conceitos básicos

O conceito relacionado ao nível mais geral de análise é o de “Sistema Agrário”, que corresponde a um modo específico de exploração de um ecossistema resultante de transformações históricas profundas e de adaptações geográficas em larga escala. Na caracterização de um sistema agrário o que importa em primeiro lugar são as tendências históricas que regem as grandes mudanças da agricultura.

Para a análise de um sistema agrário as seguintes variáveis são consideradas essenciais:

- o ecossistema cultivado, o qual é o produto histórico das transformações promovidas pelo homem sobre o ecossistema natural;
- os meios de produção que correspondem às ferramentas, máquinas e materiais biológicos (plantas cultivadas e animais domésticos) resultantes do processo de adaptação, seleção e acumulação desenvolvido historicamente pelos agricultores e outros agentes;
- a força de trabalho e as relações de produção às quais ela está submetida, ou seja, a mão-de-obra disponível e as relações de produção que regem a repartição do produto. As relações de produção (trabalho familiar, assalariamento, parceria) são importantes, pois condicionam fortemente os critérios de alocação de recursos adotados pelo agricultores;
- o modo de exploração e reprodução do ecossistema cultivado, o qual resulta da forma específica como os instrumentos de produção são aplicados localmente, em razão das características do ecossistema cultivado, da distribuição dos meios de produção, da força de trabalho disponível e das relações de produção vigentes.

Mesmo no interior de um mesmo sistema agrário, porém, a combinação dos elementos citados não é homogênea. O ecossistema cultivado, a disponibilidade da força de trabalho e dos meios de produção variam segundo o estatuto social e a acumulação de cada agricultor. Assim, se-

gundo a disponibilidade de meios de produção e de força de trabalho presentes em uma unidade de produção agropecuária, e a maneira como estes elementos são combinados para a artificialização do ecossistema cultivado, pode-se definir diferentes “Sistemas de Produção”.

A análise dos Sistemas de Produção, entretanto, dificilmente pode ser efetuada sem a sua divisão em subsistemas mais simples. Os Sistemas de Cultura e Sistemas de Criação visam, assim, a especificar a análise da produção vegetal e da produção animal no seio da unidade de produção. Os Sistemas de Cultura são, portanto, definidos a partir da forma como uma determinada gleba de terra é cultivada ao longo dos anos (rotações ou sucessões de culturas). Já os Sistemas de Criação correspondem à forma como são conduzidas as produções animais (espécies, técnicas de alimentação e de manejo, áreas utilizadas, etc.). Sobre cada uma das culturas ou criações da unidade de produção são aplicados diferentes “Itinerários Técnicos”, os quais correspondem a uma sucessão lógica de “Operações Técnicas” elementares (por exemplo, a aração, a aplicação de defensivos, etc.).

Princípios metodológicos

Uma das motivações para a elaboração dos procedimentos da ADSA foi a observação das dificuldades provocadas por uma estratégia comum adotada em estudos que visam a apoiar ações de desenvolvimento. Tais estratégias baseiam-se, essencialmente, na obtenção do máximo de informações possíveis, o que em geral ocasiona grandes dificuldades no tratamento de dados e, quase sempre, a conclusão de que os dados mais pertinentes à pesquisa não foram obtidos (Dufumier, 1996). Segue-se assim uma nova rodada (muitas vezes com recursos adicionais e com atrasos no cronograma da pesquisa) de coleta de dados, com novas dificuldades de tratamento, a qual leva a novas coletas, etc. Desta forma muitas pesquisas bastante promissoras sobre situações concretas de desenvolvimento terminam sem apresentar resultados consistentes (Michailoff, 1987).

Os procedimentos da ADSA procuram evitar este tipo de problema por meio de uma aplicação metódica e rigorosa de alguns princípios metodológicos. Tais princípios constituem-se, sinteticamente, em:

- efetuar as análises a partir dos fenômenos mais gerais para os mais específicos, por meio de uma abordagem sistêmica em vários níveis;
- analisar a realidade em cada nível de abrangência especificamente, realizando uma síntese antes de passar para a análise em nível inferior;
- privilegiar a explicação em detrimento da descrição;
- privilegiar uma visão dinâmica das situações por meio da adoção de enfoques históricos;
- atender à heterogeneidade da realidade, evitando interpretações genéricas, as quais dificultam a percepção de processos de diferenciação.

