



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES



aa/bb/cc/dd-RPQ

**AVALIAÇÃO DA FUNCIONALIDADE DE SOFTWARE
DE MODELOS NUMÉRICOS DE ESCALA GLOBAL
PARA O PROJETO MONAN – PARTE I: ANÁLISE
QUALITATIVA, QUANTITATIVA E MULTIVARIADA**

Ariane Frassoni
Julio Pablo Fernandez
Marcelo Barbio Rosa
José Roberto Rozante
Bárbara A. G. P. Yamada

Relatório referente ao Trabalho de
Definição do Núcleo Dinâmico do
MONAN-Atmosférico

URL do documento original:
<http://urlib.net/xx/yy>

INPE
São José dos Campos
2023

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELECTUAL DO INPE - CEPII (PORTARIA Nº 176/2018/SEI-INPE):

Presidente:

Dra. Marley Cavalcante de Lima Moscati - Coordenação-Geral de Ciências da Terra (CGCT)

Membros:

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação (CPG)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e Ciência Espaciais (CGCE)

Dr. Rafael Duarte Coelho dos Santos - Coordenação-Geral de Infraestrutura e Pesquisas Aplicadas (CGIP)

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

EDITORAÇÃO ELETRÔNICA:

Ivone Martins - Divisão de Biblioteca (DIBIB)

André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES



aa/bb/cc/dd-RPQ

**AVALIAÇÃO DA FUNCIONALIDADE DE SOFTWARE
DE MODELOS NUMÉRICOS DE ESCALA GLOBAL
PARA O PROJETO MONAN – PARTE I: ANÁLISE
QUALITATIVA, QUANTITATIVA E MULTIVARIADA**

Ariane Frassoni
Julio Pablo Fernandez
Marcelo Barbio Rosa
José Roberto Rozante
Bárbara A. G. P. Yamada

Relatório referente ao Trabalho de
Definição do Núcleo Dinâmico do
MONAN-Atmosférico

URL do documento original:
<http://urlib.net/xx/yy>

INPE
São José dos Campos
2023



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](#).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](#).

Informar aqui sobre marca registrada (a modificação desta linha deve ser feita no arquivo publicacao.tex).

Informar aqui sobre fontes financiadoras (a modificação desta linha deve ser feita no arquivo publicacao.tex).

AGRADECIMENTOS

Agradecemos ao Dr. José Roberto Garcia por ter contribuído com o presente trabalho por meio da obtenção e organização dos dados de reanálise do ERA5.

RESUMO

O Instituto Nacional de Pesquisas Espaciais (INPE) tem como uma de suas missões desenvolver novas soluções tecnológicas para reduzir os danos associados à ocorrência de eventos climáticos e ambientais extremos. Uma nova iniciativa busca envolver a academia, os setores privado e público, entre eles tomadores de decisão e agências meteorológicas regionais para apoiar a transformação de conhecimento científico e tecnológico em melhores produtos e serviços derivados de modelagem numérica. Como meta, tem-se o desenvolvimento de um modelo comunitário do Sistema Terrestre - o *Model for Ocean-laNd-Atmosphere predictioN* (MONAN). O INPE lidera o desenvolvimento do MONAN, que produzirá previsões operacionais adaptadas para a América do Sul, fornecendo informações climáticas em diferentes escalas espaciais e temporais. Para desenvolver um modelo do Sistema Terrestre de última geração, é necessário o uso de novas técnicas em computação de alto desempenho, um núcleo dinâmico de última geração e a análise de toda a gama de dados ambientais disponíveis. Portanto, será necessário empregar técnicas inovadoras em Inteligência Artificial (IA) e análise de dados, oferecendo oportunidades de pesquisa e inovação ao longo do fluxo de trabalho com o MONAN. O primeiro passo na construção do novo sistema envolve a escolha de um núcleo dinâmico não-hidrostático que incorporará códigos de parametrizações físicas adaptadas às características climáticas sul-americanas, que podem ser derivados de sistemas de modelagem existentes atualmente em operação ou adaptados de outros modelos numéricos. O presente trabalho tem por objetivo descrever o primeiro passo na construção do MONAN, que consiste na avaliação dos principais modelos numéricos de escala global de código aberto

disponíveis gratuitamente e que fazem parte da nova geração de modelos, visando a escolha da sua componente atmosférica. Busca-se melhor entender as principais características dos modelos analisados, seus pontos favoráveis e limitações quanto à simulação das características meteorológicas sul-americanas. Dentre os modelos avaliados quanto à qualidade de software, apenas o SHiELD/FV3 e MPAS se mostraram aptos a fornecer um núcleo dinâmico hidrostático para a composição da componente atmosférica do MONAN. Estas duas ferramentas compõem o estado-da-arte em modelagem dos processos físicos e dinâmicos da atmosfera. Verificou-se que ambos os modelos são preparados para o acoplamento com outros componentes do Sistema Terrestre e com um sistema de assimilação de dados. O SHiELD leva vantagem por ser um modelo cujo núcleo dinâmico é utilizado em modo operacional nos Estados Unidos, e diversas funcionalidades já são preparadas para o ambiente operacional. Por outro lado, a interação com a equipe de desenvolvimento do MPAS é um ponto positivo e importante a se levar em consideração, uma vez que facilita o intercâmbio do conhecimento e resolução de problemas de aspecto técnico no uso do modelo em ambiente operacional. Em termos quantitativos, o SHiELD e MPAS também apresentam destreza semelhante. ambos os modelos mostram um limite de previsibilidade de até 8 dias para o domínio global nos Hemisférios Norte e Sul. O uso de diferentes condições iniciais para o processo de inicialização do FV3-SHiELD indicou ligeiras diferenças e desempenho superior das reanálises do ERA5.

Palavras-chave: Avaliação de modelos. Modelos globais. MONAN. Previsão Numérica de Tempo. Verificação de previsões numéricas.

SOFTWARE QUALITY EVALUATION OF NUMERICAL GLOBAL MODELS FOR THE MONAN PROJECT – PART I: QUALITATIVE, QUANTITATIVE, AND MULTIVARIATE ANALYSIS

ABSTRACT

The National Institute for Space Research (INPE) aims to develop new technological solutions to mitigate the impacts of high-impact weather and environmental events through a National Program for Research, Development, and Innovation. The initiative seeks to engage academia, private and public sectors, including decision-makers and regional meteorological agencies, to support the transfer of scientific and technological knowledge into products and services derived from numerical modeling. The focus of the initiative relies on the development of a unified community-based Earth system model – the Model for Ocean-laNd-Atmosphere predictioN (MONAN). INPE, in collaboration with national and international partners, leads the initiative to develop MONAN, that will provide operational predictions tailored for South America, providing climate information at different spatial and temporal scales. To develop a state-of-the-art Earth System model, it is necessary to leverage new techniques in high-performance computing, physical and biogeochemical processes, a cutting-edge dynamical core, and advance in the application of available environmental data. This means that MONAN will employ innovative techniques in

Artificial Intelligence (AI), Machine Learning (ML), and data analysis, offering research and innovation opportunities throughout the numerical prediction workflow. The initial step towards the development of MONAN involves selecting a dynamical core that incorporates physical parameterization codes tailored to South American climatic characteristics. These codes can be derived from existing modeling systems currently in operation or adapted from other numerical models. This work aims to describe the efforts undertaken to evaluate the main open-source numerical global models freely available, which are part of the new generation of global modeling systems, to select the atmospheric component for MONAN. This work facilitates a better understanding of the key features, strengths, and limitations of the analyzed models concerning the simulation of the South American atmosphere. Among the models assessed concerning software quality, only SHiELD/FV3 and MPAS proved to be capable of providing a hydrostatic dynamical core for the composition of the atmospheric component of MONAN. These two tools represent the state-of-the-art in modeling the physical and dynamical processes of the atmosphere. It was found that both models are prepared to be coupled to other components of the Earth System and with a data assimilation system. SHiELD has an advantage as its dynamical core is operational in the United States, and various functionalities are already prepared for operational environment. On the other hand, the interaction with the MPAS development team is a positive and important aspect to consider, as it facilitates the exchange of knowledge and the resolution of technical issues in using the model in operations. The objective evaluation indicated SHiELD and MPAS with comparable skill. Both models have a predictability limit of up to 8 days for the global domain in the Northern and Southern Hemispheres. The use of different initial conditions for the initialization process of FV3-SHiELD indicated slight differences and superior performance when initialized with ERA5 reanalyses.

Keywords: Model evaluation. Global models. MONAN. Numerical Weather Prediction. Forecast verification.

LISTA DE FIGURAS

| | <u>Pág.</u> |
|--|-------------|
| 3.1 Áreas para avaliação da precipitação sobre a América do Sul | 13 |
| 4.1 Diagrama de desempenho para as previsões de precipitação de a) 24h b) 36h, c) 48h, d) 60h, e) 120h (132h) e f) 240h. As esferas representam os diferentes limiares de precipitação: 0,5mm, 2mm, 5mm, 10mm, 15mm, 20mm, 35mm, sendo a cor vermelha representativa do modelo MPAS e preto, do modelo SHiELD. | 31 |
| 4.2 Idem à Figura 4.1, exceto para a área B5 | 32 |
| 4.3 Idem à Figura 4.1, exceto para a área B1 | 33 |
| 4.4 Métricas das sub-características da Manutenibilidade com os modelos FV3, GEF e MPAS. | 34 |
| 4.5 Series temporais da ACC do geopotencial a 500 hPa dos modelos MPAS (vermelho) e SHiELD (azul), para o Hemisfério Norte (painel superior), trópicos (médio) e Hemisfério Sul (painel inferior). Previsões de 120 h (5 dias) para o período de junho 2021 a maio 2022 (74 casos). | 35 |
| 4.6 Distância de MH para o MPAS e FV3 e diferença normalizada (DifMH) para os prazos de previsão de 24h (a), 48h (b), 72h (c), 120h (d) e 240h (e) | 36 |
| 4.7 Idem que a Figura 4.6 mas para os meses de verão austral sobre a América do Sul. | 37 |
| 4.8 Idem que a Figura 4.7 mas para os meses de inverno austral sobre a América do Sul. | 38 |

LISTA DE TABELAS

| | <u>Pág.</u> |
|---|-------------|
| 2.1 Subcaracterísticas de Funcionalidade (pergunta chave: Satisfaz a neces- sidade do projeto MONAN?). | 5 |
| 3.1 Características dos modelos e das integrações numéricas realizadas. <i>wall- clock time</i> - tempo de processamento para 24h de previsão | 12 |
| 3.2 Parametrizações físicas dos modelos SHiELD e MPAS | 12 |
| 4.1 Pontuações atribuídas às métricas de avaliação da qualidade de software. | 18 |
| 4.2 Pontuações atribuídas às métricas de avaliação da qualidade de software. | 18 |

| | | |
|-----|---|----|
| 4.3 | Pontuações atribuídas às métricas de avaliação da qualidade de software. | 19 |
| 4.4 | Pontuação total atribuída às métricas de avaliação da qualidade de software. | 19 |
| 4.5 | Comportamento dos índices para as previsões dicotômicas dos modelos MPAS e SHiELD. As setas no sentido para baixo representam desempenho inferior em comparação com o outro modelo comparado, enquanto a seta para cima representa o comportamento oposto. Os símbolos = e \leq e $>$ representam o seu real significado. | 21 |
| 4.6 | Idem à Figura 4.5, exceto para o limiar menor ou igual a 35mm. | 21 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|-------|--|
| ACC | – <i>Anomaly Correlation Coefficient</i> (correlação de anomalias) |
| AS | – América do Sul |
| CPTEC | – Centro de Previsão do Tempo e Estudos Climáticos |
| CSI | – <i>Critical Success Index</i> |
| DIMNT | – Divisão de Modelagem Numérica do Sistema Terrestre |
| ECMWF | – <i>European Centre for Medium-Range Weather Forecasts</i> |
| ERA5 | – <i>the fifth generation ECMWF atmospheric reanalysis</i> |
| ETS | – <i>Equitable Threat Score</i> |
| FAR | – <i>False Alarm Ratio</i> |
| bias | – <i>Frequency Bias</i> |
| FV3 | – <i>Finite-Volume Cubed-Sphere Dynamical Core</i> |
| GAM | – Grupo de Avaliação de Modelos |
| GDPFS | – <i>Global Data-processing and Forecasting System</i> |
| GEF | – <i>Global Eta Framework</i> |
| GFDL | – <i>Geophysical Fluid Dynamics Laboratory</i> |
| GFS | – <i>Global Forecast System</i> |
| GPM | – <i>Global Precipitation Measurement Mission</i> |
| IA | – Inteligência Artificial |
| INPE | – Instituto Nacional de Pesquisas Espaciais |
| MH | – Distância de Mahalanobis |
| MONAN | – <i>Model for Ocean-laNd-Atmosphere predictioN</i> |
| MPAS | – <i>Model for Prediction Across Scales</i> |
| NASA | – <i>National Aeronautics and Space Administration</i> |
| NCAR | – <i>National Center for Atmospheric Research</i> |
| NCEP | – <i>National Centers for Environmental Prediction</i> |
| NWP | – <i>Numerical Weather Prediction</i> |
| OMM | – Organização Meteorológica Mundial |
| PNT | – Previsão Numérica de Tempo |

| | |
|--------|---|
| PQPF | – <i>Probabilistic Quantitative Precipitation Forecasting</i> |
| POD | – <i>Probability of Detection</i> |
| QPF | – <i>Quantitative Precipitation Forecasts</i> |
| SHiELD | – <i>System for High-resolution modeling for Earth-to-Local Domains</i> |
| RMSE | – Raiz do Erro Quadrático Médio |
| SR | – Razão de sucesso |
| SS | – <i>Skill Score</i> |
| SVN | – <i>Subversion</i> |
| WGNE | – <i>Working Group on Numerical Experimentation</i> |
| WIPPS | – WMO <i>Integrated Processing and Prediction System</i> |
| WMO | – <i>World Meteorological Organization</i> |

