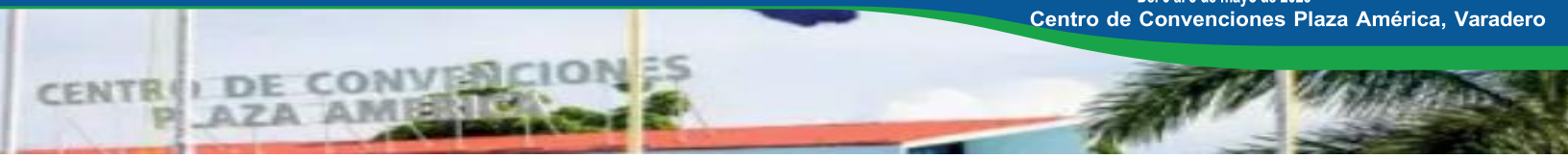


UN MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA ANALIZAR EL FENÓMENO DE ISLAS DE CALOR URBANAS EN LA CIUDAD DE SÃO PAULO, BRASIL

A MULTIPLE LINEAR REGRESSION MODEL TO ANALYZE THE URBAN HEAT ISLAND PHENOMENON IN THE CITY OF SÃO PAULO, BRAZIL

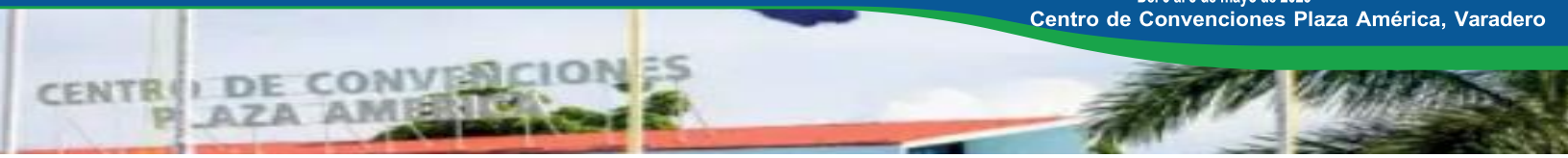
Autores:

1: Caroline Santos Silva Emolo
Identificador ORCID: 0009-0009-3518-3971
Grado científico, título académico, categoría docente y/o investigativa: Estudiante de Maestría
Centro de trabajo o estudio: Centro Estadual De Educação Tecnológica Paula Souza – Unidade De Pós-Graduação, Extensão E Pesquisa
Ocupación laboral: Docente
Cargo: Docente
Dirección postal: Rua dos Bandeirantes, 169 - Bom Retiro, São Paulo - SP - CEP: 01124-010, Brasil
Correo electrónico: caroline.silva@cpspos.sp.gov.br

2: Rosinei Batista Ribeiro
Identificador ORCID: 0000-0002-2150-4561
Grado científico, título académico, categoría docente y/o investigativa: Postdoctorado
Centro de trabajo o estudio: Centro Estadual De Educação Tecnológica Paula Souza – Unidade De Pós-Graduação, Extensão E Pesquisa
Ocupación laboral: Docente
Cargo: Docente
Dirección postal: Rua dos Bandeirantes, 169 - Bom Retiro, São Paulo - SP - CEP: 01124-010, Brasil
Correo electrónico: rosinei.ribeiro@cpspos.sp.gov.br

3: Tereza Pérez Sosa
Identificador ORCID: 0000-0002-5587-4624
Grado científico, título académico, categoría docente y/o investigativa: Doctorado
Centro de trabajo o estudio: Universidad De Matanzas – UMCC
Ocupación laboral: Docente
Cargo: Docente
Dirección postal: 2FPR+86P, Matanzas, Cuba
Correo electrónico: teresa.perez@umcc.cu

