Peer Review Statement

All papers published in this volume have been administered by the Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of proceedings journal published by UM Jember Proceeding Series.

Type of peer review: Single Anonymous
Number of submissions received: 10
Number of submissions sent to review: 10
Number of submissions accepted: 8
Acceptance rate: 80%
Average number of reviews per paper: 2
Contact person for queries:

Name: Bagus Setya Rintyarna
Email: lppm@unmuhjember.ac.id

PREFACE

The first International Conference on rural development was held on December 15, 2020 via Zoom Webinar Universitas Muhammadiyah Jember (online) due to the pandemic. ICRD 2020 aims to bring together leading academic and industry researchers, scientists, engineers and practitioners to exchange research and service ideas, methods, results and share the latest experiences, on all theoretical, experimental and applied aspects of the Social and Exact field. ICRD 2020 will provide a unique interdisciplinary and multidisciplinary forum for researchers, practitioners, and educators to present and discuss innovations, trends, practical challenges faced, and smart solutions adopted in the social and exact fields of rural development.

At this conference, there were research and service papers carried out by several researchers from various institutions and universities, which were presented in parallel oral sessions. In addition, the committee also presented four keynote speakers, namely Prof. Drs. ec. Ir. Riyanarto Sarno, M.Sc, Ph.D (Ten November Institute of Technology, Indonesia), Dr. Emy Kholifah R., M.Si (University of Muhammadiyah Jember, Indonesia), Prof. Habibullah Magsi, Ph.D (Sind Agricultural University, Pakistan) and Prof. Dr. Kittisak Jermsittiparsert, (Henan University of Economics and Law, China).

We would like to thank everyone involved for the success of the conference, especially to the organization's staff, program committee members and reviewers. They work very hard to review all papers and provide valuable suggestions for authors to improve their research. In addition, we would also like to thank the external reviewers who provided extra assistance with the review process, and the authors who contributed their research to this conference.

The conference, the first of its kind, provided a forum for researchers and scientists to communicate their latest developments and present their original results in the social and exact fields of village development. We hope that the experience of the 2020 participants is fruitful and can be remembered for a long time.

We hope to see you all again at the next conference.

TABLE OF CONTENT

International Applied Science
UMJember Proceeding Series (2022) Vol. 1, No. 1 : 14-23

ICRD 2020

Spatial Analysis For Index Identification Surface Water Pollution

Anang Aris Widodo[1*], Ike Rokhmawati[2], Anik Vega Vitianingsih[2], Fitri Marisa[3], Yunus Susilo[2]
[1]Merdeka University of Pasuruan, Indonesia
[2]University Dr, Soetomo Surabaya, Indonesia
[3]Widyagama University of Malang, Indonesia

DOI: https://doi.org/10.32528/ias.v1i1.18
*Correspondensi: Anang Aris Widodo1
Email: anangariswidodo@gmail.com

Published: Januari, 2022

Abstrak: Surface water pollution is an essential problem for policymakers, this is because water is a source of life for the environment. One of the causes of pollution is the company's contribution to waste disposal. The purpose of this research is to identify local pollution points and their pollution status and to map the location of the company to determine the possible contribution of the company to water pollution using a Web-based Geographical Information System (Web-GIS) technology. The data used in this study came from the Environmental and Sanitation Service (DLHK) Sidoarjo Regency. The Pollution Index (IP) method is used for the spatial analysis process based on parameters TSS, DO, BOD, COD, Phosphate, Total Coliform, and Fecal Coliform. The results of spatial analysis using the IP method get an accuracy value of 91% which states that this method is suitable to be recommended for further research with the same data behavior.

Keywords: spatial analysis, GIS, web-gis, water pollution, pollution index.

