

Utilização de Modelos Cognitivos na Previsão de Séries Temporais

Use of Cognitive Models in Time Series Forecast

Ademir B. S. Neto
Universidade Federal Rural de
Pernambuco
Rua Manoel de Medeiros -
Dois Irmãos, Recife
Pernambuco, Brasil
ademir.bsn@gmail.com

Maria da C. M. Batista
Universidade Federal Rural de
Pernambuco
Rua Manoel de Medeiros -
Dois Irmãos, Recife
Pernambuco, Brasil
cecamoraes@gmail.com

Paulo R. A. Firmino
Universidade Federal do Cariri
Rua Ten. Raimundo Rocha -
Cidade Universitária, Juazeiro
do Norte
Ceará, Brasil
praf62@gmail.com

Tiago A. E. Ferreira
Universidade Federal Rural de
Pernambuco
Rua Manoel de Medeiros -
Dois Irmãos, Recife
Pernambuco, Brasil
taef.first@gmail.com

RESUMO

Modelos cognitivos são usados como principais fontes de informação para modelagens de estudos onde dados empíricos são indisponíveis, escassos ou não possuem uma relevância considerável. Os modelos cognitivos são as opiniões de especialistas em relação a determinada variável de interesse com base nos seus conhecimentos e experiências. Existem algumas técnicas que buscam eduzir o conhecimento dos especialistas. Nesse artigo será utilizado o processo de educação do conhecimento. Esse processo é um conjunto de etapas que visa obter a opinião dos especialistas de maneira a melhorar fatores como acurácia das informações e reduzir possíveis vieses que os especialistas possam apresentar. Este trabalho visa comparar o desempenho de modelos cognitivos obtidos através da educação conhecimento e de modelos matemáticos em relação a sua acurácia na previsão de séries temporais.

Palavras-Chave

Modelos Cognitivos, Modelos Matemáticos, Educação do Conhecimento, ARIMA, Séries Temporais

ABSTRACT

Cognitive models are used as the main source of information to modelling problems when empirical data are unavailable, scarce or does not have an considerable relevance. The cognitive models are the expert's opinions related to determi-

ned variable based on his knowledgement and experiences. There are some techniques which aim to elicitate knowledge from experts, the one which will be used in this paper is the elicitation process. The elicitation process is a ensemble of steps that seeks to get the opinion of the expert in a way to improve the accuracy of the information and reduce some possible biases that the experts may show. This paper aim to compare the performance of cognitive models obtained through elicitation of knowledge and mathematical models in relationship to their accuracy in time series forecast.

CCS Concepts

•Applied computing → Forecasting; •Mathematics of computing → Statistical software;

Keywords

Cognitive Models, Mathematical Models, Elicitation of Knowledge, ARIMA, Time Series

1. INTRODUÇÃO

Modelos cognitivos são uma aproximação do processo cognitivo animal (predominantemente humano) para os propósitos de compreensão, predição, dentre outros. Modelos cognitivos podem ser desenvolvidos com ou sem uma arquitetura cognitiva, embora essas duas classes de modelos nem sempre são facilmente distinguíveis [14]. As opiniões de um indivíduo são frutos do processamento e síntese dos componentes da sua cognitividade. Opiniões são muitas vezes a única fonte de informação a respeito de determinado problema.

Essas modelagens frequentistas necessitam muitas vezes de um grande número de dados para descreverem os problemas. A escassez de dados tem se apresentado como um dos principais desafios à modelagem estatística de modelos baseados em dados empíricos [11]. Situações como na análise

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SBSI 2017 June 5th – 8th, 2017, Lavras, Minas Gerais, Brazil

Copyright SBC 2017.

probabilística de risco, que analisam não só a probabilidade de determinado evento ocorrer, mas também as consequências da ocorrência desses eventos, a escassez de dados se dá devido aos tempos de ocorrência dos eventos de interesse serem relativamente elevados dificultando uma base sólida de informação para a modelagem nessa área [2]. A busca por suporte à decisão diante desses tipos de sistemas usualmente envolve relações entre um grande número de variáveis, podendo resultar em sofisticados modelos estatísticos. Como consequência disso, é comum que o analista do problema necessite recorrer a opiniões de especialistas [10]. Nessa mesma linha, de acordo com MacCormack *et al* [24], devido a usual dinâmica dos sistemas, dados históricos não podem suprir os requisitos da análise probabilística de risco, tornando o uso de opiniões de especialistas um recurso indispensável. Um especialista pode ser definido como uma pessoa muito habilidosa que tem muito treino e conhecimento a respeito de um campo específico [25]. O especialista é a fonte de estimativa das informações nos modelos cognitivos. Uma pessoa pode ser considerada um especialista em determinada área quando possuir um conhecimento reconhecido por outros estudiosos dessa mesma área [1]. Uma opinião de especialista pode ser definida como um julgamento formal em algum tópico em que seus conhecimentos são requisitados. Uma opinião de especialista também pode ser entendida como um julgamento ou crença baseado em informações incertas ou conhecimentos. De modo geral, uma opinião é uma avaliação subjetiva, impressão ou estimação da quantidade ou qualidade de uma determinada variável de interesse [8]. Para uma adequada obtenção das opiniões dos especialistas a respeito das variáveis de interesse e suas dependências dos problemas é necessário a adoção de um conjunto de técnicas denominado processo de educação do conhecimento. Processos de educação de conhecimentos são um conjunto de etapas que visam educar (quantificar) o conhecimento de determinado especialista a respeito de algum problema específico. Vários autores subdividem as etapas do processo de educação do conhecimento conforme as suas avaliações da necessidade de cada etapa estar presente ou não no processo [6, 15, 13]. Por tanto, pode-se afirmar que educar opiniões significa interagir com o especialista de uma forma que ele possa transmitir fielmente suas incertezas a respeito de certa variável de interesse [27].

