

## Relação entre o poder calorífico superior e os componentes elementares e minerais da biomassa vegetal

Thiago de Paula Protásio<sup>1</sup>, Lina Bufalino<sup>1</sup>, Gustavo Henrique Denzin Tonoli<sup>1</sup>, Allan Motta Couto<sup>1</sup>, Paulo Fernando Trugilho<sup>1</sup>, Mário Guimarães Júnior<sup>2</sup>

<sup>1</sup>Universidade Federal de Lavras, Departamento de Ciências Florestais, CP 3037, CEP 37200-000, Lavras, Minas Gerais, Brasil,

<sup>2</sup>Centro Federal de Educação Tecnológica de Minas Gerais, Departamento de Eletrônica, CEP 38180.510, Araxá, Minas Gerais, Brasil

\*Autor correspondente:  
depaulaprotasio@gmail.com

### Termos para indexação:

Modelagem  
Regressão  
Correlação  
Seleção

### Index terms:

Modeling  
Regression  
Correlation  
Selection

### Histórico do artigo:

Recebido em 03 mar 2011  
Aprovado em 28 jun 2011  
Publicado em 30 jun 2011

doi: 10.4336/2011.pfb.31.66.113

**Resumo** - Os objetivos do trabalho foram avaliar a correlação, ajustar e selecionar modelos estatísticos lineares simples e múltiplos entre os componentes elementares (carbono, hidrogênio e oxigênio) e o teor de cinzas com o poder calorífico superior da biomassa vegetal; utilizar a análise de componentes principais para a criação de um índice de desempenho energético e ajustar um modelo linear entre o índice de desempenho energético e o poder calorífico superior. Utilizaram-se oito tipos de biomassa. Foram ajustadas equações referentes a três modelos estatísticos lineares simples e nove múltiplos. Os melhores modelos foram selecionados com base na significância dos seus coeficientes, no coeficiente de determinação ajustado, no erro padrão da estimativa, no coeficiente de variação, na linearidade dos parâmetros, na normalidade, na presença de heterocedasticidade e ausência de autocorrelação dos erros. Para os modelos lineares múltiplos, determinou-se o fator de inflação de variância. Encontrou-se alta correlação entre as variáveis. Os modelos 1, 3 e 11 foram considerados os mais adequados. A utilização prática do modelo 2 foi impossibilitada. A análise de componentes principais foi eficiente na obtenção de um índice de desempenho energético dos resíduos lignocelulósicos e pode ser utilizada para contornar a multicolinearidade encontrada entre as variáveis consideradas.

## Relation between higher heating value and elemental and mineral biomass plant components

**Abstract** - The aims of this work were to evaluate the correlation, to adjust and select simple and multiple linear statistical models between elemental components (carbon, hydrogen and oxygen) and ash content with higher heating value for plant biomass; to use the principle components analysis for the creation of an energetic development index and to adjust a linear model between energetic development index and higher heating value. Eight types of biomass were used. Three linear and nine multiple statistical models were adjusted. The best models were selected based on the significance of coefficients, adjusted determination coefficient, estimative standard error, coefficient of variation, linearity of parameters, normality, presence of heterocedasticity and lack of error correlation. The variance inflation factor was determined for linear multiple models. High correlation between the variables studied was found. Models 1, 3 and 11 were considered most adequate. Practical use of model 2 is not possible. Principle components analysis was efficient in obtaining an energetic development index of lignocellulosic residues and it may be used for solving multicollinearity found between variables considered.

## Introdução

O uso energético de resíduos lignocelulósicos tem ganhado destaque em vários países devido à grande quantidade desse tipo de material produzido pelo setor agroflorestal. No entanto, para o aproveitamento racional e adequado de resíduos agrícolas e florestais, é necessário o estudo de suas propriedades energéticas. O poder calorífico é um excelente parâmetro para se avaliar a potencialidade energética de combustíveis de biomassa (Brand, 2010; Friedl et al., 2005; Parikh et al., 2005), sendo definido como a quantidade de energia liberada na combustão completa de uma unidade de massa do material combustível. O poder calorífico é chamado de superior (PCS) quando a combustão se efetua a volume constante e no qual a água formada durante o processo é condensada e o calor latente do vapor d'água não é perdido (Brand, 2010; Friedl et al., 2005; Associação Brasileira de Normas Técnicas, 1984).

A modelagem com o uso de regressões lineares simples ou múltiplas pode ser fundamental para analisar a relação entre variáveis e realizar previsões, seja do PCS, dos componentes químicos elementares (carbono - C, hidrogênio - H e oxigênio - O) ou do teor de cinzas (A). Esses componentes não são os únicos presentes na biomassa vegetal, mas são os de maior relevância a serem estudados devido às relações entre esses elementos (C, H, O e A) e o poder calorífico. No entanto, a grande dificuldade de um bom ajuste para um modelo de regressão é atender a todos os pressupostos exigidos por essa análise estatística, como: a linearidade dos parâmetros, normalidade, homocedasticidade (variância constante dos erros do modelo no intervalo de valores de uma variável independente) e ausência de autocorrelação dos erros (resíduos). Além disso, a média dos erros do modelo de regressão deve ser igual a zero e as variáveis regressoras não podem apresentar acentuada colinearidade (ou multicolinearidade), sendo essa última exigência exclusiva para as regressões múltiplas (Hair Jr. et al., 2009; Werkema & Aguiar, 2006). Nesse caso, a análise de componentes principais é uma técnica que pode ser usada juntamente com a análise de regressão (Montgomery et al., 2001) como forma de contornar a multicolinearidade. Essa técnica permite ainda a interpretação direta dos componentes para a criação de índices de interesse (Mingoti, 2005).