INTRODUCTION

Water pollution is the entry or inclusion of living things and other components into water by human activities, therefore, the water quality decreases to a certain level cause water cannot function according to its use [1]. Indonesia has entered the list of 10 water-rich countries with water availability reaching 3.9 trillion m3 / year, but only 17.69% can be utilized and 25.3% of them for irrigation, domestic, urban and industrial needs [2]. Abundant water is collected in water sources of ⩾5,590 rivers and 1,035 lakes [2]. Based on PP-RI No.82 of 2001 concerning Water Quality Management and Water Pollution Control, it has been explained that water is an important component of the environment for human life and other living things [1]. Therefore, knowing and understanding environmental conditions and quality is very important. This is useful as a basis for a country's policy to utilize the potential of the environment to improve people's welfare [2]

Sidoarjo District is located in East Java Province which surrounded by two rivers, the Surabaya River and the Porong River, which is a branch of the Brantas River which has an upstream in Malang Regency. Indonesia has thousands of rivers, but the exact number is not known [2]. Some of the famous large rivers are the Ciliwung, Bengawan Solo, Kapuas and Cisadane rivers. But unfortunately, all the big rivers are in polluted condition [2]. In 2016, the three rivers were heavily polluted. Today, although Sidoarjo District is one of the big cities in East Java, but there are still many problems related to environmental conditions. An interim survey based on the results of an interview by Mr. Mohamad Edi Kurniadi, ST, MM as the Head of Environmental Management and Pollution Control at DLHK Sidoarjo District on September 27, 2019, at 11:00 WIB, it was stated that there were environmental conditions including waste problems, floods, fires, and lack of clean water. The results of Indonesian Environmental Statistics in 2018 regarding the status of river water quality in 2016 stated that the Surabaya River, Porong River, and Brantas River were

heavily polluted, so that Sidoarjo Regency is very critical of clean water, because the three large rivers surrounding it have been polluted [3].

Previous studies used the IP method to map pollution [4], to determine the status of river water quality [5]. Physical, chemical, and biological parameter data to determine the level of pollution due to waste [6]. IP and IDW methods as well as data on the rainy and dry seasons were used to predict water quality conditions spatially [7]. Water quality data for the period January-December 2017 is for determining water quality status [8], BOD and COD monitoring data are used to determine the characteristics of sea water quality at low tide [9]. Data of spatial and temporal variations in physico-chemical parameters were used to evaluate chemical physics characteristics [10]. The Water Quality Index (WQI) method with 15 parameters aims to provide an understanding of water quality [11]. The use of GIS technology aims to evaluate river water quality [12], water quality data from several river water samples are used to quickly determine the status of river water quality [13]. Based on a literature review study, the analysis process was carried out only on one river flow and there was a lack of other actions to utilize data or documents owned by the agency. This study will use the IP method to identify water quality in Sidoarjo Regency.

The purpose of this study is to provide information on the location of the river test post and the point of the surface water pollution area (river area range) with a radius of 100m based on the pollution status, and to provide information on the location of companies in Sidoarjo Regency in knowing the correlation of surface water pollution levels based on data mapping. companies (companies that contribute as a cause of water pollution). Correlation of pollution levels will be adjusted to company data based on environmental documents, namely Environmental Management Efforts and Environmental Monitoring Efforts (UKL-UPL). This research will utilize web-based Geographic Information System (GIS) technology (Web-GIS).

This study will use the IP method to determine the status of water quality which will show 4 levels of condition, namely good, lightly polluted, moderately polluted, and heavily polluted within a certain time based on river test results with physical (TSS), chemical (DO, BOD, COD, Phosphate), and microbiology (Total Coliform, Fecal Coliform), as well as company data used to find out the contribution of each company to water pollution. The status of water quality in the IP Method is in a good category if it has a value of $0 \leqslant PIj \leqslant 1.0$, the category is lightly polluted with a value of $1.0 < PIj \leqslant 5.0$, the category is moderately polluted with a value of $5.0 < PIj \leqslant 10$, and is polluted weight with PIj value> 10.

The results of this study can facilitate DLHK Sidoarjo District, as well as use supporting methods in providing information about areas that have been identified as water pollution and their pollution status, and can identify companies that contribute to water pollution based on UKL-UPL documents. So that DLHK can immediately take firm action and re-monitor the contaminated surface water flow area

Spatial dataset consists of two components, namely spatial data and attribute data [14]. The parameters used in this study will be adjusted to the results of the Decree of the State Minister for the Environment Number: 115 of 2003 [15] as in Table 1 (Source: Guidelines for the Indonesian IKLH 2013 [16]).

