

Spatiotemporal geostatistical modeling of groundwater in arid region during dry and wet seasons: a case study The Experimental Station at Hadat Ash-Sham Farm, Saudi Arabia

Jaka Budiman¹, Nassir Al-Amri¹, Anis Chabaani¹, and Amro Elfeki¹

¹King Abdulaziz University

August 24, 2020

Abstract

Saudi Arabia lies in an arid region where groundwater has become main resource. Due to the increase of water demand; therefore, it is essential to understand groundwater dynamics for the best groundwater management practice in Saudi Arabia. In Hadat Ash-Sham Farm Experimental Station, Saudi Arabia, water table data from 11 wells and rainfall data were monitored for 16 months. These water table data is analyzed using the geostatistical method with ordinary Kriging technique, to generate the best water table spatial distribution map for each month and the expected flow direction. The cross-validation technique is used to evaluate the quality of the developed water table maps. The Kriging maps show two regimes: weak spatial dependence (WSD, the ratio of the nugget to sill >75%) and strong spatial dependence (SSD, the ratio of the nugget to sill <25%). The WSD regime happens during dry seasons, while the SSD happens during wet seasons. The SSD gives better results and better accuracy compared to WSD. The root-mean-square error (RMSE) of water table varies between 0.26 – 3.4 m in the case of SSD, while it varies between 0.51-4.8 m in the case of WSD. Water table maps show groundwater flow direction in the study area is from East to West and South-East to North-West during the wet season (SSD). This direction is parallel with the relative orientation of surface stream with higher elevation to the surface stream with lower elevation, where study area is between these surface streams. While during the dry season (WSD), there is no preferred direction since there is almost no flow.

KEYWORDS

statistical modeling, Kriging technique, groundwater statistics, spatiotemporal variability, arid regions, Saudi Arabia

HIGHLIGHTS

1. Geostatistical analysis using water table data shows different of spatial dependence during dry season and wet season.
2. Water table maps using Kriging technique shows distinctive groundwater flow direction during dry and wet seasons. No preferred groundwater flow direction during dry season, while during wet season, the groundwater flow direction is similar to the surface streams orientation.
3. Cross-validation of water table maps during wet season shows better quality than water table maps during dry season.

1 | INTRODUCTION

The fast growth in the population nowadays has made increasing in freshwater demand globally. The population growth has impacted on the land-use changes and consequently influenced the availability of

freshwater resources more than the impact of climate change (Okello et al., 2015). From ancient times until today, groundwater is still a vital resource (Fitts 2012).

Saudi Arabia lies in the arid region with low precipitation and no permanent rivers or natural lakes. Therefore, groundwater is a critical resource in the arid region, especially in Saudi Arabia (SA), since it lies in the desert region and the recharge of the groundwater is mostly depending on the rainfall. The average annual rainfall in Saudi Arabia in 2014 was only 59 mm/year (World Bank Group, 2019).

Increasing water demand in Saudi Arabia requires proper management and understanding of groundwater dynamics. It is necessary to derive a better understanding of water table fluctuations, based on coupling both spatial and temporal analyses. Geostatistics is the best method to study the variability in water resources (Rouhani,1989). Delhomme (1978) used a Kriging technique to analyze precipitation data. Spatial-temporal Kriging was applied to groundwater quantity in South Georgia with low estimation variances (Rouhani & Hall, 1989). Geostatistical methods are beneficial for water resources management studies (Kumar et al., 2005). The strong dependence of the Kriging technique was shown in spatial and temporal distribution analysis of 12 years water table decreased in Iran (Ahmadi & Sedghamiz, 2007). Strong dependence was shown in spatial analysis of water quality from several parameters Ca^{2+} , Mg^{2+} , Cl^- , NO_3^- , and salinity (Kuswantoro et al., 2014). Kriging technique has the advantage of getting the variance of the estimation error at any point without having measurement on the field (Ahmed & Devi, 2008). Based on all information above, it is confirmed that Kriging technique on geostatistical analysis is suitable to evaluate the spatial and temporal correlation structures in hydrological parameters.

Utilization of geographic information system (GIS) software (e.g., ArcGIS) in analyzing the spatial distribution of groundwater has been studied by many researchers such as Mehrjardi et al. (2008) and Kuswantoro et al. (2013). In GIS, geostatistical analysis can easily create continuous surface or map, such as elevation, depth and water table, or levels of pollution. (ESRI, 2010).

The objective of this study is to analyze and map the spatial distribution of water table, the temporal water table fluctuation, and groundwater flow direction in Experimental Station at Hadat As-Sham, Wadi Al-Lusub, Saudi Arabia. Water table data from 11 wells were observed in Hadat As-Sham every week for 16 months since November 2018. Analysis of the results from this study is required to evaluate the groundwater conditions for the best groundwater resources management in the farm.

2 | STUDY AREA

Wadi Al-Lusub is a watershed located in the western part of Saudi Arabia (Figure 1a). It is the upper part of the alluvial fan of Wadi Usfan, which is called Hadat Ash-Sham area. This area was well known in the past with its agricultural activities. At present, relatively larger areas are being cultivated at Hadat Ash-Sham, the agricultural experimental station of the Faculty of Environmental Sciences, King Abdulaziz University (KAU). The geographical coordinates of the station are $39^\circ 43' 15'' - 39^\circ 43' 57''$ and $21^\circ 47' 36'' - 21^\circ 54' 15''$ with total area 789.68 m^2 . The station has a total of 14 wells distributed over the farm. However, only 11 wells are working and used for pumping and monitoring. The 11 wells are monitored at static conditions and are shown in Figure 1b. Wadi Al-Lusub flows into Wadi Usfan that runs in East-West direction, having its outlet towards the Red Sea in the west. Wadi Al-Lusub area has an arid climate with an annual average rainfall rate is 113 mm/year verified at Madrasah village (El-Hames, 2005). This area has a temperature range between $13^\circ \text{C} - 40^\circ \text{C}$.

Geologically, the Hadat Ash-Sham experimental station lies on Precambrian – Cambrian complex with Cretaceous-Tertiary sedimentary formation sequences, and Tertiary – Quaternary – Recent alluvial deposit (El-Hames, 2005).

A geomorphological study is performed using ArcGIS software to analyze the hydrological and morphological features in the Al-Lusub Basin (Figure 1c). The watershed has a total area of 803.53 km^2 , with catchment

perimeter is 275.46 km, and the catchment mainstream length is 77282.76 m. The catchment elevation ranges between 224–1295 m above sea level. The stream is flowing from North East to the South West (Figure 1c and 1d). The topographic map (Figure 1d) shows two-stream branches: one is flowing to the north of the farm, and the other is flowing to the south of the farm. Topographically, the stream in the south is relatively at a higher elevation than the stream in the north. Flash floods often attract the farm from the stream in the north.

