Air pollution distribution in Telkom University: spatial interpolation map

I D Oktaviani¹, B Erfianto² and A Rakhmatsyah³
School of Computing, Telkom University, Bandung, Indonesia

E-mail:¹oikkedian@gmail.com,²erfianto@telkomuniversity.ac.id,³kangandrian@telkomuniversity.ac.id

Abstract. Currently, air quality information in a region becomes an important thing to know. Some efforts have been conducted to show air quality in certain region. One of the efforts that has been done is information regarding air quality in several big cities in Indonesia which can be seen in the official website of the Ministry of Environment and Forestry that have several weaknesses. One of the problems to be overcome in this research is visualization of air quality data that is monitored at one point only in which that point is the placement location of air quality monitoring station. Because of that, we need an application that can display a map of air pollution distribution using the spatial interpolation method. The solution offered is the depiction of air quality by using heatmap on the map. The method used to produce heatmap with smooth result is natural cubic spline interpolation method. The production of heatmap uses API which is provided by Google Maps. The final result obtained is the map view with the coloration in the form of color gradation in accordance with the air quality value that is obtained.

1. Introduction
Air quality in this era becomes an important information that needs to be known by the surrounding community. This is due to the rampant global warming issue that is often discussed in various regions. Some elements that can be assessed as the root cause of increased air pollution, namely PM₁₀, SO₂, CO₂, O₃, and NO₂ [1]. Each element has a different impact. The danger, not everyone knows the level of air pollution in the region where he lived so that this can be used as the subject matter raised in this study. The Ministry of Environment and Forestry (Menlhk) has done a way to publicize air quality in several cities in Indonesia [2]. But apparently this is still considered less than the maximum, because the data provided by Menlhk through its official website is only data obtained from one air quality monitoring station by using one type of sensor only so that the data obtained is still considered less. Not only in urban areas alone, information on air quality is also required in areas that are often passed by motor vehicles. In this study, the area used as a place of study is the campus area of Telkom University of Bandung (Tel-U).

The developed system is a system that can display information about air quality in some Tel-U areas with data visualization in the form of heatmap using spatial interpolation method. The need to use the interpolation method because with this method, the data census conducted does not take much cost and time so that this method is considered more efficient for the development of air quality monitoring system in the Tel-U region. The use of heatmap provided by the Google Map API will provide the result of air pollution mapping that uses certain colors as a medium of information dissemination.

Based on it, this research develops an application that can display air pollution distribution map using spatial interpolation method. The scope of this research is interpolation method used is natural cubic
spline interpolation with visualization in form of heatmap to map air pollution using Google Maps API. Tests on the built system were carried out in the Tel-U area with the distance between the sensor nodes for approximately 25 meters and the types of air pollutants detected were carbon monoxide (CO), carbon monoxide (CO₂), and particles dust (PM₁₀).

After the introduction that opens the subject matter, then it will discuss the literature review which is the reference source for making this system. In the literature review will be briefly explained about spatial data and interpolation methods. Then in the next discussion will be explained about the system built. After that, the test results will be evaluated which will be presented in the evaluation chapter and conclude with a concluding chapter.

2. Related study
Air pollution is air that has been contaminated by elements or chemicals mixed into the air for example dust, gas, smoke, and other particles. Air polluted by these pollutants may be harmful if inhaled. Some of the most commonly encountered pollutants are CO, CO₂ dan PM [3]. Based on effect caused by the three, it will know the value of each kind of pollution by interpolated assistance. Interpolation is a method of calculating or search a point among some known value based on the experiment or a function. The need for interpolation because the use of limitations nodes sensors used to identify the value of the air pollution. There are different types of interpolate that can be used, namely linear interpolation, quadratic interpolation, and interpolate cubic [4]. The type of interpolation used in this research is cubic interpolation. In cubic interpolation, searching or calculating a point between 4 points. The results of cubic interpolation can look more subtle than other interpolations. In figure 1 we can see the difference of graph between interpolation.

![Differences of interpolation charts](image)

**Figure 1.** Differences of interpolation charts.

The difference between interpolation is linear interpolation connecting only two points to produce a straight line whereas quadratic interpolation connects three points with finer graphs than linear interpolation. In addition there is also interpolation of cubic interpolation that connects between 4 dots and produces finer graphs compared with linear interpolation and quadratic interpolation.

