Global NEST Journal, Vol 16, No 1, pp 87-103, 2014

AN EMPIRICAL DECOMPOSITION OF DEEP GROUNDWATER TIME SERIES AND POSSIBLE LINK TO CLIMATE VARIABILITY

SEEBOONRUANG U.

Department of Civil Engineering
Faculty of Engineering
King Mongkut's Institute of Technology
Ladkrabang, Bangkok, Thailand

Received: 30/09/13
Accepted: 05/12/13

*to whom all correspondence should be addressed:
e-mail: muggletong@hotmail.com

ABSTRACT

Deep groundwater data reflects hydrological processes, climate change and variability, as well as any anthropogenic influence. Decomposition of deep groundwater signal examines the history of the groundwater region. Detrending is a vital step in decomposition of groundwater time series because it is expected to remove anthropogenic effects and long-term cyclic patterns. Eight detrending methods were applied to long-term groundwater records monitored in the Lower Chao Phraya basin in Thailand. Detrended residuals and subsequently periodograms of the residuals were computed by applying the Fourier series analysis. The result from this study indicates that the 5th order polynomial interpolation provides the trendlines that significantly relate to the groundwater withdrawal background. The detrended residual function is imbedded with two major cyclic patterns, which can be the result from global climate variability, e.g. Indian Ocean Dipole and the El Niño Southern Oscillation. The magnitude of deep groundwater dynamics as the result from the anthropogenic effect, is much greater than that of the climate variability in this region. In addition, this study demonstrates that caution must be exercised when fitting groundwater time series with different detrending techniques can yield mistaken cyclic patterns and may infer to different climate variability phenomenon.

KEYWORDS: Climate variability, correlation, decomposition, detrending, periodogram, time series.

1. INTRODUCTION

Groundwater system not only is naturally dynamic but also changes with time in response to both climate stresses and human activities (Alley *et al.*, 2002). Pressures from human activities on groundwater resources come from the population growth, agricultural increases, and rapid expansion of urban areas (Gunatilaka, 2005).The major man-made problem on the groundwater resource is the overexploitation of the resource, especially in large cities of a number of countries in Asia, e.g. China, India, Indonesia, Nepal, the Philippines, Thailand and Vietnam (UNEP, 2002; ADB, 2007; WEPA, 2007). Bangkok, the capital city of Thailand, has experienced excessive drawdown of potentiometric surface and suffered from land subsidence due to intensive pumping of groundwater (IGES, 2007; Piancharoen, 1977).

Climate change has added to the existing pressure on groundwater resources both directly and indirectly and rendered the only available freshwater reserve less accessible (IPCC, 2007; Kundzewicz and Doll, 2007). In addition to climate change, global water resources are very sensitive to climate variability. Climate variability represents rather reversible and periodic changes in climate occurring over periods of a few years to decades (Hanson *et al.*, 2004; Gurdak *et al.*, 2007). Impacts of climate variability on groundwater are however poorly understood and the relationship between climate variability and groundwater is more complicated than that between climate variability and precipitation or surface water. In shallow and unconfined aquifer systems, the response of the change and variation may clearly be detected. Hanson *et al.* (2004) and Chiew *et al.* (1998) found shallow groundwater fluctuation response to climate variation with 2- to 6-year cycles resulting from

El Nino/Southern Oscillation (ENSO). The periodicities of 6 to 10 years of groundwater signals may be related to cycles in monsoons (Gurdak *et al.*, 2007). In addition, interdecadal variations of approximately 10-year to 25-year cycles may be related to the Pacific Decadal Oscillation (PDO) (Mantua and Steven, 2002). Very few research studies have tried to relate deep groundwater signals to climate change and climate variability (Tague *et al.*, 2008; Green *et al.*, 2011).

In order to assess and conclude these kinds of difficulties, time series and frequency analysis are typically essential. A time series record of groundwater tables or piezometric levels reflects all the natural hydrological processes, climate change and variability, as well as any human influence on the subsurface environment. Therefore, inverse analysis of groundwater time series should provide the input and the response information about history, current and future states of the system. This information has become increasingly important with growing utilization and management. As a result, a number of applications of time series methods have been applied in several groundwater studies to analyze groundwater record in response to individual stresses (Ganoulis and Morel-Seytoux, 1985; Gurdak *et al.*, 2007; Aflatooni and Mardaneh, 2011; Perry, 1994; Perry and Hsu, 2000; Yu and Chu, 2012). A systematic methodology is often required for both local and regional assessments and must allow for decomposition, time series and frequency analysis and comparison across data types and frequencies, across hydrologic settings and across different basins (Hanson *et al.*, 2004).

A concrete and logical decomposition method will isolate and quantitatively assess the impacts of individual factors on groundwater and establish the linkage and complexity among those factors, e.g. hydrologic processes and climatic variability (Hanson *et al.*, 2004; Mayer and Congdon, 2008). Typically, signal decomposition analysis may include detrending, frequency and correlation analysis. While various frequency analysis methods can extract the periodic information, detrending is the mathematical operation of removing a trend from a time series and is often applied to remove a feature that may have a long-term effect on the series such as groundwater withdrawal and application of groundwater for irrigation that may in turn distort some imbedded cyclic signals (Prinos *et al.*, 2002; Hanson *et al.*, 2004; Gardner and Heilweil, 2009). Correlation methods between original groundwater time series or extracted time series and other climatic time series can then provide the linkage between the parameters and some future prediction for integrative water resource management. This analysis requires not only long-term groundwater monitoring network (Taylor and Alley, 2001) and a systematic decomposition but also a reasonable detrending method for groundwater time series. In studying and extracting a trend, many alternative methods are available for detrending and so it is important to understand the effects of different detrending methods on the spectral properties of the residual time series. Different detrending methods may yield unique sets of periodic patterns, which in turn could result in the mistaken assessment of climate variability on groundwater signals. Therefore, this work aims to present a systematic method of groundwater signal decomposition and emphasizes certain widely used detrending methods and their consequences on the other components of the time series.

2. DECOMPOSITION THEORY AND LITERATURE REVIEW

In addition to reacting as a propagating pressure wave (Serfes, 1991), a time series is a cumulative response with variability of time lags to different signals, such as climatic signal, e.g. temperature and precipitation; hydrogeological signals, e.g. discharge and recharge; and anthropogenic signals, e.g. pumping and reservoir operation. Therefore, inverse or decomposition of groundwater time series into elemental components is a substantial step in identifying and quantifying individual natural and anthropogenic effects on historical groundwater data. Reconstruction of these components will provide means to forecast some future change due to different spatial and temporal degree of the factors on the groundwater signals. Past studies of groundwater data decomposition include recognizing the effects of climate change and variability on hydrologic cycles after eliminating anthropogenic effects from time series (Anderson and Emanuel, 2008; Hanson *et al.*, 2004; Ghanbari, 2007). Detrending methods typically are applied to eliminate the anthropogenic effects, e.g. those from a long-term groundwater development. Subsequently, a frequency analysis is performed on the detrended time series, providing the information of the cyclic response to the climate factors.

A time series data refers to observations on a variable that occurs in a time sequence and these observations are collected at discrete but not necessarily equally spaced time intervals. The classical decomposition partitions a signal into three elemental components called trend, periodicity and random or irregular components and can be written as

$$X_t=m_t+s_t+y_t \tag{1}$$

X_t is the time series data, m_t is the trend component, s_t is the periodic component, and y_t is random noise or irregular component.

The trend component typically shows a long-term increase or decrease, either linear or nonlinear, in the data; or is part of the data which exhibit rise and fall that are not of a fixed period. It underlies the tendency or the main direction, i.e. upward, downward, or upward and downward, in a time series. This is to account for the effect from the inertia of some processes. The periodicity component of a time series presents cyclic elements of the time series during the time of measurement that is influenced by any seasonal and periodic factors and recurs on a repetition basis. This component can be represented by a finite number of sinusoidal functions with different amplitudes, phases and periods. After the trend and periodicity patterns are removed from a time series, the remainder is the irregular or random component, which does not contain any fixed configuration.

There are a number of techniques available to evaluate trends within datasets. The simplest way is to apply a simple linear regression analysis (Aflatooni and Mardaneh, 2011; Polemio and Casarano, 2008) or some low-order polynomial regression (Gurdak *et al.*, 2007; Gardner and Heilweil, 2009). However, fitting a simple linear model to a groundwater time series can be questionable, especially when the groundwater data normally show a highly nonlinear pattern (Shao *et al.*, 1999). Moving average, e.g. a 12-month moving average, is an alternative detrending method to remove small-scale structures, e.g. noises, and short periods from a time series. This windowing technique removes short-term variations (fluctuations within the interval) and emphasizes long-term trends of the data. Therefore, if a moving average has a sufficient averaging interval, the smoothed data can be considered as a trend. A trend detection using moving average low pass filter (Kiely, 1999) is given by

$$MA_t=\frac{1}{2L}\left(\frac{X_{t-L}}{2}+\sum_{j=-(L-1)}^{L-1}X_{t+j}+\frac{X_{t+L}}{2}\right) \tag{2}$$

where X_t is the variable at year t and L is the interval of moving average, j varies from –(L-1) to (L-1). This window averaging method smoothes out high frequency cycles and some irregular components from the original time series (Ngongondo, 2006).However, the problem of overfitting data should not be overlooked. Overfitting occurs when a model describes some cyclic patterns instead of underlying trends.

