

## Influência da assimilação de dados na simulação dos processos físicos do modelo MONAN

### Resumo

A acurácia das simulações do tempo é intrinsecamente vinculada à qualidade dos dados assimilados e consequentemente à condição inicial e à complexidade dos modelos para representar de forma adequada os estados presentes e fornecer previsões confiáveis. O presente estudo investiga o impacto da assimilação de dados na simulação dos processos físicos dentro do modelo MONAN (Model for Ocean laNd-Atmosphere predictionN). A assimilação de dados desempenha um papel crucial na Previsão Numérica do Tempo (NWP), integrando dados observacionais com previsões de modelos para melhorar o estado inicial da atmosfera. Este processo é essencial **para reduzir incertezas na modelagem da radiação solar e terrestre**, que são componentes chave nas simulações climáticas. Ao incorporar dados de satélite e de superfície, o modelo MONAN visa aprimorar a precisão das previsões nas regiões tropicais e subtropicais, com foco na América do Sul. O estudo avalia a eficácia do modelo MONAN e examina a influência da assimilação de dados nessas simulações. Os resultados destacam a importância de dados de alta qualidade e técnicas avançadas de assimilação para melhorar a confiabilidade dos modelos climáticos.

**Palavras-chave:** Assimilação de dados, processos radiativos, modelo MONAN.

### 1 Introdução

A previsão meteorológica e as projeções climáticas tornaram-se ferramentas extremamente úteis no mundo contemporâneo e influenciam nossas atividades diárias e planejamento futuros onde consultar estas previsões tornou-se cada vez mais essencial, a ponto de fazermos isso quase que instintivamente. Todavia, anteciper as condições do tempo e clima para as próximas horas, dias e meses é um dos maiores desafios técnico-científicos enfrentados pela humanidade, no entanto diversos países têm superado esse desafio, evidenciando a importância de obter informações confiáveis e em tempo real sobre as condições atmosféricas e oceânicas atuais e futuras através de desenvolvimento de modelos atmosféricos (Bauer et al, 2015).

Observações da atmosfera terrestre, do solo e dos oceanos são essenciais para o estudo do planeta e a previsão de suas condições futuras. Esses dados são integrados por sistemas de assimilação de dados (DA do inglês *data assimilation*), operados em alguns

dos maiores supercomputadores do mundo, por organizações como a Administração Nacional da Aeronáutica e Espaço (NASA), a Administração Nacional Oceânica e Atmosférica (NOAA) e o Centro Europeu de Previsão de Tempo Médio (ECMWF). A assimilação de dados na previsão numérica do tempo (NWP) combina observações com previsões através de métodos dinâmicos/estatísticos para produzir um estado inicial ótimo da atmosfera. A quantidade, qualidade e diversidade das observações têm aumentado devido aos investimentos combinados dos setores público e privado e aos avanços tecnológicos. Embora essas melhorias forneçam mais informações globais, a maior parte delas ainda não é utilizada para informar previsões meteorológicas.

O MONAN (Model for Ocean laNd-Atmosphere predictioN) é um modelo comunitário do Sistema Terrestre Unificado que tem como principal objetivo produzir previsões que abrange a região tropical e subtropical com destaque na América do Sul em diferentes escalas temporais e espaciais (Frassoni et al., 2024). Seu desenvolvimento é gerenciado pelo Instituto Nacional de Pesquisas Espaciais (INPE) e tem sua estrutura de versão inicial baseado no núcleo dinâmico do modelo MPAS (Model for Prediction Across Scales), onde este núcleo combina parametrizações físicas obtidas tanto do próprio MPAS quanto de outras fontes ou desenvolvidas pela comunidade científica, visando melhorar a precisão das previsões meteorológicas e ambientais, aproveitando técnicas avançadas de computação de alto desempenho e dados considerados estado-da-arte para produtos de condição inicial e de contorno (Freitas, 2023).

