PAPER • OPEN ACCESS

The determination of groundwater flow system using several deterministicand classical methods in Limboto-Gorontalo Lowland, Gorontalo Province

To cite this article: Nurfaika et al 2020 IOP Conf. Ser.: Earth Environ. Sci. 485 012039

View the article online for updates and enhancements.

The determination of groundwater flow system using several deterministicand classical methods in Limboto-Gorontalo Lowland, Gorontalo Province

Nurfaika¹*, S Purnama², and Hartono²

¹Doctoral Student of Geography Department, Faculty of Geography, University Gadjah Mada, Jl. Kaliurang, Sekip Utara, Bulaksumur Sinduadi Sleman, Yogyakarta, Indonesia

² Faculty of Geography, University Gadjah Mada, Jl. Kaliurang, Sekip Utara, Bulaksumur Sinduadi Sleman, Yogyakarta, Indonesia

* Nurfaika@ung.ac.id

Abstract. The Limboto-Gorontalo lowland is the recharge area of the Gorontalo Groundwater Basin in Gorontalo Regency. There has been a considerably wide gap, with floods during rainy seasons but water crisis in dry seasons, which creates animbalanced spatial distribution of groundwater. For this reason, groundwater flow system analysisisnecessary for efficient water resource planning and management. This research was designed to compare several deterministic methods and evaluate how deterministic and classical methods could identify and map the spatial distribution of the groundwater flow system in the Limboto-Gorontalo lowland. This system was determined by monitoring and measuring 157 samples of dug wells. The obtained data were interpolated using two GIS-based deterministic methods, namely inverse distance weighted (IDW) and radial basis function (RBF), and one classical technique, i.e., manual three-point problem method (3PM). Mean absolute error (MAE) and root mean square error (RMSE) values were used to evaluate and decide the best interpolation method. After comparing MAE and RMSE values and analyzing the resulting maps, the RBF was proven to be better than the IDW method. Moreover, the maps generated by RBF and 3PM did not differ significantly. The spatial analysis results showed that the groundwater flow system in the study area originated in the hilly areas in the north and south of Lake Limboto and flew to several valleys and rivers instead of directly accumulating in this lake.

1. Introduction

The demand for groundwater continually increases with population growth and development rate. Water availability plays a crucial role in human life, and it can even be one of the inhibiting factors of economic growth in the country [1]. The lowlands around Lake Limboto in Gorontalo Regency stretch from the east to the west or toward the Paguyaman Valley. The spatial imbalance of groundwater availability is a problem frequently occurring in them. As part of the development acceleration program, water drinking and clean water provision services must reach the target of MDGs, that is, 68.87% of the entire population in the regency. According to the Strategic Plan for 2012-2017, only 46.01% of them had access to drinking water service in 2011[2]. In other words, more than half of the population (53.99%) remained unserved and had to obtain clean water from other sources, particularly groundwater.

Water level measurement is essential in hydrogeological studies. Groundwater level data can be used for various purposes, such as hydraulic gradient calculation to determine groundwater flow direction and create flow nets. Hydraulic gradient or groundwater flow direction can assist in planning drilling activities, estimating hydraulic conductivity, porosity, and water flow velocity based on Darcy's law, and modeling groundwater to predict future conditions of aquifers [3][4][5].

There have been previous hydrogeological studies using GIS and quantitative methods, namely deterministic and stochastic. For example, IDW and RBF are the deterministic methodsused to estimate the spatial distribution of the water table. Meanwhile, ordinary kriging (OK) is an example of interpolation with a stochastic simulation technique. Based on the calculated root-mean-square error (RMSE), mean absolute error (MAE), and R^2 , this research decides RBF as the best method [6]. In

another study, after comparing IDW, RBF, global polynomial interpolation (GPI), local polynomial interpolation (LPI), and OK,it is concluded that OK is the most favorable method to interpolate electrical conductivity data spatially [7]. Meanwhile, when compared with IDW (deterministic),OK, and universal kriging(stochastic), RBF is the best interpolation technique for the spatial analysis of water table data [8]. The deterministic method has a disadvantage; that is, the predicted minimum and maximum values are always below the observed minimum and maximum values of the samples [9]. Using standard error values for evaluation, the above researchers have found different best interpolation methods. Therefore, this study compares the GIS-based deterministic method with the classical method named3PM to spatially analyze the groundwater flow system in the Limboto-Gorontalo lowland accurately. This deterministic study evaluated IDW set at powers of 2, 4, and 6 and RBF with the CRS, SPT, and inverse multiquadric (IMQ) to find the most accurate interpolation method. The best method is the one with the smallest statistical error values, i.e., root-mean-square error (RMSE) and mean absolute error (MAE) [7]. For the spatial analysis of groundwater flow in this study, the water level map generated bythe best deterministic method was compared with the one made by classical interpolation (3PM).

