



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

**PREVISÃO E MONITORAMENTO DE EVENTOS EXTREMOS
CLIMÁTICOS SOBRE O BRASIL USANDO INTELIGÊNCIA
ARTIFICIAL**

Gabriele Gonçalves Vieira

Relatório de Iniciação Científica do
Programa PIBIC, orientada por
Juliana A. Anochi e Marília H. Shimizu

2023

**PREVISÃO E MONITORAMENTO DE EVENTOS EXTREMOS
CLIMÁTICOS SOBRE O BRASIL USANDO INTELIGÊNCIA
ARTIFICIAL**

Gabriele Gonçalves Vieira

Relatório de Iniciação Científica do
Programa PIBIC, orientada por
Juliana A. Anochi e Marília H. Shimizu

INPE
São José dos Campos

RESUMO

Eventos climáticos extremos são fenômenos que podem ser definidos como aqueles eventos em que uma determinada variável meteorológica apresenta valores acima ou abaixo de um limite superior ou inferior. Esses tipos de fenômenos climáticos têm um grande impacto nas regiões brasileiras, como por exemplo, o grande acúmulo de precipitação que causam enchentes e desmoronamentos, ou as secas duradouras que geram impacto nos reservatórios de água. O objetivo desse projeto é auxiliar no desenvolvimento de uma metodologia para realizar o monitoramento e previsão de eventos climáticos extremos com o uso de inteligência artificial. A metodologia empregada buscou implantar um sistema de tratamento e análise de grandes volumes de dados. Para isso, foram usados métodos estatísticos aplicados à meteorologia em dados de precipitação de alta resolução (GPCP v3.2) com foco nos estados e regiões do território brasileiro para validação da base de dados. Em seguida, esses dados foram usados para treinamento de uma rede neural que visa fornecer a previsão da precipitação. A validação da base de dados, feita através da análise dos padrões climatológicos da precipitação para as regiões brasileiras, mostrou que a precipitação do GPCP v3.2 é consistente com bases de dados anteriores. Os resultados da construção da rede neural foram avaliados através de um ensemble de previsões retrospectivas de 30 anos passados (hindcast). A análise das previsões retrospectivas através de métricas de avaliação mostrou que os resultados das previsões com a rede neural tem um performance superior aos obtidos com modelos dinâmicos. Sendo assim, a rede neural é hábil em fornecer previsões de precipitação que podem ser usadas no cálculo de métricas para classificação dos eventos extremos de precipitação, de modo a colaborar para o monitoramento e previsão de eventos meteorológicos extremos (seca profunda e chuvas intensas), e prevenir e/ou amenizar os impactos desses eventos nas regiões afetadas.

Palavras-chave: Eventos Climáticos Extremos no Brasil. Precipitação. Anomalias Climáticas. Extreme climate events in Brazil. Extreme climate indices. Precipitation. Climate anomalies.

LISTA DE FIGURAS

Figura 1: Arquitetura de uma rede neural	9
Figura 2: Código da normalização dos dados de treinamento e previsão	12
Figura 3: Código da compilação do modelo Keras	12
Figura 4: Código da previsão e cálculo do erro entre saída e previsão.....	13
Figura 5: Representação da estratégia 1.....	15
Figura 6: Representação da estratégia 2.....	15
Figura 7: Média climatológica mensal entre 1983 e 2010 da precipitação do GPCP v3.2.....	16
Figura 8: Série temporal de precipitação (mm/dia) anual do estado de São Paulo com dados do GPCP v3.2	17
Figura 9: Série temporal de precipitação (mm/dia) mensal do estado de São Paulo com dados do GPCP v3.2	17
Figura 10: - Anomalia de precipitação (mm/dia) do ano de 2019 para o estado de São Paulo .	18
Figura 11: - Avaliação do modelo de previsão (Estratégia 1).....	19
Figura 12: Erro de saída esperada x previsão (Estratégia 1)	19
Figura 13: Gráfico de Previsão x Alvo (Estratégia 1)	20
Figura 14: Gráfico Seccionado de Previsão x Alvo (Estratégia 1)	21
Figura 15: Avaliação do modelo de previsão (Estratégia 2).....	21
Figura 16: Erro de saída esperada x previsão (Estratégia 2)	22
Figura 17: Gráfico de Previsão x Alvo (Estratégia 2)	22
Figura 18: Gráfico Seccionado de Previsão x Alvo (Estratégia 2)	23
Figura 19: Precipitação observada (esquerda) e prevista pela rede neural da Estratégia 1 (direita) sobre o estado de São Paulo.	23
Figura 20: Diferença entre a precipitação prevista pelo modelo (Estratégia 1) e observada.....	24
Figura 21: Precipitação observada (esquerda) e prevista pela rede neural da Estratégia 2 (direita) sobre o estado de São Paulo.	24
Figura 22: Diferença entre a precipitação prevista pelo modelo (Estratégia 2) e observada.....	25
Figura 23: BS para a anomalia positiva (esquerda) e negativa (direita) para o trimestre DJF....	25

SUMÁRIO

1. INTRODUÇÃO	6
1.1 Objetivo Geral	7
2. DADOS E METODOLOGIA	7
2.1 Global Precipitation Climatology Project (GPCP)	7
2.2 Ferramentas computacionais	8
2.2.1 Rede Neural	9
2.3 Metodologia	10
2.4 Esquema de Previsão	13
2.4.1 Estratégia 1	14
2.4.2 Estratégia 2	15
3. RESULTADOS	15
3.1 Verificação e análise da base de dados	15
3.1.1 Média Climatológica Anual entre 1983 a 2010	16
3.1.2 Série temporal (mensal e anual) do estado de São Paulo	17
3.1.3 Anomalia de precipitação para o estado de São Paulo	18
3.3 Implementação de um modelo de previsão	18
3.3.1 Estratégia 1 - Resultados	18
3.3.2 Estratégia 2 - Resultados	21
3.4 Previsão de precipitação	23
3.4.1 Estratégia 1	23
3.4.2 Estratégia 2	24
3.4 Avaliação do modelo de previsão com o ensemble	25
4. CONCLUSÃO	26
REFERÊNCIAS	27

1. INTRODUÇÃO

Eventos climáticos extremos, como chuvas intensas e secas, são eventos que ocorrem fora de um limite determinado, seja acima ou abaixo. Esses tipos de eventos são considerados fenômenos naturais (MARENGO et al, 2013). A seca é caracterizada pelo nível de precipitação abaixo da média (MACEDO et al, 2010). Já a chuva intensa é o contrário, ou seja, quando o nível de precipitação fica acima da média. Segundo Araújo et al. (2008) e Macedo et al. (2010), ambos os fenômenos têm efeitos diretos na sociedade, na economia e no meio ambiente, afetando áreas como agricultura, saúde e o meio urbano, ocasionando inundações, erosão de solo, perda de colheita, secas de reservatórios, aumento dos índices de queimadas, baixa qualidade do ar entre outras consequências.

Nos dias atuais, com o avanço da tecnologia, vários tipos de ferramentas são criadas e disponibilizadas para a realização de tarefas que não se imaginava que seriam possíveis. A Inteligência Artificial é uma dessas ferramentas. E além de fazer coisas que os próprios humanos já fazem, como reconhecer padrões, voz, imagens, realizar diagnósticos e outras coisas, essas novas ferramentas possibilitam realizar previsões.

Prever algo que ainda está para acontecer possibilita diversas vantagens como se preparar para o acontecimento, amenizar os danos e gerenciar os riscos. Por exemplo, se for possível prever um grande volume de chuva em uma região que tem alto risco de deslizamento, é possível retirar pessoas do local e amenizar os impactos antes. Ou se for possível prever um período de seca em locais onde estão localizados reservatórios de águas, se torna possível aplicar um plano de gerenciamento de risco para evitar que os níveis das represas fiquem muito abaixo do normal.

Sendo assim, a utilização de inteligência artificial nas séries de precipitação torna possível a previsão desses eventos, tornando-se assim, uma ferramenta que pode auxiliar em estratégias emergenciais para lidar com momentos de crises ocasionados por esses fenômenos, sendo possível evitar grande impacto e os prejuízos. A utilização de dados atuais da versão 3.2 do *Global Precipitation Climatology Project (GPCP v3.2)* ocasiona uma melhor obtenção de resultados das análises, com dados atuais, que trazem os reflexos das mudanças climáticas nos últimos anos. Este relatório apresenta os resultados finais deste projeto.

