



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 INTELLECTUAL DO INPE - CEPPII (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).

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Áreas para avaliação da precipitação sobre a América do Sul	10
3.1 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.	28
3.2 Idem à Figura 3.1, exceto para a área B5	29
3.3 Idem à Figura 3.1, exceto para a área B1	30
3.4 Métricas das sub-características da Manutenibilidade com os modelos FV3, GEF e MPAS.	31
3.5 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).	32
3.6 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)	33
3.7 Idem que a Figura 3.6 mas para os meses de verão austral sobre a América do Sul.	34
3.8 Idem que a Figura 3.7 mas para os meses de inverno austral sobre a América do Sul.	35

LISTA DE TABELAS

	<u>Pág.</u>
1.1 Subcaracterísticas de Funcionalidade (pergunta chave: Satisfaz a necessidade do projeto MONAN?).	2
2.1 Características dos modelos e das integrações numéricas realizadas. <i>wall-clock time</i> significa o tempo de processamento para 24h de previsão. . . .	8
2.2 Parametrizações físicas dos modelos SHiELD e MPAS	9
3.1 Pontuações atribuídas às métricas relacionadas a opções de funcionalidades.	14
3.2 Pontuações atribuídas às métricas relacionadas a opções de funcionalidades.	15
3.3 Pontuações atribuídas às métricas relacionadas a opções de funcionalidades.	15
3.4 Pontuação total obtida a partir das métricas de avaliação das subcaracterísticas de funcionalidade de software	15
3.5 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.	17
3.6 Idem à Figura 3.5, exceto para o limiar menor ou igual a 35mm.	18

SUMÁRIO

	<u>Pág.</u>
1 Introdução	1
1.1 Análise de Funcionalidade: subcaracterísticas	2
2 Metodologia	5
2.1 Métricas de avaliação das subcaracterísticas	5
2.1.1 Avaliação da acurácia dos modelos globais	7
2.1.1.1 Desenho dos experimentos	7
2.1.1.2 Dados de referência	8
2.1.1.3 Avaliação estatística objetiva	9
2.1.1.4 Análise multivariada – Distância de Mahalanobis	11

3 Resultados	13
3.1 Pontuação das métricas de avaliação das subcaracterísticas	14
3.2 Avaliação estatística da precipitação	16
3.3 Avaliação estatística da altura geopotencial	19
3.4 Análise Multivariada de MH	20
4 Considerações Finais	22
5 Contribuição dos autores	24
REFERÊNCIAS BIBLIOGRÁFICAS	24

1 Introdução

O presente documento tem por objetivo apresentar a primeira etapa do processo de avaliação de funcionalidade de software dos modelos numéricos de escala global visando a escolha da componente atmosférica do *Model for Ocean-land-Atmosphere prediction* (MONAN).

Neste trabalho, foram considerados 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, o Grupo de Avaliação de Modelos da Divisão de Modelagem Numérica do Sistema Terrestre (GAM-DIMNT) está preocupado com os aspectos de 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 de 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 1.1, na qual são associadas a uma pergunta chave para melhor entendimento do leitor.

Tabela 1.1 - 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).

Dentre as subcaracterísticas listadas, o GAM-DIMNT se propôs a elaborar e aplicar métricas ou utilizar ferramentas existentes para as três primeiras, a saber: Adequação, Acurácia e Interoperabilidade, além das subcaracterísticas de Usabilidade, Inteligibilidade, Operacionalidade e Capacidade de Colaboração.

1.1 Análise de Funcionalidade: subcaracterísticas

A subcaracterística Adequação é a capacidade do produto de software de prover um conjunto apropriado de funções para tarefas e objetivos do usuário especificados. O MONAN é um modelo comunitário, desenvolvido com apoio de diversas instituições, sob a liderança do INPE. O MONAN envolve 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 modelo 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. Portanto, a subcaracterística Adequação deverá indicar qual núcleo dinâmico tem a capacidade de compor o modelo MONAN, visando cumprir a necessidade de atender às escalas global, regional e futuramente, a escala local.

Uma das funções requeridas para o núcleo dinâmico do MONAN é que este seja não-hidrostático. 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. Os núcleos dinâmicos que não possuíam esta funcionalidade foram considerados inaptos para as aplicações do modelo MONAN e não foram avaliados no quesito Acurácia, especialmente considerando que esta subcaracterística demanda recurso computacional.

A subcaracterística Interoperabilidade indica a capacidade do produto de software de interagir com um ou mais sistemas especificados. No caso do MONAN, espera-se que o núcleo dinâmico escolhido esteja preparado para interagir com outras componentes e/ou aplicações do Sistema Terrestre bem como com o sistema de Assimilação de Dados.

Uma das componentes essenciais para a modelagem do Sistema Terrestre é o modelo oceano-criosfera para uso em diferentes escalas de tempo e espaço. Os feedbacks entre os campos acoplados, tais como a temperatura da superfície do mar e correntes e a atmosfera, assim como fluxos de calor e momento da atmosfera para o oceano, em geral são realizados pelo uso de um código de acoplamento separado, flexibilizando vincular componentes de modelos e controlar a troca e interpolação dos campos acoplados (CHASSIGNET et al., 2018).

Outra importante componente de um sistema de modelagem numérica do Sistema Terrestre é o sistema de Assimilação de Dados, essencial para fornecer as condi-

ções iniciais para um modelo numérico a partir do uso de observações obtidas de diferentes fontes. Assimilação de Dados é também fundamental para a produção de previsões de tempo e clima por conjunto. Para explicar a natureza caótica da atmosfera e a incerteza associada à previsão, centros de meteorologia avançados utilizam um conjunto de previsões simultaneamente. A previsão por conjunto é gerada a partir do uso do melhor estado inicial possível da atmosfera adicionalmente a um conjunto de outras previsões com pequenas variações em relação ao estado inicial, limitado ao recurso computacional disponibilizado no centro de previsão. Da mesma forma, outras componentes do Sistema Terrestre passam a se tornar importantes à medida que a escala horizontal de interesse é alterada. É o caso de previsões meteorológicas em escala urbana, que requerem que o modelo seja capaz de parametrizar características urbanas, como ruas, prédios, vegetação, geração de calor antrópico entre outros.

