Beyond Mann-Kendall: Three Statistical Approaches for Groundwater Trend Analysis in a Single Well

An Opinion on Moving from Nonparametric Convention to Physics-Informed Decision Support

Steven D. Siciliano[1], PhD and Wayne R. Jones[2], PhD

About the Authors

Steven D. Sicliano, B.Sc., Ph.D.

Dr. Steven Siciliano brings his experience as one of the world's foremost soil scientists to the task of helping clients to efficiently achieve their remediation goals. Dr. Siciliano is passionate about developing and applying enhanced instrumentation for continuous site monitoring and systems that turn that data into actionable decisions for clients. He is a leading authority on remediation of cold region soils and his research in this area is changing how industry accomplishes the necessary work of remediation. Dr. Siciliano has made significant contributions to the progress of environmental and soil science with 11 book chapters and 220 scientific papers which have collectively been cited over 17,000 times. Dr. Siciliano graduated from the University of Saskatchewan (U of S) with a PhD in Toxicology. In addition to CEO & Co-founder of LiORA, he is a Professor at the U of S and NSERC Chair for In Situ Remediation.

Wayne R. Jones, B.Sc., M.Sc., Ph.D., CStat

Dr. Wayne Jones is a Principal Data Scientist in the Digital Technology Office at Shell plc. He has a BSc (Hons) degree in Mathematics from the Bangor University of Wales, a MSc in 'Mathematical Modelling for Industry' from the University of Loughborough and a PhD in Ecological Modelling from the University of Strathclyde. Wayne is a visiting Research Associate at the University of Glasgow and has been a Chartered Statistician since 2008. Over his 19-year career at Shell, he has worked across a diverse range of domains, including systematic trading strategies, fuels science, renewables production forecasting, and the analysis of environmental monitoring data. He has been developing GWSDAT for more than 15 years and is passionate about applying statistics in practice to deliver real-world impact, ensuring that rigorous methods translate into insights that can be confidently used by practitioners and regulators alike.

---

[1] Professor, Department of Soil Science, University of Saskatchewan, Saskatoon, SK, S7N 5A8, Canada

[2] Principal Data Scientist, Shell Research Ltd. Shell Centre, London SE1 7NA, United Kingdom

Introduction

The question of whether contaminant concentrations are increasing, stable, or declining forms the foundation of remediation decision making. Yet our collective reliance on a single statistical test, the Mann-Kendall method, reflects more convention than scientific optimization. This opinion piece presents three methodological approaches for evaluating within-well concentration trends: the conventional Mann-Kendall test, the penalized spline approach employed by GWSDAT, and numerical model-based data assimilation. Each approach offers distinct advantages and limitations that practitioners must understand to select appropriate methods for their sites.

The stakes of this methodological choice extend beyond academic interest. The Mann-Kendall test achieves only approximately 40% statistical power to detect real trends when sample sizes reach 50 observations under favourable conditions(*1*). For sites relying on quarterly sampling, accumulating 50 observations requires over 12 years of monitoring. The implications for site closure timelines and remediation costs merit serious consideration of alternative approaches.

The Mann-Kendall Test: Simplicity at a Cost

What is Mann-Kendall and why the Industry Chose This Method

The Mann-Kendall test became the regulatory standard for groundwater trend analysis because it requires no assumptions about the underlying data distribution(*2*). Groundwater concentration data rarely follow normal distributions, exhibiting instead skewed distributions with occasional extreme values. The nonparametric nature of the Mann-Kendall test makes it robust to these characteristics. The Interstate Technology and Regulatory Council and the U.S. Environmental Protection Agency both recommend this approach in their guidance documents, and this regulatory endorsement propagated the method throughout the environmental consulting industry(*3, 4*).

The Drawbacks with Rank-Based Analysis

The Mann-Kendall test operates by comparing the relative magnitudes of sample data rather than the actual concentration values themselves(*5*). This rank-based approach discards magnitude information that matters for remediation decisions. Knowing that concentrations are declining provides less actionable information than knowing that concentrations are declining at 15% per year and will reach regulatory thresholds in approximately four years. The test answers the question of direction but not the question of rate, and rate determines project timelines.

