



Universidade Federal Rural de Pernambuco
Departamento de Estatística e Informática



Análise do Desempenho de Modelos Cognitivos na Previsão de Séries Temporais

Ademir Batista dos Santos Neto

Recife

Dezembro de 2015

Ademir Batista dos Santos Neto

Análise do Desempenho de Modelos Cognitivos na Previsão de Séries Temporais

Orientador: Tiago Alessandro Espínola Ferreira

Coorientador: Paulo Renato Alves Firmino

Monografia apresentada ao Curso Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Recife

Dezembro de 2015

A
Deus,
Aos meus pais,
A minha família,
Aos meus orientadores,
Aos meus professores,
Aos meus amigos,
A minha namorada.

Agradecimentos

Agradeço primeiramente a Deus por ter me concedido todas as condições necessárias para que eu pudesse concluir esse trabalho e me iluminar em todo o caminho que trilhei durante a minha graduação.

Agradeço muito aos meu pais. A minha mãe Elisângela que sempre me apoiou e incentivou em todas as etapas da minha vida, me dando muito amor e transmitindo a mim suas experiências. Ao meu pai Ademir Junior que constantemente me mostrou o caminho certo a se percorrer e que a persistência leva a perfeição.

Agradeço a toda minha família pelo amor e carinho incondicionais. Em especial a meus irmãos Eduarda, Thaís, Hugo e Marcella. Agradeço muito ao meu padrasto Mauricio que sempre me guiou no caminho dos estudos. Agradeço também aos meus avós que sempre me deram muito carinho e afeto.

Também agradeço especialmente ao professor Paulo Renato, que sem ele a realização desse trabalho não seria possível. O professor Paulo sempre me apoiou e incentivou em toda minha vida acadêmica. Também agradeço muito ao professor Tiago que teve a paciência e dedicação necessárias para me orientar durante esse trabalho.

Agradeço aos meus professores que sempre estiveram transmitindo-me conhecimentos para que eu pudesse me tornar uma pessoa melhor.

Agradeço aos meus amigos, tanto os daqui quanto aos do intercâmbio no Canadá, por serem essas pessoas tão especiais em minha vida.

Agradeço também a minha namorada Vanessa, essa pessoa maravilhosa que eu tenho o privilégio de conviver e que tem me ensinado muito sobre tudo na minha vida.

Resumo

Modelos cognitivos são usados como principais fontes para a estimação das informações no sentido de arquitetarem estudos quando 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. A extração ou educação do conhecimento desses especialistas é uma área de estudo bastante específica. Existem algumas técnicas que buscam eduzir o conhecimento dos especialistas. A técnica que será abordada nesse trabalho é 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. Para isso são utilizadas algumas séries presentes em um trabalho que contém uma competição de modelos de previsão de séries temporais. Com isso, ao fim do trabalho é apresentada uma comparação entre a acurácia da previsão dos modelos cognitivos e matemáticos a partir de algumas métricas descritas no trabalho.

Palavras-chave: modelos cognitivos, modelos matemáticos, educação do conhecimento, especialistas, séries temporais, previsão, ARIMA e acurácia.

Abstract

Cognitive models are used as the main source for estimation of information in order to plan studies when empirical data is unavailable, scarce or does not have an considerable relevance. The cognitive models are the expert's opinions in relationship to determined variable based on his knowledgem and experiences. The elicitation of the knowledge from those experts is a specific study field. There are many techniques which aim to elicitate knowledge from experts. The technique which will be used in this dissertation is the elicitation process. The elicitation process is a ensemble of steps that seeks to get the opinon of the expert in a way to improve the accuracy of the information and reduce some possible biases that the experts may show.

This dissertation aim to compare the performance of cognitive models obtained through elicitation of knowledge and mathematical models in relationship to their accuracy in time séries forecast. For this are used some séries from a work which contain a time séries forecast competiton. At the end of the dissertation is showed a comparison between the accuracy of the forecast from the cognitive and mathematical models from some metrics described in this dissertation.

Keywords: cognitive models, mathematical models, elicitation of knowledge, experts, time séries, forecast, ARIMA and accuracy.

Sumário

1	Introdução	1
1.1	Justificativa	3
1.2	Objetivos	4
1.3	Organização do trabalho	5
2	Processos de educação do conhecimento	6
2.1	Treinamento	8
2.1.1	Heurísticas de comportamento e vieses	10
2.1.2	Calibração	12
2.1.3	Scoring rules	14
2.2	Métodos de educação do conhecimento	15
3	Software para coleta das opiniões dos modelos cognitivos	18
3.1	Projeto	20
3.2	Treinamento	22
3.2.1	Treinamento do método de Firmino	22
3.2.2	Treinamento do método Direto	23
3.2.3	Treinamento do método da Bisseção	24

3.3	Edução	25
3.3.1	Edução pelo método de Firmino	27
3.3.2	Edução pelo método Direto	28
3.3.3	Edução pelo método da Bissecção	29
3.4	Análise do Resultado	29
4	Séries temporais	33
4.1	Séries avaliadas no trabalho	35
4.2	Métricas de comparação	35
5	Modelos ARIMA	38
5.1	Estimação dos modelos ARIMA para as séries desse trabalho	42
6	Comparação entre os modelos	46
7	Conclusão	56
A	Coleção das séries utilizadas no trabalho	66

Lista de Tabelas

3.1	Exemplo de opinião do especialista a respeito de determinada variável de interesse	30
4.1	Informações das séries temporais utilizadas no trabalho.	36
5.1	Caracterização das estimações de parâmetros dos modelos ARIMA pela análise das funções de auto correlação e auto correlação parcial	41
5.2	Modelos ARIMA selecionados para as séries e seus respectivos coeficientes de BIC	45
6.1	Análise comparativa da série N0236	48
6.2	Análise comparativa da série N0290	49
6.3	Análise comparativa da série N1037	50
6.4	Análise comparativa da série N1268	51
6.5	Análise comparativa da série N1269	52
6.6	Análise comparativa da série N1459	53
6.7	Análise comparativa da série N1708	54
6.8	Análise comparativa da série N2355	55
A.1	Observações da série N0236	66
A.2	Observações da série N0290	66

A.3	Observações da série N1037	66
A.4	Observações da série N1268	67
A.5	Observações da série N1269	67
A.6	Observações da série N1459	67
A.7	Observações da série N1708	68
A.8	Observações da série N2355	68
A.9	Previsões dos modelos cognitivos e matemáticos para as séries	69

Lista de Figuras

2.1	Possíveis padrões de (des)calibração apresentados por especialistas	13
3.1	Fluxograma da educação utilizando <i>Mesor Elicitation</i>	21
3.2	Simulação do treinamento do método de Firmino	23
3.3	Simulação do treinamento do método direto	24
3.4	Simulação do treinamento do método da Bisseção	25
3.5	Simulação de gráfico para a predição de série temporal através do <i>Mesor Elicitation</i>	26
3.6	Simulação da aplicação do método de educação de Firmino	28
3.7	Simulação da aplicação do método de educação Direto	29
3.8	Simulação da aplicação do método de educação da Bisseção	30
3.9	Diagrama entidade relacionamento do banco de dados da aplicação	31
4.1	Manchas solares anuais de 1700 a 1986	34
5.1	Função de auto correlação da série das manchas solares apresentada na Figura 4.1	40
5.2	Função de auto correlação parcial da série das manchas solares apresentada na Figura 4.1	41
5.3	Funções de auto correlação e auto correlação parcial da série N1269	42

5.4 Representação gráfica da previsão do modelo ARIMA(1,0,0) para a série N1269 44

Capítulo 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 esses dois modelos nem sempre são facilmente distinguíveis [1]. 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.

A escassez de dados tem se apresentado como um dos principais desafios a modelagem estatística de modelos baseados em dados empíricos [2]. Essas modelagens frequentistas necessitam muitas vezes de um grande número de dados para descreverem os problemas. Situações como a análise 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 [3]. A busca por suporte a 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 [4]. Nessa mesma linha de acordo com MacCormack *et al* [5], 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. 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 [6]. 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 estimativa da quantidade ou qualidade de uma determinada variável de interesse [7]. 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 eduzir (quantificar) o conhecimento de determinado especialista a respeito de algum problema específico. Algumas características como habituar o especialista ao processo, medir os seus níveis de incerteza e analisar os dados obtidos são alguns elementos básicos que o processo de educação do conhecimento deve apresentar [8]. 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 [9], [10] e [11]. Por tanto, pode-se afirmar que eduzir opiniões significa interagir com o especialista de uma forma que ele possa transmitir fielmente suas incertezas a respeito de certa variável de interesse [12]. A utilização de opiniões de especialistas obtidas através de processos de educação do conhecimento se estende por diversas áreas. Dentre os trabalhos que citam a utilização de opiniões de especialistas, Kappel *et al* [13] utiliza opiniões de especialistas para avaliar a relação entre a ação humana e o desequilíbrio nos ambientes marinhos costeiros. Knol *et al* [14] comentam sobre a importância e necessidade de se recorrer à quantificação de opiniões para a avaliação dos efeitos a saúde da exposição de alguns fatores ambientais, como partículas de ar poluído e alimentos geneticamente modificados por exemplo. Choy *et al* [15] destacam o uso de opiniões na modelagem estatística para a ecologia. Jonhson *et al* [16] estudam as propriedades de mensuração de métodos de educação a partir de casos de estudo da área médica. Firmino *et al* [17] recorrem a opiniões para quantificar um modelo probabilístico direcionado à manutenção de linhas de transmissão de energia. Droguett *et al* [18] modelam quantitativamente as relações entre eventos críticos para a análise probabilística de risco de tecnologias de preenchimento inteligente de poços de petróleo a partir de métodos de educação do conhecimento. Firmino *et al* [19] utilizam educação do conhecimento em especialistas para modelar quantitativamente

eventos críticos ao refino de óleo de uma planta química. Na oncologia, Van der Gaag *et al* [20] recorrem a opiniões para quantificar um modelo estatístico que mapeia as causalidades existentes entre as variáveis associadas ao câncer de esôfago. Na experimentação clínica, Oremus *et al* [21] estudam a probabilidade de sucesso de dois tratamentos contra incontinência urinária feminina. Jorgensen [22] *apud* Jorgensen [23] mostra que 84% das estimações dos projetos de software conduzidos em uma grande empresa de telecomunicações foram baseados em opiniões de especialistas.

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 uma variável discreta [24]. A medida de tempo pode variar de acordo com a série que se esteja trabalhando, a escala pode estar em meses, dias, horas etc [25]. Uma das características mais importantes desse tipo de dados, que possibilita sua modelagem e previsão, é a dependência entre as observações vizinhas [26]. Os dados que compõe as séries temporais podem vir de várias áreas como economia (valores das ações na bolsa, produção na indústria, taxa de crescimento econômico), medicina (eletrocardiograma, eletroencefalograma), epidemiologia (número de novos casos de determinada doença), meteorologia (precipitação pluviométrica, temperatura diária, velocidade do vento) e outras mais [26]. 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 [27] onde os autores descrevem os modelos da família ARIMA. Já no campo da inteligência artificial outros modelos são propostos para a previsão de séries temporais como redes neurais artificiais e sistemas baseados em regras Fuzzy [28], por exemplo.

1.1 Justificativa

Vários autores como Cooke [7] classificam a educação do conhecimento como um instrumento indispensável para a obtenção de opiniões de especialistas utilizadas na modelagem de determinados problemas. Como muitos desses problemas não possuem dados empíricos associados se torna inevitável o uso de modelos cognitivos. Autores como Lawrence *et al* [29] e [30] defendem que modelos baseados em opiniões de especialistas são ao menos tão acurados quanto aqueles baseados em dados empíricos. Rowe e Wright [31] consideram que as opiniões de

especialistas são muitas vezes necessárias em problemas de previsão por conta da falta de propriedade ocasionada por procedimentos estatísticos. Embora muitos autores tenham se pronunciado a favor da utilização de opiniões de especialistas como uma fonte confiável de estimação de informações, a comunidade acadêmica por motivos como descrença, falta de poder computacional ou até mesmo desinteresse deixaram os procedimentos de coleta e registro de opiniões estacionados em relação ao seu desenvolvimento. Kynn [32] assume que é indispensável o uso de ferramentas computacionais para o procedimento aceitável de uma educação do conhecimento. Apesar dos inquestionáveis benefícios do uso de softwares para a educação do conhecimento e da maior popularização do uso de opiniões, pouco se têm estudado em relação ao desempenho dos processos de educação do conhecimento nas comparações com outros modelos matemáticos. A área de previsão de séries temporais é um campo onde atualmente há muita discussão e sempre surgem novas propostas de modelos que tentam prever de uma maneira mais acuradas os pontos futuros da série, prova disso são as competições de modelos de previsão como por exemplo a *M3-Competition* [33]. Com isso o presente trabalho faz uso de processos de educação de conhecimento utilizando ferramentas computacionais para eduzir opiniões de especialistas na temática da predição de séries temporais, com o intuito de tentar mostrar que opiniões de especialistas podem ser tão bem acuradas quanto os modelos matemáticos.

1.2 Objetivos

O principal objetivo desse trabalho é analisar qualitativamente o desempenho dos modelos cognitivos obtidos através de processos de educação do conhecimento na temática da previsão de séries temporais. Também como objetivos específicos desse trabalho destacam-se:

- Avaliar os resultados dos métodos de educação do conhecimento aplicados com a utilização de um software específico;
- Comprovar ou refutar que os métodos de educação do conhecimento podem ser utilizados como importantes ferramentas para a modelagem estatística de problemas de previsão.

1.3 Organização do trabalho

Este trabalho está organizado em 7 capítulos dos quais:

- O Capítulo 1 contém a introdução desse trabalho;
- Capítulo 2 apresenta um apanhado geral sobre educação do conhecimento e suas nuances;
- No Capítulo 3 a ferramenta que foi utilizada para a obtenção das previsões dos modelos cognitivos é apresentada com maiores detalhes em relação a sua infraestrutura e o seu funcionamento;
- O Capítulo 4 apresenta uma abordagem geral sobre séries temporais e as séries que foram utilizadas nesse trabalho;
- Os modelos da família ARIMA e as previsões obtidas a partir desses modelos são apresentados no capítulo 5;
- Capítulo 6 é a síntese sobre as comparações entre os modelos utilizando algumas métricas como *scoring rules* e MSE (*mean square error*);
- E por fim, o capítulo 7 apresenta a conclusão do trabalho.

Capítulo 2

Processos 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 eduzir conhecimento do domínio do especialista, tipicamente através de alguma iteração direta com esse especialista [12]. 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. O processo de educação do conhecimento 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 minerar informações de um conjunto de dados [34]. 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 eduzidos melhor será o modelo resultante da educação [35]. Eduzir 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 [36].

Processos de educação do conhecimento tiveram seu início formal no começo da década de 80, muito embora técnicas para educação do conhecimento já tivessem sido desenvolvidas anteriormente, como por exemplo um dos pioneiros entre os métodos de educação o método Delphi [37] que foi desenvolvido pela *RAND corporation* inicialmente para fins militares. Esse desenvolvimento em educação do conhecimento foi atribuído pelo grande interesse da comunidade científica por sistemas especialistas [35]. Esses sistemas tinham como objetivo tentar simular a performance de especialistas para tomada de decisões. No início dos estudos sobre inteligência artificial muitos dos esforços eram baseados em descobrir os princípios gerais do comportamento da inteligência humana. O trabalho de Newell e Simon [38], com uma estratégia geral de resolução de problemas, exemplifica as abordagens dos trabalhos daquela época. Contudo devido a capacidade tecnológica, desinteresse a respeito das abordagens de resolução geral de problemas e com o advento da máxima “conhecimento é poder” a educação do conhecimento passou a ganhar força no cenário das pesquisas no meio acadêmico [35]. Esses fatores também foram suportados por trabalhos como o de Chi *et al* [39] que afirmam que o desempenho das opiniões dos especialistas está atribuído ao fato de que esses especialistas possuem o domínio de fatos e regras específicas das variáveis em questão. O ato de representar e transmitir conhecimento pode ser considerado como uma atividade humana elementar, que foi perfilado ao longo de todo desenvolvimento cognitivo e social da humanidade [40]. Tratando-se apenas do conhecimento, dados provenientes de: livros, documentos, manuais e casos de estudo podem ser considerados como fontes de informação. Porém a perícia em determinado assunto deriva sobre práticas e experiências a cerca desses tópicos específicos e não da simples leitura de documentos [12]. O grande diferencial de opiniões de especialistas para os dados é a sua análise detalhada dos fatos. Devido aos seus conhecimentos e estratégias, as análises provenientes dos especialistas são mais completas do que as informações que estão contidas nos livros. De ideias como essa surgiram as primeiras tentativas formalizar a educação do conhecimento com processos direto de extração de informação [41] que logo demonstraram a necessidade de mais trabalho nessas abordagens por conta de erros e vieses que são comumente apresentados por especialistas quando entrevistados de maneira direta. Então ficou evidenciado a necessidade de processos que tentam-se obter informações de especialistas de um modo que eles pudessem expressar suas crenças de maneira mais fiel possível e amenizassem seus vieses e heurísticas de comportamento. Nesse sentido algumas novas abordagens surgiram como, por exemplo, o método proposto por Garthwaite e Al-Awadhi [42], que sugere a modelagem das opiniões dos especialistas como parâmetros de um modelo linear generalizado. Em algumas abordagens atuais os processos de educação

de conhecimentos estão sendo alinhados a atividades de gerenciamento do conhecimento corporativo. Nesse sentido existem muitas suposições, como por exemplo, que o conhecimento é uma valiosa habilidade que deve ser obtida da pessoa correta, no momento certo, de maneira ótima em um momento oportuno [43] *apud* [12]. Contudo mesmo com a utilização de educação do conhecimento em diversas áreas e o avanço dos métodos de educação muitas das etapas fundamentais para o processo de quantificação do conhecimento são negligenciadas ou realizadas de maneira precária, prejudicando o resultado final da educação.

Como mencionado anteriormente processos de educação são compostos por diversas etapas. De maneira especial, a forma com que o trabalho de Clemen e Reilly [9] divide tais etapas parece 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 eduzir 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 [44]; (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 sob o domínio da incerteza; (vi) a síntese, análise e comunicação das opiniões provenientes do processo de educação. Nas subseções a seguir, serão discutidas em detalhes as etapas de treinamento e coleta de opiniões de processo de educação para prover um melhor suporte ao entendimento da aplicação dos modelos cognitivos através da ferramenta para a educação do conhecimento.

2.1 Treinamento

A etapa de treinamento tem um papel muito importante para todo o processo de educação. Essa é a etapa onde o especialista começa a se habituar com todas as nuances envolvidas nos processos de educação do conhecimento. Nesta etapa o especialista é instruído a respeito de conceitos matemáticos de incerteza, probabilidade e também sobre estatística. Esse especialista também é conscientizado a respeito da presença de heurísticas intrínseca ao seu comportamento e depois introduzido aos métodos que de fato vão coletar suas opiniões a respeito das variáveis de interesse do problema. Cooke [35] argumenta que independente da

abordagem utilizada, o treinamento é uma etapa valiosa na aquisição do conhecimento de especialistas. A etapa de treinamento merece ser destacada dentro do processo de educação do conhecimento pois permite que o analista (indivíduo que conduz a educação) possa inferir sobre o comportamento e também sobre o atual estado de conhecimento do especialista quando convidado a expressar suas incertezas. Essa etapa também é útil para que o especialista possa aprimorar suas habilidades acerca de decidir sob níveis de incerteza. Todas essas análises são possíveis graças ao analista possuir, durante o treinamento, as respostas corretas dos questionamentos possibilitando assim o diagnóstico e reconhecimento de padrões das heurísticas adotadas pelo especialista. Vários autores como O'Hagan *et al* [44], Kynn [45], Kadane *et al* [46], Lichtenstein e Fischhoff [47] e Van Lenthe [48] fortemente recomendam uma fase de treinamento antes de serem coletadas as opiniões dos especialistas. Inclusive Van Lenthe [48] destaca a implementação da etapa de treinamento em eduções realizadas através de um software como uma maneira de elevar a acurácia das informações obtidas.

