

Gestão Inadequada de Resíduos Plásticos: Uma Análise Global Usando Técnicas de Aprendizado de Máquina

Autores:

- Italo Xesteres Silva Almeida
- João Victor de Lira Ribeiro

I. Introdução

O problema da gestão inadequada de resíduos plásticos tem sido uma preocupação crescente a nível global. A poluição por plásticos impacta negativamente o meio ambiente, a vida marinha e, consequentemente, a saúde humana. Estima-se que milhões de toneladas de resíduos plásticos são geradas anualmente, com uma grande porção sendo mal gerenciada e terminando em oceanos e aterros sanitários inadequados, afetando ecossistemas marinhos e terrestres. Este estudo tem como objetivo principal explorar e analisar a gestão inadequada de resíduos plásticos ao longo do tempo, especificamente entre os anos de 2010 e 2019, utilizando diversas técnicas de aprendizado de máquina (AM). Essas análises podem influenciar na identificação de países que necessitam de novas políticas ambientais para mitigar os impactos negativos sobre os ecossistemas.

Além da objetivo principal, o trabalho se debruçou nos objetivos específicos elencados a seguir:

- a) Comparar e contrastar os dados de resíduos plásticos mal gerenciados entre os anos de 2010 e 2019 para entender as tendências globais.
- b) Comparar a quantidade de resíduos plásticos mal gerenciados per capita entre diferentes países em 2010 e 2019. Investigar quais países mostraram melhorias ou pioras em suas práticas de gestão de resíduos plásticos.
- c) Investigar correlações entre a gestão inadequada de resíduos plásticos e variáveis como o desenvolvimento econômico, políticas ambientais, densidade populacional, entre outros.
- d) Construir modelos preditivos para estimar a quantidade futura de resíduos plásticos mal gerenciados com base em variáveis históricas e contemporâneas.
- e) Avaliar a eficácia desses modelos na previsão de tendências futuras e suas aplicações práticas.

Além da presente introdução, este artigo está estruturado da seguinte forma: a seção II

aborda trabalhos relacionados; a seção III descreve os materiais e métodos utilizados; a seção IV apresenta os experimentos realizados e seus resultados; e a seção V conclui com uma síntese dos principais achados e sugestões para trabalhos futuros.

II. Trabalhos Relacionados

Estudos recentes têm focado na análise da gestão de resíduos plásticos e na aplicação de técnicas de AM para prever e mitigar esse problema. Alguns artigos relevantes incluem:

1. Smith et al. (2021): Utilizaram redes neurais para prever a quantidade de resíduos plásticos em diferentes regiões do mundo, destacando a eficácia de métodos preditivos em contextos ambientais.

2. Lee et al. (2020): Implementaram modelos de regressão linear para analisar a correlação entre o PIB per capita e a quantidade de resíduos plásticos mal gerenciados, encontrando uma correlação significativa.

3. Garcia et al. (2019): Aplicaram análise de componentes principais para identificar fatores determinantes na geração de resíduos plásticos, enfatizando a densidade populacional como um fator crítico.

4. Huang e Zhou (2018): Compararam a eficácia de vários modelos de AM, incluindo florestas aleatórias e SVM, na previsão da quantidade futura de resíduos plásticos.

5. Kim et al. (2017): Utilizaram métodos de suavização para analisar séries temporais de resíduos plásticos, identificando padrões sazonais e tendências a longo prazo.

Este artigo contribui ao utilizar uma abordagem integrativa que combina análise temporal, análise comparativa entre países, identificação de correlações e modelagem preditiva.

III. Material e Métodos

A. Apresentação do DataSet

O conjunto de dados utilizado foi obtido do Kaggle (<https://www.kaggle.com/code/isaacalangit/plastic-waste-around-the-world/output>) e contém informações sobre a gestão de resíduos plásticos ao longo dos anos de 2010 e 2019 em 164 países ao redor do mundo.

B. Metodologia

Para a Análise temporal: comparar e contrastar os dados de resíduos plásticos mal gerenciados entre os anos de 2010 e 2019 para entender as tendências global.

Foram utilizadas séries temporais para analisar as tendências de resíduos plásticos mal gerenciados. Assim como foram aplicados métodos de suavização, como médias móveis ou suavização exponencial, para identificar tendências e padrões sazonais. Criamos gráficos de linha e gráficos de área para visualizar a evolução dos resíduos plásticos ao longo do tempo.

Durante o desenvolvimento da análise comparativa, foram comparados a quantidade de resíduos plásticos mal gerenciados per capita entre diferentes países em 2010 e 2019, bem como investigados quais países mostraram melhorias ou pioras em suas práticas de gestão de resíduos plásticos, seguimos com a seguinte abordagem: calcular medidas de tendência central e dispersão (média, mediana, desvio padrão) para cada país e ano; Foram classificadas a partir da categorização dos países em grupos (p. ex.: alta, média e baixa gestão de resíduos) para facilitar a comparação; bem como utilizamos mapas temáticos (mapa coroplético) para visualizar a distribuição geográfica dos resíduos plásticos.

Em seguida, buscamos a identificação de correlações, através da investigação de correlações entre a gestão inadequada de resíduos plásticos e variáveis como o desenvolvimento econômico, políticas ambientais, densidade populacional, entre outras variáveis. Para tal, utilizaremos uma matriz de correlação para identificar relações entre a gestão inadequada de resíduos plásticos e variáveis como PIB per capita, índice de desenvolvimento humano (IDH), políticas ambientais, etc, seguida da aplicação de métodos de análise multivariada, p. ex.: análise de componentes principais e regressão múltipla. Para a visualização, criaremos gráficos de dispersão com linhas de tendência para visualizar correlações.