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.

Inteligibilidade se traduz na 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. Tem-se também a subcaracterística de Apreensibilidade, que é a capacidade do produto de software de possibilitar ao usuário aprender sua aplicação.

Tendo em vista que o Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais (INPE) é um centro operacional de previsão numérica, a subcaracterística Operacionalidade é fundamental, e indica que o produto de software deve possibilitar ao usuário operá-lo e controlá-lo, mantendo a proteção frente a erros de usuários.

Por fim, a subcaracterística Capacidade de colaboração é interpretada como a interação e colaboração com os centros desenvolvedores do núcleo dinâmico em análise.

As referidas subcaracterísticas foram avaliadas por meio da análise de um conjunto de métricas desenvolvidas a partir da análise de características esperadas dos mode-

los. 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.

2 Metodologia

Os modelos globais avaliados na característica Funcionalidade de software foram o *Finite-Volume Cubed-Sphere Dynamical Core* (FV3) SHIELD (CHASSIGNET et al., 2018), versão 202210, o *Model for Prediction Across Scales* (MPAS) (SKAMAROCK et al., 2012), versão 6.3 e o modelo *Global Eta Framework* (GEF) (LATINOVIC et al., 2017). Na Seção 2.1 serão apresentadas as métricas utilizadas para a avaliação dos referidos modelos.

2.1 Métricas de avaliação das subcaracterísticas

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 urdano

- 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 possuísem núcleo dinâmico não-hidrostático seriam eliminados das demais etapas de avaliação.

Ao aplicar a pontuação para a métrica “a”, 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 da avaliação de acurácia.

2.1.1 Avaliação da acurácia dos modelos globais

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.

2.1.1.1 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*. 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 2.1.

Tabela 2.1 - Características dos modelos e das integrações numéricas realizadas. *wall-clock time* significa o 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

As parametrizações físicas dos modelos SHIELD e MPAS utilizadas nos experimentos são listadas na Tabela 2.2.

2.1.1.2 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*

Tabela 2.2 - 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

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.

2.1.1.3 Avaliação estatística objetiva

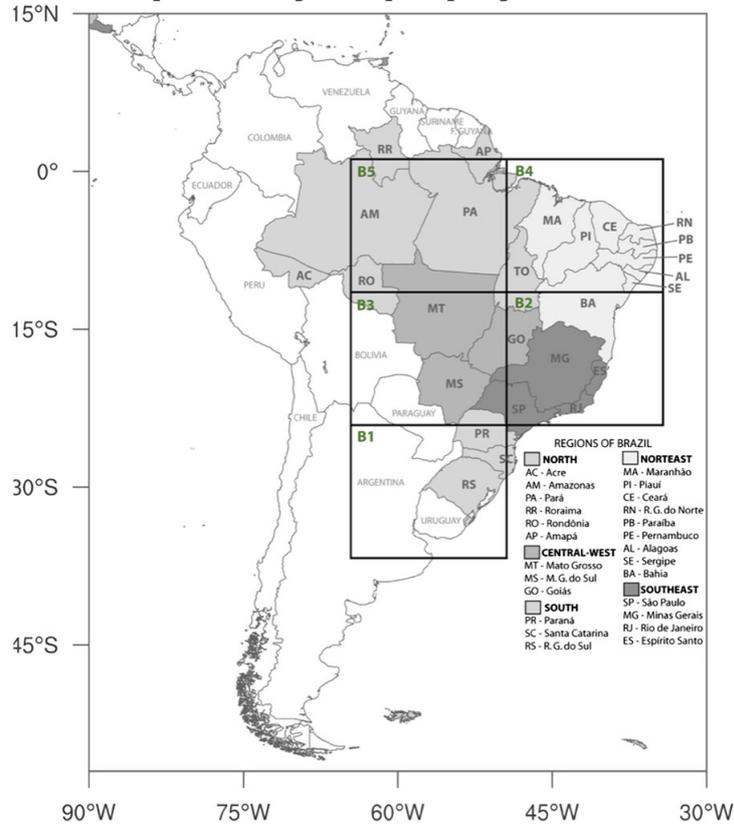
A acurácia dos modelos globais foi analisada por meio de avaliações estatísticas objetivas da precipitação e altura geopotencial, na escala de tempo, em escala global e sobre a América do Sul.

Foram consideradas diferentes áreas geográficas do globo e subdomínios do continente sul-americano. A avaliação de subdomínios da América do Sul foi realizada conforme Figueroa et al. (2016) (Figura 2.1).

A avaliação em escala global e subdomínios tem como base as recomendações da Organização Meteorológica Mundial para o *Global Data-processing and Forecasting System* (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)

Figura 2.1 - Áreas para avaliação da precipitação sobre a América do Sul



Fonte: Figueroa et al. (2016)

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.

2.1.1.4 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 distânciam de outro conjunto, dando 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:

$$A_1(l, p) = X_1(l, p) - \mu_1(l) \quad (2.1)$$

e

$$A_2(l, p) = X_2(l, p) - \mu_2(l). \quad (2.2)$$

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 2x2 e se forem três, 3x3 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:

$$Cov(l, l) = \frac{[Cov_1(l, l) \times Som_1(l, l)] + [Cov_2(l, l) \times Som_2(l, l)]}{Som_1(l, l) + Som_2(l, l)}. \quad (2.3)$$

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:

$$MH = \sqrt{\sum [\mu_1(l) - \mu_2(l)] * [\mu_1(l) - \mu_2(l)]^T * Cov(l, l)^{-1}}. \quad (2.4)$$

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:

$$DifMH = \frac{MH(MPAS) - MH(FV3)}{MH(MPAS) + MH(FV3)} * 100, \quad (2.5)$$

em que $DifMH < 0$ implica que o MPAS apresenta menor distância se comparado ao FV3.

