

INSTITUTO DE ENGENHARIA NUCLEAR

MARCOS ANTONIO GONÇALVES DA SILVA FILHO

**PREVISÃO DO DESLOCAMENTO DO PONTO DE MÁXIMA TAXA DE DOSE DE PLUMA
RADIOATIVA ATRAVÉS DE REDES NEURAIS RECORRENTES TIPO LSTM**

Rio de Janeiro

2023

MARCOS ANTONIO GONÇALVES DA SILVA FILHO

**PREVISÃO DO DESLOCAMENTO DO PONTO DE MÁXIMA TAXA DE DOSE DE PLUMA
RADIOATIVA ATRAVÉS DE REDES NEURAIS RECORRENTES TIPO LSTM**

Dissertação apresentada ao Programa de Pós-graduação em Ciência e Tecnologia Nucleares do Instituto de Engenharia Nuclear da Comissão Nacional de Energia Nuclear como parte dos requisitos necessários para a obtenção do Grau de Mestre em Ciência em Engenharia Nuclear – Profissional em Métodos Computacionais

Orientador: Prof. Dr. Claudio Marcio do Nascimento Abreu Pereira

Rio de Janeiro

2023

FILH Filho, Marcos Antonio.

Previsão do deslocamento do ponto de máxima taxa de dose de pluma radioativa através de redes neurais recorrentes tipo LSTM / Marcos Antonio Gonçalves da Silva Filho. – Rio de Janeiro: CNEN/IEN, 2023.

xi, 46f. : il. ; 31 cm.

Orientadores: Claudio Marcio do Nascimento Abreu Pereira.

Dissertação (mestrado) – Instituto de Engenharia Nuclear, PPGIEN, 2023.

1. Dispersão atmosférica de plumas radioativas. 2 Redes neurais artificiais recorrentes. 3. Fatores humanos.

PREVISÃO DO DESLOCAMENTO DO PONTO DE MÁXIMA TAXA DE DOSE DE
PLUMA RADIOATIVA ATRAVÉS DE REDES NEURAIS RECORRENTES TIPO LSTM

Marcos Antonio Gonçalves da Silva Filho

DISSERTAÇÃO SUBMETIDA AO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E
TECNOLOGIA NUCLEARES DO INSTITUTO DE ENGENHARIA NUCLEAR DA
COMISSÃO NACIONAL DE ENERGIA NUCLEAR COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM
ENGENHARIA NUCLEAR – PROFISSIONAL EM MÉTODOS COMPUTACIONAIS

Aprovada por:

Prof. Claudio Marcio do Nascimento Abreu Pereira, D.Sc.

Prof. Cláudio Henrique dos Santos Grecco, D.Sc.

Prof. Paulo Cezar Rocha Silveira, D.Sc.

Prof. Marcelo Carvalho dos Santos, D.Sc.

RIO DE JANEIRO, RJ - BRASIL
AGOSTO DE 2023

RESUMO

A liberação de materiais radioativos na atmosfera pode ocorrer devido a acidentes durante a operação de uma usina nuclear. Com a liberação de material radioativo, ocorre a formação de uma pluma de dispersão radioativa, cujo comportamento deve ser estimado para que se possa tomar as medidas de proteção radiológica necessárias para garantir a segurança da população e do meio ambiente, na região afetada. Essa estimativa é feita através de sistemas computacionais que monitoram as condições meteorológicas e os parâmetros operacionais da usina, e estimam a pluma de dispersão radioativa em tempo real.

O objetivo desse trabalho é desenvolver um modelo computacional, baseado em redes neurais artificiais recorrentes do tipo LSTM, para previsão do comportamento da pluma radioativa na atmosfera. Para treinar a rede neural, é utilizado um conjunto de dados gerados a partir de um simulador de dispersão atmosférica de plumas radioativas. Esse conjunto de dados simulados é utilizado no treinamento de diversas arquiteturas de redes neurais recorrentes, de modo a permitir selecionar a que melhor se ajuste as estimativas feitas pelo simulador. Desse modo, o modelo desenvolvido será capaz de prever o comportamento de uma pluma radioativa durante um acidente. Isso é particularmente importante em termos de segurança, pois proporcionará mais tempo e flexibilidade para que operadores e equipes de emergência tomem as decisões necessárias para garantir a segurança radiológica da região, reduzindo os riscos de erros humanos.

Palavras-Chave: pluma radioativa, dispersão atmosférica, inteligência artificial, redes neurais artificiais recorrentes, redes LSTM, fatores humanos.

ABSTRACT

Release of radioactive materials in the atmosphere can occur due to severe accidents during the operation of a nuclear power plant. As the accident progresses and the release of radioactivity in the atmosphere continues, radioactive dispersion plumes are formed. The behavior of this radioactive plume is monitored by a specialized team so that all necessary radiological protection measures can be taken for the safety of the population and environment around the nuclear power plant site. Safety monitoring is done through state-of-the-art computational systems that continuously track all meteorological conditions and plant status parameters and estimate the radioactive dispersion plume in real time.

The objective of this work is to develop a computational model, based on recurrent artificial neural network techniques, which predicts in advance what will be the future behavior of a radioactive dispersion plume. For this, a training set is generated by a radioactive plume dispersion simulator. Afterwards, simulated dataset is used to train different recurrent neural network structures to select one with the best results. The selected model must be able to predict the future behavior of a dispersion plume during a severe accident. This is important since it will provide more time and flexibility for operators and emergency team to make their decisions to guarantee the radiological safety of the region, reducing the risks of human errors.

Keywords: radioactive dispersion plume, atmospheric dispersion plume modeling, artificial intelligence, recurrent artificial neural networks, human factors.

LISTA DE FIGURAS

FIGURA 2.1 – Estrutura de neurônio	13
FIGURA 2.2 – Representação de um neurônio MP	14
FIGURA 2.3 – Representação de uma rede Perceptron	15
FIGURA 2.4 – Representação esquemática de uma rede neural MLP	16
FIGURA 2.5 – Algoritmo de retropropagação para redes MLP	18
FIGURA 2.6 – Gráfico da função sigmoide e sua derivada	20
FIGURA 2.7 – Gráfico da função ReLU e a sua derivada	20
FIGURA 2.8 – Diferença entre uma rede neural recorrente uma rede neural feedforward ..	22
FIGURA 2.9 – Representação esquemática de uma célula de uma estrutura recorrente	22
FIGURA 2.10 – Desdobramento de uma rede neural recorrente	23
FIGURA 2.11 – Componentes de uma célula LSTM	24
FIGURA 3.1 – Pluma de dispersão gaussiana a partir de uma fonte emissora	29
FIGURA 3.2 – Sucessivas bufadas provenientes de uma fonte emissora	30
FIGURA 3.3 – A mudança de direção gerada para diferentes modelagens matemáticas	31
FIGURA 3.4 – Mapa da zona de monitoramento da CNAAA	32
FIGURA 3.5 – Representação esquemática dos inputs e outputs dos diferentes módulos do simulador SCA	33
FIGURA 3.6 – Distribuição espacial das taxas de dose de corpo inteiro: (a) no ciclo 8, (b) no ciclo 11, (c) no ciclo 16, (d) no ciclo 18	35
FIGURA 4.1.1 – Ponto de maior taxa de dose de corpo inteiro presente na pluma de dispersão radioativa: (a) no ciclo 1, (b) no ciclo 2	38
FIGURA 4.1.2 – Ponto de maior taxa de dose de corpo inteiro presente na pluma de dispersão radioativa: (a) no ciclo 3, (b) no ciclo 4, (c) no ciclo 5, (d) no ciclo 6, (e) no ciclo 7, (f) no ciclo 8	39
FIGURA 4.1.3 – Ponto de maior taxa de dose de corpo inteiro presente na pluma de dispersão radioativa: (a) no ciclo 9, (b) no ciclo 10, (c) no ciclo 11, (d) no ciclo 12, (e) no ciclo 13, (f) no ciclo 14	40
FIGURA 4.1.4 – Ponto de maior taxa de dose de corpo inteiro presente na pluma de dispersão radioativa: (a) no ciclo 15, (b) no ciclo 16, (c) no ciclo 17, (d) no ciclo 18	41
FIGURA 5.1.1 – Gráfico de R^2 para: (a) abscissas do grupo de treinamento, (b) ordenadas do grupo de treinamento	46

FIGURA 5.1.2 – (a) abscissas do grupo de teste, (b) ordenadas do grupo de teste, (c) abscissas do grupo de validação, (d) ordenadas do grupo de validação.....	47
FIGURA 5.2 – Gráfico de dispersão comparativo entre: (a) resultados simulados e grupo de treinamento, (b) resultados simulados e grupo de teste, (c) resultados simulados e grupo de validação	48
FIGURA 5.3 – Representação da movimentação do ponto de máxima dose de corpo inteiro para diferentes grupos de séries temporais selecionados aleatoriamente	49

LISTA DE TABELAS

TABELA 3.1 – Parâmetros meteorológicos medidos em cada torre e ponto de medição	34
TABELA 4.1 – Parâmetros meteorológicos referente aos primeiros 18 ciclos a serem simulados	37
TABELA 4.2 – Parâmetros de entrada utilizados para o treinamento da rede neural artificial recorrente	42
TABELA 4.3 – Parâmetros de entrada normalizados utilizados para o treinamento da rede neural artificial recorrente	43
TABELA 4.4 – Organização do conjunto de dados utilizados no treinamento, validação e teste da rede neural	44
TABELA 5.1 – Resultados obtidos para treinamento, teste e validação de todas as redes testadas	45

LISTA DE SIGLAS

AS	– Acidente Severo
CFD	– <i>Computational Fluid Mechanics</i>
CNAAA	– Central Nuclear Almirante Álvaro Alberto
CNEN	– Comissão Nacional de Energia Nuclear
DRNN	– <i>Deep Rectifier Neural Network</i>
GRNN	– <i>General Regression Neural Network</i>
IA	– Inteligência Artificial
IAEA	– <i>International Atomic Energy Agency</i>
LOCA	– <i>Loss of Coolant Accident</i>
LSTM	– <i>Long Short Term Memory</i>
MLP	– <i>Multilayer Perceptron</i>
MSE	– <i>Mean Squared Error</i> (Erro Médio quadrático)
ReLU	– <i>Rectified Linear Unit</i>
RNA	– Rede Neural Artificial
SCA	– Sistema de Controle Ambiental
SDA	- Simulador de Dispersão Atmosférica
TF	– Termo Fonte

SUMÁRIO

1. INTRODUÇÃO	9
1.1. Apresentação do problema.....	9
1.2. Trabalhos relacionados	9
1.3. Justificativa	11
1.4. Objetivo.....	11
1.5. Organização do trabalho	12
2. REDES NEURAIS ARTIFICIAIS.....	13
2.1. Neurônio artificial.....	14
2.2. Neurônio perceptron	14
2.3. Rede neural tipo perceptron multicamadas.....	15
2.4. Treinamento de uma MLP	16
2.5. Rede neural profunda.....	19
2.6. Rede neural recorrente	21
2.7. Rede neural tipo LSTM	22
3. SIMULADOR DE DISPERSÃO ATMOSFÉRICA DE PLUMAS RADIOATIVAS..	27
3.1. Teoria da dispersão de plumas na atmosfera	27
3.2. Simulador do sistema de controle ambiental	32
3.2.1. Módulo termo fonte.....	33
3.2.2. Módulo campo de vento	33
3.2.3. Módulo dispersão	34
4. DESENVOLVIMENTO DO MODELO	36
4.1. Geração de dados	36
4.2. Extração dos conjuntos de treinamento, teste e validação	43
4.3. Arquiteturas das redes neurais desenvolvidas.....	44
5. RESULTADOS	45
6. CONCLUSÃO.....	51
7. REFERÊNCIAS	53

1. INTRODUÇÃO

1.1. APRESENTAÇÃO DO PROBLEMA

Emissões de materiais radioativos para o ambiente podem ocorrer em caso de acidente durante a operação de uma usina nuclear e a contínua emissão desses materiais radioativos, ao longo do tempo, ocasiona a formação das chamadas plumas de dispersão atmosférica. Quando isso ocorre, é necessário que existam mecanismos capazes de estimar o comportamento da pluma para manter a segurança radiológica da população que se encontra dentro da zona afetada.

A formação de uma pluma radioativa depende de diversos fatores que são relacionados ao tipo de acidente ocorrido, às condições operacionais da planta, as condições meteorológicas e as condições topográficas locais. O tempo de emissão e a concentração de radionuclídeos presentes na pluma de dispersão são variáveis que dependem do tipo do acidente que é iniciado na usina nuclear. Variáveis relacionadas a meteorologia e topografia locais são determinantes, não somente para a formação da pluma de dispersão, mas também para estimativa da velocidade e direção na qual a pluma irá se movimentar conforme o acidente progride.

Para estimar o comportamento dessas plumas sistemas computacionais para cálculo de dispersão atmosférica de radionuclídeos são utilizados.

1.2. TRABALHOS RELACIONADOS

Timonin & Savelieva (2005) apresentaram uma técnica de rede neural de regressão geral (GRNN) para a previsão espacial automática de níveis de contaminação radioativa. De acordo com o autor, o GRNN gera resultados com um certo grau de suavização em relação a resultados reais e, portanto, o nível de satisfação dos resultados encontrados depende se o grau de suavização é aceitável em relação ao problema e o cenário a ser modelado.

Pereira et al. (2016) desenvolveram dois modelos de rede neural artificial para estimar a dispersão de pluma radioativa considerando um acidente severo de perda de refrigerante (LOCA). O primeiro modelo considera uma estrutura de rede neural do tipo *feedforward Multilayer Perceptron* (MLP) e a segunda rede é baseada em GRNN. Os resultados encontrados

foram satisfatórios e, em Pereira et al. (2017), a rede neural artificial MLP treinada no trabalho anterior, foi incorporada em um aplicativo de telefone móvel.

Desterro et al. (2020) substituiram a rede neural artificial MLP implementada por Pereira et al. (2016) por um modelo de rede neural *Deep Rectifier Artificial Neural Networks* (DRNN) com o objetivo melhorar a acurácia do modelo. A alteração da estrutura da rede neural MLP para DRNN consiste na substituição das funções de ativação da estrutura da rede neural, que antes utilizavam funções sigmoides, por funções de ativação retificadoras. O resultado mostrou que o treinamento da rede neural utilizando processamento paralelo diminuiu de 4 horas de duração para apenas 30 minutos sem perder a acurácia dos resultados previamente encontrados.

Outros trabalhos relacionados ao desenvolvimento de técnicas de rede neural artificial foram implementados para estimar a dispersão atmosférica de outros tipos de materiais poluentes não radioativos.

Boznan et al. (1993) desenvolveram um modelo de rede neural artificial para a estimar a dispersão atmosférica de SO₂ emitida a partir de uma usina termelétrica eslovena.

Yi & Prybutok (1996) criaram um modelo de rede neural que é capaz de estimar o nível máximo de emissões de gás ozônio em zonas industrializadas.

Cao et al. (2010) estudaram a distribuição de aerossol emitida ao ambiente a partir de uma fonte pontual considerando um curto período de emissão.

Lauret et al. (2013) utilizaram redes neurais para estimar a dispersão de gás com foco nas concentrações distribuídas mais perto do solo onde o grau de complexidade da modelagem é geralmente maior.

Hossain (2014) treinou um modelo capaz de calcular as concentrações de monóxido de carbono e materiais particulados emitidas em áreas urbanas utilizando dados de meteorologia e tráfego.

Qiu et al. (2018) desenvolveram uma estimativa da fonte de emissão e predição da dispersão atmosférica de gás através de técnicas de redes neurais artificiais, enxame de partículas e algoritmo de maximização de expectativa.

Oliveira et al. (2022) criaram um modelo híbrido entre aplicações computacionais de mecânica dos fluidos (CFD) com redes neurais para estimar a dispersão de gás, onde a aplicação de redes neurais tornou o modelo CFD mais rápido e eficiente.

1.3. JUSTIFICATIVA

Em um caso de acidente em uma usina nuclear onde ocorra vazamento de material radioativo na atmosfera, é necessário que existam ferramentas para estimativa das consequências na região afetada para que medidas mitigadoras sejam tomadas o mais rápido possível para a proteção da população.

Na literatura é possível encontrar diversos trabalhos que utilizam técnicas para estimar, em tempo real, o comportamento da pluma radioativa, o que é importante para que sejam tomadas decisões mitigadoras eficazes em uma situação de acidente. Porém, não foram encontrados trabalhos que proponham aplicações de redes neurais artificiais recorrentes que sejam capazes de prever o comportamento futuro da pluma na atmosfera com antecedência. Se ferramentas de Inteligência Artificial (IA) forem capazes de prever o comportamento de uma pluma com antecedência, então, um operador da usina terá mais tempo para avaliar as consequências antes de tomar decisões mitigadoras. Assim, o ganho de tempo na tomada de decisão possibilita que medidas de segurança radiológica mais eficazes possam ser avaliadas reduzindo a possibilidade de erros humanos.