Periodicities of groundwater time series have been reported in a number of studies. Gurdak *et al.* (2007) show subsurface responses in cyclic patterns related to ENSO (2-6 years) and PDO (10-25 years), while Shah (2009) finds much shorter cycles relating to pre- and post-monsoon seasons on the groundwater levels in an irrigation area in India. Methods of spectral analysis are typically applied to study the periodic properties of a time series, e.g. economic variable or hydrologic data over frequency spectrum, i.e. in the frequency-domain. The obtained power spectral distribution describes how the monitored variable can be decomposed into a variety of sinusoidal components with various periodicities or frequencies. Singular spectrum analysis developed by Vautard *et al.* (1992) was applied for regional hydroclimatic assessment of the southwestern USA in the works of Hanson *et al.* (2004) and Gurdak *et al.* (2007). Box-Jenkings time series method developed by Pankratz (1983), was used to forecast future groundwater fluctuations (Aflatooni and Mardaneh, 2011).

Among the frequency-domain methods, the Fourier series analysis is widely used to break down a single signal into constituent sinusoids of different frequencies (Oppenheim *et al.*, 1997). Each sinusoidal component is represented by amplitude and a phase and the summation of all cyclic components determines the relative contribution of that frequency component to the entire signal. With the Fourier analysis, the signal is then transformed from a time-based to a frequency-based representation (Cohen, 1995). After separating trend and period components, the residual is termed noise which could be caused by a variety of influences including observation errors, measurement errors from data-logging equipment, and unaccounted high frequencies components, e.g. tides or seasonal variations (Bhar and Sharma, 2005; Nagpaul, 2005). These influences, either alone or in combination, are commonly small compared to other components.

3. STUDY AREA

Hydrogeologically, the Lower Chao Phraya basin is located in the Upper Tertiary to Quaternary strata of the Lower Central Plains of Thailand (Figure 1). The aquifers consist mainly of sands, gravels, and minor clay lenses of Pliocene-Pleistocene-Holocene geological age and can be zoned into eight principal confined aquifers separated by apparent thick confining clay layers (Piancharoen, 1977; Piancharoen and Chuamthaisong, 1978). This subsurface sequence is overlain by the Holocene Bangkok Clay. The depth of this aquifer system is approximately 500 meter. The main recharge to this aquifer system is from the Middle Chao Phraya basin hydraulically connected to the Lower basin. Widespread exploitation of groundwater for urban water-supply has commenced in the 1950s and the primary targets for water well construction were the 2nd, 3rd and 4th aquifers in the depth range 100-250 meter below the ground surface (Babel *et al.*, 2007). This groundwater resource exploitation caused massive drawdowns of the groundwater levels and vast land subsidence over much of the Greater Bangkok area. This aquifer system is however still identified as the largest source of groundwater in the country and is located in the most populated region in Thailand.

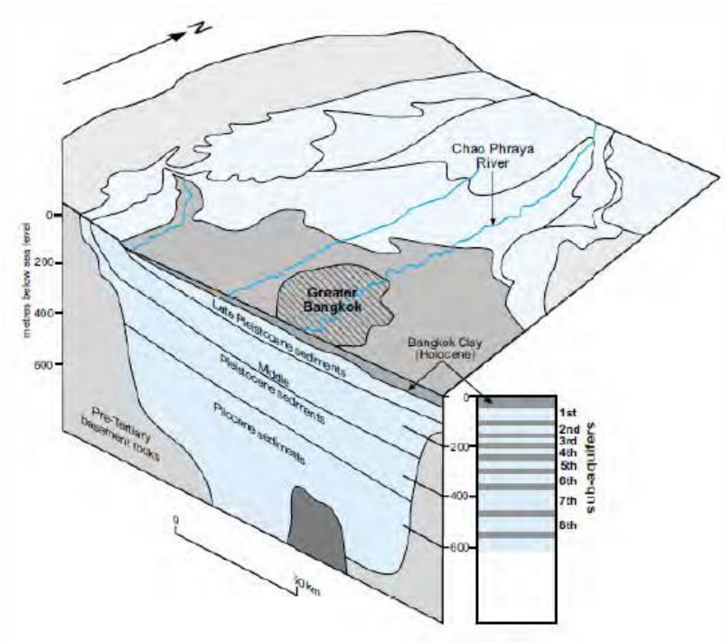


Figure 1. Hydrogeologic setting of the study area of Thailand's Lower Chao Phraya basin (Lorphensri, 2011)

There have only been a handful of studies analyzing climate and rainfall variations in Thailand (Kusreesakul, 2009; Limsakul *et al.*, 2007; Limsakul and Goes, 2008; Pratarastapornkul, 2009). These studies report that climate variability in Thailand is associated with periodicities in multiple time scales ranging from seasonal, interannual, and decadal variations. Summer monsoons, Indian Ocean Dipole (IOD), and ENSO have the most influence on climate in the region. The annual rainfall cycle in the region is in phase with the Indian summer monsoon and the western North Pacific summer monsoon. IOD also affects the annual rainfall and its impact varies depending on both the combination effect of IOD and ENSO and the strength of ENSO in the Pacific Ocean. The periodicity of IOD has been in the range of between one to 4 years since 1990. ENSO has a 3- to 7-year cycle, with a four-year average to go from El Niño to La Niña.

4. METHODOLOGY

Time series of groundwater level is cumulative response to hydrological, hydrogeological and anthropogenic inputs, such as precipitation, natural recharge from surface water sources, artificial recharge, heterogeneity of subsurface system, and pumping (Yu and Chu, 2012). To individually quantify and assess such impacts, decomposition of groundwater signal and long-term monitoring network are essential.

Available Data

Long-term continuous groundwater level data have been collected throughout parts of the Central Plain of Thailand by the Department of Groundwater Resource. The data contain records of groundwater depths from the ground surface at 77 observation sites throughout the study area, some records of which have been collected since 1921. The water level reading has been undertaken approximately once a month. The wells penetrate to the depths of between 50 – 250 m, in which groundwater was heavily withdrawn particularly between 1980 and 2004. Most bores are located in the urban and built-up lands. Up to the present, it is still unclear how climatic and anthropogenic factors could quantitatively impact this deep groundwater system, and decomposition of these groundwater signals is believed to be able to help answer this question.

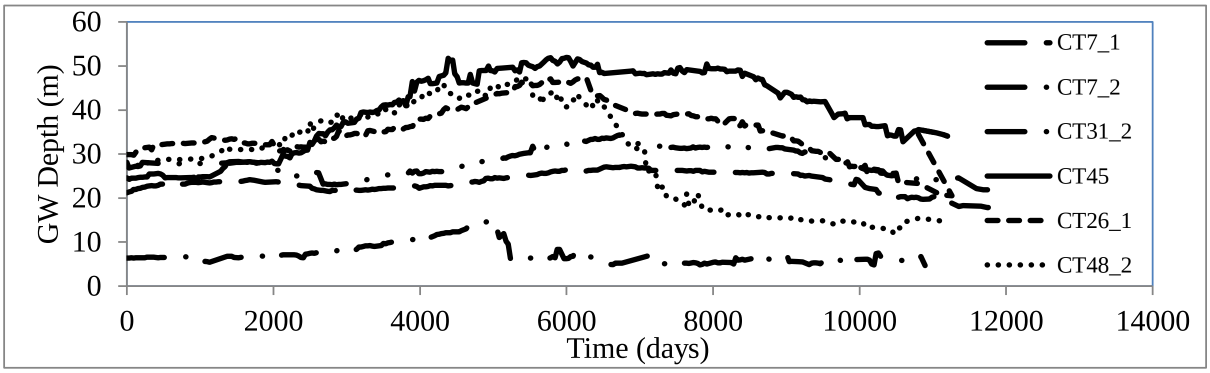


Figure 2. Selected groundwater time series from different depths and locations

The first step was to select some bore records with reasonably long-term and constantly measured observations of water levels. This is to reduce some uncertainty that may arise from interpolation when a lack of available collected data during certain periods. This selection was achieved by scanning the water level database and classifying each bore in terms of lengths of record, beginning and end of record, total number of readings, and the longest data gap. Figure 2 shows the time series of the deep groundwater at the selected sites, namely CT7_1, CT7_2, CT31_2, CT45, CT26_1, and CT48_2, from varying depths in the highly pumped aquifers mentioned above. The x-axis represents the time since the first measurement and the y-axis represents the depth of groundwater table in the observation wells. The groundwater at the selected locations behaves similarly in that the depth of groundwater continuously increased for the period of approximately 17 years and later decreased due to the previous heavy pumping. Descriptive statistical properties of the deep groundwater are shown in Table 1.