De modo geral, a etapa de treinamento é composta por questionários que introduzem o especialista a cada um dos temas relacionados a educação do conhecimento. É esperado que após as intervenções realizadas no treinamento o especialista passe a sofrer menos influências a sua crença, tais como heurísticas de comportamento inadequadas demonstradas por Tversky and Kahneman [49]. Destaca-se assim que o objetivo principal da etapa de treinamento é preparar o especialista para expressar fielmente seus níveis de incerteza. O treinamento deve possibilitar que os especialistas sejam acompanhados e avaliados diante de determinados estímulos, podendo revelar algumas características intrínsecas ao seu modo de pensar como por exemplo elevada autoconfiança ou falsa sensação de poder de controle sobre a ocorrência de eventos relacionados a sua atuação. Esses desvios são denominados heurísticas de comportamento. Heurísticas são estratégias adotadas para simplificar o problema que se tem em mãos [50]. Essas estratégias são marcas indelévels do processo cognitivo humano diante de muitas situações envolvendo incerteza. Por exemplo, mostra-se inevitável ao ser humano recorrer a heurísticas quando convidado a opinar sobre o resultado de uma operação matemática que envolva muitos cálculos, como por exemplo, estimar o valor da expressão 5 elevado a 10^a potência em um curto período de tempo. A fase de treinamento pode ser dividida em três etapas: (i) educação a matemática e estatística a respeito de como expressar incertezas em termos de probabilidade; (ii) orientação a respeito de heurísticas de comportamento e os efeitos que essas podem agregar as respostas dos especialistas; (iii) adequação ao método de educação que será utilizado no processo. Modelos de calibração são

uma maneira eficaz de validar as etapas do treinamento. Elas podem ser utilizadas para diagnosticar matematicamente vieses ou alguma deficiência que o especialista esteja apresentando durante o treinamento partindo da suposição que o analista conhece os valores reais das variáveis que estão sendo questionadas ao especialista [7]. A ideia é ajustar (calibrar) as opiniões dos especialistas a partir de medidas de erro observadas durante o treinamento. Outro fator que auxilia o especialista a entender seu desempenho é a atribuição de uma pontuação para cada resposta do especialista, através de formalismos matemáticos que utilizam a resposta do especialista e a resposta correta do experimento [51]. Dessa maneira o especialista passa a ter uma maior consciência sobre as suas respostas e sobre o seu desempenho durante a educação.

Com isso, pode se deduzir que as contribuições de software se mostram indispensáveis para o treinamento. Essa afirmativa é ressaltada por autores como Kynn [32] onde é proposto que o auxílio de ferramentas computacionais como, por exemplo, interfaces gráficas como *feedback* para o especialista é indispensável a fase de treinamento. Em uma análise das soluções para educação do conhecimento presentes na literatura, mostra-se inquestionável a contribuição que uma avaliação realizada por software pode proporcionar a diagnósticos de heurísticas e vieses presentes nas opiniões de especialistas. Modelos de intervenção em especialistas contra heurísticas inadequadas a partir de resultados de calibração obtidos através de software são expostos no trabalho de Firmino e Santana [52]. Outro fator importante é o uso de recursos de software para a educação matemática e orientação aos processos de educação. Utilizando-se esses recursos a didática de ensino se torna muito mais homogênea e a interatividade promovida por softwares contribui com o desenrolar do procedimento. Sob este escopo, mostra-se inviável a boa execução da etapa de treinamento sem o devido suporte tecnológico provido por softwares. Nas seguintes subseções serão melhor detalhados aspectos relacionados a heurísticas de comportamento, calibração de opiniões e regras de pontuação para o acompanhamento do desempenho do especialista.

2.1.1 Heurísticas de comportamento e vieses

Como mencionado anteriormente, o julgamento humano é composto por vieses e heurísticas de comportamento que formam as crenças e opiniões. As heurísticas de comportamento estão indissociavelmente ligadas a formação de opinião dos indivíduos [50]. Elas são uma simplificação do processo cognitivo diante de situações de incerteza. Nesse contexto, heurísticas

podem ser consideradas regras, princípios organizadores ou recursos simplificadores para formatar e emitir opiniões. Tais heurísticas e seus reflexos têm sido investigados por pesquisadores que transitam entre as áreas da estatística e a psicologia [49]. Algumas das heurísticas destacadas pelos autores são: disponibilidade, ancoragem e controle:

A heurística da disponibilidade está associada com a facilidade com que alguns fatos são recuperados na mente humana. Esse tipo de associação pode ocasionar uma estimativa errônea a respeito de alguma variável. Por exemplo, se alguma pessoa tiver sido recentemente traumatizada com um acidente de trânsito e dentro do processo de educação for solicitado uma estimativa sobre a possibilidade de ocorrência de acidentes de trânsito essa pessoa pode superestimar a probabilidade desse evento acontecer mesmo que sua real crença não esteja coerente com o que foi registrado;

A ancoragem é uma heurística que conduz o respondente a se vincular excessivamente a dados iniciais do problema, que muitas vezes são introduzidos pelas características do processo. Essa heurística conduz o especialista a não conseguir expressar suas reais crenças devido a recorrência demasiada de um valor que lhe foi apresentado anteriormente. Tversky e Kahneman [49] demonstram através de um experimento acerca de inferir sobre porcentagem de uma determinada variável, que os valores iniciais fornecidos aos respondentes têm efeito direto nos resultados finais dessa inferência;

O controle corresponde a heurística que faz com que o especialista tenha a impressão que possui algum controle sobre a variável em questão, quando em fato na maioria das vezes essas variáveis não estão sob o seu controle. Essa heurística está associada com a elevada autoconfiança do especialista a um ponto que seus julgamentos passam a não conter a aleatoriedade associada de alguns eventos. Um exemplo disso pode ser observado em taxa de falhas em equipamentos, que embora a ocorrência esteja dentro de um determinado intervalo de tempo, muitos fatores associados conduzem a uma incerteza mais elevada independente da experiência do especialista.

Outros autores, como Jorgensen [23], argumentam que algumas possíveis medidas podem reduzir os vieses em processos de educação do conhecimento. Dentre elas pode-se citar: (i) evitar grandes motivações em especialistas por respostas mais acuradas, grandes expectativas sobre os resultados ou associação monetária conduzem a estimativas menos acuradas [53]; (ii) perguntar aos especialistas variáveis para justificar ou criticar as estimativas a serem feitas, a ideia é que os especialistas estejam cientes de que suas opiniões possam ser avaliadas

por critérios relacionados a seu próprio domínio; (iii) evitar informações irrelevantes ou não confiáveis, é importante para a acurácia do processo desconsiderar qualquer informação que não agrega valor as opiniões. Para ilustrar a importância de utilizar apenas informações relevantes autores como Whitecotton [54] afirmam que opiniões de especialistas tem desempenho tão bom quanto modelos matemáticos em previsões financeiras apenas quando informações irrelevantes não forem incluídas ao processo; (iv) usar dados empíricos (quando estiverem disponíveis) para auxiliar o especialista a realizar uma avaliação mais analítica e conseqüentemente reduzir seus vieses; (v) escolher especialistas que tenham boa experiência e que saibam expressar bem suas opiniões em situações de incerteza.

2.1.2 Calibração

Opiniões de especialistas podem, apesar de serem recursos valiosos, apresentar vieses e heurísticas de comportamento. Um julgamento que não esteja coerente com as reais crenças dos especialistas pode ser entendido como descalibrado. Por exemplo, um especialista que tem uma autoconfiança elevada pode representar suas incertezas com margens de erro menores do que os ideais para suas estimativas [55]. Em um contexto de educação do conhecimento, 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. Abordagens de intervenção são comentados em [52]. Quando um especialista é bem calibrado a acurácia das suas opiniões é aprimorada. A calibração das opiniões é necessária uma vez que a base para o julgamento humano sobre incerteza está ligada a heurísticas simplificadoras de problemas complexos que muitas vezes conduzem a vieses [56].

Cooke [7], 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 2.1 ilustra alguns dos possíveis padrões de (des)calibração do especialistas.

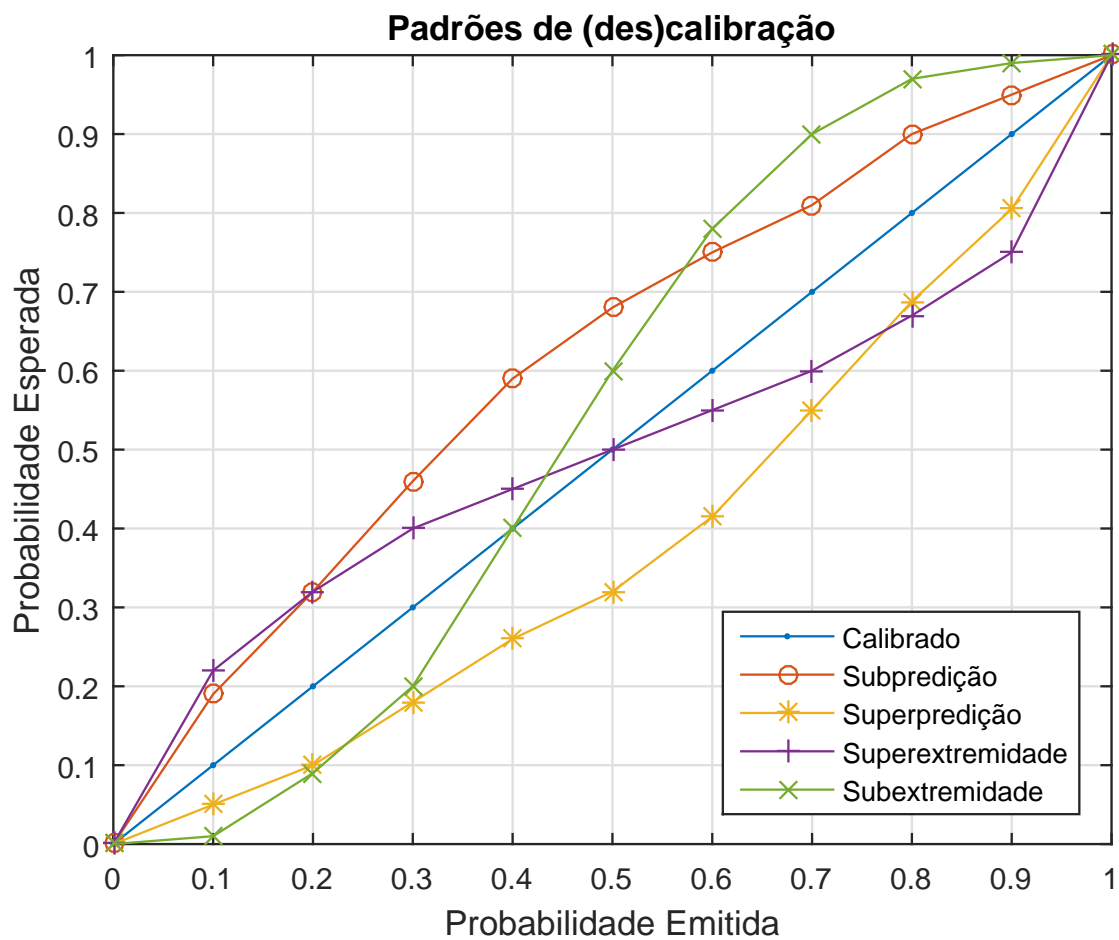


Figura 2.1: Possíveis padrões de (des)calibração apresentados por especialistas

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 inferiores 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.

2.1.3 Scoring rules

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 [7]. Um dos *scoring rules* mais conhecidos é o *score* quadrático (Q_i) introduzido por Brier [51]. 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. Outras regras de pontuação também são disponibilizadas na literatura como logarítmico (L_i) [57], esférico (S_i) [58], chi-quadrado (C_i) [59] e kullback (K_i) [6]. Autores como Bickel [60] e Nau [61] se propõe a comparar as propriedades de alguns desses *scores*. A seguir serão apresentados os formalismos que representam os *scores* citados. Para isso entende-se que 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.

$$Q_i(p) = \sum_{i=1}^n (p_i - d_i)^2 \quad (2.1)$$

$$L_i(p) = \ln \left(\sum_{i=1}^n p_i d_i \right) \quad (2.2)$$

$$S_i(p) = \frac{\sum_{i=1}^n p_i d_i}{\sqrt{\sum_{i=1}^n p_i^2}} \quad (2.3)$$

$$C_i(p) = - \sum_{i=1}^n \frac{(p_i - d_i)^2}{d_i} \quad (2.4)$$

$$K_i(p) = - \sum_{i=1}^n d_i \cdot \ln \left(\frac{d_i}{p_i} \right) \quad (2.5)$$

Alguns autores como O’Hagan [44] destacam que ao depender da natureza do problema, os *scoring rules* não conseguem diferenciar adequadamente o desempenho de certas distribuições, especialmente se o vetor d não for um vetor canônico. Trabalhos como [62] propõem que o vetor canônico d seja substituído pela real distribuição de probabilidades da variável relacionada a questão exposta quando isso for possível. Assim essa proposta apresenta uma maior capacidade de diferenciar opiniões em determinadas situações.

2.2 Métodos de educação do conhecimento

Existem uma variedade de técnicas para se extrair informações de especialistas. As mais usuais são aquelas que por meio de interação direta tentam eduzir o conhecimento do especialista de alguma maneira, como por exemplo entrevistando o especialista [12]. As técnicas estruturadas de iteração com os especialistas são chamadas de métodos de educação do conhecimento. Os métodos de educação tem como objetivo medir matematicamente as crenças do especialista a respeito de determinada variável de interesse. Nessa fase, o especialista é convidado a expor suas opiniões sobre as variáveis de interesse do problema. Existe uma grande variedade de métodos de educação na literatura como exposto em: [37], [42], [45], [48],

[63], [64], [65] e [66]. Um dos pioneiros dentre os métodos de educação do conhecimento foi o método Delphi [37]. Esse método inicialmente desenvolvido para fins militares, tornou-se muito conhecido e utilizado em diversos trabalhos relacionados a coleta de informações de especialistas. Gordon e Helmer [67] utilizaram o método Delphi em um estudo de previsões de longo termo na área de inovações tecnológicas. Outro método bastante conhecido é o método da bisseção desenvolvido por Raiffa [63], esse método foi muito utilizado em diversas áreas como por exemplo na previsão climática [68]. Os métodos de educação podem ser divididos em três classes: diretos, indiretos e paramétricos:

Métodos Diretos convidam o especialista a responder diretamente sobre suas incertezas acerca das variáveis de interesse. Nessa abordagem o especialista dá sua opinião sobre cada intervalo da sua distribuição relativa a variável de interesse. De modo a facilitar a compreensão das perguntas, o analista geralmente utiliza gráficos de barras indicando onde se localizam as probabilidades atribuídas pelo especialista ajudando-o a compreender melhor a modelagem de suas opiniões. Johnson *et al* [16] utilizam essa classe de métodos para estudar incerteza de especialistas quanto a probabilidade de pacientes com esclerodermia e hipertensão sobreviverem após três anos sendo tratados com e sem anticoagulantes. Nessa classe de métodos, ao fim da educação, se obtém uma distribuição de probabilidades com a opinião do especialista a respeito da variável de interesse do problema. Em relação ao seu desempenho, esses métodos apresentam a melhor eficiência dentre os métodos de educação, isso quer dizer que a educação utilizando essa classe de métodos consome menos tempo. Contudo a acurácia das opiniões não é tão elevada, especialmente se o especialista não estiver habituado a conceitos de probabilidade [6]. O método Delphi é um exemplo de método de educação direto. Nesse método, a partir de um problema geral, o especialista é convidado a expor suas opiniões a respeito de determinada variável a partir de um conjunto de perguntas que estão ligadas entre si e estão direcionadas ao problema geral;

Métodos indiretos são uma classe de métodos onde as incertezas dos especialistas a respeito da variável de interesse do problema não são medidas diretamente. Nessa classe de métodos o especialista é induzido a refletir sobre os valores da distribuição de sua incerteza. Dessa forma o especialista tem uma maior reflexão ao determinar em que área da distribuição a variável de interesse se encontra. Os métodos indiretos possuem a vantagem de proporcionar ao especialista uma abordagem mais elaborada ao invés de simplesmente pergunta-lhe qual é o valor de determinada variável, proporcionando a esse especialista uma maior maturidade quanto suas crenças e incertezas. Dessa maneira as opiniões dos especialistas tendem

a ser mais bem acuradas para essa classe de métodos. Contudo sua eficiência é prejudicada uma vez que o especialista necessita inferir mais sobre uma mesma variável do que em métodos diretos. O método da bisseção de Raiffa [63] é um dos mais conhecidos dentre os métodos indiretos. Nesse método tendo um intervalo $[a, b]$ o especialista é induzido a estimar um valor md cuja a probabilidade da resposta da questão estar no intervalo $[a, md]$ seja igual a dela estar em $[md, b]$, formando assim dois intervalos de indiferença quanto a preferência do especialista. De fato, a obtenção da região de indiferença é uma das características mais marcantes dessa classe de métodos. Uma das áreas onde esse tipo de abordagem vem sendo bastante utilizada é a teoria das decisões [65];

Métodos paramétricos são aqueles que os especialistas são eduzidos a fim de se encontrar parâmetros relacionados matematicamente as incertezas do especialista sobre a variável de interesse. Nessa classe de métodos as incertezas dos especialistas são modeladas de acordo com uma determinada distribuição de probabilidades. Algumas das distribuições usadas são Normal, Lognormal e Beta. Uma das deficiências dessa classe de métodos é que eles requerem que além do especialista possuir razoável conhecimento a respeito de probabilidades, ele também deve ter domínio a respeito de distribuições de probabilidades. O método paramétrico de Garthwaite e Al-Awadhi [42] utiliza as opiniões dos especialistas para gerar um modelo linear generalizado que represente a distribuição da variável de interesse.

Jorgensen [23], afirma que independentemente da escolha do método, a sua simplicidade deve ser um dos fatores mais importantes a serem considerados. Uma vez que quando o especialista não compreende o método, os resultados obtidos não serão acurados. Os métodos de educação apesar de serem boas alternativas para a educação de especialistas apresentam algumas dificuldades. Dentre elas o que mais é ressaltado é a relação entre eficiência e acurácia dos métodos. Usualmente métodos de educação independentemente da classe que estejam associados são processos que necessitam de muito tempo para que o especialista possa expressar suas incertezas. Desse cenário surgem as abordagens de educação através de software, que tentam simplificar a compreensão dos especialistas e ao mesmo tempo propõem uma maior eficiência ao processo. Autores como Kynn [32] afirmam que a utilização de softwares pode amenizar muitas das dificuldades encontradas na educação do conhecimento.