A radiação solar que atinge o topo da atmosfera é influenciada pelos movimentos de rotação e translação da Terra. A atmosfera tem um papel crucial na absorção e dispersão de uma parte significativa dessa radiação antes que ela chegue à superfície terrestre, além de absorver a radiação emitida pela própria Terra. A radiação de onda curta descendente ( $R_{sd}$ ) representa a quantidade de energia solar que atinge a superfície da Terra sendo ela a força de entrada necessária para diferentes tipos de modelos climáticos (Kay et al., 2015) e modelos de superfície terrestre (Ek et al., 2003 ; Liang et al., 1994), no entanto, uns dos maiores desafios na simulação de radiação solar e terrestre se dá em fornecer estimativas precisas de  $R_{sd}$  em larga escala contínua (Wang e Dickinson, 2013).

A precisão das análises realizadas pelas técnicas de DA está intimamente ligada à qualidade da NWP (Kalnary, 2003), uma vez que os métodos de DA ajudam a reduzir as incertezas associadas aos modelos. De acordo com Talagrand (1997), o objetivo da

assimilação de dados é utilizar todas as informações disponíveis para definir o estado do fluxo atmosférico (ou ocenânico) com a máxima precisão.

## **2 Metodologia e Dados**

### **2.1 Dados de Satélite**

Os dados utilizados neste estudo são derivados do conjunto de dados GridSat-B1 (Globally Gridded Satellite-B1), no qual fornece informações históricas de satélites meteorológicos geoestacionários, cobrindo um período longo e contínuo, desde 1980 até o presente. Esses dados são usados principalmente para o monitoramento climático e a análise de eventos extremos ao longo do tempo. A sigla "B1" refere-se a um formato de dados específico utilizado pelo National Oceanic and Atmospheric Administration (NOAA) para armazenar imagens de satélite de geossíncrono global.

Os canais visíveis capturam a luz solar refletida pela superfície da Terra e pelas nuvens, operando na faixa de comprimento de onda entre 0,4 e 0,7  $\mu\text{m}$ . Embora sejam limitados à observação durante o dia, os dados visíveis são fundamentais para a identificação da cobertura de nuvens, a determinação da refletividade de superfícies, e a análise da estrutura das tempestades e sistemas convectivos. Esses dados são também utilizados para validar e complementar as observações de outros canais espectrais.

Já os canais de vapor d'água observam a radiação emitida na faixa de comprimento de onda em torno de 6,7  $\mu\text{m}$ . Estes canais são sensíveis à presença de vapor d'água na troposfera média a superior, permitindo monitorar o transporte de umidade, a circulação atmosférica e a identificação de jatos de vapor d'água, por exemplo. Esses dados são essenciais para a compreensão da dinâmica das tempestades e a análise da distribuição de umidade em grandes escalas.

Por fim, os canais de infravermelho (IR) são amplamente utilizados na meteorologia para monitorar a temperatura da superfície e a cobertura de nuvens, tanto durante o dia quanto à noite. Esses dados são obtidos na faixa de comprimento de onda entre 10 e 12  $\mu\text{m}$ , onde a emissão térmica da Terra e das nuvens é mais pronunciada. A temperatura do brilho detectada pelo IR é crucial para a estimativa da altura das nuvens, a identificação de topos de tempestades severas, e a avaliação da estrutura térmica da atmosfera.

## **2.2 Dados de Superfície**

Os dados de superfície provenientes de estações meteorológicas globais, obtidos através do Sistema Global de Telecomunicações (GTS). O GTS é uma rede internacional que coleta e distribui dados essenciais como temperatura, pressão, umidade e vento de estações distribuídas globalmente. Esses dados são disponibilizados em intervalos regulares, garantindo alta frequência temporal e ampla cobertura espacial.

A qualidade e a homogeneidade dos dados do GTS são asseguradas por protocolos internacionais de padronização, o que minimiza erros e inconsistências, tornando-os altamente confiáveis para a modelagem numérica e análise climática (World Meteorological Organization, 2017, Seção 6.1). Estes dados serão fundamentais para validar modelos, analisar eventos meteorológicos específicos e estudar padrões climáticos regionais.

## **2.3 ShereDA**

Esferas convolucionais são uma arquitetura especializada de redes neurais convolucionais, projetadas para lidar com dados definidos em superfícies esféricas, como o globo terrestre. Ao contrário das redes convolucionais tradicionais, que são geralmente aplicadas em dados definidos em superfícies planas, esferas convolucionais são utilizadas para processar dados como os provenientes de satélites meteorológicos ou modelagem climática, onde a superfície da Terra é tratada como uma esfera.

