

APPLIED PYTHON DATA SCIENCE IN NICKEL EXPLORATION: PART 1 – STATISTICAL VALIDATION OF NICKEL EXPLORATION ASSAY DATA

Achmad Adyatma Ardi, ST

Geologist, Data science enthusiast

adyatmaardi@gmail.com

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Keywords: ..., ..., ...

1. PROJECT OVERVIEW

1.1 Background and Motivation

In nickel laterite exploration, assay data represent one of the most critical foundations for geological modeling, resource estimation, and decision-making throughout the exploration-to-mining workflow. The integrity and comparability of assay data directly affect how accurately grade distributions, geochemical domains, and laterite profiles can be defined. However, in practical exploration projects, assay datasets are often obtained from multiple drilling campaigns, analytical laboratories, and time periods, each potentially introducing variability in measurement precision and accuracy.

Ensuring that such datasets are statistically consistent is therefore a prerequisite before any higher-level geochemical or spatial analyses can be conducted. Without prior validation, merging datasets may lead to systematic bias, erroneous grade interpretation, and unreliable domain boundaries. This issue becomes particularly significant when exploration data are used to model transitional zones between lithological units such as **limonite (LIM)**, **saprolite (SAP)**, **rocky-saprolite (RSAP)**, and **bedrock (BRK)**—each of which exhibits distinct geochemical behavior and variability in nickel (Ni), iron (Fe), magnesium oxide (MgO), and silica (SiO_2) concentrations.

To address these challenges, this study applies a structured **data science–driven validation framework** to assess the statistical relationship between two independent assay datasets collected from the same exploration area but analyzed at different laboratories and time intervals. The motivation is twofold:

1. To ensure **data reproducibility and analytical equivalence** between laboratories, and
2. To establish a **validated baseline dataset** suitable for subsequent multivariate geochemical and spatial modeling (to be continued in Part 2 and Part 3 of this series).

By integrating statistical testing, effect size analysis, and twin-hole comparison into a reproducible Python-based workflow, this work demonstrates how applied data science techniques can enhance the **transparency, traceability, and confidence** in nickel exploration datasets. The approach exemplifies how quantitative validation supports geological interpretation—bridging the gap between geoscience domain knowledge and modern computational analysis.

1.2 Project Objectives

The primary objective of this project is to **evaluate the statistical consistency and analytical reliability** between two independent *nickel laterite assay datasets* obtained from different sampling periods and laboratories within the same exploration area. Ensuring the comparability of these datasets is essential for building a unified geochemical database that supports subsequent stages such as **multivariate geochemical analysis, spatial validation, and resource modeling**.

To achieve this, the study employs a two-tiered analytical framework:

1. **Global Statistical Analysis** — evaluates the overall agreement and distributional characteristics across major geochemical elements (Ni, Fe, MgO, SiO₂). This includes tests of central tendency ratios, distribution normality, homogeneity of variance, and inferential comparison (e.g., *t-test*, *Mann–Whitney U test*), complemented by effect size and bootstrap confidence interval estimation.
2. **Twin-Hole Statistical Validation** — provides a localized assessment of reproducibility using *paired drillholes* (twin holes) situated within close spatial

proximity. The analysis applies paired comparison techniques (*Paired t-test*, *Wilcoxon signed-rank test*) and agreement-based measures (Pearson's r , Spearman's ρ , Concordance Correlation Coefficient) to evaluate consistency at finer scales.

Together, these complementary approaches aim to:

- Quantify the degree of **agreement** and **bias** between datasets,
- Determine whether historical and recent assay data can be **integrated** within the same analytical framework, and
- Establish a **reproducible data science workflow** for future applications in exploration and geochemical validation.

1.3 Workflow Overview

The analytical workflow developed in this project follows a **modular, data-driven structure** designed to ensure transparency, reproducibility, and scalability across different stages of exploration data validation. The workflow integrates statistical and computational procedures within a unified Python-based environment, enabling automated analysis and clear traceability from data preparation to interpretation.

The entire process is organized into **three main phases** (Figure I):

1. **Data Preparation** — involves importing, cleaning, and integrating assay datasets from multiple sources. Boolean attributes such as `is_old` and `is_twin` are generated to facilitate time-based and spatial-based comparisons. Data are stored in a PostgreSQL database to maintain structural consistency and support efficient querying
2. **Statistical Analysis** — the core analytical phase, subdivided into two complementary modules:
 - Global Statistical Analysis, which quantifies the overall differences and similarities between datasets using descriptive, inferential, and effect size metrics.
 - Twin-Hole Analysis, which validates local reproducibility by comparing paired assays from spatially adjacent drillholes, incorporating correlation, bias, and agreement measures.

3. **Visualization and Interpretation** — combines graphical analysis (histograms, KDE, ECDF, boxplots, scatter, and Bland-Altman plots) with statistical outcomes to provide an intuitive understanding of dataset relationships. These visualizations form the basis for interpretive insights into analytical consistency, potential bias, and geochemical reliability

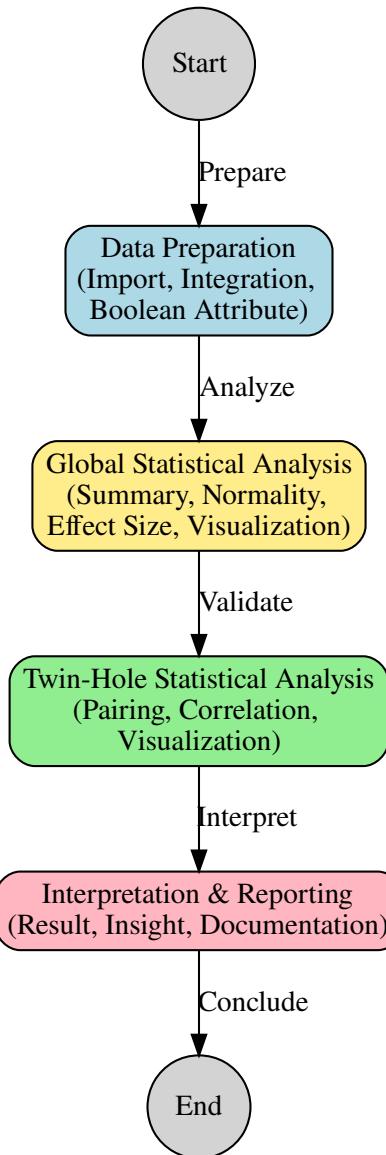


Figure I. Project Workflow

Each module is implemented as a Python function or workflow script within a Miniconda-controlled environment, ensuring version stability and reproducibility. The modular design allows for **independent execution, easy modification, and future**

integration with subsequent project components such as multivariate and spatial analyses (Part 2 and Part 3 of the series).

2. DATA PIPELINE DESIGN

This section outlines the design of the data pipeline for the **Applied Data Science in Nickel Exploration** project, focusing on **modularity, reproducibility, and structured data management** to support statistical validation, twin-hole analysis, and spatial/geostatistical modeling.

2.1 Data Source and Schema (Collar, Geology, Assay)

The pipeline integrates two primary datasets:

- **Collar Data:** Contains drill hole coordinates, Hole_ID, end of holes (EOH), number of increment samples per hole (`inc_sample`), date of collection (`date_collection`), and other relevant metadata
- **Assay data (with merged lithology):** Contains chemical analysis of nickel laterite core samples, including `depth_from`, `depth_to`, `core_length`, `core_recovery`, major elements (Ni, Fe, MgO, SiO₂), and lithology (`lith`) merged from geological logging

Boolean attributes were introduced for temporal distinction (`is_old`) and twin-hole identification (`is_twin`). These datasets were ingested into a **PostgreSQL database**, providing a structured and reproducible foundation for downstream analysis.

2.2 Data Integration Workflow

Data integration follows an **ETL (Extract, Transform, Load)** approach:

1. **Extract:** Import CSV datasets, with format conversions to ensure compatibility with PostgreSQL and Python I/O operations.
2. **Transform:** Cleaning of invalid entries (#DIV/0!, #N/A, blanks) using standardized null values (`pd.NA`). Column types were explicitly defined for schema consistency. Lithology information was merged into the assay table to reduce redundancy.
3. **Load:** Cleaned and integrated datasets were ingested into **PostgreSQL** using Python scripts with **DBeaver** as the SQL interface

Validation steps ensured consistency between collar and assay data, including increment sample counts and metadata integrity.

2.3 Preprocessing and Boolean Attributes

Preprocessing prepares the data for downstream analysis:

- **Data Cleaning:** Handle missing values, and duplicates
- **Normalization / Standardization:** Standardize chemical concentration scales for comparability
- **Boolean Attributes:**
 - `is_old`: 1 for samples collected in 2021 (older dataset), 0 for samples collected in 2023-2024 (newer dataset)
 - `is_twin`: 1 for twin-hole samples, 0 otherwise

Boolean attributes were generated using **Excel formulas and VBA macros**, which also automated the creation of `samp_id` (sample id column) based on increment sample counts. These attributes simplify **filtering, grouping, and twin-hole comparison logic**

2.4 Database Implementation (PostgreSQL + DBeaver + Python)

The PostgreSQL database serves as the central repository for structured exploration data:

- **PostgreSQL:** Stores collar and assay tables (with merged lithology) while maintaining relational integrity
- **DBeaver:** GUI tool for schema visualization, query execution, and data validation
- **Python (pandas, SQLAlchemy, psycopg2):** Connects to PostgreSQL for ETL automation and modular analysis

This implementation enables:

1. Efficient query and extraction for global and twin-hole analyses.
2. Seamless integration between preprocessing scripts and database queries.
3. A **reproducible and scalable workflow** suitable for future exploration campaigns

3. STATISTICAL VALIDATION FRAMEWORK

3.1 Global Statistical Analysis

Global statistical analysis evaluates the overall characteristics of the combined assay datasets, ensuring comparability between temporal datasets (2021 vs. 2023-2024) prior to twin-hole or spatial analyses. (See Figure II)