Be Mindful of how Autocorrelation Inflates Error Rates

Autocorrelation in groundwater data increases the probability of detecting trends when none exist (Type I error) and missing trends when they do exist (Type II error)(*6*). For example, for a groundwater time series with an autocorrelation coefficient of 0.9, typical of many aquifer systems due to slow hydraulic response, the variance of the test statistic becomes 15 times larger than for uncorrelated data. Groundwater systems exhibit precisely the autocorrelation structure that increase the rate of False Positive or False Negative tests by a rank based test like Mann-Kendall.

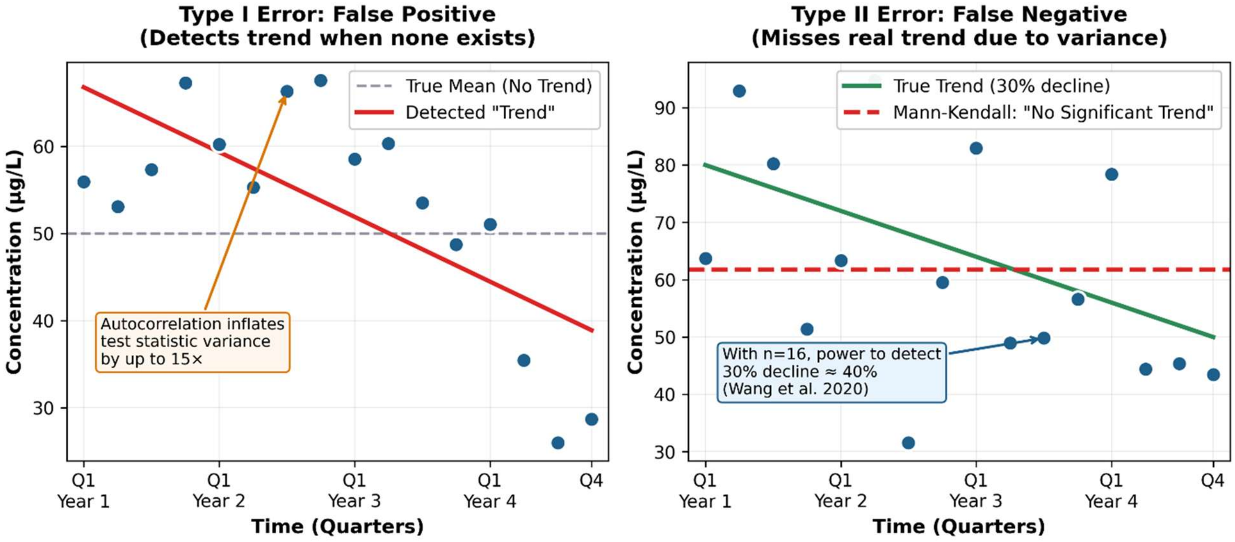

Figure 1. *Mann-Kendall test vulnerability to Type I errors (left panel: detecting false trends in autocorrelated stationary data) and Type II errors (right panel: missing real trends obscured by variance). With quarterly sampling and n=16 observations, power to detect a 30% decline reaches only approximately 40%.*

Penalized Splines: Borrowing Strength Across Time

What is the GWSDAT Approach of Using Penalized Splines

The GroundWater Spatio-Temporal Data Analysis Tool represents a fundamentally different approach to trend analysis(*7*). Rather than testing for monotonic trends at individual wells independently, GWSDAT models variation in groundwater solute concentration as a function of spatial coordinates and time simultaneously using penalized splines following the methodology of Eilers and Marx(*8*). This spatiotemporal framework allows the model to borrow strength across both space and time, improving estimation accuracy when data are sparse.

Key Benefits of Penalized Splines are their Robustness to Data Reduction

Spatiotemporal analysis such as that using penalized spines is robust to sparse data compared to kriging. For example, as data is progressively removed, the performances of spatial-only methods deteriorated while spatiotemporal methods showed a much gentler rate of decline(*9*). This robustness to data loss has direct implications for monitoring network optimization, as spatiotemporal methods can maintain estimation accuracy with smaller sample sizes.

Unlike Mann-Kendall, Penalized Splines Capture Non-Monotonic Behaviour

A critical advantage of the penalized spline approach is that the trend estimate is not constrained to be monotonic(*7*). Real groundwater plumes exhibit complex behaviour: concentrations may initially increase following a release, stabilize as the plume reaches quasi-equilibrium, and then decline as attenuation processes dominate. The Mann-Kendall test cannot represent this behaviour, while penalized splines capture the full trajectory of concentration change over time.