2. Methods

2.1. Study area

As written in [12], the study area was, geomorphologically,part of the Limboto-Gorontalo valley, but in [13], it wasreferred to as the Limboto plains. The Limboto-Gorontalo alluvial plains werewithin the administrative borders of Gorontalo Regency, and they extended from the east to the west, i.e., toward the Paguyaman valley. This study focused on the Limboto-Gorontalo alluvial plain,which layover the Lake Sediment formation unit. Based on the 1:100,000 Geologic Map of Limboto [14] and the 1:50,000 Geologic Map of Limboto [15], this landform consisted of one formation unit, namely, the lake sediment formed in the Quaternary Period (Holocene-Pleistocene). The lake sediment wascomposed of claystone, sandstone, and gravel; and according to borehole data, it was94 m thick [16]. The groundwater flow system analysis was limited to unconfined aquifers because the map output wasa medium-scale map (1:50,000) and the focus of the study wasthe eastern part of the Limboto-Gorontalo alluvial plain (the lowlands around Lake Limboto). The map of the research location waspresented in Figure 1.

2.2. Data collection

The groundwater level was measured at 157 samples of dug wells. This study employed a systemic random sampling technique with the grid system to ensure that the evenly distributed samples (i.e., training wells) were representative of the population [17]. The research area was divided into $1x1km^2$ grids; the sampling within each grid was determined randomly based on the existence of dug wells in the field. The map of the training wells is presented in Figure 1. The water table measurement included pinpointing the coordinates of the dug wells and measuring the surface elevation and depth of groundwater. The coordinates and surface elevations were determined with Garmin GPSMAP 64s and by contour analysis on the 1:25.000 Indonesian Topographical Map. Meanwhile, the depth of the water table was identified with a rolling measuring tape.

2.3. Interpolation methods

Although environmental conditions sometimes constrain data collection in the field and create limited samples, a studycan still proceed with representative data by applying a spatial prediction method. An interpolation is an approach to address data limitation by predicting the unknown values (predicted values) with the values of the surrounding sample points (observed values) [18].

The deterministic and geostatistical methods often employ interpolation. There are two types of deterministic interpolation, namely, global and local. Global interpolation uses the entire dataset for prediction. In contrast, local interpolation uses only a few data for prediction and based on measured sample points, e.g., inverse distance weight (IDW) and radial basis function (RBF) [19].



2.3.1. Inverse Distance Weight (IDW)

Inverse distance weighted (IDW) interpolation is an approach frequently used in GIS [19]. It is included in the deterministic method that predicts samples with unknown values from the sample points around them. The samples are weighted based on distance; therefore, the nearest point with known value has a higher weight value [10]. The formula for the IDW method is as follows (equation 1).

$$z(\mathbf{x}_{0}) = \frac{\sum_{i=1}^{n} z(x_{1}).d_{i0}^{-r}}{\sum_{i=1}^{n} d_{i0}^{-r}} (1)$$

where $z(x_0)$ is a function of the neighboring observation at the location (x_0) , n is representing the total number of sample data, its the nearest neighbors (i= 1,2,3...n), r is an exponent that determines the weight assigned to each observation, and d is the distance separating the location of prediction (x_0) from the location of observation (x_i) . In the GIS software, the IDW method has several power settings. Power is the function that determines the interpolation results. A power coefficient of 2 is the default setting of IDW interpolation in ArcGIS [9].

