



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

AVALIAÇÃO DA HABILIDADE DE MODELOS DE PREVISÃO NUMÉRICA DE TEMPO BASEADA EM DIAGRAMAS ESTATÍSTICOS

Carlos José Ribeiro Júnior

Relatório de Iniciação Científica do
programa PIBIC, orientada pelo Dr.
Carlos Frederico Bastarz.

URL do documento original:

[<http://urlib.net/>](http://urlib.net/)

INPE
São José dos Campos
2018

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6923/6921

Fax: (012) 3945-6919

E-mail: <pubtc@sid.inpe.br>

**COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO
DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):****Presidente:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Membros:

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Dr. Amauri Silva Montes - Coordenação Engenharia e Tecnologia Espaciais (ETE)

Dr. André de Castro Milone - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Dr. Joaquim José Barroso de Castro - Centro de Tecnologias Espaciais (CTE)

Dr. Manoel Alonso Gan - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dra. Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Clayton Martins Pereira - Serviço de Informação e Documentação (SID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Marcelo de Castro Pazos - Serviço de Informação e Documentação (SID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

AVALIAÇÃO DA HABILIDADE DE MODELOS DE PREVISÃO NUMÉRICA DE TEMPO BASEADA EM DIAGRAMAS ESTATÍSTICOS

Carlos José Ribeiro Júnior

Relatório de Iniciação Científica do
programa PIBIC, orientada pelo Dr.
Carlos Frederico Bastarz.

URL do documento original:

[<http://urlib.net/>](http://urlib.net/)

INPE
São José dos Campos
2018

Dados Internacionais de Catalogação na Publicação (CIP)

Sobrenome, Nomes.

Cutter Avaliação da Habilidade de Modelos de Previsão Numérica de
Tempo Baseada em Diagramas Estatísticos / Carlos José Ribeiro
Júnior. – São José dos Campos : INPE, 2018.
xi + 32 p. ; ()

Dissertação (Iniciação Científica) – Instituto Nacional de Pes-
quisas Espaciais, São José dos Campos, 2017.
Orientador : Carlos Frederico Bastarz.

1. Palavra chave. 2. Palavra chave 3. Palavra chave. 4. Palavra
chave. 5. Palavra chave I. Título.

CDU 000.000



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).

**ATENÇÃO! A FOLHA DE
APROVAÇÃO SERÁ IN-
CLUIDA POSTERIORMENTE.**

Iniciação Científica

RESUMO

Com a constante evolução dos modelos de previsão numérica de tempo, torna-se necessária também a melhoria e o aprimoramento na sua avaliação objetiva. Tendo-se em vista esta necessidade, a Divisão de Modelagem e Desenvolvimento do CPTEC veio a criar o Sistema Comunitário de Avaliação de modelos Numéricos de Tempo E Clima (SCANTEC), contendo em sua estrutura as métricas estatísticas básicas para a avaliação desses modelos. Estas métricas incluem o Viés, a Raiz do Erro Quadrático Médio (REQM) e o Coeficiente de Correlação de Anomalias (CCA). O objetivo principal desse trabalho é incrementar a forma de avaliação dos modelos numéricos de previsão de tempo, por meio de aplicação de diagramas estatísticos e ferramentas de classificação a partir das métricas estatísticas básicas do SCANTEC. Após o estudo necessário para o desenvolvimento deste trabalho, com foco nos conceitos de “precisão” e “acurácia”, foi possível desenvolver por meio da linguagem de programação Python, scripts que geram diferentes diagramas estatísticos, como por exemplo, um diagrama polar baseado no diagrama de Taylor que geralmente é utilizado para representar até três métricas em apenas um único gráfico. Além disso, foi elaborado também, através do Scorecard, um diagrama de habilidades que por sua vez considera o cálculo da diferença do REQM dos membros avaliados em relação a uma referência, classificando-os em função da sua “acurácia” utilizando um gradiente de cores e possibilitando uma nova forma complementar de se avaliar as tabelas geradas pelo SCANTEC, a qual se mostrou bastante útil na avaliação de um conjunto de previsões.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Exemplos gráficos sobre uma forma de demonstrar os termos de precisão e acurácia. Fonte: Revista digital “Negócio Digital” < http://www.ndig.com.br/item/2014/05/preciso-no--a-mesma-coisaex-que-exatido >	8
2.2 Os gráficos na figura representam o VIES (cuja tabela foi omitida no texto) o REQM e o CCA em três níveis de pressão distintos (925, 850 e 500 hPa) no período compreendido entre os dias 5 e 15 de agosto de 2014. Esses gráficos são feitos a partir das tabelas (por exemplo, Tabelas 2.1 e 2.2) geradas pelo SCANTEC. Fonte: produção do próprio autor.	13
2.3 Diagrama de Taylor representando uma simulação com cinco modelos distintos. São representados 5 avaliações A, B, C, D e E. O experimento A mostra $CCA = 0,5$ (linha azul), $REQM = 1$ (linha verde), e Desvio Padrão(S) = 1(linha preta). O ideal seria $REQM = 0$, $CCA = 1$ e $S = 1$. Fonte: < http://www.lamma.ufrj.br/sites/spo/atlasul/scripts/taylor/diagrama.php >.	15
2.4 A figura representa uma das formas do Diagrama Radar ou Diagrama de Kiviat, com valores representativos para uma exemplificação. Fonte: produção do próprio autor.	16
2.5 Exemplo de uma série temporal em que estão sendo mostradas as curvas referentes ao REQM (indicado por RMSE na figura de cima) e CCA (na figura de baixo) de um conjunto de previsões. Neste caso, para amenizar a quantidade de informações, estão sendo mostradas apenas a média do conjunto (curva vermelha), o membro controle (curva preta) e os membros com maiores e menores valores da série (em azul e laranja, respectivamente). Fonte: produção do próprio autor.	18
2.6 Exemplo de um Diagrama (ou Gráfico) de Habilidades em que a informação do RMSE de todos os membros está sendo mostrada para todos os tempos de previsão. O Diagrama de Habilidades pode ser utilizado como uma informação complementar às curvas da série temporal da Figura 2.5. Fonte: produção do próprio autor.	19

3.1	Diagramas de 24 a 120 horas dos modelos testados NCEP, EnSRF e EnKF, com base no modelo de referência CTRL. Esse diagrama está representando como se comporta a variação de uma das variáveis, VTMP-500 (Temperatura Virtual em 500 hPa) durante os horários de previsão (24, 48, 72, 96 e 120 horas). Nos eixos estão demonstrados os valores do REQM e no ângulo os valores do CCA. Fonte: produção do próprio autor.	22
3.2	Aplicação do Diagrama Polar na amostra do conjunto de previsões da temperatura do ar (em 850 hPa) no período de 2014110100 a 2015022812, fornecida por Bastarz et al. (2016). No detalhe, destacam-se as diferenças entre os membros. Fonte: produção do próprio autor.	26
3.3	Diagrama de Habilidades da temperatura do ar (em 850 hPa), complementar à Figura 3.2, válido para o dia 2014110100. Fonte: produção do próprio autor.	27
3.4	Diagrama de Habilidades da temperatura do ar (em 850 hPa), complementar à Figura 3.2, válido para o período de 2014110100 a 2015022812. Fonte: produção do próprio autor.	28

LISTA DE TABELAS

	<u>Pág.</u>
2.1 Exemplo da Tabela da REQM gerada pelo SCANTEC.	11
2.2 Exemplo da Tabela da CCA gerada pelo SCANTEC.	12
3.1 Exemplo da Tabela de Precisão dos experimentos avaliados.	23
3.2 Exemplo da Tabela de Acurácia dos experimentos avaliados.	24

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
1.1 Objetivo Geral	2
1.2 Objetivos Específicos	2
2 DADOS E METODOLOGIA	5
2.1 Modelos de Previsão Numérica de Tempo	5
2.2 Previsões por Conjuntos	6
2.3 Avaliação Objetiva de Modelos de PNTC	6
2.3.1 Precisão e Acurácia	7
2.3.2 Métricas Estatísticas Básicas	9
2.3.2.1 Viés	9
2.3.2.2 Raiz do Erro Quadrático Médio (REQM)	9
2.3.2.3 Coeficiente de Correlação de Anomalia (CCA)	10
2.3.3 SCANTEC	11
2.3.4 Diagramas Estatísticos	14
2.3.4.1 Diagrama de Taylor	14
2.3.4.2 Diagrama de Kiviat	16
2.3.4.3 Diagrama de Habilidade	17
3 RESULTADOS	21
4 CONCLUSÕES	29
4.1 Sugestões para Trabalhos Futuros	30
REFERÊNCIAS BIBLIOGRÁFICAS	31

1 INTRODUÇÃO

Um Modelo de Previsão Numérica de Tempo, de forma simplificada, pode ser descrito como um conjunto equações matemáticas regidas pelas leis da física que expressam o comportamento da atmosfera, e por meio deles é possível prever o estado futuro da atmosfera (MOURA, 1996). Geralmente cada centro de previsão tem seu próprio modelo, podendo ser global (circulação geral), regional ou modelo de área limitada. Mas, antes de aprofundar o assunto principal desse trabalho, que é a avaliação objetiva destes modelos, é necessário o entendimento de algumas noções básicas sobre tempo e clima, que é onde os modelos de previsão numérica são aplicados.

De acordo com Sampaio e Dias (2014), podem ser distinguidos dois elementos que se entrelaçam entre si: o tempo e o clima. O termo *tempo* representa as variações das condições meteorológicas instantâneas de um determinado local e em um determinado instante. Já o termo *clima* representa um estado médio do tempo da atmosfera em um ponto qualquer do espaço e a forma como este tempo (meteorológico) evolui ao longo de um determinado período.

Saindo do plano conceitual e indo em direção ao presente, cuja influência da tecnologia é notável, percebe-se que, com o acúmulo de informações sobre a simulação do estado atual e o futuro da atmosfera produzido nos diversos centros operacionais de Previsão Numérica de Tempo e Clima (PNTC), torna-se necessária uma avaliação estatística elaborada com o objetivo de se identificar, a partir da análise de sua habilidade de previsão, quais modelos melhor cumprem com as suas devidas finalidades (SAPUCCI et al., 2011).

