

Grafos de Quantis de Séries Temporais com Incertezas

Jonathan G. Ribeiro* Erivelton G. Nepomuceno**
Andriana S. L. O. Campanharo***

* *GCOM - Grupo de Controle e Modelagem, Universidade Federal de São João del-Rei, MG, (e-mail: jgreletrical@gmail.com).*

** *Maynooth University, Maynooth, Ireland (e-mail: erivelton.nepomuceno@mu.ie)*

*** *Departamento de Bioestatística, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, São Paulo (e-mail: andriana.campanharo@unesp.br)*

Abstract: The Quantile Graph (*QG*) method proved to be a promising technique for mapping time series into complex networks, due to the fact that, not only it made possible to use network statistics to characterize time series and time series statistics to characterize networks, but also it is a method with low computational cost, and it is simple to implement. However, in many cases, there is a possibility in which there is some sort of uncertainty in the time series, which implies in some adaptations of the *QG* method. Based on this matter, the goal of this paper is to analyze the consequences caused by the uncertainties which are present in the time series and to define the impact of these changes in the topology and behavior of the constructed network. Uncertainty is considered by changing the threshold of the quantiles and using Monte Carlo simulation. The results express the network's metrics values for a time series with different values of uncertainty, mapped using the proposed method. Network's metrics are compared with no-uncertainty *QG* scenario. The main results found are that the proposed method is able to express the changes in the constructed network caused by different types of uncertainties.

Resumo: O método Grafos de Quantis (*QG*) provou-se uma técnica promissora para o mapeamento de séries temporais em redes complexas, por possibilitar a utilização de estatísticas em redes complexas para a caracterização de séries temporais e vice-versa. Além do método ser de fácil implementação e baixo custo computacional. No entanto, em várias situações há a possibilidade de existirem incertezas nas séries temporais, desta forma, faz-se necessária a adaptação do método *QG* para tais casos. Com base nessa observação, o objetivo deste trabalho foi analisar os efeitos causados pelas incertezas presentes nas séries temporais, e determinar o impacto dessas mudanças na topologia e no comportamento das redes associadas. O impacto causado pelas incertezas é determinado pela mudança nos quantis e utilizando método de Monte Carlo. Os resultados obtidos apresentam os valores de diferentes métricas topológicas, para uma série temporal com diferentes valores de incerteza, mapeada utilizando o método proposto. Os valores das métricas calculadas são comparados aos valores obtidos utilizando o método *QG*. Os principais resultados encontrados mostram que o método proposto consegue replicar nas redes construídas, diferentes mudanças nas séries temporais, causadas por diferentes tipos de incertezas.

Keywords: Complex Networks; Time Series; System Identification; Quantile Graphs; Uncertainty

Palavras-chaves: Redes Complexas; Séries Temporais; Identificação de sistemas; Gráfos de quantis; Incertezas.

1. INTRODUÇÃO

Séries temporais recebem uma atenção especial devido a sua importância em distintas áreas de estudo, como: engenharia, física, fisiologia, biologia, ciências sociais, etc (Gao et al., 2021; de Bézenac et al., 2020; Zhang, 2007; Elgar et al., 2020). Devido a essa importância, modelos que consigam representar e prever os resultados das séries temporais têm sido pesquisados há décadas (Gamboa,

2017), principalmente na abordagem de reconhecimento de padrões e identificação de comportamentos a partir de séries temporais (Wang et al., 2017).

No início do século XXI, o estudo dos processos dinâmicos presentes em séries temporais ganhou mais uma ferramenta, a utilização de redes complexas, introduzida em (Zhang and Small, 2006). Nesse sentido, diversos modelos de representações de séries temporais em redes complexas foram posteriormente propostos, baseados em diferentes

conceitos como, a correlação (Zhang and Small, 2006), grafos de visibilidade (Lacasa et al., 2008), análise de recorrência (Marwan et al., 2009), probabilidade de transição (Nicolis et al., 2005), reconstrução do espaço de fases (Gao and Jin, 2009). A utilização de redes complexas mostra-se como uma ferramenta promissora, pois tais redes são capazes de modelar sistemas que possuem um número elevado de componentes interagindo uns com os outros de maneira complexa.

O Método *QG* (*quantile graph*) (Campanharo et al., 2011) apresenta algumas características bem interessantes para a análise de sistemas dinâmicos a partir de séries temporais. Primeiramente, o método é de fácil implementação, possui apenas um parâmetro livre que é o número de quantis/vértices, dessa forma o método *QG* é uma alternativa eficaz para a análise de séries temporais extensas.

Os métodos citados anteriormente utilizam-se de simulações computacionais ou de medições de sistemas reais para obter as séries. Porém, não há informação quanto a análise dos erros decorrentes dessas simulações ou das medições, ou seja, a série temporal utilizada para construção de uma rede complexa pode conter diferentes incertezas, que podem comprometer a topologia da rede obtida no que diz respeito ao mapeamento dessas incertezas. O propósito deste estudo é desenvolver uma metodologia baseada no método *QG* para mapear séries que apresentam incertezas. Na abordagem proposta, é utilizado o método de Monte Carlo unido ao método *QG* para determinar os impactos causados pelas incertezas na topologia da rede associada.

O presente trabalho é organizado da seguinte forma. A Seção 2, aborda os conceitos preliminares necessários para a implementação da metodologia. A Seção 3, apresenta o método proposto. A Seção 4, apresenta os resultados do método proposto, aplicado ao sistema apresentado na Seção 2. As conclusões são apresentadas na Seção 5.