1.1 Objetivo Geral

O objetivo desse projeto é auxiliar no desenvolvimento de uma metodologia para realizar o monitoramento e previsão de eventos climáticos extremos com o uso de inteligência artificial.

2. DADOS E METODOLOGIA

2.1 Global Precipitation Climatology Project (GPCP)

A observação, coleta e análise de dados de precipitação possibilita a utilização em diversas aplicações como monitoramento de seca, flash de inundações, previsão de colheitas, previsão de doenças e estudo de salinidade oceânica. Porém, esses dados ainda sofrem com escassez de informações em regiões de baixa densidade populacional e em países subdesenvolvidos (HUFFMAN et al, 2022). O GPCP surgiu como uma base de dados desenvolvida pela Global Water and Energy Exchange (GEWEX), no projeto World Climate Research Programme (WCRP) a fim de servir como um registro global de precipitação para auxiliar estudos climáticos e outras atividades.

A versão utilizada no desenvolvimento do projeto é do GPCP v3.2, sucessor do GPCP v2, trazendo dados de registros mais recentes. A escolha da utilização dos dados do GPCP v3.2 se deu por conta de ser a última versão desse conjunto de dados e trazer uma maior resolução espacial do que o GPCP v2. O período dos dados é de Janeiro de 1983 até Dezembro de 2020, contendo uma resolução temporal mensal. A resolução espacial é de $0.5^{\circ} \times 0.5^{\circ}$. O arquivo é em formato netCDF, sendo possível acessá-lo através do comando `ncdump` ou outras linguagens como Java, Python, GrADS, Matlab, entre outras. Os dados são formados pelas seguintes três dimensões: Latitude, sendo de -90.0 a 90.0; Longitude sendo de -180.0 a 180.0 e Time, sendo um período de 01-01-1983 a 31-12-2020. Os campos contidos no arquivo de dados podem ser conferidos abaixo:

Tabela 1 - Descrição dos dados GPCP v3

Nome do Campo	Descrição	Unidade
latitude	latitude do centro do elemento da grade	graus_norte
longitude	longitude do centro do elemento da grade	graus_leste
time	data e hora de início. dos dados	minutos desde 01-01-1979 00:00:00
sat_gauge_precip	estimativa combinada de precipitação por satélite	mm/dia
sat_gauge_error	estimativa de erro aleatório de precipitação combinada por satélite	mm/dia
satellite_precip	estimativa de precipitação apenas por satélite	mm/dia
satellite source	fonte da estimativa do satélite contribuinte	valores de index
gauge_precip	precipitação de manômetro ajustada à perda de vento	mm/dia
probability_liquid_phase	probabilidade de precipitação da fase líquida	por cento
gauge_relative_weight	ponderação relativa dos medidores em sat_gauge_precip	por cento
quality index	número equivalente de medidores	sem unidade

2.2 Ferramentas computacionais

Para o desenvolvimento de uma metodologia eficaz e atual, foram selecionadas ferramentas com suporte para Inteligência Artificial, pensando na quantidade de opções

oferecidas e a performance com um volume de dados grande. As ferramentas utilizadas são descritas a seguir.

O CDO (Climate Data Operators) é um conjunto de comandos para processar dados em formato NetCDF, assim como o NCL (NCAR Command Language), que também é uma ferramenta para leitura de dados meteorológicos. Ambos foram utilizados para a análise e tratamento de dados.

O Python é uma linguagem de programação com fácil usabilidade e de código aberto, com várias bibliotecas de suporte para análise de dados e inteligência artificial, e foi utilizada para a implementação do modelo de Inteligência Artificial. Especificamente as bibliotecas Pandas, desenvolvida para análises de dados, e scikit-learn, desenvolvida para machine learning.

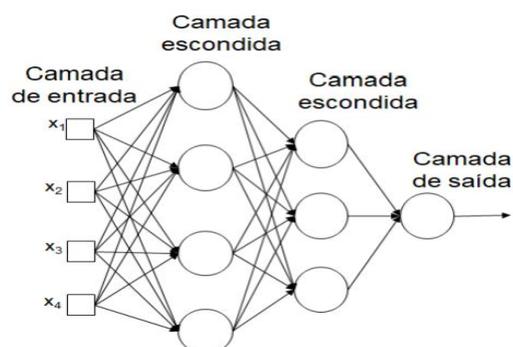
2.2.1 Rede Neural

As Redes Neurais, também conhecidas como Redes Neurais Artificiais (Artificial Neural Networks, ANNs), assim como o nome já apresenta, são modelos matemáticos baseados nas redes neurais que compõem o sistema nervoso humano. Desse modo, através desse método de inteligência artificial, é possível ensinar computadores a trabalhar com informações de forma semelhante ao cérebro humano, tornando-se possível um algoritmo reconhecer padrões, imagens e áudios com grande precisão e solucionar problemas de níveis mais complexos.

2.2.1.1 Arquitetura de uma rede neural

A estrutura de uma rede neural é composta por neurônios interligados por camadas, representada na figura 1 abaixo.

Figura 1: Arquitetura de uma rede neural



A primeira camada é chamada de camada de entrada, que é por onde as informações entram na rede neural para serem processadas e encaminhadas para a próxima camada. A segunda camada é chamada de camada escondida e/ou camada oculta e tem como função analisar os resultados que chegam de outras camadas e caminhar para a próxima. Em uma rede neural é possível ter várias camadas escondidas. A última camada é a camada de saída, responsável por fornecer o resultado final do processamento e análise das informações.

2.2.1.2 Modelo Keras

Keras é uma API de machine learning para Python focado em trazer resultados de forma rápida. Tem como característica ser simples, flexível e poderoso. A estrutura do Keras é formada por camadas e modelos. Para o desenvolvimento desse projeto foi utilizado o modelo Sequential, que é o modelo mais simples.

2.2.1.3 Configurações do projeto

Para o desenvolvimento do projeto, foram utilizadas três camadas intermediárias e em cada uma havia 35, 25 e 1 neurônio, respectivamente. Como função de ativação, foi utilizada a ReLu, que é a mais comum e é uma função não linear. A principal vantagem é que ela não ativa todos os neurônios de uma vez. Para medir e analisar os erros, foi definido para utilização como métrica de avaliação o Erro Quadrático Médio (EQM), e como otimizador foi selecionado o Adam.

Para o treinamento de uma rede neural são utilizadas épocas. Os dados passarão pela rede neural pelo número de vezes determinadas na época, por exemplo, no caso desse projeto foram determinadas 100 épocas. Assim, os dados passaram 100 vezes pela rede neural e a cada vez, os pesos foram atualizados para o mais próximo das características que determinados como parâmetros de aprendizagem.

2.3 Metodologia

A metodologia do projeto é composta por três partes. A primeira é responsável pela verificação e análise do GPCP v3.2. O objetivo é a validação da base de dados. Nesse caso, como a base de dados representa um grande volume de dados, a metodologia foi definida e adequada para que fosse possível realizar o processamento desses dados. Por ser uma base recente, foi avaliado se as informações do GPCP v3.2 representavam o

padrão esperado de dados, através de cálculos de média climatológica, anomalia e séries temporais.

A segunda parte é referente à preparação dos dados para entrada no modelo de Inteligência Artificial, através da aplicação de filtros e tratamento dos dados. Nesta etapa, foram definidos alguns filtros para os dados do GPCP, utilizando o CDO e o Python. A região analisada foi o estado de São Paulo. A variável de precipitação GPCP usada foi `sat_gauge_precip`.

Também foi necessário fazer um tratamento para a divisão da variável `time` da base de dados, separando o valor em ano, mês e dia, criando assim novas variáveis em um dataframe. Utilizando o CDO, foi gerado uma média de precipitação trimestral, utilizando filtros dos meses referente a cada estratégia, gerando um arquivo `.csv` para utilizar como entrada em um Jupyter Notebook para trabalhar com python.

E por fim, a terceira parte é a implementação e avaliação do modelo de Inteligência Artificial para previsão da precipitação. Pode-se dividir a implementação do modelo de previsão em três partes também. A primeira é referente a normalização das entradas e saídas, tanto dos dados de treinamento quanto os de previsão, o código pode ser visto na figura 2.

