



Anuário do Instituto de Geociências

ISSN: 0101-9759

ISSN: 1982-3908

Universidade Federal do Rio de Janeiro

Tenorio, Ricardo Bruno de Araújo; Fernandez, José
Henrique; Mendes, David; Silva, José Pedro da
Estimação da Cobertura de Gelo Marinho nos Mares Antárticos de
Weddell, Belingshausen e Amundsen com Redes Neurais Artificiais
Anuário do Instituto de Geociências, vol. 45, 2022, pp. 1-17
Universidade Federal do Rio de Janeiro

DOI: https://doi.org/10.11137/1982-3908_2022_45_40763

Disponível em: <https://www.redalyc.org/articulo.oa?id=695375056011>

- Como citar este artigo
- Número completo
- Mais informações do artigo
- Site da revista em [redalyc.org](https://www.redalyc.org)

redalyc.org

Sistema de Informação Científica Redalyc
Rede de Revistas Científicas da América Latina e do Caribe, Espanha e Portugal
Sem fins lucrativos acadêmica projeto, desenvolvido no âmbito da iniciativa
acesso aberto

Estimação da Cobertura de Gelo Marinho nos Mares Antárticos de Weddell, Belingshausen e Amundsen com Redes Neurais Artificiais

Estimation of Sea Ice Cover in Weddell, Bellingshausen and Amundsen Antarctic Seas Using Artificial Neural Networks

Ricardo Bruno de Araújo Tenorio , José Henrique Fernandez , David Mendes ,
José Pedro da Silva Júnior 

Universidade Federal do Rio Grande do Norte, Natal, RN, Brasil

E-mails: ricardobatenorio@gmail.com; jhenrix@gmail.com; david.mendes22@gmail.com; josepedro_junior@hotmail.com

Autor Correspondente: Ricardo Bruno de Araújo Tenorio; ricardobatenorio@gmail.com

Resumo

O gelo marinho desempenha um papel fundamental na regulação térmica das regiões polares. Observações de satélites evidenciam que na Antártica o gelo apresentava, na série histórica, tendências positivas em cobertura e extensão. Em 2019 houve um padrão de inversões entre os valores da normal climatológica e dos dados de reanálise. Nesse contexto, este estudo teve como principal objetivo avaliar o potencial de previsibilidade de cobertura de gelo marinho com a aplicação de técnicas de RNAs em 3 mares que banham o continente Antártico, a saber: Weddell, Bellingshausen e Amundsen. Para tanto, foram utilizados como previsores a temperatura da superfície do mar, a temperatura do ar a 2 metros, a velocidade do vento a 10 metros, o albedo e os fluxos de calor latente e sensível, no período de 1979 a 2019. Os dados foram particionados em 70% para treinamento e 30% para testes. Modelos SARIMAX serviram como valores de referência para aferição da precisão das previsões com RNAs. Em todos os meses com anomalias absolutas superiores a 15% de concentração, o modelo de RNA CNN-LSTM superou os modelos MLP e SARIMAX.

Palavras-chave: Aprendizado profundo; CNN-LSTM; TensorFlow

Abstract

Sea ice plays a fundamental role in the thermal regulation of polar regions. Satellite observations show that the Antarctic ice, in the historical series, presented positive trends in coverage and extension. In 2019, there was a pattern of inversions between climatological normal and observed data values. In this context, this study aimed to assess the potential for predictability of sea ice concentration with the application of RNA techniques in 3 Antarctic Seas, namely Weddell, Bellingshausen, and Amundsen. The sea surface temperature, the air temperature at 2 meters, the wind speed at 10 meters, the albedo, and the latent and sensitive heat fluxes, from 1979 to 2019, were used as predictors. The data set was split 70% for training and 30% for testing. SARIMAX models were used as a baseline for quantifying the precision of forecasts with ANNs. In all months with absolute anomalies above 15% of concentration, the CNN-LSTM ANN model surpassed the MLP and SARIMAX models.

Keywords: Deep learning; CNN-LSTM; TensorFlow

1 Introdução

O gelo marinho, formado a partir do congelamento da água do mar, é um isolante térmico natural na interface oceano-atmosfera. Por restringir as trocas de calor, massa e momentum entre o mar e o ar adjacente, e por refletir grande parte da radiação solar incidente, esse fator climático é uma variável fundamental para se entender o clima nas regiões polares (Parkinson 2002; Serreze, Holland & Stroeve 2007). Seu impacto e importância vão além dos processos físicos. O ciclo de formação do gelo marinho afeta diretamente plantas e animais aquáticos polares (árticos/antárticos) de micro-organismos aos grandes cetáceos (Boetius et al. 2015; Chambault et al. 2018).

Segundo Armour et al. (2016) a superfície do Oceano Antártico tem sido aquecida a uma taxa de $0,02^{\circ}\text{C}/\text{década}$ desde 1950, enquanto que, para a média global da Temperatura da Superfície do Mar (TSM), o gradiente tem sido de $0,08^{\circ}\text{C}/\text{década}$ para o mesmo período. Uma explicação para este comportamento diferenciado entre o Oceano Austral e os demais oceanos é que ao sul da Corrente Circumpolar Antártica, a extensa cobertura de gelo marinho por si só é um fator de retardamento na taxa de aquecimento oceânico, funcionando fisicamente como um isolante térmico contra as forças radiativas (Hutchinson et al. 2013).

Observações satelitais, desde os anos de 1970, indicam um Ártico com gelo marinho cada vez mais fino e mais jovem (Cavalieri et al. 1997; Cavalieri & Parkinson, 2012) associado a um declínio em sua extensão (Kirchmeier-Young et al. 2017; Stroeve et al. 2012). Por outro lado, algumas regiões da Antártica (Mar de Weddell e Mar de Ross) têm apresentado um ligeiro aumento na extensão desta variável ambiental para o período de 1979 à 2016 (Meehl et al. 2019). Indo de encontro com essas observações estão os resultados de simulações numéricas com Modelos Climáticos Globais, que apresentam uma média de decréscimo na extensão do gelo para ambos os polos (Chemke & Polvani 2020; Gagné, Gillett & Fyfe 2015). Embora a Antártica, a partir de 2016, tenha apresentado uma tendência de perda de gelo, as simulações numéricas ainda apresentam taxas mais acentuadas que os valores observados (Rouch et al. 2020).

Atualmente, com a crescente e consistente alta performance dos *hardwares* de uso pessoal, as aplicações de técnicas de Aprendizado de Máquina têm se mostrado cada vez mais versáteis e potentes, como na Visão Computacional e em Processamento de Linguagem Natural (Young et al. 2018). Neste contexto, esta pesquisa teve por objetivo a avaliação da aplicação de Redes Neurais Artificiais, subconjunto da Inteligência Artificial (IA), na previsão

de concentração e área de cobertura de gelo marinho nos mares antárticos de Weddell, Bellingshausen e Amundsen.

2 Dados e Métodos

2.1 Área de Estudo

A área estudada (Figura 1) está situada entre as longitudes 0° e 120°W e entre as latitudes 60°S e 75°S , compreendendo os mares de Weddell ($0^{\circ} - 60^{\circ}\text{W}$), Bellingshausen ($60^{\circ}\text{W} - 90^{\circ}\text{W}$), e Amundsen ($90^{\circ}\text{W} - 120^{\circ}\text{W}$). Estas regiões foram selecionadas por apresentarem características distintas quanto aos padrões de tendência na extensão e concentração de gelo marinho, o que permitiu avaliar as previsões das redes neurais em diferentes situações.

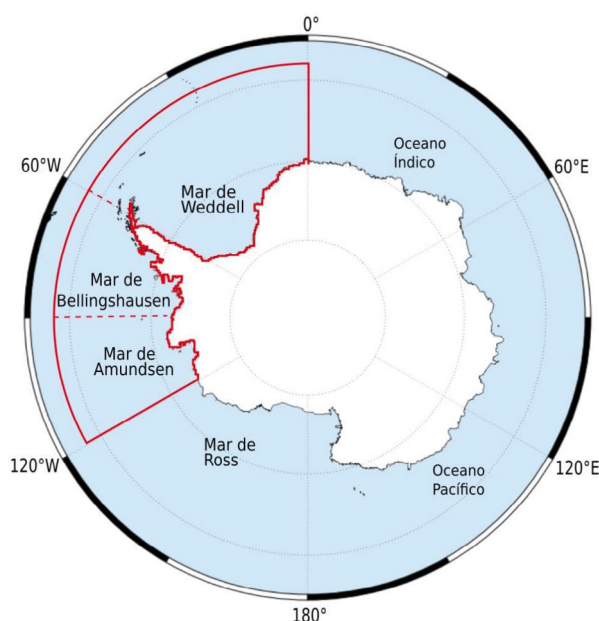


Figura 1 Área de estudo, cobrindo 1/3 da região marítima da Antártica Ocidental, mostra os mares antárticos de Weddell, Bellingshausen e Amundsen com seus limites destacados em vermelho. Entre os mares de Weddell e Bellingshausen encontra-se a Península Antártica.

Se analisada como um todo, anualmente, a Antártica tem seus pontos de mínimo, em extensão de gelo, em fevereiro (sempre abaixo de $5 \times 10^6 \text{km}^2$) e, de máximo, em setembro (em torno de $17 \times 10^6 \text{km}^2$). O continente austral também apresentou uma tendência positiva em sua extensão de gelo de $(+22.4 \pm 4.3) \times 10^3 \text{km}^2 \text{ano}^{-1}$ entre os anos de 1979-2014.

No entanto, a tendência observada para essa variável nos mares de Bellingshausen/Amundsen, durante toda a série

histórica analisada (1979-2019), foi a única negativa ($-3.7 \pm 1.8 \times 10^3 \text{ km}^2 \text{ ano}^{-1}$) em todo o oceano austral (Parkinson 2019). Este padrão anômalo nesta região foi associado ao intenso aquecimento (1979-1997) e resfriamento (1999-2014) da temperatura do ar na Península Antártica (Vaughan et al. 2003; Turner et al. 2016) e à Baixa no Mar de Amundsen, que consiste em uma área de baixa pressão no extremo sul do Pacífico Sul, limitado a leste pelo mar de Ross e a oeste pelo mar de Bellingshausen, resultante da interação dos ventos de oeste com as altas topografias da Terra de Vitória (Raphael et al. 2016). Por outro lado, a formação de gelo marinho no mar de Weddell foi de tendência positiva ($7.0 \pm 3.7 \times 10^3 \text{ km}^2 \text{ ano}^{-1}$) entre 1979 e 2018, entretanto, de 2014 a 2018, Parkinson (2019) observou um intenso declínio para este setor ($4.0 \pm 3.5 \times 10^3 \text{ km}^2 \text{ ano}^{-1}$). As condições oceanográficas de superfície nesta região são predominantemente causadas pelo Giro de Weddell, o qual transporta icebergs e massas de gelo para o norte da Península Antártica (Oliva et al. 2017).