Uma área bastante estudada ao longo dos anos é da previsão de séries temporais. Séries temporais são definidas como um conjunto de observações organizadas em uma ordem cronológica, onde o tempo geralmente é o índice de ordenamento, sendo um parâmetro discreto [19]. A medida de tempo pode variar de acordo com a série que se esteja trabalhando, a escala pode estar em meses, dias, horas, etc [4]. Uma das características mais importantes desse tipo de dados, que possibilita sua modelagem e previsão, é a dependência entre as observações vizinhas [9]. A previsão de séries temporais é baseada tradicionalmente em modelos estatísticos baseados nas dependências entre as observações da série. Um trabalho bastante comentado na literatura a respeito da previsão de séries temporais é o livro de Box e Jenkins [3] onde os autores descrevem os modelos da família ARIMA.

Neste contexto apresentado, o principal objetivo desse trabalho é analisar quantitativamente o desempenho dos modelos cognitivos obtidos através de processos de educação do conhecimento na temática da previsão de séries temporais através de uma comparação com previsões de modelos matemáticos.

2. FUNDAMENTAÇÃO TEÓRICA

Essa Seção aborda as fundamentações teóricas de alguns tópicos relevantes para esse trabalho, são eles: processo de educação do conhecimento, séries temporais e modelos ARIMA. Estes tópicos serão descritos nas subseções a seguir.

2.1 Processo de Educação do Conhecimento

Em processos de educação do conhecimento, cada indivíduo é considerado como uma fonte de informação inerentemente ruidosa, a ser decodificada em quantidades úteis a modelagem estatística. Com isso a proposta passa a ser como quantificar (eduzir) a opinião do especialista, considerando-se não apenas suas crenças, mas também, os ruídos intrínsecos a cognição humana. Para alcançar este objetivo, disciplinas da estatística, psicologia e computação têm se alinhado, conduzindo ao desenvolvimento de técnicas que visam quantificar o conhecimento de especialistas. Processos de educação do conhecimento são um conjunto de técnicas e métodos que tentam educar conhecimento do domínio do especialista, tipicamente através de alguma iteração direta com esse especialista [27]. O objetivo geral do processo de educação do conhecimento é representar as opiniões dos especialistas de forma que elas possam ser implementadas em um modelo computacional. O processo de educação do conhecimento visa, através de um conjunto de etapas, auxiliar o especialista a poder expressar suas incertezas sobre seu conhecimento a respeito de determinada variável. Esse processo se difere da mineração de dados, pois se trata de um processo de construção ou aquisição da informação. Este é um processo criacional e mais complexo do que mineração de informações de um conjunto de dados [12]. A escolha correta das etapas dos processos de educação tem uma grande interferência no resultado final do processo uma vez que quanto mais adequadamente os dados forem educados melhor será o modelo resultante [7]. Educar um modelo a partir de conhecimento de especialistas é inquestionavelmente uma tarefa complexa uma vez que os conhecimentos dos especialistas partem de fontes de incerteza aleatória, epistêmica, tácita e abstrata. Isso envolve conhecimento sobre conhecimento que necessita de metacognição [18].

Clemen e Reilly [6] dividem o processo de educação em etapas de maneira a contemplar com maior abrangência as fases críticas à quantificação do conhecimento. Seguindo estes autores, as principais etapas de um processo de educação são: (i) a avaliação da real necessidade de educar opiniões com a constatação da insuficiência de dados empíricos sobre o problema, uma vez que a educação do conhecimento pode se tornar um processo custoso a depender do assunto que esteja sendo tratado; (ii) a seleção de especialistas sobre o problema de interesse a partir de critérios como provas tangíveis do seu conhecimento [25]; (iii) o treinamento dos especialistas para o processo de educação; (iv) o estudo das variáveis pertinentes ao problema e suas relações e as suas respectivas dependências; (v) a coleta de opiniões dos especialistas sobre o domínio da incerteza; (vi) a síntese, análise e comunicação das opiniões provenientes do processo de educação.

A etapa de treinamento é muito importante para todo o processo pois essa é a etapa onde o especialista começa a se habituar com todas as nuances envolvidas nos processos de educação do conhecimento. Dentro do treinamento existem dois componentes principais para a avaliação e correção das estimativas dos especialistas, são elas a função de calibração e os *scoring rules*. Em um contexto de educação do conheci-

