

Fernández S, Galavís M, Velasco G

La interpolación como herramienta de aproximación a variables meteorológicas

(12 Mayo 2021)

Resumen— Se presenta un ejemplo de aplicación de interpolación en el cálculo de variables meteorológicas relacionadas con dos estaciones ubicadas en el territorio Brasileño. Para ello se emplearon los métodos de spline cúbico y lineal y el cálculo de los errores máximos, mínimos, absolutos, medios y el emc con el fin de precisar su estimación.

Palabras clave— *Clima, Interpolación Lineal, Spline, Temperatura, Brasil, Errores.*

I. INTRODUCCIÓN

Tener la mayor certidumbre posible respecto al comportamiento del clima es una de las más antiguas inquietudes que ha acompañado a la humanidad y sobre la cual aún hoy en día se siguen construyendo herramientas y desarrollando una ciencia como lo es la climatología. Lo anterior, no es para menos, pues la importancia de estos pronósticos radica en la planificación óptima de actividades humanas que van desde el cubrimiento de necesidades básicas como la siembra de alimentos y la construcción de viviendas, hasta la celebración de un evento de cualquier índole[1].

El presente documento proporciona un ejemplo de aplicación del subcampo del análisis numérico: interpolación, en la estimación en primera instancia de una variable climática y en la realización en segunda instancia, de una proyección de datos de una estación a partir de otra estación. Lo anterior, se hace posible por el conjunto de valores indexados por tiempo y espacio de 15 estaciones pluviométricas de Brasil en 2013, un mapa de estas y sus respectivas coordenadas.

II. ACERCAMIENTO AL PROBLEMA

Para el primer problema, en cuanto a la selección de la estación y variable, se tuvo en cuenta como criterios la cantidad de información disponible de dicha base y el cambio entre los datos de sus variables climatológicas buscando una que presentará cambios pequeños entre estos. El resultado de esta evaluación fue la selección de la temperatura interna de la estación de Quixada. Para el segundo problema, la

selección de las bases se realizó conforme a la visualización de un mapa de las estaciones y una verificación posterior con las coordenadas proporcionadas. Este procedimiento tuvo como resultado la selección de la base de Quixeramobim y Quixada.

En relación al manejo de los datos proporcionados, para ambos problemas se decidió usar interpolación. La interpolación es una herramienta que permite llevar a cabo una aproximación a $f(x)$ para un valor de x en específico desde la construcción de una curva o superficie que une los puntos sobre los cuales se hizo la medición y se tiene información de su valor[2]. En particular, en este trabajo se proponen los métodos de interpolación lineal y spline.

Por un lado, la interpolación lineal es reconocida por ser el método más simple usado hoy en día. Este método consiste en considerar como función de interpolación la recta que une dos puntos. Conocida la ecuación de la recta, se facilita ubicar el punto desconocido. Para un caso en el que se lleguen a tener más de dos puntos, estas rectas se concatenan.

$$y = y_0 + (x - x_0) \frac{y_1 - y_0}{x_1 - x_0}, \quad (1)$$

(1) Función correspondiente a la interpolación lineal de dos puntos (x_0, y_0) y (x_1, y_1)

Para la interpolación lineal se usó el módulo de interpolate ofrecido por la librería científica de scipy, en específico, el paquete `interp1d` que recibe como parámetros dos arreglos x y y del mismo tamaño que son utilizados para aproximarse a alguna función. Esta última es la que se devuelve para encontrar nuevos valores de y a partir de nuevos valores de x .