SUMÁRIO

| | <u>Pág.</u> |
|--|-------------|
| 1 Introdução | 2 |
| 2 Qualidade de produto de software | 3 |
| 3 Avaliação da qualidade de software | 6 |
| 3.1 Métricas de avaliação da qualidade de software | 8 |
| 3.1.1 Avaliação da acurácia dos modelos globais | 10 |
| 3.1.1.1 Desenho dos experimentos | 10 |
| 3.1.1.2 Dados de referência | 11 |
| 3.1.1.3 Avaliação estatística objetiva | 12 |
| 3.1.1.4 Análise multivariada – Distância de Mahalanobis | 15 |
| 4 Resultados | 17 |
| 4.1 Métricas de avaliação da qualidade de software – Pontuação | 17 |
| 4.2 Avaliação estatística da precipitação | 17 |
| 4.3 Avaliação estatística da altura geopotencial | 22 |
| 4.4 Análise Multivariada de Mahalanobis | 24 |
| 5 Considerações Finais | 25 |
| 6 Contribuição dos autores | 27 |
| REFERÊNCIAS BIBLIOGRÁFICAS | 28 |

1 Introdução

Para uma compreensão abrangente do sistema acoplado oceano-superfície-atmosfera-criosfera/gelo marinho em diferentes escalas espaciais e temporais, é essencial integrar e expandir o conhecimento sobre a dinâmica atmosférica, os processos oceânicos que modulam ou impactam a atmosfera e os possíveis mecanismos de retroalimentação. O Centro de Previsão do Tempo e Estudos Climáticos (CPTEC), sediado no Centro de Ciência do Sistema Terrestre do Instituto Nacional de Pesquisas Espaciais (INPE), busca aumentar sua liderança nacional e internacional em ciência e tecnologia. Em um mundo mais quente e em transformação, o INPE visa desenvolver novas estratégias nacionais de resposta à sociedade brasileira, oferecendo soluções eficazes para reduzir problemas associados à ocorrência de eventos climáticos, climáticos e ambientais de alto impacto por meio de um Programa Nacional de Pesquisa, Desenvolvimento e Inovação. A iniciativa busca envolver diferentes partes interessadas, como academias, setores públicos, formuladores de políticas e agências meteorológicas regionais, para apoiar a transferência de conhecimento científico para serviços em uma abordagem do Sistema Terrestre.

Com o objetivo de fornecer uma gama mais ampla de produtos numéricos meteorológicos e ambientais mais precisos, o foco da iniciativa é o desenvolvimento de um modelo unificado baseado na comunidade do sistema terrestre - o *Model for Ocean-laNd-Atmosphere predictioN* (MONAN). O MONAN produzirá previsões contínuas adequadas para a América do Sul, fornecendo informações úteis para diferentes setores econômicos e sociais, por meio de previsões mais confiáveis em diferentes escalas espaciais e temporais. O INPE lidera o desenvolvimento do MONAN, planejado para substituir os modelos atmosféricos atualmente aplicados. Um corpo científico de direção, composto por destacados cientistas nacionais, é responsável pela gestão do desenvolvimento e operação do MONAN. Para desenvolver um modelo terrestre de última geração, é necessário aproveitar as novas técnicas em computação de alto desempenho, processos físicos e biogeoquímicos, um núcleo dinâmico de última geração e tornar-se centrado em dados. Isso significa que o MONAN utilizará técnicas inovadoras em Inteligência Artificial (IA), Aprendizado de Máquina (ML) e volume de dados, oferecendo ótimas oportunidades ao longo do fluxo de trabalho da previsão numérica. O primeiro passo na construção do MONAN envolve a escolha de um núcleo dinâmico que incorpora códigos computacionais que incorporam física adaptada às características climáticas sul-americanas. Esses códigos podem ser derivados de sistemas de modelagem existentes em operação ou adaptados de outros modelos modernos.

Para identificar qualitativa e quantitativamente o desempenho dos modelos avaliados em prover a base do MONAN, formulou-se um processo de avaliação em ponto de grade que tem como base as recomendações da WMO para o *Global Data-processing and Forecasting System* (GDPFS) ([OMM, 2019](#)), atualmente denominado *WMO Integrated Processing and Prediction System* (WIPPS), e para o *Verification and Intercomparison of QPFs and PQPFs from Operational NWP Models* ([WGNE, 2018](#)). As métricas, variáveis dos modelos avaliadas foram extraídas das referidas documentações. Para avaliar de forma quantitativa os modelos selecionados, fez-se uso da suíte de avaliação de modelos disponível no CPTEC. A área de interesse é a América do Sul, de modo que avaliações em regiões específicas serão apresentadas. Aspectos de grande escala serão também abordados, tendo em vista a escala global dos modelos analisados.

O presente documento tem por objetivo apresentar a primeira etapa do processo de avaliação, considerando os critérios estabelecidos pela norma 9126 ([WIKIPEDIA, 2022](#)), que medem a qualidade do produto de software. Existe uma série de Atributos de Qualidade que pode ser analisada, dividida em seis características principais, sendo estas ainda subdivididas em subcaracterísticas ([WIKIPEDIA, 2022](#)). Entre os Atributos de Qualidade, o Grupo de Avaliação de Modelos da Divisão de Modelagem Numérica do Sistema Terrestre (GAM-DIMNT) está preocupado com o Aspecto da Funcionalidade e da Usabilidade.

2 Qualidade de produto de software

A norma ISO/IEC 9126 é um padrão internacional que trata da qualidade de produtos de software. Seu propósito é estabelecer um conjunto de parâmetros para uniformizar a avaliação da qualidade desses produtos, integrando-se ao modelo de qualidade das normas da família 9000 ([WIKIPEDIA, 2022](#)).

A ISO/IEC 9126, intitulada “Engenharia de Software - Qualidade do Produto”, é composta por diversas partes, incluindo:

- a) Modelo de qualidade;
- b) Métricas externas;
- c) Métricas internas;
- d) Métricas de qualidade em uso.

No contexto brasileiro, a norma correspondente é a NBR 13596, que foi substituída pela NBR ISO/IEC 9126-1, por sua vez, substituída pela Norma ISO/IEC 25010:2011 ([WIKIPEDIA](#), 2022).

Esta seção da NBR ISO/IEC 9126 descreve um modelo de qualidade para produtos de software, dividido em duas partes: a) qualidade interna e externa e b) qualidade em uso. Na primeira parte, são especificadas seis características para qualidade interna e externa, cada uma subdividida em subcaracterísticas. Essas subcaracterísticas são evidenciadas externamente quando o software é utilizado em um sistema computacional, resultando de atributos internos do software. Essa parte da norma não detalha o modelo além do nível de subcaracterísticas.

A segunda parte do modelo especifica quatro características para qualidade em uso, mas não aprofunda o modelo além do nível de características. A qualidade em uso representa, para o usuário, o impacto combinado das seis características de qualidade do produto de software.

Essas características são aplicáveis a todos os tipos de software, incluindo programas de computador. Elas fornecem terminologia consistente para abordar a qualidade de produtos de software e uma estrutura para especificar requisitos e realizar comparações entre esses produtos.

A norma 9126 foca na qualidade do produto de software, propondo Atributos de Qualidade, distribuídos em seis características principais, com cada uma delas divididas em subcaracterísticas. Entre elas, destaca-se por sua importância para o presente trabalho as características de **Funcionalidade** e **Usabilidade**.

Funcionalidade

Entende-se por Funcionalidade a capacidade de um software de satisfazer as necessidades de seu usuário, sejam elas declaradas ou implícitas, por meio do provimento de funcionalidades, dentro de um determinado contexto de uso (ISO/IEC 9126, Wikipedia, 2022).

O Aspecto da Funcionalidade possui as seguintes subcaracterísticas ([WIKIPEDIA](#), 2022):

- Adequação: Capacidade do produto de software de prover um conjunto apropriado de funções para tarefas e objetivos do usuário especificados.

- Acurácia: Capacidade do produto de software de prover, com o grau de precisão necessário, resultados ou efeitos corretos ou conforme acordados.
- Interoperabilidade: Capacidade do produto de software de interagir com um ou mais sistemas especificados.
- Conformidade: Capacidade do produto de software de estar de acordo com normas, convenções ou regulamentações previstas em leis e prescrições similares relacionadas à funcionalidade.
- Segurança: Capacidade do produto de software de proteger informações e dados, de forma que pessoas ou sistemas não autorizados não possam lê-los nem modificá-los e que não seja negado o acesso às pessoas ou sistemas autorizados.

As subcaracterísticas mencionadas são resumidas na Tabela 2.1, na qual são associadas a uma pergunta chave para melhor entendimento do leitor.

Tabela 2.1 - Subcaracterísticas de Funcionalidade (pergunta chave: Satisfaz a necessidade do projeto MONAN?).

| Subcaracterística | Pergunta Chave |
|--------------------|---|
| Adequação | Propõe-se a fazer o que é apropriado? |
| Acurácia | Faz o que foi proposto de forma correta? |
| Interoperabilidade | É capaz de interagir com os sistemas especificados? |
| Conformidade | Está de acordo com as normas, leis, entre outros? |
| Segurança | Evita acesso não autorizado a programas e dados? |

Fonte: Adaptado de Bianchi (2022).

Usabilidade

A Usabilidade é definida como a capacidade do produto de software de ser compreendido, aprendido, operado e atraente ao usuário, quando usado sob condições especificadas. Note que este conceito é bastante abrangente e se aplica mesmo a programas que não possuem uma interface para o usuário final. Por exemplo, um programa *batch* executado por uma ferramenta de programação de processos também pode ser avaliado quanto a sua usabilidade, no que diz respeito a ser facilmente compreendido, aprendido etc.

O Aspecto da Usabilidade possui as seguintes subcaracterísticas ([WIKIPEDIA, 2022](#)):

- Inteligibilidade: É a capacidade do produto de software de possibilitar ao usuário compreender se o software é apropriado e como ele pode ser usado para tarefas e condições de uso específicas.
- Apreensibilidade: É a capacidade do produto de software de possibilitar ao usuário aprender sua aplicação.
- Operacionalidade: É a capacidade do produto de software de possibilitar ao usuário operá-lo e controlá-lo.
- Proteção frente a erros de usuários: Se refere a como o produto consegue prevenir erros dos usuários.
- Atratividade: É a capacidade do produto atrair um potencial usuário para o sistema, o que pode incluir desde a adequação das informações prestadas para o usuário até os requintes visuais utilizados na sua interface gráfica (se houver).
- Acessibilidade: Refere-se à prática inclusiva de fazer softwares que possam ser utilizados por todas as pessoas que tenham deficiência ou não.
- Conformidade: Relacionado à capacidade do produto de apresentar normalidade quanto às convenções, guias ou regras/regulamentações relacionadas à usabilidade.

3 Avaliação da qualidade de software

Os modelos globais selecionados para avaliação foram o *System for High-resolution modeling for Earth-to-Local Domains–Finite-Volume Cubed-Sphere Dynamical Core* (SHiELD–FV3) ([CHASSIGNET et al., 2018](#)), desenvolvido pelo *Geophysical Fluid Dynamics Laboratory* (GFDL), versão 202210, o *Model for Prediction Across Scales–Atmosphere* (MPAS-A, a partir de agora denominado apenas como MPAS) ([SKA-MAROCK et al., 2012](#)), desenvolvido pelo *National Center for Atmospheric Research* (NCAR), versão 6.3, e o modelo *Global Eta Framework* (GEF) ([ZHANG; RANČIĆ, 2007; LATINOVIC et al., 2017](#)). Os candidatos são os principais modelos de código aberto, disponíveis gratuitamente, que embarcam núcleos dinâmicos da nova geração de sistemas de modelagem global.

Para avaliar os modelos de acordo com a proposta de qualidade de software, um conjunto de métricas foi desenvolvido a partir da análise de características esperadas

ou desejadas dos modelos. As subcaracterísticas selecionadas como potenciais para impactar a qualidade dos softwares analisados foram Adequação, Acurácia, Interoperabilidade, Inteligibilidade e Operacionalidade. Além destas, tornou-se necessário elaborar uma nova subcaracterística, denominada de Capacidade de Colaboração, definida a seguir::

- Capacidade de colaboração: É interpretada como a interação e colaboração com o(s) grupo(s) de modelagem responsável(ies) pelo desenvolvimento dos modelos numéricos selecionados.

Vale destacar que a avaliação da qualidade de software fazendo uso das métricas desenvolvidas é um processo subjetivo, pois nem todas as métricas estão estabelecidas na literatura. Entretanto, conforme será apresentado ao longo do texto, é possível encontrar subsídios na literatura específica dos modelos para identificar a melhor resposta para o que cada métrica se propõe a analisar.

O MONAN deverá ser um modelo comunitário, desenvolvido com apoio de diversas instituições, sob a liderança do INPE, que envolverá a participação de diferentes atores, como os setores acadêmico e público, formadores de políticas públicas e agências de meteorologia regionais. Os produtos numéricos a serem gerados com o MONAN serão utilizados como base para a oferta de uma série de serviços ambientais, em uma abordagem do Sistema Terrestre, em diferentes escalas espacial e temporal. A Adequação deverá indicar qual núcleo dinâmico embarcado nos modelos avaliados tem a capacidade de compor o MONAN, visando cumprir a necessidade de atender às escalas global, regional e futuramente, à escala local.

Centros de Previsão Numérica de Tempo (PNT) operacionais têm utilizado modelos numéricos abaixo do limite de espaçamento horizontal de 10 km (WGNE, 2022), o que não mais permite que a aproximação hidrostática seja utilizada (PIELKE, 1984; STEPPELER et al., 2003). Com o crescimento contínuo do poder computacional, centros de meteorologia e grupos de modelagem numérica têm empenhado esforços para permitir a simulação da circulação geral da atmosfera em domínios não hidrostáticos (STEVENS et al., 2019). Para a região tropical, onde grande parte do território brasileiro está localizado, representar numericamente a convecção tropical e todos os processos físicos associados é fundamental, e para tal, a aproximação não-hidrostática é uma exigência.

