

EFEITO DA ESCOLHA DA ORDEM DO MODELO SOBRE A CAUSALIDADE DE GRANGER

VINICIUS B. VASCONCELOS, CELSO J. MUNARO

Departamento de Engenharia Elétrica, Universidade Federal do Espírito Santo

Av Fernando Ferrari, 514, Vitória-ES-CEP 29.075-910

E-mails: belmuds.eng@gmail.com, munaro@ele.ufes.br

Abstract— Granger causality has been applied in many areas, one of them is fault diagnosis in industrial processes. This paper shows how the Granger causality tests are sensitive to different lag lengths chosen for the estimated models. The statistical significance using F distribution for different orders is analysed and compared with the order provided by several information criteria from literature. The analysis is performed using 7 signals with well known causality relations from an industrial process. It was found that 7 from a total of 42 causal relationships were ambiguous. The information about the frequency of causality detection in the range of orders selected resulted in just one ambiguous relation. Although a higher computational effort is required, the proposal allowed reducing considerably the uncertainty on the detected causality relations.

Keywords— Granger causality, f statistic, Bonferroni correction, VAR models, order selection.

Resumo— A causalidade de Granger tem sido aplicada em muitas áreas, entre elas a de diagnósticos de falhas em processos industriais. Este trabalho mostra como os testes de causalidade de Granger são sensíveis à escolha das ordens dos modelos utilizados. A significância estatística usando a distribuição F para diferentes ordens é analisada e comparada com a ordem dada por critérios de seleção bem estabelecidos na literatura, baseados na teoria da informação. As análises são feitas com 7 sinais de um processo industrial com relações de causalidade conhecidas. Observou-se que das 42 relações de causalidade inferidas, 7 foram ambíguas. A utilização da frequência de detecções de causalidade para as diferentes ordens resultou na ambiguidade de apenas uma das relações de causalidade. Embora exija um maior esforço computacional, a proposta permitiu remover a incerteza da maioria das relações de causalidade.

Palavras-chave Causalidade de Granger, estatística f , correção de Bonferroni, modelos VAR, seleção de ordem de modelos.

1 Introdução

A causalidade de Granger (Granger, 1969) tem sido utilizada em várias áreas da ciência, como por exemplo, na área das ciências sociais (Russo, 2009), neurociências (Seth, et al., 2015) e no campo da engenharia de processos industriais (Yang et al, 2014). Aplicações em processos industriais requerem em geral que se obtenha a topologia dos sinais considerados, permitindo indicar assim a origem de distúrbios originados de falhas.

Para reduzir erros na aplicação do método de Granger, várias técnicas de identificação de sistemas para melhorar os modelos e para verificar a significância estatística das relações de causalidade foram propostas. Condições para pré-processamento dos sinais, utilização de critérios para escolha da ordem do modelo, técnicas de validação, estatísticas de testes na inferência da causalidade de Granger e correções para testes de múltiplas comparações são apresentadas em (Seth, 2010). Um estudo comparativo da análise condicional e par a par (Ding et al., 2006) na aplicação de redes de genes regulatórios, usando a inferência estatística da distribuição f com diferentes ordens de modelos, é feita por Tam et al. (2012) e por Siyal et al. (2014). O leitor interessado em outros métodos de detecção de causalidade pode consultar Yang et. al, (2014).

O objetivo deste trabalho é avaliar o efeito da escolha da ordem de modelos, assim como a utilização de métodos de seleção de ordem de modelos baseados na teoria da informação, para avaliar seu efeito na inferência da causalidade de Granger, usan-

do para isto sinais de uma planta industrial com relações de causalidade conhecidas.

Este trabalho é organizado como segue: a causalidade de Granger juntamente com a inferência estatística e métodos de seleção de modelos baseados na teoria da informação são apresentados na Seção 2. Um estudo de caso com dados de uma planta industrial é mostrado na Seção 3. A Seção 4 destaca alguns pontos importantes observados nas análises e indica novos caminhos que podem ser utilizados para reduzir a incerteza nas relações da causalidade detectadas.