In the calculation of interpolation process required spatial data in the form of latitude and longitude of the sensor node. Spatial data is obtained from measurements containing information about location and measurement. This data is presented in the form of geographical position of the object, location, relation with other objects, using the coordinate point and area. Spatial data can be either discrete or continuous data. Spatial data is one of the dependent data models, since spatial data is collected from different locations indicating dependency between data measurement and location of [5]. Spatial data has two important parts that make it different from other data, ie location information (spatial) and descriptive (attribute) information [6].

3. Spatial interpolation modeling
In this research, a model with the expected output of a heatmap with the help of points on the map to produce a smoother appearance. The process of obtaining a heatmap starts from the collection of sensor data which is then stored in the database to be processed and can produce a heatmap as illustrated in the figure 2. Data processing is done by using natural cubic spline interpolation method. Figure 3 below shows the color comparison on the map when using the linear interpolation method and cubic interpolation [7].

In the picture we can see the difference in color distribution. The linear interpolation describes a relatively distinct color spreading spread with cubic interpolation. It is based on the use of the amount of data used. A linear interpolation using two origin points as a parameter will result in a value that is in the middle of two origin values. While cubic interpolation using four point of origin as a parameter so it will produce a value closer to the original value. Based on this, the results displayed by the cubic interpolation will look more subtle because of the accuracy of the value obtained from the calculation so that if the value of interpolation obtained closer to the original value, then the method can be said to be a good interpolation method. If the cubic interpolation method is applied to the grid of the sensor such as the Figure 3, then the distribution of the data will look more real due to the sensor involved in the calculation of the two axes, ie the x and the y axis.

Spatial interpolation map creation requires several types of parameters to then be processed into an interpolation point on the map. Required parameters are specified spatial data and air quality data received by the server. Spatial data used in the form of latitude and longitude from the location of air quality monitoring station. While air quality data used in the form of ISPU value of each type of air pollution, ie CO, CO2, and PM10.
The spatial data as well as the air quality data that has been obtained can not be fully displayed directly on the map. The calculation process using interpolation method is done to process the data as described in the Figure 4. There are two calculation processes, namely when determining the point of interpolation on the map and when determining the value of air quality at each point of interpolation. The parameters used to determine the interpolation point on the map are consistently generated latitude data. While the parameters used to determine the value of air quality at each point of interpolation in the form of an integer that represents the number of air quality monitoring stations, namely 1, 2, 3, 4, and 5. The resulting output is longitudinal data for laying the interpolation point on the map and data air pollution at each point of interpolation. Figure 5 describes the resulting interpolation point by 50 dots and the Figure 6 describes the distance between air quality monitoring stations as far as 25 meters with the position of each node sensor straight lined.

![Figure 4. System flow.](image)

![Figure 5. Interpolation point.](image)

![Figure 6. Sensor topology.](image)

The interpolation calculation process has several steps [8]. First, the set of data that has been known value. In this case, the data already known are the latitude and longitude data of each air quality monitoring station that has been accepted by the web server. In addition to latitude and longitude data, the data quality CO, CO2, and PM10 into data that has been known value and will be searched for interpolation value. The value of this air quality will always be updated every 30 minutes so the map visualization changes every 30 minutes once adjusting the air quality data stored on the server. After the data needed to do the interpolation calculations have been collected, the next step is to find the value of interpolation. There are several calculation processes using the natural cubic spline interpolation method. The first step is to identify the values of \( h, b, v, \) and \( u \). As shown on equation (1) below.
\[ h_i = t_{i+1} - t_i, \quad i = 0, ..., n - 1 \]

\[ b_i = \frac{1}{h_i} (y_{i+1} - y_i), \quad i = 0, ..., n - 1 \]

\[ v_i = 2(h_{i-1} + h_i), \quad i = 0, ..., n - 1 \]

\[ u_i = 6(b_i - b_{i-1}), \quad i = 0, ..., n - 1 \]

\[ z_0 = z_n = 0, \quad i = 0, ..., n - 1 \]  \hspace{1cm} (1)