Parameters Total Suspended Solid (TSS) is a solid in the solution but not dissolved, cannot turn into a cloudy solution, cannot immediately settle. Example: mud, clay, metal, fungus [15]. Dissolved Oxygen (DO) parameter is the dissolved oxygen content in river water for measuring water quality parameters [15]. Biochemical Oxygen Demand (BOD) parameters are the amount of dissolved oxygen required by microorganisms in the air to break down organic matter [15]. Parameter of Chemical Oxygen Demand or Chemical Oxygen Demand (COD) is the amount of chemical compounds needed for oxygen in the air to break down organic matter [15]. Phosphate parameter is any compound in the form of dissolved, suspended

in the organism cell [15]. Parameters Total Coliform or Total Coliform Bacteria are bacteria found in soil and air environments that affect surface water, as well as those that affect the disposal of animal and human waste [15]. The parameter of Fecal Coliform or Coliform Bacteria Content is the total coli group, but more specifically the indication is an accurate indication of animal or human waste contamination [15].

Table 1. Water Pollution Index Spatial Data

Spatial Data	Quality standards
TSS	50
DO	4
BOD	3
COD	25
Fosfat	0
Total Coliform	1000
Fecal Coliform	5000
Firm Location	

METHOD

Spatial data analysis resulted from spatial data modeling. Spatial data sets were used as baseline data [17]. The process stages in spatial data modeling to identify the surface water pollution index in Sidoarjo Regency are shown in the flowchart of Figure 1. This stage provides an overview of how the system works. Starting from inputting all the data needs, then the spatial data modeling process is carried out by determining the AI method that matches the behavior of the data obtained from the recording process, and the final process is displaying the spatial data modeling results file according to its function and purpose in GIS [14].

The first step is defining spatial data and attribute data in the spatial shapefile dataset (* .shp). The dataset includes maps of Sidoarjo Regency, including TSS, DO, BOD, COD, Phosphate, Total Coliform and Fecal Coliform which are included in the Pollution Index calculation method. The Pollution Index method already has a value classification with good coverage categories, lightly polluted, moderate polluted and heavily polluted as in Table 2 (Source: KepMenLH Number 115 of 2003 [15]). It can be concluded that the good category will be indicated by a green dot description, the lightly polluted category will be shown in yellow, the moderate polluted category will be shown in orange, while the heavily polluted category will be shown in red.

Table 2: Category Result Of Water Pollution

Pollution Index value	Water Criteria
$0 \leq PIj \leq 1{,}0$	Good
$1{,}0 < PIj \leq 5{,}0$	Lightly Polluted
$5{,}0 < PIj \leq 10$	Moderately Polluted
$PIj > 10$	Heavily Polluted