2. CONCEITOS PRELIMINARES

2.1 Sistema

Dentre os circuitos que apresentam comportamento caótico, é possível encontrar um modelo que pode ser descrito por um sistema de equações diferenciais de ordem 3. A solução desta equação pode ser encontrada em (Sprott, 1997). A função não linear J chamada “jerk” descreve a derivada de terceira ordem de x , que corresponde à primeira derivada de aceleração em um sistema mecânico (Sprott, 2011). Este sistema é descrito da seguinte forma:

$$\ddot{x} = J(\ddot{x}, \dot{x}, x) \quad (1)$$

onde a função J é dada por $J = -A\ddot{x} - x \mp \dot{x}^2$, que tem sido considerada como a equação diferencial ordinária mais simples com uma não linearidade quadrática que apresenta um resultado caótico. O termo não linear de J é \dot{x} . Nessa equação, A é o parâmetro de bifurcação com comportamento caótico na faixa de $2.0168 < A < 2.0577$ (Sprott, 2011). A Figura 1 apresenta o esquema e a implementação do circuito, na Figura 2 se encontra a representação da série temporal utilizada neste trabalho.

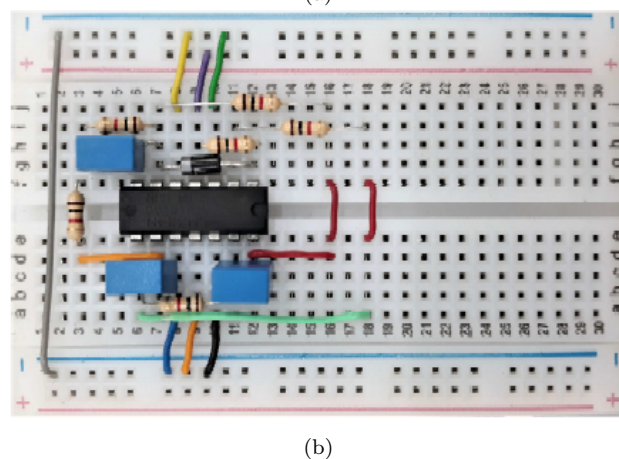
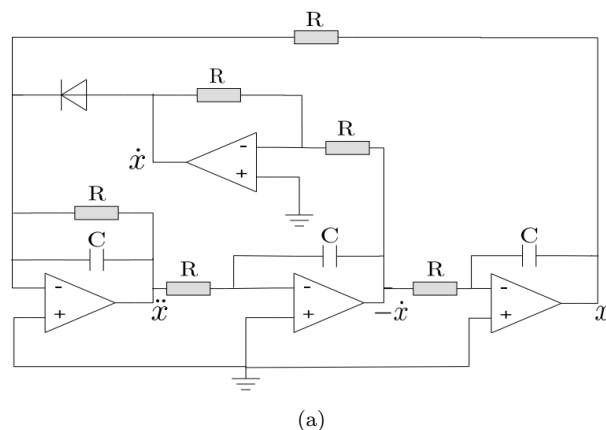


Figura 1. Esquema e implementação do circuito. (a) Circuito Caótico Jerk proposto por Sprott (2011). Os componentes utilizados são: 4 resistores de $1\text{ k}\Omega$, amplificadores operacionais (LM324), 3 capacitores de $1\mu\text{F}$ e 1 diodo 1N4001. (b) Implementação do circuito Jerk (Figura 1a). (Nazaré et al., 2020)

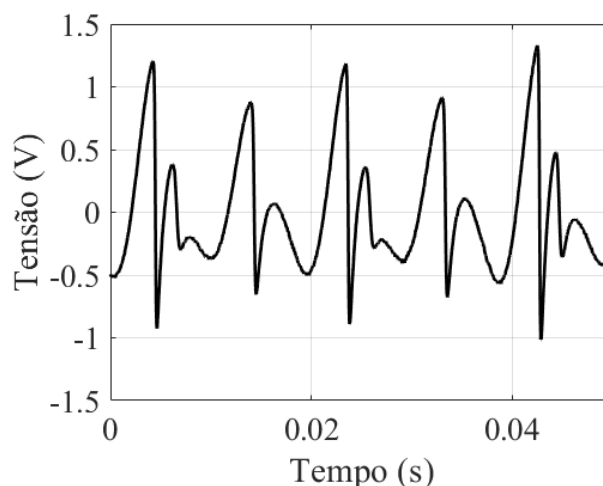


Figura 2. Série temporal do Circuito Caótico Jerk medida em \ddot{x} (Nazaré et al., 2020)

2.2 Aritmética Intervalar

A partir de um modelo numérico simples, cada grandeza no domínio dos números reais pode ser representada por

intervalos numéricos, sendo possível realizar as operações de adição, subtração, multiplicação, de tal modo que cada intervalo computado contenha o valor desconhecido da grandeza.

Cada variável pode ser definida pela Equação (2), em que a variável intervalar X possui limites fechados mínimo \underline{X} e máximo \overline{X} (Moore et al., 2009).

$$X = [\underline{X}, \overline{X}] = \{x \in R : \underline{X} \leq x \leq \overline{X}\}. \quad (2)$$

Segundo (Alefeld and Herzberger, 2012) diz-se que um intervalo é degenerado se $\underline{X} = \overline{X}$. Assim, um número real pode ser representado por um intervalo degenerado como $x = [x, x]$. Além disso, se $x \in [\underline{X}, \overline{X}]$, define-se o tamanho e o ponto médio do intervalo pelas Equações (3) e (4).

$$\omega(X) = \underline{X} - \overline{X}, \quad (3)$$

$$m(X) = \underline{X} + \frac{\omega(X)}{2}, \quad (4)$$

Além dessas operações básicas que são utilizadas neste trabalho, outras podem ser definidas e encontradas na literatura (Moore et al., 2009).

2.3 Incertezas

Neste trabalho, o cálculo das incertezas δ da série temporal é feito por meio da equação:

$$\delta(i) = X(i) \times erro \quad i = 1, \dots, T, \quad (5)$$

$\delta(i)$ é o valor da incerteza para cada ponto da série temporal, $X(i)$ é cada amostra da série e *erro* é o erro percentual, presente na série, com variação de 0% a 20%.

