

## EFFECTOS DE LA PANDEMIA EN LA REDUCCIÓN ECONÓMICA Y LA CONTAMINACIÓN

ALONSO NAHIR RAMÍREZ <sup>a,\*</sup>, HORTENSIA J. REYES CERVANTES <sup>a</sup>, FERNANDO VELASCO LUNA <sup>a</sup>

<sup>a</sup>Facultad de Ciencias Físico-Matemáticas

Benemérita Universidad Autónoma de Puebla, Dirección Avenida San Claudio y 18 Sur, Colonia San Manuel, Ciudad Universitaria, C.P. 72570, Puebla, México.

e-mail: alonso.nahir@alumno.buap.mx

Una de las principales vías del gobierno mexicano para contener la propagación del virus SARS-CoV-2 fue el confinamiento de la población. Al disminuir la movilidad social y el nivel de producción industrial también se reduce el grado de actividad económica. El objetivo de esta investigación es modelar mediante una regresión lineal la reducción de la actividad económica y su relación con los niveles de contaminación en el ambiente en la Ciudad de México para los años 2019, 2020 y 2021. El enfoque de esta investigación es cuantitativo, mientras que el alcance es de carácter descriptivo debido a que no se encontraron suficientes bases de datos que involucren las variables aleatorias.

**Keywords:** Contaminación, economía, SARS-COV-2, regresión lineal.

### 1. Introducción

En estos tiempos de calentamiento global hemos visto las consecuencias de la contaminación en todo el mundo con sus altos niveles de temperatura, sequía y nivel del mar. La contaminación ambiental es la causante de numerosas enfermedades o empeoramiento de ellas [10]. Es por ello la importancia de encontrar posibles causas de esta contaminación. En este trabajo se busca la relación entre la actividad económica y la contaminación del ambiente a través de un modelo de regresión lineal sobre los niveles de presencia de agentes contaminantes e índices económicos para los periodos mensuales de los años 2019, 2020 y 2021 en la Ciudad de México.

Durante la pandemia por SARS-COV-2 podemos ver una reducción significativa de la economía en el sentido de una contracción de la

producción y comercialización para la mayoría de los sectores lo cual se refleja en una reducción de la circulación de moneda y un aumento en la tasa de desempleo [2]. Los años 2019, 2020 y 2021 son de utilidad para disminuir el error aleatorio fuera de nuestra capacidad de medición al representar un antes, durante y después de la pandemia.

### 2. Modelo de Regresión Lineal

Un modelo de regresión donde interviene más de una variable regresora se llama modelo de regresión múltiple [6]. Si definimos a la variable aleatoria  $Y$  en función de los parámetros  $\beta_k$ , las variables independientes  $x_k$  y el error  $\varepsilon$ , al hacer  $n$  observaciones mutuamente independientes  $y_1, y_2, \dots, y_n$ , para  $i = 1, \dots, n$  podemos escribir la  $i$ -ésima observación como

$$y_i = \beta_0 + \beta_1 x_{i,1} + \dots + \beta_k x_{i,k} + \varepsilon_i.$$

\*Corresponding author

Para que este modelo sea aplicable es importante verificar las siguientes suposiciones, donde  $X$  es la matriz que se define a continuación.

$$\begin{aligned} E(\varepsilon_i) &= 0 \\ \text{Var}(\varepsilon_i) &= \sigma^2 \\ \varepsilon_i &\sim N(0, \sigma^2) \\ \text{Cov}(\varepsilon_i, \varepsilon_j) &= 0 \quad \text{para } i \neq j \\ &\exists (X^T X)^{-1}. \end{aligned} \quad (1)$$

**Definición 1.** (*Distribución normal*) Se dice que la variable aleatoria continua  $X$  tiene una distribución normal con parámetros  $\mu \in R$  y  $\sigma^2 > 0$ , se denota  $X \sim N(\mu, \sigma^2)$ , si su función de densidad está dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Como consecuencia directa de estas suposiciones podemos ver que  $E(Y) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$ . Los estimadores por mínimos cuadrados de los parámetros  $\beta_i$  son insesgados y de mínima varianza. Para este caso multiparamétrico se trabajará con matrices tales que

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} x_0 & x_{11} & x_{12} & \dots & x_{1k} \\ x_0 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad y \quad a = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix}.$$

Donde  $x_0 = 1$ . Note que de esta manera podemos reescribir el modelo como  $Y = X\beta + \varepsilon$ .