No contexto do sistema de assimilação de dados SphereDA, essa rede convolucional esférica é usada para combinar estados do modelo de previsão, chamados de "Background", com observações de estações meteorológicas e dados de satélites geostacionários e estações de superfície nesta versão. A arquitetura esférica permite uma integração geométrica mais natural e permite que as bordas sejam contínuas desses dados, respeitando as características geométricas da Terra e garantindo que o modelo resultante mantenha um equilíbrio adequado entre vento e pressão ao longo do processo de assimilação.

## **2.4 Configuração física do modelo MONAN/MPAS**

O modelo foi executado em sua versão global com resolução espacial de 30 km e 55 níveis verticais. As análises iniciais de todos os experimentos utilizaram 13 níveis verticais e uma resolução horizontal de 1°, sendo a melhoria da resolução realizada por

interpolação. O passo de tempo da simulação foi de 180 segundos para todas as simulações. Abaixo estão descritas as configurações físicas:

**Tabela 1 Parâmetros de Esquematização Utilizados**

Parâmetro	Esquema
Convecção	Grell-Freitas
Microfísica	WSM6
Superfície terrestre	Noah
Camada limite	YSU
Camada de superfície	Monin-Obukhov
Radiação de ondas longas (LW)	RRTMG
Radiação de ondas curtas (SW)	RRTMG
Fração de nuvens para radiação	Xu-Randall
Arrasto de ondas gravitacionais por orografia	YSU

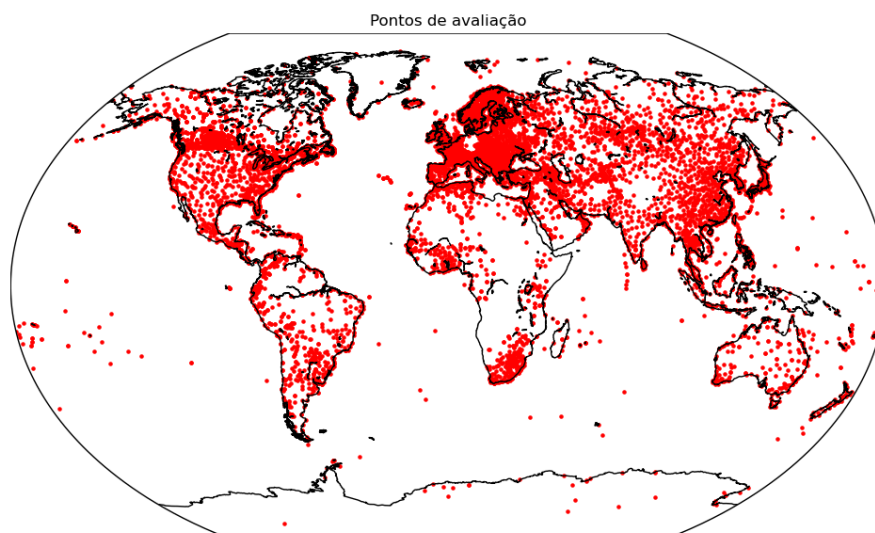
**2.5 Avaliação estatística dos resultados**

A avaliação estatística do modelo é uma etapa crucial para determinar sua capacidade de reproduzir os fenômenos atmosféricos com precisão. Nesta análise, focamos no cálculo do Erro Médio Quadrático (MSE) entre as previsões do modelo e as observações de superfície. O MSE é uma métrica amplamente utilizada para quantificar a diferença entre os valores previstos pelo modelo e os valores observados, oferecendo uma medida da precisão da representação dos fenômenos pelo modelo. A equação que define o MSE é dada por:

$$MSE = 1/N \sum (y_o - y_i)^2$$

onde  $N$  é o número de pontos de dados,  $y_o$ , representa os valores observados, e  $y_i$  são os valores previstos pelo modelo. Esta métrica é calculada para o ponto mais próximo das previsões do modelo em relação às observações de superfície, garantindo uma comparação direta entre ambos.