O objetivo dessa etapa é adequar os dados de acordo com o formato aceito pelo modelo e as bibliotecas de previsão. Para realizar as normatizações foi utilizada a linguagem Python, especificamente a biblioteca *scikit-learn*.

Figura 2: Código da normalização dos dados de treinamento e previsão

```
X_dev, y_dev = shuffle(X_TRAIN, Y_TRAIN, random_state=0)
X_train, X_val, y_train, y_val = train_test_split(X_dev, y_dev, test_size=0.10)

#normalizacao input
scaler = MinMaxScaler()
scaler.fit(X_train)
X_train_stand = scaler.transform(X_train)
X_val_stand = scaler.transform(X_val)

#normalizacao output
scaler = MinMaxScaler()
scaler.fit(y_train)
Y_train_stand = scaler.transform(y_train)
Y_val_stand = scaler.transform(y_val)

#normalizacao input test(gen)
scaler = MinMaxScaler()
scaler.fit(X_test)
X_test_stand = scaler.transform(X_test)

#normalizacao output test(gen)
scaler = MinMaxScaler()
scaler.fit(Y_test)
Y_test_stand = scaler.transform(Y_test)

print(X_val_stand.shape)
print(Y_val_stand.shape)
```

A segunda parte é o treinamento do modelo. Então, é onde se aplica as configurações abordadas no tópico 2.2.1.3 no modelo Keras, utilizando os dados de entrada definidos em cada estratégia, visto na figura 3.

Figura 3: Código da compilação do modelo Keras

```
def get_model():
    model = Sequential()
    model.add(Dense(35, input_dim=5, activation='relu')) #input_dim=numero de entrada (num variaveis)
    model.add(Dense(25, activation='relu'))
    model.add(Dense(1, activation='linear'))
    return model

# define the keras model
model=get_model()

# compile the keras model
model.compile(loss='mean_squared_error', optimizer='adam', metrics=['mse'])

history=model.fit(X_train_stand,
                  Y_train_stand.ravel(),
                  validation_data=(X_val_stand,Y_val_stand),
                  batch_size=10,
                  epochs=100,
                  verbose=2,
                  callbacks=[callbacks.EarlyStopping(monitor='val_loss',
                                                    patience=100,
                                                    restore_best_weights=True)])
```

Na terceira etapa, após o treinamento do modelo, o objetivo é realizar e avaliar as previsões. Nessa parte, é realizado previsões através do modelo Keras, e logo após, utilizando a biblioteca scikit-learn, é realizada a desnormalização dos dados gerados para realizar o cálculo de erro e plotar a previsão x alvo (Figura 4). O erro é calculado através da fórmula $ERROR = saída_esperada - previsão$.

Figura 4: Código da previsão e cálculo do erro entre saída e previsão

```
# realiza previsoes
z_pred=model.predict(X_test_stand)

#realiza a desnormalizacao
z_pred = scaler.inverse_transform(z_pred)
#print(z_pred)

from scipy import signal

# calcula o erro entre saida alvo e previsao
ERROR = Y_test - z_pred

#plota erro
ERROR_SMOOTH=signal.savgol_filter(ERROR,99,3,axis=0)
plt.plot(ERROR_SMOOTH, '--', color='black')
plt.hlines(ERROR.mean(),xmin=0,xmax=ERROR.shape[0],linestyle='-',color='blue')
plt.hlines(ERROR.mean()+ERROR.std(),xmin=0,xmax=ERROR.shape[0],linestyle='--',color='blue')
plt.hlines(ERROR.mean()-ERROR.std(),xmin=0,xmax=ERROR.shape[0],linestyle='--',color='blue')

ERROR.mean(),ERROR.std()
```

As três etapas são aplicadas de maneira igual nas duas estratégias. O que diferencia são os dados de entrada e saída esperada.

A avaliação da previsão foi realizada através de um ensemble de 10 previsões retrospectivas de 30 anos passados (ou hindcast). Após a obtenção do hindcast, foram calculados os valores do Brier Score (BS). O BS compara a frequência da previsão de cada ano com o que foi observado. Nesse caso, tanto a previsão como a observação precisam ser categorizadas. A categoria escolhida para a avaliação foi a anomalia positiva e negativa. A fórmula do BS, exibida abaixo, compara em cada ano previsto a probabilidade de previsão na categoria i (p_i) com a observação na mesma categoria ($a_i = 1$ se a categoria i ocorreu e $a_i = 0$ se não ocorreu).

$$BS = \frac{1}{N} \sum_{i=1}^N (p_i - a_i)^2, \quad (0 \leq BS \leq 1)$$

O valor de BS fica entre 0 e 1. Menores valores de BS indicam maior destreza na previsão.

2.4 Esquema de Previsão

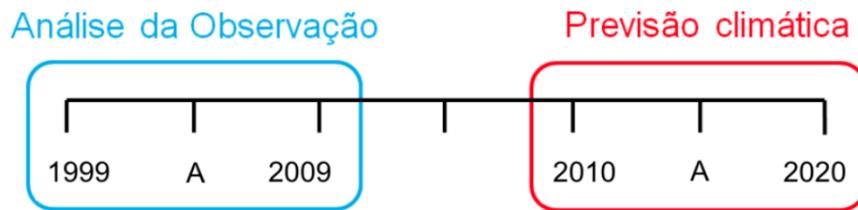
Para o treinamento do modelo de previsão, foi utilizada a análise de observação e previsão. Então, o arquivo de entrada contém os valores que o modelo utilizou para

aprender, analisar e observar, e o arquivo de saída contém os valores esperados como previsão. Foram criadas duas estratégias de previsão para o modelo.

2.4.1 Estratégia 1

Na estratégia 1, representada na figura 5, o arquivo de entrada contém como dados a média trimestral de precipitação dos meses de Março, Abril e Maio, e do período entre 1999 a 2009. Para essa estratégia, foram aplicados os filtros de ano, separando em décadas. O arquivo de saída contém como dados esperados a média trimestral dos meses de Março, Abril e Maio, e do período entre 2010 a 2020.

Figura 5: Representação da estratégia 1

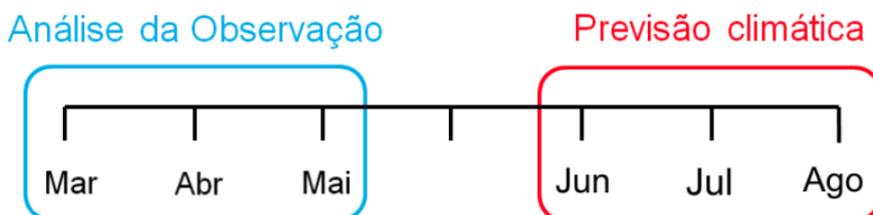


O objetivo da estratégia 1 é treinar nosso modelo com os dados do período de um período de 10 anos e esperar que ele seja capaz de prever os 10 anos seguintes, ou seja, inserindo os dados de um trimestre do período de 1999 a 2009, é esperado que seja possível prever os valores referentes ao mesmo trimestre dos anos 2010 a 2020.

2.4.2 Estratégia 2

Na estratégia 2, representada na figura 6, o arquivo de entrada contém como dados a média trimestral de precipitação dos meses de Março, Abril e Maio e o arquivo de saída contém como dados esperados a média trimestral dos meses de Junho, Julho e Agosto. Para essa estratégia, foram gerados dois arquivos .csv distintos, sem a necessidade de aplicação do filtro de ano.

Figura 6: Representação da estratégia 2



Para a estratégia 2 foi utilizado o período completo de dados, de 1983 até 2020. O objetivo foi treinar nosso modelo com os dados de um trimestre e esperar que ele seja capaz de prever o trimestre seguinte, ou seja, inserindo os dados de um trimestre de Março, Abril e Maio dos anos de 1983 a 2020, é esperado que seja possível prever os valores referentes ao trimestre de Junho, Julho e Agosto do mesmo período de anos.