Conhecidas as incertezas, é possível calcular os intervalos I da série temporal X utilizando:

$$I = [\underline{I}, \overline{I}] = \begin{cases} \overline{I} = X(i) + \delta(i) \\ \underline{I} = X(i) - \delta(i) \end{cases} \quad i = 1, \dots, T. \quad (6)$$

2.4 Propriedades topológicas

A análise, caracterização e diferenciação de redes complexas dependem da utilização de medições que sejam capazes de expressar suas características topológicas mais relevantes. Nesse trabalho foram utilizados 5 caracterizadores baseados na matriz de adjacências A e na matriz de transição de Markov W , a saber, o Comprimento do menor caminho médio (L), Comprimento do salto médio (Δ), Coeficiente de agrupamento (CC), Modularidade (M_o) e Grau de intermediação (BC). As implementações desses caracterizadores foram disponibilizadas por (Pineda et al., 2020; ?), de forma gratuita.

Comprimento do menor caminho médio: Em uma rede direcionada e com peso, o comprimento do menor caminho entre dois vértices n_i e n_j é o caminho que minimiza a soma dos pesos das arestas que formam aquele caminho (Costa et al., 2007). O valor do comprimento médio do menor

caminho L de uma rede conectada é calculado a partir da média dos valores de L para todos os pares de vértices da rede, a única exceção são as auto-ligações, que não são consideradas. Os menores caminhos entre os vértices são expressos por uma matriz de distâncias D , na qual os elementos d_{ij} representam o valor do menor caminho entre os vértices n_i e n_j . Com a matriz D construída o menor caminho médio de uma rede complexa conectada pode ser obtida pela equação:

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}, \quad (7)$$

N sendo o número de vértices da rede e $\sum_{i \neq j} d_{ij}$ a soma de todos os elementos de D .

Comprimento do salto médio: A partir da matriz de adjacência A , é possível obter a matriz de transição de Markov W , de tal forma que $\sum_j w_{ij} = 1$. Portanto, cada elemento $w_{ij} = a_{ij} / \sum_{j=1}^N a_{ij}$ representa a probabilidade de transição do vértice n_i para o vértice n_j . Assim, é possível realizar um caminho aleatório sobre a rede R e calcular o comprimento do salto médio Δ (Campanharo and Ramos, 2016), definido como:

$$\Delta = \frac{1}{S} \sum_{s=1}^S \delta_s(i, j), \quad (8)$$

s são os saltos de comprimentos $\delta_t(i, j) = |i - j|$ onde $i, j = 1, \dots, N$ são os índices dos vértices definidos por W . O comprimento do salto médio, pode ser calculado com base na expressão:

$$\Delta = \frac{1}{N} tr(GW^T), \quad (9)$$

W^T é a transposta de W , G é a matriz $N \times N$ com os elementos $g_{i,j} = |i - j|$ e tr é o traço da matriz resultante.

Coeficiente de agrupamento: Em redes complexas, a tendência de uma rede em criar vizinhanças fortemente conectadas pode ser calculada pelo coeficiente de agrupamento (Fagiolo, 2007). O coeficiente de agrupamento foi determinado com base em estruturas de triângulos formadas por vértices da rede, em outras palavras, se o vértice n_i está conectado aos vértices n_j e n_k , existe uma grande chance dos vértices n_j e n_k também estarem conectados (Latora et al., 2017). Na literatura existem várias tentativas de determinar um coeficiente de agrupamento destinado a redes com peso. Neste trabalho o método escolhido é o que foi apresentado em (Clemente and Grassi, 2018), em que o coeficiente de agrupamento do vértice n_i é dado por:

$$CC_i = \frac{[(A + A^T)(A_b + A_b^T)]_{ii}}{2[s_i^{tot}(d_i^{tot} - 1) - 2s_i^{\leftrightarrow}]}, \quad (10)$$

A_b é a matriz de adjacências binária, A é a matriz de adjacências com peso, d_i^{tot} é o grau de conectividade do vértice n_i , s_i^{tot} é a força de conectividade do vértice n_i e $2s_i^{\leftrightarrow}$ é a força de conectividade relacionada às arestas bilaterais entre o vértice n_i e seus vizinhos.

Grau de intermediação (Betweenness Centrality): O grau de intermediação é uma medida de centralidade baseada

nos menores caminhos. Para cada par de vértices n_i e n_j em uma rede conectada, existe ao menos um caminho que minimiza a distância de conexão entre esses vértices. Assim o grau de intermediação de um vértice n_u quantifica o número de caminhos que minimizam a soma dos pesos das arestas entre dois vértices, que passam pelo vértice n_u . O grau de intermediação (BC) de um vértice n_u é dado por (Newman, 2005):

$$BC_{n_u} = \sum_{ij} \frac{\sigma(n_i, n_u, n_j)}{\sigma(n_i, n_j)}, \quad (11)$$

$\sigma(n_i, n_u, n_j)$ é o número de menores caminhos entre os vértices n_i e n_j que passam pelo vértice n_u , $\sigma(n_i, n_j)$ é o número total de menores caminhos entre os vértices n_i e n_j , e a soma total é calculada para todos os pares distintos de vértices n_i e n_j (Newman, 2005; Costa et al., 2007). O grau de intermediação para a rede completa (BC), é definida como a média do grau de intermediação de todos os vértices.

Modularidade: A modularidade é uma medida topológica da rede que quantifica a capacidade da rede de se dividir em comunidades (módulos). Redes complexas que possuem alto grau de modularidade possuem vértices fortemente conectados com seus vizinhos que estão dentro do mesmo módulo, porém, possuem pouca conectividade com vértices localizados em outros módulos (Guimera et al., 2004). Para uma comunidade c a modularidade M_c pode ser definida por:

$$M_c = \frac{1}{2m} \sum_{i,j} \left[w_{ij} - \frac{s_i s_j}{2m} \right] \phi(c_i, c_j), \quad (12)$$

w_{ij} é o peso da aresta entre os vértices n_i e n_j , $s_i = \sum_j w_{ij}$ é a soma total dos pesos das arestas associadas ao vértice n_i , c_i é a comunidade na qual o vértice n_i foi alocado, $m = \frac{1}{2} \sum_{i,j} w_{ij}$ e a função $\phi(c_i, c_j)$ é 1 quando $c_i = c_j$ e 0 quando $c_i \neq c_j$.