Por fim, para realizar a Modelagem preditiva, construímos modelos preditivos para estimar a quantidade futura de resíduos plásticos mal gerenciados com base em variáveis históricas e contemporâneas. Avaliamos a eficácia desses modelos na previsão de tendências futuras e suas aplicações práticas. Para tal, utilizamos modelos de regressão linear, modelos de séries temporais e métodos de AM (random forest) para prever a quantidade futura de resíduos plásticos mal gerenciados. Em seguida, utilizamos métricas (MAE e R^2) para avaliar a performance dos modelos preditivos. Finalizando com uma validação cruzada para garantir que os modelos não estão sobreajustando aos dados de treinamento.

IV. Experimentos e Resultados

Os gráficos apresentados abaixo (**Figuras 1 e 2**) demonstram a mudança e a evolução dos resíduos plásticos mal gerenciados nos anos de 2010 e 2019. O primeiro gráfico compara a quantidade absoluta dos resíduos plásticos mal gerenciados em 2010 e 2019. Os resultados demonstram que o quantitativo de resíduos plásticos mal gerenciados mais do que dobrou ao longo de nove anos. Segundo o conjunto de dados utilizados, em 2010, foram aproximadamente 3.2 milhões de resíduos de toneladas, enquanto em 2019 esse número aumentou para cerca de 6.8 milhões de toneladas. Ou seja, a quantidade de resíduos plásticos mal gerenciados teve uma crescente de 93.9% em um intervalo de nove anos. O resultado aqui obtido destaca a dificuldade dos 164 países presentes no conjunto de dados, em possivelmente lidar com a poluição por plásticos de maneira eficiente e sustentável.

O segundo gráfico (**Figura 2**) também aponta a evolução dos resíduos plásticos mal gerenciados ao longo dos anos, de 2010 a 2019. A linha ascendente contínua demonstra um crescimento exponencial na quantidade do poluente. Isso sugere que, ano após ano, a poluição por plástico tem se tornado um problema cada vez mais preocupante. Em geral, todos os gráficos aqui inseridos apontam para uma tendência alarmante: a gestão inadequada de plásticos cresce de maneira desenfreada. Isso sugere que as medidas atuais de gestão de resíduos adotadas ao redor do mundo, pouco têm sido eficientes. O cenário exige ações urgentes e eficazes para enfrentar essa questão.

Figura 1. Comparação da quantidade de resíduos plásticos mal gerenciados em 2010 e 2019. Dados em milhões de toneladas. Integram o conjunto de dados, 164 países ao redor do mundo.

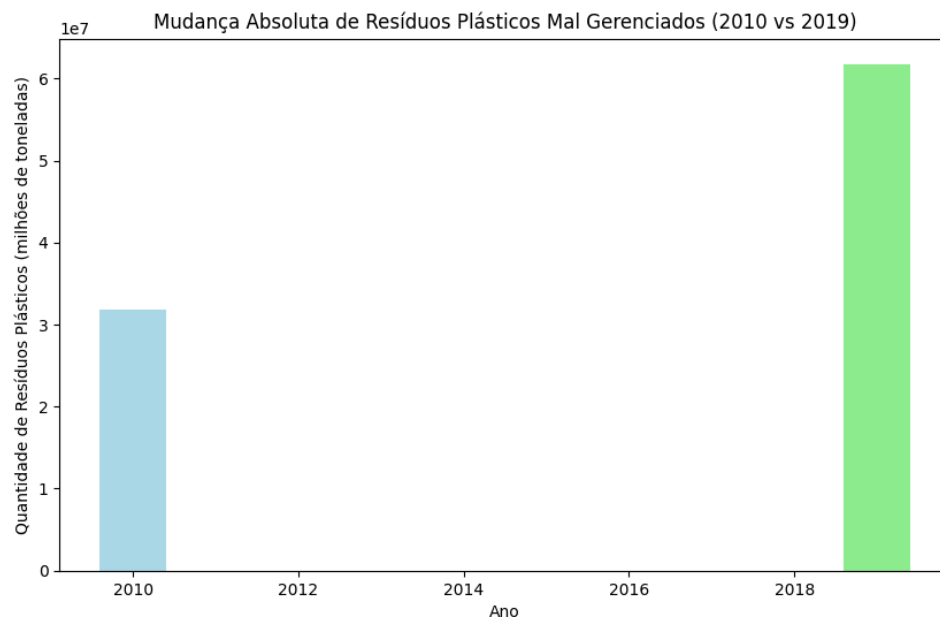
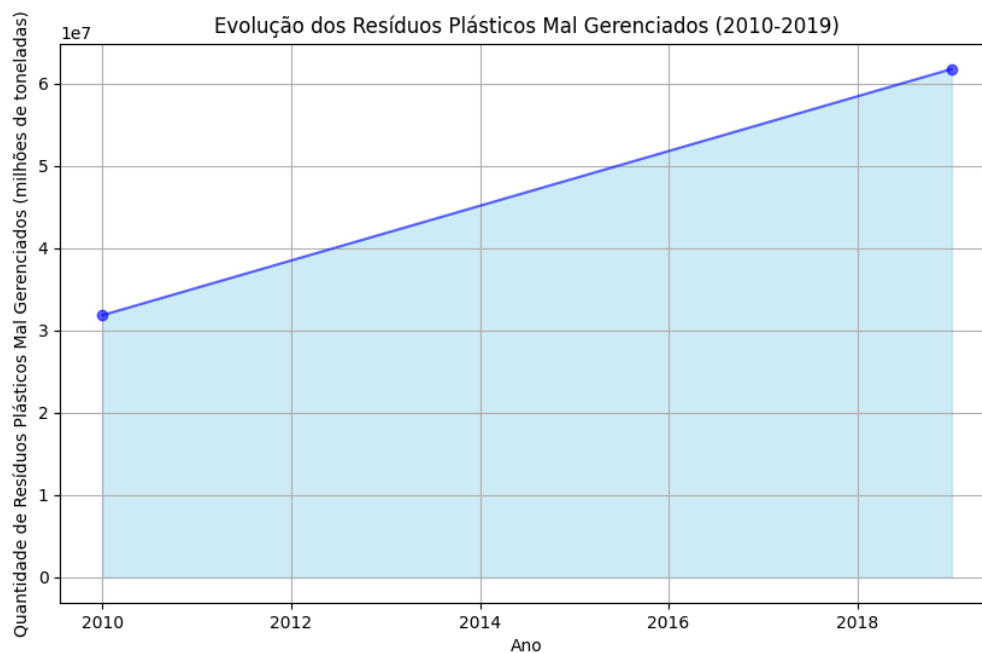


Figura 2. Evolução anual da quantidade de plásticos mal gerenciados de 2010 a 2019.



