Advanced Geomatics

http://publish.mersin.edu.tr/index.php/geomatics/index

e-ISSN: 2791-8637

Least-Squares Spectral Analysis of Hourly Tide Gauge Data – A Case Study

Ramazan Alpay Abbak*[1]

[1]*Department of Geomatics Engineering, Konya Technical University, 42075, Konya,*

Keywords

Least-squares,
Period,
Periodicity,
Spectral analysis,
Tide gauge data.

Research Article

Received : 11.12.2022
Revised : 14.02.2023
Accepted : 09.03.2023
Published : 31.03.2023

Abstract

Tide gauge observations are samples of geodetic time series realized depending on the time. These observations like other experimental time series might have trends, short gaps, datum shifts and unequally spaced data which usually make disturbing effects to the analysis. In other methods (e. g. classical Fourier Transform), trend is removed before the analysis, the others (i.e. short gap and unequally spaced data) are taken over by filling any interpolation techniques. In this case the editing may produce well-composed time series, but it may obliterate the useful information in the series or even introduce artificial signals. This means that unwanted results take place during the process. There is an alternative method, called the Least-Squares Spectral Analysis (LSSA) which can bypass these problems without editing or pre-processing. In the present study, hourly sea level observations obtained from the Antalya tide gauge in Turkey were analyzed by using the LSSA method. Consequently, five hidden periodicities were successfully determined from the sea level observations containing difficulties mentioned above.

1. Introduction

The fundamental aim of geodetic studies is to determine the shape and gravity field of the Earth. In this context, several geodetic observations, such as the distances, directions, height differences and gravity surveys are performed by different techniques. These geodetic observations are always effected from all physical forces or process on the Earth surface. Therefore, some observations are made depending on time or space in order to examine the physical process comprehensively (Abbak, 2007). For example, tide gauge data changing with the gravity effect of Sun and Moon, is the most well-known time series in geodesy. As being the independent observations, determining unknown parameters in random process (i. e. time series) is subject to investigate in geodetic statistics (e. g. Zeray Öztürk et al., 2020; Zeray Öztürk and Abbak, 2020).

The time series have many drawbacks bothering the analysis. Some of them are the presence of datum shift, trend, short gap, unequally spaced and weighted data in the time series. Several synthetic methods are used to bypass these difficulties in practice. For example, (i) trend is removed before the analysis; (ii) if the data has two or more datum shifts, every datum is analyzed separately, which means the disorder of complete analysis; (iii) provided that the time series has short gaps or unequally spaced data, the gaps and certain values are predicted and filled by the harmonic analysis or any interpolation technique (e. g. Craymer, 1998). The whole temporary solutions discussed above cause to a series of new problems in the analysis. Thus an alternative technique is developed without the corruption of originality of the data, which is called “Least Squares Spectral Analysis (LSSA)”.

The LSSA method was first released and developed by Vaníček between 1969 and 1971 (Vaníček, 1969a; 1971). Therefore, the method was also named as “Vaníček Spectral Analysis” in geodetic literature (Taylor and Hamilton; 1972). Afterwards a wide range of researchers successfully applied the technique in geodesy and related fields such as electronic distance measurement, tidal data, superconducting gravimeter data, mesospheric temperature estimating and star positioning (Craymer, 1998; Abbasi, 1999; Omerbashich, 2003; Espy and Witt, 1996; Mantegazza, 1997; Abbak, 2005).

The main aim of the LSSA method is to find the hidden periodicity in the time series consisting of the difficulties

* **Corresponding Author**

*(raabbak@ktun.edu.tr) ORCID ID 0000-0002-6944-5329

Cite this;

Abbak, R. (2023). Least-Squares Spectral Analysis of Hourly Tide Gauge Data – A Case Study. Advanced Geomatics, 3(1), 23-27.

mentioned above, and thus clarify the physical events that cause these periodicities. For this purpose, in this study an investigation was conducted in order to determine phenomena causing to the periodic changes in the sea level observations.

The paper starts with a brief review of the computational scheme of the LSSA method. Then, the structure of sea level observations is explained as an input data. Next, the LSSA method is applied to tide gauge data for the determination of the periodicities in the sea level changes. Finally, the results obtained from the current analysis are compared with astronomic counterparts.