4: Eliane Antonio Simões
Identificador ORCID: 0000-0002-0738-2625



Grado científico, título académico, categoría docente y/o investigativa: Doctorado
Centro de trabajo o estudio: Centro Estadual De Educação Tecnológica Paula Souza – Unidade De Pós-Graduação, Extensão E Pesquisa
Ocupación laboral: Docente
Cargo: Docente
Dirección postal: Rua dos Bandeirantes, 169 - Bom Retiro, São Paulo - SP - CEP: 01124-010, Brasil
Correo electrónico: eliane.simoes@cpspos.sp.gov.br

RESUMEN

Este estudio propone un modelo de Regresión Lineal Múltiple para analizar las Islas de Calor Urbanas (ICU) en São Paulo, Brasil, combinando datos de tres regiones estratégicas. Al integrar datos meteorológicos y ambientales, el modelo predice variaciones de temperatura basadas en variables como humedad, contaminantes y precipitación. Los resultados validan el modelo y proporcionan herramientas para la toma de decisiones en políticas públicas y planificación urbana ante la crisis climática. Este enfoque se presenta como una herramienta clave para comprender el impacto de las islas de calor y mejorar las estrategias urbanas, contribuyendo a la sostenibilidad y la innovación en el contexto del cambio climático.

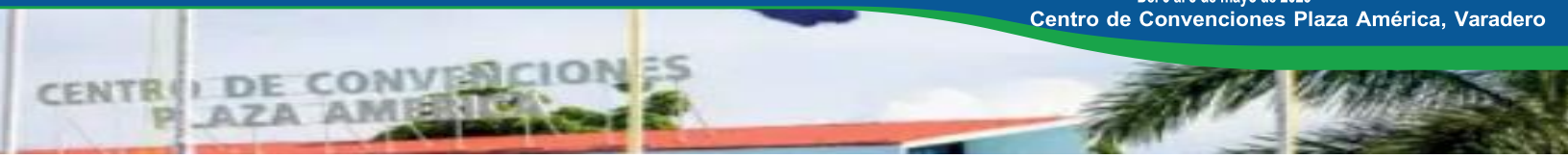
PALABRAS CLAVE

Innovación, Islas de Calor Urbanas, Modelo de Regresión Lineal Múltiple, Planificación Urbana, Sostenibilidad.

ABSTRACT

This study proposes a Multiple Linear Regression model to analyze the Urban Heat Island (UHI) phenomenon in São Paulo, Brazil, combining data from three strategic regions. By integrating meteorological and environmental data, the model predicts temperature variations based on variables such as humidity, pollutants, and precipitation. The results validate the model and provide tools for decision-making in public policies and urban planning in the face of the climate crisis. This approach is presented as a key tool to understand the impact of UHIs and improve urban strategies, contributing to sustainability and innovation in the context of climate change.
KEYWORDS
Innovation, Multiple Linear Regression Model, Sustainability, Urban Heat Islands, Urban Planning.

INTRODUCCIÓN



La urbanización acelerada ha transformado el paisaje natural e intensificado los desafíos ambientales y climáticos (Gomes & Pinto, 2020). En São Paulo, Brasil, este proceso ha agravado la desigualdad centro-periferia (Carlos, 2009; Espíndola et al., 2020) y problemas como las Islas de Calor Urbanas (ICU), caracterizadas por el aumento de la temperatura en áreas urbanas en comparación con las rurales (Oke, 2002).

En este contexto, el objetivo de este estudio es analizar cómo las variables ambientales influyen en la formación de las ICU en el municipio de São Paulo. Para ello, se aplica la Regresión Lineal Múltiple (RLM) utilizando datos meteorológicos y de calidad del aire (2019-2023) de tres regiones estratégicas cercanas a estaciones automáticas. Este enfoque permite identificar patrones e investigar los principales factores que contribuyen a este fenómeno (Lopes & Hora, 2023).