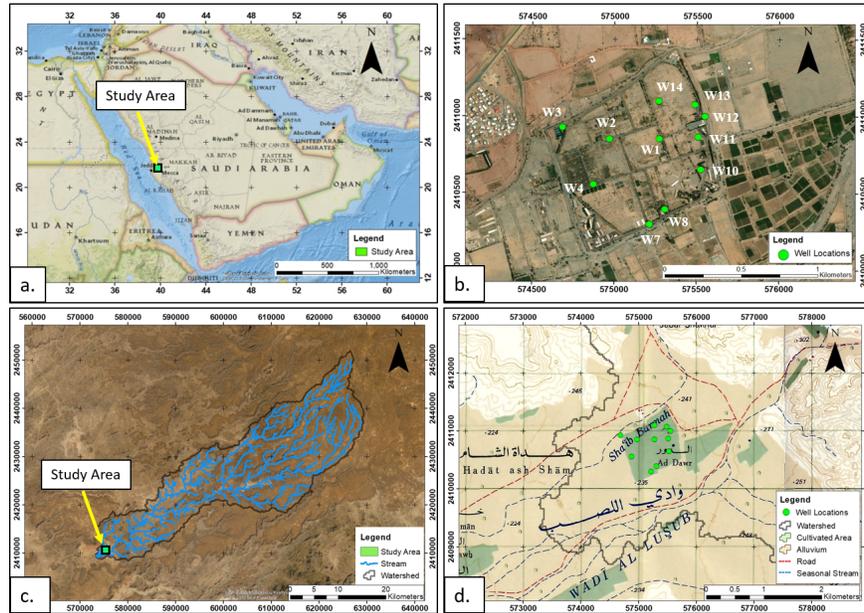


FIGURE 1 Hadat Ash-Sham experimental station location: (a) general areas, (b) distribution of well locations, (c) Watershed of the study area, (d) topographic map of the study area

3 | MATERIALS AND METHODS

The water table from the 11 wells at Hadat Ash-Sham station, together with rainfall information from three nearest rainfall stations, was observed every week from October 2019 until February 2020 (Figure 2). The coordinates of the wells were also recorded to support spatial analysis. Average monthly water table data is utilized for spatiotemporal geostatistical analysis at a monthly time scale, with ordinary Kriging technique, using ArcGIS software. It is obvious from Figure 2, that the water table for the period of November 2018 up to May 2019 shows that most of the wells do not change since it was almost dry period. The exception occurs on well 4, 7, and 8, where those three wells seem to be influenced by the rainfall-event in January 2019.

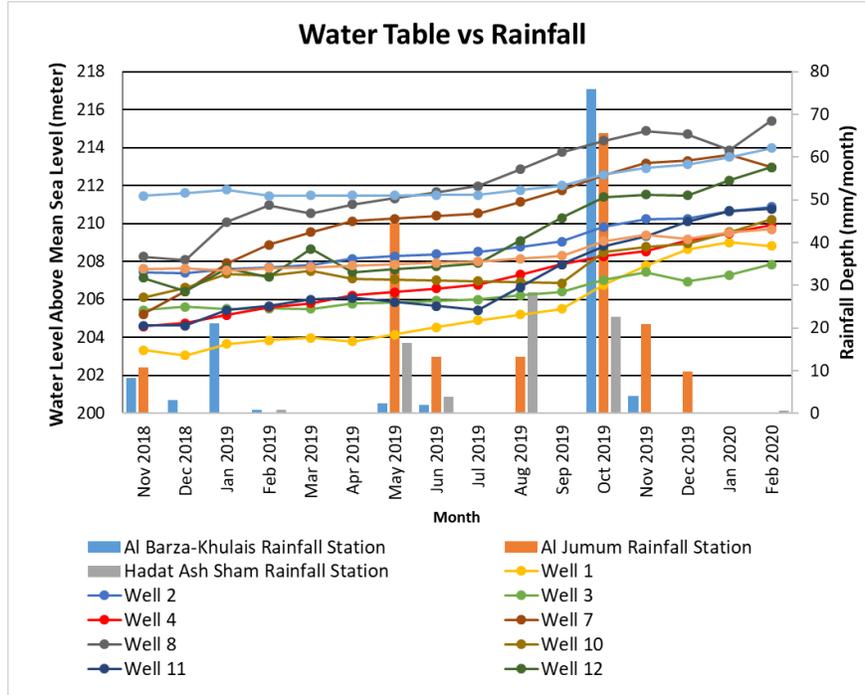


FIGURE 2. Water table data from 11 observation wells vs rainfall rate every month

The flowchart of this study can be seen in Figure 3. The coefficient of skewness can be derived from the statistical tools in GIS. Based on the coefficient of skewness, the original data from each well is transformed using logarithmic transformation. The purpose of the logarithmic transformation is to gain a better coefficient of skewness and to stabilize the variance. The limit of the excellent coefficient of skewness is ± 2 (Rubin, 2012), therefore if it is outside the threshold, the data is highly skewed, and logarithmic of transformation must be considered. In this case, the original data is within the range ± 2 ; however, by the log-transformation, it is slightly improved. Therefore, it has been decided to use the log-transformation for better predictions.

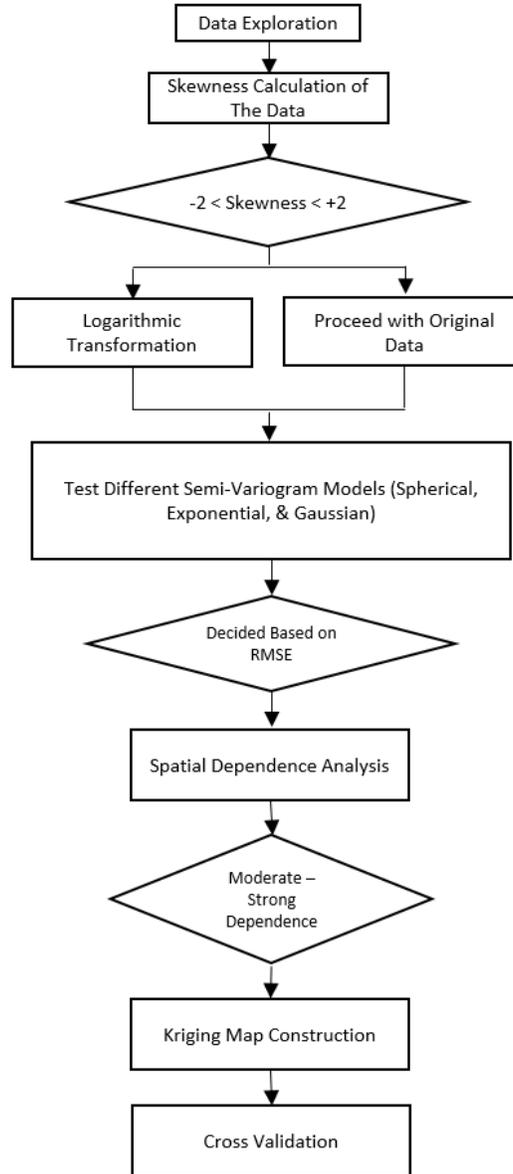


FIGURE 3 Flow chart for geostatistical analysis in this study

The semi-variogram is the fundamental tool of geostatistics to analyze whether spatial data is correlated and how far the correlation can be reliable. The common equation to calculate the semi-variogram model is by Matheron’s method of moments (MoM) prediction (Oliver and Webster 2015):

$$\hat{\gamma}(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \{z(x_i) - z(x_i + h)\}^2 \quad (1)$$

where $z(x_i)$ and $z(x_i + h)$ are the observed values of z at places x_i and $x_i + h$ and $m(h)$ is the number of paired comparisons at lag h .