As regressões lineares simples e múltiplas envolvendo os componentes químicos elementares (carbono, hidrogênio e oxigênio), o teor de cinzas e o poder calorífico superior (PCS) são de grande valia, pois a determinação do PCS é um processo de custo elevado, requer instrumentos especiais e analistas altamente treinados (Majumder et al., 2008). Existem diversos modelos na literatura que relacionam o PCS com as análises químicas, imediata e elementar de combustíveis sólidos (Erol et al., 2010; Akkaya et al., 2009; Majumder et al., 2008; Friedl et al., 2005; Parikh et al., 2005; Sheng & Azevedo, 2005; Thipkhumthod et al., 2005; Demirbas & Demirbas, 2004; Demirbas, 2001; Cordero et al., 2001; Demirbas, 1997). No entanto, os pressupostos exigidos na análise de regressão são pouco explorados na literatura científica, sendo que esse procedimento pode ser fundamental para a avaliação da qualidade do ajuste e a confiabilidade do modelo. Além disso, esses modelos não foram desenvolvidos para os combustíveis de biomassa específicos do Brasil, o que pode resultar em estimativas pouco precisas.

Sheng & Azevedo (2005) afirmam que as correlações baseadas na análise elementar são as mais acuradas, quando comparadas àquelas realizadas com a análise imediata, porque esse tipo de análise fornece apenas uma composição empírica da biomassa. Os autores ainda mencionam que as correlações baseadas na composição química (lignina e extrativos) não são viáveis devido à variação das propriedades dos componentes. Contudo, a adição do teor de cinzas nos modelos pode contribuir para a capacidade preditiva das equações ajustadas (Majumder et al., 2008).

Portanto, no presente trabalho, objetivou-se:

Avaliar a correlação linear entre os teores de carbono, hidrogênio, oxigênio e cinzas e o poder calorífico superior.

Ajustar e selecionar modelos lineares simples e múltiplos entre o poder calorífico superior e os componentes químicos elementares (carbono, hidrogênio e oxigênio) e o teor de cinzas.

Aplicar a análise de componentes principais para a criação de um índice de desempenho energético da biomassa analisada.

Ajustar um modelo linear entre o índice de desempenho energético e o poder calorífico superior, visando verificar a associação entre essas variáveis.

## Material e métodos

Foram utilizados oito tipos de biomassa vegetal compostos por maravalhas de eucalipto (*Eucalyptus urophylla* S.T. Blake com 6 anos de idade), maravalhas de cedro australiano (*Toona ciliata* M. Roem com 18 anos de idade), maravalhas de pinus (*Pinus* sp. com 35 anos de idade), casca de arroz, bagaço de cana-de-açúcar, resíduos do processamento de café (casca e pergaminho), resíduos da colheita do milho (palha, sabugo, caule e folhas) e polpa celulósica residual de bambu (*Bambusa vulgaris* Schrad. ex J.C. Wendl). A biomassa vegetal avaliada foi amostrada aleatoriamente. Para os ensaios realizados foram utilizadas cinco frações de cada resíduo que foram devidamente moídos, totalizando 40 observações por variável.

Para a análise elementar, as amostras foram moídas e peneiradas, sendo utilizada a fração retida pela peneira de 270 mesh. A quantificação dos teores de carbono, hidrogênio, nitrogênio e enxofre em relação à massa seca dos resíduos lignocelulósicos foi realizada em um analisador universal da marca Elementar (modelo Vario Micro Cube). O analisador utiliza como gás de arraste e ignição o hélio e o oxigênio, respectivamente. As amostras de 2 mg foram colocadas em cápsulas de estanho e completamente incineradas a 1.200 °C. O teor de oxigênio foi obtido por diferença (Eq. 1), conforme diretrizes de Bech et al. (2009):

$$O = 100 - C - H - N - S - A \quad (1)$$

em que: O é o teor de oxigênio (%), C é o teor de carbono (%), H é o teor de hidrogênio (%), N é o teor de nitrogênio (%), S é o teor de enxofre (%) e A é o teor de cinzas (%).

Para a quantificação do poder calorífico superior (PCS), foi utilizado um calorímetro digital da marca IKA C-200, sendo o ensaio padronizado com a norma ABNT NBR 8633 (Associação Brasileira de Normas Técnicas, 1984). As amostras foram classificadas em peneiras de 40/60 mesh (American Society for Testing Materials, 1977). Para a determinação do poder calorífico superior, as frações das amostras retidas na peneira de 60 mesh foram secas em estufa a 103±2°C até massa constante. O teor de cinzas foi determinado segundo as diretrizes previstas na norma ABTCP M11/77 (Associação Brasileira Técnica de Celulose e Papel, 1974) com o uso de um forno tipo mufra.

Na Tabela 1 encontram-se os modelos estatísticos lineares simples e múltiplos propostos que foram ajustados com base no método dos mínimos quadrados

ordinários (MQO). Os modelos foram escolhidos com base na retirada e adição de variáveis visando à melhoria e determinação das equações que melhor representam a relação entre o poder calorífico superior (PCS) e os teores de carbono (C), hidrogênio (H) e oxigênio mais cinzas (O+A).

**Tabela 1.** Modelos estatísticos propostos entre o poder calorífico superior (PCS) e os componentes químicos elementares (C, H e O) e o teor de cinzas (A).

Nº do modelo	Modelos estatísticos
1	$PCS = \beta_0 + \beta_1 C + \varepsilon$
2	$PCS = \beta_0 + \beta_1 H + \varepsilon$
3	$PCS = \beta_0 + \beta_1 (O+A) + \varepsilon$
4	$PCS = \beta_0 + \beta_1 C + \beta_2 H + \varepsilon$
5	$PCS = \beta_0 + \beta_1 C + \beta_2 H + \beta_3 C*H + \varepsilon$
6	$PCS = \beta_0 + \beta_1 C + \beta_2 (O+A) + \varepsilon$
7	$PCS = \beta_0 + \beta_1 C + \beta_2 (O+A) + \beta_3 [C*(O+A)] + \varepsilon$
8	$PCS = \beta_0 + \beta_1 H + \beta_2 (O+A) + \varepsilon$
9	$PCS = \beta_0 + \beta_1 H + \beta_2 (O+A) + \beta_3 [H*(O+A)] + \varepsilon$
10	$PCS = \beta_0 + \beta_1 C + \beta_2 H + \beta_3 (O+A) + \varepsilon$
11	$PCS = \beta_0 + \beta_1 C + \beta_2 H + \beta_3 (O+A) + \beta_4 C*H + \varepsilon$
12	$PCS = \beta_0 + \beta_1 C + \beta_2 H + \beta_3 (O+A) + \beta_4 [C*(O+A)] + \varepsilon$

$\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$ : coeficientes dos modelos;  $\varepsilon$ : erro do modelo; C: Carbono (%); H: Hidrogênio (%); O: Oxigênio (%); A: cinzas (%); PCS: poder calorífico superior (kcal kg<sup>-1</sup>).