Capítulo 3

Software para coleta das opiniões dos modelos cognitivos

Apesar de ser uma das melhores opções quando dados empíricos são insuficientes, não confiáveis ou inexistentes, eduzir opiniões de especialistas não é um processo trivial [69]. Transformar opiniões de especialistas contendo vieses e heurísticas através de técnicas que irão transformar essas opiniões em dados que serão utilizados na modelagem do problema não é uma tarefa simples. Esse tipo de modelagem exige muito do analista que intermedeia a educação do conhecimento e principalmente do especialista que tem que traduzir seu conhecimento através de probabilidades, exigindo conceitos que muitas vezes vão além do seu domínio. A utilização de softwares visa atenuar esses problemas relacionados com a educação do conhecimento tornando o processo muito mais simples e amigável para o especialista além de poder avaliar o comportamento desse especialista durante o processo [32].

Além disso, softwares também auxiliam calibração e treinamento do especialista que é um fator extremamente importante para a educação pois reduz vieses e heurísticas apresentados pelo especialista [69]. Com isso alguns softwares relacionados a educação do conhecimento surgiram. Deville e Knowl [70] argumentam que existem softwares em pelo menos 5 linhas distintas na educação do conhecimento, esses softwares são voltados especificamente para: educação do conhecimento, colaboração entre especialistas e formação de consenso, agregação de resultados, seleção de especialistas e caracterização de incertezas. Mais especificamente os softwares voltados para a obtenção de opiniões de especialistas têm sido desenvolvidos ao longo de muitos anos de pesquisa, alguns exemplos desses software são apresentados em:

[11], [32], [48], [66], [71] e [72]. A partir das ideias de eduzir especialistas de maneira acurada e eficiente com todos os benefícios que o uso do software traz ao processo de educação surgiu a proposta de desenvolvimento do *Mesor Elicitation*¹.

A ferramenta *Mesor Elicitation* 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 [9]. Essa é uma aplicação *stand alone* desenvolvida na plataforma .NET com a interface gráfica e integração dos módulos desenvolvidos em C#, os módulos de calibração e dos métodos de educação foram desenvolvidos em C++ e a persistência é feita utilizando MySQL. Essa ferramenta implementa tanto o treinamento do especialista a respeito dos métodos de educação e amenização de vieses como a educação do conhecimento através de três métodos distintos. Nessa ferramenta também é possível saber o desempenho em termos de acurácia e tempo consumido (eficiência) de cada um dos métodos implementados, dando margem para uma avaliação qualitativa desses métodos de educação.

Os métodos de educação implementados no *Mesor Elicitation* são o método indireto de Firmino [2], método direto e o método da bisseção [63]. As abordagens para todos os três métodos serão explicadas posteriormente nas seguintes subseções. Os softwares também possuem suporte para a educação matemática e sobre probabilidades através de materiais didáticos que vão auxiliar o analista a explicar os conceitos básicos necessários para que o especialista possa ter suas opiniões eduzidas por um processo de educação.

O software também provê suporte matemático para identificação de vieses através da calibração de suas opiniões no treinamento. O grande diferencial dessa ferramenta para as demais voltadas para educação do conhecimento é a preocupação com o treinamento do especialista. A ideia é reduzir os vieses, levando as opiniões dos especialistas a serem mais bem acuradas, melhorando a eficiência do processo e também a simplicidade com que o especialista compreende o processo. Autores como Kynn [73] alertam que muitos poucos trabalhos se propõem a considerar os vieses e heurísticas intrínsecos a opinião de especialistas. O'Hagan [69] afirma que é de grande importância uma etapa de treinamento para que o especialista possa reduzir seus vieses e também se habituar aos processos de educação do conhecimento.

Outro fator que é muito esquecido pelos estudos é a eficiência do processo, poucos autores

¹O software *Mesor Elicitation* foi desenvolvido pelos pesquisadores Paulo Renato, Ademir Batista e Nielson Santana. Registrado pelo INPI através do número BR512013000.

como Spaccasassi e Deleris [72] consideram que a presença do especialista não é muitas vezes um recurso disponível então é interessante aproveitar o melhor possível o tempo com ele, evitando fadiga e desinteresse. Outro fator é a simplicidade pois se os especialistas não compreenderem bem os processos eles não terão possibilidade de expressar adequadamente seu conhecimento sobre as variáveis de interesse.

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 eduzidas, esse cadastro informa o projeto que esse especialista está associado, seu número identificador, sua formação e o método de educação e treinamento que esse especialista será submetido. Após o cadastro dos especialistas é iniciado 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 *scoring rules* e caso seja necessário o software irá indicar algum padrão de descalibração que o especialista esteja apresentando para que o analista possa intervir. 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 3.1 demonstra a ordem de execução das funcionalidades do software.

Nas seguintes subseções serão apresentados a ferramenta para a utilização de modelos cognitivos na predição de séries temporais assim como o funcionamento dos métodos de educação e treinamento.

3.1 Projeto

Inicialmente é cadastrado o projeto sobre o qual serão associadas as perguntas que serão utilizadas para eduzir as variáveis de interesse do problema. O projeto cadastrado foi o de comparação de métodos de educação do conhecimento através da predição de séries temporais. Nesse projeto os especialistas são convidados a prever os valores futuros de algumas séries

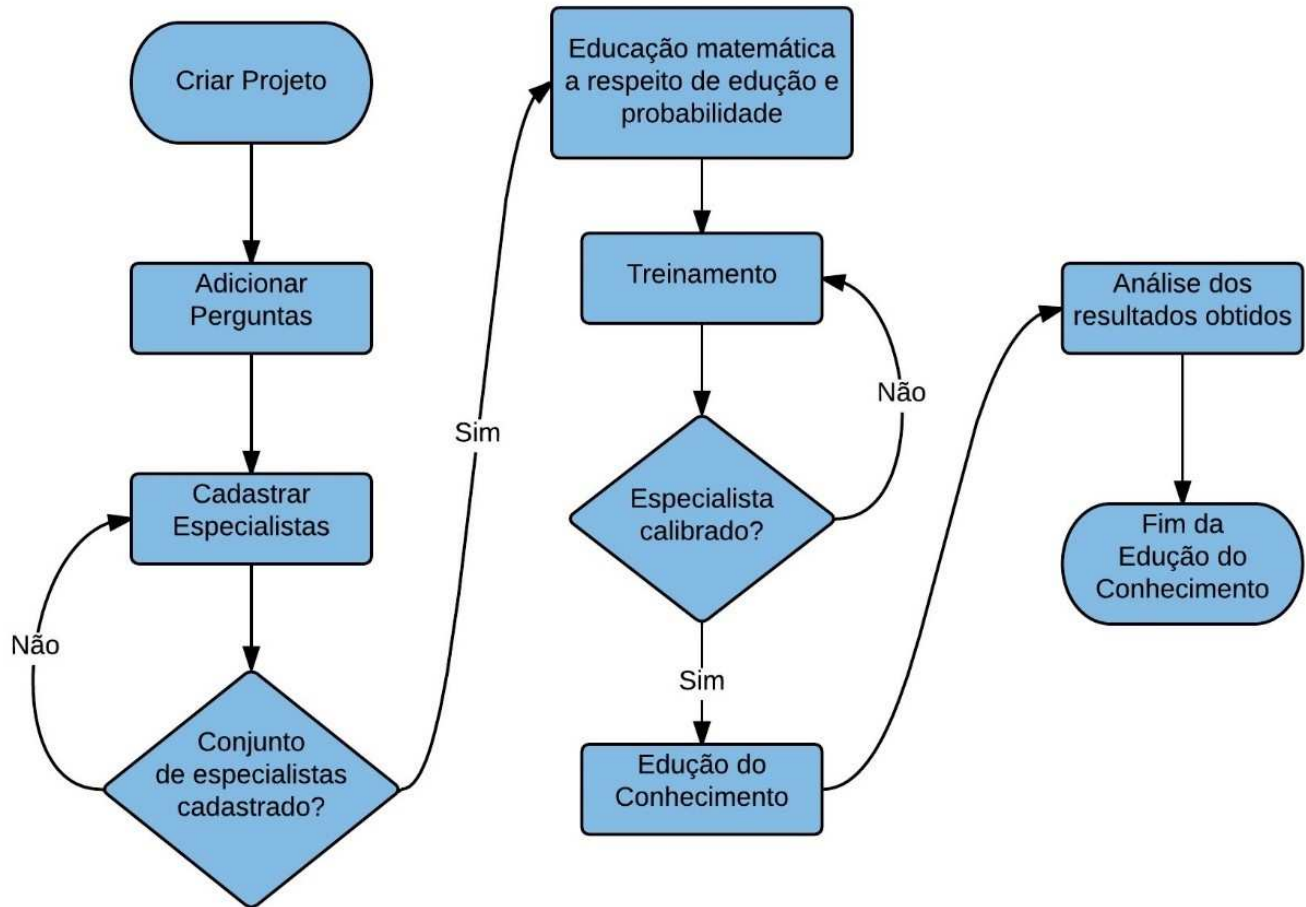


Figura 3.1: Fluxograma da educação utilizando *Mesor Elicitation*

temporais presentes na *M3-Competition* [33] através de algum método de educação. Foram selecionadas 25 perguntas para o treinamento do especialista inspiradas nas reflexões presentes no trabalho de Tversky e Kahneman [49] para detecção de vieses e heurísticas de comportamento inadequados usualmente apresentados pelo especialista. Dentre essas questões estão presentes perguntas que envolvem: álgebra, geografia, probabilidade e previsão de séries temporais. Após o treinamento o especialista passa a responder questões específicas sobre previsão de séries temporais. Ao todo foram entrevistados 33 participantes na sua maioria alunos de diferentes programas da Universidade Federal Rural de Pernambuco. Após o especialista ser cadastrado, selecionado o método de treinamento e o método de educação o especialista será educado sobre probabilidade e incerteza através de um material didático. Ao fim dessas etapas o especialista é conduzido para a etapa de treinamento.

3.2 Treinamento

Durante a etapa de treinamento o especialista é perguntado sobre diversas questões que 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 [44]. No treinamento, o software possui as respostas corretas das questões aplicadas, dessa maneira é possível intervir nos especialistas contra padrões de descalibração e heurísticas de comportamento indesejadas. Durante o treinamento a educação pode ser interrompida a qualquer momento caso o software detecte algum padrão de descalibração presente nas repostas do especialista, dando oportunidade ao analista intervir e tentar corrigir esse padrão apresentado pelo especialista. Além disso, durante todo o treinamento do especialista é exibido um gráfico com seu desempenho avaliado em uma nota dentro do intervalo $[0,10]$ gerada por uma normalização do *score* quadrático, dado na Equação (3.1), onde o vetor p representa a distribuição de probabilidades fornecido pelo especialista e d o vetor canônico que representa a resposta correta da pergunta onde $d_i \in d$ assume o valor 1, na posição da resposta correta no vetor d .

$$Q_i(p) = 5 \cdot \left[2 - \sum_{i=1}^n (p_i - d_i)^2 \right] \quad (3.1)$$

As seguintes subseções demonstram como são aplicados o treinamento para cada método de educação implementado.

3.2.1 Treinamento do método de Firmino

O treinamento do método de Firmino [2] 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 o *Mesor Elicitation* traça a curva de calibração das respostas do especialista e também começa a desenhar o gráfico que contem as suas pontuações. A Figura 3.2 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” através do treinamento do método de Firmino.

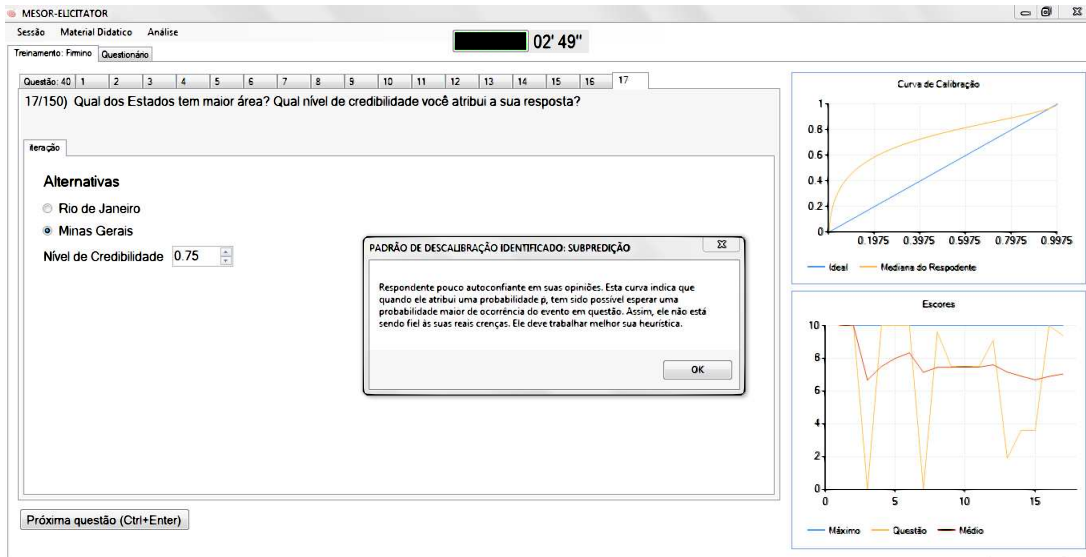


Figura 3.2: Simulação do treinamento do método de Firmino

Nessa simulação um padrão de descalibração foi detectado a partir das respostas do especialista. Nesse caso a linha amarela no gráfico superior direito da Figura 3.2 indica que o especialista está sendo pouco autoconfiante com suas respostas, atribuindo uma baixa credibilidade a suas opiniões e acertando o valor correto da variável em mais questões do que proporcionalmente ele acredita com base no nível de credibilidade que ele atribui a suas respostas. Esse padrão deve ser evitado porque o especialista deve atribuir uma credibilidade condizente com suas crenças para não comprometer o resultado da educação. No gráfico inferior direito da Figura 3.2 é apresentado o gráfico dos *scoring rules*, a linha amarela representa a *scores* pontual de cada questão respondida e a linha vermelha é a média dos seus *scores* até determinada questão.

3.2.2 Treinamento do método Direto

No treinamento do método direto o especialista é provido de uma pergunta e três alternativas. Essa proposta visa habituar o especialista com a aplicação da educação utilizando o método direto. Dessa forma o especialista deve atribuir um nível de credibilidade para cada uma das alternativas. Caso ele considere ter total confiança na resposta correta da questão ele pode atribuir a credibilidade igual a 1 para essa alternativa e como esse conjunto de alternativas

é uma distribuição de probabilidades a credibilidade para as demais alternativas seria de 0. A Figura 3.3 ilustra uma simulação do treinamento do método direto para a questão sobre qual dos estados é o mais populoso dentre as alternativas “Goiás”, “Tocantins” e “Bahia”.

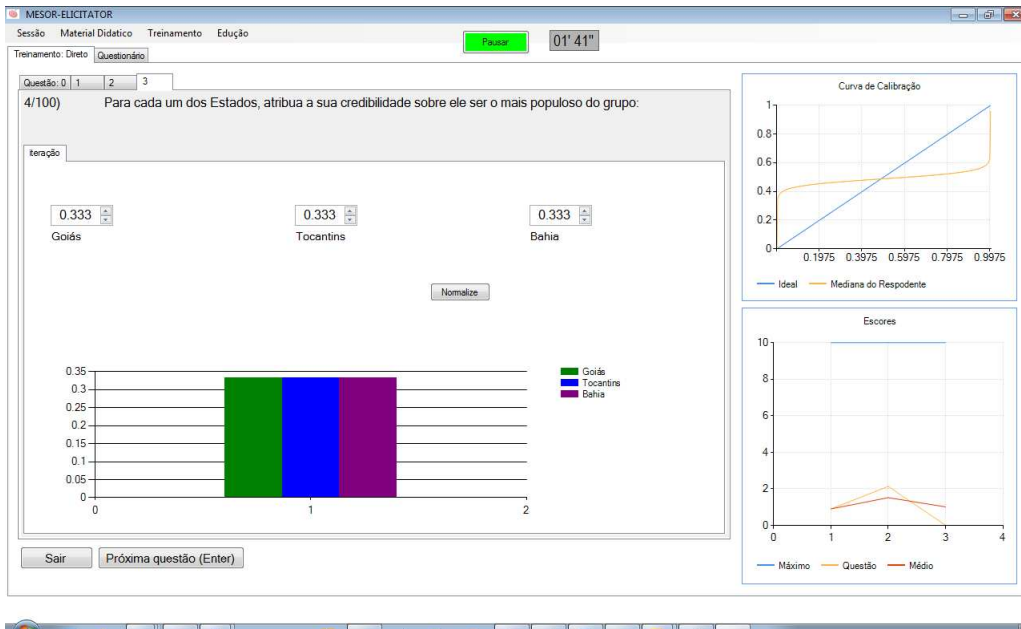


Figura 3.3: Simulação do treinamento do método direto

Os gráficos da curva de calibração e o dos *scores* apresentados na Figura 3.3 seguem os mesmos padrões explicados para o treinamento do método de Firmino na Seção 3.2.1. Contudo nessa simulação o especialista apresentou um padrão de descalibração de superextremidade apresentando baixa autoconfiança quando atribui credibilidades menores que 50% e muito autoconfiante quando atribui credibilidades maiores que 50%. No treinamento do método direto o especialista é auxiliado por um gráfico de barras que representa a distribuição de probabilidades para a variável em questão. Com isso o especialista tem um auxílio gráfico para suas crenças e também fica habituado a ver suas crenças representadas graficamente da mesma forma que será feito quando suas opiniões forem eduzidas através do método direto.

3.2.3 Treinamento do método da Bissecção

No método da bissecção é perguntado qual é o intervalo de credibilidade que o especialista acredita que a variável de interesse pode se encontrar. Em outras palavras, o método da bissecção solicita um valor máximo e mínimo que a variável pode assumir onde o especialista

afirma isso com a maior credibilidade possível. Após esse intervalo de confiança ser estabelecido, assim como ocorre na proposta de Raiffa [63], o especialista é induzido a apontar suas preferências entre dois intervalos: o valor mínimo do intervalo de confiança e a média desse intervalo ou a média do intervalo de confiança e o valor máximo do intervalo. Quando o especialista escolhe um dos intervalos ele automaticamente cria um novo intervalo o processo é repetido até que o especialista não tenha mais preferência por nenhum dos intervalos. A Figura 3.4 ilustra uma simulação do treinamento do método da bisseção para a variável sobre a média de altura dos homens brasileiros.