The semi-variogram model contains three popular models, such as the gaussian model, the spherical model, and the exponential model (Lloyd, 2010).

One of the conventional measures of model performance is the root-mean-square error (RMSE) (Chai & Draxler, 2014). RMSE is a method, which used to derive the information of the best model from various semi-variogram models. The most excellent semi-variogram model can be selected from the lowest RMSE value. If the estimations are perfect, then the RMSE value should be zero. The RMSE value is obtained from the cross-validation step of the Kriging results.

The ratio of nugget and to the sill or dependence is used to see how well the data variance correlated with distance. The classification of this ratio is as follows: if the ratio is [?]25%, the variable is classified as strongly spatially dependent, if the ratio is between 25% and 75%, the variable is classified as moderately spatially dependent, and if the ratio is [?]75% the variable is classified as weakly spatially dependent (Cambardella et al., 1994).

Based on RMSE calculation and emphasized by the ratio of the nugget to the sill, the best semi-variogram model is selected. A selected semi-variogram will be used to generate a kriging map of the water table.

Ordinary Kriging based on the assumption that variation is random and spatially dependent. Kriging predicts value within some distances from sparse sample data. The estimation equation is given by:

$$\hat{z}(x_0) = \sum_{i=1}^N \lambda_i z(x_i) \quad (2)$$

where $\hat{z}(x_0)$ is the estimated value at the unsampled point x_0 , $z(x_i)$ is the observation data, N is the number of observations, and λ_i is the weights.

The next step is to find the weights that minimize the kriging variance conditional on the unbiasedness condition that the weights have to sum up 1,

$$\sum_{i=1}^N \lambda_i \gamma(x_i - x_0) + \psi(x_0) = \gamma(x_j - x_0) \text{ for all } j \quad (3)$$

$$\sum_{i=1}^N \lambda_i = 1 \quad (4)$$

Where $\psi(x_0)$ is the Lagrange multiplier which will be utilized to reach minimization.

Cross-validation is a method to evaluate the quality of the kriging map in predicting the unknown values. The basic procedure of cross-validation is one by one estimation method called “leave-one-out.” The method is performed considering all the data, then removing them one by one, and estimating the removed one by the rest of the data. It compares the value of measured and predicted, where the difference is called a prediction error (ESRI, 2010). The calculation of the prediction errors is utilized as the assessment of the best model for map production. The model performance can be assessed using the following criteria:

The mean prediction errors (MPE) equation is given by:

$$MPE = \frac{1}{n} \sum_{i=1}^n (\hat{z}_i - z_i) \quad (5)$$

Root-mean-square prediction errors (RMSE) equation is given by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{z}_i - z_i)^2} \quad (6)$$

The average standard error (ASE),

$$ASE = \sqrt{\frac{1}{n} \sum_{i=1}^n \hat{\sigma}_i} \tag{7}$$

The root-mean-square standardized errors (RMSSE) equation is given by:

$$RMSSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left\{ \frac{(\hat{z}_i - z_i)}{\hat{\sigma}_i} \right\}^2} \tag{8}$$

where z_i is the observed data, \hat{z}_i is the predicted data, $\hat{\sigma}_i$ is the prediction standard error for location i , and n is the number of sampling points.

The range of values of these criteria are:

- mean prediction error value close to 0, indicates the predictions are unbiased,
- root-mean-square standardized error prediction value is close to 1, indicates the standard errors are accurate,
- root-mean-square error and average standard error need to be as small as possible, so the predictions do not deviate too much from the measure values,
- on the QQ plot, using the root-mean-square standardized error (RMSSE), the values must be as close as possible with the 45 degrees straight line.

4 | RESULTS AND DISCUSSIONS

Figure 2 shows that in Hadat Ash-Sham observation wells, the water table was increasing from November 2018 to February 2020. The average water level elevation in November 2018 was 206.35 meters above the mean sea level (m asl), and the average water level elevation in February 2020 was 211.3 m asl; therefore, the total water increasing was 4.94 m. The highest increase was 8.38 m, which occurred in well 7. Based on rainfall observation from the nearest rainfall stations, the increasing water table is related to the increasing in rainfall within the given period.

TABLE 1 Correlation coefficient matrix of the water table in observation wells

| Well Code | W1 | W2 | W3 | W4 | W7 | W8 | W10 | W11 | W12 | W13 | W14 |
|------------|------|------|------|------|------|------|------|------|------|------|------|
| W1 | 1.00 | | | | | | | | | | |
| W2 | 0.99 | 1.00 | | | | | | | | | |
| W3 | 0.96 | 0.98 | 1.00 | | | | | | | | |
| W4 | 0.97 | 0.99 | 0.95 | 1.00 | | | | | | | |
| W7 | 0.88 | 0.91 | 0.86 | 0.95 | 1.00 | | | | | | |
| W8 | 0.90 | 0.93 | 0.90 | 0.96 | 0.96 | 1.00 | | | | | |
| W10 | 0.91 | 0.89 | 0.89 | 0.86 | 0.78 | 0.78 | 1.00 | | | | |
| W11 | 0.98 | 0.97 | 0.95 | 0.96 | 0.88 | 0.90 | 0.92 | 1.00 | | | |
| W12 | 0.96 | 0.96 | 0.95 | 0.95 | 0.86 | 0.91 | 0.89 | 0.98 | 1.00 | | |
| W13 | 0.95 | 0.93 | 0.94 | 0.89 | 0.74 | 0.79 | 0.95 | 0.95 | 0.93 | 1.00 | |
| W14 | 0.98 | 0.99 | 0.99 | 0.95 | 0.85 | 0.89 | 0.91 | 0.97 | 0.96 | 0.96 | 1.00 |

The correlation coefficient of all water table in observation wells shows that most values are close to 1, which

means all wells have a positive and a high correlation (Table 1). The lowest correlation coefficient value is 0.74 occurred in the correlation of well 7 and well 13, where both wells have the longest distance from each other.

Logarithmic data transformation is utilized to get a better skewness value, as shown in Table 2. Figure 4 shows that skewness values are slightly better after logarithmic data transformation.

