

XV COBREAP - CONGRESSO BRASILEIRO DE ENGENHARIA DE AVALIAÇÕES E PERÍCIAS – IBAPE/SP – 2009

TRABALHO DE AVALIAÇÃO

Resumo

Um estudo foi realizado com a finalidade de obter um modelo matemático capaz de estimar o valor de imóveis rurais irrigáveis, localizados na Microrregião de Petrolina/PE. O método empregado para a modelagem foi a regressão linear múltipla, e a variável dependente o valor do imóvel - VTI/ha (Y), e as independentes, a data de negociação ou coleta de informação de oferta (X₁), nível de infraestrutura (X₂), elasticidade da oferta (X₃), área do imóvel (X₄), percentual de área irrigável (X₅), e área equivalente de produção vegetal (X₆). O modelo proposto foi: $Y=9,22005 \times 10^{-48} \times X_1^{10,86709827} \times e^{0,0879898X_2} \times (0,7211X_3+1) \times X_4^{-0,348251052} \times X_5^{0,498900517} \times X_6^{0,120011808}$, e apresentou bom ajustamento com os dados observados, sendo aceitas as hipóteses estatísticas que permitiram sua validação, mostrando-se adequado para estimativa do valor de imóveis irrigáveis com área inferior a 1.500 ha, com o propósito de avaliação por interesse social ou por outros motivos, sendo mais recomendado para avaliação de imóveis de forma massiva ou por motivos fiscais ou tributários. Os fatores que compõem o valor do imóvel podem também ser utilizados na avaliação pelo processo estatístico descritivo de homogeneização por fatores, quando a metodologia de avaliação utilizada for o método comparativo direto de dados de mercado.

Fator irrigação, fator área, produção vegetal, fatores de homogeneização

INTRODUÇÃO

As mesorregiões do Sertão e São Francisco Pernambucanos ocupam mais da metade do estado e possuem características muito específicas. Os imóveis rurais localizados neste território podem assumir valores distintos quando reúnem condições propícias à prática da irrigação (imóveis irrigáveis) em relação àqueles limitados exclusivamente a atividades dependentes de chuvas (imóveis de sequeiro). Além dessa, outras fontes de variação podem estar influenciando o Valor do Imóvel (VTI), como a distribuição de suas terras por classe de capacidade de uso; elasticidade de valores de oferta; dimensão da área; variação da época de negociação ou coleta das informações de ofertas; nível de infraestrutura para o sistema produtivo; percentual da área com potencial para irrigação e área de produção vegetal, no caso de imóveis irrigáveis. Na Microrregião de Petrolina, que abrange os municípios de Petrolina, Lagoa Grande, Santa Maria da Boa Vista, Cabrobó, Orocó, além de Afrânio, Dormentes e Terra Nova, possuem imóveis com características que possibilitam seu uso para a prática da agricultura irrigada, sendo esta região grande produtora e exportadora de frutas, com destaque para a viticultura e mangicultura, além da vitivinicultura.

O método comparativo direto de dados de mercado consiste em obter uma amostra representativa de dados de mercado de imóveis com características, tanto quanto possível, semelhantes às do bem avaliando, usando-se toda a evidência possível (ABNT, 2004). Ao utilizar este método, deve-se fazer o tratamento estatístico dos dados pesquisados. São duas as modalidades de tratamento estatístico das amostras pesquisadas: a metodologia científica, que emprega a estatística inferencial que, segundo BAPTISTELLA (2005), vem sendo mais utilizada por sua maior confiabilidade, e a homogeneização por fatores, que emprega a estatística descritiva. A utilização de tratamento científico (Anexo A da NBR 14.653-3/2004), com emprego de regressão linear múltipla, amplia as condições de atingir melhores níveis de fundamentação e precisão das avaliações, além de permitir a avaliação de forma massiva, em casos de desapropriação por utilidade pública, por exemplo, na construção de obras públicas como ferrovias, estradas, canais de irrigação, etc. Outra aplicação dos modelos é a estimativa do valor venal de imóveis rurais; avaliação de imóveis em processos de desapropriação por interesse social e utilidade pública; avaliação de imóveis na modalidade compra e venda; embasamento técnico para defesa de Entidades Públicas em processos judiciais; e outras aplicações como: reavaliação de ativos de empresas; partilha oriunda de herança, meações ou divórcios; lançamento de impostos; e hipotecas imobiliárias. Para fixação de valores relativos a impostos como o ITBI e ITR, e outras contribuições como taxas cartoriais, são utilizadas comumente tabelas de preços referenciais, plantas genéricas de valores ou apenas a declaração dos envolvidos na transação, que no caso de imóveis rurais, a última opção é a mais adotada.

A avaliação técnica só é considerada legal quando expedida por engenheiro legalmente habilitado. A Lei Federal nº 5.194/66, relativa ao exercício da profissão de engenheiro, arquiteto e engenheiro agrônomo, garante a estes profissionais a exclusividade de avaliar um imóvel, sendo o engenheiro agrônomo responsável pela área rural. As referências bibliográficas que possam auxiliar estes técnicos na tomada de decisão são escassas na região, prejudicando uma análise mais

confiável dos dados de pesquisa, podendo induzi-los a erro e na própria convicção do valor do bem avaliando.

No que concerne a avaliação de imóveis rurais em processos de desapropriação por interesse social, a Lei Federal nº 8.629/03, quando se refere à avaliação do imóvel declarado de interesse social, explicita que a mesma deve fazer referência ao imóvel em sua totalidade, e deste total é deduzido o valor das benfeitorias para serem indenizadas separadamente. Entretanto, é prática comum por parte dos avaliadores, no caso do uso do método comparativo direto de dados de mercado, a retirada dos valores relativos às benfeitorias da amostra para então utilizar os artifícios de estatística descritiva para proceder à homogeneização, obtendo-se assim o valor da terra nua (VTN), que a este é somado o valor das benfeitorias obtendo-se o valor total do imóvel. Esta prática, embora aceita pela Norma Brasileira, não é adequada, pois o mercado imobiliário valoriza o imóvel e não a terra e suas benfeitorias distintamente, como é comum em imóveis urbanos. As benfeitorias existentes no imóvel, de acordo com a vocação de exploração, seu nível de tecnologia e o uso que o comprador pretende dar, constituem um fator que exerce influência na valorização do imóvel.

O trabalho tem como objetivo estabelecer um modelo através de regressão linear múltipla, para estimativa do VTI/ha em função das variáveis que estão influenciando o seu valor (parâmetros significativos) para a Microrregião de Petrolina/PE.

METODOLOGIA

Foram utilizadas efetivamente 32 amostras de pesquisas de preços coletadas pelos técnicos do INCRA SR-29 no período de 2001 a 2008, na Microrregião de Petrolina no Sertão do São Francisco, nos municípios de Petrolina, Santa Maria da Boa Vista e Lagoa Grande. As pesquisas foram realizadas, na sua maioria, a partir de levantamento em cartórios e corretoras com informações confirmadas através de entrevista com o comprador e/ou vendedor e visita ao imóvel. Para a modelagem, foram utilizados apenas dados de imóveis utilizados com agricultura irrigada ou com potencial para esta atividade.

As hipóteses foram formuladas a partir da experiência dos técnicos do INCRA envolvidos nos trabalhos de avaliação de imóveis, em que foram postuladas que: exercem influência positiva no valor do imóvel – VTI/ha, a nota agrônômica; a época da coleta de informações; o nível de infraestrutura; a oferta em relação ao negócio efetivamente realizado; o percentual da área com potencial para irrigação; e, a produção vegetal; e negativa para a dimensão da área. A influência das variáveis foi estudada através de modelagem dos dados por regressão linear múltipla.

A técnica estatística de regressão linear múltipla é usada para estudar a relação entre uma variável dependente e várias variáveis independentes. Em engenharia de avaliações geralmente trabalha-se com modelos de regressão múltipla, tendo em vista a multiplicidade de fatores que interferem nos preços de um bem (DANTAS, 2003).

O modelo genérico é dado pela expressão (1), segundo LEVINE *et al.* (2005), quando aplicado a uma amostra de tamanho n ,

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i, \rightarrow i = 1, 2, \dots, n \quad (1)$$

onde:

y_i = variável dependente ou explicada $i=1, 2, \dots, n$, também chamada de resposta;

β_0 = intercepto ou termo independente de variável;

β_1 = inclinação de Y em relação a variável X_1 , mantendo constante as variáveis X_2, X_3, \dots, X_k ;

β_2 = inclinação de Y em relação a variável X_2 , mantendo constante as variáveis X_1, X_3, \dots, X_k ;

β_k = inclinação de Y em relação a variável X_k , mantendo constante as variáveis X_1, X_3, \dots, X_{k-1} ;

ε_i = erro aleatório em Y, para a observação $i, i=1, 2, \dots, n$.

O modelo de regressão linear clássico (1) depende de algumas suposições que descrevem as formas do modelo e também ditam os procedimentos adequados de estimação e inferência:

I – A variável dependente y_i é função linear das variáveis independentes X_{ki} ($i= 1, 2, \dots, k$);

II – os valores das variáveis independentes são fixos;

III – $E(\varepsilon_i) = \phi$, onde, ϕ representa um vetor de zeros;

IV – os erros são homocedásticos, isto é, $E(\varepsilon_i)^2 = \sigma^2$;

V – os erros são independentes e têm distribuição normal.

A NBR 14.653-3:2004 (ABNT, 2004), no seu Anexo A, além desses, impõe outros condicionantes para a construção de modelos com uso de regressão linear múltipla, sejam eles: número mínimo de dados efetivamente utilizados (n) com respeito ao número de variáveis independentes (k), para evitar a micronumerosidade; investigar a correlação entre as variáveis independentes para evitar a multi-colinearidade; e, examinar a possível presença de pontos influenciantes.

A estimativa dos parâmetros foi realizada pelo método dos mínimos quadrados, na forma explicitada abaixo de acordo com REIS (2000):

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + e_i, \rightarrow i = 1, 2, \dots, n \quad (2)$$

Fazendo i variar de 1 até n , obtemos as n equações seguintes:

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \dots + \beta_k x_{k1} + e_1 \\ Y_2 &= \beta_0 + \beta_1 x_{12} + \beta_2 x_{22} + \dots + \beta_k x_{k2} + e_2 \\ &\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ Y_n &= \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} + \dots + \beta_k x_{kn} + e_n \end{aligned} \quad (3)$$

Essas equações podem ser apresentadas compactamente na notação matricial por:

$$Y = X\beta + \varepsilon \quad (4)$$

onde:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}; X = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix}; \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}; \varepsilon = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (5)$$

Y é um vetor $nx1$ de observações aleatórias;
X é uma matriz $nx(k+1)$ de quantidades fixas conhecidas;
 β é um vetor $(k+1)x1$ de parâmetros desconhecidos;
 ε é um vetor aleatório $nx1$.

