Times Series Regression and Spectral Analysis on Nitrate Concentrations in Long Island, New York

Thomas Plitnick[1], tplitnick1@pride.hofstra.edu
Richard Lagattolla[1], rlagattolla1@pride.hofstra.edu
Katerina G. Tsakiri[2], ktsakiri@rider.edu
Charalambos Petrocheilos[3], charalambos.petrocheilos@postgrad.plymouth.ac.uk
Antonios E. Marsellos[1], antonios.e.marsellos@hofstra.edu

[1]*Department of Geology, Environment and Sustainability, Hofstra University, Hempstead, NY*
[2]*Department of Information Systems and Supply Chain Management, College of Business Administration, Rider*
University, Lawrence, NJ
[3]*School of Computing, Electronics, and Mathematics at Plymouth University, United Kingdom*

---

1. Introduction:

Long Island, New York encompasses an area of 1,401 mi^2 and contains four counties. Of the four counties Nassau and Suffolk counties make up the largest portion by area. Long Island sits on top of multiple large sand and gravel aquifers which support the local population vast need for potable water. The geologic construct of Long Island has established a series of three main large aquifers to cover the region: the Loyd, Magothy and Upper Glacial aquifers (NYSDEC, 2019). The extensive coverage and varying depth of these aquifers allow for Long Island residents to use these underground reservoirs to support the growing need for more potable water. Nitrate levels studies on the surface or in the water has been an increasing issue in Long Island and is of major concern to the health of the public due to the use of aquifers as a source of public drinking water (Tamborski et al., 2017; NYSDEP, 2014; Bleifuss, 1998; Bleifuss et al., 1997; Flipse, 1984). An increase in nitrate concentration can be attributed to certain variables that affect the long-term correlation to the changing nitrate levels. As there are increased levels of anthropogenic nitrate over the decades nitrate accumulation in surface soils increases as time progresses. It is important to understand what atmospheric and hydrogeological variables have the most effect on the concentration of nitrate near or within the surface soil (Kovač, 2018). Examining Long Island from the east end to west end there were three United States Geological Survey (USGS) water data stations and three USGS groundwater wells that data were obtained from (USGS, 2019) along with one atmospheric data station from the National Oceanic and Atmospheric Administration (NOAA, 2019). Regression and spectral analysis was completed in relation to atmospheric and hydrogeological factors such as pH, salinity, depth to groundwater, water temperature, tides, and precipitation daily data for a period of seven years from 2012 to 2019 to investigate the long-term trend of nitrate maximum values and the nitrate periodicities.

2. Methodology

In order to evaluate the effects of pH, salinity, depth to groundwater level, precipitation, tidal height, and water temperature on the concentration of nitrate, water data were obtained from the United States Geological Survey (USGS, 2019) (water data locations; USGS-01311143, USGS-01304200, USGS-01304562 and relatively near well locations; USGS-403517073430702N, USGS-405149072532201S, USGS-405830072331502S. All atmospheric data were obtained from the National Oceanic and Atmospheric Administration (NOAA, 2019). Once the data were obtained, variables were inserted into the Konstanz Information Miner (KNIME), an open source data analytics and data mining software, reporting and integration platform. In this study, the raw data was mined from their original url sources. All data used was in daily values, and missing data were replaced with the previous daily value. A periodicity of 177 days was selected as it was identified by spectral analysis using time series analysis in R software (TSA package; Chan and Ripley, 2018). A low pass filter, named as Kolmogorov-Zurbenko filter was applied in order to determine the long-term cycles in the data. In addition, a linear regression analysis was conducted in order to identify any possible similarities between the variables (Yang et al. 2010). The filtered data were transferred into Statistical Package for the Social Sciences (SPSS) where further investigation of the linear regression model was performed.