2 Causalidade de Granger e testes estatísticos de significância

2.1 Causalidade Condicional de Granger e inferência estatística

Considere duas séries temporais $x_1(t)$ e $x_2(t)$ onde $t = 1, 2, \dots, T$. Este sistema bivariável admite um modelo vetor autoregressivo (VAR) que inclui todas as variáveis (modelo irrestrito):

$$x_1(t) = \sum_{l=1}^p A_{11,l} x_1(t-l) + \sum_{l=1}^p A_{12,l} x_2(t-l) + \varepsilon_1(t) \quad (1)$$

$$x_2(t) = \sum_{l=1}^p A_{21,l} x_1(t-l) + \sum_{l=1}^p A_{22,l} x_2(t-l) + \varepsilon_2(t)$$

Neste modelo, p é a ordem do modelo e representa a quantidade de atrasos utilizados, $A_{i,j}$ são os

coeficientes do modelo e os erros de predição das duas séries temporais são representados por $\varepsilon_1(t)$ e $\varepsilon_2(t)$. Existe um total de $m = T - p$ pares de equações que pode ser escrito na notação matricial como:

$$Y = XB + \Sigma \quad (2)$$

Sendo:

$$Y_{(mx2)} = \begin{bmatrix} x_{1,p+1} & x_{2,p+1} \\ x_{1,p+2} & x_{2,p+2} \\ \vdots & \vdots \\ x_{1,T} & x_{2,T} \end{bmatrix}$$

$$X_{(m+2p)} = \begin{bmatrix} x_{1,p} & x_{2,p} & \dots & x_{1,1} & x_{2,1} \\ x_{1,p+1} & x_{2,p+1} & \dots & x_{1,2} & x_{2,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{1,T-p} & x_{2,T-p} & \dots & x_{1,T} & x_{2,T} \end{bmatrix}$$

$$B_{(2px2)} = \begin{bmatrix} A'_1 \\ A'_2 \\ \vdots \\ A'_p \end{bmatrix}; \Sigma_{(mx2)} = \begin{bmatrix} \varepsilon_{1,p+1} & \varepsilon_{2,p+1} \\ \varepsilon_{1,p+2} & \varepsilon_{2,p+2} \\ \vdots & \vdots \\ \varepsilon_{1,T} & \varepsilon_{2,T} \end{bmatrix}$$

A condição $m \geq np$, onde n é a quantidade de variáveis, é necessária para evitar a singularidade do produto matricial $X'X$, de modo que os coeficientes da matriz B possam ser estimados utilizando o método dos mínimos quadrados:

$$\hat{B} = (X'X)^{-1}X'Y \quad (3)$$

O erro de predição do modelo de x_1 é dado por

$$RSS_1 = (y - X\hat{\beta})'(y - X\hat{\beta}) \quad (4)$$

onde y e $\hat{\beta}$ são os vetores da primeira coluna das matrizes Y e $\hat{B}$, respectivamente. Para medir a influência da variável x_2 na x_1 , um modelo VAR de $x_1(t)$ sem incluir valores passados de $x_2(t)$ é obtido (modelo restrito):

$$x_1(t) = \sum_{l=1}^p A_{11,l}x_1(t-l) + \varepsilon_{12}(t) \quad (5)$$

A soma quadrada dos resíduos do modelo representado na equação 5 é calculada (RSS_{21}) e a inferência de causalidade de $x_2 \rightarrow x_1$ pode ser calculada utilizando a estatística F (Hamilton, 1984; Hastie et al., 2009, p. 48):

$$f_{x_2 \rightarrow x_1} = \frac{\frac{(RSS_{21} - RSS_1)}{p}}{\frac{RSS_1}{(m - np)}} \quad (6)$$

Esta estatística é desenvolvida sob a hipótese nula de que não existe causalidade na direção testada, no caso, na direção $x_2 \rightarrow x_1$. A hipótese nula segue a distribuição $F(p, m - np)$, sendo p o grau de liberdade do numerador e $m - np$ o grau de liberdade do denominador, desta forma o p -valor correspondente pode ser calculado. Caso $f_{x_2 \rightarrow x_1} > F^\alpha$ ou o p -valor seja menor que o nível de significância α utilizado no teste, rejeita-se a hipótese nula de que x_2 não causa x_1 (Silveira Bueno, 2011; Tam et al., 2012).