**Teorema 1.** Sea  $Y = X\beta + \varepsilon$  un modelo de regresión lineal en la forma matricial, los estimadores de mínimos cuadrados están dados por:

$$\hat{\beta} = (X^T X)^{-1} X^T Y.$$

Para el estudio de modelos de regresión lineal múltiple nos puede resultar conveniente hacer pruebas de hipótesis sobre la combinación

lineal de los parámetros, fijando un nivel de confianza  $\alpha$ , tal que

$$H_0 : a^T \beta = c \text{ vs } H_a : \neg H_0.$$

Utilizando el estadístico  $T = \frac{a^T \hat{\beta} - c}{S \sqrt{a^T (X^T X)^{-1} a}}$  que tiene una distribución  $t_\alpha$  con  $[n - (k + 1)]$  grados de libertad.

Es de particular interés conocer si al menos un parámetro  $\beta$  tiene un valor significativo, es decir, algún valor diferente de cero. Para ello tenemos la hipótesis

$$H_0 : \beta_i = 0 \text{ vs } H_a : \neg H_0.$$

Para esta prueba de hipótesis utilizamos el estadístico  $F = \frac{SS_R/k}{SS_{Res}/(n-k-1)}$  que tiene una distribución  $F_{k, n-k-1}$  [5]. Esta prueba es usualmente conocida como significancia del modelo. Recordemos que la suma de cuadrados debidos a la regresión  $SS_R$  y la suma de cuadrados residuales  $SS_{Res}$  se calculan de la forma matricial

$$SS_R = \hat{\beta}^T X^T Y - \frac{(\sum_{i=1}^n y_i)^2}{n} \quad (2)$$

$$SS_{Res} = Y^T Y - \hat{\beta}^T X^T Y.$$

*Dem.* (Teorema 1) Supongamos un modelo de regresión lineal simple. SSE tiene un mínimo en los valores que satisfacen las ecuaciones  $\partial SSE / \partial \hat{\beta}_0 = 0$  y  $\partial SSE / \partial \hat{\beta}_1 = 0$ . Esto es:

$$\begin{aligned} \frac{\partial SSE}{\partial \hat{\beta}_0} &= \frac{\partial [\sum_{i=1}^n [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2]}{\partial \hat{\beta}_0} \\ &= - \sum_{i=1}^n 2[y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)] \\ &= -2(\sum_{i=1}^n y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^n x_i) = 0 \end{aligned}$$

y

$$\begin{aligned} \frac{\partial SSE}{\partial \hat{\beta}_1} &= \frac{\partial [\sum_{i=1}^n [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)]^2]}{\partial \hat{\beta}_1} \\ &= - \sum_{i=1}^n 2[y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)] x_i \\ &= -2(\sum_{i=1}^n x_i y_i - \hat{\beta}_0 \sum_{i=1}^n x_i - \hat{\beta}_1 \sum_{i=1}^n x_i^2) = 0 \end{aligned}$$

Resolviendo este sistema de ecuaciones lineales llegamos a que:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.$$

El caso multiparamétrico es una extensión de esta idea [13]. ■

Para el caso de regresión simple donde las variables aleatorias  $X$  y  $Y$  se distribuyen normal bivariada, una prueba análoga para la significancia del modelo es  $H_0 : \rho = 0$  frente a  $H_a : \rho \neq 0$  donde  $\rho$  es el coeficiente de correlación de la distribución. El estimador de máxima verosimilitud de  $\rho$  está dado por el coeficiente de correlación muestral

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}.$$

Más general, para probar  $H_0 : \rho = \rho_0$  podemos utilizar una prueba  $Z$  en la cual

$$Z = \frac{(1/2) \ln\left(\frac{1+r}{1-r}\right) - (1/2) \ln\left(\frac{1+\rho_0}{1-\rho_0}\right)}{\frac{1}{\sqrt{n-3}}}.$$

En el análisis de varianza ANOVA (analysis of variance), Cuadro 1. Es de interés conocer cuanta de la variación de los datos ha sido explicada por la línea de regresión. Es un resumen de los datos necesarios para desarrollar la prueba  $F$  antes descrita. En la primera columna se denota la fuente de variación, después los grados de libertad, luego la suma de cuadrados, seguido de los cuadrados medios y por último la prueba que resulta del cociente de los cuadrados medios. En algunos casos se añade otra columna con el valor de significancia observado (p-value). Para la regresión simple es de la siguiente manera.

**Definición 2.** (*p-value*) Sea  $T(X)$  una estadística de prueba tal que los valores grandes de  $T$  dan evidencia a favor de la hipótesis alternativa. Para cada posible valor de  $X$ ,

$$x = (x_1, x_2, \dots, x_n) \in \mathcal{X},$$

Table 1. Análisis de varianza.