3. Results

To examine the atmospheric and hydrogeological variables that affect the nitrate concentration in Long Island daily data was retrieved from three USGS well water stations, three USGS water data stations and one weather data station. To determine the variables that have the maximum correlation with nitrate concentration in Long Island, NY three locations were selected. Moving from west to east across Long Island the three stations were: Hog Island Channel, Peconic River and Orient Harbor, the potentiometric contour lines were examined to study how the flow and elevation of groundwater can have an effect on nitrate concentration in the shallow subsurface (Figure 1). The linear regression has derived an adjusted coefficient of determination, R^2 value of 0.81 with the variables used in the model which are the maximum pH value, minimum salinity and the water temperature value at the Hog Island Channel location. Data examined further east at Peconic River the linear regression model has derived an adjusted coefficient of determination, R^2 of 0.41, and furthest east at Orient Harbor the adjusted coefficient of determination, R^2 value is 0.35. Using the periodogram in spectral analysis, we have determined two main periodicities at each USGS station. Stations at Hog Island Channel, Peconic River, and Orient Harbor show the first peak periodicities at 640, 800, and 2400 days, respectively, while the second peak periodicities are 160, 240, 600 days, respectively (Table 1, Figure 2).

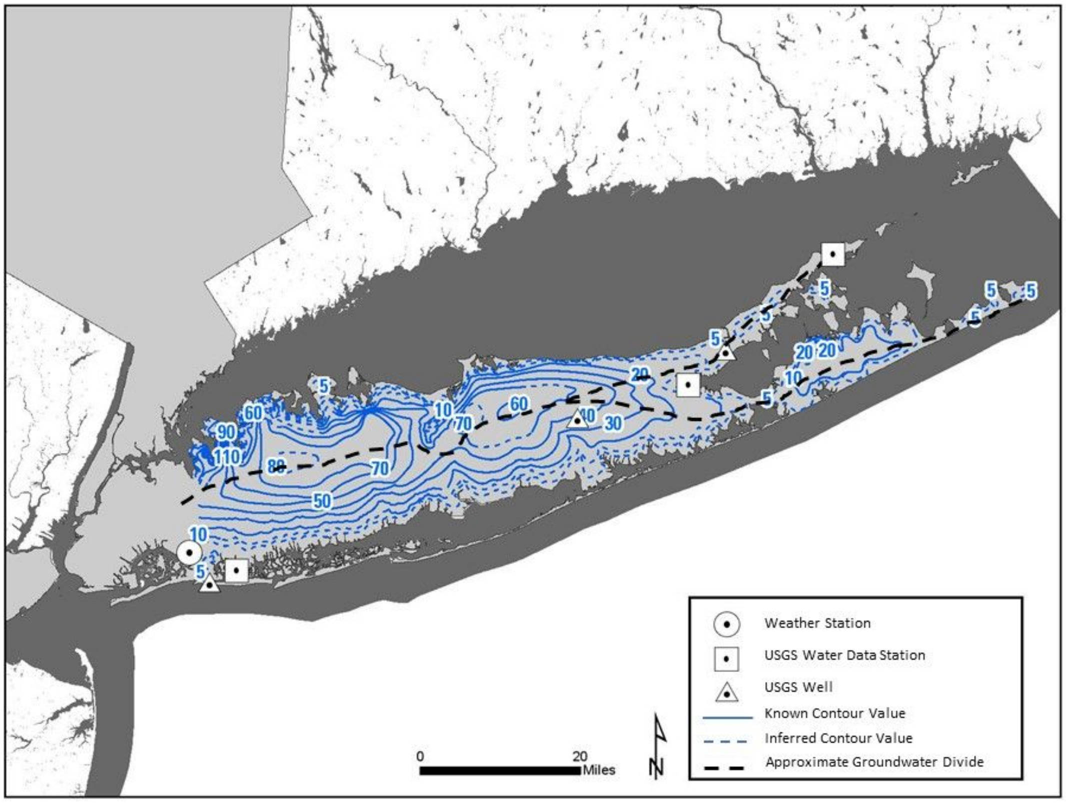

Figure 1: Potentiometric contour lines for Nassau and Suffolk County. Long Island's groundwater divide runs west to east and is approximated on this map in dashed black lines. Map contour lines obtained from the USGS ScienceBase, 2018.