3 Resultados

Nesta Seção, são apresentados os resultados obtidos com a aplicação das pontuações estabelecidas na norma ISO, com aplicação da análise multivariada e os resultados da avaliação estatística dos campos de precipitação e altura geopotencial.

3.1 Pontuação das métricas de avaliação das subcaracterísticas

Um dos grandes diferenciais entre os modelos MPAS e FV3, reside na métrica “b” apresentada na Seção 2.1, que dispõe sobre interação com grupo oficial do modelo. As interações com colaboradores do NCAR, desenvolvedores do MPAS, foram consideradas mais acessíveis que as interações com os colaboradores do GFLD, 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 2.1 são indicadas nas Tabelas 3.1 – 3.3. As pontuações recebidas pelos modelos globais indicam que o MPAS e SHiELD atendem aos requisitos formulados, com pontuação total similar.

Tabela 3.1 - Pontuações atribuídas às métricas relacionadas a opções de funcionalidades.

Métricas de avaliação

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 de 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 RMRSP, 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

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 3.4.

Tabela 3.2 - Pontuações atribuídas às métricas relacionadas a opções de funcionalidades.

Métricas de avaliação

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 l= 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 sazonal 7. Previsão sazonal 8. Projeções climáticas 9. Eventos extremos/raros	Aplicações 1- https://www.mdpi.com/2073-4433/13/4/591 2- https://gmd.copernicus.org/preprints/gmd-2019-298/gmd-2019-298.pdf 4. https://repository.library.noaa.gov/view/noaa/28398/noaa_28398_DS1.pdf 7 https://link.springer.com/article/10.1007/s00382-021-05797-7 https://agupubs.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

Tabela 3.3 - Pontuações atribuídas às métricas relacionadas a opções de funcionalidades.

Métricas de avaliação

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 sug-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

Tabela 3.4 - Pontuação total obtida a partir das métricas de avaliação das subcaracterísticas de funcionalidade de software

Modelo	Pontuação total
SHIELD	70
MPAS	72

3.2 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 3.1 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 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 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 liminares 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

Tabela 3.5 - 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	↓	↑	↑	> 1	↑	↓	↓
48	≈ 1	↑	↑	↑	> 1	↓	↓	↓
72	≈ 1	↑	↑	↑	> 1	↓	↓	↓
98	≈ 1	↑	↑	↑	> 1	↓	↓	↓
120	≈ 1	↑	↑	↑	> 1	↓	↓	↓
144	≈ 1	↑	↑	↑	> 1	↓	↓	↓
240	≈ 1	↓	↑	=	> 1	↑	↓	=

de chuva observados. Este índice considera apenas aquelas previsões corretas, sendo sensível ao número de acertos. Ele também é fortemente dependente 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 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 3.5 e 3.6 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 grande parte da região amazônica, de grande importância para o clima global. Diferente da análise global, que considera como referência

Tabela 3.6 - Idem à Figura 3.5, exceto para o limiar menor ou igual a 35mm.

$\leq 35mm$								
	MPAS				SHiELD			
h	BIAS	POD	SR	CSI	BIAS	POD	SR	CSI
24	< 1	↓	↑	↑	< 1	↑	↓	↓
48	< 1	↓	↑	=	< 1	↑	↓	=
72	< 1	↓	↑	↓	< 1	↑	↓	↑
96	< 1	↓	↑	↓	< 1	↑	↓	↑
120	< 1	↓	↑	=	< 1	↑	↓	=
144	< 1	↓	↑	=	< 1	↑	↓	=
240	< 1	↓	↑	=	≈ 1	↑	↓	=

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 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.

3.3 Avaliação estatística da altura geopotencial

Na Figura 3.4a e 3.4b, 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 3.4a). 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 3.4a e 3.4b, painéis centrais) e BIAS (Figuras 3.4a e 3.4b, 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 3.4a) 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 3.4b) 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 3.4a, 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 3.5 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

3.4 Análise Multivariada de MH

Na Figura 3.6, são apresentadas as médias dos resultados obtidos da análise multivariada de MH para o período completo de simulação.

Em 24h (Figura 3.6a) 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 3.6b), 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 extratropicos 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 3.6c e 3.6e), 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 3.6e). O campo DifMH também mostrou um aumento, com o MPAS se distanciando mais do campo de observação em relação ao FV3.

Na Figura 3.7 apresenta-se o comportamento dos modelos na estação de verão do hemisfério austral. Em 24h (Figura 3.7a), 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 3.7b) 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 3.7c e 3.7d) há um aumento do MH em ambos os modelos, em especial nos extratropicos. 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 3.7e e Figura 3.7f), observa-se o aumento expressivo do MH nos extratropicos 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 3.8a), 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 3.8b), há um pequeno aumento do MH, forçado pelo aumento nos extratropicos. 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 3.8c), o MH aumenta nos extratropicos, 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 3.8d), 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 3.8e), há um aumento do valor de MH, principalmente nos subtropicos e extratropicos, 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 quando 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.

4 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, decidiu-se por não realizar a avaliação nos demais quesitos da característica de funcionalidade de software.

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 conseqüentemente inflar os valores de probabilidade de detecção, o que pode induzir a uma interpretação equivocada de seu desempenho. Uma vez que o SHiELD 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.

5 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.

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.