Temos da equação (4) que:

$$\varepsilon = Y - X\beta \quad (6)$$

Minimizando a soma dos quadrados dos erros por $\varepsilon'\varepsilon$:

$$\varepsilon'\varepsilon = [e_1 \quad e_2 \quad \cdots \quad e_n] \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \sum_{i=1}^n e_i^2 \quad (7)$$

Seja:

$$\begin{aligned} Z &= \varepsilon'\varepsilon \\ Z &= (Y - X\beta)'(Y - X\beta) \\ Z &= (Y' - X'\beta')(Y - X\beta) \\ Z &= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta \end{aligned}$$

Sendo as matrizes $Y'X\beta$ e $\beta'X'Y$ de dimensões $1x1$ e uma sendo a transposta da outra, temos:

$$Y'X\beta = \beta'X'Y$$

logo:

$$Z = Y'Y - 2\beta'X'Y + \beta'X'X\beta$$

A função Z apresenta ponto de mínimo para valores de β que tornem sua diferencial identicamente nula:

Diferenciando Z, vem:

$$\partial Z = -2(\partial\beta')X'Y + (\partial\beta')(X'X\beta) + \beta'X'X(\partial\beta)$$

Sendo $(\partial\beta')(X'X\beta) = \beta'X'X(\partial\beta)$ por serem matrizes de dimensões 1x1 e uma ser transposta da outra:

$$\partial Z = -2(\partial\beta')X'Y + 2(\partial\beta')X'X\beta = 2(\partial\beta')(X'X\beta - X'Y)$$

Fazendo $\partial Z \equiv \phi$, temos:

$$2\partial\hat{\beta}'(X'X\hat{\beta} - X'Y) \equiv \phi; \text{ como } \partial\hat{\beta}' \neq \phi, \text{ temos:}$$

$$X'X\hat{\beta} - X'Y = \phi \Rightarrow X'X\hat{\beta} = X'Y \quad (8)$$

Onde $\hat{\beta}$ é o vetor das estimativas dos parâmetros, isto é:

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \quad (9)$$

O sistema de equações (8) é denominado sistema de equações normais, e sua solução nos fornece as estimativas dos parâmetros constituintes do vetor $\hat{\beta}$.

$X'X$ é não singular, logo existe a matriz $(X'X)^{-1}$, pré-multiplicando ambos os membros da equação (8) por $(X'X)^{-1}$, temos:

$$(X'X)^{-1}(X'X\hat{\beta}) = (X'X)^{-1}(X'Y)$$

Fazendo $X'X = S$, temos:

$$S^{-1}S\hat{\beta} = S^{-1}X'Y$$

Portanto:

$$\hat{\beta} = S^{-1}X'Y \quad (10)$$

A primeira etapa dos cálculos para a obtenção dos parâmetros foi a construção das matrizes:

$$S = X'X$$

$$S = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ X_{11} & X_{12} & X_{13} & \cdots & X_{1n} \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{k1} & X_{k2} & X_{k3} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} 1 \\ X_{11} \\ X_{12} \\ X_{13} \\ \vdots \\ X_{1n} \end{bmatrix} \begin{bmatrix} X_{21} \\ X_{22} \\ X_{23} \\ \vdots \\ X_{2n} \end{bmatrix} \begin{bmatrix} \cdots \\ X_{k2} \\ X_{k3} \\ \vdots \\ X_{kn} \end{bmatrix} \quad (11)$$

$$S = \begin{bmatrix} n & \sum X_{1i} & \sum X_{2i} & \cdots & \sum X_{ki} \\ \sum X_{1i} & \sum X_{1i}^2 & \sum X_{1i} X_{2i} & \cdots & \sum X_{1i} X_{ki} \\ \sum X_{2i} & \sum X_{1i} X_{2i} & \sum X_{2i}^2 & \cdots & \sum X_{2i} X_{ki} \\ \vdots & \vdots & \vdots & & \vdots \\ \sum X_{ki} & \sum X_{1i} X_{ki} & \sum X_{2i} X_{ki} & \cdots & \sum X_{ki}^2 \end{bmatrix}$$

A matriz é simétrica, isto é, a primeira linha é igual primeira coluna, a segunda linha é igual a segunda coluna, e por fim, a linha ordem (k+1) é igual a coluna ordem (k+1); então $S = S^{-1}$

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ X_{11} & X_{12} & X_{13} & \cdots & X_{1n} \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{k1} & X_{k2} & X_{k3} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{1i} Y_i \\ \sum X_{2i} Y_i \\ \vdots \\ \sum X_{ki} Y_i \end{bmatrix} \quad (12)$$

A análise de variância da regressão:

A soma de quadrados do resíduo é dada por:

$$SQR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (13)$$

isto é, a soma de quadrado dos desvios entre os vetores observados e os estimados pela equação de regressão.

De forma matricial, temos:

$$SQR = \hat{\varepsilon}' \hat{\varepsilon} \quad (14)$$

$$\hat{\varepsilon} = Y - X\beta \Rightarrow \hat{\varepsilon}' = Y' \hat{\beta}' X'$$

$$SQR = (Y' - \hat{\beta}' X') (Y - X\hat{\beta}) = Y'Y - \hat{\beta}' X'Y - Y'X\hat{\beta} + \hat{\beta}' X'X\hat{\beta}$$

Sendo as matrizes $Y'Y\hat{\beta}$ e $\hat{\beta}' X'Y$ de dimensões 1x1 e sendo uma a transposta da outra, temos:

$$Y'X\hat{\beta} = \hat{\beta}' X'Y$$

logo:

$$SQR = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'Y$$

$$SQR = \left[Y'Y - \frac{(\sum Y_i)^2}{n} \right] - \left[\hat{\beta}'X'Y - \frac{(\sum Y_i)^2}{n} \right] \quad (15)$$

$$SQR = SQTotal - SQRegressão \quad (16)$$

$$SQTotal = Y'Y - C \quad (17)$$

$$SQRegressão = \hat{\beta}'X'Y - C \quad (18)$$

$$C = \frac{(\sum Y_i)^2}{n} \quad (19)$$

Tabela 1- Tabela de análise de variância

Fonte de variação	Graus de liberdade	Soma de quadrados	Quadrado médio	F
<i>Regressão</i>	<i>K</i>	<i>SQRegressão</i>	<i>QMRegressão= SQRegressão/k</i>	<i>QMRegressão/QM R</i>
<i>Ind. Da Regressão</i>	<i>n-k-1</i>	<i>SQR</i>	<i>QMR=SQR/(n-k-1)</i>	
<i>Total</i>	<i>n-1</i>			

Foram utilizados para validação do modelo:

a) **Coefficiente de correlação linear múltiplo (R) e coeficiente de determinação (R²):** o coeficiente de correlação traduz numericamente o quanto as variáveis estão linearmente relacionadas entre si (NADAL *et. al.* 2003). É fornecido matricialmente pela raiz quadrada da expressão:

$$R^2 = \frac{\hat{\beta}'X'Y - C}{Y'Y - C} \quad (20)$$

O valor de R encontra-se no intervalo de $-1 \leq R \leq 1$, já o coeficiente de determinação indica numericamente o valor de avaliações que está sendo explicado pelo modelo, encontra-se entre o intervalo de $0 \leq R^2 \leq 1$.

Segundo BAPTISTELLA (2005), a desvantagem do coeficiente de determinação é que para uma mesma amostra ele cresce na medida em que aumentam o número de variáveis independentes incluídas no modelo, não levando em conta o número de graus de liberdade perdidos a cada parâmetro estimado. Para corrigir esta deficiência é preferível utilizar o conceito de Coeficiente de Determinação Ajustado, que para um modelo com *k* variáveis independentes, ajustado a uma amostra de *n* elementos, é calculado através da seguinte expressão:

$$R_{aj}^2 = 1 - \left[(1 - R^2) \frac{n-1}{n-k-1} \right] \quad (21)$$

b) **Teste de existência da regressão:** consistiu em se estudar a probabilidade dos parâmetros de regressão $\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k$ serem iguais a zero o que, neste caso, não existiria regressão. O teste é efetivado através da distribuição de Fischer-Snedecor. O coeficiente F calculado é obtido pela expressão matricial:

$$F_c = \frac{\hat{\beta}' X' Y - C}{(Y' Y - C) - (\hat{\beta}' X' Y - C)} \times \frac{n-k-1}{k} \quad (22)$$

A hipótese básica é aceita, ou seja, de que haja regressão de y em $X_1, X_2, X_3, \dots, X_k$, com nível de significância $\alpha=1\%$, no caso deste trabalho, se:

$$F_c > F \quad (23)$$

c) **Teste da significância dos regressores:** utilizou-se neste caso a distribuição T de Student (unicaldal), tendo como hipótese básica que os regressores são diferentes de zero ao nível de significância $\alpha=5$ e 10% . Os valores calculados para o modelo de melhor ajuste são dados pela expressão que relaciona cada regressor com seu desvio padrão:

$$T_i = \frac{\hat{\beta}_i}{s(\hat{\beta}_i)} \quad (24)$$

com base nos valores obtidos da distribuição de Student (T) com entradas 2α e graus de liberdade $n-k-1$, aceita-se a hipótese básica que cada um dos coeficientes β_i são diferentes de zero se:

$$T_i > T \quad (25)$$

d) **Teste de Durbin-Watson:** existe autocorrelação serial quando os termos de resíduo são correlacionados com os valores anteriores ou posteriores da mesma série. A má especificação do modelo de regressão, em função de resíduos na forma do modelo ou por exclusão de variáveis independentes importantes para a análise é uma das causas de autocorrelação (NADAL *et. al.* 2003). Isto ocorre principalmente em aplicações envolvendo séries temporais (JOHNSTON, 1977). Segundo DANTAS (2005), o conceito de independência dos resíduos está ligado à independência dos dados de mercado. A situação ideal é aquela onde cada transação se realiza independentemente da outra. Isto é, o conhecimento do preço e condições de uma não interfira na outra. A existência de autocorrelação serial nos termos de perturbação aleatória pode ser verificada com o auxílio da razão de Von Neumann, ou estatística de Durbin-Watson (IMAPE, 1998), onde a hipótese básica é a

existência de autocorrelação entre os resíduos, que pode ser calculada pela expressão:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (26)$$

onde e_i é o i -ésimo resíduo do modelo, ordenado crescentemente em relação aos valores ajustados, considerando-se uma amostra de tamanho n .