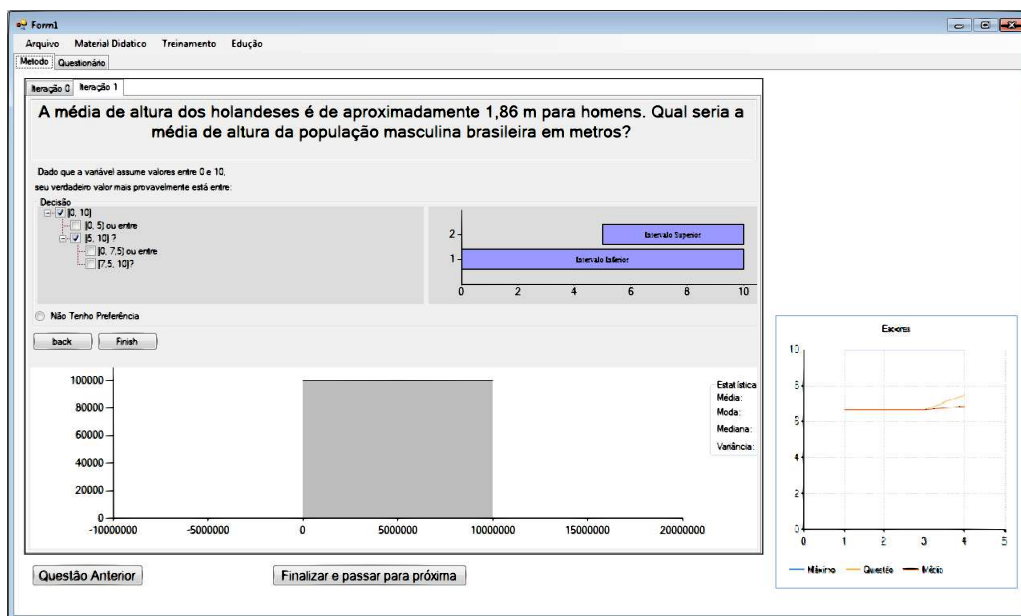


Figura 3.4: Simulação do treinamento do método da Bisseção

Pelo exemplo é possível observar que a cada vez que o especialista escolhe um intervalo de sua preferência o espaço entre o valor máximo e mínimo diminui, até o ponto em que o especialista não tem mais preferência entre nenhum dos intervalos. O gráfico dos *scores* segue as mesmas definições apresentadas na Seção 3.2.1, porém nessa simulação não foi exibido o gráfico da curva de calibração uma vez que essa curva só é apresentada ao respondente em caso dele apresentar algum padrão de descalibração.

3.3 Educação

Todas as etapas do processo de educação são muito relevantes. De fato, autores como Scholten [66] afirmam que é muito importante para qualquer processo de educação do conhecimento

os métodos implementados exibem a série da variável em questão como é demonstrado na Figura 3.5, mas cada método implementa uma maneira distinta de perguntar qual será o próximo valor da série nesse instante. O *Mesor Elicitation*, como ocorre no treinamento, implementa três métodos de educação, são eles: Firmino, Direto e Bisseção. Nas subseções a seguir será detalhado como são obtidas as opiniões dos especialistas para cada um dos métodos implementados.

3.3.1 Educação pelo método de Firmino

O método indireto de educação de Firmino foi apresentado por Firmino em 2008 [19]. Primeiramente o especialista deve definir o intervalo 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 Figura 3.6 ilustra uma simulação de uma educação utilizando o método de Firmino.

A medida que o especialista vai escolhendo os intervalos de sua preferência o software vai modelando sua distribuição de probabilidade e exibindo em um gráfico de barras. Uma opção interessante é que caso o especialista não goste da forma distribuição que suas opiniões tomaram ele pode voltar para pensar melhor entre o intervalo escolhido anteriormente e sobre o nível de credibilidade atribuído a sua resposta. A medida que o especialista vai escolhendo entre os intervalos esses vão se estreitando até o ponto em que o especialista não tenha mais preferência para opinar. Nesse ponto, o especialista forneceu sua opinião para a variável de interesse em questão.

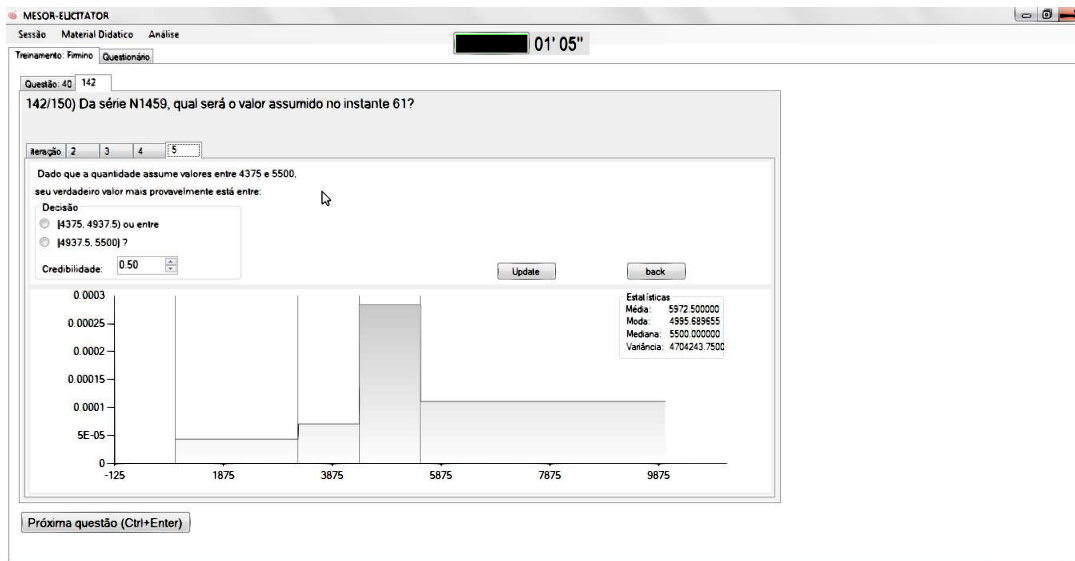


Figura 3.6: Simulação da aplicação do método de educação de Firmino

3.3.2 Educação pelo método Direto

O método direto de educação é um método onde o especialista é convidado a inferir diretamente sobre a probabilidade da variável em questão estar em determinado intervalo dentro de uma distribuição. Esse método preza pela eficiência uma vez que o especialista apenas tem que julgar qual é a probabilidade da variável em questão pertencer a determinado intervalo. Nesse método, assim como no método de Firmino, o especialista é convidado a estabelecer um intervalo de confiança onde o valor mínimo representa o menor valor possível que a série pode assumir no instante analisado e o valor máximo é o maior valor que a variável pode assumir nesse instante de acordo com as crenças do especialista. Além desses valores uma outra variável é solicitada ao especialista, o número de partições que ele deseja dividir o intervalo de confiança. Após isso o especialista vai atribuir uma credibilidade para cada partição do intervalo respeitando as leis da probabilidade onde todas as credibilidades somadas devem resultar em 1. A Figura 3.7 ilustra uma simulação da educação pelo método direto.

Para cada atualização dos níveis de credibilidade do especialista o gráfico com a distribuição de sua opinião é alterado. Quando o especialista estiver satisfeito com sua distribuição essa vai ser sua opinião a respeito da variável de interesse.

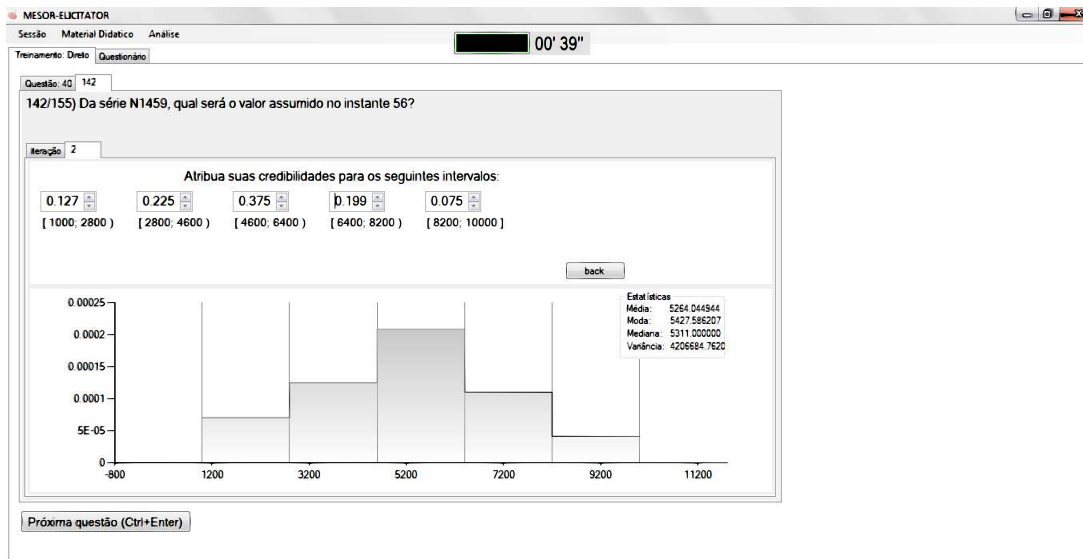


Figura 3.7: Simulação da aplicação do método de educação Direto

3.3.3 Educação pelo método da Bissecção

No método da bissecção, assim como proposto por Raiffa [63], o especialista baseia sua opinião inferindo em que intervalo a variável de interesse mais provavelmente está. A abordagem da educação pelo método da bissecção é muito semelhante ao que é proposto no seu treinamento. Primeiramente, a exemplo dos outros dois métodos, o especialista estabelece um intervalo de confiança com os valores máximo e mínimo que a série pode assumir no ponto analisado. Depois desse intervalo estabelecido, o especialista é convidado a escolher um intervalo entre o valor mínimo e a média do intervalo de confiança ou a média do intervalo de confiança e o valor máximo. Após isso, o intervalo escolhido pelo especialista toma o lugar do intervalo de confiança e o processo é repetido até que o especialista não tenha mais preferência entre os intervalos. A Figura 3.8 ilustra uma simulação da aplicação do método da bissecção.

3.4 Análise do Resultado

As opiniões dos especialistas sobre as variáveis de interesse são armazenadas no banco de dados do *Mesor Elicitation*. Após isso o analista possui a opinião do especialista sobre a variável de interesse, algumas métricas avaliadas durante a entrevista e também alguns *scores* dos especialistas. Nesse projeto todas as perguntas envolvidas possuem as respostas corretas para que seja possível a comparação entre os modelos cognitivos e matemáticos

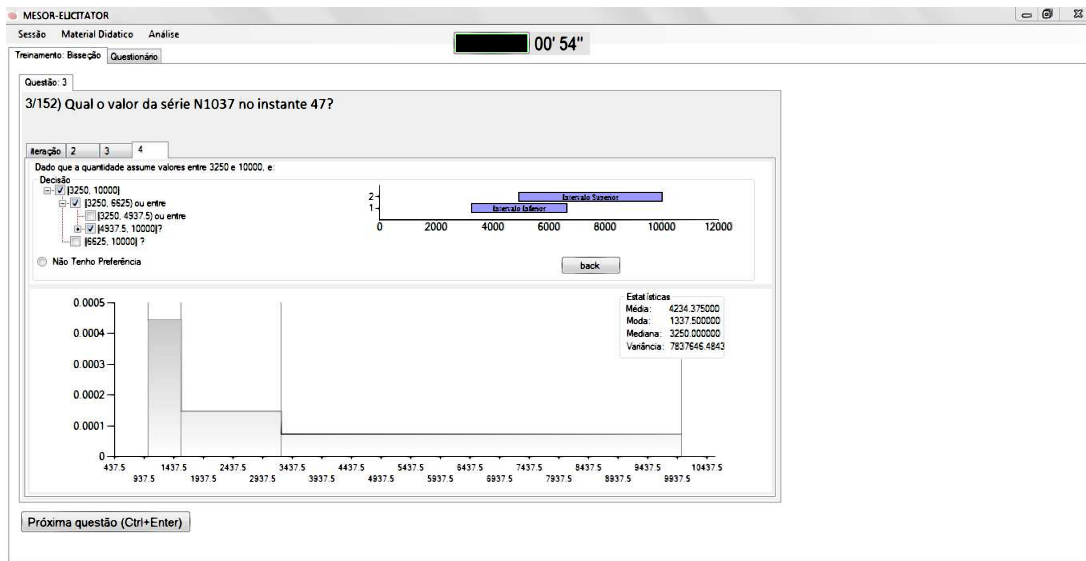


Figura 3.8: Simulação da aplicação do método de educação da Bissecção

de predição. Sendo assim os *scores* apresentados na Seção 2.1.3 são calculados para cada questão respondida pelo especialista. É apresentado na Tabela 3.1 um exemplo de respostas armazenada no banco de dados da aplicação para as perguntas sobre qual valor a série N1037 vai assumir nos instantes: 45, 46, 47, 48 e 49.

id_estatisticas	id_pergunta	id_processos	media	moda	mediana	variância
12341	474	881	4948.5	5329.26	5171.95	699181.08
12351	475	881	5034	5351.85	5254.62	696177.33
12361	476	881	5318	5292.68	5355.96	592209.33
12371	477	881	5291	5281.25	5337.96	608652.33
12381	478	881	5362	5260.86	5349.62	457289.33

Tabela 3.1: Exemplo de opinião do especialista a respeito de determinada variável de interesse

Da Tabela 3.1, são exibidos a moda, média, mediana e variância, para as perguntas com id “474”, “475” “478” “477” “478” ligadas ao processo “881”. A média representa o valor estimado pelo respondente para a variável, enquanto que a mediana representa o valor cuja a probabilidade da variável não ultrapassá-lo é de 50%. Por sua vez, a moda representa o valor mais provável a ser assumido pela variável de acordo com o método de educação. Por fim, a variância reflete o nível de incerteza do respondente, de forma que quanto menor ele for, menos incerteza o respondente expressa. Portanto para a pergunta “Qual será o valor

assumido pela série N1037 no instante 47”, os modelos cognitivos acreditam ser “5318”, já que essa é a média da sua distribuição de probabilidades da opinião desse especialista.

A estrutura de persistência de dados da aplicação foi desenvolvida em MySQL [74]. O diagrama entidade relacionamento do banco de dados da aplicação é exibido na Figura 3.9.

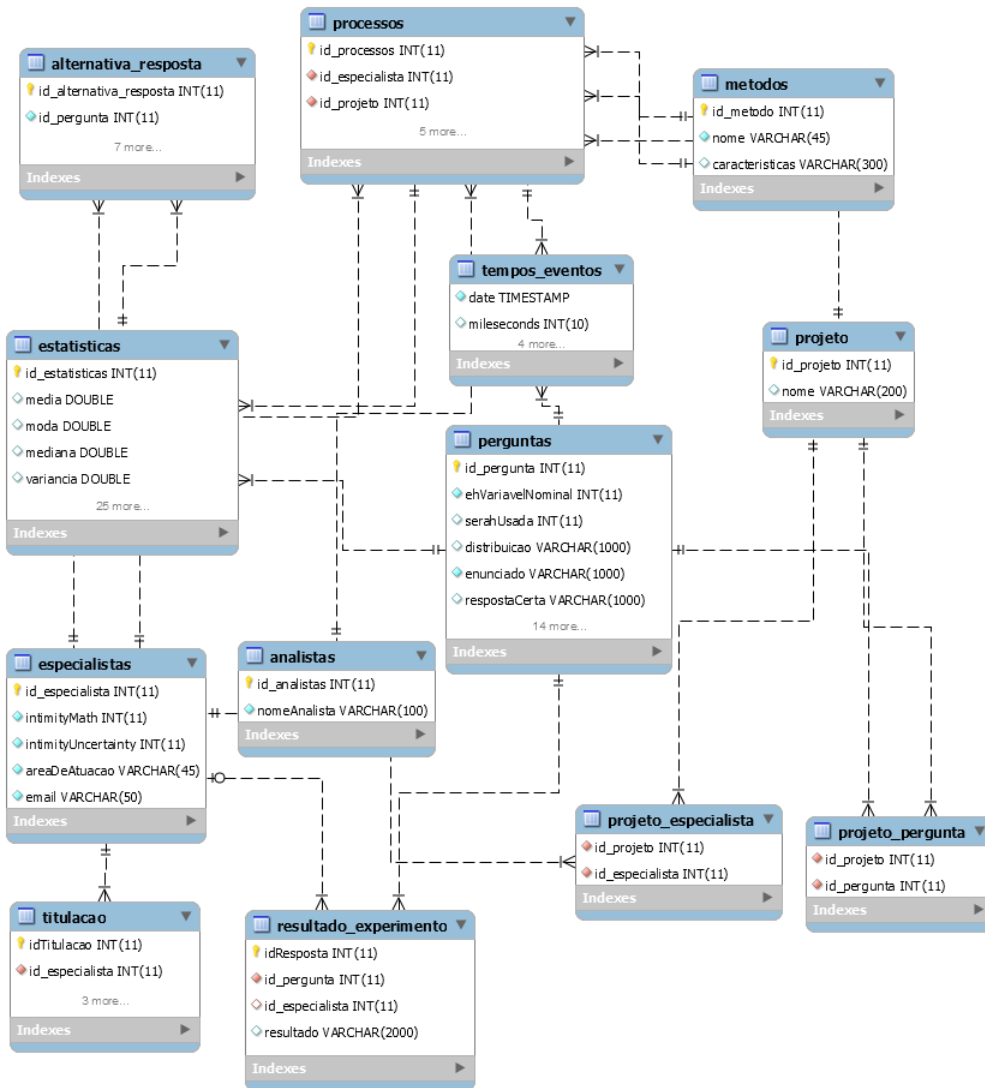


Figura 3.9: Diagrama entidade relacionamento do banco de dados da aplicação

Cada uma das entidades apresentadas nesse esquema representa:

- alternativa_resposta: Armazena as alternativas de resposta para cada pergunta registrada no banco;
- analista: Registro dos analistas cadastrados na aplicação;

-
- especialistas: Armazena os especialistas da aplicação e suas informações que foram solicitadas em seu cadastro na aplicação;
 - estatísticas: Contém as respostas dos especialistas para determinada pergunta além de outras informações como *scores* das suas respostas;
 - metodos: Contém os métodos de educação do conhecimentos implementados na aplicação;
 - perguntas: Armazena as perguntas cadastradas na aplicação;
 - processos: Armazena variáveis que relacionam projetos, especialistas, métodos de educação e analistas envolvidos em determinado processo;
 - projeto: Registro dos projetos cadastrados na aplicação;
 - projeto_especialista: Relaciona os projetos que o especialista esteja ligado;
 - projeto_pergunta: Relaciona as perguntas a um determinado projeto;
 - resultado_experimento: Contém a opinião do especialista para uma determinada pergunta;
 - tempos_eventos: Registra o tempo consumido pelo especialista em cada uma das perguntas;
 - titulacao: Armazena os títulos dos especialistas exemplo (graduação, mestrado, doutorado etc).

Capítulo 4

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 entre si. Essa característica permite que seja possível elaborar modelos para prever por exemplo qual seria o próximo valor assumido por determinada série temporal [26]. Os primeiros modelos de séries temporais surgiram nos Estados Unidos em 1939 na área da econometria [24]. Desde então as séries temporais passaram a ser usadas em várias áreas como economia (bolsa de valores), medicina (eletrocardiograma), meteorologia (temperaturas mensais), etimologia (casos de dengue em determinada área), fenômenos naturais (manchas solares) etc. Os objetivos em analisar séries temporais podem ser vários, como por exemplo: descrever as propriedades da série, explicar o comportamento de uma série através da análise de outra, previsão dos pontos futuros da série e estudar o fenômeno que controla a série [26]. Contudo um dos principais objetivos no estudo de uma série é a sua previsão. Para isso devem ser considerados muitas variáveis intrínsecas a série e também fenômenos do ambiente externo que possam estar relacionados a ela. Considere Y_i uma série temporal onde: $Y_i = y_i \in R | i = 1, 2, 3, \dots, N$, y_i é uma observação ordenada cronologicamente dada a ocorrência dos eventos e N é o número de observações. Para exemplificar séries temporais será apresentado a série *sunspots*. Essa série representa as manchas solares na superfície do sol. Manchas solares são regiões do sol que tem temperatura menor do que a média local que é de 5780K. As manchas solares são catalogadas desde 1700 e essa é uma série muito comentada nos livros a respeito de análises de séries temporais [75]. A Figura 4.1 apresenta essa série medida em anos onde o eixo horizontal representa os anos das amostras e o eixo

vertical representa a quantidade de manchas solares observadas.