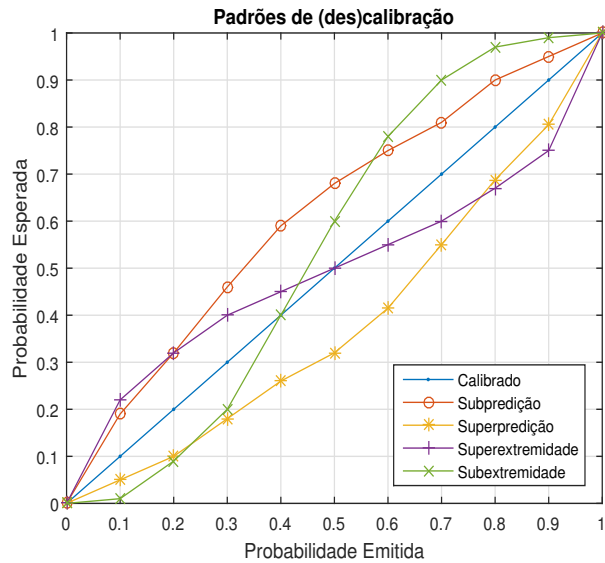


Figura 1: Possíveis padrões de (des)calibração apresentados por especialistas

mento, calibração pode ser entendido como um modelo que ajusta as opiniões dos especialistas em funções dos vieses que podem ser apresentados pelo especialista. Uma abordagem matemática de calibração se baseia na construção de um modelo onde as respostas dos especialistas sejam processadas e então seja possível analisar o padrão de calibração que o especialista esteja utilizando para formular seus julgamentos. Dessa maneira o analista pode intervir no especialista a fim de que os efeitos de possíveis vieses indesejáveis sejam amenizados. Cooke [8] afirma que modelos matemáticos de calibração dão suporte ao analista contra heurísticas inadequadas adotadas pelo especialista. Em termos matemáticos, os modelos de calibração se destinam a encontrar uma relação entre a probabilidade emitida (atribuída pelo especialista) e sua respectiva probabilidade esperada (estimada a partir da frequência de ocorrências) de um determinado evento. Tais probabilidades serão representadas pelas siglas PE e PS , respectivamente. Para qualquer PE no intervalo $[0, 1]$, diz-se que o especialista está calibrado quando $PE=PS$. Isto significa que para um determinado conjunto de eventos para os quais o especialista atribuiu uma probabilidade PE , espera-se que a frequência relativa de ocorrências dos eventos associados (PS) aproxime-se de PE à medida que o número de amostras crescer. A Figura 1 ilustra alguns dos possíveis padrões de (des)calibração do especialistas.

Uma curva de subpredição indica que o especialista está pouco autoconfiante em suas opiniões. Esse padrão indica que as probabilidades atribuídas por ele são na verdade menores do que as da ocorrência do evento analisado, indicando que o especialista necessita ser mais autoconfiante em seus julgamentos. O padrão de superpredição indica que o especialista está muito autoconfiante. As probabilidades atribuídas pelos especialistas são maiores que as reais probabilidades atribuídas aos eventos. Isso indica que o especialista deve considerar melhor seus julgamentos. Subextremidade é o padrão que indica que o especialista apresenta elevada autoconfiança, quando emite probabilidades inferior

res a 40% e baixa autoconfiança, quando infere sobre problemas com probabilidade superior a 40%. Superextremidade ocorre quando o especialista é pouco autoconfiante quando atribui probabilidades inferiores a 50% e muito autoconfiante ao emitir probabilidades superiores a 50%. Todos esses padrões devem ser evitados pois o especialista não está sendo fiel a suas reais crenças. O padrão ideal é quando $PE = PS$ (calibrado) onde o especialista atribui uma probabilidade condizente com suas crenças e ajustada as probabilidades de ocorrência do evento, indicando que ele está calibrado.

Já os *scoring rules* ou regras de pontuação são formalismos matemáticos utilizados para avaliar previsões. A ideia básica geral é atribuir uma nota ao especialista a partir do confronto da distribuição de probabilidades que expressa suas incertezas e a resposta correta da questão [8]. Um dos *scoring rules* mais conhecidos é o *score* quadrático (Q_i) introduzido por Brier [5]. Nessa regra de pontuação o especialista é recompensado por quanto mais próximo conseguir se aproximar do valor real da pergunta, ou seja, quanto mais as opiniões dos especialistas se assimilarem com a frequência relativa de cada evento melhor é a nota atribuída. Por outro lado, se um especialista não tiver tanta certeza sobre uma certa variável e emitir uma grande probabilidade ele será penalizado em sua avaliação. A Equação 1 apresenta a fórmula *score* quadrático onde o vetor $p = (p_1, p_2, \dots, p_n)$ corresponde a valores da variável atribuídos pelo especialista e a resposta correta é representada pelo vetor canônico d , onde $d_i \in d$ assume o valor 1, indicando i como a posição da resposta correta em d , e os demais valores de d , $d_{j \neq i}$, são iguais a zero.

$$Q_i(p) = \sum_{i=1}^n (p_i - d_i)^2 \quad (1)$$

2.2 Modelos de Predição de Séries Temporais

Uma série temporal é qualquer conjunto de observações a respeito de um fenômeno que são organizados em relação a algum índice de ordenamento. A característica que diferencia uma série temporal de uma amostra qualquer de dados é que os dados observados possuem dependência cronológica entre si. Essa característica permite que seja possível elaborar modelos para prever, por exemplo, qual seria o próximo valor assumido por determinada série temporal [9]. Uma característica muito importante de uma série temporal é a estacionariedade. Esse é um atributo elementar no que se diz respeito ao nível de complexidade do processo, influenciando diretamente na modelagem para a previsão da série. Um processo é dito estacionário se a sua média e variância não variam ao longo do tempo. Logo, sendo y_t um ponto da série, μ a média, σ a variância e γ a covariância do processo, este é estacionário se: $E[y_t] = \mu_t = \mu$; média constante para qualquer tempo t , sua variância é constante representado por: $E[(y_t - \mu)^2] = E[(y_{t+s} - \mu)^2] = \sigma_y^2$, sendo s uma constante qualquer, e a sua covariância também é constante descrita por: $E[y_t, y_{t+s}] = [(y_t - \mu)(y_{t+s} - \mu)] = \gamma_{(|s|)}$ [19], a qual depende apenas da diferença de tempo s (uma constante qualquer). Para uma série que apresente tendências ou sazonalidades não há estacionariedade porque essas sazonalidades ou tendências vão afetar o valor da média da série temporal em diferentes períodos [16].