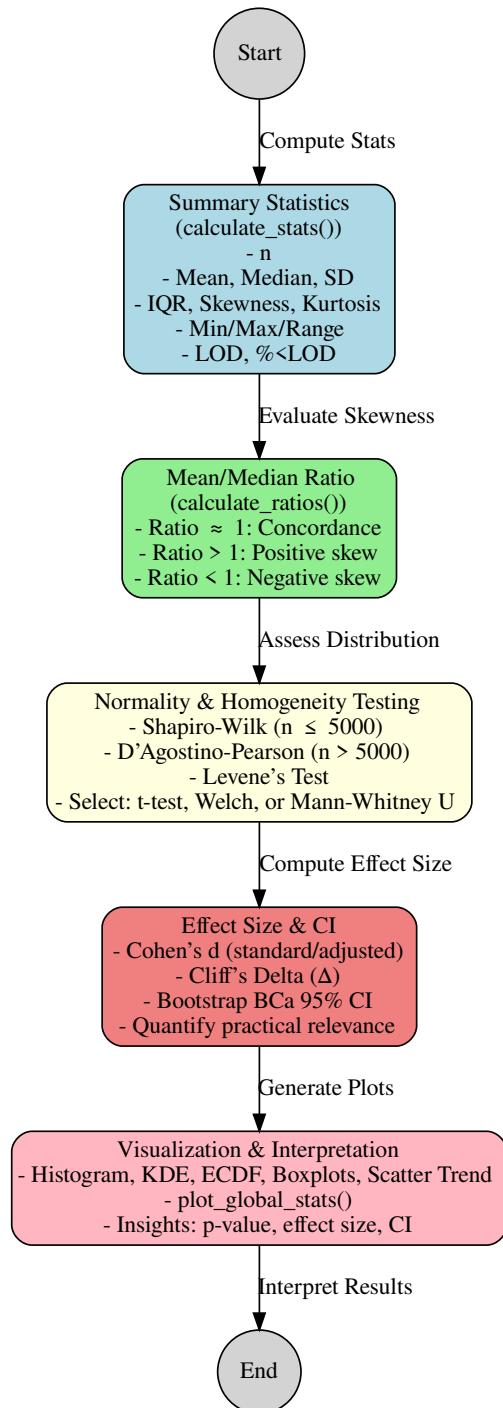


Figure II. Global Statistical Analysis Workflow

3.1.1 Summary Statistics

- **Objective:** Provide a quantitative overview of dataset quality and variability.
- **Methodology:** The `calculate_stats()` function computes:
 - Number of valid observations (n)
 - Central tendency and dispersion (mean, median, standard deviation)
 - Distributional shape (interquartile range, skewness, kurtosis)
 - Minimum, maximum, and range
 - Estimated limit of detection (LOD, 1st percentile)
 - Proportion of values below LOD (%<LOD)

These metrics establish a **statistical baseline** for interpreting normality, correlation, and distributional differences across all elements (Ni, Fe, MgO, SiO₂)

3.1.2 Mean/Median Ratio

- **Objective:** Detect skewness and potential biases in dataset alignment.
- **Methodology:** Calculated via the `calculate_ratios()` function.
- **Interpretation:**
 - Ratio ≈ 1 → Strong agreement between datasets.
 - Ratio > 1 → Positive skew, potential high-value outliers.
 - Ratio < 1 → Negative skew.

This step evaluates **central tendency concordance** and informs data transformation or outlier handling

3.1.3 Normality and Homogeneity Testing

- **Normality:** Shapiro-Wilk test for n ≤ 5000, D'Agostino–Pearson test for larger datasets.
- **Homogeneity:** Levene's test to assess equality of variance between older and newer datasets.
- **Test selection logic** (see Table 1):

Table I. Statistical Test Decision Based on Distribution and Homogeneity

Normality	Homogeneity	Statistical Test
Normal & Homogeneous	✓	Independent t-test
Normal & Heterogeneous	X	Welch's t-test
Non-normal	-	Mann-Whitney U test

- **Automation:** The `stat_test_selector_df()` function dynamically selects the appropriate inferential test based on each element's characteristics, ensuring reproducibility and methodological consistency

3.1.4 Effect Size and Confidence Intervals

- **Objective:** Quantify the magnitude and practical relevance of differences between datasets
- **Metrics:**
 - Cohen's d (standard) for normal and homogeneous variance
 - Cohen's d (adjusted) for normal but heterogeneous variance
 - Cliff's Delta (Δ) for non-normal distributions
- **Interpretation Thresholds** (see Table 2):

Table II. Interpretation of Effect Size Using Cohen's d and Cliff's Δ

Magnitude	Cohen's d	Cliff's Δ	Interpretation
Small	0.2	0.147	Minor difference
Medium	0.5	0.33	Moderate difference
Large	0.8	0.474	Strong difference
Very Large	>1.0	>0.6	Systematic difference

- **Bootstrap CI:** Bias-Corrected and Accelerated (BCa) 95% confidence intervals are computed using `bootstrap_ci()`, providing a robust estimation of uncertainty and effect stability

3.1.5 Visualization and Interpretation

- **Objective:** Communicate statistical insights clearly and intuitively
- **Techniques:**

1. **Histogram** – Compare relative probability distribution with fixed bin intervals.
 2. **Kernel Density Estimation (KDE)** – Smoothed overlay to visualize subtle shifts.
 3. **Empirical Cumulative Distribution Function (ECDF)** – Assess cumulative patterns across datasets.
 4. **Boxplots** – Summarize median, IQR, and outlier spread
 5. **Scatter trend plots** – Evaluate directional alignment and global trend consistency between datasets.
- **Implementation:** All plots are generated via modular Python function (`plot_global_stats()`), ensuring reproducibility, scalability, and publication quality visualization.
 - **Outcome:** Integrates three dimensions of insights – **statistical significance (p-value)**, **practical significance (effect size)**, and **uncertainty (confidence interval)**
 - producing results that are both statistically robust and geochemically interpretable.

3.2 Twin-Hole Statistical Analysis

Twin-hole analysis is designed to **evaluate the reproducibility and agreement of assay measurement** from closely spaced or paired drill holes. By comparing twin-hole samples, we can detect systematic biases, assess precision, and validate laboratory or field consistency.

3.2.1 Paired Data Preparation and Compositing

- **Objective:** Identify twin-hole pairs and prepare composited assay data to enable meaningful statistical comparison between corresponding intervals
- **Methodology:**
 1. **Twin-hole identification:**
 - Identify twin holes based on naming pattern (e.g., E07-003 and E07-003i) using string matching logic (regex expression)

- Assign a boolean flag `is_twin` and shared `base_id` to link each pair
- Record collar coordinates (Easting, Northing, Elevation) for both holes.

2. Compositing increment samples:

- Aggregate multiple increment samples within equivalent depth intervals for each twin-hole
- Calculate **weighted averages** for key chemical assays (Ni, Fe, MgO, SiO₂), **recovery**, and **core length**
- Preserve the dominant lithological layer (`lith`) within each composite using the mode or majority rule.

3. Final paired dataset:

- Produce an **interval-aligned twin-hole dataset** containing composited assay data and spatial references
- This dataset serves as the input for subsequent **paired statistical comparison** (e.g., correlation, regression, Bland-Altman analysis)

- **Outcome:**

A structured, twin-hole-based dataset aligned depth intervals, composited assay values, and spatial references — providing a consistent foundation for statistical validation and reproducibility.

3.2.2 Paired Statistical Testing

- **Objective:** Quantify the consistency and detect significant differences between twin-hole samples
- **Methodology:**
 - Assess normality (Shapiro-Wilk for $n \leq 5000$, D'Agostino-Pearson test for larger datasets)
 - Select test based on normality:
 - **Paired t-test:** Normally distributed variables
 - **Wilcoxon signed-rank test:** Non-normal variables

- **Interpretation:**
 - $p\text{-value} > 0.05 \rightarrow$ No significance difference
 - $p\text{-value} \leq 0.05 \rightarrow$ Potential systematic discrepancy; further review needed

3.2.3 Correlation & Agreement Analysis

- **Objective:** Measure linear association and agreement between twin-hole measurement
- **Methodology:**
 - **Pearson correlation (r):** for normal data
 - **Spearman rank correlation (ρ):** for non-normal or skewed data
 - **Bland-Altman analysis:** evaluate systematic bias and limits of agreement
- *Interpretation Thresholds:*

Table III. Interpretation of Correlation and Agreement Metrics

Metric	Threshold	Interpretation
r / ρ	0.0-0.3	Weak correlation
	0.3-0.5	Moderate correlation
	0.5-0.7	Strong correlation
	0.7-1.0	Very strong correlation
Bland-Altman	± 1 SD	Acceptable agreement
	± 2 SD	Minor deviation
	± 3 SD	Significant deviation

- **Outcome:** High correlation and narrow Bland-Altman limits indicate strong reproducibility between twin-hole samples

3.2.4 Graphical Validation

- **Objective:** Visualize and interpret the degree of agreement between twin-hole assays through multiple complementary statistical plots, providing intuitive support for quantitative validation.
- **Techniques:**
 1. **Scatter Plots with Identity and Regression Lines**

- Plot paired assay results (e.g., Ni₁ vs. Ni₂) to assess alignment along the 1:1 line.
- Include linear regression fit, R², and RMSE
- Deviation from the identity line indicates analytical bias or grade offset.

2. Bland-Altman Plots

- Plot mean values against their differences to visualize agreement limits
- Thresholds applied:
 - ±1 SD → Acceptable agreement
 - ±2 SD → Minor deviation
 - ±3 SD → Significant discrepancy
- Supports interpretation of analytical precision.

3. Boxplots and ECDFs

- Compare overall distributional similarity between twin-hole datasets.
- Highlight skewness, median, shifts, and dispersion differences

4. Paired Difference Histograms / Density Plots

- Visualize the spread and symmetry of assay differences ($\Delta = \text{Hole_A} - \text{Hole_B}$).
- A centered, symmetric distribution around zero indicates strong reproducibility.
- Asymmetric or wide distribution suggest potential bias or measurement inconsistency.