1.4. OBJETIVO

Esse trabalho tem como objetivo desenvolver um modelo de previsão do comportamento de plumas radioativas utilizando redes neurais artificiais. É escolhida a técnica de redes neurais recorrentes (do inglês *Recurrent Neural Networks* - RNN), pois as RNN são especialmente adequadas para lidar com sequências de dados e capturar dependências temporais complexas.. Mais especificamente, são utilizadas as redes neurais recorrentes do tipo *Long Short Term Memory* (LSTM) (Goller & Kuchler, 1996).

É desenvolvido um modelo de previsão para a posição do ponto da pluma onde a taxa de dose de radiação é máxima. Para a geração dos dados necessários para o treinamento da rede neural recorrente, é utilizado um simulador, baseado no simulador do Sistema de Controle Ambiental (SCA) da Central Nuclear Almirante Álvaro Alberto (CNAAA), doravante denominado Simulador de Dispersão Atmosférica (SDA), que implementa o modelo matemático básico de dispersão atmosférica para os diferentes tipos de acidentes, porém utilizando dados típicos de reatores PWR. Diversas arquiteturas de redes neurais artificiais recorrentes são testadas para

ajustar o melhor modelo que consiga prever a movimentação dessas plumas. Ao fim, é selecionada a melhor arquitetura de rede capaz de prever a movimentação do ponto de máxima taxa de dose de corpo inteiro ao longo do tempo.

1.5. ORGANIZAÇÃO DO TRABALHO

Esse trabalho é dividido em seções, onde as duas seções seguintes têm como objetivo fundamentar a teoria que envolve redes neurais artificiais recorrentes e modelos de dispersão atmosférica respectivamente. Na seção de modelagem de dispersão atmosférica, também são introduzidos o funcionamento e a metodologia referente ao simulador utilizado para gerar os dados de treinamento da rede. A Seção 4 introduz as arquiteturas de rede neural recorrente que foram testas e geração dos dados. A Seção 5 apresenta os resultados encontrados a partir do treinamento das redes testadas e verifica qual rede foi a mais eficiente na previsão do ponto de máxima dose de corpo inteiro. As seções finais consistem na apresentação de conclusão, sugestões para trabalhos futuros e referências bibliográficas utilizadas.

2. REDES NEURAIS ARTIFICIAIS

A rede neural artificial é uma vertente da inteligência artificial utilizada para o aprendizado de máquina. A RNA consiste em criar um modelo matemático que seja capaz de reproduzir o comportamento da transferência de informações do sistema nervoso humano.

O neurônio é uma unidade básica do sistema nervoso humano que possui a capacidade de transmitir informações através de impulsos elétricos. A Figura 2.1 apresenta a estrutura esquemática de um neurônio que é constituído de um corpo celular com diversos dendritos (canais de entrada) e um axônio (canal de saída). Os neurônios estão interligados uns aos outros através de conexões entre axônios e dendritos, denominadas sinapses. É nessa interconexão que as informações são transmitidas ao longo do sistema nervoso. Estima-se que o cérebro humano possua dezenas de bilhões de neurônios conectados formando uma complexa rede neural.

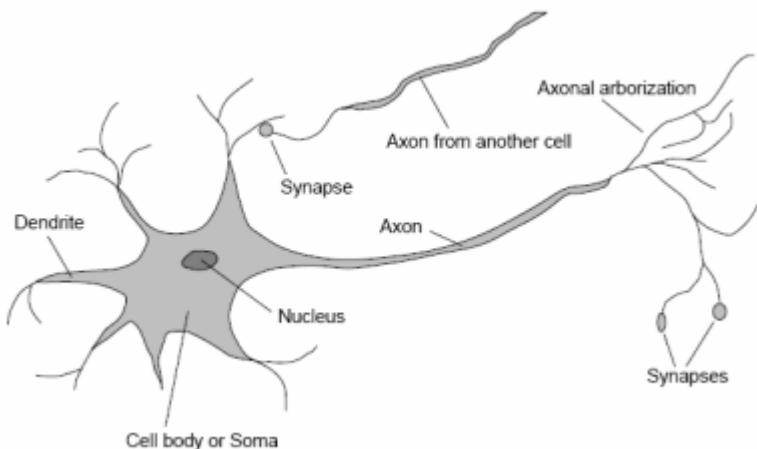


Figura 2.1 – Estrutura de um neurônio.
Fonte: Domínio <http://www.ic.uff.br/>.

A transmissão de informação através do sistema nervoso é feita a partir de impulsos elétricos recebidos pelos dendritos de um neurônio. Se um impulso elétrico ultrapassar certo limiar, então o neurônio irá liberar uma substância neurotransmissora que irá fluir através de seu axônio até atingir as suas sinapses onde haverá as conexões com os dendritos dos neurônios adjacentes. O neurotransmissor pode ser inibido ou excitado nas atividades pós-sinápticas, aumentando ou diminuindo a intensidade com que a informação a ser transmitida será propagada ao longo dessa rede neural até o cérebro.

2.1. NEURÔNIO ARTIFICIAL

McCulloch & Pitts (1943) foram os primeiros cientistas a desenvolver uma metodologia capaz de reproduzir matematicamente o funcionamento de um neurônio humano. O modelo consiste em uma rede bastante simples de dois neurônios conectados a um terceiro neurônio, onde as informações binárias são transferidas para o terceiro neurônio somente se extrapolado um dado limiar de ativação. O neurônio artificial criado por McCulloch & Pitts (1943) ficou conhecido na literatura como o neurônio MP (Figura 2.2).

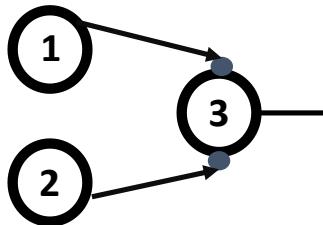


Figura 2.2 – Representação de um neurônio MP.

Fonte: o autor (2023).

2.2. NEURÔNIO PERCEPTRON

O modelo de 14eurônio MP desenvolvido por McCulloch & Pitts (1943) até então era bastante simples, mas era uma primeira tentativa de reproduzir o funcionamento de um neurônio. Porém, ainda havia a necessidade de melhor implementar a inibição ou excitação das atividades pós-sinápticas na transferência de informações entre neurônios. Rosenblatt (1958) desenvolveu um modelo matemático que acrescentava um valor de peso w_i para cada neurônio x_i permitindo com que esse peso pudesse ser ajustado de modo a transferir as informações interneurônios nas proporções adequadas como uma forma de atividade pós-sináptica. O modelo desenvolvido por Rosenblatt (1958) se sobrepôs ao modelo anteriormente criado e ficou reconhecidamente chamado de Perceptron (Figura 2.3).

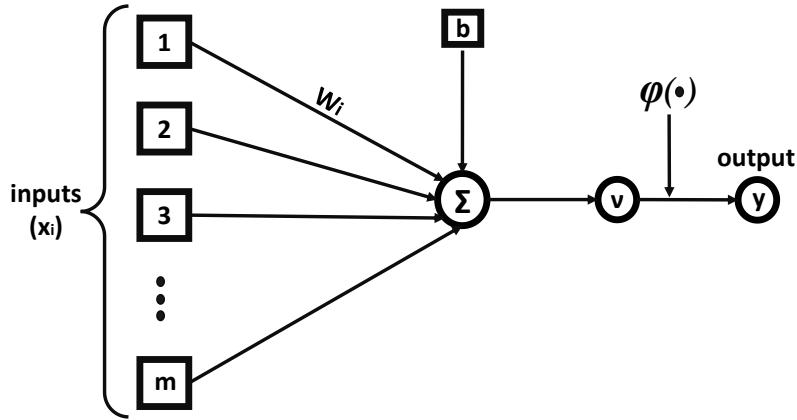


Figura 2.3 – Representação de uma rede neural Perceptron.

Fonte: o autor (2023).

É possível calcular a saída de um Perceptron quando os valores de entrada são introduzidos. Considerando que x_i representa os valores de entrada, w_i representam os pesos sinápticos e b é o valor de viés (bias) introduzido, temos que v representa a pré-ativação do Perceptron, que é dado pela Equação 2.1.

$$v = \sum_{i=1}^n x_i w_i + b \quad (2.1)$$

Então, o resultado y do neurônio de saída é dado por $y = \varphi(v)$, onde φ representa uma função de ativação conhecida. Existem diversas funções de ativação que são comumente utilizadas, como, por exemplo: função degrau, linear, sigmoide logística e tangente hiperbólica.

2.3. REDE NEURAL TIPO PERCEPTRON MULTICAMADAS

O modelo desenvolvido para o Perceptron recebeu muitas críticas devido a dificuldade de se obter boas respostas relacionadas a problemas não-lineares, como por exemplo, a função XOR. Isso ocorria porque selecionar apenas uma certa quantidade de neurônios em apenas uma camada ainda restringia bastante a capacidade da rede neural de aprender problemas mais complexos. Para superar essa limitação, Rumelhart et al. (1986) desenvolveram um modelo de rede neural artificial Perceptron multicamadas (MLP – *Multi-layer Perceptron*). O modelo propunha a utilização de camadas de neurônios (Figura 2.4). Desta forma foi possível criar modelos de redes neurais que fossem capazes de resolver problemas mais complexos.

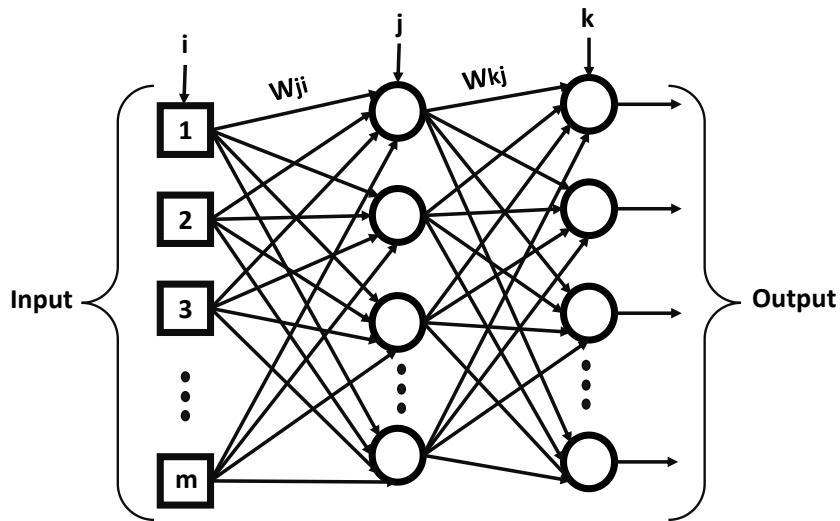


Figura 2.4 – Representação esquemática de uma rede neural MLP.
Fonte: o autor (2023).

2.4. TREINAMENTO DE UMA MLP

Considerando que uma rede neural tem como objetivo o aprendizado, é necessário que exista um treinamento. Ou seja, é necessário apresentar um grupo de exemplos do problema que se deseja resolver para que a rede seja capaz de aprender a resolvê-los. Esse treinamento consiste em introduzir dados de entrada (parâmetros do problema) com os seus respectivos dados de saída (resultados), para que a rede tenha a capacidade de ajustar matematicamente os valores de seus pesos sinápticos de modo a criar um modelo que seja capaz de reproduzir aquele problema.

O ajuste de peso no treinamento de uma rede neural é comumente feito através do algoritmo de retropropagação (backpropagation) de erro. Esse algoritmo consiste no reajuste dos pesos a partir do valor do erro quadrático entre os dados de saída desejados e os dados de saída calculados pela rede neural. O algoritmo de retropropagação de erro pode ser descrito da seguinte forma:

1. Inicializar todos os pesos de todas as camadas com valores aleatórios não nulos.
2. Propagar as entradas x_i até a sua saída com os respectivos pesos atuais.
3. Calcular o erro entre a saída calculada pela rede e o valor a ser aprendido.
4. Ajustar todos os pesos w_i da rede de forma a minimizar o erro.
5. Repetir o processo 3 e 4 para todas as camadas ocultas até a primeira camada.

6. Repetir o processo 1 a 5 até que um critério de parada seja atingido, como um número pré-definido de épocas de treinamento ou quando o erro alcança um valor aceitável.

É possível formular um cálculo matemático para o ajuste dos pesos feito pelo algoritmo de retropropagação para uma rede neural do tipo MLP.

Considerando um neurônio k da camada de saída da rede neural MLP da Figura 2.4, pode-se definir e_k como o erro instantâneo dada pela Equação 2.2.

$$e_k = d_k - y_k \quad (2.2)$$

Onde d_k é o valor desejado como resultado e y_k é o valor efetivamente calculado pela propagação da rede neural com os pesos sinápticos atuais.

Então calcula-se o gradiente local para o neurônio k da camada de saída que é dado por:

$$\delta_k = e_k \varphi_k'(v_k) \quad (2.3)$$

$\varphi'(v_k)$ refere-se a derivada da função de ativação do neurônio k da camada de saída.

Portanto, o ajuste do peso w_{kj} é calculado pela seguinte expressão:

$$\Delta w_{kj} = \eta \delta_k y_j \quad (2.4)$$

Δw_{kj} representa a diferença entre o peso a ser atualizado menos o peso atual e η é a taxa de aprendizado.

Agora é possível notar que o cálculo do ajuste de peso de uma camada de saída é, em parte, calculado pelo erro instantâneo que depende do valor de saída desejado (d) e o valor de saída efetivamente calculado pela rede neural (y) para aquele conjunto de pesos atuais. Porém, considerando a conexão sináptica entre um neurônio de uma camada escondida e uma anterior, não haverá mais conexão direta com a saída da rede neural. Portanto, o cálculo do erro total de um neurônio de uma camada interna não será mais dado em função de um erro instantâneo, mas de um somatório de todos os erros das camadas posteriores multiplicado pelo seu respectivo peso. Ou seja, considerando um neurônio j da camada escondida da rede neural MLP apresentada na Figura 2.4, temos que o cálculo do erro é calculado pela Equação 2.5.

$$\delta_j = \varphi_j'(v_j) \sum_{k \in C} \delta_k w_{kj} \quad (2.5)$$

E, consequentemente, o ajuste do peso w_{ji} é dado pela seguinte expressão:

$$\Delta w_{ji} = \eta \delta_j y_i \quad (2.6)$$

Finalmente, é possível escrever a equação global do ajuste de peso w_{kj} e w_{ji} apresentados na Equação 2.7 e Equação 2.8 respectivamente.

$$\Delta w_{kj} = \eta (e_k \varphi'_k(v_k)) y_i \quad (2.7)$$

$$\Delta w_{ji} = \eta \left(\varphi'_j(v_j) \sum_{k \in C} (e_k \varphi'_k(v_k) w_{kj}) \right) y_i \quad (2.8)$$

De forma esquemática, a Figura 2.5 abaixo apresenta um formato esquemático para o algoritmo de retropropagação considerando uma rede neural perceptron multicamadas. A estrutura da rede da figura é organizada em uma camada de entrada de dois neurônios, uma camada intermediária escondida com três neurônios e uma camada de saída com dois neurônios e considerando uma função de ativação qualquer φ .

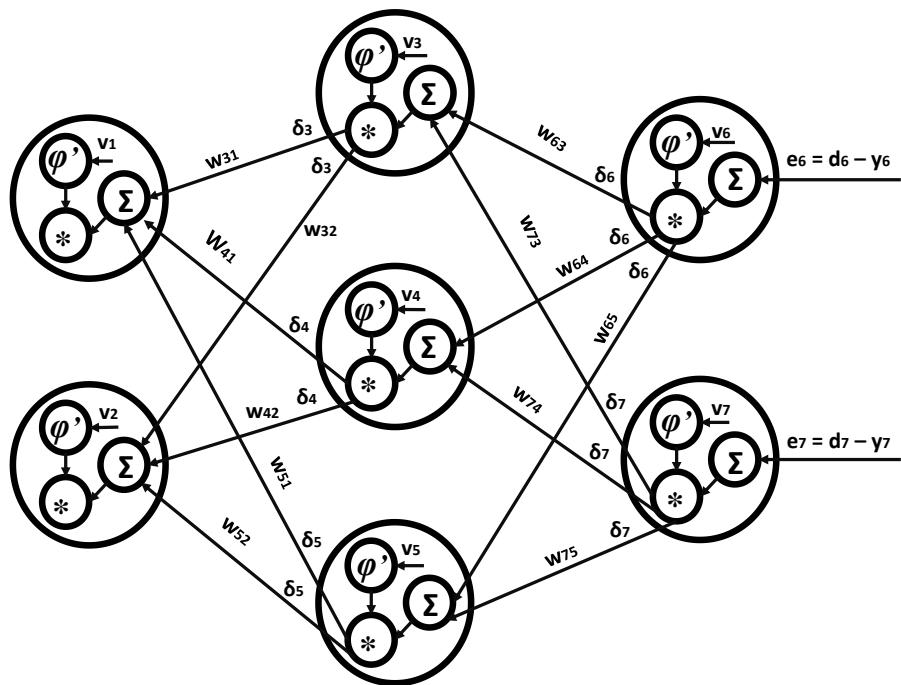


Figura 2.5 – Algoritmo de retropropagação para redes MLP.
Fonte: o autor (2023).