Segundo Willmott (1981) a maioria dos modelos são baseados em procedimentos de ajustes de curvas, onde uma porção da variabilidade de uma variável obrigatoriamente está ligada a variações exibidas em uma ou mais variáveis independentes. Diante disso para se determinar, por exemplo, a probabilidade de precipitação, são analisadas as interações entre pelo menos: a pressão atmosférica, a umidade, a temperatura do ar e a velocidade do vento. Com esse conjunto de variáveis e com base na acurácia e precisão de cada modelo, pode-se definir a capacidade de algumas previsões na identificação da possibilidade ou não de precipitação.

Mikhail e Ackermann (1976) apresentam *acurácia* como sendo o grau de proximidade de uma estimativa com seu parâmetro (ou valor verdadeiro), enquanto que *precisão* expressaria o grau de consistência da grandeza medida com sua média. Os autores também acrescentam que acurácia reflete a proximidade de uma grandeza estatística

ao valor do parâmetro para o qual ela foi estimada, enquanto que a precisão está diretamente ligada com a dispersão da distribuição das observações.

Para a avaliação de um modelo são utilizadas métricas baseadas em índices estatísticos como o Viés, a Raiz do Erro Quadrático Médio (REQM) e o Coeficiente de Correlação de Anomalias (CCA), que por sua vez, fornecem uma visão geral da habilidade de previsão desse modelo. Sua avaliação pode ser tanto para variáveis contínuas (o vento, a pressão, a temperatura e a umidade) quanto para variáveis discretas (a precipitação). As observações utilizadas na avaliação dos modelos requerem um tratamento especial, pois se trata de combinar, em uma grade regular, as informações de modelos com uma grade irregular de dados observados. Ou seja, mapeia-se e divide-se o globo terrestre em pontos de grade regulares, onde cada ponto é equidistante um do outro, enquanto que temos as estações meteorológicas que fornecem as observações do tempo, que são distribuídas de forma irregular no tempo e no espaço (SAPUCCI et al., 2011).

Diante do estudo teórico e prático sobre a avaliação de modelos numéricos de previsão de tempo, encontrou-se uma oportunidade para se contribuir com a classificação dos modelos quanto à sua acurácia no contexto da previsão numérica de tempo por conjunto. A previsão por conjuntos é uma metodologia em que, ao invés de se utilizar uma única previsão numérica de tempo para se determinar o estado futuro da atmosfera, utiliza-se um conjunto destas previsões de forma a se verificar os possíveis estados futuros da atmosfera. Nesse sentido, uma verificação da acurácia dos membros do conjunto de previsões com relação a uma referência mostra-se necessária.

1.1 Objetivo Geral

Este trabalho tem como objetivo incrementar a análise dos modelos numéricos de previsão de tempo, permitindo identificar com maior clareza aqueles experimentos que se destacam em relação a uma referência, por meio de aplicação de diagramas estatísticos e ferramentas de classificação. Busca-se também verificar se em um conjunto de previsões, existe uma previsão (ou um subconjunto delas) que mais contribuem para que a sua média seja mais próxima de uma referência.

1.2 Objetivos Específicos

Para essa pesquisa, foi necessário o cumprimento das etapas listada a seguir:

- Investigar métricas básicas de avaliação da habilidade de modelos de previsão numérica de tempo e quais podem ser aplicadas em diagramas estatísti-

cos;

- Investigar diagramas estatísticos mais sofisticados para avaliação de desempenho de modelos de previsão, tais como diagrama de Taylor;
- Implementar os diagramas estatísticos definidos na etapa anterior em linguagem Python para a avaliação dos resultados;
- Elaboração e aplicação de uma ferramenta classificatória para definir a precisão e acurácia de modelos numéricos em relação a uma referência;
- Avaliar o desempenho de algumas amostras das previsões do modelo global do CPTEC utilizando um diagrama polar, ferramentas de classificação e um diagrama de habilidades, desenvolvidos de forma a identificar a acurácia e a precisão entre diferentes experimentos;
- Verificar a contribuição de um ou mais membros de um conjunto de previsões do Sistema de Previsão por Conjunto global do CPTEC na acurácia da sua média com relação a uma referência.

2 DADOS E METODOLOGIA

Para o desenvolvimento desta pesquisa de iniciação científica, foram realizados estudos sobre a previsão numérica de tempo, os modelos utilizados para esta finalidade e as metodologias de avaliação objetiva das previsões numéricas de tempo. Como o tema central deste trabalho trata sobre a avaliação eficiente e objetiva das previsões do modelo, maior enfoque foi dado na acurácia e na precisão dos modelos. Estes conceitos são apresentados e discutidos ao longo do tempo a fim de dar suporte ao desenvolvimento da metodologia proposta.

2.1 Modelos de Previsão Numérica de Tempo

Modelos de previsão numérica de tempo são equações matemáticas regidas pelas leis da física que, por sua vez, expressam o comportamento da atmosfera. A partir deles é possível prever, em um determinado período de tempo e de uma forma aproximada, o estado futuro da atmosférica (VOGT, 2002). Diversos centros de previsões espalhados ao redor do mundo trabalham constantemente para elaborar e melhorar seus modelos. A divisão atual de modelagem do CPTEC vem desenvolvendo seu próprio modelo de previsão numérica global, baseando-se no modelo americano do National Centers for Environmental Prediction (NCEP).

Um modelo tem como característica principal a capacidade para calcular e prognosticar (prever), ou seja, projetar no tempo os processos físicos de formação de nuvens, chuva, ventos, as condições de solos e vegetações, temperatura, umidade e pressão. Com todas estas variáveis e com suas fórmulas matemáticas é possível, de forma aproximada, prever o tempo de um determinado local. Diz-se forma aproximada porque os modelos numéricos não são perfeitos e infalíveis diante dos elementos da análise científica, mas porque apenas apresentam resultados aproximados, estruturados pelas leis que regem a dinâmica e a física da atmosfera. Dentro desse contexto é importante destacar a importância que os modelos de previsão possuem para a determinação da sua própria condição inicial a partir das técnicas de assimilação de dados, as quais também são realizadas no CPTEC.

De acordo com Vogt (2002), o processo de construção da previsão de tempo ocorre da seguinte forma: recebem-se os dados observados; interpola-se os dados em um grade regular (ou ao contrario, pois também é possível interpolar o modelo para o um ou mais pontos das observações), gera-se as condições iniciais, realiza-se o modelo e, finalmente, preparam-se os produtos para o meteorologista analisar e fazer a disseminação das previsões. De acordo com esse processo pode-se concluir

que a previsão de tempo é um processo complexo e que envolve os conhecimentos acumulados de diversas áreas que estão representados nos modelos numéricos. Ao longo dos anos, estes desenvolvimentos abordaram desde a criação dos primeiros modelos (simplificados) até o avanço tecnológico dos computadores com seus altos graus de processamento, o que permitiu os modelos se tornarem cada vez mais confiáveis na sua habilidade de previsão.

Até o momento as previsões citadas nesse trabalho são previsões determinísticas, isto é, previsões geradas a partir de uma única condição inicial. A fim de que fosse possível ter-se uma primeira ideia sobre os resultados dos modelos de previsão numérica de tempo e, como forma de validação da aplicação da metodologia, realizou-se um estudo prático com algumas previsões do modelo de circulação geral da atmosfera (global) do CPTEC. Para este propósito foram utilizados arquivos provenientes de duas simulações diferentes com o modelo global do CPTEC: “3DVar” e “NCEP”. Os experimentos “3DVar” e “NCEP” foram realizadas por [Bastarz \(2017\)](#), em que no primeiro foram utilizadas as análises do ciclo de assimilação de dados do sistema de assimilação de dados global do CPTEC, enquanto que na segunda foram utilizadas as análises do NCEP, a partir das quais se obtiveram as previsões do modelo global do CPTEC.

2.2 Previsões por Conjuntos

Além dos dados de previsões determinísticas, foram também utilizadas para a aplicação da metodologia algumas amostras das previsões oriundas do sistema de previsões por conjunto global do CPTEC ([BASTARZ et al., 2016](#)). Estas previsões constituem um conjunto de 15 membros, sendo 1 previsão chamada de controle (com prefixo “NMC”) e outras 14 previsões obtidas a partir da perturbação da análise do membro controle. Mais detalhes sobre esta metodologia podem ser obtidos em [Mendonça e Bonatti \(2009\)](#).

2.3 Avaliação Objetiva de Modelos de PNTC

Com a criação de diversos modelos com diferentes características e conseqüentemente, diferentes níveis de habilidade de previsão, tornou-se necessário que uma avaliação de tais modelos fosse realizada afim de se identificar quais modelos melhor representam o estado a atmosfera. A avaliação é feita a partir de métricas estatísticas que fornecem uma visão geral da habilidade dos modelos. Nas seções a seguir, são apresentados algumas das métricas estatísticas mais utilizadas nas avaliações dos modelos, bem como alguns diagramas que podem ser utilizados nesta finalidade. Na Seção 2.3.4.3

são mostradas ideias para a utilização de um diagrama estatístico para a verificação da habilidade de previsões numéricas de tempo.

2.3.1 Precisão e Acurácia

Antes de entender os conceitos de precisão e acurácia, precisa-se compreender o que são os erros aleatório e sistemático. Sempre que se tenta medir qualquer grandeza haverá um erro associado, então pode-se entender que o “erro” é a diferença do valor esperado em relação ao valor medido. Com esse conceito podemos dizer que os erros aleatórios são medidas que interagem individualmente, ou seja, fora de padrão, sendo que cada resultado sofre uma interferência diferente (aleatória). Os erros aleatórios são causados por fatores que não podem ser completamente determinados, sendo de natureza desconhecida ou porque mudam constantemente ou porque são probabilísticos. Já os erros sistemáticos, todos os resultados são deslocados de forma semelhante, gerando assim uma tendência. Com isso os resultados variam em torno de um valor diferente do real. Suas causas são geralmente identificáveis (DAMASCENO et al., 2008).

De forma geral, a precisão é a dispersão do conjunto de valores que se obtém a partir das medições repetidas de uma determinada grandeza: quanto menor for a dispersão, maior será a precisão (WALTHER; MOORE, 2005). A precisão pode ser então representada pela Equação 2.1:

$$P = \frac{1}{n} \sum_{i=1}^n (E_i - \bar{E})^2 \quad (2.1)$$

onde:

- n : números de pontos de grade;
- \bar{E} : média das previsões;
- E_i : previsão;
- i : índice do tempo ou ponto de grade.

A acurácia é o termo utilizado para indicar a qualidade de uma grandeza observada ou um determinado valor estimado (MONICO et al., 2009), além de fazer referência à proximidade do valor medido em relação ao valor real ou valor de uma referência.