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

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

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. 3, 5

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. 9

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>. 5

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://library.wmo.int/doc_num.php?explnum_id=11599>>. 9

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>>>. 9

PERSSON, A. **User Guide to ECMWF forecast products**. 2001. Online. Disponível em: <<https://ghrc.nsstc.nasa.gov/uso/ds_docs/tcsp/tcspecmwf/ECMWFUserGuideofForecastProductsm32.pdf>>. 19

PIELKE, R. **Mesoscale meteorological modeling**. Orlando, FL.: Academic Press, 1984. 3

ROEBBER, P. Visualizing multiple measures of forecast quality. **Wea. Forecasting**, v. 24, p. 601–608, 2009. 10

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. 9, 18

SKAMAROCK, W. C.; KLEMP, J. B.; DUDA, M. G.; AL. et. A multiscale nonhydrostatic atmospheric model using centroidal voronoi tessellations 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>. 5

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>. 3

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>. 3

TAYLOR, K. Summarizing multiple aspects of model performance in a single diagram. **J. Geophys. Res.**, v. 106 (D7), p. 7183–7192, 2001. 10

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

WGNE. 2022. Disponível em:

http://wgne.meteoinfo.ru/wp-content/uploads/2021/11/wgne_table_2021_v4.xls.

Acesso em: 08/06/2022. 3

WIKIPEDIA. 2022. https://pt.wikipedia.org/wiki/ISO/IEC_9126. Acesso em

07/06/2022. 1

Teses e Dissertações (TDI)

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

Manuais Técnicos (MAN)

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

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.

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.

Propostas e Relatórios de Projetos (PRP)

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

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

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

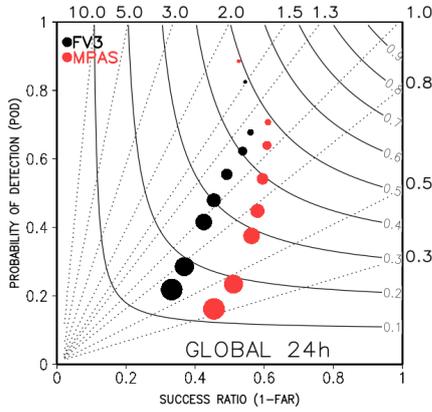
Programas de Computador (PDC)

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.

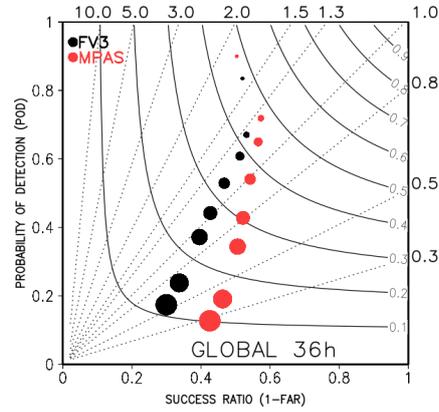
Pré-publicações (PRE)

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

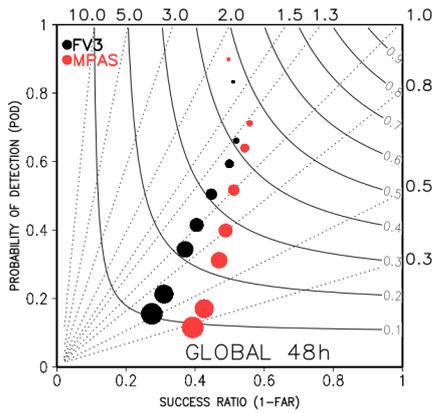
Figura 3.1 - Diagrama de desempenho para as previsões de precipitação de a) 24h b) 36h, c) 48h, d) 60h, e) 120h (132h) e f) 240h. As esferas representam os diferentes limiares de precipitação: 0,5mm, 2mm, 5mm, 10mm, 15mm, 20mm, 35mm, sendo a cor vermelha representativa do modelo MPAS e preto, do modelo SHIELD.



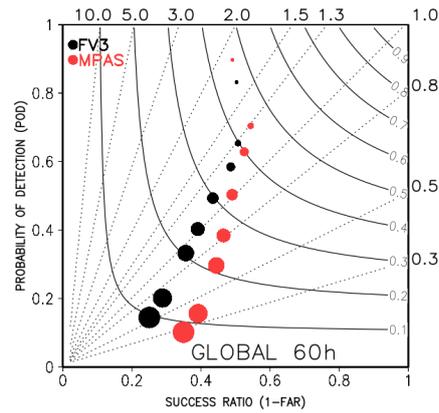
(a)



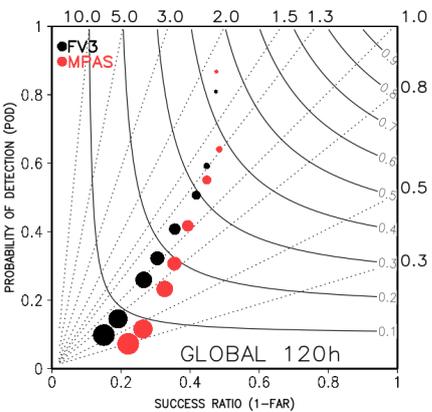
(b)



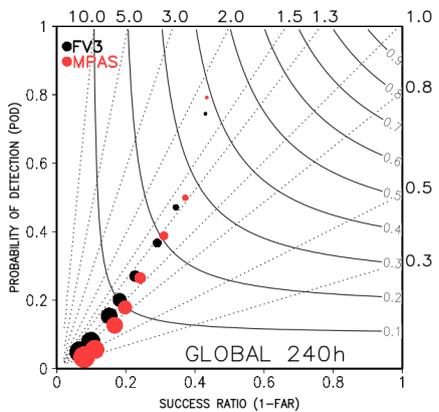
(c)



(d)



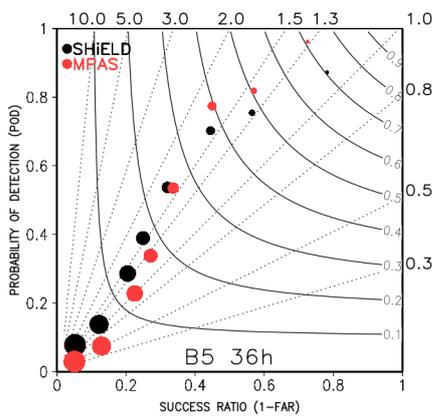
(e)



