



CALIBRACIÓN Y CUANTIFICACIÓN DEL ERROR DE MEDICIÓN DE UN SENSOR DE BAJO COSTO

Alfaro Rivera, Valentina [2211702]
Perez Rojas, Nathalia Alexandra [2200023]
Universidad Industrial de Santander
Escuela de física
17 de diciembre de 2021

Índice

1. Introducción:	1
1.1. Sobre el material particulado PM_{25}	1
1.2. Sensores de bajo costo	2
2. Metodología	2
3. Resultados	2
3.1. Preselección de datos	2
3.2. Primer método de calibración	2
3.3. Segundo método de calibración	3
3.4. Verificación	4
4. Análisis de resultados y conclusiones	4

Resumen

Como un acercamiento al concepto de métrica mediante la calibración, se cuantificó el error de un sensor de $PM_{2,5}$ de bajo costo ubicado en el colegio Normal Superior (de dónde? ni idea) con ayuda de medidas de referencia del mismo lugar. Este error fue obtenido a partir de la distancia euclidiana para medidas puntuales del sensor para diferentes ventanas de tiempo de 3, 8 y 12 horas para tres días consecutivos escogidos aleatoriamente por mes entre el periodo de prueba de Abril hasta Agosto del año 2019. A partir del error (descripción de resultado del error, quizá de tipo zero) se dispuso una aproximación lineal para la calibración del sensor. Finalmente a modo de verificación se dispusieron dos métodos con los que se predijo la cantidad de $PM_{2,5}$ para un día aleatorio por en el mismo periodo de prueba, diferentes a los anteriormente seleccionados, de lo que se obtuvo una disminución considerable del error cuantificado inicialmente (descripción o valores(?) de ese error).

Palabras clave: error de medición, verificación, sensor, calibración.

1. Introducción:

1.1. Sobre el material particulado PM_{25}

$PM_{2,5}$ se refiere a la materia particulada (PM) atmosférica que tiene un diámetro de menos de 2.5 micrómetros, que es aproximadamente el 3 % del diámetro de un cabello humano.[1].”

Así como el auge en investigaciones sobre las consecuencias de las nanopartículas está poco a poco tomando lugar en seminarios y coloquios, también la $PM_{2,5}$ se estuvo dando su propio lugar. Ambas generan un problema ambiental pues los nano polímeros,

provenientes principalmente del plástico en las basuras, pueden prevalecer en los mares así como la materia particulada, provenientes de muchas más fuentes como plantas de energía, vehículos de motor, aviones, quema de madera residencial, incendios forestales, quemas agrícolas, erupciones volcánicas y tormentas de polvo, entre otros, que también pueden prevalecer en el aire, ambas pueden ser inhaladas o digeridas/bebidas por los animales y el mismo ser

humano.

Así como ya se han investigado indicios de las consecuencias de las nanonapartículas (ver por ejemplo [2]), son incluso más los problemas expuestos que generan las PM como por ejemplo el estudio "Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution" [3] donde se sugiere que la exposición a largo plazo a $PM_{2,5}$ puede conducir a depósitos de placa en las arterias, lo que causa inflamación vascular y endurecimiento de las arterias que eventualmente puede conducir a un ataque cardíaco y un derrame cerebral o también, la Asociación Estadounidense del Corazón también advirtió sobre el impacto de las $PM_{2,5}$ en la salud y la mortalidad del corazón, Sin embargo, es bueno saber que las autoridades de salud pública han tomado cartas en el asunto y por ello se tiene mayor información al respecto: En un día muy claro y no brumoso, la concentración de $PM_{2,5}$ puede ser tan baja como $5 \text{ g} / \text{m}^3$ o menos. La concentración de 24 horas de $PM_{2,5}$ se considera insalubre cuando se eleva por encima de $35,4 \text{ g} / \text{m}^3$.