2. Least Squares Spectral Analysis

2.1. Basic approach

The LSSA is an application of least-squares approach in the time series analysis. This method uses least squares principle to minimize the norm of the error vector of the time series, estimating the periodic components as well as non-periodic ones (Amerian and Voosoghi, 2011).

The experimental time series consist of the signal and noise. The noise disturbing on observation is divided into two components: random and systematic. An ideal noise called "white noise" is uncorrelated with any dependent variable. Nevertheless, in practice a systematic noise called coloured noise occurs and is correlated with one or more variables. Systematic noise can be defined by mathematical functions, but their magnitudes are not known in the time series. It is categorized by two types: periodic and non-periodic. The main aim is to able to determine these systematic noises (both periodic and non-periodic).

2.2. Mathematical Basics

The problem is following:

1. Vector of observation time: $t_i,\ i = 1,\ 2,\ \ldots,\ n$
2. Vector of observables: $\boldsymbol{f}(t_i)$
3. Vector of frequencies whose spectral values are desired: $\omega_j, j = 1,\ 2,\ \ldots,\ m$

The $\boldsymbol{s}(\omega_j)$ vector, the spectral value of the ω_j frequency, is sought by spectral analysis. There are several ways of computing a spectrum from a time series. Here we simply state the problem in the Least Squares Approximation (LSA). Thus the time series can be modelled by simple formula,

$$\boldsymbol{g} = \boldsymbol{A}\boldsymbol{x} \tag{1}$$

where $\boldsymbol{A}$ is the design matrix modelling the physical relationship between the observations and the vector of unknown parameters $(\boldsymbol{x})$. The parameters of the design matrix are summarized by:

$$\boldsymbol{A} = \begin{bmatrix} cos\ \ \omega_j t_1 & sin\ \ \omega_j t_1 \\ cos\ \ \omega_j t_2 & sin\ \ \omega_j t_2 \\ \vdots & \vdots \\ cos\ \ \omega_j t_n & sin\ \ \omega_j t_n \end{bmatrix} \tag{2}$$

For computing these parameters by least-squares approximation, differences between $\boldsymbol{f}$ and $\boldsymbol{g}$ functions will be minimum. Using the least-squares notation (e. g. Vanícek and Krakiwsky 1986), it is written,

$$\hat{\boldsymbol{r}} = \boldsymbol{f} - \hat{\boldsymbol{g}} = \boldsymbol{f} - \boldsymbol{A}(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{P}\boldsymbol{A})^{-1}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{P}\boldsymbol{f} \tag{3}$$

where $\hat{\boldsymbol{r}}$ is the residual vector, $\hat{\boldsymbol{g}}$ is the best approximation of $\boldsymbol{g}$ by the least-squares. Then spectral values can be calculated by the ratio,

$$\boldsymbol{s} = \frac{\boldsymbol{f}^{\boldsymbol{T}}\hat{\boldsymbol{g}}}{\boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{f}} \tag{4}$$

If the spectral value is desired to compute for a frequency ω_j, the result is

$$\boldsymbol{s}(\omega_j) = \frac{\boldsymbol{f}^{\boldsymbol{T}}\hat{\boldsymbol{g}}(\omega_j)}{\boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{f}} \tag{5}$$

In the geodetic literature related with the parameter estimation using the least-squares, it is essential important to evaluate the results statistically. Thus a great superiority of the LSSA method is that the significance of peaks in the spectrum can be tested statistically. Pagiatakis (1999) introduces a test as follows: a statistical test of null hypothesis $H_0 : s(\omega_j) = 0$ can be tested using a decision function by

$$s(\omega_j)\begin{cases} \leq (1 + (\alpha^{-2m/v} - 1)^{-1})^{-1}; & accept \quad H_0 \\ > 1 + (\alpha^{-2m/v} - 1)^{-1})^{-1}; & reject \quad H_0. \end{cases} \tag{6}$$

where α is the significance level (usually 5%), m is the number of frequencies participated in simultaneously estimation of the LSSA, and $v = n - 2m$ is the degrees of freedom including n equal to the number of data.