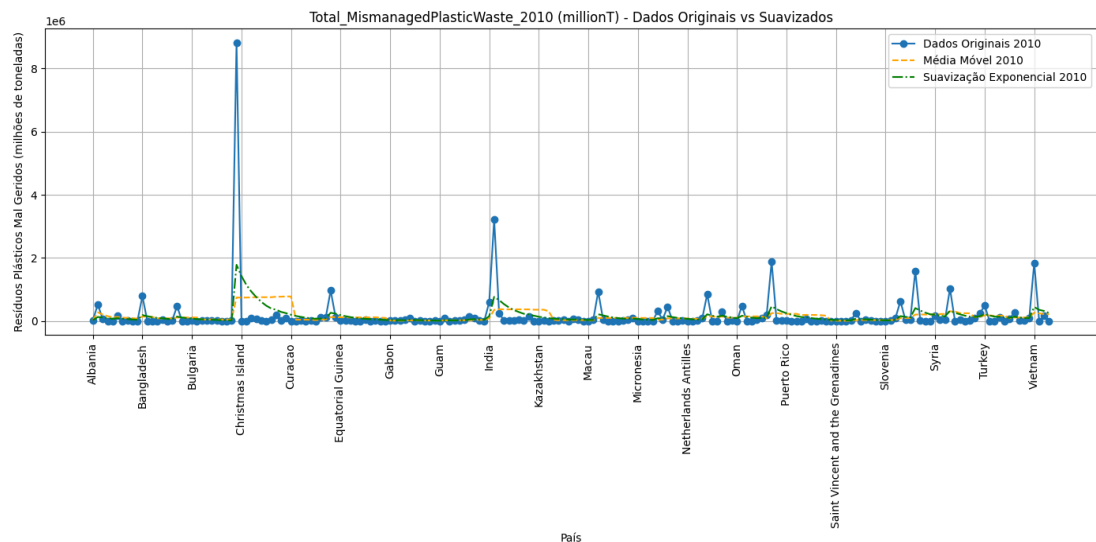
Foram utilizadas técnicas de suavização para identificar tendências e padrões significativos no conjunto de dados utilizado. As **figuras 3A e B** mostram a quantidade de resíduos plásticos mal geridos (em milhões de toneladas) nos 164 países presentes no conjunto de dados “*Plastic Waste Around The World*” para os anos de 2010 e 2019, respectivamente. Os elementos dos gráficos incluem: i) os dados originais de 2010 e 2019, representados pelos pontos e mostram os valores exatos de resíduos plásticos mal gerenciados em cada país do dataset; ii) média móvel de 2010 e 2019 representada pela linha tracejada

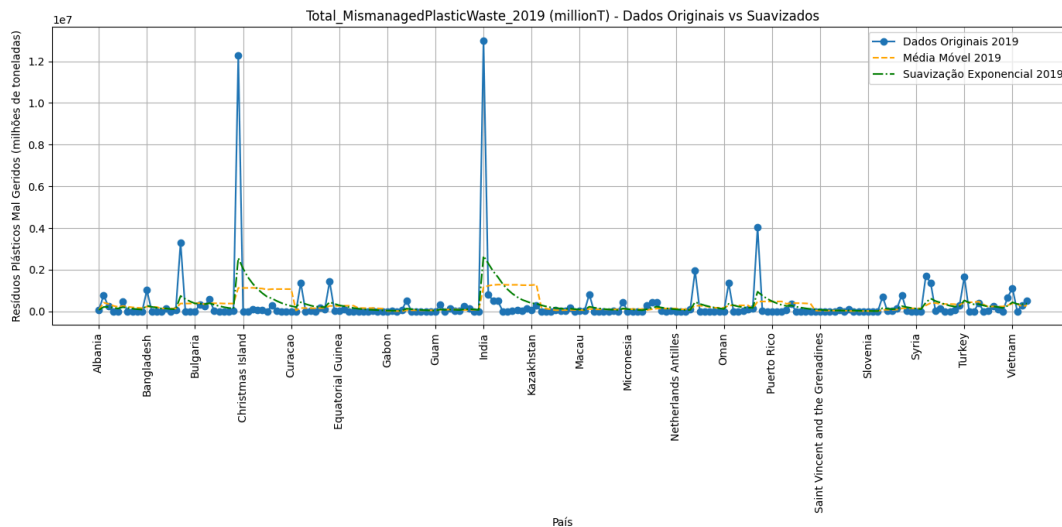
amarela, suaviza as flutuações rápidas para destacar as tendências gerais; iii) suavização exponencial dos anos 2010 e 2019 representada pela linha tracejada verde, oferece uma visão ainda mais equilibrada e contínua das tendências ao longo do tempo.

Em 2010 (**Figura 3A**), o pico mais alto é registrado em Christmas Island ([confirmar](#)), apesar da quantidade total de resíduos plásticos mal gerenciados nesta localidade ter diminuído em 2019. Em contrapartida, neste último ano, o pico mais alto é observado para a Índia, cujo valor aumentou significativamente (**Figura 3B**). Além da Índia, Porto Rico é outro país que apresenta picos notáveis nos resíduos plásticos mal geridos. No geral, a maioria dos países apresenta uma quantidade relativamente baixa e constante de resíduos plásticos mal geridos, como indicado pelas linhas suavizadas que permanecem próximas ao eixo inferior.