2.3.2. Radial Basis Function (RBF)

RBF belongs to the exact interpolator group that controls the suitability of each surface based on the measured samples, and the basis of its calculation is the distance between the interpolated points and the known sample points [6]. Conceptually, RBF is similar to an artificial nervous system in which

samples are measured by minimizing the total curvature of the surface [20]. The mathematical formulation of the RBF method is as follows (equation 2).

$$Z(\mathbf{x}) = \sum_{i=1}^{m} a_i f_i(\mathbf{x}_k) + \sum_{j=1}^{n} b_{i_j}(d_j) \quad \text{for } \mathbf{k} = 1, 2, \dots, n$$

$$\sum_{i=1}^{m} a_i f_i(\mathbf{x}_k) = 0 \quad \text{for } \mathbf{k} = 1, 2, \dots, m.$$
(2)

RBF consists of five functions, namely thin-plate spline (TPS), spline with tension (SPT), completely regularized spline (CRS), multiquadric (MQ), and inverse multiquadric (IMQ) functions [21]

2.3.3. The Classical Three-Point Method (3PM)

In general, groundwater movesfollowing the hydraulic gradient, that is, from the head (hydrostatic pressure) to a lower point [5]. Relative to the spatial and temporal evaluation in the field, the three-point problem method is a faster, more straightforward, and more cost-effective technique to estimate hydraulic gradient and, at the same time, groundwater flow. It determines the magnitude and direction of the hydraulic gradient from three data points (head data), which, in this study, represented three dug or bore wells [22].

2.4. Cross-validation and comparison methods

Cross-validation and comparisonwere carried out to select the best method for generating a groundwater contour map and determining the groundwater flow system. Cross-validation aims to assess whether or not an analysis results in independent data sets. Many studies have applied this validation model, e.g., [23] partitions the research data sets into 80% training data and 20% testing data, while [24] uses 65% of the research samples as training data and the other 35% as testing data. In this research, the cross-validation divided the 157 samples of wells into two sets, namely 109 wells (70%) as training data and 48 wells (30%) as testing data. The spatial distribution of the training and testing data is depicted in Figure 5. The interpolation results were later tested for accuracy based on their error values, namely MAE and RMSE, which were computed using equations 3 and 4.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |Z_i - Z|(3)$$

$$RMSE = \sqrt{\frac{\sum (Z_{\bar{l}} - Z)^2}{n}} (4)$$

whereZi is the predicted value, Z is the observed value, and n is the total number of actual data. MAE and RMSE show the level of data accuracy in relation to the coordinate system. The smaller the MAE and RMSE, the closer the predicted values to observed values. In other words, the smallest MAE and RMSE signify the best deterministic method [10][11]

The products of the deterministic and classical methods are water table maps consisting ofequipotential lines that are depicted with color gradation. The maps generated from the selected deterministic method and the classical method (3PM) were compared by observing the features they displayed. These maps were further processed to produce flow net maps containing information for the spatial analysis of the groundwater flow system.

3. Results and Discussion

3.1. Water table interpolation

According to Muehrcke[1992, cited in 18], data samples may be limited due to some encountered hurdles or unfavorable environmental conditions during data collection in the field. Nevertheless, a spatial prediction technique named interpolation has been extensively used to address this flaw by estimating the unknown value of one point with the values of the surrounding points. Both deterministic and geostatistical methods are the most applied interpolation techniques [18]. [19]divides interpolation into two, namely global and local interpolation. Global interpolation uses the entire dataset for prediction, while local interpolation relies on only a few data for estimation, e.g., Inverse Distance Weighted (IDW) and Radial Basis Function (RBF) [19].

The deterministic method produces a surface model from the sampling points based on the similarity of values between neighboring samples (inverse distance weighted) or the radial basis

IOP Publishing

function. Among the 157 well samples, the lowest water table was 1 m, whereas the highest one was 64 m. Accordingly, the average water table was 16 m with a rather wide standard deviation, i.e., 13 m, which potentially created variations in the interpolation results.

3.1.1. The results of the Inverse Distance Weighted (IDW) interpolation

Deterministic and stochastic interpolation methods differ from each other. The former depends on mathematical computation, while the latter is a statistical approach. The stochastic method in the ArcGIS application is included in the geostatistical feature. The IDW method is commonly called the weighted average method because it predicts values based on distance and weight. In this study, the IDW interpolation involved simulating several changes in weight values, which were identical to the power setting that shapes the results of the interpolation in ArcGIS. Then, it was set at powers of 2, 4, and 6. This technique was selected based on the cross-validation procedure that categorized 70% of the total data as training data (interpolation samples) and the remaining 30% as testing data (test samples). The results were later subjected to validation by MAE and RMSE. The spatial distribution of the wells acting as either training or testing data is presented in Figure 5.