O *autoregressive integrated moving average* (ARIMA) é um modelo estatístico linear para a previsão de séries tem-

porais formalizado por Box e Jenkins [3]. Esse é um dos modelos mais comentados na literatura para previsão de séries temporais. O modelo ARIMA é uma integração de dois modelos, o auto regressivo ($AR(p)$) e o de médias móveis ($MA(q)$). O modelo é denominado ARIMA(p,d,q), onde:

- p é o número de termos auto regressivos;
- d é o número de diferenciações necessárias para se obter a estacionariedade;
- q é o número de termos de médias móveis.

O modelo auto regressivo para variável de interesse é baseado em uma combinação linear dos valores passados da variável, que nesse caso são os pontos da série. Então um modelo auto regressivo de ordem p pode ser definido pela Equação 2:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (2)$$

onde ε_t representa o ruído branco da série, ϕ_p os parâmetros auto regressivos e y_t é o ponto da série temporal no tempo t .

O modelo de médias móveis ao invés de usar os valores passados da série temporal para gerar uma regressão, utiliza os erros da previsão (ou choques aleatórios) para formar essa regressão. O modelo de médias móveis de ordem q pode ser definido pela Equação 3:

$$y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3)$$

onde ε_t representa o ruído branco da série, θ_q parâmetros do modelo de médias móveis e novamente y_t é o ponto da série temporal no tempo t .

Algumas séries não são estacionárias, mas apresentam um comportamento homogêneo na sua não-estacionariedade, por conta disso é necessário haver uma quantidade de diferenciações d para que a série passe a apresentar um comportamento estacionário. Essa diferenciação é representada pela Equação 4.

$$\nabla y_t = y_t - y_{t-1} \quad (4)$$

onde ∇y_t é a derivação do ponto no tempo t da série, em uma primeira aproximação.

Uma função muito importante para estimar os parâmetros dos modelos ARIMA é a Função de Auto Correlação. A Função de Auto Correlação utiliza os coeficientes de auto covariância para gerar uma função que revela a correlação entre quaisquer valores da série [4]. Essa função é definida pela Equação 5.

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \frac{E[(y_t - \mu)(y_{t+\tau} - \mu)]}{E[(y_t - \mu)^2]} \quad (5)$$

Outra função utilizada para estimar os modelos da previsão é a Função de Auto Correlação Parcial. A Função de Auto Correlação Parcial, entre quaisquer dois pontos da série, é a correlação que permanece se o impacto de todas os outros pontos da série fosse eliminado. Sendo ϕ_{kj} o j -ésimo coeficiente em uma função auto regressiva de ordem k é possível escrever [4],

$$\rho_j = \phi_{k1} \rho_{j-1} + \dots + \phi_{k(k-1)} \rho_{j-k+1} + \phi_{kk} \rho_{j-k}; \quad (6)$$

com $j = 1, 2, 3, \dots, k$. A função de Auto-Correlação Parcial é então definida como sendo o último elemento do parâmetro ϕ de ordem k , ou seja, ϕ_{kk} .

3. METODOLOGIA

A metodologia desse trabalho é um estudo comparativo de caráter quantitativo a fim de avaliar a acurácia de modelos cognitivos aplicados na previsão de séries temporais com relação a modelos matemáticos. Nas seguintes subseções serão abordadas as séries utilizadas nesse trabalho, como foram obtidas as estimações para os modelos cognitivos e dos modelos matemáticos.

3.1 Séries Temporais Avaliadas no Trabalho

As séries utilizadas neste trabalho são provenientes da *M3-Competition* [22]. A *M3-Competition* é uma competição de modelos de previsão de séries temporais. Os autores fornecem um conjunto de séries temporais, informando apenas os pontos da série, a ordem cronológica que foram registradas as amostras, uma categoria a qual as séries pertencem e as previsões corretas para a série. Com isso vários modelos são propostos para prever essas séries temporais e depois são analisados estatisticamente pelos autores. Para o trabalho de comparação, foram selecionadas aleatoriamente 8 séries¹ dentre as 3003 disponíveis e essas 8 séries temporais têm suas informações exibidas na Tabela 1.

3.2 Previsões dos modelos Cognitivos

Apesar de ser uma das melhores opções quando dados empíricos são insuficientes, não confiáveis ou inexistentes, educar opiniões de especialistas não é um processo trivial [26]. Transformar opiniões de especialistas contendo vieses através de técnicas que irão converter essas opiniões em dados que serão utilizados na modelagem do problema não é uma tarefa simples. A utilização de *software* visa atenuar problemas relacionados com a educação do conhecimento, como por exemplo: o processo ser ineficiente e de pouca iteratividade, tornando então esse processo muito mais simples e amigável para o especialista além de poder avaliar o comportamento desse especialista durante o processo [20]. A ferramenta *Mesor Elicitation*² é um pacote de *software* que contém algumas das principais etapas para educação do conhecimento de acordo com a divisão das etapas propostas por Clemen e Reilly [6]. Essa ferramenta implementa tanto o treinamento do especialista a respeito do processo de educação, quanto a etapa de extração da informação em si.

De maneira geral o funcionamento da aplicação se dá através de um conjunto de passos que vão da escolha do projeto até a síntese das respostas dos especialistas. Inicialmente deve ser criado um novo projeto e adicionado todas as perguntas associadas a esse novo de projeto, essas serão as variáveis de interesse do problema. Após o cadastro do projeto, deve ser cadastrado o conjunto de especialistas que terão suas opiniões educadas. Esse cadastro informa o projeto que esse especialista está associado, seu número identificador, sua formação e o método de educação e treinamento ao qual esse especialista será submetido. Após o cadastro dos especialistas é iniciada a fase de treinamento com um número específico de questões, apresentadas de uma maneira similar a uma aplicação de um método de educação específico. O especialista terá um *feedback* do seu desempenho através de

¹As séries temporais estão disponíveis no repositório do endereço eletrônico: <https://forecasters.org/resources/time-series-data/m3-competition/>

²O *software Mesor Elicitation* foi desenvolvido pelos pesquisadores Paulo Renato, Ademir Batista e Nielson Santana. Registrado pelo INPI através do número BR512013000.