3. Analyses of the Tide Gauge Data

3.1. Input data

By using the LSSA method, this study investigates tidal components in the sea level observation acquired from the tide gauge. The tide gauge is an observatory, which continuously records the instant sea level. The systems in the tide gauge are divided into two categories: analog and digital. Analog system is based on the drum which records the sea level (Fig. 1). The plotter over the drum reflects the sea level changes to the paper on the drum in real time. The data obtained from analog system is not suitable for the numerical analysis due to its graphical structure. Today this system is out of the date owing to the mechanical reasons and digitization errors.

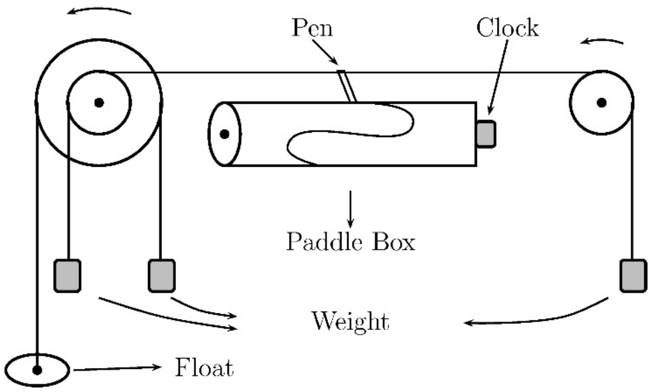

Figure 1. Analog system on tide gauge

Before the analysis, we should consider the structure of the sea level data more closely. The sea level observations are decomposed to three components (Ainscow et al., 1985):

$$\text{Observation} = \text{mean sea level} + \text{tides} + \text{atmospheric residuals} \tag{7}$$

where the components are related with the different physical process and uncorrelated with each other. The mean sea level is determined by averaging hourly data at least one-year long. The tides are the periodic sea level changes which are the consistent with some geophysical effects and are of the certain amplitude and phase. Atmospheric residuals are irregular variations related with weather conditions, which is so-called "surge effect".

Tides mentioned above are explained by Newton's gravity force. This force is summarized by

$$F = G\frac{m_1 m_2}{r^2} \tag{8}$$

where G is the Newton gravitational constant, r is the distance between m_1 and m_2. Provided that the Earth and Moon masses, and also distance between Earth and Moon are inserted to equation, the result is the gravitational force between Earth and Moon. The gravitation between the Earth and Sun can be achieved similarly. Both gravitations affect sea level changes in same periods with the movements of the Sun and Moon. Accordingly, tidal components are shown in Table 1.

Table 1: Short period Primary tidal components (ORNL, 2011)

Tidal components	Period (Sun time)	Description	Explanation
M2	12.42	Principal lunar	Semidiurnal
S2	12.00	Principal solar	Semidiurnal
N2	12.66	Larger Lunar elliptic	Semidiurnal
K2	11.97	Luni-solar	Semidiurnal
K1	23.93	Luni-solar diurnal	Diurnal
O1	25.82	Principal lunar diurnal	Diurnal
P1	24.07	Principal solar diurnal	Diurnal
Q1	26.87	Larger Lunar elliptic	Diurnal

After these explanations, we concern the source of error in the tide gauge data. Generally, three types of systematic noise contaminate time series from our point of view. Firstly, there can be a step function due to sudden changes in the tide gauge datum (possibly caused by vertical displacements). The dates of such step functions are usually well-documented, but their magnitudes may not be documented properly. Secondly, there may be a gradual change in the tide gauge datum due to sea level rise or land subsidence around the tide gauge, which is simply modelled by a linear trend. Thirdly, there are data gaps resulted from the equipment failures.

3.2. Numerical Application

In the application process, the observations obtained from the Antalya tide gauge in 1990 in Turkey were analyzed using the LSSA method. These observations belong to analog system of the tide gauge. Therefore, the Turkish General Directorate of Mapping (TGDM) digitalized them and then corrected for the outliers (TGDM, 1991).