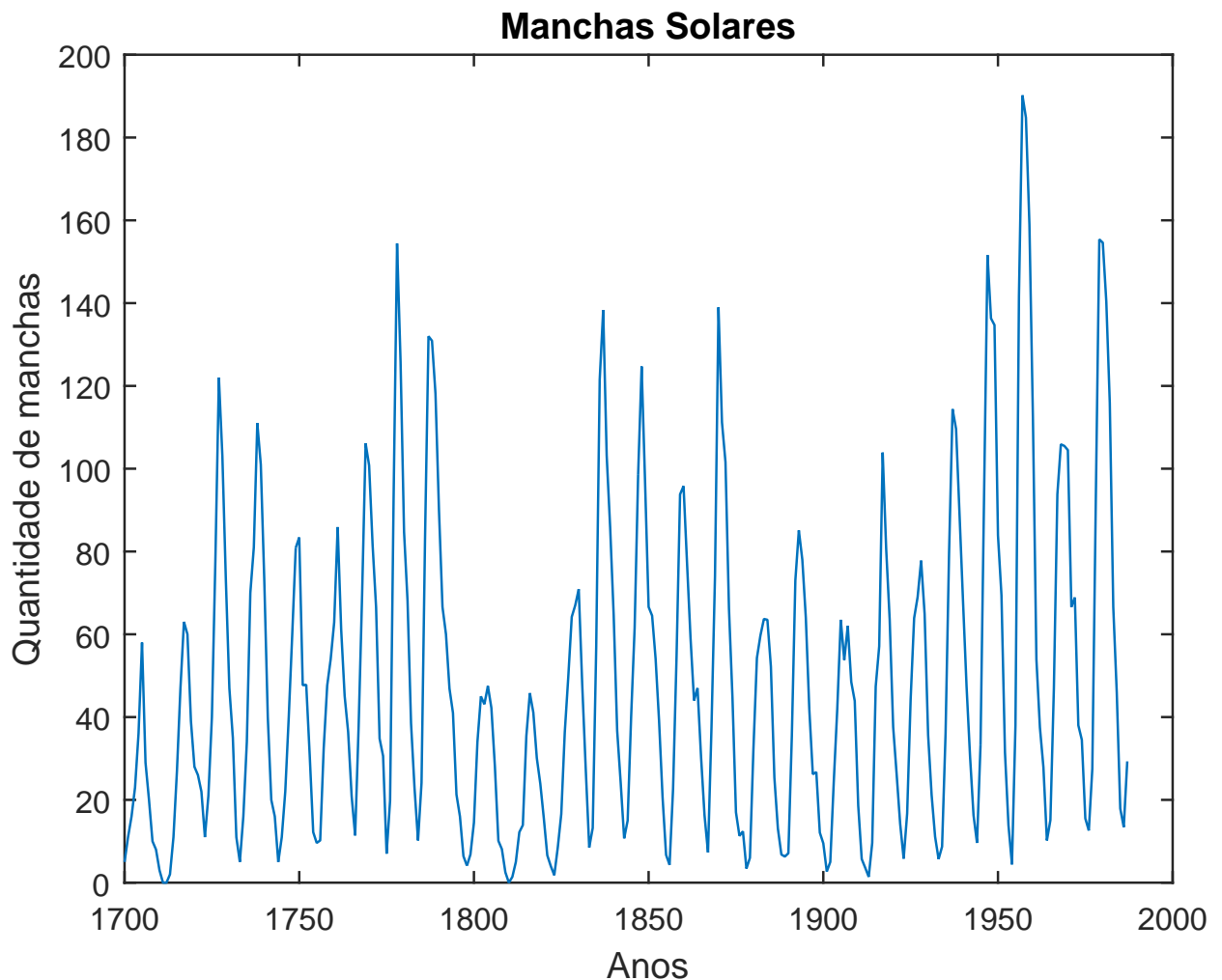


Figura 4.1: Manchas solares anuais de 1700 a 1986

Algumas propriedades de séries temporais são muito importantes especialmente no que se diz respeito a sua predição. A sazonalidade de uma série é um certo padrão que tende a se repetir a cada determinado período de tempo. Na série das manchas solares é possível observar que existe um comportamento cíclico, com um crescimento precedido de uma queda no número de manchas solares durante os anos. Outra característica de uma série temporal é a tendência. Uma série pode exibir uma tendência de crescimento ou decréscimo de seus valores. Observando a série temporal das manchas solares é possível inferir que a quantidade de manchas solares apresenta uma tendência crescente ao longo dos anos. Uma característica muito importante de uma série temporal é a sua estacionariedade. Esse é um atributo elementar no que se diz respeito 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 ele é estacionário se: $E[y_t] = \mu_t = \mu$; média constante para qualquer tempo t , sua variância é constante representado por $E[(y_t - \mu)^2] = E[(y_{t-s} - \mu)^2] = \sigma_y^2$ e a sua covariância também é constante descrita por $E[y_t, y_s] = [(y_t - \mu_t)(y_s - \mu_s)] = \gamma_{(|x-s|)}$ [24]. 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 [76].

4.1 Séries avaliadas no trabalho

As séries utilizadas nesse trabalho são provenientes da *M3-Competiton* [33]. 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 que as séries pertencem e as predições 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 8 séries temporais que tem seus dados exibidos na Tabela 4.1. Para um maior detalhamento dos dados utilizados, todos os valores dos pontos das séries temporais desse trabalho estão discriminados no Apêndice A, Tabela A.1 até Tabela A.8.

4.2 Métricas de comparação

Um dos principais objetivos no estudo de séries temporais é a previsão dos pontos futuros dessas séries. Existem vários modelos matemáticos que se destinam a predição de séries temporais como por exemplo modelos ARIMA [27], redes neurais artificiais [77], modelos bayesianos [78], TAEF [79] e outros mais. A acurácia desses modelos é um fator crítico para a análise de séries temporais [33]. Por conta disso muitas métricas para avaliação das previsões são utilizadas. Além dos *scoring rules* já apresentadas, outras métricas serão utilizadas na comparação do desempenho dos modelos e serão mais bem detalhadas a seguir.

	Quant.	Pontos	Valor	Valor	Valor	Ordem	Tipo
	Pontos	Previstos	Mínimo	Médio	Máximo	Cronológica	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 4.1: Informações das séries temporais utilizadas no trabalho.

Uma das medidas mais utilizadas para comparação da acurácia em previsões de séries é o *Mean Square Error* (MSE). Essa medida é definida por:

$$MSE = \frac{1}{N} \sum_{j=1}^N (d_j - p_j)^2 \quad (4.1)$$

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 predição. 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:

$$MAPE = \frac{100}{N} \sum_{j=1}^N \left| \frac{(d_j - p_j)}{d_j} \right| \quad (4.2)$$

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 predição. Quanto menor for o valor do MAPE melhor é a previsão.

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

$$POCID = 100 \frac{\sum_{j=1}^N D_j}{N} \quad (4.3)$$

Onde

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

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 predição 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:

$$ARV = \frac{1}{N} \frac{\sum_{j=1}^N (p_j - d_j)^2}{(p_j - \bar{d})^2} \quad (4.4)$$

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 predição 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 predição 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 predição para o próximo ponto.

Capítulo 5

Modelos ARIMA

O *autoregressive integrated moving average* (ARIMA) é um modelo estatístico linear para a previsão de séries temporais formalizado no livro de Box e Jenkins [27]. Esse é um dos modelos mais comentados na literatura para predição de séries temporais lineares e estacionárias. O modelo ARIMA é uma integração de dois modelos, o auto regressivo e o de médias móveis. O modelo é classificado em ARIMA(p,d,q), onde:

- p é o número de termos auto regressivos;
- d é o número de diferenciações necessárias para 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 como:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (5.1)$$

Onde ε_t representa o ruído branco da série, ϕ_p os parâmetros auto regressivos e y_t é um determinado ponto na série temporal.

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 como:

$$y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (5.2)$$

Onde ε_t representa o ruído branco da série, θ_q parâmetros do modelo de médias móveis e y_t é um determinado ponto na série temporal.

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

$$\nabla y_t = y_t - y_{t-1} \quad (5.3)$$

Onde ∇y_t é a derivação de determinado ponto 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 [25]. Essa função é definida por:

$$\rho(\tau) = \frac{\gamma_\tau}{\gamma_0} = \frac{E[(y_t - \mu)(y_{t+\tau} - \mu)]}{E[(y_t - \mu)^2]} \quad (5.4)$$

A Figura 5.1 apresenta a função de auto correlação para a série temporal das manchas solares apresentada no Capítulo 4 obtida através do MatLab [80].

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 e ϕ_{kk} o último elemento, então de acordo com [25] temos:

$$\rho_j = \phi_{k1}\rho_{j-1} + \dots + \phi_{k(k-1)}\rho_{j-k+1} + \phi_{kk}\rho_{j-k}; \text{ com } j = 1, 2, 3, \dots, k \quad (5.5)$$

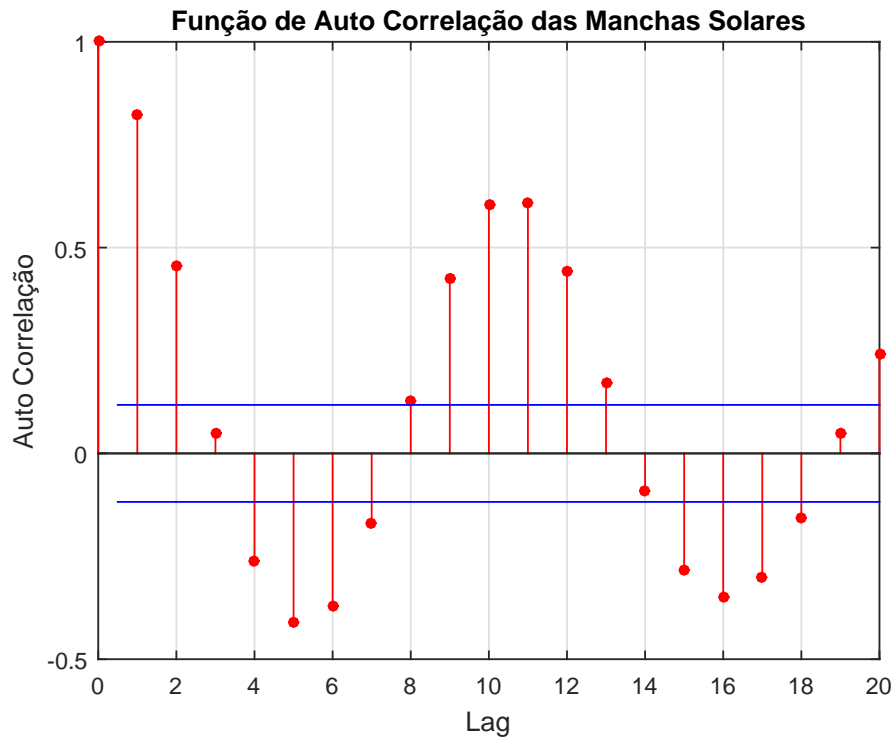


Figura 5.1: Função de auto correlação da série das manchas solares apresentada na Figura 4.1

As equações de Yule-Walker levam a essa expressão, que depois de resolvidas para um caso específico de um modelo AR(2) por exemplo resulta em:

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \quad (5.6)$$

A Figura 5.2 apresenta a Função de Auto Correlação Parcial da série temporal das manchas solares apresentada no Capítulo 4 obtida no MatLab [80].

Para verificar o ajuste do modelo a série é necessário utilizar algum critério para verificar a adequação dos parâmetros estimados. Os critérios mais utilizados são o *Akaike Information Criterion* (AIC) e o *Bayesian Information Criterion* (BIC), calculados como: $AIC = \ln \sigma_t^2 + \frac{2}{T}(p + q)$ e $BIC = \ln \sigma_t^2 + \frac{\ln T}{T}(p + q)$, onde T é o número de observações usadas, p e q os parâmetros estimados para o modelo e σ_t^2 quadrado da variância da série em determinado

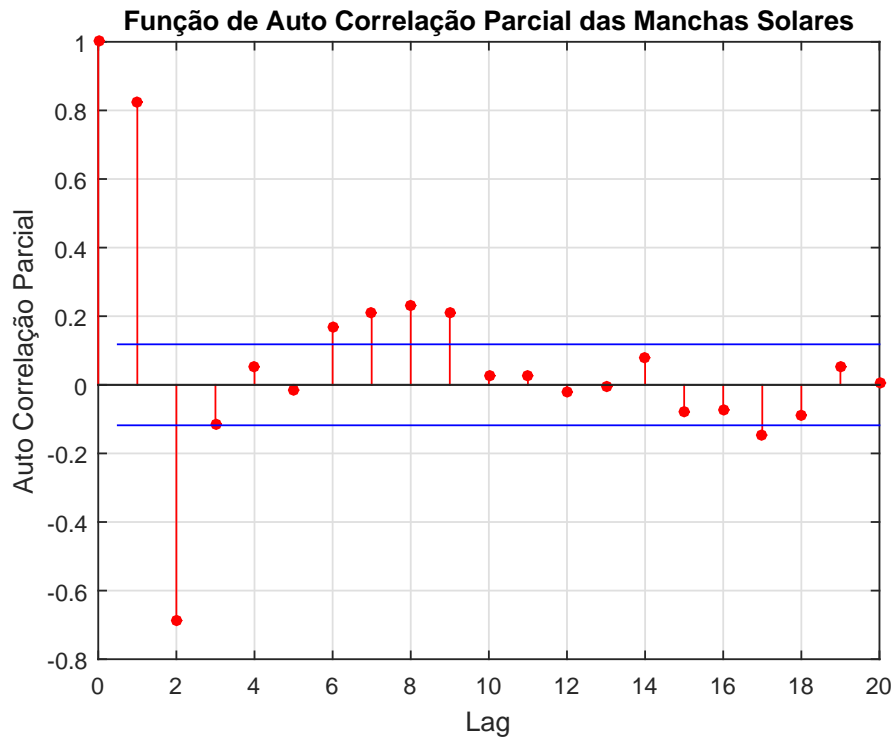


Figura 5.2: Função de auto correlação parcial da série das manchas solares apresentada na Figura 4.1

ponto t . Os valores do AIC e o BIC deverão ser os menores possíveis.

Para avaliar qual dos modelos ARIMA melhor se ajusta a série é necessário realizar uma análise das funções de auto correlação e auto correlação parcial da série. Esses modelos partem da suposição que a série seja estacionária, caso contrário é necessário uma quantidade de diferenciações d para os parâmetros serem estimados. A Tabela 5.1 apresenta os comportamentos das funções para os seus respectivos modelos e parâmetros.

Modelo	Função de Auto correlação	Função de Auto correlação Parcial
AR(p)	Apresenta um declínio gradual	Apresenta uma grande queda entre os lags da função (<i>cut off</i>) após o lag p
MA(q)	Apresenta uma grande queda entre os lags da função (<i>cut off</i>) após o lag q	Apresenta um declínio gradual
ARMA(p,q)	Apresenta um declínio gradual	Apresenta um declínio gradual

Tabela 5.1: Caracterização das estimações de parâmetros dos modelos ARIMA pela análise das funções de auto correlação e auto correlação parcial

5.1 Estimação dos modelos ARIMA para as séries desse trabalho

A série N1269 será utilizada para ilustrar o procedimento adotado para estimar os parâmetros dos modelos ARIMA e suas respectivas previsões. Todos os procedimentos demonstrados aqui foram realizados no MatLab versão 8.5.0. Primeiramente é obtido as funções de auto correlação e auto correlação parcial da série apresentadas na Figura 5.3:

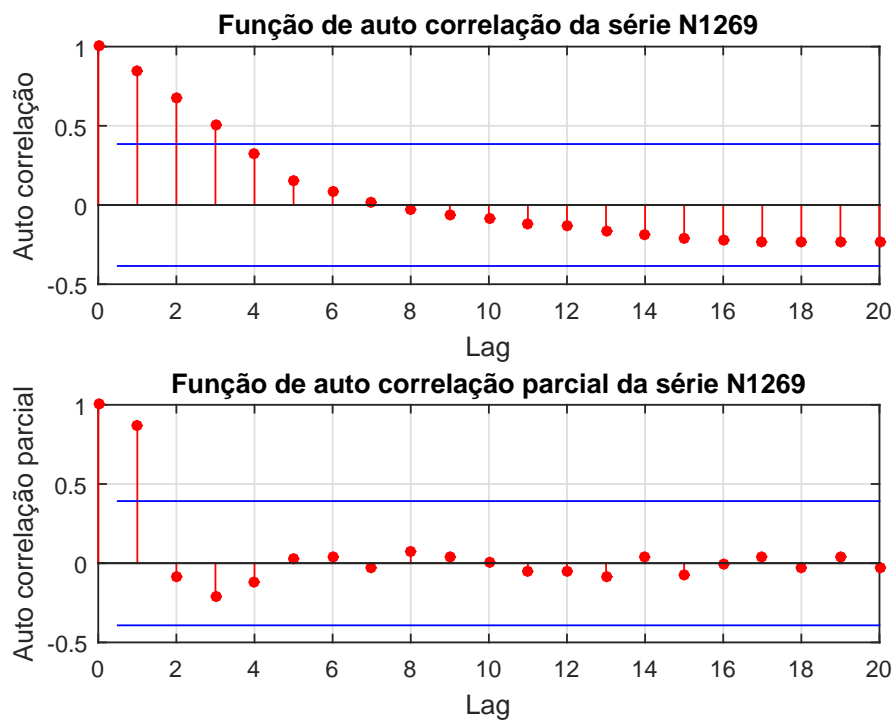


Figura 5.3: Funções de auto correlação e auto correlação parcial da série N1269

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^1 1 da função de auto correlação parcial. Para comprovar esse modelo, é utilizado um pseudocódigo que gera os valores da métrica BIC para os modelos dessa série. Esse pseudocódigo, baseado no pseudocódigo presente na documentação do MatLab [80], é dado por:

$$Y = [9439.4, 9272.4, 8917.3, 8373.7, 8069.6, 6499.9, 6183.4, 5825, 5550.7,$$

¹Lag é um elemento da janela de observação na série temporal.

```
5426.8, 5363.8, 5249.8, 5281.4, 5311, 5258.3, 5142.8, 5074.4, 5039, 5058.9,
5070.7, 5119.6, 5120.8, 5111, 5117.2, 5084.9, 5107.9, 5126.2];
```

```
LOGL = zeros(3,3); %Initialize
PQ = zeros(3,3);
for p = 1:3
    for q = 1:3
        mod = arima(p,0,q);
        [fit,~,logL] = estimate(mod,Y,'print',false);
        LOGL(p,q) = logL;
        PQ(p,q) = p+q; %add ''-2'' here to get other column
    end
end
LOGL = reshape(LOGL,9,1);
PQ = reshape(PQ,9,1);
[~,bic] = aicbic(LOGL,PQ+1,27);
reshape(bic,3,3)
```

Esse pseudocódigo executa métodos da classe ARIMA do Matlab, mais especificamente *setLagOp* e *estimate*. Desse código é gerado uma matriz com os valores do BIC para cada um dos modelos ARIMA, sendo possível comparar qual modelo possui o menor valor de BIC.