MATERIALES Y MÉTODOS

La metodología involucró la recopilación de datos meteorológicos y ambientales de tres regiones de São Paulo, denominadas A1, A2 y A3, correspondientes a los años 2019 a 2023. Los datos fueron obtenidos a través de la Companhia Ambiental do Estado de São Paulo (CETESB), del Instituto Nacional de Meteorologia (INMET) y del Instituto de Astronomia, Geofísica e Ciências Atmosféricas de la Universidade de São Paulo (IAG-USP), con la ubicación de estas áreas presentada en la Tabla 1 y la Figura 1.

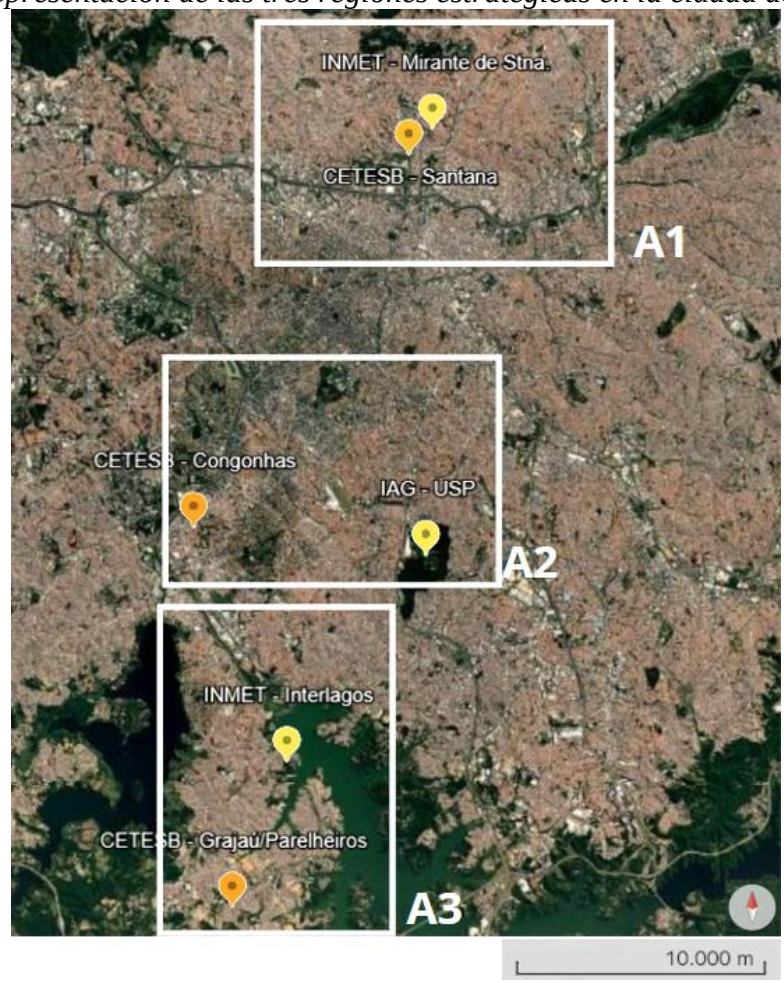
Tabla 1: Localización de las estaciones automáticas adoptadas para la obtención de datos.

Área de Análisis	Estación Automática CETESB	Estación meteorológica correspondiente
"A1"	Santana (Av. Santos Dumont, 1019, Parque de Material Aeronáutico, Santana, SP)	INMET – Mirante de Santana (Praça Vaz Guaçu – Jardim São Paulo, SP)
"A2"	Congonhas (Alameda dos Tupiniquins, 1571, E.M.Prof.J.C. da Silva Borges, Planalto Paulista, SP)	IAG – USP (Av. Miguel Estéfano, 4200 – Água Funda – SP)