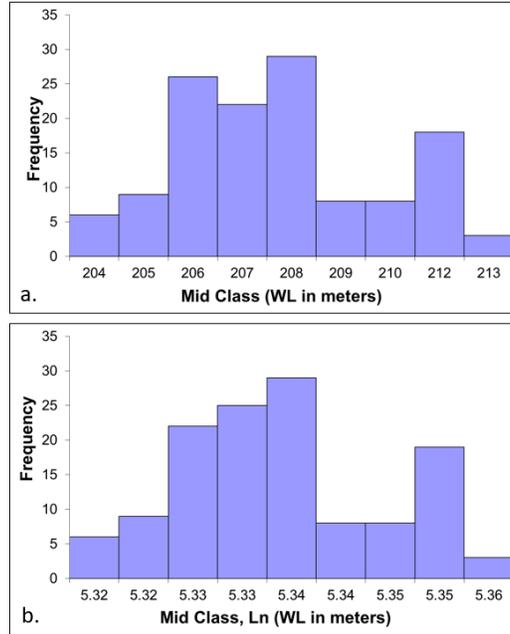


FIGURE 4 Histogram of all data from 11 wells and 16 months data: a. original data, b. logarithmically transformed data

TABLE 2 Statistical analysis of the water table in observation wells every month

| Time | Transformation | Min | Max | Mean | Standard Deviation | Skewness | Kurtosis |
|----------|----------------|--------|-------|--------|--------------------|----------|----------|
| Nov 2018 | None | 203.3 | 211.4 | 206.4 | 2.2 | 0.79 | 3.28 |
| | Log | 5.3 | 5.3 | 5.3 | 0.01 | 0.76 | 3.24 |
| Dec 2018 | None | 203.05 | 211.6 | 206.56 | 2.2325 | 0.71 | 3.66 |
| | Log | 5.3 | 5.3 | 5.3 | 0.010 | 0.67 | 3.62 |
| Jan 2019 | None | 203.6 | 211.7 | 207.2 | 2.2 | 0.43 | 2.72 |
| | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.41 | 2.70 |
| Feb 2019 | None | 203.8 | 211.4 | 207.4 | 2.3 | 0.40 | 2.34 |
| | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.38 | 2.33 |
| Mar 2019 | None | 203.9 | 211.4 | 207.6 | 2.2 | 0.10 | 2.12 |
| | Log | 5.3 | 5.3 | 5.3 | 0.010 | 0.09 | 2.11 |
| Apr 2019 | None | 203.7 | 211.4 | 207.7 | 2.3 | 0.20 | 2.19 |
| | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.18 | 2.20 |
| May 2019 | None | 204.1 | 211.4 | 207.8 | 2.3 | 0.30 | 2.05 |
| | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.28 | 2.06 |
| Jun 2019 | None | 204.5 | 211.6 | 207.9 | 2.3 | 0.37 | 1.97 |
| | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.36 | 1.97 |
| Jul 2019 | None | 204.9 | 211.9 | 208.04 | 2.3 | 0.43 | 1.94 |

| | | | | | | | |
|----------|------|--------|-------|--------|-------|------|------|
| Aug 2019 | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.41 | 1.94 |
| | None | 205.2 | 212.8 | 208.5 | 2.4 | 0.46 | 2.01 |
| Sep 2019 | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.45 | 2.00 |
| | None | 205.5 | 213.7 | 209.05 | 2.5 | 0.43 | 2.05 |
| Oct 2019 | Log | 5.3 | 5.3 | 5.3 | 0.012 | 0.41 | 2.04 |
| | None | 206.7 | 214.3 | 209.9 | 2.4 | 0.43 | 1.99 |
| Nov 2019 | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.42 | 1.98 |
| | None | 207.4 | 214.8 | 210.6 | 2.4 | 0.56 | 2.05 |
| Dec 2019 | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.55 | 2.04 |
| | None | 206.94 | 214.7 | 210.5 | 2.3 | 0.40 | 2.12 |
| Jan 2020 | Log | 5.3 | 5.3 | 5.3 | 0.011 | 0.38 | 2.12 |
| | None | 207.2 | 213.8 | 210.8 | 2.1 | 0.10 | 1.85 |
| Feb 2020 | Log | 5.3 | 5.3 | 5.3 | 0.010 | 0.09 | 1.86 |
| | None | 207.8 | 215.4 | 211.2 | 2.3 | 0.37 | 2.11 |
| | Log | 5.3 | 5.3 | 5.3 | 0.010 | 0.35 | 2.10 |

Using logarithmically transformed data, the semi-variogram model generated from each well. Root-mean-square error (RMSE) calculation is used to evaluate the best model fitted for each month, which indicated the smallest RMSE value. Table 3 shows that the exponential semi-variogram model type is the best for the experimental semi-variogram.

TABLE 3 RMSE value of semi-variogram for each month

| No. | Month | Lag Size (m) | RMSE for Models using Logarithmic Transformation of the Data | RMSE for Exponential |
|-----|----------|--------------|--|----------------------|
| | | | Spherical | |
| 1 | Nov 2018 | 156 | 2.3469 | 2.3469 |
| 2 | Dec 2018 | 156 | 2.3983 | 2.3983 |
| 3 | Jan 2019 | 156 | 2.5225 | 2.5225 |
| 4 | Feb 2019 | 156 | 2.5953 | 2.5953 |
| 5 | Mar 2019 | 156 | 2.5280 | 2.5280 |
| 6 | Apr 2019 | 156 | 2.6570 | 2.6570 |
| 7 | May 2019 | 156 | 2.6662 | 2.6662 |
| 8 | Jun 2019 | 79 | 2.6997 | 2.6858 |
| 9 | Jul 2019 | 93 | 2.7025 | 2.6661 |
| 10 | Aug 2019 | 120 | 2.4030 | 2.2597 |
| 11 | Sep 2019 | 102 | 2.3503 | 2.1350 |
| 12 | Oct 2019 | 115 | 2.1126 | 1.9436 |
| 13 | Nov 2019 | 155 | 2.0706 | 1.9804 |
| 14 | Dec 2019 | 420 | 2.0693 | 1.9712 |
| 15 | Jan 2020 | 420 | 1.7714 | 1.7573 |
| 16 | Feb 2020 | 155 | 2.0934 | 1.9913 |

TABLE 4 The most suitable semi-variogram based on spatial dependence

| Time | Data Transformation | Model | Nugget | Partial Sill | Sill | Range (m) | Nugget/Sill |
|----------|---------------------|--------------|----------|--------------|----------|-----------|-------------|
| Nov 2018 | Log Transformation | Sph\Exp\Gau* | 0.000117 | 0.00 | 0.000117 | 0 | 100.00% |
| Dec 2018 | Log Transformation | Sph\Exp\Gau* | 0.000116 | 0.00 | 0.000116 | 0 | 100.00% |
| Jan 2019 | Log Transformation | Sph\Exp\Gau* | 0.000123 | 0.00 | 0.000123 | 0 | 100.00% |
| Feb 2019 | Log Transformation | Sph\Exp\Gau* | 0.000125 | 0.00 | 0.000125 | 0 | 100.00% |
| Mar 2019 | Log Transformation | Sph\Exp\Gau* | 0.000120 | 0.00 | 0.000120 | 0 | 100.00% |