Table 1. Descriptive statistical properties of the selected groundwater

	CT7_1	CT7_2	CT31_2	CT45	CT26_1	CT48_2
Mean	23.61	28.20	7.21	40.32	35.22	29.54
Standard Deviation	2.28	3.07	2.25	8.95	6.33	11.39
Kurtosis	-0.27	-1.08	1.87	-1.19	-0.34	-1.48
Skewness	-0.54	0.02	1.61	-0.46	-0.11	-0.11
Range	9.37	12.45	9.97	27.74	26.77	35.22
Count	394	392	364	374	376	374

Groundwater level measured from CT48_2 has highest standard deviation with maximum data range. The distribution of the selected groundwater generally has low and wide peak (negative kurtosis) and is left-skewed with relatively few values of low groundwater depth (negative skewness). The groundwater record at CT31_2 is the exception with positive skewness and kurtosis. The total number of measurements of each data series is almost identical.

Prior to the decomposition process, each time series was interpolated into monthly basis using the simple linear interpolation. This is because groundwater level records were somewhat measured at irregular intervals that ranged from monthly to bimonthly and were not recorded during certain time intervals. The interpolation interpolates an irregularly sampled time series into a regular spacing, as required by many methods for time series analysis.

Detrending is the mathematical removal of a potential trend from a time series. Several detrending methods were used in this study to observe and assess the effects of various detrending techniques on the overall successively decomposed components. The first detrending technique followed the moving average low-pass filter developed by Kiely (1999) as shown in Eq. (2). The low pass filter averaged the selected monthly groundwater time series according to Eq. (2) over the intervals of 3, 5, 7, and 10 years. The trend component was also achieved using low to medium order polynomial fitting regression (Gardner and Heilweil, 2009). In order to assess the sensitivity of different trends on a signal, 2nd, 3rd, 4th, and 5th order polynomials were fitted to the groundwater signals using the least square error methods. These nonlinear regression lines are the trendlines and the component of the signals with low frequencies (high periodicities). In addition, these nonlinear trends in the data may represent a combination of parts of larger climatic cycles and periods of anthropogenic effects, detrending of the data should eliminate the parts of anthropogenic effects and climatic periodicity longer than half of the monitoring period. The groundwater trendlines were then correlated with the pumping record in each location using the cross correlation technique to examine the validity of the hypothesis and of the detrending methods. Furthermore, a simple technique was proposed in order to explore some detrending methods that might be subjected to overfitting data.

The detrended residuals for each series were obtained by subtracting the fitting trends from the interpolated data over the periods of record. These corresponding residuals should exhibit some climate variability with short and medium length cycles, i.e. less than 15 years, which is less than half of the record periods.

Autocorrelation was applied to the detrended residuals of time series. Autocorrelation (Davis, 1986) is a simple classical technique for time series analysis and refers to the correlation of a time series with its own past and future values. Positive correlation indicates a specific persistence in a physical system, a tendency for the system to remain in the same state from one time to the next. The autocorrelation function of a time series with a lag time k is defined by

$$R_{xx}(k)=\frac{1}{\sigma^2(n-k)}\sum_{t=1}^{n-k}\left(X_t-\bar{X}\right)\left(X_{t+k}-\bar{X}\right) \tag{3}$$

where X_t is the variable at measured time t, n is the total observations, $\bar{X}$ is the mean of the time series, and σ^2 is the time series variance. The function can tell whether the data are long enough such that real periodicities begin to show up and the function is more commonly applied to examine the degree of smoothness.

The detrended residuals were then identified for periodicity patterns in frequency response domain. While groundwater hydrograph is a function of time, a frequency response representation is a way of representing the same hydrograph signal as a frequency function. Frequency domain or spectral analysis is usually applied to assess the periodic variation of precipitation, streamflow, or groundwater levels (Perry, 1994; Perry and Hsu, 2000; Gurdak *et al.* 2007). The purpose of spectral analysis is to estimate the power (strength) of periodic components at all possible frequencies. One approach to spectral analysis is the Fourier analysis in which a time series is decomposed into a complete set of sine and cosine components with different coefficients. These components are sinusoidal; each with a certain amplitude and phase, and power of each periodicity is proportional to amplitude squared. The program called PAleontological STatistics (PAST) was used in this study (Hammer and Harper, 2006; Hammer *et al.*, 2001; Harper, 1999) to perform the Fourier analysis. It can be shown that any evenly spaced time series of length N can be represented precisely and completely as a sum of N/2-1 sinusoids. The first of these sinusoids has a period of N samples, the second has a period of N/2 samples and until the last sinusoid has a period of only two samples.

The result from the Fourier series analysis on a time series signal is called the power spectral density or periodogram, which represents the strength of different frequencies or inverse of periodicities of the signal. To compare the periodograms from particular detrending methods, the first and second moments of the complete periodograms were computed. These moments indicate the periodic means and variances of the spectral distributions. Subsequently, two normal distributions were fitted to the spectral distributions. The attempt was incorporated into this analysis because the shape of most of the resulting distributions is similar to two normal distributions superposed. The simple superposition of two normal distributions with different means and variances was applied such that the best fit was obtained when the sum of the square errors between the actual spectral distribution and the superposed bi-normal distribution was minimized. The result of this matching was reported in terms of the values of two means and two variances.

5. RESULTS AND DISCUSSION

5.1. Detrending results

Examples of fitting the groundwater time series with different detrending methods are shown in Figure 3 for well CT7_1 and Figure 4 for well CT7_2.

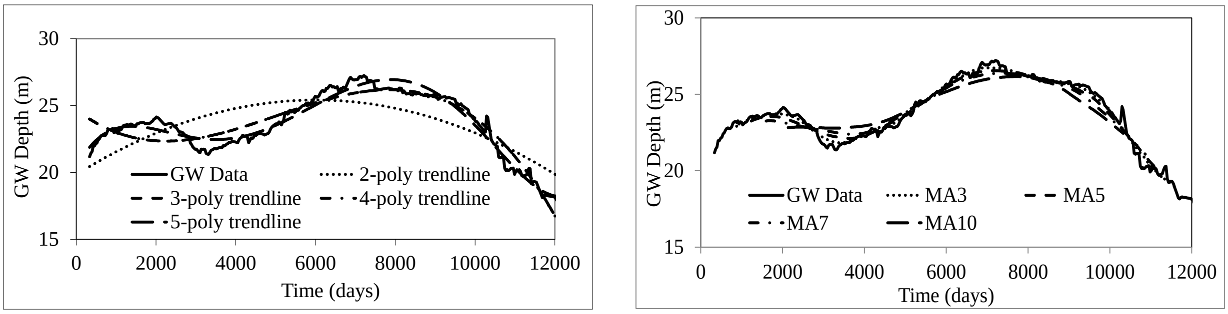


Figure 3. Groundwater time series at CT7_1 isdetrended with a) low-medium order polynomial interpolation and b) moving average techniques

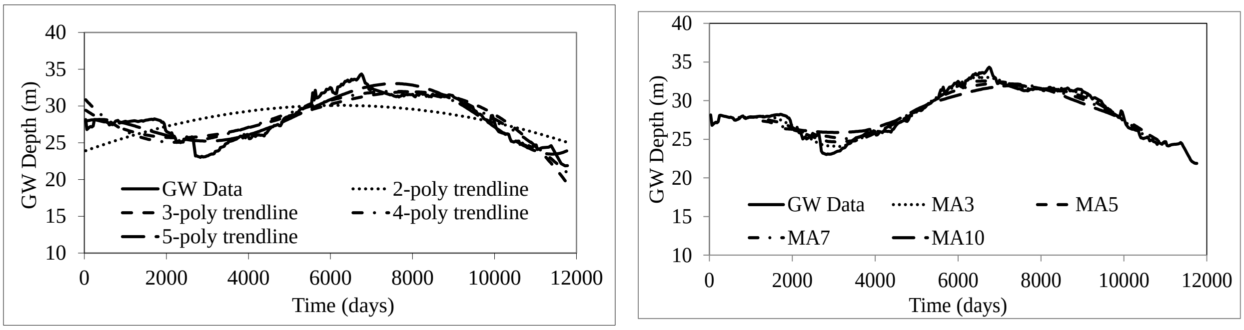


Figure 4. Groundwater time series at CT7_2 is detrended with a) low-medium order polynomial interpolation and b) moving average techniques

The 2ndorder polynomial regression is off the actual data of both wells, while the other fitting techniques perform qualitatively quite well. Pearson's correlation between the data at the selected six well locations and the eight trendlines is shown in Figure 5.