463.9842	470.5759	469.7425	461.6183
387.0964	393.6880	414.1053	397.3057
464.1220	470.7136	469.7425	420.6970

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 e o gráfico (Figura 5.4) com as previsões do modelo escolhido:

```
y = [ 9439.4, 9272.4, 8917.3, 8373.7, 8069.6, 6499.9, 6183.4, 5825, 5550.7,
5426.8, 5363.8, 5249.8, 5281.4, 5311, 5258.3, 5142.8, 5074.4, 5039, 5058.9,
5070.7, 5119.6, 5120.8, 5111, 5117.2, 5084.9, 5107.9, 5126.2];
```

```
Mdl = arima(1,0,0);  
EstMdl = estimate(Mdl,y);  
[Y,YMSE] = forecast(EstMdl,5,'Y0',y);  
  
figure  
h1 = plot(y,'Color',[.7,.7,.7]);  
hold on;  
h2 = plot(28:32,Y,'r:','LineWidth',2);  
legend([h1 h2],'série N1269','Previsão',...  
'Location','North')  
title('Previsão para a série 1269')
```

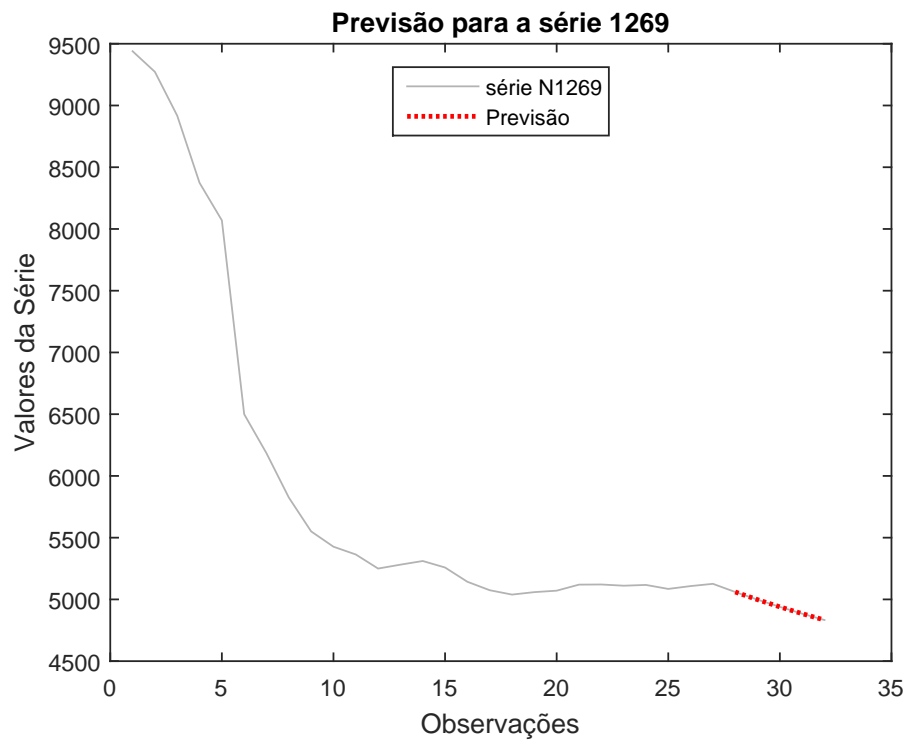


Figura 5.4: Representação gráfica da previsão do modelo ARIMA(1,0,0) para a série N1269

A Tabela 5.2 apresenta os modelos estimados por esse procedimento para as séries utilizadas nesse trabalho e seus respectivos coeficientes de BIC.

Série	Modelo 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 5.2: Modelos ARIMA selecionados para as séries e seus respectivos coeficientes de BIC

Capítulo 6

Comparação entre os modelos

Para a comparação da acurácia dos modelos cognitivos obtidos através da educação do conhecimento e dos modelos matemáticos obtidos através dos modelos ARIMA é apresentado nas próximas tabelas (Tabela 6.1 até Tabela 6.8) a representação gráfica das séries temporais com suas respectivas estimações dos modelos além das métricas para a previsão de cada série apresentadas na Seção 4.2. Os valores corretos das predições, previsão dos modelos cognitivos e previsão dos modelos ARIMA estão presentes no Apêndice A na Tabela A.9. Os gráficos gerados para as comparações foram gerados no MatLab.

As comparações entre os modelos foram realizadas de maneira que um modelo deve apresentar um desempenho considerável em relação ao outro para ser considerado superior em algum critério. As séries que apresentaram resultados muito similares para determinado critério foram desconsideradas na comparação. Na análise do *Mean Squared Error* (MSE) as opiniões dos especialistas obtiveram um resultado superior, com 5 séries (N0236, N0290, N1459, N1708 e N2355) e os modelos matemáticos apresentaram resultado superior em três séries (N1037, N1268 e N1269). No *Mean Absolute Percentage Error* (MAPE) as opiniões dos especialistas obtiveram resultado inferior aos modelos ARIMA, os especialistas obtiveram um melhor em duas séries (N0290 e N2355) já os modelos ARIMA obtiveram melhor resultado em 5 séries (N0236, N1037, N1268, N1269 e N1459). Também na análise *Prediction of Change in direction* (POCID) os modelos matemáticos obtiveram melhor resultado, sendo 2 séries melhor avaliadas para os especialistas (N0236 e N1459) contra 3 dos modelos matemáticos (N0290, N1037 e N1269). Ainda sobre essa métrica três observações, predição dos modelos matemáticos e cognitivos para a série N2355 e a predição dos modelos cognitivos para a

série N0290, apresentaram POCID igual a 50% o que significa que as predições tem acurácia equivalente a um *random walk*¹. Na medida *Average Relative Variance* (ARV) apenas uma das previsões obteve resultado superior a 1, a previsão dos modelos cognitivos para a série N1708. Na avaliação do *scoring rule* quadráticos houve uma superioridade dos especialistas, onde eles obtiveram um melhor resultado em 5 séries (N0236, N0290, N1037, N1708 e N2355) enquanto que os resultados dos modelos matemáticos foram superiores em outras 3 séries (N1268, N1269 e N1459).

Os resultados dos modelos cognitivos foram promissores para a área de previsões de séries temporais, confirmando a afirmativa de Lawrence [30] onde ele propõe que modelos baseados em opiniões de especialistas são ao menos tão bons quantos 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 das séries contribuiu no desempenho de ambos modelos 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 tenha conhecimentos específicos na área de previsão de séries temporais, com acurácia em um patamar considerável em comparação aos modelos ARIMA.

¹*Random walk* ou caminho aleatório é uma formalização matemática que consiste numa sucessão de passos aleatórios.