Da estatística de Durbin-Watson para o nível de confiança α com graus de liberdade $n-k-1$, obtêm-se du que é o limite superior de variação e di , o limite inferior, assim: se $d < du < 4-du$ a hipótese é aceita, se $4-di < d < di$ é rejeitada e se ocorrer que $di < d < du$ e $4-du < d < 4-di$, o teste é inconclusivo.

e) **Análise gráfica dos resíduos:** ao se colocar em um gráfico os resíduos e as variáveis explicativas é possível a verificação da existência de uma multicolinearidade, ou seja, uma relação exata entre as variáveis, se o coeficiente de correlação apresenta-se muito próximo da unidade. Se o gráfico demonstrar que os resíduos não estão alinhados então a correlação é meramente casual e os resíduos não mostram nenhuma tendência. Ao se analisar graficamente as distribuições (resíduos versus valor estimado) verifica-se a existência de homocedasticidade, ou seja, a hipótese de variância constante, que é aceita quando não há nenhuma tendência dos resíduos em relação ao valor estimado, neste caso denominado de heterocedasticidade (NADAL *et. al.* 2003). Outro artifício utilizado para testar a hipótese de homocedasticidade ou heterocedasticidade é o teste de Pesaran-Pesaran (PEIXOTO, 2007), em que é testada a existência de regressão entre os quadrados dos resíduos e_i^2 e os valores preditos \hat{y}_i pelo modelo ao nível de significância $\alpha=5\%$, para o caso deste estudo. Analisando graficamente os resíduos padronizados e_i^* , encontrados dividindo-se cada resíduo e_i pelo desvio padrão do modelo, conforme expressão (27), é possível inferir se os erros têm distribuição normal. A hipótese de normalidade pode ser confirmada se 68% dos resíduos estão no intervalo $[-1;+1]$, 90% entre $[-1,64;+1,64]$ e 95% entre $[-1,96;+1,96]$. A possível presença de pontos atípicos ou outlier pode ser detectada através de análise gráfica e o avaliador poderá estudar a possibilidade de sua retirada do modelo, haja vista que os mesmos exercem forte influência na estimação dos parâmetros da regressão.

A condição de comportamento linear das variáveis independentes em relação a Y_i , também pode ser avaliada analisando os gráficos de $Y_i=f(X_k)$, indicando também qual a melhor opção de linearização, no caso de os dados assumirem outra tendência.

O desvio padrão do modelo geral, com parâmetros ajustados a n dados Y_i por uma equação de média \hat{Y}_i , é dado pela expressão (DANTAS, 2003):

$$s_e = \sqrt{\frac{\sum (Y_i - \hat{Y}_i)^2}{n - p}} \quad (27)$$

onde o denominador $(n - p)$ representa o número de graus de liberdade do modelo, pois cada parâmetro estimado provoca a perda de um grau de liberdade.

f) **Teste de Kolmogorov-Smirnov:** este teste avalia se duas amostras tem distribuições semelhantes, ou melhor dizendo, se foram extraídas de uma mesma população. Se apresentarem grandes diferenças provavelmente estas não se devem ao acaso. É um teste que detecta diferenças em relação à tendência central, dispersão e simetria (NADAL *et. al.* 2003). A violação da normalidade pode estar ligada a alguns aspectos relacionados ao modelo, tais como: omissão de variáveis explicativas importantes, inclusão de variável explicativa irrelevante para o modelo, utilização de relação matemática incorreta para análise entre as variáveis do modelo (PEIXOTO, 2007). O teste não paramétrico de Kolmogorov-Smirnov foi utilizado para avaliação da normalidade, e a estatística de teste é dada por:

$$d_{\max} = \frac{y_i}{n} - Z_i \quad (28)$$

onde: n = tamanho da amostra, sendo $i=1,2,3,\dots,n$; Z_i = probabilidade acumulada da distribuição normal padronizada, considerando os valores de $h_i = (e_i/s)$, onde e_i são os resíduos ordenados de forma crescente e s é o desvio padrão dos e_i .

Aceita-se a hipótese de que os resíduos se distribuem normalmente se $d_{\max} < d_{\text{crítico}}$. O valor de $d_{\text{crítico}}$ é obtido a partir de tabela, com nível de significância de 5% e $n=32$ (tamanho da amostra). O valor de Z_i foi obtido usando a função estatística do Microsoft Office Excel $\text{dist.normP}(h_i)$.

Intervalo de confiança

Para estimativa de um bem avaliando, ou de um paradigma adotado, que possui características $X_{h1}, X_{h2}, \dots, X_{hk}$. Como \hat{Y}_i é um estimador não-tendencioso para $E(Y_i)$, é bastante que se substituam no modelo as características do bem a avaliar, obtendo-se a estimativa pontual correspondente (DANTAS, 2005). O intervalo de confiança a um nível $1-\alpha$ para \hat{Y}_h é dado por:

$$\left[\hat{Y}_h - t_{(1-\alpha/2; n-k-1)} \times s(\hat{Y}_h); \hat{Y}_h + t_{(1-\alpha/2; n-k-1)} \times s(\hat{Y}_h) \right] \quad (29)$$

De forma matricial, o desvio padrão de Y_h , para um imóvel com uma matriz de características $X'_h = [1 \ X_{h1} \ X_{h2} \ \dots \ X_{hk}]'$, é dada por:

$$S(\hat{Y}_h) = S_e \sqrt{\frac{1}{n} \times X'_h (X'X)^{-1} X_h} \quad (30)$$

onde, S_e é dado pela expressão (27).

Variáveis envolvidas na modelagem

Os preços dos imóveis são valores reais estritamente positivos, enquanto que o campo da distribuição normal é todo o intervalo dos reais, isto faz com que esta não seja apropriada para dados imobiliários. Segundo DANTAS (2005), uma solução que tem sido utilizada é a mudança para a escala logarítmica, uma vez que o logaritmo do preço abrange a reta real, a distribuição log-normal tem se mostrado bastante aderente a dados imobiliários.

a) Variável dependente:

Valor do imóvel [Y]: logaritmo na base e do valor unitário do imóvel, calculado por hectare – VTI(R\$)/ha.

b) Variáveis independentes utilizadas no modelo:

Data da transação ou coleta da informação [X₁]: logaritmo na base e da data de coleta da informação da oferta ou negociação, em valor numérico relativo a contagem de dias desde 1º de janeiro de 1900 (a transformação de data em valor numérico é feita facilmente utilizando o aplicativo Microsoft Office Excel). As informações foram coletadas em 2001, 2002, 2005, 2007 e 2008.

Infraestrutura [X₂]: foram associados valores de 0 a 3 em função do nível de infraestrutura. Os imóveis utilizados com agricultura irrigada normalmente possuem infraestrutura como energia elétrica, casa de morador, galpões, cercas e equipamentos de irrigação. Alguns imóveis utilizados neste trabalho não possuíam nenhuma benfeitoria, embora tivesse potencial para irrigação.

Tabela 2- Nível de infraestrutura

Nível de infraestrutura	Valor	
Inexistente	0	-
Mínima	1	Equipamento de irrigação por aspersão convencional ou por superfície
Regular	2	Equipamento de irrigação por aspersão convencional ou microaspersão, casa de morador, cercas, drenos
Boa	3	Equipamento de irrigação por microaspersão e/ou gotejamento, casa de morador, galpões, cercas, drenos, outras instalações

Obs.: Todas os imóveis utilizados na modelagem tinham rede de energia elétrica

Elasticidade da oferta [X₃]: foram atribuídos valor 0 para os imóveis efetivamente negociados e 1 para imóveis ofertados.

Área do imóvel [X₄]: logaritmo na base e da área do imóvel em hectare. A amostra contém imóveis de 5 a 975 ha.

Percentual da área do imóvel com potencial para irrigação [X₅]: logaritmo na base e do percentual da área do imóvel com potencial para irrigação. O potencial do imóvel para irrigação foi estimado com base no potencial dos solos, e a existência de fonte perene de água com volume e qualidade adequados para utilização. As amostras utilizadas neste trabalho estão localizadas em áreas próximas ao Rio São Francisco ou em perímetros irrigados públicos administrados pela Companhia de Desenvolvimento dos Vales do São Francisco e Parnaíba – CODEVASF.

Produção vegetal [X₆]: logaritmo na base e da área equivalente de produção vegetal obtida pelo produto da área cultivada em hectare acrescida da unidade e o índice de valoração da produção vegetal.

O valor econômico das produções vegetais foi calculado através do método da capitalização da renda para cada ano de vida útil da cultura, conforme expressão:

$$Ve = RL \times Fa(1 - r) \quad (31)$$

$$Fa = \frac{(1 + i)^n - 1}{(1 + i)^n \times i} \quad (32)$$

onde,

Ve é o valor econômico;

RL é a renda líquida obtida para uma safra média, renda bruta deduzida dos custos diretos e indiretos;

r é a taxa de risco do empreendimento;

Fa é o fator de valor presente para pagamento único para uma série uniforme de rendimentos futuros para *n* períodos de capitalização; e,

i é a taxa de capitalização anual.

As variáveis envolvidas na estimativa do valor econômico podem sofrer variação em função das tendências de mercado, como aumentos dos preços de insumos e serviços, variação de preços de venda da produção, quedas na produção, alterações de câmbio, além de turbulências como crises econômicas internacionais. Para minimizar os efeitos acima, ao menor valor econômico foi atribuído à unidade e de forma proporcional aos demais valores relativos às culturas e ao tempo de vida útil (índice de valoração da produção vegetal) (Tabela 3).