Os melhores modelos foram selecionados com base na significância de todos os seus coeficientes, no coeficiente de determinação ajustado (R<sup>2</sup> ajust), no erro padrão da estimativa (S<sub>xy</sub>), no coeficiente de variação experimental, na linearidade das variáveis (Teste de Harvey-Collier), na normalidade (Teste Shapiro Wilk), presença de variâncias desiguais, ou seja, heterocedasticidade (Teste Goldfeld-Quandt) e ausência de autocorrelação (Teste Durbin-Watson) dos erros dos modelos ajustados.

Verificou-se o sinal algébrico dos coeficientes da regressão, uma vez que as variáveis independentes podem ser consideradas por sua contribuição individual à variável estatística dependente. O parâmetro  $\beta_j$  representa a alteração esperada na variável resposta quando a variável  $x_j$  sofre um acréscimo unitário e se todas as demais variáveis explicativas forem mantidas constantes (Hair Jr. et al., 2009; Werkema & Aguiar, 2006).

Para os modelos lineares múltiplos, determinou-se o fator de inflação de variância (VIF), pois esse é um indicador do efeito que as outras variáveis independentes têm sobre o erro padrão de um coeficiente de regressão. Valores de VIF altos (geralmente acima de 10) indicam um alto grau de colinearidade ou multicolinearidade (Hair Jr. et al. 2009).

Os melhores modelos foram selecionados e comparados estatisticamente por análises de variância a 5% de significância, ou seja, foram comparados aos pares os modelos 4 e 5, 6 e 7, 8 e 9, 10 e 11, 1 e 11 e as equações 3 e 11.

Também foi avaliada a correlação linear de Pearson entre as variáveis dos modelos propostos. Utilizou-se o teste t a um nível de 5% de significância para verificar se o coeficiente de correlação é igual a zero e, conseqüentemente, avaliar a existência de uma relação linear entre as variáveis.

### Análise de componentes principais

Foi utilizada a técnica multivariada de componentes principais com o objetivo de construir um índice de desempenho energético (IDE) da biomassa e verificar a associação desse índice com o poder calorífico superior por meio do ajuste de um modelo linear simples. Para

essa técnica estatística foram considerados apenas os teores de carbono (C), hidrogênio e oxigênio mais cinzas (O+A). A utilização dessa análise multivariada apresenta a vantagem de que as componentes principais obtidas são compostas por variáveis aleatórias não correlacionadas, facilitando, portanto, a interpretação conjunta dessas variáveis (Mingoti, 2005).

Avaliaram-se os autovalores das componentes principais que refletem a variância estimada por elas, a porcentagem de variação explicada pelas componentes e as correlações entre as variáveis originais e as componentes principais.

Todas as análises estatísticas foram efetuadas utilizando o *software* estatístico R versão 2.11.0 (R Development Core Team, 2008) e na avaliação dos pressupostos da análise de regressão utilizou-se os pacotes *lmtest* (Hothorn et al., 2010) e *fmsb* (Nakazawa, 2010).

### Resultados e discussão

Na Tabela 2 encontram-se os valores médios encontrados para os teores de carbono, hidrogênio, oxigênio e cinzas e para o poder calorífico superior dos combustíveis de biomassa vegetal analisados.

**Tabela 2.** Valores médios encontrados para os teores de carbono, hidrogênio, oxigênio e cinzas e para o poder calorífico superior da biomassa analisada, bem como os coeficientes de variação encontrados (entre parênteses).

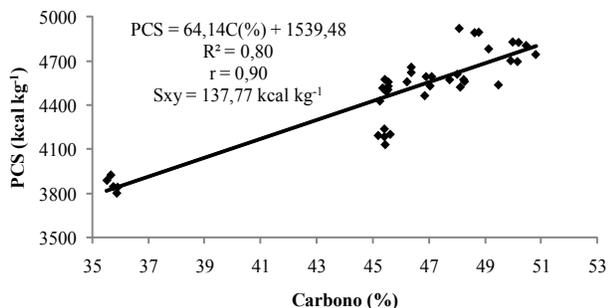
Resíduos Lignocelulósicos	C	H	O	A	PCS
Maravalhas de eucalipto	48,20 <sup>(1,38)</sup>	6,36 <sup>(1,27)</sup>	45,04 <sup>(1,66)</sup>	0,25 <sup>(10,62)</sup>	4572,2 <sup>(0,57)</sup>
Maravalhas de pinus	49,00 <sup>(1,61)</sup>	6,65 <sup>(1,82)</sup>	43,76 <sup>(2,08)</sup>	0,39 <sup>(11,70)</sup>	4864,2 <sup>(1,18)</sup>
Maravalhas de cedro	50,00 <sup>(0,73)</sup>	6,26 <sup>(1,01)</sup>	42,34 <sup>(0,95)</sup>	1,01 <sup>(2,68)</sup>	4756,2 <sup>(0,61)</sup>
Bagaço de cana-de-açúcar	46,80 <sup>(2,35)</sup>	6,34 <sup>(1,79)</sup>	45,44 <sup>(2,51)</sup>	0,93 <sup>(4,78)</sup>	4511,8 <sup>(1,26)</sup>
Polpa celulósica de bambu	45,20 <sup>(0,34)</sup>	6,43 <sup>(1,85)</sup>	46,67 <sup>(0,64)</sup>	1,56 <sup>(8,70)</sup>	4192,4 <sup>(0,91)</sup>
Processamento de café	46,50 <sup>(0,82)</sup>	6,38 <sup>(0,99)</sup>	39,86 <sup>(1,09)</sup>	4,92 <sup>(0,98)</sup>	4606,4 <sup>(0,80)</sup>
Colheita do milho	45,20 <sup>(0,26)</sup>	6,43 <sup>(2,13)</sup>	39,32 <sup>(0,35)</sup>	6,83 <sup>(0,39)</sup>	4515,4 <sup>(1,28)</sup>
Casca de arroz	36,00 <sup>(0,44)</sup>	5,12 <sup>(1,55)</sup>	41,56 <sup>(0,67)</sup>	16,78 <sup>(0,53)</sup>	3863,8 <sup>(1,22)</sup>