Por otro lado, los splines son ampliamente reconocidos por proporcionar resultados próximos a los reales a partir del uso de polinomios de bajo grado evitando de esta manera oscilaciones indeseadas. El spline cúbico en especial es el spline más empleado gracias al excelente ajuste a los puntos tabulados; esto mediante el uso de cuatro constantes, lo cual le concede flexibilidad para asegurar que la primera y segunda derivada sean continuas permitiendo una unión suave entre sus segmentos[3].

$$S_i(x) = \alpha_i + \beta_i(x - x_i) + \gamma_i(x - x_i)^2 + \delta_i(x - x_i)^3 \quad (2)$$

(2) Donde $S_i(x)$ es la función cúbica para el segmento,

x_i es el primer punto del segmento ,
 x es el punto de evaluación de la función,
 $\alpha_i, \beta_i, \gamma_i, \delta_i$ son los coeficientes de la función

Para la interpolación por spline cúbico, se usaron los paquetes `splrep` y `splev`. Con `splrep` se calcula la matriz de tuplas que contiene los coeficientes del spline. `Splrep` maneja por defecto el spline de grado tres por lo que para este ejercicio sólo recibe dos arreglos x y y que representan los datos de la curva $f(x)$. Una vez se obtienen los coeficientes, se utiliza `splev` que recibe un arreglo de puntos, para los que se quiere calcular el spline, y la matriz de coeficientes calculado con `splrep`.

Determinamos no usar la interpolación polinómica por el volumen de los datos debido a que esta usa un solo polinomio de alto grado para la totalidad de estos, lo cual se presta para grandes oscilaciones no siendo útil para el propósito del ejercicio.

III.SOLUCIÓN PROBLEMA 1: INTERPOLACIÓN DE LOS DATOS DE LA ESTACIÓN DE QUIXADA

Para el desarrollo de este punto de los 720 datos que se tenía de la temperatura interna de la base de Quixada, se tomó el 80% (576) para realizar los métodos de interpolación. A continuación se presenta el resultado de la interpolación lineal y spline cúbico:

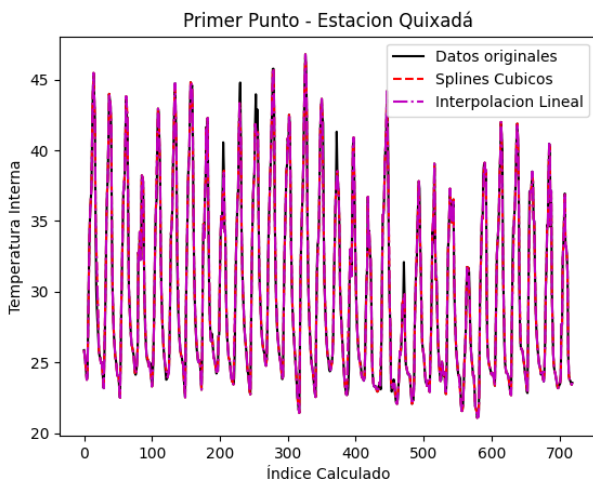


Figura 1. Gráfica interpolación temperatura interna Quixada

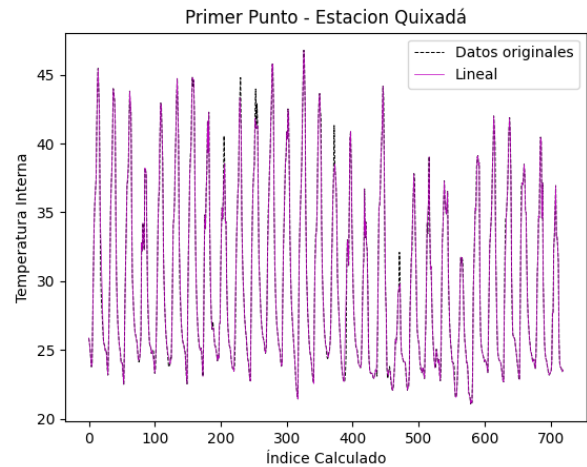


Figura 2. Gráfica interpolación lineal vs datos originales para la temperatura interna Quixada