Table 1: Periodicity, frequency, and spectral density are presented for the groundwater stations located at Hog Island Channel, NY, Peconic River, NY and Orient Harbor, NY.

Hog Island

Frequency (cycles/time unit)	Spectral	Period (days)
0.00156	2.86	640
0.00625	2.85	160

Peconic River

Frequency (cycles/time unit)	Spectral	Period (days)
0.00125	6.66	800
0.00417	3.23	240

Orient Harbor

Frequency (cycles/time unit)	Spectral	Period (days)
0.000417	3.17	2400
0.00167	3.12	600

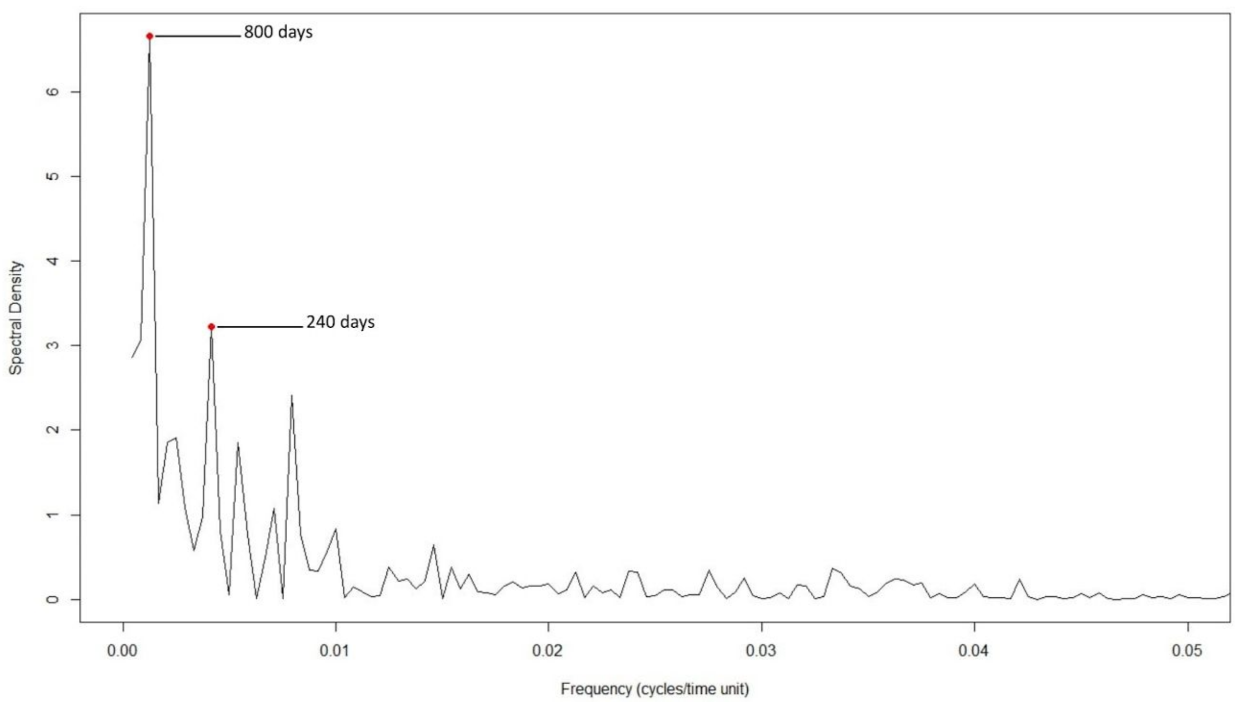

Figure 2: Periodograms showing the major cycles of groundwater for Hog Island region. Periodicities (in days) are showing with red bullet on each corresponding spectral density peak.