Tabela 3- Índice de valoração da produção vegetal em função do tempo de vida útil das culturas perenes comumente cultivadas na microrregião de Petrolina

Cultura	Vida útil (anos)								
	17	16	15	14	13	12	11	10	9
Acerola					16,0375	15,4653	14,8245	14,1068	13,3029
Coco	8,5405	8,3658	8,1702	7,9510	7,7055	7,4306	7,1227	6,7779	6,3916
Goiaba					28,2207	27,2139	26,0862	24,8233	23,4087
Manga	7,9740	7,8108	7,6282	7,4235	7,1944	6,9377	6,6502	6,3282	5,9676
Uva de mesa					68,1627	65,7309	63,0072	59,9567	56,5401
Uva s/ semente					38,6943	37,3138	35,7676	34,0359	32,0964

Cultura	Vida útil (anos)							
	8	7	6	5	4	3	2	1
Banana					5,2360	4,1404	2,9134	1,5392
Acerola	12,4026	11,3942	10,2649	9,0000	7,5833	5,9966	4,2195	2,2292
Coco	5,9591	5,4746	4,9320	4,3242	3,6435	2,8812	2,0273	1,0711
Goiaba	21,8244	20,0501	18,0627	15,8369	13,3441	10,5520	7,4249	3,9226
Mamão						2,7885	1,9622	1,0366
Manga	5,5637	5,1114	4,6048	4,0373	3,4018	2,6900	1,8929	1,0000
Maracujá						3,6880	2,5950	1,3710
Uva de mesa	52,7135	48,4278	43,6277	38,2517	32,2305	25,4868	17,9338	9,4744
Uva s/ semente	29,9242	27,4912	24,7664	21,7145	18,2964	14,4682	10,1806	5,3784

RESULTADOS E DISCUSSÃO

As hipóteses formuladas foram aceitas, exceto para influência das classes de capacidade de uso da terra e condições de acesso ao imóvel, que compõe a nota agrônômica, calculada pela escala de França (1983) citada por INCRA (2006). Esta variável não possui importância significativa no modelo provavelmente devido à condição de semelhança quanto às propriedades do solo das amostras utilizadas no modelo, ou a escala utilizada para cálculo da nota agrônômica não foi a adequada. Os solos com aptidão para irrigação estão associados às melhores classes de capacidade de uso, e o percentual deles de forma crescente exerce influência positiva na variável dependente valor da terra. Então, mesmo de forma indireta as classes de capacidade de uso da terra exercem sim influência sobre a variável dependente.

A matriz dos regressores encontrada pelo método dos mínimos quadrados foi:

$$\hat{\beta} = \begin{bmatrix} -108,3027035 \\ 10,86709827 \\ 0,08789898 \\ 0,542987955 \\ -0,348251052 \\ 0,498900517 \\ 0,120011808 \end{bmatrix} \quad (33)$$

ou seja, o modelo de regressão com eixos transformados resultante foi:

$$Y = e^{-108,3027035 + 10,86709827 \ln X_1 + 0,08788998 X_2 + 0,542987955 X_3 - 0,348251052 \ln X_4 + 0,498900517 \ln X_5 + 0,120011808 \ln X_6} \quad (34)$$

e o modelo resultante com eixos não transformados foi:

$$Y = 9,22005 \times 10^{-48} \times X_1^{10,86709827} \times e^{0,08788998 X_2} \times (0,7211 X_3 + 1) \times X_4^{-0,348251052} \times X_5^{0,498900517} \times X_6^{0,120011808} \quad (35)$$

O coeficiente de determinação (R^2) foi de 0,98285091, ou seja, 98,28% do valor de mercado dos imóveis está sendo explicado pelo modelo, e o coeficiente de determinação ajustado foi de (R^2_{aj}) foi de 0,978936593. O coeficiente de correlação linear múltiplo foi de 0,991388971, indicando uma correlação fortíssima entre as variáveis explicativas e a explicada. O desvio padrão global foi de 0,177173215.

Pela Tabela 4, observa-se a análise de variância, onde o coeficiente F calculado foi maior que o tabelado, aceitando-se a hipótese de existência de regressão.

Tabela 4- Análise de variância

CV	GL	SQ	MQ	F	Significação
Regressão	6	44,9793461	7,496557684	238,817283	**
Resíduo	25	0,784758706	0,031390348		
Total	31	45,76410481			

CV= causa da variação; GL= grau de liberdade; SQ= soma de quadrados; MQ= quadrado médio; F= coeficiente F calculado; ** significativo ao nível de 1%.

Os regressores foram submetidos ao teste T de Student, onde todos foram diferentes de zero ao nível de significância $\alpha=5\%$, exceto para infraestrutura que foi diferente de zero apenas a $\alpha=10\%$ (Tabela 5).

Tabela 5- Coeficientes de regressão do modelo, desvio padrão dos regressores e estatística T de Student (Stat t)

	Coeficientes	Erro padrão	Stat t	Signif.
Interseção (X_0)	-108,3027035	15,7474886	-6,877458764	*
Ln dia (X_1)	10,86709827	1,482388339	7,330803933	*
Infraestrutura (X_2)	0,08788998	0,048046538	1,829455026	“**”
Elasticidade (X_3)	0,542987955	0,072542895	7,485060418	*
Ln área (X_4)	-0,348251052	0,028389154	-12,26704598	*
Ln % área irrigável (X_5)	0,498900517	0,072984936	6,83566425	*
Ln produção vegetal (X_6)	0,120011808	0,020135768	5,960130544	*

* significativo ao nível de 5%; “**” significativo ao nível de 10%.

O teste de Durbin-Watson se mostrou inconclusivo para a hipótese de existência de autocorrelação entre os resíduos. A estatística d calculada foi igual a 2,39036, situado entre os valores de $4-du$ e $4-di$, sendo du e di , para k variáveis explicativas e n amostras a 5% de significância, iguais a 1,05 e 1,92, respectivamente. Segundo Cuthbertson *et al.* (1992) citado por GONZÁLEZ & FORMOSO (2000), é comum em modelos econômicos, principalmente com análise

de séries temporais, o aparecimento de relações seriais entre as medidas dos resíduos, provocando dificuldade de análise, e ocorrendo autocorrelação, os estimadores obtidos por Mínimos Quadrados são não-viesados, mas são ineficientes, e os modelos não são plenamente válidos, havendo restrições para serem empregados na inferência de valores. Quase todas as anomalias estatísticas para obtenção de um bom modelo podem ser controladas, mas no caso da autocorrelação permanecem as dificuldades. Como mencionado acima, a hipótese de autocorrelação não foi rejeitada, entretanto, também não foi aceita. A autocorrelação serial implica na existência de uma relação onde um erro ε_t , que pode ser determinado em função dos erros anteriores, que por sua vez, deveria ser aleatório, com média nula e desvio padrão constante (Dubin, 1998; Maddala, 1988, citados por GONZÁLES & FORMOSO, 2000). Entretanto, a análise dos erros através da Figura 3 indica que os mesmos são aleatórios, o que poderia indicar a ausência de autocorrelação serial.

A análise de correlação, de modo geral, revelou fraca correlação entre as variáveis independentes. No entanto, para X_4 versus X_5 a correlação foi inversa e acima de 0,5, o que segundo DANTAS (2005), merece maior atenção já que gera degeneração do modelo e limita sua utilização (Tabela 6). Todavia, a dimensão da área do imóvel é, na amostra utilizada, normalmente inversa ao percentual de área irrigável, pois os imóveis pequenos são, na maioria, lotes em projetos de irrigação públicos sendo sua área, na totalidade ou quase toda ela, utilizada para irrigação. Por outro lado, as áreas de maior dimensão, localizadas às margens do Rio São Francisco, possuem parte da área sem nenhuma condição de aproveitamento com esta atividade, seja por limitações edáficas ou legais, limitando desta forma o uso para irrigação.

A análise dos gráficos da distribuição dos resíduos e as variáveis independentes X_4 e X_5 não revelou qualquer tendência, o que indica que a correlação é meramente casual, afastando a hipótese da existência de colinearidade (Figura 1).

Tabela 6- Correlação entre o logaritmo na base e das variáveis envolvidas no modelo; valor da terra (VTI/ha), dia da transação (dia), nível de infraestrutura (infraest.), elasticidade da oferta (elast), dimensão da área (área), percentual de área irrigável (% área irrig.), e produção vegetal (prod. veg.)

	Ln VTI/ha (Y)	Ln dia (X ₁)	infraest. (X ₂)	elast (X ₃)	Ln área (X ₄)	Ln % área irrig. (X ₅)	Ln prod. veg. (X ₆)
Ln VTI/ha (Y)	1,0000						
Ln dia (X ₁)	0,5237	1,0000					
infraest. (X ₂)	0,3549	-0,0195	1,0000				
elast (X ₃)	0,4711	0,3667	-0,0533	1,0000			
Ln área (X ₄)	-0,8105	-0,2965	-0,1519	-0,1529	1,0000		
Ln % área irrig. (X ₅)	0,7522	0,0419	0,4634	0,0728	-0,6487	1,0000	
Ln produção veg. (X ₆)	0,6112	0,2953	0,4486	0,3588	-0,1792	0,4387	1,0000

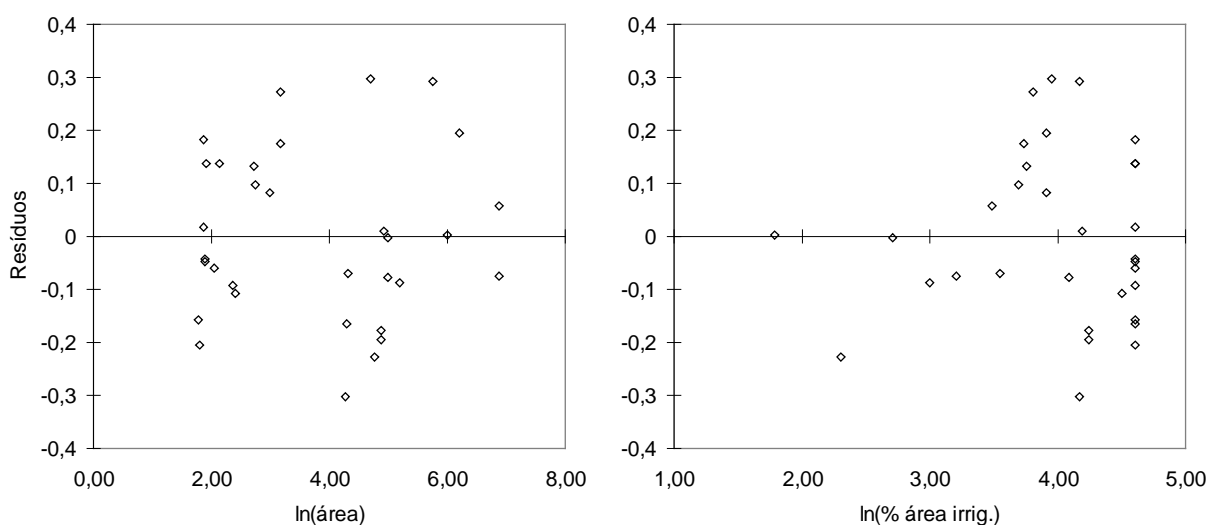


Figura 1- Distribuição dos resíduos em função do logaritmo na base e da área e do percentual de área com potencial para irrigação.