2.2 Correção de Bonferroni para o teste de múltiplas comparações

A inferência da causalidade de Granger possui um total de $M = n(n - 1)$ testes de hipóteses independentes a serem considerados. Desta forma é necessário utilizar algum tipo de correção para o teste de múltiplas comparações, pois a medida que se aumenta o número de hipóteses a serem testadas, aumenta-se a probabilidade de se cometer um erro do tipo I, que é caracterizado pela rejeição da hipótese nula quando ela é verdadeira (falsos positivos), através da equação (Hastie et al., 2009, p. 686):

$$P(\text{pelo menos 1 falso positivo}) = 1 - (1 - \alpha)^M \quad (7)$$

A correção de *Bonferroni* é utilizada quando não se possui um número muito grande de testes de hipóteses independentes a serem realizados. É obtida encontrando o valor crítico do teste (α_c), que é a razão do nível de significância nominal (α) pela quantidade total de testes (M). Desta forma o valor crítico para cada teste individual será α/M (McDonald, 2014). Assim, apenas são consideradas significantes causalidades na qual o p -valor seja inferior ou igual a α_c ou $f > F^{\alpha_c}$ para a direção testada.

2.3 Métodos de seleção de ordem de modelos baseados na teoria da informação e análise de resíduos

Para fazer a estimação de um modelo utilizando a estrutura VAR na análise de Granger é necessário fornecer a ordem que será utilizada. Uma das maneiras de se escolher a ordem de um modelo é utilizando os critérios de informação, sendo usados neste trabalho os seguintes critérios: *Akaike Information Criterion* (AIC), *Bayesian Information Criterion* (BIC) e *Hannan-Quinn* (HQ), cujas expressões para o cálculo são (Lütkepohl, 2005):

$$AIC_{(p)} = \ln |\Sigma| + \frac{2pn^2}{T} \quad (8)$$

$$BIC_{(p)} = \ln |\Sigma| + \frac{\ln(T)pn^2}{T} \quad (9)$$

$$HQ_{(p)} = \ln |\Sigma| + \frac{2 \ln(\ln(T))pn^2}{T} \quad (10)$$

Onde $\ln |\Sigma|$ é o logaritmo neperiano do determinante da matriz de covariância dos resíduos dos modelos restritos e mede a adequação do modelo. Desta forma, quanto menor a variância dos resíduos, melhor será o modelo. Porém, isto é obtido utilizando mais parâmetros na estimação e isto é penalizado na segunda parcela das equações. Assim é selecionada a ordem do modelo que minimiza o critério. Após o modelo ser estimado, é necessário verificar a autocorrelação dos resíduos gerados, uma vez que caso os resíduos não sejam ruído branco há informações ainda não captadas pelo modelo, e eles podem ser eventualmente melhorados (Silveira Bueno, 2011). Neste trabalho é utilizada a função de autocorrelação, a função de autocorrelação parcial e a estatística

de *Ljung-Box* sobre os resíduos gerados para verificar se eles são estatisticamente ruído branco.

2.4 Efeito da ordem nos testes de significância

A ordem escolhida para o modelo possui uma influência na inferência estatística da distribuição f quando se observa a equação 6. O aumento da ordem do modelo faz aumentar o grau de liberdade do numerador e diminui o grau de liberdade do denominador para um dado tamanho de vetor de dados T . A equação 6 pode ser reescrita da seguinte forma:

$$f_{x2 \rightarrow x1} = \frac{RSS_{21} - RSS_1}{RSS_1} \cdot \frac{m - np}{p} = f_1 \cdot f_2 \quad (11)$$

O termo f_1 depende dos resíduos, que tendem a diminuir com a ordem. O termo f_2 depende da ordem e também diminui quando ela aumenta. Portanto, o produto $f_1 f_2$ tende a diminuir com o aumento da ordem dos modelos.