Before the numerical analysis, at first the observations were plotted in month by month. Thus the datum shift and short gaps in time series are determined easily. While the observations are examined carefully, it is clear that there is a variety of difficulties in the time series (see, Fig. 2)

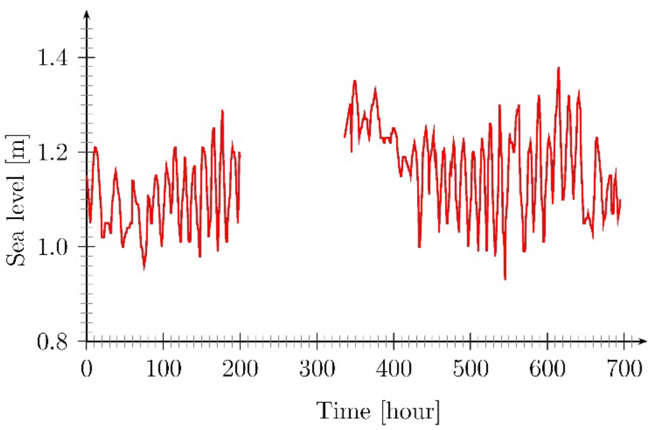

Figure 2: A sample data from the Antalya tide gauge (April 1990)

The software developed by Wells et al. (1985), was utilized for the analysis. This software was encoded in FORTRAN Programming Language by the LSSA algorithm. The interested reader can be found this program at the web page of LSSA (2022). The program starts with input data and some settings (i. e. whether some known constituents such as the datum shift, linear trend etc., are present or not).

The hourly sea level data obtained from every month was analysed in a sequence. In order to predict the periods as possible as accurate, two spectral bands were used for this process: minimum, maximum and number of periods were selected as firstly 11.5h, 13.5h and 500, secondly, 23.5h, 26.5h and 500. As a result of this analysis, five periods were successfully estimated in the sea level observations (Table 2).

Table 2: The periods estimated by the LSSA method [hour]

Months	Period #1 (S2)	Period #2 (M2)	Period #3 (N2)	Period #4 (K1)	Period #5 (O1)
January	11,988	12,418		23,989	25,694
February	11,999	12,438	12,778	24,097	25,516
March	11,947	12,394	12,579		25,744
April	12,119	12,490	12,829	23,940	24,986
May	11,947	12,498	12,795	23,876	25,834
June	11,943	12,382	12,551	24,011	
July	12,051	12,406	12,604	24,032	25,564
August	12,040	12,414	12,583	23,869	25,968
September	11,995	12,394	12,637	23,876	25,818
October	11,973	12,406	12,612	23,826	25,475
November	12,014	12,450	12,787	23,742	25,959
December	11,976	12,422	12,720	23,862	
Average	12,00	12,43	12,68	23,92	25,66
Stand. Dev.	0,05	0,04	0,10	0,10	0,29

In order to check whether the results are compatible with astronomical values or not, a test was carried out by using the student distribution (t test). A null hypothesis $H_0 : \bar{x} = \mu_0$ can be tested using a decision function by

$$t_{test} = \frac{|\bar{x} - \mu_0|}{\sigma_x / \sqrt{n}} \begin{cases} < t_{\alpha,\, f}; & accept\ H_0 \\ \geq\ t_{\alpha,\, f}; & reject\ H_0 \end{cases} \tag{9}$$

where α is the significance level (5%), f is the degree of freedom, σ_x is the standard deviation of the periodicity, $\bar{x}$ is the average value of the periodicity, and μ_0 is the astronomical value of the periodicity. As a result of the test, the periods derived from the analysis are well-consistent with the astronomical ones in a confidence level of 95%. This result shows that the LSSA method is successfully on determination of the hidden periodicity in the geodetic time series induced by the disturbing effects.

Although there are totally eight diurnal and semidiurnal tidal components, in the current study five of them can be determined by the LSSA method. The reasons of incomplete frequencies are evaluated that the other components might have low amplitudes, or that the dock position of the tide gauge might be unsuitable to sense the other components.

4. Conclusion

This paper has summarized the theory of the LSSA approach to spectral analysis of the time series. The LSSA is an efficient method to determine the periodicities of the experimental time series that have the difficulties. In other words, the LSSA can be carried out directly to the experimental time series without editing or pre-processing. Furthermore, all unknown parameters are found simultaneously, while a rigorous test is implemented to assess the significance of the determined peaks in the spectrum.

In order to exhibit the powerfulness of the method, a numerical application was accomplished in the sea level observations which were affected by many disturbances. As a result of the numerical analysis, five consistent periods were determined in those series. Consequently, the numerical analyses show that the LSSA method gives us reasonable results in the experimental time series. Thus it is strongly recommended that the method should be used in any geodetic time series, e. g. GNSS time series.

Acknowledgement

Author thanks to the Turkish General Directorate of Mapping for providing the data in this study.