A análise gráfica da distribuição dos resíduos em função dos valores estimados pelo modelo não revelou a existência de nenhuma tendência, indicando a hipótese de variância constante ou homocedasticidade (Figura 2). O teste de Pesaran-Pesaran revelou a ausência de heterocedasticidade, onde os valores dos quadrados dos resíduos (e^2) versus os valores estimados pelo modelo estão dispersos, rejeitando-se a existência de regressão pelo teste F ao nível de 5% (Figura 3). Também, na análise gráfica foi possível inferir que os erros padronizados estão distribuídos cerca de 62% entre -1 e 1 ; 88% entre $-1,64$ e $1,64$; e 100% entre $-1,96$ e $1,96$, portanto a distribuição é aproximadamente normal.

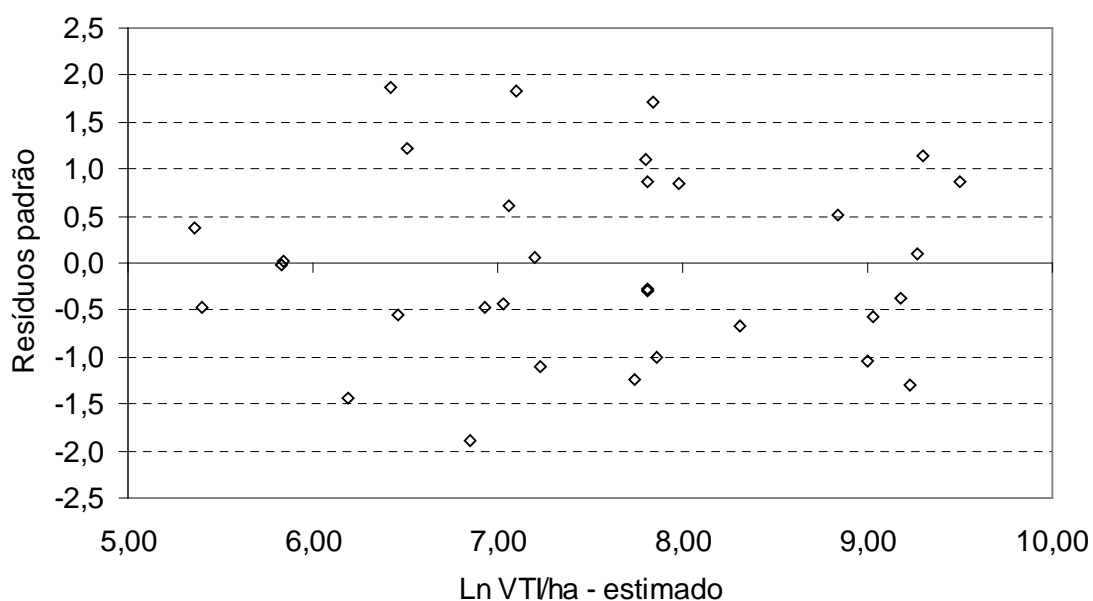
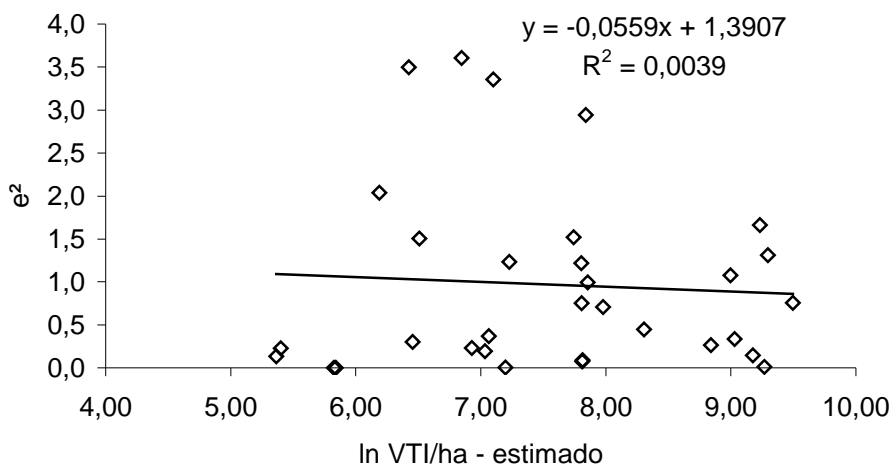


Figura 2- Distribuição dos resíduos padronizados em função dos valores estimados pelo modelo.

A hipótese de normalidade foi aceita com base na estatística d_{max} de Kolmogorov-Smirnov, onde $d_{max} < d_{crítico}$, sendo seus valores 0,183 e 0,234, respectivamente.

A Figura 4 mostra que os valores estimados estão dispostos em linha diagonal indicando uma linearidade satisfatória, estando os valores dispersos próximos à bisetriz do primeiro quadrante, indicando que as previsões se aproximam dos valores reais.



$\rho=0,73$, não significativo ao nível de 5%.

Figura 3- Diagrama dos resíduos versus os valores estimados pelo modelo.

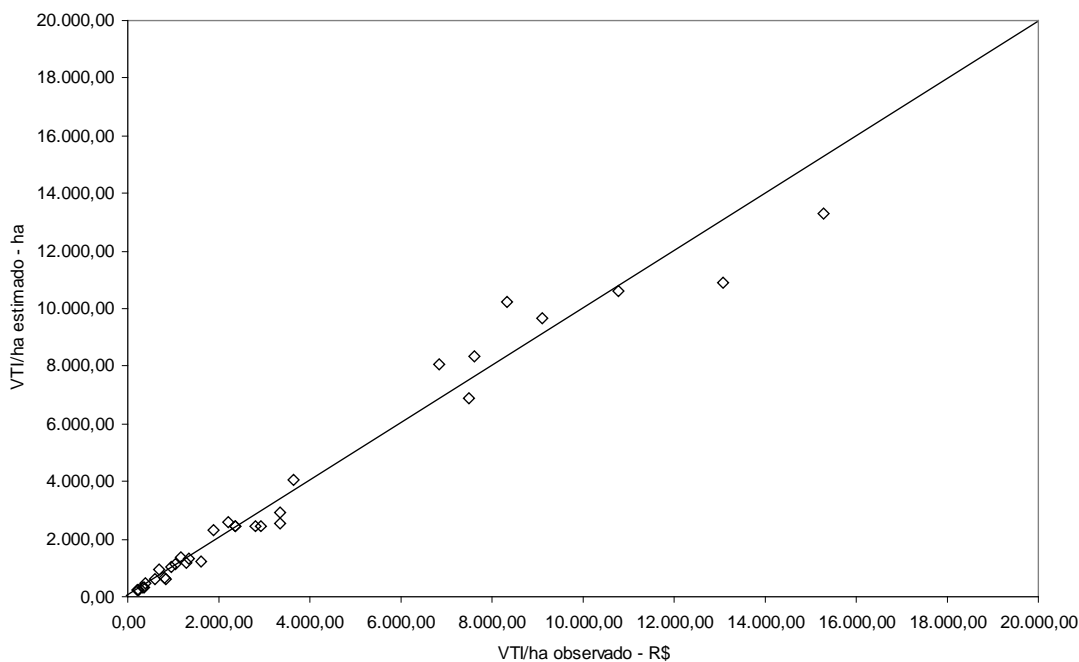


Figura 4- Relação entre os valores de imóveis irrigáveis (VTI/ha) estimados pelo modelo e observados na Microrregião de Petrolina/PE.

Pela Tabela 7, observa-se que em 19% dos casos, os erros dos valores preditos ficaram na faixa de 0 a 5% para mais ou para menos do valor observado e 60% dos casos ficaram entre 5 e 20% do valor observado. É importante observar que em apenas 3% dos casos o erro é superior a 30%, indicando o bom ajustamento do modelo em relação aos valores reais.

Tabela 7- Ocorrência de casos por faixa de erro no modelo de regressão

Erro	Nº de casos	%
0 a 5%	6	19
>5% a 10%	9	28
>10% a 15%	4	13
>15% a 20%	6	19
>20% a 25%	3	9
>25% a 30%	3	9
>30% a 36%	1	3
Total de casos	32	100

O valor da terra (VTI/ha) estimado pelo modelo variou em função do tempo do primeiro semestre de 2001 ao segundo semestre de 2008 em cerca de 115%, a uma taxa média semestral de 5,24%. Mantendo as demais variáveis explicativas fixas, a relação que representa a variação do VTI/ha estimado em função do tempo (valores numéricos ordenados desde 1º de janeiro de 1900) é a dada pela Figura 5. A taxa de valoração na data X_i pode ser obtida pela expressão:

$$\frac{dy}{dx} = 2,1734 \times 10^{-49} X_i^{9,867} \quad (36)$$

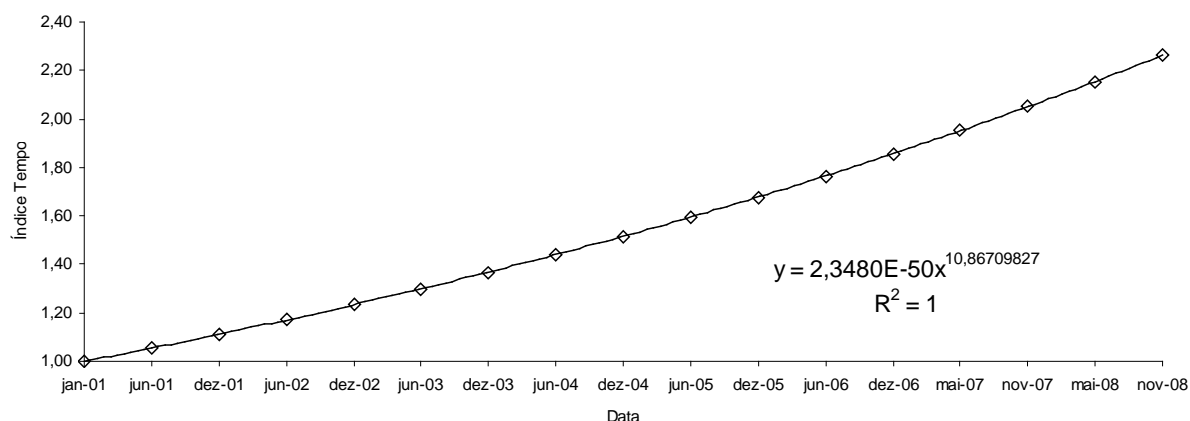


Figura 5- Variação do índice de valoração de imóveis irrigáveis em função do tempo de 2001 a 2008 para a Microrregião de Petrolina/PE.