2.2 Dados

Este estudo focou-se na utilização de modelos regressivos de IA para previsão da Concentração de Gelo Marinho (CGM) em algumas regiões do continente antártico. Portanto, foram utilizadas reanálises mensais de CGM como alvo, tendo as variáveis ambientais (i) temperatura da superfície do mar (TSM), (ii) temperatura do ar a 2 metros (T2M), (iii) velocidade do vento a 10 metros, (iv) albedo e (v e vi) fluxos de calor latente e sensível como previsores. O conjunto total de dados foi composto por mais 5 variáveis, a saber: temperatura do ponto de orvalho, temperatura de evaporação, pressão ao nível do mar, radiação UV na superfície e precipitação. A seleção dos previsores foi realizada utilizando-se o algoritmo de Floresta Aleatória (Ho 1995). A escolha desse algoritmo, para seleção de atributos, foi arbitrária. Floresta Aleatória é definida por Breiman (2001) como um classificador composto por um conjunto de classificadores estruturados em árvores $\{h(x, \theta_k), k = 1, \dots\}$ para $\{\theta_k\}$ sendo os vetores aleatórios independentes e identicamente distribuídos. Cada árvore tem um voto para as classes de entrada x . Além disso, é utilizado o método de agregação *bootstrap* (ou *bagging*) na amostragem para treinar cada árvore. Em regressão, o valor previsto com essa técnica é a média das previsões de todas as árvores. (Chen, Li & Wang 2012; Strobl et al. 2007). O nível de importância de cada predictor é apresentado na Tabela 1.

A base de dados foi disponibilizada pelo Era-5 do *European Centre for Medium-Range Weather Forecasts* (ECMWF) com resolução espacial de $0,5^\circ \times 0,5^\circ$ de latitude

e longitude, e o período estudado de 1979 a 2019. Os dados foram particionados em 70% (1979 – 2007) destinado ao treinamento da rede, e 30% para testes dos modelos regressivos, sendo ainda o último ano (2019) separado para validação das estimativas encontradas. Salienta-se que, por se tratar de séries temporais, a preservação da ordem dos elementos é fundamental, de maneira que não se aplicaria técnicas de particionamento que não levam em conta dados sequenciais como acontece com validações cruzadas. A justificativa para a escolha do ano de 2019 foi que, a partir de 2014, houve duas situações distintas, sob o ponto de vista de médias anuais, entre as áreas estudadas: (i) no mar de Weddell, até 2014, houve acresção na CGM e declínio entre 2015 e 2018, e (ii) nos mares de Bellingshausen/Amundsen houve retração até 2014 e acresção entre 2015 e 2018 (Meehl et al. 2019; Parkinson 2019; Turner et al. 2020). Dessa forma, o objetivo era a análise da sensibilidade das previsões das RNAs após períodos significativos de inversões nas estruturas temporais das séries em questão.

Tabela 1 Valores de importâncias relativas das variáveis climáticas selecionadas, utilizando a extração de características com o algoritmo de Florestas Aleatórias.

Variável	Nível de Pertinência (Importância Relativa)
TSM	0.26
Vento a 10 m	0.19
Albedo	0.17
Temperatura a 2m	0.15
Fluxo de Calor Latente	0,14
Fluxo de Calor Sensível	0,09

A última fase de pré-processamento dos dados foi a normalização de todas as variáveis para o intervalo de 0 a 1. Os valores utilizados como limites máximo ($X_{\max}$) e mínimo ($X_{\min}$) para transformação foram aqueles reservados para os dados da base de treinamento, por meio da equação:

$$\text{norm} = a + \frac{X - X_{\min}(b - a)}{X_{\max} - X_{\min}} \quad (1)$$

Onde $a = 0$ e $b = 1$, *norm* é a variável normalizada e X o vetor a ser normalizado. Em seguida, a normalização foi aplicada aos dados de teste. Esses dados foram utilizados também para treinamento e previsão dos Modelos Sazonais Autorregressivos Integrados de Médias Móveis com variáveis Exógenas (SARIMAX, em inglês), usados como valores de base para aferição de acurácia dos modelos de IA.

2.3 Métodos

2.3.1. Redes Neurais Artificiais

O conjunto de modelos de IA, utilizado nesta pesquisa, foram as Redes Neurais Artificiais (RNA) – que são conjuntos densamente conectados de unidades básicas de processamento denominadas de neurônio artificial – as quais fazem parte dos algoritmos de aprendizado de máquina supervisionados que são baseados na capacidade do cérebro humano em aprender determinados padrões em conjuntos de dados por meio da experiência e, por sua vez, realizar generalizações com as informações aprendidas (Haykin 2007, p. 27; Yegnanarayana 2005, p. 30). O algoritmo consiste basicamente em fornecer uma entrada (previsores) para o treinamento da RNA que efetuará, grosso modo, o somatório do produto entre as entradas e os respectivos pesos (iniciados aleatoriamente). Esse resultado passará por uma função de ativação – podendo ser uma função degrau unitário, ou sigmoide, por exemplo – que determinará se a saída será excitatória ou inibitória a depender do limiar utilizado (Rosenblatt 1958). Por se tratar de um método supervisionado, a saída é avaliada, utilizando cálculos de erros, com os alvos previamente selecionados como resultados esperados. Se o valor do erro não for satisfatório, os pesos serão atualizados e o processo inicia-se novamente.

O processo de ajuste dos pesos é chamado de *backpropagation*. Existem diversas arquiteturas de RNA's para os mais variados campos de aplicação indo desde classificação de dados, visão computacional, a processamento de linguagem natural. Neste trabalho foram usadas, para a previsão de CGM com séries temporais multivariadas, os *Perceptrons Multicamadas* (MLP) e as *Redes Neurais Convolucionais acopladas às Memórias de Longo e Curto Prazo* (CNN-LSTM). Para a implementação da modelagem de CGM com as RNAs foi utilizada a biblioteca de código aberto *TensorFlow*, desenvolvida pelo *Google* (Abadi et al. 2015).

2.3.2. Perceptron Multicamadas

O *Perceptron Multicadas* (Figura 2) é um tipo de RNA *feedforward* na qual os nós (ou neurônios) de uma camada são todos conectados aos nós da camada seguinte. O ajuste dos hiperparâmetros, isto é, a quantidade de neurônios, *batch size*, tipos de funções de ativação, número de camadas ocultas, tamanho das entradas e camadas de *dropout*, que compõem cada modelo de IA foi baseado no esquema

de validação móvel (Hu et al., 1999) com alterações de combinações simples desses parâmetros. O monitoramento de função de custo (no caso o *MSE*) foi realizado com a utilização de uma *callback* do *Tensorflow*, a *EarlyStopping*. Com essa *callback*, o treinamento é interrompido após a estagnação (intervalo previamente definido no método *fit*) do parâmetro monitorado, contribuindo para redução de *overfitting*. O *EarlyStopping* foi aplicado para as duas arquiteturas, isto é, CNN-LSTM e MLP. É importante salientar que foi encontrado um modelo com combinações específicas para cada ponto de grade (lat x lon). Tratam-se, portanto, de modelos unidimensionais e distintos entre si. As composições espaciais (2D) das previsões são interpolações dos resultados de cada modelo individual.

2.3.3. CNN-LSTM

As Redes Neurais Convolucionais (CNN, em inglês) são amplamente utilizadas para tratamento de imagens (Lawrence et al. 1997; LeCun et al. 1998). Sua aplicação na área de detecção e classificação de objetos deve-se ao fato desse tipo de RNA ser capaz de extrair características de uma matriz com a aplicação de camadas de pré-processamento, em que se efetuam convoluções entre os dados originais e um filtro (ou grupo de filtros) denominado de *kernel*, para detectar os atributos mais relevantes, reduzindo assim os ruídos. Após as convoluções é aplicado mais um filtro denominado de *Pooling* para evitar o excesso de aprendizado na base de treinamento, promovendo, dessa forma, a preservação das estruturas espaciais e a redução de dimensionalidade do problema (Brownlee 2016). O fluxo de processamento de uma CNN é esquematizado na Figura 3.

Duas camadas de convolução, cada uma com 64 filtros (de *kernel* 3x3), e uma camada de *polling* (2x2), foram utilizadas para extração de características. As saídas das CNN's compõem a camada visível das LSTM (Figura 3), que são um tipo de Redes Neurais Recorrentes (RNR), ou seja, são capazes de propagar uma cópia de si mesma para as camadas seguintes. Elas utilizam *portões* de liberação (adição) de informações na memória o que permite um aprendizado de dependências de longo prazo nas séries temporais evitando o desaparecimento do gradiente comum às RNR (Evermann, Rehse & Fettke 2017; Hochreiter & Schmidhuber 1997). Nas RNN comuns há apenas uma estrutura de repetição e só existe o estado interno h_t , expresso por:

$$h_t = f(Wx_t + Uh_{t-1} + b) \quad (2)$$

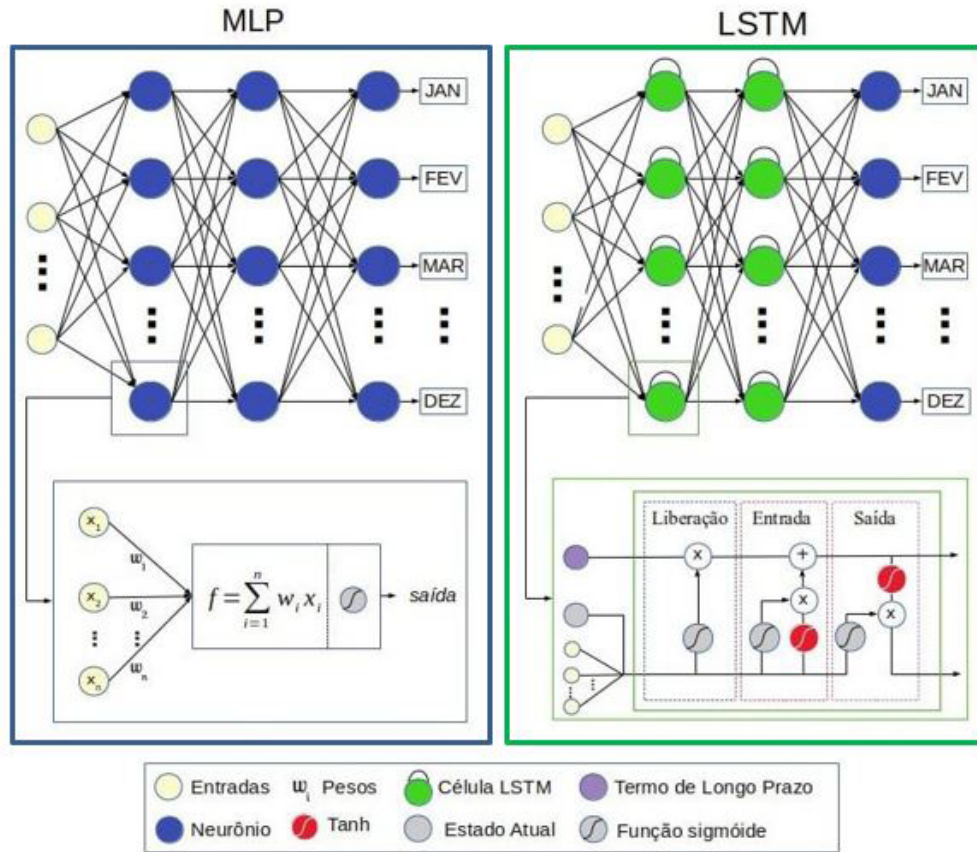


Figura 2 Esquemas das arquiteturas MLP (bloco azul à esquerda) e LSTM (bloco verde à direita) de redes neurais artificiais. No bloco azul é apresentada a rede MLP com destaque para a estrutura interna de sua unidade básica: o neurônio, com suas entradas, respectivos pesos (w_i), e função de ativação (sigmoide). No bloco verde, apresenta-se a rede neural recorrente LSTM, com destaque para sua unidade básica: a célula de memória, com respectivas entradas, portões de liberação, entrada e saída. Fonte: Adaptado de Olah (2015).

Onde f é a função de ativação, W e U são os vetores dos pesos para as entradas e para o estado escondido h_{t-1} ; x_t e b são as entradas e a unidade de bias, respectivamente. Na Figura 2 (bloco verde) nota-se o primeiro *portão* de liberação ou esquecimento. Ele controla qual parte da célula de memória C_{t-1} será esquecida:

$$f_t = f(W_f x_t + U_f h_{t-1} + b_f) \quad (3)$$

Onde f_t é o resultado do *portão* de liberação; W_f e U_f são os vetores dos pesos dentro deste *portão* e para o estado escondido h_{t-1} ; x_t são as entradas e b_f a unidade de bias. Após este *portão*, uma potencial atualização, c , para o estado atual da célula é estimada:

$$c = f(W_c x_t + U_c h_{t-1} + b_c) \quad (4)$$

Onde W_c e U_c são os pesos para entrada e para o estado escondido da potencial atualização; b_c é o *bias* para

a atualização. O segundo *portão* é o que define qual parte da informação atualizará o estado atual:

$$i_t = f(W_i x_t + U_i h_{t-1} + b_i) \quad (5)$$

Onde i_t é o resultado do *portão* de entrada; W_i e U_i são os pesos nesse *portão* para entradas e para o estado escondido, b_i é o *bias* nesse *portão*. Finalmente, o estado da memória C_t é atualizado:

$$C_t = f_t \odot C_{t-1} + i_t \odot c \quad (6)$$

2.4 SARIMAX

O modelo Autorregressivo Integrado de Médias Móveis (ARIMA, em inglês) não sazonal é uma técnica estatística que combina elementos autorregressivos (AR) – isto é, os valores da série em um dado tempo t podem ser encontrados por uma função linear dos valores passados

nos tempos $t-1, t-2, \dots, t-p$ – e de médias móveis (MA) para previsões de séries temporais estacionárias. Uma série temporal é dita estacionária quando sua média e variância são constantes ao longo do tempo, portanto, se a série for não estacionária torna-se necessária a aplicação de diferenciações (componente Integrada, I, do modelo) até que a condição de estacionariedade seja satisfeita (Box & Jenkins 1976). Cada um desses elementos é denominado de parâmetro do modelo. Estes parâmetros são definidos por *ARIMA* (p, d, q):

- **p**: Número de passos de observações passadas;
- **d**: Número de vezes que foram aplicadas diferenciações;
- **q**: Tamanho da janela da média móvel.

O modelo SARIMAX é uma extensão do ARIMA com uma componente sazonal (S) e a utilização de variáveis exógenas (séries temporais multivariadas) e pode ser definido por *SARIMAX* (p, d, q)(P, D, Q)[S], em que P corresponde ao termo autorregressivo sazonal (AR); D às diferenciações sazonais e Q representa a média móvel sazonal. Nesta pesquisa o termo S foi definido em 12 meses, e as variáveis exógenas foram as mesmas utilizadas como previsores das RNAs. A métrica adotada para seleção de cada parâmetro dos modelos pontuais SARIMAX foi o Critério de Informação de Akaike (AIC) (Akaike, 1974).

2.5 Cálculos dos Erros, Anomalias e Área de Cobertura

A avaliação quantitativa de desempenho das previsões dos modelos foi realizada por meio dos cálculos do Erro Médio Absoluto (MAE em inglês) e da Raiz do Erro Quadrático Médio (RMSE em inglês). O MAE (Wilks 2006, p. 278) é encontrado por:

$$MAE = \frac{1}{N} \sum_{i=1}^N (|y_i - x_i|) \quad (7)$$

Onde N é o número de pontos totais do experimento, y_i são os resultados previstos, e x_i os dados observados. A vantagem de se utilizar o erro absoluto em vez da simples diferença entre valores previstos e observados é que, além de evidenciar a tendência do erro, o MAE é menos afetado por erros individuais de mesma magnitude e de sinais opostos, que no somatório anulam-se (Wilks 2006, p. 279). O RMSE é definido por:

$$RMSE = \left[\frac{1}{N} \sum_{i=1}^N (y_i - x_i)^2 \right]^{1/2} \quad (8)$$

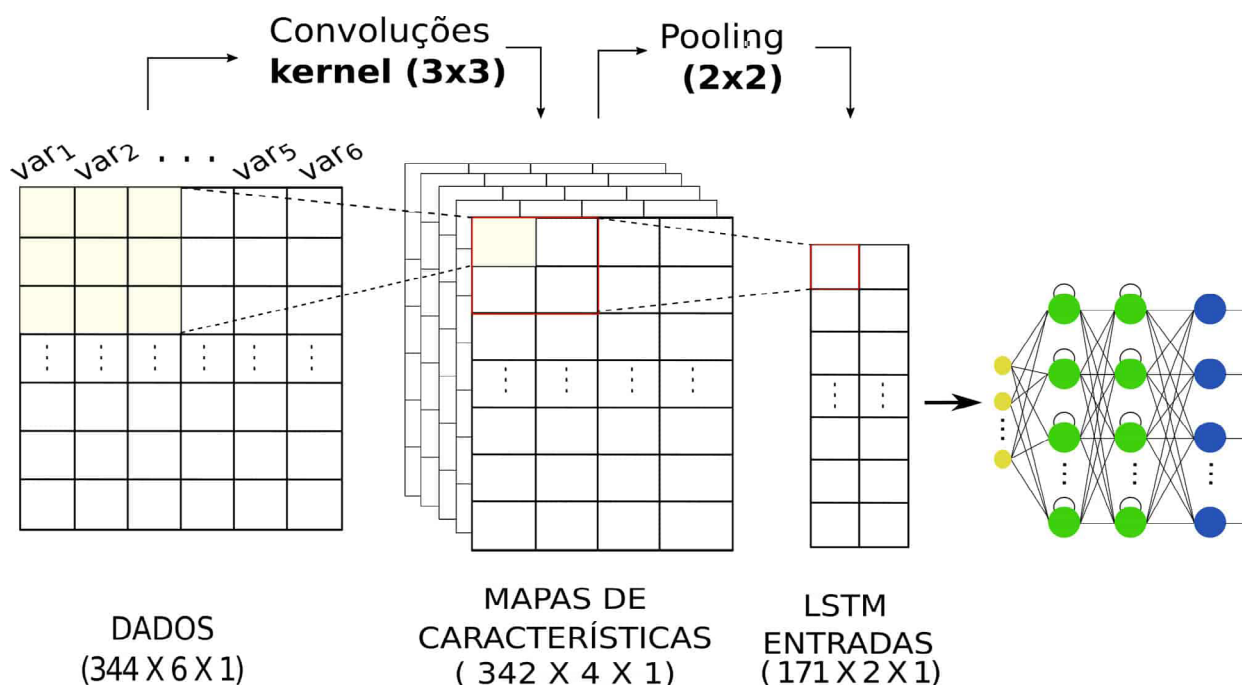


Figura 3 Modelo Esquemático das camadas de convoluções acopladas à rede LSTM. Os dados de entrada são representados por uma matriz em que cada coluna é uma série temporal de uma variável climática. Nessas entradas são aplicadas convoluções com 64 filtros **3x3** que extraem os mapas de características. Em seguida é aplicado o filtro de *pooling* que armazena os maiores valores percorrendo o mapa de característica em uma janela **2x2**. Essas informações são finalmente utilizadas como entradas para o modelo LSTM.

As anomalias de CGM foram calculadas pela diferença ponto a ponto entre a CGM mensal, sazonal e anual com a normal climatológica (1981-2010).

A área de cobertura (em milhões de km², e corresponde ao total de concentração de gelo marinho acima de 15% em cada pixel) é definida por:

$$A = \sum_{i=1}^n a_i c_i \quad (9)$$

Em que a_i é a área do i -ésimo pixel, e c_i é o valor da concentração acima de 15% para o i -ésimo pixel (Gloersen et al. 1992, p. 290).

3 Resultados e Discussões

Na Figura 4, observam-se as diferenças mensais de concentração de gelo marinho entre os valores das reanálises para o ano de 2019 e a normal climatológica (1981-2010).

Os valores negativos (positivos) representam um desvio abaixo (acima) da média. No mar de Weddell, identifica-se uma área com intensa retração de gelo, especificamente entre as latitudes 60°S e 70°S, para os meses de janeiro a junho e para o mês de dezembro. Esta retração diverge da tendência positiva de área com gelo da série temporal (1979-2019), observada na Figura 5, para o mesmo período de ablação. Nesta região os meses de julho a novembro apresentam acréscimo de gelo entre as latitudes 55°S e 60°S. As anomalias positivas, nas bordas do mar de Weddell, deste período não são compensadas por nenhuma anomalia negativa. Há um aumento, portanto, de concentração de gelo no período de formação.