Most of the regression techniques achieve the target data reasonably well with the average correlation of approximately 0.9 with the exception of the 2nd order polynomial regression. The fitting method performs poorly on the time series of CT7_1, CT7_2, and especially CT31_2. In fact, all of the regression techniques provide poorer fitting to the CT31_2 time series than other time series. This can be explained by examining the distribution of the groundwater depth at CT31_2 by which the behavior of the groundwater system is quite irregular with high skewness and high kurtosis numbers (Table 1). However, the 2nd order polynomial method can execute satisfactorily at wells CT45, CT26_1 and CT48_2. On average, the 3-, 5-, 7- and 10-year moving averages yield the coefficients of Pearson's correlation with the data of 0.97, 0.97, 0.96 and 0.93, respectively, while the 5th, 4th, 3rd, and 2nd order polynomial regressions provide the coefficients of 0.94, 0.91, 0.88 and 0.75, respectively.

The various trendlines were correlated with the groundwater withdrawal rates obtained from the Thailand Department of Groundwater Resources.

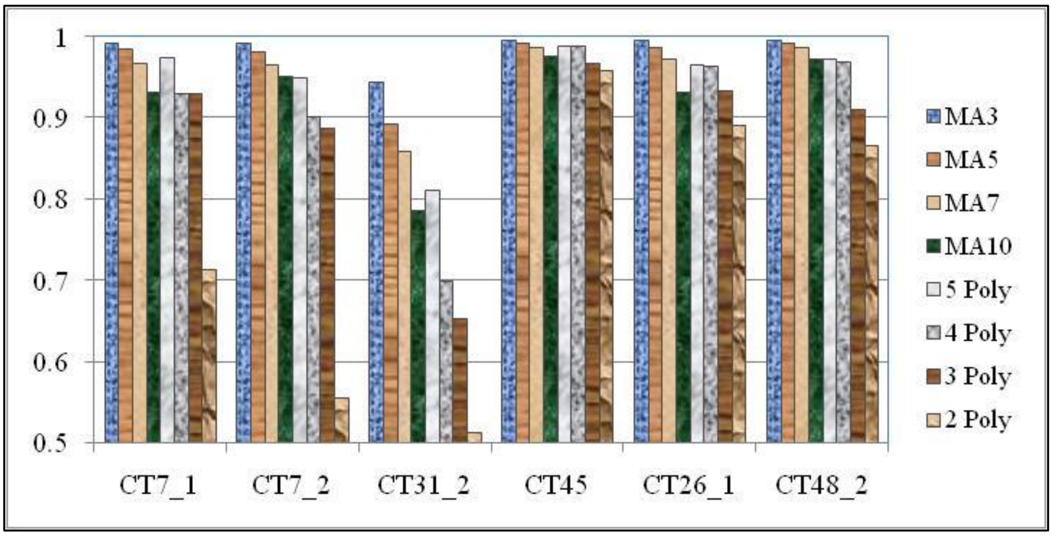


Figure 5. Pearson's correlation between actual data and various trends

Figure 6 shows the correlation results between the trendlines from different detrending methods and pumping rates at the six selected sites.

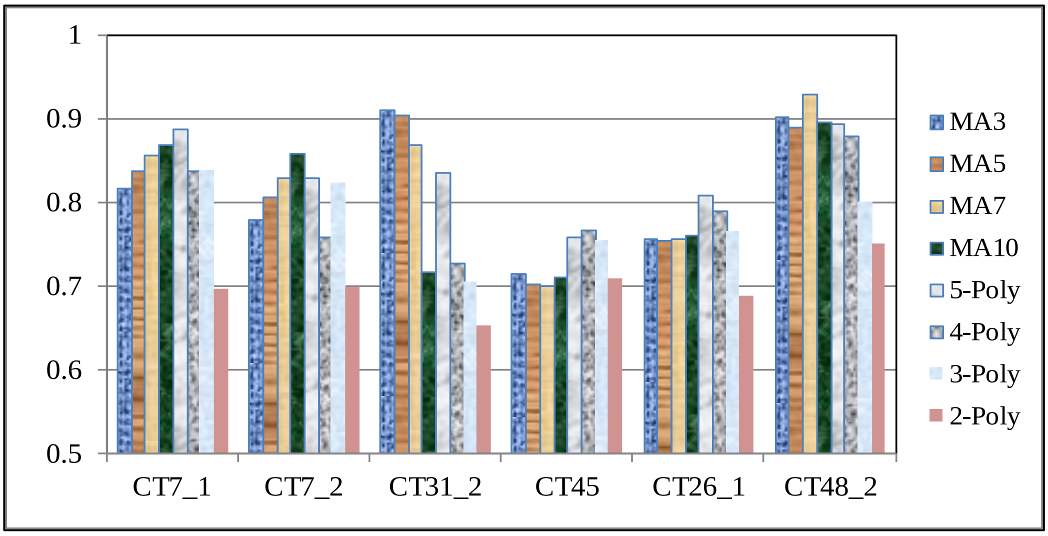


Figure 6. Pearson's correlation between pumping rate and various trends

The obtained trendlines correlate well with the pumping rates. This implies that the groundwater fluctuation in this region is highly influenced by the withdrawal and the detrending the data therefore can remove the anthropogenic effect from the groundwater time series. The 2nd order polynomial interpolation incurs less fitting to yield the trendlines that respond to the anthropogenic effect as shown in the lowest correlation coefficients. The moving average techniques with the different length windows extract trends of groundwater signals that reflect the pumping activity satisfactorily. This can be seen from their comparable correlation coefficients. On average, the 3-, 5-, 7- and 10-year moving averages yield the coefficients of Pearson's correlation with the data of 0.81, 0.82, 0.83, and 0.80, respectively, while the 5th, 4th, 3rd, and 2nd order polynomial regressions provide the coefficients of 0.84, 0.79, 0.78, and 0.70, respectively. In conclusion, the 5th order polynomial interpolation and the 7-year window moving average technique can perform better than the other detrending techniques. Therefore, these detrending techniques can remove the anthropogenic effect, which is groundwater withdrawal, from the groundwater signal. The sensitivity analysis shows that various detrending techniques can yield different trendlines, which some of them do not associate well with the withdrawal activity.

5.2. Descriptive statistics of the detrended residuals

In the previous step, detrending was performed to yield what is termed the trendlines of the selected time series. The trends present the main tendency of groundwater behavior in the past, and different detrending techniques yield different trendlines. Subsequently, the actual time series data of all the selected wells were subtracted by the regression trends to obtain the detrended residuals. For instance, Figures 7 and 8 respectively show the detrended residuals of CT7_1 and CT7_2. The residuals from all the detrending methods show similar up and down cyclic phase but with distinct values of residuals. The amplitudes of the 5th order polynomial and the 3-year moving average methods are smallest, while those of the 2nd order polynomial and the 10-year moving average methods are largest.

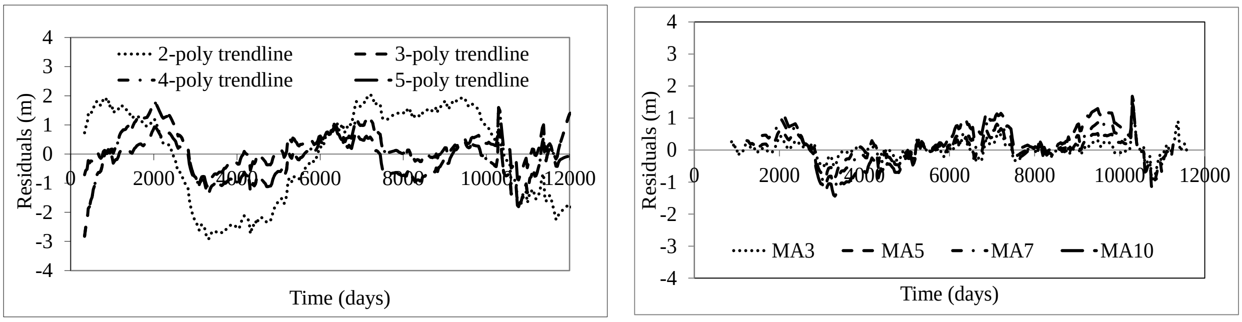


Figure 7. Detrended residuals at CT7_1 with a) low-medium order polynomial interpolation and b) moving average techniques

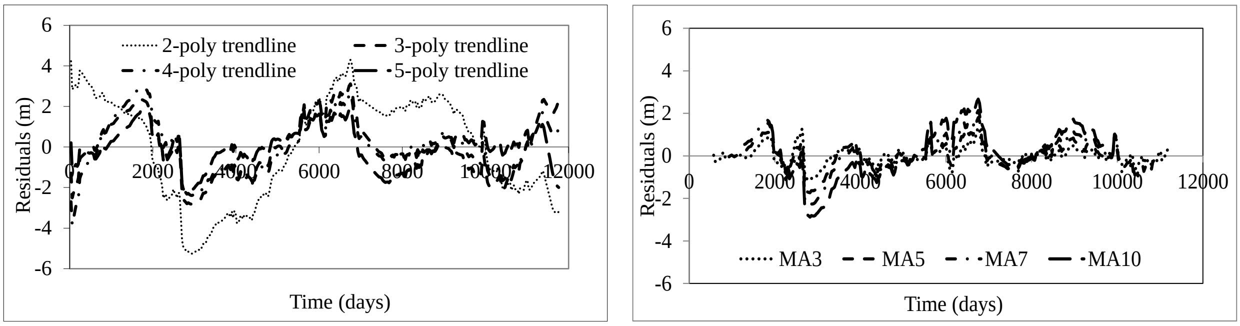


Figure 8. Detrended residuals at CT7_2 with a) low-medium order polynomial interpolation and b) moving average techniques

The descriptive statistics of the detrended residuals from the different regression techniques of the selected time series are computed. The wider the window for moving average, the higher the mean, i.e. the mean of the residuals from the 10-year moving average is greater than that of the 3-year moving average. On the other hand, all polynomial regressions provide the residuals with zero mean. The moving average window increases and so the standard deviation. On the contrary, the lower the order of polynomial interpolation, the larger the standard deviation is. This is because the higher orders of polynomial interpolation can yield better fitting due to their greater flexibility and thus typically produce less deviation from the data. On average, the 3- to 10-year moving averages provide less discrepancy than the low order polynomial interpolations. The ultimate goal of this study, however, is not to obtain the best fitting to the original data, but to link the models to anthropogenic effects.