Através desses conceitos pode-se determinar a precisão dos modelos. Através da análise do erro aleatório os modelos mais precisos tem como característica a menor diferença em relação a uma média de um conjunto de previsões, e determina-se que a sua acurácia está relacionado a ambos os erros (aleatórios e sistemáticos). Com isso, pode-se dizer que o termo acurado, por si só, envolve também a medida de precisão (MONICO et al., 2009).

Na Figura 2.1, são apresentados, de forma aproximada, os conceitos de precisão e acurácia. No primeiro exemplo (da esquerda para a direita), não podemos identificar a tendência como sendo apenas erros aleatórios, sendo portanto, nem preciso, nem acurado. No segundo, pode-se notar uma tendência, mesmo que os pontos estejam longe do alvo (indicado pelo ponto vermelho), eles estão próximos entre si. Com isso, é possível identificar a precisão no exemplo.

Já na terceira imagem da Figura 2.1, nota-se que os acertos estão ao redor do centro, isso mostra que ao analisar a média da soma dos valores em relação ao alvo, seu resultado estará próximo ao centro, com isso é possível dizer que esse modelo é acurado, mas não é preciso. Na última imagem é possível concluir que além dos pontos estarem próximos entre si, eles estão próximos do alvo, isso mostra que esse modelo é preciso e acurado.

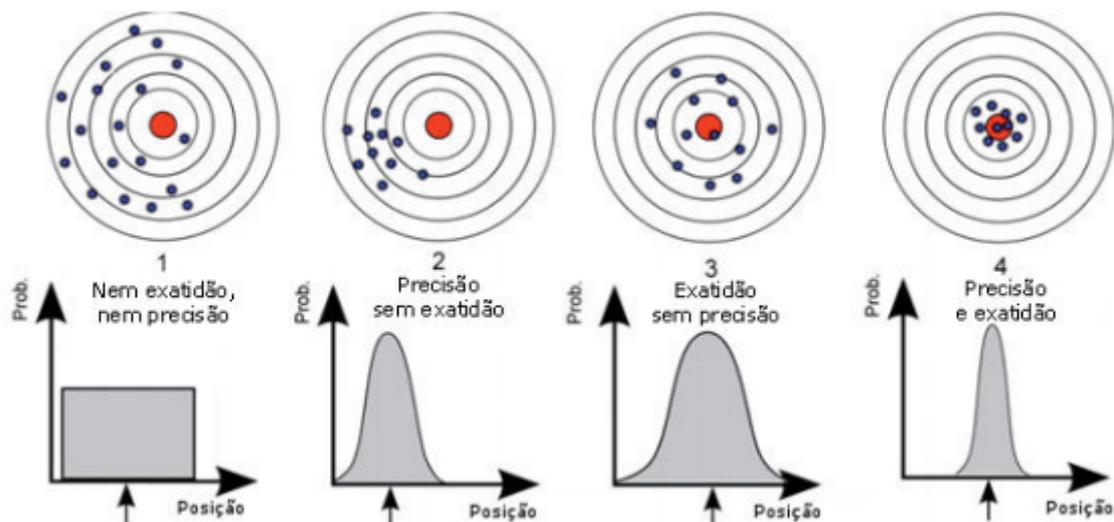


Figura 2.1 - Exemplos gráficos sobre uma forma de demonstrar os termos de precisão e acurácia. Fonte: Revista digital “Negócio Digital” <<http://www.ndig.com.br/item/2014/05/preciso-no-a-mesma-coisaex-que-exatido>>

Com uma avaliação sobre as definições de precisão e acurácia, para aplicação nesta pesquisa, pode-se entender que, a análise da precisão é voltada para a correção do Viés (que por sua vez pode contribuir para a correção do mesmo nos modelos de previsão) e que o estudo da acurácia pode contribuir para identificar a habilidade do modelo nas previsões. Os resultados gerados a partir destes estudos podem contribuir para a melhoria e uniformização da avaliação dos modelos no CPTEC.

2.3.2 Métricas Estatísticas Básicas

Métricas Estatísticas são frequentemente utilizadas para avaliação de variáveis contínuas. A avaliação de variáveis discretas também pode ser feita utilizando-se as métricas básicas, porém devido às suas características, prefere-se a utilização de outras metodologias que não fizeram parte do escopo deste estudo. Essas métricas básicas são utilizadas no CPTEC para a avaliação a partir do software SCANTEC, apresentado na Seção 2.3.3.

2.3.2.1 Viés

O viés, em síntese, é a comparação da previsão média com a observação. O viés trata do desvio médio em relação ao valor verdadeiro ou valor de referência. Podendo ser maior, menor ou igual ao valor referenciado. Se o viés for próximo do valor de referência, mais precisa será a capacidade do modelo de previsão. O viés é expresso pela Equação 2.2:

$$VIES = \frac{1}{n} \sum_{i=1}^n (E_i - A_i) \quad (2.2)$$

onde:

- n : números de pontos de grade;
- A_i : valor observado;
- E_i : valor previsto;
- i : índice do tempo ou ponto de grade.

2.3.2.2 Raiz do Erro Quadrático Médio (REQM)

A Raiz do Erro Quadrático Médio (REQM) representa uma medida da diferença média entre os valores previstos e observados. Devido ao termo quadrático na equação,

o REQM é mais sensível a grandes erros na previsão e é definido como a média do quadrado da diferença entre o valor estimado e o valor de referência. Sua expressão é dada pela seguinte Equação 2.3:

$$REQM = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_i - A_i)^2} \quad (2.3)$$

onde:

- n : números de pontos de grade;
- A_i : valor observado;
- E_i : valor previsto;
- i : índice do tempo ou ponto de grade.

2.3.2.3 Coeficiente de Correlação de Anomalia (CCA)

O Coeficiente de Correlação de Anomalia (CCA) pode ser descrito como uma medida adimensional (sem grandeza física ou unidade) e geralmente positiva, sendo frequentemente utilizada para verificar a correlação, espacial ou temporal, entre os campos meteorológicos previstos, os campos de análise e o intervalo da extensão das previsões válidas dos modelos. A anomalia mede, portanto, o “desvio” do valor previsto com respeito à sua medida climática ou histórica e o desvio do valor analisado com respeito à sua medida climática. O CCA pode ser calculado através da Equação 2.4:

$$CCA = \frac{\sum_{i=1}^n [(P_i - C_i) - (\overline{P_i - C_i})][(A_i - C_i) - (\overline{A_i - C_i})]}{\sqrt{\sum_{i=1}^n [(P_i - C_i) - (\overline{P_i - C_i})]^2 \sum_{i=1}^n [(A_i - C_i) - (\overline{A_i - C_i})]^2}} * 100 \quad (2.4)$$

onde:

- P : previsões;
- A : análises;
- C : climatologia;

- i : índice do tempo ou ponto de grade.

2.3.3 SCANTEC

O Sistema Comunitário de Avaliação de modelos Numéricos de Tempo E Clima (SCANTEC, Sapucci et al. (2011)), foi criado com o objetivo de normalizar e facilitar as avaliações das previsões numéricas produzidas pelo CPTEC. O sistema reúne diversas métricas estatística básicas (como o Viés, a Raiz do Erro Quadrático Médio e o Coeficiente de Correlação de Anomalias, apresentados anteriormente) mais utilizadas na verificação da habilidade das previsões dos modelos.

O sistema é baseado em ferramentas livres e pode ser executado em diversas plataformas. Com sua estrutura escrita em fortran-90, o SCANTEC é capaz de realizar os cálculos estatísticos e escrever os resultados contendo as estatísticas básicas das variáveis, níveis e tempos de previsão selecionados pelo usuário.

Após sua execução o SCANTEC gera-se como resultado campos espaciais (os quais não foram considerados no presente estudo) e tabelas com as métricas estatísticas básicas. Exemplos dessas tabelas estão representadas nas Tabelas 2.1 e 2.2, onde são representadas como atributos as variáveis contínuas em suas colunas e os horários de previsões em suas linhas.

A Tabela 2.1 representa os valores do REQM de algumas variáveis contínuas gerada pelo SCANTEC, sendo também uma das tabelas utilizadas para fazer a avaliação que gerou os resultados plotados nos gráficos da Figura 3.1. No contexto da avaliação de modelos de previsão numérica de tempo, os valores de REQM tendem a crescer com o tempo (coluna “Previsão”), indicando que os erros do modelo são maiores à medida em que se distanciam da condição inicial (tempo “000”).

Tabela 2.1 - Exemplo da Tabela da REQM gerada pelo SCANTEC.

Previsão	VTMP-925	VTMP-850	VTMP-500	TEMP-850	TEMP-500	TEMP-250	UMES-925	UMES-850	UMES-500	UVEL-850	UVEL-500	UVEL-250
000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
006	1.031	0.897	0.521	0.888	0.528	0.535	0.783	0.828	0.370	1.491	1.375	1.835
012	1.302	1.221	0.740	1.207	0.747	0.743	0.985	1.022	0.465	1.984	1.964	2.684
018	1.562	1.508	0.921	1.484	0.928	0.935	1.121	1.183	0.529	2.382	2.493	3.309
024	1.716	1.719	1.084	1.687	1.089	1.111	1.213	1.280	0.573	2.693	2.924	3.833
030	1.921	1.941	1.245	1.900	1.249	1.278	1.305	1.393	0.613	2.970	3.308	4.312
036	2.043	2.111	1.400	2.064	1.402	1.433	1.376	1.472	0.642	3.213	3.650	4.767
042	2.186	2.275	1.551	2.220	1.552	1.579	1.432	1.543	0.669	3.440	3.973	5.222
048	2.296	2.424	1.702	2.363	1.702	1.725	1.479	1.593	0.692	3.649	4.277	5.674
054	2.432	2.575	1.859	2.509	1.856	1.864	1.524	1.646	0.716	3.853	4.576	6.114
060	2.517	2.699	2.011	2.628	2.007	2.002	1.561	1.687	0.736	4.052	4.866	6.530
066	2.626	2.826	2.164	2.749	2.158	2.132	1.589	1.723	0.753	4.249	5.159	6.931
072	2.722	2.952	2.314	2.871	2.306	2.260	1.613	1.752	0.769	4.442	5.445	7.321
078	2.838	3.082	2.463	2.996	2.454	2.380	1.637	1.777	0.782	4.622	5.720	7.681
084	2.920	3.198	2.604	3.107	2.593	2.491	1.659	1.799	0.795	4.792	5.977	8.025
090	3.021	3.314	2.739	3.219	2.727	2.593	1.680	1.821	0.808	4.943	6.221	8.344
096	3.112	3.428	2.868	3.330	2.855	2.692	1.697	1.838	0.819	5.076	6.441	8.636
102	3.216	3.541	2.990	3.440	2.976	2.785	1.713	1.857	0.829	5.200	6.645	8.915
108	3.289	3.638	3.107	3.533	3.091	2.872	1.729	1.874	0.837	5.319	6.825	9.176
114	3.376	3.735	3.222	3.627	3.205	2.953	1.743	1.890	0.846	5.431	7.000	9.428
120	3.456	3.831	3.328	3.719	3.310	3.028	1.754	1.904	0.853	5.537	7.164	9.675

A Tabela 2.2 representa os valores do CCA (Coeficiente de Correlação de Anomalias) de algumas variáveis contínuas gerada pelo SCANTEC que, conforme dito anteriormente, tem a mesma importância na confecção dos diagramas polares da Figura 3.1.