Be Mindful that Penalized Splines Struggle to Estimate Variance

Penalized spline methods do face challenges in variance estimation that affect confidence interval construction(*10*). The smoothing parameter that controls the flexibility of the fitted curve must be selected carefully, as inappropriate choices can lead to either overfitting or underfitting. Evers and colleagues addressed this challenge by developing a Bayesian framework for automatic smoothing parameter selection, allowing fully data-driven methods for fitting flexible models without requiring manual specification of smoothing levels(*10*).

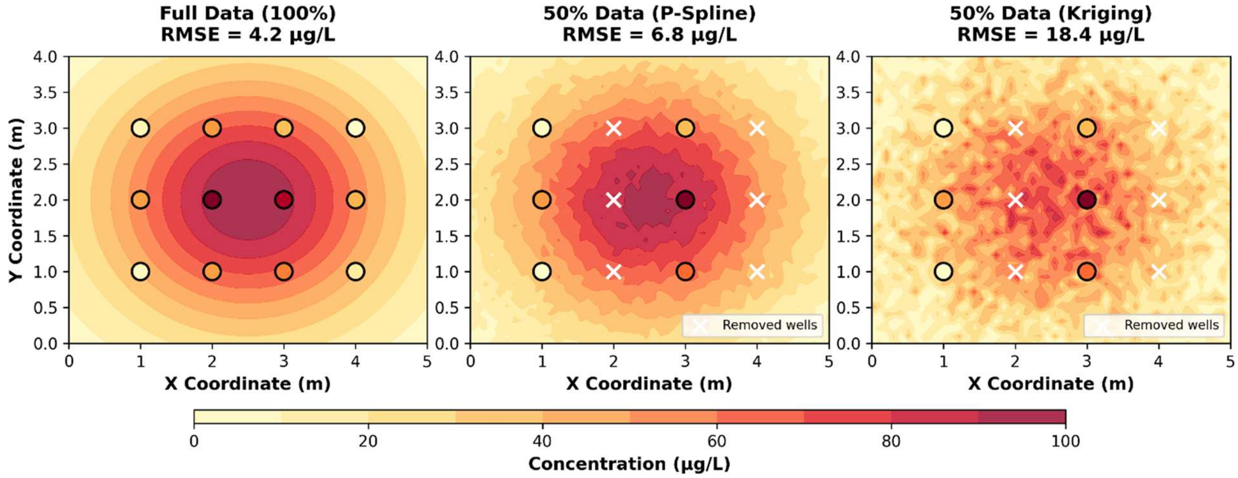

Figure 2. *Comparison of spatiotemporal penalized spline methods versus spatial kriging under data reduction. When 50% of monitoring wells are removed (marked with X), spatiotemporal methods maintain plume delineation accuracy (middle panel) while kriging degrades substantially (right panel). Based on findings from McLean et al. (2019).*

Numerical Models: Physics-Constrained Extrapolation

How Numerical Models use Pattern Recognition to Process Understanding

Both Mann-Kendall testing and penalized spline smoothing operate by identifying patterns in historical observations without incorporating physical understanding of the processes that created those patterns. Numerical groundwater models coupled with data assimilation techniques offer a fundamentally different paradigm(*11*). These approaches use physics-based simulation to infill sparse monitoring data, creating continuous concentration fields that enable application of more powerful parametric statistical methods.

How Numerical Models use Sensors to Improve Outputs

One way to infill sparse data is to use an Ensemble Kalman Filter. For example, using the Ensemble Kalman Filter improved characterization of the hydraulic head field substantially(*12*). The approach updates model states together with parameters through an augmented state vector, enabling continuous updating and prediction at locations where monitoring wells do not exist. Without data assimilation, ensemble standard deviation of hydraulic head exceeded 1.4 metres after two years; with assimilation, uncertainty remained bounded and decreased with each observation.