5.3. Model overfitting Detection

It may appear possibly to closely match almost any groundwater signal pattern using any available detrending method. Generally, when the fitting model has excess degree of freedom, e.g. 5th order polynomial interpolation or 3-years window moving averaging, it is expected that their data fitting will appear to perform better than the other less-sophisticated fitting techniques with less degree of freedom. It is therefore important to demonstrate if the results are meaningful from the true underlining processes or artifacts resulting from overfitting. Overfitting is a concern in decomposition a signal using a detrending method because when some models too closely fit the signal and so may damp out some significant short term variation in any signal to produce a smoothed pattern. A simple method was applied here to investigate when an overfitting is an issue for any detrending on a groundwater time series and when a more complicated model is not resulting in better fitting.

Figure 9 shows the correlations between the obtained trendlines and the actual groundwater time series and the correlations between the trendlines and the corresponding pumping rates. 'MA' represents the moving averaging techniques and "Poly" represents the polynomial interpolation. The correlations between the CT7_2 groundwater data and polynomial trendlines increase and so those trendline correlations with the withdrawal rates are. This means that when the technique has more degree of freedom, i.e. more sophisticated and higher order of interpolation, the technique yields the trendlines with higher correlation to the pumping rates. This is an indicator of matching between the trends and the actual underlining force and says that the trendlines represent the true anthropogenic effect. This phenomenon is also true when applying both moving average and polynomial interpolation techniques to other time series.

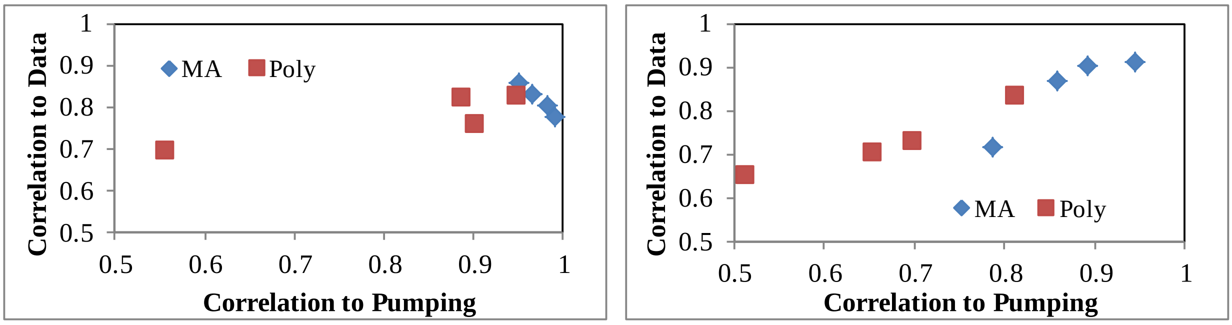


Figure 9. Correlations of trendlines with data and those with pumping for CT7_2 and CT31_2

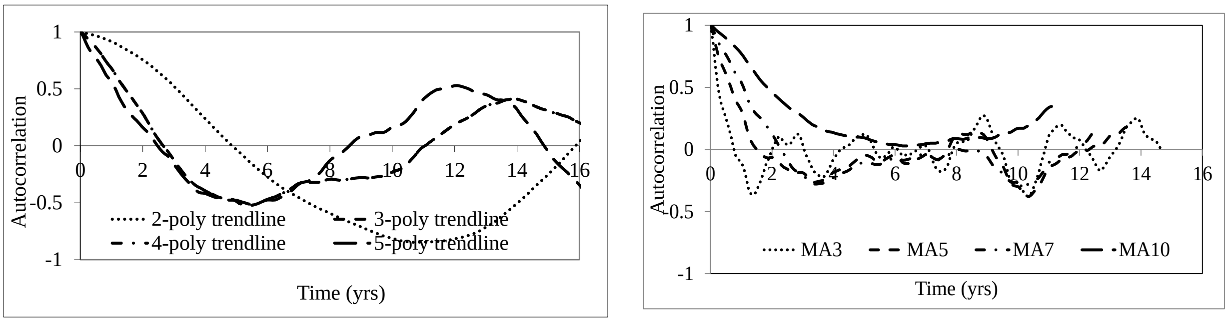
On the other hand, when the moving average trendlines have a better fit to the original CT7_2 groundwater data, they are less associating with the rates of withdrawals. This suggests that the more sophisticated detrending technique, i.e. 3-yr window moving average, overfits the data and has poor matching performance to the underlying anthropogenic factor, as it may exaggerate minor fluctuations in the data. The latter phenomenon is an indicator of overfitting the data while less agreeing with the pumping rates. Therefore, the potential for overfitting does not depend on the methods used but rather depends on the data shape, and the actual linkage between the data and some physical parameters, which is the withdrawal rate in this case. In general, polynomials can mathematically model underlying processes, although significance may be lost due to overfitting with high order polynomials.

5.4. Autocorrelation of the detrended residuals

The detrended residuals from different detrending techniques at the selected well locations were then autocorrelated. Figures 10 and 11 show examples of the autocorrelation functions of the detrended residuals at CT7_1 and CT7_2. The residuals from both detrending methods during the record period are strongly dependent upon the values of the previous months and the linkage diminishes with increasing time lag. This is called "memory effect" (Polemio and Casarano, 2008; Tomasella *et al.*, 2008). The autocorrelation functions from the low order polynomial regression detrending methods are shown in Figures 10a and 11a. The autocorrelation functions of the detrended residuals obtained from the 2nd order polynomial interpolation have longer memory effect than some other higher order interpolation. Those autocorrelation functions from 3rd, 4th, and 5th order interpolations show similar memory effects with maximum correlation coefficients at the periods of approximately 5 and 12 years. This is an initial indicator of periodic patterns in groundwater time series and needs to be confirmed by a spectral analysis result.

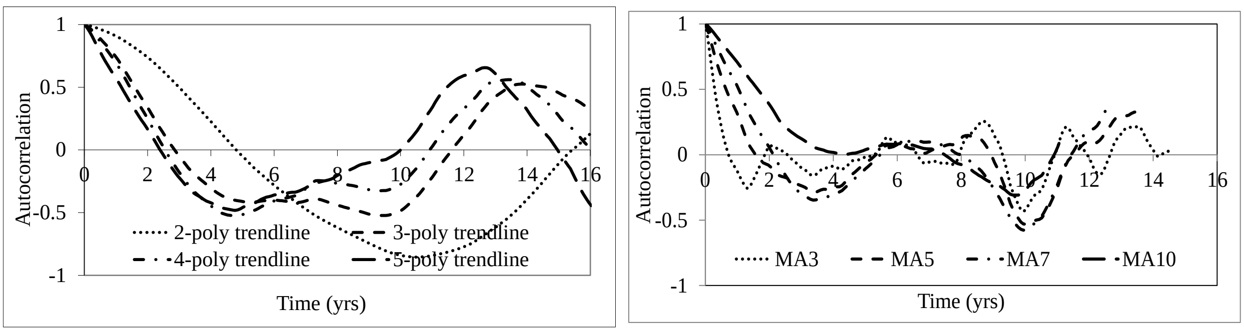
The autocorrelation functions from the moving average detrending methods are shown in Figures 10b and 11b. The autocorrelation functions from the moving average methods have short memory effect with insignificant correlation coefficients. The autocorrelation functions of the detrended residuals obtained from the wider window moving average are smoother than those from the narrow window moving average. However, the plotted curves do not clearly demonstrate any periodicity at any lag time as the coefficients of the autocorrelation functions are relatively small.