Assim, os estudos devem inicialmente se concentrar nos aspectos mais gerais da realidade a ser estudada e só passar a detalhes mais específicos após uma síntese que permita formular quais são as variáveis mais pertinentes a serem analisadas (ou questões mais importantes a serem respondidas), no nível imediatamente inferior. Tal síntese é efetuada pela organização e análise da coerência das informações obtidas, sendo retidas apenas aquelas consideradas imprescindíveis para explicar a realidade observada, e não para descrevê-la, no nível de abrangência em questão.

Principais etapas

A partir dos princípios metodológicos descritos anteriormente a ADSA é aplicada seguindo-se algumas etapas, caracterizadas concisamente a seguir (Dufumier, 1996; Garcia Filho, 1999).

Etapa 1 – Caracterização geral e identificação das principais heterogeneidades do processo de desenvolvimento local

Esta etapa compreende a caracterização agroecológica e socioeconômica da região de estudo (IDH, nível e distribuição da renda, condições gerais de vida da população, principais atividades econômicas, etc.). São realizadas análises de dados secundários e, principalmente, uma leitura da paisagem a partir da qual são definidas zonas relativamente homogêneas do ponto de vista das condições para a atividade agropecuária (clima, solo, infra-estrutura, etc.).

Etapa 2 – Análise da Evolução Histórica

Nesta etapa é realizada uma análise da história da região, especialmente no que diz respeito a sua agricultura. Concomitantemente a esta análise histórica, procura-se reconstituir os principais processos de diferenciação social e trajetórias de acumulação de capital relacionados à agricultura. As informações são obtidas a partir de fontes secundárias e, principalmente, por meio de entrevistas com pessoas que, durante as últimas décadas, acompanharam as transformações sofridas pela agricultura local.

Etapa 3 – Tipologia das Unidades de Produção

Nesta etapa as unidades de produção da região são agrupadas em tipos. As informações para a realização desta tipologia são obtidas por meio de entrevistas com agricultores, os quais são selecionados progressivamente ao longo do estudo, adotando-se como ponto de partida os processos de diferenciação (em especial as trajetórias de acumulação, estagnação e descapitalização) identificados na etapa anterior.

Os tipos são definidos de forma a representar as categorias sociais e formas de organização da produção (sistemas de produção) necessárias para a compreensão da dinâmica da agricultura da região. Nesta etapa também são realizadas a caracterização técnica e a avaliação econômica dos sistemas de produção, objetivando esclarecer a capacidade de reprodução social de cada tipo. Para cada tipo de unidade de produção são elaborados modelos lineares representando a variação dos resultados econômicos (valor agregado e renda) em função da área disponível.

Os modelos de valor agregado global permitem a avaliação da capacidade de geração de riquezas para a sociedade como um todo dos sistemas de produção de cada tipo, segundo sua escala de produção. Já os modelos de renda global permitem a observação da escala mínima de produção necessária a cada tipo para que este possa atingir um nível de renda compatível com a sua reprodução social.

São formulados também modelos de composição da renda por meio dos quais é possível observar como se combinam os diferentes sistemas de cultura e de criação de cada tipo, assim como quais destes sistemas possuem o maior potencial de geração de renda.

Etapa 4 – Definição de Linhas Estratégicas de Desenvolvimento

Inicialmente procura-se avaliar as possibilidades de melhorar as condições para a reprodução econômica das unidades de produção em função do tipo de sistema de produção adotado. Após, a partir da caracterização técnica e das avaliações econômicas da etapa anterior, procura-se identificar atividades ou técnicas que contribuam para um aumento da produtividade e da renda, levando em conta os estrangulamentos anteriormente detectados em cada tipo de sistema de produção analisado. Com base nestes resultados são discutidas, com os agentes locais de desenvolvimento, alternativas de ação técnica, organizacional, gerencial e de políticas públicas para o desenvolvimento dos diferentes tipos de unidades de produção, bem como estratégias de intervenção no processo de desenvolvimento regional.