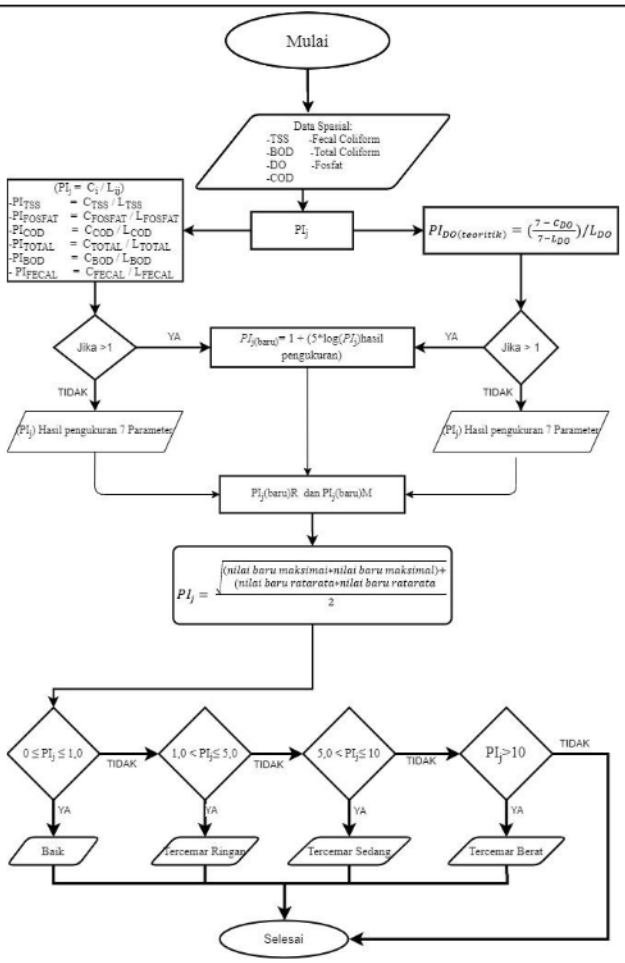

Figure 1. Flowchart of Spatial Data Modeling with Pollution Index Method

The Pollution Index Method

The steps for the Pollution Index Method are as follows [15]:

The pollution index method PI_i in (1).

$$PI_j = \frac{\sqrt{(C_i/L_{ij})_M^2 + (C_i/L_{ij})_R^2}}{2} \qquad (1)$$

Where :

PI_i = Pollution index of Designation to-j

C_i = The concentration of water quality parameter to-i measurement result

L_{ij} = The concentration of water quality to-i listed in the designation quality standard to-j

(C_i / L_{ij})M = Maximum Value C_i / L_{ij}

(C_i / L_{ij})R = Average Value C_i / L_{ij}

The initial step for the value of PI_j in the equation (1) is calculated with 2 conditions, namely having no theoretical value and having a theoretical value. In conditions that have no theoretical value, parameters TSS, BOD, COD, Phosphate, Total Coliform and Fecal Coliform are shown in Equation (2). Meanwhile, conditions that have a theoretical value are found in the DO parameter shown in Equation (3).

Has no theoretical value

PI_j Value is calculated based on C_i value which is the test result for each parameter and L_{ij} variable is the quality standard for each parameter.

$$PI_j = C_i / L_{ij} \quad (2)$$

Has theoretical value

Value of $PI_{j\,(teoritik)}$ is calculated based on C_i than the result test of L_{ij} that designation of each parameter and C_{im} variable that maximum limit of saturated DO value (7).

$$PI_{DO\,(teoritik)} = \left(\frac{7 - C_{DO}}{7 - L_{DO}}\right) / L_{DO} \quad (3)$$

Determine $PI_{j(new)}$ that also has 2

Condition value based on the final result $PI_j \leq 1.0$ as equation (4) and $PI_j \geq 1.0$ on equation (5)

If result measurement value $PI_j \leq 1.0$ based on equation (1) and (2), then in this second step the final measurement results will be written down immediately, without any calculations.

$$PI_{j(hasil\ pengukuran)} \quad (4)$$

If result measurement $PI_j \geq 1.0$ in equations (1) and (2), then the next calculation will be based on equation (5). The calculation of the logarithm with the variable P, which is a predetermined constant, is 5. Value of $PI_{j(baru)}$ = 1,0 is a critical value, because this value is expected to meet the Water Allocation Quality Standard.

(C_i / L_{ij})baru = 1.0+P.log (C_i / L_{ij}) result measurement (5)

Determine the maximum and average values based on the results of equations (4) and (5). The value of PI_M in equation (6) based on the variable M is the maximum value. The value of PI_R in equation (7) based on the variable M is the maximum value.

$$PI_M = PI_{j\,(baru)\,M} \quad (6)$$

$$PI_R = PI_{j\,(baru)\,R} \quad (7)$$

Determine the final result of PIj based on equation (1) which will then be determined pollution category based on Table 2.

Confusion Matrix Method

The Confusion Matrix method is used to determine accuracy, precision and recall values between the actual data and the classification results [18] [19].