Observa-se que o índice de valoração de imóveis irrigáveis mais que duplicou em oito (8) anos o que não foi confirmado por SILVA (2007), onde o mesmo não encontrou variação significativa do valor da terra nua (VTN/ha) em área de sequeiro entre os anos de 2001 a 2005 para as microrregiões do sertão pernambucano. Os

dados publicados pelo Instituto FNP (FNP, 2007) e (FNP, 2009) também indicam uma tendência de estabilidade nos preços de terra nua de áreas agrícolas de Projetos de Irrigação públicos e ribeirinhas entre 2004 e 2009. Nos valores do imóvel (VTI/ha) estão inclusos os valores relativos às benfeitorias existentes e as produções vegetais que sofrem influência direta das oscilações do mercado, principalmente o aumento de preços de materiais de construção, insumos agrícolas, mão de obra, combustíveis, etc. No mesmo período, segundo IBGE (2009), o índice para construção civil para o Estado de Pernambuco acumulou acréscimo de 114%, o índice nacional de preços ao consumidor amplo (IPCA) 71%, e o índice nacional de preços ao consumidor (INPC) 76%. Contudo, o crescente incremento do valor da terra ao longo do período estudado não foi resultado do aumento de índices inflacionários, haja vista que os valores relativos a infraestrutura normalmente não superaram 30% do valor final do imóvel, embora a área equivalente à produção vegetal, não presente em todas as amostras, possa alcançar valores que permitam sua duplicação. Então, o fator que provavelmente exerce maior influência sobre o valor dos imóveis, principalmente os que estão localizados em projetos públicos de irrigação, são os relativos à consolidação da agricultura irrigada no vale do São Francisco, sobretudo com as culturas da uva e manga, que respondem por grande parte das exportações de frutas do Brasil. Essa expansão se iniciou no final dos anos noventa e se intensificou desde então até a época atual. Segundo IBRAF (2009), o volume das exportações de uva sem sementes no Brasil cresceu cerca de 118% de 2003 a 2008 e os ganhos superaram 190%, e no município de Petrolina, segundo MDIC (2009), o volume exportado cresceu quatro (4) vezes e a receita em quase cinco (5) vezes de 2004 a 2008, embora os volumes de manga tenham se mantido estáveis, a alta rentabilidade da atividade agrícola-exportadora certamente contribuiu para o aquecimento do mercado de terras e sua valorização como ativo de produção.

A aplicação de fatores que explicam o comportamento do mercado de terras fora do intervalo estudado deve ser seguida de muitos cuidados, haja vista, que o que está influenciando o valor da terra não é o tempo e sim outras características inerentes ao mercado, como, nesse caso, a expansão da fruticultura e valorização dos produtos em função da demanda externa e não há garantias de que o comportamento destas características se mantenha nos mesmos patamares do período estudado.

Os valores estimados relativos a infraestrutura variaram de forma exponencial, conforme Figura 6, a uma taxa média de 9,19% a cada mudança de nível, ou seja, inexistente, mínima, regular e boa, indicando esta valoração a cada nível, cumulativamente. Uma melhor análise da taxa de variação a cada nível de infraestrutura pode ser conseguida derivando-se a função em X_i . Embora os índices de infraestrutura pareçam bem ajustados ao modelo, refletindo este ajuste para a amostra utilizada na modelagem, mas na prática podemos ter dificuldades com este fator, pois a infraestrutura para imóveis com irrigação pode assumir valores mais elevados devido à existência de instalações de alta tecnologia, sobretudo em imóveis de pequena dimensão, desta forma, o valor da infraestrutura pode ultrapassar o do imóvel. A infraestrutura necessária à unidade produtiva para atender as necessidades da exploração, não precisa necessariamente, ser de alta tecnologia, entretanto, há casos de produtores que empregam alto nível tecnológico e possuem inclusive unidades de processamento de frutas dentro do imóvel, o que

aumenta de forma considerável o valor da avaliação. Portanto, a utilização do modelo proposto deve seguir de cuidados, como os que acabamos de mencionar, além de outros aspectos relativos às características dos imóveis utilizados na modelagem.

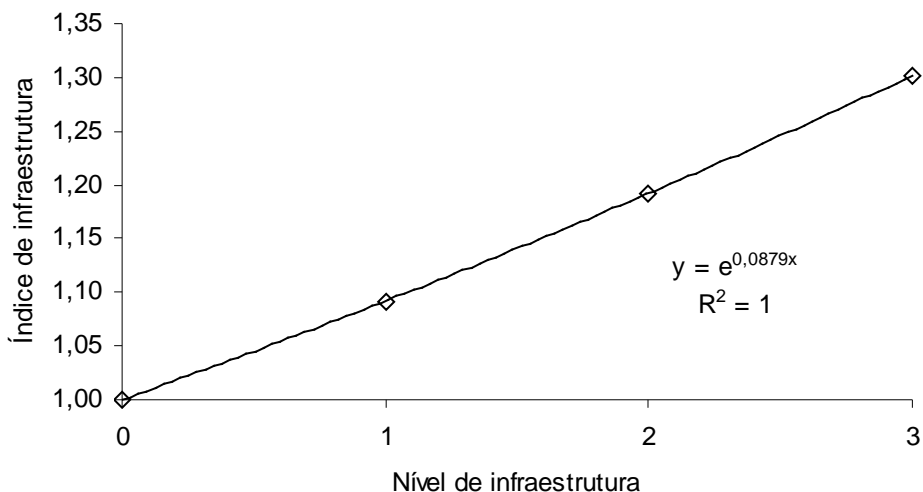


Figura 6- Índice de valoração da infraestrutura, calculado com base no VTI/ha estimado, em função do seu nível, para imóveis irrigáveis da Microrregião de Petrolina/PE.

A exponencial do coeficiente da variável explicativa X_3 fornece a valoração do imóvel ofertado em relação ao efetivamente negociado. Esta variação foi de 72%, ou seja, o negócio foi fechado em média por 58% dos preços de oferta. O fator encontrado de 0,58, foi menor que o proposto por SILVA (2007), que recomenda 0,75, no entanto, este fator foi estimado para imóveis sem nenhuma benfeitoria, indicando que o mesmo serve apenas para homogeneizar preços de terra nua (VTN), e a região de estudo foi muito mais ampla que a proposta nesta modelagem. Segundo LIMA (2005), este valor é menor também que as recomendações usuais para imóveis rurais que são de 80% do valor ofertado, embora sejam relatados que esses valores podem ser menores, entre 80 e 50% dos valores ofertados.

A relação entre o VTI/ha estimado e a área dos imóveis é inversa, confirmando a hipótese pré-formulada e a tendência geral dos avaliadores de que quanto maior a área menor seu valor unitário. Mantendo fixas as demais variáveis explicativas, a relação que representa o VTI/ha estimado e a área em hectare pode ser observada na Figura 7. As amostras usadas para construção do modelo possuem área variando de 5 a 975 ha, no entanto, a NBR: 14.653-3:2004 (ABNT, 2004) permite a extrapolação em 50%, ampliando a utilização do modelo para avaliação de imóveis com área de até aproximadamente 1.500 ha. A taxa de variação do valor do imóvel em função da sua área foi elevada para áreas até cerca de 100 ha, diminuindo após este valor até aproximadamente 500 ha, daí por diante assumindo tendência quase linear. As taxas de variação de y em X_i podem ser obtidas pela expressão:

$$\frac{dy}{dx} = -0,348251052 X_i^{-1,348251052} \quad (37)$$

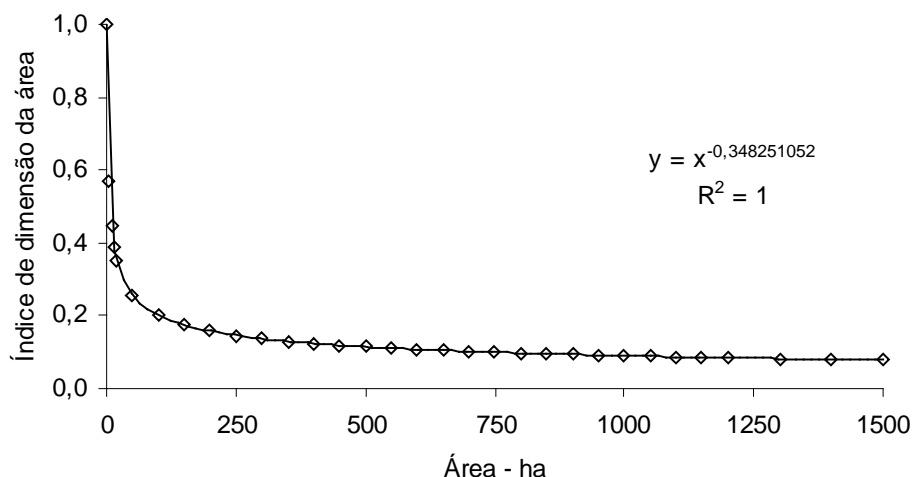


Figura 7- Índice de dimensão da área, calculado com base nos VTI/ha estimados pelo modelo, em função da área em hectare, para imóveis irrigáveis da Microrregião de Petrolina/PE.

Os valores da terra (VTI/ha) estimados pelo modelo crescem de maneira potencial em relação à percentagem de área com potencial para a prática de irrigação, tendo o índice de irrigação variando de 1 a aproximadamente 9,95, sendo este último para imóveis 100% irrigáveis (Figura 8). Os índices aqui encontrados são mais elevados que os propostos por SILVA (2007), que propôs uma relação linear com taxa angular de 0,0792 e a reta de regressão passando sobre a unidade no eixo da ordenada. Ressaltamos que os índices estudados por este autor são para estimativa do valor da terra nua utilizando amostras provenientes de ofertas ou negócios realizados entre os anos 2001 a 2005 na microrregião de Petrolina/PE.

O VTI/ha estimado pelo modelo aumentou de forma potencial também no caso da relação com a área equivalente da produção vegetal. O índice calculado com base nos VTI/ha estimados pelo modelo, cresceu de forma mais acentuada para áreas equivalentes inferiores a 100 ha, e de forma menos aguda a partir de cerca de 200 ha (Figura 9).