1.2. Sensores de bajo costo

Los sensores de bajo costo tienen el fin de monitorear la calidad del aire (conocer la cantidad de $PM_{2,5}$ y PM_{10}) con una frecuencia diaria y por horas y así poder implementar las necesarias políticas públicas para mejorar la calidad de este, sin embargo, estos sensores poseen ciertos inconvenientes ya que no son muy precisos a la hora de arrojar medidas (por datos faltantes) y su baja resolución espaciotemporal, por ello, es necesario realizar después de cierto tiempo de uso la calibración de estos sensores. [?]

Para realizar la calibración de estos, se toma otros equipos de referencia los cuales arrojan los datos para realizar la comparación con los datos de los sensores de bajo costo, los primeros meses se puede observar que la distancia entre los puntos es muy mínima, lo que significa que el sensor esta funcionando bien, no obstante, después de cierta cantidad de tiempo, estos valores empiezan a ser más imprecisos y a tener una mayor distancia, por tanto, a cada valor hay que aplicarle su respectiva calibración.

2. Metodología

Con las medidas de $PM_{2,5}$ tomadas por un sensor de bajo costo [5] durante el periodo de Noviembre de 2018 a Agosto de 2019, se preseleccionó otro periodo más corto por la falta de continuidad en los datos entre Noviembre de 2018 hasta Abril de 2019, siendo el subperiodo a trabajar de Mayo a Agosto de 2019.

Para poder manipular la información se necesita de un dato en común; el tiempo y como las medidas

del sensor de bajo costo no fueron tomadas en la misma hora que la información del patrón de referencia se decidió calibrar por dos métodos; por interpolación y manualmente, donde uno toma en cuenta dicha diferencia en el tiempo y la otra no, respectivamente.

En la calibración por ambos métodos se trabajan los datos para ventanas de tiempo móviles comunes en las que se promedian los datos contenidos. Dichas ventanas estarán dadas por la siguiente suceción:

$$a_0 = x_0 \quad , \quad b_0 = x_0 + \delta$$

$$\epsilon = \frac{b_n - a_n}{2}$$

$$a_{n+1} = a_n + \epsilon \quad , \quad b_{n+1} = b_n + \epsilon$$

Note que se requiere únicamente información (x_0) de uno de los dos conjuntos (medidas del sensor $f(\bar{x}_0)$ o del patrón $f(x_0)$), cosa que bajo ambos métodos pueden ser interpretados como una ventana común entre los valores del patrón y los tomados por el sensor.

Una vez calibrado el sensor se procede a verificar con el fin de notar una disminución significativa del error y así poder determinar el alcance del modelo.

3. Resultados

3.1. Preselección de datos

Bajo la primera preselección con la mejor continuidad en los datos de Abril a Agosto de 2019, lo siguiente se aplicó con ayuda de Python:

- Identificación de las horas y fechas a partir de una sola unidad de medida de minutos. Esto con el fin de poder obtener una mejor manipulación de los datos a la hora de identificarlos así tal que el tiempo inicial empieza el primero de Abril a las 00:00:00 a.m. finalizando el 31 de Agosto a las 23:59:59 p.m.
- Borrados de información con algún contenido nulo.
- Promediado de los valores tanto medidos por el sensor como los obtenidos del patrón para tres ventanas móviles diferentes, cada una con un δ de 5, 11 y 17 horas.

3.2. Primer método de calibración

A los datos obtenidos a partir de las ventanas móviles, se le aplica el una regresión lineal, que consiste en encontrar varias rectas que permitan definir un α y un β a partir de las respectivas operaciones. α es la diferencia entre las pendientes cuyo valor es 0.56 aproximadamente y el β es la diferencia entre

los puntos de corte de dos rectas encontradas anteriormente de 3.78 aproximadamente. Con los valores de estos dos operadores se pueden calibrar los datos experimentales. Seguido de esto, se grafican los puntos y se hace una comparación para la cuantificación del error.