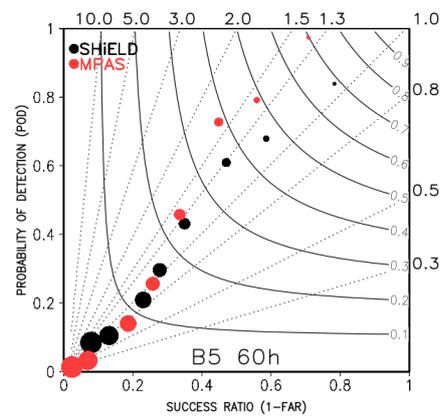
(f)

Fonte: Produção própria.

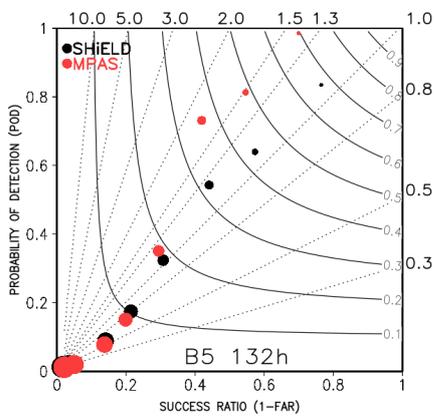
Figura 3.2 - Idem à Figura 3.1, exceto para a área B5



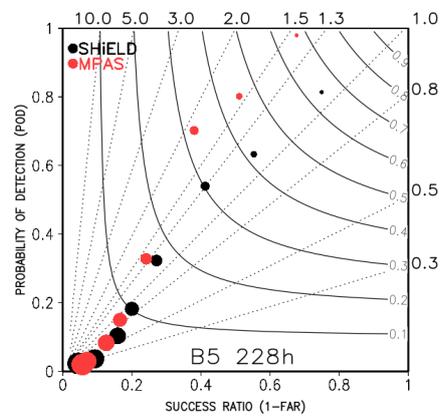
(a)



(b)



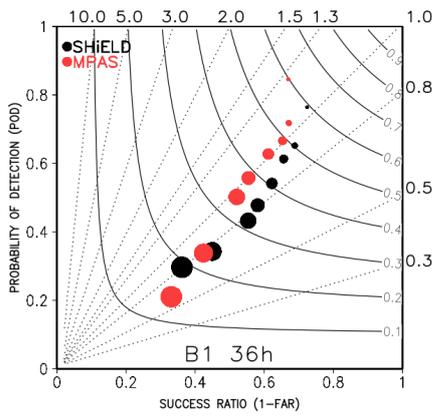
(c)



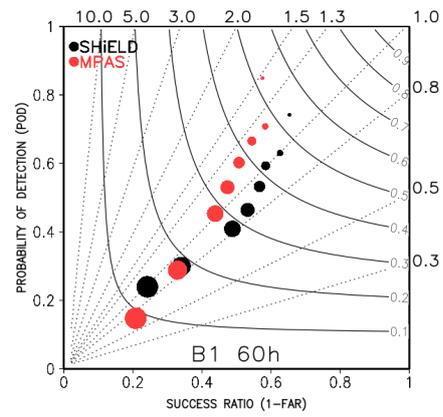
(d)

Fonte: Produção própria.

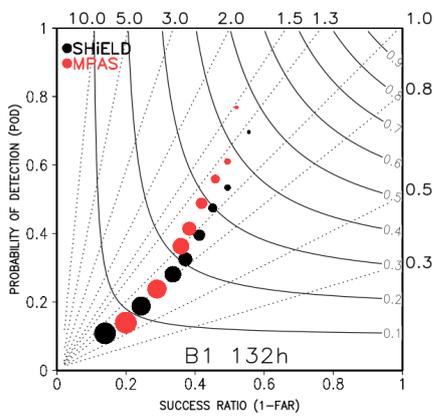
Figura 3.3 - Idem à Figura 3.1, exceto para a área B1



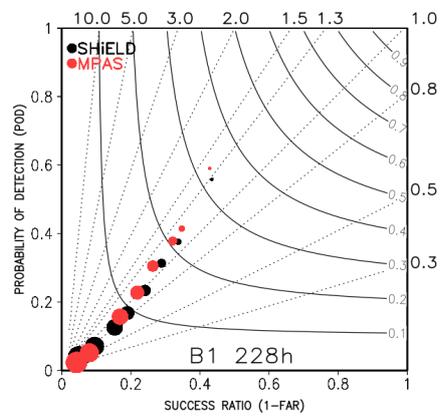
(a)



(b)

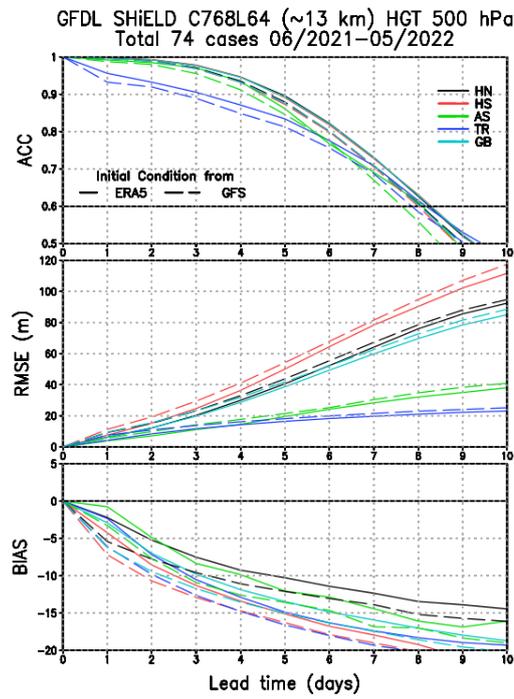


(c)