5. Q-Q (Quantile-Quantile) Plots

- Evaluate distributional equivalence of paired datasets
- Deviations from the 45° reference line highlight non-normality or transformation requirements

6. Vertical Depth Plots (Depth-Aligned Profiles)

- Overlay or mirror composited assay profiles for twin holes along a common depth axis (depth increasing downward).
 - X-axis: assay values for old and new holes (overlaid or side-by-side); Y-axis: depth
 - Purpose: verify vertical continuity and identify depth-specific deviations (e.g., interval mismatch, recovery issues).
- **Implementation:** All plots are generated using modular Python functions (`plot_twin_scatter()`, `plot_bland_altman()`, `plot_distribution_comparison()`, `plot_depth_profile()`) built with `matplotlib`, `seaborn`, and `scipy.stats`, ensuring reproducibility and visual consistency.
- **Interpretation:**
 - Tight clustering near the 1:1 line and narrow Bland-Altman spread → Strong statistical reproducibility
 - Broadened variance or directional bias → Indicates compositional or sampling heterogeneity
 - ± 3 SD thresholds → Require geological re-evaluation or re-assay validation.
 - Vertical depth profile disagreements at specific intervals → check compositing rules, depth alignment, and core recovery

4. INSIGHTS AND INTERPRETATION

4.1 Global Statistical Insights

Objective: To evaluate the uniformity, distribution, and comparability between datasets (Old Vs. New)

4.1.1 Central Tendency and Dispersion

The statistical evaluation of the old and new assay datasets highlights clear improvements in the consistency and stability of analytical results across Ni, Fe, MgO, and SiO₂. Central tendency indicators (mean and median) show a **systematic downward shift** in the new dataset—particularly for Ni and Fe—while dispersion indicators (standard deviation and IQR) generally **narrow**.

These findings are reinforced by the **mean–median ratio**, which shows improved alignment and distribution symmetry in the updated dataset, indicating reduced skewness and better resistance to the influence of extreme values. A comparative assessment of the old and new assay datasets demonstrates clear improvements in distributional behavior, supporting enhanced analytical reliability. Key descriptive statistics—including mean, median, standard deviation, and interquartile range—were used to evaluate the shifts in central tendency and dispersion across the four major elements (Ni, Fe, MgO, SiO₂). These statistical measures reflect both the precision of the analytical workflow and the degree to which the data represent true geological variability. The following interpretation summarizes the observed behavior and practical implications for QAQC and resource evaluation (See Table IV and Figure III):

- **Nickel (Ni)**

The new dataset exhibits lower mean and median values with a tighter IQR, indicating reduced grade inflation and less influence from high-grade anomalies. This supports more conservative and defensible grade estimation, minimizing risk in resource classification decisions.

- **Iron (Fe)**

A downward shift in central tendency accompanied by reduced SD suggests greater analytical control and fewer extreme deviations. This stabilizes boundary definition between ore and waste where Fe content is a primary geological discriminator, contributing to more reliable tonnage calculations.

- **Magnesium Oxide (MgO)**

Minor central tendency shifts combined with a slightly wider IQR indicate that MgO variability is primarily geological rather than analytical. Continued monitoring is recommended to ensure representative sampling, particularly within heterogeneous saprolite domains that influence metallurgical performance.

- **Silicon Dioxide (SiO₂)**

The reduction in median and suppression of outliers point to improved data precision and fewer preparation-induced variances. This strengthens geological domaining at transitional boundaries and supports more consistent metallurgical behavior modeling.

Table IV. Summary of Descriptive and Distributional Statistics for Old and New Assay Datasets

Element	Ni	Fe	Mgo	SiO ₂
LOD_old	0.24	4.82	0.42	21.83
LOD_new	0.23	5.30	0.85	6.30
% <LOD_old	0.71	1.00	0.95	0.97
% <LOD_new	1.00	1.03	1.00	1.03
Count <LOD_old	42.00	59.00	56.00	57.00
Count <LOD_new	33.00	34.00	33.00	34.00
n_old		5886.00		
n_new		3305.00		
Mean_old	0.91	16.23	22.78	38.71
Mean_new	0.77	14.26	22.05	33.05
Mean ratio	0.84	0.88	0.97	0.85
Median_old	0.83	10.89	23.68	40.94
Median_new	0.63	9.23	26.30	36.49
Median ratio	0.76	0.85	1.11	0.89
std_old	0.53	12.49	10.92	7.47
std_new	0.50	10.92	11.04	9.52
IQR_old	0.89	13.59	14.37	7.00
IQR_new	0.81	11.14	16.79	6.33
skewness_old	0.59	1.34	-0.07	-0.31
skewness_new	0.84	1.46	-0.81	-1.38
kurtosis_old	-0.49	0.67	0.63	4.05
kurtosis_new	-0.20	0.99	-0.79	1.17
min_old	0.08	0.51	0.02	7.62
max_old	3.36	57.31	88.33	112.78
range_old	3.28	56.80	88.31	105.16
min_new	0.18	3.47	0.64	3.16
max_new	2.68	50.82	36.45	64.10
range_new	2.50	47.34	35.81	60.94

The comparative boxplots visually validate the improvements observed in the summary statistics. For **Ni** and **Fe**, the new assay dataset demonstrates **tighter dispersion, lower median values, and a reduction in extreme high-end outliers**, indicating enhanced reproducibility and reduced analytical bias toward elevated grades. **MgO** and **SiO₂** reflect **similar consolidation of variability**, with narrower interquartile ranges and fewer anomalies, suggesting improved sampling consistency while still capturing genuine geological variability within laterite profiles. (See Figure III)

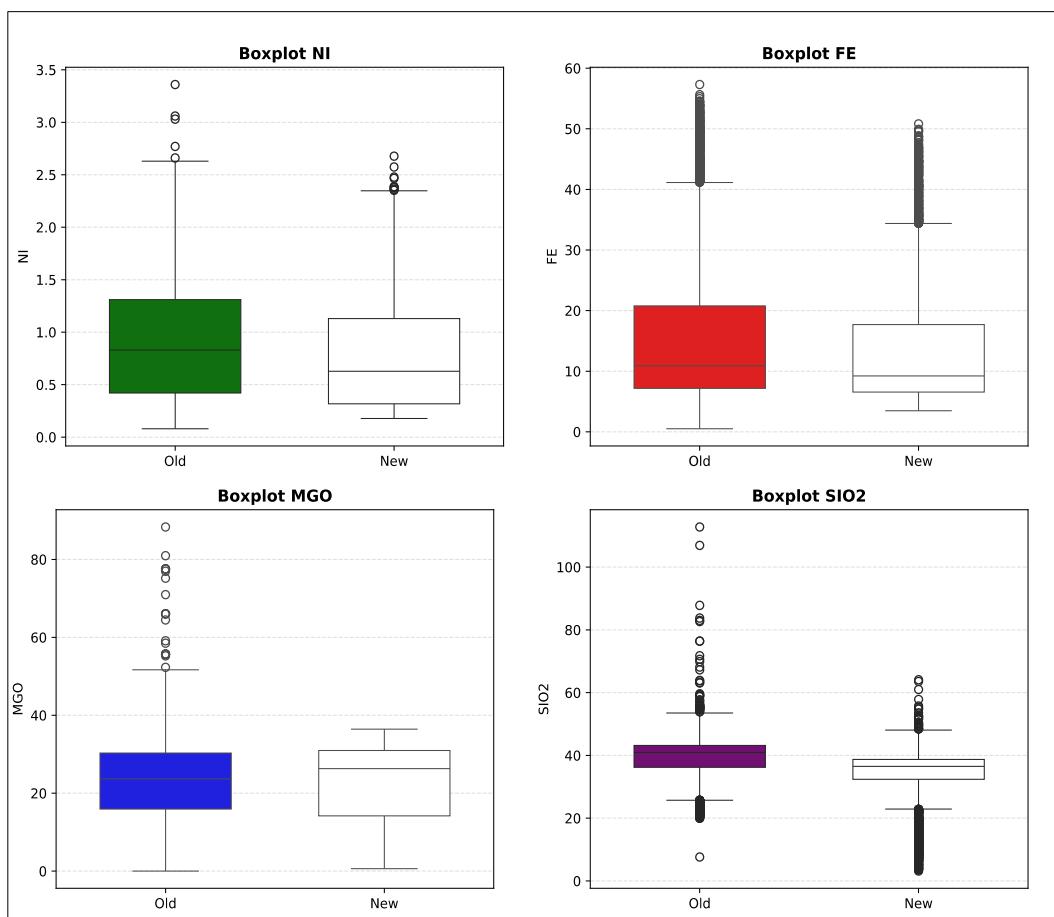


Figure III. Comparative Boxplot of Elemental Distribution (Ni, Fe, MgO, and SiO₂): Old vs. New Assay Dataset

Distribution diagnostics further support these trends. **Skewness and kurtosis values remain close to neutral**, signifying balanced distributions without excessive tail behavior. The mild positive skew in Ni and Fe, along with slight negative skew in MgO

and SiO_2 , aligns with expected element-enrichment patterns in limonite and saprolite zones rather than indicating methodological distortion.

Collectively, these outcomes confirm that the new assay workflow delivers **higher internal consistency, reduced random variability, and more reliable central tendency behavior**. This upgraded data quality strengthens confidence in subsequent **grade estimation, geological domaining, and metallurgical characterization**, making the dataset more robust for decision-making in resource development.

4.1.2 Normality and Variance Homogeneity

The statistical evaluation compares the old and new assay datasets for Ni, Fe, MgO, and SiO_2 to assess analytical consistency, precision improvements, and suitability for inferential comparison. The descriptive summary previously demonstrated that the new dataset exhibits **lower SD, narrower IQR, and reduced coefficient of variation (CV)** across most elements—indicating **improved precision and more constrained dispersion** in the latest analytical workflow. (See Table V)

To validate these observations, formal assumptions of statistical testing were examined using **appropriate normality tests based on sample size**:

- **D'Agostino–Pearson** was applied for the **old dataset ($n > 5,000$)** due to strong sensitivity in detecting deviations from normality.
- **Shapiro–Wilk** was applied to the **new dataset ($n < 5,000$)** as it provides **higher power for smaller samples**.