Para mostrar esta situação, foram escolhidos dois sinais de um processo industrial que possuem relações de causalidade bem definidas (*PIC400* e *LIC400*), respectivamente de uma malha de controle de pressão e uma de vazão, e inferida a causalidade nas duas direções. A utilização do AIC estimou uma ordem igual a 18 para os modelos, já os critérios BIC e HQ estimaram ordens iguais a 8 e 10 respectivamente. Para a análise do efeito de *PIC400* sobre *LIC400*, a análise de resíduos indicou ruído branco para ordens $p \geq 2$. Para o modelo que relaciona o efeito de *LIC400* sobre *PIC400*, os resíduos se tornaram ruído branco para ordens $p \geq 8$. Baseando-se nisso, variou-se a ordem dos modelos de 8 a 30, já que este intervalo garante tanto ordem mínima encontrada pelo critério BIC quanto a validação dos resíduos dos modelos. A partir da ordem 30, o resultado da indicação ou não da causalidade não mais sofreu variações. A Figura 1 ilustra o comportamento da estatística f e do limiar do teste F na inferência de causalidade nas direções testadas entre os dois sinais. Para todas as análises utilizadas neste trabalho se considerou a correção de *Bonferroni* para os testes de múltiplas comparações e $\alpha = 5\%$.

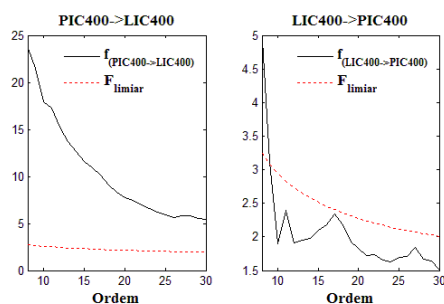


Figura 1. Estatística f e limiar variando a ordem do modelo

O comportamento dos termos f_1 e f_2 da equação 11 para esta análise é mostrado na Figura 2

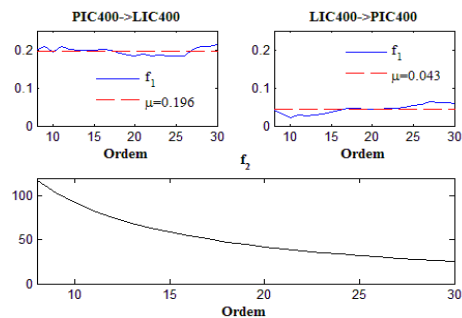


Figura 2. Comportamento dos termos f_1 e f_2 da estatística f com o aumento da ordem do modelo

O termo f_2 possui um efeito grande para baixas ordens ($f_2 > 100$), e vai diminuindo com o aumento da ordem. O valor do termo f_1 na direção *PIC400* $\rightarrow$ *LIC400* é maior do que na direção *LIC400* $\rightarrow$ *PIC400*. A utilização de ordens baixas pode fornecer um resultado significativo na inferência da causalidade devido ao peso do termo f_2 na equação 11 quando a magnitude de f_1 é pequena. Isto pode ser observado na direção *LIC400* $\rightarrow$ *PIC400* da Figura 1, onde houve uma indicação de causalidade quando a ordem utilizada foi de 8 e 9.

A relação de causalidade na direção *PIC400* $\rightarrow$ *LIC400* é bem definida como pode ser constatado pela Figura 1, que indica sempre uma relação de causalidade com o aumento da ordem do modelo utilizado. Ao inferir a causalidade na direção *LIC400* $\rightarrow$ *PIC400* com a ordem dada pelo critério BIC a inferência estatística mostra uma relação de causalidade significativa. Ao usar a ordem fornecida pelos critérios de HQ e AIC, a inferência da causalidade não foi significativa, resultado que permaneceu até a ordem 30. Estes resultados mostram a divergência que a análise de causalidade pode indicar com a variação da ordem do modelo e da escolha do critério a ser utilizado para a sua estimação, mesmo que os resíduos dos modelos sejam ruído branco.

3 Aplicação em um processo industrial

Esta seção utiliza sinais provenientes de 7 malhas de controle de uma central termoeletrica que foram coletados por 5000 segundos com tempo de amostragem de 5s. O diagrama de instrumentação da planta é mostrado na Figura 3, com as malhas de controle utilizadas para gerar os sinais analisados. As relações de causa e efeito entre as variáveis do processo, apontada pelos operadores, são mostradas no mapa causal da Figura 4.

