MDAG.com Internet Case Study 85

The Profound Enduring Ignorance of the Errors in Applying the Mann-Kendall Statistical Test to Minesite Drainage and Water Quality

by K.A. Morin
© 2025 Kevin A. Morin
www.mdag.com/case_studies.html

1. Introduction

It is interesting to observe professionals propagate major errors when such errors are obvious even to non-professionals with some commonsense. This was discussed to some extent in MDAG Case Study #83 (Morin, 2025a) about the ongoing incorrect usage of acronyms related to metal leaching and acid rock drainage (ML-ARD). This includes “group identity” for defense, as in “everyone else does it this way so I do also”. This was seen further in MDAG Case Study #84 (Morin, 2025b) on common oversimplifications for toxicity and bioavailability that can be expected to be almost always wrong for ML-ARD.

This MDAG Case Study #85 looks closely at the Mann-Kendall statistical test as typically applied to full-scale minesite drainage and water quality at minesites and even in environmentally undisturbed water bodies. Faced with simple as well as with more detailed explanations of why the various types of the Mann-Kendall test are not applicable to full-scale minesite drainage and water quality, and why it can provide wrong results, several professionals have essentially stated and written without justification, “We will continue to use it”. Some even incorrectly state that using a variation like the “seasonal Mann-Kendall” is more valid which is shown below also to be wrong.

In most cases, the application of the various types of the Mann-Kendall test to full-scale minesite drainage and water quality is an indication that the professional is not aware of, or is ignoring, the common errors of this statistical test. Does this reflect a lack of understanding of water-quality statistics and a disregard for findings of the U.S. Environmental Protection Agency?

2. The Mann-Kendal Statistical Analysis

Google AI explains,

> “The Mann-Kendall (MK) test is a non-parametric statistical method used to determine if a time series has a *monotonic (consistently increasing or decreasing) trend* [emphasis added here].... The test compares every data point to all preceding points to see if later values tend to be larger or smaller than earlier values. The MK test is widely used in environmental sciences for analyzing climate and hydrological data to detect significant trends in variables like temperature, rainfall, or contaminant levels over time.”

Thus, Mann-Kendall test indicates whether a consistently increasing or consistently decreasing trend exists, and not whether there is a general but non-consistent trend as is often found with full-scale minesite drainage and water quality (discussed in more detail below).

In agreement, the United States Environmental Protection Agency ("U.S. EPA") has stated that the Mann-Kendall test is not an appropriate statistical test where fluctuations occur with time and should be used for "monotonic" or "no seasonality" trends. For example:

https://www.epa.gov/sites/default/files/2015-10/documents/monitoring_chap4_1997.pdf
https://www.epa.gov/sites/default/files/201605/documents/tech_notes_6_dec2013_trend.pdf

If not already obvious, later sections in this MDAG Case Study show that drainages at minesites and even environmentally undisturbed water bodies typically display variable concentrations across many time intervals such as decades, years, seasons, months, days, hours, etc. There may be major general trends through these time intervals but rarely is there a monotonic consistent trend.

This fact, combined with the statements from the U.S. EPA, provide the first confirmation that the Mann-Kendall test should not be applied to full-scale minesite drainage and water quality, but it is.

3. A Simple Example of a Mann-Kendall Test Applied to Full-Scale Water Quality

The Mann-Kendall statistical test is "categorical", meaning it is a "yes-or-no" or "up-or-own" type of test, independent of how high a value might climb or how low it might decrease.

Here is a simple example. A time-sequential sampling of water from flowing surface water or groundwater, across any arbitrary time interval (hours, seasons, years, etc.), provides the following aqueous concentrations: 1 mg/L, 30 mg/L, 20 mg/L, 100 mg/L, 90 mg/L, 130 mg/L, and then 120 mg/L (Figure 1).

Is there a general trend of increasing concentrations in Figure 1? Obviously, the commonsense answer is "yes", a major increase, perhaps even environmentally alarming. However, the Mann-Kendall test as typically applied to full-scale ML-ARD and water quality says "no change" because there were three increases of concentration balanced by three decreases of concentrations.