A Teoria da Evidência e a ADSA

A breve revisão realizada nos itens anteriores permite apontar para uma grande coerência entre, por um lado, os princípios metodológicos e as etapas dos procedimentos da ADSA e, por outro, os conceitos relacionados à produção de conhecimento baseados na TE. Neste sentido des-

taca-se o conceito de quadro de discernimento, o qual permite que as hipóteses sobre as quais se pretende reunir evidências sejam inicialmente definidas de forma relativamente imprecisa, podendo ser especificadas de acordo com o avanço da pesquisa. Tal procedimento contrasta com os métodos estatísticos normalmente adotados que exigem que todas as hipóteses pertinentes ao desenvolvimento da agricultura da região sejam definidas *a priori*.

Alguns exemplos podem contribuir para ilustrar a forma como a TE pode ser utilizada para interpretar os procedimentos adotados na ADSA.

O primeiro exemplo está relacionado a uma leitura da paisagem. Suponhamos que foram identificadas três zonas em 70% de uma região a ser estudada. Observou-se que 10% da região percorrida correspondem à zona 1, 30% à zona 2 e os 60% restantes correspondem à zona 3. A partir destes dados pode-se calcular os valores das funções de alocação de massa, de credibilidade, de plausibilidade e de ignorância, os quais são mostrados na Tabela 1.

Tabela 1: Valores das funções da TE em um zoneamento parcial de uma região

Hipótese	Alocação de Massa	Credibilidade	Plausibilidade	Ignorância
Zona 1	7%	7%	37%	30%
Zona 2	21%	21%	51%	30%
Zona 3	42%	42%	72%	30%
Zona 1, 2 ou 3	30%	100%	100%	

Fonte: Elaboração do autor.

Segundo a TE, os dados descritos na Tabela 1 indicam que a proporção exata da zona 1 em relação a toda a região é de, no mínimo, 7% e, no máximo, de 37%, enquanto que a da zona 2 encontra-se entre os limi-

tes de 21% e 51% e a zona 3 entre 42% e 72%. Isto porque, dada a inexistência de qualquer distribuição estatística confiável, não é possível que as proporções observadas em uma parte da região possam ser extrapoladas para a totalidade da mesma.

Esta posição se justifica plenamente se considerarmos que uma região é um sistema complexo, cujos processos sociais, ecológicos e geológicos de diferenciação não podem ser descritos por uma distribuição aleatória. Assim, a única maneira de realizar uma leitura da paisagem adequada é percorrer toda a região, sendo para isto necessário redefinir o quadro de discernimento para que as zonas possam ser definidas de forma suficientemente abrangente. Isto, por sua vez, exige que o zoneamento esteja estritamente focado nas questões mais relevantes a serem respondidas no atual estágio da pesquisa. Um exemplo do que poderia ser esta redefinição é mostrado na Tabela 2.

Tabela 2: Valores das funções da Teoria da Evidência em um zoneamento de uma região após a redefinição do quadro de discernimento

Hipótese	Alocação de Massa	Credibilidade	Plausibilidade	Ignorância
Zona 1	40%	40%	40%	0%
Zona 2	60%	60%	60%	0%
Zona 1 ou 2	0%	100%	100%	

Fonte: Elaboração do autor.

Outro exemplo, bastante simples, mostra-se interessante para ilustrar a forma como, de acordo com a TE, evidências podem ser buscadas para a sustentação de uma hipótese. Suponhamos que alguém esteja viajando e deseje saber se o alarme da sua casa disparou. Esta pessoa então telefona para um vizinho, que lhe responde que o alarme não está tocando, porém como este vizinho mora relativamente longe da sua casa, esta pessoa atribui uma credibilidade bastante limitada à informação