Smooth autocorrelation function with high autocorrelation coefficient of a time series generally implies that periodicity can be more easily detected than the function with many high frequencies. In this case, we might not be able to isolate the recurrence pattern (cycles) of groundwater signals from the residuals of the moving averages. On the other hand, the periodicity might apparently be identified from the detrended residuals obtained by the polynomial approximations as the residuals' autocorrelation functions are relatively smooth with high autocorrelation function values. However, as it will be shown later, the periodicity extracted from each method will be somewhat different. Thus, caution is required because the total data record is only slightly more than 30 years. Therefore, any periodic pattern beyond 15 years may simply be an artifact of the data sets, requiring the original data that are long enough such that real periodicities begin to show up.



(a) (b)

Figure 10. Autocorrelation function of the detrended residuals at CT7_1 with a) low-medium order polynomial interpolation and b) moving average techniques



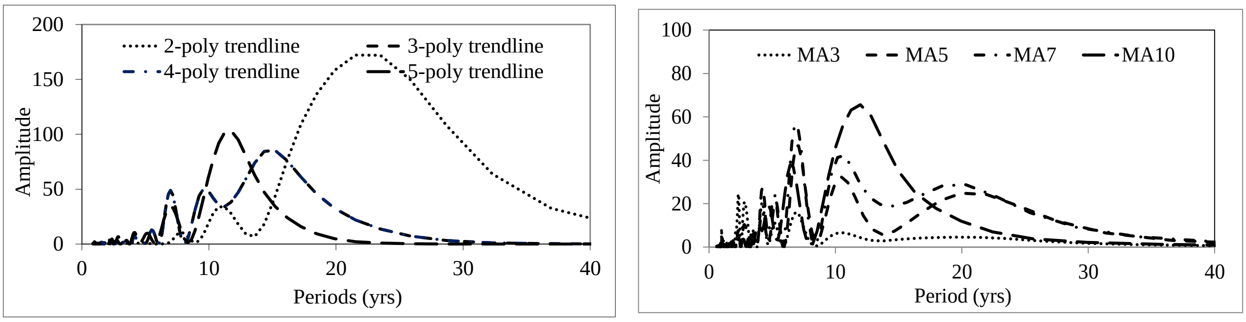
(a) (b)

Figure 11. Autocorrelation function of the detrended residuals at CT7_2 with a) low-medium order polynomial interpolation and b) moving average techniques

5.5. Spectral analysis of the detrended residuals

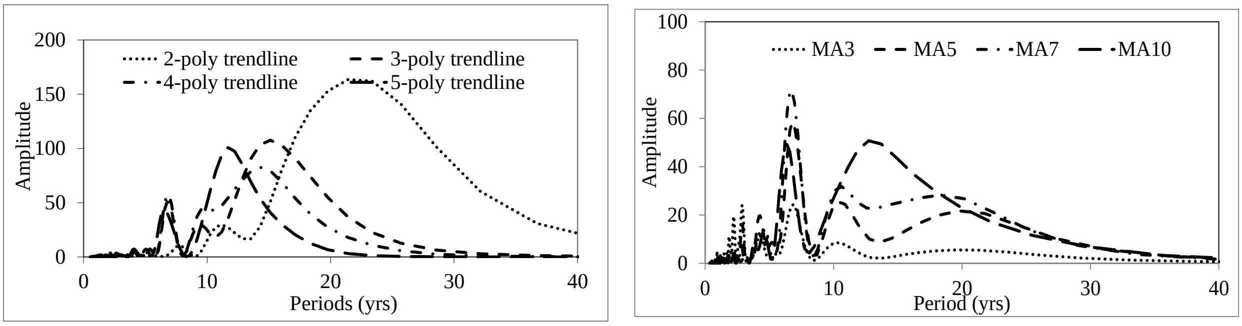
Examples of periodogram obtained from the Fourier analysis on the detrended residuals are shown in Figures 12 and 13. A periodogram is the spectral density estimating the power or strength of periodic components at all possible frequencies.

The patterns of the periodograms of the detrended residuals obtained from both methods are composed of one lower peak at the small periodic recurrence and one taller peak at the larger cycle. The polynomial regression yields more smooth and apparent modes of its periodograms than the moving average detrending. The periodicities obtained from the lower order polynomial regression are generally higher than those from the higher order approximation. This is because detrending a time series by applying a high order interpolation removes low frequency components with high periodicities. On the other hand, the narrow window moving averaging provides lower distributed periodicities than that with the wider window.



(a) (b)

Figure 12. Periodgrams of the detrended residuals at CT7_1 with a) low-medium order polynomial interpolation and b) moving average techniques



(a) (b)

Figure 13. Periodograms of the detrended residuals at CT7_2 with a) low-medium order polynomial interpolation and b) moving average techniques

Different detrending methods provide unique strength and spectral distribution of periodicities and thus difference in their corresponding first and second moments of the distributions. Physically, these first moment values refer to the average period of recurring pattern of the groundwater behavior in which the groundwater fluctuates with the cycles according to these values. The lower the order of polynomial approximation is, the higher the means and the greater variances of the periodicities are. For example, the groundwater at the location CT7_1 shows different periodicities from the 2nd, 3rd, 4th, and 5th order polynomial approximations. The mean cycles of groundwater at the locations are respectively 26.85, 15.35, 15.35, and 8.70 years, and the variations associated with these means referred to as variances of the distributions are 131.32, 26.56, 26.56, and 20.16, respectively. On the other hand, there is no distinct correlation between the first and second moments and the window width of the moving average.

Most of the obtained periodograms appear quite bimodal with two strongest peaks of periodicities; thus, two normal distributions with different means and variances are superimposed together to assemble the best fit in minimizing the sum of squared errors to the actual periodogram obtained from the Fourier analysis. The first normal distribution fitting is attempted to fit the longer cycle and the second distribution fitting is aimed for the shorter cycle. The means of the first and second normal distributions for the different detrending methods are shown in Figures 14 and 15, respectively.

In general, the wider the window of the moving average or the lower the order of the polynomial approximation, the greater the means of the two periodic normal distributions are. For example, the periodic means of the first normal distributions at CT7_1 provided by the 2nd, 3rd, 4th, and 5th order polynomial regressions are 23.39, 15.10, 15.10, and 11.98 years, respectively; the means provided by the 3-, 5-, 7-, and 10-year moving averages are 6.39, 10.24, 13.39 and 12.42 years, respectively.

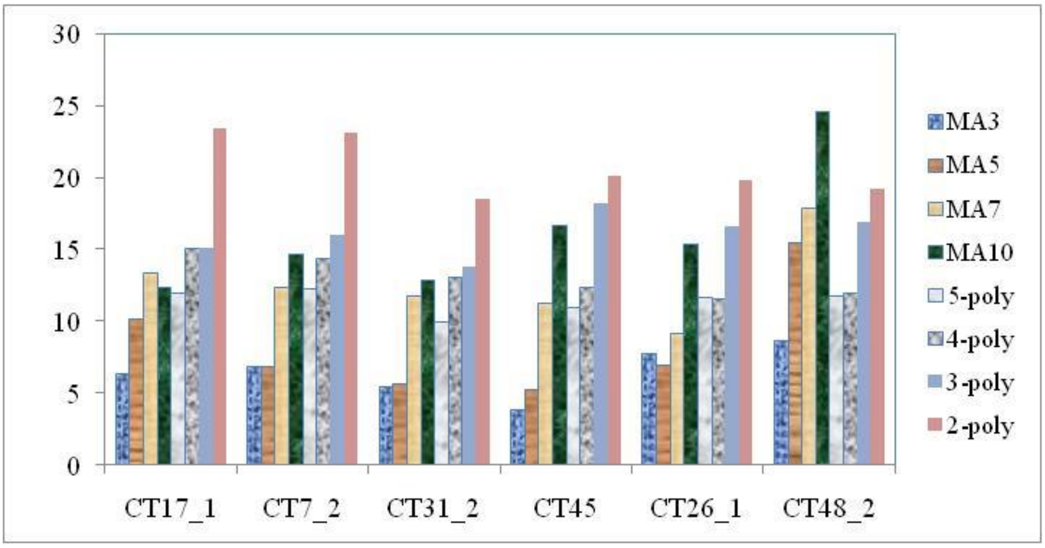


Figure 14. Means of the first normal distributions of the detrended residuals from different regressions of selected time series