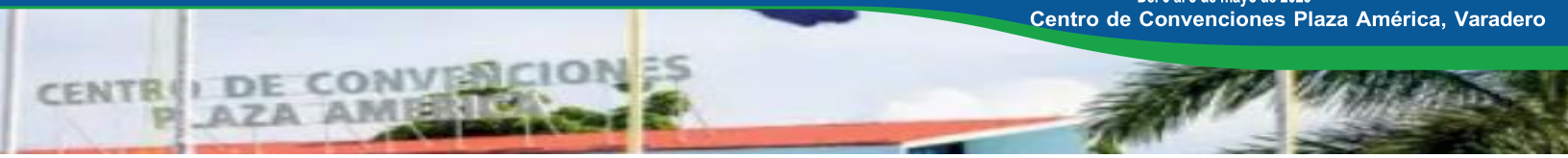
"A3"	Grajaú/ Parelheiros (Av. Paulo Guilguer Reimberg, 2448, E.E. Pres. Tancredo A. Neves, Jd Novo Horizonte, SP)	INMET – Interlagos (SESC Interlagos – SP)

Fuente: *los autores.*

Figura 1: *Representación de las tres regiones estratégicas en la ciudad de São Paulo.*



Fonte: *Los autores.*



Se recopilaron datos promedio mensuales de las siguientes variables: temperatura del aire (°C), definida como la variable dependiente del modelo estadístico, y humedad relativa del aire (%), emisión de contaminantes del tipo material particulado – MP2,5 (µg/m³) y precipitación promedio mensual (mm), consideradas variables independientes.
Los datos de A1, A2 y A3 se combinaron en un único conjunto, permitiendo la elaboración de un modelo unificado que representa la dinámica de las ICU en São Paulo. Después de la exportación a Excel® y la estandarización, se utilizó el software Eviews® para la construcción del modelo estadístico de RLM.
El modelo fue ajustado mediante el método de los mínimos cuadrados ordinarios, que minimiza la suma de los cuadrados de los residuos, proporcionando el mejor ajuste lineal entre las variables (Myers et al., 2016). La validez estadística se evaluó mediante pruebas fundamentadas y definidas por Myers et al. (2016), como la prueba F (significancia global del modelo), R^2 y R^2 ajustado (proporción de la variación explicada), multicolinealidad (correlación entre variables independientes) y heterocedasticidad (verificación de la constancia de la varianza de los residuos).

RESULTADOS

Se consideró para el modelo estadístico:

- Variable dependiente: Temperatura media mensual (°C) – T.
- Variables independientes: Humedad media mensual (%) – U; Emisión de contaminante MP2,5 (µg/m³) – EP; Precipitación media mensual (mm) – PT.

Con esto, se obtuvieron las ecuaciones del modelo de RLM: la ecuación inicial y la ecuación estimada, que refleja la relación práctica fundamentada en los datos observados, como se muestra en la Figura 2. Además, la Tabla 2 presenta la verificación de las variables significativas, utilizando el método de mínimos cuadrados ordinarios.

Figura 1: Ecuación inicial y ecuación del modelo estimado.

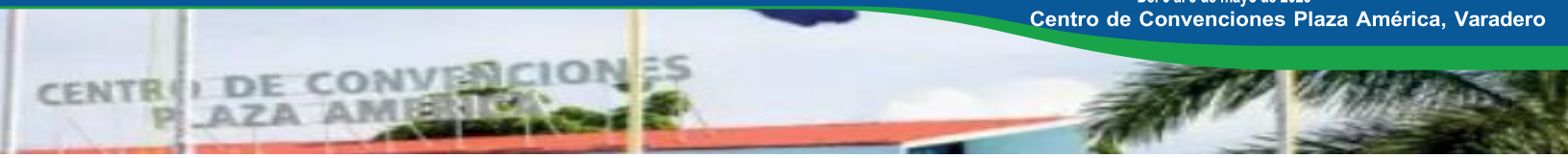
$$T = \beta_1 + \beta_2 \cdot U + \beta_3 \cdot EP + \beta_4 \cdot PT + U_i$$

$$\hat{T} = b_1 + b_2 \cdot U + b_3 \cdot EP + b_4 \cdot PT$$

Fuente: los autores.