After the values of \( h, b, v, \) and \( u \) are known, then next is to compute with the matrix to get the value of \( z \). The calculation method with the matrix used is the multiplication of the values of \( v \) and \( h \) to \( z \) so as to generate the value \( u \). In this case, the values of \( v, h, \) and \( u \) are known. Thus, to be able to know the value of \( z \), a calculation is done using the inverse matrix. The \( z_0 \) and \( z_n \) are 0 so all you need to know is the value from \( z_1 \) to \( z_{n-1} \) in equation (2).

\[
\begin{bmatrix}
  v_1 & h_1 \\
  h_1 & v_2 & h_2 \\
  & h_2 & v_3 & h_3 \\
  & & & \vdots & \ddots & \ddots \\
  & & & & h_{n-2} & v_{n-1} \\
 \end{bmatrix}
\begin{bmatrix}
  z_1 \\
  z_2 \\
  \vdots \\
  \vdots \\
  z_{n-2} \\
  z_{n-1} \\
\end{bmatrix}
= 
\begin{bmatrix}
  u_1 \\
  u_2 \\
  \vdots \\
  \vdots \\
  u_{n-2} \\
  u_{n-1} \\
\end{bmatrix}
\]  \hspace{1cm} (2)

When all values of the required variables have been collected then look for the final interpolation value done. Here is the formula used to determine the interpolation point.

\[ S_i(x) = \frac{z_{i+1}}{6h_i} (x - t_i)^3 + \frac{z_i}{6h_i} (t_{i+1} - x)^3 + \left( \frac{y_{i+1}}{h_i} - \frac{z_{i+1}}{6h_i} \right) (x - t_i) + \left( \frac{y_i}{h_i} - \frac{z_i}{6h_i} \right) (t_{i+1} - x) \]  \hspace{1cm} (3)

From the equation, \( S_i(x) \) or which can be called with \( y \) is the interpolation value of the longitude and the value of the air quality sought. The parameter that becomes the reference in equation (3) is the variable \( x \). The value of \( x \) used is the latitude data already generated with consistent values using the iteration. The algorithm of the interpolation calculation process is as follows.
In the algorithm there is a description of some function calls on the API provided by Google Maps API. APIs used in making this heatmap map include the use of Google maps, latitude and longitudinal plots, the use of pop ups as a description, heatmap staining, and other matters involving the map itself. The use of the API produces the output of a map displayed by heatmap staining in accordance with a predetermined index and displaying the sensor information at each air quality monitoring station. On line 8 to line 15 is the process of searching the values of \( h, b, v \), and \( u \) with the formula according to equation (1). The process of calculating equation (2) is found in lines 16 and 17. And the interpolated end result of the \( S_i(x) \) value can be generated on the 18th line until the API redirects from Google Maps to display the heatmap.

4. Result and discussion

The air quality monitoring test is carried out in the motorcycle area of the Faculty of Engineering Tel-U on the time span from 11:00 to 13:00. The area used is on the guardrail of the motorcycle parking area from the inside area to the outside area. The spread of air pollution can be seen on the map which is updated every two minutes. This is because the user can continue to monitor the development of pollution of air polluted by air quality monitoring station every thirty minutes. Figure 7 displays the data received by the server and has been converted to json format.

Algorithm 1. Cubic interpolation algorithm.

In the algorithm there is a description of some function calls on the API provided Google Maps API. APIs used in making this heatmap map include the use of Google maps, latitude and longitudinal plots, the use of pop ups as a description, heatmap staining, and other matters involving the map itself. The use of the API produces the output of a map displayed by heatmap staining in accordance with a predetermined index and displaying the sensor information at each air quality monitoring station. On line 8 to line 15 is the process of searching the values of \( h, b, v \), and \( u \) with the formula according to equation (1). The process of calculating equation (2) is found in lines 16 and 17. And the interpolated end result of the \( S_i(x) \) value can be generated on the 18th line until the API redirects from Google Maps to display the heatmap.