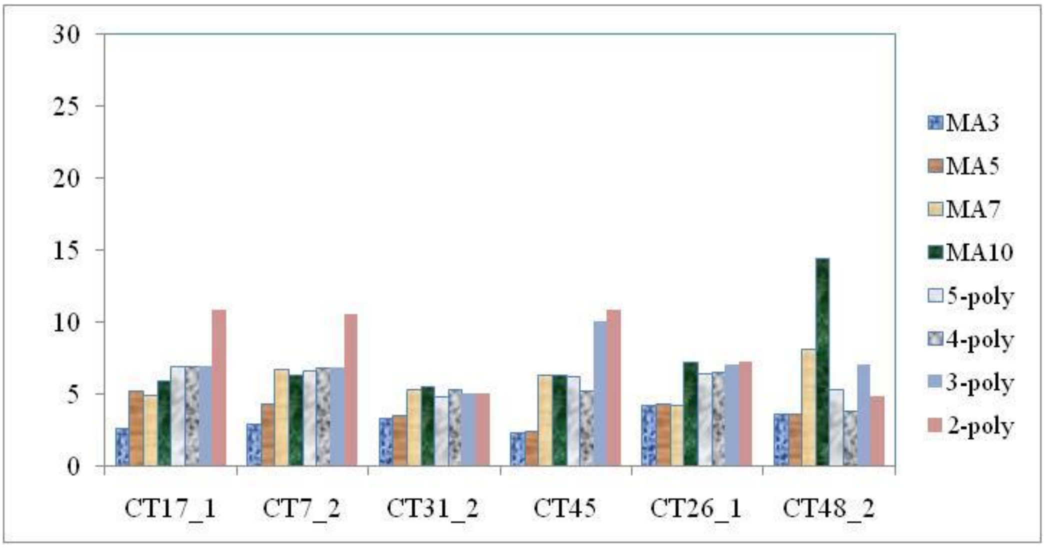


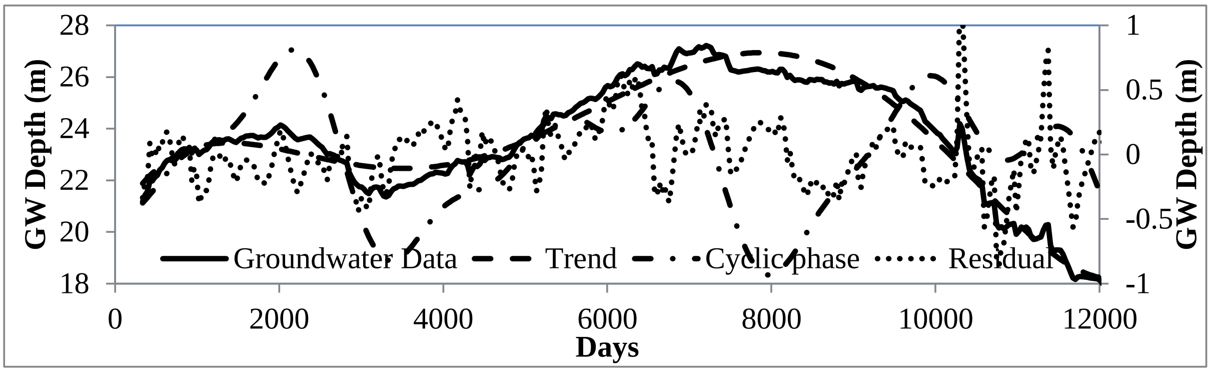
Figure 15. Means of the second normal distributions of the detrended residuals from different regressions of selected time series

On average, the polynomial interpolation yields approximately 30% higher means of both normal distributions than the moving averaging technique. The first hump with large periodic mean is more distributed than the second hump with small periodic mean. These two means of the bimodal periodogram distributions indicate the average cyclic patterns of the deep groundwater in the Lower Chao Phraya Basin that the cyclic patterns could be from the climate variability phenomenon, e.g. ENSO and IOD. This needs some further research investigation. However, different detrending techniques show large variability of the period patterns. For example, the groundwater analysis may show cycles of 2 and 5 years when the 3-year window moving average technique is applied, while the result may conclude that the groundwater has approximately 10 and 20 years of cycles when the 2nd order polynomial interpolation is used. Therefore, caution should be exercised when apply different detrending techniques as different techniques yield different cyclic patterns and may infer to different climate variability phenomenon. An applicable method of detrending that apply to decompose a set of groundwater time series should generate a trend of groundwater that correlates with anthropogenic effects such as pumping record and that does not overfitting the original groundwater time series such that all the important cyclic patterns are removed from the corresponding detrended residuals.

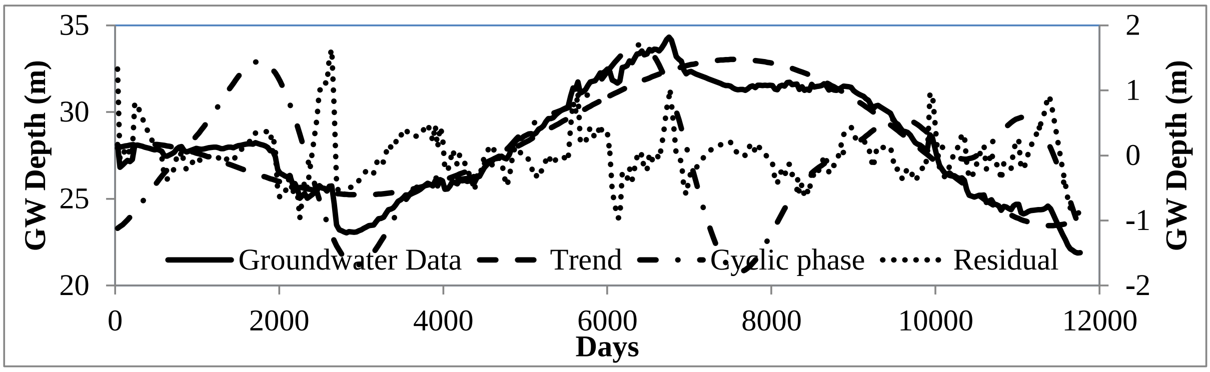
Uncertainty in this analysis can be associated with the length and interval of the groundwater time series data. When the time series is not long enough, some true imbedded large cyclic signals with periodicity longer than half of the monitoring period can be mistakenly alleged as a part of a trend. Only when the data is long enough, the true periodic signals as the result of climatic effects can be established.

5.6. Groundwater components from decomposition and implication for future management

Figure 16 shows the different components as proposed in this study and these components are trend, cyclic phase and residual component. The trend component is human relating showing long term fluctuation as shown in Figure 6 and this work shows that this trend very likely associates with groundwater withdrawal during the heavily pumped period. On the other hand, the cyclic component possibly links to climate variability with short and medium length cycles, i.e. less than 15 years. Fourier analysis establishes two main periodicities in dynamics of deep groundwater in the Lower Chao Phraya Basin and these cyclic patterns can be attributed to the general cycles of global climate, Indian Ocean Dipole (IOD) and the El Niño-Southern Oscillation (ENSO).



(a)



(b)

Figure 16. Trend, cyclic and residual components of groundwater time series at a) CT7_1 and b) CT7_2

From the observation on the magnitude of the groundwater dynamics, human relating factor has influenced far beyond the climate variability on this deep groundwater region. However, changes in climate variability can be reflected in the deep groundwater hydrographs. As depictured in Figure 16, the magnitude of the groundwater change as the result from climate variability in this area ranges between 1and 2 meters, while the extent of the anthropogenic effect on this groundwater is on the order of 10 meter during the period of observation. Therefore, the change in land use and human activities affect the groundwater more than climate variability in this region. We have selected only two cycles of groundwater in this study. More cyclic patterns however can be identified from the periodograms and this can considerably reduce the magnitude of the residual component.

6. SUMMARY

A total of eight detrending methods, 2nd, 3rd, 4th, 5th order polynomial regressions as well as 3-, 5-, 7- and 10-year moving averages, were applied to six long-term monthly groundwater records monitored at stations in the Lower Chao Phraya basin in Thailand where the groundwater is dominantly influenced by anthropogenic effect, i.e. pumping. The correlation of the trendlines to the groundwater withdrawal data reveals that the 5th order polynomial interpolation performs well yielding on average the best fitting of the trendlines linking to the human effect.

The Fourier analysis produces the periodograms of the detrended time series from the different detrending techniques. The periodogram patterns of the detrended residuals obtained from all the methods are comprised of two bimodal distributions. The two means of the bimodal periodogram distributions indicate the average cyclic patterns of the deep groundwater in the Lower Chao Phraya Basin that the cyclic patterns could be from the climate variability phenomenon, e.g. ENSO and IOD.

The novelty of this study are the application of deep groundwater data from the area, which has had long history of heavy pumping, and the sensitivity analysis of different detrending methods highlighting some commonly concerned overfitting issue and uncertainty derived from the empirical detrending approaches. Caution must be exercised when applying different detrending methods for decomposition. Different detrending methods can yield dissimilar periodic patterns of the original data, which may infer to different climate variability phenomenon. In the case of many anthropogenic activities influencing groundwater, it may be able to separate the original time series into a number of trendlines, each representing individual anthropogenic impact. The consequence key issue is to validate that those particular trendlines are actually the results from specific human impacts. Once this is justified, the prediction of the future groundwater behavior related to various human activities may be executed.

ACKNOWLEDGEMENTS

This research has been fully financially supported by a research grant from Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL). The author is deeply grateful to Prof. Dr. Jason J. Gurdak from San Francisco State University for helpful advice and recommendation. Special gratitude is paid to all research assistants and colleagues in the Department of Civil Engineering, KMITL.

