NOISE REMOVAL AND FREQUENCY DETECTION OF GROUNDWATER LEVELS USING R: AN APPROACH TO IDENTIFY UNDERGROUND EROSION

[1]TSAKIRI K.G., [2]MARSELLOS A.E.

[1]Dept. Information Systems, Analytics, and Supply Chain Management, Rider University, Lawrenceville, NJ
[2]Dept. Geology, Environment, and Sustainability, Hofstra University, Hempstead, NY
E-mail: [1]ktsakiri@rider.edu; [2]Antonios.e.marsellos@hofstra.edu

Abstract - This paper presents a novel approach for detecting underground erosion, which can lead to increased landslides and sinkhole events, by analyzing the frequency of groundwater level (GWL) oscillations. We utilized spectral analysis and noise removal techniques in R to identify high-frequency GWL fluctuations as indicators of underground erosion. Our method involves applying a Kolmogorov-Zurbenko (KZ) filter and spectral analysis to strengthen spectral resolution and detect erosion-prone areas. This study emphasizes the importance of filtering before frequency detection, focusing on low frequencies (greater than annual) and high frequencies (less than semi-annual), to enhance the accuracy and effectiveness of geological hazard predictions.

Keywords - Groundwater, Spectral Analysis, Underground Erosion, Noise Removal.

I. INTRODUCTION

Groundwater plays a critical role in shaping the subsurface geology, with fluctuations influencing both natural and anthropogenic processes. Recent studies on Long Island, New York, have revealed a correlation between high-frequency GWL oscillations and occurrences of underground erosion, which can lead to sinkholes and landslides. This paper aims to integrate these findings and apply advanced spectral analysis techniques using R to improve the detection of erosion-prone areas. Previous studies, [1]-[3] have investigated the relationship between GWL fluctuations and geological hazards. Reference [1] and [2] focused on high-frequency GWL fluctuations and their correlation with sinkhole occurrences, while reference [3] examined the relationship between GWL fluctuations and landslides on Long Island's North Shore (New York, U.S.A.). These studies underscore the importance of understanding GWL dynamics to predict and mitigate geological hazards. Groundwater fluctuations occur naturally due to various factors, including seasonal changes, precipitation levels, and human activity [4]. Underground erosion, driven by groundwater movement, can create fractures that lead to sinkholes and landslides [5]. This study leverages advanced statistical techniques to analyze GWL frequency data and improve the detection of erosion-prone areas.

II. METHODOLOGY

A. Data Collection: Groundwater data from 58 USGS monitoring stations across Long Island were collected. The study focused on identifying high-frequency and low-frequency GWL oscillations, using data from previous studies [2]-[3] and additional sources (e.g., USGS National Water Information System).

B. Noise Removal and Spectral Analysis in R

We have applied the R programming language to remove noise and enhance the detection of high-frequency GWL fluctuations.

To address missing data in the groundwater level (GWL) time series, we employed the `na_interpolation` function from the (imputeTS) package [6].

This method ensures continuous data by estimating and filling gaps based on the existing data trends, which is crucial for accurate subsequent analysis. To enhance the detection of high-frequency oscillations noise removal has taken place. We applied the Kolmogorov-Zurbenko (KZ) filter using the `kz` function from the (kza) package [7].

The KZ filter, known for its ability to smooth data while preserving significant fluctuations, was applied with a window size of 27 days and repeated three times (kz(27, 3)). This iterative process effectively reduces noise and highlights meaningful variations in the GWL data.