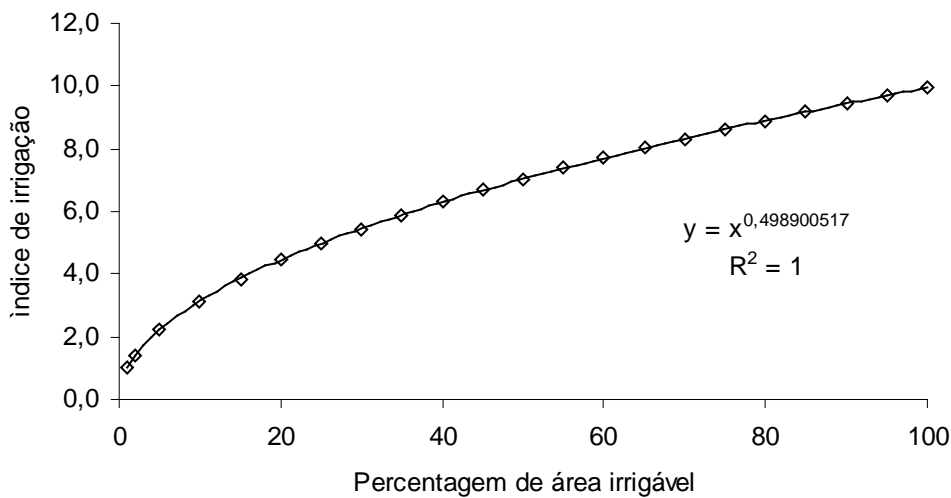


Figura 8- Índice de irrigação, calculado com base nos VTI/ha estimados pelo modelo, em função da percentagem de área com potencial para irrigação da microrregião de Petrolina/PE.

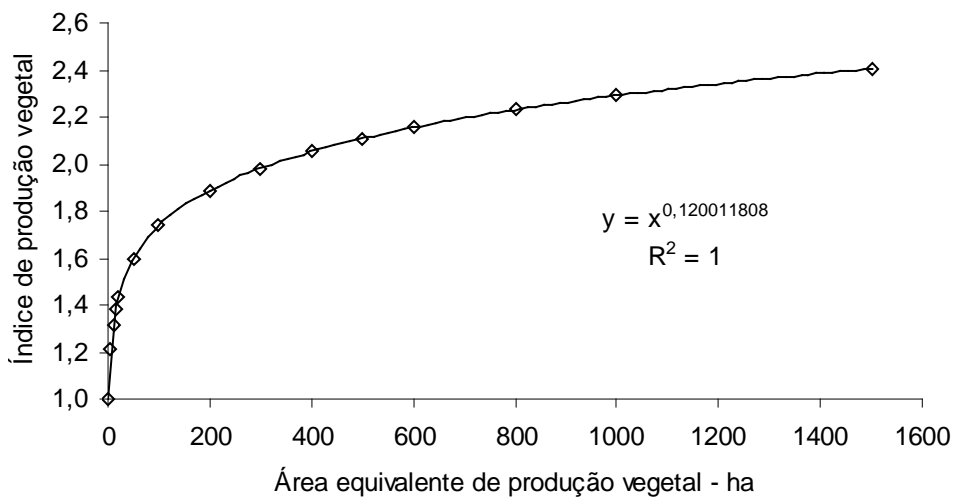


Figura 9- Índice de produção vegetal, calculado com base nos VTI/ha estimados pelo modelo, em função da área equivalente, obtida através do produto da área cultivada e os índices de valoração em função da cultura e do tempo de vida útil (Tabela 3).

A variação do índice de produção vegetal em relação à área equivalente é dada pela expressão:

$$\frac{dy}{dx} = 0,120011808 X_i^{-0,879988192} \quad (38)$$

Os índices obtidos neste estudo, com base na modelagem de dados por regressão linear múltipla, podem ser utilizados na avaliação de imóveis irrigáveis

também pelo processo estatístico descritivo de homogeneização por fatores (Anexo B da NBR: 14.653-3:2004) quando utilizado o método comparativo direto de dados de mercado, recomendando-se que todos os fatores sejam utilizados na mesma homogeneização. Isto se justifica pelo caráter multiplicativo do modelo, em que a estimativa da variável dependente é resultante do produto dos índices referentes a cada variável explicativa que estão interferindo no seu valor. Para facilitar a aplicação dos fatores sugerimos a adoção de intervalos de cada variável explicativa e seus índices médios correspondentes (Anexo I).

O intervalo de confiança onde está contido o VTI/ha estimado, e que através do qual pode ser obtido o campo de arbítrio para a avaliação, poderá ser calculado conforme expressão (29) e os componentes da fórmula obtidos pela (27) e (30). A matriz $S=X'X$ e sua inversa são:

$$S = X'X = \begin{bmatrix} 3556,39260 & 569,26257 & 179,36451 & 1271,27896 & 1340,06324 & 527,00205 \\ 569,26257 & 112 & 28 & 197,18617 & 223,39381 & 109,21048 \\ 179,36451 & 28 & 17 & 60,12027 & 68,04978 & 38,85577 \\ 1271,27896 & 197,18617 & 60,12027 & 539,02469 & 455,84514 & 168,23681 \\ 1340,06324 & 223,39381 & 68,04978 & 455,84514 & 521,20899 & 220,61840 \\ 527,00205 & 109,21048 & 38,85577 & 168,23681 & 220,61840 & 225,96539 \end{bmatrix} \quad (39)$$

$$S^{-1} = (X'X)^{-1} = \begin{bmatrix} 0,031317353 & 0,004235766 & -0,023392266 & -0,20888093 & -0,06393784 & 0,006912719 \\ 0,004235766 & 0,073466908 & 0,020378881 & -0,05692495 & -0,03606514 & -0,009440071 \\ -0,023392266 & 0,020378881 & 0,161911989 & 0,012177957 & 0,027810541 & -0,01935397 \\ -0,020888093 & -0,005692495 & 0,012177957 & 0,021457429 & 0,036901879 & -0,002631372 \\ -0,063937838 & -0,03606514 & 0,027810541 & 0,036901879 & 0,151715506 & -0,013834086 \\ 0,006912719 & -0,009440071 & -0,01935397 & -0,002631372 & -0,01383408 & 0,011659743 \end{bmatrix}$$

(40)

CONCLUSÃO

Com base nos resultados e nas condições em que as pesquisas foram realizadas ao longo de oito anos na Microrregião de Petrolina/PE, podemos concluir que:

- Foram aceitas as hipóteses pré-formuladas de influência positiva da variação da data de transação ou coleta de informação de oferta, do nível de infraestrutura, da elasticidade da oferta, da percentagem da área irrigável, e da área equivalente de produção vegetal, além de influência negativa para a dimensão da área;
- A hipótese de influência das classes de capacidade de uso da terra e das condições de acesso ao imóvel mensurada pela nota agrônômica, calculada com base na escala de França, foi rejeitada pelo teste t de Student ao nível de 10% de significância;
- O modelo proposto apresentou bom ajustamento com os dados observados, sendo aceitas as hipóteses estatísticas que permitem sua validação;
- O modelo é adequado para estimativa do valor de imóveis irrigáveis com área inferior a 1.500 ha e energia elétrica instalada, com o propósito de avaliação para desapropriação por interesse social ou compra e venda, por utilidade pública, e

para ações judiciais que envolvem partilha de bens, mas os modelos preditores, em geral, são mais úteis para avaliação de imóveis de forma massiva ou por motivos fiscais, tributários ou para cobrança de emolumentos cartorários;

- A técnica estatística da regressão linear múltipla mostrou-se uma ferramenta eficiente no estudo do mercado de imóveis rurais;
- Os fatores que compõem o valor do imóvel podem ser utilizados na avaliação de imóveis irrigáveis pelo processo estatístico descritivo de homogeneização por fatores, quando a metodologia de avaliação utilizada for o método comparativo direto de dados de mercado, recomendando-se que todos os fatores que compõem o VTI/ha, determinados neste trabalho, sejam utilizados na mesma homogeneização;
- Outras pesquisas necessitam ser feitas para melhor avaliar e conhecer o comportamento do mercado de imóveis rurais nas microrregiões do Sertão de Pernambuco, além da validação da ferramenta estatística de análise de regressão linear múltipla e do teste de outras formas de análise.

BIBLIOGRAFIA

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Avaliação de bens parte 3:** imóveis rurais. NBR 14.653-3. Rio de Janeiro, 2004. 27 p.

BAPTISTELLA, M. **O uso de redes neurais e regressão linear múltipla na engenharia de avaliações:** determinação dos valores venais de imóveis urbanos. Universidade Federal do Paraná, Dissertação de Mestrado, Curitiba, 2005.

DANTAS, R.A. **Engenharia de avaliação:** uma introdução à metodologia científica. 2ª.ed. São Paulo: Pini, 2005.

GONZÁLEZ, M.A.S.; FORMOSO, C.T. Análise conceitual das dificuldades na determinação de modelos de formação de preços através de análise de regressão. **Engenharia Civil – UM**, nº 8, 2000, p. 65 – 75.

IBGE. Indicadores de preços. www.ibge.gov.br. Acesso em 15 de abril de 2009.

INSTITUTO BRASILEIRO DE FRUTAS – IBRAF. Estatísticas. www.ibraf.org.br. Acesso em 01 de junho de 2009.

IMAPE. **Fundamentos de avaliação patrimoniais e perícias de engenharia.** 1.ed. São Paulo. Editora Pini, 1998.

INCRA. **Manual para obtenção de terras e perícia judicial.** Brasília, 2006.

INSTITUTO FNP. **Análise do mercado de terras:** relatório bimestral, nº 15. São Paulo, 2007.

INSTITUTO FNP. **Análise do mercado de terras**: relatório bimestral, nº 28. São Paulo, 2009.

JOHNSTON, J. **Métodos econométricos**. São Paulo: Atlas, 1977.

LEVINE, D.M.; BERENSON, M.L.; STEPHAN, D. **Estatística**: teoria e aplicações. Rio de Janeiro: LTC, 2005.

LIMA, M.R. de C. **Avaliação de propriedades rurais**: manual básico. 2ª ed. ver. e atual. São Paulo: Liv. e Ed. Universitária de Direito, 2005.

MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO EXTERIOR – MDIC. Balança comercial brasileira, município de Petrolina. www.mdic.gov.br/sítio/ Acesso em 01 de junho de 2009.

NADAL, C.A. JULIANO, K.A.; RATTON, E. Testes estatísticos utilizados para a validação de regressões múltiplas aplicadas na avaliação de imóveis urbanos. **Bol. Ciênc. Geod.**, séc. Artigos, Curitiba, v.9, nº2, p.243-262, 2003.

PEIXOTO, A.P.N. **Regressão linear simples**. UNIFESP, São Paulo, 2007.

REIS, F.P. **INF 661 – Estatística aplicada**. Centro de ciências exatas e tecnologia, Viçosa, 2000.

SILVA, J.A.M. Fatores de homogeneização e valor da terra nua (VTN) ao longo do tempo para o sertão pernambucano e norte baiano. In: Congresso Brasileiro de Engenharia de Avaliações e Perícias, 14. IBAPE-BAHIA, **Anais...** Salvador, 2007. (CD-ROM)

Anexo I – Fatores de homogeneização sugeridos por intervalo das variáveis explicativas para avaliação de imóveis irrigáveis localizados nos municípios de Petrolina, Lagoa Grande e Santa Maria da Boa Vista/PE.