Os resultados próximos de 0 indicam uma melhor capacidade do modelo de representar os fenômenos atmosféricos com fidelidade. Foram avaliados três parâmetros essenciais: a temperatura a 2 metros, a temperatura do ponto de orvalho e a direção do vento. Essas variáveis foram analisadas em relação a várias estações meteorológicas, com o modelo sendo inicializado de diversas maneiras. A comparação dessas diferentes inicializações permite uma compreensão mais profunda de como as condições iniciais afetam a performance do modelo ao longo do tempo. Na Figura 01 mostra o local das estações que foram avaliadas.



**Figura 01** - Localização das estações meteorológicas em superfície.

### 3 Resultados

#### 3.1 Impacto na previsão

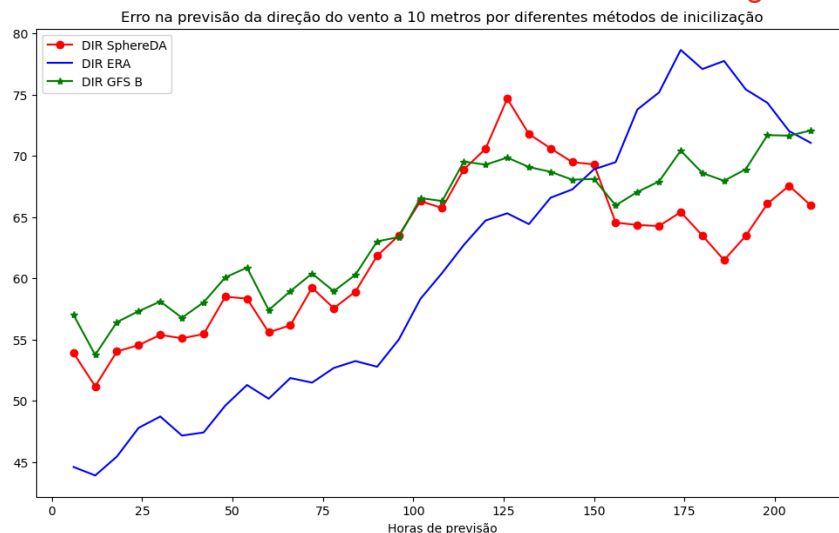
A Figura 02 mostra como diferentes métodos de assimilação de dados afetam o erro na previsão da temperatura do ponto de orvalho e na direção do vento. Três métodos de inicialização são comparados: Sphere, ERA e GFS. Podemos observar que as linhas mostram que o erro não é constante, mas flutua ao longo do tempo indicando que a precisão da previsão varia dependendo do método utilizado. No início, as três linhas estão próximas, sugerindo que os métodos geram erros semelhantes, mas à

é necessário descrever os dados na metodologia. todos possuem a mesma resolução horizontal e vertical?

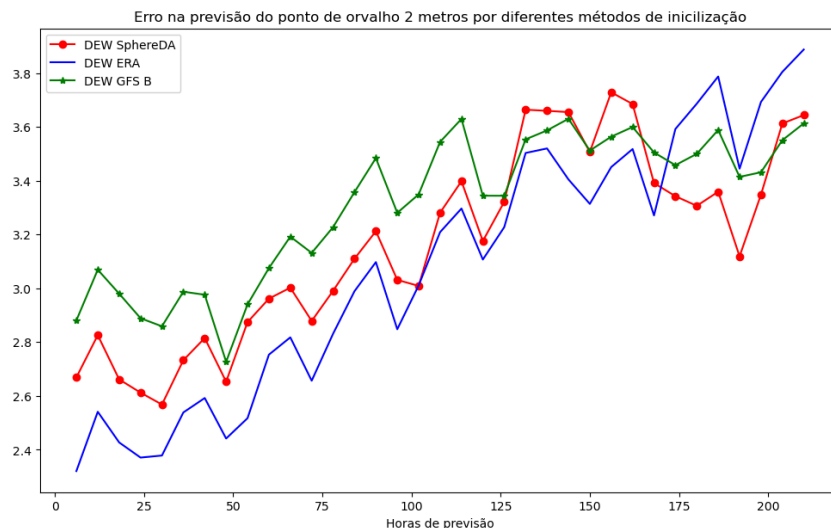
medida que o tempo avança, observam-se diferenças nos níveis de erro, com alguns métodos se destacando em certos momentos.