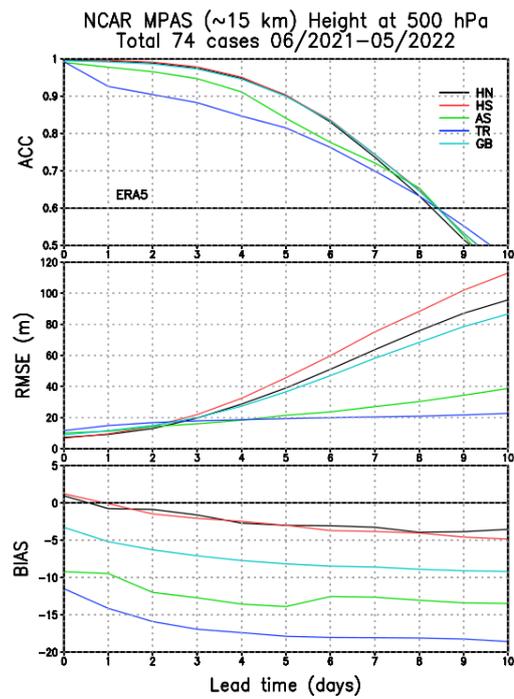


(d)

Figura 3.4 - Métricas das sub-características da Manutenibilidade com os modelos FV3, GEF e MPAS.



(a)



(b)

Fonte: Produção própria.

Figura 3.5 - 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).

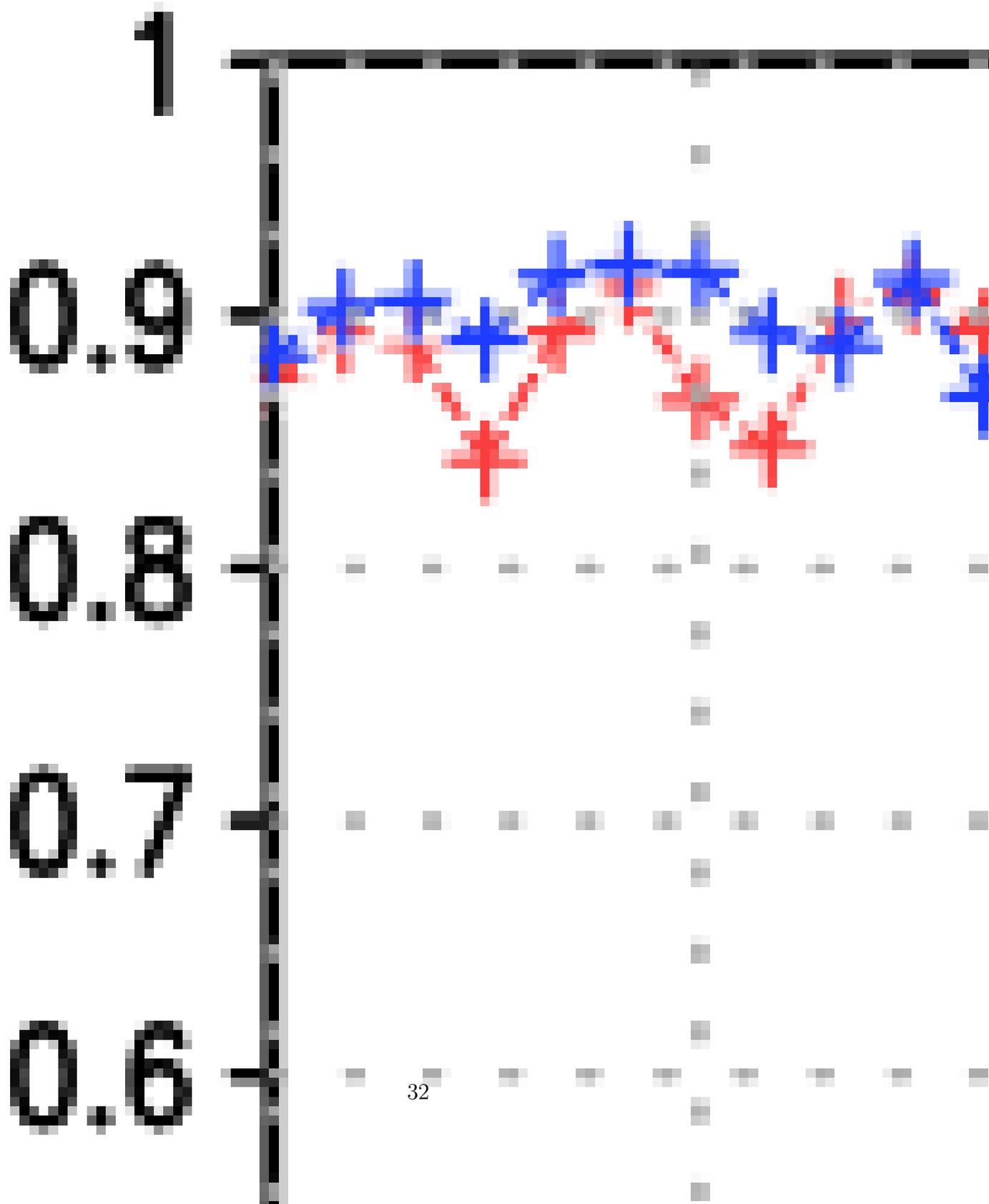


Figura 3.6 - Distância de MH para o MPAS e FV3 e diferença normalizada (DifMH) para os prazos de previsão de 24h (a), 48h (b), 72h (c), 120h (d) e 240h (e)