C: carbono (%); H: hidrogênio (%); O: oxigênio (%); A: teor de cinzas (%); PCS: poder calorífico superior (kcal kg<sup>-1</sup>).

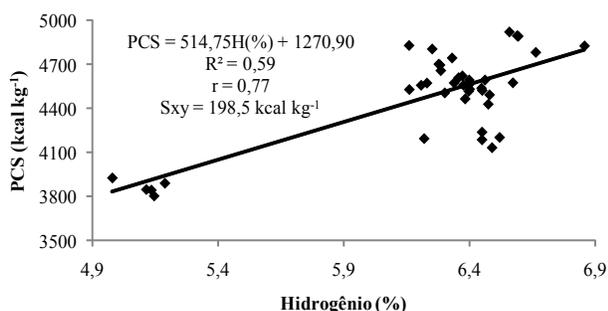
Observam-se, de maneira geral, baixos coeficientes de variação amostral, o que sugere alta precisão das análises efetuadas. Os valores encontrados para os teores dos componentes elementares (C, H e O) e cinzas (A) e para o poder calorífico superior (PCS) estão coerentes ao relatado na literatura para vários tipos de combustíveis

de biomassa (Lima et al., 2011; Brand, 2010; Kumar et al., 2010; Bech et al., 2009; Friedl et al., 2005; Parikh et al., 2005; Cordero et al., 2001).

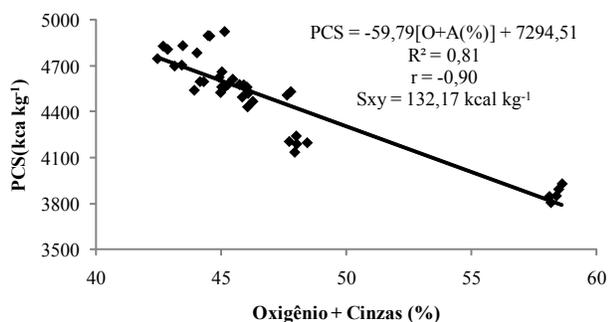
Nas Figuras 1, 2 e 3 encontram-se as relações lineares simples estabelecidas entre o poder calorífico superior e os teores de C, H e O+A, respectivamente.



**Figura 1.** Representação gráfica da relação entre o poder calorífico superior (PCS) e o teor de carbono (modelo 1).



**Figura 2.** Representação gráfica da relação entre o poder calorífico superior (PCS) e o teor de hidrogênio (modelo 2).



**Figura 3.** Representação gráfica da relação entre o poder calorífico superior (PCS) e o teor de oxigênio+cinzas (modelo 3).

Todas as correlações lineares foram significativas ( $p$ -valor  $< 0,001$ ) pelo teste  $t$ , o que comprova a existência de uma relação linear entre as variáveis estudadas. Foi encontrada correlação positiva entre o PCS e o teor de H e C, ou seja, há uma tendência de maiores valores de PCS estarem associados a maiores valores percentuais de H e C. Demirbas & Demirbas (2004) encontraram correlações altamente significativas entre os teores de C e H com o PCS, corroborando o presente trabalho. Quanto à correlação encontrada entre o teor de O+A e

o PCS, observou-se que não há relação positiva entre essas variáveis, atendendo assim a expectativa prévia. É necessário mencionar, no entanto, que o oxigênio é um elemento essencial para promover a combustão do material, apesar de não contribuir para um aumento do PCS. Huang et al. (2009) afirmam que altos valores de oxigênio decrescem o poder calorífico. Nordin (1994), de forma semelhante, afirma que alta relação entre O e H diminui o poder calorífico do material combustível. Brand (2010) afirma que altos teores de cinzas contribuem para a redução do poder calorífico, visto que os materiais minerais (cinzas) não participam do processo de combustão, mas são contabilizados na massa do combustível submetido ao processo de queima. Dessa forma, a associação entre o oxigênio e as cinzas nos modelos estatísticos propostos é válida, pois ambos contribuem de forma negativa para o PCS.

Os modelos lineares ajustados 1, 2 e 3 apresentaram todos os coeficientes significativos ( $p$ -valor  $< 0,001$ ). Entretanto, o modelo 2 apresentou baixo coeficiente de determinação e elevado erro padrão da estimativa, quando comparado aos modelos 1 e 3. Dessa forma, não foram encontrados parâmetros estatísticos satisfatórios para o modelo 2. A provável causa é o baixo teor desse elemento em relação aos demais componentes avaliados.