O modelo a ser avaliado deverá também estar preparado para interagir com outras componentes do Sistema Terrestre e com outros softwares. O sistema terrestre

é composto por várias partes, como a atmosfera, os oceanos, as camadas de gelo, a terra e outros. Cada uma dessas componentes é modelada separadamente para simular os processos físicos específicos que ocorrem em cada parte. A avaliação de qualidade de software indicará se o modelo numérico está preparado para ser conectado a um acoplador, responsável por facilitar a comunicação bidirecional entre as componentes de um modelo do sistema terrestre, garantindo que as informações relevantes sejam trocadas entre eles em intervalos regulares durante uma simulação. Ele sincroniza os diferentes modelos de cada componente e garante que suas interações sejam adequadamente representadas no sistema como um todo.

Outra importante componente de um sistema de modelagem numérica é o sistema de Assimilação de Dados, essencial para fornecer as condições iniciais a partir do uso de observações obtidas de diferentes fontes. A Assimilação de Dados é também fundamental para a produção de previsões de tempo e clima por conjunto. A avaliação de qualidade de software indicará se o modelo numérico está preparado para ser integrado a um sistema de Assimilação de Dados.

Por fim, o CPTEC/INPE é um centro operacional de previsão numérica, e requer que os softwares selecionados estejam o mais preparados possível para atuar operacionalmente. É fundamental que o software avaliado possibilite ao usuário operá-lo e controlá-lo, mantendo a proteção frente a erros de usuários.

3.1 Métricas de avaliação da qualidade de software

Uma série de métricas e ferramentas foram consideradas na avaliação de Funcionalidade de software dos modelos globais. São elas:

- a) Núcleo dinâmico não-hidrostático;
- b) Facilidade de interação com grupo desenvolvedor do modelo;
- c) Formato de saída dos dados dos modelos;
- d) Topo do modelo;
- e) Capacidade de regionalização;
- f) Coordenada vertical;
- g) Manual técnico disponível e de fácil acesso;
- h) Endendimento do uso do modelo;

- i) Estabilidade numérica;
- j) Capacidade de aplicações em:
 - Ondas oceânicas
 - Clima urbano
 - Qualidade do ar
 - Agricultura
 - Previsão de Tempo
 - Previsão subsazonal
 - Previsão sazonal
 - Projeções climáticas
 - Eventos extremos/raros
- k) Parametrizações físicas adequadas para a América do Sul;
- l) Possuir acoplamento com outras componentes:
 - Superfície continental
 - Oceano/criosfera
 - Ondas
 - Química atmosférica
 - Hidrologia
 - Assimilação de Dados
- m) Possuir acoplador
- n) Nível de acurácia dos modelos adequados para o projeto MONAN.

Exceto para os itens “a” e “n”, todas as demais métricas foram avaliadas seguindo os critérios de pontuação definidos pelas normas ISO, a saber:

- Três pontos para o modelo que melhor atende as métricas listadas
- Um ponto para o modelo que atende de forma insatisfatória as métricas listadas

Os modelos que atenderam a métrica receberam três pontos, enquanto que aqueles que não atenderam, receberam um ponto. Para casos de empate, os melhores avaliados receberam três pontos. Analogamente, os empatados com desempenho inferior receberam um ponto.

Para o item “a”, foram também aplicadas as pontuações estabelecidas na normas ISO, porém, estabeleceu-se que a aproximação não-hidrostática para o projeto MONAN é de fundamental importância, e portanto, não seria justificável avaliar os modelos, em termos de acurácia, caso não pudessem atender a esta métrica. Isto porque a demanda por recurso computacional (tempo de processamento e armazenamento) seria alta para atender a métrica “n” para os três modelos avaliados. Assim, os modelos que não possuíssem núcleo dinâmico não-hidrostático seriam eliminados das demais etapas de avaliação.

3.1.1 Avaliação da acurácia dos modelos globais

A primeira etapa que atecedeu o desenho dos experimentos e a definição da metodologia de análise de acurácia foi a avaliação dos modelos selecionados com base da métrica “a”.

O núcleo dinâmico dos modelos MPAS e SHIELD/FV3 é nativamente não-hidrostático, enquanto que o GEF possui um núcleo dinâmico hidrostático. o modelo GEF foi o que recebeu menor pontuação, sendo considerado inapto para as aplicações a que se destina o MONAN. Como resultado, o modelo GEF foi eliminado das seguintes etapas de avaliação.

Para a análise da Acurácia (item “n”), foi utilizado um conjunto de métricas e ferramentas estatísticas com base em experimentos numéricos. A seguir, serão detalhados os experimentos numéricos realizados bem como as métricas estatísticas aplicadas nesta etapa do trabalho.

3.1.1.1 Desenho dos experimentos

Foram realizados dois experimentos numéricos para cada um dos modelos avaliados. A diferença entre os experimentos reside na fonte da condição inicial. Foram utilizadas as análises do Global Forecast System (GFS) no *National Centers for Environmental Prediction* (NCEP) e as reanálises ERA5 do *European Centre for Medium-Range Weather Forecasts* (ECMWF) (*the fifth generation ECMWF atmospheric reanalysis*). Para reduzir o custo computacional, os modelos foram integrados a cada 5 dias a partir do dia 01/06/2021. Para cada data de inicialização,

os modelos foram integrados até 240h (10 dias), com saídas a cada 6h. As parametrizações físicas utilizadas foram aquelas já embarcadas nos modelos numéricos. Nesta etapa do trabalho, foi realizado nenhum teste de sensibilidade ou mudanças nas parametrizações físicas. As integrações foram realizadas no cluster EGEON do CPTEC/INPE. As previsões dos modelos foram interpoladas à grade 0.25 x 0.25 do ERA5 para as comparações. A seguir são apresentadas as características dos experimentos realizados para avaliação.

- a) Espaçamento de grade horizontal de 15 km para o modelo MPAS e 13 km para o SHiELD (núcleo dinâmico – FV3); 55 e 64 níveis verticais, respectivamente, sendo 0,2 hPa assumido como o topo dos modelos;
- b) Condições iniciais da reanálise ERA5 e GFS;
- c) Período: 01/06/2021 a 01/06/2022;
- d) Análise de 74 casos para da experimento
- e) Prazo de previsão de até 10 dias (240 h), com início às 00:00UTC;
- f) Resolução temporal das saídas: 6 h.

As características e especificações das integrações numéricas realizadas são apresentadas na Tabela 3.1.

As parametrizações físicas dos modelos SHiELD e MPAS utilizadas nos experimentos são listadas na Tabela 3.2.

3.1.1.2 Dados de referência

A reanálise do ERA5 foi utilizada como dado de referência no processo de avaliação. A exceção é a precipitação, que foi avaliada considerando os dados de precipitação do *Global Precipitation Measurement Mission* (GPM) da *National Aeronautics and Space Administration* dos Estados Unidos (NASA). Para as áreas avaliadas sobre a América do Sul, foi utilizado o dado MERGE ([ROZANTE et al., 2020](#)), um produto de precipitação que combina dados de precipitação estimados por sensoriamento remoto e dados in situ.

Como os modelos foram inicializados às 00:00 UTC, a precipitação do modelo foi acumulada a partir das primeiras 12h de integração considerando os prazos de previsão de 36h, 60h, ..., até 240 dias.

Tabela 3.1 - Características dos modelos e das integrações numéricas realizadas. *wall-clock time* - tempo de processamento para 24h de previsão

| Características | SHiELD | MPAS |
|----------------------------------|-------------|-----------|
| Computador/cluster (XE, XC, etc) | EGEON | EGEON |
| Resolução Horizontal (Km) | 15 | 13 |
| Coordenada Vertical | σ -P | Z |
| Domínio | GLB | GLB |
| Passo de Tempo (s) | 150 | 90 |
| No. de Pontos de Grade | 3.538.944 | 2.611.442 |
| Fonte das Condições Iniciais | ERA5 | ERA5 |
| Topo do Modelo | 20 Pa | 30 km |
| Prazo da Previsão (h) | 240 | 240 |
| Compilador | intel | intel |
| Formato dos Outputs | nc | nc |
| Frequência dos Outputs (h) | 6/3 | 6 |
| Espaço total para Rodada (Gb) | 50 | 70 |
| Espaço total Pós-Processado (Gb) | 12 | 6 |
| Nº de núcleos de processadores | 1156 | 1024 |
| <i>wall-clock time</i> | 15min | 20min |

Fonte: Produção própria.

Tabela 3.2 - Parametrizações físicas dos modelos SHiELD e MPAS

| Parametrizações físicas | SHiELD | MPAS |
|----------------------------|-----------|-------------|
| Radiação Atmosférica | RRTMG | RRTMG |
| Superfície Continental | NOAH | NOAH |
| Fluxos superficiais ar-mar | SLAB | - |
| Microfísica de Nuvens | GFDL | WSM6 |
| Convecção Rasa/Profunda | SAS | New Tiedtke |
| Camada limite Planetária | EDMF | YSU |
| Gravity Wave Drag | GWD | YSU |
| Química | Desligada | Desligada |
| Aerossóis | Desligada | Desligada |

Fonte: Produção própria.

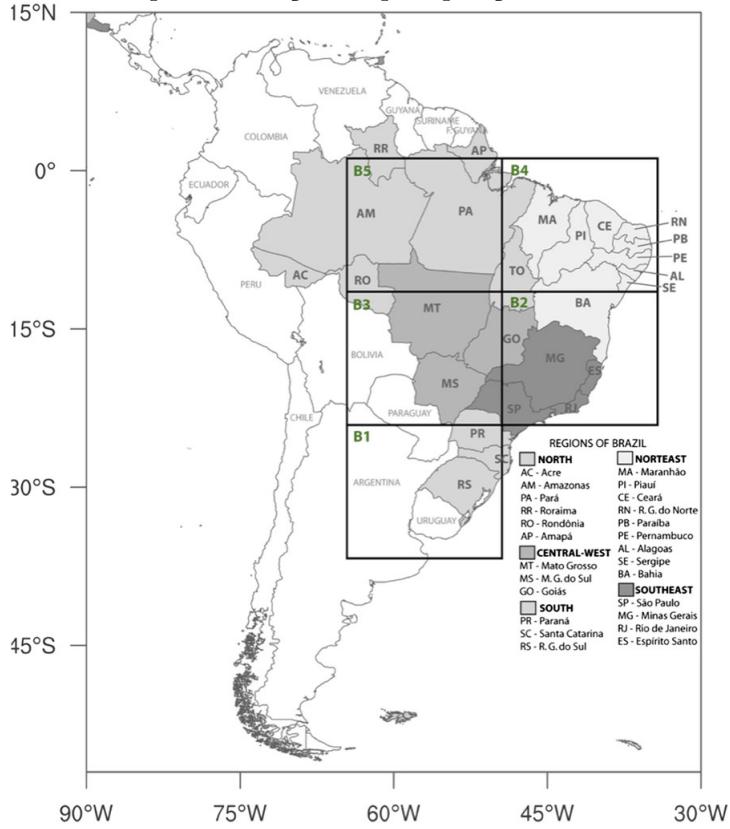
3.1.1.3 Avaliação estatística objetiva

A acurácia dos modelos globais foi analisada por meio de avaliações estatísticas objetivas da precipitação e altura geopotencial, na escala de tempo, em escala global e sobre a América do Sul.

Foram consideradas diferentes áreas geográficas do globo e subdomínios do conti-

nente sul-americano. A avaliação de subdomínios da América do Sul foi realizada conforme Figueroa et al. (2016) (Figura 3.1).

Figura 3.1 - Áreas para avaliação da precipitação sobre a América do Sul



Fonte: Figueroa et al. (2016)

A avaliação em escala global e subdomínios tem como base as recomendações da Organização Meteorológica Mundial para o WIPPS, antigo GDPFS (OMM, 2019), e para o *Verification and Intercomparison of QPFs and PQPFs from Operational NWP Models* (OMM; WWRP, 2009). As métricas, variáveis dos modelos avaliadas e as áreas para avaliação foram extraídas das referidas documentações, e forneceram as principais informações para o presente estudo.

As previsões de precipitação, consideradas como previsões dicotômicas (sim/não) foram analisadas por meio de métricas baseadas na tabela de contingência. São elas:

- *Frequency Bias* (bias);
- *Equitable Threat Score* (ETS);

- *Probability of Detection* (POD);
- *Critical Success Index* (CSI);
- *False Alarm Ratio* (FAR).

Os índices ETS, bias, POD, FAR (ou o seu oposto, a razão de sucesso, SR, ou 1-FAR), e CSI foram analisados por meio da plotagem do diagrama de performance ([ROEBBER, 2009](#)), que resume as medidas SR, POD, bias e CSI. Em uma abordagem que é conceitualmente semelhante ao diagrama de [Taylor \(2001\)](#), é possível explorar a relação geométrica entre POD, FAR, bias score e CSI. As previsões com melhores desempenhos apresentam POD, SR, bias e CSI próximo de 1, e a previsão perfeita seria aquela posicionada no canto superior direito do diagrama. Os desvios em uma dada direção indicam as diferenças relativas no POD e SR, e consequentemente no bias e CSI. É possível visualizar rapidamente as diferenças nos desempenhos dos modelos avaliados. Na posição de 45 graus no diagrama são encontrados os melhores resultados, mantendo as previsões sem viés e simultaneamente aumentando a detecção e a redução do alarme falso.