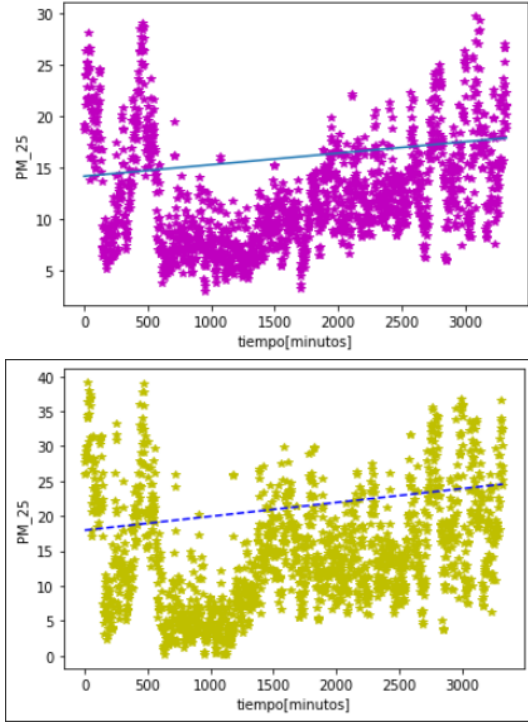


Figura 1: Puntos teoricos (morados) y experimentales (amarillo) con sus respectivas rectas.

Aplicando el respectivo α y β se obtiene:

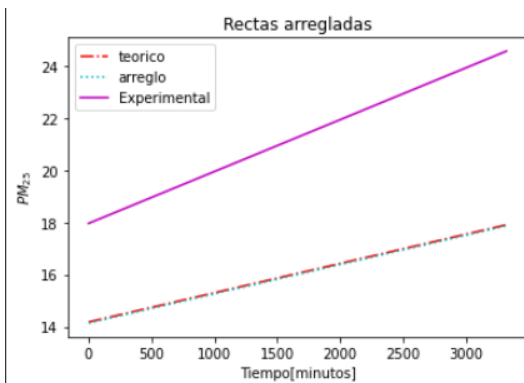


Figura 2: Comparación entre las rectas de las rectas de datos teóricos, experimentales y calibrados.

Ahora, se cuentan los datos por fuera de diferentes valores de la tolerancia aceptada con los que se generan diferentes errores tabulados en la siguiente tabla:

Tolerancia	Método 1	
	Desvíos	error %
3	503	30.24
4	330	19.88
5	223	13.40
6	136	8.18

Tabla 1: Errores porcentuales obtenidos para diferentes tolerancias.

3.3. Segundo método de calibración

Con las ventanas móviles previamente definidas se desea aplicar un segundo método de calibración que consiste en un barrido de parámetros para corregir los errores de tipo zero y multiplicación teniendo en cuenta diferentes tolerancias para cada barrido de modo que el usuario al final del programa pueda definir para una determinada tolerancia, los mejores parámetros a la hora de realizar la calibración.

Previo a la aplicación del código se cuantifica el error como el error absoluto entre la medida del sensor y los valores de referencia, que es lo mismo a obtener la distancia euclidiana entre cada par de datos correspondientes a la misma toma respecto al tiempo. Esto es porque dicha distancia será clave en el código a la hora de obtener la mejor correlación y entonces se aplica el código del que se obtienen los mejores parámetros para la calibración en orden de multiplicación y luego suma. Estos se muestran en las siguientes tablas:

Tolerancia	Ajuste de multiplicación		
	Parámetro	Desvíos	error %
5	0,7	318	19,12
6	0,7	218	13,11
7	0,6	141	8,48
8	0,6	92	5,53
9	0,6	55	3,31
10	0,6	35	2,10

Tabla 2: Errores porcentuales obtenidos para el ajuste de error por multiplicación.

Una vez calibrados dos datos con el mejor parámetro de 0.6 escogido según la tabla anterior, se grafican los puntos en la siguiente figura, notando que hace falta un ajuste por error de tipo zero.