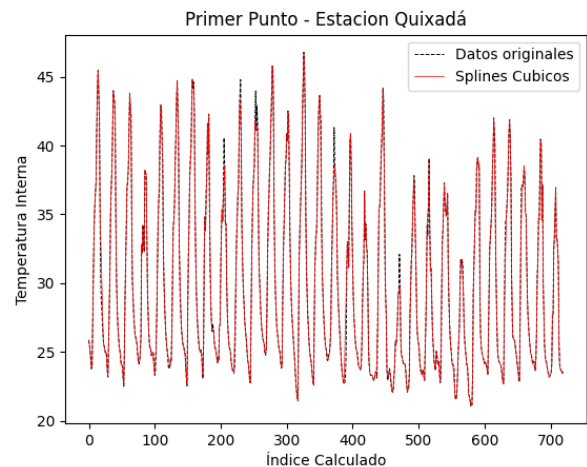


Figura 3. Gráfica interpolación spline cúbico vs datos originales para la temperatura interna Quixada

Al ser este subcampo del análisis numérico una herramienta para estimar valores, se presentan errores en los cálculos de estos con respecto a los originales. Para verificar la precisión de los resultados arrojados, se realizó el cálculo de los errores máximo, mínimo, medio, absoluto y error medio cuadrático, empleando el 20% restante de la información que se disponía de la temperatura interna para calcular en estos puntos la interpolación y de esta manera comparar la diferencia entre la aproximación producto de esta operación y el valor conocido de los datos de la estación. Los resultados de los errores obtenidos para el primer punto se muestran en las siguientes figuras:

```

ERROR MAXIMO SPLINE 4.721748068471605
ERROR MINIMO SPLINE 0.0024532923507791793
ERROR MEDIA SPLINE 0.6172843894536232
ERROR ABSOLUTO SPLINE 0.5185477571663826
ERROR MEDIO CUADRADO SPLINE 0.975598316500103

```

Figura 4. Valores error interpolación spline cúbico temperatura interna Quixada

```

ERROR MAXIMO LINEAL 4.34
ERROR MINIMO LINEAL 0.0
ERROR MEDIA LINEAL 0.7310393518518521
ERROR ABSOLUTO LINEAL 0.6863062950102885
ERROR MEDIO CUADRADO LINEAL 1.1742162074229312

```

Figura 5. Valores error interpolación lineal temperatura interna Quixada

IV. Análisis solución problema 1:

Analizando las figuras que muestran la comparación de los datos originales con cada método de interpolación empleado, se observa que en varios puntos de la gráfica la línea correspondiente a la interpolación lineal es diferente a la línea dada por los datos originales, es allí donde se genera el error dado por este método de interpolación. El comportamiento de este error en la gráfica presenta el mismo efecto empleando el método de Splines cúbicos.

Con base en la observación planteada en el párrafo anterior, determinar cuál método presenta una mejor estimación para este problema resulta difícil utilizando únicamente las gráficas por lo que es necesario acudir al cálculo de los errores.

Al momento de calcular el error mínimo del spline, este resulta ser mayor que el error mínimo dado por el método lineal, en el cual el valor es 0, lo que indica que existen datos hallados por medio de la interpolación lineal que son completamente iguales a los datos originales; por el contrario, en el spline cúbico no existe ningún dato hallado por medio de la interpolación que coincida a la perfección con el dato original. Esta misma situación se repite con el error máximo donde el de la interpolación lineal es menor al de la interpolación por spline cúbico.

Sin embargo, esta conducta se presenta solo en estos dos errores pues, tanto el error medio como el error absoluto y el error medio cuadrático, siendo estos últimos errores frecuentemente usados en la comparación de pronósticos, demuestran que el cálculo producto de la interpolación por spline cúbicos es más preciso en la serie de datos que el del método de interpolación lineal por lo que para la resolución de este problema es el mejor.