Fuente	df	SS	MS	F-value
Regresión	1	$\hat{\beta}_1 S_{xy}$	$\hat{\beta}_1 S_{xy}/1$	$\frac{\hat{\beta}_1 S_{xy}}{Se^2}$
Error	n-2	$S_{yy} - \hat{\beta}_1 S_{xy}$	$\frac{S_{yy} - \hat{\beta}_1 S_{xy}}{n-2} = Se^2$	
Total	n-1	$S_{yy}$		

se define el p-value como

$$p(x) = \max_{\theta \in \Theta_0} P[T(X) \leq t(x)]$$

donde  $t(x)$  es el valor de la estadística  $T(X)$  en  $x$ .

Otro estadístico útil para medir el ajuste del modelo es  $R^2$ . Este mide la proporción de la variación total sobre la media  $\bar{y}$  explicado por la regresión. Es un coeficiente de correlación estimado que toma valores entre 0 y 1. Mientras más cercano esté al 1 existe una mayor correlación, es decir, un mejor ajuste del modelo.

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = \frac{S_{xy}^2}{S_{xx} S_{yy}}.$$

La prueba estadística Goldfeld-Quandt es útil para probar, dado un valor de confianza  $\alpha$ , la homogeneidad en las varianzas (homocedasticidad) [2]. Supongamos dos submuestras de tamaño  $N_1$  con  $K_1$  parámetros y  $N_2$  con  $K_2$  parámetros. Los valores estimados de las varianzas de estas submuestras están dados por:

$$\hat{\sigma}_1^2 = \frac{SSE_1}{N_1 - K_1} \text{ y } \hat{\sigma}_2^2 = \frac{SSE_2}{N_2 - K_2}.$$

La prueba de hipótesis es de la forma

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \text{ vs } H_a : \frac{\sigma_1^2}{\sigma_2^2} \neq 1.$$

El estadístico consiste en un cociente de varianzas estimadas:

$$GQ = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \sim F_{(N_1 - K_1, N_2 - K_2)}.$$

De este modo, si no se rechaza  $H_0$ , entonces las varianzas de las submuestras son iguales, es decir, existe homocedasticidad [2].

### 3. Enfermedades Ocasionadas por la Contaminación Ambiental

Los cambios ecológicos influyen fuertemente en los patrones de enfermedad. La inestabilidad climática y el calentamiento progresivo están desempeñando un papel creciente en la emergencia de nuevas enfermedades infecciosas, así como en la reaparición y redistribución de olvidadas infecciones, además de las desigualdades económicas donde influyen fuertemente en los patrones de enfermedad [8].

**3.1. Cambio Climático y Enfermedades.** El clima de la tierra ha estado siempre en un estado de cambio. La climatología moderna reconoce que el cambio es una característica inherente y fundamental del clima. Los diferentes elementos climáticos como la temperatura, humedad, lluvia y viento presentan variaciones anuales, pero más importante, variaciones mayores que pueden medirse en función a décadas, siglos, milenios y millones de años. Estos cambios son totalmente naturales, sin embargo existen evidencias crecientes que en los últimos años una parte del calentamiento actual puede ser atribuible a las actividades humanas. Las mayores consecuencias del cambio climático son sobre el propio hombre, y paradójicamente es el ser humano uno de los principales causantes de este.

Esta aseveración puede constatar al observar, por ejemplo, que la sequía provoca al hombre escasez de agua, dificultad de higiene, disminución de la producción alimentaria, incendios forestales y la consecuente contaminación del aire con la resultante de efectos directos sobre la salud individual como aumento de la incidencia de enfermedades respiratorias y alérgicas, enfermedades diarreicas y desnutrición entre otras. La variación climática en cuanto al incremento de lluvias provoca inundaciones, contaminación del agua e incremento de criaderos de insectos vectores como el *Aedes aegypti*, que resulta vector de enfermedades como el dengue, fiebre amarilla o chikungunya [10].

**3.2. Muertes Infantiles y Contaminación.** Las condiciones insalubres del entorno, tales

como la contaminación del aire en espacios cerrados y en el exterior, la exposición al humo de tabaco ajeno, la insalubridad del agua, la falta de saneamiento y la higiene inadecuada, causan la muerte de 1,7 millones de niños menores de cinco años.