| | | | | | | | |
|----------|--------------------|--------------|----------|----------|----------|------|---------|
| Apr 2019 | Log Transformation | Sph\Exp\Gau* | 0.000129 | 0.00 | 0.000129 | 0 | 100.00% |
| May 2019 | Log Transformation | Sph\Exp\Gau* | 0.000129 | 0.00 | 0.000129 | 0 | 100.00% |
| Jun 2019 | Log Transformation | Exponential | 0.000125 | 0.000005 | 0.000131 | 359 | 95.86% |
| Jul 2019 | Log Transformation | Exponential | 0.000115 | 0.000020 | 0.000135 | 758 | 85.41% |
| Aug 2019 | Log Transformation | Exponential | 0.000046 | 0.000110 | 0.000156 | 1144 | 29.33% |
| Sep 2019 | Log Transformation | Exponential | 0 | 0.000183 | 0.000183 | 1159 | 0.00% |
| Oct 2019 | Log Transformation | Exponential | 0 | 0.000168 | 0.000168 | 1190 | 0.00% |
| Nov 2019 | Log Transformation | Exponential | 0 | 0.000184 | 0.000184 | 1860 | 0.00% |
| Dec 2019 | Log Transformation | Exponential | 0 | 0.000264 | 0.000264 | 4119 | 0.00% |
| Jan 2020 | Log Transformation | Exponential | 0 | 0.000232 | 0.000232 | 4804 | 0.00% |
| Feb 2020 | Log Transformation | Exponential | 0 | 0.000158 | 0.000158 | 1529 | 0.00% |

*the semi-variogram model shows the same model for spherical, exponential, and gaussian model, which is flat/straight line.

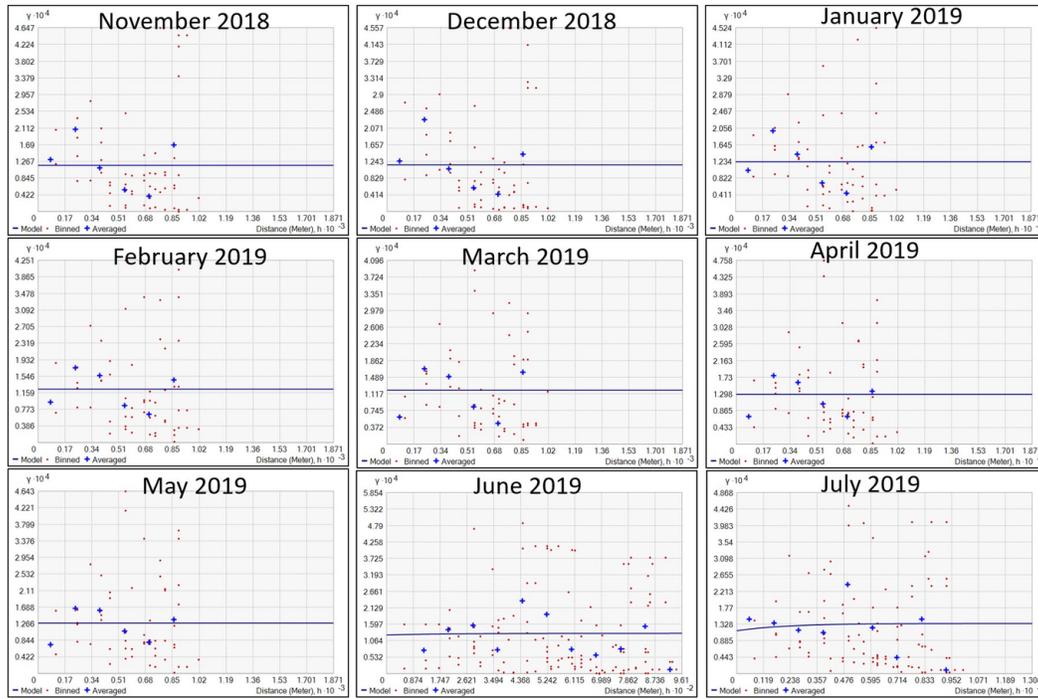


FIGURE 5 Semi-variogram with weak spatial dependence during the period of November 2018 - July 2019

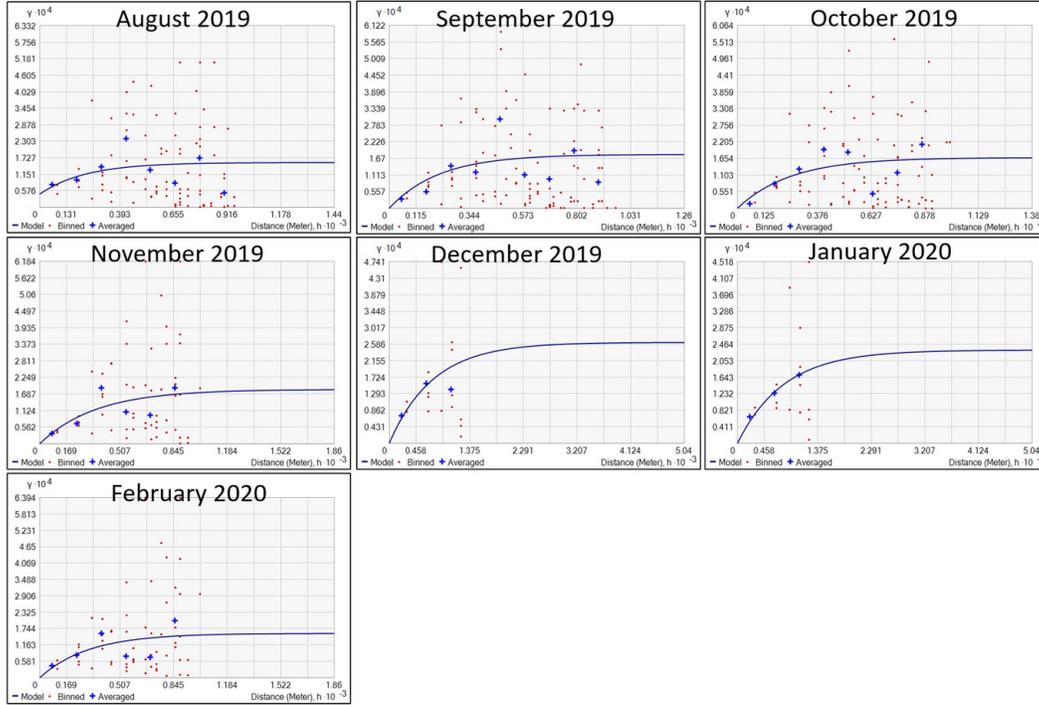


FIGURE 6 Semi-variogram with strong spatial dependence during the period of August 2019 - February 2020

Table 4 shows the spatial dependence of water table data from November 2018 until July 2019 are classified as weak spatial dependence (>75%). A flat semi-variogram indicates weak dependence, with most range value is 0 m, and it shows no spatial correlation (Figure 5).

Moderate – strong spatial dependence is shown from water table data from August 2019 until February 2020, which indicates spatial correlation with the correlation range average within this period is 2258 m (Figure 6). From this point forward, the analysis will be separated into two groups data: weak spatial dependence (November 2018 – July 2019) and strong spatial dependence (August 2019 – February 2020).