fornecida. Vamos supor também que o dono da casa atribui valores à credibilidade dos vizinhos, tal como um sistema “expert” baseado na TE faria. Assim, o dono da casa atribui 60% de credibilidade à informação de que o alarme não está tocando, fornecida por meio do primeiro telefonema. Ele resolve então telefonar para um segundo vizinho que, por morar mais perto, pode lhe fornecer uma informação com maior credibilidade, que ele atribui ser de 90%. Este segundo vizinho afirma que o alarme está tocando. Pela regra de Dempster descrita em (14) é possível calcular o quanto o dono da casa pode crer que o alarme está tocando a partir da combinação das evidências que ele reuniu da seguinte forma:

$$\frac{0,9 * 0,4}{(0,9 * 0,4) + (0,1 * 0,6) + (0,1 * 0,4)} \cong 0,783$$

É interessante observar que a combinação das afirmações dos dois vizinhos, que corresponderia a $(0,6 * 0,9)$ não consta nos cálculos, pois tais afirmações são contraditórias (o alarme não pode estar tocando e não estar tocando ao mesmo tempo).

O importante a reter deste exemplo é que a forma como o dono da casa procurou as evidências baseou-se em uma “teoria” que ele estabeleceu entre a distância que os vizinhos moram da sua casa e a credibilidade das informações fornecidas pelos mesmos, e não por qualquer procedimento definido segundo a teoria da amostragem usualmente adotada em análises estatísticas. Ele simplesmente procurou a informação onde ela poderia lhe ser fornecida com a maior credibilidade e, evidentemente, se ele pudesse telefonar para mais vizinhos próximos a sua casa, a credibilidade das informações sobre o disparo ou não do alarme tenderia a crescer, considerando que a sua “teoria” seja adequada à situação. Ele poderia, contudo, também facilmente perceber que tal teoria é inadequada. Isto seria caracterizado pelo surgimento de evidências altamente

conflitantes.⁴ Por exemplo, a consulta de dois vizinhos que moram a uma mesma distância da casa poderia indicar evidências contrárias porque um dos vizinhos tem problemas de audição devido à idade avançada. Neste caso, o dono da casa deveria rever a sua teoria, levando em consideração outras características dos vizinhos. É interessante observar que ao final deste processo iterativo de reunião de evidências e refinamento da teoria o dono da casa não apenas concluiria se o alarme está tocando ou não, mas disporia também de uma “teoria” mais adequada. Em outras palavras, ao procurar saber se o alarme disparou o dono da casa poderia aprender muito sobre o comportamento dos seus vizinhos.

Este exemplo pode contribuir para que se esclareça a forma como são realizadas as tipologias de unidades de produção na ADSA. Como decorre da definição das funções de alocação de massa (6 e 7), de credibilidade (8), de plausibilidade (9) e de ignorância (13), se

$$I(A) > 0 \quad (16)$$

então

$$Bel(A) < Pl(A) \quad (17)$$

o que implica em

$$\sum Bel(A) < 1 \quad (18)$$

e

$$\sum Pl(A) > 1 \quad (19)$$

A expressão (16) indica a existência de incerteza na definição de uma tipologia a partir de conceitos relacionados à TE, sendo que as expressões (18) e (19) permitem especificar como tal incerteza se manifes-

⁴ Isto faria também com que a aplicação da regra de Dempster desse resultados contra-intuitivos, o que é considerado um problema na construção de sistemas inteligentes (*experts*), gerando uma abundante literatura sobre regras de combinação de evidências (Yager et al, 1994). No entanto, a nosso ver, tais resultados contra-intuitivos gerados pela regra de Dempster apenas indicam a inadequação da combinação de evidências muito conflitantes.

ta. Assim, a expressão (18) (que indica que, na presença de incerteza a função de credibilidade é subaditiva) indica que as informações são insuficientes para suportar a pertinência de certas unidades de produção a algum dos tipos definidos. Já a expressão (19) (que indica que na presença de incerteza a função de plausibilidade é super-aditiva) indica que as informações são insuficientes para definir que certas unidades de produção pertencem a apenas um dos tipos definidos. Em outras palavras, o conceito de plausibilidade serve para indicar a possibilidade do surgimento de ambigüidades (Strat, 1994).