Tabla 2: *Verificación de las variables estadísticamente significativas.*

Variable dependiente: T
Método: Mínimos Cuadrados Ordinarios
Fecha: 03/10/24 Hora: 15:12
Muestra: 1 180
Observaciones: 180



T=C(1)+C(2)*U+C(3)*EP+C(4)*PT

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	20.64300	0.414568	49.79398	0.0000
C(2)	1.334986	0.432449	3.087036	0.0023
C(3)	3.513369	0.437044	8.038939	0.0000
C(4)	-1.746440	0.410471	-4.254723	0.0000
R-cuadrado	0.322628	Media de la variable dependiente		19.89733
R-cuadrado ajustado	0.311081	Desv. Estándar – variable dependiente		2.387409
Error estándar de la regresión	1.981576	Criterio de info. de Akaike		4.227634
Suma de los cuadrados de los residuos	691.0895	Criterio de Schwarz		4.298589
Log-verosimilitud	-376.4871	F-statistic		27.94250
Estadística Durbin-Watson	1.238479	Prob(F-statistic)		0.000000

Fuente: los autores.

Los coeficientes fueron estadísticamente significativos (Prob < 0,05), lo que indica la influencia de las variables independientes sobre la temperatura. El coeficiente de determinación ajustado (R^2 = 0,3226) sugiere que aproximadamente el 32 % de la variación de la temperatura es explicada por el modelo. Se presenta la ecuación de regresión estimada resultante, como se muestra en la Figura 3.

Figura 2: Ecuación de regresión ajustada.

$$\hat{T} = 20.64 + 1.33 \cdot U + 3.51 \cdot EP - 1.75 \cdot PT$$

Fuente: los autores.

La ecuación resultante demostró una relación directa entre U y T, entre EP y T, y una relación inversa entre PT y T. Pruebas adicionales confirmaron la validez del modelo estadístico, con base en Myers et al. (2016), según el diagrama de flujo de la Figura 4 y el correlograma de la Figura 5.

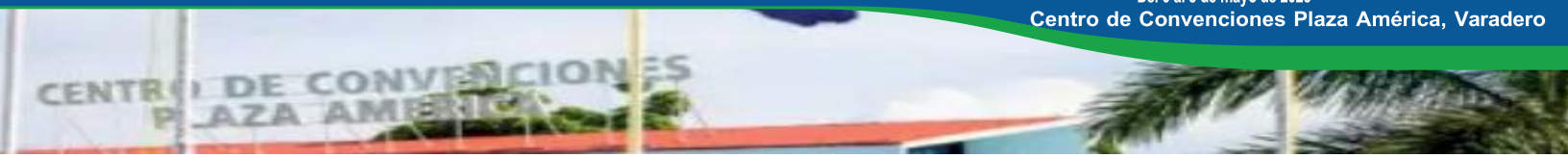


Figura 3: *Resultados de las pruebas de validación del modelo estadístico.*

Ausencia de multicolinealidad: la matriz de correlación entre las variables independientes demostró que no existe una relación lineal fuerte entre ellas, lo que indica la ausencia de multicolinealidad.

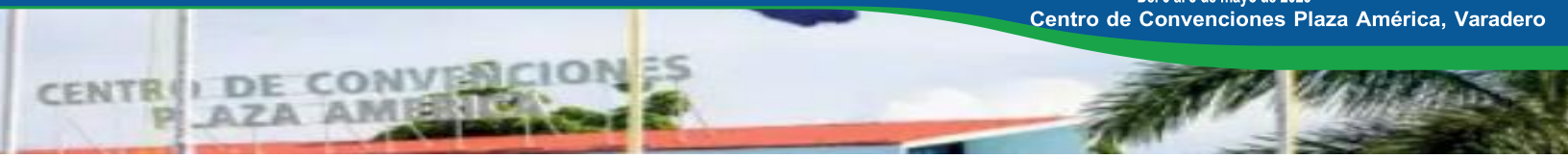
Valor esperado de los residuos igual a cero: al evaluar el modelo, se obtuvieron residuos $E(U_i)$ muy cercanos o iguales a cero.