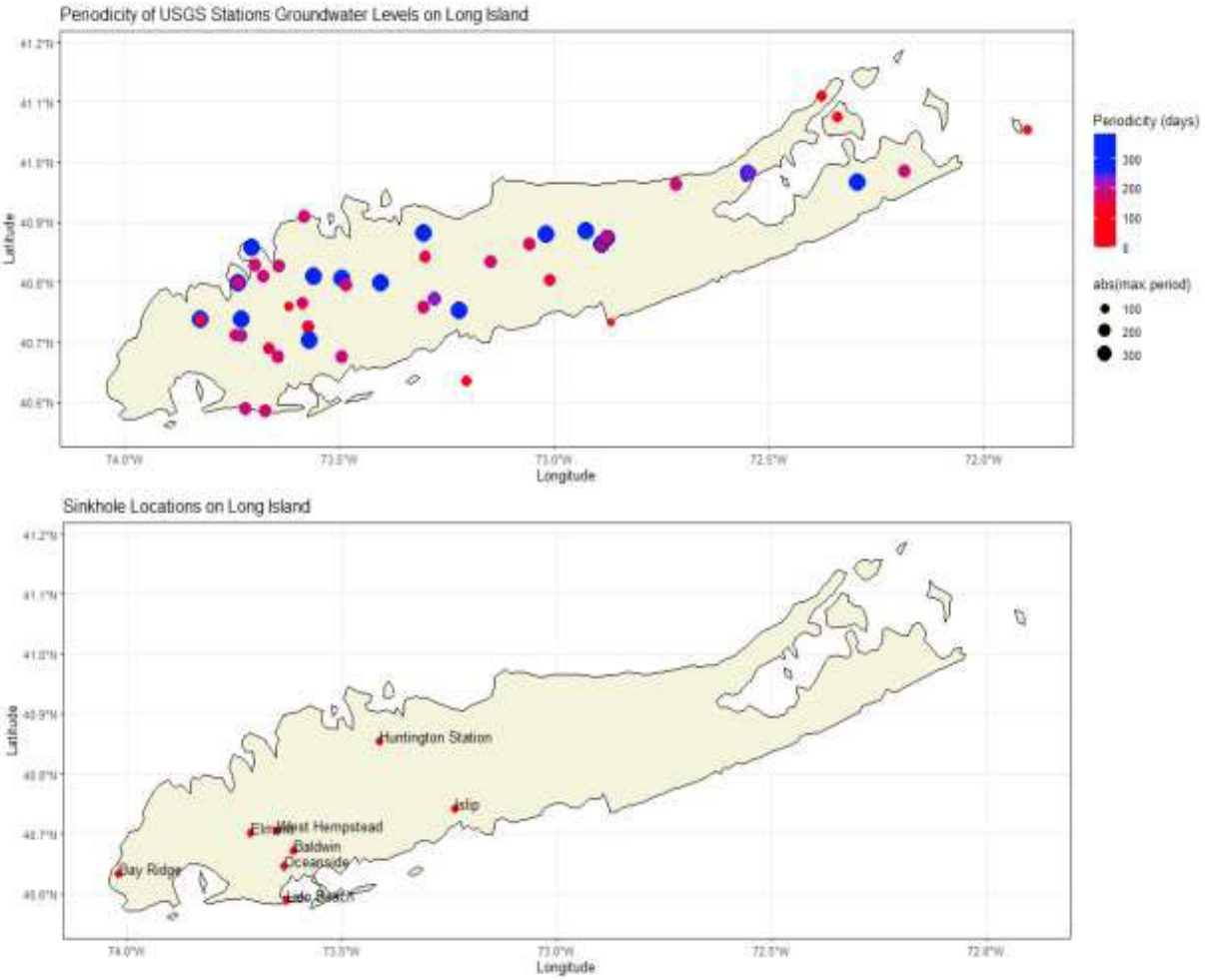

Figure 1: (Top) A GIS map showing the periodicity in days of USGS stations GWL across Long Island, NY. Higher frequency of GWL oscillations indicate lower periodicity (red dots), while higher periodicity represents the lower frequency of GWL oscillations (blue dots). (Bottom) Red dots show recent sinkhole locations across Long Island, New York.

Spectral analysis is required for the identification of dominant frequencies within the GWL data. We utilized the `periodogram` function from the (TSA) package [8]. This function calculates the periodogram, which reveals the frequency components of the time series. To improve spectral resolution and facilitate the detection of short-term variations, we applied the zero-tapping technique by replicating the GWL data and adding zero-padding at both ends. This method minimizes spectral leakage and enhances the accuracy of frequency detection.