The results shows **all variables return $p < 0.05$** → Non-normal distribution persists in both datasets. This reflects **natural geochemical skewness** in lateritic systems influenced by mineralogical transitions, weathering profiles, and compositional boundaries. Variance homogeneity was further evaluated using **Levene's test**, which demonstrates:

- **Ni, Fe, and SiO_2 = heterogeneous variances** → Differences in analytical variability still exist
- **MgO = homogeneous variance ($p > 0.05$)** → Stable measurement across assay generations

The distributional assessment shows that all analytes exhibit statistically significant deviations from normality, accompanied by unequal variances between the old and new

datasets (with MgO as the only exception showing variance homogeneity). As a result, the assumptions required for parametric hypothesis testing—specifically normality and homoscedasticity—are not satisfied. To ensure methodological rigor and avoid biased inference, a non-parametric approach is more appropriate. Therefore, the **Mann-Whitney U test** was selected as the recommended inferential method, as it is robust to skewed distributions and variance imbalances, and directly evaluates median differences without relying on Gaussian assumptions. This testing strategy ensures that performance comparisons between assay generations are statistically valid and reliable for downstream analytical workflows.

Table V. Results of Normality (Shapiro-Wilk / D'Agostino-Pearson) and Homogeneity (Levene's) Test for Old and New Assay Datasets

Element	Old Assays Element				New Assays Element			
	Ni	Fe	MgO	SiO ₂	Ni	Fe	MgO	SiO ₂
Normality Method	D'Agostino-Pearson						Shapiro-Wilk	
Stat	402.01	1126.10	62.34	637.99	0.89	0.76	0.84	0.80
p-value	0	0	0	0	0	0	0	0
Normal distribution ?	×	×	×	×	×	×	×	×
n	5886						3305	
Outlier detection method	IQR (1.5 ^x)						IQR (1.5 ^x)	
Outlier Count	5	443	23	663	10	296	0	574
Variance Stat (W)	31.00	39.37	2.28	46.08	31.00	39.37	2.28	46.08
Variance p	0	0	0.13	0	0	0	0.13	0
Homogeneous	×	×	✓	×	×	×	✓	×
Recommended Test	Mann-Whitney U test						Mann-Whitney U test	

In addition to the formal statistical tests, visual distribution diagnostics were performed using fixed-interval histograms and Kernel Density Estimation (KDE) overlays to compare the distribution shapes of the Old and New assay datasets (Figure X). The results clearly show **right-skewed distributions** across all analytes, which aligns with the outcomes of the normality tests. (See Figure IV)

The New dataset consistently exhibits:

1. narrower spread and tighter grade clustering
2. reduced occurrence of extreme high-grade tails

3. more centralized density around the modal values

These visual patterns support the earlier statistical finding that the New dataset achieves **improved analytical precision** with lower dispersion (as reflected by SD, IQR, and CV). Despite the persistent natural geochemical variability typical of lateritic profiles, the improved distribution compactness indicates enhanced data reliability for downstream geostatistical modeling and resource evaluation.

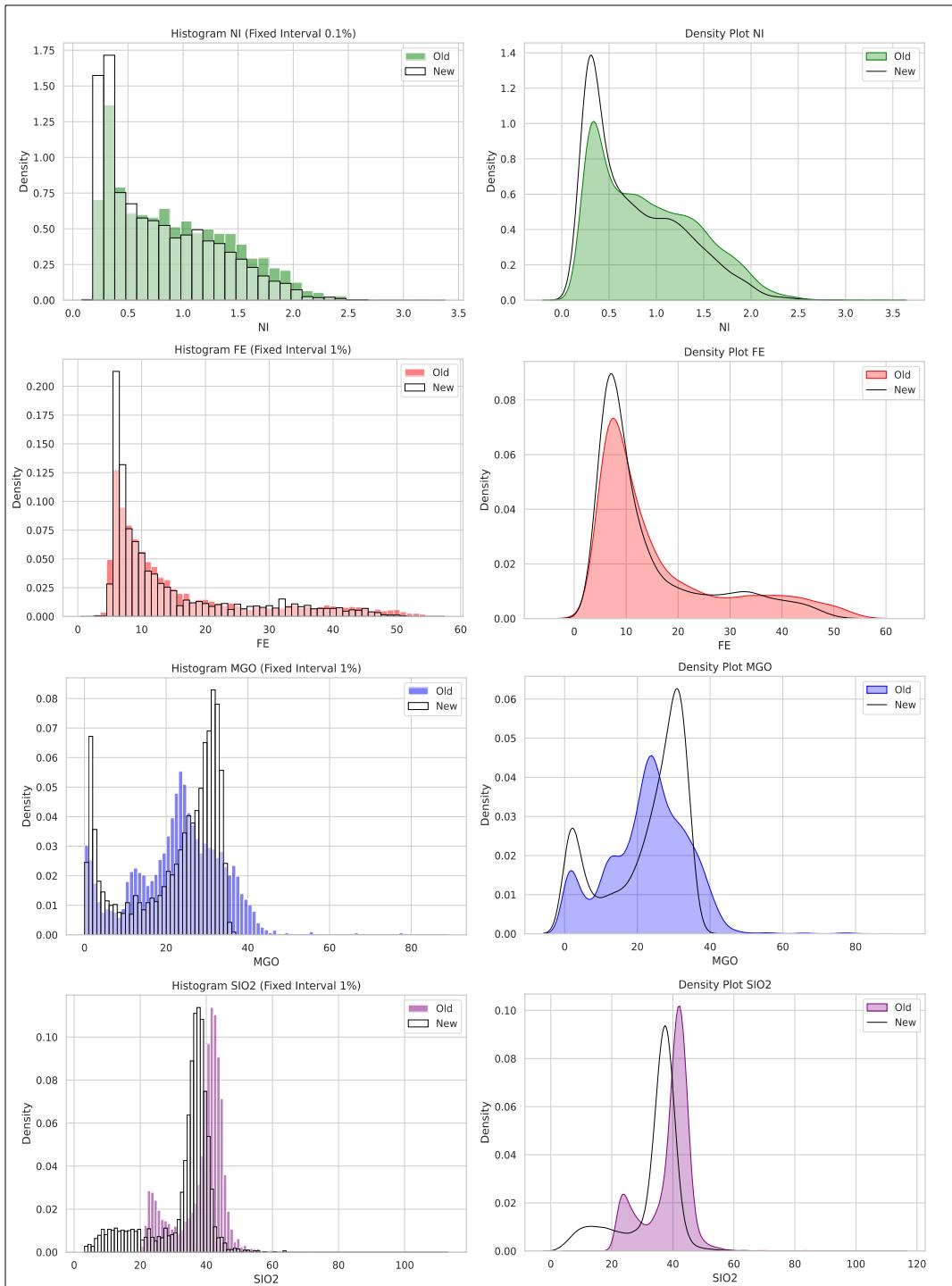


Figure IV. Comparative Boxplot of Elemental Distribution (Ni, Fe, MgO, and SiO₂): Old vs. New Assay Dataset

Overall, the coherence between statistical indicators and visual histogram–KDE diagnostics confirms that variance reduction and precision improvements are achieved in the latest assay workflow, while still capturing the geological characteristics inherent to nickel laterite systems.

4.1.3 Inferential Comparison and Effect Size

Comparative analysis using Cliff's Delta and 95% confidence interval (CI) provides a deeper inference regarding the consistency between Base and Twin drilling assay datasets for Ni, Fe, MgO, and SiO₂. Since the Mann–Whitney U test was applied due to non-normal data behavior, statistical significance was inferred through confidence intervals: CIs that do not cross zero indicate meaningful differences between the two populations. This approach ensures that the evaluation covers both statistical reasoning and operational relevance through effect size magnitude—an essential consideration in resource classification and ore blending decisions. (See Table VI)

1. Nickel (Ni) demonstrates a **statistically significant** difference with a **medium-magnitude** effect size. This suggests that the grade deviation is sufficiently impactful to influence resource evaluation, economic cut-off assessment, and ore quality consistency during mining operations.

Table VI. Effect Size Comparison Between Old and New Assay Datasets for Ni, Fe, MgO and SiO₂

Element	Ni	Fe	MgO	SiO ₂
Metode	Cliff's delta			
Mean Effect	0.144	1.974	0.722	5.654
Effect Size	0.172	0.104	-0.023	0.476
Magnitude	Medium	Small	Small	Large
95% CI	[0.149, 0.196]	[0.080, 0.129]	[-0.046, 0.003]	[0.455, 0.495]
CI Significance	Significant	Significant	Not Significant	Significant

2. Iron (Fe) also shows significant differences, but the **small effect size** indicates **limited operational impact**, meaning discrepancies detected statistically are not expected to materially affect downstream processing or material classification.
3. Magnesium Oxide (MgO) exhibits **no statistically nor practically significant difference**. With an effect size approaching zero and CI overlapping zero, Base and Twin datasets present highly consistent distributions—supporting data reliability for metallurgical performance analysis.
4. Silica (SiO₂) shows the **strongest deviation**, evidenced by a **large effect size** and CI well above zero. This reinforces SiO₂ as a key controlling element in laterite

grade variability, potentially affecting both lithological interpretation and processing efficiency due to its known impact on smelting behaviors.

Overall, the inferential results confirm that only Ni and SiO₂ require closer attention during grade reconciliation due to their practical impacts, while Fe and MgO show sufficient reliability between Base and Twin datasets for further geological modeling.