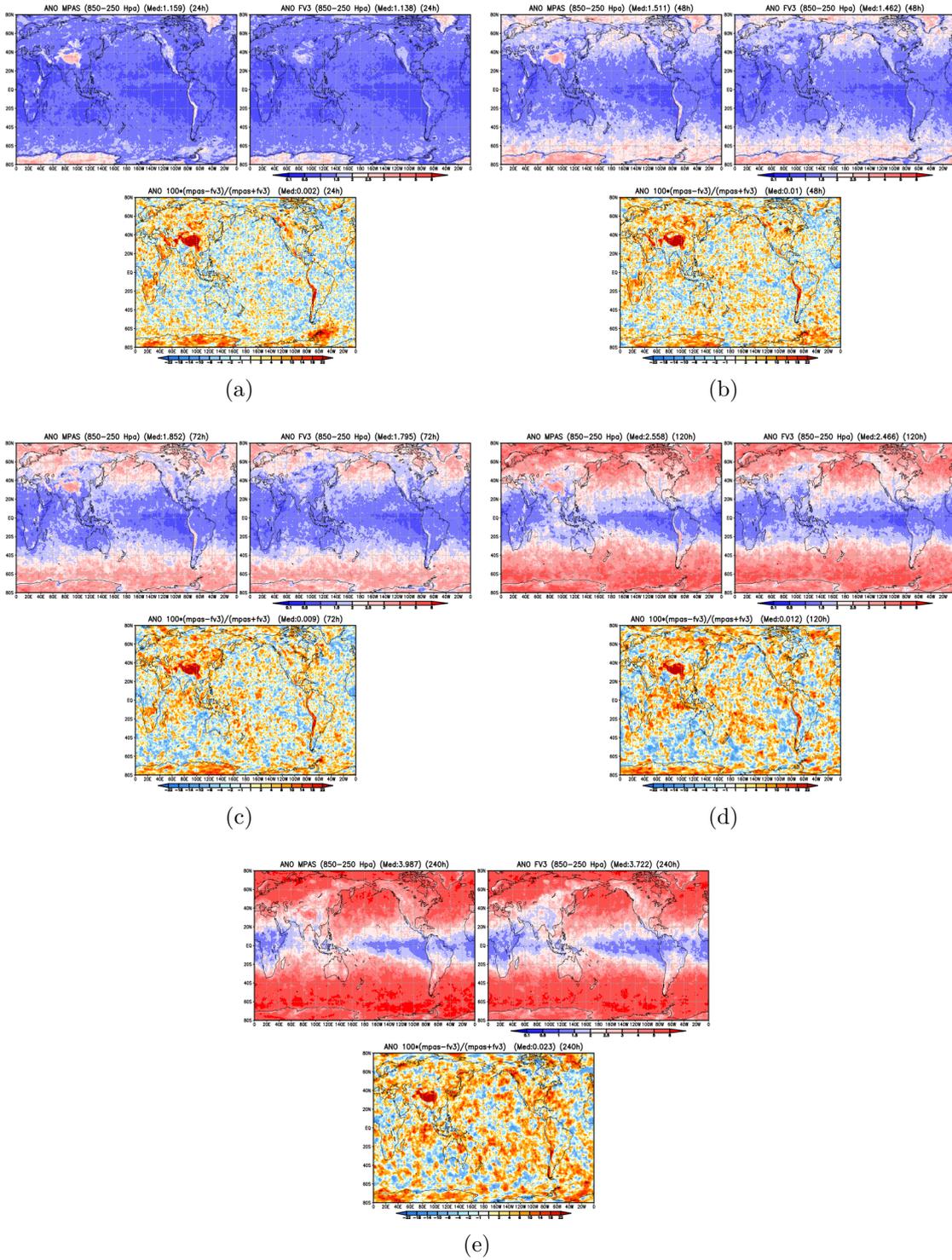
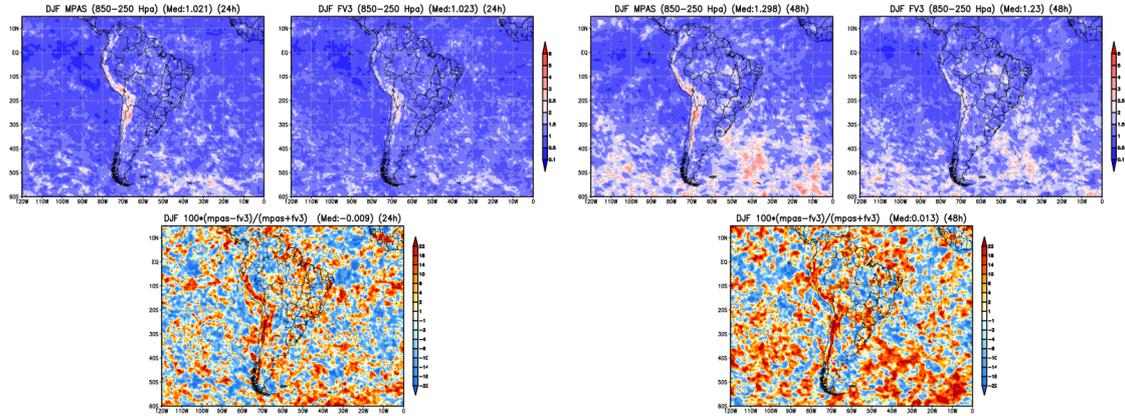
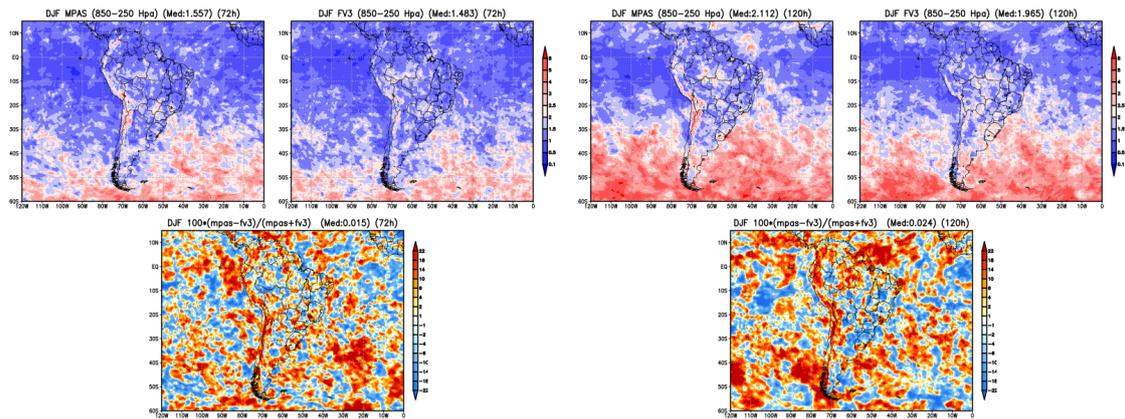