Observando as Equações 2.7 e 2.8 acima, fica claro que, conforme o processo de retropropagação progride, a quantidade de termos referente ao gradiente local da derivada da

função de ativação ($\varphi'(v)$) tende a aumentar. A influência matemática que o acúmulo desses termos gera no cálculo do ajuste de pesos por retropropagação ocasionou diversas críticas a esses modelos que se sucederam ao longo dos anos.

2.5. REDE NEURAL PROFUNDA

O desenvolvimento do MLP, apesar de inovador, ainda assim recebeu críticas sobre a inconsistência de resultados quando os problemas complexos demandavam estruturas de redes neurais com muitas camadas (redes neurais profundas). Hochreiter (1991, 1998) verificou que a rede neural profunda possuía problemas de desaparecimento de gradiente e esse efeito ficava cada vez mais evidente conforme a quantidade de camadas com gradientes locais aumentavam. Glorot & Bengio (2010) fizeram um estudo de aplicações de funções de ativação sigmoide em estruturas de redes MLP profundas e verificaram que nas camadas mais profundas o desempenho durante o treinamento era ruim ou próximo de zero.

Para verificar o efeito da função sigmoide no desaparecimento de gradiente e, consequentemente, as inconsistências no treinamento de redes neurais profundas é necessário entender o comportamento dessa função. A Equação 2.9 define a função sigmoide $\varphi(x)$ e a Figura 2.6 apresenta o gráfico da função com a sua respectiva derivada $\varphi'(x)$.

É possível verificar que a função sigmoide possui uma derivada que tende a zero quando x tende a $+\infty$ ou $-\infty$. E, como foi visto anteriormente, a derivada da função de ativação que compõe o gradiente local é importante para o cálculo dos reajustes de pesos no treinamento de uma rede neural. No processo de retropropagação de erro e correção de pesos, são feitas, a cada camada, multiplicações pela derivada da função. Com o aumento do número de camadas, cresce proporcionalmente a quantidade de multiplicações sucessivas pela derivada da função. Assim sendo, como a derivada possui valores menores do que a unidade (podendo ser relativamente bem menores) a multiplicação dos termos tende a zero para uma grande quantidade de camadas e, consequentemente, acarreta pouco ou praticamente nenhum treinamento da rede ao longo das épocas.

$$\varphi(x) = \frac{1}{1 + e^{-x}} \quad (2.9)$$

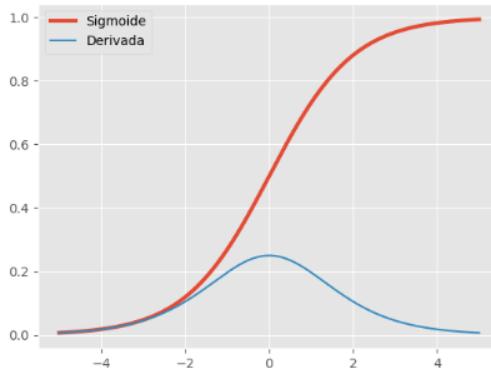


Figura 2.6 – Gráfico da função sigmoide e sua derivada.
Fonte: Domínio <https://matheusfacure.github.io/>.

Com o objetivo de resolver o problema das redes neurais profundas, Glorot et al. (2011) desenvolveu um grupo de funções de ativação chamadas de funções retificadas. Dentre elas, a função mais conhecida é a chamada ReLU (*Rectified Linear Unit*). Essa função de ativação proposta por Glorot et al. (2011) possui características capazes de superar o problema de desaparecimento de gradiente que tendem surgir em redes neurais profundas. A função de ativação ReLU é definida pela Equação 2.10 e a Figura 2.7 apresenta o comportamento gráfico da função e de sua respectiva derivada.

$$\varphi(x) = \max(0, x) \quad (2.10)$$

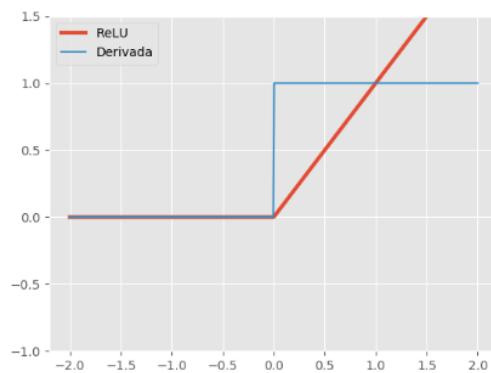


Figura 2.7 – Gráfico da função ReLU e sua derivada.
Fonte: Domínio <https://matheusfacure.github.io/>.

É possível verificar na Figura 2.7 que a função ReLU possui uma derivada constante para valores positivos, enquanto a função sigmoide possui um comportamento decrescente. Dessa forma, é possível garantir que o reajuste de pesos do treinamento de uma rede neural com algoritmo de retropropagação não sofrerá efeitos de desaparecimento de gradiente com tanta

velocidade conforme a quantidade de camadas na estrutura da rede aumenta. Por outro lado, para valores negativos o gradiente é nulo, causando uma esparsidade no modelo, o que traz benefícios em termos de eficiência no treinamento. O modelo desenvolvido de redes neurais que possui uma função de ativação retificadora ficou conhecida como redes neurais profundas retificadas (do inglês *Deep Rectifier Neural Network* - DRNN) .

2.6. REDE NEURAL RECORRENTE

As estruturas de redes neurais apresentadas são conhecidas como redes neurais do tipo *feedforward*. Por exemplo, é possível verificar na Figura 2.4 que uma rede neural do tipo MLP é uma rede do tipo *feedforward* pois ela sempre propaga os valores na direção da entrada (esquerda) para a saída (direita), não havendo retorno (realimentação) de nenhuma informação anterior. Verifica-se que esse tipo de estrutura *feedforward* não mantém nenhuma memória do que foi calculado anteriormente e todas as suas informações são consideradas independentes a cada rodada.

Apesar de existirem inúmeras aplicações para redes neurais *feedfowards*, muitos problemas dependem de informações passadas para que se possa complementar um raciocínio e chegar a uma conclusão. Um exemplo claro é a leitura de um livro, onde é necessário que o leitor mantenha em memória o que foi lido anteriormente para entender o contexto da frase presente no meio do texto. Caso contrário, a informação contida naquela frase não fará sentido lógico no contexto do texto completo. Caso contrário, a informação daquela frase no contexto do texto lido não fará nenhum sentido lógico. Outro exemplo é quando um indivíduo está assistindo um filme e procura predizer o que irá acontecer no final. Note que só será possível prever o final de um filme depois de assisti-lo até certo ponto, a partir do qual conclusões podem ser tomadas sobre um possível desfecho. Ou seja, nota-se que existe uma dependência temporal de todo um histórico de informações para que uma conclusão possa ser tomada.

Em geral, problemas que envolvem séries temporais e sequências, não conseguem ser adequadamente solucionados com redes neurais artificiais do tipo *feedforward*. Para superar isso, Goller & Kuchler (1996) desenvolveram as conhecidas redes neurais artificiais recorrentes. Essas arquiteturas de redes neurais possuem a capacidade de serem retroalimentadas com informações calculadas anteriormente. Assim como as arquiteturas *feedforward*, que recebem os dados de entrada, a rede neural artificial recorrente possui um

componente extra que retorna informações anteriores junto aos dados de entrada. Essa retroalimentação particular das redes neurais artificiais recorrentes são chamadas de *feedbacks*. Dessa forma, o modelo da rede incorpora a capacidade de manter informações passadas de modo a tomar decisões para o passo seguinte. A Figura 2.8 apresenta a diferença na arquitetura entre as redes neurais artificiais recorrentes e a rede neural artificial *feedforward*.

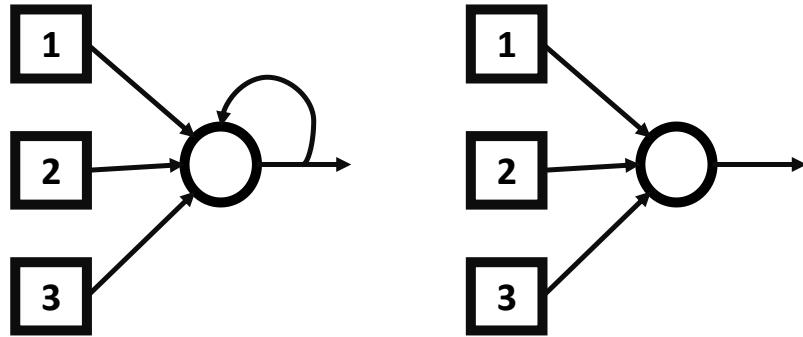


Figura 2.8 – Diferença entre uma rede neural recorrente e uma rede neural *feedforward*.
Fonte: o autor (2023).

De forma simplificada, pode-se definir uma célula recorrente como na Figura 2.9, onde h_t representa o estado atual da célula, x_t e y_t são as entradas e saídas da célula respectivamente. A Equação 2.11 apresenta a relação geral de recorrência para o mecanismo de memorização temporal da célula, onde h_{t-1} é o estado anterior da célula. Note que f_w representa a aplicação da função de ativação relacionada ao respectivo peso w daquela célula.

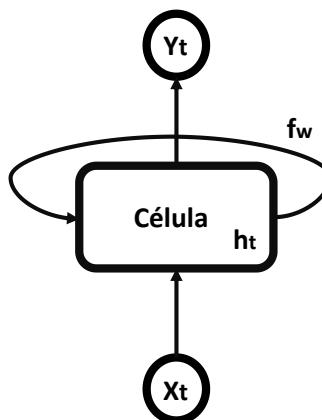


Figura 2.9 – Representação esquemática de uma célula de uma estrutura recorrente.
Fonte: o autor (2023).

$$h_t = f_w(x_t, h_{t-1}) \quad (2.11)$$

Uma rede neural recorrente pode ser representada temporalmente por uma sequência de células recorrentes alinhadas, como representa a Figura 2.10. É importante notar que além dos pesos dos dados de entrada W_{xh} e os pesos dos dados de saída W_{hy} da célula, agora, a rede neural recorrente, possui um novo componente de pesos W_{hh} que conecta uma célula a outra para criar o aspecto de recursividade.

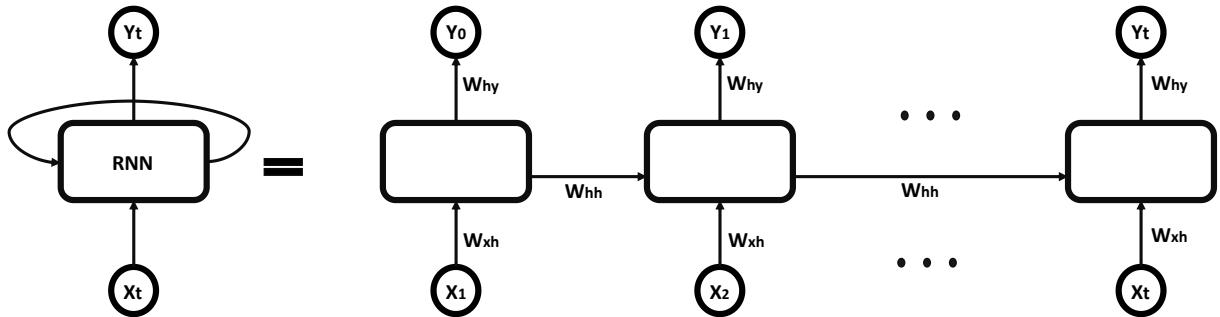


Figura 2.10 – Desdobramento temporal de uma rede neural recorrente.

Fonte: o autor (2023).

Considerando que a função de ativação f_w pode ser considerada uma tangente hiperbólica, podemos reescrever a equação do estado da célula h_t da seguinte forma:

$$h_t = \tanh (\mathbf{W}_{hh}^T h_{t-1} + \mathbf{W}_{xh}^T x_t) \quad (2.12)$$

E consequentemente, teremos que a saída y_t da célula recorrente é dada por:

$$y_t = \mathbf{W}_{hy}^T h_t \quad (2.13)$$

O algoritmo de retropropagação para os ajustes de pesos e treinamento da rede neural artificial recorrente segue o raciocínio similar ao de retropropagação das redes neurais *feedforward*. No caso de redes neurais recorrentes, o ajuste de pesos é nomeado de retropropagação temporal (*backpropagation through time*). Porém, redes neurais recorrentes possuem ainda mais pesos a serem ajustados devido a interconexão entre as células recorrentes. Isso acentua, mais uma vez, o problema de perda de gradiente discutido na Seção 2.4 para o caso de redes neurais profundas. Agora, considerando que a rede neural possui uma memória temporal, o problema de perda de gradiente é visto como uma forma de perda de memória. Ou seja, quanto maior for a série

temporal que se deseja modelar e, consequentemente, quanto maior for a quantidade de células recorrentes introduzidas na rede neural, maior será a quantidade de pesos disponíveis para serem ajustados pelo algoritmo de retropropagação. Portanto, verifica-se que a rede neural recorrente pode chegar a um limite de memória que, dependendo do problema, pode se configurar uma perda de memória de curto prazo.

2.7. REDES NEURAL TIPO LSTM

Para contornar o problema devido a perda de capacidade de memória de uma rede neural recorrente, Hochreiter & Schmidhuber (1997) desenvolveram uma estrutura celular recorrente chamada de memória de curto a longo prazo (do inglês *Long Short-Term Memory* - LSTM). Para entender como esse tipo de célula desenvolvida possui a capacidade de superar o problema do desaparecimento de gradiente, é necessário entender todos os componentes da estrutura de uma célula LSTM.

A implementação da LSTM desenvolvida consiste em uma célula recorrente que possui filtros capazes de avaliar o nível de relevância das informações dos dados de entrada na rede. Através de ponderações matemáticas, a célula LSTM possui terminais que são capazes de manter informações julgadas relevantes para solucionar o problema e excluir aquelas as informações consideradas menos relevantes. Dessa forma, os portões da célula LSTM tornam possível estender as memórias temporais. A Figura 2.11 apresenta uma célula LSTM com seus respectivos terminais de avaliação de informação.

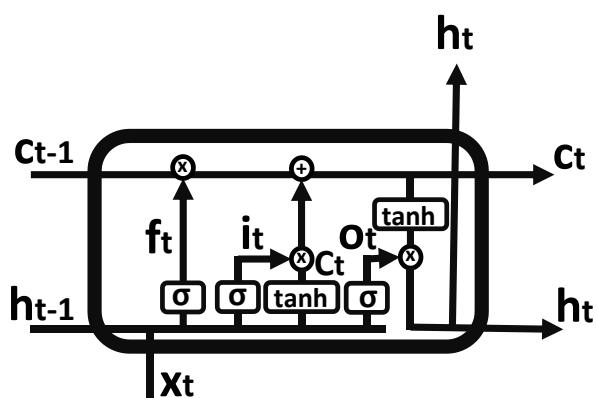


Figura 2.11 – Componentes de uma célula LSTM.
Fonte: o autor (2023).

A célula LSTM possui três grupos de dados de entrada: os valores atuais das entradas (x_t), as informações do estado anterior da célula (h_{t-1}) e as informações globais (c_{t-1}). A diferença entre os dados de estado anterior da célula e os dados de entrada globais é que os dados de estado anterior são apenas temporários e substituídos por novas informações que serão atualizados pela célula atual, enquanto as informações globais se mantêm ao longo de toda a série temporal. Além disso, a saída global não interfere nos cálculos realizados pelos terminais da célula, ele apenas acumula de informações ao longo do tempo, como uma espécie de memória interna.

A LSTM possui três portões internos característicos, são eles: o terminal de esquecimento (f_t), o terminal de entrada (i_t) e o terminal de saída (o_t).