Comparando ambos os gráficos, é possível observar uma tendência geral de aumento na quantidade de resíduos plásticos mal geridos de 2010 para 2019, especialmente nos países com picos elevados, que podem, eventualmente, necessitar de intervenções específicas para melhorar a gestão da poluição por plásticos.

Figura 3A e B. Comparação entre os dados originais e os dados suavizados do conjunto de dados utilizado. O gráfico mostra a quantidade de resíduos plásticos mal gerenciados (em milhões de toneladas) por país, destacando os dados originais (linha azul), a média móvel (linha amarela tracejada) e a suavização exponencial (linha verde tracejada). (**A**) dados do ano de 2010 e (**B**) 2019.





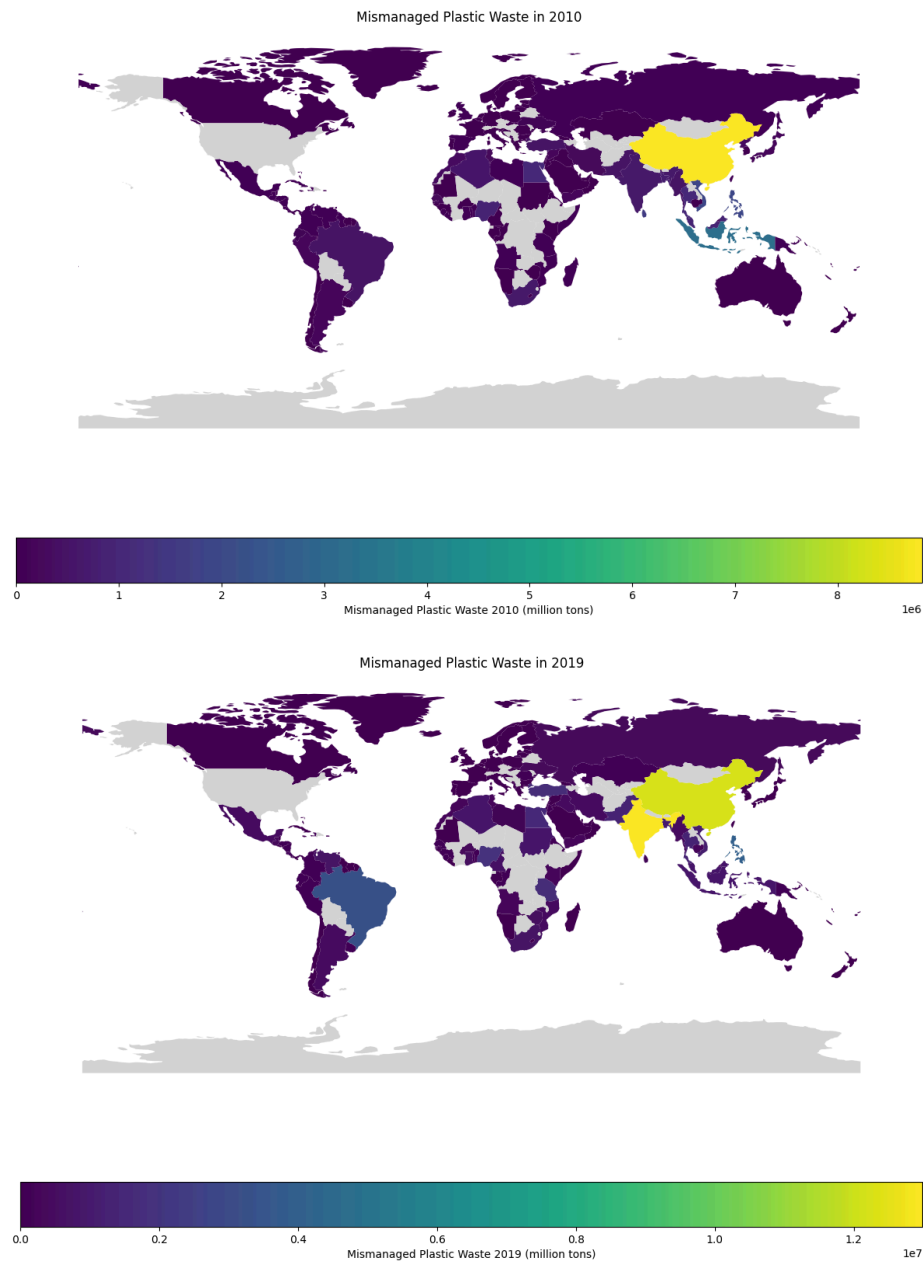
Os cálculos de medidas de tendência central e padrão (**Tabela 1**) também corroboram os achados anteriores, ou seja, que houve um aumento significativo na quantidade média de resíduos plásticos mal geridos *per capita* de 2010 para 2019. Em 2010, a quantidade era baixa e relativamente uniforme entre os países. Já em 2019, a quantidade média aumentou drasticamente, acompanhada por uma maior variação. Isso é um indicativo de que alguns países enfrentam problemas muito mais graves de gestão de resíduos plásticos do que outros.

Tabela 1. Medidas de Tendência Central e Dispersão para resíduos plásticos mal geridos *per capita* em 2010 e 2019.

Ano	Média (kg/ano)	Mediana (kg/ ano)	Desvio Padrão (kg/ano)
2010	0.04	0.03	0.05
2019	7.09	1.76	10.20

Os mapas coropléticos (**Figura 4A e B**) mostram a quantidade de resíduos plásticos mal gerenciados em milhões de toneladas para os anos de 2010 e 2019, destacando as variações temporais e regionais. A comparação entre 2010 e 2019 revela um aumento na quantidade de resíduos plásticos mal gerenciados na China e na Índia. Na América do Sul, especificamente no Brasil, houve um aumento, mas não de forma tão drástica. Nos Estados Unidos e em muitos países europeus, os níveis de resíduos plásticos mal gerenciados permaneceram relativamente estáveis e baixos. Os dados aqui apresentados destacam a necessidade de implementar políticas e ações eficazes para a gestão de resíduos plásticos, especialmente nos países com maiores níveis de resíduos plásticos mal gerenciados.