2.5 O método QG (grafo de quantis)

O método QG, proposto por Campanharo *et al.* (Campanharo et al., 2011) mapeia uma série temporal $X = \{x(t) | t \in \mathbb{N}, x(t) \in \mathbb{R}\}$ em uma rede $\mathcal{G} = \{\mathcal{N}, \mathcal{L}\}$, com um conjunto de vértices \mathcal{N} e de arestas \mathcal{L} . O método QG considera uma discretização de X baseada em quantis que não é sensível à distribuição de seus valores. Através de uma escolha adequada de Q (número de quantis), cada quantil da série temporal possui um vértice correspondente e as arestas da rede construída representam as transições entre os quantis. Para calcular os quantis Q de uma série temporal, é feito um ordenamento crescente dos valores da série temporal, assim podendo dividir os valores em Q intervalos de tamanhos idênticos (Campanharo et al., 2011). O cálculo de cada quantil é dado por:

$$q_i = X \left(\frac{T \times i}{Q} \right), \quad (13)$$

para cada $i = 1, 2, \dots, Q$. Com os q_i calculados, os valores de X são alocados em seus respectivos quantis q_i . Em seguida, atribui-se cada quantil q_i para um vértice $n_i \in \mathcal{N}$.

Quando dois pontos x_t e x_{t+k} estão contidos nos quantis q_i e q_j respectivamente, uma aresta direcionada é criada na rede complexa, sendo $t = 1, 2, \dots, T$ e as diferenças de tempos $k = 1, \dots, k_{max} < T$, sendo T o número de elementos da série temporal. Os pesos associados as arestas da rede w_{ij}^k contidos na matriz de adjacências com peso A_k , possuem valor equivalente à frequência de ocorrência da transição de q_i e q_j . Dessa forma, quando uma transição é repetida, seu peso w_{ij}^k é incrementado em $w_{ij}^k = w_{ij}^k + 1$ Campanharo et al. (2020). De modo a alcançar um valor aceitável para Q a aproximação $Q \approx 2T^{1/3}$ é utilizada (Campanharo et al., 2011). A rede complexa obtida pelo método QG é conectada, com peso, com direção e possui Q vértices.

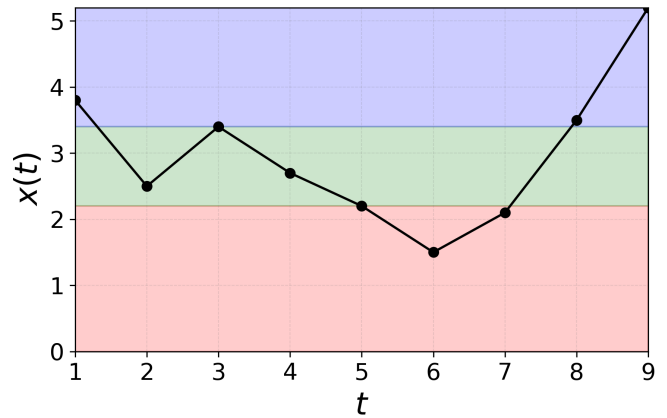


Figura 3. Exemplo do método QG para uma série com $T = 9, Q = 3$ e $k = 1$. Os intervalos que delimitam os quantis são: $[x(1), x(4)[$, $[x(4), x(7)[$, $[x(7), x(9)]$, *i.e.*, $[1,5, 2,5[$, $[2,5, 3,4[$, $[3,4, 5,2]$. Os quantis são mapeados em uma rede com $N = 3$ vértices, com as arestas sendo $(1, 1, 2), (1, 3, 1), (2, 1, 1), (2, 2, 2), (3, 2, 1), (3, 3, 1)$.

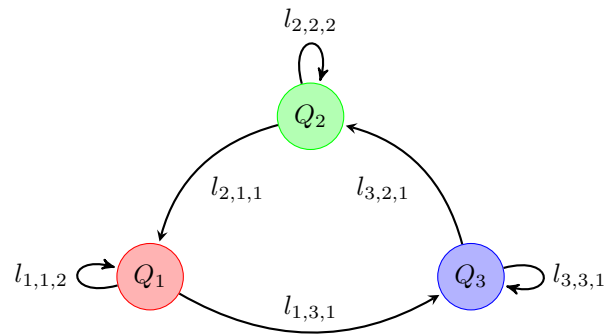


Figura 4. Rede construída a partir das ligações apresentadas na Figura 3, as cores das ligações indicam o vértice de partida da aresta.

Nas Figuras 3 e 4 está ilustrado o processo de construção uma rede complexa a partir de uma série temporal, utilizando o método QG. A série temporal X com $T = 9$ amostras, dividida em $Q = 3$ quantis e $k = 1$. Como $Q = 3$ quantis, os intervalos que delimitam os quantis são: $[x(1), x(3)]$, $[x(4), x(5)]$, $[x(7), x(9)]$, *i.e.*, $[1,5, 2,2]$, $[2,5, 3,4]$, $[3,5, 5,2]$. Teremos 3 elementos da série contidos em cada um dos quantis, sendo $Q_1 = \{1, 5, 2, 1, 2, 2\}$,

$Q_2 = \{2, 5, 2, 7, 3, 4\}$, $Q_3 = \{3, 5, 3, 8, 5, 2\}$. As arestas resultantes das transições entre os quantis, para $k = 1$, são apresentadas na forma (n_i, n_j, w_{ij}^k) , com vértice de partida n_i , vértice de chegada n_j e peso w_{ij}^k equivalente à frequência de transições entre q_i e q_j . Por fim, as arestas obtidas: $(1, 1, 2)$, $(1, 3, 1)$, $(2, 1, 1)$, $(2, 2, 2)$, $(3, 2, 1)$, $(3, 3, 1)$.

3. O MÉTODO QG COM INCERTEZAS (UQG):

Nesta seção é apresentada a metodologia para adaptar o método QG proposto por Campanharo et al. (2011) para incorporar incertezas nas séries temporais, assim criando um novo método chamado Grafos de quantil com incerteza UQG (*Uncertain Quantile Graph*).