É no terminal de esquecimento (f_t) que a rede neural elimina as informações que são julgadas não relevantes para solucionar o problema. O cálculo é feito através de ponderações realizadas com o auxílio da função sigmoide:

$$f(t) = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.14)$$

Uma vez que o dado de entrada passa pelo terminal de esquecimento, a informação que é aceita pela célula segue para o terminal de entrada (i_t). Este refere-se ao terminal onde o dado de entrada é de fato tratado para seguir adiante como output da célula da rede neural. O terminal de entrada consiste em duas etapas, onde a primeira etapa consiste em calcular a informação que será atualizada para a próxima célula com o auxílio da função sigmoide (Equação 2.15) e uma segunda etapa que filtra as informações que serão acumuladas no terminal global (Equação 2.16).

$$i(t) = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2.15)$$

$$C(t) = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (2.16)$$

O terminal de saída (o_t) consiste na construção dos dados de saída do estado atual de uma célula dentro da rede neural. Ela consiste em um filtro realizado com o auxílio da função sigmoide (Equação 2.17).

$$o(t) = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (2.17)$$

Finalmente, a célula LSTM possui dois dados de saída: um local (h_t) que é alterado por cada célula e um global (c_t) que é continuamente memorizado ao longo das células. O cálculo de

saída local e o dado de saída global são dados pela Equação 2.18 e Equação 2.19 respectivamente.

$$c_t = f_t * c_{t-1} + i_t * C_t \quad (2.18)$$

$$h_t = o_t * \tanh(c_t) \quad (2.19)$$

Com o objetivo de verificar como as redes neurais criadas por sucessivas células do tipo LSTM possuem a capacidade de contornar o problema do desaparecimento de gradiente, primeiro é importante reescrever a equação para o cálculo de c_t , referente a saída global da célula LSTM, com suas devidas substituições.

$$c_t = (\sigma(W_f[h_{t-1}, x_t])) * c_{t-1} + \sigma(W_i[h_{t-1}, x_t]) * \tanh(W_c[h_{t-1}, x_t]) \quad (2.20)$$

O cálculo do gradiente ao longo das células LSTM é calculado aplicando $\frac{\partial c_t}{\partial c_{t-1}}$ na Equação 2.20 e, portanto, podemos reescrevê-la da seguinte forma:

$$\frac{\partial c_t}{\partial c_{t-1}} = (\sigma(W_f[h_{t-1}, x_t])) + \frac{\partial c_t}{\partial c_{t-1}} (\sigma(W_i[h_{t-1}, x_t]) * \tanh(W_c[h_{t-1}, x_t])) \quad (2.21)$$

Verifica-se que, considerando sucessivas células, o valor do gradiente tende a ser aproximado pela seguinte expressão:

$$\frac{\partial c_t}{\partial c_{t-1}} \cong (\sigma(W_f[h_{t-1}, x_t])) \quad (2.22)$$

Portanto, verifica-se que o gradiente possui um valor, diferente de zero, ao qual irá se aproximar quando sucessivas células LSTM são adicionadas a uma estrutura de uma rede neural. Com isso, o problema do desaparecimento de gradiente consegue ser superado em aplicações de redes neurais recorrentes.

3. SIMULADOR DE DISPERSÃO ATMOSFÉRICA DE PLUMAS RADIOATIVAS

Existem ferramentas que simulam a dispersão de poluentes e materiais radioativos na atmosfera. Cada simulador pode utilizar uma solução analítica específica que varia de acordo com a complexidade do problema que se deseja modelar. Diversas formulações analíticas para o cálculo da dispersão de plumas na atmosfera foram desenvolvidas ao longo dos anos.

Nesse capítulo serão descritos alguns modelos analíticos que são mais utilizados em simuladores de dispersão de plumas e será apresentado o simulador SDA, que foi utilizado na geração de dados para o treinamento da rede neural recorrente desse trabalho.

3.1. TEORIA DA DISPERSÃO DE PLUMAS NA ATMOSFERA

A solução analítica para a dispersão atmosférica é descrita principalmente por dois modelos: eulerianos e lagrangianos. O modelo euleriano consiste na solução de uma equação diferencial baseada na conservação de massa com um sistema de referência fixo. O modelo lagrangiano considera que a trajetória de uma partícula é feita a partir de um campo turbulento que depende de condições iniciais e físicas. No caso de dispersão atmosférica, esse campo considerado no modelo lagrangiano, geralmente refere-se a um campo de vento.

Considerando uma metodologia euleriana, para um dado material de concentração $C(x,y,z,t)$ na atmosfera, a modelagem analítica é feita a partir da equação de conservação de massa dada por:

$$\frac{\partial C}{\partial t} = -\nabla(C\vec{v}) + \nabla(D_c\nabla C) + S_c \quad (3.1)$$

Onde o primeiro termo consiste em quantificar a variação temporal da concentração C de um material no meio, o segundo termo e terceiro termo referem-se a advecção e difusão respectivamente e S_c é o termo fonte de emissão.

Como o campo de vento (\vec{v}) é calculado pelos dados meteorológicos que são medidos em campo e a fonte (S_c) está sempre em constante monitoramento, então pode-se concluir que única variável da equação acima é a concentração do material C .

Para seguir adiante com os cálculos da equação diferencial para descrever a dispersão da concentração do material C , é necessário introduzir termos referentes aos efeitos de turbulência. A teoria da turbulência de Reynolds, verifica que, para descrever o efeito de turbulência, é necessário sobrepor dois efeitos: o efeito médio temporal e o efeito de flutuação turbulenta. Ou seja, em teoria, isso consiste na substituição das variáveis do problema por variáveis referentes a média temporal e de flutuação.

Considerando uma variável genérica (ϕ), temos que a média temporal ($\bar{\phi}$) e a flutuação turbulenta (ϕ^*) são dadas respectivamente pelas equações abaixo.

$$\bar{\phi} = \frac{1}{\Delta t} \int_0^{\Delta t} \phi(t) dt \quad (3.2)$$

$$\phi^* = \frac{1}{\Delta t} \int_0^{\Delta t} \phi^{*(t)} dt = 0 \quad (3.3)$$

Com as duas definições acima, é possível verificar que $\phi = \bar{\phi} + \phi^*$, garantindo que o efeito da variável ϕ é, de fato, a sobreposição desses dois efeitos.

Voltando a equação governante, de acordo com a teoria de Reynolds, a média temporal e flutuação turbulenta são introduzidas nas variáveis de velocidade (v) e concentração (C). Portanto, a equação governante média pode ser reescrita da seguinte forma:

$$\begin{aligned} \frac{\partial \bar{C}}{\partial t} + \frac{\partial \bar{C}^*}{\partial t} = & -\nabla(\bar{C}\bar{v}) - \nabla(\bar{C}^*\bar{v}) - \nabla(\bar{C}\bar{v}^*) - \nabla(\bar{C}^*\bar{v}^*) + \nabla(D_c \nabla \bar{C}) \\ & + \nabla(D_c \nabla \bar{C}^*) + \bar{S}_c \end{aligned} \quad (3.4)$$

Devido a relação entre média temporal e flutuação turbulenta, algumas propriedades podem ser estabelecidas de modo a reduzir a equação acima, tais como $\bar{\phi}^* = 0$ e $\bar{\phi}\phi^* = 0$. Aplicando essas propriedades na equação acima, temos que:

$$\frac{\partial \bar{C}}{\partial t} = -\nabla(\bar{C}\bar{v}) - \nabla(\bar{C}^*\bar{v}^*) + \nabla(D_c \nabla \bar{C}) + \bar{S}_c \quad (3.5)$$

Explicitando os componentes turbulentos para cada direção (u^*, v^*, w^*) implícitos pelo campo de vento v^* e considerando uma difusão isotrópica, é possível reescrever a equação acima no formato conhecido para dispersão atmosférica de um material de concentração C .

$$\frac{\partial \bar{C}}{\partial t} = -\nabla(\bar{C}\bar{v}) + D_c \nabla^2 \bar{C} + \bar{S}_c - \frac{\partial(\bar{u}^* \bar{C}^*)}{\partial x} - \frac{\partial(\bar{v}^* \bar{C}^*)}{\partial y} - \frac{\partial(\bar{w}^* \bar{C}^*)}{\partial z} \quad (3.6)$$

Existem diversos modelos matemáticos desenvolvidos com o objetivo de aproximar o movimento de uma pluma de dispersão atmosférica solucionando a equação diferencial acima. Dentre eles se destacam os modelos gaussianos e os modelos de bufadas (*puffs*), onde cada modelo desenvolvido possui as suas próprias premissas e simplificações. Portanto, cada abordagem é indicada para diferentes tipos de aplicações e níveis de complexidade do problema a ser modelado.

O modelo gaussiano (Gifford, 1960) considera que a dispersão de uma pluma é proveniente de uma fonte de emissão estacionária e o material emitido possui escoamento estacionário e homogêneo. Ou seja, trata-se de um modelo de dispersão que considera que a emissão é contínua e sempre em uma mesma direção. A Figura 3.1 apresenta um exemplo esquemático de uma fonte emissora criando uma pluma de dispersão que poderia ser aproximada por um modelo gaussiano.

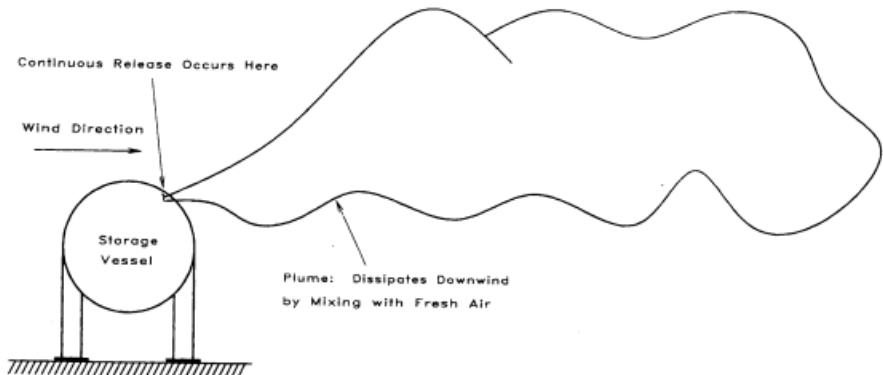


Figura 3.1 – Pluma de dispersão gaussiana a partir de uma fonte emissora.
Fonte: Crowl & Louvar (2011).

A equação que rege a dispersão atmosférica por um modelo gaussiano é derivada da equação da conservação de massa sob a hipótese que a turbulência é homogênea e constante, o fluxo de emissão é constante, o contaminante é estável e a topografia é constante. Com isso, a equação geral do modelo gaussiano que rege a dispersão atmosférica de uma pluma é apresentada na Equação 3.7 a seguir.

$$C(x, y, z) = \frac{Q}{2 \pi \sigma_y \sigma_z v} e^{-\frac{1}{2} \left(\frac{y}{\sigma_y} \right)^2} \left(e^{-\frac{1}{2} \left(\frac{z-H}{\sigma_z} \right)} + \alpha e^{-\frac{1}{2} \left(\frac{z+H}{\sigma_z} \right)} \right) \quad (3.7)$$

Onde C é a concentração de um poluente no ponto de coordenadas x, y, z ; Q é a taxa de emissão do poluente da fonte emissora; σ_y é o coeficiente de dispersão horizontal da concentração da pluma em função da direção do vento e da distância da fonte; σ_z é o coeficiente de dispersão vertical da concentração da pluma em função da direção do vento e da distância da fonte; v é a velocidade média do vento na altura da fonte emissora; z é ponto de amostragem em relação ao solo; y é o ponto de amostragem na direção y ; H é a altura geométrica da fonte emissora; α é o coeficiente de reflexão no solo.

O modelo Gaussiano, apesar de ser um grande avanço para as estimativas do comportamento das plumas radioativas, ainda era muito restrito a problemas mais simples devido as simplificações feitas. Com isso, foi desenvolvido o modelo de dispersão por bufadas (Sullivan et al., 1993; Ehrhardt & Weis, 1996) que pode ser considerado como um modelo híbrido entre o modelo gaussiano e langrangiano. Isso acontece pois considera-se que os cálculos das concentrações são feitos considerando distribuições gaussianas enquanto a trajetória de movimentação da pluma segue um modelo lagrangiano. Esse modelo é conhecido como bufadas pois considera que toda a liberação é feita de uma única vez e, considerando a ocorrência de sucessivas bufadas, é possível recriar um perfil de uma pluma. A Figura 3.2 apresenta um exemplo esquemático de uma fonte emissora gerando sucessivas bufadas.

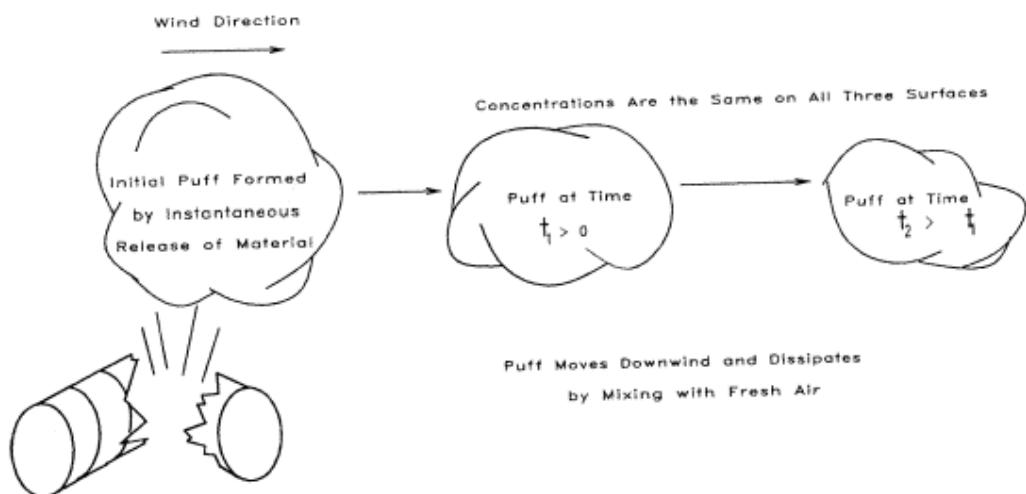


Figura 3.2 – Sucessivas bufadas provenientes de uma fonte emissora.
Fonte: Crowl & Louvar (2011).

A pluma de dispersão no modelo de bufadas é recriada a partir de uma sequência de bufadas. Portanto, é possível escrever a equação que rege o modelo de bufadas para a concentração $C(x,y,z)$ de um certo material emitido como a somatória de todas as bufadas ocorridas.

$$C(x, y, z) = \frac{Q\Delta t}{2\pi^2} \sum_{k=1}^N \frac{1}{\sigma_{xk}^2 \sigma_{yk}^2 \sigma_{zk}^2} e^{\left(-\frac{(x_k-x)^2}{2\sigma_{xk}^2} - \frac{(y_k-y)^2}{2\sigma_{yk}^2} - \frac{(z_k-z)^2}{2\sigma_{zk}^2}\right)} \quad (3.8)$$

Onde $Q\Delta t$ refere-se ao termo fonte; N é o número total de bufadas; (x_k, y_k, z_k) referem-se à posição da k -ésima bufada; σ_{ik} refere-se ao desvio da i -ésima direção da distribuição gaussiana da k -ésima bufada.

A diferença entre os métodos gaussianos e o método de bufadas é que o modelo matemático gaussiano considera sempre que a emissão é feita de forma contínua seguindo uma trajetória fixa. Já o modelo matemático de bufadas não considera uma emissão contínua e sim uma liberação completa e imediata de todo o gás onde, a cada bufada, é possível haver a mudança de direção. De forma prática, é possível assimilar o modelo gaussiano a contínua emissão de gases através de uma chaminé, enquanto o de bufadas caracteriza o rompimento de um vaso de pressão onde gás seria imediatamente expelido. A Figura 3.3 apresenta como a diferença entre os métodos gaussiano e bufadas facilitam a mudança de direção de uma pluma. Portanto, é possível notar que o modelo de bufadas é a modelagem matemática mais indicada para problemas onde a mudança de direção do vento impacta os resultados do problema, enquanto o modelo gaussiano é mais indicado para cenários onde mudança de direção não é decisivo.

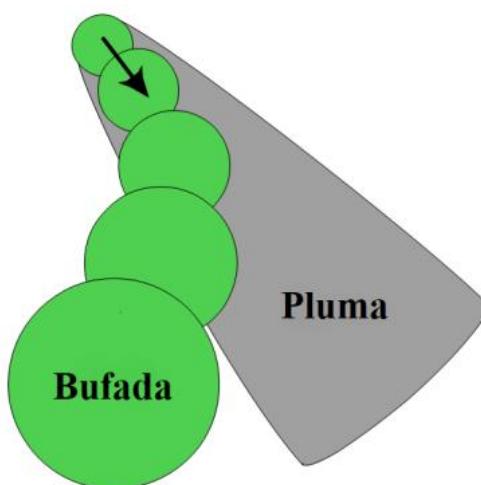


Figura 3.3 – A mudança de direção gerada para diferentes tipos de modelagem matemática.
Fonte: Lagzi et. al (2013).

3.2. SIMULADOR DE DISPERSÃO ATMOSFÉRICA

O Simulador de Dispersão Atmosférica (SDA), utilizado neste trabalho, baseou-se no simulador do Sistema de Controle Ambiental (SCA) (PEN/COPPE/UFRJ, 1987) que foi desenvolvido para fazer o acompanhamento atmosférico ao redor das usinas nucleares de Angra I e Angra II da CNAAA. O SDA implementa o modelo matemático de cálculo de dispersão baseado em bufadas com trajetória lagrangiana, como visto anteriormente.