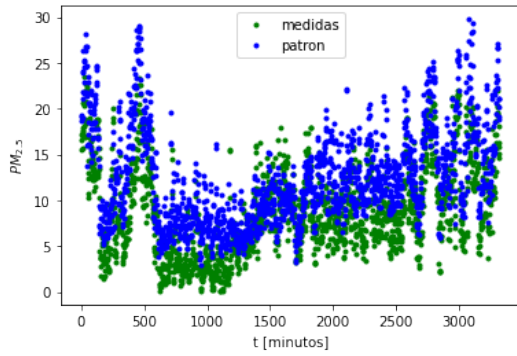


Figura 3: Comparación entre los valores del patrón y los medidos con un primer ajuste.

Tolerancia	Ajuste de error zero		
	Parámetro	Desvíos	error %
3	4	517	31,09
4	3	344	20,69
5	4	223	13,41
6	3	139	8,36

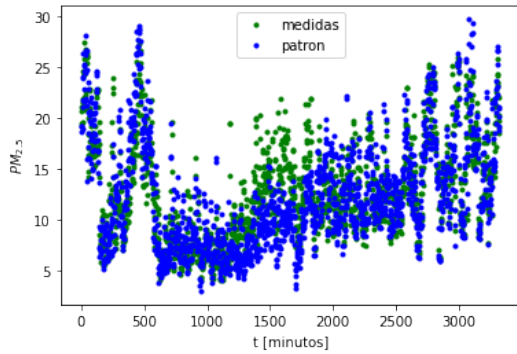
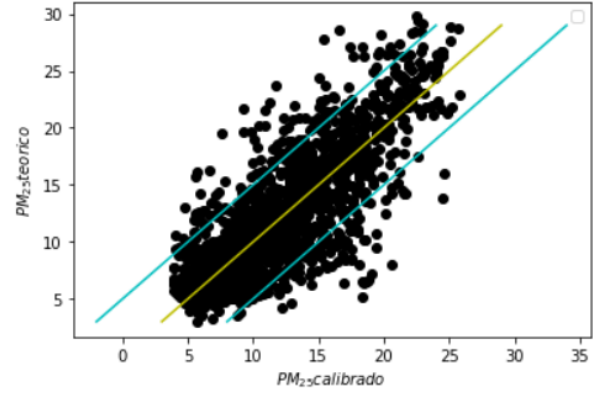


Figura 4: Comparación entre los valores del patrón y los medidos calibrados.

3.4. Verificación

Sea, $f(\epsilon) = \alpha \bar{f}(\epsilon)$ la relación que se espera una vez hecha la calibración, es obvio notar que también α tendría que ser 1, entonces para concluir los cálculos se grafica dicha relación y con ayuda de Python

y la función `.curvefit()`, se grafica $f(\epsilon_i)$ contra $\bar{f}(\epsilon_i)$ para ϵ_i los valores en cada uno de los rangos de las ventanas móviles comparando los dos métodos presentados anteriormente y así determinar también el conjunto mínimo que se debería tener en cuenta para una calibración del sensor.



b)

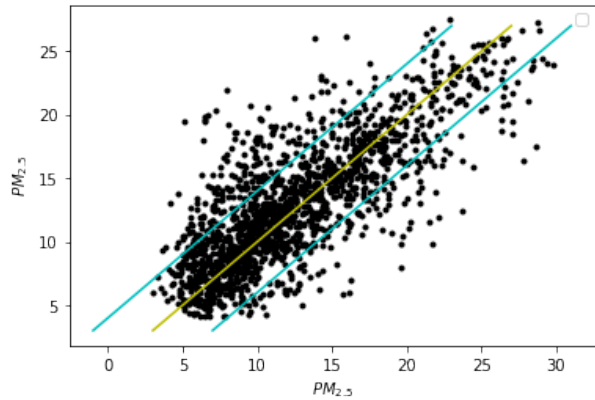


Figura 5: Vista inmediata de la comparación entre el patrón y las medidas calibradas del sensor para una ventana móvil con rango de 5 horas por dos métodos; a) de regresión lineal y b) manual.