O método QG mapeia uma série temporal, porém, considerando a existência de incertezas na série, fazem-se necessárias algumas modificações. Considerando a série temporal apresentada na Figura 3, teremos 3 elementos em cada um dos quantis. Caso a série possua alguma incerteza, cada ponto $x(t)$ da série se transforma em um intervalo $I(t)$. Nesse sentido, os intervalos I são calculados com base na Equação 6 e o valor médio de cada ponto da série temporal é calculado com base na Equação 4. Com o valor médio $m(X)$ é possível aplicar o método QG pela primeira vez e obter uma estimativa dos valores que delimitam os quantis (Figura 5).

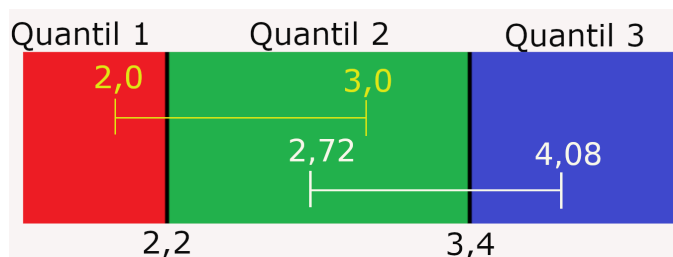


Figura 5. O intervalo $I(2) = [2, 3]$ (em amarelo), está localizado no Q_1 e no Q_2 simultaneamente. O intervalo $I(3) = [2,72, 4,08]$ (em branco), está localizado no Q_2 e no Q_3 simultaneamente

Mapeando os valores médios $m(I)$, são obtidos os valores delimitadores dos 3 quantis. O primeiro quantil Q_1 possui $[1,5, 2,2]$ como seus delimitadores, ou seja, se um valor da série é menor que 2,2, ele pertence a Q_1 . Q_2 possui $]2,2, 3,4]$ como seus delimitadores, ou seja, se um valor da série é menor que 3,4 e maior que 2,2, ele pertence a Q_2 . Q_3 possui $]3,4, 5,2]$ como seus delimitadores, ou seja, se um valor da série é menor que 5,2 e maior que 3,4, ele pertence a Q_3 .

Com $\delta = 20\%$ para a série X , todos os pontos da série se tornam intervalos com 20% de incerteza, o que resulta em alguns intervalos localizados em mais de um quantil simultaneamente. O ponto $x(2) = 2,5$, por exemplo, se transforma no intervalo $I(2) = [2, 3]$ (Figura 5), o limite inferior $\underline{I}(2) = 2 < 2,2$ assim $\underline{I}(2) \in Q_1$ e o limite superior $\bar{I}(2) = 3$ e $2,2 < \bar{I}(2) < 3,4$ assim $\bar{I}(2) \in Q_2$. O ponto $x(3) = 3,4$, se transforma no intervalo $I(3) = [2,72, 4,08]$ (Figura 5), o limite inferior $\underline{I}(3) = 2,72$ e $2,2 < \underline{I}(3) < 3,4$ assim $\underline{I}(3) \in Q_2$, o limite superior $\bar{I}(3) = 4,08$ e $\bar{I}(3) > 3,4$ assim $\bar{I}(3) \in Q_3$. Em resumo, utilizando o método QG, os

intervalos podem estar localizados em mais de um único quantil.

Para o caso apresentado no parágrafo anterior no qual a série é intervalar, os intervalos $I(2)$ e $I(3)$, estão localizados em dois quantis diferentes simultaneamente, existem 4 possibilidades de arestas, sendo: $(1, 2, 1)$, $(1, 3, 1)$, $(2, 2, 1)$, $(2, 3, 1)$. Desta forma, se faz necessário um procedimento para definir em qual quantil se encontra cada intervalo, e assim se defina apenas uma aresta para cada transição entre os intervalos.

Com o objetivo de definir a qual quantil pertence cada intervalo, é utilizado o método de Monte Carlo. Primeiramente é criada uma nova série temporal Y :

$$Y(i) = \underline{I}(i) + (\bar{I}(i) - \underline{I}(i)) \times rand, \quad i = 1, \dots, T, \quad (14)$$

sendo $rand$ um número aleatório entre 0 e 1. O valor de $Y(i)$ é estipulado como um valor aleatório entre $\underline{I}(i)$ e $\bar{I}(i)$. Com a série Y construída, a mesma é mapeada utilizando o método QG. O processo é repetido n vezes, e a cada iteração é calculada a matriz de adjacências com peso A . Após n iterações a solução é definida como $A = mediana([A_i, \dots, A_n])$.

Algoritmo 1 Mapeamento UQG

Input: Série temporal X com T amostras, e o número de quantis Q , k_{max} e n

Output: matriz de adjacências com peso A

- 1: **for** $i = 1, \dots, T$ **do**
- 2: Calcular as incertezas $\delta(i)$, $I(i) = [\underline{I}(i), \bar{I}(i)]$, $m(I(i))$.
- 3: **end for**
- 4: **for** $i = 1, \dots, n$ **do**
- 5: **for** $j = 1, \dots, T$ **do**
- 6: $Y(j) \leftarrow \underline{I}(j) + (\bar{I}(j) - \underline{I}(j)) \times rand()$
- 7: **end for**
- 8: Aplicar o método QG na série Y
- 9: **end for**
- 10: $A \leftarrow mediana([A_i, \dots, A_n])$
- 11: **return** A

4. RESULTADOS

Para a série X com $T = 1000$ (Figura 2), foi definido $Q = 2 \times (T^{1/3}) \approx 20$. Em seguida, a série X foi mapeada com base no procedimento descrito na Seção 3. O algoritmo foi aplicado para 5 situações distintas, cada uma associada a uma das incertezas, a saber, $\delta = 0\%, 5\%, 10\%, 15\%, 20\%$.