Os cálculos de dispersão são feitos em uma zona de controle limitada ao redor da CNAAA que possui 11 km de extensão vertical e 17 km de extensão horizontal conforme mostra a Figura 3.4. A zona de monitoramento controlada é dividida em uma malha tridimensional não-divergente refinada com 215 divisões na direção vertical e 335 divisões na direção horizontal. Além disso, por ser específico para a zona de monitoramento da CNAAA, o simulador SCA, e consequentemente, o SDA utilizado neste trabalho, já possui todo o mapeamento topográfico implementado internamente.

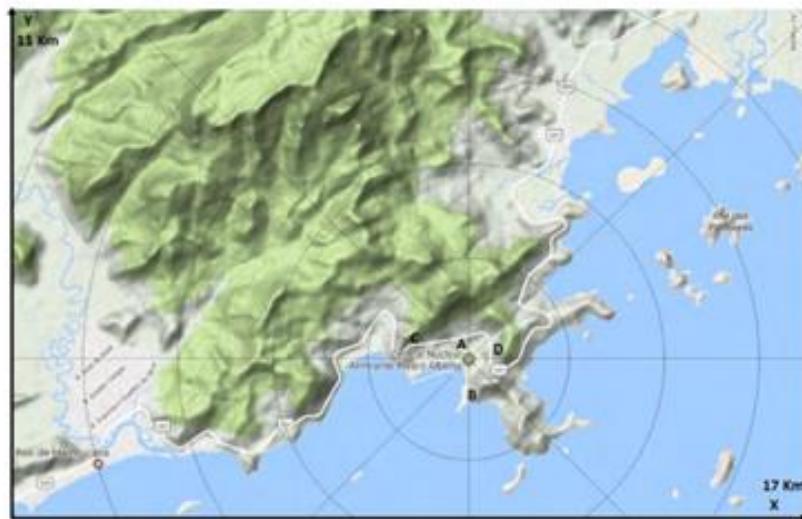


Figura 3.4 – Mapa da zona de monitoramento da CNAAA.
Fonte: o autor (2023).

O simulador opera através de ciclos com intervalos de 15 minutos, onde cada ciclo se refere ao progresso do acidente selecionado e é associado a uma respectiva condição meteorológica. Portanto, para realizar uma simulação no SDA é necessário introduzir os arquivos com as informações meteorológicas e selecionar qual será o acidente que será simulado ao longo dos ciclos.

O SDA compreende 3 módulos básicos que subdividem os cálculos para a estimativa da dispersão. São eles: módulo do cálculo do termo fonte, módulo do cálculo do campo de vento e módulo do cálculo da dispersão de doses. O fluxograma da Figura 3.5 resume as interdependências entre os 3 módulos do simulador.

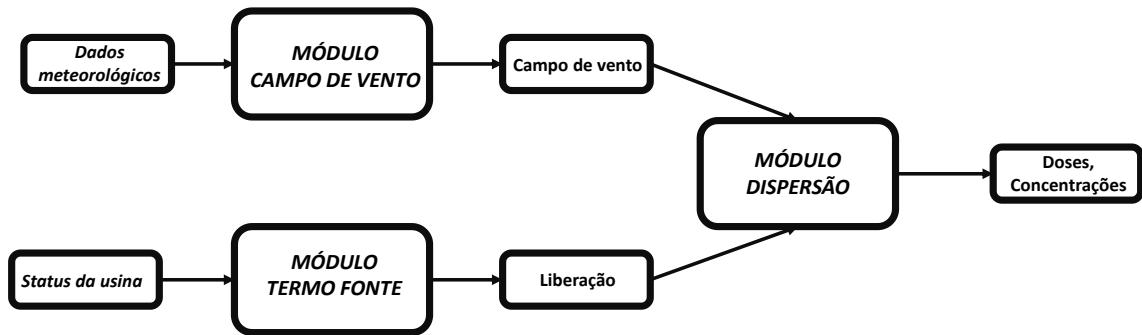


Figura 3.5 – Representação esquemática das entradas e saídas dos diferentes módulos do SDA.
Fonte: o autor (2023).

3.2.1. MÓDULO TERMO FONTE

O módulo termo fonte tem como objetivo calcular a quantidade e o tipo de material radioativo que é disponibilizado para liberação para o meio ambiente. Para isso, é importante identificar qual o tipo de acidente está ocorrendo na usina, pois a quantidade e o tipo de material radioativo que é liberado possuem características diretas ao tipo de acidente em andamento. Além disso, a liberação depende do status dos diversos componentes eletromecânicos da usina, uma vez que o projeto de uma usina nuclear prevê diversas barreiras em série para evitar que o material radioativo chegue à atmosfera.

3.2.2. MÓDULO CAMPO DE VENTO

O módulo campo de vento tem como função reunir todas as informações meteorológicas da usina. Na CNAAA os registros meteorológicos são feitos a partir de 4 torres (nomeadas torre A, B, C, D) localizadas em diferentes pontos da central nuclear. A torre A possui 3 pontos de

medição em diferentes cotas enquanto as demais torres possuem apenas um único ponto de medição. Os parâmetros medidos em cada torre estão resumidos na Tabela 3.1.

Tabela 3.1 – Parâmetros meteorológicos medidos em cada torre e ponto de medição

Torre	Pontos de Medição	Parâmetros Meteorológicos
A	3	Velocidade do vento, direção do vento, estabilidade do vento e temperatura
B	1	Velocidade do vento, direção do vento e estabilidade
C	1	Velocidade do vento, direção do vento e estabilidade
D	1	Velocidade do vento, direção do vento e estabilidade

Fonte: o autor (2023).

A zona de monitoramento é discretizada em uma malha que aplica um modelo tridimensional, não-divergente que considera a topografia complexa do local. Portanto, o objetivo do módulo campo de vento é calcular a distribuição espacial discretizada do vento dentro dessa zona de monitoramento.

3.2.3. MÓDULO DISPERSÃO

Como pode ser visto na Figura 3.5, o módulo termo fonte e campo de vento são módulos independentes, onde o primeiro depende especificamente das condições de funcionamento da usina nuclear e o segundo depende das condições meteorológicas e topográficas locais. Em uma simulação feita no SDA, esses dois módulos não dependem de informações entre si para realizar a sua tarefa. Porém, o módulo dispersão depende diretamente do campo gerado pelo módulo do campo de vento e a liberação calculada pelo módulo termo fonte. Uma vez que os dois módulos são finalizados, então o simulador é capaz de iniciar o processo de estimativa de dispersão da pluma pelo módulo dispersão.

O módulo de dispersão é a sessão do simulador que calcula a movimentação da pluma de dispersão ao longo de cada ciclo conforme o acidente progride e de acordo com as condições meteorológicas e topográficas locais calculados nos dois módulos anteriores. Agora, as equações referentes ao modelo matemático de bufadas com trajetória lagrangiana é feito para todos os elementos da malha discretizada.

O SDA calcula e arquiva o histórico da concentração média de cada radioisótopo liberado durante a simulação. Ao final dos cálculos do módulo de dispersão, arquivos de saída são criados com informações das distribuições espaciais de taxas de dose de corpo inteiro, pulmão e tireoide para todos os ciclos. As sequências da Figura 3.6 abaixo apresentam 4 ciclos resultantes de uma simulação feita no SDA que considera um acidente severo hipotético iniciado em 3 de novembro de 2016 às 8 horas e 30 minutos na usina de Angra II.

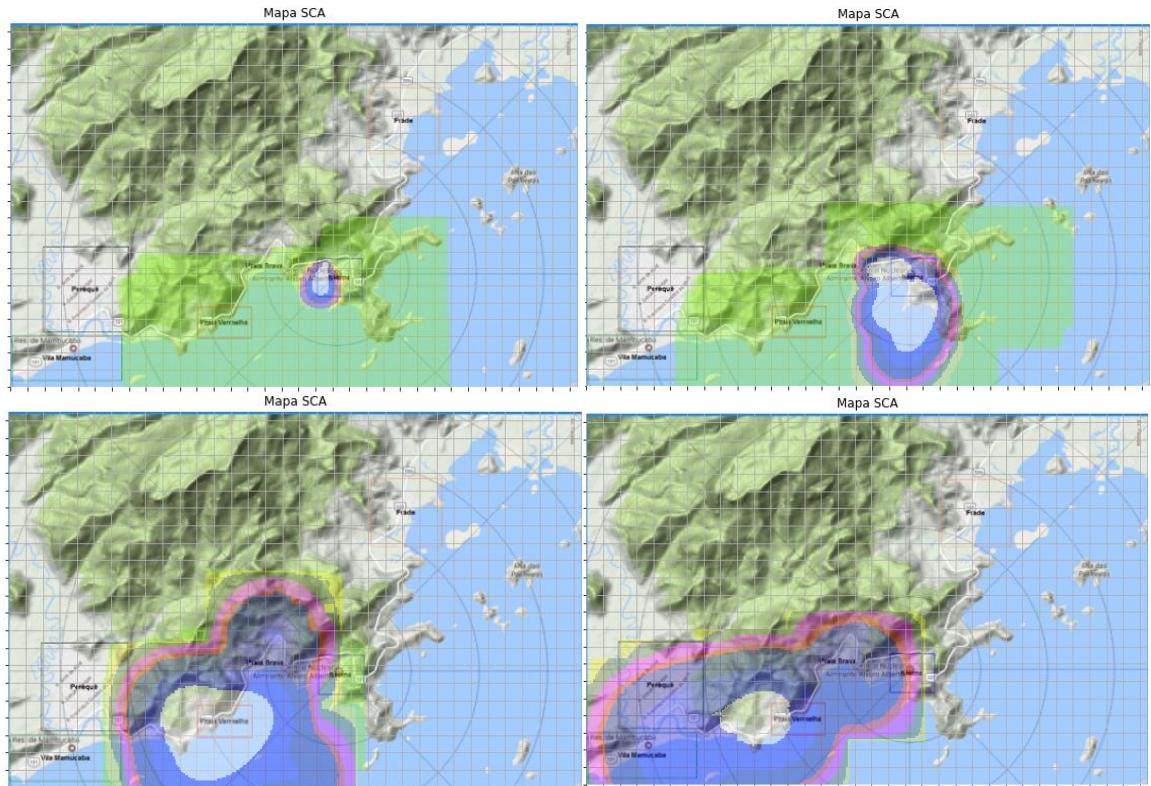


Figura 3.6 – Distribuição espacial das taxas de dose de corpo inteiro: (a) no ciclo 8, (b) no ciclo 11, (c) no ciclo 16, (d) no ciclo 18.

Fonte: o autor (2023).

4. DESENVOLVIMENTO DO MODELO

4.1. GERAÇÃO DOS DADOS

O objetivo deste trabalho é criar um modelo de rede neural artificial recorrente, mais especificamente, do tipo LSTM, que seja capaz de prever o movimento do ponto de máxima taxa de dose de corpo inteiro presente na pluma radioativa em caso da ocorrência de um acidente severo durante a operação de um reator nuclear. Para tal, primeiramente, é necessário reunir informações meteorológicas que são utilizadas como dados de entrada para o simulador de dispersão atmosférica. Em seguida, a simulação é feita através de um Simulador de Dispersão Atmosférica (SDA) baseado no SCA, programa utilizado no monitoramento das usinas nucleares Angra I e Angra II para estimar as respectivas plumas de dispersão. Ao fim de cada ciclo de 15 minutos, a pluma é atualizada e é extraído o ponto de máxima dose de corpo inteiro para cada pluma gerada e então é formado o grupo de amostras que será utilizado no treinamento da rede neural.

Os registros meteorológicos que são utilizados na simulação, são registros reais arquivados no histórico meteorológico da Central Nuclear Almirante Álvaro Alberto (CNAAA) datados de novembro de 2016. Os arquivos contêm informações de 4 torres meteorológicas contendo informações de velocidade, direção e estabilidade do vento atualizadas, entre outras, a cada quarto de hora, seguindo os critérios de medição da Tabela 3.1.

A Tabela 4.1 apresenta as informações dos arquivos meteorológicos já compiladas do histórico da usina de Angra II. Note que *da1*, *va1*, *ta1* e *sa1* referem-se as medições de direção do vento, velocidade do vento, temperatura e estabilidade do vento do ponto de medição 1 da torre A respectivamente; *da2*, *va2*, *ta2* e *sa2* referem-se as medições do ponto de medição 2 da torre A; *da3*, *va3*, *ta3* e *sa3* referem-se as medições do ponto de medição 3 da torre A; *db1*, *vb1*, *tb1* e *sb1* referem-se ao ponto de medição 1 da torre B; *dc1*, *vc1*, *tc1* e *sc1* referem-se ao ponto de medição 1 da torre C; *dd1*, *vd1*, *td1* e *sd1* referem-se ao ponto de medição 1 da torre D;

Tabela 4.1 – Parâmetros meteorológicos referente aos primeiros 18 ciclos a serem simulados.

Data e hora	da1	va1	ta1	sa1	da2	va2	ta2	sa2	da3	va3	ta3	sa3	dh1	vb1	sb1	dc1	vc1	sc1	dd1	vd1	sd1
03/11/2016 02:45:00	16.051	1.679	22.133	5.5992	0	0	22.027	0	0	0	20.503	0	0	0	0	44.716	0.73629	25.608	92.343	0.8526	22.352
03/11/2016 03:30:00	21.259	1.6777	22.001	6.6106	0	0	21.968	0	0	0	20.326	0	0	0	0	141.5	0.52291	10.134	85.148	0.9335	21.008
03/11/2016 03:15:00	18.196	1.6806	21.966	6.1634	0	0	21.873	0	0.22531	0.22531	20.328	3.4242	0	0	0	128.8	0.47531	12.02	63.044	0.7772	23.376
03/11/2016 03:30:30	18.706	1.5202	21.868	11.128	260.04	0.48115	21.845	9.0484	0.67772	0.67772	20.259	9.0473	0	0	0	103.9	0.54792	6.5752	78.681	0.6312	26.443
03/11/2016 03:45:00	22.37	1.7447	22.181	6.008	36.157	0.17671	21.892	14.6	0.42004	0.42004	20.328	12.401	0	0	0	112.57	1.3549	12.019	67.248	0.8109	32.695
03/11/2016 04:00:00	7.7789	1.1737	21.713	22.616	260.8	0.10236	21.718	3.6173	0.16873	0.16873	20.358	44.446	0	0	0	157.78	0.9325	22.633	33.938	1.4026	20.298
03/11/2016 04:15:00	18.483	1.4476	21.741	8.2873	270.67	0.13727	21.68	3.4102	0.36164	0.36164	20.31	9.8136	0	0	0	125.76	0.66736	10.726	31.522	1.3811	18.506
03/11/2016 04:30:00	17.105	1.3903	21.771	8.6941	0	0	21.612	0	0.17149	0.17149	20.261	7.2634	0	0	0	64.931	0.73987	49.374	38.939	1.9722	18.979
03/11/2016 04:45:00	14.048	1.0248	21.872	16.968	225.34	0.15745	21.473	5.4175	0.45934	0.45934	20.143	43.818	0	0	0	68.485	1.1738	49.518	52.377	2.1151	21.711
03/11/2016 05:00:00	17.019	0.95737	21.562	32.132	240.16	0.93455	21.291	19.862	0.91275	0.91275	20.299	66.034	0	0	0	125.12	0.94032	20.712	47.903	2.3509	25.072
03/11/2016 05:15:00	13.997	1.3076	21.845	11.333	238.55	0.54872	21.368	7.4935	0.89629	0.89629	20.532	21.072	0	0	0	129.63	0.81306	62.471	36.791	1.9516	19.563
03/11/2016 05:30:00	12.094	1.138	21.651	15.596	247.56	0.26218	21.39	10.541	0.90794	0.90794	20.371	20.981	0	0	0	77.502	0.36189	40.54	39.607	2.1596	17.849
03/11/2016 05:45:00	17.409	1.2412	21.503	11.891	237.13	0.37701	21.302	4.8748	0.89539	0.89539	20.234	22.516	0	0	0	280.3	0.81417	27.519	34.096	1.7831	16.53
03/11/2016 06:00:00	14.956	1.0661	21.821	10.585	240.14	0.14896	21.291	4.6783	0.38022	0.38022	20.047	20.799	0	0	0	239.7	0.7955	23.444	38.382	2.2524	20.019
03/11/2016 06:15:00	15.866	1.2325	21.712	8.4696	279.49	0.05948	21.337	1.1833	0.38538	0.38538	19.999	5.4996	0	0	0	234.53	1.0271	17.385	49.204	2.4277	20.078
03/11/2016 06:30:00	17.228	1.3318	21.777	10.188	257.71	0.27295	21.446	1.7265	1.1661	1.1661	19.972	5.7149	0	0	0	232.15	1.0562	26.307	42.661	2.481	18.21
03/11/2016 06:45:00	14.64	1.251	22.082	13.331	277.54	0.08669	21.633	1.1142	1.3574	1.3574	19.987	78.042	0	0	0	317.51	1.1585	41.009	50.197	2.8099	19.947
03/11/2016 07:00:00	340.74	0.87459	22.224	44.78	23.293	1.0002	22.211	60.508	2.8794	2.8794	21.196	22.606	0	0	0	301.08	1.225	15.422	49.486	3.0183	23.24

Fonte: o autor (2023).