Avaliando-se o sinal algébrico dos coeficientes dos modelos, pode-se observar que todos estão adequados à expectativa prévia, ou seja, relações diretas entre o PCS e os teores de C e H e inversa com o teor de O+A. Os valores positivos de  $\beta_1$  encontrados para os modelos 1 e 2 podem ser interpretados como um aumento no PCS de 64,14 e 514,75 kcal kg<sup>-1</sup> quando se aumenta 1% de C e H, respectivamente. Jenkins (1989) relata que a cada variação unitária de C (%) há um acréscimo de aproximadamente 93,15 kcal kg<sup>-1</sup> no PCS, diferindo em 29 kcal kg<sup>-1</sup> do encontrado no presente trabalho. Essa diferença deve-se provavelmente aos tipos de combustíveis de biomassa utilizados nos ajustes dos modelos. No modelo 3, o coeficiente  $\beta_1$  apresentou sinal algébrico negativo, ou seja, há uma diminuição de 59,79 kcal kg<sup>-1</sup> quando ocorre uma variação unitária no teor de O+A. Nota-se ainda que a alteração no PCS é muito maior quando ocorre a variação no teor de H.

Entretanto, a quantidade desse elemento na biomassa é inferior ao teor de carbono elementar.

Na Tabela 3 estão apresentadas as relações funcionais lineares múltiplas ajustadas, bem como o coeficiente de determinação ajustado e o erro padrão da estimativa.

**Tabela 3.** Parâmetros das regressões lineares múltiplas ajustadas para a predição indireta do poder calorífico superior (PCS).

Nº	Coefficientes	tc	Fc	R <sup>2</sup> <sub>ajust</sub>	Sxy (kcal kg <sup>-1</sup> )
4	$\beta_0$	1658,08	5,31**	0,79	137,12
	$\beta_1$	69,89	6,37**		
	$\beta_2$	-61,33	-0,60 <sup>ns</sup>		
5	$\beta_0$	8041,93	2,22**	0,80	131,53
	$\beta_1$	-89,36	-0,98 <sup>ns</sup>		
	$\beta_2$	-1147	-1,84 <sup>ns</sup>		
	$\beta_3$	26,73	1,77 <sup>ns</sup>		
6	$\beta_0$	5739,28	2,634*	0,81	41,87
	$\beta_1$	17,60	0,717 <sup>ns</sup>		
	$\beta_2$	-43,89	-1,938 <sup>ns</sup>		
7	$\beta_0$	820,08	0,326 <sup>ns</sup>	0,85	73,94
	$\beta_1$	166,11	3,163**		
	$\beta_2$	76,98	1,757 <sup>ns</sup>		
	$\beta_3$	-3,54	-3,117**		
8	$\beta_0$	9607,63	8,353**	0,83	125,36
	$\beta_1$	-221,17	-2,045*		
	$\beta_2$	-79,62	-7,473**		
9	$\beta_0$	-2524,50	-0,647 <sup>ns</sup>	0,86	110,38
	$\beta_1$	1833,20	2,845**		
	$\beta_2$	151,90	2,097*		
	$\beta_3$	-39,70	-3,224**		
10	$\beta_0$	8549,31	3,348**	0,82	124,97
	$\beta_1$	11,12	0,466 <sup>ns</sup>		
	$\beta_2$	-213,99	-1,938 <sup>ns</sup>		
	$\beta_3$	-68,93	-2,716*		
11	$\beta_0$	21462,25	4,677**	0,86	109,59
	$\beta_1$	-270,43	-3,022**		
	$\beta_2$	-2044,5	-3,930**		
	$\beta_3$	-93,65	-3,564**		
	$\beta_4$	43,72	3,239**		
12	$\beta_0$	833,26	0,201 <sup>ns</sup>	0,84	116,47
	$\beta_1$	165,92	2,328*		
	$\beta_2$	-0,56	-0,004 <sup>ns</sup>		
	$\beta_3$	76,78	1,129 <sup>ns</sup>		
	$\beta_4$	-3,54	-2,29*		

Nº: número do modelo; tc: t calculado; Fc: F calculado; R<sup>2</sup><sub>ajust</sub>: coeficiente de determinação ajustado; Sxy: erro padrão da estimativa; e\*: significativo a 5% de probabilidade; e\*\*\*: significativo a 1% de probabilidade; e<sup>ns</sup>: não significativo a 5% de probabilidade.

Observa-se que todos os modelos lineares múltiplos apresentaram-se estatisticamente significativos pelo teste F (p-valor < 0,01) e com elevado coeficiente de determinação ajustado. Entretanto, apenas os modelos

8 e 11 apresentaram todos os coeficientes significativos pelo teste t (p-valor < 0,05). Observou-se ainda que o modelo 11 apresentou erro padrão da estimativa menor e um coeficiente de determinação maior que o modelo 8.

Avaliando-se os sinais algébricos dos coeficientes do modelo 6, observa-se que são coerentes ao encontrado na literatura e explicam com veracidade a influência do teor de carbono e oxigênio + cinzas no PCS (Brand, 2010; Huang et al., 2009; Demirbas & Demirbas, 2004; Nordin, 1994). Entretanto, esses coeficientes não foram significativos pelo teste t a 5% de significância.

De forma geral, os demais modelos lineares múltiplos ajustados apresentaram pelo menos um dos sinais algébricos dos coeficientes não coerentes com o que foi observado para os modelos lineares. Esse resultado indica que os modelos lineares múltiplos propostos não são inteiramente confiáveis para explicar as relações e possíveis interações entre as variáveis.

Os modelos 10, 11 e 12 apresentaram, de maneira geral, coeficientes com sinais algébricos opostos àqueles que seriam esperados. O melhor modelo múltiplo ajustado por Sheng & Azevedo (2005) para predição do poder calorífico superior de combustíveis da biomassa foi:  $PCS (MJ/kg) = -1,3675 + 0,3137C + 0,7009H + 0,0318O$ . No entanto, também é possível observar que o sinal algébrico do coeficiente  $\beta_3$  obtido não foi adequado, pois aumentos no teor de O não contribuem para aumento do PCS (Huang et al., 2009; Nordin, 1994). Demirbas et al. (1997) ajustaram o modelo estatístico  $PCS (kJ/g) = 33,5C + 142,3H - 15,4O$  que estabelece relação direta entre o PCS e os teores de C e H e inversa com o teor de O, ou seja, apresenta sinais algébricos coerentes ao relatado na literatura (Huang et al., 2009; Demirbas & Demirbas, 2004; Nordin, 1994) e diferente do encontrado no presente trabalho.