	Quant. Pontos	Pontos Previstos	Valor Mínimo	Valor Médio	Valor Máximo	Ordem Cronológica	Tipo da Variável
N0236	38	3	1674	3450.6	5303	Anual	Industria
N0290	17	4	3960.5	4321	5259	Anual	Negócios
N1037	44	6	2958.5	4198	5250	Trimestral	Negócios
N1268	45	4	3549.6	4473	5637.8	Trimestral	Negócios
N1269	27	5	5039	5907.2	9439.4	Trimestral	Finanças
N1459	51	8	750	2911.8	6600	Mensal	Mercado
N1708	108	5	680	2012	5060	Mensal	Mercado
N2355	108	2	2053.4	3324.9	4285.6	Mensal	Negócios

Tabela 1: Informações das séries temporais utilizadas no trabalho.

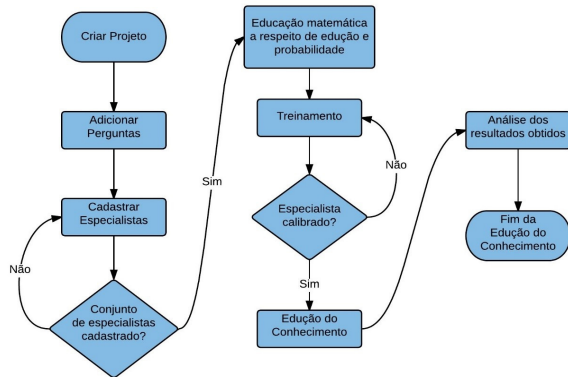


Figura 2: Fluxograma das etapas do software Mesor Elicitation

scoring rules e caso seja necessário, o *software* irá indicar algum padrão de descalibração que o especialista esteja apresentando no decorrer do processo. Depois do treinamento, o especialista irá passar por um conjunto de questões que é a etapa de coleta de opiniões utilizando os métodos de educação do conhecimento. E, por fim, uma síntese das opiniões pode ser gerada pela ferramenta apontando os resultados das opiniões dos especialistas para a variável de interesse. O fluxograma apresentado na Figura 2 demonstra a ordem de execução das funcionalidades do *software*.

Durante a etapa de treinamento o especialista é perguntado sobre diversas questões que, por exemplo, vão de conhecimentos gerais sobre geografia como qual o estado mais populoso dentre “A” e “B”, até perguntas mais específicas sobre o projeto como qual é o valor assumido por determinada série no momento “X”. Essas questões são apresentadas de uma maneira similar ao modo que serão apresentadas na etapa de educação, habituando o especialista aos métodos de educação como ressaltado por O’Hagan [25]. O treinamento se baseia em perguntar ao especialista qual das alternativas é a que ele acredita ser a correta e atribuir um nível de credibilidade para sua resposta. A partir desse nível de credibilidade e da resposta do especialista a *Mesor Elicitation* traça a curva de calibração das respostas do especialista e também começa a desenhar o gráfico que contém as suas pontuações. A Figura 3 demonstra a simulação de um especialista respondendo a pergunta “Qual estado tem maior área?”, tendo como opções de resposta “Rio de Janeiro” e “Minas Gerais”

Na etapa de educação o especialista deve definir o intervalo

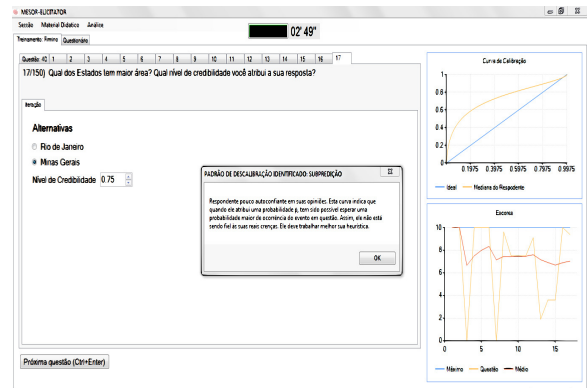


Figura 3: Simulação do treinamento para educação do conhecimento no software Mesor Elicitation

de confiança da variável em questão. Para esse projeto o especialista deve estimar, com maior credibilidade possível, o valor mínimo e máximo que a série pode assumir no instante analisado. Após definido esse intervalo o especialista é convidado a responder duas questões. A primeira o especialista decide entre qual dos intervalos é mais provável que a variável de interesse esteja. A segunda questão é atribuir um nível de credibilidade entre 50% e 100% para sua escolha, tendo em conta que 50% significa uma escolha aleatória e 100% indica máxima confiança na sua resposta. Na primeira iteração o intervalo de confiança emitido pelo especialista é dividido em dois, sendo o primeiro o valor mínimo e a média do intervalo, e o segundo a média do intervalo de confiança e o valor máximo fornecido pelo especialista. Após a primeira interação os intervalos subsequentes são gerados a partir da preferência sobre os intervalos do especialista e do nível de credibilidade atribuído por ele a sua escolha. A cada escolha o especialista é atualizado com algumas variáveis estatísticas de sua resposta como: média, moda, mediana e variância de suas opiniões. A Figura 4 ilustra uma simulação de uma educação para a previsão de um ponto da série temporal N1459.