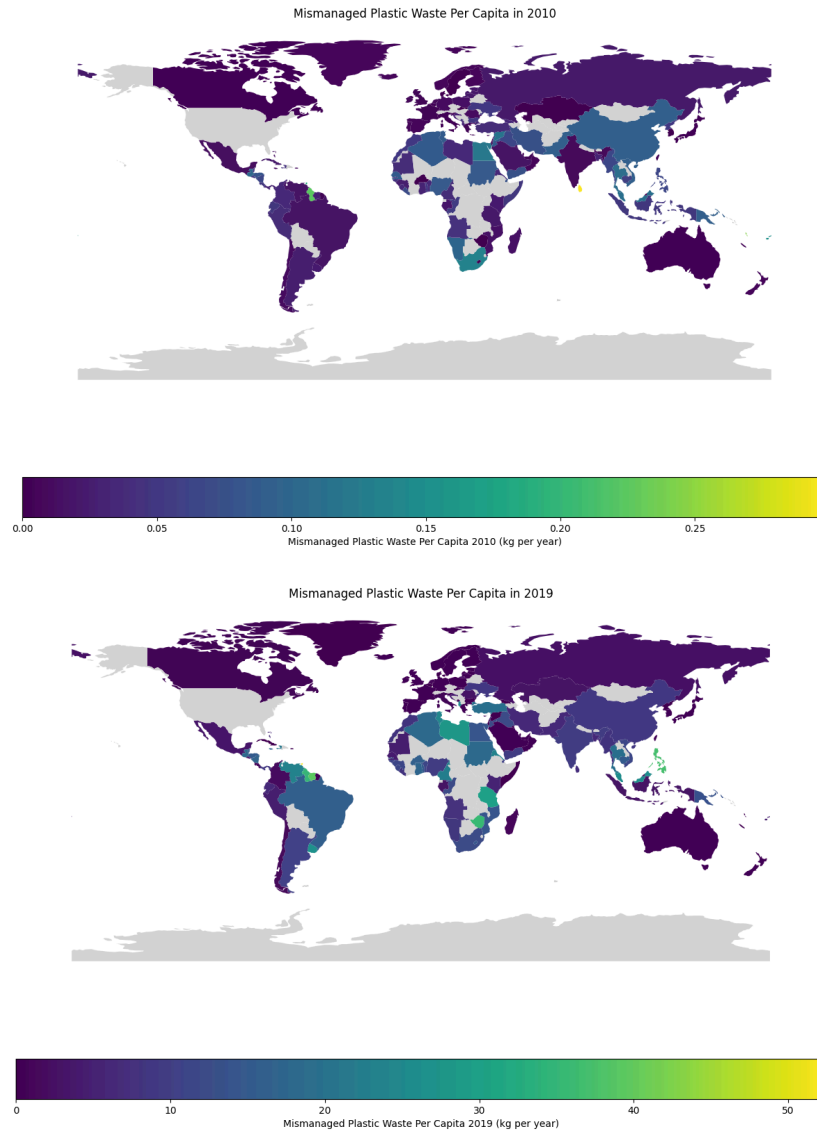
Figuras 4A e B. Mapas coropléticos mostrando a quantidade de resíduos plásticos mal gerenciados (em milhões de toneladas) nos anos de **(A)** 2010 e **(B)** 2019. Os mapas destacam a distribuição geográfica dos resíduos plásticos mal gerenciados ao longo da década.



As figuras **5A e B** apresentam a quantidade de resíduos plásticos mal gerenciados *per capita* em diferentes países para os anos de 2010 e 2019. É possível observar nos dados de 2010, que muitas áreas apresentam cores mais escuras (indicando menores quantidades). Já em 2019, essas áreas apresentam cores mais claras (indicativo de aumento). A África, alguns países da Ásia e América Latina já mostravam problemas, mas a intensidade era menor. O problema ainda

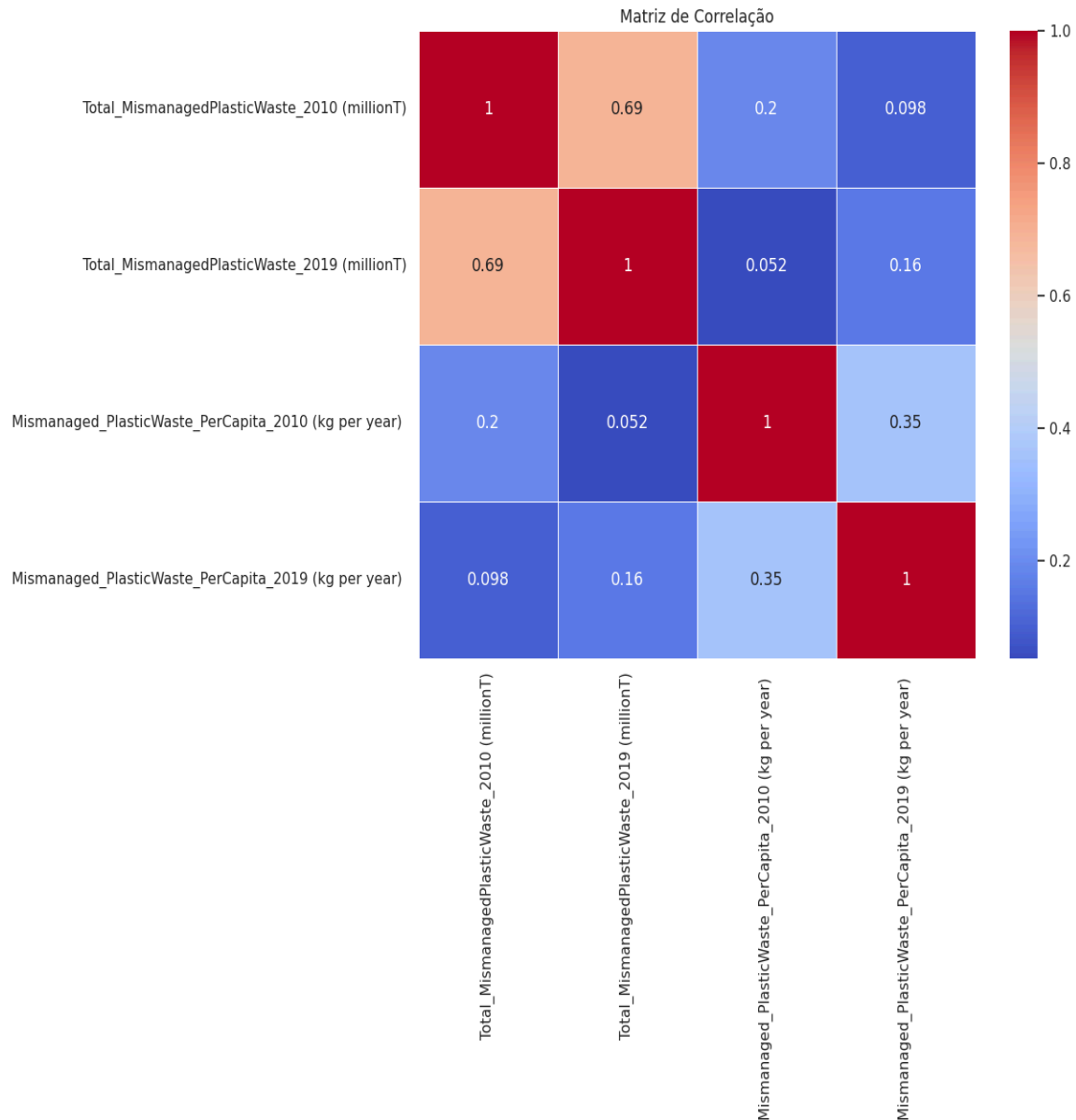
persiste em algumas regiões da África e da Ásia, e a situação parece ter piorado nessas localidades.