As métricas estatísticas utilizadas para avaliar as previsões de altura geopotencial no domínio global foram aquelas indicadas para a avaliação de variáveis contínuas. São elas:

- Coeficiente de correlação de anomalia (do inglês (*Anomaly Correlation Coefficient*, ACC));
- Viés;
- Raiz do Erro Quadrático Médio (RMSE).

O ACC é um índice utilizado para avaliar os limites de previsibilidade das previsões dos modelos numéricos de tempo. Valor de 1 indica previsões perfeitas e usa-se o valor de 0.6 como limite útil das previsões dos modelos. Abaixo deste limiar se considera não existir utilidade prática das previsões numéricas. A ACC é uma métrica com variações ruidosas nas diferentes escalas temporais e espaciais, então normalmente em escala global são definidas áreas para avaliar valores médios para verificar o desempenho dos modelos: Global (Glb), Hemisfério Norte e Sul (NH, NS), Trópicos (TRP) e América do Sul. Em geral, são avaliadas as variáveis: altura geopotencial, temperatura e vento no níveis de 200hPa, 500hPa e 850hPa. Destaca-se que as

previsões de MPAS fornecem como saída altura geométrica, a qual foi convertida para geopotencial, portanto, é necessário cautela em sua análise. Foram utilizados dados da reanálise ERA5 como referência e a climatologia do período de 1991-2021.

3.1.1.4 Análise multivariada – Distância de Mahalanobis

Além das métricas estatísticas apresentadas, foi realizada uma análise multivariada dos campos de circulação atmosférica utilizando a Distância de Mahalanobis (MH). A MH é uma medida que se assemelha a distância euclidiana, porém leva em conta a covariância dos do conjunto de dados. Em outros termos, a distância MH é uma correlação multivariável que permite comparar dois tipos de população representada por um conjunto de variáveis. MH é útil para se obter uma medida objetiva de quanto um conjunto de variáveis, i.e. população, se distancia de outro conjunto, dando uma visão uma visão holística dos resultados. MH não tem por objetivo indicar a causa do problema, mas sim o de quantificar o quanto este problema está afetando os resultados, a partir de valores numéricos. Quanto menor for MH, a população em teste (i.e. a previsão) estará se aproximando da população de controle (i.e. a análise).

A aplicação da MH propicia uma análise geral sobre a destreza dos modelos sobre os domínios analisados, fazendo uso de um conjunto de campos de circulação atmosférica que possuem uma correlação multivariável que permite comparar modelo e dado de referência.

A distância MH de uma população é definida matematicamente, por um conjunto de variáveis (x) e que possam ser categorizadas em p categorias. Dada uma coordenada de ponto de grade (i,j) com múltiplas variáveis (temperatura, geopotencial, vento zonal etc.), pode-se definir um vetor $X_1(l, p)$, onde l representa a variável prevista e p sua categoria, que pode ser o perfil vertical, a variação no tempo ou qualquer outro tipo de categoria. No presente estudo, levar-se-á apenas a distribuição vertical das variáveis. Define-se, com os mesmos critérios, uma população que servirá de comparação, ou seja, a análise do modelo, definido por $X_2(l, p)$. Os vetores X_1 e X_2 tem que ter as mesmas categorias, porém, não necessariamente todas as categorias de um mesmo vetor precisam estar definidas.

Faz-se a média de cada categoria, ou seja, $\mu_1(l)$ e $\mu_2(l)$. Em seguida, calcula-se a anomalia de cada par de vetores, tal como segue:

$$A_1(l, p) = X_1(l, p) - \mu_1(l) \quad (3.1)$$

e

$$A_2(l, p) = X_2(l, p) - \mu_2(l). \quad (3.2)$$

A seguir, calcula-se a covariância das anomalias de cada um dos vetores. As matrizes de covariância geradas ($Cov_1(l, l)$) e ($Cov_2(l, l)$) e as respectivas somatórias ($Som_1(l, l)$) e ($Som_2(l, l)$) terão o tamanho do número de variáveis usadas. Se duas variáveis, ter-se-á uma matriz 2×2 e se forem três, 3×3 e assim sucessivamente. Em seguida faz-se a média ponderada dos dois vetores, onde o fator ponderador é a somatória definida acima. Gerando uma covariância conjunta definida como segue:

$$Cov(l, l) = \frac{[Cov_1(l, l) \times Som_1(l, l)] + [Cov_2(l, l) \times Som_2(l, l)]}{Som_1(l, l) + Som_2(l, l)}. \quad (3.3)$$

Calcula-se a matriz inversa da covariância conjunta, ou seja, $Cov(l, l)^{-1}$. MH será definido pela raiz quadrada do somatório do produto matricial entre a matriz da diferença das médias $\mu_1(l)$ e $\mu_2(l)$ e transposta $[\mu_1(l)]^T$ e $[\mu_2(l)]^T$, e da matriz inversa das covariâncias. Ou seja:

$$MH = \sqrt{\sum [\mu_1(l) - \mu_2(l)] * [\mu_1(l) - \mu_2(l)]^T * Cov(l, l)^{-1}}. \quad (3.4)$$

A distância MH é calculada para cada ponto de grade do modelo e é sempre positivo.

A fim de condensar a análise de várias variáveis, a distância MH foi calculada para as previsões feitas a cada cinco dias, totalizando 74 casos para os prazos de previsão de 24h até 240h, com passo de tempo de 24h. As variáveis utilizadas foram temperatura, vento meridional e vento zonal em três níveis verticais (850, 500 e 250). A média de cada horário tanto para o ano como para as estações de DJF e JJA foi calculada. A análise do domínio global é apresentada e a análise regional fica restrita à América do Sul, com foco nas estações de verão (DJF) e inverno (JJA).

Definiu-se uma diferença ponderada para comparação entre os modelos MPAS e FV3, definida como:

$$DifMH = \frac{MH(MPAS) - MH(FV3)}{MH(MPAS) + MH(FV3)} * 100, \quad (3.5)$$

em que $DifMH < 0$ implica que o MPAS apresenta menor distância se comparado

ao FV3.

4 Resultados

Nesta Seção, são apresentados os resultados obtidos a partir das pontuações estabelecidas na norma ISO, bem como os resultados da avaliação estatística objetiva e multivariada.

4.1 Métricas de avaliação da qualidade de software – Pontuação

Um dos grandes diferenciais entre os modelos MPAS e FV3, reside na métrica “b” apresentada na Seção 3.1, que dispõe sobre interação com grupo de desenvolvimento do modelo. As interações com colaboradores do NCAR, desenvolvedores do MPAS, foram consideradas mais acessíveis que as interações com os colaboradores do GFDL, desenvolvedores do FV3. A possibilidade de interação e troca de experiências é fundamental para o melhor entendimento das particularidades do software e também da possibilidade de estabelecimento de acordos de cooperação bilateral. Portanto, para esta métrica, o FV3 recebeu pontuação 1 enquanto que o MPAS recebeu pontuação 3.

As pontuações atribuídas às métricas listadas na Seção 3.1 são indicadas nas Tabelas 4.1 – 4.3. As pontuações recebidas pelos modelos globais indicam que o MPAS e SHiELD atendem aos requisitos formulados, com pontuação total similar.

Somadas as pontuações atribuídas, tem-se 66 pontos para o SHiELD e MPAS. Adicionando 3 pontos para os modelos SHiELD e MPAS referente à métrica “a”, e 1 e 3 pontos para o SHiELD e MPAS referente à métrica “b”, respectivamente, a totalização para cada um dos modelos é, respectivamente, de 70 e 72 para o SHiELD e MPAS. Estes resultados são apresentados na Tabela 4.4.

4.2 Avaliação estatística da precipitação

O desempenho das previsões de precipitação produzidas pelos modelos MPAS e SHiELD é analisado a partir dos diagramas de desempenho apresentados na Figura 4.1 para o domínio global. Optou-se por apresentar o desempenho dos modelos por prazo de previsão, e portanto, cada gráfico apresenta a nuvem de desempenho, sendo cada ponto um limiar de precipitação.

À medida em que os limiares de precipitação aumentam, o desempenho do modelo

Tabela 4.1 - Pontuações atribuídas às métricas de avaliação da qualidade de software.

| Métrica | Descrição | Impactos Positivos | Impactos Negativos | Ferramentas | Unidade de medida | SHIELD | MPAS |
|--|--|---------------------------------------|--------------------|---|---|--------|------|
| Formato de saída dos dados dos modelos | O formato de saída do modelo é importante para determinação de consumo de armazenamento e de leitura do dado por outros softwares utilizados no CPTEC | Interoperabilidade Adequação | - | Manual técnico e/ou artigos científicos | Quantidade de formatos Preferencialmente grib, bin, nc, grib2 | 1 | 1 |
| Topo do modelo | Capacidade do modelo de ser configurado para um nível mais alto da atmosfera | Acurácia Adequação | - | Manual técnico e/ou artigos científicos | De 1 a 3 Quanto mais alto for o topo, maior a nota | 3 | 1 |
| Capacidade regionalização | Devido ao aspecto comunitário do MONAN, a regionalização sobre áreas de domínio de centros regionais é necessária Dados de saída dos modelos são facilmente utilizados por outros sistemas | Acurácia Interoperabilidade Adequação | - | Manual técnico e/ou artigos científicos | Regionalização sobre RMSP, RMRJ, Amazônia, Sul do Brasil, NEB | 3 | 3 |
| Coordenada vertical híbrida | O tipo de coordenada vertical é importante para o acoplamento entre o núcleo dinâmico e o sistema de AD | Acurácia Interoperabilidade Adequação | - | Manual técnico e/ou artigos científicos | V (1) ou F (0) | 3 | 3 |
| Total | | | | | | 10 | 8 |

Fonte: Produção própria.

Tabela 4.2 - Pontuações atribuídas às métricas de avaliação da qualidade de software.

| Métrica | Descrição | Impactos Positivos | Ferramentas | Unidade de medida | SHIELD | MPAS |
|---|--|------------------------------|---|-------------------------------|--|--|
| Manual técnico disponível e de fácil acesso | Manual técnico de fácil acesso, disponível publicamente | Adequação Interoperabilidade | Internet | V (3) e F (1) | 1 | 3 |
| Entendimento uso do modelo | Tempo necessário para aprendizado do uso do modelo | Interoperabilidade | Manual técnico | Menos tempo - 3 pontos | 1 | 3 |
| Estabilidade numérica | Capacidade do modelo de ser executado sem a ocorrência de instabilidade numérica | Acurácia | Acompanhamento das simulações | 0 => 3 se != 0 3 => se = 0 | 3 | 3 |
| Capacidade de aplicações 1. Ondas oceânicas 2. Clima urbano 3. Qualidade do ar 4. Agricultura 5. Previsão de tempo 6. Previsão subsazonal 7. Previsão sazonal 8. Projeções climáticas 9. Eventos extremos/raros | Aplicações 1. https://www.mdpi.com/2073-4433/13/4/591 2. https://md.copernicus.org/preprints/gmd-2019-298/gmd-2019-298.pdf 4. https://repository.library.noaa.gov/view/noaa/29398/noaa_29398_DS1.pdf 7. https://link.springer.com/article/10.1007/s00382-021-05797-7 https://aoupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020MS002223 | Acurácia Adequação | Manual técnico e/ou artigos científicos | Pontos por aplicação | 1. 3 2. 3 3. 3 4. 3 5. 3 6. 3 7. 3 8. 3 9. 3 | 1. 3 2. 3 3. 3 4. 1 5. 3 6. 3 7. 3 8. 3 9. 3 |
| Total | | | | | 32 | 34 |

Fonte: Produção própria.

diminui, já que se torna mais difícil que o modelo preveja valores elevados de precipitação que ocorrem em um número menor de pontos de grade (eventos). O modelo SHIELD tende a superestimar os eventos de precipitação abaixo de 15mm, indicado pelo bias acima de 1. Para casos acima de 35mm, o modelo gera subestimativa (bias inferior a 1). O modelo apresenta desempenho perfeito para eventos de precipitação

Tabela 4.3 - Pontuações atribuídas às métricas de avaliação da qualidade de software.

| Métrica | Descrição | Impactos Positivos | Ferramentas | Unidade de medida | SHIELD | MPAS |
|--|---|--------------------------------|---|-------------------------------------|--|--|
| Parametrizações físicas adequadas para a América do Sul | Parametrizações físicas desenvolvidas pelo CPTEC para melhor representar processos sub-grade nos trópicos | Acurácia Adequação | Artigos | 3 pontos para cada parametrização | 3 | 3 |
| Possuir acoplamento com: 1. Superfície 2. Oceano/criosfera 3. Ondas 4. Química atmosférica 5. Hidrologia 6. Assimilação de dados | Possuir acoplamento as componentes de interesse 5- https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2022MS003089 | Interoperabilidade | Manual técnico e/ou artigos científicos | V (3) e F (1) para cada acoplamento | 1. 3 2. 3 3. 3 4. 3 5. 3 6. 3 | 1. 3 2. 3 3. 3 4. 3 5. 3 6. 3 |
| Acoplador | O acoplador aumenta a interoperabilidade e promove a reutilização de código. O acoplador conecta domínios físicos em um ESM, ligando as saídas de uma componente e transformando-as nas entradas necessárias para executar outra componente | Acurácia Interoperabilidade | Manual técnico e/ou artigos científicos | V (3) e F (1) | 3 | 3 |
| Total | | | | | 24 | 24 |

Fonte: Produção própria.

Tabela 4.4 - Pontuação total atribuída às métricas de avaliação da qualidade de software.

| Modelo | Pontuação total |
|--------|-----------------|
| SHiELD | 70 |
| MPAS | 72 |

Fonte: Produção própria.

de até 20mm (bias próximo de 1). O MPAS também superestima eventos de precipitação leve, porém, de até 5mm. Para os casos de precipitação acima de 10mm, ocorre subestimativa da precipitação global. Para as previsões de 36h, houve o aumento da superestimativa dos eventos de precipitação até 15mm e aumento da subestimativa dos eventos de precipitação acima de 20mm por parte do SHiELD, enquanto que o MPAS apresentou desempenho semelhante, porém, o aumento da superestimativa ocorre para eventos de até 5mm. Já a subestimativa que ocorre com o MPAS foi reduzida para eventos acima de 10mm, sendo que o bias para eventos de até 10mm é próximo de 1.