Based on the results of the cross-validation analysis and MAE and RMSE calculations, the power of 4 produced the lowest MAE, i.e., 6.3%. However, there was no significant difference between the error values of the three IDW methods. The calculated error values of the three methods are listed in Table 1

Table 1. The MAE and RMSE values of the IDW	
internalation results with three nower settings	

interpolation results with three power settings				
IDW Methods	MAE (%)	RMSE (%)		
Power of 2	6.401	25.300		
Power of 4	6.397	25.300		
Power of 6	6.401	25.300		

Figure 6 shows the comparison of the water table maps generated by the IDW interpolations using three different power settings (2, 4, and 6).



IOP Publishing

The cross-validation analysis and MAE and RMSE calculation revealed that among the three power settings, the power of 4 yielded the best IDW interpolation. However, when observed from the appearances of the resulted maps, the power of 2 produced a smoother contour line structure, though it was not significantly different from the powers of 4 and 6.

3.1.2. The results of the Radial Basis Function (RBF) interpolation

The RBF method takes into account surface controls generated from the measured values, and it is computed based on the distance between interpolated points and sample points with known values. Among the five functions of RBF, only three were used in this research, namely CRS, SPT, and IMQ [21]. The cross-validation analysis and MAE and RMSE calculation results proved that the CRS, SPT, and IMQ functions produced similar values, with SPT having the smallest MAE. The calculated error values are listed in Table 2, while the RBF interpolation results are visualized in Figure 7. Except for SPT, the RBF interpolations using the CRS and IMQ functions showed similar results. Although not significant, the error values of SPT were lower than the other two functions.

Table 2. The MAE and RMSE values of the RBF
interpolation results using three functions

RBF Method	ds MAE (%)RMS	MAE (%)RMSE (%)	
CRS	6.396672	25.30088	
SPT	6.394281	25.30088	
IMQ	6.395980	25.30088	



Referring to the MAE and RMSE values and the resulted maps, the IDW processed with a power of 4 (MAE= 6.397) was the best interpolation method. Meanwhile, among the three RBF methods (CRS, SPT, and IMQ), the SPT function generated the best interpolation with MAE= 6.394. Because the SPT function (RBF) had smaller MAE than the IDW with a power of 4, the former created a prediction that wascloser to the observed values than the latter. Therefore, in the spatial analysis of the groundwater flow system, the deterministic RBF with SPT function was compared with the three-point problem method (3PM).

3.1.3. Three-Point Problem Method (3PM)

3PM is a classical interpolation method with manual calculation and drawing. It predicts values from three adjacent well samples with known water tables. In this study, 3PM was employed to calculate and determine the hydraulic gradient or flow direction. The output is a map consisting of equipotential lines (i.e.,groundwater contours) and hydraulic gradients (groundwater flow lines). The results of the 3PM method are presented in Figure 8.

Most studies on spatial analysis of groundwater level generally use and compare modern or GISbased interpolation methods. In this study, the said spatial analysis was carried out using not only GISbased deterministic methods but also aclassical technique. Figure 8 below compares the groundwater flow systems produced by two deterministic methods (RBF with SPT function) and one classical method (3PM). There were no significant differences between the maps generated by the RBF with the SPT function and the 3PM method. Similarities were detected in the equipotential line structure or

contour line. For instance, the red squares in Figures 8a and 8b mark closed contours, i.e., signs of shallow water tables, at the same locations.



3.2. Groundwater flow system

The spatial analysis of the groundwater flow system was carried out based on the two water table mapsabove. Both maps were processed to create flow nets. The equipotential lines or groundwater contours were observed while drawing the flow lines (the black arrows in Figures 7a and 7b). The flow line is a representation of the flow system or, in other words, shows the direction of the groundwater movement in an area. Based on the flow net analysis (Figures 7a and 7b), the groundwater flow pattern in the Limboto-Gorontalo lowlands is assumed to originate in the hilly areas in the north and south of the lake and mostly flow to the areas with shallow water tables (red squares in Figures 7a and 7b) before eventually discharged into the lake. In other words, it accumulates in several regions with low hydraulic heads or pressures.