Tabela 2.2 - Exemplo da Tabela da CCA gerada pelo SCANTEC.

Previsão	VTMP-925	VTMP-850	VTMP-500	TEMP-850	TEMP-500	TEMP-250	UMES-925	UMES-850	UMES-500	UVEL-850	UVEL-500	UVEL-250
000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
006	0.970	0.978	0.992	0.977	0.992	0.996	0.987	0.978	0.955	0.970	0.990	0.992
012	0.955	0.963	0.983	0.960	0.983	0.991	0.980	0.966	0.928	0.947	0.980	0.983
018	0.938	0.947	0.975	0.943	0.974	0.986	0.974	0.955	0.905	0.924	0.967	0.973
024	0.928	0.934	0.965	0.930	0.965	0.980	0.970	0.947	0.887	0.902	0.954	0.964
030	0.912	0.918	0.955	0.914	0.954	0.974	0.965	0.937	0.870	0.881	0.941	0.954
036	0.901	0.906	0.944	0.901	0.943	0.967	0.961	0.929	0.856	0.860	0.928	0.944
042	0.889	0.893	0.931	0.888	0.930	0.960	0.958	0.922	0.842	0.839	0.915	0.932
048	0.880	0.880	0.918	0.875	0.917	0.952	0.955	0.917	0.829	0.819	0.901	0.920
054	0.865	0.866	0.903	0.860	0.902	0.944	0.952	0.911	0.816	0.797	0.886	0.906
060	0.855	0.853	0.887	0.847	0.886	0.936	0.950	0.906	0.803	0.775	0.870	0.893
066	0.842	0.839	0.870	0.833	0.869	0.927	0.948	0.902	0.793	0.752	0.854	0.878
072	0.832	0.825	0.852	0.819	0.851	0.918	0.946	0.898	0.782	0.729	0.836	0.864
078	0.816	0.808	0.833	0.801	0.832	0.909	0.945	0.895	0.773	0.706	0.819	0.849
084	0.803	0.791	0.814	0.785	0.813	0.900	0.943	0.893	0.764	0.683	0.801	0.834
090	0.788	0.775	0.795	0.768	0.794	0.892	0.942	0.890	0.755	0.662	0.784	0.820
096	0.777	0.758	0.776	0.751	0.775	0.884	0.941	0.888	0.747	0.642	0.767	0.806
102	0.760	0.740	0.757	0.733	0.756	0.876	0.940	0.885	0.739	0.623	0.751	0.792
108	0.747	0.723	0.739	0.715	0.738	0.869	0.938	0.883	0.733	0.604	0.737	0.779
114	0.733	0.707	0.721	0.698	0.720	0.862	0.937	0.881	0.726	0.586	0.722	0.766
120	0.722	0.691	0.704	0.682	0.703	0.855	0.937	0.879	0.720	0.569	0.708	0.752

Os dados avaliados pelo SCANTEC podem ser representados em três gráficos diferentes, com suas respectivas métricas mais utilizadas para analisar os modelos de PNTC, conforme mostrado na Figura 2.2:

SCANTEC - Estatísticas Básicas

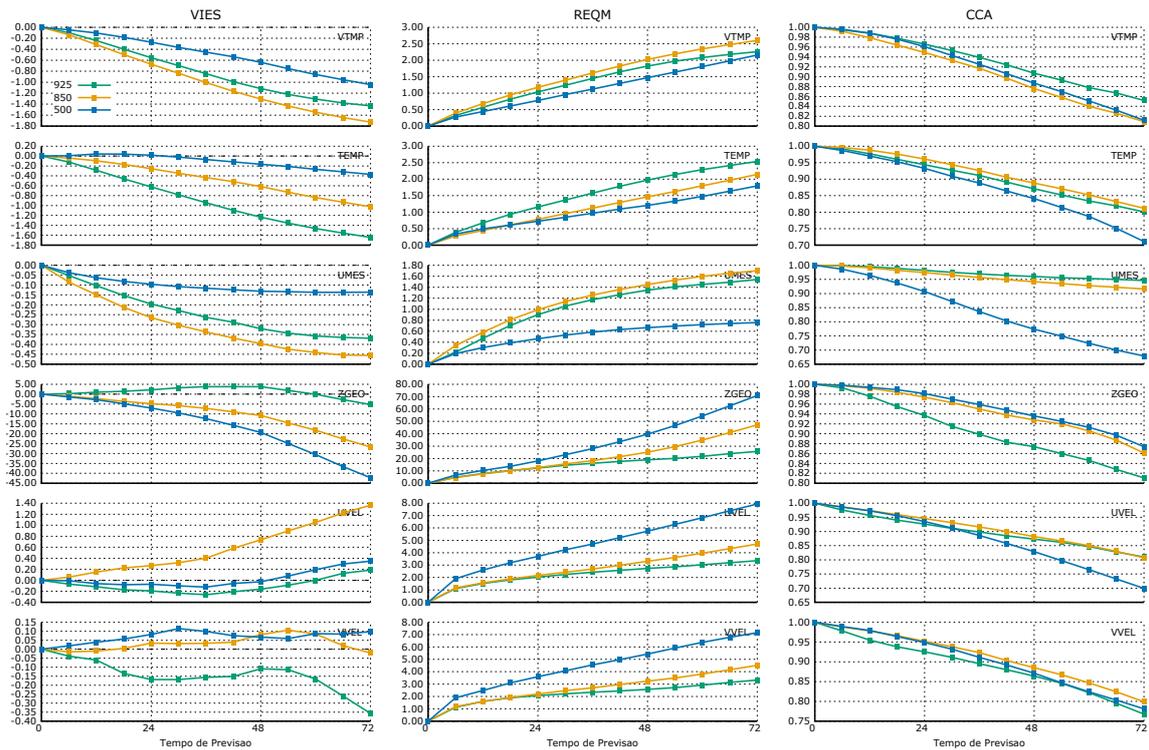


Figura 2.2 - Os gráficos na figura representam o VIES (cuja tabela foi omitida no texto) o REQM e o CCA em três níveis de pressão distintos (925, 850 e 500 hPa) no período compreendido entre os dias 5 e 15 de agosto de 2014. Esses gráficos são feitos a partir das tabelas (por exemplo, Tabelas 2.1 e 2.2) geradas pelo SCANTEC. Fonte: produção do próprio autor.

As análises dos presentes gráficos tem suas limitações, uma vez que essa análise deve ser feita com os três gráficos simultaneamente a fim de se inferir sobre as características de erros aleatórios (Viés), sistemáticos (REQM) e a correlação das análises e previsões com uma climatologia (CCA). Com dois ou mais experimentos, a dificuldade aumenta, pois pode-se querer ou avaliar todos os três níveis indicados em um mesmo gráfico, ou fazer um gráfico para cada um dos níveis considerados. Como este tipo de avaliação requer o cuidado de se distinguir as variações e tendências que cada curva mostra, considerando-se um conjunto de previsões, esta pode não ser a forma mais adequada de se realizar uma avaliação objetiva. Por este motivo, decidiu-se propor uma metodologia em que podem ser classificadas as previsões quanto à sua acurácia e precisão (utilizando uma ferramenta classificatória) e diagramas estatísticos que podem mostrar a variação destas características ao longo do tempo. Nas seções a seguir, serão apresentadas as metodologias que darão suporte ao desenvolvimento do trabalho.

2.3.4 Diagramas Estatísticos

Nesta seção são apresentados dois diagramas mais comumente utilizados para a avaliação de diferentes características de um determinado modelo. O primeiro deles, na Seção 2.3.4.1, pode ser empregado para resumir a correlação, o erro aleatório e sistemático de diversos experimentos numéricos. O segundo, na Seção 2.3.4.2, pode ser empregado para ordenar, em termos da magnitude das grandezas, quais parâmetros são mais evidentes ou relevantes em uma comparação. Outros dois diagramas estatísticos: o primeiro na Seção 2.3.4.1 e o segundo na Seção 2.3.4.3, são também apresentados e representam o desenvolvimento da metodologia do presente trabalho e dão suporte ao objetivo de se estudar as características de precisão e acurácia dos experimentos utilizados.

2.3.4.1 Diagrama de Taylor

Esse diagrama foi desenvolvido por Taylor et al. (2000) para visualizar informações do padrão de similaridade entre modelo e observações, além de caracterizar a relação estatística entre esses dois componentes. Cada ponto no espaço bidimensional do Diagrama de Taylor representa três estatísticas diferentes ao mesmo tempo: o REQM, o CCA e o S. O Diagrama de Taylor exibe estatísticas úteis para avaliar a variação de uma variável simulada por um modelo em relação à variável observada. Como no exemplo a seguir:

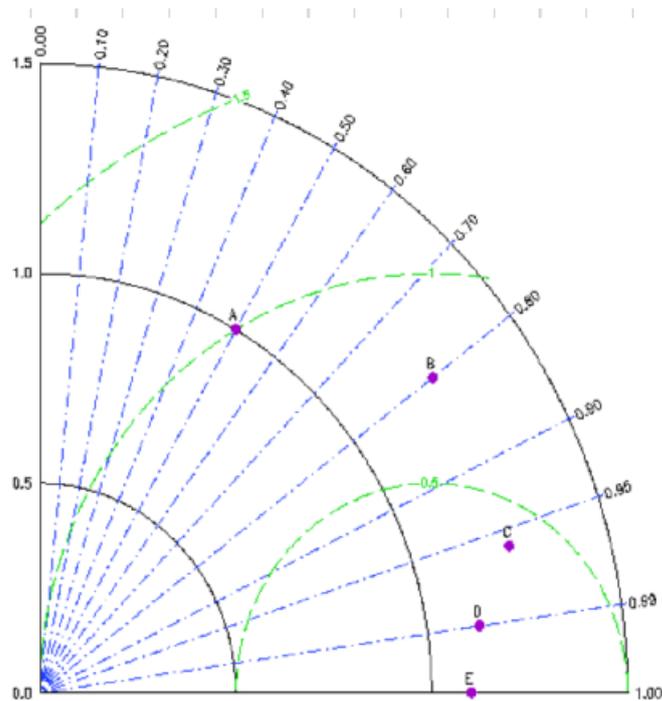


Figura 2.3 - Diagrama de Taylor representando uma simulação com cinco modelos distintos. São representados 5 avaliações A, B, C, D e E. O experimento A mostra CCA = 0,5(linha azul), REQM = 1(linha verde), e Desvio Padrão(S) = 1(linha preta). O ideal seria REQM = 0, CCA = 1 e S = 1. Fonte: <<http://www.lamma.ufrj.br/sites/spo/atlasul/scripts/taylor/diagrama.php>>.