Continuando a linha de implementação de auxílio gráfico defendida por autores como James *et al* [17], o *Mesor Elicitation* apresenta um gráfico contendo a projeção das previsões do especialista para a série em questão. A cada previsão do especialista o *software* imprime o ponto e promove algumas variáveis como valores: máximo, mínimo, moda e média para que o especialista se sinta mais confortável com suas previsões. Esse gráfico pode ser acessado a qualquer

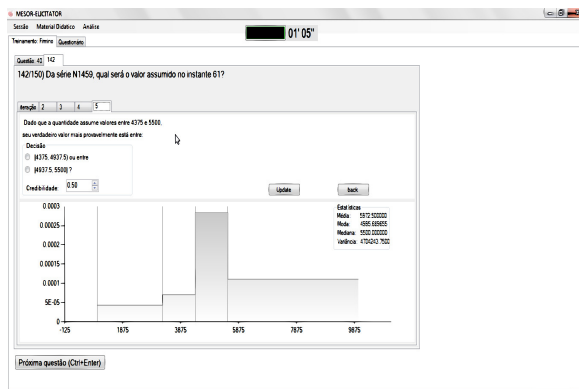


Figura 4: Simulação do treinamento para educação do conhecimento no software *Mesor Elicitation*

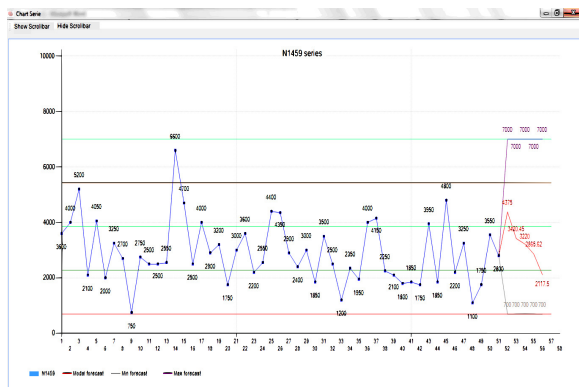


Figura 5: Simulação de gráfico para a previsão de série temporal através do *Mesor Elicitation*

momento pela interface do *software* conforme o especialista sinta necessidade de buscar um recurso gráfico. A Figura 5 ilustra um exemplo de série temporal contendo as previsões do especialista apresentada através de um gráfico.

3.3 Previsões dos Modelos Matemáticos

Para se obter a previsão do modelo matemático ARIMA para as séries temporais primeiramente é necessário obter as funções de auto correlação e auto correlação parcial das séries. A série N1269 será utilizada para ilustrar o procedimento adotado para estimar as funções de correlação e correlação parcial. Todos os procedimentos demonstrados aqui foram realizados no MatLab versão 8.5.0. A Figura 6 mostra a função de auto correlação e auto correlação parcial dessa série.

Através de uma análise gráfica do comportamento de declínio não suave da função de auto correlação é possível observar que a série possui um comportamento estacionário. Após analisar a função, foi estimado um Modelo ARIMA(1,0,0) ou AR(1) por conta do declínio gradual das funções de auto correlação e do queda no lag^3 1 da função de auto correlação parcial, comportamento típico dos modelos AR(1) [4]. Para validar esse modelo, é utilizado um pseudocódigo presente na documentação do MatLab [23] que gera os valores

³Lag é um elemento da janela de observação na série temporal.

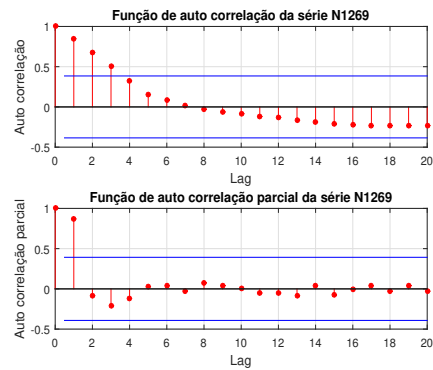


Figura 6: Funções de auto correlação e auto correlação parcial da série N1269

Série	ARIMA	BIC
N0236	(1,1,1)	575.5983
N0290	(0,2,0)	191.6063
N1037	(1,1,1)	569.7846
N1268	(0,1,0)	465.5548
N1269	(1,0,0)	387.0964
N1459	(2,0,2)	868.5259
N1708	(2,1,2)	1683.7
N2355	(2,1,1)	1489.2

Tabela 2: Modelos ARIMA selecionados para as séries e seus respectivos coeficientes de BIC

da métrica BIC (*Bayesian Information Criterion*) para os modelos dessa série. Dessa forma é possível avaliar que a estimativa $p = 1$ e $q = 0$ é a mais adequada para essa série. Com o modelo definido pode-se utilizar esse pseudocódigo para gerar as previsões dos modelos ARIMA para as séries. A Tabela 2 apresenta os modelos estimados por esse procedimento para as séries utilizadas nesse trabalho e seus respectivos coeficientes de BIC.

4. RESULTADOS

O experimento foi baseado em comparar modelos cognitivos obtidos através de educação do conhecimento com modelos matemáticos obtidos através da estimação de modelos ARIMA. Para isso foram entrevistados 33 alunos de diversos programas da Universidade Federal Rural de Pernambuco para que fosse possível obter as estimações dos modelos cognitivos. Como o estudo tem caráter experimental, nenhum dos entrevistados possuía conhecimentos prévios em previsão de séries temporais. Cada um desses entrevistados passou por todo o processo de educação a fim de obter suas estimações para as séries estudadas. Já os modelos matemáticos foram obtidos através do MatLab.