Author contributions

R. Alpay Abbak designed the study, determined the analysis plan, investigated the results, and wrote the manuscript.

Conflicts of interest

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

Research and publication ethics were complied with in the study.

References

Abbak, R. A. 2005. Least Squares Spectral Analysis of Sea Level Observations. Master Thesis, Institution of Natural and Applied Sciences, Selcuk University.

Abbak, R. A. (2007). Wavelet Analysis of Time Series in Geodesy. PhD Seminar, Institution of Natural and Applied Sciences, Selcuk University.

Abbasi M., 1999, "Comparison of Fourier, Least Squares and Wavelet Spectral Analysis Methods Test on Persian Gulf Tidal Data", Master's Thesis, Surveying Engineering Department, K. N. Toosi University of Technology, Tehran, Iran.

Ainscow B., Blackman D., Kerrigde J., Pugh D., and Shaw S. 1985, "Manual on Sea Level Measurement and Interpretation", Volume I, UNESCO.

Amerian Y., Voosoghi B., 2011, "Least Squares Spectral Analysis for Detection of Systematic Behaviour of Digital Level Compensator", Journal of Geodetic Science, Vol: 1(1), pp: 35—40.

Craymer M. R. 1998, "The Least Squares Spectrum, Its Inverse Transform and Autocorrelation Function: Theory and Applications in Geodesy", PhD thesis, Graduate Department of Civil Engineering, University of Toronto.

Espy P. J. and Witt G. 1996, "Observations of a Quasi 16 Day Wave in the Polar Summer Mesospheric Temperature", Journal of Geophysical Research, Vol: 102, pp: 1—32.

LSSA, 2022, "Least-Squares Spectral Analysis Program", Accessed date: September 2022, http://www3.sympatico.ca/craymer/software/lssa.

Mantegazza L., 1997, High Azimuthal Number Pulsation Modes in Fast Rotating ⍰ Scuti Stars: The Case of HD 101158⍰V837 Cen, Astronomy and Astrophysics, Vol: 23 pp: 844—852

Omerbashich M., 2003, "Earth-Model Discrimination Method", PhD Thesis, Geodesy and Geomatics Engineering, University of New Brunswick.

ORNL, 2022, Oak Ridge National Laboratory, http://www.phy.ornl.gov/csep/CSEP/-OM/NODE31.html

Pagiatakis S. D., 1999, "Stochastic Significance of Peaks in The Least-Squares Spectrum", Journal of Geodesy, Vol: 73, pp: 67—78.

Taylor J. and Hamilton S., 1972, "Some Tests of the Vaníček Method of Spectral Analysis", Astrophysics and Space Science, Volume: 17, pp: 357—367.

TGDM, 1991, "The Hourly sea level records for the year 1990 at Erdek, Mentes/Izmir, Bodrum, Antalya tide-gauges", Turkish General Diretorate of Mapping, Ankara.

Vaníček P., 1969a, "Approximate Spectral Analysis by Least Squares", Astrophysics Space Science, Volume: 4, pp: 387—391.

Vaníček P., 1969b, "New analysis of the earth pole wobble". Studia Geophysica et Geodaetica, Volume: 13, pp: 225—230.

Vaníček P. 1971, "Further Development and Properties of the Spectral Analysis by Least Squares", Astrophysics Space Science, Vol: 12, pp: 10—33.

Vaníček P. and Krakiwsky E. J. 1986, Geodesy Concepts: Part 1, Elsevier Book Company, 2nd Edition, 748 p., Amsterdam

Wells D. E., Vaníček P. and Pagiatakis S. D. 1985. "Least Squares Spectral Analysis Revisited", Technical Report 84, Geodesy and Geomatics Engineering, University of New Brunswick.

Zeray Ozturk, E. and Abbak, R. A. 2020. "PHCSOFT: A Software package for computing physical height changes from GRACE based global geopotential models". Earth Science Informatics, 13(4):1499–1505.

Zeray Ozturk, E., Godah, W., and Abbak, R. A. 2020. "Estimation of physical height changes from GRACE satellite mission data and WGHM over Turkey". Acta Geodaetica et Geophysica, 55(2):301–317.

© Author(s) 2023.
This work is distributed under https://creativecommons.org/licenses/by-sa/4.0/