O Diagrama de Taylor é um diagrama polar no qual podem ser representados, nos eixos x e y , os valores da Raiz do Erro Quadrático Médio. No ângulo de inclinação pode ser representado a Correlação de Anomalia, enquanto que o Desvio Padrão pode ser representado pelas linhas pretas, tendo seu ponto de origem o zero e sendo referenciado por um ponto pré determinado (ponto de referência).Essas estatísticas estão relacionadas pela Equação 2.5:

$$E'^2 = \sigma_f^2 + \sigma_r^2 - 2\sigma_f\sigma_r R \quad (2.5)$$

onde:

- R é o CCA entre a previsão e a observação;
- E' é o REQM;
- σ_f e σ_r são os Desvios Padrão da previsão e da observação;

2.3.4.2 Diagrama de Kiviat

Criado na década de 70 por Phil Kiviat, o diagrama radar ou diagrama de teia pode propiciar, em suas diversas características, a representação gráfica que é realizada a partir da comparação dos valores agregados de várias séries de dados. É importante ressaltar que a estrutura desse diagrama permite colocar lado a lado as diferentes series de dados (HACKSTADT; MALONY, 1995). Com isso, pode-se dizer que há possibilidade de empregar o diagrama radar a esse projeto, uma vez que sua representação das variáveis pelo seus múltiplos modelos é realizada de uma forma clara e objetiva. Abaixo um exemplo de um diagrama radar:

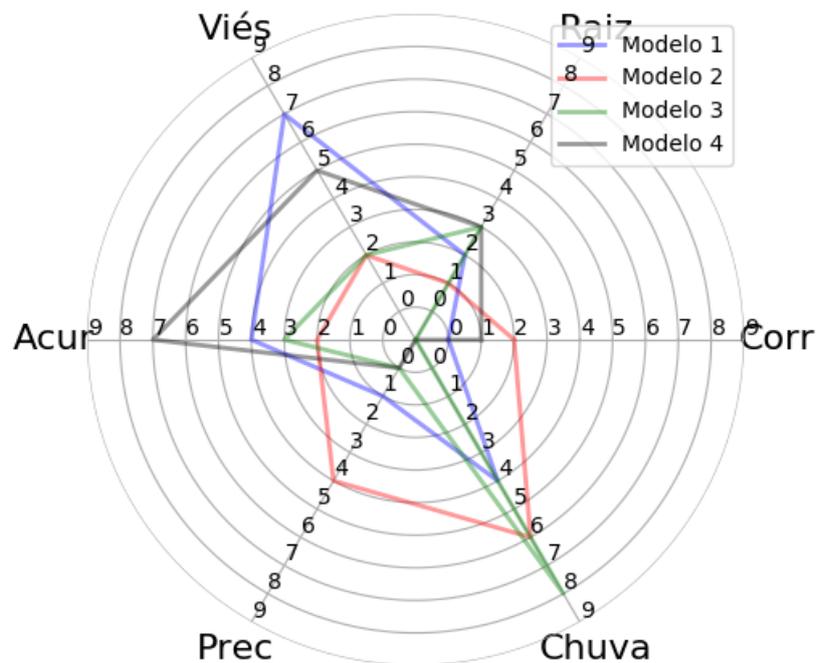


Figura 2.4 - A figura representa uma das formas do Diagrama Radar ou Diagrama de Kiviat, com valores representativos para uma exemplificação. Fonte: produção do próprio autor.

Outros diagramas foram desenvolvidos ao longo dos anos com o propósito de se mostrar diferentes características da habilidade dos modelos de previsão numérica de tempo. No escopo do presente trabalho, buscou-se desenvolver metodologias simples baseadas nos diagramas de Taylor de forma que fosse possível amostrar algumas das características de alguns modelos e previsões produzidos pelo CPTEC. As seções a seguir mostram as metodologias desenvolvidas para isso.

2.3.4.3 Diagrama de Habilidade

Como foi mostrado nas Seções 2.3.2 e 2.3.3, existe alguma dificuldade em se avaliar e comparar diversos experimentos utilizando-se os gráficos de linha padrão a partir das tabelas estatísticas do SCANTEC. Nas Seções 2.3.4.1 e 2.3.4.2, foram apresentadas formas diferentes que buscam sintetizar em um só gráfico as diferentes características dos modelos em avaliação, porém elas não podem mostrar a variação destas informações ao longo do tempo. A partir disso, propõem-se uma forma alternativa que busca sintetizar e complementar a avaliação objetiva de forma a incluir a informação da variação temporal utilizando para isto um tipo de diagrama estatístico, o qual chamamos de “Diagrama de Habilidades” ou “Gráfico de Habilidades”. Este diagrama mostra as informações das estatísticas básicas do SCANTEC de forma visual em que pode-se facilmente identificar padrões temporais dos diversos membros de um conjunto de previsões ou de diferentes experimentos. Para exemplificar, utilizando a mesma amostra de dados contendo o conjunto de previsões e sua referência, na Figura 2.5 demonstra-se um gráfico de uma série temporal representando a variação da REQM e da CCA ao longo do tempo. De forma complementar, a Figura 2.6 apresenta as informações da REQM, mas com a vantagem de não possuir várias linhas se cruzando em diversos pontos. São duas formas aplicadas para analisar a mesma série de dados, uma complementando a outra.

AVALIAÇÃO SÉRIE TEMPORAL

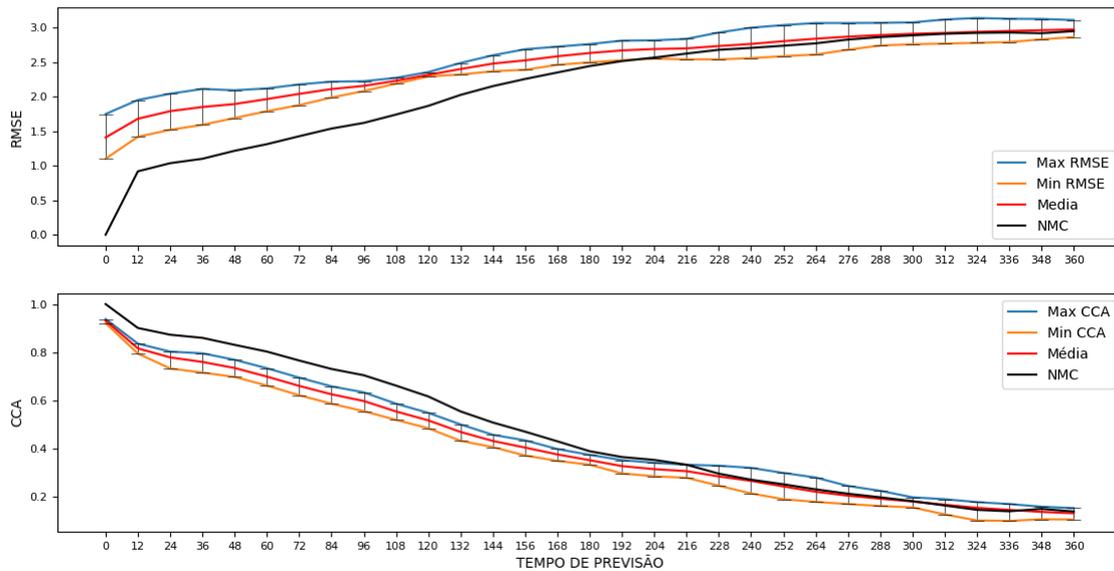


Figura 2.5 - Exemplo de uma série temporal em que estão sendo mostradas as curvas referentes ao REQM (indicado por RMSE na figura de cima) e CCA (na figura de baixo) de um conjunto de previsões. Neste caso, para amenizar a quantidade de informações, estão sendo mostradas apenas a média do conjunto (curva vermelha), o membro controle (curva preta) e os membros com maiores e menores valores da série (em azul e laranja, respectivamente). Fonte: produção do próprio autor.

Através da figura 2.5 observa-se que os membros representados dentro de um intervalo entre o máximo e o mínimo já apresentam um erro inicial, isso é dado pois ao calcular as estatísticas básicas dos membros através do SCANTEC os membros são analisado em relação a análise no modelo controle listado com “NMC”, tendo-se uma penalidade inicial nesse conjunto de previsões, logo que a simulação não foi realizado com membros em relação a sua própria análise. Também nota-se que a partir da metade do tempo de previsão, ou seja, 180 horas a uma uniformidade no geral, pois nesse período de tempo as condições iniciais dos modelos já não tem ação isso é dado pelas limitação e características dos modelos. Podendo ser notado claramente na figura 2.6, onde o valor inicial do controle é a menor medida desse gráfico de habilidade.