Homoscedasticidad: para evaluar la homoscedasticidad del modelo, se realizó la prueba de White, que presentó un valor de probabilidad = 0,058097, mayor que el nivel de significancia de $\alpha = 0,05$, lo que indica la ausencia de heterocedasticidad.

Ausencia de autocorrelación de los residuos: para evaluar este supuesto, se aplicaron la prueba de Jun y Box y el correlograma. Los valores de probabilidad mayores que el nivel de significancia de $\alpha = 0,05$, junto con el correlograma, confirman la ausencia de autocorrelación entre los residuos.

Normalidad de los residuos: se aplicó la prueba de Jarque-Bera. El valor de probabilidad = 0,484546, mayor que el nivel de significancia de $\alpha = 0,05$, indicó que los residuos siguen una distribución normal.

Fuente: los autores.



Date: 10/03/24 Time: 15:37
Sample: 1 180
Included observations: 180

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.100	0.100	1.8289	0.176
		2	-0.077	-0.088	2.9269	0.231
		3	0.071	0.090	3.8574	0.277
		4	0.056	0.032	4.4505	0.348
		5	-0.093	-0.092	6.0798	0.299
		6	-0.003	0.020	6.0816	0.414
		7	-0.045	-0.072	6.4692	0.486
		8	-0.122	-0.099	9.2906	0.318
		9	-0.046	-0.023	9.6935	0.376
		10	0.008	-0.005	9.7066	0.467
		11	-0.074	-0.060	10.765	0.463
		12	-0.123	-0.108	13.736	0.318
		13	-0.033	-0.039	13.954	0.377
		14	0.037	0.026	14.223	0.433
		15	-0.050	-0.057	14.721	0.472
		16	0.098	0.110	16.657	0.408
		17	0.142	0.092	20.697	0.240
		18	-0.077	-0.103	21.899	0.236
		19	-0.015	-0.002	21.948	0.287
		20	0.172	0.106	28.004	0.109
		21	-0.039	-0.073	28.311	0.132
		22	-0.089	-0.035	29.967	0.119
		23	-0.010	-0.047	29.986	0.150
		24	0.166	0.186	35.791	0.058
		25	-0.087	-0.097	37.378	0.053
		26	0.030	0.067	37.568	0.066
		27	0.015	-0.015	37.614	0.084
		28	-0.094	-0.063	39.518	0.073
		29	0.078	0.165	40.852	0.071

Figura 5: *Prueba de Jun y Box y correlograma del modelo estadístico.*

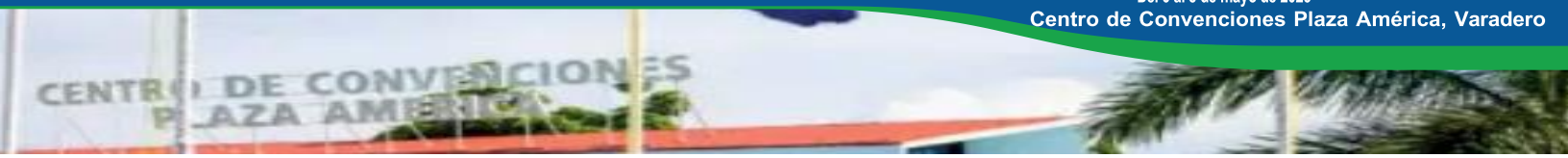
Fonte: os autores.

CONCLUSIONES

La regresión lineal múltiple (RLM) confirmó la influencia de la humedad, la contaminación y la precipitación en la temperatura (Siqueira et al., 2020; Lisbôa et al., 2020). Dado los coeficientes estadísticamente significativos ($p < 0,05$), el R^2 de 0,3226 sugiere que otros factores también impactan el fenómeno (Myers et al., 2016). Además, la RLM, al ser aplicada en software, permite una interpretación matemática directa de los resultados, facilitando la predicción de variaciones térmicas en contextos urbanos (Sitompul, 2022).