Figura 3.7 - Idem que a Figura 3.6 mas para os meses de verão austral sobre a América do Sul.



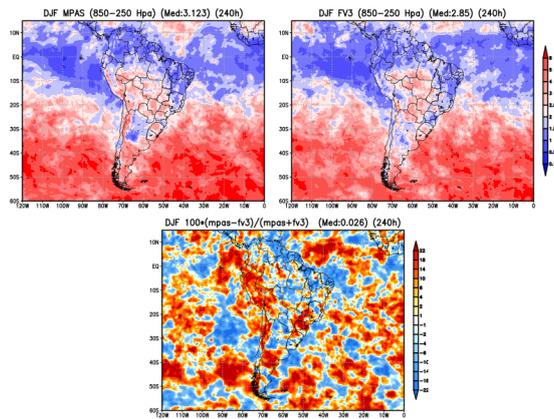
(a)

(b)



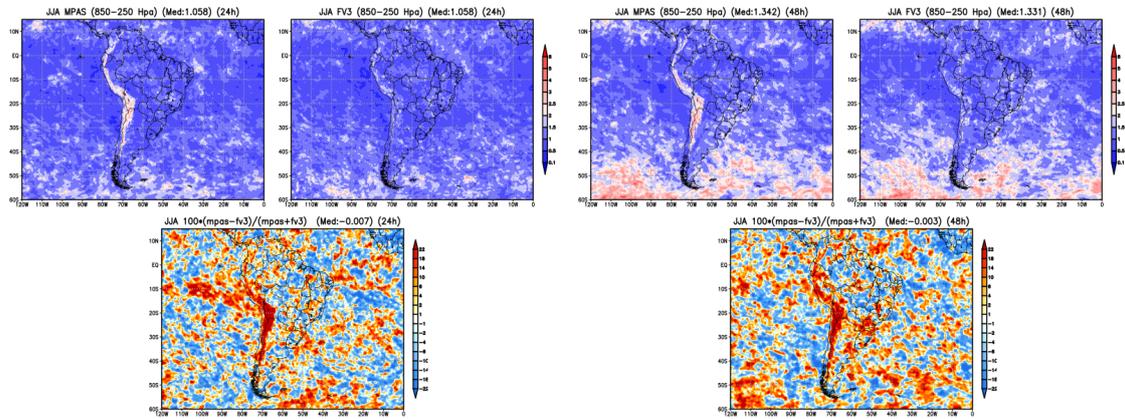
(c)

(d)



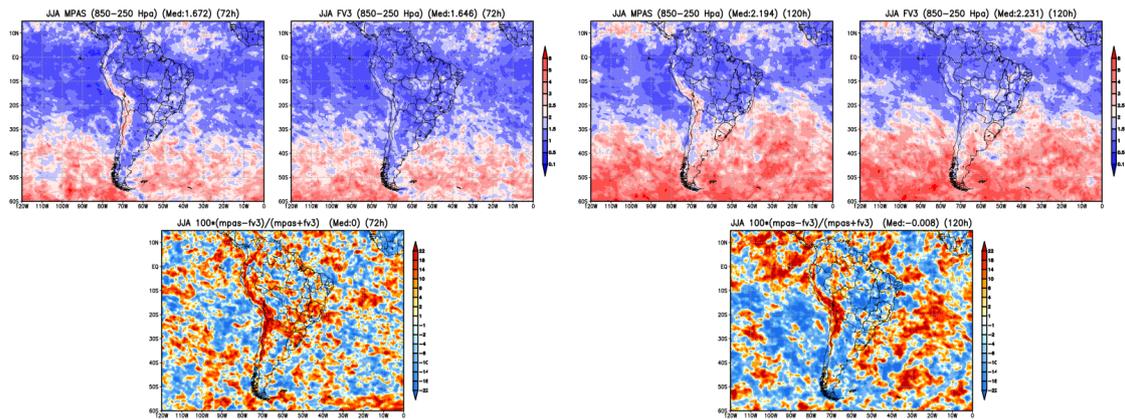
(e)

Figura 3.8 - Idem que a Figura 3.7 mas para os meses de inverno austral sobre a América do Sul.



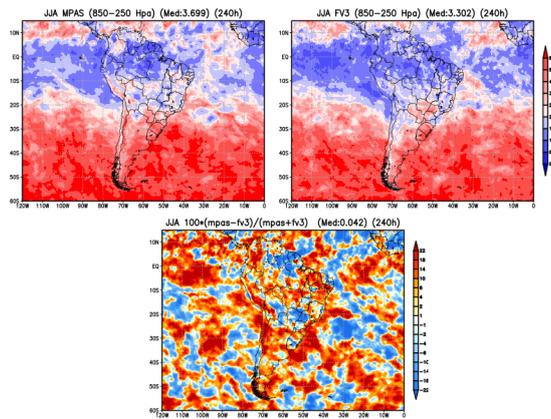
(a)

(b)



(c)

(d)



(e)