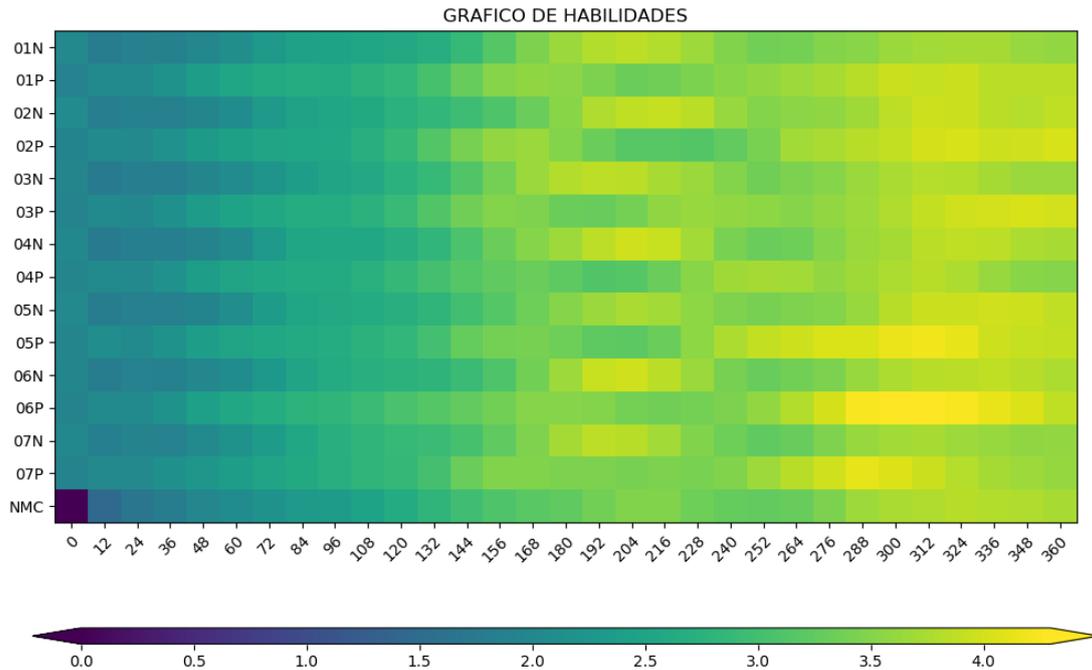


Figura 2.6 - Exemplo de um Diagrama (ou Gráfico) de Habilidades em que a informação do RMSE de todos os membros está sendo mostrada para todos os tempos de previsão. O Diagrama de Habilidades pode ser utilizado como uma informação complementar às curvas da série temporal da Figura 2.5. Fonte: produção do próprio autor.

Na Figura 2.6, aplica-se, o diagrama de habilidade para apresentar os mesmos conjuntos de dados, que foi necessário para gerar a figura 2.5, nota-se uma colaboração do diagrama de habilidade e sua característica de cores para a análise da Figura 2.5, onde pode-se, notar uma diferença, ou seja, um erro inicial como também demonstra a Figura 2.5. Essa característica, ficou clara a partir da análise desse diagrama, nota-se também uma semelhança quando a análise feita a partir das 180 horas de previsão, quando a condição inicial já não interfere nos resultados dos modelos de previsão como apresentados nesses diagramas. Com isso, pode-se dizer, que houve uma contribuição para a avaliação dos modelos.

3 RESULTADOS

Após todo levantamento teórico sobre meteorologia, softwares e ferramentas utilizadas no CPTEC, pôde-se desenvolver um script que possibilita a leitura das tabelas geradas pelo SCANTEC com os valores de VIÉS, REQM e CCA. Baseando-se no Diagrama de Taylor (vide Seção 2.3.4.1), desenvolveu-se um tipo de diagrama polar que tem por objetivo apresentar e elencar as métricas REQM e CCA de diversos experimentos. Esta é uma forma objetiva e gráfica de se representar dois atributos em um mesmo gráfico, utilizando diretamente as tabelas do SCANTEC. Aqui não houve a intenção de se substituir o Diagrama de Taylor, mas aplicar as técnicas de programação como um exercício para a análise e teste das metodologias estudadas e propostas.

O script desenvolvido fornece os diagramas polares de cada horário de previsão para cada experimento (CTRL, NCEP, EnSRF e EnKF), mostrando assim os valores de duas métricas em um mesmo diagrama. Com o apoio da linguagem de programação Python gerou-se os diagramas polares da Figura 3.1 com o objetivo de tornar a análise mais resumida e objetiva. Os experimentos avaliados foram obtidos a partir dos resultados de [Bastarz \(2017\)](#). O experimento CTRL representa uma simulação com o modelo global do CPTEC utilizando como condições iniciais as análises do sistema de assimilação de dados do CPTEC (3DVar); o experimento NCEP representa a mesma simulação, mas utilizando como condição inicial as análises do NCEP. Os experimentos EnSRF e EnKF, representam as simulações com o modelo global do centro, mas utilizando dois tipos diferentes de condições iniciais, extraídas de um conjunto de análises distintos.

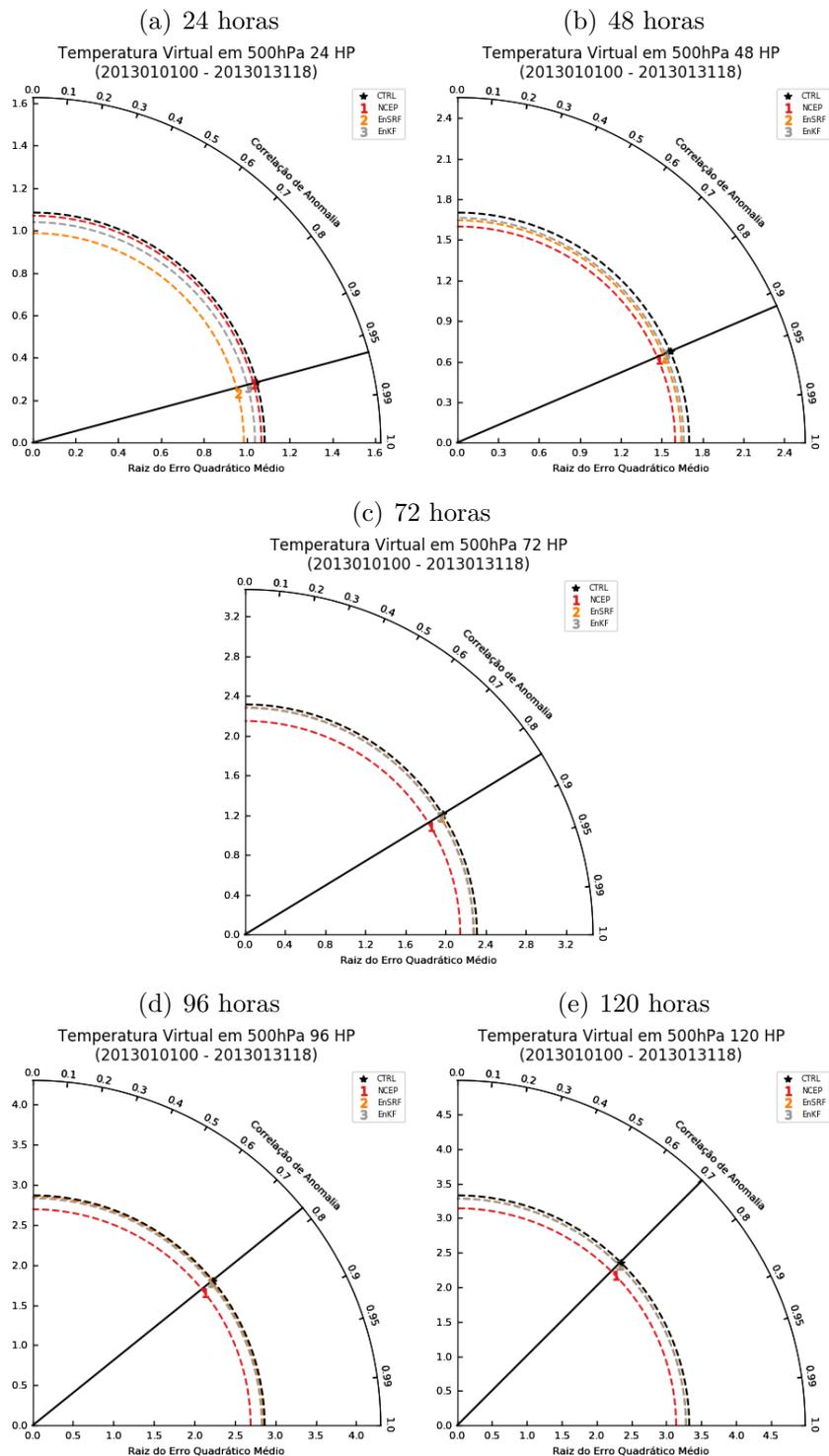


Figura 3.1 - Diagramas de 24 a 120 horas dos modelos testados NCEP, EnSRF e EnKF, com base no modelo de referência CTRL. Esse diagrama está representando como se comporta a variação de uma das variáveis, VTMP-500 (Temperatura Virtual em 500 hPa) durante os horários de previsão (24, 48, 72, 96 e 120 horas). Nos eixos estão demonstrados os valores do REQM e no ângulo os valores do CCA. Fonte: produção do próprio autor.

Com as leituras das tabelas incorporadas à lógica de programação, de início foi possível ter uma ideia de como calcular a precisão e acurácia dos modelos, uma vez que o script é capaz de calcular as diferenças entre os valores dos experimentos em relação a um valor de referência. Portanto, este script além de ter a habilidade de gerar os diagramas polares, serve também como uma ferramenta classificatória quanto à acurácia e precisão dos modelos/experimentos em avaliação.

Na primeira etapa da pesquisa, para se ter uma forma de classificar os modelos, determinou-se que a precisão é quando os valores da REQM de um modelo tem menor variação em relação a uma referência, assim, quanto menor a variação entre os valores, mais preciso é o modelo. Na Tabela 2.1, há uma representação dos resultados do script, onde pode-se ver que às 6 horas o experimento EnKF é o mais preciso e o experimento menos preciso é o NCEP, uma vez que o EnKF apresenta a menor diferença da REQM em relação ao CRTL (modelo considerado como referência).

Tabela 3.1 - Exemplo da Tabela de Precisão dos experimentos avaliados.

Horas	Modelo Mais Preciso	Modelo Menos Preciso
6	EnKF	NCEP
12	NCEP	EnSRF
18	NCEP	EnSRF
24	NCEP	EnSRF
30	EnKF	EnSRF
36	EnKF	EnSRF
42	EnKF	NCEP
48	EnKF	NCEP
54	EnKF	NCEP
60	EnKF	NCEP
66	EnSRF	NCEP
72	EnSRF	NCEP
78	EnSRF	NCEP
84	EnSRF	NCEP
90	EnSRF	NCEP
96	EnSRF	NCEP
102	EnSRF	NCEP
108	EnSRF	NCEP
114	EnSRF	NCEP
120	EnSRF	NCEP

O mesmo critério para calcular a diferença do REQM foi utilizado para determinar a

acurácia, usando, dessa vez, os valores do CCA. Assim, por definição, quanto mais próximos os valores de correlação de um modelo com a sua referência, mais acurado é esse modelo. A Tabela 2.2 apresenta a classificação dos modelos testados em mais e menos acurado para cada horário de previsão. Nota-se que nas 6 horas de previsão o experimento EnKF é o mais acurado, pois apresenta a menor diferença do CCA em relação a CTRL (modelo considerado como referência). Nessa simulação, em 18 horas de previsão, tem-se dois modelos mais acurados, devendo-se ao fato de que as diferenças entre esses modelos em relação à referência são as mesmas.