V. SOLUCIÓN PROBLEMA 2: ESTIMACIÓN DE DATOS DE LA ESTACIÓN QUIXERAMOBIM

El proceso que se siguió para resolver este problema en un principio fue el mismo que el del primer problema puesto que se seleccionó el 80% (250) de la información de la temperatura interna de la estación de Quixeramobim. Luego de esto, se buscaron estos mismos datos en la estación de Quixada, estación que se utilizó para aproximar los datos de la primera mediante la interpolación. A continuación se presentan los resultados obtenidos:

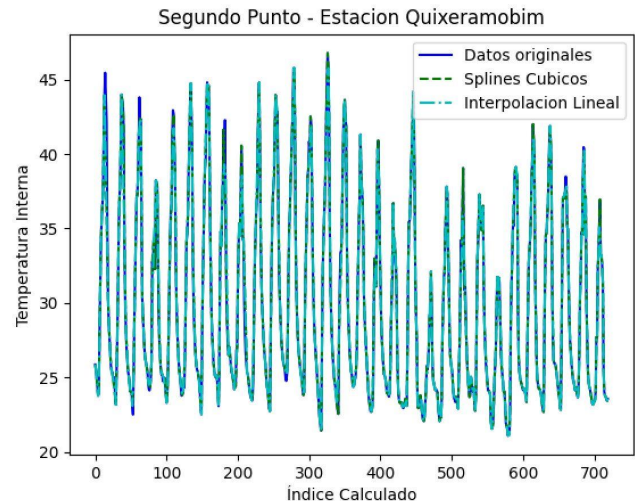


Figura 6. Gráfica aproximación temperatura interna Quixeramobim con interpolación de la temperatura interna en Quixada.

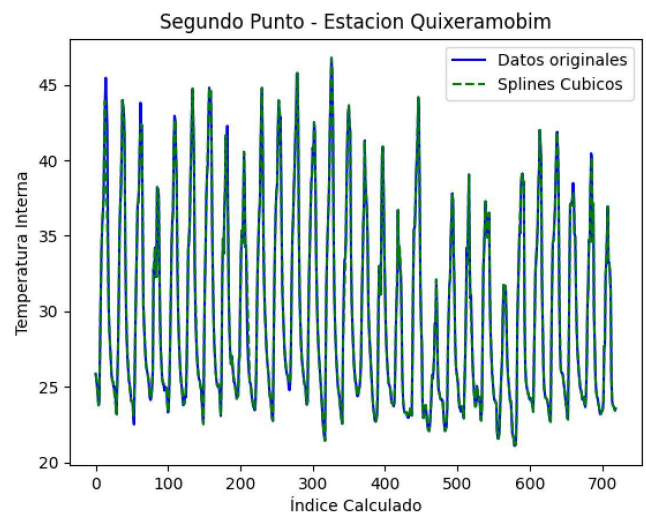


Figura 7. Gráfica aproximación temperatura interna Quixeramobim con interpolación spline cúbico de la temperatura interna en Quixada.

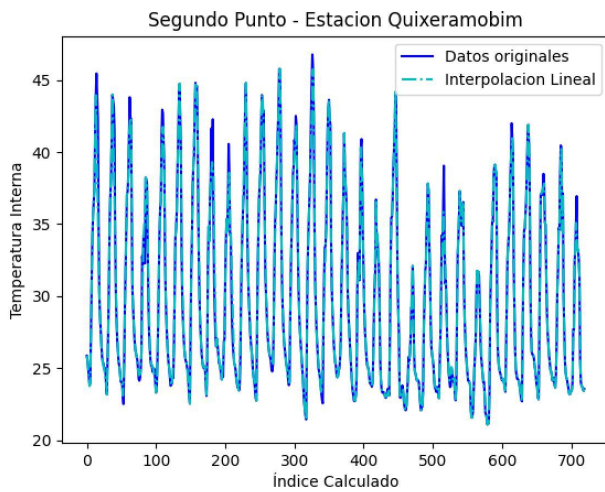


Figura 8. Gráfica aproximación temperatura interna Quixeramobim con interpolación lineal de la temperatura interna en Quixada.