É possível identificar que o método que apresenta o menor erro ao longo do tempo **não é o ERA**, em grande parte do gráfico, o erro associado a este método é consistentemente mais baixo em comparação com os outros dois métodos, portanto ele pode ser considerado o mais preciso. Então toda esta análise, destaca a importância de escolher o método de assimilação de dados adequado para melhorar a precisão das previsões de temperatura que é fundamental para simulações climáticas.

**Esta figuras são erros somente sobre a America do Sul o no globo.**



**Por que o erro a partir de 120 horas tornam-se mais estáveis**

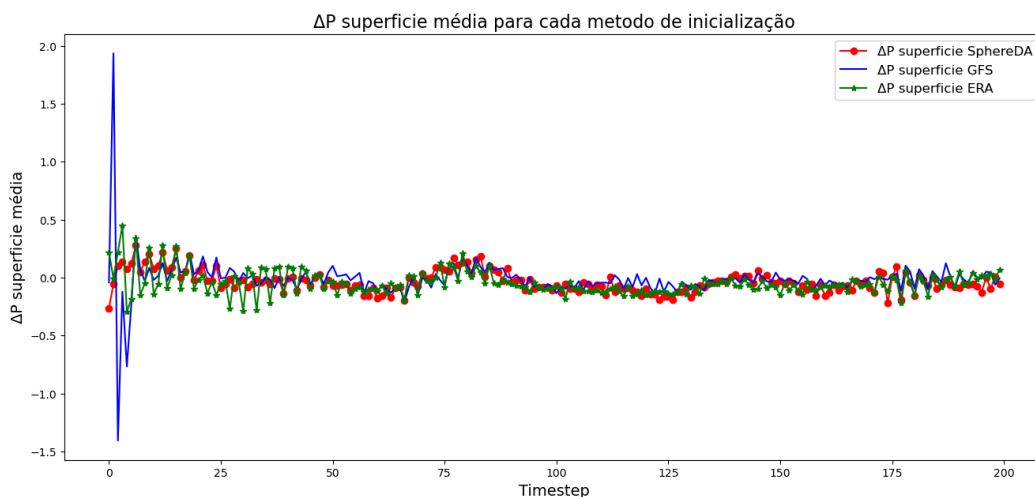


**Figura 02 - Comparação do erro na previsão da temperatura do ponto de orvalho a 2m e na direção do vento a 10m através de três métodos de inicialização: Sphere, ERA e GFS.**

### 3.2 Estabilidade na inicialização para cada método

A Figura 03 apresenta uma análise da tendência da pressão de superfície ao longo do tempo para três métodos de inicialização (SphereDA, GFS e ERA), fornecendo uma visão clara sobre a eficácia da assimilação de dados. É notável que inicialmente todos os métodos exibem variações significativas, refletindo o ajuste dinâmico necessário nos primeiros passos de tempo. Na medida que se avança no tempo, essas variações tendem a se estabilizar em torno de zero, indicando que os métodos conseguiram assimilar os dados de forma eficaz para tentar alcançar um estado de equilíbrio.

Acompanhar a estabilidade do modelo numérico por meio da tendência de pressão em superfície a cada timestep é fundamental para poder garantir a confiabilidade das previsões. A estabilização da pressão de superfície nos sugere que o modelo está se comportando de maneira consistente, sem grandes perturbações, o que é muito importante para as previsões das simulações seguintes. É possível observar que o método SphereDA demonstrou uma capacidade superior de estabilização, onde alcançou essa condição mais rapidamente do que em comparação com os outros métodos, sugerindo uma assimilação de dados mais eficiente e potencialmente uma previsão estável, fator muito importante na avaliação do procedimento de assimilação de dados. De maneira geral, monitorar as tendências é crucial para validar a qualidade da inicialização e também a performance do modelo ao longo de toda a simulação.



**Figura 03** - Tendência da pressão de superfície ao longo do tempo para três métodos de inicialização: SphereDA, GFS e ERA.