Measurements can be made based on equations. (8), (9) and (10).

Accuracy = $(TP + TN)/(TP + FP + TN + FN)$ (8)

Precision = $TP/(TP + FP)$ (9)

Recall = $TP/(TP + FN)$ (10)

Where:

• TP is positive value data that is properly labeled by the classifier, let the TP value be the true positive number.

• TN is Negative value data correctly labeled by the classifier, let the TN value be the true negative number.

• FP is Negative data which is incorrectly labeled as positive, let the FP value be the number of false positives.

• FN represents positive value data which is incorrectly labeled as negative, let the FN value be the number of false negatives.

RESULTS AND DISCUSSION

The results of trials that have been carried out from 21 river test posts and 45 company location data in Sidoarjo Regency in 2019, data were obtained from the Environmental and Cleanliness Service of Sidoarjo Regency. The results of modeling the spatial data from the river test post with the category of surface water pollution index classification using the Pollution Index method as in Table 3 and Figure 2.As for the results of the company location mapping as in Table 4 and Figure 3.

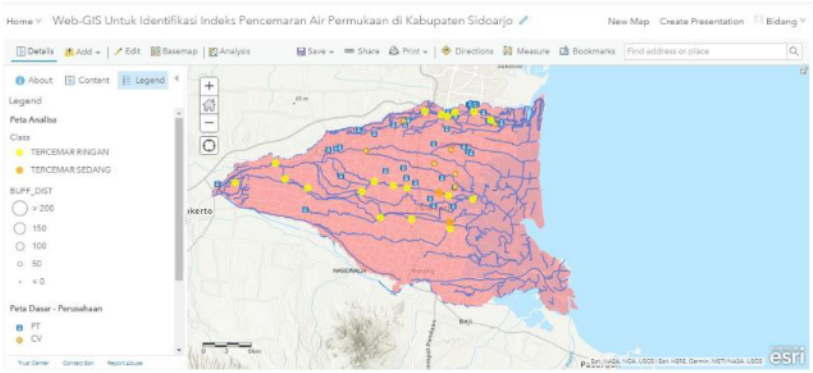

Figure 2. Classification with IP Method

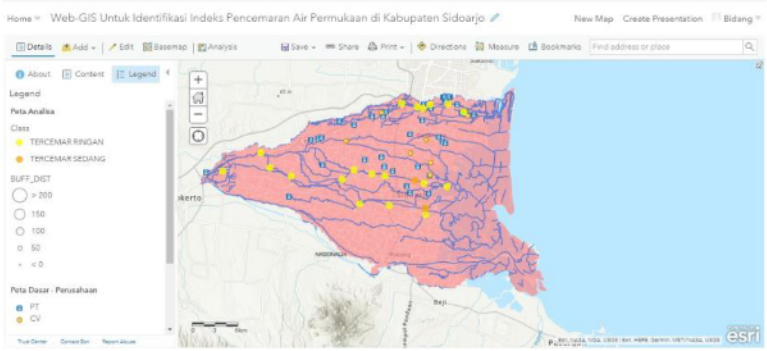

Figure 3. The results of the company location mapping

Table 3: Classification With Ip Method

Class	Data in 2019
Good	0
Lightly Polluted	19
Moderately Polluted	2
Heavily Polluted	0

Table 4: The Results Of The Company Location Mapping

Type	Data in 2019
PT	38
CV	7

PI_i value of equation (1) calculated with 2 conditions, namely having no theoretical value and having a theoretical value. In conditions that have no theoretical value, parameters TSS, BOD, COD, Phosphate, Total Coliform and Fecal Coliform are shown in Equation (2). Meanwhile, conditions that have a theoretical value are found in the DO parameter shown in Equation (3). $PI_{j\,(teoritik)}$ will shaped in to PIT_DO. C_{im} value is saturated DO with value of 7.