4.1.4 Agreement and Correlation

The scatter trend comparison demonstrates a **strong linear agreement** between the Old and New datasets across the four major assay elements (Ni, Fe, MgO, and SiO₂). Although down-sampling was applied to optimize visualization, the statistical characteristics and element-specific trends remain fully representative of the full population. (See Figure V)

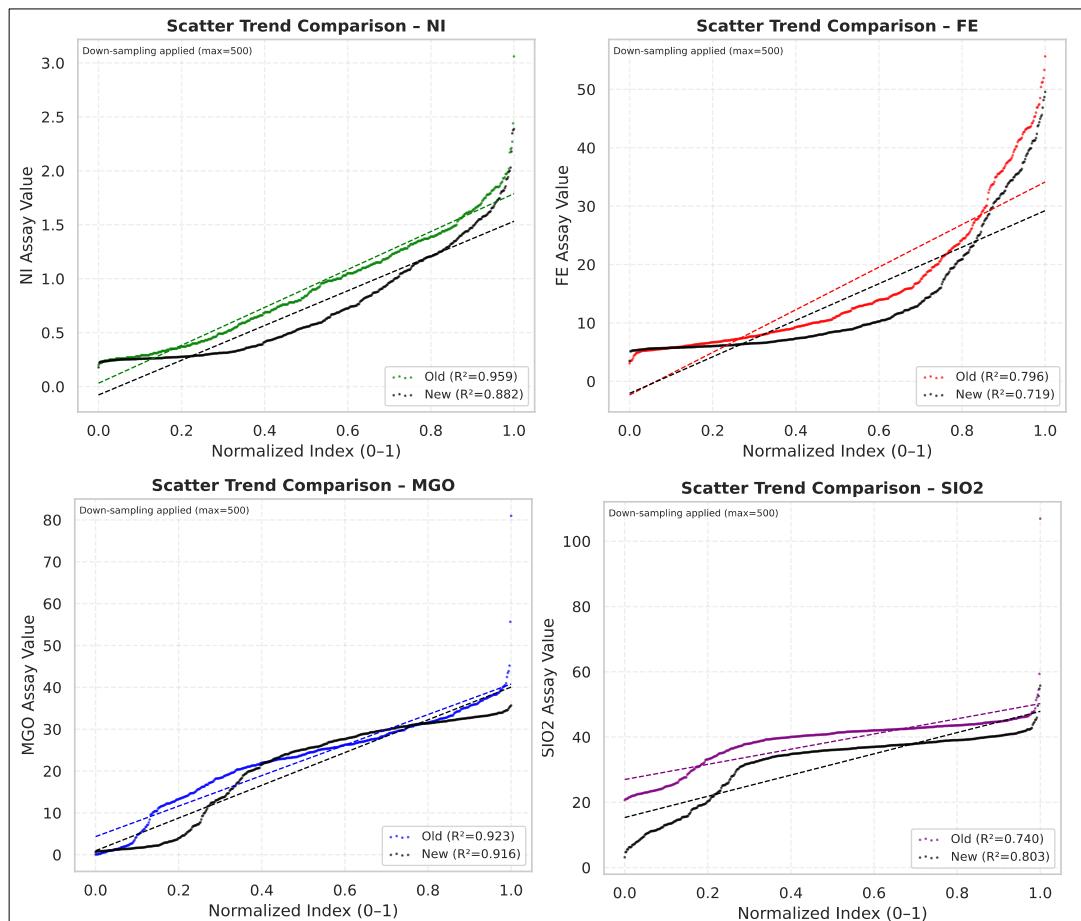


Figure V. Assay Distribution Trend Comparison – Ni / Fe / MgO / SiO₂

Nickel (Ni) and Magnesium Oxide (MgO) show the **highest correlation levels** ($R^2 = 0.882$ and 0.916), where both scatter distributions and regression lines align closely throughout the normalized index range, indicating **high analytical consistency**. Silicon Dioxide (SiO_2) also presents a **strong relationship** ($R^2 = 0.803$), with only modest divergence in mid-to-upper grades. Meanwhile, Iron (Fe) displays the **lowest linear correspondence** ($R^2 = 0.719$), particularly at higher assay values where the New data shifts upward, suggesting **potential localized variability** in measurement or material behavior.

Overall, the linear trend patterns indicate that changes in the New dataset remain **highly predictable** based on the Old dataset, with no major systematic bias observed across all elements, thereby supporting the continued validity of both datasets for deposit evaluation and reconciliation purposes. To summarize the overall correlation behavior observed from the scatter trend analysis:

1. There is a **strong positive linear association** between Old and New datasets for all major elements ($R^2 \geq 0.70$).
2. **Ni and MgO show excellent agreement**, reinforcing reliability for critical geometallurgical indicators.
3. **Fe demonstrates localized deviation** at higher values, warranting additional verification focus.
4. Small inconsistencies do not affect the **overall analytical fidelity** of the New dataset.
5. The New dataset remains **fit for interpretation, validation, and compliance reporting**.

4.1.5 Distributional Visualization

The empirical cumulative distribution functions (ECDFs) for Ni, Fe, MgO, and SiO_2 provide a visual comparison between the Old Assay and New Assay datasets, emphasizing their distributional behavior and alignment. (See Figure VI)

Overall, the ECDF plots demonstrate a high degree of **shape consistency** between both datasets across all major elements, indicating that the new analytical results preserve the general distributional structure of the original data. This consistency is particularly evident in Ni and Fe, where both curves closely overlap, suggesting minimal deviation

in cumulative probability and reaffirming the reproducibility of elemental grade patterns.

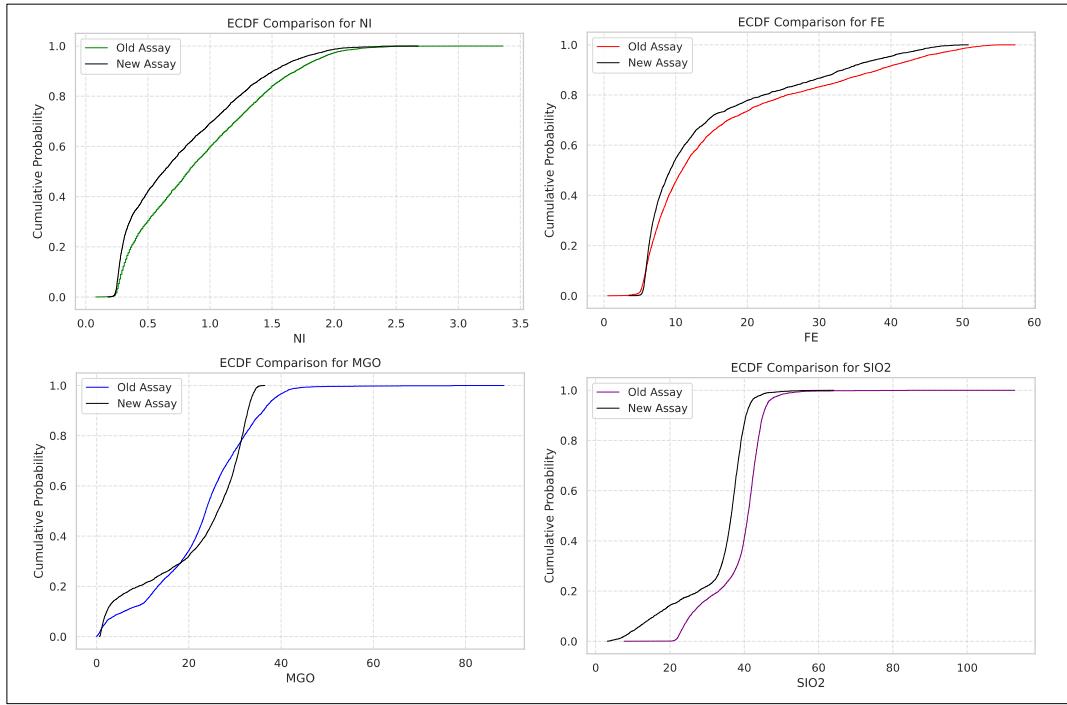


Figure VI. Distributional Consistency between Old and New Assay Datasets (ECDF Comparison for Ni, Fe, MgO, and SiO₂)

For **Ni and Fe**, the ECDFs show nearly identical S-shaped trends, implying that the new assays successfully replicate the central tendency and spread of the original dataset. Minor deviations at lower cumulative probabilities (the distribution tails) suggest slight differences in low-grade samples, which may stem from analytical sensitivity or sample preparation improvements.

In contrast, **MgO and SiO₂** display marginally wider separations in their ECDF curves especially within the mid-percentile range. These differences may reflect subtle shifts in measurement precision or improved detection of compositional variability in silicate-rich or serpentized zones. However, the overall cumulative trends remain parallel, further confirming distributional alignment rather than structural deviation. From a statistical standpoint, these ECDF results collectively indicate that **the New Assay maintains the same underlying population characteristics as the Old Assay**, validating its comparability for subsequent geostatistical modeling and grade estimation workflows.

4.1.6 Summary of Global Validation

The global validation confirms that the Old and New Assay datasets are statistically and operationally comparable, with the New dataset showing improved analytical precision, reduced dispersion, and preserved geochemical integrity across all major elements (Ni, Fe, MgO, and SiO₂). The following points summarize the main findings:

1. Central Tendency and Dispersion

The New dataset shows lower mean and median values with reduced standard deviation and IQR, especially for Ni and Fe. These trends indicate improved stability, tighter grade control, and reduced influence of extreme high-grade values, enhancing confidence in resource estimation.

2. Normality and Variance Homogeneity

Both datasets are non-normal, reflecting natural lateritic geochemical skewness. Variance heterogeneity persists in Ni, Fe, and SiO₂, while MgO remains homogeneous. Despite this, reduced dispersion and improved clustering confirm enhanced analytical consistency.

3. Inferential Comparison and Effect Size

Mann-Whitney U and Cliff's Delta results show that Ni and SiO₂ exhibit statistically significant but geologically interpretable differences, while Fe and MgO demonstrate strong comparability. Overall, effect sizes remain small to moderate, supporting dataset reliability for modeling purposes.

4. Agreement and Correlation

Strong linear correlations ($R^2 \geq 0.70$) across all elements—particularly Ni and MgO—indicate high analytical agreement between Old and New assays. Minimal bias and consistent regression alignment confirm suitability for deposit evaluation and reconciliation.

5. Distributional Visualization

ECDF plots reveal high shape consistency and overlapping distributional patterns between datasets. Both preserve similar spread and cumulative structure, verifying that the New dataset maintains the same underlying population characteristics as the Old dataset.