Antes de aplicar o método de causalidade de Granger, é necessário avaliar se os dados possuem a covariância estacionária (Seth, 2010). Esta análise foi feita utilizando o teste aumentado de *Dickey-Fuller*, que verifica a existência de raízes unitárias em uma série temporal, verificando-se a necessidade de diferenciar os dados uma vez. A Figura 5 mostra os sinais no tempo após a diferenciação.

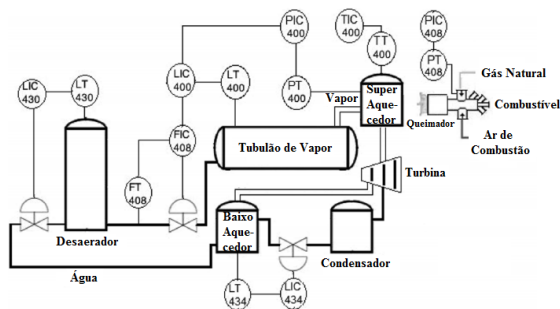


Figura 3. Diagrama de instrumentação das malhas de controle selecionadas

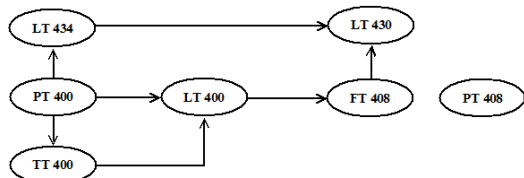


Figura 4. Mapa causal fornecido pelos operadores do processo com os operadores

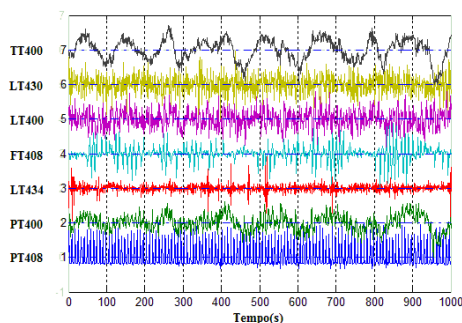


Figura 5. Sinais no tempo provenientes das 7 malhas após a diferenciação

Ao aplicar os critérios de seleção da ordem dos modelos nos sinais diferenciados das 7 malhas, as ordens encontradas pelos critérios AIC, BIC e HQ foram respectivamente 14, 7 e 7. A Tabela 1 mostra a

ordem do modelo para o qual os resíduos são ruído branco e o intervalo das ordens escolhidas para inferir a causalidade de Granger tendo determinado sinal como saída.

Tabela 1. Ordens de estimação dos modelos cujos resíduos se tornam ruído branco e intervalos utilizados para inferência da causalidade de Granger.

Análise de resíduos						
PIC408	PIC400	LIC434	FIC408	LIC400	LIC430	TIC400
$p \geq 15$	$p \geq 6$	$p \geq 8$	$p \geq 10$	$p \geq 2$	$p \geq 3$	$p \geq 7$
Intervalos das ordens para inferência da causalidade de Granger						
PIC408	PIC400	LIC434	FIC408	LIC400	LIC430	TIC400
$15 \leq p$	$7 \leq p$	$8 \leq p$	$10 \leq p$	$7 \leq p$	$7 \leq p$	$7 \leq p$
≤ 30	≤ 30	≤ 30	≤ 30	≤ 30	≤ 30	≤ 30

A menor ordem estimada pelos critérios foi 7 (BIC e HQ). Utilizando, por exemplo, a inferência da causalidade tendo a LIC434 como saída, se pode verificar pela Tabela 1 que a partir da ordem 8, os resíduos são ruído branco. Portanto para a análise deste sinal foi escolhido um intervalo de $8 \leq p \leq 30$, pois este intervalo garante a validação dos resíduos do modelo e utiliza a menor ordem proposta pelos testes. O mesmo procedimento foi realizado para as outras saídas. A Tabela 1 mostra também que a ordem estimada pelos critérios não garante a validação dos resíduos, como pode ser constatado para o caso da PIC408 cujos resíduos só se tornaram ruído branco para ordens acima de 15, superior às ordens estimadas por todos os critérios utilizados. Outro caso são os sinais da LIC434 e FIC408 que caso utilizassem a ordem indicadas pelos critérios BIC e HQ, não garantiriam que os resíduos são ruído branco.