4.1 Métricas de Comparação

Existem algumas métricas de comparação para se avaliar modelos de previsão. Uma das medidas mais utilizadas para comparação da acurácia em previsões de séries é o *Mean Square Error* (MSE). Essa medida é definida por

$$MSE = \frac{1}{N} \sum_{j=1}^N (d_j - p_j)^2 \quad (7)$$

onde N representa o número de observações da série, d_j é o valor real da série no ponto j e p_j é o valor da previsão. Quanto menor for o valor do MSE melhor é a previsão.

Outra medida bastante utilizada é o *Mean Absolute Percentage Error* (MAPE). Essa medida é uma porcentagem do erro da previsão, definida por:

$$MAPE = \frac{100}{N} \sum_{j=1}^N \left| \frac{d_j - p_j}{d_j} \right| \quad (8)$$

Nessa medida N representa o número de observações na série, d_j é o valor real da série no ponto j e p_j é o valor da previsão. Quanto menor for o valor do MAPE melhor é a previsão.

A medida *Prediction of Change in Direction* (POCID) é definida por:

$$POCID = 100 \frac{\sum_{j=1}^N D_j}{N} \quad (9)$$

Onde:

$$D_j = \begin{cases} 1 & (d_j - d_{j-1})(p_j - p_{j-1}) > 0 \\ 0 & \text{caso contrário} \end{cases}$$

Nessa medida N representa o número de observações na série, d_j é o valor real da série no ponto j , p_j é o valor da previsão e D_j é a porcentagem das estimativas corretas sobre a direção da previsão. Essa medida avalia a porcentagem de decisões corretas das estimativas em relação ao valor da série crescer ou decrescer no próximo ponto. Quanto maior for a porcentagem do POCID melhor é a previsão.

Por fim, a medida *Average Relative Variance* (ARV) é representado por essa equação:

$$ARV = \frac{1}{N} \frac{\sum_{j=1}^N (p_j - d_j)^2}{\sum_{j=1}^N (p_j - \bar{d})^2} \quad (10)$$

Nessa medida, N representa o número de observações na série, p_j é o valor da previsão para o ponto j da série, d_j é o valor real da série no ponto j e \bar{d} é a média da série temporal. Se o valor do ARV for igual a 1 a previsão tem acurácia equivalente a calcular a média da série como sendo a previsão para o próximo ponto. Se o valor do ARV for maior que 1 a acurácia da previsão é inferior do que calcular a média da série como sendo a previsão para o próximo ponto. Porém se o valor do ARV for menor do que 1 a previsão é melhor do que apenas extrair a média da série como sendo a previsão para o próximo ponto.

4.2 Comparação entre os Modelos

Para fins de comparação foram realizadas avaliações nas previsões dos modelos estudados e então calculado os resultados de cada uma das métricas apresentadas. Um sumário dos resultados é apresentado na Tabela 3. Das 32 comparações entre as métricas avaliadas, 4 obtiveram resultados

Série		MSE	MAPE	POCID	ARV
N0236	Cognit	16546	33.38%	66%	0.114
	Matem	25813	2.49%	33%	0.273
N0290	Cognit	10147	1.38%	75%	0.174
	Matem	295191	30.09%	75%	0.202
N1037	Cognit	43881	51.4%	16%	0.221
	Matem	102079	5.83%	48%	0.173
N1268	Cognit	89132	4.26%	50%	0.269
	Matem	2572	0.65%	50%	0.314
N1269	Cognit	12870	1.9%	0%	0.39
	Matem	581	0.38%	80%	0.09
N1459	Cognit	1676111	59.76%	24%	0.35
	Matem	981045	47.41%	12%	0.903
N1708	Cognit	1584692	69.15%	20%	2.164
	Matem	2620559	66.70%	20%	0.809
N0236	Cognit	13212	2.24%	50%	0.539
	Matem	132480	8.82%	50%	0.572

Tabela 3: Comparação dos modelos cognitivos (Cognit) e matemáticos (matem) pelas métricas apresentadas.

iguais na métrica POCID. Dentre as que tiveram resultados diferentes, os modelos cognitivos tiveram resultados superiores em 14 das 28 comparações. De modo geral, os resultados dos modelos cognitivos foram promissores para a área de previsões de séries temporais, confirmando a afirmativa de Lawrence [21] onde ele propõe que modelos baseados em opiniões de especialistas são ao menos tão bons quanto modelos cognitivos. Contudo, como pode ser observado nas métricas das séries N1269 e N1708, os modelos cognitivos não obtiveram um bom resultado. Essas séries apresentam um comportamento decrescente e a série N1708 apresenta também muita oscilação. O que indica que modelos cognitivos podem possuir um viés de se resguardarem quando a avaliação da série é decrescente e o fator oscilação influencia negativamente na previsão. A pouca quantidade de pontos de algumas das séries contribuiu no desempenho do modelo cognitivo e inviabilizou a utilização de outros modelos matemáticos, como por exemplo, redes neurais artificiais que necessitam uma quantidade considerável de dados de entrada para gerar a previsão. O resultado mostra ainda o potencial de aplicações de apoio a educação, como o *Mesor Elicitation*, que podem deixar as opiniões de especialistas, que não tenham conhecimentos específicos na área de previsão de séries temporais, com acurácia em um patamar considerável em comparação aos modelos ARIMA.

5. CONCLUSÃO

É inegável a importância da utilização de opiniões de especialistas em problemas onde dados empíricos sejam escassos, inexistentes ou até mesmo irrelevantes. A mente humana é um processador complexo e com a ajuda necessária na parte matemática e computacional o que se pode obter através dela são resultados surpreendentes. Contudo, é importante salientar que existem diversos fatores que contribuem para que as informações provenientes dos modelos cognitivos apresentem ruídos. Esses fatores podem ter várias origens como vieses dos especialistas ou por vezes pode estar relacionado ao seu próprio conhecimento. Porém, é evidente que opiniões de especialistas são recursos viáveis para modelagem de vários problemas incluindo os que envolvem

a previsão de séries temporais.