To visualize and identify significant oscillations, we normalized the frequency spectrum using the `normalize` function from the (BBmisc) package [9]. This step scales the spectral density, allowing for a clearer interpretation of the frequency components and ensuring that the significant oscillations are easily discernible in the plots.

III. RESULTS

The application of the Kolmogorov-Zurbenko (KZ) filter and spectral analysis techniques to the groundwater level (GWL) data has yielded significant insights into the frequency characteristics and potential indicators of underground erosion. The spectral analysis revealed distinct high-frequency GWL oscillations in several monitoring stations across Long Island. These oscillations are characterized by shorter periods, typically less than semi-annual, indicating the high-frequency fluctuations in groundwater levels (Figure 1; red dots). Stations exhibiting these high-frequency oscillations are particularly noteworthy as they align with areas known for underground erosion and geological instability such as landslides and sinkholes.

The normalized periodograms (Figure 2) for each monitoring station represent the period (in days) along the x-axis. The spectral density has been normalized (for comparison between different stations) and represented along the y-axis. The peaks in these periodograms correspond to the dominant frequencies of the GWL oscillations. Stations of red color (Figure 1) for example, show prominent peaks at periods less than 180 days, suggesting high-frequency oscillations that may contribute to increased underground erosion.

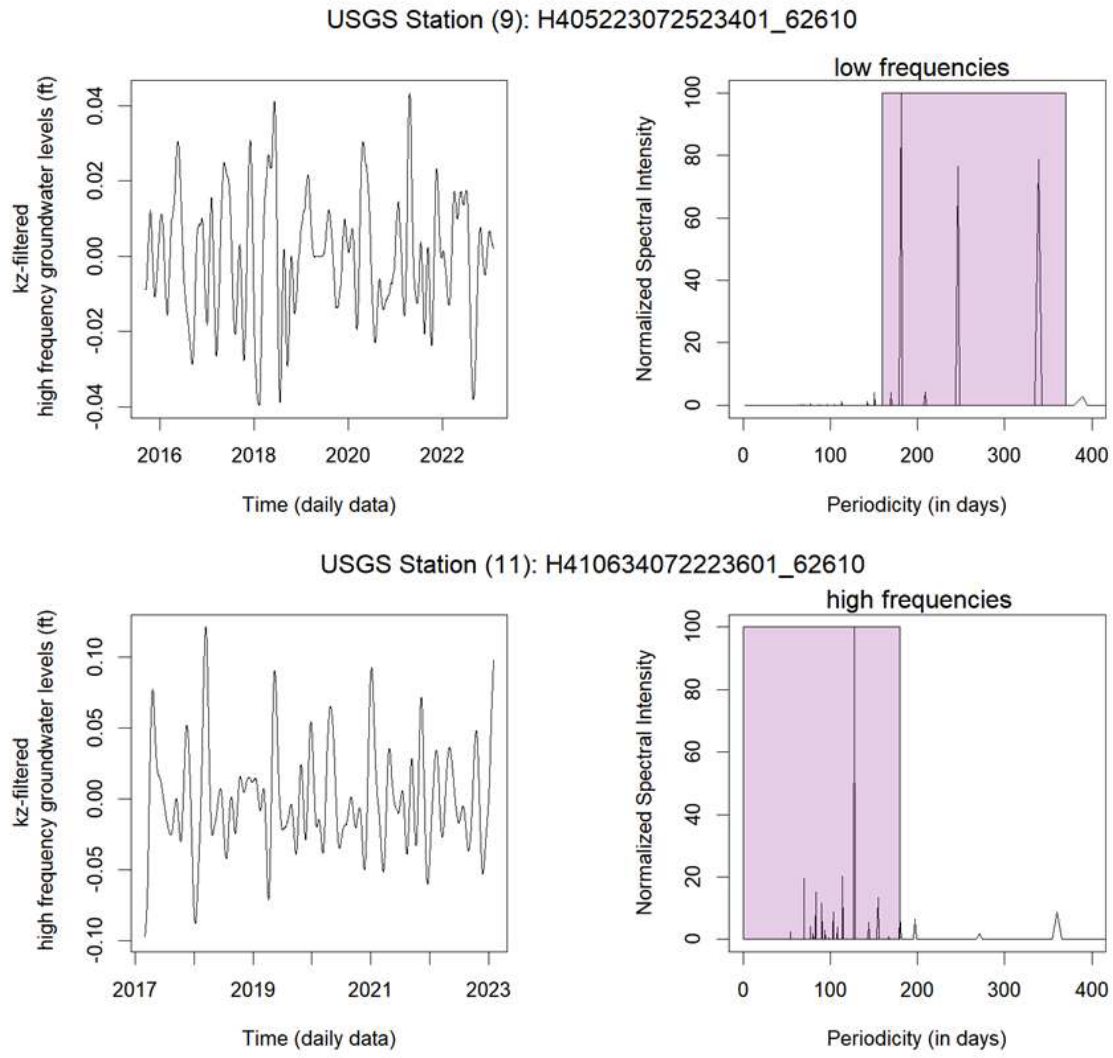

Figure 2: Groundwater fluctuations and frequency levels at two different USGS monitoring stations that show low frequencies (top) and high frequencies (bottom). The main peaks in frequency are highlighted in purple.

The GIS map (Figure 1) integrates the periodogram results with the geographical locations of the monitoring stations. Red dots indicate stations with significant high-frequency oscillations, while blue dots represent stations with lower frequency oscillations. The clustering of red dots in areas such as southern Long Island and the North Shore aligns with regions where underground erosion has been historically reported. The spatial analysis of the monitoring stations (Figure 1) shows where high-frequency GWL oscillations are predominantly located in areas that mostly are previously reported sinkhole and landslide occurrences. This moderate spatial correlation underscores the potential of using high-frequency GWL oscillations as indicators for predicting geological hazards.

IV. DISCUSSION