A Key Benefit is that Numerical Models Enable Forward Simulation

The most powerful capability of numerical model approaches is forward simulation. -Kendall methods are not appropriate to predict future concentrations(*13*). Statistical methods extrapolate from past patterns, and when those patterns change due to new sources, flow modifications, or remediation activities, the extrapolations fail. Numerical models incorporate transport physics (advection, dispersion, reactions) and can simulate how plumes will evolve under different scenarios, enabling site closure timeline predictions that statistical methods cannot provide.

Be Mindful of The Investment Needed for Physics Based Models

Numerical model approaches require greater technical investment than statistical methods. Model development, calibration, and validation demand hydrogeological expertise and computational resources that exceed what most routine monitoring programs support. However, for sites where remediation decisions depend on quantifying trend magnitude and projecting future conditions, the literature supports model-based approaches as superior alternatives to standalone statistical testing(*14*).

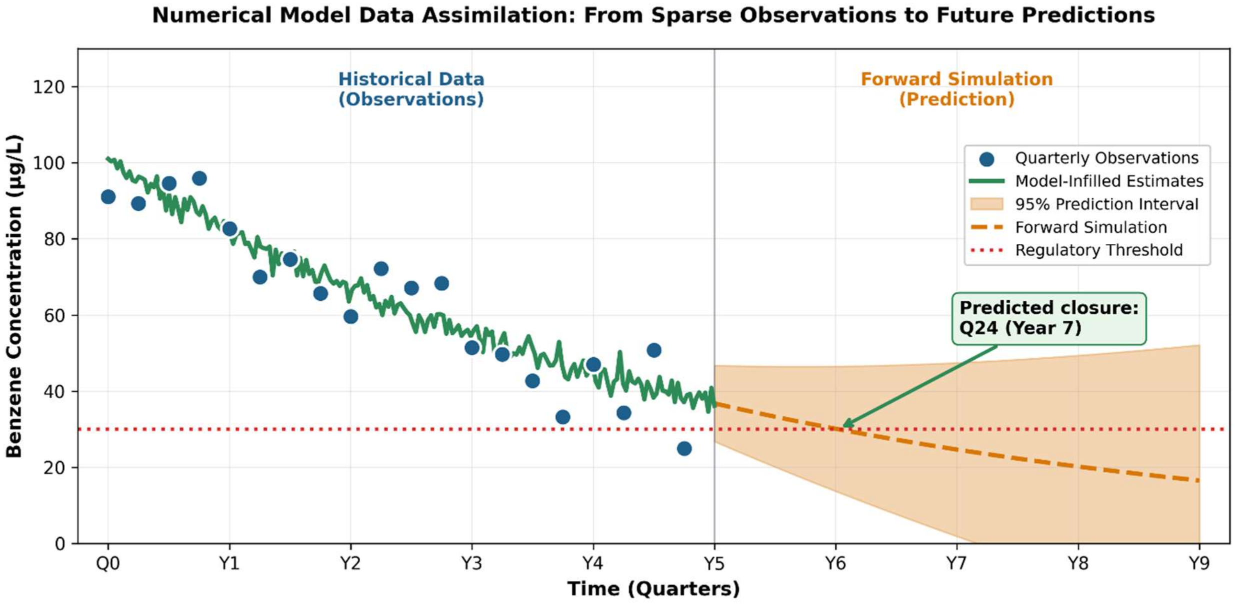

Figure 3. *Numerical model data assimilation converts sparse quarterly observations (blue points) into continuous concentration estimates (green line) and enables forward simulation with quantified uncertainty (orange envelope). Unlike statistical methods, model-based approaches can predict when concentrations will reach regulatory thresholds.*

Comparative Summary

Table 1 summarizes the key characteristics of each approach, providing practitioners with a framework for method selection based on site-specific requirements and constraints.