O resultado da inferência da estatística f na análise de causalidade de Granger para os 7 sinais é mostrado na Figura 6, sendo que as curvas vermelhas pontilhadas são o limiar do teste F^{α_c} e as curvas pretas contínuas são da estatística f .

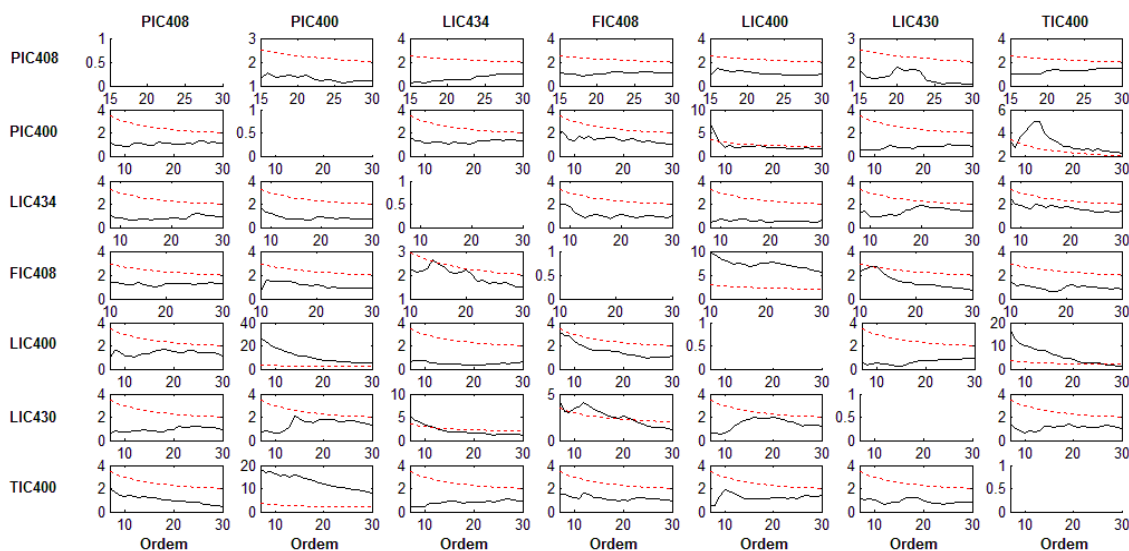


Figura 6. Inferência de causalidade com o aumento da ordem do modelo para os 7 sinais

A utilização dos critérios para a escolha da ordem do modelo juntamente com a verificação da sua qualidade pela análise de resíduos fornecem resultados bastante consistentes na maioria dos casos quando se utiliza a estatística f para inferir a causalidade entre os sinais. Porém, ao se analisar algumas relações de causalidade, estas relações passam a se tornar duvidosas quando se avalia o comportamento da estatística de teste em relação ao seu limiar com o aumento da ordem de estimação dos modelos.

Como pode ser observado na Figura 6, existem relações de causalidade bem consistentes como, por exemplo, a relação da $PIC408 \rightarrow LIC434$, que indica sempre ausência de causalidade. Outra relação consistente, por exemplo, é a relação na direção $PIC400 \rightarrow TIC400$, que indica sempre relações de causalidade significantes. Porém, algumas direções de causalidade possuem uma variação no comportamento quando se varia a ordem do modelo e são estas direções que merecem uma atenção especial. Quanto se utiliza os critérios BIC e HQ para analisar a relação de causalidade na direção $LIC434 \rightarrow LIC430$, o resultado indica uma relação significativa de causalidade. Porém, ao se utilizar os resultados fornecidos pelo critério de Akaike essa relação de causalidade deixa de existir. A direção $FIC408 \rightarrow LIC430$, indica uma causalidade para ordem igual a 7, não indica para ordem igual a 8, passa a indicar novamente quando se utiliza a ordem do critério de Akaike e alterna o resultado com o aumento da ordem. Na direção $LIC400 \rightarrow PIC400$, percebe-se uma indicação de causalidade quando se utiliza as ordens fornecidas pelos critérios BIC e HQ e uma ausência de indicação quando se utiliza a ordem fornecida pelo critério de Akaike. Esta variação na indicação de causalidade também é percebida nas seguintes direções com o aumento da ordem do modelo: $LIC430 \rightarrow FIC408$, $TIC400 \rightarrow PIC400$, $LIC434 \rightarrow FIC408$ e $TIC400 \rightarrow LIC400$. As relações de causalidade nestas 7 direções passam a se tornar incertas quando este tipo de comportamento acontece mesmo utilizando as ordens indicadas pelos critérios de informações utilizados com o teste de análise sobre os resíduos.