Este mesmo período de aumento em concentração em 2019 não segue a redução de cobertura da área de gelo, que pode ser analisado por meio da Figura 5, na qual exibe tendências negativas para a área com gelo. No mar de Bellingshausen, os meses de janeiro, fevereiro, e o período de setembro a novembro apresentam intensas anomalias positivas, contrastando com as linhas de tendência, as quais para esse mar eram negativas, exceto para setembro (Figura 6). No mar de Amundsen, as anomalias positivas de concentração são observadas apenas de janeiro a junho, alcançando os maiores valores (>50%) em fevereiro e maio. De julho a dezembro são apresentadas as maiores retrações, com o mês de setembro atingindo valores menores que -50% (Figura 4). O mesmo padrão de inversões entre as linhas de tendência (1979-2019) de extensão e as anomalias de concentração de gelo marinho

(2019) observados em Weddell e Bellingshausen aparece no mar de Amundsen.

São apresentados na Tabela 2 os valores dos erros de CGM estimados pelos modelos estudados no mar de Weddell. Para esta região, os menores erros encontrados de janeiro a junho e dezembro foram obtidos com a CNN-LSTM, para os meses de julho, setembro e outubro, nos quais a CNN-LSTM foi superada pelo MLP e/ou SARIMAX, a diferença entre os erros (*MAE* e *RMSE*) não fica acima de 5%, sendo que nesses meses não houve anomalias de concentração superior (inferior) a 25% (-25%), como apresentado na Figura 4, implicando afirmar que: os períodos de maiores anomalias foram melhor estimados utilizando-se os recursos de acoplamento CNN-LSTM. Em maio, por exemplo, o *MAE* ficou 10% menor que o encontrado com MLP e 20% menor que o do SARIMAX.

3.1 Mar de Weddell

Na Figura 7, estão as distribuições espaciais das diferenças entre o modelo CNN-LSTM e as reanálises. Para os meses de janeiro a março, observa-se que o modelo foi capaz de generalizar as fortes anomalias negativas desta região, superestimando (valores negativos) a concentração entre 15% e 18%.

As intensas anomalias negativas de abril e maio, observadas na Figura 4, são refletidas nas previsões do modelo como regiões de superestimava, cabendo ressaltar que a diferença entre real e observado ficou entre 15% e 21%. Em junho e dezembro, as fortes retrações em Weddell (Figura 4) não atingiram variações absolutas superiores a 15% entre o estimado e o real, evidenciando a capacidade adaptativa do modelo para previsão de concentração de gelo em eventos extremos. Na Figura 8, observam-se as diferenças entre os valores estimados com o modelo MLP e os valores das reanálises.

Em janeiro, assim como ocorreu com o CNN-LSTM, houve uma boa generalização das estimativas, que pode ser percebida pela não coincidência entre as linhas da normal (linha roxa) com a prevista (linha vermelha). De fevereiro a maio, o comportamento da não coincidência entre as curvas da normal com o previsto repete-se, embora, as estruturas de intensas anomalias negativas presentes na Figura 4, para estes meses, estejam com diferenças entre -0,18 e -0,36. A anomalia negativa em junho no mar de Weddell também é visível com o modelo MLP, com diferenças menores que -0,3 estrutura esta ainda que presente nas previsões CNN-LSTM (Figura 7) apresentaram valores não inferiores a -0,15 para este último modelo.

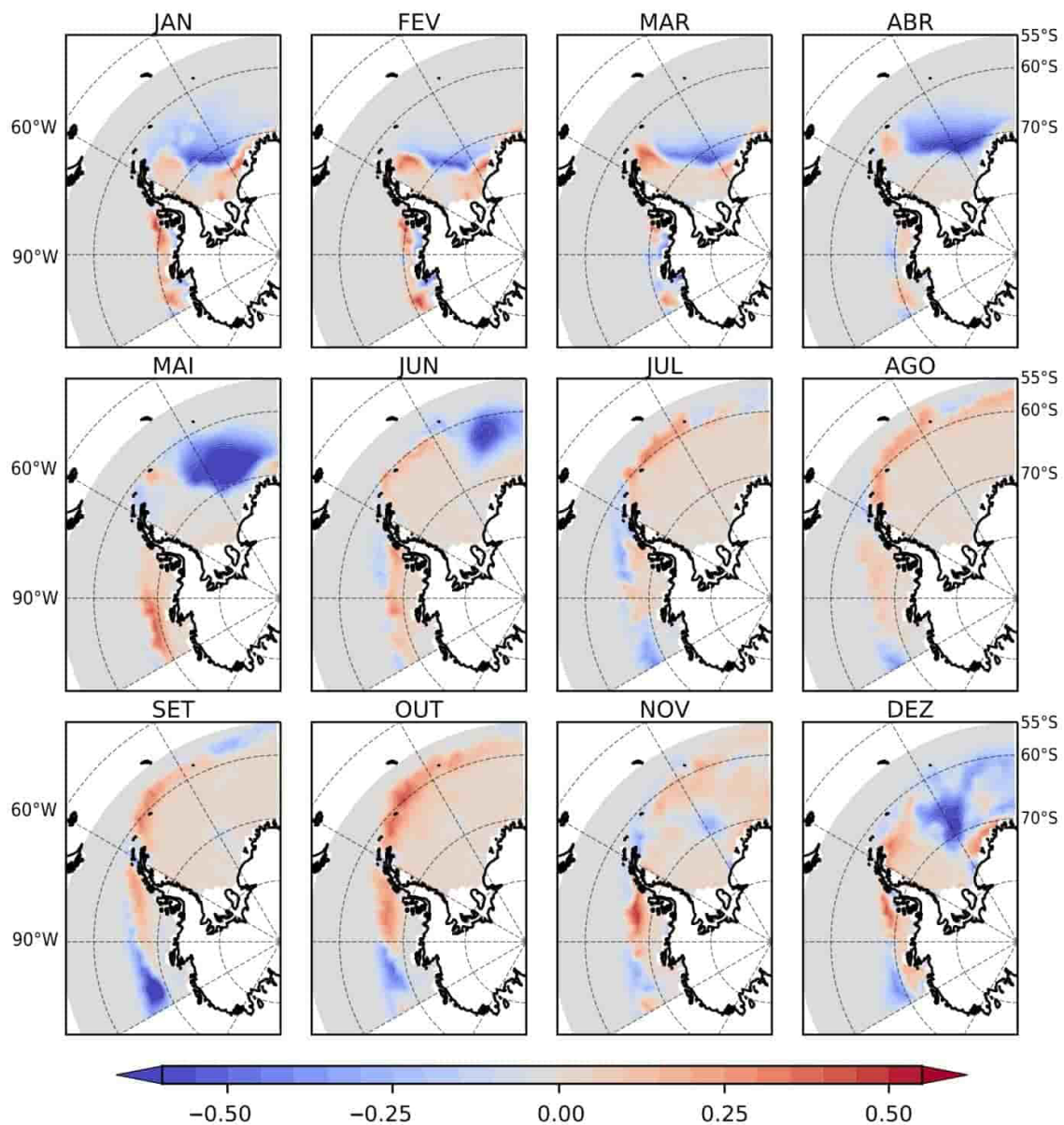


Figura 4 Diferenças mensais entre os valores das reanálises em 2019 e a normal climatológica (1981-2010) de concentração de gelo marinho. Valores em azul (vermelho) representam retração (acresção) de gelo.

Em julho e setembro a rede MLP apresentou os melhores resultados (Tabela 2) de *MAE* 0,043 e 0,049, no entanto ao comparar com os outros dois modelos este desempenho ficou 4% superior ao CNN-LSTM e 6% ao SARIMAX. Estes valores próximos podem estar relacionados ao padrão dos dados em 2019 com distribuições espaciais de anomalias absolutas abaixo de 25% (Figura 4) em uma região bem delimitada nos limites do mar de Weddell. Implicando afirmar que julho, agosto, setembro e novembro foram meses de estabilidade em relação à normal. A Figura 9 apresenta as diferenças entre os valores das reanálises e dos estimados

com o modelo SARIMAX. No mar de Weddell todas as anomalias negativas de janeiro a junho nas reanálises não foram bem estimadas, com diferenças atingindo -0,45. Neste período, há uma superposição entre as linhas de gelo estimadas (vermelhas) e da normal (roxas), sendo que, para alguns meses como janeiro, há uma superestimação dessas linhas, indicando que o modelo não só superestimou os dados como também previu concentrações acima da normal climatológica. Em outubro o modelo SARIMAX apresentou o menor *MAE*, de 0,57, sendo este valor apenas 4% melhor que o obtido com a CNN-LSTM.

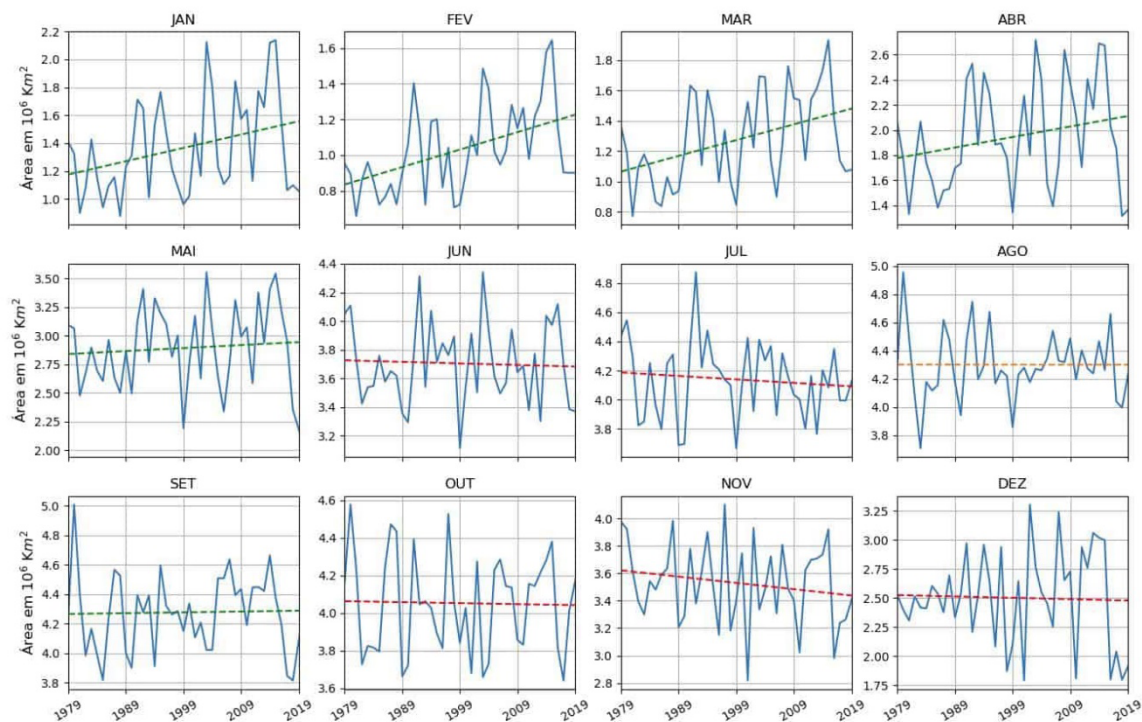


Figura 5 Área de cobertura mensal de gelo marinho, concentração acima de 15%, em milhões de km^2 no período de 1979-2019 para o mar de Weddell. As linhas tracejadas são as linhas de tendência linear: linhas verdes para tendências positivas e vermelhas para negativas.