The test results from 7 parameters such as Table 1 will be C_TSS, C_ DO, C_BOD, C_COD, C_Fosfat, C_TotalC and C_FecalC. As for the quality standards of the 7 parameters, it will be L_TSS, L_ DO, L_BOD, L_COD, L_Fosfat, L_TotalC and L_FecalC. Equation (2), namely:

PI_TSS = C_TSS/L_TSS
PI_BOD = C_BOD/L_BOD
PI_COD = C_COD/L_COD
PI_Fosfat = C_Fosfat/L_Fosfat
PI_FecalC = C_ FecalC /L_ FecalC
PI_TotalC = C_ TotalC /L_ TotalC

Equation (3) described :

$$\text{PIT\_DO} = \left(\frac{7 - C_{DO}}{7 - L_{DO}}\right) / L_{DO}$$

Value of $PI_{j(baru)}$ has 2 condition value final result measurement $PI_i \leq 1.0$ as equation (4) and $PI_i \geq 1.0$ in equation (5). $PI_{j(new)}$ to be PIB_TSS, PIB_DO, PIB_BOD, PIB_COD, PIB_Fosfat, PIB_FecalC dan PIB_TotalC. The value of P is a constant that has a value of 5. $PI_{j(baru)}$ will be determined by looking at the final result of Equations (2) and (3). Equation (4), namely:

PIB_TSS = PIB_TSS(result measurement)
PIB_DO = PIB_DO(result measurement)
PIB_BOD = PIB_BOD(result measurement)
PIB_COD = PIB_COD(result measurement)
PIB_Fosfat = PIB_Fosfat(result measurement)
PIB_FecalC = PIB_FecalC(result measurement)
PIB_TotalC = PIB_ TotalC (result measurement)

Equation (5) is:

PIB_TSS = 1.0+5*log PIB_TSS(result measurement)

PIB_DO = 1.0+5*log PIB_DO(result measurement)

PIB_BOD = 1.0+5*log PIB_BOD(result measurement)

PIB_COD = 1.0+5*log PIB_COD(result measurement)

PIB_Fosfat = 1.0+5*log PIB_Fosfat(result measurement)

PIB_FecalC = 1.0+5*log PIB_ FecalC (result measurement)

PIB_TotalC = 1.0+5*log PIB_ TotalC (result measurement)

The maximum and average values are determined based on Equations (6) and (7). The determination will be seen based on the results of equations (4) and (5). Next, look for the final result PIj based on Equation (1) to determine the classification of pollution categories based on Table 2.

The IP method testing based on Equation (1) to Equation (7) in Figure 2 is carried out with a spatial dataset such as Table 2. The figure shows that the orange polluted category has a smaller number, namely 2 test posts. The results of the moderate polluted category will be given an example as in Figure 4. In this figure there is a test location at Pos VI Pucang Bridge Jl. A. Yani who is in the Pucang Canal with a test value of TSS 40.50, DO 4.20. BOD 13.70, COD 47.84, Phosphate 4.66, Fecal Coliform 191.50 and Total Coliform 1.155. Based on the data from the test results, it has given the final result an average value of 5.23, a maximum final value of 61.38 and PI_i is 5.77.