La Dra. Margaret Chan, Directora General de la OMS, señala que "la insalubridad del medio ambiente puede ser letal, especialmente para los niños pequeños, que son vulnerables a la contaminación del aire y el agua". La contaminación del aire en espacios cerrados y en el exterior y la exposición al humo de tabaco ajeno aumenta el riesgo que corren los bebés y niños en edad preescolar de contraer neumonías en su infancia y enfermedades respiratorias crónicas durante toda la vida. La contaminación del aire también puede aumentar el riesgo de sufrir cardiopatías, accidentes cerebro-vasculares y cáncer a lo largo del ciclo de vida [9].

Las cinco principales causas de muerte en los niños menores de cinco años guardan relación con el medio ambiente.

- 570 000 niños menores de cinco años fallecen como consecuencia de infecciones respiratorias (entre ellas las neumonías) causadas por la contaminación del aire en espacios cerrados y en el exterior y la exposición al humo de tabaco ajeno.

- 361 000 niños menores de cinco años fallecen a causa de enfermedades diarreicas debidas al acceso insuficiente a agua salubre, saneamiento e higiene.

- 270 000 niños fallecen en el transcurso del primer mes posterior al parto por diversas causas entre ellas la prematuridad que podrían prevenirse proporcionando acceso a agua potable y a instalaciones de saneamiento e higiene en los centros de salud, y reduciendo la contaminación del aire.

- 200 000 defunciones por paludismo de niños menores de cinco años podrían evitarse actuando sobre el medio ambiente, por ejemplo, reduciendo el número de criaderos de mosquitos o cubriendo los depósitos de agua.

- 200 000 niños menores de cinco años mueren a causa de lesiones o traumatismos involuntarios relacionados con el medio ambiente, como envenenamientos, caídas y

ahogamientos.

Paludismo: enfermedad causada por un parásito Plasmodium, el cual es transmitido por la picadura de un mosquito infectado.

#### 4. Aplicación de los Datos al Modelo de Regresión Lineal

Una limitante para el estudio es la inexistencia de datos que midan el nivel de contaminación ambiental tales como toneladas de residuos sólidos o de agentes químicos en el agua. Es por ello que sólo se trabaja con datos del Sistema de Monitoreo Atmosférico (SIMAT) que presenta información de las partículas suspendidas en el aire SO<sub>2</sub>, PM<sub>25</sub>, PM<sub>10</sub>, O<sub>3</sub>, CO y NO<sub>2</sub> para la Ciudad de México.

Por la parte económica, se obtienen datos del Instituto Nacional de Estadística y Geografía (INEGI) a través de la encuesta mensual de servicios y la encuesta nacional de empresas constructoras [7]. Los tres sectores que se presentan son

- Suministro de bienes y servicios de alojamiento temporal y de preparación de alimentos y bebidas.
- Construcción.
- Transporte.

Los dos primeros medidos en ingresos totales (moneda) y en número de personal ocupado por mes con límites superior e inferior (útiles para omitir la correlación temporal), error estándar y variación, para la Ciudad de México. La variable transporte contabiliza el ingreso registrado en casetas de carreteras y puentes de la República Mexicana.

En la Fig. 1 podemos encontrar un diagrama de caja y bigotes para los datos de ozono y monóxido de carbono, que resultan los de mayor interés. Para el ozono vemos un segundo cuantil cercano a la media y un dato extremo no muy lejano al máximo. Aún más centrado a la media, tenemos los datos del nivel monóxido de carbono, sin datos atípicos (extremos). En la Fig. 2 se encuentran las tres variables aleatorias económicas. Note que la variable con un intervalo más compacto es para los bienes y servicios. Para

ver todas las variables a detalle y el ajuste de escala revisar apéndice (8).

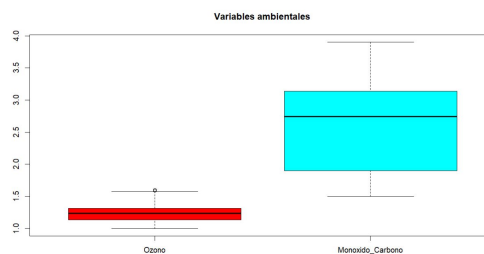


Fig. 1. Diagrama de caja; Ozono y monóxido de carbono.

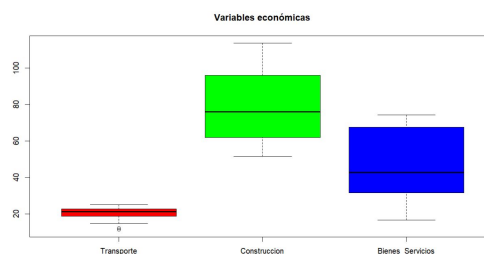


Fig. 2. Diagrama de caja; Transporte, construcción, bienes y servicios.