Para gerar o conjunto de dados utilizados para o treinamento, teste e validação da rede neural artificial, é utilizado o simulador SDA.

Como resultado da simulação, é gerado um arquivo de saída referente a cada ciclo com o valor da taxa de dose de corpo inteiro para todas as coordenadas da zona de interesse da CNAAA. Com isso, é possível obter um mapa de distribuição espacial de taxas de dose de corpo inteiro geradas pelo simulador.

Nesse trabalho foram realizadas 34 simulações de dispersão da pluma radioativa considerando um acidente severo hipotético na usina de Angra II. Cada simulação é feita por um período de 18 ciclos de 15 minutos cada (aproximadamente 4,5 horas de simulação), totalizando 612 simulações de plumas de dispersão radioativa. Foi criado um programa que extrai as coordenadas do ponto que possui a máxima taxa de dose de corpo inteiro de cada amostra.

As Figuras 4.1.1, 4.1.2, 4.1.3. 4.1.4 apresentam os 18 ciclos simulados quando um acidente severo é iniciado na usina de Angra II em 14 de novembro de 2016 às 19 horas e 30 minutos. A figura exibe uma sequência de mapas plotados em forma de *heatmap*, onde cada cor refere-se à uma faixa de taxa de dose e é destacado o ponto de maior taxa de dose de corpo inteiro.

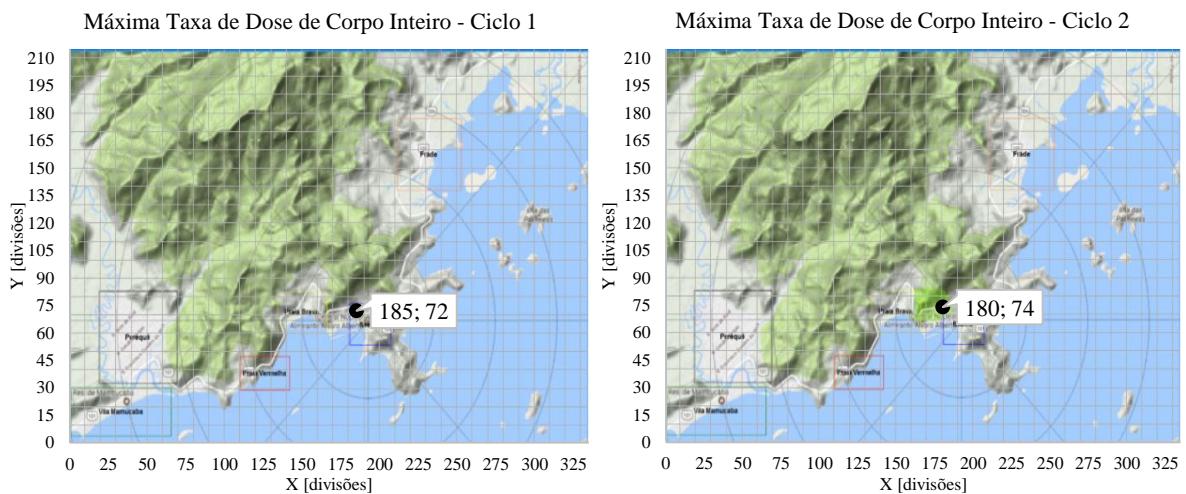


Figura 4.1.1 – Ponto de maior taxa de dose de corpo inteiro presente na pluma de dispersão radioativa: (a) no ciclo 1, (b) no ciclo 2.
Fonte: o autor (2023).

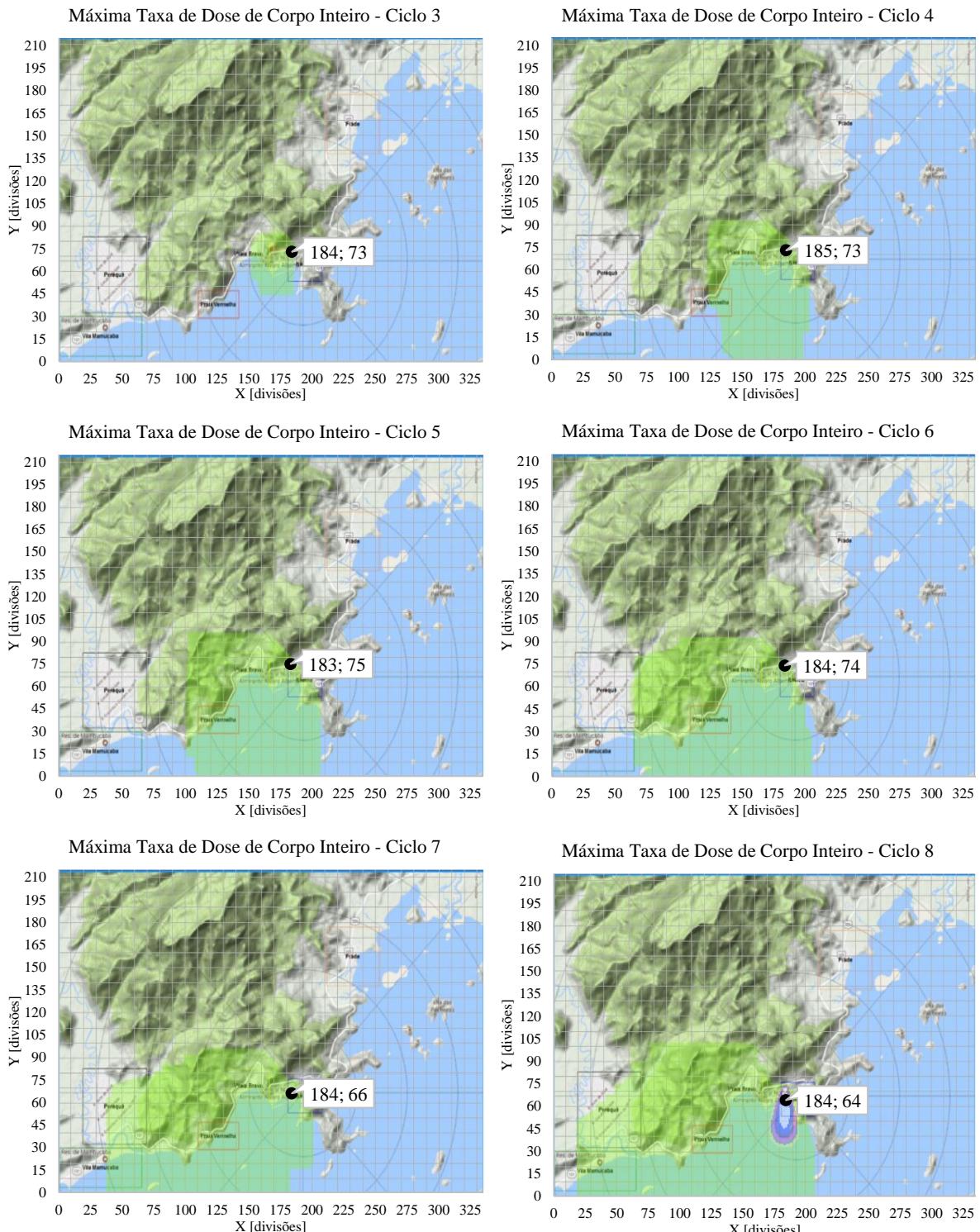


Figura 4.1.2 – Ponto de maior taxa de dose de corpo inteiro presente na pluma de dispersão radioativa: (a) no ciclo 3, (b) no ciclo 4, (c) no ciclo 5, (d) no ciclo 6, (e) no ciclo 7, (f) no ciclo 8.

Fonte: o autor (2023).

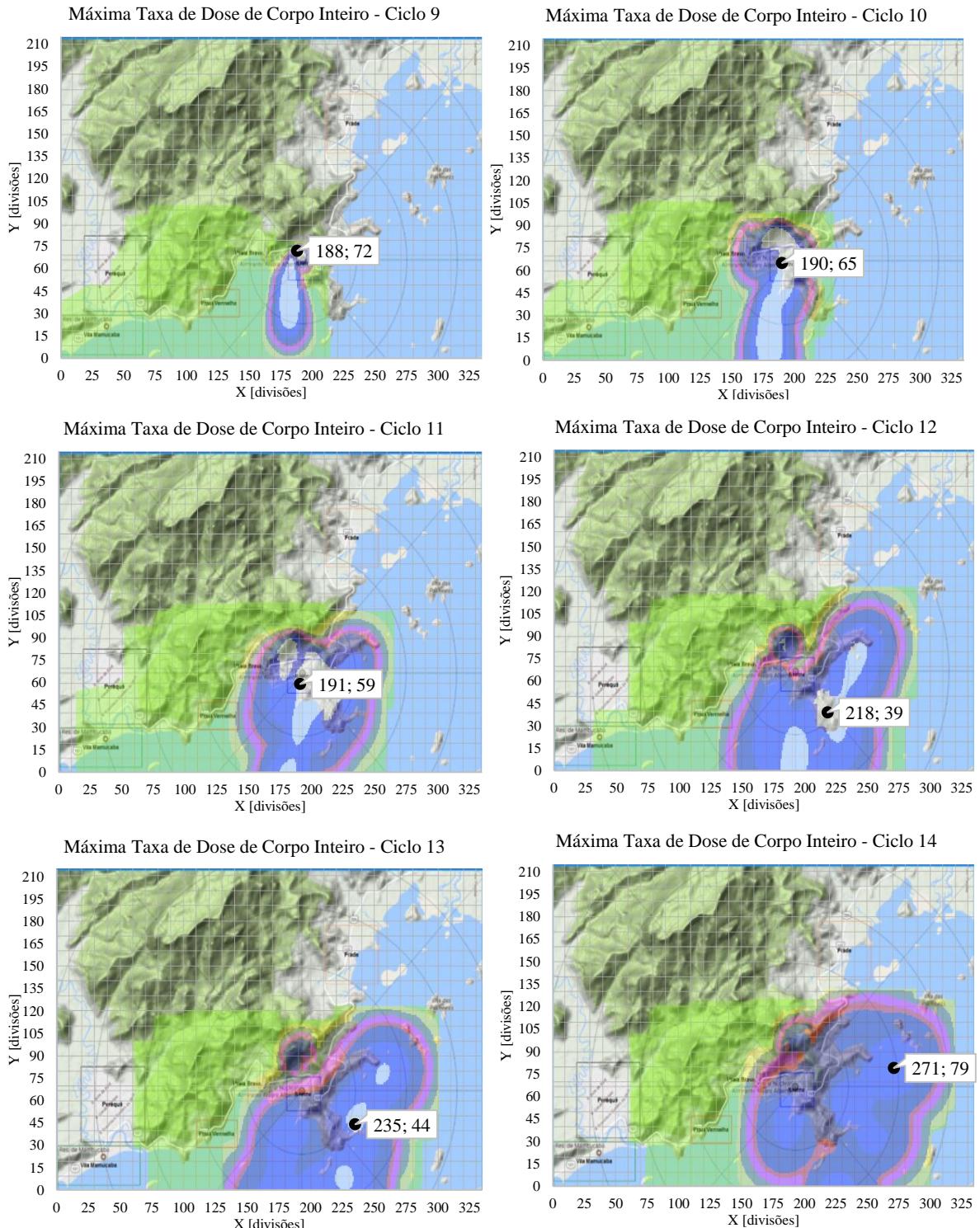


Figura 4.1.3 – Ponto de maior taxa de dose de corpo inteiro presente na pluma de dispersão radioativa: (a) no ciclo 9, (b) no ciclo 10, (c) no ciclo 11, (d) no ciclo 12, (e) no ciclo 13, (f) no ciclo 14.

Fonte: o autor (2023).

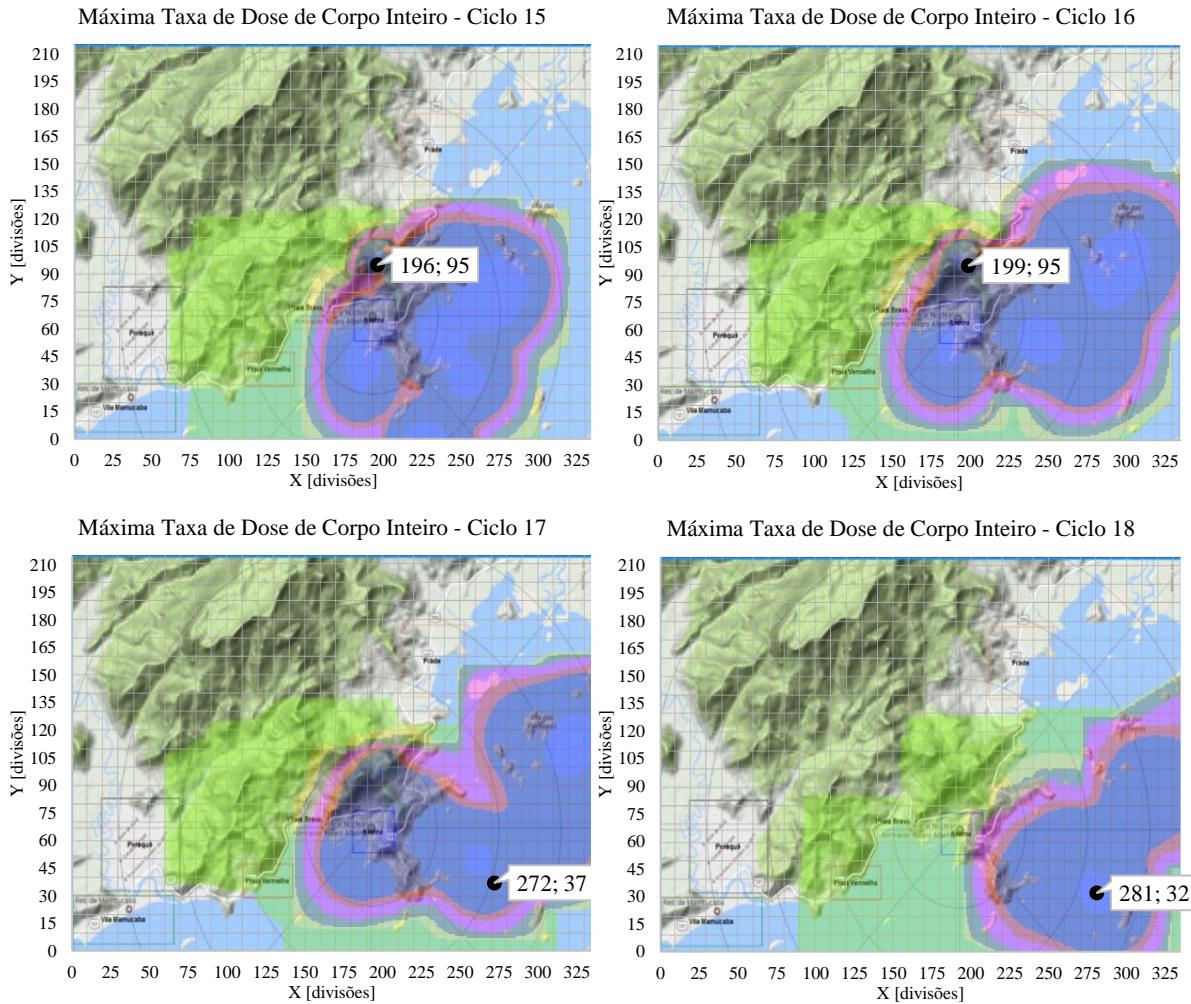


Figura 4.1.4 – Ponto de maior taxa de dose de corpo inteiro presente na pluma de dispersão radioativa: (a) no ciclo 15, (b) no ciclo 16, (c) no ciclo 17, (d) no ciclo 18.
Fonte: o autor (2023).

De posse dos dados, foram selecionados 4 parâmetros de entrada para a rede neural a ser treinada, são eles:

- abscissa do ponto de maior dose de corpo inteiro (Xm),
- ordenada do ponto de maior dose de corpo inteiro (Ym),
- velocidade do vento do ponto de medição 1 da torre A ($va1$) e
- direção do vento do ponto de medição 1 da torre A ($da1$).

Portanto, para gerar o arquivo de entrada para o treinamento da rede neural, é necessário extrair a coordenada espacial que possui a maior taxa de dose de corpo inteiro para cada ciclo de cada simulação realizada e reuni-las as informações de velocidade e direção do vento do ponto de medição 1 da torre A que estão presentes nos arquivos dos registros meteorológicos.