### 3.3 Fontes de incertezas na modelagem da radiação solar e terrestre



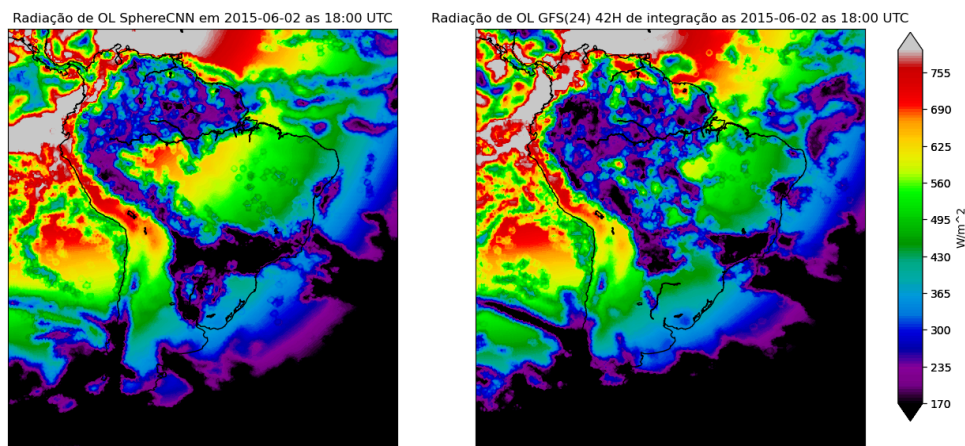
Um dos papéis da modelagem é representar os processos físicos da interação da radiação com a atmosfera através do esquema de transferência radiativa nos Modelos de Circulação Geral da Atmosfera (MCGA) e proceder com a parametrização de radiação em cada camada atmosférica do modelo, obtendo o saldo de radiação de onda curta e de onda longa a fim de obter o saldo líquido de radiação.

A quantidade de radiação em cada camada vai depender se é uma atmosfera limpa, com nuvens, com aerossóis ou qualquer outro composto, pois estes interagem com a radiação mudando a quantidade de radiação em cada camada do modelo e impactando no balanço global de energia (aquecimento ou resfriamento da atmosfera).

Pela equação de conservação de calor, incertezas serão geradas devido a dificuldade de definir a distribuição vertical e horizontal da nuvem. A radiação na presença de nuvens tende a desestabilizar a atmosfera, pois quando a radiação de onda curta interage com a nuvem, parte dela é refletida e parte é absorvida, aquecendo o topo da nuvem e resfriando abaixo dela, por exemplo. Alguns processos que ocorrem nessa interação não estão bem representados nas parametrizações (como a mudança de fase da água, que envolve a parametrização de microfísica de nuvens).

Para saber como a radiação será atenuada na atmosfera por aerossóis, gases e/ou nuvens, se faz necessário resolver a equação de transferência de radiação, que requer informações sobre propriedades ópticas dos gases e particulados, tais como: coeficientes de extinção, albedo de espalhamento único, função de fase de espalhamento, etc.

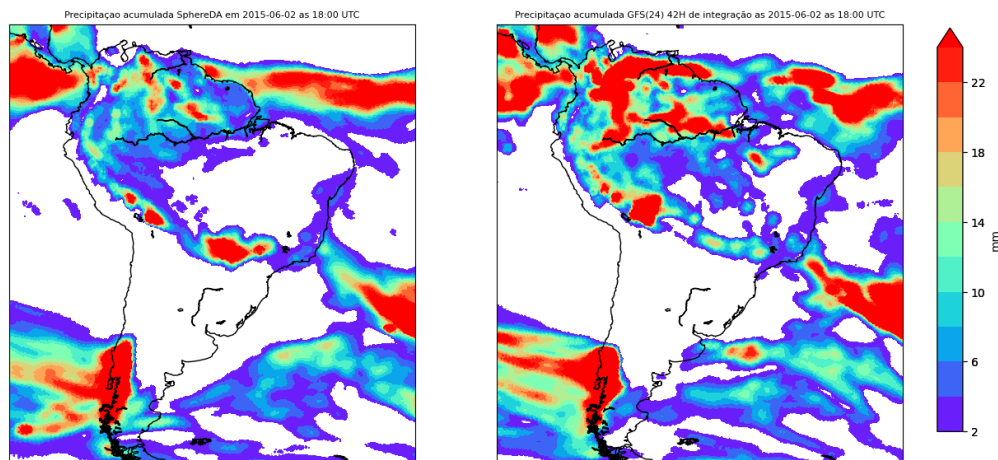
Portanto, o que gera as maiores incertezas na estimativa da radiação e nas taxas de aquecimento e resfriamento é a distribuição vertical e horizontal da nebulosidade e a parametrização adequada da microfísica de nuvens, que se propaga para o escoamento médio da atmosfera, afetando a representação da circulação atmosférica.