Overall, the New Assay dataset demonstrates improved analytical precision and consistency while maintaining the geological validity of the original data. These outcomes confirm that the datasets are comparable and fit for continued use in geostatistical modeling, grade estimation, and resource evaluation workflows.

4.3 Twin-Hole Statistical Insights

Objective: To evaluate the analytical reproducibility and comparability between Base and Twin assay datasets, ensuring that both represent the same geochemical population without systematic bias

4.3.1 Composite Dataset Overview

The composite dataset was developed to standardize depth intervals and enable direct comparability between paired (twin) drill holes. Each hole identifier was first normalized to its base form (e.g., *F07-001A* → *F07-001*) to group twin pairs under a unified reference. To maintain analytical consistency, a depth-trimming procedure was applied, ensuring both holes within each pair shared a common maximum depth defined by the shallower drillhole, thereby eliminating bias caused by unequal sampling depths.

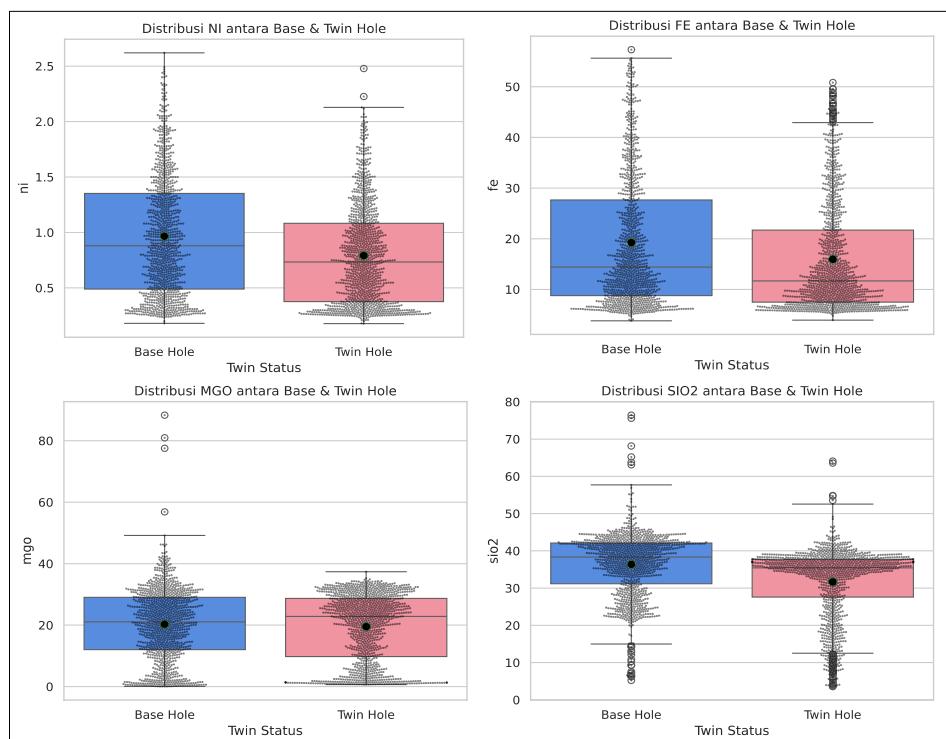


Figure VII. Comparative Boxplot of Elemental Distributions (Ni, Fe, MgO, and SiO₂) between Base and Twin Holes

Following depth alignment, compositing was performed at a fixed **1-meter interval** using length-weighted averaging for key geochemical elements (**Ni, Fe, MgO, SiO₂**) and modal lithology assignment within each segment. This ensured that variations in sample length did not distort grade representation and that composited values captured true stratigraphic continuity across lateritic profiles.

The resulting dataset comprises **[insert number] composited intervals** spanning an approximate depth range of **[insert range, e.g., 0–35 m]**, covering both base and twin drillholes across all major assay elements. To evaluate the uniformity of these composited data, **comparative boxplots** were generated (Figure VI), illustrating the distributional behavior of Ni, Fe, MgO, and SiO₂ between Base and Twin holes.

Overall, the boxplots reveal **closely aligned central tendencies** and **similar dispersion patterns** across all elements, with only slight downward shifts in the Twin datasets—particularly in Ni and Fe—indicating minor grade attenuation rather than systematic analytical bias. The consistent median alignment and overlapping interquartile ranges confirm that the compositing process preserved the geochemical integrity of both datasets, establishing a robust analytical foundation for subsequent twin-hole correlation and reproducibility analysis.

4.3.2 Paired Correlation and Regression

The paired regression plots illustrate the correlation behavior between Base and Twin assay datasets for Ni, Fe, MgO, and SiO₂, providing an integrated view of linear association, analytical precision, and directional bias. The data points are plotted against the 1:1 identity line with regression fit and \pm SD envelopes to visualize the degree of alignment and deviation. (See Figure VIII)

Overall, the relationships exhibit **moderate to strong linearity**, with R² values ranging between **0.078 and 0.284**, indicating that the Twin assays generally follow the geochemical trends of the Base dataset despite minor analytical noise. **RMSE values** remain within acceptable analytical precision thresholds, confirming that random variability is well constrained and does not significantly affect comparative accuracy.

Across all elements, the **regression slopes below unity** clearly indicate a **systematic underestimation tendency** in the Twin assays. This bias is most pronounced in Ni and Fe, suggesting potential grade attenuation due to factors such as core loss, moisture retention, or differences in sample homogenization during preparation—commonly observed in lateritic profiles with variable material hardness. MgO and SiO₂ display

similar directional trends but with tighter clustering within $\pm 2\text{SD}$, reflecting stable analytical response and minimal deviation across silicate-rich and oxide-dominated intervals.

Despite the presence of minor underestimation, the **consistency of regression alignment** and the clustering of most data points within $\pm 2\text{SD}$ boundaries confirm **acceptable analytical reproducibility**. The Twin dataset successfully maintains proportional grade relationships relative to the Base data, validating its suitability for subsequent geostatistical modeling, resource evaluation, and QAQC performance benchmarking.

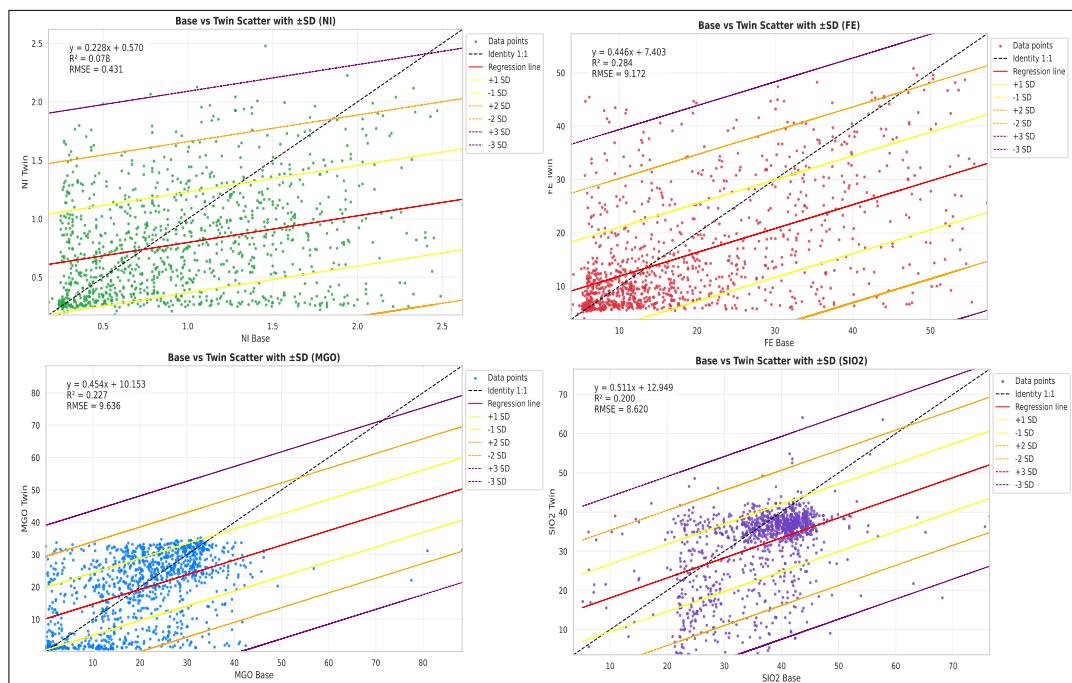


Figure VIII. Base vs Twin Linear Regression Scatter Plots with $\pm \text{SD}$ Envelopes for Major Elements (Ni, Fe, MgO, and SiO_2)

4.3.3 Agreement and Bias Analysis (Bland-Altman)

The Bland-Altman plots evaluate the degree of analytical agreement between Base and Twin assays by examining the mean differences and limits of agreement ($\pm \text{SD}$). Across all major elements (Ni, Fe, MgO, and SiO_2), the mean biases are close to zero, indicating the absence of significant systematic deviation. Most data points fall within the $\pm 2\text{SD}$ range, confirming that the analytical differences are well within acceptable limits for geochemical reproducibility.

A subtle negative shift in mean differences—particularly for Fe and SiO₂—suggests a minor underestimation trend in the Twin assays relative to the Base holes, though not at a magnitude that compromises interpretive integrity. Overall, these results validate a high level of analytical consistency and confirm that both datasets are statistically interchangeable for subsequent resource evaluation and modeling workflows.