To be more exact, the Mann-Kendall test as used for full-scale water quality says there is no continuous monotonic trend in Figure 1, and it cannot reliably assess a general trend like Figure 1 that can eventually lead to toxicity and adverse environmental impacts. Thus, the Mann–Kendall test provides no reliable advance warning that full-scale environmental conditions are worsening in Figure 1 and instead implies there is no worsening. Water-quality professionals typically ignore this detailed but environmentally critical difference.

The usage of the Mann-Kendall test for water quality at minesites can allow erroneously "unexpected" yet actually foreseeable environmental disasters, but it is still used anyway.

4. A More Complex Example of a Mann-Kendall Test Applied to Full-Scale Water Quality

If Sections 2 and 3 are not sufficient to avoid the usage of the Mann-Kendal statistical test to full-scale minesite drainage and water quality, then additional proof is that the Mann-Kendall test can yield all three outcomes (no change, continuous increase, and continuous decrease) at the same time simply depending on when samples were collected! This is ignored by professionals using the various types of the Mann-Kendall test for full-scale minesite drainage and water quality.

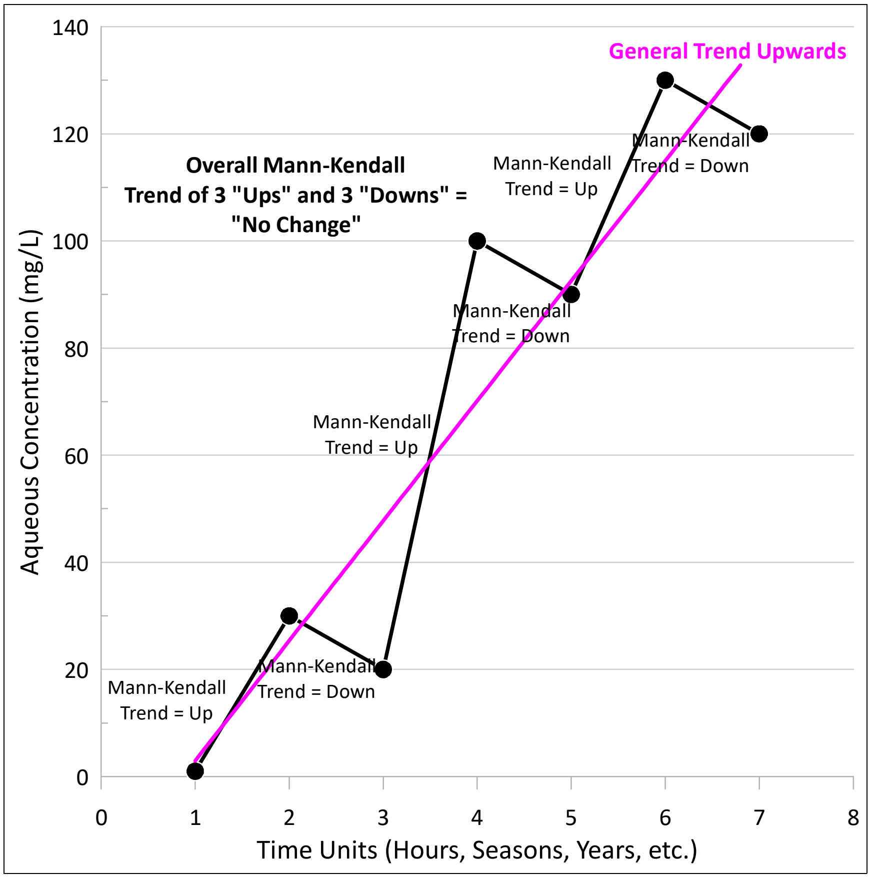

Figure 1. A simplified example of full-scale aqueous concentrations across several arbitrary time units showing a general trend upwards, but the Mann-Kendall statistical test indicating "no change" due to a lack of a continuous monotonic trend.

The proof of this starts with the recognition that full-scale minesite drainage and water quality are highly dynamic, changing significantly across all reasonable time intervals including from hour to hour, from season to season, and from year to year. Aqueous concentrations are not constant between two human-determined water samples, but instead are variable and dynamic.

Strangely, and in contradiction, the typical human notion is that a water sample collected and analyzed once a month (or once a week, etc.) reliably characterizes all water quality across that time interval. That is so wrong, but is a prevalent assumption. The selected sampling frequency is often thought to define actual water quality, when in fact the selected sampling frequency is only a small subset of reality that often does not reliably characterize dynamic water quality (e.g., Morin et al., 1994). The existing paradigm that water quality is defined by the human-selected frequency of monitoring is wrong.