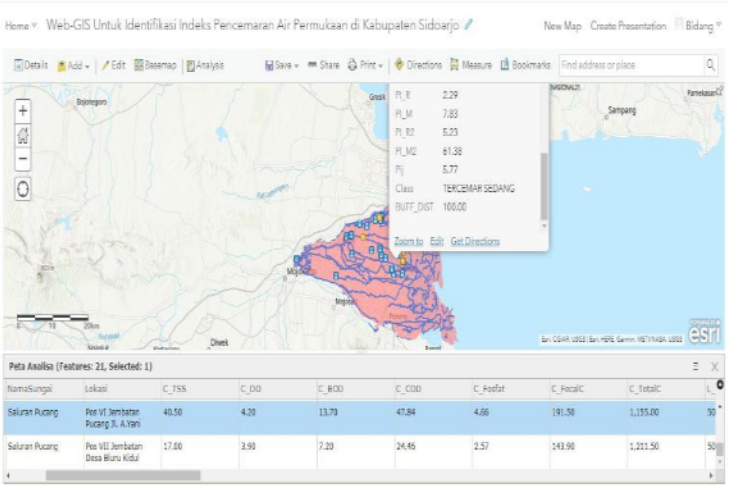

Figure 4. Result of Moderately Polluted Category IP Method

Meanwhile, for the data 19 test posts were categorized as lightly polluted which were marked in yellow. The results of the moderate polluted category will be given an example as in Figure 5. In this figure there is a test location at Post III Dsn Bridge. Tambaksari Tambakrejo Waru Village which is in the Buntung Channel with a TSS test value of 11.00, DO 4.05. BOD 6.50, COD 21.45, Phosphate 0.45, Fecal Coliform 81.50 and Total Coliform 295.00. Based on the data from the test results, it has given the final result an average value of 0.97, a maximum final value of 7.49 and PI_i of 2.06.

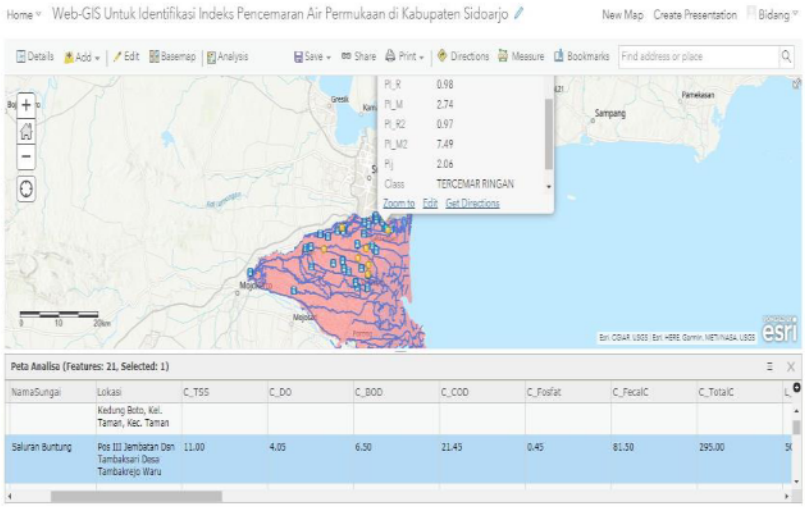

Figure 5. Results of the Light Polluted Category IP Method

.CONCLUSION

The analysis results with the Pollution Index method based on the parameters of physics (TSS), chemistry (DO, BOD, COD, Phosphate), and microbiology (Total Coliform, Fecal Coliform) at 21 river test posts gave classification results of 19 posts with lightly polluted status and 2 posts. has a medium tainted status. At that post point with a radius of 100m, no company location points enter the circular area. The results of calculations using the Pollution Index method are able to provide an accuracy value of 91%, 34% precision and 49% recall. The development of further research is the need for research trials using other methods that are considered supportive of the use of parameters and the addition of parameters in testing. It is necessary to add the location of the test post to find out the differences in accuracy, precision and recall values. The test results using the Confusion Matrix to test the feasibility of using the IP method with Web-GIS technology to identify the surface water pollution index obtained the Accuracy 91%, 34% Precision and 49% Recall.

REFERENCES

[1] Presiden Republik Indonesia, *Peraturan Pemerintah Republik Indonesia Nomor 82 Tahun 2001*. 2001, pp. 1–5.

[2] W. S. Purba, P. A. Safitri, and R. Andianti, *Statistika Lingkungan Hidup Indonesia (SLHI) 2017*. Badan Pusat Statistika Indonesia, 2017.

[3] P. A. Safitri, W. S. Purba, and M. Zulkifli, *Statistik Lingkungan Hidup Indonesia (SLHI) 2018*. Badan Pusat Statistika Indonesia, 2018.

[4] A. Erliza *et al.*, "Identifikasi Pencemaran Air Di Sepanjang Aliran Sungai Utama DAS Batang Arau Kota Padang," *Kapita Sel. Geogr.*, vol. 2, no. 5, pp. 29–34, 2019, doi: 10.24036/ksgeo.v2i5.239.