REFERENCES

Aflatooni M. and Mardaneh M. (2011), Time series analysis of groundwater fluctuations due to temperature and rainfall change in Shiraz Plain, International Journal of Water Resources and Environmental Engineering, **3** (9), 178-188.

Alley W.M., Healy R.W., LaBaugh J.W. and Reilly T.E. (2002), Flow and storage in groundwater systems, *Science*, **296**, 1985–1990.

Anderson W. and Emanuel R. (2008), Effect of interannual and interdecadal climate oscillations on groundwater in North Carolina, *Geophysical Research Letters*, **35**, 1-4.

Asian Development Bank (ADB) (2007), Recent advances in water resources development and management in developing countries in Asia, Asian Water Development Outlook 2007, Discussion Paper, Manila, Philippines.

Babel M.S., Das Gupta A., Domingo N.D. and Kamalamma A. (2007), Sustainable groundwater management in Bangkok. Report of Bangkok Case Study on Sustainable Water Management Policy (SWMP), Water Engineering and Management, Asian Institute of Technology, Thailand.

Bhar L.M. and Sharma V.K. (2005), Time Series Analysis, Indian Agricultural Statistics Research Institute, New Delhi: 1-15.

Chiew F.H.S., Piechota T.C., Dracup J.A. and McMahon T.A. (1998), El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting, *Journal of Hydrology*, **204**, 138-149.

Cohen L. (1995), Time-frequency Analysis, Prentice Hall PTR.

Davis J.C. (1986), Statistics and Data Analysis in Geology, John Wiley and Sons.

Ganoulis J. and Morel-Seytoux H. (1985), Application of stochastic methods to the study of aquifer systems, Technical Documents in Hydrology, UNESCO, Paris.

Gardner P.M. and Heilweil V.M. (2009), Evaluation of the effects of precipitation on ground-water levels from wells in selected alluvial aquifers in Utah and Arizona, 1936-2005, U.S. Geological Survey Scientific Investigations Report 2008-5242.

Ghanbari R.N. (2007), Climate Signals in Groundwater and Surface Water Systems: Spectral Analysis of Hydrologic Processes. University of Wisconsin – Milwaukee.

Green T.R., Taniguchi M., Kooi H., Gurdak J.J., Allen D.M., Hiscock K.M., Treidel H. and Aureli A. (2011), Beneath the surface of global change: Impacts of climate change on groundwater, Paper 851. USDA-ARS/UNL Faculty.

Gunatilaka A. (2005), Groundwater woes of Asia, *Asian Water*, January/February, 19-23.

Gurdak J.J., Randall T.H., McMahon P.B., Bruce B.W., McCray J.E., Thyne G.D. and Reedy R.C. (2007), Climate variability controls on unsaturated water and chemical movement, High Plains Aquifers USA, *Vadose Zone Journal*, **6**, 533-547.

Hammer O. and Harper D.A.T. (2006), Paleontological Data Analysis, Blackwell.

Hammer O., Harper D.A.T. and Ryan P.D. (2001), PAST: Paleontological Statistics Software Package for Education and Data Analysis, *Paleontologia Electronica*, **4**(1), 1-9.

Hanson R.T., Newhouse M.W. and Dettinger M.D. (2004), A methodology to assess relations between climatic variability and variations in hydrologic time series in the southwestern United States, *Journal of Hydrology*, **287**, 252-269.

Harper D.A.T. (1999), Numerical Palaeobiology, John Wiley and Sons.

Institute for Global Environmental Strategies (IGES) (2007), Sustainable groundwater management in Asian cities. Report of research on sustainable water management policy.

Intergovernmental Panel on Climate Change (IPCC) (2007), Climate Change 2007 – Impacts, adaptation and vulnerability working group II contribution to the Fourth Assessment Report of the IPCC.

Kiely G. (1999), Climate change in Ireland from precipitation and stream flow observations, *Advances in Water Resources*, **23**, 141-151.

Kundzewicz Z.W. and Doll P. (2008), Will groundwater ease freshwater stress under climate change? Proceedings of International Conference on Groundwater and Climate in Africa, Kampala, June 2008.

Kusreesakul K. (2009), Spatio-temporal Rainfall Changes in Thailand and Their Connection with Regional and Global Climate Variability, Master Thesis, Prince Songkhla University, Thailand.

Limsakul A. and Goes J. (2008), Empirical evidence for interannual and longer period variability in Thailand surface air temperature, *Atmospheric Research*, **87**, 89-102.

Limsakul A., Limjirakan S. and Suthamanuswong B. (2007), Spatio-temporal changes in total annual rainfall and the annual number of rainy days in Thailand, *Journal of Environmental Research*, **29**(2), 1-21.

Lorphensri O. (2011), Controlling groundwater abstraction and related environmental degradation in Metropolitan Bangkok-Thailand, Asia Pacific Water Forum (APWF) Regional Water KnowledgeHub for Groundwater Management, June 2011, Bangkok.

Mantua N.J. and Steven R.H. (2002), The Pacific Decadal Oscillation, *Journal of Oceanography*, **58**, 35-44.

Mayer T.D. and Congdon R.D. (2008), Evaluating Climate Variability and Pumping Effects in Statistical Analyses, *Ground Water*, **46**(2), 212-227.

Nagpaul P.S. (2005), Time Series Analysis in WinIDAMS, New Delhi, India. http://portal.unesco.org/ci/en/files/18650/11133194701TimeSeriesAnal.pdf/TimeSeriesAnal.pdf

Ngongondo C.S. (2006), An analysis of long-term rainfall variability, trends and groundwater availability in the Mulunguzi river catchment area, Zomba mountain, Southern Malawi, *Quaternary International*, **148**, 45-50.

Oppenheim A.V., Willsky A.S. and Hamid S. (1997), Signals and Systems, Prentice Hall.

Pankratz A. (1983), Forecasting with univariate Box Jenkings Models: Concepts and Cases, Whiley Series in Probability and Statistics.

Perry C.A. (1994), Solar-irradiance variations and regional precipitation fluctuations in the western United States, *International Journal of Climatology*, **14**, 969-983.

Perry C.A. and Hsu K.J. (2000), Geophysical, archaeological, and historical evidence support a solar-output model for climate change, *Proceedings of National Academy of Science*, **97**(23), 1244-1248.

Piancharoen C. (1977), Groundwater and land subsidence in Bangkok, Thailand, IAHS Publication No. 121, 355-364.

Piancharoen C. and Chuamthaisong C. (1978), Groundwater of Bangkok Metropolis, Thailand, IAH Memoire Vol. XI, Budapest, 510-528.

Polemio M. and Casarano D. (2008), Climate change, drought and groundwater availability in southern Italy. In: Climate change and groundwater, Dragoniand W. and Sukhija B.S. (eds), The Geological Society Special Publications, 288, 39-51.

Pratarastapornkul T. (2009), An apparent effect of Indian Ocean Dipole phenomenon on annual rainfall variability of Thailand, Proceedings of the fifth IOGOOS annual Meeting, Phuket, Thailand. (in Thai with abstract in English)

Prinos S.T., Leitz A.C. and Irvin R.B. (2002), Design of a real-time ground-water level monitoring network and portrayal of hydrologic data in southern Florida, U.S. Geological Survey Water-Resources Investigations Report 01-4275, 108 p.

Serfes M.E. (1991), Determining the mean hydraulic gradient of groundwater affected by tidal fluctuations, *Ground Water*, **29**(4), 549-555.

Shah T. (2009), Climate change and groundwater: India's opportunities for mitigation and adaptation, *Environmental Research Letters,* 4 035005, doi:10.1088/1748-9326/4/3/035005.

Shao Q., Campbell N.A., Ferdowsian R. and O' Connell D.O. (1999), Analyzing trends in groundwater levels, CSIRO Mathematical and Information Sciences Report No. CMIS 99/37, Agriculture Western Australia.

Tague C., Grant G., Farrell M., Choate J. and Jefferson A. (2008), Deep groundwater mediates streamflow response to climate warming in the Oregon Cascades, *Climate Change*, **86**, 189-210.

Taylor C.J. and Alley W.M. (2001), Groundwater level monitoring and the importance of long-term water-level data, U.S. Geological Survey Circular 1217.

Tomasella J., Hodnett M.G., Cuartas L.A., Nobre A.D., Waterloo M.J. and Oliveira S.M. (2008), The water balance of an Amazonian micro-catchment: the effect of interannual variability of rainfall on hydrological behavior, *Hydrological Process*, **22**, 2133–2147.

United Nations Environment Programme (UNEP) (2002), Global Environment Outlook 3. Nairobi, Kenya.

Vautard R., Yiou P. and Ghil M. (1992), Singular-spectrum analysis: A toolkit for short, noisy, chaotic signals, *Physica D*, **58**, 95-126.

Water Environment Partnership in Asia (WEPA) (2007), http://www.wepa-db.net/policies/state/china/groundwater.htm.

Yu H.L. and Chu H.J. (2012), Recharge signal identification based on groundwater level observations, *Environmental Monitoring and Assessment*, **184**, 5971-5982.