A Tabela 4.2 apresenta algumas combinações dos referidos parâmetros que foram selecionados para serem dados de entrada do treinamento da rede neural artificial recorrente.

Tabela 4.2 – Parâmetros utilizados para o treinamento da rede neural artificial recorrente.

Xm	Ym	va1	da1
1.86E+02	7.20E+01	1.69E+00	2.39E+01
1.88E+02	7.10E+01	1.63E+00	1.61E+01
1.87E+02	7.10E+01	1.52E+00	1.32E+01
1.83E+02	6.30E+01	1.94E+00	1.83E+01
1.86E+02	7.20E+01	2.02E+00	1.85E+01
1.83E+02	7.20E+01	1.92E+00	1.59E+01
1.84E+02	7.10E+01	1.53E+00	1.91E+01
1.84E+02	7.20E+01	1.38E+00	2.25E+01
1.82E+02	7.20E+01	1.69E+00	1.66E+01
1.79E+02	7.00E+01	1.68E+00	1.61E+01
1.76E+02	7.10E+01	1.68E+00	2.13E+01
1.69E+02	7.10E+01	1.68E+00	1.82E+01
1.65E+02	6.50E+01	1.52E+00	1.87E+01
1.55E+02	7.00E+01	1.74E+00	2.24E+01
1.58E+02	4.00E+01	1.17E+00	7.78E+00
1.50E+02	4.40E+01	1.45E+00	1.85E+01
1.40E+02	3.40E+01	1.39E+00	1.71E+01
1.29E+02	4.50E+01	1.02E+00	1.40E+01
1.89E+02	7.40E+01	1.33E+00	1.72E+01
1.89E+02	7.40E+01	1.25E+00	1.46E+01
1.85E+02	7.30E+01	8.75E-01	3.41E+02
1.86E+02	7.50E+01	7.75E-01	3.50E+02
1.84E+02	7.50E+01	5.55E-01	4.76E+01
1.81E+02	7.50E+01	6.08E-03	9.92E+01
1.87E+02	7.20E+01	3.39E-01	1.98E+02
1.86E+02	7.20E+01	4.81E-01	1.76E+02
1.86E+02	7.20E+01	3.28E-01	1.93E+02
1.64E+02	7.50E+01	7.78E-01	2.06E+02
1.84E+02	6.80E+01	6.93E-01	2.04E+02
1.84E+02	6.80E+01	8.29E-01	2.09E+02

Fonte: o autor (2023).

Todos os dados de treinamento, validação e teste foram normalizados, utilizando o método Min-Max, em um intervalo de [0,1]. A Tabela 4.3 apresenta as mesmas combinações dos parâmetros ilustrados na Tabela 4.2, agora normalizados de acordo com o método Min-Max.

Tabela 4.3 – Parâmetros normalizados utilizados para o treinamento da rede neural artificial recorrente.

Xm	Ym	va1	da1
0.528662	0.336449	0.562216	0.06655
0.535032	0.331776	0.541095	0.044649
0.531847	0.331776	0.503636	0.036718
0.519108	0.294393	0.642812	0.050941
0.528662	0.336449	0.672301	0.051524
0.519108	0.336449	0.638927	0.044057
0.522293	0.331776	0.509149	0.05315
0.522293	0.336449	0.458506	0.062512
0.515924	0.336449	0.562282	0.046119
0.506369	0.327103	0.557567	0.044605
0.496815	0.331776	0.557135	0.059077
0.474522	0.331776	0.558098	0.050566
0.461783	0.303738	0.504832	0.051983
0.429936	0.327103	0.579384	0.062165
0.43949	0.186916	0.389765	0.021617
0.414013	0.205607	0.480723	0.051363
0.382166	0.158879	0.461694	0.047534
0.347134	0.21028	0.340318	0.039038
0.538217	0.345794	0.442267	0.047876
0.538217	0.345794	0.415435	0.040684
0.525478	0.341122	0.290436	0.946895
0.528662	0.350467	0.257381	0.971766
0.522293	0.350467	0.184342	0.132313
0.512739	0.350467	0.002019	0.275559
0.531847	0.336449	0.112496	0.55123
0.528662	0.336449	0.159639	0.490399
0.528662	0.336449	0.10896	0.535306
0.458599	0.350467	0.258344	0.571516
0.522293	0.317757	0.230243	0.566153
0.522293	0.317757	0.27535	0.58127

Fonte: o autor (2023).

4.2. EXTRAÇÃO DOS CONJUNTOS DE TREINAMENTO, TESTE E VALIDAÇÃO

Ao fim da extração da coordenada de máxima dose de corpo inteiro, o conjunto de 612 amostras é separado em grupos de treinamento, teste e validação. Foram utilizados 80% das amostras para o conjunto de treinamento e os 10% para o conjunto de teste e 10% para o conjunto de validação.

4.3. ARQUITETURAS DAS REDES NEURAIS DESENVOLVIDAS

Diversas arquiteturas de rede neural artificial recorrente do tipo LSTM foram treinadas com objetivo de prever a posição do ponto de maior taxa de dose com antecedência de 15 minutos, ou seja, 1 ciclo de simulação. Foram variados: o número de épocas, a quantidade de observações passadas (*timesteps*) através do método de janelas deslizantes e o tamanho do lote (*batch*).

Todas as estruturas de rede testadas possuem uma camada densa de saída de dois neurônios com função de ativação linear. Com isso, diferentes combinações de camadas internas escondidas foram treinadas considerando camadas LSTM e densas, com funções de ativação ReLU, para variadas quantidades de neurônios. A Tabela 4.4 ilustra como os dados ficaram organizados.

Tabela 4.4 – Organização do conjunto de dados utilizados no treinamento, validação e teste da rede neural.

Rede	Épocas	Step	Batches	Camada LSTM 1		Camada LSTM 2		Camada Densa	
				Neurônios	Função de Ativação	Neurônios	Função de Ativação	Neurônios	Função de Ativação
1	300	2	1	10	relu	x	x	x	x
2	300	2	1	4	relu	x	x	x	x
3	300	2	1	4	relu	4	relu	x	x
4	300	2	1	10	relu	10	relu	x	x
5	600	2	1	10	relu	10	relu	x	x
6	300	2	1	4	relu	4	relu	4	relu
7	300	2	1	4	relu	x	x	4	relu
8	300	2	1	10	relu	x	x	10	relu
9	600	2	1	10	relu	x	x	10	relu
10	300	2	3	4	relu	x	x	x	x
11	600	2	3	4	relu	x	x	x	x
12	300	2	3	10	relu	x	x	x	x
13	600	2	3	10	relu	x	x	x	x
14	600	2	3	5	relu	5	relu	x	x
15	600	2	3	10	relu	10	relu	x	x
16	600	2	3	10	tanh	x	x	x	x
17	600	3	3	10	relu	x	x	x	x
18	600	5	3	10	relu	x	x	x	x

Fonte: o autor (2023).

5. RESULTADOS

A função de erro (*loss function*) utilizada durante o treinamento foi o erro quadrado médio (MSE). A Tabela 5.1 apresenta um resumo das redes testadas para as diversas combinações de estruturas de camadas escondidas e seus respectivos resultados de erro quadrático médio para cada grupo de treinamento, teste e validação.

Tabela 5.1 – Resultados obtidos para treinamento, teste e validação todas as redes testadas

Rede	Resultados		
	Treinamento	Validação	Teste
1	0.0024	0.0015	0.0021
2	0.0026	0.0015	0.0020
3	0.0024	0.0017	0.0020
4	0.0023	0.0016	0.0026
5	0.0021	0.0017	0.0044
6	0.0115	0.0080	0.0170
7	0.0030	0.0012	0.0032
8	0.0023	0.0014	0.0025
9	0.0022	0.0015	0.0035
10	0.0026	0.0015	0.0019
11	0.0026	0.0014	0.0021
12	0.0025	0.0015	0.0020
13	0.0025	0.0014	0.0019
14	0.0024	0.0015	0.0025
15	0.0025	0.0015	0.0019
16	0.0026	0.0015	0.0019
17	0.0024	0.0012	0.0022
18	0.0020	0.0037	0.0022

Fonte: o autor (2023).

De acordo com os resultados dos erros quadráticos médios da Tabela 5.1, as redes neurais, em geral, mantiveram resultados consistentes (variação relativamente pequena) para a maioria das estruturas de redes testadas. É possível verificar que o melhor resultado, considerando o conjunto de teste, é encontrado com as redes 10, 13, 15 e 16, todas elas sem camada densa na

saída e com *timesteps* igual a 2. Os resultados mais discrepantes encontrados são aqueles quando uma camada escondida densa é combinada com uma camada LSTM, onde o erro quadrático médio foi relativamente superior as demais redes testadas. Visto que os resultados para camadas densas não foram razoáveis, nenhuma estrutura de rede com camada densa foi levada em consideração no processo de seleção.

Foi escolhida a estrutura de rede número 10 que possui *timesteps* de tamanho 2 e uma camada LSTM com 4 neurônios. Essa rede foi treinada com 300 épocas, batches com 3 padrões de treinamento e função de ativação ReLU. É possível verificar que existem redes que encontram resultados semelhantes ao da rede escolhida, como por exemplo as redes número 13, 15 e 16. É possível inferir que, se os resultados da rede neural número 10 são boas estimativas, então pode-se concluir que os resultados também serão bons para a rede de número 13.

Uma vez que a rede selecionada é a rede número 10, a Figura 5.1 apresenta os resultados de R^2 calculados quando são comparados os resultados entre a simulação e a rede neural para abscissas e ordenadas do ponto de máxima dose de corpo inteiro nos grupos de treinamento, teste e validação.

Os resultados das coordenadas do ponto de máxima dose de corpo inteiro simulados são comparados com os resultados calculados pela rede neural artificial escolhida. A Figura 5.2 apresenta de forma visual a distribuição espacial do conjunto de dados simulados e o conjunto de dados calculados pela rede neural para os grupos de treinamento, teste e validação respectivamente.

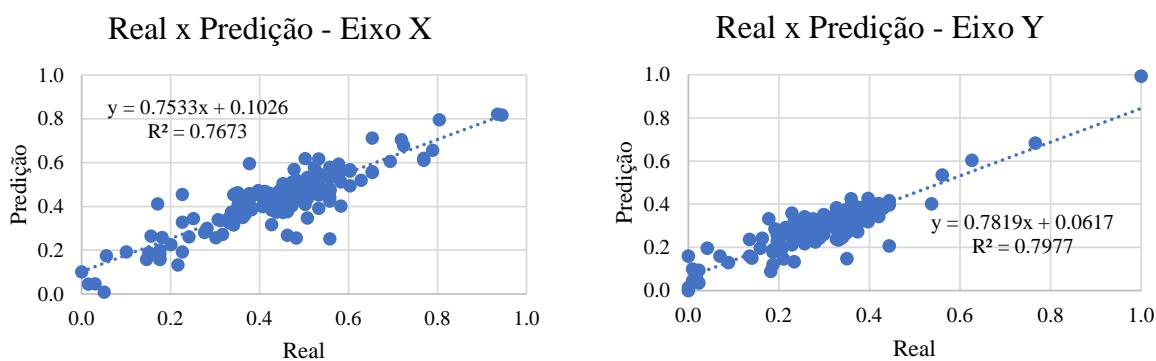


Figura 5.1.1 – Gráfico de R^2 para: (a) abscissas do grupo de treinamento, (b) ordenadas do grupo de treinamento.

Fonte: o autor (2023).

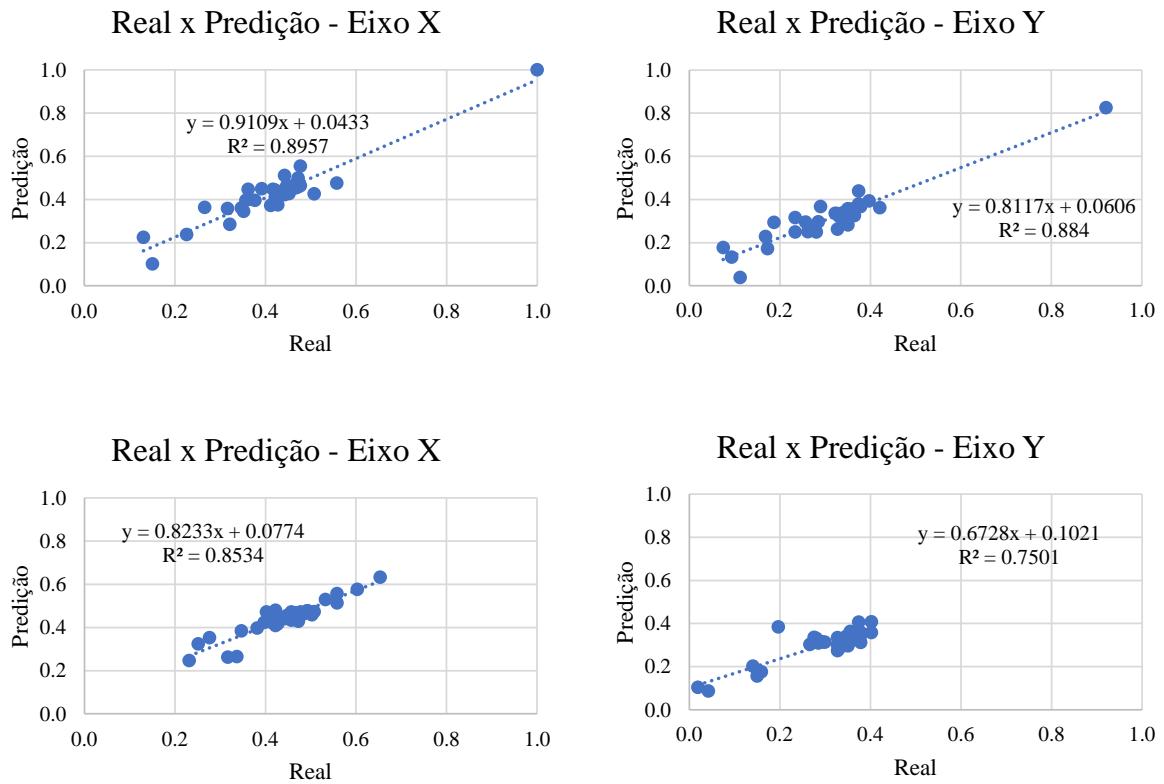
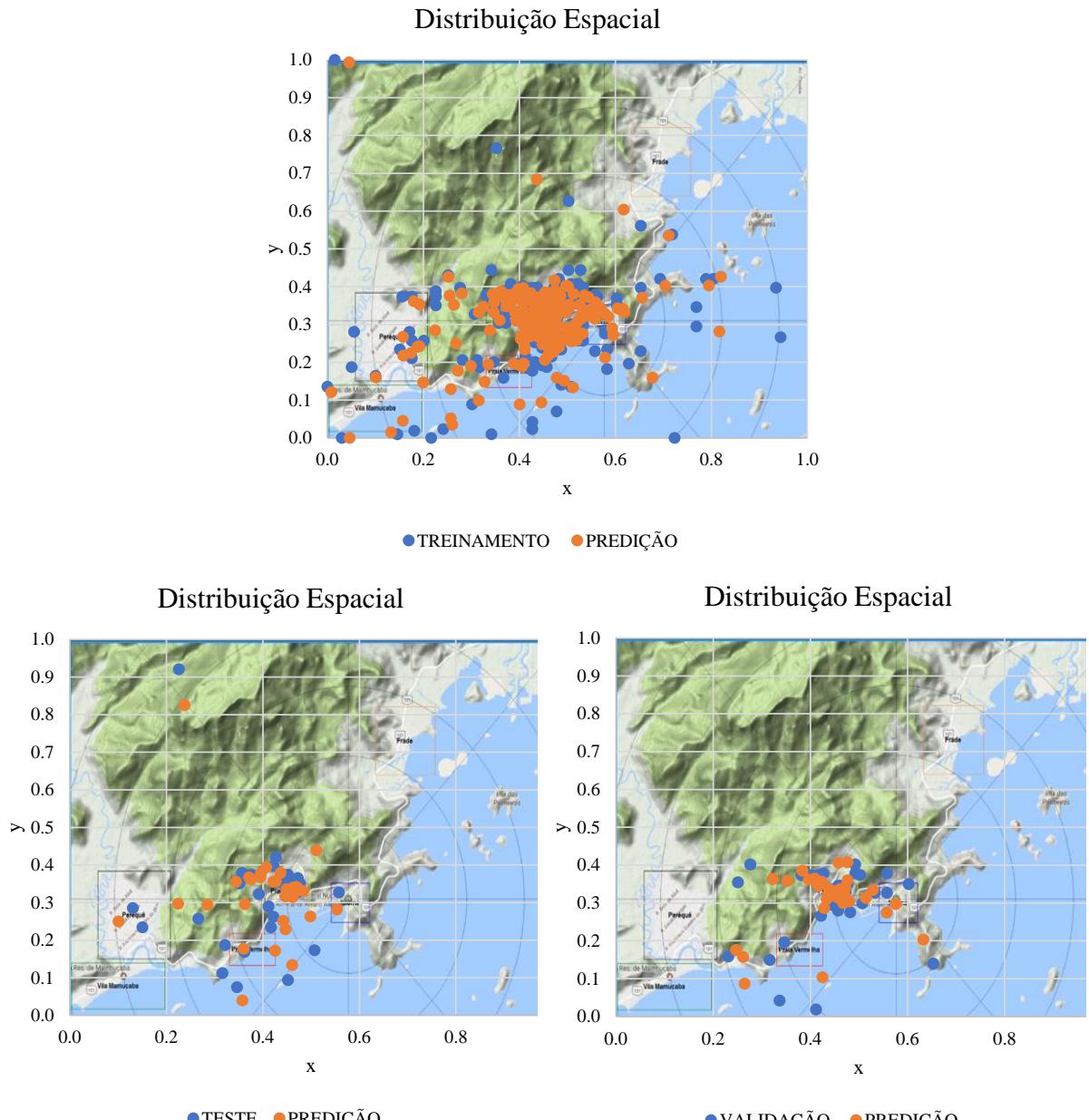


Figura 5.1.2 – Gráfico de R² para: (a) abscissas do grupo de teste, (b) ordenadas do grupo de teste, (c) abscissas do grupo de validação, (d) ordenadas do grupo de validação.