Para previsões a partir de 60h, não se observou mudanças importantes no bias do MPAS, enquanto o bias do SHiELD apresentou maior superestimativa dos eventos de precipitação até 20mm. Para eventos acima de 35mm, a partir de 96h o seu bias se aproxima de 1.

Até 132h de previsão, observou-se que a subestimativa do MPAS foi reduzida, indicado pela movimentação dos pontos no diagrama em direção às previsões sem bias, e mantendo, simultaneamente, maior detecção e reduções de falsos alarmes (mudanças em POD e FAR) se comparado ao SHiELD. Por outro lado, a probabilidade de detecção é maior para limiares mais baixos de precipitação. O SHiELD tende a prever mais eventos de precipitação que não ocorrem mas apresenta melhor desempenho em prever os eventos que realmente ocorrem para os limiares maiores de precipitação.

O CSI representa o quão os eventos de chuva previstos correspondem aos eventos de chuva observados. Este índice considera apenas aquelas previsões corretas, sendo sensível ao número de acertos. Ele também é fortemente depende da frequência climatológica dos eventos, e gera pontuações mais baixas para eventos mais raros, pois alguns acertos podem ocorrer aleatoriamente.

A análise do CSI indica que, embora o modelo MPAS tenha um menor número de acertos se comparado ao SHiELD (POD maior), ele apresenta maior CSI para todos os limiares de precipitação para até 240h de integração (exceto para de eventos de até 0,5mm). Este resultado indica que os eventos de chuva previstos pelo MPAS têm maior correspondência aos eventos de chuva observados se comparado ao SHiELD. Este resultado pode estar associado ao SHiELD prever eventos de precipitação aleatoriamente, o que estaria associado à superestimativa observada pelo modelo.

Nas Tabelas 4.5 e 4.6 são mostrados os resumos, em termos qualitativos, da análise realizada dos diagramas de desempenho dos modelos MPAS e SHiELD para o domínio global, considerando os limiares de até 10mm e 35mm, respectivamente.

Na análise dos diagramas de desempenho para as regiões da América do Sul, destaca-se a área B5, que corresponde à grade parte da região amazônica, de grande importância para o clima global. Diferente da análise global, que considera como referência os dados do GPM, aqui são considerados os dados do MERGE (ROZANTE et al., 2020).

Ambos os modelos apresentaram superestimativa dos eventos de precipitação até 15mm ao longo do tempo de integração, sendo o MPAS o modelo que apresentou uma maior frequência de eventos de precipitação, que se traduz em maior número de falsos alarmes. Menor número de eventos superestimados foi observado nas simulações de até 15mm para até 136h de previsão. A partir deste prazo de integração, houve redução de superestimativa e as previsões passaram a tender para bias próximo

Tabela 4.5 - Comportamento dos índices para as previsões dicotômicas dos modelos MPAS e SHiELD. As setas no sentido para baixo representam desempenho inferior em comparação com o outro modelo comparado, enquanto a seta para cima representa o comportamento oposto. Os símbolos = e \leq e $>$ representam o seu real significado.

| $\leq 10mm$ | | | | | | | | |
|-------------|-------------|--------------|------------|------------|--------|--------------|--------------|--------------|
| | MPAS | | | | SHiELD | | | |
| h | BIAS | POD | SR | CSI | BIAS | POD | SR | CSI |
| 24 | < 1 | \downarrow | \uparrow | \uparrow | > 1 | \uparrow | \downarrow | \downarrow |
| 48 | ≈ 1 | \uparrow | \uparrow | \uparrow | > 1 | \downarrow | \downarrow | \downarrow |
| 72 | ≈ 1 | \uparrow | \uparrow | \uparrow | > 1 | \downarrow | \downarrow | \downarrow |
| 98 | ≈ 1 | \uparrow | \uparrow | \uparrow | > 1 | \downarrow | \downarrow | \downarrow |
| 120 | ≈ 1 | \uparrow | \uparrow | \uparrow | > 1 | \downarrow | \downarrow | \downarrow |
| 144 | ≈ 1 | \uparrow | \uparrow | \uparrow | > 1 | \downarrow | \downarrow | \downarrow |
| 240 | ≈ 1 | \downarrow | \uparrow | = | > 1 | \uparrow | \downarrow | = |

Fonte: Produção própria.

Tabela 4.6 - Idem à Figura 4.5, exceto para o limiar menor ou igual a 35mm.

| $\leq 35mm$ | | | | | | | | |
|-------------|-------|--------------|------------|--------------|-------------|------------|--------------|--------------|
| | MPAS | | | | SHiELD | | | |
| h | BIAS | POD | SR | CSI | BIAS | POD | SR | CSI |
| 24 | < 1 | \downarrow | \uparrow | \uparrow | < 1 | \uparrow | \downarrow | \downarrow |
| 48 | < 1 | \downarrow | \uparrow | = | < 1 | \uparrow | \downarrow | = |
| 72 | < 1 | \downarrow | \uparrow | \downarrow | < 1 | \uparrow | \downarrow | \uparrow |
| 96 | < 1 | \downarrow | \uparrow | \downarrow | < 1 | \uparrow | \downarrow | \uparrow |
| 120 | < 1 | \downarrow | \uparrow | = | < 1 | \uparrow | \downarrow | = |
| 144 | < 1 | \downarrow | \uparrow | = | < 1 | \uparrow | \downarrow | = |
| 240 | < 1 | \downarrow | \uparrow | = | ≈ 1 | \uparrow | \downarrow | = |

Fonte: Produção própria.

a 1 para todos os prazos de integração mais longos. O modelo SHiELD apresentou comportamento inverso, subestimando todos os eventos de precipitação acima de 2mm e até 35mm para até 132h, quando passou a tender ao bias de 1 a partir deste prazo. Devido à superestimativa dos eventos previstos de precipitação, o MPAS apresenta POD mais elevado comparado ao SHiELD para eventos de até 10mm ao longo de todo o tempo de integração. Para os eventos de precipitação cujos limiares são acima de 20mm, o SHiELD apresentou POD mais elevado, acompanhado de maior frequência de SR a partir de 60h de integração. Este modelo também apresentou maior probabilidade de ter eventos previstos de precipitação com limiares superiores a 20mm correspondentes aos observados (CSI mais elevado), não indicando, portanto, resultado de inflação devido ao POD.

A região 1 também é de sensível importância, tendo em vista que abrange parte da região sul da América do Sul onde é influenciada por sistemas frontais e pela formação de sistemas convectivos de mesoescala.

Para previsões de 36h, o MPAS apresentou bias próximo de 1 para os eventos com limiares de até 20mm. Para estes casos, o SHiELD apresentou subestimativa. Para até 180h, o MPAS tende a superestimar estes casos, enquanto que ocorre redução da subestimativa pelo SHiELD. Os valores de POD acompanharam o desempenho observado no bias, enquanto SR apresentou maior frequência para o modelo SHiELD. Como o SR indica qual a fração das previsões de precipitação foi realmente observada, fornecendo informações da probabilidade de um evento observado ser realmente previsto, o que ocorre é que devido aos falsos alarmes, o SR acaba sendo reduzido, sendo, portanto, também uma consequência da superestimativa da precipitação de limiares mais baixos. Em termos de CSI, ambos os modelos se comportaram de forma semelhante ao longo do tempo de integração, sendo que para alguns prazos, o MPAS apresentou maior CSI, enquanto para outros prazos, o SHiELD apresentou os valores mais elevados.

4.3 Avaliação estatística da altura geopotencial

Na Figura 4.4a e 4.4b, painéis superiores, são apresentados os coeficientes de anomalia de correlação (ACC) médios da altura geopotencial em 500hPa em função do horizonte de previsão (até 10 dias) para os modelos FV3-SHiELD e MPAS, respectivamente, inicializados com as reanálises ERA5. Adicionalmente, o FV3-SHiELD foi inicializado usando as análises do GFS/NCEP (linhas tracejadas da Figura 4.4a). Ambos os modelos mostram um limite de previsibilidade de até 8 dias para as regiões GBL, HN e HS e ligeiramente menor para as regiões TPR e América do Sul.

Destaca-se que o ACC mede as correlações entre anomalias, sendo mais sensível às semelhanças entre os padrões das previsões e análises que seus valores absolutos. Por exemplo, um erro de fase pode afetar a correta interpretação do desempenho dos modelos (PERSSON, 2001).

O RMSE (Figuras 4.4a e 4.4b, painéis centrais) e BIAS (Figuras 4.4a e 4.4b, painéis inferiores) dos modelos também foram avaliados. Conforme aumenta o horizonte de previsão, o RMSE também cresce, em maior grau no HS que no HN e na escala global. Menores valores são observados em TPR e AS, lembrando que as variações de geopotencial nessas regiões são menores às que acontecem em latitudes médias e altas. Em geral, o BIAS indica subestimativa da altura geopotencial em 500 hPa nas previsões do SHiELD (Figura 4.4a) conforme se afasta da condição inicial. De forma semelhante ao RMSE, o BIAS apresenta menores valores nos TPR e AS comparado a HN, HS e GLB. Entretanto, os valores de BIAS do MPAS (Figura 4.4b) diferem do SHiELD, com menores valores no HN e HS. É possível que o método de conversão da altura geométrica para a altura geopotencial tenha interferido nos resultados do MPAS, o que pode justificar esse comportamento.

As previsões do SHiELD inicializado com dados do GFS/NCEP ((Figura 4.4a, linhas tracejadas), apresentaram um desempenho ligeiramente inferior àquelas com as condições iniciais do ERA5. Esse resultado evidencia que as reanálises do ECMWF são superiores à análise do NCEP, independente do modelo considerado. Testes de significância estatística indicaram que não são diferentes as médias das previsões de ambos os modelos.

As séries temporais da ACC das previsões de 120h (5 dias) são mostradas na Figura 4.5 para as regiões HN (painel superior), HS (painel inferior) e TRP (painel central) comparando os modelos MPAS e SHiELD. Pode-se notar que ambos modelos apresentam pouca variação sazonal ao longo do período. Por outro lado, como esperado, encontrou-se maior correlação média nas regiões de médias e altas latitudes (HN e HS) em comparação aos trópicos (TRP). Entre os modelos, as diferenças médias são mínimas no HN e HS e o SHiELD é ligeiramente superior nos trópicos. Também, as análises das séries temporais das previsões para além de cinco dias mostraram comportamento semelhante entre ambos modelos com valores menores de ACC, mas em geral não diferem segundo a análise estatística de significância ao nível de 95

4.4 Análise Multivariada de Mahalanobis

Na Figura 4.6, são apresentadas as médias dos resultados obtidos da análise multivariada de MH para o período completo de simulação.

Em 24h (Figura 4.6a) observa-se valores baixos ($< 1,5$) de MH em quase todo o globo tanto para o MPAS como para o FV3. Este padrão só é rompido sobre regiões de mais altas latitudes e sobre regiões continentais elevadas. Nesta última, é evidente a diferença entre o MPAS e o FV3, com o último mostrando menor distância de MH. A diferença entre ambos mostra que o MH do MPAS é maior sobre os Andes, as montanhas Rochosas, sobre o platô tibetano, no sul da África, Antártida e no Mar de Weddell. Esta diferença pode ser relacionada a processos físicos não satisfatoriamente resolvidos em regiões de topografia elevada.

Em 48h (Figura 4.6b), a distância MH tende a aumentar nas regiões polares, principalmente sobre o continente Antártico e no oceano Ártico. O valor MH também aumenta rapidamente nos extratrópicos sobre os oceanos. Este aumento ocorre em ambos os modelos, sendo que a diferença entre eles se mantém constante, apesar do valor médio do MH (0,01) ter subido.

O valor de MH tende a aumentar nas horas seguintes de integração (Figuras 4.6c e 4.6e), a partir das altas latitudes em ambos os modelos, mas em especial sobre os oceanos. O MH global médio cresce rapidamente, com valor médio de 3,987 para o MPAS e 3,722 para o FV3 (Figura 4.6e). O campo DifMH também mostrou um aumento, com o MPAS se distanciando mais do campo de observação em relação ao FV3.

Na Figura 4.7 apresenta-se o comportamento dos modelos na estação de verão do hemisfério austral. Em 24h (Figura 4.7a), observa-se uma grande similaridade em ambos os modelos. Porém, com um valor médio de MH ligeiramente menor para o MPAS. No campo de diferença, observa um melhor desempenho do MPAS em relação ao FV3 sobre a Bacia Amazônica Ocidental e na região de formação de CCM no norte da Argentina e Paraguai. O DifMH apresentou um valor negativo (-0,009), ou seja, o MPAS apresentou distância de MH ligeiramente melhor que o FV3.

Em 48h (Figura 4.7b) há um aumento do valor de MH em ambos os modelos, mas em especial no MPAS sobre o Atlântico Sul. Nas regiões tropicais não se observa uma diferença expressiva entre os modelos. Porém, no campo DifMH se observa que

o MPAS se distanciou menos no campo original sobre a Bacia Amazônica e parte do Sudeste e Centro-Oeste e sobre a Argentina. Em 72h e 120h (Figuras 4.7c e 4.7d) há um aumento do MH em ambos os modelos, em especial nos extratrópicos. Porém, com o MPAS se distanciando mais lentamente do campo original sobre grande parte do território brasileiro, onde a DifMH tende a ser negativa. No oceano, o distanciamento tende a ser mais aleatório, com o MPAS se distanciando mais sobre o Atlântico Sul. Em 240h (Figuras 4.7e e Figura 4.7f), observa-se o aumento expressivo do MH nos extratrópicos conforme já observado no campo global. A diferença está no aumento expressivo do MH sobre o Brasil central em ambos os modelos. Porém, com o MPAS mostrando um distanciamento menor em relação ao FV3 sobre grande parte do território brasileiro e maior sobre os oceanos subtropicais e extratropicais.