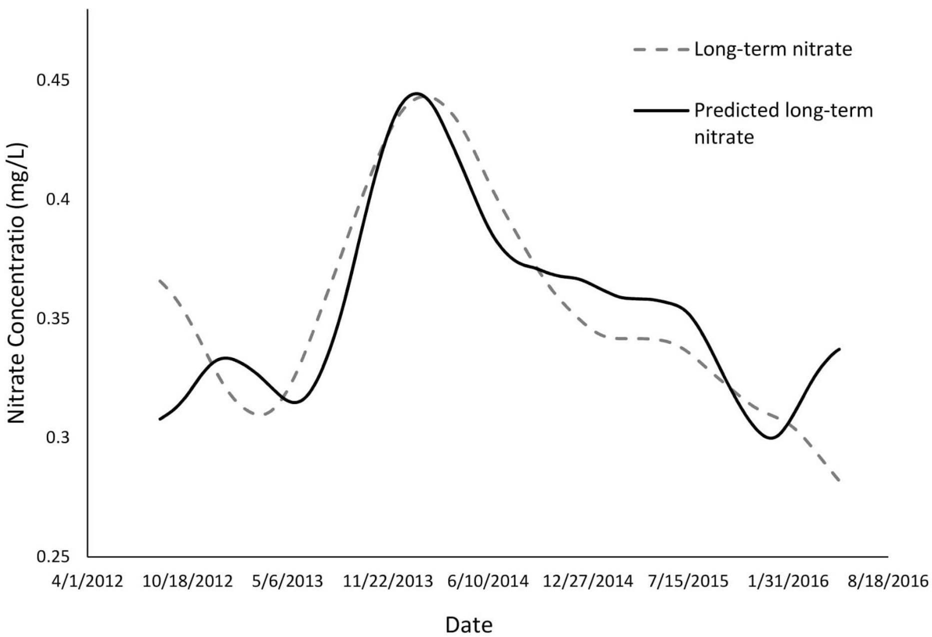

Figure 3: The raw nitrate concentration data overlain with the predicted long-term nitrate concentration at Hog Island Channel from April of 2012 to August of 2016. With a long-term correlation value of 0.81 the predicted long-term trend line follows closely to that of the raw data.

4. Discussion

It is interesting that in the other two monitoring stations located further east across Long Island the periodicity of the nitrate concentration increases. Due to the various time ranges of USGS nitrate concentration records from station to station across Long Island, we have only used data for the past seven years to maintain a common time span prior to analysis. Data limitation to 7 years does not allow exploration of the large nitrate cycles, such as the 2400-days period which cannot be explored until more data are obtained. This limitation has an impact in the time series regression model and subsequent R^2 values. The prediction model performs well in Hog Island channel station (Figure 3). The higher R^2 value obtained at Hog Island Channel is facilitated by the shorter nitrate periodicity. As the periodicity increases eastwards from Hog Island at each location the linear regression model and associated explanation is less informative (Table 1).

Peconic river and Orient Harbor stations in comparison to Hog Island channel are located closer to the groundwater divide (Figure 1) and the associated watershed areas to accumulate nitrate are smaller. We think that near the groundwater divide the accumulation of nitrate may be less, and therefore nitrate may require longer time for a full period. Periodicities are also larger as distance from the station to the groundwater divide is shorter. Shorter distance from the divide yields a smaller land coverage which allows a smaller nitrate accumulation.

Due to the location of the groundwater divide, Orient Harbor is located at a strategic location where water runoff disperses to the north, while Hog Island Channel region sits furthest south of the groundwater divide and can accumulate a nitrate concentration at a shorter time due to a larger exposed land area. Peconic River is located near groundwater divide which splits into two separate divides running through the north and south fork of Long Island. A lower correlation value at this station can be attributed to the nitrate being pulled away north and south of the divide. Peconic River sits at a higher potentiometric value than the other two locations, this may cause nitrate to not have adequate time to build up significant nitrate levels due to groundwater intensive moving while in low potentiometric values higher concentrations of nitrate may encouraged.