Com o uso do Algoritmo (1) foi possível extrair a matriz de adjacências com peso A e a matriz de transições de Markov W (Normalização de A), para cada implementação de Y . Resultando em duas listas com $n = 1000$ elementos, contendo os valores de A e W . Em seguida, foi calculado a $mediana([A_i, \dots, A_n])$ e $mediana([W_i, \dots, W_n])$, assim determinando os valores de A e W para série X com intervalos. Dessa forma, foi possível calcular os valores dos caracterizadores para cada uma dos 5 valores de δ .

As Figuras 6a e 6b apresentam a matriz de transição de Markov W calculada utilizando o método UQG, com $\delta = 0\%$ e $\delta = 20\%$, respectivamente. Na comparação entre

Tabela 1. Comparação entre o método QG e o método UQG .

δ	Método UQG					Método QG				
	0%	5%	10%	15%	20%	0%	5%	10%	15%	20%
L	22,13	18,41	16,22	9,53	6,69	22,13				
Δ	0,39	0,44	0,55	0,68	0,82	0,39				
CC	0,41	0,45	0,48	0,52	0,57	0,41				
M_o	0,75	0,74	0,72	0,69	0,66	0,75				
BC	63,75	61,84	58,57	55,75	53,52	63,75				

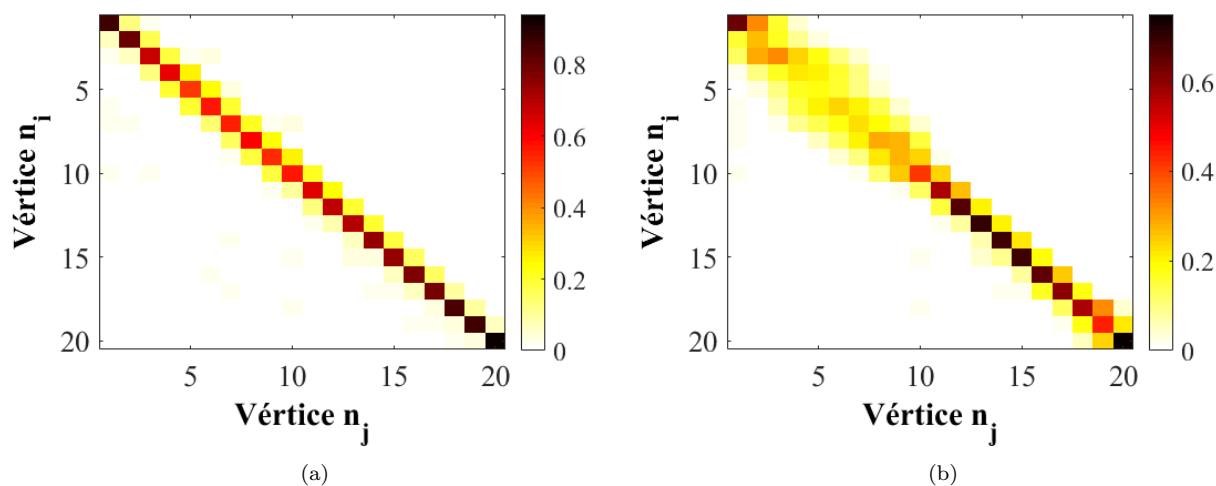


Figura 6. Matriz de transição de Markov W com $Q = 20$ e $\delta = 20\%$, (a) utilizando o método QG , (b) utilizando o método UQG

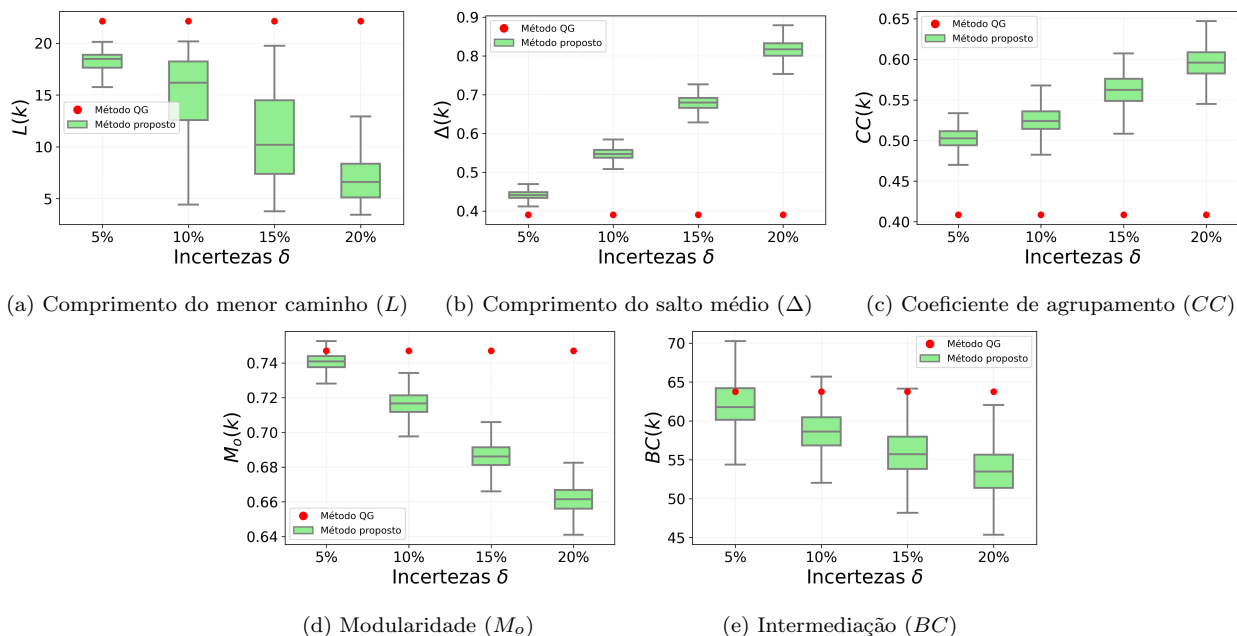


Figura 7. Boxplot do menor caminho ($L(k)$) com $k_{max} = 3$, Salto ($\Delta(k)$) com $k_{max} = 1$, Intermediação ($BC(k)$) com $k_{max} = 1$, Modularidade ($M_o(k)$) com $k_{max} = 1$, Coeficiente de agrupamento ($CC(k)$) com $k_{max} = 1$. Os boxplot foram calculados utilizando o método UQG , com incertezas iguais a 5%, 10%, 15%, 20%. Os k_{max} foram escolhidos para maximizar as variações entre os caracterizadores (Pineda et al., 2020). A linha vermelha é o valor do caracterizador calculado utilizando o método QG .