[5] R. Ermawati and L. Hartanto, "Pemetaan Sumber Pencemar Sungai Lamat Kabupaten Magelang," *J. Sains &Teknologi Lingkung.*, vol. 9, no. 2, pp. 92–104, 2017, doi: 10.20885/jstl.vol9.iss2.art3.

[6] Perdana, "Menganalisis Pencemaran Daerah Aliran Sungai (DAS) Akibat Limbah Domestik di Kota Medan," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2018, doi: 10.1017/CBO9781107415324.004.

[7] N. Lusiana and B. Rahadi, "Prediksi Distribusi Pencemaran Air Sungai Das Brantas Hulu Kota Batu

Pada Musim Hujan Dan Kemarau Menggunakan Metode Spasial Inverse Distance Weighted," *ECOTROPHIC J. Ilmu Lingkung. (Journal Environ. Sci.*, vol. 12, no. 2, p. 212, 2018, doi: 10.24843/ejes.2018.v12.i02.p10.

[8] S. Anwar, B. Hariono, M. J. Wibowo, and M. M. Dyah Utami, "Penentuan Status Mutu Air Metode Storet DAS Kali Curah Macan," *J. Ilm. Inov.*, vol. 18, no. 2, pp. 95–99, 2018, doi: 10.25047/jii.v18i2.1160.

[9] R. Arya, S. Sariffuddin, and A. Bilqis, "Mapping of Pollution Load Capacity of Tidal Manggar Watershed, Balikpapan City," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 313, no. 1, 2019, doi: 10.1088/1755-1315/313/1/012021.

[10] K. K. Vadde, J. Wang, L. Cao, T. Yuan, A. J. McCarthy, and R. Sekar, "Assessment of water quality and identification of pollution risk locations in Tiaoxi River (Taihu Watershed), China," *Water (Switzerland)*, vol. 10, no. 2, 2018, doi: 10.3390/w10020183.

[11] Z. Wu, X. Wang, Y. Chen, Y. Cai, and J. Deng, "Assessing river water quality using water quality index in Lake Taihu Basin, China," *Sci. Total Environ.*, vol. 612, pp. 914–922, 2018, doi: 10.1016/j.scitotenv.2017.08.293.

[12] Ş. Şener, E. Şener, and A. Davraz, "Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey)," *Sci. Total Environ.*, vol. 584–585, pp. 131–144, 2017, doi: 10.1016/j.scitotenv.2017.01.102.

[13] H. Effendi, "River Water Quality Preliminary Rapid Assessment Using Pollution Index," *Procedia Environ. Sci.*, vol. 33, pp. 562–567, 2016, doi: 10.1016/j.proenv.2016.03.108.

[14] A. V. Vitianingsih, A. Choiron, D. Cahyono, A. Umam, and S. Suyanto, "Spatial Data Modeling on GIS for Classification of Measles-prone Region Using Multiple Attribute Decision Making," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 3, 2019, doi: 10.22266/ijies2019.0630.11.

[15] Menteri Negara Lingkungan Hdup, *Keputusan Menteri Negara Lingkungan Hidup Nomor : 115 Tahun 2003 Tentang Pedoman Penentuan Status Mutu Air*. 2003, pp. 1–15.

[16] D. Gardera *et al.*, *Indeks Kualitas Lingkungan Hidup Indonesia Tahun 2013*. Jakarta: Kementerian Lingkungan Hidup, 2013.

[17] M. S. Anang Widhi Nirwansyah, S.Pd., *Dasar Sistem Informasi Geografi dan Aplikasinya Menggunakan ARCGIS 9.3*, Pertama. Yogyakarta: Deepublish, 2017.

[18] S. Adinugroho and Y. A. Sari, *Implementasi Data Mining Menggunakan Weka*, Pertama. Malang: UB Press, 2018.

[19] J. Han, M. Kamber, and J. Pei, *Data mining: Data mining concepts and techniques*, Ketiga. Waltham, USA: Elsevier Inc, 2012.