Tabela 3.2 - Exemplo da Tabela de Acurácia dos experimentos avaliados.

Horas	Modelo Mais Acurado	Modelo Menos Acurado
6	EnKF	NCEP
12	NCEP	EnSRF
18	NCEP/EnKF	EnSRF
24	NCEP	EnSRF
30	NCEP	EnSRF
36	NCEP	EnSRF
42	EnKF	EnSRF
48	EnKF	EnSRF
54	EnKF	NCEP
60	EnKF	NCEP
66	EnKF	NCEP
72	EnKF	NCEP
78	EnKF	NCEP
84	EnKF	NCEP
90	EnKF	NCEP
96	EnKF	NCEP
102	EnKF	NCEP
108	EnKF	NCEP
114	EnKF	NCEP
120	EnKF	NCEP

Analisando as tabelas geradas, pode-se identificar qual é o melhor modelo para cada horário de previsão, de acordo com os critérios estabelecidos. Por exemplo, com 6 horas de previsão o experimento EnKF apresentou ser o modelo mais hábil, pois é o mais preciso e o mais acurado dentre todos os outros experimentos verificados.

A construção e a aplicação de uma ferramenta classificatória da precisão e da acurácia, mostrou-se como um interessante exercício de verificação da habilidade dos modelos.

As mesmas ideias apresentadas até aqui podem ser estendidas para a avaliação de um conjunto de previsões (tal como aquele mencionado na Seção 2.2). Porém, para se incrementar a forma como um conjunto de previsões pode ser avaliado a partir das tabelas de estatísticas básicas do SCANTEC, buscou-se fazer esta avaliação aplicando-se os diagramas polares (tal como mostrado na Figura 3.1), além de uma outra forma interessante que permite melhorar a compreensão sobre a distribuição dos erros desse conjunto de previsões ao longo do tempo.

A aplicação do diagrama polar idealizado (baseado no diagrama de Taylor), tal como demonstrado na Figura 3.2, mas para a amostra do conjunto de previsões, mostrou-se não ser muito eficiente. Uma hipótese pode ser considerada para isto: a de que os membros da amostra do conjunto de previsões estão relativamente próximos entre si e, em alguns casos, apresentando valores de erro muito próximos. Isto pode ser um indicativo de que, pela pequena variação no tempo, tem-se um conjunto de previsões com pequeno espalhamento, isto é, os membros deste conjunto podem ser muito semelhantes entre si. Por outro lado, como destacado no Figura 3.2, pode-se observar dois subconjuntos de previsões, o que é uma característica do método de geração dos membros do conjunto (BASTARZ et al., 2016).

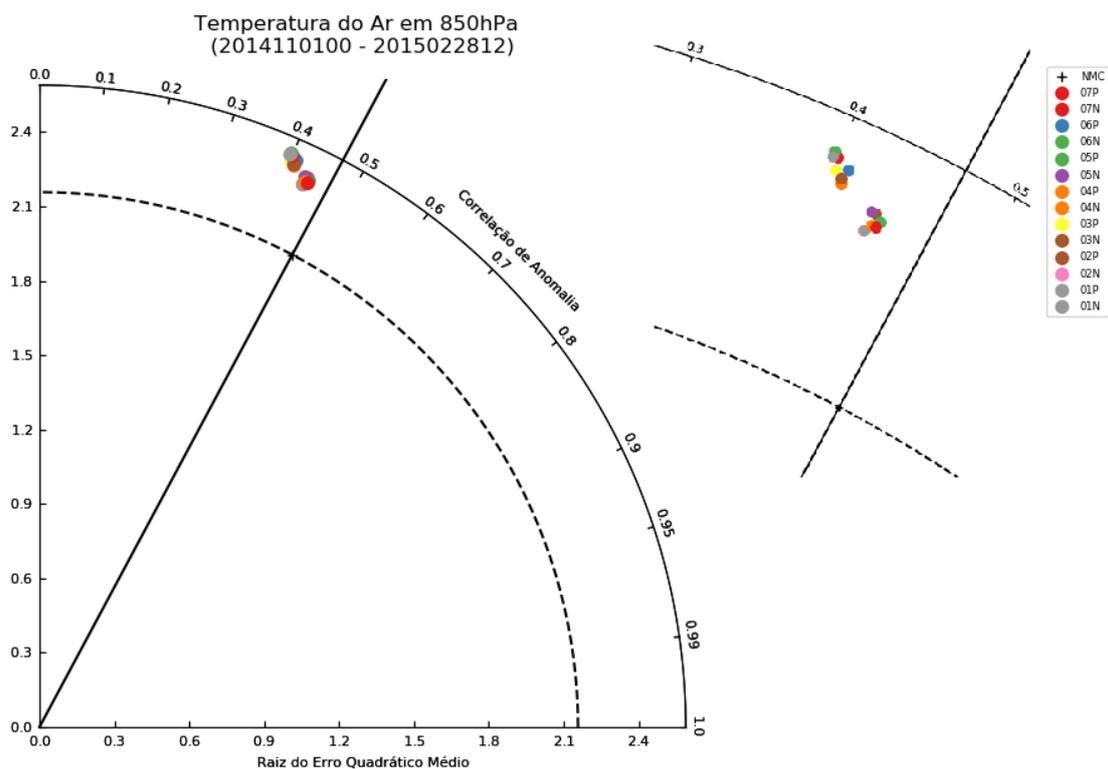


Figura 3.2 - Aplicação do Diagrama Polar na amostra do conjunto de previsões da temperatura do ar (em 850 hPa) no período de 2014110100 a 2015022812, fornecida por Bastarz et al. (2016). No detalhe, destacam-se as diferenças entre os membros. Fonte: produção do próprio autor.

Com o aprofundamento teórico sobre precisão e acurácia, identificou-se que apenas com a análise do REQM é possível classificar os modelos, já que o termo acurado, por si só, como dito anteriormente, envolve a precisão. Desse modo, foi possível criar um diagrama de habilidades, a fim de contribuir para a análise do conjunto de previsões.

Para a classificação dos modelos em relação a sua acurácia, foi utilizado um diagrama baseado no Balanced ScoreCard (KAPLAN; NORTON, 1995). Com base nos estudos da análise do REQM do conjunto de previsões em relação a uma referência, foi elaborado um script que calcula a diferença entre esses campos e plota os diagramas de habilidades representados nas Figuras 3.3 e 3.4. Nessas figuras é possível representar, de uma forma gráfica, a classificação da habilidade desse conjunto de previsões, sendo representado no eixo “y” os membros e no eixo “x” o tempo de previsão (horas).

Na Figura 3.3 nota-se que os membros que possuem as cores mais escuras ao longo do tempo são os membros mais acurados. Além disso, nota-se também que para uma

análise de um único dia os membros de classificação “N” (isto, cujas perturbações são subtraídas do membro controle), mesmo não sendo um regra no geral, formam um conjunto mais acurado. Como demonstrado na Figura 3.3 a seguir:

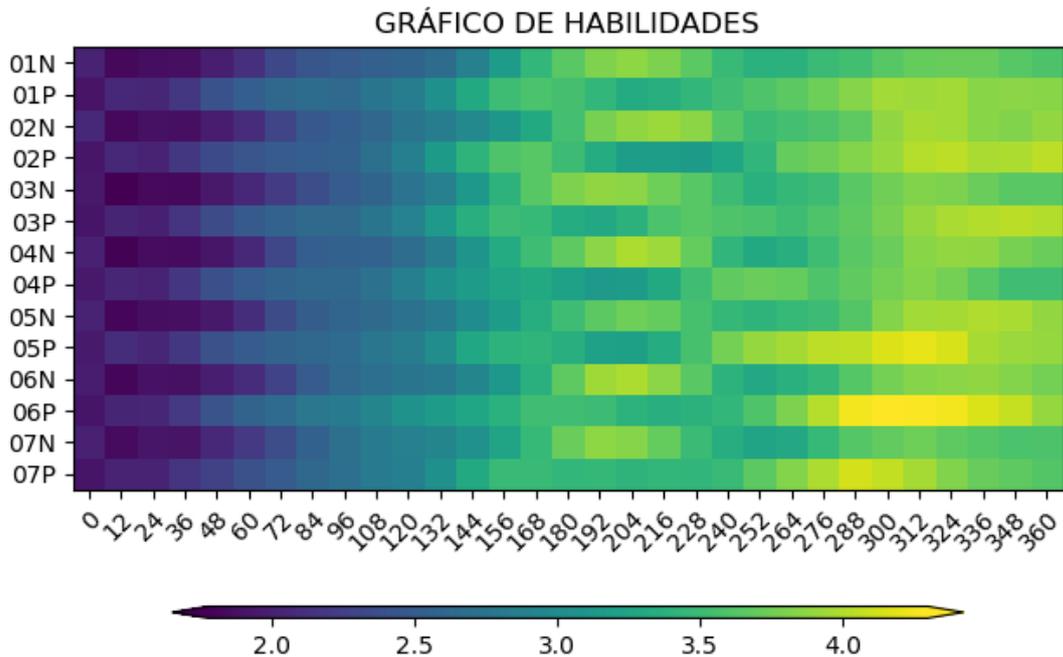


Figura 3.3 - Diagrama de Habilidades da temperatura do ar (em 850 hPa), complementar à Figura 3.2, válido para o dia 2014110100. Fonte: produção do próprio autor.

A Figura 3.4 segue o mesmo padrão da Figura 3.3, sendo diferenciada desta apenas pelo período de realização, sendo de 01/11/2014 até 28/02/2015. Com um período longo de previsão, a Figura 3.4 demonstra um gradiente mais uniforme ao longo do tempo e com poucas variações, diferenciando-se da Figura 3.3. Nota-se também uma semelhança quanto aos membros de classificação “N”, onde mantém-se uma melhor classificação quanto à sua acurácia.

