Research Statement

Xiulin Xie
Department of Statistics

My research interests lie broadly in statistical methodologies with real-world applications, including but not limited to biostatistics, public health business, and environmental science. Recently, big data is a popular term that is used to describe data-rich and/or complex applications, and it often takes the form of *data streams* in the sense that observations of certain processes keep being collected sequentially over time. My research has mainly focused on developing a new sequential learning framework for online monitoring of different types of data streams.

Data streams are used in many applications, including disease screening, spatial-temporal disease surveillance and environmental monitoring, and so forth. The crucial question in such applications is how to monitor the longitudinal pattern of the related processes. To sequentially monitor a process, a major statistical tool is statistical process control (SPC) chart, whose major goal is to check whether a process has a significant distributional shift over time. However, traditional SPC charts are developed mainly for monitoring production lines in the manufacturing industry under the assumptions that process observations at different observation times are independent and identically distributed with a parametric (e.g., normal) distribution when the process is stable. However, these assumptions are rarely satisfied in real applications, especially those involving big data, due to data complexity.

To address these challenges, I have developed several flexible and robust SPC methods. For instance, my collaborators and I introduced nonparametric methods based on data categorization and ranking to eliminate the restrictive normality assumption (Xie and Qiu 2023a, 2022b). We also proposed robust transformation-based charts as an alternative approach (Xie and Qiu, 2022a, 2024). To handle nonstationary environments where the in-control distribution evolves over time (e.g., daily PM2.5 levels), we developed new SPC methods for dynamic process monitoring (Xie and Qiu 2023a, 2023b; Xie et al., 2024). Furthermore, to deal with the serial correlation that often arises in longitudinal data, we constructed control charts that do not rely on independence assumptions or parametric time-series models (Xie and Ha 2025, Xie and Qiu 2025a).

My research has also investigated additional challenges in monitoring data streams. In high-dimensional settings, process shifts may be sparse, impacting only a small subset of variables, or may manifest as gradual drifts over time. To address these issues, I have proposed LASSO-based control charting methods for high-dimensional processes, which are capable of identifying the specific variables that have shifted (Xie 2025b). Building on this line of work, Yi et al. (2025) and Liao et al. have further developed methods for effective drift detection.

Currently, I am working on methods for **dynamic disease screening**, with the overarching goal of enabling early detection of health deterioration in individuals. For example, in the SHARe Framingham Heart Study of the National Heart, Lung, and Blood Institute, one important task is to monitor multiple health indicators (e.g., blood pressure, cholesterol level) so that irregular longitudinal patterns can be detected promptly, enabling timely medical intervention to prevent severe cardiovascular outcomes such as stroke. In such applications, none of the assumptions of conventional SPC methods are valid, necessitating the development of robust new methodologies.

Overall, my research aims to create rigorous statistical methods for sequential data monitoring that are both theoretically sound and practically useful in critical domains. In particular, my work on dynamic disease screening and health monitoring contributes directly to the mission of the Institute for Successful Longevity (ISL): promoting healthy aging and timely medical interventions to enhance the well-being of older adults.

References

[1] Liao, R., Yi, F., and **Xie, X.** (2025). Control charts for detecting linear drifts in multivariate process mean and covariance matrix. (Under review in *Applied Stochastic Models in Business and Industry*).

[2] Qiu, P., and **Xie, X.** (2022). Transparent sequential learning for statistical process control of serially correlated data. *Technometrics*, 64, 487–501.

[3] **Xie, X.**, and Qiu, P. (2022a). Robust monitoring of multivariate processes with short-ranged serial data correlation. *Quality and Reliability Engineering International*, 38, 4196–4209.

[4] **Xie, X.**, and Qiu, P. (2022b). Machine learning control charts for monitoring serially correlated data. In *Control Charts and Machine Learning for Anomaly Detection in Manufacturing* (Ed. Kim Phuc Tran), pp. 131–147, Springer.

[5] **Xie, X.**, and Qiu, P. (2023a). Control charts for dynamic process monitoring with an application to air pollution surveillance. *Annals of Applied Statistics*, 17, 47–66.

[6] **Xie, X.**, and Qiu, P. (2023b). Dynamic process monitoring using machine learning based control charts. In *Artificial Intelligence for Smart Manufacturing: Methods, Applications, and Challenges*, pp. 65–82, Springer.

[7] **Xie, X.**, and Qiu, P. (2024). A general framework for robust monitoring of multivariate processes. *Technometrics*, 66, 40–54.

[8] **Xie, X.**, Qian, N., and Qiu, P. (2024). Online monitoring of air quality using PCA-based sequential learning. *Annals of Applied Statistics*, 18, 600–625.

[9] **Xie, X.** (2025a). An adaptive CUSUM chart for robust monitoring of multivariate processes. *Journal of Quality Technology*, In press.

[10] **Xie, X.** (2025b). Adaptive LASSO-based robust multivariate process monitoring and diagnostics. *Quality and Reliability Engineering International*, In press.

[11] **Xie, X.**, and Ha, J. (2025). Online monitoring of irregularly spaced serially correlated univariate processes. (Under review in *Naval Research Logistics*).

[12] Yi, F., **Xie, X.**, and Qiu, P. (2025). A change-point-detection chart for detecting process mean drifts with an application for monitoring the shape of the Salton Sea. In *Frontiers in Statistical Quality Control 14*, Springer.