De las variables que presentan más tendencia en el tiempo, se encuentran bienes y servicios, construcción, transporte y nivel de partículas PM<sub>10</sub>. En los siguientes gráficos se muestran, donde los meses se contabilizan de 1 (enero 2019) a 36 (diciembre 2021). Se puede notar como se ve un cambio de tendencia a la mitad del lapso a causa de las restricciones gubernamentales por COVID19 en la Fig. 3, Fig. 5 y Fig. 6. Es periodo corresponde a mayo, junio y julio del 2020.

Los modelos de regresión lineal se hicieron sobre las variables ya descritas, utilizando las variables económicas como variables explicativas  $X_i$  y los niveles de contaminación como variables dependientes  $Y_i$ . Para cada variable  $Y_i$  se elaboró un modelo simple (sólo una variable explicativa) conjugando cada una de las 5 variables económicas, además de dos modelos

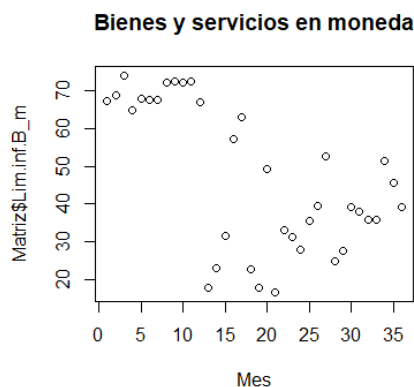


Fig. 3. Suministro de bienes y servicios en el tiempo.

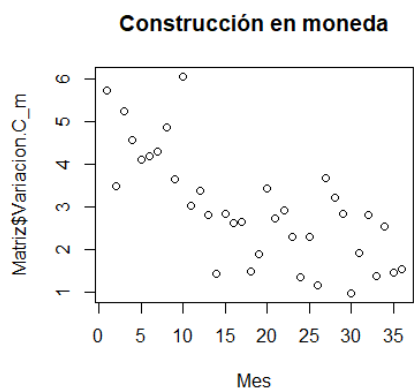


Fig. 4. Construcción en el tiempo.

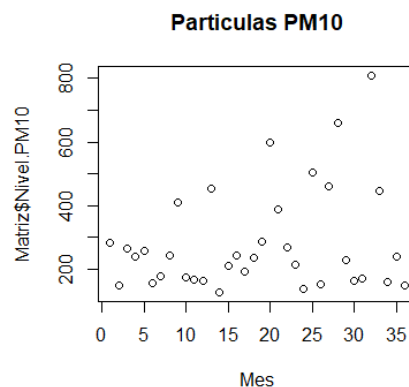


Fig. 5. Partículas PM10 en el tiempo.

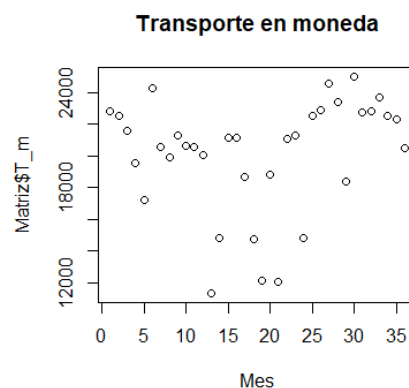


Fig. 6. Transporte en el tiempo.

multiparamétricos. El primero con todas las variables existentes para el caso de las partículas y con las variables económicas medidas en moneda; el segundo únicamente con la variable *CO* contra las variables económicas medidas en moneda.

## 5. Resultados

El mayor coeficiente de correlación al cuadrado  $R^2 = 0.436$  pertenece al modelo de la Fig. 7 con el nivel de monóxido de carbono (Nivel.CO) como variable *Y* y suministro de bienes y servicios en moneda como variable *X*.

Un segundo modelo con la mitad de desempeño y mejor al resto es del modelo de la Fig. 8 con un  $R^2 = 242$  que tiene como variable *Y* al nivel de monóxido de carbono y en el eje *X* a la cantidad monetaria por construcción.

Para los modelos multiparamétricos de la Fig. 9 y la Fig. 10, el primero no resultaba eficiente de ninguna manera y el segundo da una única variable significativa al  $\alpha = 0.001$  y resulta ser, nuevamente, la variable por bienes y servicios. De esta manera, los modelos multiparamétricos nos llevan de vuelta al modelo simple de la Fig. 7.