The application of the KZ filter (kz(27, 3)) has proven effective in removing noise from the GWL data, thereby enhancing the clarity of the spectral analysis. This filtering process is crucial for accurately identifying high-frequency oscillations that might otherwise be obscured by noise. By removing low-frequency components and noise, the KZ filter helps isolate the significant oscillations that are indicative of underground erosion processes.

To improve the spectral resolution, we employed the zero-tapping technique, which involves replicating the GWL signal and adding zero-padding. This method effectively reduces spectral leakage and enhances the detection of short-term variations. The impact of this technique on the periodogram is significant as previously has been reported [10], showing sharper and more distinct peaks compared to the periodogram without zero-tapping. The enhanced spectral resolution allows for more accurate identification of significant frequencies, particularly in the higher frequency range. This improvement is critical for detecting the subtle oscillations that contribute to underground erosion and subsequent geological hazards. The results of our analysis support the

hypothesis that high-frequency GWL oscillations are correlated with increased underground erosion, leading to a higher likelihood of sinkhole and landslide events. The spatial distribution of high-frequency oscillations aligns with known geological hazards, reinforcing the validity of using spectral analysis for hazard prediction.

The detection of frequencies in groundwater level (GWL) oscillations is crucial, especially amidst the increasing impacts of climate change, such as higher rainfall and elevated groundwater activity. With warmer temperature fluctuations, particularly in higher latitude regions, the risks of landslides become more severe. This heightened risk is primarily due to increased underground water mobility during periods that were previously characterized by prolonged winter cold temperatures and minimal water movement. Understanding which areas of Long Island are prone to such conditions is essential. This knowledge can help prevent the loss of life, mitigate damage to roads and structures, and reduce environmental degradation. By accurately detecting and analyzing GWL frequencies, we can better predict and manage the risks associated with these climate-induced changes.

Our findings support the hypothesis that high-frequency GWL oscillations can be indicators of underground erosion. The use of the Kolmogorov-Zurbenko (KZ) filter and signal replication significantly enhances the resolution of spectral analysis, making it easier to detect short-term variations. Without these preprocessing steps, the detection of high-frequency oscillations would be challenging, reducing the accuracy of geological hazard predictions. By improving the spectral resolution, these techniques allow for a more precise identification of areas at risk, thereby enhancing our ability to predict and mitigate geological hazards related to climate change.