3. RESULTADOS

3.1 Verificação e análise da base de dados

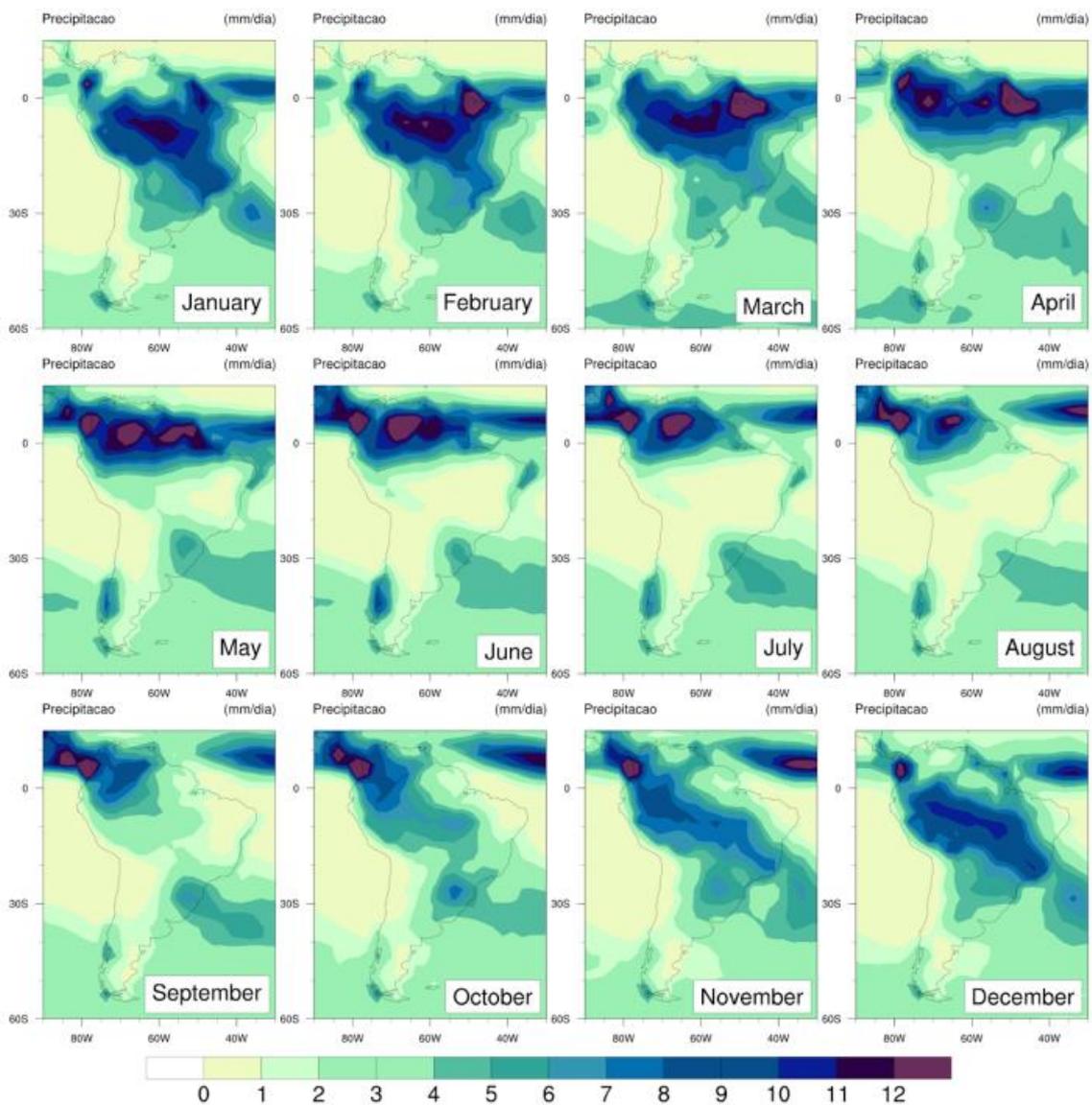
Para essa parte do projeto, foram realizados scripts, utilizando o NCL, que aplicavam filtros, médias e cálculos que também seriam utilizados na preparação dos dados para o

modelo de previsão, a fim de realizar a validação da base de dados e determinar se seria possível utilizá-la no projeto.

3.1.1 Média Climatológica Anual entre 1983 a 2010

Através de um script NCL, foi selecionado o período de 1983 a 2010 e plotada a média climatológica anual para a região da América do Sul (Figura 7). Os resultados mostram que o GPCP v3.2 representa as principais características da variação anual da precipitação na região, como o desenvolvimento do sistema de monção da América do Sul entre a primavera e o verão e a estação seca nos meses de inverno na região central do Brasil.

Figura 7: Média climatológica mensal entre 1983 e 2010 da precipitação (mm/dia) do GPCP v3.2



3.1.2 Série temporal (mensal e anual) do estado de São Paulo

Utilizando NCL, foi selecionado o período de 1983 a 2010 e plotada a série temporal de precipitação anual (Figura 8) e mensal (Figura 9) para a região aproximada do estado de São Paulo através de gráficos de linhas. Em ambas as figuras é possível notar uma grande variabilidade mensal e anual da precipitação sobre o estado de São Paulo, com destaque para anos com extremos chuvosos (2000) e secos (1983).

Figura 8: Série temporal de precipitação (mm/dia) anual do estado de São Paulo com dados do GPCP v3.2

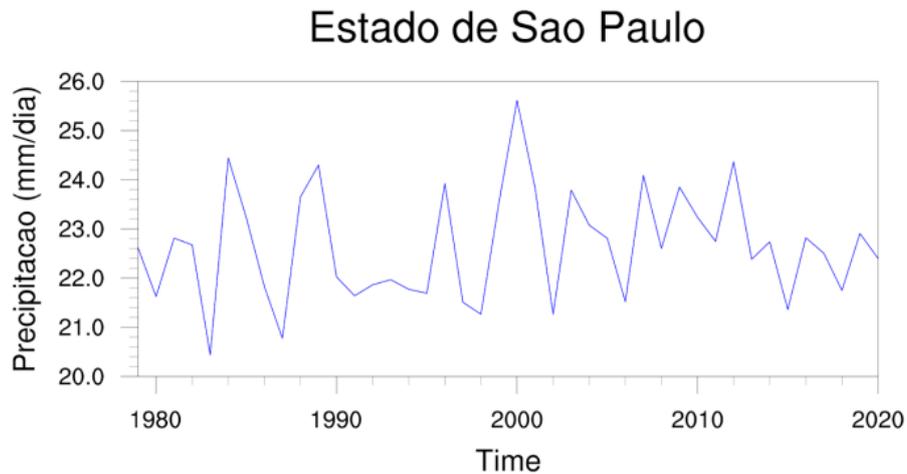
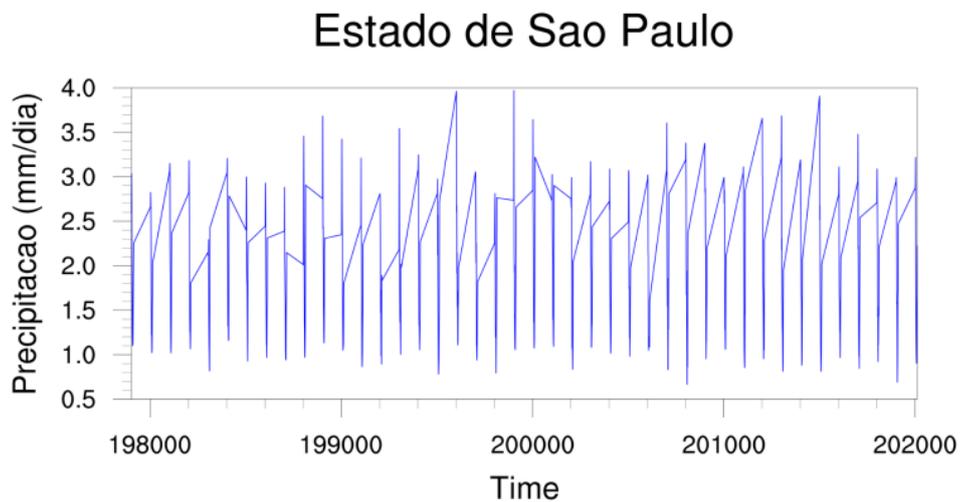