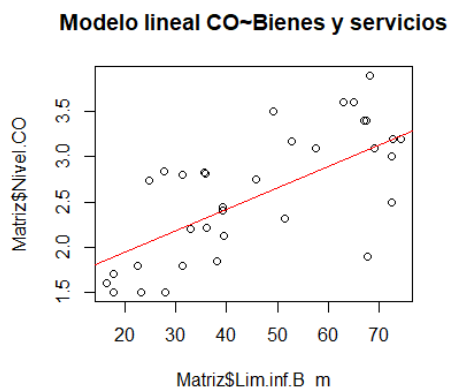


Fig. 7. Modelo m19.

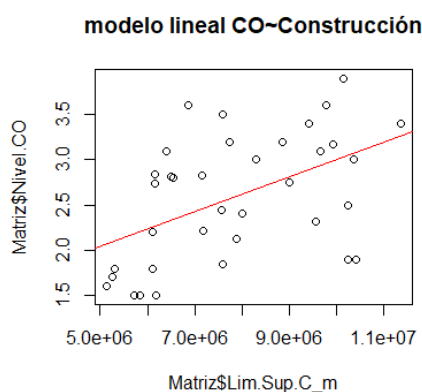


Fig. 8. modelo mc20.

```

Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.435e+02  3.523e+02   2.111  0.0427 *
T_m          1.994e-02  1.884e-02   1.058  0.2978
Lim.Sup.B_m  -2.678e+00  4.599e+00  -0.582  0.5645
Lim.Sup.C_m  -1.883e-05  6.162e-05  -0.306  0.7620
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 344.4 on 32 degrees of freedom
Multiple R-squared:  0.06278,    Adjusted R-squared:  -0.02508
F-statistic: 0.7145 on 3 and 32 DF,  p-value: 0.5506

```

Fig. 9. ANOVA del modelo CO + SO<sub>2</sub> + O<sub>3</sub> + PM<sub>10</sub> + NO<sub>2</sub> ~ Transporte + Bienes.Servicios + Construcción.

```

Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.005e+00  5.450e-01   1.845  0.07431 .
T_m          3.222e-05  2.915e-05   1.105  0.27725
Lim.Sup.B_m  2.433e-02  7.115e-03   3.419  0.00173 **
Lim.Sup.C_m -1.010e-07  9.533e-08  -1.059  0.29742
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5328 on 32 degrees of freedom
Multiple R-squared:  0.4604,    Adjusted R-squared:  0.4098
F-statistic:  9.1 on 3 and 32 DF,  p-value: 0.0001674

```

Fig. 10. ANOVA del modelo CO ~ Transporte + Bienes.Servicios + Construcción.

En una prueba de significancia para  $\beta_1$  en el modelo de la Fig. 7 resulta en un valor de significancia observado 0.0000116, es decir, se rechaza la hipótesis nula de no significancia  $H_0 : \beta_1 = 0$  para casi cualquier valor de  $\alpha$ . En particular, el modelo arroja un valor de  $\beta_1 = 0.0238$ , de esta manera, por cada unidad de cambio en el monto de bienes y servicios, el nivel de partículas en el aire de CO aumenta un 2.32 % (escala 1:1'000'000). Esta información se encuentra en la tabla ANOVA, en la Fig. 11.

```

Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.469200   0.235240   6.246 4.14e-07 ***
Lim.inf.B_m  0.023801   0.004638   5.132 1.16e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5283 on 34 degrees of freedom
Multiple R-squared:  0.4365,    Adjusted R-squared:  0.4199
F-statistic: 26.33 on 1 and 34 DF,  p-value: 1.16e-05

```

Fig. 11. ANOVA del modelo CO ~ Bienes.Servicios.

Una vez seleccionado el mejor modelo, otro supuesto importante a evaluar es la normalidad de los errores con media cero y varianza constante  $\varepsilon \sim N(0, \sigma^2)$ . Para ello, nos valemos del gráfico de la Fig. 12 y la prueba Goldfeld-Quant (GQ) en la Fig. 13. Note que los valores se presentan de manera aleatoria al rededor del cero y en la prueba GQ nos da una significancia observada  $\alpha = 0.3078$  suficientemente alta para no rechazar la hipótesis nula de varianzas homogéneas en los errores. Esto nos dice que nuestro modelo, a demás de ser eficiente, es consistente.



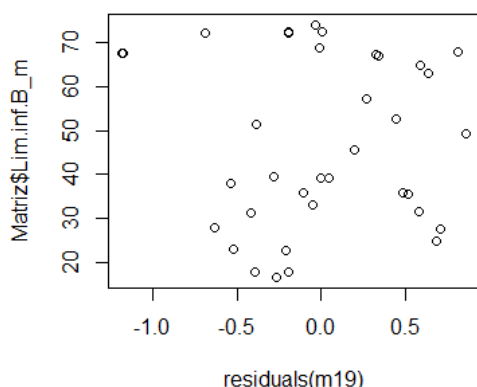


Fig. 12. Valores residuales del modelo  $CO \sim \text{Bienes.Servicios}$ .