Fonte: o autor (2023).

Os resultados calculados possuem uma boa aproximação com resultados da simulação feita pelo SDA. É possível verificar que a maior densidade de pontos se encontra na região da localização do reator nuclear, o que é esperado, uma vez que esse é o ponto onde os radionuclídeos liberados pelo acidente se concentram. Porém, dependendo das condições climáticas e conforme o acidente começa a cessar, a pluma de dispersão de radioatividade tem a tendência de se movimentar na direção em que as condições meteorológicas indicam, movimentando também o ponto de máxima dose de corpo inteiro. Portanto, é importante que a rede neural aprenda a identificar quando esse ponto deve iniciar o seu deslocamento e qual direção esse ponto deve seguir.



A fim de verificar comportamento da movimentação do ponto de máxima dose de corpo inteiro, é necessário avaliar diversas sequências temporais da movimentação desse ponto na zona de monitoramento da central nuclear. A Figura 5.3 destaca alguns exemplos de sequências com resultados calculados por ambos o simulador e a rede neural artificial.

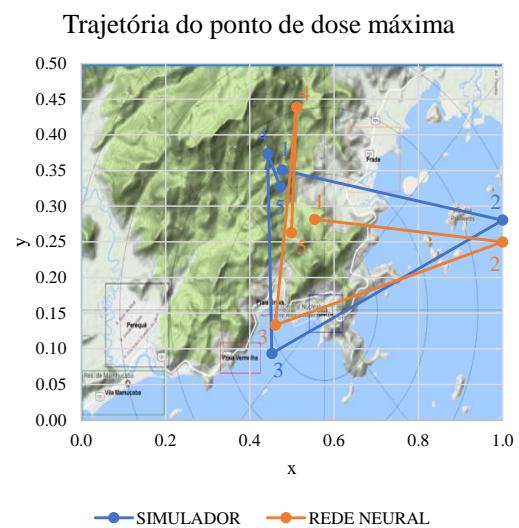
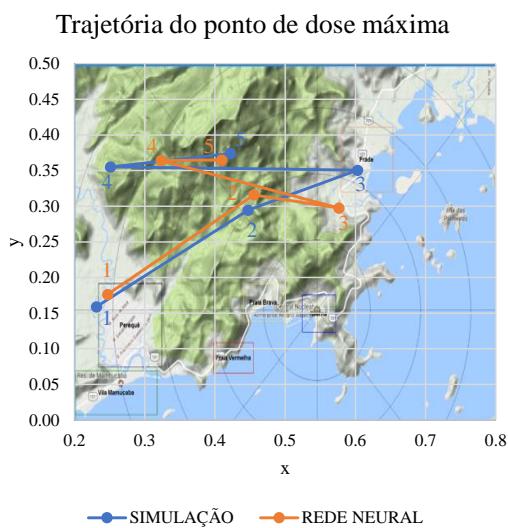
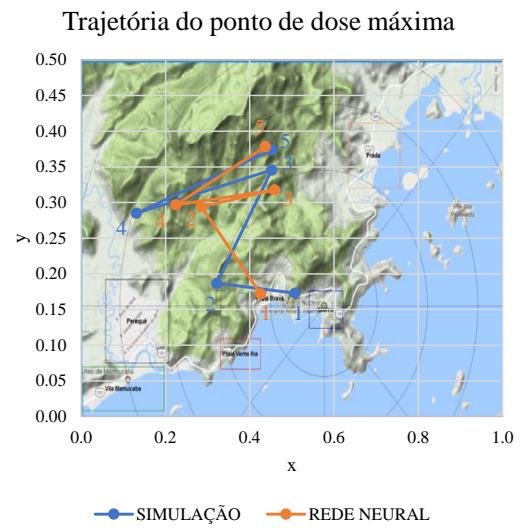
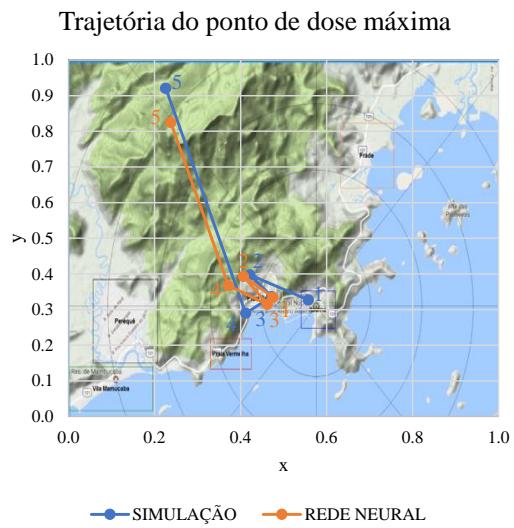
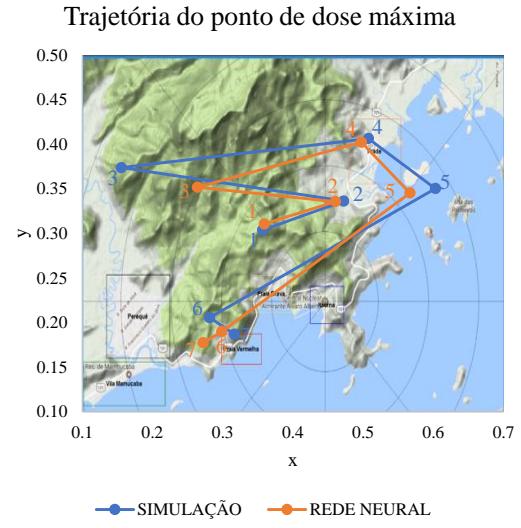
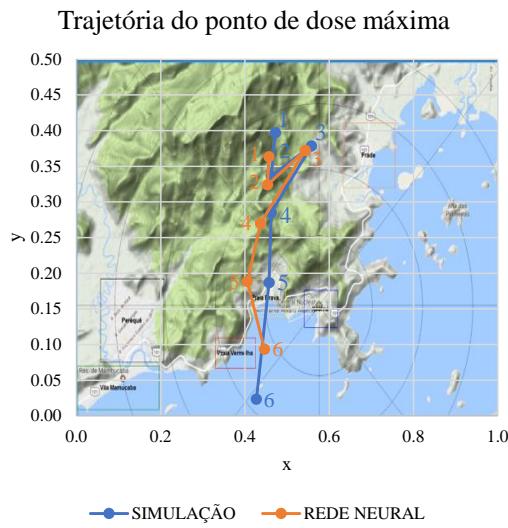


Figura 5.3 – Representação da movimentação do ponto de máxima dose de corpo inteiro para diferentes grupos de séries temporais selecionados aleatoriamente.

Fonte: o autor (2023).

É possível verificar que a rede neural artificial consegue prever de forma consistente os padrões de mudança de direção que o ponto de máxima taxa de dose de corpo inteiro registra ao longo do tempo conforme a progressão do acidente severo e as condições meteorológicas. Os valores absolutos das posições, entretanto, ora apresentam boa acurácia (se aproximam do ponto real), ora possuem uma discrepância maior. Contudo, a predição da trajetória, de uma forma global, foi considerada razoável.

6. CONCLUSÃO

Esse trabalho apresenta uma abordagem para utilização de técnicas de inteligência artificial, através de redes neurais artificiais recorrentes, tipo LSTM, para estimar o movimento da pluma de radioativa em caso de acidentes severos em usinas nucleares. A capacidade de prever o movimento da dispersão radioativa é um grande avanço nos estudos de nível de segurança em usinas nucleares e mitigação de acidente.

A rede neural é treinada por um conjunto de dados de 34 cenários de 18 ciclos considerando apenas 4 dados de entrada, sendo eles as coordenadas do ponto máximo simulado, a velocidade do vento registrado pelo ponto de medição 1 da torre A e a direção do vento registrado pelo ponto de medição 1 da torre A. O treinamento desses dados gerou um modelo de rede neural relativamente simples de apenas uma camada LSTM com 4 neurônios e função de ativação ReLU que apresentou bons resultados para o erro quadrático médio quando comparados com os resultados simulados.

É possível verificar através dos gráficos apresentados (Figura 5.3) que o modelo de rede neural encontrado possui a capacidade de prever, com aproximação razoável, a movimentação do ponto de máxima dose de corpo inteiro quando um acidente severo é iniciado no reator nuclear. Destaca-se que a rede neural conseguiu prever as mudanças de direções que esse ponto pode sofrer de acordo com as condições meteorológicas locais e a progressão do acidente severo.

Apesar da estimativa ser considerada consistente, o erro encontrado comparando os resultados da rede neural com os simulados se devem principalmente ao alto grau de não linearidade das condições meteorológicas. Esse trabalho utilizou dados meteorológicos reais dos registros oficiais da CNAAA, o que dificulta o aprendizado da rede neural durante o treinamento, porém reproduz com realidade as condições ambientais da região.

Uma sugestão para melhorar a estimativa seria necessário incluir os registros meteorológicos das demais torres presentes na CNAAA e incluir mais dados de treinamento para suprir o aumento da quantidade de variáveis de entrada.

Esse trabalho apresenta uma primeira investigação de previsão do movimento do ponto de maior taxa de dose de uma pluma radioativa. Apesar dos resultados encontrados conseguirem prever de forma consistente os resultados encontrados pelo simulador, é possível estender o nível de complexidade do problema. Isso é possível aumentando a quantidade de informações de entradas como, por exemplo, incluir a medição das demais torres, incluir mais tipos de

acidentes severos ao problema e estender a estimativa para a previsão da pluma inteira, ou zonas de níveis de taxa de dose relevantes, ao invés de apenas um único ponto.

Investigações futuras é a utilização de outras arquiteturas de redes neurais, como as redes do tipo Transformers, que têm se mostrado altamente eficientes em problemas de predição com dados sequenciais (Phuong & Hutter, 2022; Wen et al, 2022). Essa abordagem pode proporcionar melhorias significativas na previsão do comportamento da pluma radioativa, especialmente em relação à dependência temporal das informações.

Explorando essas possibilidades, é possível aprimorar a capacidade de previsão do movimento da pluma e ampliar a aplicabilidade desse estudo em cenários mais complexos e realistas envolvendo a dispersão de materiais radioativos.

7. REFERÊNCIAS

- BOZNAR, M., LESJAK, M., MLAKAR, P. **A neural network-based method for the short-term prediction of ambient SO₂ concentrations in highly polluted industrial areas of complex terrain.** Atmospheric Environment B 27 (2), 221-230, 1993.
- CAO, X., ROY, G., HURLEY, W. J., ANDREWS, W. S. **Modeling the concentration distributions of aerosol puffs using artificial neural networks.** Boundary-Layer Meteorology 136 (1), 83–103, 2010.
- CROWL, D. A., LOUVAR J. F. **Chemical process safety.** International Series in the Physical and Chemical Engineering Sciences, 2011.
- DESTERRO, F. S. M., SANTOS, M. C., GOMES, K. J., HEIMLICH, A., SCHIRRU, R., PEREIRA, C. M. N. A. **Development of a deep rectifier neural network for dose prediction in nuclear emergencies with radioactive material releases.** Progress in Nuclear Energy 118, 103-110, 2020.
- EHRHARDT, J., WEIS, A. **Development of RODOS, A comprehensive real-time on-line decision support for nuclear emergencies management in europe final report.** FI3P-CT92-0036, FZKA-5772, 1996.
- GIFFORD, F. A. **Atmospheric dispersion calculations using the generalized Gaussian plume model.** Nuclear Safety, 2(2), pp. 56–59, 1960.
- GLOROT, X.; BENGIO, Y. **Understanding the difficulty of training deep feedforward neural networks.** In: International conference on artificial intelligence and statistics, v. 9, pp. 249–256, 2010.
- GLOROT, X.; BORDES, A.; BENGIO, Y. **Deep sparse rectifier neural networks.** In: International conference on artificial intelligence and statistics, v. 15, pp. 315-323, 2011.
- GOLLER, C.; KUCHLER, A. **Learning task-dependent distributed representations by backpropagation through structure.** In: International conference on artificial intelligence and statistics, pp. 347-352, 1996.
- HOCHREITER, S. **The vanishing gradient problem during learning recurrent neural nets and problem solutions.** International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, v. 6, n. 02, pp. 107-116, 1998.
- HOCHREITER, S., SCHMIDHUBER, J., 1997. **Long short-term memory.** Neural Computation. v. 9, n. 8, pp. 1735-1780, 1997.
- HOSSAIN K. M. A. **Predictive ability of improved neural network models to simulate pollutant dispersion.** International Journal of Atmospheric Sciences 2014, 141923, 2014.
- LAGZI I., MESZAROS, R., GELYBO, G., LEELOSSY A. **Atmospheric chemistry.** [S.l.]: Eötvös Loránd University, 2013.

LAURET, P., HEYMESA, F., APRINA, L., JOHANETB, A., DUSSERREA, G., MUNIERC, L., LAPÉBIEC, E. **Near field atmospheric dispersion modeling on an industrial site using neural networks.** Chemical Engineering Transaction 31, 2013.

McCULLOCH, W. S.; PITTS, W. **A logical calculus of the ideas immanent in nervous activity.** The Bulletin of Mathematical Biophysics, v. 5, n. 4, p. 115-133, 1943.

OLIVEIRA, J. P. S., ALVES, J. V. B., CARNEIRO, J. N. E., MEDRONHO, R. A., SILVA, L. F. L. R. **Coupling a neural network technique with CFD simulations for predicting 2-D atmospheric dispersion analyzing wind and composition effects.** Journal of Loss Prevention in the Process Industries 80, 104930, 2022

PEREIRA, C. M. N. A., SCHIRRU, R., GOMES, K. J., CUNHA, J. L. **Artificial neural networks for radiation dose prediction in nuclear emergencies: preliminary investigations.** Advances in Computer Science Research. Atlantis Press ISSN 2352-538X, 2016.

PEREIRA, C. M. N. A., SCHIRRU, R., GOMES, K. J., CUNHA, J. L. **Development of a mobile dose prediction system based on artificial neural networks for NPP emergencies with radioactive material releases.** Annals of Nuclear Energy 105, 219-225, 2017.

PHUONG M., HUTTER M. **Formal algorithm for transformers.** doi:10.48550/arXiv.2207.09238, 2022.

QIU, S., CHEN, B., WANG, R., ZHU, Z., WANG, Y., QIU, X. **Atmospheric dispersion prediction and source estimation of hazardous gas using artificial neural network, particle swarm optimization and expectation maximization.** Atmospheric Environment 178, 158-163, 2018.

ROSENBLATT, F. **The perceptron: a probabilistic model for information storage and organization in the brain.** Psychological Review, v. 65, n. 6, p. 386, 1958.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. **Learning representations by back-propagating errors.** Nature, v. 323, n. 6088, pp. 533-536, 1986.

SULLIVAN, T. J., ELLIS, J. S., FOSTER C. S. **Atmospheric Release Advisory Capability: Real-Time Modeling of Airborne Hazardous Materials.** Bulletin of the American Meteorological Society, 74[12], 2343, 1993.

TIMONIN, V., SAVELIEVA, E. **Spatial prediction of radioactivity using general regression neural networks.** Applied GIS, vol.1, n.2, Monash University EPress, 2005.

WEN, Q., ZHOU, T., ZHANG, C., CHEN, W., MA, Z., YAN, J., SUN L. **Transformers in time series: a survey.** doi: arXiv:2202.07125v5, 2022.

YI, J., PRYBUTOK, R. **A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area.** Environmental Pollution 92 (3), 349-357, 1996.