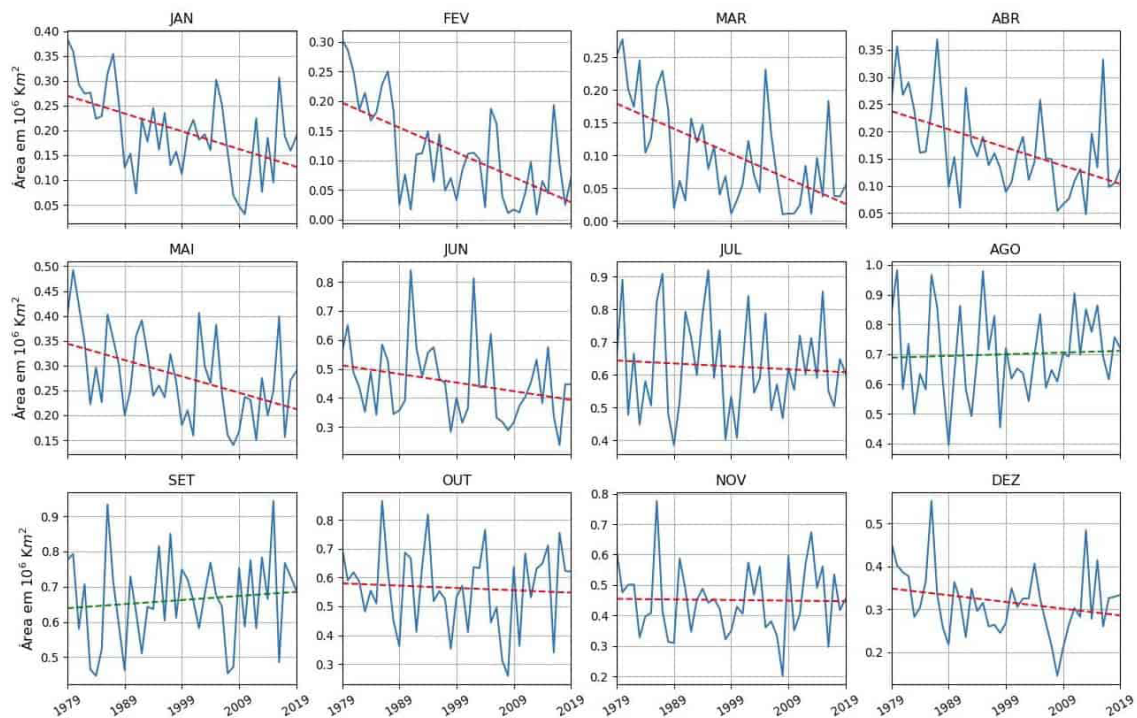


Figura 6 Área de cobertura mensal de gelo marinho, concentração acima de 15%, em milhões de km^2 no período de 1979-2019 para o mar de Bellingshausen. As linhas tracejadas são as linhas de tendência linear: linhas verdes para tendências positivas e vermelhas para negativas.

Tabela 2 Valores mensais, sazonais e anual de MAE e RMSE de concentração de gelo marinho para o mar de Weddell, em 2019, encontrados com os modelos de RNA e SARIMAX. Os menores erros estão destacados em azul.

	WEDDELL					
	CNN-LSTM		MLP		SARIMAX	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
ANNUAL	0,066	0,119	0,076	0,132	0,099	0,172
VERÃO	0,044	0,092	0,064	0,141	0,113	0,195
OUTONO	0,089	0,165	0,099	0,167	0,150	0,268
INVERNO	0,050	0,094	0,061	0,107	0,058	0,103
PRIMAVERA	0,059	0,099	0,060	0,101	0,058	0,105
JANEIRO	0,043	0,080	0,060	0,014	0,125	0,205
FEVEREIRO	0,045	0,104	0,069	0,138	0,101	0,184
MARÇO	0,057	0,110	0,068	0,131	0,111	0,209
ABRIL	0,081	0,154	0,092	0,155	0,158	0,279
MAIO	0,128	0,216	0,141	0,222	0,181	0,316
JUNHO	0,080	0,153	0,089	0,161	0,082	0,159
JULHO	0,045	0,074	0,043	0,082	0,046	0,072
AGOSTO	0,043	0,069	0,051	0,085	0,046	0,078
SETEMBRO	0,051	0,082	0,049	0,080	0,052	0,089
OUTUBRO	0,059	0,108	0,065	0,115	0,057	0,110
NOVEMBRO	0,066	0,106	0,066	0,105	0,068	0,115
DEZEMBRO	0,094	0,144	0,118	0,177	0,165	0,250

Na Figura 10 estão os valores previstos e as reanálises de cobertura de gelo. O período de janeiro a maio, no mar de Weddell, foi o que apresentou as maiores diferenças entre o esperado e previsto. O modelo SARIMAX foi o que apresentou as maiores superestimativas atingindo valores de $1,412 \times 10^6 \text{ km}^2$ em maio, contra $0,480 \times 10^6 \text{ km}^2$ e $0,719 \times 10^6 \text{ km}^2$ do CNN-LSTM e MLP respectivamente.

3.2 Mares de Bellingshausen e Amundsen

Na Tabela 3, podem-se analisar os erros anuais, sazonais e mensais dos modelos regressivos de IA. Anualmente, o melhor resultado foi o da CNN-LSTM, com MAE 14% e 29% inferiores ao MLP e SARIMAX nesta ordem. Sazonalmente, o modelo CNN-LSTM foi o que obteve os menores erros em todas as estações. Em abril, o modelo SARIMAX superou (pelo MAE) o CNN-LSTM em 3%. Em janeiro e fevereiro, o padrão de anomalia positiva (Figura 4) contrastando com a tendência de redução (Figura 6) da série temporal foi espacialmente melhor estimada com o CNN-LSTM como pode ser identificado na Figura 7. As outras três anomalias positivas (> 30%) de outubro a dezembro também tiveram melhores previsões com o modelo CNN-LSTM, tanto em distribuição espacial

de diferenças quanto em MAE e RMSE. Na Figura 10, o menor valor de área de cobertura (reanálises) foi em março com $0,071 \times 10^6 \text{ km}^2$ e o maior em agosto com $0,683 \times 10^6 \text{ km}^2$. Em relação aos outros mares, a cobertura de gelo em Bellingshausen foi a que obteve as melhores estimativas, com os piores valores em setembro o qual apresentou diferenças absolutas entre observado e previsto de $0,027 \times 10^6 \text{ km}^2$ para o CNN-LSTM, $0,047 \times 10^6 \text{ km}^2$ para o MLP e $0,085 \times 10^6 \text{ km}^2$ para o SARIMAX.

Os meses com as maiores anomalias positivas, acima de 25%, também foram melhor previstos com a CNN-LSTM, o que pode ser confirmado não somente pelos menores erros (Tabela 3) como também pelas distribuições espaciais das diferenças entre estimados e esperados (Figuras 7, 8 e 9). Em setembro houve o evento com menores taxas de concentração (Figura 4) em Amundsen, mais uma vez a arquitetura CNN-LSTM superou os outros modelos na previsão dessa estrutura. Cabe ressaltar que esse evento extremo foi o único que apresentou menor diferença de desempenho entre as previsões, sendo o MAE e o RMSE do CNN-LSTM apenas 5% inferior ao MLP e 9% ao SARIMAX, em contrapartida, para estes dois últimos modelos as curvas de linha de gelo (> 15%) do previsto e da normal foram mais coincidentes que o primeiro.