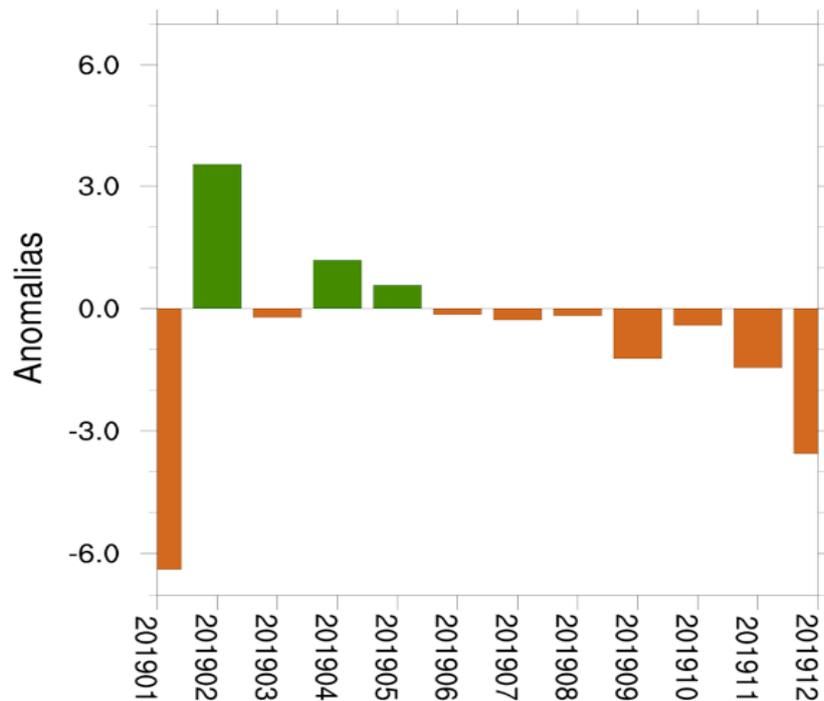
Figura 9: Série temporal de precipitação (mm/dia) mensal do estado de São Paulo com dados do GPCP v3.2



3.1.3 Anomalia de precipitação para o estado de São Paulo

Utilizando NCL, foi criado um gráfico de barras com a anomalia de precipitação (Figura 10) para o estado de São Paulo. As barras em verde representam valores positivos e as alaranjadas representam valores negativos.

Figura 10: - Anomalia de precipitação (mm/dia) do ano de 2019 para o estado de São Paulo



3.3 Implementação de um modelo de previsão

3.3.1 Estratégia 1 - Resultados

O resultado da aplicação da estratégia 1 se adequou ao esperado, que é o aprendizado eficiente, erros dentro da média e a previsão correta ou próxima de alguns valores. Como pode ser visto na figura 11, o modelo teve um processo de aprendizado eficaz. No gráfico da Figura 12 mostra que os erros ficaram dentro da média esperada.

Figura 11: - Avaliação do modelo de previsão (Estratégia 1)

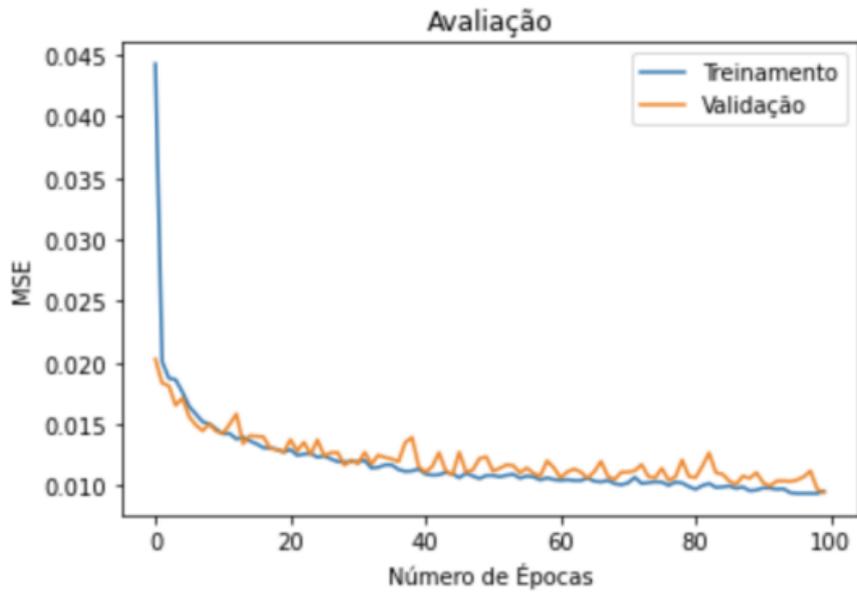
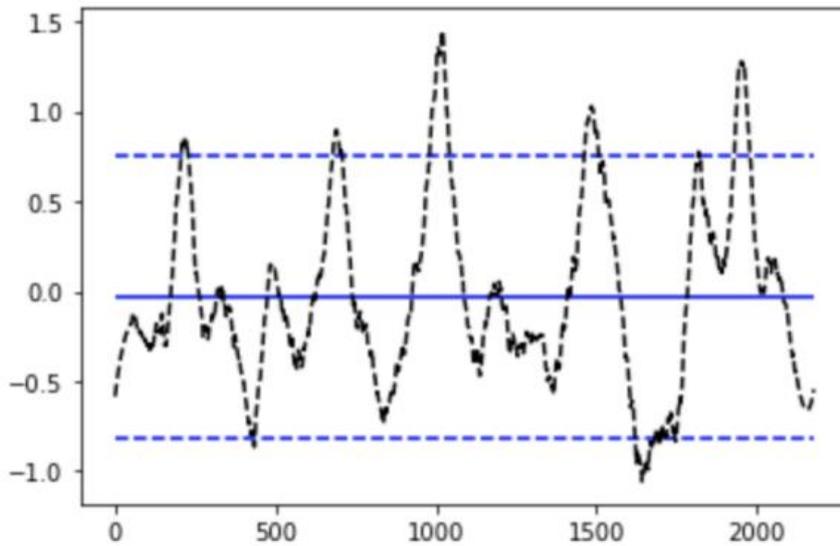


Figura 12: Erro de saída esperada x previsão (Estratégia 1)



Os gráficos apresentados nas figuras 13 e 14 apresentam valores referentes a previsão, em laranja, e o valor esperado, em azul. Sendo a figura 14 uma parte seccionada do gráfico da figura 13, podendo ser analisado os valores mais de perto.

Figura 13: Gráfico de Previsão x Alvo (Estratégia 1)

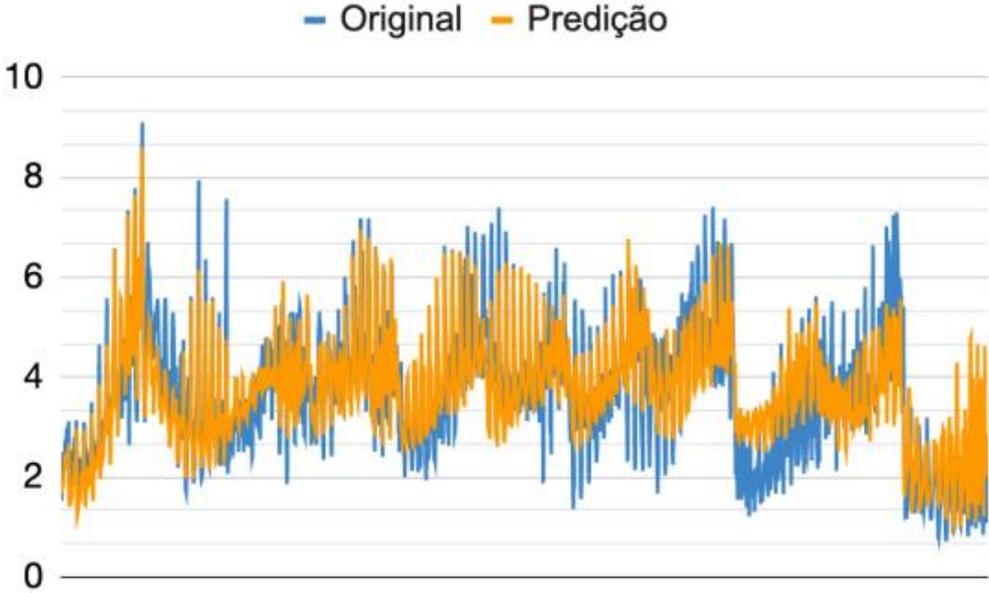
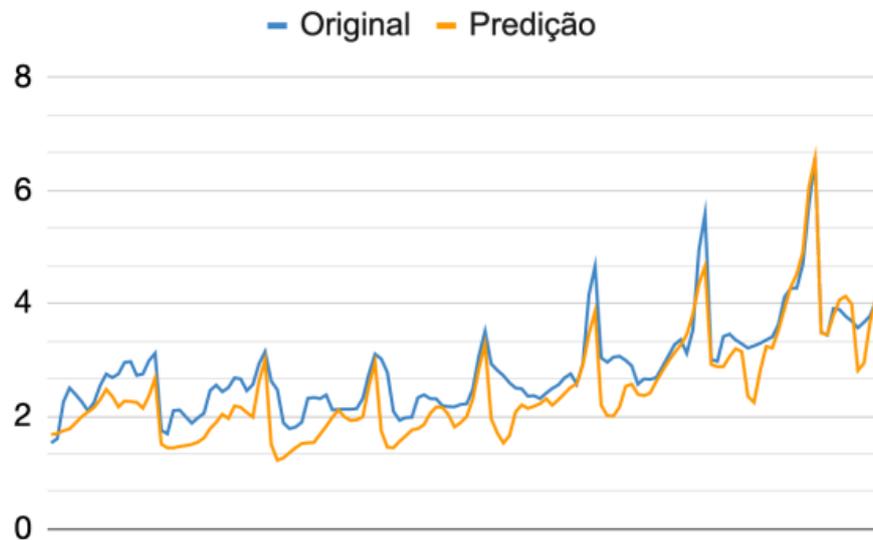