Cross-validation for all data is done to see the accuracy of the Kriging map, based on the predicted and measured data (Table 5). For weak spatial dependence, the mean error is negative, which means that the predicted value is lower than the actual measurement. For strong spatial dependence, the mean error is positive, which means that the predicted value is higher than the actual measurement. But overall, all the mean error values are close to 0, which means the predictions are unbiased. Most of the root-mean-square standardized error values close to 1, which indicates the standard errors are normally distributed. RMSE value for weak spatial dependence data is higher than RMSE value for strong spatial dependence data. It demonstrates that weak spatial dependence has a higher variance compared to strong spatial dependence. Similar circumstances have occurred for average standard error, where the weak spatial dependence data has a higher value compared with the strong spatial dependence.

TABLE 5 Cross-validation results of Kriging map for the period of November 2018 - February 2020

| Spatial Dependence | Months | Mean Error (m) | RMSSE | RMSE (m) | Average Standard Error (m) |
|--------------------|----------|----------------|-------|----------|----------------------------|
| Weak | Nov 2018 | -0.19 | 1.00 | 2.35 | 2.36 |
| | Dec 2018 | -0.17 | 1.02 | 2.40 | 2.35 |
| | Jan 2019 | -0.16 | 1.04 | 2.52 | 2.42 |

| Spatial Dependence | Months | Mean Error (m) | RMSSE | RMSE (m) | Average Standard Error (m) |
|--------------------|----------|----------------|-------|----------|----------------------------|
| Strong | Feb 2019 | -0.12 | 1.06 | 2.60 | 2.45 |
| | Mar 2019 | -0.16 | 1.05 | 2.53 | 2.40 |
| | Apr 2019 | -0.11 | 1.07 | 2.66 | 2.49 |
| | May 2019 | -0.10 | 1.07 | 2.67 | 2.49 |
| | Jun 2019 | -0.20 | 1.06 | 2.69 | 2.54 |
| | Jul 2019 | -0.16 | 1.06 | 2.67 | 2.54 |
| | Aug 2019 | 0.00 | 0.93 | 2.26 | 2.48 |
| | Sep 2019 | 0.15 | 0.88 | 2.13 | 2.47 |
| | Oct 2019 | 0.22 | 0.85 | 1.94 | 2.30 |
| | Nov 2019 | 0.29 | 0.92 | 1.98 | 2.15 |
| | Dec 2019 | 0.35 | 1.03 | 1.97 | 1.90 |
| | Jan 2020 | 0.33 | 0.98 | 1.76 | 1.74 |
| Feb 2020 | 0.28 | 0.95 | 1.99 | 2.10 | |

QQ plot is also being generated for the period August 2019 – February 2020 using normal values and root-mean-square standardized error (Figure 7). Most values are slightly close to the straight line. It indicates that the prediction error for weak spatial dependence (November 2018 – July 2019) and strong spatial dependence (August 2019 – February 2020) are close to the normal distribution.

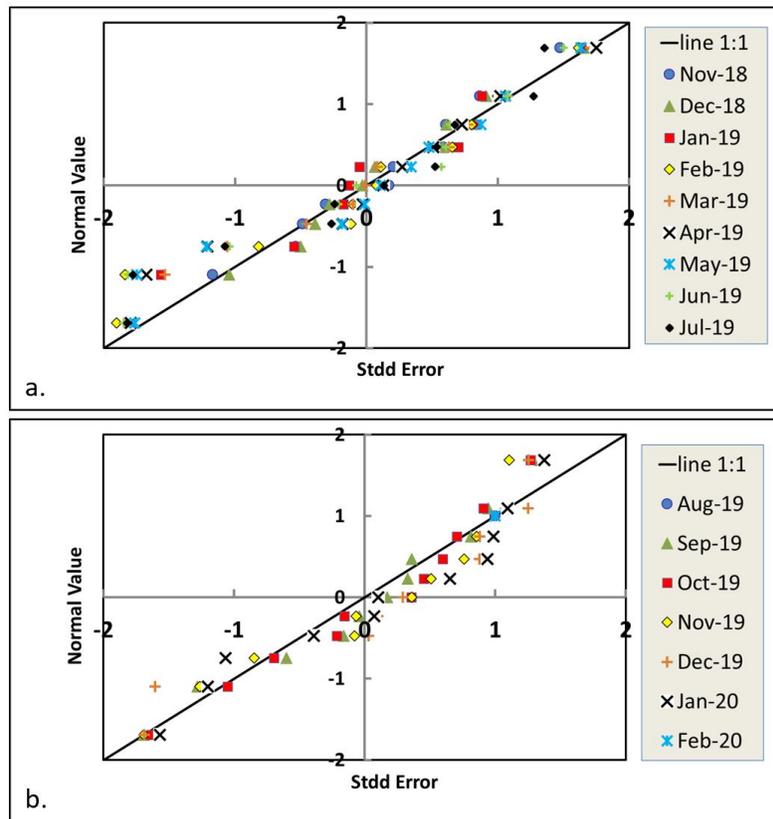


FIGURE 7 QQ pot of the RMSSE and the normal distribution values: (a) the period of November 2018 – July 2019 (weak spatial dependence), (b) the period of August 2019 - February 2020 (strong spatial dependence)

Cross-validation using correlation of coefficient and RMSE for each well was done to observe the variance of each well during November 2018 – July 2019 (weak spatial dependence), and August 2019 – February 2020 (strong spatial dependence) (Table 6). Most wells with strong spatial dependence have a high correlation of coefficient (0.76 – 0.97). On the other hand, most of the weak spatial dependence data have a relatively low positive correlation of coefficient (0.37 – 0.98), and even sometimes, it becomes negative like in well 13. The highest correlation happens in well 4, well 7, and well 8 with a coefficient 0.98, 0.95, and 0.96, respectively. It indicates that during the weak spatial dependence period, these three wells were influenced by the rainfall that occurred in January 2019.

TABLE 6 Correlation of coefficient and RMSE between observed and predicted WL for strong spatial dependence, weak spatial dependence, and all data compiled

| | Data with Strong Spatial Dependence | Data with Strong Spatial Dependence | Data with Weak |
|----------------|-------------------------------------|-------------------------------------|----------------------------|
| | Correlation of Coefficient | RMSE | Correlation of Coefficient |
| Well 1 | 0.97 | 2.54 | 0.86 |
| Well 2 | 0.94 | 1.76 | 0.88 |
| Well 3 | 0.97 | 2.84 | 0.37 |
| Well 4 | 0.94 | 1.39 | 0.98 |
| Well 7 | 0.92 | 0.67 | 0.95 |
| Well 8 | 0.76 | 3.39 | 0.96 |
| Well 10 | 0.95 | 2.02 | 0.55 |
| Well 11 | 0.96 | 0.88 | 0.73 |
| Well 12 | 0.97 | 0.28 | 0.62 |
| Well 13 | 0.95 | 2.52 | -0.25 |
| Well 14 | 0.97 | 1.27 | 0.46 |

The graph of the observed water level and predicted water level is displayed in Figure 8. It shows that the data points with weak spatial dependence are relatively far from the line of 1:1 compared to the case of strong spatial dependence. Therefore, the accuracy of the kriging map using data of strong spatial dependence (August 2019 – February 2020) is more accurate compared to weak spatial dependence (November 2018 – July 2019).