Com base na avaliação dos coeficientes dos modelos de regressão múltipla ajustados, pode-se perceber a dificuldade de se modelar estatisticamente o comportamento ou incremento do PCS da biomassa vegetal. Isso ocorreu provavelmente devido às altas correlações observadas entre os componentes elementares (multicolinearidade). Segundo Hair Jr. et al. (2009) o impacto da multicolinearidade é reduzir o poder preditivo de qualquer variável independente na medida em que ela é associada com as outras variáveis independentes. Os autores ainda afirmam que quando a colinearidade aumenta, a variância única explicada por conta de cada variável independente diminui e o percentual da previsão compartilhada aumenta.

Os modelos 4 e 5 não apresentaram diferença significativa (p-valor = 0,08), ou seja, a interação entre o C e o H pode ser estatisticamente desconsiderada.

Já os modelos 6 e 7 apresentaram-se estatisticamente diferentes (p-valor = 0,003), o que pode ser atribuído a interação entre o C e o O+A. Os modelos 8 e 9 também diferiram estatisticamente (p-valor = 0,003), sendo a interação entre os teores de H e O+A relevante para a modelagem do PCS em função dos componentes elementares. Comparando-se estatisticamente o modelo 1 com o modelo 11, observou-se diferença significativa entre eles (p-valor = 0,001) proporcionada pela adição das demais variáveis independentes no modelo 11, mesmo sendo observada relativa multicolinearidade entre as variáveis. De forma análoga, comparou-se o modelo 3 e o modelo múltiplo 11 e observou-se diferença estatística significativa (p-valor = 0,004) entre essas relações funcionais. Além disso, o modelo 11 apresentou coeficiente de determinação maior que os modelos 1 e 3, sendo que esse resultado é um indicativo da capacidade preditiva e da superioridade do modelo 11 em explicar a variabilidade do PCS dos resíduos lignocelulósicos.

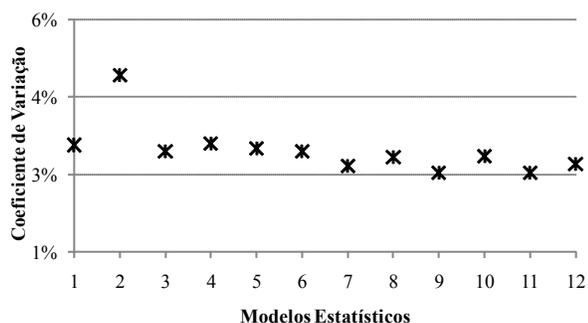
Avaliando a linearidade das variáveis, foi possível verificar que para todos os modelos propostos, com exceção do modelo 2, o ajuste linear é adequado segundo o Teste de Harvey-Collier (p-valor > 0,05). As variáveis dependentes e independentes apresentaram comportamento linear, reforçando a relação linear encontrada entre o PCS e os componentes elementares (C, H, O) e o teor de cinzas (A) da biomassa. Diante desse resultado, a utilização prática do modelo 2 é impossibilitada.

Com base nos valores do fator de inflação de variância (VIF) e nas altas correlações lineares encontradas, foi observado que há efeito da multicolinearidade nas regressões múltiplas propostas, sendo que isso pode ter afetado de maneira direta os sinais algébricos dos modelos. Contudo, o VIF se mostrou estatisticamente aceitável (abaixo de 10 para todos os modelos múltiplos ajustados), sendo que o modelo 4 apresentou o menor índice (5,08) e o modelo 11 apresentou o maior índice (7,96). A multicolinearidade (relação entre várias variáveis independentes) ocorre quando qualquer variável independente é altamente correlacionada com um conjunto de outras variáveis independentes.

Observou-se que os modelos estatísticos propostos de número 6, 7, 8, 9 e 10 não apresentaram resíduos da regressão com distribuição normal a 5% de significância pelo teste Shapiro-Wilk, uma das suposições que são realizadas no ajuste. As demais relações funcionais ajustadas atenderam a esse pressuposto da análise de regressão.

Avaliando-se os modelos quanto ao teste de Durbin-Watson, observou-se que apenas os modelos 9 e 11 não apresentaram correlação serial dos erros ( $p$ -valor  $< 0,05$ ), sendo que esse resultado pode proporcionar estimativas não eficientes e aumento do erro padrão da estimativa. Quando foram avaliadas a homocedasticidade dos erros da regressão (teste de Goldfeld-Quandt), observou-se que todos os modelos ajustados apresentaram erros com variância constante ( $p$ -valor  $> 0,05$ ).

Na Figura 4 encontram-se os coeficientes de variação encontrados para os modelos estatísticos ajustados.



**Figura 4.** Coeficiente de variação experimental dos modelos estatísticos ajustados.

O modelo 2 apresentou o maior coeficiente de variação experimental. Esse resultado pode ser um indicativo da baixa capacidade preditiva desse modelo. Além disso, o modelo 2 apresentou baixo coeficiente de determinação, elevado erro padrão da estimativa e ausência de linearidade dos parâmetros. Notou-se ainda que o modelo 9 apresentou o menor valor para o coeficiente de variação experimental, mas não o menor erro padrão da estimativa quando comparado aos demais ajustes. Os demais modelos apresentaram coeficientes de variação semelhantes.

O modelo 6 apresentou o menor erro padrão da estimativa, um elevado coeficiente de determinação ajustado e sinais algébricos consistentes. Entretanto, apresentou erros da regressão com correlação e sem distribuição normal, não atendendo assim dois pressupostos da análise de regressão. Além disso, não apresentou todos os seus coeficientes significativos.