Figura 14: Gráfico Seccionado de Previsão x Alvo (Estratégia 1)

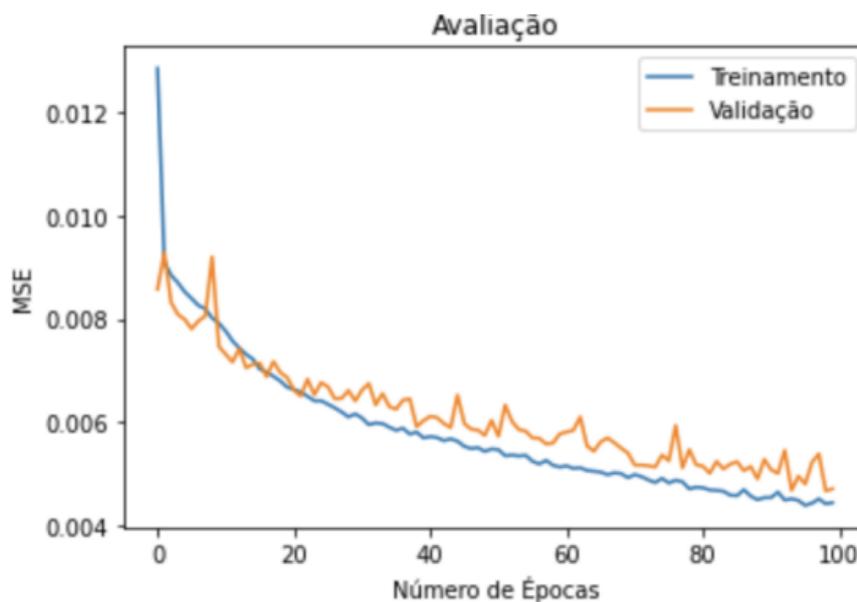


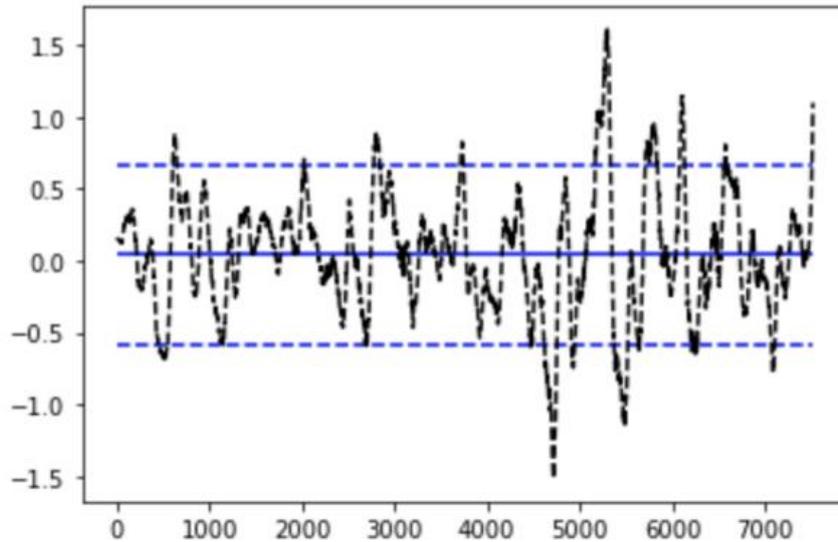
É possível observar através dos gráficos que o valor previsto acompanhou bem o valor esperado. Na figura 14 é visto que em muitos pontos o modelo foi capaz de chegar bem próximo ao esperado e até mesmo atingiu o valor esperado.

3.3.2 Estratégia 2 - Resultados

O resultado da aplicação da estratégia 2, assim como na estratégia 1, se adequou ao esperado. O modelo apresentou um processo de aprendizado eficaz (Figura 15) e os erros ficaram dentro da média esperada, apesar do maior volume de dados (Figura 16).

Figura 15: Avaliação do modelo de previsão (Estratégia 2)





É possível observar através dos gráficos que o valor previsto acompanhou bem o valor esperado (Figura 17), e comparado à estratégia 1, acompanhou de maneira mais próxima. Na figura 18 é possível ver que houve muito mais pontos onde o modelo atingiu o valor esperado.

Figura 17: Gráfico de Previsão x Alvo (Estratégia

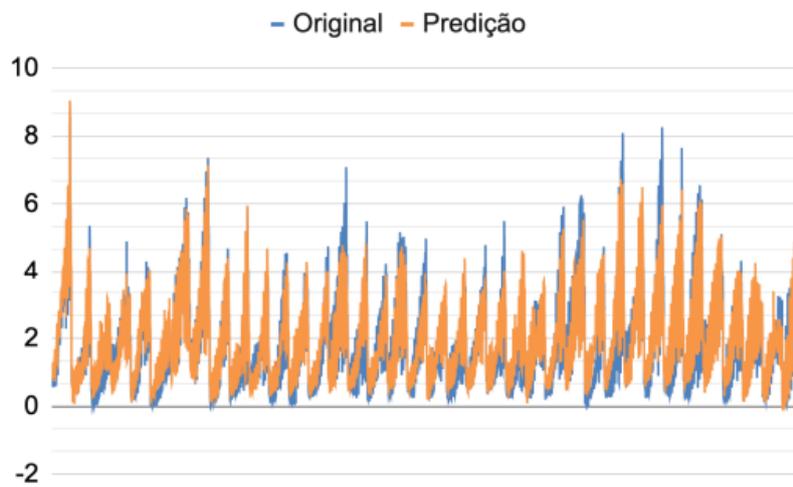
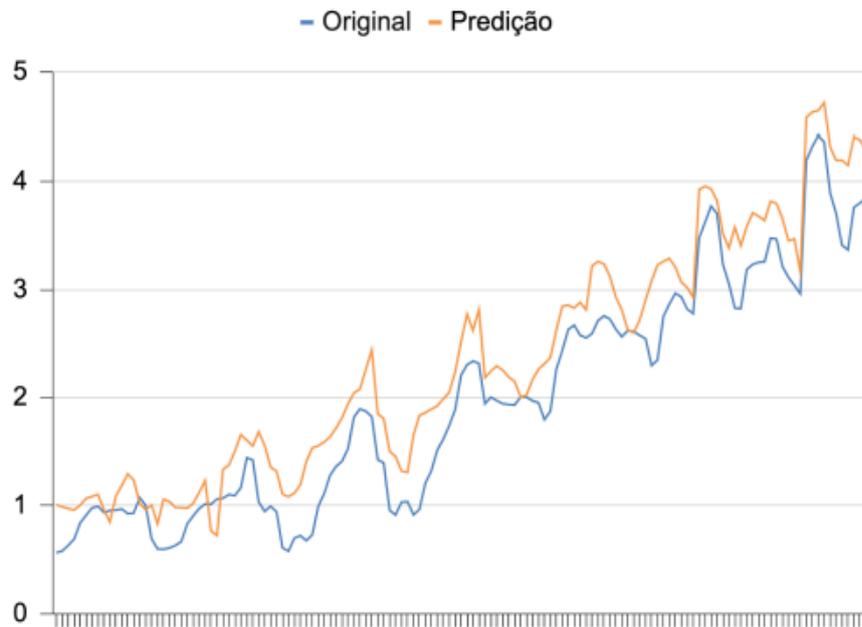


Figura 18: Gráfico Seccionado de Previsão x Alvo (Estratégia 2)



3.4 Previsão de precipitação

A seguir são mostrados os resultados para a previsão da precipitação sobre o estado de São Paulo para o trimestre Dezembro, Janeiro e Fevereiro (DJF).