deve colocar o grafico do ERA. O campo da magnetude não ajuda muito  
é melhor colocar o BIAS ou RMSE

**Figura 04** - Radiação de onda longa ( $\text{W/m}^2$ ) por dois métodos de inicialização:

Sphere e GFS. Colocar o bias ou rmse é mais robusto



**Figura 05** - Precipitação acumulada (mm) por dois métodos de inicialização:

Sphere e GFS. Colocar o bias ou rmse é mais robusto

Qual dos graficos está mais próximo da observação?

#### 4 Conclusão

Este estudo demonstrou que a assimilação de dados desempenha um papel crucial na melhoria da precisão das simulações no modelo MONAN. Ao incorporar dados de satélite e de superfície, verificou-se uma redução significativa das incertezas associadas. Os resultados obtidos enfatizam a importância de dados de alta qualidade e técnicas avançadas de assimilação para a representação precisa de processos físicos. A análise revela que a implementação de técnicas de assimilação de dados aprimora o estado inicial da atmosfera, resultando em previsões mais confiáveis e robustas. Além disso, o estudo evidenciou que a complexidade do modelo e a qualidade dos dados assimilados são fatores determinantes para a acurácia das simulações do tempo, destacando a necessidade de contínuo avanço nas técnicas de assimilação para enfrentar os desafios da previsão numérica em diferentes escalas temporais e espaciais. A aplicação das metodologias discutidas neste trabalho tem o potencial de contribuir significativamente para o desenvolvimento de previsões mais eficientes e precisas. Avaliações mais profundas com o impacto da radiação devem ser feitas a fim de se obter os reais parâmetros de melhora e redução do erro ao longo da integração relacionados aos processos radioativos e outros processos físicos do modelo.

## 5 Referências

- Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567), 47-55. <<https://doi.org/10.1038/nature14956>>.
- Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., Tarpley, J. D. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, 108, 8851, <<https://doi.org/10.1029/2002JD003296>>.
- Frassoni, A., Fernandez, J., Rosa, M., Rozante, J., Yamada, B. Avaliação da funcionalidade de modelos numéricos de escala global para o projeto MONAN - Parte 1: Análise qualitativa, quantitativa e multivariada. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2024. <<http://urlib.net/8JMKD3MGP3W34T/4AR4SDS>>.
- Freitas, S. Model for Ocean-Land-Atmosphere Prediction (MONAN): Avaliação dos candidatos ao núcleo dinâmico do componente atmosférico (MONAN-ATM) - Recomendação ao comitê científico do MONAN. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2023. <<http://urlib.net/8JMKD3MGP3W34T/49MKML8>>.
- Kay, J. E., Deserto, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J. F., Lourenço, D., Lindsay, K., Middleton, A., Muñoz, E., Neale, R., Oleson, K., Polvani, L., Vertenstein, M. The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. 2015, *Bull. Amer. Meteor. Soc.*, 96, 1333–1349, <<https://doi.org/10.1175/BAMS-D-13-00255.1>>.
- Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. 1994, *J. Geophys. Res.*, 99, 14 415–14 428, <<https://doi.org/10.1029/94JD00483>>.
- Wang, K., Dickinson, R. E. Global atmospheric downward longwave radiation at the surface from ground-based observations, satellite retrievals, and reanalyses. 2013, *Rev. Geophys.*, 51, 150–185, <<https://doi.org/10.1002/rog.20009>>.

TALAGRAND, O. Assimilation of observations, an introduction (gt special issue data assimilation in meteorology and oceanography: Theory and practice). **Journal of the Meteorological Society of Japan**, v. 75, n. 1B, p. 191-209, 1997.

KALNAY, E. **Atmospheric Modeling, Data Assimilation, and Predictability**, v. 54. p. 341, 2003.

World Meteorological Organization. (2017). *Guide to the Global Observing System (WMO-No. 488)*. Geneva: WMO. Retrieved from <https://library.wmo.int/viewer/35699>.