Nos meses de inverno austral para 24h (Figura 4.8a), observa-se um comportamento similar em ambos os modelos, com o MH de 1,058 apesar do maior distanciamento do MPAS sobre os Andes. Em termos de DifMH, o MPAS mostrou um menor distanciamento, o que é visível no predomínio de valores negativos sobre o todo o campo de diferença. Em 48h (Figura 4.8b), há um pequeno aumento do MH, forçado pelo aumento nos extratrópicos. Porém o DifMH se manteve negativamente favorável ao MPAS, em especial sobre o território brasileiro, embora tenha apresentado maior distância no Sul do Brasil. Em 72h (Figura 4.8c), o MH aumenta nos extratrópicos, com uma DifMH positiva favorável ao FV3 no Sul e Sudeste do Brasil, e negativa favorável ao MPAS no centro-norte do Brasil. Em 120h (Figura 4.8d), a distância de MH aumenta, porém a DifMH tende a ser negativa e favorável ao MPAS sobre grande parte da América do Sul. Para 240h (Figura 4.8e), há um aumento do valor de MH, principalmente nos subtrópicos e extratrópicos, porém com o FV3 apresentando uma distância menor que o MPAS no geral (DifMH média de 0,042).

Em resumo, o MPAS mostrou ter um MH ligeiramente maior que o FV3, tanto no campo global quanto no regional sobre a América do Sul. O MPAS também mostrou um maior distanciamento sobre as regiões de topografia elevada (Andes e Platô do Tibete) e sobre a Antártida, Groelândia e Oceano Ártico. Porém, o MPAS mostrou ter um MH menor que o FV3 sobre a América do Sul, em especial sobre a Bacia Amazônica, tanto no verão astral quanto no inverno, com alguns momentos em que este menor distanciamento foi dominante sobre todo o continente sul americano.

5 Considerações Finais

O modelo GEF foi avaliado somente no quesido Adequação. Concluiu-se que este modelo não pode ser aplicado, a curto prazo, para as ações a que se destina o MONAN devido ao seu núcleo dinâmico ser hidrostático. Desta forma, a avaliação da acurácia e dos demais quesitos não foi realizada.

Os modelos avaliados nos demais aspectos compõem o estado-da-arte em modelagem dos processos físicos e dinâmicos da atmosfera. Por meio da análise de diversas métricas relacionadas às características de funcionalidade de software, foi possível identificar que ambos os modelos são amplamente utilizados pela comunidade americana de meteorologia para diversas aplicações. Os modelos são aptos a serem acoplados a outros componentes do sistema terrestre e possuem funcionalidades que permitem que sejam executados fazendo uso de um sistema de assimilação de dados. O SHIELD leva vantagem por ser um modelo cujo núcleo dinâmico é utilizado em modo operacional nos Estados Unidos, e diversas funcionalidades já são preparadas para o ambiente operacional. Por outro lado, a interação com a equipe de desenvolvimento do MPAS é um ponto positivo e importante a se levar em consideração, uma vez que facilita o intercâmbio do conhecimento e resolução de problemas de aspecto técnico no uso do modelo em ambiente operacional.

Em relação à avaliação de acurácia dos modelos, destaca-se o que segue.

A partir da análise da distância de MH, o MPAS mostrou ter um MH ligeiramente maior que o FV3, tanto no campo global quanto no regional sobre a América do Sul. Por outro lado, o MPAS mostrou um maior distanciamento sobre as regiões de topografia elevada (Andes e Platô do Tibete) e sobre a Antártida, Groelândia e Oceano Ártico. O MPAS apresentou MH menor que o FV3 sobre a América do Sul, em especial sobre a Bacia Amazônica, tanto no verão quanto no inverno, sendo que este menor distanciamento foi dominante sobre todo o continente sul americano.

Em relação aos índices estatísticos das previsões de precipitação, identificou-se que o MPAS apresentou melhor desempenho, de modo geral, para limiares de precipitação de até 10mm, enquanto que o SHIELD é ligeiramente superior em termos de frequência de ocorrência do total de previsões de precipitação para limiares de até 35mm, apesar de apresentar maior ocorrência de falsos alarmes. Entretanto, uma característica fundamental é a tendência do SHIELD de superestimar os eventos de chuva, e consequentemente inflar os valores de probabilidade de detecção, o que pode induzir a uma interpretação equivocada de seu desempenho. Uma vez que o SHIELD

ELD apresentou desempenho inferior ao MPAS no índice de razão de sucesso, que é interpretado como a razão entre as previsões corretas em relação ao número total de eventos previstos (inclusive os falsos alarmes), pode-se interpretar que, ainda que gere subestimativa da frequência de eventos totais, o MPAS é capaz de prever com maior confiabilidade os eventos de precipitação que realmente ocorrem, para todos os limiares de precipitação.

As análises da destreza da altura geopotencial indicaram que ambos os modelos mostram um limite de previsibilidade de até 8 dias para o domínio global nos Hemisférios Norte e Sul. O uso de diferentes condições iniciais para o processo de inicialização do FV3-SHiELD indicou ligeiras diferenças e desempenho superior das reanálises do ERA5, resultado este esperado. As reanálises do ERA5 são o estado-da-arte em termos de representação do estado da atmosfera em períodos passados. Ela baseia-se na abordagem de assimilação de dados do ECMWF, combinando um conjunto extremamente grande de observações com o modelo de previsão meteorológica. Com esta metodologia, o ERA5 fornece uma das reconstruções mais precisas do clima passado.

As análises apresentadas neste documento focaram nos aspectos de funcionalidade de software, na avaliação objetiva da precipitação e na análise multivariada de um conjunto de variáveis atmosféricas. Entretanto, ressalta-se que ainda é necessário realizar análises objetivas de outras variáveis meteorológicas e estudos de casos para ampliar o conhecimento sobre os erros dos modelos avaliados. Esta etapa deverá ser realizada e publicada em um novo relatório técnico dedicado a este fim.

6 Contribuição dos autores

Ariane Frassoni: Concebeu e projetou a análise; desenvolveu e aplicou a avaliação da característica de funcionalidade de software e suas subcaracterísticas; realizou a análise da avaliação quantitativa de precipitação; escreveu o relatório.

Bárbara A. G. P. Yamada: preparou o modelo de publicação do INPE; organizou e formatou as tabelas do relatório; realizou a revisão final do relatório.

José Roberto Rozante: desenvolveu e aplicou as ferramentas de avaliação quantitativa de precipitação; contribuiu para a avaliação quantitativa de precipitação.

Julio Pablo Fernandez: preparou e executou os modelos numéricos; organizou os dados dos modelos numéricos; realizou a análise da avaliação quantitativa de altura

geopotencial.

Marcelo Bario Rosa: desenvolveu e aplicou a técnica de análise multivariada de Mahalanobis; realizou a análise multivariada de Mahalanobis.

REFERÊNCIAS BIBLIOGRÁFICAS

Coupled atmosphere-ocean modelling. In: CHASSIGNET, E.; PASCUAL, A.; TINTORÉ, J.; VERRON, J. (Ed.). **New Frontiers In Operational Oceanography**. Florida, USA: CreateSpace Independent Publishing Platform, 2018. (Florida State University Libraries), cap. 16, p. 445–463. ISBN –10 1720549974. Disponível em:
https://www.godae.org/godae-data/School/Chapter16_Harris_et_al.pdf. 6

FIGUEROA, S. N.; BONATTI, J. P.; KUBOTA, P. Y.; GRELL, G. A.; MORRISON, H.; BARROS, S. R. M.; FERNANDEZ, J. P. R.; RAMIREZ, E.; SIQUEIRA, L.; LUZIA, G.; SILVA, J.; SILVA, J. R.; PENDHARKAR, J.; CAPISTRANO, V. B.; ALVIM, D. S.; ENORÉ, D. P.; DINIZ, F. L. R.; SATYAMURTI, P.; CAVALCANTI, I. F. A.; NOBRE, P.; BARBOSA, H. M. J.; MENDES, C. L.; PANETTA, J. The brazilian global atmospheric model (bam): Performance for tropical rainfall forecasting and sensitivity to convective scheme and horizontal resolution. **Weather and Forecasting**, American Meteorological Society, Boston MA, USA, v. 31, n. 5, 2016. 13

LATINOVIC, D.; CHOU, S. C.; RANCIC, M. Seasonal range test run with global eta framework. **Adv. Sci. Res.**, v. 14, p. 247–251, 2017.
[Https://doi.org/10.5194/asr-14-247-2017](https://doi.org/10.5194/asr-14-247-2017). 6

OMM, O. M. M. **Manual on the Global Data-processing and Forecasting System: Annex IV to the WMO Technical Regulations**. Geneva, Switzerland: [s.n.], 2019. Updated in 2022. Disponível em: <<<https://wmo.int/activities/wmo-integrated-processing-and-prediction-system-wipps>>>. 3, 13

OMM, W. M. O.; WWRP, W. W. R. P. **Recommendations for the Verification and Intercomparison of QPFs and PQPFs from Operational NWP Models**. Geneva, Switzerland: [s.n.], October 2009. Revision 2, updated in 2008. Disponível em:

<<<https://filecloud.wmo.int/share/s/03mFumNQRpeMRYSDZusmkA>>>. 13

PERSSON, A. **User Guide to ECMWF forecast products**. 2001. Online. Disponível em: <<https://ghrc.nsstc.nasa.gov/uso/ds_docs/tcsp/tcsp_ecmwf/ECMWFUserGuideofForecastProducts32.pdf>>. 23

PIELKE, R. **Mesoscale meteorological modeling**. Orlando, FL.: Academic Press, 1984. 7

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

ROEBBER, P. Visualizing multiple measures of forecast quality. **Wea. Forecasting**, v. 24, p. 601–608, 2009. [14](#)

ROZANTE, J. R.; GUTIERREZ, E. R.; FERNANDES, A. de A.; VILA, D. A. Performance of precipitation products obtained from combinations of satellite and surface observations. **International Journal of Remote Sensing**, Taylor & Francis, v. 41, n. 19, p. 7585–7604, 2020. [11, 20](#)

SKAMAROCK, W. C.; KLEMP, J. B.; DUDA, M. G.; AL. et. A multiscale nonhydrostatic atmospheric model using centroidal voronoi tesselations and c-grid staggering. **Monthly Weather Review**, v. 140, n. 9, p. 3090–3105, 2012. <Https://doi.org/10.1175/MWR-D-11-00215.1>. [6](#)

STEPPELER, J.; HESS, R.; SCHÄTTLER, U. e. a. Review of numerical methods for nonhydrostatic weather prediction models. **Meteorol Atmos Phys**, v. 82, p. 287–301, 2003. <Https://doi.org/10.1007/s00703-001-0593-8>. [7](#)

STEVENS, B.; SATOH, M.; AUGER, L. e. a. Dyamond: the dynamics of the atmospheric general circulation modeled on non-hydrostatic domains. **Prog Earth Planet Sci**, v. 6, 2019. <Https://doi.org/10.1186/s40645-019-0304-z>. [7](#)

TAYLOR, K. Summarizing multiple aspects of model performance in a single diagram. **J. Geophys. Res.**, v. 106 (D7), p. 7183–7192, 2001. [14](#)

WGNE. 2018. Disponível em: <https://community.wmo.int/activity-areas/wwrp/wwrp-working-groups/wwrp-forecast-verification-research>. Acesso em: 08/06/2018. [3](#)

_____. 2022. Disponível em:
http://wgne.meteoinfo.ru/wp-content/uploads/2021/11/wgne_table_2021_v4.xls.
Acesso em: 08/06/2022. [7](#)

WIKIPEDIA. 2022. Https://pt.wikipedia.org/wiki/ISO/IEC_9126. Acesso em 07/06/2022. [3, 4, 5](#)

ZHANG, H.; RANČIĆ, M. A global eta model on quasi-uniform grids. **Quarterly Journal of the Royal Meteorological Society**, v. 133, n. 623, p. 517–528, 2007. Disponível em:
[<<Https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.17>>](Https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.17). [6](#)

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o International Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