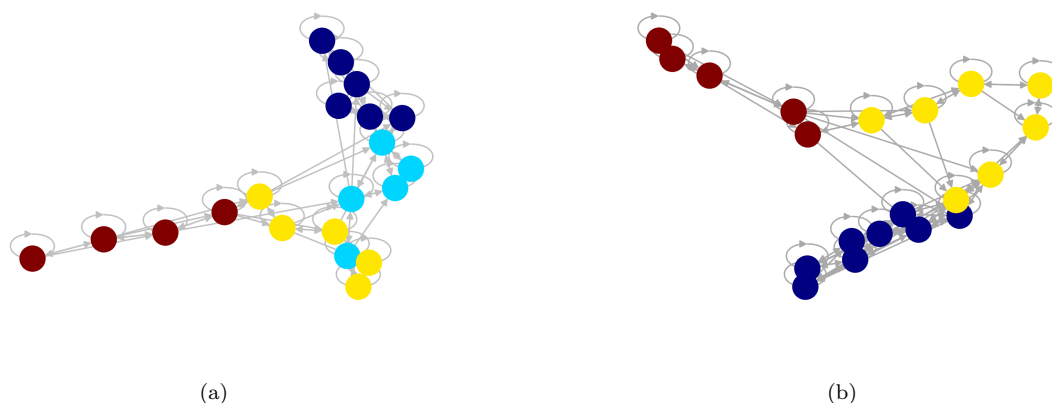


Figura 8. Redes complexas construídas com $Q = 20$, $T = 1000$ e $\delta = 20\%$, (a) para $\delta = 0\%$. (b) para $\delta = 20\%$.

a Figura 6a e a Figura 6b, percebe-se que a matriz W da primeira figura é basicamente formada por sua diagonal principal e por elementos bem próximos dela, indicando que a topologia da rede apresenta número elevado de auto-ligações, ou seja, o vértice de chegada é o mesmo de partida. Já na segunda matriz W , as ligações que se encontravam na diagonal principal agora estão mais espalhadas, indicando uma mudança na topologia da rede, que pode ser confirmada pela análise de seus caracterizadores (Figura 7).

Com A e W calculadas para as diferentes δ , é possível calcular os valores das métricas da rede para os diferentes casos. A Figura 7 apresenta o *boxplot* para 5 caracterizadores diferentes. Os *boxplot* demonstram que, quando a série é mapeada utilizando o método QG , os valores dos caracterizadores são constantes, independentemente do valor de δ . Porém, quando a série é mapeada utilizando o método UQG , os valores dos caracterizadores variam conforme a magnitude das incertezas, ou seja, com o aumento de δ os valores dos caracterizadores tendem a se distanciar dos valores obtidos pelo método QG .

Quanto ao comportamento dos caracterizadores utilizados nesta análise, a Figura 7 apresenta suas variações à medida que as incertezas se ampliam. O comprimento do menor caminho médio L apresentou um comportamento de queda em seu valor, correspondente ao aumento das incertezas, esse comportamento pode ser explicado por meio das mudanças na topologia da rede, apresentadas na Figura 6, onde a matriz W , para $\delta = 0\%$, possui a maioria de se seus elementos próximos da diagonal principal, isso indica que os vértices possuem ligações mais fortes com poucos vértices, já W para $\delta = 20\%$, é uma matriz mais esparsa, indicando que os vértices da rede construída possuem ligações mais fracas, porém as ligações são com vértices mais diversificados. Essa diversificação das ligações faz com que, por exemplo, surjam um caminho direto entre o vértice n_i e n_j , que antes passavam por um terceiro vértice n_k .

O grau de Intermediação BC também apresentou queda, como BC de um vértice determina o número de menores caminhos que passam por ele, conforme a rede diversifica suas ligações é esperado que alguns *hubs* diminuam sua relevância. A Modularidade M_o também apresentou queda,

a modularidade da rede é maior quando os vértices da mesma possuem ligações fortes com um número pequeno de vértices, assim criando diferentes comunidades na rede, explicando a queda do valor de M_o a medida que as ligações da rede se diversificam. O comprimento do salto médio Δ aumentou, como foram rearranjadas as ligações entre os vértices, existem novos caminhos que podem ser percorridos, resultando em uma maior probabilidade de transição entre os vértices. O Coeficiente de agrupamento CC também apresenta valores crescentes à medida que δ aumenta, pois o coeficiente de agrupamento é medido pela formação de estruturas triangulares entre os vértices, com a formação de novas ligações entre os vértices, o aumento do CC é esperado. Outra constatação feita foi que o comportamento da rede mediante ao aumento das incertezas δ se provou similar ao comportamento de redes mundo pequeno quando submetidas a uma diminuição da probabilidade de conexão p (Watts and Strogatz, 1998).

A Tabela 1 apresenta a comparação entre os valores dos caracterizadores da rede, utilizando o método QG e o método UQG , para $\delta = 0\%$, 5% , 10% , 15% , 20% . À medida que o valor de δ aumenta, a diferença entre os resultados dos dois métodos também se amplia. A Figura 8 apresenta as redes construídas para $\delta = 0\%$ e $\delta = 20\%$. Quando $\delta = 0\%$ a rede possui 4 comunidades diferentes, representadas pelas 4 cores dos vértices, à medida que as incertezas se ampliam, o número de vértices nas comunidades é alterado, até que para $\delta = 20\%$, uma das comunidades deixa de existir, o que é mais uma demonstração do impacto das incertezas no mapeamento da série temporal.