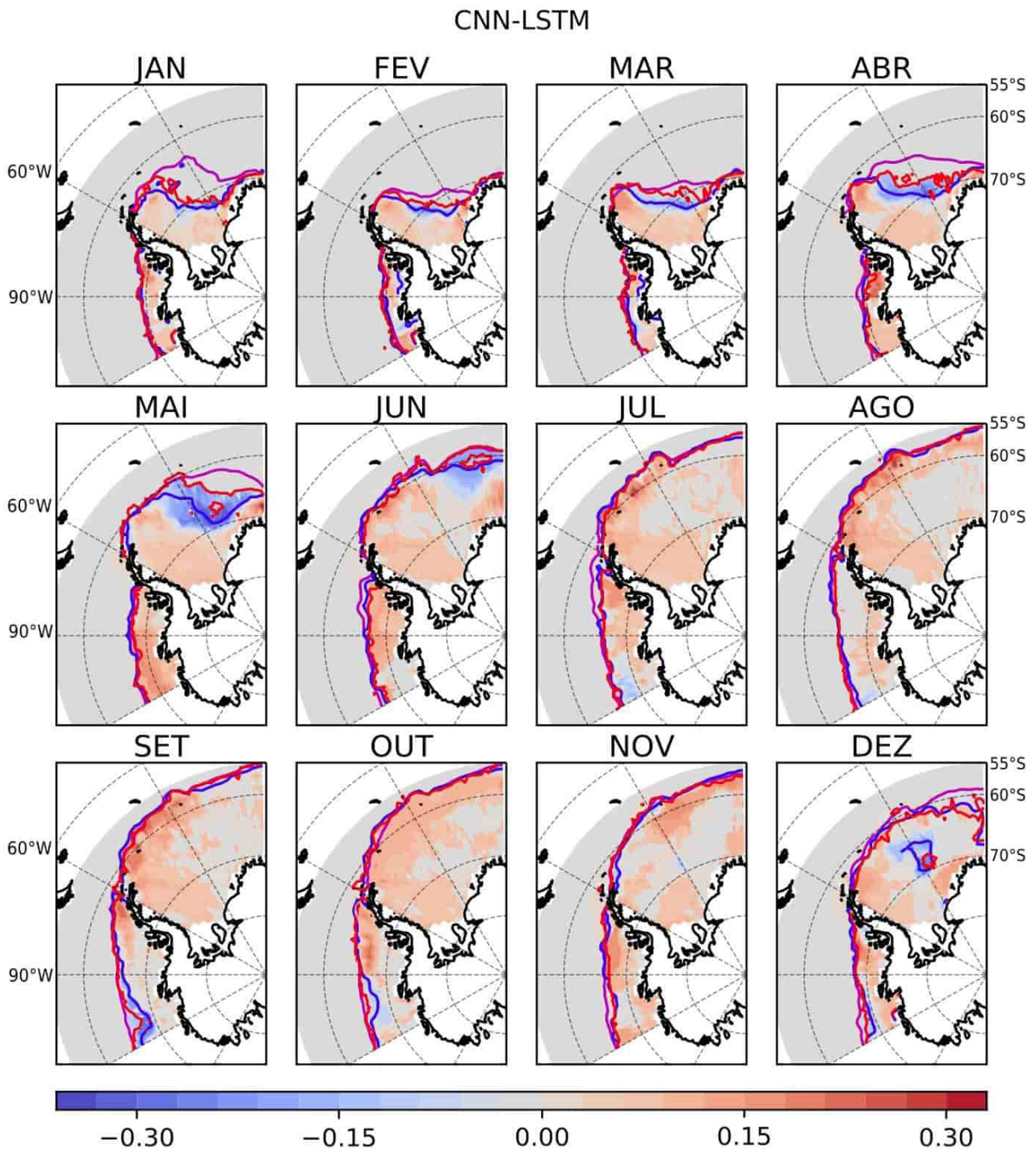


Figura 7 Distribuições espaciais das diferenças de concentração de gelo marinho entre os valores das reanálises e os estimados com CNN-LSTM. As linhas azuis, vermelhas e roxas representam, respectivamente, os valores de concentração maiores que 15% dos das reanálises (2019), estimados (2019) e da normal climatológica (1981-2010).

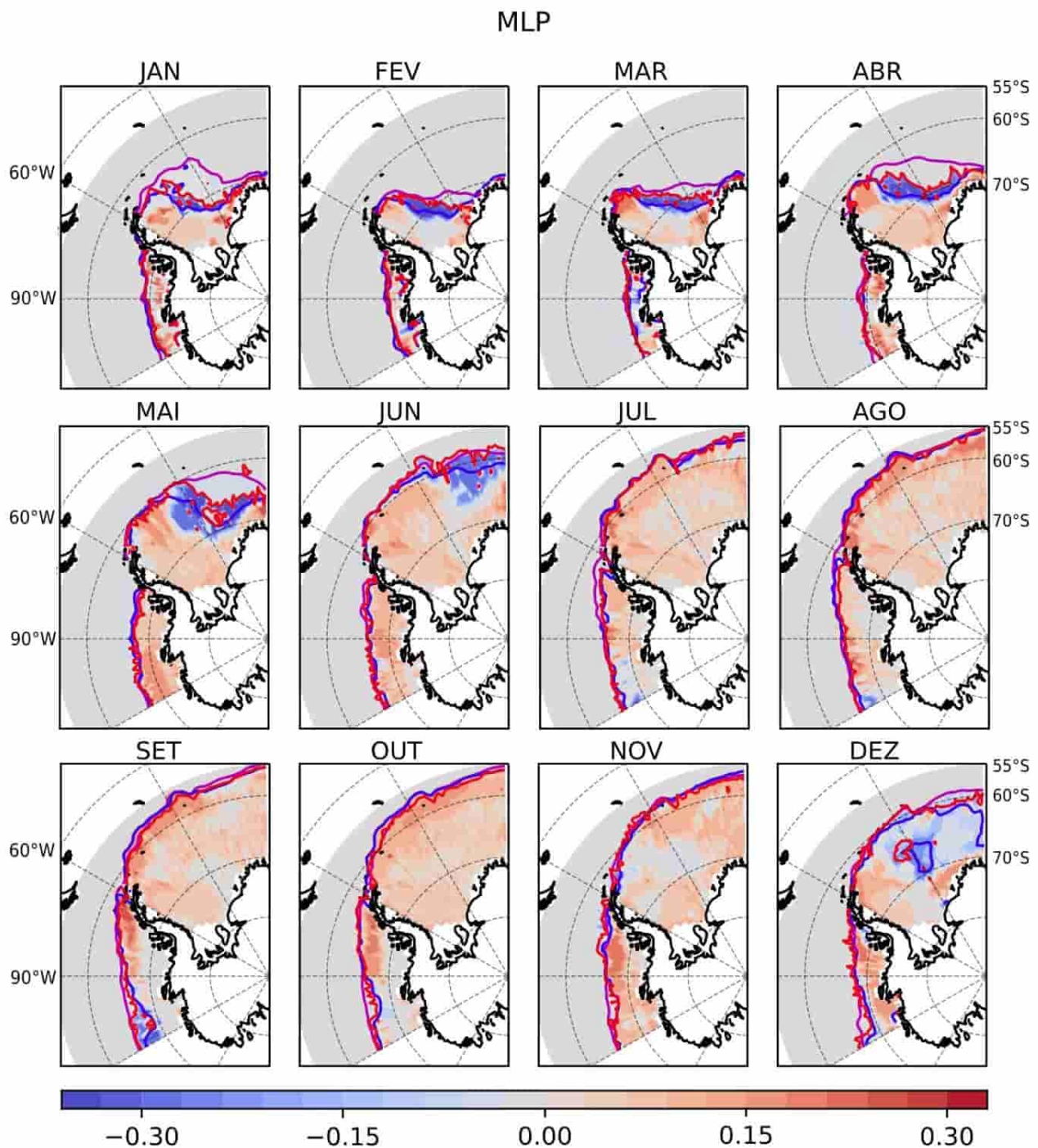


Figura 8 Distribuições espaciais das diferenças de concentração de gelo marinho entre os valores das reanálises e os estimados com MLP. As linhas azuis, vermelhas e roxas representam, respectivamente, os valores de concentração maiores que 15% nas reanálises (2019), estimados (2019) e da normal climatológica (1981-2010).

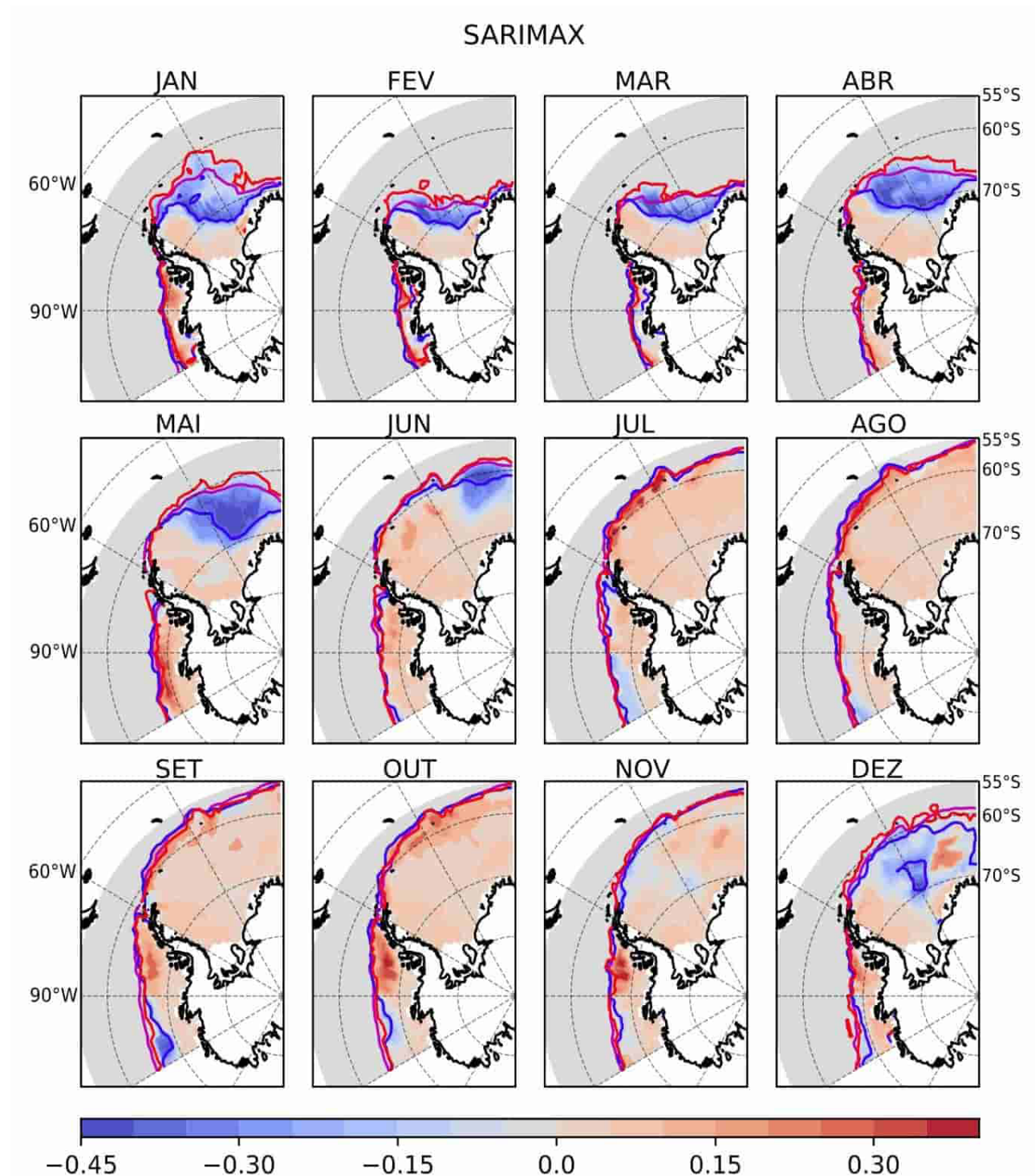


Figura 9 Distribuições espaciais das diferenças de concentração de gelo marinho entre os valores observados e os estimados com SARIMAX. As linhas azuis, vermelhas e roxas representam, respectivamente, os valores de concentração maiores que 15% dos dados observados (2019), estimados (2019) e da normal climatológica (1981-2010).