3.4.1 Estratégia 1

A Figura 19 mostra que o modelo foi capaz de representar a variabilidade da precipitação sobre o estado de São Paulo entre os anos de 2010 a 2020, com máximos de chuva sobre a região litorânea e menores valores na região central do estado. Os erros da previsão (Figura 20) foram menores que 1 mm/dia na maior parte do estado.

Figura 19: Precipitação observada (esquerda) e prevista pela rede neural da Estratégia 1 (direita) sobre o estado de São Paulo.

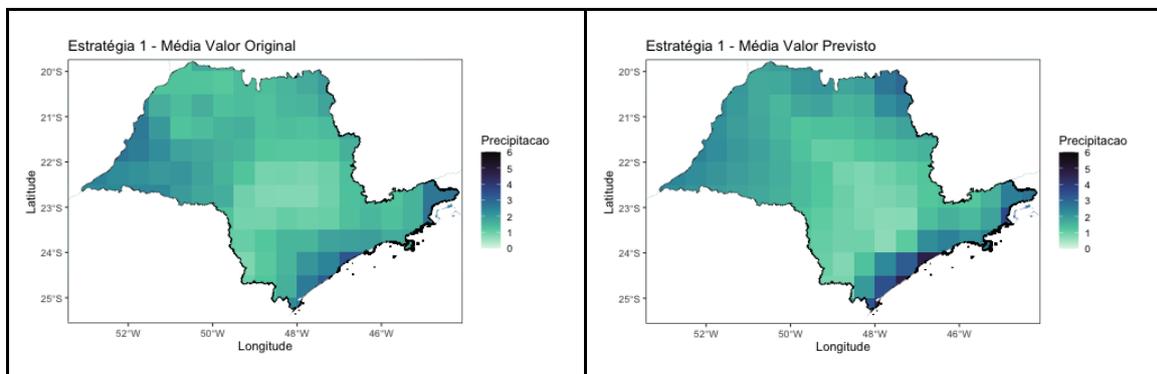
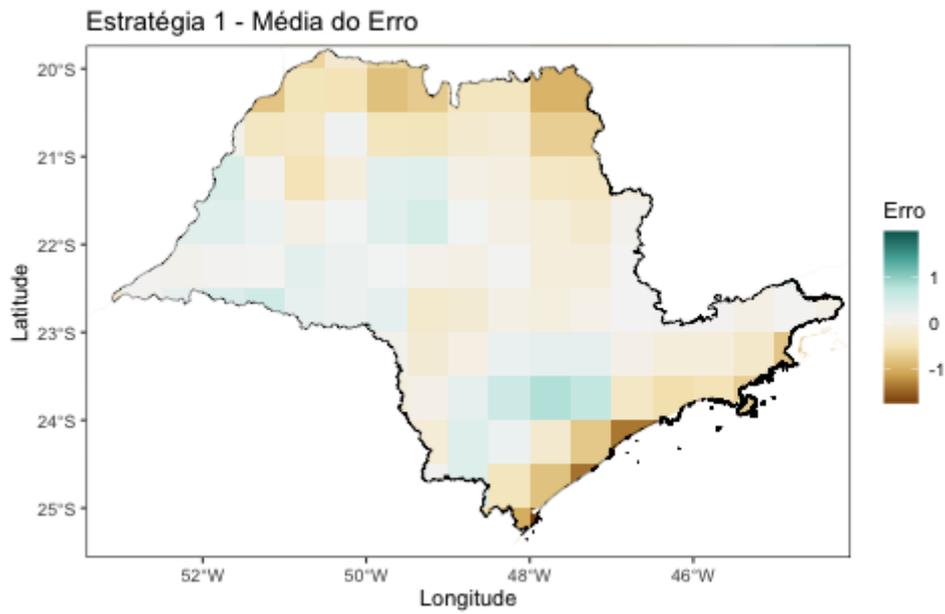


Figura 20: Diferença entre a precipitação prevista pelo modelo (Estratégia 1) e observada.



3.4.2 Estratégia 2

A Figura 21 mostra que o modelo obteve uma boa representação da precipitação sobre o estado de São Paulo para o verão (DJF) de 2020, com máximos de chuva sobre o oeste e sul do estado e menores valores ao norte. Os erros da previsão (Figura 21) foram menores que 1 mm/dia na maior parte do estado.

Figura 21: Precipitação observada (esquerda) e prevista pela rede neural da Estratégia 2 (direita) sobre o estado de São Paulo.

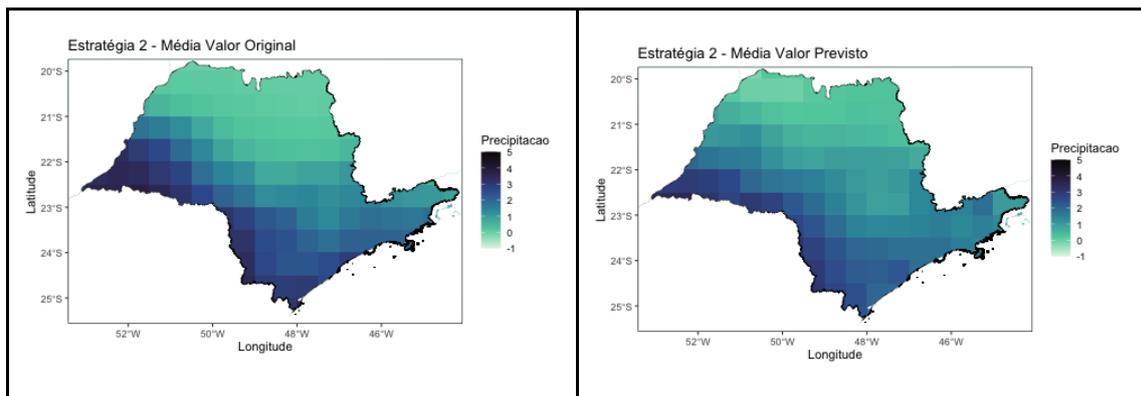
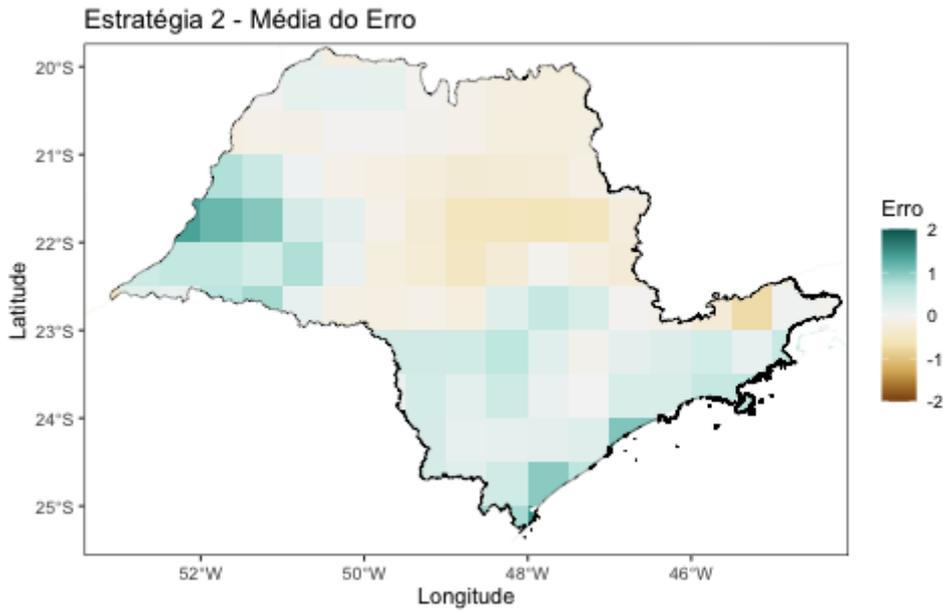


Figura 22: Diferença entre a precipitação prevista pelo modelo (Estratégia 2) e observada.