Furthermore, this methodology is not limited to Long Island; it can be applied to any other region worldwide that exhibits significant groundwater fluctuations. By implementing these techniques globally, we can better understand and address the risks posed by underground erosion and other related geological hazards in diverse environments. This broader application underscores the versatility and importance of advanced spectral analysis in managing the impacts of climate change on geological stability.

V. CONCLUSION

The application of advanced spectral analysis techniques, including the Kolmogorov-Zurbenko (KZ) filter and zero-tapping, has significantly enhanced our ability to detect high-frequency groundwater level (GWL) oscillations. These oscillations are indicative of underground erosion and correlate with areas prone to sinkholes and landslides. Our findings highlight the importance of filtering and signal replication in improving spectral resolution and detecting short-term variations. This study suggests another methodology for predicting geological hazard events that are increasing in frequency due to climate change.

High-frequency GWL oscillations are indicative of areas prone to underground erosion. The use of filtering and signal replication is crucial for enhancing spectral resolution and detecting short-term variations. Furthermore, continuous monitoring and advanced signal processing are essential for understanding and mitigating the impacts of climate change on geological hazards. By applying these techniques, we can improve our predictive capabilities and implement more effective mitigation strategies to address the growing risks associated with climate change-induced geological hazards. This approach may enhance spectral analysis to detect and manage underground erosion, thereby safeguarding infrastructure and communities against the adverse effects of climate change.

REFERENCES

[1] Rienzo, A., Roscoe, S.L., Mahoney, L., Weinstein, P., Tsakiri, K.G., Petrocheilos, C., Marsellos, A.E. (2019). Time series analysis comparing climatic averages and the water levels of aquifers in Albany, NY and Queens, NY. 26th Conference on the Geology of Long Island and Metropolitan New York. April 13th 2019, p.1-8. https://pbisotopes.ess.sunysb.edu/lig/Conferences/abstracts19/abstracts%202019/Rienzo%20et%20al,%202019%20LIG.pdf

[2] Hope, J. H., Marsellos, A. E., & Tsakiri, K. G. (2024). High-Frequency of Groundwater Level Fluctuations, Underground Erosion, Potential Sinkhole Occurrences Across Long Island, NY. 31st Conference on the Geology of Long Island and Metropolitan New York.

[3] Badger, T., Parag, A., Pelletier, A., DeRocchis, S., Marsellos, A.E., & Tsakiri, K.G. (2024). Investigating Groundwater and Landslides Relationship Using R: A Study on Long Island's North Shore. 31st Conference on the Geology of Long Island and Metropolitan New York.

[4] Tamiru, A., Khuliso, M., Haile, M., & Molla, D. (2018). Understanding the groundwater-level fluctuations for better management of groundwater resource: A case in the Johannesburg region. Groundwater for Sustainable Development, 7, 1-7. https://doi.org/10.1016/j.gsd.2018.02.004.

[5] Matthew, C. D., & Adam, M. M. (2018). Sinkhole formation mechanisms and geostatistical-based prediction analysis in a mantled karst terrain. CATENA, 165, 333-344. https://doi.org/10.1016/j.catena.2018.02.010.

[6] Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time Series Missing Value Imputation in R. The R Journal, 9(1), 207-218. doi:10.32614/RJ-2017-009.

[7] Close, B., Zurbenko, I., & Sun, M. (2020). kza: Kolmogorov-Zurbenko Adaptive Filters. R package version 4.1.0.1. https://CRAN.R-project.org/package=kza.

[8] Chan, K., & Ripley, B. (2022). TSA: Time Series Analysis. R package version 1.3.1. https://CRAN.R-project.org/package=TSA.

[9] Bischl, B., Lang, M., Bossek, J., Horn, D., Richter, J., & Surmann, D. (2022). BBmisc: Miscellaneous Helper Functions for B. Bischl. R package version 1.13. https://CRAN.R-project.org/package=BBmisc.

[10] Tsakiri, K.G., and Marsellos, A.E. (2024). "Signal Detection in High-Noise Time Series Data Using R." Peer-reviewed Proceedings of the International Conference on Time Series and Forecasting, 2024. Gran Canaria, Spain. Web. https://itise.ugr.es/.

★★★