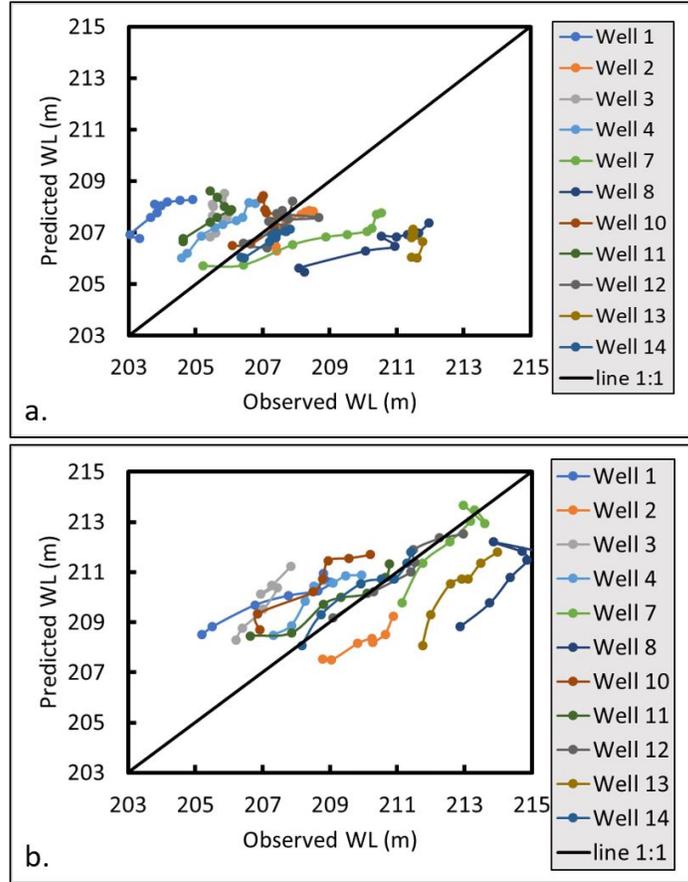


FIGURE 8 Observed water level vs predicted water level graph: (a) weak spatial dependence, (b) strong spatial dependence

It is known that the data during the period of November 2018 – July 2019 has weak spatial dependence. The water table map using the Kriging technique from this period is displayed in Figure 9. Random groundwater flow direction occurred, where it flows toward the outside of the farm or into the existing stream in the North and South of the study area. During June and July 2019, which is the transition into the high rainfall season, the groundwater direction is started to change toward the Northern stream.