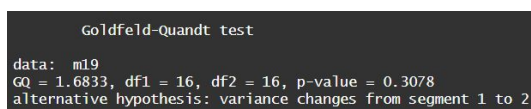


Fig. 13. Prueba estadística Goldfeld-Quandt.

## 6. Conclusiones

Después de buscar la relación entre la contaminación y el nivel de actividad económica, a pesar de las limitantes de acceso (o existencia) de datos para tener indicadores cuantitativos del desarrollo ambiental, los modelos de regresión lineal no arrojan correlación entre la actividad de Transporte con ninguna partícula suspendida en el aire, posiblemente debido a que no una variable sólo para la Ciudad de México. Para el caso de la actividad del sector de construcción, sólo se encontró una ligera relación con los niveles de monóxido de carbono ( $CO$ ). El modelo más eficiente resulta en una correlación entre la actividad económica por suministro de bienes y servicios al sector de hospedaje y producción de alimento con la presencia de monóxido de carbono para la Ciudad de México. Dicho modelo arroja que por cada

millón de pesos mexicanos, los niveles de ésta partícula aumentarán un 2.32 %. No se encontró relación de esta variable económica con ninguna otra partícula medida por el SIMAT. Esto nos conduce de manera inmediata a pensar que para reducir los niveles de contaminación tiene como consecuencia la disminución de los niveles de actividad económica, sin embargo, puede ser más eficiente entender a los procesos económicos como metodologías con posible modificación y no como métodos a priori.

Para tener mayor capacidad predictiva sobre la contaminación es necesario disponer de información cuantitativa fiable y suficiente. Es por ello que se sugiere a la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) o a las dependencias pertinentes impulsar el monitoreo ambiental no sólo de la calidad del aire, sino también de otros agentes, como la contaminación del agua, del suelo, contaminación acústica o contaminación lumínica [12].

## 7. Agradecimientos

Agradezco al cuerpo evaluador de la Semana Internacional de la Probabilidad y la Estadística de la Benemérita Universidad Autónoma de Puebla por la revisión y aceptación de esta investigación. A la doctora Hortensia J. Reyes y al doctor Fernando Velasco Luna por brindarme los conocimientos necesarios y el asesoramiento para el desarrollo de este trabajo. Por último, agradecemos al Instituto Nacional de Estadística y Geografía, así como al Sistema de Monitoreo Atmosférico por recolectar, ordenar y ofrecer datos a todo público.

## 8. Apéndice

**8.1. Variables del modelo.** En la Fig. 14 se muestran las variables presentes en el modelo, incluyendo las no significativas. Estas son:

- $T_m$
- $Lim.inf.B_p$
- $Lim.Sup.B_p$
- $Lim.inf.B_m$
- $Lim.Sup.B_m$



- Lim.inf.C.p
- Lim.Sup.C.p
- Lim.inf.C.m
- Lim.Sup.C.m
- Nivel.SO2
- Nivel.PM25
- Nivel.PM10
- Nivel.O3
- Nivel.CO
- Nivel.NO2

Periodo					
Length:36	Min. :11345	Min. :28.21	Min. :117.2	Min. :16.58	
Class :character	1st Qu.:11867	1st Qu.:32.35	1st Qu.:121.8	1st Qu.:31.44	
Mode :character	Median :21122	Median :37.30	Median :124.2	Median :47.62	
	Mean :20131	Mean :50.46	Mean :126.0	Mean :47.03	
	3rd Qu.:22603	3rd Qu.:78.17	3rd Qu.:129.2	3rd Qu.:67.48	
	Max. :25021	Max. :80.58	Max. :141.3	Max. :74.17	
Lim.Sup.B.m	Lim.inf.C.p	Lim.Sup.C.p	Lim.inf.C.m	Lim.Sup.C.m	
Min. :37.03	Min. :29067	Min. :42739	Min. :4624343	Min. :5146813	
1st Qu.:52.98	1st Qu.:37039	1st Qu.:46843	1st Qu.:5577764	1st Qu.:6180559	
Median :64.92	Median :41023	Median :50234	Median :6830346	Median :7586444	
Mean :70.95	Mean :54181	Mean :57528	Mean :6937552	Mean :7838907	
3rd Qu.:94.30	3rd Qu.:44045	3rd Qu.:69225	3rd Qu.:8423071	3rd Qu.:9574508	
Max. :106.29	Max. :54060	Max. :99650	Max. :9062383	Max. :11356139	
Nivel.SO2	Nivel.PM25	Nivel.PM10	Nivel.O3	Nivel.CO	Nivel.NO2
Min. :44.0	Min. :54.00	Min. :128.0	Min. :100.0	Min. :1.500	Min. :128.0
1st Qu.:71.0	1st Qu.:66.00	1st Qu.:168.2	1st Qu.:113.8	1st Qu.:1.900	1st Qu.:168.2
Median :104.0	Median :81.50	Median :237.5	Median :123.5	Median :2.745	Median :237.5
Mean :113.2	Mean :83.25	Mean :283.7	Mean :124.2	Mean :2.589	Mean :283.7
3rd Qu.:146.5	3rd Qu.:86.00	3rd Qu.:312.8	3rd Qu.:131.2	3rd Qu.:3.118	3rd Qu.:312.8
Max. :257.0	Max. :162.00	Max. :809.0	Max. :160.0	Max. :3.900	Max. :809.0