O mapa causal mostrado na Figura 7 resume os resultados obtidos nas análises. As setas pretas indicam relações de causalidade consistentes e as setas vermelhas pontilhadas indicam relações de causalidade duvidosas devido a essas variações de indicação de causalidade com o aumento da ordem do modelo.

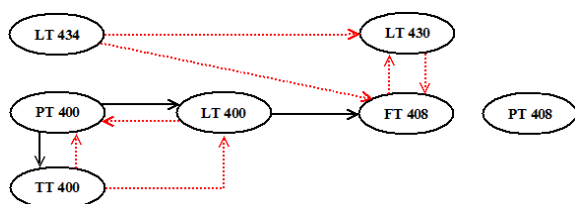


Figura 7. Mapa causal obtido pela análise da Figura 6

Uma proposta sugerida para diminuir essa ambiguidade nos resultados da análise é utilizar a fre-

quência em que as relações de causalidade duvidosas apareceram no intervalo de ordem utilizado. A Tabela 2 mostra a frequência (medida em porcentagem) em que as direções de causalidade duvidosas tiveram indicação ou não de causalidade.

Tabela 2. Frequências de detecções e ausência de detecções de causalidade para as relações duvidosas

Direção	Presença de causalidade (%)	Ausência de causalidade (%)
$LIC434 \rightarrow FIC408$	4,76	95,24
$LIC434 \rightarrow LIC430$	25	75
$FIC408 \rightarrow LIC430$	58,33	41,67
$LIC400 \rightarrow PIC400$	12,5	87,5
$LIC430 \rightarrow FIC408$	4,76	95,24
$TIC400 \rightarrow PIC400$	91,67	8,33
$TIC400 \rightarrow LIC400$	83,33	16,67

Na Tabela 2, os maiores valores de cada direção estão apresentados em negrito, porém a direção $FIC408 \rightarrow LIC430$ apresentou valores muito próximos, o que não permite concluir sobre a existência ou não de uma relação de causalidade, sendo necessário recorrer ao diagrama do processo para obter esta informação. Um novo mapa causal baseado nos resultados da Figura 6 e da Tabela 2 foi obtido e é mostrado na Figura 8.

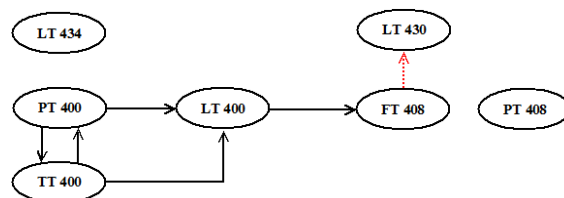


Figura 8. Novo mapa causal obtido após a utilização da frequência de causalidade

Estas relações encontradas podem ser validadas pelo conhecimento que se tem do processo: a diminuição da pressão no superaquecedor ($PT400$) gera uma nova referência para o nível da água no tubulão ($LT400$) para a produção de mais vapor. Esta malha ajusta o ponto de operação da vazão da água ($FIC408$) que vem do reservatório do desaerador cujo nível é controlado pela $LIC430$, sendo percebido aqui que a relação duvidosa na direção $FIC408 \rightarrow LIC430$, mostrada na Tabela 2, de fato é significativa. Ao diminuir a pressão no superaquecedor, a temperatura ($TT400$) também é reduzida pela lei dos gases ideais. O vapor condensado ao deixar a turbina, alimenta o tanque cujo nível é controlado por $LIC434$, e este tanque alimenta o desaerador ($LT430$). A malha $PIC408$ regula a pressão de gás natural da chama piloto para garantir que os combustíveis sejam inflamados no interior da fornalha, desta forma esta variável não está relacionada com nenhuma outra. A análise encontrou a relação de causalidade na direção $TIC400 \rightarrow PIC400$ como significativa, isto pode ser explicada pela equação dos gases ideais, já que um aumento da temperatura de um gás em um recipiente fechado com volume constante gera um aumento da pressão deste gás no reservatório (superaquecedor). A Figura 3 mostra a distância física entre a variável $LT434$ com o tubulão e com as