3.4 Avaliação do modelo de previsão com o ensemble

A Figura 23 mostra o BS para a anomalia positiva e negativa derivado do ensemble de previsões retrospectivas com o modelo de rede neural. Os valores de BS estão bem próximos de 0, o que indica um alto nível de destreza das previsões. Como comparação, a Figura 24 exibe o BS para um modelo dinâmico no mesmo trimestre, em que é possível notar que os valores de BS para o estado de São Paulo são muito maiores, indicando um nível menor de destreza. Cabe ressaltar que o recurso computacional empregado para a execução de previsões com modelos dinâmicos é muito maior do que o empregado pelo modelo de rede neural implementado neste estudo.

Figura 23: BS para a anomalia positiva (esquerda) e negativa (direita) para o trimestre DJF.

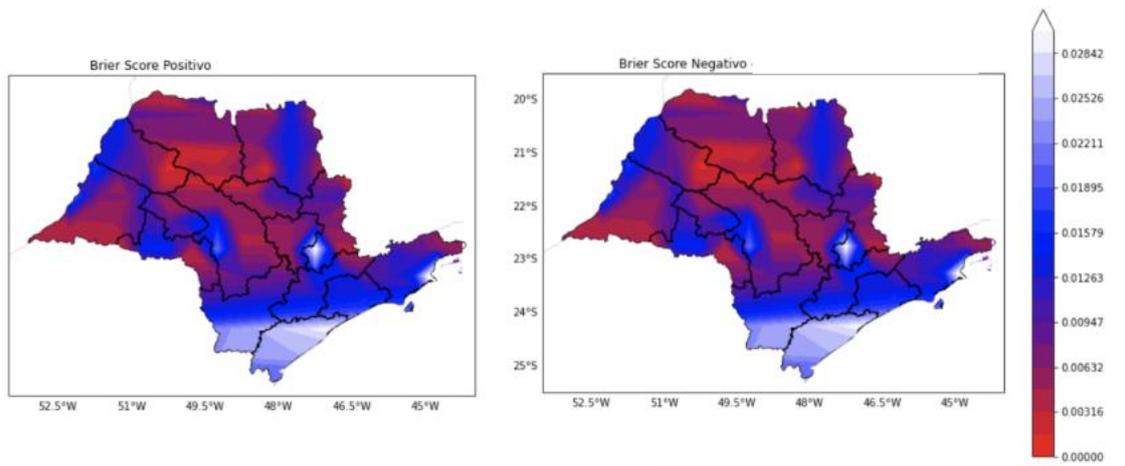
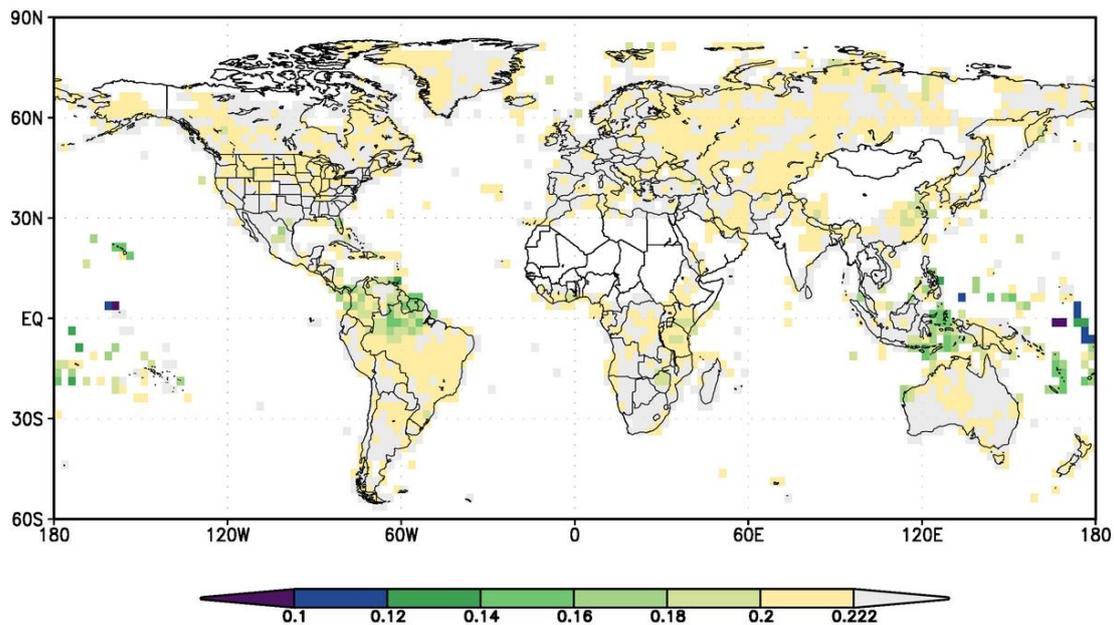


Figura 24 - BS obtido pela previsão sazonal do International Research Institute (IRI) para DJF.



4. CONCLUSÃO

A base de dados GPCP v3.2 correspondeu bem ao padrão de precipitação esperado, sendo assim, possível utilizar para a implementação do modelo de previsão. A utilização de dados atuais irá trazer uma maior fidelidade ao modelo e números novos que possibilitam novas análises ainda não feitas, tornando-se um diferencial do projeto.

Para a implementação do projeto, conclui-se que os parâmetros utilizados foram satisfatórios. Mesmo utilizando apenas 100 épocas, que é um tempo relativamente pequeno para treinamento, o modelo foi capaz de aprender e prever corretamente o valor esperado.

Apesar de ambas as estratégias propostas apresentarem resultados eficazes e dentro do esperado, a estratégia 2 se mostrou mais eficiente entre as duas. É possível observar que os valores previstos foram mais próximos do esperado. Vale ressaltar que essa estratégia utilizou um maior volume de dados do que a estratégia 1.

A avaliação das previsões obtidas pelo modelo de rede neural implementado através de um ensemble de previsões retrospectivas mostrou que as previsões realizadas pelo nosso modelo apresentou um alto nível de destreza, sendo superior aos modelos dinâmicos.

Para o aprimoramento das previsões realizadas com o uso de inteligência artificial sugere-se aumentar a área de abrangência, aplicar novas estratégias, testar novos

parâmetros para a rede neural utilizada e testar novas redes neurais. Os resultados apresentados podem colaborar para o monitoramento e previsão de eventos meteorológicos extremos (seca profunda e chuvas intensas), e prevenir e/ou amenizar os impactos desses eventos nas regiões afetadas.

REFERÊNCIAS

MARENGO J.A; VALVERDE M.C; OBREGON G.O. Observed and projected changes in rainfall extremes in the Metropolitan Area of São Paulo. *Clim Res* 57(1): 61-72, 2013.

MACEDO, M.J.H; GUEDES, R.V.S; SOUSA, F.A. S; DANTAS, F.R.C. Análise do índice padronizado de precipitação para o estado da Paraíba, Brasil. *Revista Ambiente & Água*, v. 5, n. 1, p. 204-214, 2010. ARAÚJO, L.E.; SOUSA, F. DE A.S.; RIBEIRO, M.A.; SANTOS, A.S.;

MEDEIROS, P.C. Análise estatística de chuvas intensas na bacia hidrográfica do Rio Paraíba. *Revista Brasileira de Meteorologia*, v. 23, n. 2, p. 162-169, 2008. HUFFMAN, G.J.; A. BEHRANGI, D.T.; BOLVIN, E.J; GPCP Version 3 Satellite-Gauge (SG) Combined Precipitation Data Set, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), 2022.