El modelo estadístico permitió comprender los efectos de las variables ambientales sobre la temperatura del aire considerando tres regiones distintas de São Paulo. La RLM demostró una buena capacidad de generalización, reforzando su aplicabilidad en el análisis de las Islas de Calor Urbanas.

REFERENCIAS BIBLIOGRÁFICAS



Carlos, A. F. A. (2009). A metrópole de São Paulo no contexto da urbanização contemporânea. *Estudos Avançados, 23,* 303–314. https://doi.org/10.1590/S0103-40142009000200021. Acceso en: 19 feb. 2025.

Companhia Ambiental do Estado de São Paulo. (n.d.). *QUALAR – Qualidade do ar*. https://cetesb.sp.gov.br/ar/qualar/. Acceso en: 15 oct. 2024.

Espíndola, I. B., & Ribeiro, W. C. (2020). Cidades e mudanças climáticas: Desafios para os planos diretores municipais brasileiros. *Cadernos Metrópole, 22(48),* 365-396. https://doi.org/10.1590/2236-9996.2020-4802. Acceso en: 19 feb. 2025.

Gomes, M. F., & Pinto, W. S. (2020). Justiça socioambiental e processo de urbanização das cidades / Socio-environmental justice and city urbanization process. *Revista de Direito da Cidade, 12(1),* 582–608. https://www.e-publicacoes.uerj.br/rdc/article/view/39931. Acceso en: 20 feb. 2025.

Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG/USP). (n.d.). *Dados solares*. http://www.estacao.iag.usp.br/sol_dados.php. Acceso en: 15 oct. 2024.

Instituto Nacional de Meteorologia – INMET. (n.d.). *Catálogo de Estações Automáticas*. https://portal.inmet.gov.br/paginas/catalogoaut. Acceso en: 15 sep. 2024.

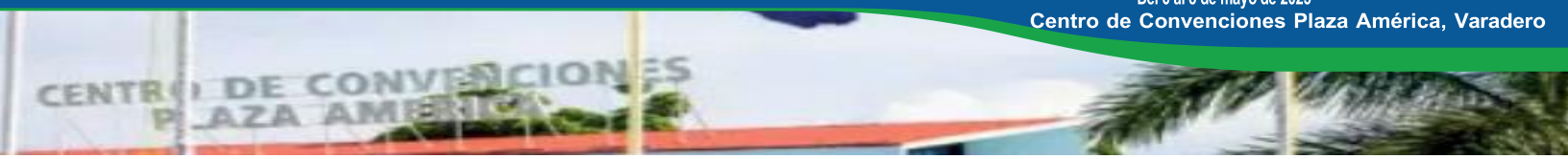
Lisbôa, E. G., et al. (2020). Indicadores de Desenvolvimento Sustentável: Uma Análise Quantitativa utilizando o modelo de Regressão Linear Múltipla. *Brazilian Journal of Development, 6(3),* 15179–15195.

Lopes, E. da S., & Hora, K. E. R. (2023). Modelo de regressão para identificação de ilha de calor: um mapeamento sistemático. *PARC: Pesquisa em Arquitetura e Construção, 14(00),* e023026. https://periodicos.sbu.unicamp.br/ojs/index.php/parc/article/view/8668386. Acceso en: 23 ene. 2025.

Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. (2016). *Response surface methodology: Process and product optimization using designed experiments*. John Wiley & Sons.

Oke, T. R. (2002). *Boundary layer climates* (2nd ed.). Routledge.

Siqueira, I. S., et al. (2020). Análise dos condicionantes Sócio-Ambientais na incidência de dengue na cidade de Belém/PA: Aplicação do Modelo de Regressão Linear Múltipla. *Brazilian Journal of Development,* 6(10), 80979–80991.



Sitompul, P. (2022). Analysis of Ordinary Least Square and Geographically Weighted Regression on the Human Development Index of North Sumatra 2021. *Formosa Journal of Applied Sciences, 1(*6), 981–1000.