Com base na TE pode-se argumentar que na ADSA um tipo é estabelecido principalmente a partir da sua plausibilidade (ou seja, pelas evidências que não excluem a sua existência), o que sempre acarreta alguma incerteza em relação as suas características específicas. Isto porque um tipo corresponde a unidades de produção em que ocorrem processos que lhes asseguram características comuns, porém não necessariamente idênticas. Neste sentido formalmente um tipo corresponderia a uma intersecção de diferentes unidades de produção (correspondente aos processos que lhes são comuns) e não de um conjunto mais abrangente do qual as unidades de produção seriam partes integrantes. E a identificação dos processos que determinam o que é plausível para a definição dos tipos é realizada principalmente⁵ pela análise da formação histórica da região, especialmente ao se estudar os processos de diferenciação social.

A identificação desses processos constitui-se, entretanto, apenas no primeiro elemento de uma “teoria” dos tipos. A posterior busca de evidências para o refinamento e caracterização dos mesmos deve então ser realizada por meio da análise de dados coletados em entrevistas com agricultores e, é importante salientar, por meio do concomitante refinamento dos pressupostos utilizados para a definição dos tipos (ou seja, da própria “teoria” dos tipos), até o ponto em que as ambigüidades tenham

⁵Não exclusivamente, porém, iniciando-se já nas primeiras etapas da pesquisa.

sido reduzidas a um nível que não impeça uma interpretação adequada do papel dos diferentes tipos no processo de desenvolvimento da agricultura local. Assim sendo, a elaboração de uma tipologia não implica necessariamente a eliminação de todas as suas ambigüidades, mas apenas daquelas que impedem uma compreensão adequada dos processos de diferenciação e de reprodução social das unidades de produção da região. É que a complexidade destes processos, especialmente no que diz respeito ao seu caráter emergente e contexto-dependente, dificilmente permite que estes sejam totalmente discriminados. Muitos deles podem mesmo coexistir em uma mesma unidade de produção, cada um predominando de acordo com uma configuração específica que o sistema agrário exhibe em dados momentos. Por esta razão a tentativa de eliminar totalmente as ambigüidades de uma tipologia pode levar à exigência de grandes volumes de dados sem a obtenção de resultados satisfatórios. A redução controlada das ambigüidades, desse modo, é uma das principais características que asseguram a eficiência dos procedimentos adotados para a realização de tipologias no âmbito da ADSA.

O entendimento de que um tipo de unidade de produção corresponde a certa interpretação dos processos de diferenciação social e de acumulação implica que os elementos que presumivelmente constituem as unidades de produção representadas por um tipo não podem ser tomados como o ponto de partida para uma tipologia. Ao contrário, a descrição de tais elementos, como os relacionados às categorias sociais e aos sistemas de produção, constituem-se em uma forma de caracterizar um tipo cuja concepção antecede esta descrição. Em outras palavras, na ADSA, a construção de uma tipologia não corresponde a uma combinação de dados, mas sim a uma progressiva desagregação de uma visão mais abrangente teoricamente estabelecida. Evidentemente, posteriormente a análise dos elementos constituintes das unidades de produção contribui para refinar a definição do tipo, mas na ADSA não é a coleta de

elementos singulares o ponto de partida para a elaboração da tipologia. De acordo com a interpretação da TE de Fioretti (2001), discutida anteriormente, tanto o quadro de discernimento (no caso, a tipologia), como os elementos focais (no caso, cada um dos tipos) constituem-se em conjuntos que expressam uma visão global do seu objeto, a qual é resultado do conhecimento disponível sobre o mesmo e não uma agregação de elementos singulares que, a princípio, representariam todo o conhecimento possível sobre o objeto de estudo.⁶ Isto implica que a validação de uma tipologia realizada no quadro de uma ADSA só pode ser efetuada pela análise da sua coerência em relação aos processos de desenvolvimento da agricultura local e não pelas características dos dados obtidos nas unidades de produção (correlações, dispersão em relação a uma medida de tendência central, etc.).