Os modelos 9 e 11 apresentaram os maiores coeficientes de determinação ajustado. No entanto, o modelo 9 apresentou menor erro padrão da estimativa, mas não apresentou resíduos com distribuição normal. Já o modelo 11 atendeu a todos os pressupostos da análise

de regressão avaliados, sem, entretanto, ter apresentado sinais algébricos dos coeficientes condizentes. Isso ocorreu provavelmente devido à multicolinearidade entre as variáveis do modelo. Diante disso, o modelo 11 deve ser utilizado apenas para fins preditivos.

Analisando-se os modelos lineares é possível visualizar que o modelo estatístico 1 apresentou considerável coeficiente de determinação ajustado, sinais algébricos consistentes, normalidade e homocedasticidade dos erros da regressão. Entretanto, apresentou elevada correlação serial dos erros do modelo. Já o modelo linear 3 apresentou coeficiente de determinação ajustado maior e menor erro padrão da estimativa, porém também apresentou correlação serial dos erros da regressão, ou seja, a predição do poder calorífico superior não é independente em cada ponto considerado. Isso ocorreu, provavelmente, devido à proximidade dos valores encontrados para as amostras retiradas ao acaso de cada biomassa vegetal avaliada. O modelo 3 também apresentou normalidade e homocedasticidade dos erros. No entanto, esses dois modelos lineares simples podem ser utilizados para predições e avaliações preliminares da biomassa.

#### Análise de componentes principais

Na Tabela 4 encontram-se os resultados obtidos para a análise de componentes principais considerando-se como variáveis originais os teores de carbono, hidrogênio e oxigênio elementares e o teor de cinzas.

**Tabela 4.** Autovalores e proporção da variância total explicada e acumulada pelos três componentes principais obtidos a partir das variáveis originais carbono, hidrogênio e oxigênio mais cinzas.

Componentes	Autovalores	Variância Explicada (%)	Variância Acumulada (%)
$\hat{Y}_1$	2,847	94,89	94,89
$\hat{Y}_2$	0,135	4,51	99,40
$\hat{Y}_3$	0,018	0,60	100,00

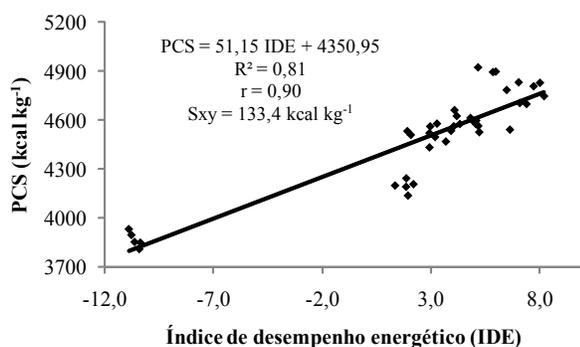
A primeira componente principal obtida apresentou o maior autovalor e, conseqüentemente, uma alta variância estimada. Essa componente explica grande parte da variância total dos dados e pode ser interpretada como um índice de desempenho energético (IDE) dos resíduos lignocelulósicos, pois apresenta correlações positivas com os teores de C (0,98) e H (0,95) e negativa com o teores de O+A (-0,99). Quanto maior o

valor desse índice (escore) melhor será o desempenho dos resíduos lignocelulósicos para a produção de bioenergia. As demais componentes principais podem ser desconsideradas, pois contribuem pouco para a variância total original. O índice de desempenho energético (IDE) ou componente principal 1 pode ser representado da seguinte forma (Eq. 2):

$$\text{IDE} = 0,580 C + 0,505 H - 0,639 (O + A) \quad (2)$$

em que: O é o teor de oxigênio (%), C é o teor de carbono (%), H é o teor de hidrogênio (%), A é o teor de cinzas (%).

Na Figura 5 encontra-se o modelo de regressão ajustado entre o poder calorífico superior e o índice de desempenho energético (escores da componente principal 1).



**Figura 5.** Relação funcional entre o poder calorífico superior (PCS) e o índice de desempenho energético (IDE).

Analisando-se o modelo de regressão ajustado entre o PCS e o IDE, nota-se a alta relação entre as variáveis, ou seja, quanto maior o índice de desempenho energético maior será o poder calorífico superior e melhor será a biomassa para a produção energética. O modelo ajustado apresentou todos os coeficientes significativos ( $p$ -valor  $< 0,001$ ), alto coeficiente de determinação e baixo erro padrão da estimativa. A análise dos erros do modelo indicou que os erros do modelo apresentaram normalidade ( $p$ -valor = 0,0842) e homocedasticidade ( $p$ -valor = 0,2640), mas a 5% de significância os resíduos apresentaram autocorrelação. O ajuste do modelo linear é estatisticamente aceitável, pois as variáveis apresentaram comportamento linear ( $p$ -valor = 0,9752). A análise de regressão utilizando-se como variável independente o IDE apresenta a grande vantagem de contornar a multicolinearidade observada nas variáveis, uma vez que a componente principal 1 (IDE) é composta por variáveis aleatórias não correlacionadas. Desse

modo, esse modelo pode ser utilizado para a classificação da biomassa estudada para a produção energética, bem como para avaliações preliminares da biomassa considerada nos ajustes.

As altas correlações encontradas entre o PCS, os componentes elementares (C, H e O) e o teor de cinzas (A) reforçam a expectativa da possibilidade de utilização de modelos simples ou múltiplos para a predição do PCS e das demais variáveis (C, H, O e A). Isso pode ser muito útil na diminuição dos custos e tempo das análises e em avaliações preliminares de vários tipos de biomassa, como na seleção de resíduos lignocelulósicos para a produção de calor, vapor ou eletricidade em caldeiras ou gaseificadores.

## Conclusões

Foi obtida alta correlação linear entre o poder calorífico superior e os teores de carbono, hidrogênio e oxigênio mais cinzas.

Os modelos 1, 3 e 11 foram considerados os mais adequados e podem ser utilizados para predições de interesse. A utilização prática do modelo 2 é impossibilitada.