El propósito de este ejercicio era conocer la aproximación al valor de la temperatura interna de la estación pluviométrica de Quixeramobim en Brasil a partir de los datos de la estación de Quixada, la más próxima a ella. El resultado de esta interpolación puede resultar muy importante debido a que de ser los valores finales muy próximos, se podría por ejemplo, retirar la base y moverla a un lugar que actualmente no esté cubierto. Para tomar decisiones como la anterior resulta bastante útil el cálculo de los errores entre los valores reales conocidos y los resultantes tras la aplicación del método. Estos fueron los cálculos del error obtenidos :

```
ERROR MAXIMO SPLINE 4.835418783683394
ERROR MINIMO SPLINE 0.004684827874456232
ERROR MEDIA SPLINE 0.7376370450447101
ERROR ABSOLUTO SPLINE 0.649870423479586
ERROR MEDIO CUADRADO SPLINE 1.1749690902628702
```

Figura 9. Valores error interpolación spline cúbico temperatura interna estación Quixada vs datos originales Quixeramobim.

```
ERROR MAXIMO LINEAL 6.484000000000002
ERROR MINIMO LINEAL 0.015000000000000568
ERROR MEDIA LINEAL 0.7552893518518518
ERROR ABSOLUTO LINEAL 0.6712467849794237
ERROR MEDIO CUADRADO LINEAL 1.2553428908206516
```

Figura 10. Valores error interpolación lineal temperatura interna estación Quixada vs datos originales Quixeramobim.

VI. Análisis solución problema 2

En este problema en primera instancia, se evidencia un aumento en los resultados obtenidos de los errores de ambos métodos en comparación a los del problema 1, no obstante

este aumento en la magnitud era de esperarse puesto que se está realizando una aproximación a otra base.

Respecto a las gráficas, podemos evidenciar el mismo fenómeno que se presentó en el primer problema donde tanto la línea correspondiente a la interpolación lineal como la de splines cúbicos muestra diferencias con respecto a la línea original en distintos puntos dejando en evidencia los errores que se presentan tras usar estos métodos.

En vista de que a partir de las gráficas nuevamente resulta difícil concluir cuál método de interpolación ofrece mejores resultados, es necesario recurrir al cálculo de errores en el que la interpolación lineal presentó los valores más altos en cada uno de los evaluados en el problema. De esta manera y a pesar de que ambos métodos tienen una muy buena aproximación, spline cúbico es el que logra realizar la estimación más apropiada.

Para finalizar y con base en los buenos resultados obtenidos por ambos métodos teniendo en cuenta que la media del error es 0.73 para el spline cúbico y 0.75 para la lineal, resultaría conveniente considerar mover la base a otro sitio ya que a partir de la estación de Quixada se pueden realizar pronósticos con errores con valores reducidos..

V. CONCLUSIONES

En conclusión, la interpolación es una buena herramienta para realizar estimaciones de variables meteorológicas siendo el método que mejor se ajusta a los problemas presentados el de splines cúbicos, debido a que tiende a presentar un valor menor en los errores calculados, además se evidencia que se obtienen resultados con menor error cuando los datos usados en la interpolación se recogen de la estación en cuestión. Al momento de estimarlos con otra estación, aunque sea similar, el error crece, como se esperaba que pasara.

El trabajo de estos problemas también demuestra la importancia que tienen los errores cuando en las gráficas es difícil comparar el rendimiento entre dos resultados. De estos también se resalta que el hecho de que un método presente el menor error máximo y error mínimo no implica necesariamente una mejor aproximación.

VI. REFERENCIAS

- [1] S. Montaña, "Armando el rompecabezas del estado del tiempo en Bogotá", Pesquisa, pp. 4,5, 2014.
- [2] "Interpolación." [Online]. Available: <http://www.prof.uniandes.edu.co/~gprieto/classes/compufis/interpolacion.pdf>. [Accessed: 12-May-2021].
- [3] J. Vidal and G. Trincada, "Aplicación de interpolación 'spline' cúbica en la estimación de volumen," *Revistas electrónicas UACh*.