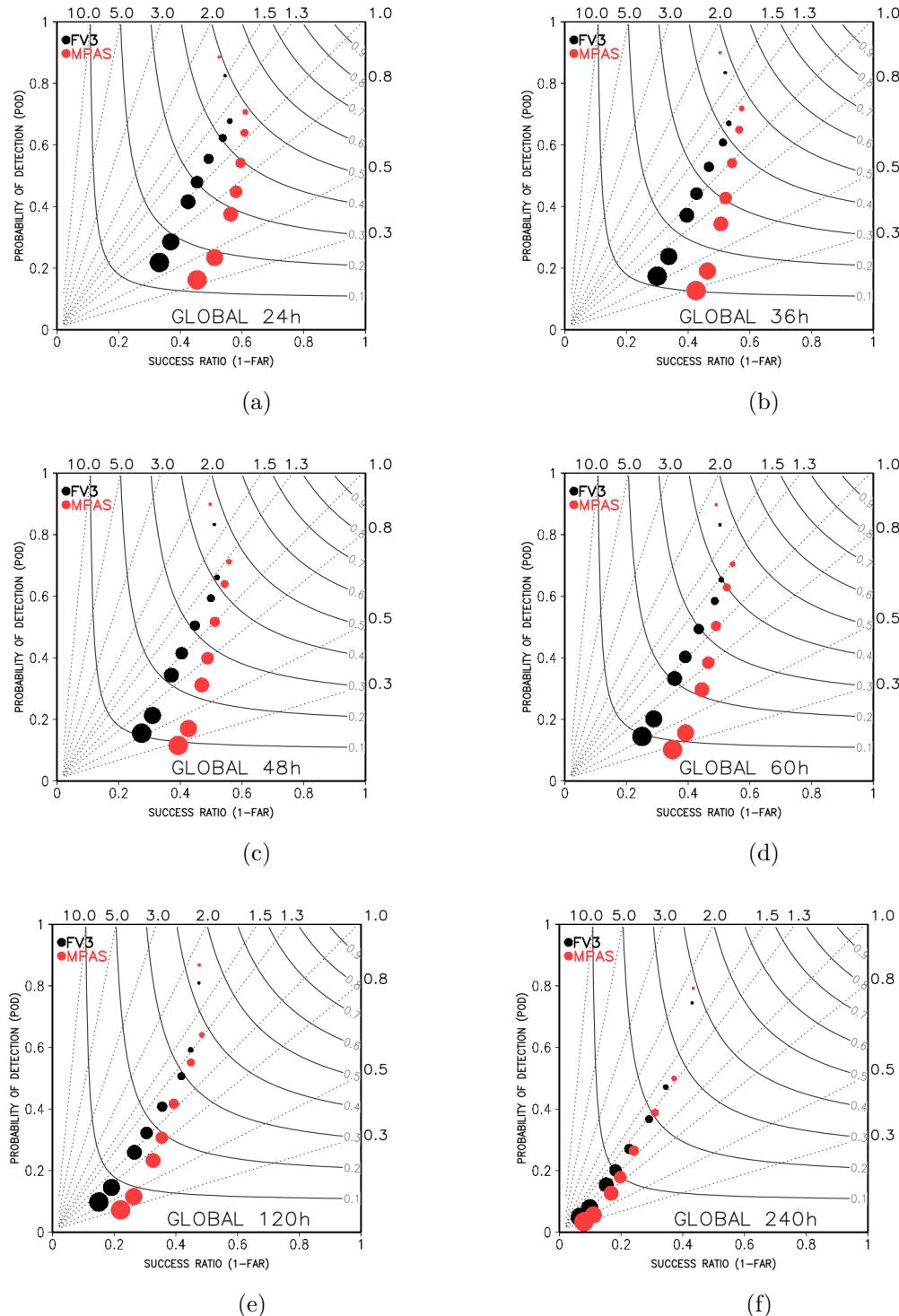
Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Programas de Computador (PDC)

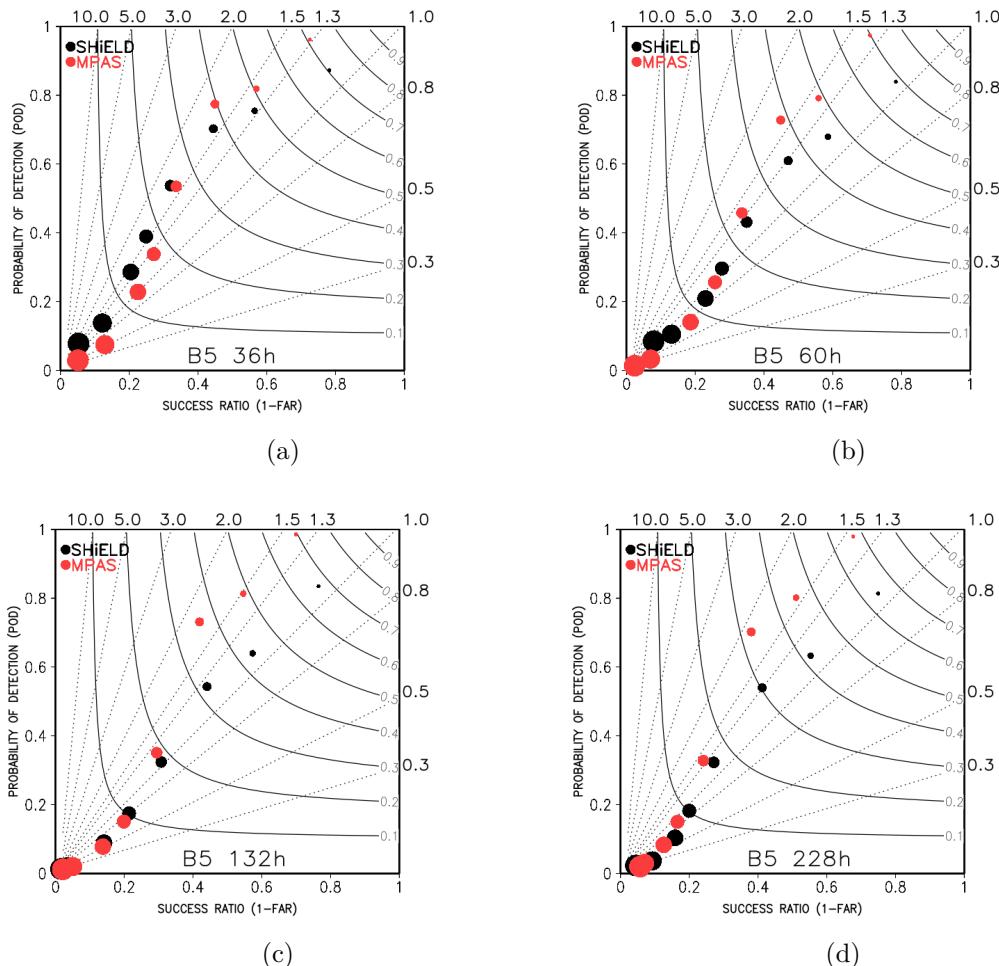
São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Figura 4.1 - Diagrama de desempenho para as previsões de precipitação de a) 24h b) 36h, c) 48h, d) 60h, e) 120h (132h) e f) 240h. As esferas representam os diferentes limiares de precipitação: 0,5mm, 2mm, 5mm, 10mm, 15mm, 20mm, 35mm, sendo a cor vermelha representativa do modelo MPAS e preto, do modelo SHiELD.



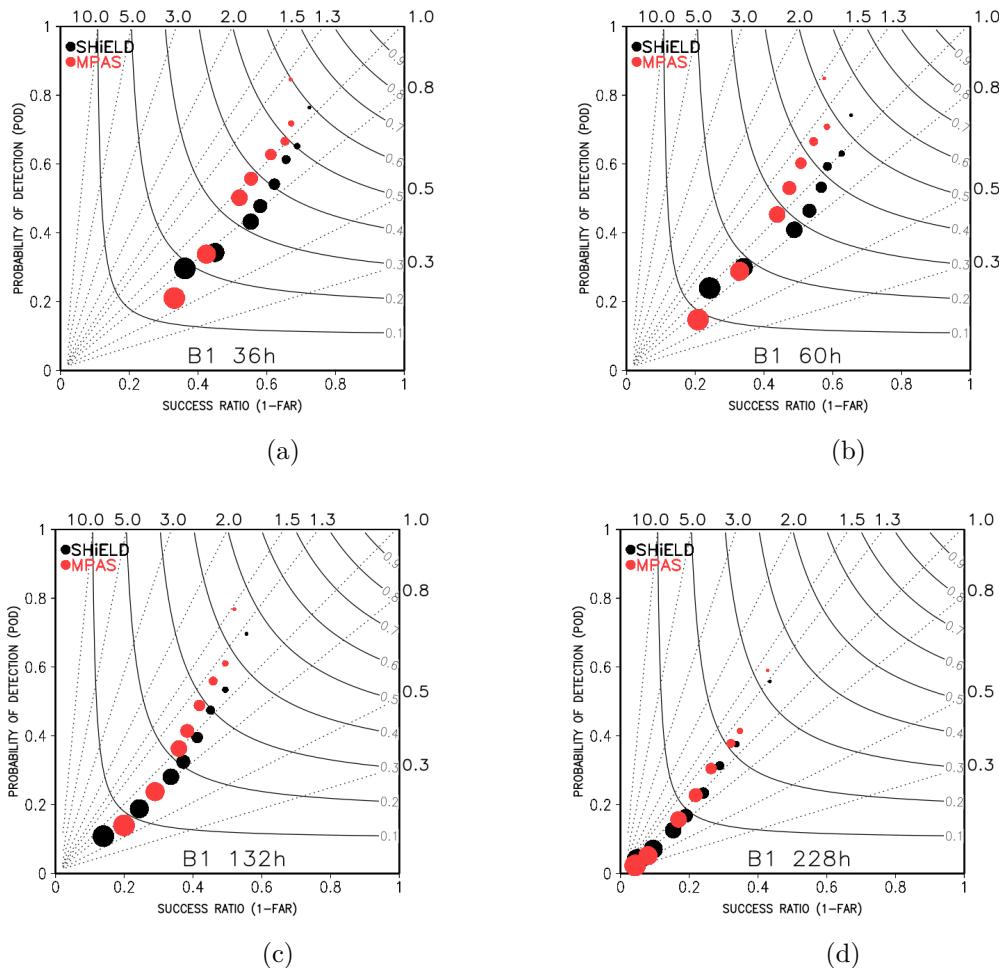
Fonte: Produção própria.

Figura 4.2 - Idem à Figura 4.1, exceto para a área B5



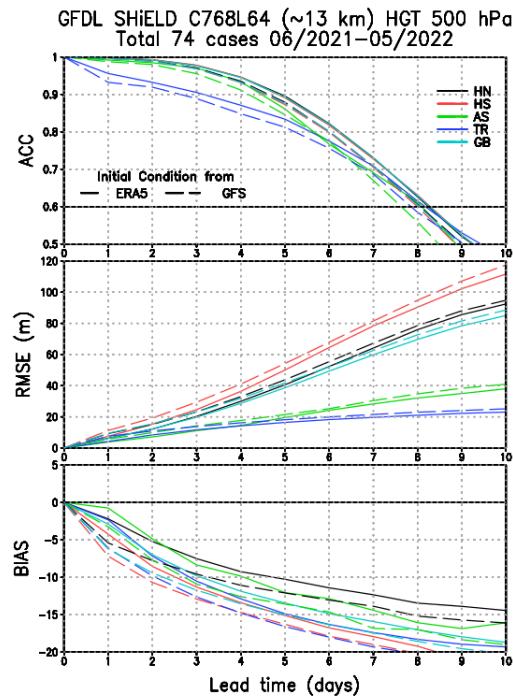
Fonte: Produção própria.

Figura 4.3 - Idem à Figura 4.1, exceto para a área B1

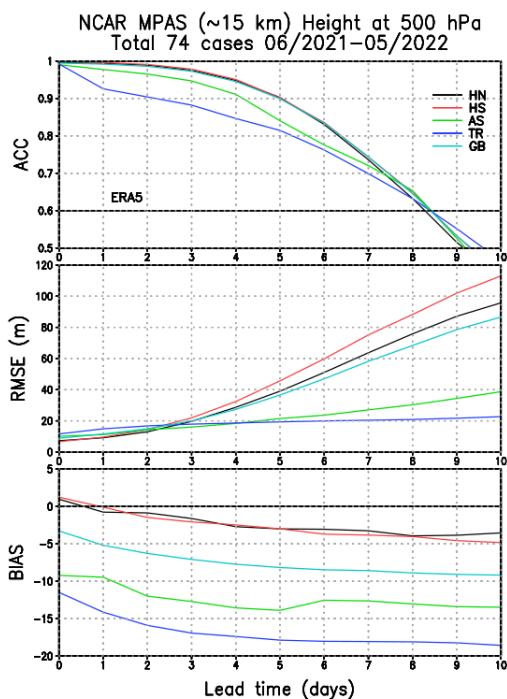


Fonte: Produção própria.

Figura 4.4 - Métricas das sub-características da Manutenibilidade com os modelos FV3, GEF e MPAS.



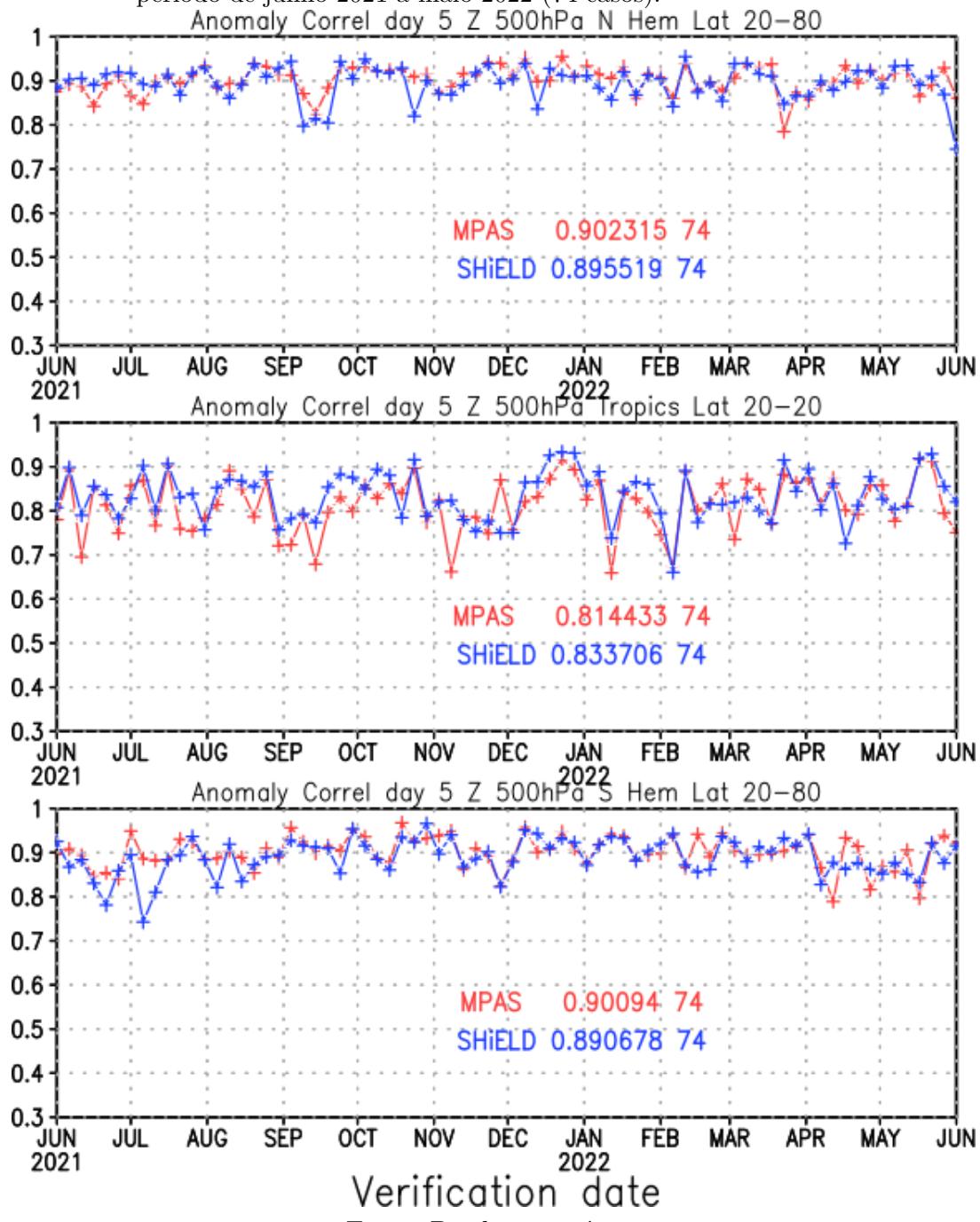
(a)