High-frequency and long-duration monitoring of full-scale drainages at minesites and of water quality in environmentally undisturbed water bodies consistently show that aqueous concentrations are significantly variable and dynamic across all measured lengths of time intervals. For decades, this has been a major MDAG research topic (e.g., Morin and Hutt, 1997, 2001, and 2007; Morin, 2016a, 2016b, 2017, 2018, 2019, and 2020), with only a few diagrams out of hundreds shown below as Figures 2 to 4.

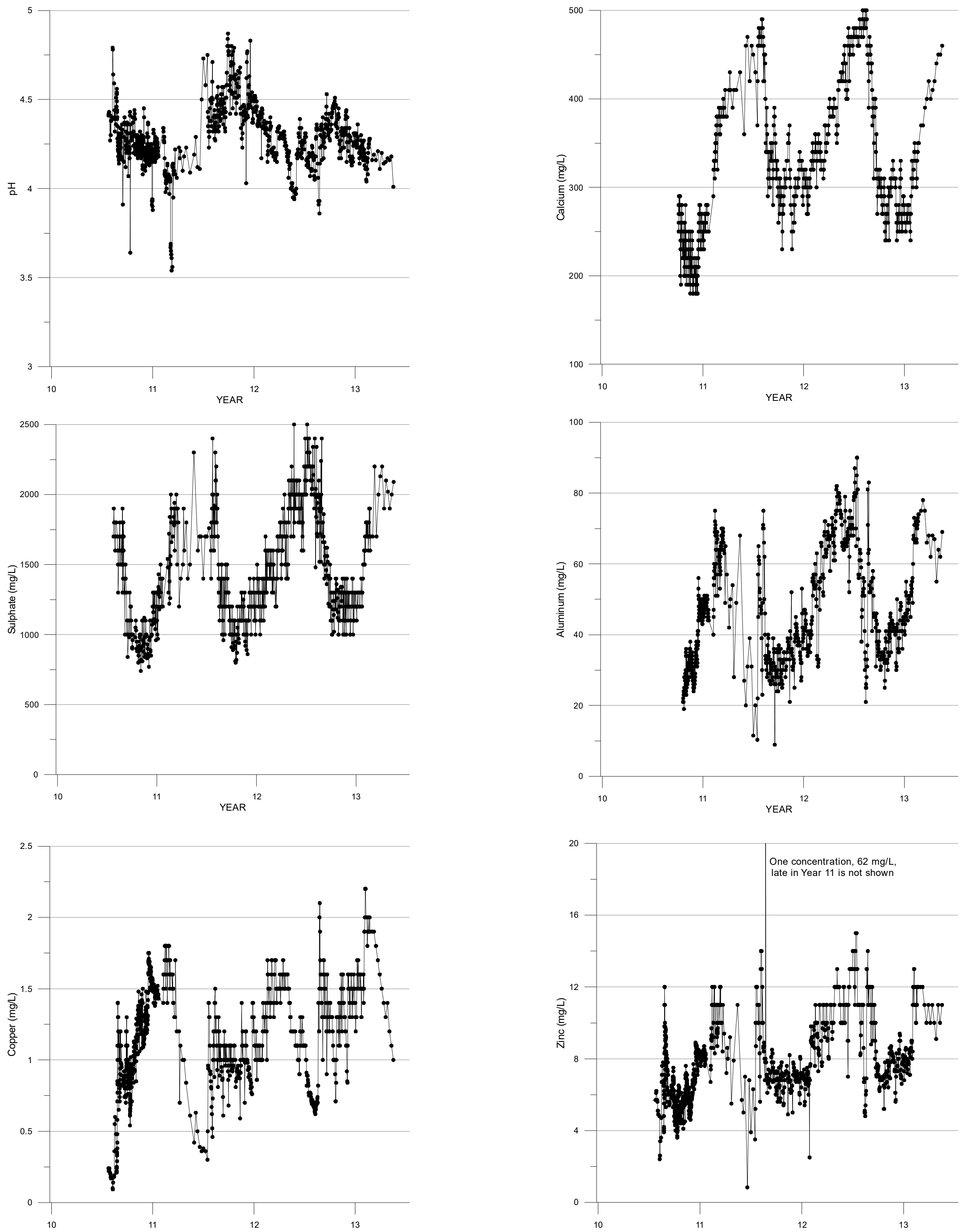

Figure 2. Examples of temporal trends in full-scale acidic minesite-drainage chemistry at one monitoring location, collected as frequently as every four hours; adapted from Morin (2019). Note: y-axis is arithmetic, not logarithmic like Figure 3.