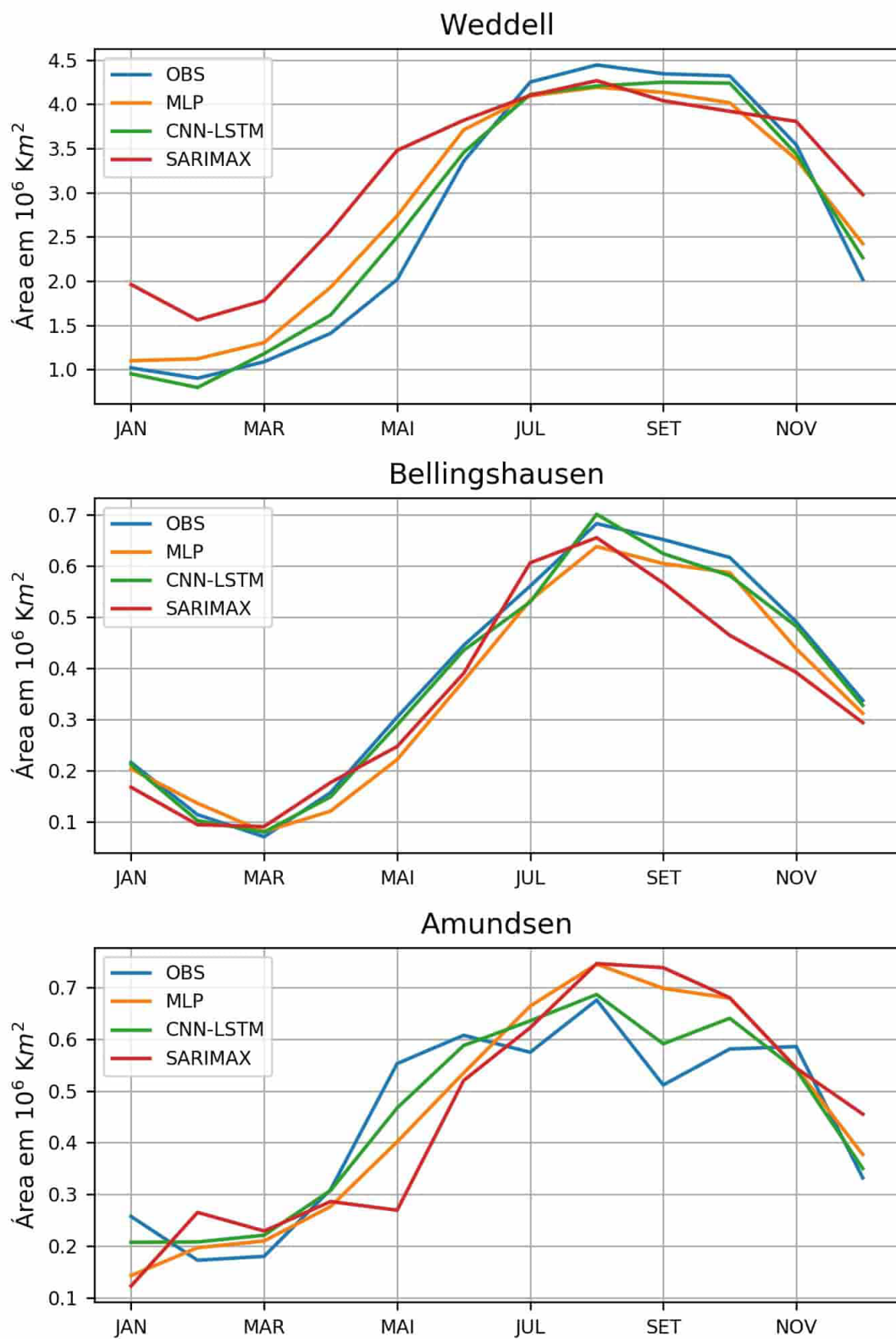


Figura 10 Área de gelo marinho observada e estimada, em 2019, em milhões de km², para os mares de Weddell, Bellingshausen e Amundsen.

Tabela 3 Valores mensais, sazonais e anual de MAE e RMSE de concentração de gelo marinho para os mares de Bellingshausen e Amundsen, em 2019, encontrados com os modelos de RNA e SARIMAX. Os menores erros estão destacados em azul.

	BELLINGSHAUSEN/AMUNDSEN					
	CNN-LSTM		MLP		SARIMAX	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
ANUAL	0,070	0,110	0,082	0,115	0,100	0,161
VERÃO	0,063	0,105	0,071	0,108	0,114	0,202
OUTONO	0,076	0,114	0,089	0,129	0,097	0,164
INVERNO	0,064	0,095	0,070	0,102	0,082	0,121
PRIMAVERA	0,077	0,116	0,081	0,118	0,111	0,168
JANEIRO	0,061	0,098	0,071	0,097	0,108	0,191
FEVEREIRO	0,064	0,111	0,072	0,118	0,122	0,219
MARÇO	0,052	0,090	0,076	0,115	0,096	0,171
ABRIL	0,068	0,109	0,069	0,111	0,066	0,116
MAIO	0,107	0,144	0,122	0,161	0,131	0,206
JUNHO	0,062	0,090	0,076	0,104	0,102	0,155
JULHO	0,070	0,106	0,089	0,101	0,084	0,118
AGOSTO	0,058	0,089	0,069	0,104	0,060	0,092
SETEMBRO	0,079	0,127	0,086	0,129	0,107	0,163
OUTUBRO	0,067	0,101	0,071	0,101	0,106	0,153
NOVEMBRO	0,085	0,122	0,087	0,124	0,120	0,187
DEZEMBRO	0,071	0,117	0,096	0,125	0,103	0,157

Em se tratando de área com gelo, o menor valor observado foi de $0,172 \times 10^6 \text{ km}^2$ em fevereiro, enquanto o máximo foi em agosto com $0,676 \times 10^6 \text{ km}^2$. Foi em maio que se atingiram as maiores subestimativas sendo as diferenças entre as reanálises e estimados de $0,084 \times 10^6 \text{ km}^2$ para o CNN-LSTM, $0,151 \times 10^6 \text{ km}^2$ para o MLP, e $0,280 \times 10^6 \text{ km}^2$ com o SARIMAX. Em setembro, em Amundsen, todos os modelos superestimaram a cobertura de gelo em $0,079 \times 10^6 \text{ km}^2$ para o CNN-LSTM, em $0,187 \times 10^6 \text{ km}^2$ para o MLP, e $0,227 \times 10^6 \text{ km}^2$ para o SARIMAX.

4 Conclusões

Este estudo teve como objetivo avaliar o potencial de previsibilidade de CGM com a aplicação de técnicas de RNAs. Analisando apenas os erros de CGM para todos os mares estudados observou-se que a CNN-LSTM foi superada somente nos meses com anomalias inferiores a $\pm 25\%$, sendo que as diferenças máximas entre os erros (RMSE e MAE) nestes eventos não atingiram valores acima de 5%, como o que ocorreu nos meses de julho, setembro e outubro no mar de Weddell. Com estes resultados foi possível inferir que o modelo CNN-LSTM foi o que

previu com maior precisão os períodos com as maiores anomalias de CGM. Com as análises das distribuições espaciais das diferenças de CGM pode-se afirmar que as delimitações dos campos anômalos também foram melhores previstas com o acoplamento CNN-LSTM, por exemplo, a anomalia negativa de junho no mar de Weddell ($MAE < 0,15$) e a positiva, em setembro, no mar de Amundsen ($MAE < 0,10$).

Como consequência dos valores previstos para as CGM as áreas de cobertura que atingiram os melhores resultados também foram as obtidas com CNN-LSTM. No mar de Weddell, como exemplo, de janeiro a maio todos os modelos superestimaram a área de cobertura de gelo marinho, no entanto, enquanto em maio as diferenças entre real e estimado foram de aproximadamente $0,5 \times 10^6 \text{ km}^2$ para a rede Convolutiva, de $0,75 \times 10^6 \text{ km}^2$ para a MLP e de $1,5 \times 10^6 \text{ km}^2$ para o modelo SARIMAX.

Portanto, com esta pesquisa foi possível avaliar a precisão de duas arquiteturas de RNAs, as CNN-LSTM e MLP, na previsão de séries temporais multivariadas de gelo marinho, para o ano de 2019, nos mares de Weddell, Amundsen e Bellingshausen, que correspondem a 1/3 da região marítima da Antártica Ocidental.

5 Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Código de Financiamento 001.