Assim, a busca de evidências para a elaboração da tipologia na ADSA, como o procedimento adotado pelo dono da casa para saber com mais precisão se o alarme estava tocando, nada tem a ver com os processos de amostragem baseados na Teoria da Probabilidade. De fato, o procedimento adotado na ADSA implica o tratamento de um volume de dados muito menor do que os que são necessários para a aplicação de outros métodos estatísticos.⁷ Assim, com um volume de dados relativamente reduzido, pois coletado para responder questões específicas e bem definidas surgidas ao longo da pesquisa, é possível pela ADSA chegar a um conhecimento bastante esclarecedor dos processos de desenvolvimento presentes na agricultura de uma região.

⁶Tal como é pressuposto nos métodos probabilistas (Fioretti, 2001).

⁷Nos quais o próprio volume de dados (como o tamanho da amostra) parece servir como uma garantia de confiabilidade da tipologia, o que explica a estranheza muitas vezes provocada pelos procedimentos da ADSA entre pesquisadores habituados aos métodos baseados na Teoria da Probabilidade.

Conclusões

A interpretação dos procedimentos adotados na ADSA no quadro da TE indica que a ADSA não apresenta qualquer contradição com os princípios que devem reger uma análise estatística rigorosa, especialmente quando se leva em conta a complexidade do desenvolvimento da agricultura, cujo estudo não pode ser realizado isolando-a do seu contexto histórico e geográfico. Neste caso, típico de pesquisa em condições não controladas, a aplicação generalizada de métodos baseados na Teoria da Probabilidade pode comprometer seriamente o rigor científico da pesquisa, pois implicaria uma definição *a priori* de todas as variáveis a serem analisadas e hipóteses a serem consideradas. Já a adoção de procedimentos coerentes com a TE, ao contrário, permite que a ADSA apresente uma grande eficiência, tanto do ponto de vista da sua adequação ao seu objeto como em relação aos recursos necessários para a sua execução.

Isto não significa, no entanto, que a TE permita uma formalização dos procedimentos da ADSA de modo que esta possa ser aplicada de forma mecânica. Ao contrário, a maior contribuição da TE para a interpretação dos procedimentos da ADSA parece residir no fato de ela oferecer um quadro de análise interessante ao processo, contínuo e progressivo, de construção do conhecimento na ADSA, o qual se constitui na sua principal característica. Se a TE, contudo, permite formalizar o quanto as evidências reunidas suportam ou tornam plausíveis certas hipóteses e como, e em que condições, novas evidências podem ser incorporadas ao longo do estudo, tanto a criação e a modificação dos quadros de discernimento, que podem ser refinados, generalizados, ampliados ou reduzidos, conforme as necessidades da pesquisa (Shafer, 1976, p. 281), assim como o surgimento de novas hipóteses (Fioretti, 2001, 2004), não podem ser formalizados. Por isso é provável que a maior contribuição da TE ao aperfeiçoamento da ADSA esteja relacionada mais aos seus aspectos conceituais do que à definição de procedimentos quantitativos.