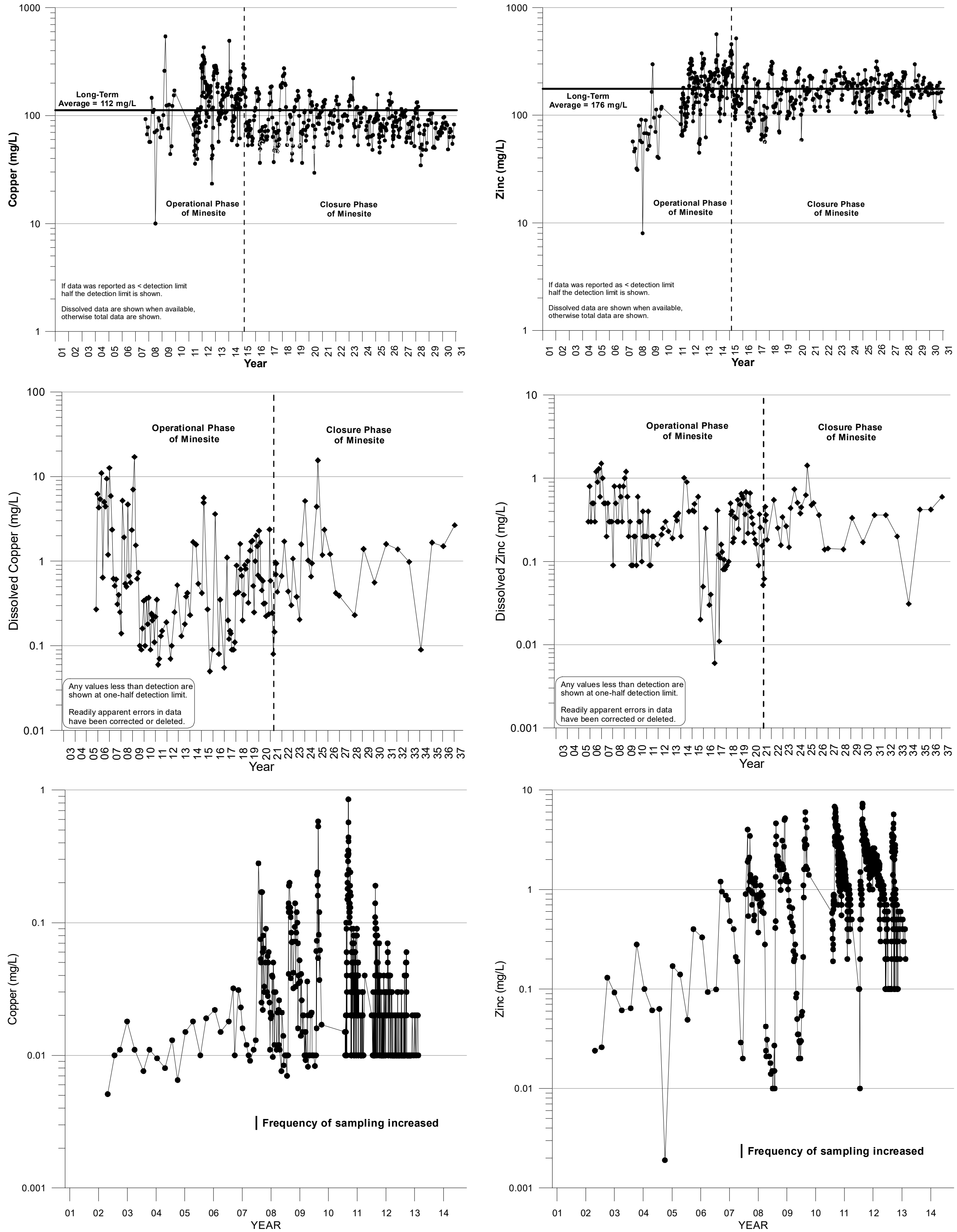

Figure 3. Examples of temporal trends in full-scale minesite-drainage chemistry, spanning at least one decade; adapted from Morin (2019). Upper row (Minesite 1): acidic drainage; Middle (Minesite 2) and Lower (Minesite 3) rows: near-neutral drainage. Left column: copper; Right column: zinc. Note: y-axis is log10 (concentration).