Characteristic	Mann-Kendall	P-Splines (GWSDAT)	Numerical Models
Statistical Approach	Nonparametric rank-based	Bayesian smoothing	Data assimilation
Minimum Data Required	8+ observations	Variable (fewer needed)	Site-specific calibration
Handles Autocorrelation	Poorly (inflates errors)	Implicitly modelled	Explicitly modelled
Trend Shape	Monotonic only	Non-monotonic capable	Process-based
Magnitude Information	Lost (ranks only)	Preserved	Preserved with physics
Forward Prediction	Not possible	Limited extrapolation	Full simulation capability
Technical Complexity	Low	Moderate	High
Regulatory Acceptance	Widely accepted	Growing acceptance	Case-specific

Conclusions and Recommendations

There is a clear hierarchy of methodological sophistication with corresponding tradeoffs for analysis of groundwater concentrations. Mann-Kendall testing offers simplicity and regulatory acceptance but suffers from limited statistical power with sparse quarterly data and inflated error rates with autocorrelated groundwater data. Penalized spline approaches provide superior spatiotemporal integration and can capture non-monotonic trends while achieving equivalent accuracy with smaller sample sizes. Numerical model-based approaches enable physics-constrained interpolation and forward simulation but require substantially greater technical investment.

For routine regulatory compliance where detecting the presence or absence of monotonic trends suffices, Mann-Kendall testing remains defensible despite its limitations. For sites requiring accurate plume delineation and trend characterization with limited monitoring data, spatiotemporal penalized spline methods like GWSDAT offer superior performance. For sites where remediation decisions depend on predicting future conditions and estimating closure timelines, numerical model-based data assimilation provides capabilities that statistical methods cannot match.

The choice of statistical method shapes what we can know about contaminated sites and therefore what decisions we can defend. Moving beyond convention to embrace the full spectrum of available tools will accelerate site closure timelines and reduce remediation costs across our industry.

Contact

Questions or feedback may be directed to **inquiry@joinliora.com**

References

1. Wang, F., Shao, W., Yu, H., Kan, G., He, X., Zhang, D., Ren, M., and Wang, G. (2020). Re-evaluation of the Power of the Mann-Kendall Test for Detecting Monotonic Trends in Hydrometeorological Time Series. Frontiers in Earth Science, 8, 14.
2. Hirsch, R.M., Slack, J.R., and Smith, R.A. (1982). Techniques of trend analysis for monthly water quality data. Water Resources Research, 18(1), 107-121.
3. Interstate Technology & Regulatory Council (ITRC). (2013). Groundwater Statistics and Monitoring Compliance: Statistical Tools for the Project Life Cycle (GSMC-1). Washington, D.C.
4. U.S. Environmental Protection Agency (EPA). (2009). Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities: Unified Guidance (EPA 530/R-09-007). Office of Resource Conservation and Recovery.
5. Yue, S., Pilon, P., and Cavadias, G. (2002). Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. Journal of Hydrology, 259(1-4), 254-271.
6. Hamed, K.H. and Ramachandra Rao, A. (1998). A modified Mann-Kendall trend test for autocorrelated data. Journal of Hydrology, 204(1-4), 182-196.
7. Jones, W.R., Spence, M.J., Bowman, A.W., Evers, L., and Molinari, D.A. (2014). A software tool for the spatiotemporal analysis and reporting of groundwater monitoring data. Environmental Modelling & Software, 55, 242-249.
8. Eilers, P.H.C. and Marx, B.D. (1996). Flexible smoothing with B-splines and penalties. Statistical Science, 11(2), 89-121.

9. McLean, M.I., Evers, L., Bowman, A.W., Bonte, M., and Jones, W.R. (2019). Statistical modelling of groundwater contamination monitoring data: A comparison of spatial and spatiotemporal methods. Science of the Total Environment, 652, 1339-1346.

10. Evers, L., Molinari, D.A., Bowman, A.W., Jones, W.R., and Spence, M.J. (2015). Efficient and automatic methods for flexible regression on spatiotemporal data, with applications to groundwater monitoring. Environmetrics, 26(6), 431-441.

11. Ross, M.C., Worden, C., and Hughes, R. (2018). The Ensemble Kalman Filter for Groundwater Plume Characterization: A Case Study. Groundwater, 56(6), 914-929.

12. Hendricks Franssen, H.J., and Kinzelbach, W. (2008). Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem. Water Resources Research, 44(9), W09408.

13. McConnell, L., Detloff, B., and Truex, M. (2022). Forecasting Groundwater Contaminant Plume Development Using Statistical and Machine Learning Methods. Groundwater Monitoring & Remediation, 42(3), 34-43.

14. Visser, A., Dubus, I., Broers, H.P., et al. (2009). Comparison of methods for the detection and extrapolation of trends in groundwater quality. Journal of Environmental Monitoring, 11(11), 2030-2043.