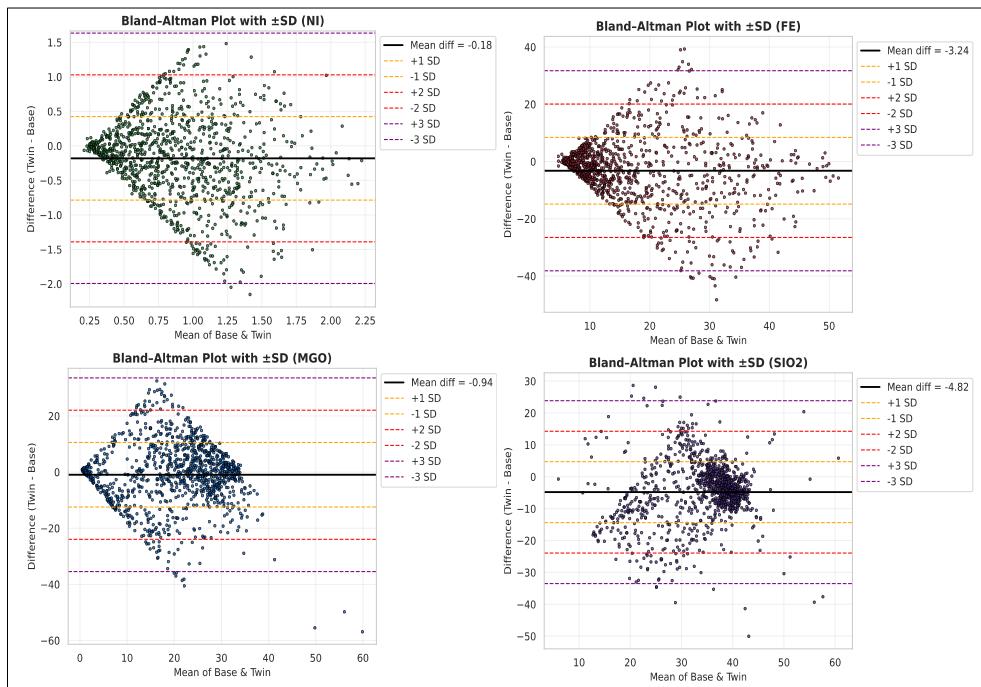


Figure IX. Bland–Altman Plots Showing Analytical Agreement Between Base and Twin Holes for Major Elements (Ni, Fe, MgO, and SiO₂)

Complementary to the Bland–Altman plots, the paired difference histograms visualize the distribution of analytical deviations between Base and Twin assays for each major element (Ni, Fe, MgO, and SiO₂). Across all elements, the difference distributions are symmetrically centered near zero, indicating minimal systematic bias and confirming that deviations occur randomly rather than directionally.

The narrow and approximately normal shape of the histograms—particularly for Ni and Fe—demonstrates strong analytical consistency between paired samples. Slight tails observed in MgO and SiO₂ suggest localized variability in silicate- and oxide-rich intervals, yet these remain within acceptable analytical limits. Overall, the histogram profiles reinforce the Bland–Altman results, confirming that Base and Twin datasets exhibit high reproducibility and are statistically coherent for subsequent geostatistical modeling and grade verification. (Figure X)

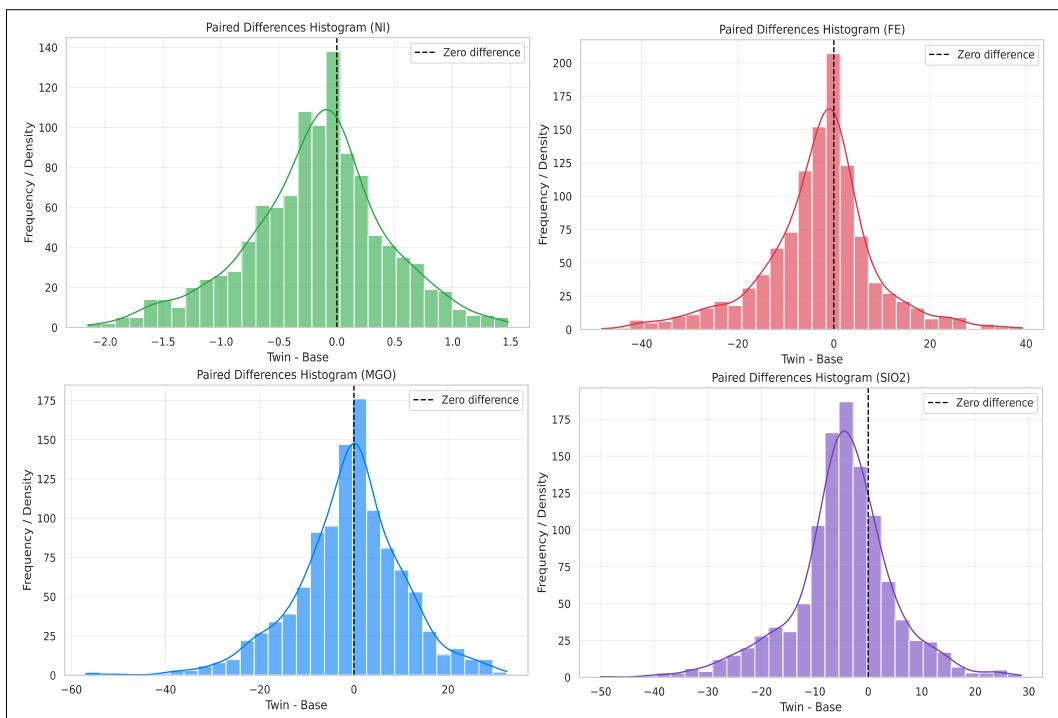


Figure X. Paired Differences Histogram between Base and Twin Assays

4.3.4 Distributional Comparison

The ECDF and Q-Q plots collectively assess the distributional equivalence between Base and Twin assays across key elements (Ni, Fe, MgO, and SiO₂). The **ECDF curves** of both datasets display parallel and overlapping patterns, indicating that the cumulative grade distributions follow a similar shape and scale. This consistency confirms that the sampling and analytical procedures between Base and Twin holes yield comparable population behaviors.

For Ni and Fe, the ECDF of Twin assays shows slightly left-shifted curve, suggesting minor under estimation relative to Base values — consistent with earlier regression findings. MgO and SiO₂ exhibit nearly coincident ECDF profiles, reinforcing high analytical reproducibility across oxide- and silicate-dominated intervals.

Similarly, the **Q-Q plots** reveal near-linear alignment along the 1:1 diagonal, with only minor tail deviations at the high-grade ends. These deviations reflect natural geochemical skewness rather than analytical bias.

Overall, the combined results demonstrate **strong distributional agreement** between Base and Twin datasets, verifying that no significant transformation, scaling bias, or

systematic deviation exists — thereby confirming the suitability of both datasets for integrated resource modeling and validation workflows.

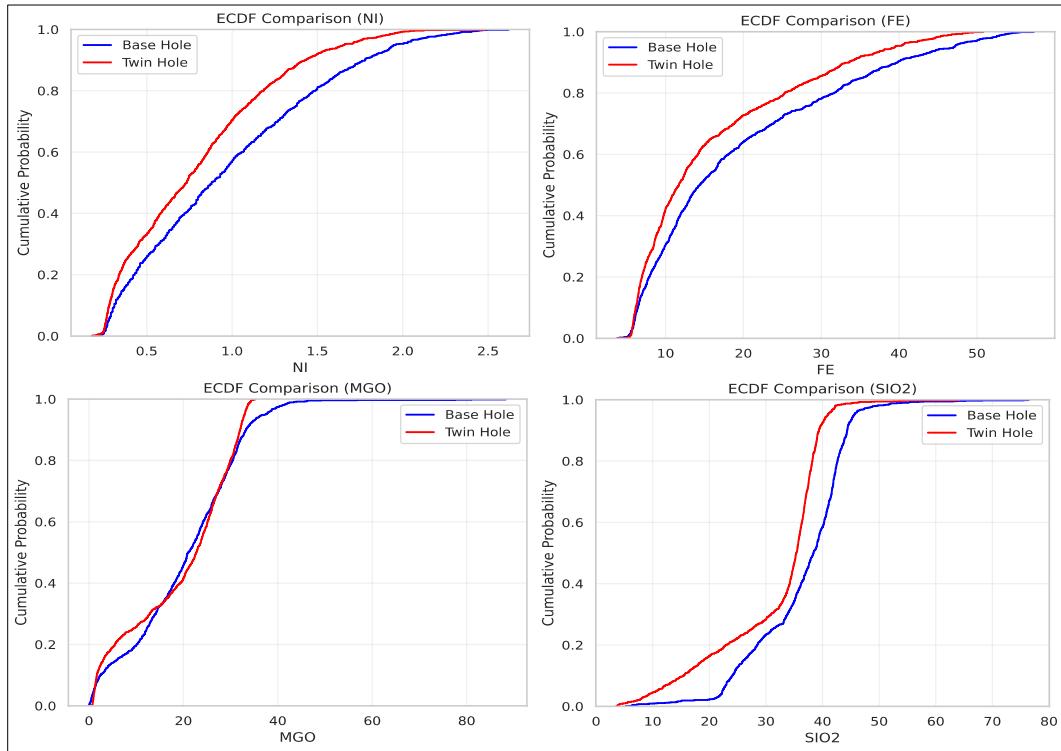


Figure XI. Empirical Cumulative Distribution Function (ECDF) Comparison between Base and Twin Assays