Result: Heatmap with interpolation
read co
for as much as the sensor point do
  if the value of co is as required then
    call Google Maps API weight function and set the value
  end
end
for three times do
  calculate the value of h and b for CO
  calculate the value of h and b for Peta
  if the number of iteration is one then
    calculate the value of v and u for CO
    calculate the value of v and u for Peta
  end
end
calculate the value of z for CO
calculate the value of z for Peta
for three times do
  for as much as the interpolation point do
    calculate the value of CO interpolation using interpolation formula
    calculate the value of map interpolation using interpolation formula
    plotting lat lng interpolation using Google Maps API
    if the value of CO interpolation is as required then
      call Google Maps API weight function and set the value
    end
  end
end

Algorithm 1. Cubic interpolation algorithm.

4. Result and discussion

The air quality monitoring test is carried out in the motorcycle area of the Faculty of Engineering Tel-U on the time span from 11:00 to 13:00. The area used is on the guardrail of the motorcycle parking area from the inside area to the outside area. The spread of air pollution can be seen on the map which is updated every two minutes. This is because the user can continue to monitor the development of pollution of air polluted by air quality monitoring station every thirty minutes. Figure 7 displays the data received by the server and has been converted to json format.

```json
{"detail": [{"id": "1", "lat": "-6.9760731", "lng": "107.6316371", "co": "5.70", "co2": "495.65", 
  "dust": "0.001", "id": "2", "lat": "-6.9760731", "lng": "107.631862", "co": "9.77", "co2": "218.61", 
  "dust": "0.001", "id": "3", "lat": "-6.9760731", "lng": "107.632087", "co": "16.51", 
  "dust": "0.001", "id": "4", "lat": "-6.9760731", "lng": "107.632937", "co": "7.33", "co2": "1434.15", 
  "dust": "0.001", "id": "5", "lat": "-6.9760731", "lng": "107.632537", 
  "co": "4.88", "co2": "886.05", "dust": "156.00", "timestamp": "2018-06-16 01:56:10"}]
```

Figure 7. The data in JSON format.
After the data successfully received by the server, then performed the calculation by the method of cubic interpolation. Figure 8 is the result of the interpolated calculation that has been done.

Based on these interpolation calculations, the interpolation points on the map can be constructed along with the heatmap staining. In the heatmap, the colors shown to show the quality of CO and PM$_{10}$ follow the rules set by the Air Pollution Standards Index (ISPU) there are:

i. Green, represents the value of air quality both with a range of values 0-50.
ii. Blue, represents the value of medium air quality with a range of 51-100 values.
iii. Yellow, represents the value of unhealthy air quality with a range of 101-199 values.
iv. Red, represents a very unhealthy air quality value with a range of 200-299 values.
v. Black, represents the value of dangerous air quality with a range of 300-500 values.

While the colors shown to show the quality of CO$_2$ that uses the original ppm value to carbon dioxide [9] there are:

i. Green, represents the value of air quality both with a range of 0-450 values.
ii. Blue, represents the value of medium air quality with a range of 451-600 values.
iii. Yellow, represents the value of unhealthy air quality with a range of 601-1000 values.
iv. Red, represents the value of very unhealthy air quality with a range of 1001-2500 values.
v. Black, represents the value of dangerous air quality with a range of 2501-5000 values.

The color change will be visible when air pollution data reaches a certain value limit. Each type of air pollution will display the visual map differently according to the sensor data results as shown in figure 9.

![Figure 8. Interpolation result.](image-url)

(a) Map of carbon monoxide
(b) Map of carbon dioxide
Comparison if without the use of interpolation aid, the map can only display a heatmap with a disconnected heatmap view on some data with a low value. This underlies the importance of using the aid of interpolation points so that air pollution can be seen in more detail in areas where air quality monitoring stations. In this research, test is done to show map without using interpolation method, using 25 points, 50 point and 75 point interpolation point. The result will look much different. The density level of the interpolated point may affect the color of the heatmap shown as can be seen in figure 10. Selection of the number of interpolation points used can also affect the speed of computing and data accuracy.

Similarly, the comparison made to the use of the weight function in the Google Maps API can affect the degree of color gradation in the heatmap. The weight function performs a color plot on the marker of a set number of values. In this test the comparison of the value of weight is used. The figure 11 below shows the test results for the weight function with the decimal point of the title and the federation.
In the air quality monitoring station marker on the map there is information that the user can see. Users simply select one of the stations then the station will appear a pop up with information in it. The information displayed is latitude, longitude, pollution type data, and the last data update time as can be seen in figure 12.