A análise de componentes principais foi eficiente na obtenção de um índice de desempenho energético dos resíduos lignocelulósicos e pode ser utilizada para contornar a multicolinearidade encontrada entre o poder calorífico superior e os teores de carbono, hidrogênio e oxigênio mais cinzas.

## Agradecimentos

Os autores expressam seus sinceros agradecimentos aos Professores Doutores Daniel Furtado Ferreira e Renato Ribeiro de Lima, da Universidade Federal de Lavras, pelo suporte estatístico concedido, fundamental para a execução desse trabalho.

## Referências

- AKKAYA, A. V. Proximate analysis based multiple regression models for higher heating value estimation of low rank coals. **Fuel Processing Technology**, v. 90, n. 2, p. 165–170, 2009.
- AMERICAN SOCIETY FOR TESTING MATERIALS. **Standard method for chemical analyses of wood charcoal**. Philadelphia, 1977. 1042 p.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 8633: carvão vegetal – determinação do poder calorífico**. Rio de Janeiro, 1984.

- ASSOCIAÇÃO BRASILEIRA TÉCNICA DE CELULOSE E PAPEL. **M-11/77**: teor de cinzas. São Paulo, Brasil, 1974.
- BECH, N.; JENSEN, P. A.; DAM-JOHANSEN, K. Determining the elemental composition of fuels by bomb calorimetry and the inverse correlation of HHV with elemental composition. **Biomass and Bioenergy**, v. 33, n. 3, p. 534–537, 2009.
- BRAND, M. A. **Energia de biomassa florestal**. Rio de Janeiro: Interciência, 2010. 131 p.
- CORDERO, T.; MARQUEZ, F.; RODRIGUEZ-MIRASOL, J.; RODRIGUEZ, J. J. Predicting heating values of lignocellulosics and carbonaceous materials from proximate analysis. **Fuel**, v. 80, n. 1, p. 1567-157, 2001.
- DEMIRBAS, A. Calculation of higher heating values of biomass fuels. **Fuel**, v. 76, n. 5, p. 431-434, 1997.
- DEMIRBAS, A.; DEMIRBAS, H. A. Estimating the calorific values of lignocellulosic fuels. **Journal Energy, Exploration & Exploitation**, v. 20, n. 1, p. 105-111, 2004.
- DEMIRBAS, A. Relationships between lignin contents and heating values of biomass. **Energy Conversion & Management**, v. 42, n. 2, p. 183-188, 2001.
- EROL, M.; HAYKIRI-ACMA H.; KUÇUKBAYRAK, S. Calorific value estimation of biomass from their proximate analyses data. **Renewable Energy**, v. 35, n. 1, p. 170–173, 2010.
- FRIEDL, A.; PADOUVAS, E.; ROTTER, H.; VARMUZA, K. Prediction of heating values of biomass fuel from elemental composition. **Analytica Chimica Acta**, v. 544, n. 1-2, p. 191–198, 2005.
- HAIR JR., J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise multivariada de dados**. Porto Alegre: Bookman, 2009. 688 p.
- HOTHORN, T.; ZEILEIS, A.; MILLO, G.; MITCHELL, D. Testing Linear Regression Models. R package version 2.11.1. Disponível em: <http://cran-r.c3sl.ufpr.br/>. Acesso em: 13 jun. 2011.
- HUANG, C.; HAN, L.; YANG, Z.; LIU, X. Ultimate analysis and heating value prediction of straw by near infrared spectroscopy. **Waste Management**, v. 29, n. 6, p. 1793–1797, 2009.
- JENKINS, B. M. Physical properties of biomass. In: KITANI, O.; HALL, C. W. (Eds.). **Biomass Handbook**. New York: Gordon & Breach, 1989. p. 860-891.
- KUMAR, R.; PANDEY, K. K.; CHANDRASHEKAR, N.; MOHAN, S. Effect of tree-age on calorific value and other fuel properties of Eucalyptus hybrid. **Journal of Forestry Research**, v. 21, n. 4, p. 514–516, 2010.
- LIMA, E. A.; SILVA, H. D.; LAVORANTI, O. J. Caracterização dendroenergética de árvores de *Eucalyptus benthamii*. **Pesquisa Florestal Brasileira**, Colombo, v. 31, n. 65, p. 09-17, 2011.
- MAJUMDER, A. K.; JAIN, R.; BANERJEE, P.; BARNWAL, J. P. Development of a new proximate analysis based correlation to predict calorific value of coal. **Fuel**, v. 87, n. 13-14, p. 3077-3081, 2008.
- MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: UFMG, 2005. 297 p.
- MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**. New York: John Wiley, 2001. 640 p.
- NAKAZAWA, M. Functions for medical statistics book with some demographic data. R package version 2.11.1. Disponível em: <http://cran-r.c3sl.ufpr.br/>. Acesso em: 13 jun. 2011.
- NORDIN, A. Chemical elemental characteristics of biomass fuels. **Biomass and Bioenergy**, v. 6, n. 5, p. 339–347, 1994.
- PARIKH, J.; CHANNIWALA, S. A.; GHOSAL, G. K. A correlation for calculating HHV from proximate analysis of solid fuels. **Fuel**, v. 84, n. 5, p. 487–494, 2005.
- R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2008. Disponível em: <http://www.R-project.org>. Acesso em: 10 dez. 2010.
- SHENG, C.; AZEVEDO, J. L. T. Estimating the higher heating value of biomass fuels from basic analysis data. **Biomass and Bioenergy**, v. 28, n. 5, p. 499-507, 2005.
- THIPKHUNTHOD, P.; MEEYOO, V.; RANGSUNVIGIT, P.; KITIYANAN, B.; SIEMANOND, K.; RIRKSOMBOON, T. Predicting the heating value of sewage sludges in Thailand from proximate and ultimate analyses. **Fuel**, v. 84, n. 7-8, p. 849–857, 2005.
- WERKEMA, C.; AGUIAR, S. **Análise de regressão: como entender o relacionamento das variáveis de um processo**. Belo Horizonte: Werkema, 2006. 306 p.