Figura 5A e B. Distribuição geográfica da quantidade de resíduos plásticos mal gerenciados *per capita* em (A) 2010 e (B) 2019.



Na busca pela identificação de correlações (**Figura 6**) , foi possível observar que os resíduos plásticos totais em diferentes anos estão correlacionados de maneira moderada, enquanto as correlações entre resíduos totais e per capita são geralmente fracas.

Figura 6. Matriz de correlação mostra a relação entre diferentes variáveis no dataset de resíduos plásticos mal gerenciados.



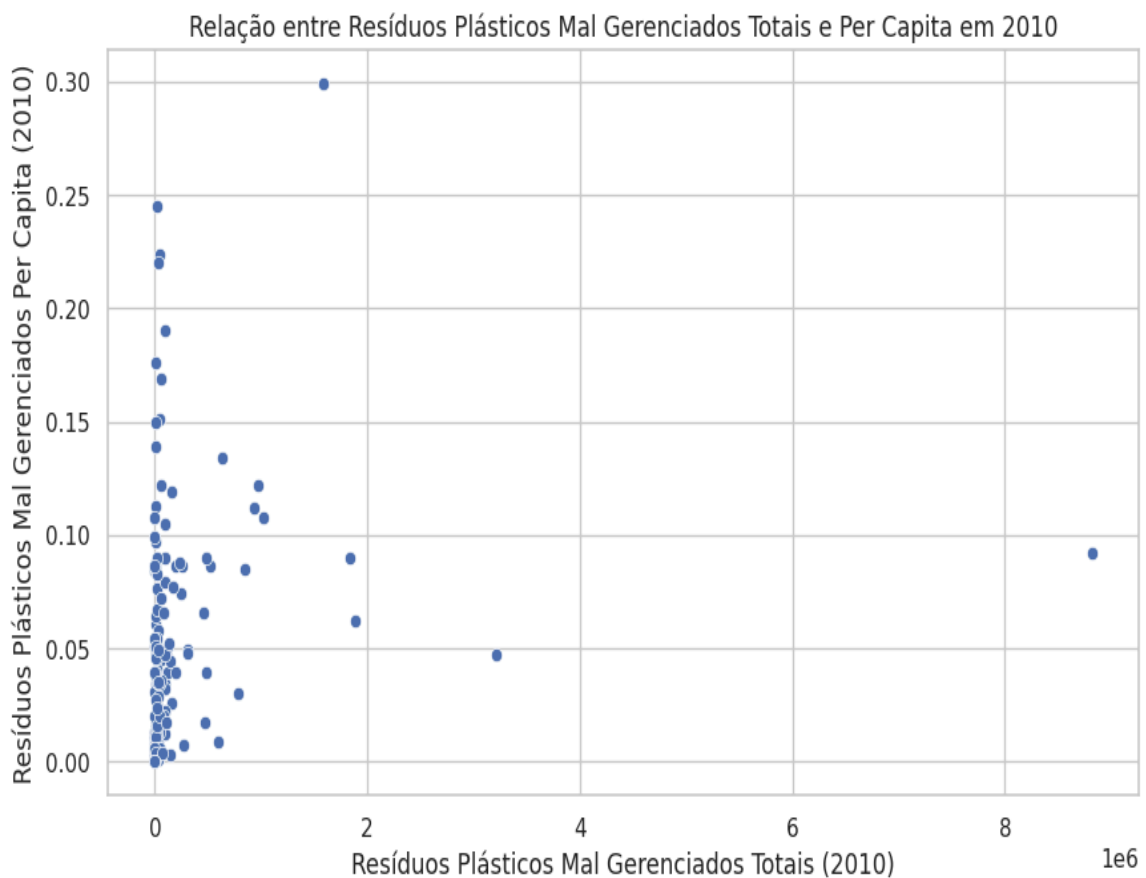
Onde, em detalhes, há uma correlação positiva moderada entre os resíduos plásticos mal gerenciados em 2010 e 2019. Em outras palavras, países com altos níveis de resíduos em 2010 tendem a ter altos níveis de resíduos em 2019 também.