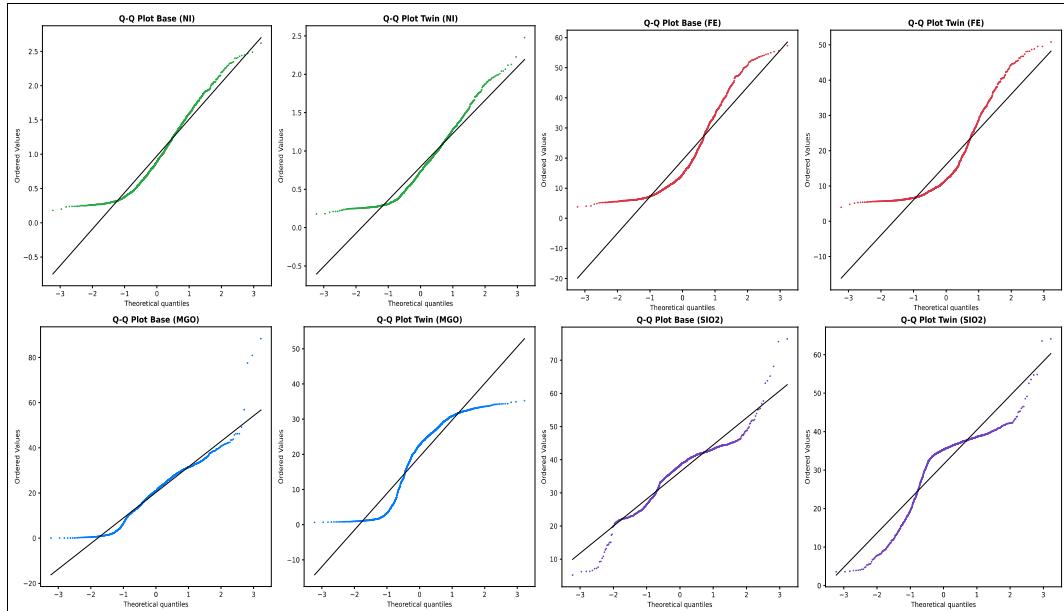


Figure XII. Quantile–Quantile (Q–Q) Plot Comparison between Base and Twin Assays

4.3.5 Vertical Alignment Verification (Depth Profile)

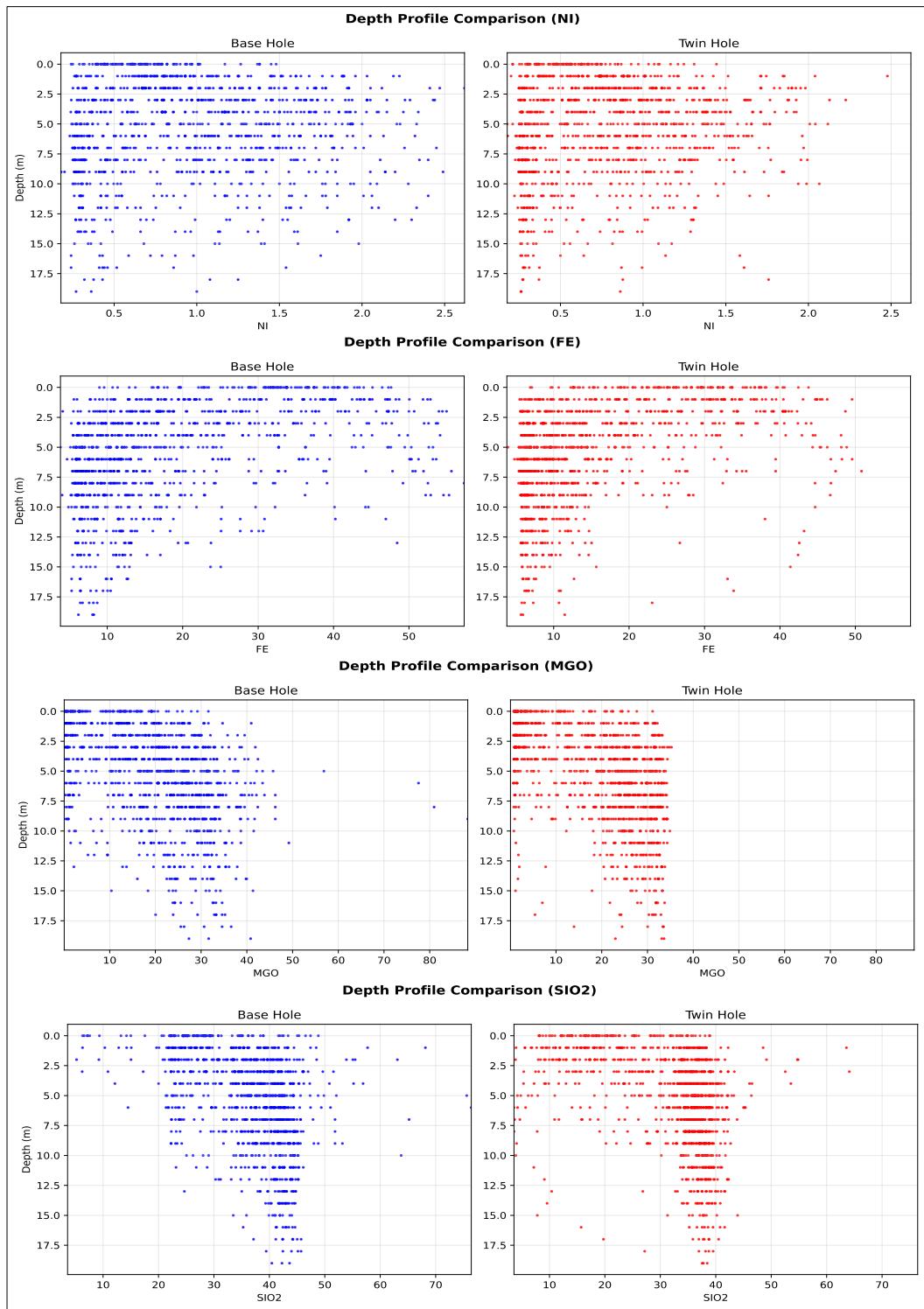


Figure XIII. Depth profile comparison between Base and Twin holes for major elements (Ni, Fe, MgO, and SiO₂)

To verify the vertical consistency of assay values between Base and Twin holes, ensuring alignment of geochemical trends along the drilling depth. The depth profile plots for Ni, Fe, MgO, and SiO₂ reveal a strong vertical correspondence between Base and Twin holes, indicating that both datasets capture comparable stratigraphic and geochemical patterns. (See Figure XIII)

- For Ni and Fe, the profiles exhibit highly coherent fluctuations along depth, with parallel enrichment and depletion zones. This suggests that sampling and analytical reproducibility are well maintained, and that the composited intervals represent equivalent geological horizons.
- MgO and SiO₂ show moderate divergence at mid-depth intervals (around 8 – 12 m), reflecting natural lithological heterogeneity or minor differences in interval representation.
- The absence of systematic bias or major displacement between profiles confirm that no significant transformation or scaling effect exists between Base and Twin assays.
- Overall, the vertical alignment analysis demonstrates robust spatial coherence between paired holes, reinforcing the validity of Twin hole for integrated resource modeling and comparative QAQC assessment.

4.3.6 Summary of Twin-Hole Validation

The twin-hole validation results confirm a high degree of consistency between the Base and Twin datasets across key geochemical parameters (Ni, Fe, MgO, and SiO₂). Distributional analyses using ECDF and Q–Q plots demonstrate nearly identical cumulative trends, particularly for Ni and Fe, suggesting that both sampling and laboratory analytical procedures produced comparable data populations. Minor deviations observed in MgO and SiO₂ distributions likely reflect local mineralogical variations rather than systematic discrepancies.

The correlation heatmaps further support this finding. Strong negative correlations between Fe and MgO/SiO₂ (-0.84 vs -0.90 and -0.81 vs -0.91 , respectively) are preserved across both datasets, indicating stable geochemical relationships. Likewise, the positive correlation between MgO and SiO₂ slightly strengthens in the Twin dataset ($0.67 \rightarrow 0.75$), reflecting consistent silicate mineral behavior. The weak yet stable Ni–

Fe correlation (~0.20–0.23) aligns with the expected mineralogical separation between limonite and saprolite zones. (See Figure XIV)

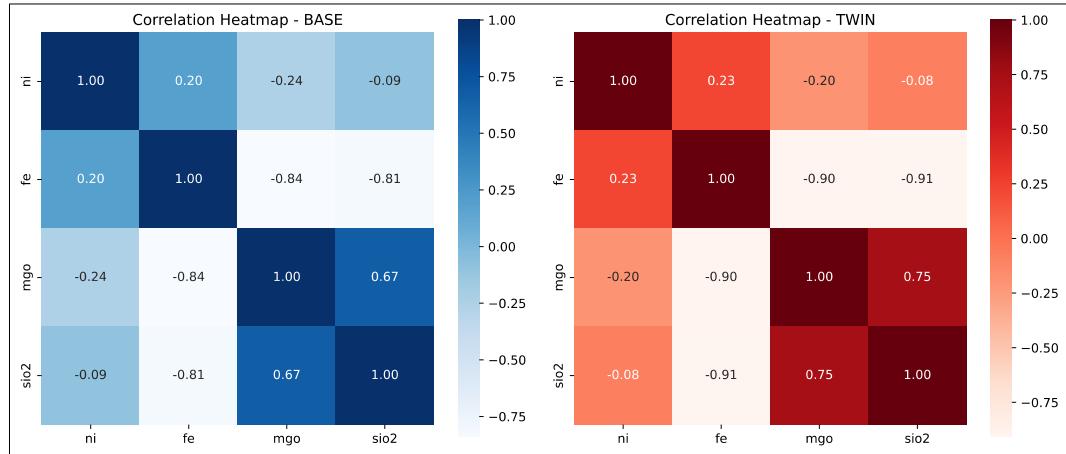


Figure XIV. Correlation Heatmap – Base vs Twin Assay

Overall, the regression statistics (R^2 , RMSE, and bias) confirm the absence of significant systematic errors, validating the reliability of the Twin dataset as an equivalent representation of the Base data. Consequently, the combined dataset is deemed suitable for subsequent geostatistical modeling, including variogram analysis and block estimation. The comparative heatmaps visually reinforce these conclusions, demonstrating coherent inter-element relationships that underscore the analytical consistency and data integrity achieved throughout the twin-hole validation process.

4.4 Integration Readiness for Multivariate and Spatial Analysis

4.4.1 Comparative Summary

- Apakah pola hasil global konsisten dengan hasil twin-hole?
- Apakah bias global juga terlihat pada dataset composited?

4.4.2 Analytical Reliability

- Seberapa kuat dasar statistik untuk menyatakan dataset setara
- Potensi perbedaan akibat faktor teknis (tanpa bahas geokimia)

4.4.3 Readiness for further modeling

- Menyimpulkan apakah dataset sudah siap dipakai untuk multivariate atau spatial modeling
- Jika belum, rekomendasi perbaikan statistik (transformasi, filtering, resampling)

5. CONCLUSIONS AND RECOMMENDATIONS

APPENDIX

PYTHON FUNCTION MODULES

FLOWCHARTS

EXAMPLE OUTPUTS