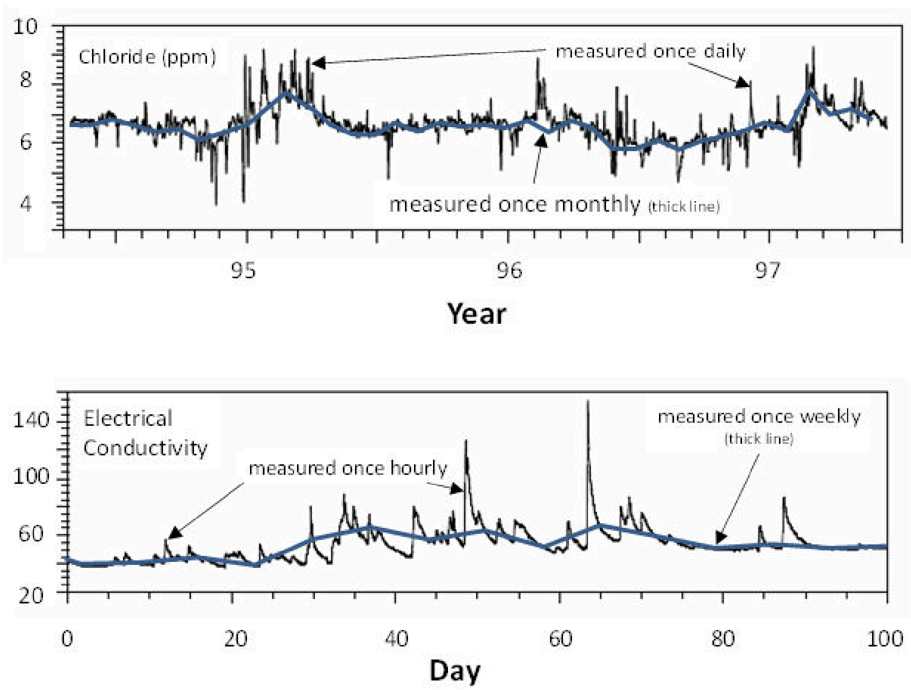

Figure 4. Comparison of high-frequency and low-frequency monitoring in a full-scale non-mining catchment of aqueous chloride (top) and electrical conductivity (bottom) adapted from Kirchner et al., 2004, showing the significant loss of detected dynamic variation and loss of reality in less frequent analyses.

Figures 2 and 3 show that full-scale aqueous concentrations are significantly variable and dynamic across all time intervals. Figure 4 also shows that monthly analyses do not detect the range and variability detected by daily sampling (top of Figure 4), and that weekly analyses do not detect the range and variability detected by hourly sampling (bottom of Figure 4). Further to this, the fractal distributions in these figures (Morin, 2016a, 2016b, 2017, 2018, 2019, and 2020), by definition, also mean that hourly sampling cannot likely detect the range and variability available from more frequent sampling by the minute and by the second. The more frequent the monitoring, the closer to true water quality and its variability that theoretically are never fully measured due to fractal trends.

Thus, it should be obviously unreasonable to expect reliable temporal trends to be reported by the Mann-Kendall statistical test applied to some subset of the data (e.g., monthly monitoring) in Figures 2 to 4.

To make this clearer through major simplification, Figure 5 shows aqueous concentrations varying one cycle a day in a simple "sawtooth" pattern, with no clear long-term increase or decrease. Since one "peak" and one "valley" are repeated every day, then twice-daily sampling could allow the Mann-Kendall test (see Trend 1 in Figure 5) to detect the half-day "up" trend and the half-day "down" trend only if the twice-daily sampling occurred close to the real peaks and valleys; or (see Trend 2) to detect no trend at all if twice-daily sampling occurred around the center of each up and down trend; or (not shown) many other variations. More serious problems and even significant errors arise when the sampling frequency is long than a half day (Figure 6), which is typical of almost all minesite drainages and water-quality monitoring at weekly, monthly, and quarterly intervals.