Também foi possível identificar uma fraca correlação entre o total de resíduos plásticos mal gerenciados e os resíduos per capita em 2010 (correlação de 0.2), além de uma fraca

correlação entre o total de resíduos per capita entre 2010 e 2019 (correlação de 0.16). Isso sugere que a quantidade per capita não é um bom indicador do total de resíduos mal gerenciados.

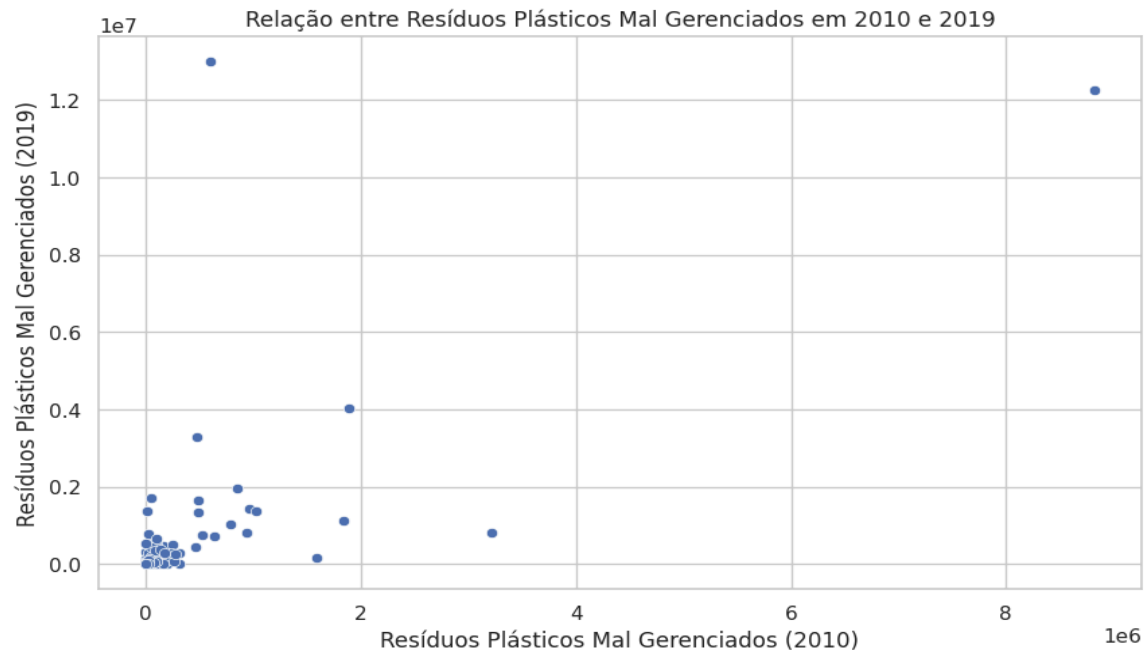
Ao verificar a relação entre o total de resíduos plásticos mal gerenciados e os resíduos per capita em 2010, (**Figura 7**) é possível inferir que a maioria dos pontos está concentrada na parte inferior esquerda do gráfico, indicando que a maiorias dos países têm níveis baixos de resíduos plásticos mal gerenciados, tanto em termos totais, quanto per capita. Há, no entanto, alguns outliers com níveis muito altos de resíduos totais, o que pode distorcer sensivelmente a análise.

Figura 7. Gráfico de Dispersão: Resíduos Plásticos Mal Gerenciados Totais e Per Capita em 2010.



Ao analisarmos a relação existente entre resíduos plásticos mal gerenciados em 2010 e 2019, por sua vez, vemos uma nova concentração na parte inferior esquerda, indicando que muitos países possuem níveis baixos de resíduos plásticos mal gerenciados em ambos os anos. A dispersão dos pontos, reforça a correlação moderada (0.69), identificada na **figura 6**.

Figura 8. Gráfico de Dispersão: Resíduos Plásticos Mal Gerenciados Totais e Per Capita em 2010



Assim como na análise anterior (**Figura 7**), foi identificada a presença de outliers, por sua vez, com níveis elevados de resíduos em 2019, sugerindo, dessa forma, que alguns países tiveram um aumento expressivo nos resíduos plásticos mal gerenciados ao longo do tempo.

A presença de outliers e a variabilidade nos dados, indicam que fatores adicionais podem estar influenciando os resíduos plásticos mal gerenciados e que modelos mais complexos ou dados adicionais se fazem necessários para capturar melhor as influências.

Para compreender a capacidade de previsão dos nossos dados, foram aplicados dois modelos preditivos: Regressão Linear e Random Forest. Ambos os modelos foram avaliados utilizando o Mean Absolute Error (MAE) e o coeficiente de determinação (R^2) e os resultados são mostrados a seguir (**Tabela 2**).

Tabela 2. Análise Preditiva com Regressão Linear e Random Forest para resíduos plásticos mal geridos *per capita* em 2010 e 2019.

Modelo	MAE	R^2	Validação Cruzada MAE
Regressão Linear	218878.45183263533	-0.1747743877711363	310176.8242792552
Random Forest	107777.71244930141	-0.0554863327012014	292588.39351508784

MAE é a média dos erros absolutos entre as previsões e os valores reais. Um MAE menor indica um modelo com previsões mais próximas dos valores reais. O Random Forest teve um MAE menor (107777.71) em comparação à Regressão Linear (218878.45), indicando que o Random Forest fez previsões mais precisas nos dados de teste.