variáveis ligadas diretamente a este processo, o que pode explicar o motivo para o método não ter apontado nenhuma relação de causalidade desta variável com as demais.

4 Conclusões

Este trabalho teve como proposta mostrar como a escolha da ordem do modelo influencia nos resultados da inferência da estatística f na análise da causalidade de Granger. Foi feita uma comparação de três diferentes critérios de seleção da ordem do modelo baseados na teoria da informação e mostrado como a escolha de um determinado critério pode gerar resultados diferentes nas relações de causalidade quando aplicada a sinais mais complexos pertencentes a um processo industrial, mesmo que os resíduos de estimação dos modelos sejam ruído branco.

A utilização apenas de algum tipo de critério de seleção da ordem de modelos pode não gerar uma topologia confiável, sendo necessários então outros meios para se chegar a esta informação. Pelo método aqui proposto, do total das 42 relações de causalidade, sete foram incertas após a inferência da causalidade de Granger, destas sete, apenas uma continuou ambígua após a aplicação da proposta sugerida em detectar a frequência de ocorrência ou não de causalidade no intervalo de ordem selecionado. As análises e resultados deste trabalho evidenciam os cuidados necessários e indicam procedimentos que podem aumentar a confiança sobre as relações de causalidade detectadas.

Agradecimentos

Agradecemos a Capes e a Petrobrás pelo apoio financeiro por meio da concessão de bolsas e apoio à pesquisa.

Referências Bibliográficas

- Bueno, R. L. S., 2011. *Econometria de Séries Temporais*. Cengage Learning, No. 2, 360 pp.
- Ding, M; Chen, Y; Bressler, S. L., October 2006. *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*. Wiley-VCH, No. 1, pp 437-460.
- Granger, C. W. J., August 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, Vol. 37, No. 3, pp. 424-438.
- Hamilton, J. D., 1994. *Time Series Analysis*. Princeton University Press, No.1, p 305.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer-Verlag New York, No. 2.

- Lutkepohl, H., 2005. *New Introduction to Multiple Times Series Analysis*. Springer-Verlag Berlin Heidelberg, No. 1, 764 pp.
- McDonald, J. H., 2014. *Handbook of Biological Statistics*. Sparky House Publishing, Baltimore, Maryland, No. 3, pp 257-259.
- Russo, F., 2009. *Causality and Causal Modelling in the Social Sciences: Measuring Variations*. New York Springer, Vol 5, No 1. 235 pp.
- Seth, A. K; Barret, A. B; Barnett, L., February 2015. Granger Causality Analysis in Neuroscience and Neuroimaging. *The Journal of Neuroscience*, 35 (8), pp 3293-3297.
- Seth, A. K., February 2010, A MATLAB Toolbox for Granger Causal Connectivity Analysis. *Journal of Neuroscience. Methods*, Vol. 186, pp 262-273.
- Siyal, M. Y., Furqan, M. S., Monir, S, M, G., December 2014. Granger Causality: Comparative Analysis of Implementations for Gene Regulatory Networks. *The 13th International Conference on Control, Automation, Robotics and Vision*, At Singapore.
- Tam, G. H. F., Chang, C., Hung, Y. S., August 2012. Application of Granger causality to gene regulatory network discovery. *Proc. IEEE 6th International Conference on Systems Biology*, Xi'an, China, pp. 232- 239.
- Yang, F., Duan, P., Shah, S.L., Chen, T., 2014. *Capturing Connectivity and Causality in Complex Industrial Processes*. Springer International Publishing, No 1. 99 pp.