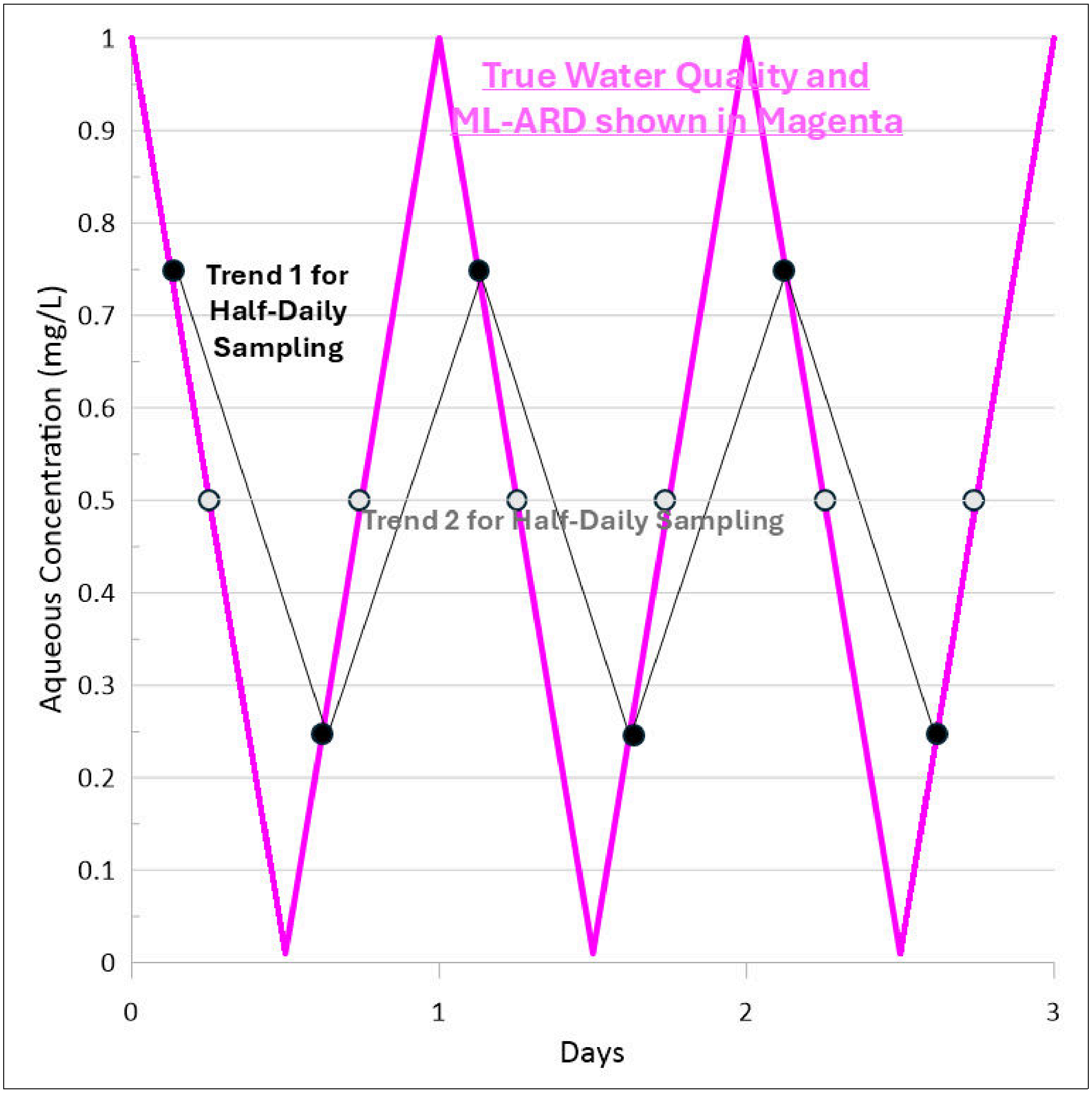

Figure 5. A hypothetical, very simplified water-quality trend with a one-day cycle, showing that exactly when the half-day samples are collected determines which trend is "identified".