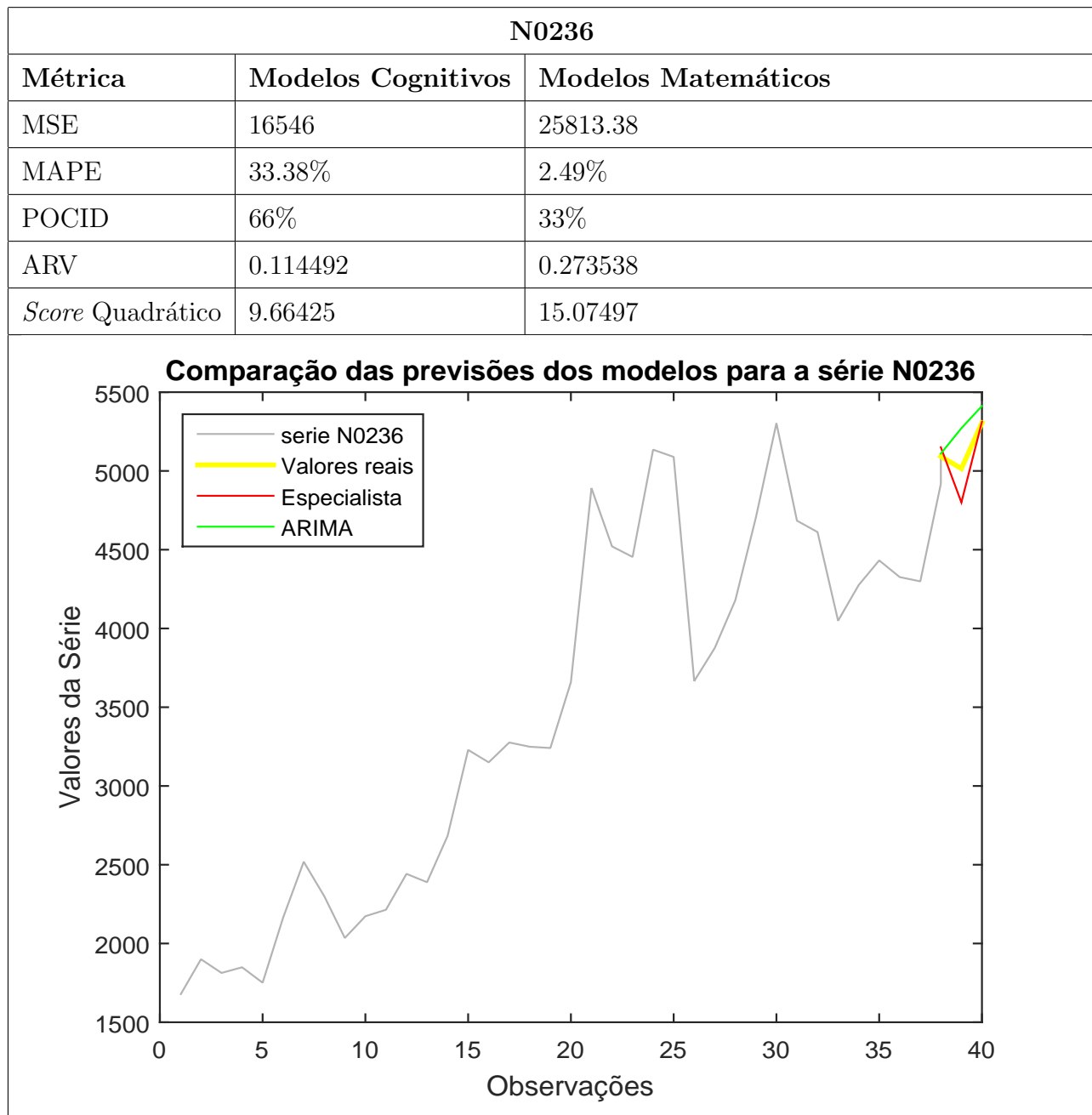


Tabela 6.1: Análise comparativa da série N0236

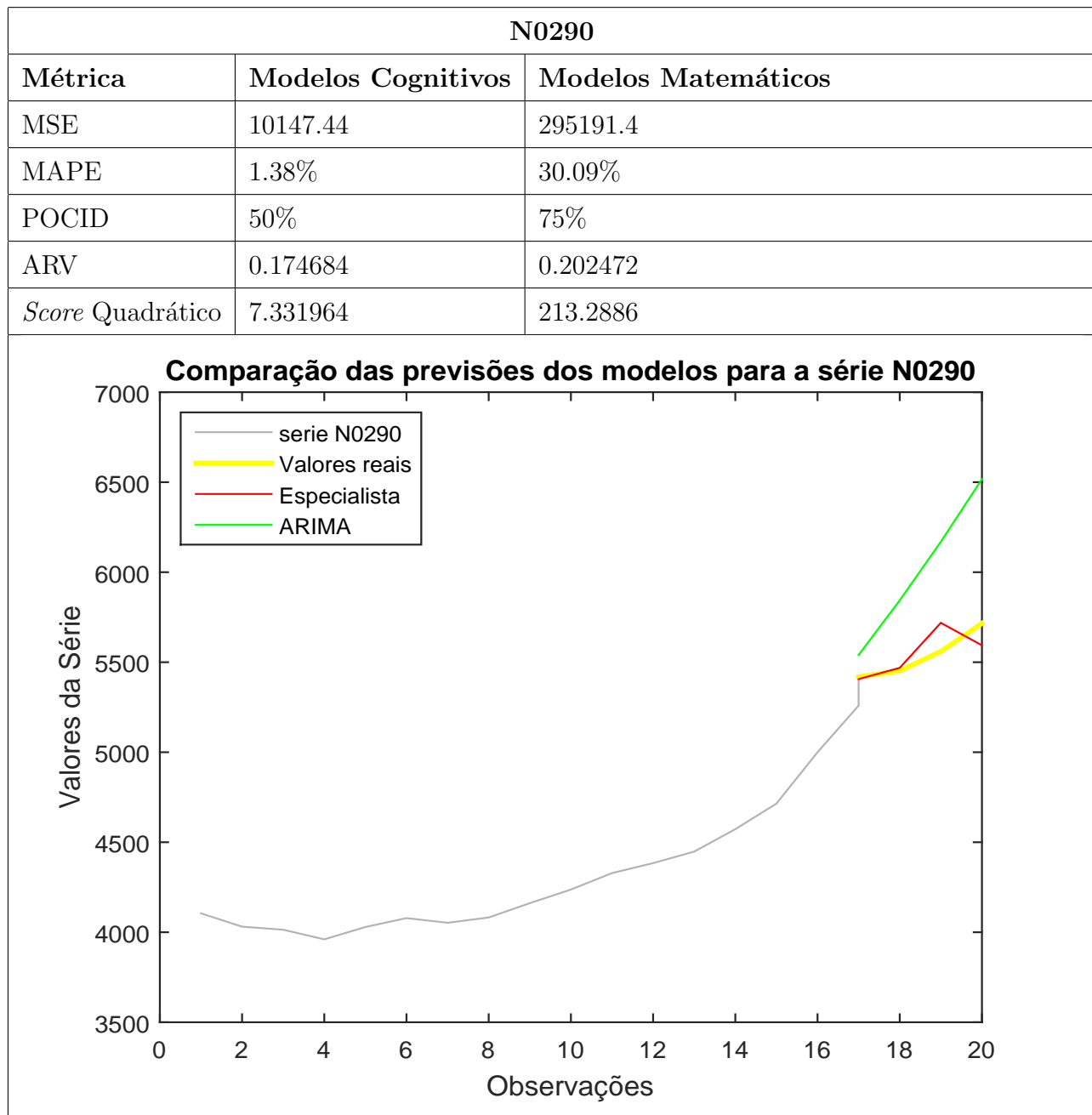


Tabela 6.2: Análise comparativa da série N0290

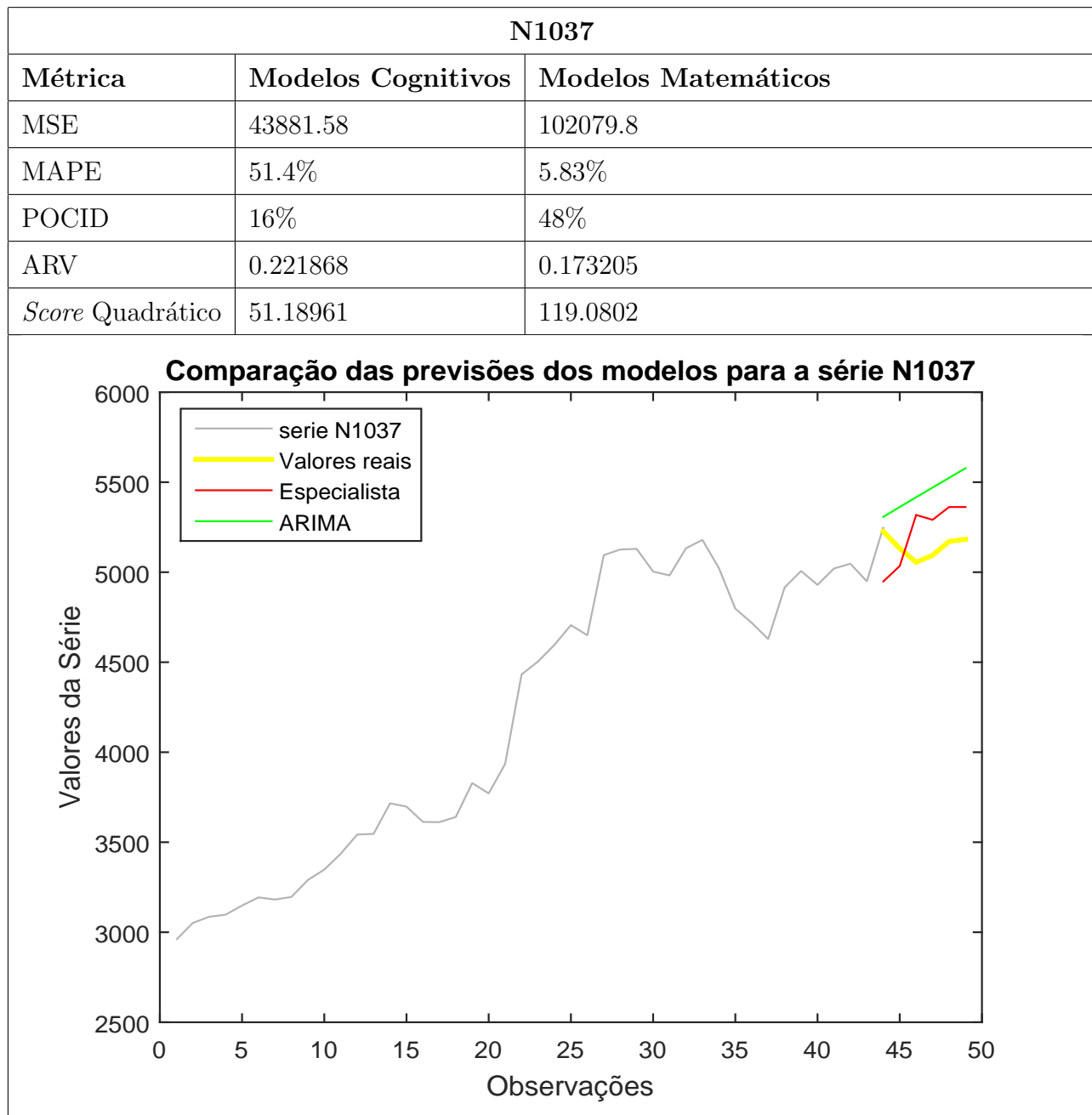


Tabela 6.3: Análise comparativa da série N1037

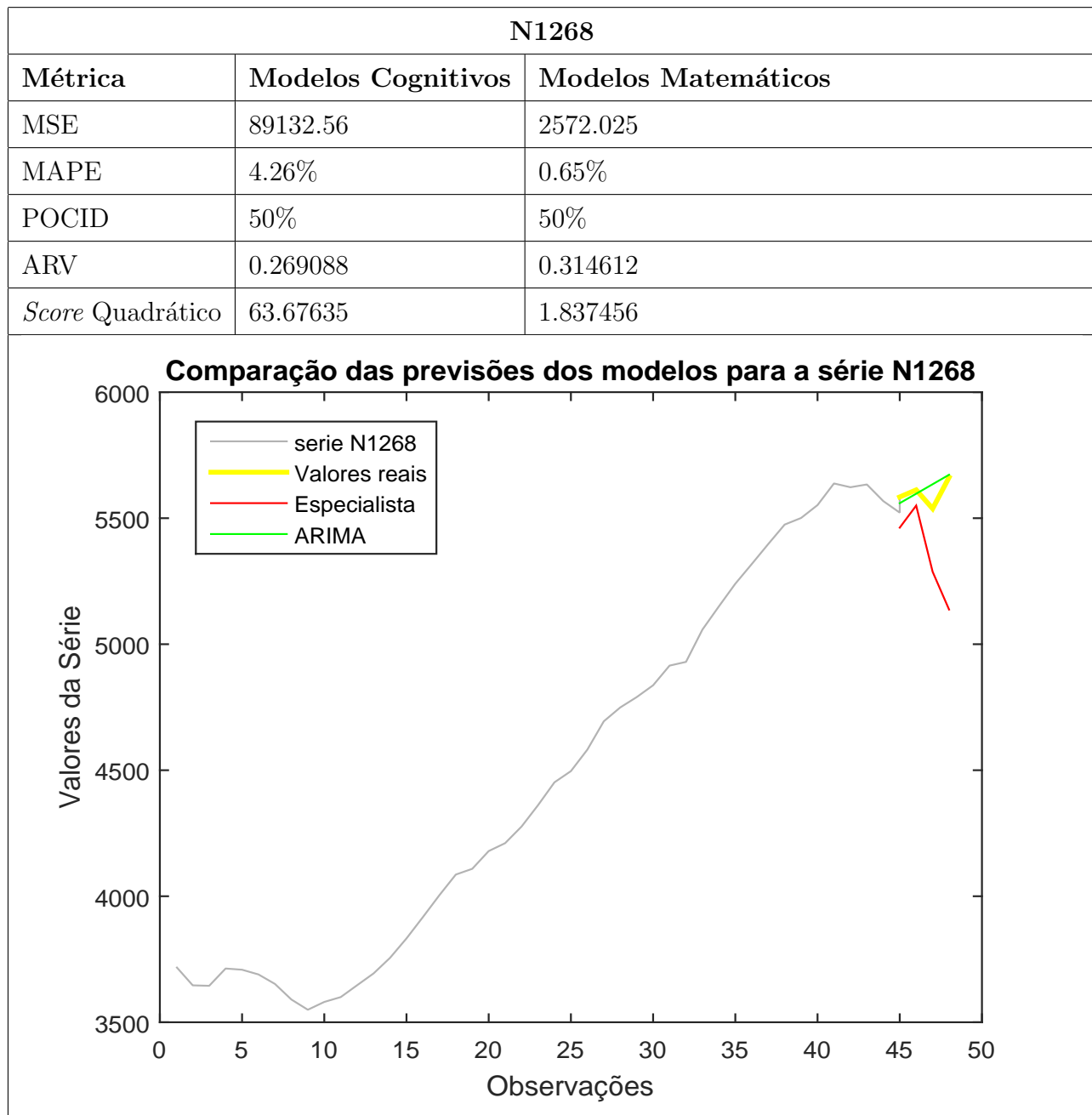


Tabela 6.4: Análise comparativa da série N1268

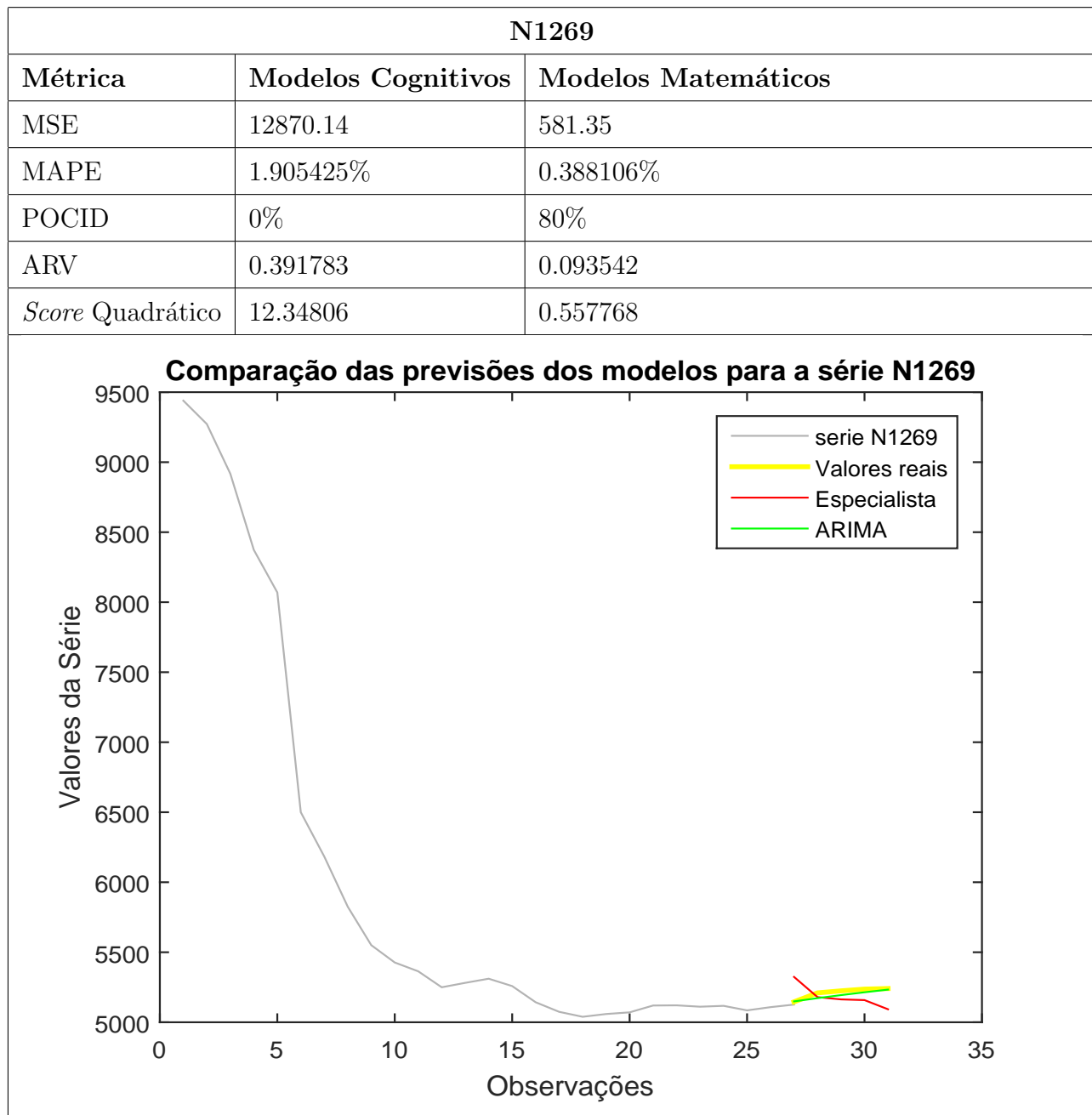


Tabela 6.5: Análise comparativa da série N1269

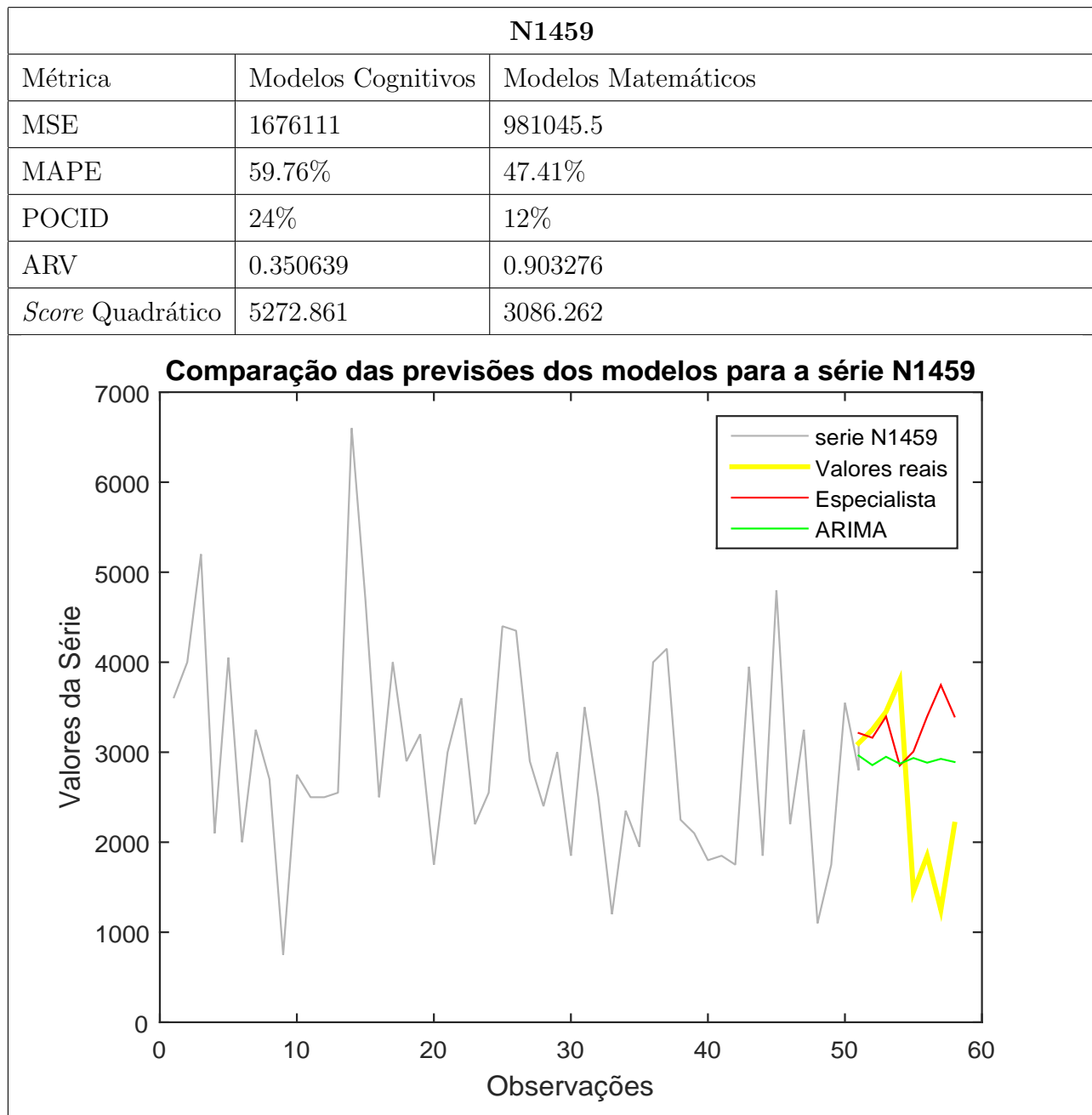


Tabela 6.6: Análise comparativa da série N1459

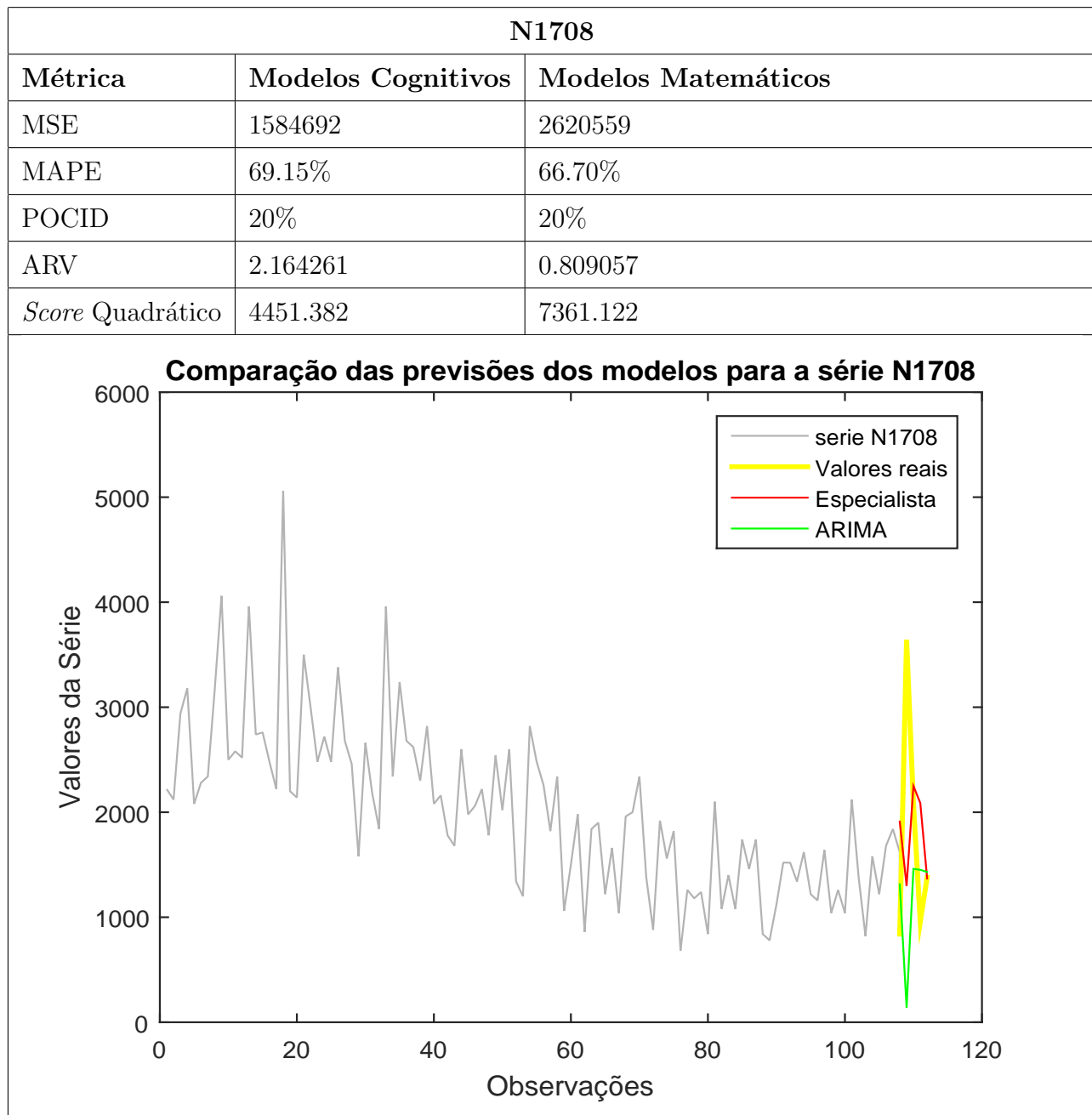


Tabela 6.7: Análise comparativa da série N1708

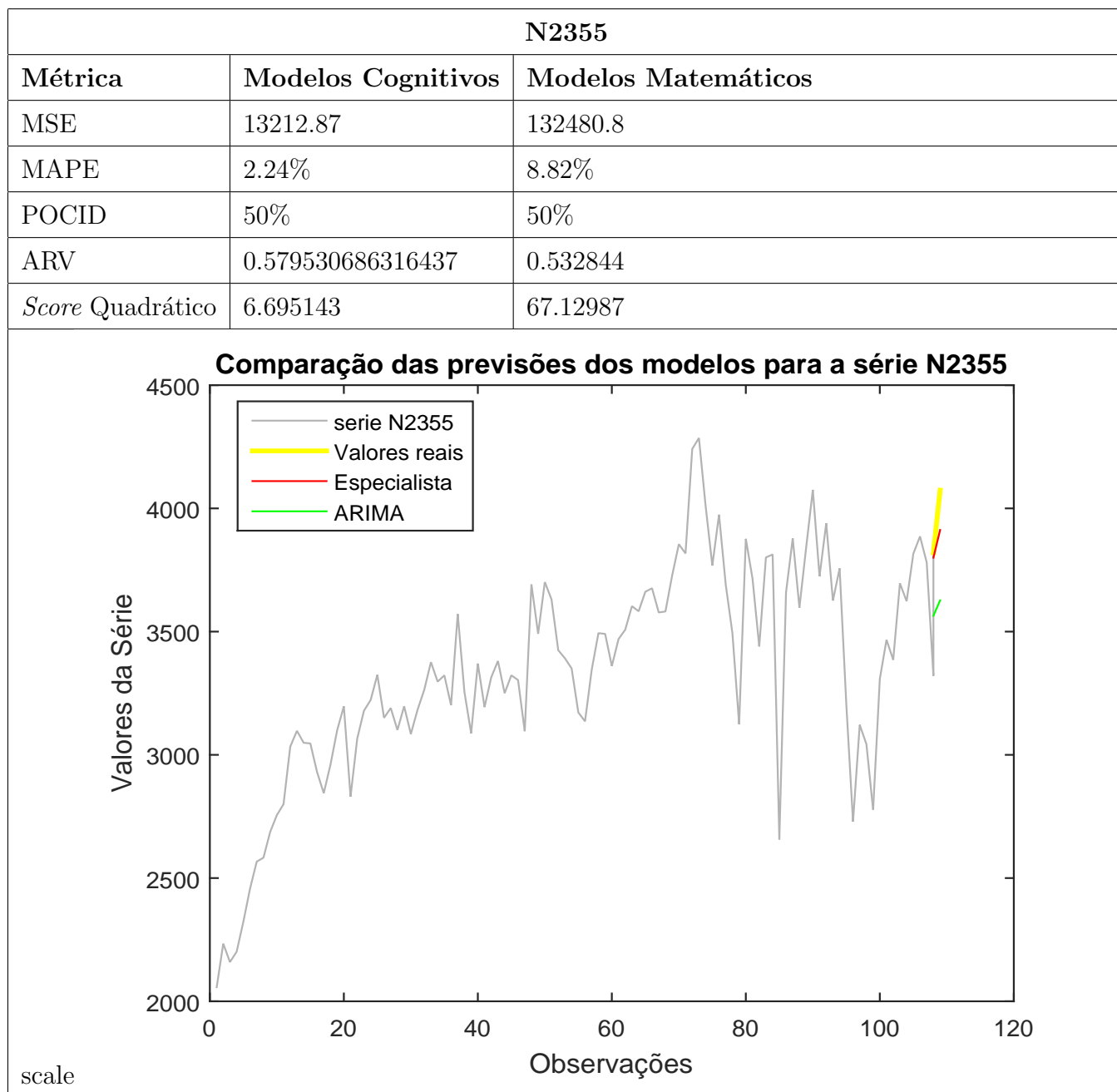


Tabela 6.8: Análise comparativa da série N2355

Capítulo 7

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 incrível 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. Contudo é evidente que opiniões de especialistas são recursos viáveis para modelagem de vários problemas incluído 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 predições para ambos modelos. Além disso o tamanho das séries não favoreceu o uso de outros modelos matemáticos. Contudo a comparação entre os modelos foi realizada de maneira condizente e os resultados obtidos foram satisfatórios. Esses resultados demonstraram que as opiniões dos especialistas têm 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.

Como sugestão para possíveis trabalhos futuros pode-se citar:

- Agregação de modelos cognitivos e matemáticos através de funções de cópulas;

- Ajuste de predições de modelos matemáticos por especialistas através de educação do conhecimento;
- Utilização de opiniões de especialistas em modelagem de sistemas complexos.