Referências

- COUTTO FILHO, M. B.; SOUZA, J. C. S. de; SCHILLING, M. T. Sobre o problema da integração generalizada de dados. *Revista Controle & Automação*, vol. 8, número 1, p. 24-43, 2007.
- DUFUMIER, M. *Agricultures et paysanneries des Tiers Mondes*. Paris: Editions Karthala, 2004.
- DUFUMIER, M. *Les Projets de Développement Agricole Manuel d'expertise*. Paris: Éditions Karthala, 1996.
- FIORETTI, G. A mathematical theory of evidence for G.L.S. Shackle. ICER – International Centre for Economic Research, série *ICER Working Papers* número 3, 2001. Disponível em: <<http://edirc.repec.org/data/icerrit.html>>. Acesso em: 14 set. 2007).
- FIORETTI, G. Evidence theory: a mathematical framework for unpredictable hypotheses. *Metroeconomica*, vol. 55, n. 4, p. 345-366, 2004.
- GARCIA FILHO, D. P. *Guia metodológico – diagnóstico de sistemas agrários*. Brasília, Incra/FAO, 1999. Disponível em: <<http://www.incra.gov.br/fao/>>. Acesso em: 4 set. 2002.
- GLEICK, J. *Chaos: Making a New Science*. New York: Viking-Penguin, 1987.
- GRIBBIN, J. *Deep simplicity: bringing order to chaos and complexity*. New York: Random House, 2004.
- HARVEY, D. L.; REED, M. Social science as the study of complex systems. In: KIEL, D. L.; ELLIOT, E. (Ed.). *Chaos theory in the social sciences: foundations and applications*. Michigan: University of Michigan Press, 2004. p. 295-323.
- KLIR, G. Measures of uncertainty in the Dempster-Shafer Theory of Evidence. In: YAGER, R. R.; FEDRIZZI, M.; KACPRZYK, J. (Ed.). *Advances in the Dempster-Shafer Theory of Evidence*. New York: John Wiley & Sons, 1994. p. 35-49.
- KRUSE, R.; KLAWONN, F. Mass distribution on L-fuzzy sets and families of frames of discernment. In: YAGER, R. R.; FEDRIZZI, M.; KACPRZYK, J. (Ed.). *Advances in the Dempster-Shafer Theory of Evidence*. New York: John Wiley & Sons, 1994. p. 239-250.
- MAZOYER, M.; ROUDART, L. *Histoire des Agricultures du Monde*. Du Néolithique à la Crise Contemporaine. Paris: Éd. du Seuil, 1997.
- MICHAÏLOFF, S. *Les apprentis sorciers du développement: mythes technocratiques face a la pauvreté rurale*. Paris: Economica, 1987.

PRIGOGINE, I.; STENGERS, I. *La nouvelle alliance*. Métamorphose de la science. Paris: Ed. Gallimard, 1986.

SAFFIOTTI, A. Issues of knowledge representation in Dempster-Shafer's Theory. In: YAGER, R. R.; FEDRIZZI, M.; KACPRZYK, J. (Ed.). *Advances in the Dempster-Shafer Theory of Evidence*. New York: John Wiley & Sons, 1994. p. 415-439.

SHAFER, G. *A Mathematical Theory of Evidence*. Princeton: Princeton University Press, 1976. (Cópia autorizada demandada à Books on Demand, University of Michigan).

SILVA NETO, B. Abordagem sistêmica, sistemas agrários e complexidade. In: MOTA, D. M.; SCHMITZ, H.; VASCONCELOS, H. E. (Org.). *Agricultura familiar e abordagem sistêmica*. Aracaju: Sociedade Brasileira de Sistemas de Produção, 2005. p. 81-103.

SILVA NETO, B. Desenvolvimento sustentável: uma abordagem baseada em sistemas dissipativos. *Ambiente & Sociedade*, vol. XI, n. 1, p. 15-31, jan./jun. 2008.

SILVA NETO, B. Análise-diagnóstico de situações de desenvolvimento: uma interpretação baseada na teoria da complexidade e no realismo crítico. *Desenvolvimento em Questão*, 9: 33-58, jan./jun. 2007.

SILVA NETO, B. Complexidade e desenvolvimento. *Desenvolvimento em Questão*, (4): 9-32, jul./dez. 2004.

SMETS, P. *The Combination of Evidence in the Transferable Belief Model*. IEEE Trans. Pami 12, p. 447-458, 1990.

STRAT, T. M. Decision analysis using belief functions. In: YAGER, R. R.; FEDRIZZI, M.; KACPRZYK, J. (Ed.). *Advances in the Dempster-Shafer Theory of Evidence*. New York: John Wiley & Sons, 1994. p. 276-309.

YAGER, R. R.; FEDRIZZI, M.; KACPRZYK, J. (Ed.). *Advances in the Dempster-Shafer Theory of Evidence*. New York: John Wiley & Sons, 1994.

ZHANG, L. Representation, independence, and combination of evidence in the Dempster-Shafer theory. In: YAGER, R. R.; FEDRIZZI, M.; KACPRZYK, J. (Ed.). *Advances in the Dempster-Shafer Theory of Evidence*. New York: John Wiley & Sons, 1994. p. 51-69.

Recebido em: 28/7/2008

Aceito em: 22/10/2008