The heatmap that appears will have a different pattern when zooming or downlining the map. This heatmap pattern still follows the value of the air quality monitoring station and the interpolated value at each interpolation point. Figure \ref{fig:zoom} shows a map sighting if maximum magnification is performed with good air quality status so it can display the green color on the heatmap. Differences are seen when the map has been down three times giving green and blue on the map. If the map is minimized again, it will cause a color change to display the color black.
The advantages of using interpolation method is to save the use of sensors to receive air quality data. The value of air quality at a point can be represented by the use of air quality values from four air quality monitoring stations and the value has a result that is close to one point to another. This can be seen from the color gradation seen between one station with other station caused by interpolation value to air quality. The involvement of interpolation points can help the user to know the air quality in the area between the two stations. Unlike the case if not done interpolation calculation process. The less spaced the interpolated points are raised then the color display on the map looks more faded, so even with the number of interpolated points raised too big, the color display on the map looks darker resulting in a mismatch on the actual air quality score. Based on this, in this study the point of interpolation generated as much as 50 points.

While based on the use of the value used in the weight function provided by the Google Maps API, the smaller the decimal value used, the more invisible the color will be. Conversely, if the decimal value used is higher, the resulting color gets thicker resulting in a mismatch of actual air quality values. The use of decimal places with two digits behind the comma can produce a map with color according to the level of pollution that has been determined.

5. Conclusion
The development of a system to complement air quality monitoring produces a web application with a visualization of maps in which there is a heatmap. The presence of heatmap indicates that the data has been successfully sent by the air quality monitoring station and successfully received by the server. The color display on heatmap is supported by several parameters such as spatial data in the form of latitude and longitude of air quality monitoring stations as well as the value of the air quality itself. The making of this heatmap involves the interpolation calculation process using the natural cubic spline interpolation method which is then equipped with color gradation which shows the air quality in the area other than the air quality monitoring station. Although the computation process is larger than other methods, this method can produce smoother visuals with more accurate interpolation values, which are close to the point value of origin. The use of the number of interpolation points can make the map look different. The more the number of interpolation points used, the heatmap display will be smoother but use the computing time is getting bigger too. So that 50 interpolation points used are considered sufficient to display a map with density points that make the map look smoother with faster computation process when compared with the use of more interpolation points. In further development, it is expected that this air quality monitoring system can be used dynamically with random sensor station locations.

References
[1] Kusumaatmadja S 1997 ISPU (Indeks Standar Pencemar Udara) Universitas Islam Indonesia: http://www.cets-uui.org/BML/Udara/ISPU/ISPU%20(Indeks%20Standar%20Pencemar%20Udara).htm Accessed: October 12, 2017
[2] Kehutanan K L 2015 INDEKS STANDARD PENCEMARAN UDARA (ISPU) http://iku.menlhk.go.id/ Accessed: September 26, 2017

[3] WHO 2000 Air Quality Guidelines – Second Edition WHO Regional Office for Europe

[4] Haryanto A Interpolasi, Definisi, dan Macam-macamnya Academia: http://www.academia.edu/8738105/Interpolasi_DefinisidanaMacam-macammnya Accessed: October 12, 2017

[5] Cressie N A 1993 Statistics For Spatial Data (New York: John Wiley and Sons, Inc.)

[6] Atie Puntodewo S D 2003 Sistem Informasi Geografis untuk Pengelolaan Sumber Daya Alam (Bogor: CIFOR)

[7] Nugroho A S 2017 Pemetaan Radiasi Gas Butane berbasis Sensor Array Mapping of Butane Radiation using Array Sensors

[8] Kvarving A M 2008 Natural cubic splines NTNU: https://www.math.ntnu.no/emner/TMA4215/2008h/cubicsplines.pdf

[9] ToolBox E 2008 Carbon Dioxide Concentration - Comfort Levels The Engineering ToolBox: https://www.engineeringtoolbox.com/co2-comfort-level-d_1024.html Accessed: May 16, 2018