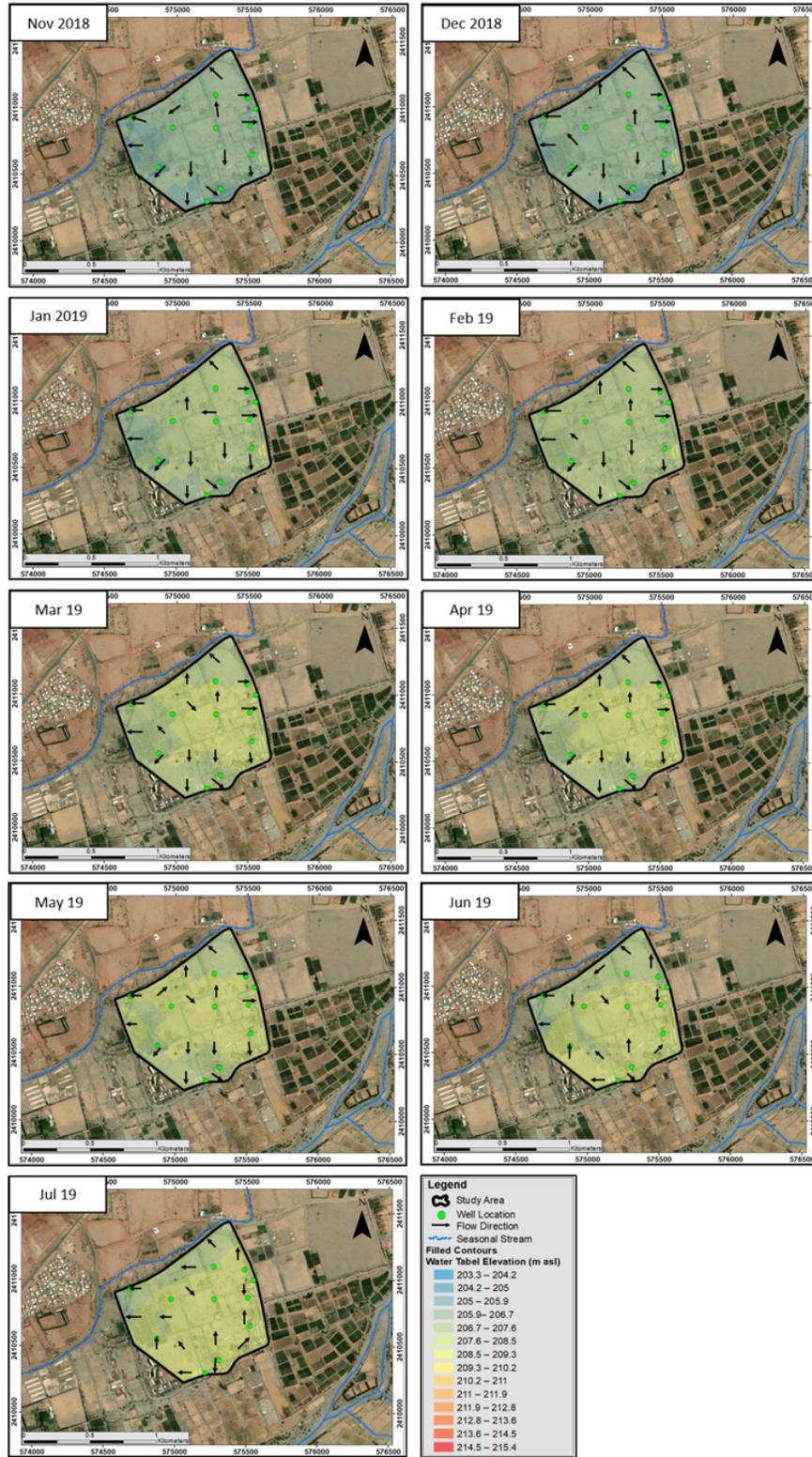


FIGURE 9 Water table map for the period of November 2018 - July 2019

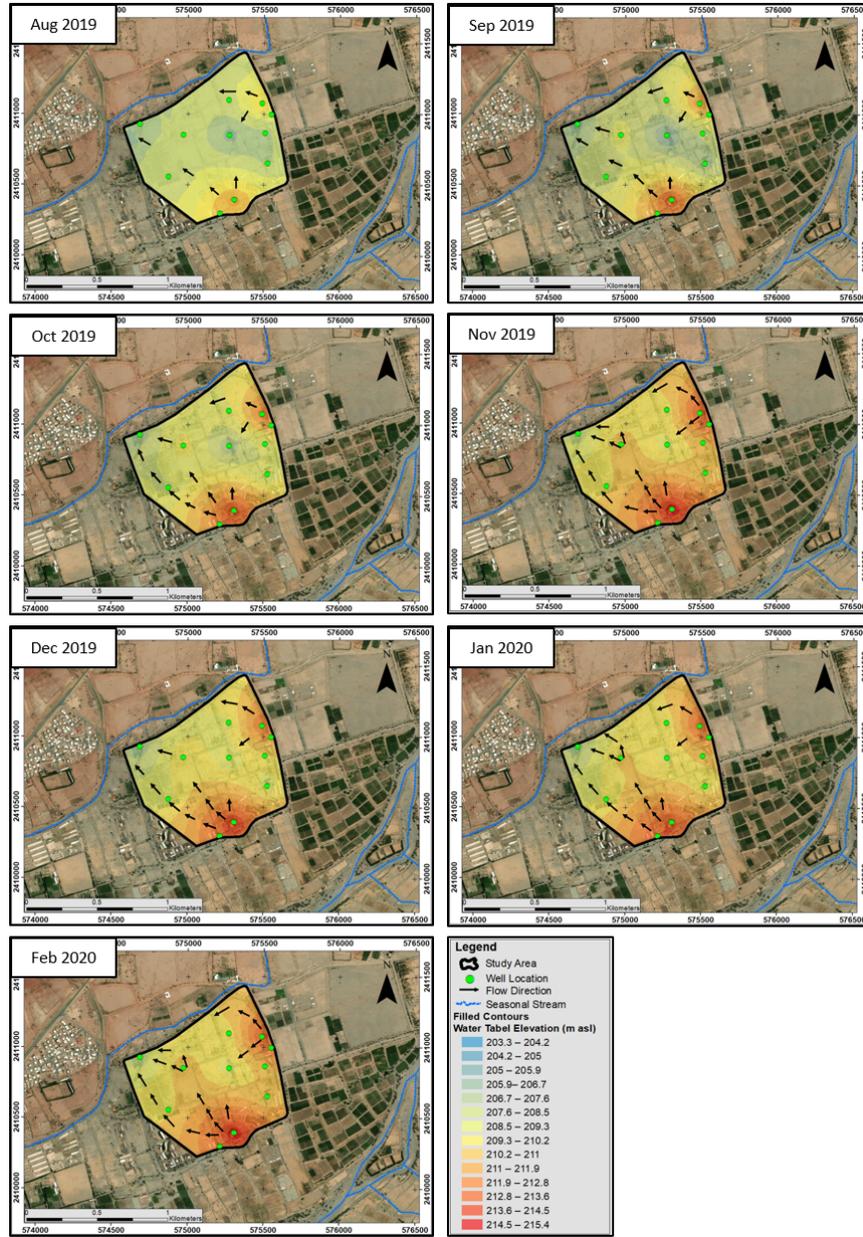


FIGURE 10 Water table map and water direction line for the period of August 2019 - December 2020

The water table maps for strong spatial dependence are displayed in Figure 10. Color change during the period of August 2019 – February 2020 indicates that the water table is increasing during this period. The shift in the water table also shows the direction of the groundwater flow. Groundwater flow direction is perpendicular to the hydraulic head (water table); therefore, during high rainfall season, it was flowing from East to West direction and South East to North West direction toward the Northern stream. It is known that the stream at the Southern of the study area is relatively at a higher elevation compared with the stream in the north of the study area. All information above shows that water from the Southern stream was infiltrating and recharging groundwater in the farm, and it was flowing toward the Northern stream.

5 | CONCLUSION

Based on water table measurement in 16 months, the water table was increasing and directly proportional to rainfall data. It can be concluded that in the arid region, for shallow groundwater aquifers water table fluctuation is highly influenced by rainfall.

The best-fitted-semi-variogram is the exponential model type, which has the lowest RMSE value. Within this given period of water table measurements, the first nine months of data (dry season, November 2018 – July 2019) has a weak spatial dependence ($>75\%$), and during wet season or the second period (wet season) data from August 2019 – February 2020, has strong spatial dependence ($<25\%$).

Based on the cross-validation analysis of the Kriging map using data of weak and strong spatial dependences, the following conclusions can be drawn:

- A negative mean error has occurred in data with weak spatial dependence, which means that the predicted value is lower than the actual measurement. A positive mean error has occurred in data with strong spatial dependence, which means that the predicted value is higher than the actual measurement.
- Most of the root-mean-square standardized error is close to 1, which indicates the errors are following a normal distribution. It is also confirmed by the QQ plot that shows the prediction error for both weak spatial dependence (November 2018 – July 2019) and strong spatial dependence (August 2019 – February 2020) are close to normal distribution.
- The RMSE for weak spatial dependence data is higher than the RMSE value for strong spatial dependence data, which means that the weak spatial dependence has a higher variance compared to the strong spatial dependence. A similar result occurs for average standard error, where the data with weak spatial dependence has a higher value compared to strong spatial dependence.
- Most correlation coefficients for wells data with strong spatial dependence are close to one, which means that the data from August 2019 – February 2020 has a positive and high correlation (0.76 – 0.97). On the other hand, most of the wells data with weak spatial dependence have a relatively low correlation of coefficient (0.37 – 0.98). It indicates that during high rainfall season, the water table rises in all wells and provides high correlations, while during the dry season, the water table drops at each well individually, and the correlation at some wells are poor and can reach a negative value.
- The graph of actual measurement and prediction shows that in the weak spatial dependence, the points are far from the line of 1:1 compared to strong spatial dependence. Therefore, the accuracy of the Kriging map using data of strong spatial dependence (August 2019 – February 2020) is more reliable compared to the case of weak spatial dependence (November 2018 – July 2019).

Based on the results of the water table map, the water table is increasing from August 2019 to February 2020. The groundwater is flowing from East to West direction and South-East to North-West direction, which is almost similar to the stream direction on the surface. This result corresponds to the streams elevations where it is assumed that the groundwater is flowing from the higher stream elevation (Southern stream) to the lower stream elevation (Northern stream). On the other hand, a Kriging map using weak spatial dependence data (November 2018 – July 2019) shows that the flow direction is going outside the farm in random directions. It happens because the water tables in the wells do not change in time due to there is no feeding from the rainfall, and therefore, there is no specified direction on the flow.

ACKNOWLEDGMENT

The authors would like to express gratitude to the technical staff of the Department of Hydrology and Water Resources Management, Faculty of Meteorology, Environment, and Arid Land Agriculture, King Abdulaziz University, who measured and collected the data at Hadat Ash-Sham Farm Experimental Station in 16 months.

APPROVAL DECLARATION

Authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

DATA AVAILABILITY STATEMENT

In accordance with the “DFG Guidelines on the Handling of Research Data”, we will make all data available upon request. The data set will be archived for at least 10 years after publications.

ORCID

Jaka Satria Budiman <https://orcid.org/0000-0002-1680-173X>