6 Referências

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, T., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y. & Zheng, X. 2015, *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. <https://arxiv.org/pdf/1603.04467.pdf>
- Akaike, H. 1974, 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716-23. <https://doi.org/10.1109/TAC.1974.1100705>
- Armour, K.C., Scott, J., Donohoe, A., Newsom, E.R. & Marshall J.C. 2016, 'Southern Ocean warming delayed by circumpolar upwelling and equatorward transport', *Nature Geoscience*, vol. 9, no. 7, pp. 549-54. <https://doi.org/10.1038/ngeo2731>
- Boetius, A., Anesio, A.M., Deming, J.W., Mikucki, J.A. & Rapp, J.Z. 2015, 'Microbial ecology of the cryosphere: sea ice and glacial habitats', *Nature Reviews Microbiology*, vol. 13, no. 11, pp. 677-90. <https://doi.org/10.1038/nrmicro3522>
- Box, G.E.P. & Jenkins, G.M. 1976, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, CA.
- Breiman, L. 2001, 'Random forests', *Machine learning*, vol. 45, no. 1, pp. 5-32. <https://doi.org/10.1023/A:1010933404324>
- Brownlee, J. 2016, *Deep Learning with Python: Develop Deep Learning Models on Theano and TensorFlow using Keras*, Machine Learning Mastery.
- Cavaleri, D.J., Gloersen, P., Parkinson, C.L., Comiso, J.C. & Zwally, H.J. 1997, 'Observed hemispheric asymmetry in global sea ice changes', *Science*, vol. 278, no. 5340, pp. 1104-6. <https://doi.org/10.1126/science.278.5340.1104>
- Cavaleri, D. & Parkinson, C.L. 2012, 'Arctic sea ice variability and trends, 1979-2010', *The Cryosphere*, vol. 6, no. 4, pp. 881-9. <https://doi.org/10.5194/tc-6-881-2012>
- Chambault, P., Albertsen, C.M., Patterson, T.A., Hansen, R.G., Tervo, O., Laidre, K.L. & Heide-Jørgensen, M.P. 2018, 'Sea surface temperature predicts the movements of the Arctic cetacean: the bowhead whale', *Scientific reports*, vol. 8, no. 1, pp. 1-12. <https://doi.org/10.1038/s41598-018-27966-1>
- Chemke, R. & Polvani, L.M. 2020, 'Using multiple large ensembles to elucidate the discrepancy between the 1979-2019 modeled and observed Antarctic Sea ice trends', *Geophysical Research Letters*, vol. 47, no. 15, pp. e2020GL088339. <https://doi.org/10.1029/2020GL088339>
- Chen, J., Li, M. & Wang, W. 2012, 'Statistical uncertainty estimation using Random Forests and its application to drought forecast', *Mathematical Problems in Engineering*, vol. 2012, no. 915053. <https://doi.org/10.1155/2012/915053>
- Evermann, J., Rehse, J.R. & Fettke, P. 2017, *XES Tensorflow: Process prediction using the Tensorflow deep-learning framework*, ArXiv, vol. 1. <https://arxiv.org/pdf/1705.01507.pdf>
- Gagné, M., Gillett, N. & Fyfe, P. 2015, 'Observed and simulated changes in Antarctic sea ice extent over the past 50 years', *Geophysical Research Letters*, vol. 42, no. 1, pp. 90-5. <https://doi.org/10.1002/2014GL062231>
- Gloersen, P., Campbell, W.J., Cavalieri, D.J., Comiso, J.C., Parkinson, C.L. & Zwally, H.J. (eds) 1992, *Arctic and antarctic sea ice: satellite passive-microwave observations and analysis*, NASA, Washington D.C.
- Haykin, S. 2007, *Redes Neurais: Princípios e Prática*, 2nd edn, Bookman, Porto Alegre, RS.
- Ho, T.K. 1995, 'Random decision forests', *3rd international conference on document analysis and recognition 1995*, IEEE, Murray Hill, NJ, pp. 278-82. <https://doi.org/10.1109/ICDAR.1995.598994>
- Hochreiter, S. & Schmidhuber, J. 1997, 'Long Short-Term Memory', *Neural computation*, vol. 9, no. 8, pp. 1735-80. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hutchinson, D.K., England, M.H., Santoso, A. & Hogg, A.M. 2013, 'Interhemispheric asymmetry in transient global warming: The role of Drake Passage', *Geophysical Research Letters*, vol. 40, no. 8, pp. 1587-93. <https://doi.org/10.1002/grl.50341>
- Hu, M.Y., Zhang, G., Jiang, C.X. & Patuwo, B.E. 1999, 'A Cross-Validation Analysis of Neural Network Out-of-Sample Performance in Exchange Rate Forecasting', *Decision Sciences*, vol. 30, no. 1, pp. 197-216. <https://doi.org/10.1111/j.1540-5915.1999.tb01606.x>
- Kirchmeier-Young, M.C., Zwiers, F.W. & Gillett, N.P. 2017, 'Attribution of extreme events in Arctic sea ice extent', *Journal of Climate*, vol. 30, no. 2, pp. 553-71. <https://doi.org/10.1175/JCLI-D-16-0412.1>
- Lawrence, S., Giles, C.L., Tsoi, A.C. & Back, A.D. 1997, 'Face recognition: A convolutional neural-network approach', *IEEE transactions on neural networks*, vol. 8, no. 1, pp. 98-113. <https://doi.org/10.1109/72.554195>
- LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. 1998, 'Gradient-based learning applied to document recognition', *Proceedings of the IEEE*, vol. 86, no. 21, pp. 2278-324. <https://doi.org/10.1109/5.726791>
- Meehl, G.A., Arblaster, J.M., Chung, C.T.Y., Holland, M.M., DuVivier, A., Thompson, L., Yang, D. & Bitz, C.M. 2019, 'Sustained ocean changes contributed to sudden Antarctic ice retreat in late 2016', *Nature Communications*, vol. 10, no. 1, pp. 1-9. <https://doi.org/10.1038/s41467-018-07865-9>
- Olah, C. 2015, *Understanding LSTM Networks*, Colah's Blog, weblog, viewed 13 July 2021, <<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>>
- Oliva, M., Navarro, F., Hrbáček, F., Hernández, A., Nýlt, D., Pereira, P., Ruiz-Fernández, J. & Trigo, R. 2017, 'Recent regional climate cooling on the Antarctic Peninsula and

- associated impacts on the cryosphere', *Science of The Total Environment*, vol. 580, pp. 210-23. <https://doi.org/10.1016/j.scitotenv.2016.12.030>
- Parkinson, C.L. 2002, 'Trends in the length of the Southern Ocean sea-ice season', *Annals of Glaciology*, vol. 34, no. 1, pp. 435-40. <https://doi.org/10.3189/172756402781817482>
- Parkinson, C.L. 2019, 'A 40-y record reveals gradual Antarctic Sea ice increases followed by decreases at rates far exceeding the rates seen in the Arctic', *Proceedings of the National Academy of Sciences*, vol. 116, no. 29, pp. 14414-23. <https://doi.org/10.1073/pnas.1906556116>
- Raphael, M.N., Marshall, G.J., Turner, J., Fogt, R.L., Schneider, D., Dixon, D.A., Hosking, J.S., Jones, J.M. & Hobbs, W.R. 2016, 'The Amundsen Sea Low: Variability, Change and Impact on Antarctic climate', *Bulletin of the American Meteorological Society*, vol. 97, no. 1, pp. 111-21. <https://doi.org/10.1175/BAMS-D-14-00018.1>
- Rosenblatt, F. 1958, 'The Perceptron: A probabilistic model for information storage and organization in the brain', *Psychological Review*, vol. 65, no. 6, pp. 386-408. <https://doi.apa.org/doi/10.1037/h0042519>
- Rouch, L.A., Dörr, J., Holmes, C.R., Massonnet, F., Blockley, E.W., Notz, D., Rackwö, T., Raphael, M.N., O'Farrell, S.P., Bailey, D.A. & Bitz, C.M. 2020, 'Antarctic Sea ice area in CMIP6', *Geophysical Research Letters*, vol. 47, no. 9, pp. e2019GL086729. <https://doi.org/10.1029/2019GL086729>
- Serreze, M.C., Holland, M.M. & Stroeve, J. 2007, 'Perspectives on the Arctic's Shrinking sea-ice cover', *Science*, vol. 315, no. 5818, pp. 1533-36. <https://doi.org/10.1126/science.1139426>
- Strobl, C., Boulesteix, A.L., Zeileis, A., & Hothorn, T. 2007, 'Bias in random forest variable importance measures: Illustrations, sources and a solution', *BMC bioinformatics*, vol. 8, no. 1, pp. 1-21. <https://doi.org/10.1186/1471-2105-8-25>
- Stroeve, J.C., Kattsov, V., Barret A., Serreze, M., Pavlova, T., Holland, M. & Meier, W. 2012, 'Trends in Arctic Sea ice extent from CMIP5, CMIP3 and observations', *Geophysical Research Letters*, vol. 39, no. 16, pp. L16502. <https://doi.org/10.1029/2012GL052676>
- Turner, J., Lu, H., White, I., King, J.C., Phillips, T., Scott H.J., Bracegirdle, T.J., Marshall, G.J., Mulvaney, R. & Deb, P. 2016, 'Absence of 21st century warming on Antarctic Peninsula consistent with natural variability', *Nature*, vol. 535, no. 7612, pp. 411-5. <https://doi.org/10.1038/nature18645>
- Turner, J., Guarino, M.V., Arnatt, J., Jena, B., Marshall, G.J., Phillips, T., Bajish, C.C., Clem, K., Wang, Z., Andersson, T., Murphy, E.J. & Cavanagh, R. 2020, 'Recent Decrease of Summer Sea Ice in the Weddell Sea, Antarctica', *Geophysical Research Letters*, vol. 47, no. 11, pp. e2020GL087127. <https://doi.org/10.1029/2020GL087127>
- Yegnanarayana, B. 2005, *Artificial Neural Networks*, Prentice-Hall of India, Nova Deli.
- Young, T., Hazarika, D., Poria, S. & Cambria, E. 2018, 'Recent trends in deep learning based natural language processing', *IEEE Computational intelligence magazine*, vol. 13, no. 3, pp. 55-75. <https://arxiv.org/abs/1708.02709>
- Vaughan, D.G., Marshall, G.J., Connolley, W.M., Parkinson, C., Mulvaney, R., Hodgson, D.A., King, J.C., Pudsey, C.J. & Turner, J. 2003, 'Recent rapid regional climate warming on the Antarctic Peninsula', *Climate Change*, vol. 60, no. 3, pp. 243-74. <https://doi.org/10.1023/A:1026021217991>
- Wilks, D.S. 2006, *Statistical Methods in the Atmospheric Sciences*, 2nd edn, Academic Press, London.

Contribuições dos Autores

Ricardo Bruno de Araújo Tenorio: Conceituação; análise formal; metodologia; validação; redação – rascunho original; visualização.
José Henrique Fernandez: Análise formal; validação; redação – rascunho original; redação – revisão e edição; supervisão; visualização.
David Mendes: Conceituação; análise formal; metodologia; validação; supervisão; visualização. **José Pedro da Silva Júnior:** Conceituação; análise formal; metodologia; validação; redação – rascunho original; visualização.

Conflito de interesse

Os autores não possuem potenciais conflitos de interesse.

Declaração de disponibilidade de dados

Os dados do modelo estão disponíveis gratuitamente mediante solicitação. Os conjuntos de dados de referência podem ser baixados de: <https://www>.

Como citar:

Tenorio, R.B.A., Fernandez, J.H., Mendes, D. & Silva Júnior, J.P. 2022, 'Estimação da Cobertura de Gelo Marinho nos Mares Antárticos de Weddell, Be-lingshausen e Amundsen com Redes Neurais Artificiais', *Anuário do Instituto de Geociências*, 45:40763. https://doi.org/10.11137/1982-3908_45_40763

[ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5](https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5). Scripts e código estão disponíveis mediante solicitação.

Todos os dados incluídos neste estudo estão disponibilizados publicamente na literatura.

Financiamento

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Editora chefe

Dra. Claudine Dereczynski

Editora Associada

Dra. Fernanda Vasconcellos