Fig. 14. Variables existentes del modelo

Para la visualización de los diagramas de cajas se hicieron los siguientes ajustes de escala:

- Transporte/1000
- Construcción/100000
- Ozono/100

**8.2. Código en R..** Puede acceder al archivo en R con el código utilizado y a las bases de datos en formato csv y xls en el siguiente enlace: <https://mega.nz/folder/CQgxHbKb#jwjlWA9toSMwyNqdFr5prQ>

## Referencias

- [1] De Gregorio, J. Macroeconomía: teoría y políticas. Pearson eBooks, 2007.
- [2] Hill, R. C., Griffiths, W. E., & Lim, G. C. Principles of Econometrics. John Wiley & Sons, 2018.
- [3] Draper, N. R., & Smith, H. Applied Regression Analysis. John Wiley & Sons, 2014.
- [4] Wackerly, D., Mendenhall, W. M., & Scheaffer, R. L. Estadística matemática con aplicaciones, 2008.
- [5] Gil, S. I. Métodos estadísticos: un enfoque interdisciplinario, 1984.
- [6] Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining. Introducción al Análisis de Regresión Lineal(1.a ed.), 2002.
- [7] INEGI (2023). Consulta. Instituto Nacional de Estadística e informática, 2022. [https://www.inegi.org.mx/app/tabulados/pxw ebclient/default.html?pxq=BISE\\_BISE.T3x P6x1G\\_230614101923\\_c8ab0272-bbc7-4a4b-b44c-76e9b5a38c9a](https://www.inegi.org.mx/app/tabulados/pxw ebclient/default.html?pxq=BISE_BISE.T3x P6x1G_230614101923_c8ab0272-bbc7-4a4b-b44c-76e9b5a38c9a)
- [8] ONU. (2023). Las consecuencias de la contaminación ambiental: 1,7 millones de defunciones infantiles anuales, según la OMS. Organización de las Naciones Unidas, 2017. [https://www.paho.org/hq/index.php?option=com\\_content&view=article&id=13042:cost-of-polluted-environment-1-7-million-child-deaths-year&Itemid=0&lang=es#gsc.tab=0](https://www.paho.org/hq/index.php?option=com_content&view=article&id=13042:cost-of-polluted-environment-1-7-million-child-deaths-year&Itemid=0&lang=es#gsc.tab=0)
- [9] Arbo, Antonio, Sanabria, Gabriela, & Martínez, Celia. (2022). Influencia del Cambio Climático en las Enfermedades Transmitidas por Vectores.. Revista del Instituto de Medicina Tropical, 17(2), 23-36, 2000. 2022.<https://doi.org/10.18004/imt/2022.17.2.4>
- [10] WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: World Health Organization. Organización Mundial de la Salud, 2021. <https://www.who.int/publications/i/item/9789240034228>
- [11] Coronavirus y cambio climático: Por qué la pandemia no es realmente tan buena para el medio ambiente. (2018,

11 mayo). BBC News Mundo, 2023.  
<https://www.bbc.com/mundo/noticias-52596472>

[12] La contaminación del medio ambiente, causas, consecuencias y soluciones. (2017). Aupece.org, 2023.  
<https://aupeace.org/contaminacion-ambiental/>

[13] Vázquez A. J, Naranjo A, L., Fuentes G. R., Chávez C. M., Inferencia Estadística, Facultad de Ciencias, UNAM, 2019.