Elasticidade		Dimensão da área		Época da coleta da amostra		% de área irrigável		Área equivalente de Produção Vegetal		Nível de infraestrutura			
Tipo	índice	Área (ha)	índice	Período	índice	Intervalo (%)	índice	Intervalo - Área (ha)	índice	Nível	índice		
NR	1,00	1	1,00	01/01/2001 a 30/06/2001	1,02	>0 - 5	1,73	>1 - 5	1,12	0	1,00		
OF	1,72	>1 - 2	0,87	01/07/2001 a 31/12/2001	1,08	>5 - 10	2,73	>5 - 10	1,27	1	1,09		
		>2 - 5	0,73	01/01/2002 a 30/06/2002	1,14	>10 - 15	3,53	>10 - 15	1,35	2	1,19		
		>5 - 10	0,50	01/07/2002 a 31/12/2002	1,2	>15 - 20	4,17	>15 - 20	1,41	3	1,30		
		>10 - 20	0,40	01/01/2003 a 30/06/2003	1,27	>20 - 25	4,72	>20 - 30	1,47				
		>20 - 30	0,33	01/07/2003 a 31/12/2003	1,34	>25 - 30	5,22	>30 - 40	1,53				
		>30 - 50	0,28	01/01/2004 a 30/06/2004	1,41	>30 - 35	5,68	>40 - 50	1,58				
		>50 - 100	0,23	01/07/2004 a 31/12/2004	1,48	>35 - 40	6,10	>50 - 60	1,62				
		>100 - 150	0,19	01/01/2005 a 30/06/2005	1,56	>40 - 45	6,49	>60 - 80	1,67				
		>150 - 200	0,17	01/07/2005 a 31/12/2005	1,64	>45 - 50	6,86	>80 - 100	1,72				
		>200 - 300	0,15	01/01/2006 a 30/06/2006	1,73	>50 - 60	7,38	>100 - 150	1,78				
		>300 - 500	0,13	01/07/2006 a 31/12/2006	1,82	>60 - 70	8,02	>150 - 200	1,86				
		>500 - 1000	0,10	01/01/2007 a 30/06/2007	1,92	>70 - 80	8,62	>200 - 300	1,94				
		>1000 - 1500	0,09	01/07/2007 a 31/12/2007	2,02	>80 - 90	9,17	>300 - 400	2,02				
						01/01/2008 a 30/06/2008	2,12	>90 - 100	9,70	>400 - 500	2,08		
						01/07/2008 a 31/12/2008	2,23			>500 - 700	2,15		
				01/01/2009 a 30/06/2009	2,34			>800 - 1000	2,26				
				01/07/2009 a 31/12/2009	2,46			>1000 - 1200	2,32				
				01/01/2010 a 30/06/2010	2,59			>1200 - 1500	2,37				
				01/07/2010 a 31/12/2010	2,72								
				01/01/2011 a 30/06/2011	2,85								
				01/07/2011 a 31/12/2011	3,00								
				01/01/2012 a 30/06/2012	3,14								
				01/07/2012 a 31/12/2012	3,30								

NR= imóvel negociado; OF= imóvel ofertado

Anexo II – Características da amostra utilizada para a modelagem

Item	Data	Nome do imóvel	Município/PE	Área (ha)	Valor Total - R\$	NR/OF	Infraestrutura	Produção vegetal	% área irrigável
1	04/06/01	Fazenda Formiga	Lagoa Grande	137,0000	185.000,00	NR	3	88,2554	66,10
2	08/11/01	Fazenda Ribeiro	Lagoa Grande	315,0000	511.000,00	NR	3	321,6466	65,00
3	08/11/01	Fazenda Gado Bravo	Lagoa Grande	72,0000	50.000,00	NR	1	2,2975	65,00
4	21/03/01	Fazenda Santa Rita	Petrolina	108,4973	90.000,00	NR	1	1,00	52,00
5	21/03/01	Fazenda Pedrinhas	Petrolina	148,0000	140.000,00	OF	1	1,00	59,90
6	21/03/01	Lote Agrícola	Petrolina	132,0000	250.000,00	OF	2	159,2050	70,00
7	21/03/01	Fazenda Pedrinhas	Petrolina	130,0000	150.000,00	OF	3	1,00	70,00
8	21/03/01	Lote Projeto de Irrigação Maria Teresa	Petrolina	490,0000	400.000,00	OF	2	1,00	50,00
9	23/03/01	Lote 1642 N-10 P.I.S.N.C.	Petrolina	5,8440	12.865,00	NR	2	1,00	100,00
10	23/03/01	Lote 1253 N-08 P.I.S.N.C.	Petrolina	6,6000	15.567,50	NR	2	1,00	100,00
11	23/03/01	Lote 1409 N-10 P.I.S.N.C.	Petrolina	6,7396	18.991,20	NR	2	1,00	100,00
12	23/03/01	Lote 883 N-11 P.I.S.N.C.	Petrolina	15,4841	19.915,15	NR	2	1,00	40,33
13	23/03/01	Lote C-2024 P.I. Maria Teresa	Petrolina	6,6531	15.712,41	NR	2	1,00	100,00
14	05/09/02	Lote 19 P.I. Maria Teresa	Petrolina	116,0000	45.000,00	OF	0	1,00	10,00
15	08/03/05	Lote 298 N-03 P.I.S.N.C.	Petrolina	6,5000	70.000,00	OF	2	59,2684	100,00
16	08/03/05	Lote Projeto de Irrigação Maria Teresa	Petrolina	6,0000	50.000,00	OF	2	35,1620	100,00
17	08/03/05	Área na Estrada das Pedrinhas	Petrolina	15,0000	50.000,00	OF	1	1,00	43,00
18	08/03/05	Lote Projeto de Irrigação Maria Teresa II	Petrolina	24,0000	70.000,00	OF	1	1,00	42,00
19	08/03/05	Lote C-2410 R-5 P.I. Maria Teresa	Petrolina	7,7000	70.000,00	OF	2	44,8570	100,00
20	08/03/05	Lote Projeto de Irrigação Bebedouro	Petrolina	10,5000	80.000,00	OF	1	68,5849	100,00
21	14/03/05	Lote N-25 P.I. Maria Teresa	Petrolina	24,0000	80.000,00	OF	1	1,00	45,00
22	01/05/05	Lote C-2409 R-5 P.I. Maria Teresa	Petrolina	6,5000	85.000,00	OF	1	137,7102	100,00
23	03/03/08	Lote 1960 P.I. Maria Teresa	Petrolina	20,0000	150.000,00	OF	2	62,0256	50,00
24	06/03/08	Lotes C-2201, C-2202, C-2205, C-2206 e As-18	Petrolina	73,0000	500.000,00	OF	3	261,6671	100,00
25	06/03/08	Lote 655 N-06 P.I.S.N.C.	Petrolina	8,5000	130.000,00	OF	3	32,8584	100,00
26	08/03/08	Lote 2402 P.I. Maria Teresa	Petrolina	11,0000	40.000,00	NR	2	1,00	90,00
27	14/08/01	Fazenda Caraibas II	Sta. M ^a da Boa Vista	148,0000	50.000,00	NR	2	1,00	15,00
28	04/06/01	Fazenda Nossa Senhora do Carmo	Sta. M ^a da Boa Vista	975,0000	200.000,00	NR	2	1,00	24,80
29	04/06/01	Fazenda Nova Esperança	Sta. M ^a da Boa Vista	972,0000	220.000,00	NR	0	1,00	32,80
30	11/02/05	Fazenda Poço do Icó	Sta. M ^a da Boa Vista	406,0000	140.000,00	OF	1	1,00	6,00
31	23/10/07	Fazenda Barra do Jacaré	Sta. M ^a da Boa Vista	75,5891	80.000,00	NR	1	1,00	34,81
32	23/10/07	Fazenda Malhada Real	Sta. M ^a da Boa Vista	180,0846	105.000,00	NR	1	1,00	20,00

Anexo III – Matrizes utilizadas na modelagem

Y - ln(VTI/ha)	Xi - ln(Dia - Juliano)	X2 - Benfeitoria	X3 - Elast	X4- ln(área)	X5- ln(% área irrig.)	X6 - ln(Produção Veg.(ha))
7,20813	10,51992	3	0	4,91998	4,19117	4,48023
7,39155	10,52414	3	0	5,75257	4,17439	5,77345
6,54311	10,52414	1	0	4,27667	4,17439	0,83182
6,72084	10,51789	1	0	4,68673	3,95124	0,00000
6,85219	10,51789	1	1	4,99721	4,09268	0,00000
7,54641	10,51789	2	1	4,88280	4,24850	5,07019
7,05086	10,51789	3	1	4,86753	4,24850	0,00000
6,70481	10,51789	2	1	6,19441	3,91202	0,00000
7,69685	10,51794	2	0	1,76542	4,60517	0,00000
7,76587	10,51794	2	0	1,88707	4,60517	0,00000
7,94373	10,51794	2	0	1,90800	4,60517	0,00000
7,15942	10,51794	2	0	2,73981	3,69710	0,00000
7,76712	10,51794	2	0	1,89508	4,60517	0,00000
5,96083	10,53220	0	1	4,75359	2,30259	0,00000
9,28445	10,55631	2	1	1,87180	4,60517	4,08208
9,02802	10,55631	2	1	1,79176	4,60517	3,55997
8,11173	10,55631	1	1	2,70805	3,76120	0,00000
7,97820	10,55631	1	1	3,17805	3,73767	0,00000
9,11503	10,55631	2	1	2,04122	4,60517	3,80348
8,93841	10,55631	1	1	2,35138	4,60517	4,22807
8,11173	10,55646	1	1	3,17805	3,80666	0,00000
9,47860	10,55771	1	1	1,87180	4,60517	4,92515
8,92266	10,58431	2	1	2,99573	3,91202	4,12755
8,83190	10,58439	3	1	4,29046	4,60517	5,56707
9,63522	10,58439	3	1	2,14007	4,60517	3,49221
8,19874	10,58444	2	0	2,39790	4,49981	0,00000
5,82257	10,52183	2	0	4,99721	2,70805	0,00000
5,32364	10,51992	2	0	6,88244	3,21084	0,00000
5,42203	10,51992	0	0	6,87936	3,49043	0,00000
5,84304	10,55566	1	1	6,00635	1,79176	0,00000
6,96447	10,58096	1	0	4,32531	3,54990	0,00000
6,36829	10,58096	1	0	5,19343	2,99573	0,00000