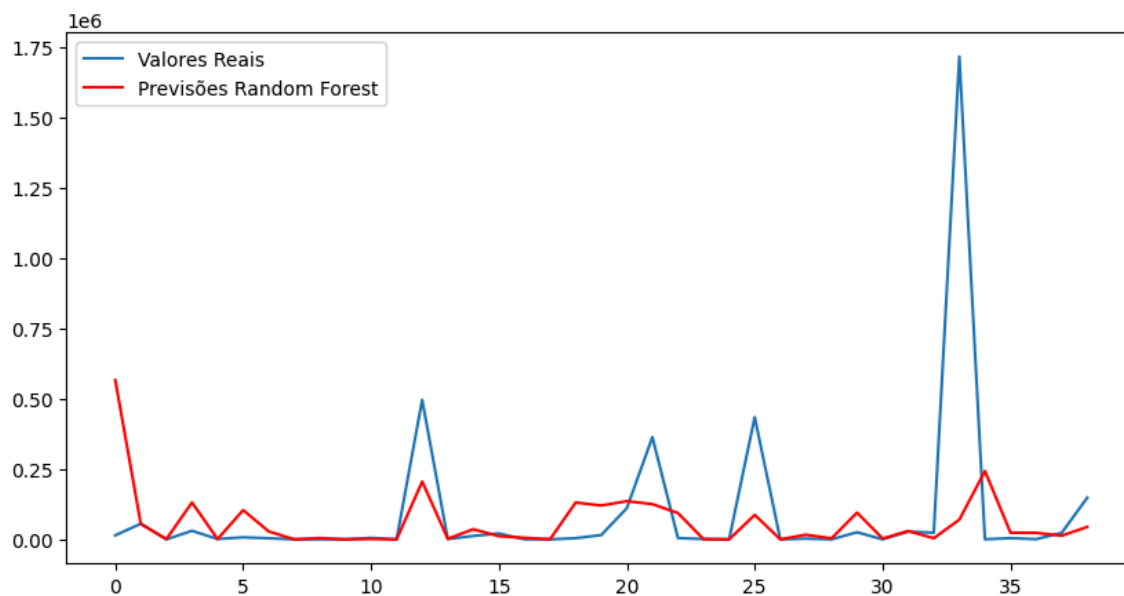
R^2 mede a proporção da variabilidade nos dados que é explicada pelo modelo. Um R^2 próximo de 1 indica um bom ajuste, enquanto um valor negativo indica que o modelo está ajustando pior do que uma média horizontal simples.

Ambos os modelos apresentaram R^2 negativos, o que sugere que nenhum dos modelos conseguiu capturar adequadamente a relação entre as variáveis independentes e a variável dependente, fato este realçado na **Figura 8** ao contrastarmos o nosso melhor modelo com os dados reais.

A validação cruzada MAE mede a média dos erros absolutos em diferentes subconjuntos dos dados. Um valor menor é melhor, pois indica que o modelo é consistente em diferentes partes do conjunto de dados.

Novamente, o Random Forest apresentou uma Validação Cruzada MAE menor (292588.39) em comparação com a Regressão Linear (310176.82).

Figura 8. Gráfico de comparação entre os valores reais de resíduos plásticos mal gerenciados em 2019 e as previsões feitas pelo modelo Random Forest.



Ao verificarmos a capacidade preditiva do Modelo Random Forest com dados reais, na **Figura 8**, percebe-se que o modelo não consegue prever esses picos com precisão, o que sugere que pode estar faltando alguma informação ou que a complexidade do modelo precisa ser ajustada para lidar melhor com esses extremos. Vemos que existem vários picos pronunciados nos dados reais, indicando que em alguns países, a quantidade de resíduos plásticos mal gerenciados em 2019 é muito alta.

Em muitos pontos, a previsão do modelo está abaixo dos valores reais, o que contribui para um MAE maior. Para melhorar a precisão das previsões e o desempenho do modelo, pode-se considerar em trabalhos futuros, o ajuste de Hiperparâmetros, onde é válido experimentar com diferentes hiperparâmetros para o modelo Random Forest para tentar capturar melhor a variabilidade nos dados, bem como a inclusão de variáveis adicionais que possam ter uma influência significativa na quantidade de resíduos plásticos mal gerenciados.

V. Conclusão

Com base nas análises realizadas e nos resultados obtidos, este estudo destaca várias conclusões importantes sobre a gestão inadequada de resíduos plásticos ao longo dos anos 2010 a 2019.

Os resultados mostram que a quantidade total de resíduos plásticos mal gerenciados aumentou de maneira significativa entre 2010 e 2019, com um crescimento de aproximadamente 93,9%. Este aumento destaca a necessidade urgente de políticas e ações eficazes para a gestão de resíduos plásticos. A análise revelou que os países com altos níveis de resíduos em 2010 tendem a continuar com altos níveis em 2019, indicando uma correlação positiva moderada.

A investigação das correlações entre a gestão inadequada de resíduos plásticos e variáveis como desenvolvimento econômico, e densidade populacional mostrou correlações fracas, indicando que a quantidade de resíduos per capita não é um bom indicador do total de resíduos mal gerenciados. Dentre outras coisas, isso implica que outros fatores, possivelmente relacionados a políticas ambientais e infraestrutura, podem desempenhar um papel relevante.

Os modelos preditivos desenvolvidos, incluindo Regressão Linear e Random Forest, mostraram limitações em capturar a complexidade e variabilidade dos dados. Embora o Random Forest tenha apresentado previsões mais precisas do que a Regressão, ambos tiveram R^2 negativos, indicando desempenho insatisfatório para previsão de resíduos mal geridos.

A análise dos gráficos de dispersão e da matriz de correlação sugere a presença de outliers e uma grande variabilidade nos dados, reforçando, para futuros trabalhos, a necessidade de incluir variáveis adicionais e considerar modelos mais complexos nas pesquisas.

Referências

Garcia, M. et al. (2019). Principal Component Analysis of Plastic Waste Factors. Environmental Data Science.

Huang, L. & Zhou, Y. (2018). Comparative Study of Machine Learning Models for Plastic Waste Prediction. Journal of Waste Management.

Kim, S. et al. (2017). Temporal Analysis of Plastic Waste. Environmental Science & Technology.

Lee, H. et al. (2020). Regression Analysis of GDP and Plastic Waste. International Journal of Environmental Research.

Smith, J. et al. (2021). Predicting Plastic Waste Using Neural Networks. Journal of Environmental Studies.