(b)

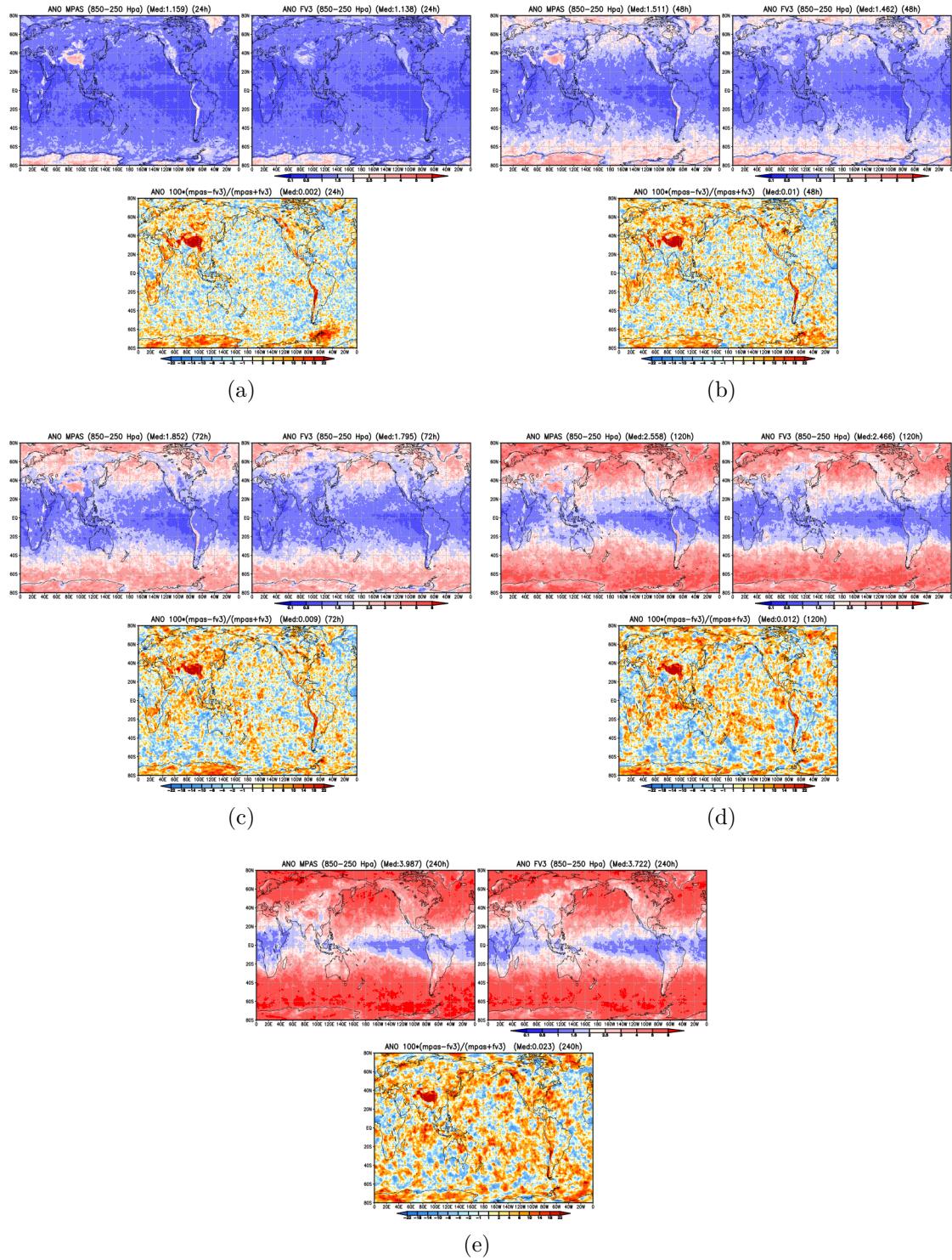
Fonte: Produção própria.

Figura 4.5 - Séries temporais da ACC do geopotencial a 500 hPa dos modelos MPAS (vermelho) e SHIELD (azul), para o Hemisfério Norte (painel superior), trópicos (médio) e Hemisfério Sul (painel inferior). Previsões de 120 h (5 dias) para o período de junho 2021 a maio 2022 (74 casos).



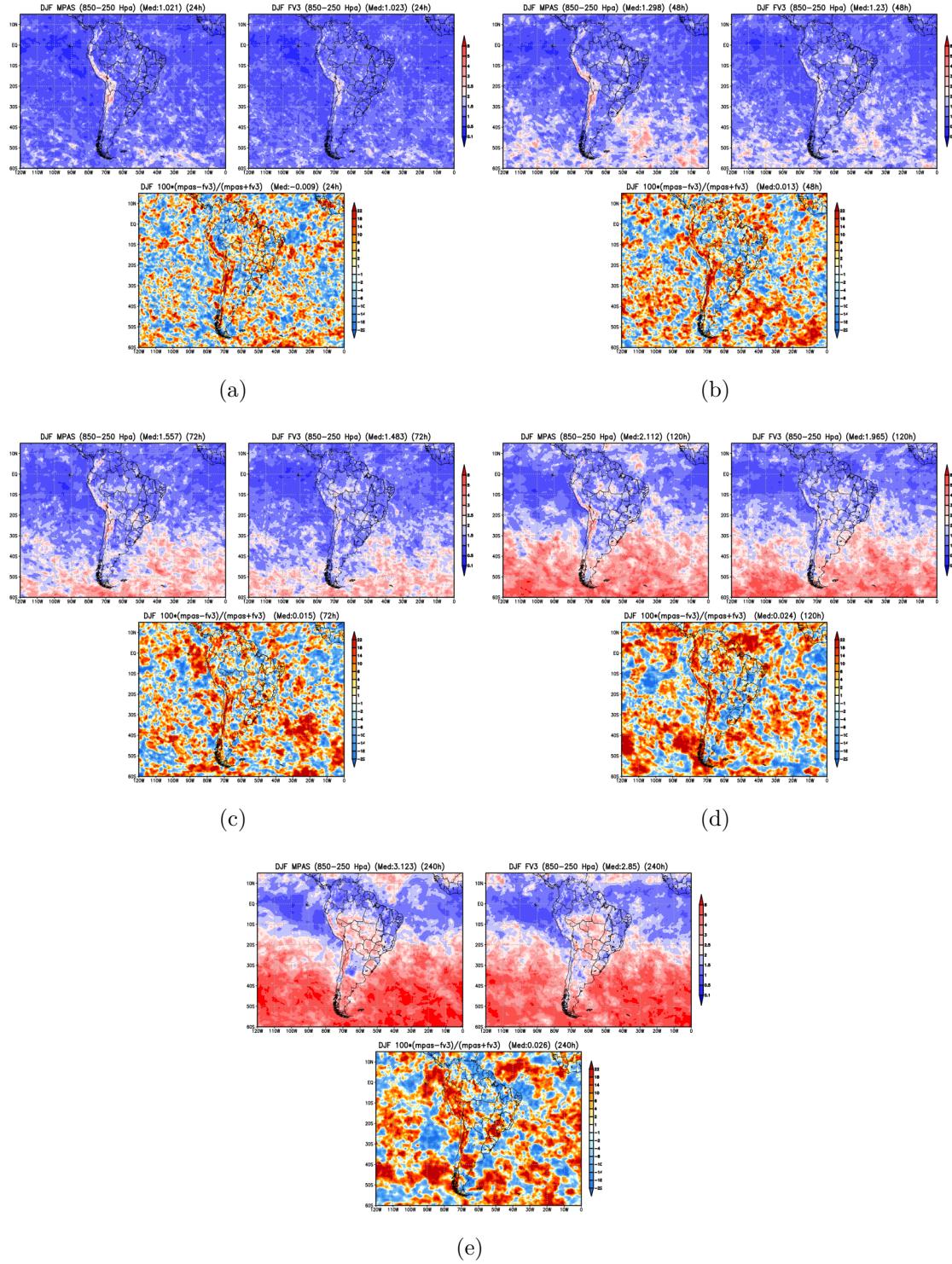
Fonte: Produção própria.

Figura 4.6 - Distância de MH para o MPAS e FV3 e diferença normalizada (DifMH) para os prazos de previsão de 24h (a), 48h (b), 72h (c), 120h (d) e 240h (e)



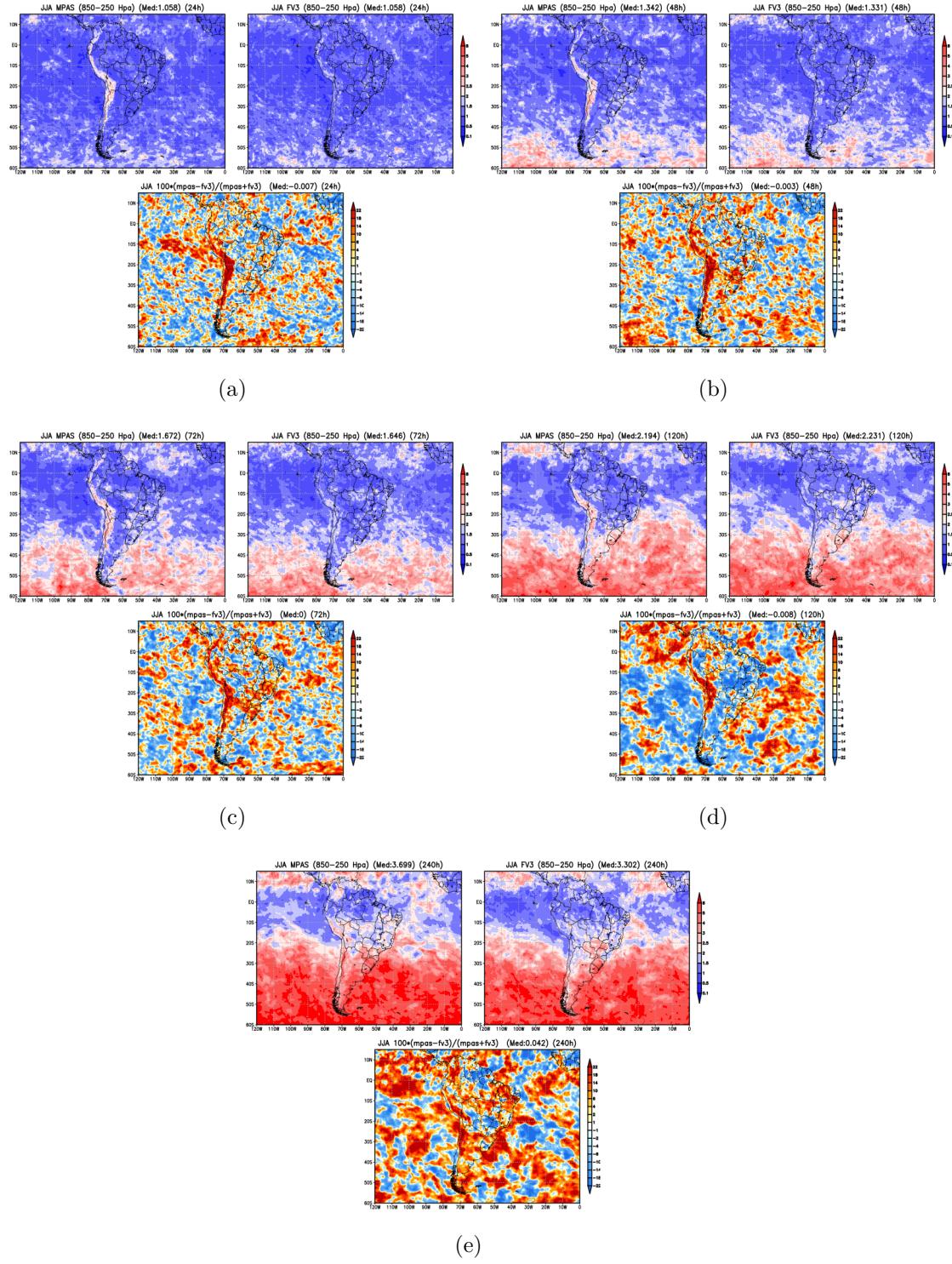
Fonte: Produção própria.

Figura 4.7 - Idem que a Figura 4.6 mas para os meses de verão austral sobre a América do Sul.



Fonte: Produção própria.

Figura 4.8 - Idem que a Figura 4.7 mas para os meses de inverno austral sobre a América do Sul.



Fonte: Produção própria.