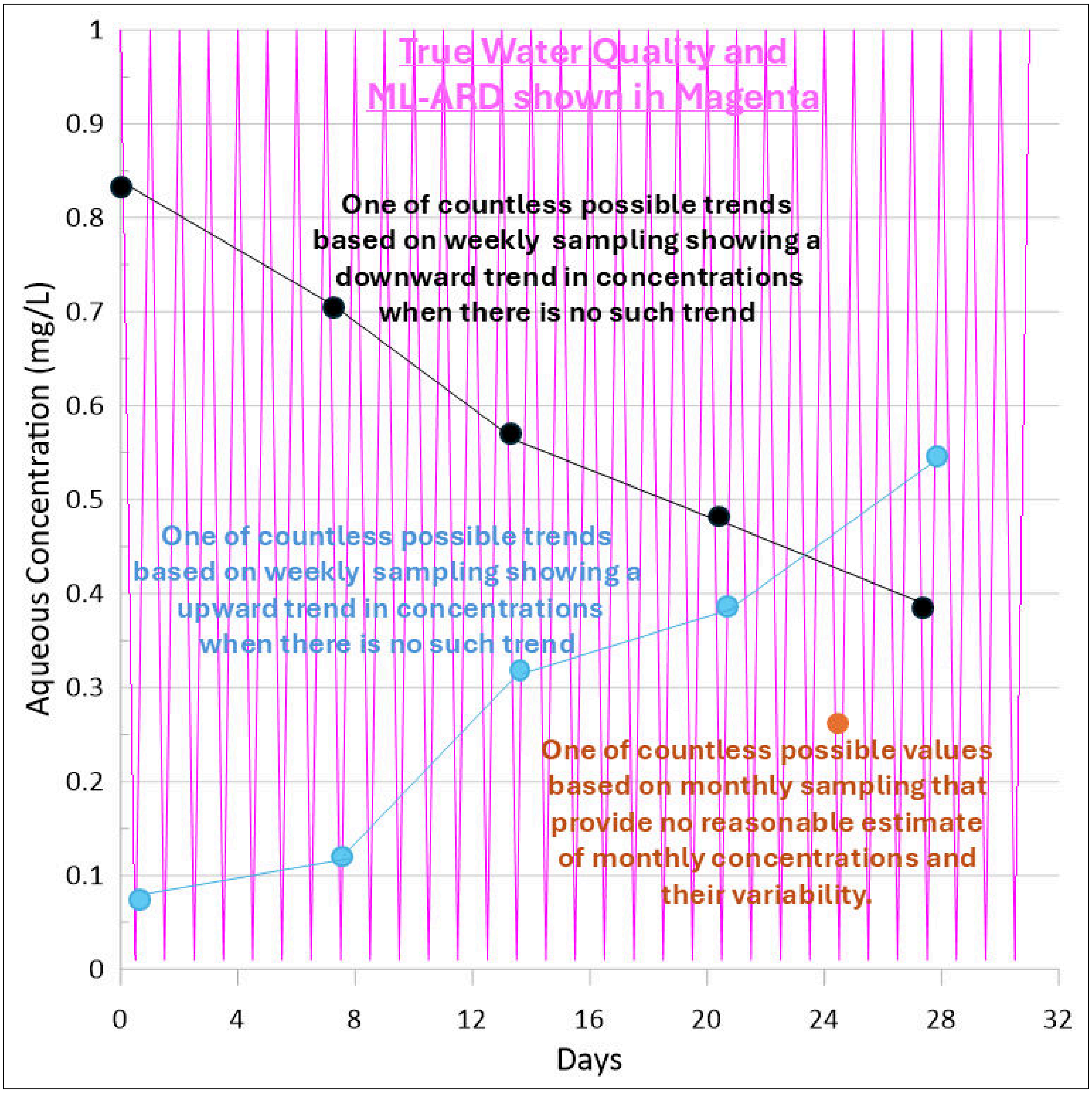

Figure 6. Similar to the one-day cycle of water quality in Figure 5, but showing the countless possible, erroneous trends that can result from weekly and monthly sampling.

Figure 6 shows that sampling less frequently than the significant variability, which is arbitrarily set at every half day in Figure 6, can result in misinterpretations and errors. As explained above, the "significant variability" may have no minimum duration due to fractal trends, so even hourly sampling may be insufficient to avoid these errors.