5. Conclusion

Based on a linear regression and spectral analysis performed, a high correlation value of 81% obtained at Hog Island Channel in western Long Island between the nitrate concentration and other atmospheric and hydrogeological variables. As the linear regression model was applied in the other two stations located at the eastern Long Island there was an observable drop in the correlation between nitrate concentration and variables such as precipitation, pH, salinity, water temperature, tidal height and the depth to groundwater. This decrease in correlation can be attributed to a decrease in surface area in eastern Long Island along with each stations location relative to the groundwater divide running west to east across Long Island. We think, that due to the reduction in watershed exposed surface area from western to eastern Long Island this may have had a noticeable effect on the periodicity of the nitrate variable at each USGS monitoring water station.

6. References

Bleifuss, P.S., Hanson, G.N., Schoonen, M., 1997. A geochemical study of the effects of land use on nitrate contamination in the Long Island aquifer system. Long Island Geologists, SUNY Stony Brook University.

Bleifuss, P. S., 1998. “Tracing Sources of Nitrate in the Long Island Aquifer System. Thesis. State University of New York at Stony Brook.

Chan, K.S., Ripley, B., 2018, TSA: Time Series Analysis. R package version 1.2. https://CRAN.R-project.org/package=TSA

Flipse, W. J., et al., 1984, “Sources of Nitrate in Ground Water in a Sewered Housing Development, Central Long Island, New York.” *Ground Water*, vol. 22, no. 4, pp.

418–426., doi:10.1111/j.1745-6584.1984.tb01412.x.\
Kovač, Zoran, et al. "Estimation of Nitrate Trends in the Groundwater of the Zagreb Aquifer." *Geosciences*, vol. 8, no. 5, 2018, p. 159., doi:10.3390/geosciences8050159.

Munster, Hanson, G.,N., Bokuniewicz, H., A geochemical study to determine the sources of nitrate in the groundwater of Suffolk County, Long Island, NY

National Centers for Environmental Information, and Ncei. "Local Climatological Data Station Details." *Local Climatological Data Station Details: JFK INTERNATIONAL AIRPORT, NY US, WBAN:94789 | Climate Data Online (CDO) | National Climatic Data Center (NCDC)*,Feb.2019,www.ncdc.noaa.gov/cdo-web/datasets/LCD/stations/WBAN:94789/detail.

NYSDEP, 2014. Hog Island channel at Island Park, NY. Monitoring report per agreements: AM07147 & AM08775. New York State Department of Environment al Conservation. Division of Water, Bureau of Water Resource Management. p.26.

Tamborski, J., Bokuniewicz, H., Cochran, J.K., 2017. Submarine groundwater discharge driven nitrogen fluxes to Long Island Sound, NY: Terrestrial vs. marine sources, Geochim Cosmochim Ac, Vol. 218, p. 40-57. DOI: 10.1016/j.gca.2017.09.003

Yang, W., and Zurbenko, I., 2010, Kolmogorov–Zurbenko filters. Wiley Interdisciplinary Reviews: Computational Statistics 2, no. 3: 340-351.

"U.S. Geological Survey." *ScienceBase*, 2018, www.sciencebase.gov/catalog/item/5995caeae4b0fe2b9fea76a6.

USGS 01311143 Hog Island Channel at Island Park NY, accessed Feb. 2019, waterdata.usgs.gov/nwis/inventory/?site_no=01311143.

USGS 01304200 Orient Harbor at Orient NY, accessed Feb. 2019, waterdata.usgs.gov/nwis/inventory/?site_no=01304200

USGS 01304562 Peconic River at County Highway 105 at Riverhead NY, accessed Feb. 2019, waterdata.usgs.gov/nwis/inventory/?site_no=01304562

USGS Current Conditions for USGS 403517073430702 N 6702. 1, Feb. 2019, waterdata.usgs.gov/nwis/uv?site_no=403517073430702.

USGS Current Conditions for USGS 405149072532201 S 5517. 1, Feb. 2019, waterdata.usgs.gov/nwis/uv?site_no= 405149072532201

USGS Current Conditions for USGS 405830072331502 S 6558. 2, Feb. 2019 waterdata.usgs.gov/nwis/uv?site_no= 405830072331502