In previous research, several GIS-based interpolation methods were merely compared using standard error values [6][8][23]. In contrast, this study analyzed groundwater flow systems based on groundwater contour maps that were generated using some GIS-based interpolation methods and evaluated or cross-checked with the results of the 3PM manual interpolation. As explained in Chapter 3.1.2 and Chapter 3.1.3, RBF with the SPT function was the interpolation method that produced the smallest MAE and RMSE values and the same flow line structure as the groundwater contour map created by manual interpolation. Consequently, RBF with the SPT function is the best method to map the groundwater flow systems in the study area.

4. Conclusions

Cross-validation by MAE and RMSE has shown that IDW set at a power coefficient of 4 is better than powers of 2 and 6 and that RBF with SPT function is better compared to the CRS and IMQ functions. Because the SPT function has smaller MAE than the IDW with a power of 4, it creates the best prediction among the tested deterministic interpolation methods in this study. The maps generated by the RBF method with SPT function and the 3PM method have no significant differences in groundwater contour, and there are spots with closed contours formed at the same locations in both maps.

Acknowledgments

Authors would like to thank all parties, particularly the students of the Geography and Geology Departments in the State University of Gorontalo, for their assistance during groundwater data collection in the field.

References

[1] Schouten M 2006 Integrated Water Resources Management Institutions and Livelihoods Under Stress(New York: CRC Press)

- [2] Public Works Office of Gorontalo Province2012 Strategy Plan in 2017-2022
- [3] Todd D K and Mays L W 2005Groundwater Hydrology (New York: John Willey and Sons)
- [4] Hiscock K M 2005 Hydrogeology Principles and Practice(Australia: Blackwell)
- [5] Hudak P F 2000 Principles of Hydrogeology second edition(New York:CRC Press)
- [6] Arslan H 2014 Environ Monit Assess. 186(8) 5123-34
- [7] Seyedmohammadi J,Esmaeelnejad L, Shabanpour M 2016 Model Earth Syst Environ. 2(4) 1-10
- [8] Adhikary P P, Dash C J 2017 Appl Water Sci7(1)339-48
- [9] Bronowicka, Mielniczuk J, Obroslak R, Przystupa W 2019Int J Environ Res. 13 (4) 679-87
- [10] Varouchakis EA, Hristopulos DT 2013 Environ Monit Assess. 185(1) 1–19
- [11] Xiao Y, Gu X, Yin S, Shao J, Cui Y, Zhang Q 2016 SpringerPlus. 5(1) 425
- [12] Katili J A 1970*GeolRundsch*. 59(2):581–600
- [13] Trail D S, John T V, Bird M C, Obial R C, PetzelB A, AbiongD B, Parwoto, and Sabagio1974 The general geological survey of Block 2 Nort Sulawesi, Indonesia(Internal Report.) (Jakarta-Indonesia: PT. Tropic Endeavour)
- [14] Bawono S, Partoyo E, Wirosujono S, Situmorang R I, and Spanjoro R J 1990Geological Map of the Limboto Quadrangle, Sulawesi, 1:100.000 (Bandung-Indonesia: Geological Research and Development Centre)
- [15] Sidarto, Kusdji DK 2010 (Geological Research and Development Centre)
- [16] Bachri S and Apandi T 1997 Geological Research and Development Centre)
- [17] Yang F, Cao S, Liu X, Yang K 2008 J Hydrodyn. 20(3) 339-46
- [18] DeMers MN 2009 Fundamentals of Geographic Information Systems (Hoboke: John)
- [19] Christopher DL2007Local models for spatial analysis (New York: CRC Press)
- [20] Mitasova H, Mitas L, 1993Math Geol. 25641-655
- [21] Xie Y, Chen T, Lei M, Yang J, Guo Q, Song B, et al 2011 Chemosphere.82(3) 468-76
- [22] Beljin M, Ross RR, Acree S D 2014 3PE: Environmental Protection Agency) p 60
- [23] Ahmadi M K, Dezfouli ABZ 2012 Journal of Academic and Applied Studies. 2(11) 56-69
- [24] Tutmez B and Hatipoglu Z 2010 Ecological Informatics.5 311–315