Based on Figure 6, the Mann-Kendall test could arbitrarily say there is an increasing trend, or a decreasing trend, or no trend, depending on the sampling frequency and on the day and time that sampling starts. The probability of these false Mann-Kendall results significantly increases, from "very likely" in very simplified Figure 6, to "almost certainly" for the full-scale reality in Figures 2 to 4.

These arbitrary results and common erroneous findings from the Mann-Kendall test are dependent on the real variations (e.g., Figures 2 to 4), the sampling frequency, and the start date/time of sampling. These are in addition to the problems and errors discussed above in Sections 2 and 3 above. This includes the U.S. EPA stating that the Mann-Kendall test is not an appropriate statistical test where fluctuations occur with time and should be used for "monotonic" or "no seasonality" trends. This includes incorrect statements that a variation of the test, like "seasonal Mann-Kendall", is valid for full-scale water quality and minesite drainage.

Faced with these simple as well as more detailed explanations of why the various types of the Mann-Kendall test are not applicable to full-scale minesite drainage and water quality, and can provide wrong results, several professionals have essentially stated and written without justification, "We will continue to use it". In most cases, this is an indication that the professional is not aware of, or is ignoring, the common errors of this statistical test. Does this reflect a lack of understanding of water-quality statistics and a disregard for findings of the U.S. Environmental Protection Agency?

5. References

Kirchner, J.W., X. Feng, C. Neal, and A.J. Robson. 2004. The fine structure of water-quality dynamics: the (high-frequency) wave of the future. Hydrological Processes, 18, p. 1353-1359. DOI: 10.1002/hyp.5537

Morin, K.A. 2025a. PAG, PAF, NAG (Non-PAG), and NAF - the Acronyms. MDAG Internet Case Study #83, www.mdag.com/case_studies.html

Morin, K.A. 2025b. Oversimplification of Toxicity and Bioavailability in ML-ARD Studies. MDAG Internet Case Study #84, www.mdag.com/case_studies.html

Morin, K.A. 2020. MDAG-com Case Study 63 - Planetary and large-scale processes affecting minesite-drainage flows and chemistries. MDAG Internet Case Study #63, www.mdag.com/case_studies.html

Morin, K.A. 2019. Application of Spectral Analysis and Wavelet Transforms to Full-Scale Dynamic Drainages at Minesites. Springer Nature (SN) Applied Sciences 1:1058, published online August 2019. DOI: 10.1007/s42452-019-1102-3

Morin, K.A. 2018. Wavelet Transforms of Drainage from Highly Reactive Geologic Materials. Morwijk-MDAG Publishing, Surrey, Canada. ISBN 978-0-9952149-3-4. 277 p.

Morin, K.A. 2017. A Case Study of High-Frequency Time Series for Individual Chemical Elements in Waste-Rock Drainage. MDAG Internet Case Study #46, www.mdag.com/case_studies/cs46.html

Morin, K.A. 2016a. Spectral Analysis of Drainage from Highly Reactive Geologic Materials. Morwijk-MDAG Publishing, Surrey, Canada. ISBN 978-0-9952149-1-0. 331 p.

Morin, K.A. 2016b. Fractal 1/f temporal trends in minesite drainage from waste-rock dumps. IN: 14th Experimental Chaos and Complexity Conference, May 16-19, Banff Center, Banff, Canada.

Morin, K.A., and N.M. Hutt. 2007. Scaling and Equilibrium Concentrations in Minesite-Drainage Chemistry. MDAG Internet Case Study #26, www.mdag.com/case_studies/cs26.html

Morin, K.A., and N.M. Hutt. 2001. Environmental Geochemistry of Minesite Drainage: Practical Theory and Case Studies, Digital Edition. PDF Version, MDAG Publishing, Vancouver, Canada. ISBN 0-9682039-1-4

Morin, K.A., and N.M. Hutt. 1997. Environmental Geochemistry of Minesite Drainage: Practical Theory and Case Studies. Original Paper Version with Limited Availability, MDAG Publishing, Vancouver, Canada. ISBN 0-9682039-0-6

Morin, K.A., I.A. Horne, and D. Riehm. 1994. High-frequency geochemical monitoring of toe seepage from mine-rock dumps, BHP Minerals' Island Copper Mine, British Columbia. IN: Proceedings of the Third International Conference on the Abatement of Acidic Drainage, Pittsburgh, Pennsylvania, USA, April 24-29, Volume 1, p.346-354.