Foram utilizados nesse trabalho séries temporais de eventos reais de diversas áreas. Com isso as tendências e sazonalidades dessas séries são mais complexas de se encontrar dificultando suas previsões para ambos modelos. Além disso, o tamanho das séries não favoreceu o uso de outros modelos matemáticos de maior complexidade. Contudo, a comparação entre os modelos foi realizada de maneira condizente e os resultados obtidos demonstraram que os modelos cognitivos são equivalentes aos matemáticos, onde para os experimentos realizados o “especialista” humano não necessariamente detinha um bom conhecimento a respeito dos fenômenos representados pelas séries temporais avaliadas. Esses resultados demonstraram que as opiniões dos especialistas (ou apenas humanos empenhados a resolverem um certo problema) têm a capacidade de serem utilizadas para prever pontos em séries temporais e também abrem novas perspectivas como a análise das previsões de modelos matemáticos através de educação do conhecimento e do uso de opiniões em outros problemas.

6. AGRADECIMENTOS

Os autores gostariam de agradecer a UFRPE (Universidade Federal Rural de Pernambuco) por fornecerem a estrutura física necessária para a realização desse trabalho. Os autores também gostariam de agradecer a FACEPE (Fundação de Amparo a Ciência no Estado de Pernambuco) pelo suporte financeiro fornecido para a realização desse trabalho.

7. REFERÊNCIAS

- [1] B. M. Ayyub. *Elicitation of expert opinions for uncertainty and risks*. CRC press, 2001.
- [2] W. Barbosa dos Santos, A. L. Drogue, et al. Análise probabilística de riscos via redes bayesianas: Uma aplicação na construção de poços multilaterais. 2005.
- [3] G. Box. P. and Jenkins, G. M. *Time series analysis: forecasting and control*, 1970.
- [4] G. Box, G. M. Jenkins, and G. Reinsel. *Time series analysis: Forecasting & control*. 1994.
- [5] G. W. Brier. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3, 1950.
- [6] R. Clemen and T. Reilly. *Making hard decisions with decisiontools*. 2001.
- [7] N. J. Cooke. Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, 41(6):801–849, 1994.
- [8] R. M. Cooke. *Experts in uncertainty: opinion and subjective probability in science*. 1991.
- [9] R. S. Ehlers. Análise de séries temporais. *Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná*, 2007.
- [10] R. J. Ferreira, P. E. L. d. Oliveira, P. R. A. Firmino, and E. A. L. Drogue. Mapa estratégico para programas de recursos humanos: avaliando o desempenho de redes bayesianas. *Gestão & Produção*, 17(1), 2010.
- [11] P. R. A. Firmino and E. L. Drogue. An expert opinion elicitation method based on binary search and bayesian intervals. *International Journal of Risk Assessment and Management*, 18(3-4):336–362, 2015.
- [12] K. M. Ford and J. R. Adams-Webber. Knowledge acquisition and constructivist epistemology. In *The psychology of expertise*, pages 121–136. Springer, 1992.
- [13] P. H. Garthwaite, J. B. Kadane, and A. O’Hagan. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470):680–701, 2005.
- [14] J. Han. A goms-based granular computing model for human-computer interaction design. In *Service Operations, Logistics, and Informatics (SOLI), 2011 IEEE International Conference on*, pages 243–248. IEEE, 2011.
- [15] R. Hodge, M. Evans, J. Marshall, J. Quigley, and L. Walls. Eliciting engineering knowledge about reliability during design-lessons learnt from implementation. *Quality and Reliability Engineering International*, 17(3):169–179, 2001.
- [16] R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2014.
- [17] A. James, S. L. Choy, and K. Mengersen. Elicitor: an expert elicitation tool for regression in ecology. *Environmental Modelling & Software*, 25(1):129–145, 2010.
- [18] G. Keren. Calibration and probability judgements: Conceptual and methodological issues. *Acta Psychologica*, 77(3):217–273, 1991.
- [19] G. Kirchgässner, J. Wolters, and U. Hassler. *Introduction to modern time series analysis*. Springer Science & Business Media, 2012.
- [20] M. Kynn. Designing elicitor: Software to graphically elicit expert priors for logistic regression models in ecology. Available from www.winbugs-development.org.uk, 2006.
- [21] M. Lawrence, P. Goodwin, M. O’Connor, and D. Önkal. Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, 22(3):493–518, 2006.
- [22] S. Makridakis and M. Hibon. The m3-competition: results, conclusions and implications. *International journal of forecasting*, 16(4):451–476, 2000.
- [23] M. MatLab. The language of technical computing. *The MathWorks, Inc.* <http://www.mathworks.com>, 2012.
- [24] K. P. McCormack, M. Bronzo, and M. P. V. Oliveira. Uma abordagem probabilística para a avaliação de riscos em cadeias de suprimento. *Revista Produção Online*, 10(3):577–598, 2010.
- [25] A. O’Hagan, C. E. Buck, A. Daneshkhan, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley, and T. Rakow. *Uncertain judgements: eliciting experts’ probabilities*. John Wiley & Sons, 2006.
- [26] A. O’Hagan. Probabilistic uncertainty specification: Overview, elaboration techniques and their application to a mechanistic model of carbon flux. *Environmental Modelling & Software*, 36:35–48, 2012.
- [27] N. Shadbolt and P. R. Smart. *Knowledge elicitation: Methods, tools and techniques*. 2015.