Finalmente se obtiene los α de 0.99 y 0.93 para la calibración aplicada en el primer y segundo método respectivamente con tolerancia de 4 $PM_{2,5}$

4. Análisis de resultados y conclusiones

Como primera medida de calibración se analiza respecto al tiempo la mejor forma de realizar una calibración para el sensor, las calibraciones necesitan entonces medir lo mismo en el mismo tiempo para poder hacer una comparación. El método por regresión lineal tiene en cuenta este facto en la variación de los tiempos puesto que la recta encontrada representa la relación de los datos respecto en exactamente cada instante de tiempo mientras que el segundo método se concentra en la comparación por cada par de datos sin importar el desfase en minutos o segundos que posean respecto a la hora exacta en la que se tiene el patrón.

Al finalizar las calibraciones por ambos métodos se encuentra que este desfase de tiempo no representa un agravio potencial dado que la diferencia de los errores por desvíos no supera el 1%, siendo este de: 0,81%.

Además, cabe resaltar que de esta forma se puede interpretar al tiempo con un factor de identificación en los datos, sin influencia mayor a la relación que este tiempo presenta respecto a los factores medio ambientales que pueden tomar partido a la hora en que el sensor hace la toma de su medida. En este orden de ideas, la distancia euclidiana en el segundo método, es fácilmente reemplazable en el desarrollo del código por un valor absoluto ya que de los pares ordenados (x_i, y_i) con x_i representante de los tiempos y y_i de la medida del material particulado, solo se tiene en cuenta el factor y_i , no se tiene en cuenta la variable del tiempo.

Durante el proceso de calibración al poner en práctica la transformación de las medidas a partir de operaciones básicas, se puede notar que la idea de distancia ayuda además de ayudar a cuantificar el error, en este caso por la cantidad de desvíos o mejor, por la cantidad de valores que se salen del rango de tolerancia, a preservar una forma de comparación para el proceso de calibración.

Referencias

- [1] C. Linares, J. Díaz: “Las PM_{2,5} y su impacto sobre la salud. El caso de la ciudad de Madrid”. Ecosostenible. 2008;35:32-37. <https://www.ecologistasenaccion.org/17842/que-son-las-pm25-y-como-afectan-a-nuestra-salud/>
- [2] Pope CA 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA. 2002 Mar 6;287(9):1132-41. doi: 10.1001/jama.287.9.1132. PMID: 11879110; PMCID: PMC4037163.
- [3] Brook RD, Rajagopalan S, Pope CA 3rd, Brook JR, Bhatnagar A, Diez-Roux AV, Holguin F, Hong Y, Luepker RV, Mittleman MA, Peters A, Siscovick D, Smith SC Jr, Whitsel L, Kaufman JD; American Heart Association Council on Epidemiology and Prevention, Council on the Kidney in Cardiovascular Disease, and Council on Nutrition, Physical Activity and Metabolism. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation. 2010 Jun 1;121(21):2331-78. doi: 10.1161/CIR.0b013e3181dbece1. Epub 2010 May 10. PMID: 20458016.
- [4] Perez. N Alfaro. V 2021. CódigoAsignación2. <https://github.com/hoanhhan101/ultimate-go>. (2021). <https://github.com/NathaliaAPR/Repositorio-de-tareas/blob/6c07053d0e2362549944f822db0c83885975561d/codigos/asignaciC3%B3n2.ipynb>
- [5] Roncancio, L. (2019). Evaluación del desempeño de sensores de bajo costo como complemento para el monitoreo de la calidad del aire de Bogotá y como herramienta para la determinación del grado de exposición de una población caso de estudio. *Universidad nacional de Colombia* <https://repositorio.unal.edu.co/handle/unal/77273>