5. CONCLUSÃO

Na análise da série temporal do circuito *jerk* com incertezas constantes, foi demonstrado que conforme as incertezas aumentavam as diferenças entre os resultados do método QG e do método UQG se acentuavam. Os caracterizadores calculados utilizando o método QG tradicional se mantinham constantes para qualquer valor de incerteza. Porém, quando o método UQG foi aplicado, os valores obtidos variavam proporcionalmente em relação a incrementos nas incertezas. Portanto o método UQG provou-se capaz de construir redes que refletem as variações nas incertezas das séries. Outro ponto a ser destacado, diz respeito a

magnitude das mudanças causadas pelas variações nas incertezas. Incrementos de 5% nos valores das incertezas causaram mudanças expressivas nos caracterizadores da rede construída, o que demonstra que as incertezas constantes são muito impactantes, pois todas as amostras da série são afetadas de maneira semelhante, desta forma, qualquer mudança nos valores dessas incertezas resultam em grandes mudanças na rede construída.

Em resumo, os resultados mostraram que o método *UQG* proposto é uma alternativa viável para os casos em que existem incertezas presentes nas séries temporais analisadas, pois se demonstrou capaz de computar os impactos causados pelas incertezas na topologia da rede construída.

Em trabalhos futuros, sugere-se criar uma operação de mapeamento inverso, ou seja, um mapeamento bilateral para séries intervalares.

6. AGRADECIMENTOS

Os autores agradecem a Universidade Federal de São João del-Rei pelo apoio financeiro

REFERÊNCIAS

- Alefeld, G. and Herzberger, J. (2012). *Introduction to interval computation*. Academic press.
- Campanharo, A.S., Doescher, E., and Ramos, F.M. (2020). Application of quantile graphs to the automated analysis of eeg signals. *Neural Processing Letters*, 52(1), 5–20.
- Campanharo, A.S. and Ramos, F.M. (2016). Hurst exponent estimation of self-affine time series using quantile graphs. *Physica A: Statistical Mechanics and its Applications*, 444, 43–48.
- Campanharo, A.S., Siner, M.I., Malmgren, R.D., Ramos, F.M., and Amaral, L.A.N. (2011). Duality between time series and networks. *PloS one*, 6(8), e23378.
- Clemente, G.P. and Grassi, R. (2018). Directed clustering in weighted networks: A new perspective. *Chaos, Solitons & Fractals*, 107, 26–38.
- Costa, L.d.F., Rodrigues, F.A., Travieso, G., and Villas Boas, P.R. (2007). Characterization of complex networks: A survey of measurements. *Advances in physics*, 56(1), 167–242.
- de Bézenac, E., Rangapuram, S.S., Benidis, K., Bohlke-Schneider, M., Kurle, R., Stella, L., Hasson, H., Gallinari, P., and Januschowski, T. (2020). Normalizing kalman filters for multivariate time series analysis. *Advances in Neural Information Processing Systems*, 33, 2995–3007.
- Elgar, F.J., Stefaniak, A., and Wohl, M.J. (2020). The trouble with trust: Time-series analysis of social capital, income inequality, and covid-19 deaths in 84 countries. *Social Science & Medicine*, 263, 113365.
- Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E*, 76(2), 026107.
- Gamboa, J.C.B. (2017). Deep learning for time-series analysis. *arXiv preprint arXiv:1701.01887*.
- Gao, Z., Dang, W., Wang, X., Hong, X., Hou, L., Ma, K., and Perc, M. (2021). Complex networks and deep learning for eeg signal analysis. *Cognitive Neurodynamics*, 15(3), 369–388.
- Gao, Z. and Jin, N. (2009). Complex network from time series based on phase space reconstruction. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(3), 033137.
- Guimera, R., Sales-Pardo, M., and Amaral, L.A.N. (2004). Modularity from fluctuations in random graphs and complex networks. *Physical Review E*, 70(2), 025101.
- Lacasa, L., Luque, B., Ballesteros, F., Luque, J., and Nuno, J.C. (2008). From time series to complex networks: The visibility graph. *Proceedings of the National Academy of Sciences*, 105(13), 4972–4975.
- Latora, V., Nicosia, V., and Russo, G. (2017). *Complex networks: principles, methods and applications*. Cambridge University Press.
- Marwan, N., Donges, J.F., Zou, Y., Donner, R.V., and Kurths, J. (2009). Complex network approach for recurrence analysis of time series. *Physics Letters A*, 373(46), 4246–4254.
- Moore, R.E., Kearfott, R.B., and Cloud, M.J. (2009). *Introduction to interval analysis*. SIAM.
- Nazaré, T.E., Nepomuceno, E.G., Martins, S.A., and Butusov, D.N. (2020). A note on the reproducibility of chaos simulation. *Entropy*, 22(9), 953.
- Newman, M.E. (2005). A measure of betweenness centrality based on random walks. *Social networks*, 27(1), 39–54.
- Nicolis, G., Cantu, A.G., and Nicolis, C. (2005). Dynamical aspects of interaction networks. *International Journal of Bifurcation and Chaos*, 15(11), 3467–3480.
- Pineda, A.M., Ramos, F.M., Betting, L.E., and Campanharo, A.S. (2020). Quantile graphs for eeg-based diagnosis of alzheimer’s disease. *Plos one*, 15(6), e0231169.
- Sprott, J. (1997). Some simple chaotic jerk functions. *American Journal of Physics*, 65(6), 537–543.
- Sprott, J.C. (2011). A new chaotic jerk circuit. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 58(4), 240–243.
- Wang, W., Tang, M., Stanley, H.E., and Braunstein, L.A. (2017). Unification of theoretical approaches for epidemic spreading on complex networks. *Reports on Progress in Physics*, 80(3), 036603.
- Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of ‘small-world’ networks. *nature*, 393(6684), 440–442.
- Zhang, G.P. (2007). A neural network ensemble method with jittered training data for time series forecasting. *Information Sciences*, 177(23), 5329–5346.
- Zhang, J. and Small, M. (2006). Complex network from pseudoperiodic time series: Topology versus dynamics. *Physical review letters*, 96(23), 238701.