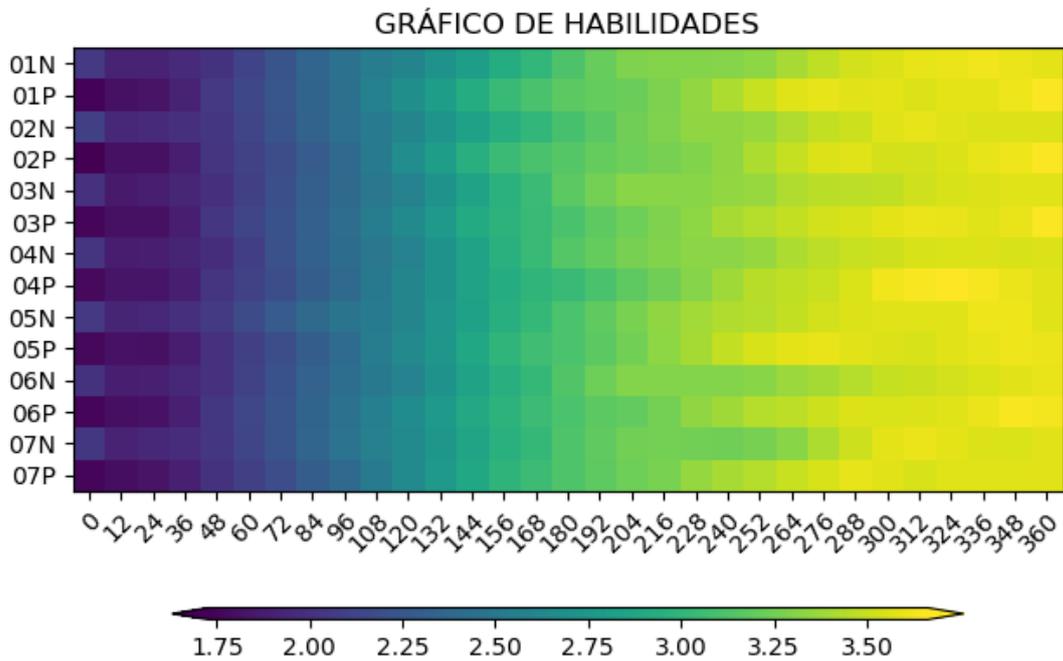


Figura 3.4 - Diagrama de Habilidades da temperatura do ar (em 850 hPa), complementar à Figura 3.2, válido para o período de 2014110100 a 2015022812. Fonte: produção do próprio autor.

A metodologia implementada nesse trabalho pode ser aplicada para avaliação dos modelos, logo que, as características dos diagramas estatísticos listados nesse trabalho, proporciona uma análise simplificada da habilidade dos modelos estudados.

4 CONCLUSÕES

Com o levantamento teórico dos conceitos necessários para a execução desse trabalho, pôde-se adquirir o entendimento das métricas estatísticas básicas de uma forma simplificada: o viés, uma medida do erro aleatório; a REQM, uma medida do erro sistemático e o CCA, uma medida do desvio do valor previsto em relação a sua média climática ou climatológica. Esses conceitos deram base de como um modelo de previsão numérica de tempo pode ser avaliado e com o SCANTEC que, por sua vez, faz os cálculos dessas estatísticas, gerando tabelas e campos espaciais (não explorados neste trabalho), que formaram a fonte de dados utilizada para esse trabalho.

Contudo, após a conceituação e aplicação da linguagem Python, foi desenvolvido um script que possibilitou explorar novas formas para avaliar os experimentos dos modelos utilizados. Uma destas formas de avaliação foi a partir da utilização do diagrama de Taylor que em sua estrutura compreende até três métricas estatísticas diferentes em apenas um gráfico, dando assim uma visão rápida dos modelos analisados. Após a aplicação desse diagrama, percebeu-se que outras formas de se representar a habilidade dos modelos era possível. Utilizando-se a idéia de um Scorecard, foi possível desenvolver um gráfico de habilidades e, com a utilização de um gradiente de cores, foi possível complementar a avaliação de séries temporais, antes realizadas apenas com base em gráficos de linhas.

A partir da aplicação da metodologia construída na primeira etapa da pesquisa e dos estudos sobre a avaliação de modelos de forma determinística, deu-se início a aplicação dessa metodologia na avaliação de uma amostra de um conjunto de previsões. Para analisar a amostra desse conjunto de previsões foi desenvolvido um gráfico com a série temporal dos seus erros, onde estão representadas (por meio de curvas) o modelo de referência, a maior variação do conjunto de previsão (positivo e negativo) em relação a sua média e a própria média.

Juntamente com a aplicação da metodologia, fora aplicado, na avaliação das previsões por conjunto, o diagrama de habilidades de forma complementar, permitindo, assim, a classificação dos membros do conjunto com relação à sua acurácia a partir de uma análise visual SCANTEC. Em suma, a aplicação dessa metodologia contribui para a avaliação dos modelos, possibilitando novas formas de avaliar a habilidade dos modelos por meio dos diagramas estatísticos e das ferramentas classificatórias.

4.1 Sugestões para Trabalhos Futuros

Como sugestão para a melhoria dos desenvolvimentos realizados neste trabalho, sugere-se os seguintes tópicos que podem refinar a metodologia e os gráficos produzidos para a avaliação do modelo de PNTC.

- Metodologia:
 - Validar a aplicação da metodologia para um conjunto de dados maior e mais completo;
 - Aprimorar o critério de classificação quanto à acurácia e precisão;
 - Desenvolver uma metodologia para se calcular a área do gráfico ocupada por um atributo de forma que este valor possa ser utilizado em uma comparação direta entre os diferentes atributos;
 - Evoluir a metodologia desenvolvida para que seja aplicável na avaliação de variáveis discretas;
- Gráficos:
 - Inserir os valores das diferenças nas quadrículas do gráfico de habilidades a fim de facilitar a interpretação das características dos membros;
 - Incorporar a metodologia de avaliação das previsões por conjunto dentro do contexto do SCANTEC (incorporação dos scripts desenvolvidos);
- Script:
 - Criação de uma interface gráfica, onde incorpora todos os métodos para avaliar os modelos de previsão numérica de tempo, com foco em facilitar a utilização desse sistema ao usuário;
 - Criação de uma plataforma web, para disponibilizar essa ferramenta e suas novas formas de avaliar os modelos, contribuindo-se para que outros centros de previsão e pessoas interessadas utilizem dessa metodologia.

REFERÊNCIAS BIBLIOGRÁFICAS

- BASTARZ, C. F. **Assimilação de dados global híbrida por conjunto-variacional no CPTEC**. 275 p. Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2017-07-18 2017. Acesso em: 25 jul. 2018. 6, 21
- BASTARZ, C. F.; SAPUCCI, L. F.; BONATTI, J. P.; GONÇALVES, L. G. G. d. **Sistema de Modelagem por Conjunto (SMC) (Versão Inicial V0.0)**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2016. 98 p. Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m21b/2016/08.17.14.20>>. Acesso em: 25 jul. 2018. viii, 6, 25, 26
- DAMASCENO, J. C.; ANTONIO, J.; OLIVEIRA, S. P. de. **Avaliação de dados de medição—Guia para a expressão de incerteza de medição**. 2008. 7
- HACKSTADT, S. T.; MALONY, A. D. Visualizing parallel programs and performance. **IEEE Computer Graphics and Applications**, IEEE, v. 15, n. 4, p. 12–14, 1995. 16
- KAPLAN, R. S.; NORTON, D. P. Putting the balanced scorecard. **Performance measurement, management, and appraisal sourcebook**, Human Resource Development, p. 66, 1995. 26
- MENDONÇA, A. M.; BONATTI, J. P. Experiments with eof-based perturbation methods and their impact on the cptec/inpe ensemble prediction system. **Monthly Weather Review**, v. 137, n. 4, p. 1438–1459, 2009. Disponível em: <<https://doi.org/10.1175/2008MWR2581.1>>. 6
- MIKHAIL, E.; ACKERMANN, F. **Observations and least squares**. IEP, 1976. (IEP series in civil engineering). ISBN 9780700224814. Disponível em: <<https://books.google.com.br/books?id=Yx3vAAAAMAAJ>>. 1
- MONICO, J. F. G.; POZ, A. P. D.; GALO, M.; SANTOS, M. C. D.; OLIVEIRA, L. C. D. Acurácia e precisão: revendo os conceitos de forma acurada. **Boletim de Ciências Geodésicas**, v. 15, n. 3, 2009. 7, 8
- MOURA, A. D. Von Neumann e a previsão numérica de tempo e clima. **Estudos Avançados**, scielo, v. 10, p. 227 – 236, 04 1996. ISSN 0103-4014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40141996000100021&nrm=iso>. 1
- SAMPAIO, G.; DIAS, P. L. da S. Evolução dos modelos climáticos e de previsão de tempo e clima. **Revista USP**, n. 103, p. 41–54, 2014. 1
- SAPUCCI, L. F.; MATTOS, J. G. F. d.; SANTOS, A. F. d.; GONÇALVES, L. G. G. d.; BASTARZ, C. F.; SCHUBERT, D. V. Potencialidades do scamtec (sistema de avaliação de modelos numérico de previsão de tempo e clima) na avaliação de previsões de precipitação. In: CONGRESSO BRASILEIRO DE

AGROMETEOROLOGIA., 2011, Guarapari-SP Rio de Janeiro RJ. **Anais...** [S.l.]: Sociedade Brasileira de Agrometeorologia, 2011. Setores de Atividade: Pesquisa e desenvolvimento científico. Acesso em: 28 abr. 2017. 1, 2, 11

TAYLOR, K.; DIAGNOSIS, L. L. N. L. P. for C. M.; INTERCOMPARISON. **Summarizing Multiple Aspects of Model Performance in a Single Diagram**. Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, 2000. (PCMDI report). Disponível em: <<https://books.google.com.br/books?id=moBDGwAACAAJ>>. 14

VOGT, C. Modelagem matemática na previsão do tempo e do clima. feb 2002. Disponível em: <<http://www.comciencia.br/dossies-1-72/reportagens/modelagem/mod06.htm>>. 5

WALTHER, B. A.; MOORE, J. L. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. **Ecography**, Wiley Online Library, v. 28, n. 6, p. 815–829, 2005. 7

WILLMOTT, C. J. On the validation of models. **Physical geography**, Taylor & Francis, v. 2, n. 2, p. 184–194, 1981. 1