Referências Bibliográficas

- [1] Jianchao 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.
- [2] Paulo Renato A Firmino and Enrique López Droguett. 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.
- [3] Wagner Barbosa dos Santos, Andrés López Droguett, et al. Análise probabilística de riscos via redes bayesianas: Uma aplicação na construção de poços multilaterais. 2005.
- [4] Ricardo José Ferreira, Paulo Estevão Lemos de Oliveira, Paulo Renato Alves Firmino, and Enrique Andrés López Droguett. Mapa estratégico para programas de recursos humanos: avaliando o desempenho de redes bayesianas. *Gestão & Produção*, 17(1), 2010.
- [5] Kevin P McCormack, Marcelo Bronzo, and Marcos Paulo Valadares Oliveira. Uma abordagem probabilística para a avaliação de riscos em cadeias de suprimento. *Revista Produção Online*, 10(3):577–598, 2010.
- [6] Bilal M Ayyub. *Elicitation of expert opinions for uncertainty and risks*. CRC press, 2001.
- [7] Roger M Cooke. Experts in uncertainty: opinion and subjective probability in science. 1991.
- [8] Nick F Pidgeon, Barry A Turner, and David I Blockley. The use of grounded theory for conceptual analysis in knowledge elicitation. *International Journal of Man-Machine Studies*, 35(2):151–173, 1991.

-
- [9] RT Clemen and T Reilly. Making hard decisions with decisiontools. 2001.
- [10] 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.
- [11] Paul H Garthwaite, Joseph B Kadane, and Anthony O’Hagan. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470):680–701, 2005.
- [12] Nigel Shadbolt and Paul R Smart. Knowledge elicitation: Methods, tools and techniques. 2015.
- [13] Carrie V Kappel, Benjamin S Halpern, Kimberly A Selkoe, and Roger M Cooke. Eliciting expert knowledge of ecosystem vulnerability to human stressors to support comprehensive ocean management. In *Expert Knowledge and Its Application in Landscape Ecology*, pages 253–277. Springer, 2012.
- [14] AB Knol, P Slottje, JP van der Sluijs, and E Lebrete. The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental health: a global access science source*, 9:19–19, 2009.
- [15] Samantha Low Choy, Rebecca O’Leary, and Kerrie Mengersen. Elicitation by design in ecology: using expert opinion to inform priors for bayesian statistical models. *Ecology*, 90(1):265–277, 2009.
- [16] Sindhu R Johnson, George A Tomlinson, Gillian A Hawker, John T Granton, Haddas A Grosbein, and Brian M Feldman. A valid and reliable belief elicitation method for bayesian priors. *Journal of clinical epidemiology*, 63(4):370–383, 2010.
- [17] Paulo Renato Firmino, RDCS Menezes, EL Droguett, DC de Lemos Duarte, et al. Eliciting engineering judgments in human reliability assessment. In *Reliability and Maintainability Symposium, 2006. RAMS’06. Annual*, pages 512–519. IEEE, 2006.
- [18] Enrique López Droguett, Carlos Magno Couto Jacinto, Regilda Menêzes, Paulo Renato Firmino, Andréa Pontual, Gabriel Sotomayor, and Pauli AA Garcia. Probability assessment of offshore oil multilateral wells construction process in brazil (psam-0173). In *Proceedings of the Eighth International Conference on Probabilistic Safety Assessment & Management (PSAM)*. ASME Press, 2006.

-
- [19] PRA FIRMINO, RLM SALES FILHO, and EL DROGUETT. An expert opinion elicitation method based on bayesian intervals estimation and computational algorithm: an application to oil refinery risk analysis. In *INTERNATIONAL PROBABILISTIC SAFETY ASSESSMENT AND MANAGEMENT CONFERENCE*, volume 9, 2008.
- [20] Linda C van der Gaag, Silja Renooij, CLM Witteman, Berthe MP Aleman, and Babs G Taal. Probabilities for a probabilistic network: a case study in oesophageal cancer. *Artificial Intelligence in medicine*, 25(2):123–148, 2002.
- [21] Mark Oremus, Jean-Paul Collet, Jacques Corcos, and Stanley H Shapiro. A survey of physician efficacy requirements to plan clinical trials. *Pharmacoepidemiology and drug safety*, 11(8):677–685, 2002.
- [22] Magne Jørgensen. An empirical evaluation of the mkii fpa estimation model. In *Norwegian Informatics Conference*, pages 7–18, 1997.
- [23] Magne Jørgensen. A review of studies on expert estimation of software development effort. *Journal of Systems and Software*, 70(1):37–60, 2004.
- [24] Gebhard Kirchgässner, Jürgen Wolters, and Uwe Hassler. *Introduction to modern time series analysis*. Springer Science & Business Media, 2012.
- [25] George Box, Gwilym M Jenkins, and Gregory Reinsel. *Time series analysis: Forecasting & control*. 1994.
- [26] Ricardo S Ehlers. Análise de séries temporais. *Laboratório de Estatística e Geoinformação*. Universidade Federal do Paraná, 2007.
- [27] George Box. P. and jenkins, gm. *Time series analysis: forecasting and control*, 1970.
- [28] José Luis Aznarte and José Manuel Benítez. The links between statistical and fuzzy models for time series analysis and forecasting. In *Time Series Analysis, Modeling and Applications*, pages 1–30. Springer, 2013.
- [29] Michael J Lawrence, Robert H Edmundson, and Marcus J O’Connor. An examination of the accuracy of judgmental extrapolation of time series. *International Journal of Forecasting*, 1(1):25–35, 1985.

-
- [30] Michael Lawrence, Paul Goodwin, Marcus O'Connor, and Dilek Önkal. Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, 22(3):493–518, 2006.
- [31] Gene Rowe and George Wright. Expert opinions in forecasting: the role of the delphi technique. In *Principles of forecasting*, pages 125–144. Springer, 2001.
- [32] Mary Kynn. Designing elicitor: Software to graphically elicit expert priors for logistic regression models in ecology. Available from www.winbugs-development.org.uk, 2006.
- [33] Spyros Makridakis and Michele Hibon. The m3-competition: results, conclusions and implications. *International journal of forecasting*, 16(4):451–476, 2000.
- [34] Kenneth M Ford and Jack R Adams-Webber. Knowledge acquisition and constructivist epistemology. In *The psychology of expertise*, pages 121–136. Springer, 1992.
- [35] Nancy J Cooke. Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, 41(6):801–849, 1994.
- [36] Gideon Keren. Calibration and probability judgements: Conceptual and methodological issues. *Acta Psychologica*, 77(3):217–273, 1991.
- [37] Norman Dalkey and Olaf Helmer. The use of experts for the estimation of bombing requirements. a project delphi experiment. *The Rand Corporation*, 1951.
- [38] Allen Newell and Herbert A Simon. *GPS, a program that simulates human thought*. Defense Technical Information Center, 1961.
- [39] Michelene TH Chi, Robert Glaser, and Marshall J Farr. *The nature of expertise*. Psychology Press, 2014.
- [40] Brian R Gaines. Knowledge acquisition: Past, present and future. *International Journal of Human-Computer Studies*, 71(2):135–156, 2013.
- [41] Marianne LaFrance. Excavation, capture, collection, and creation: Computer scientists' metaphors for eliciting human expertise. *Metaphor and Symbolic Activity*, 7(3-4):135–156, 1992.
- [42] PH Garthwaite and SA Al-Awadhi. Quantifying opinion about a logistic regression using interactive graphics. *Submitted for publication*, 2003.

-
- [43] Ikujiro Nonaka and Hirotaka Takeuchi. *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford university press, 1995.
- [44] Anthony O'Hagan, Caitlin E Buck, Alireza Daneshkhah, J Richard Eiser, Paul H Garthwaite, David J Jenkinson, Jeremy E Oakley, and Tim Rakow. *Uncertain judgements: eliciting experts' probabilities*. John Wiley & Sons, 2006.
- [45] Mary Kynn. Eliciting expert knowledge for bayesian logistic regression in species habitat modelling. 2005.
- [46] Joseph B Kadane, James M Dickey, Robert L Winkler, Wayne S Smith, and Stephen C Peters. Interactive elicitation of opinion for a normal linear model. *Journal of the American Statistical Association*, 75(372):845–854, 1980.
- [47] Sarah Lichtenstein and Baruch Fischhoff. Training for calibration. *Organizational Behavior and Human Performance*, 26(2):149–171, 1980.
- [48] Jelle Van Lenthe. Eli: An interactive elicitation technique for subjective probability distributions. *Organizational Behavior and Human Decision Processes*, 55(3):379–413, 1993.
- [49] Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases. *science*, 185(4157):1124–1131, 1974.
- [50] Daniel Kahneman. *Rápido e devagar: duas formas de pensar*. Editora Objetiva, 2012.
- [51] Glenn W Brier. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3, 1950.
- [52] Paulo Renato A Firmino and Nielson A Santana. A bayesian approach for calibrating probability judgments. In *XI BRAZILIAN MEETING ON BAYESIAN STATISTICS: EBEB 2012*, volume 1490, pages 135–142. AIP Publishing, 2012.
- [53] Joan E Sieber. Effects of decision importance on ability to generate warranted subjective uncertainty. *Journal of Personality and Social Psychology*, 30(5):688, 1974.
- [54] Stacey M Whitecotton, D Elaine Sanders, and Kathleen B Norris. Improving predictive accuracy with a combination of human intuition and mechanical decision aids. *Organizational Behavior and Human Decision Processes*, 76(3):325–348, 1998.

- [55] Robert T Clemen and Kenneth C Lichtendahl. Debiasing expert overconfidence: A bayesian calibration model. In *Sixth International Conference on Probabilistic Safety Assessment and Management (PSAM6)*. Citeseer, 2002.
- [56] Thomas Gilovich, Dale Griffin, and Daniel Kahneman. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press, 2002.
- [57] Robert L Winkler. Scoring rules and the evaluation of probability assessors. *Journal of the American Statistical Association*, 64(327):1073–1078, 1969.
- [58] Thornton B Roby. Belief states: A preliminary empirical study. Technical report, DTIC Document, 1964.
- [59] Olivier Thas. *Comparing distributions*. Springer, 2010.
- [60] J Eric Bickel. Some comparisons among quadratic, spherical, and logarithmic scoring rules. *Decision Analysis*, 4(2):49–65, 2007.
- [61] Robert F Nau. Should scoring rules be "effective"? *Management Science*, 31(5):527–535, 1985.
- [62] Paulo Renato A Firmino and Ademir B dos Santos Neto. Essays on probability elicitation scoring rules. In *XI BRAZILIAN MEETING ON BAYESIAN STATISTICS: EBEB 2012*, volume 1490, pages 143–152. AIP Publishing, 2012.
- [63] Howard Raiffa. *Decision analysis: introductory lectures on choices under uncertainty*. 1968.
- [64] Franciscus Petrus Antonius Coolen. *Statistical modeling of expert opinions using imprecise probabilities*. Eindhoven University of Technology, 1994.
- [65] Pavlo Blavatsky. Error propagation in the elicitation of utility and probability weighting functions. *Theory and Decision*, 60(2-3):315–334, 2006.
- [66] Lisa Scholten, Andreas Scheidegger, Peter Reichert, and Max Maurer. Combining expert knowledge and local data for improved service life modeling of water supply networks. *Environmental Modelling & Software*, 42:1–16, 2013.
- [67] Theodore J Gordon and Olaf Helmer. Report on a long-range forecasting study, 1964.

-
- [68] Cameron R Peterson, Kurt J Snapper, and Allan H Murphy. Credible interval temperature forecasts. *Bulletin of the American Meteorological Society*, 53(10):966–970, 1972.
- [69] Anthony 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.
- [70] JLA Devilee and AB Knol. Software to support expert elicitation: An exploratory study of existing software packages. *RIVM letter report 630003001*, 2012.
- [71] Allan James, Samantha Low Choy, and Kerrie Mengersen. Elicitor: an expert elicitation tool for regression in ecology. *Environmental Modelling & Software*, 25(1):129–145, 2010.
- [72] Carlo Spaccasassi and Lea Deleris. A web-based tool for expert elicitation in distributed teams. In *Proceedings of the 8th Bayesian Modelling Applications Workshop of UAI*, volume 2011. Citeseer, 2011.
- [73] Mary Kynn. The ‘heuristics and biases’ bias in expert elicitation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(1):239–264, 2008.
- [74] AB MySQL. Mysql 5.0 reference manual (2008).
- [75] Gu Shenshi, Wang Zhiqian, and Chen Jitai. The fractal research and predicating on the times series of sunspot relative number. *Applied Mathematics and Mechanics*, 20(1):84–89, 1999.
- [76] Rob J Hyndman and George Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2014.
- [77] B Farley and W Clark. Simulation of self-organizing systems by digital computer. *Transactions of the IRE Professional Group on Information Theory*, 4(4):76–84, 1954.
- [78] John Geweke and Charles Whiteman. Bayesian forecasting. *Handbook of economic forecasting*, 1:3–80, 2006.
- [79] Tiago AE Ferreira, Germano C Vasconcelos, and Paulo JL Adeodato. A new intelligent system methodology for time series forecasting with artificial neural networks. *Neural Processing Letters*, 28(2):113–129, 2008.

- [80] Manual MatLab. The language of technical computing. *The MathWorks, Inc.*
<http://www.mathworks.com>, 2012.

Apêndice A

Coleção das séries utilizadas no trabalho

N0236									
1674	1900	1813	1849	1751	2165	2519	2300	2035	2173
2214	2442	2389	2683	3229	3150	3276	3249	3241	3659
4891	4521	4454	5135	5089	3666	3876	4179	4704	5303
4684	4612	4049	4275	4432	4326	4299	4918		

Tabela A.1: Observações da série N0236

N0290									
4105	4031	4014	3960.5	4029	4078.5	4052.5	4082	4161.5	4237
4328.5	4384	4448	4572.5	4714.5	5000	5259			

Tabela A.2: Observações da série N0290

N1037									
2958.5	3051.5	3086	3097.5	3148.5	3193.5	3181.5	3196	3290	3348
3436	3542.5	3546.5	3716	3698	3613	3611.5	3640.5	3829	3771
3935	4432	4504	4596.5	4706	4649.5	5095	5126	5129.5	5003.5
4982	5133	5179.5	5023	4797.5	4718.5	4628.5	4915.5	5006.5	4929.5
5020.5	5047	4949	5250						

Tabela A.3: Observações da série N1037

N1268

3719.8	3646.4	3644.6	3713.2	3708.4	3689.6	3652	3590.2	3549.6	3580.6
3599.8	3647	3693.8	3755.6	3832.6	3917.4	4004	4086	4108.8	4179.2
4210.6	4276.6	4361.2	4452	4496.4	4581.6	4694	4749	4790	4837
4915	4929.8	5058	5150	5240	5318	5397.2	5474.6	5500.8	5552
5637.8	5622.8	5633.8	5567.8	5522					

Tabela A.4: Observações da série N1268

N1269

9439.4	9272.4	8917.3	8373.7	8069.6	6499.9	6183.4	5825	5550.7	5426.8
5363.8	5249.8	5281.4	5311	5258.3	5142.8	5074.4	5039	5058.9	5070.7
5119.6	5120.8	5111	5117.2	5084.9	5107.9	5126.2			

Tabela A.5: Observações da série N1269

N1459

3600	4000	5200	2100	4050	2000	3250	2700	750	2750
2500	2500	2550	6600	4700	2500	4000	2900	3200	1750
3000	3600	2200	2550	4400	4350	2900	2400	3000	1850
3500	2500	1200	2350	1950	4000	4150	2250	2100	1800
1850	1750	3950	1850	4800	2200	3250	1100	1750	3550
2800									

Tabela A.6: Observações da série N1459

N1708									
2220	2120	2940	3180	2080	2280	2340	3160	4060	2500
2580	2520	3960	2740	2760	2480	2220	5060	2200	2140
3500	3000	2480	2720	2480	3380	2680	2460	1580	2660
2180	1840	3960	2340	3240	2680	2620	2300	2820	2080
2160	1780	1680	2600	1980	2060	2220	1780	2540	2020
2600	1340	1200	2820	2480	2260	1820	2340	1060	1500
1980	860	1840	1900	1220	1660	1040	1960	2000	2340
1380	880	1920	1560	1820	680	1260	1180	1240	840
2100	1080	1400	1080	1740	1460	1740	840	780	1120
1520	1520	1340	1620	1220	1160	1640	1040	1260	1040
2120	1380	820	1580	1220	1680	1840	1620		

Tabela A.7: Observações da série N1708

N2355									
2053.4	2233.8	2159.2	2201.2	2322.6	2457.4	2566.8	2582.6	2686.8	2755.4
2800	3033.2	3097.4	3049	3046.2	2930.4	2844.6	2959.6	3099.4	3197.2
2830.4	3066.2	3179	3222.4	3324.6	3150.2	3190	3100.8	3197.2	3084.6
3182.6	3263.8	3376.2	3297	3322.6	3201.4	3571.4	3254.6	3087.4	3370.4
3193.6	3314.8	3380.6	3250.4	3322.6	3303.6	3095.8	3691.4	3491	3701.2
3630.2	3424.6	3391.8	3350.4	3171.8	3136	3345.8	3493.8	3490.8	3360.4
3469.6	3507.8	3603.2	3582.6	3661.8	3676	3577.8	3581.6	3727.2	3854.6
3817.2	4241	4285.6	4009.4	3767.8	3974.4	3691	3493.2	3124.2	3876.6
3714.8	3439.8	3801	3813	2656.6	3658.8	3878.8	3597	3840.4	4075.4
3724	3939.8	3626	3756.6	3204.4	2728.8	3121.8	3043.2	2776.8	3308.2
3466.8	3385.4	3696	3623.4	3816	3886.4	3780.6	3320.6		

Tabela A.8: Observações da série N2355

N0236	N0290	N1037	N1268	N1269	N1459	N1708	N2355
Valores Reais da Série							
5092	5419	5226	5584.2	5145.7	3100	840	3820.2
5015	5452	5132	5612.2	5209.8	3250	3640	4073.8
5305	5560	5055	5538	5224.2	3450	2080	
	5716	5094	5662	5236.6	3800	960	
		5171		5240.7	1450	1380	
		5182.5			1850		
					1250		
					2200		
Previsões dos Modelos Cognitivos							
5150	5406.25	4948.5	5462.5	5323.75	3213.5	1911.25	3800
4800	5468.75	5034	5550	5179.5	3160.710938	1300	3912.5
5312.5	5718.75	5318	5287.5	5163.3	3397.0625	2250	
	5593.75	5291	5137.5	5158.17	2852.28125	2087.5	
		5362		5092.813	3007.875	1368.75	
		5362			3397.0625		
					3746.742188		
					3397.0625		
Previsões dos Modelos ARIMA							
5109.9823	5540.2	5307.2	5559.6	5149	2962.9	1313.7	3564.6
5271.268	5843.6	5362.3	5597.1	5172.2	2856.3	139.9	3627
5411.9742	6169.2	5416.5	5634.7	5193.9	2949.3	1460.7	
	6517	5470.3	5672.2	5214.1	2872.1	1451.5	
		5523.9		5233.1	2935.6	1431	
		5577.5			2883.5		
					2926.3		
					2891.2		

Tabela A.9: Previsões dos modelos cognitivos e matemáticos para as séries