CIPM MRA
Comparison reports

GULFMET.TF-S2

Time scale difference

SUPPLEMENTARY COMPARISON

This report is published by the BIPM.

Any further distribution of this Report must be cited as:
A. Alassaf *et al* 2026 *CIPM MRA Comparison reports* **05002**

https://doi.org/10.59161/LCBK1480

Supplementary Comparison of UTC Time Scales Between UTC(SASO), UTC(UME), UTC(UZ), UTC(NIS), UTC(EMI), UTC(AZ) Using Global Navigation Satellite Systems (GNSS)

Khalid S Al Dawood[1], Waleed M. Al Harbi[1], Assaf Alassaf[1*], Ramiz Hamid[1], Adem Gedik[2], Sheroz Ismatullaev[3], Lyubov Gazieva[3], Vohobjon Nishonov[3], Mohamed El Hawary[4], Aly Ibrahim Mostafa[4], Shamkhal Abbasov[5], Nazrin Aliyeva[5], Jon Bartholomew[6]

[1] National Measurement and Calibration Center (NMCC), Saudi Standards, Metrology and Quality Organization (SASO), Riyadh, Saudi Arabia
[2] National Metrology Institute of Türkiye (TÜBİTAK UME), Kocaeli, Türkiye
[3] Uzbekistan National Institute of Metrology (UzNIM), Tashkent, Uzbekistan
[4] National Institute of Standards (NIS), Giza, Egypt
[5] Azerbaijan Metrology Institute (AzMI), Baku, Azerbaijan
[6] Emirates Metrology Institute (EMI), Abu Dhabi, United Arab Emirates
Corresponding Author*: a.assaf@saso.gov.sa

ABSTRACT

Under the framework of the Gulf Cooperation Council Standardization Organization's Regional Metrology Organization (GULFMET), a supplementary comparison of UTC time scales was conducted among national metrology institutes from Saudi Arabia (SASO-NMCC), Türkiye (TÜBİTAK UME), Uzbekistan (UzNIM), Egypt (NIS), the United Arab Emirates (EMI), and Azerbaijan (AzMI). The comparison employs the GNSS Common-View technique based on GPS L3P data to evaluate the performance and alignment of the participating UTC(k) realizations and to strengthen regional traceability to UTC. This exercise represents a cross-RMO comparison coordinated by GULFMET, involving institutes affiliated with EURAMET, COOMET, and AFRIMETS. The study presents the comparison methodology, measurement results, and experience gained from organizing a regional time-scale comparison, emphasizing the importance of interregional cooperation in reinforcing global time and frequency networks. In addition to the primary comparison objectives, an extended set of statistical analysis methods was applied to systematically profile the participating systems. Time-domain metrics such as mean offset, standard deviation, and linear drift were used to distinguish short-term stability from long-term frequency behavior using a common reference time scale (UTCr). These methods provide a consistent and transparent framework for comparative performance assessment across laboratories. The results demonstrate both the technical outcomes of the comparison and the value of enhanced statistical profiling for characterizing time-scale behavior. The work also supports the development of improved data handling, analysis, and reporting practices within GULFMET, contributing to a more integrated and collaborative regional time-keeping community.

1. INTRODUCTION

Precise intercomparison of national time scales is fundamental for maintaining the global coherence of Coordinated Universal Time (UTC) and ensuring traceability among timing laboratories worldwide. At present, more than 80 laboratories contributing to UTC maintain local realizations of the National Time Scale (denoted UTC(k)) on ensembles of atomic frequency standards, primarily commercial cesium beam clocks, cesium fountain primary frequency standards, and hydrogen masers used as flywheel oscillators. These laboratories perform continuous time transfer using Global Navigation Satellite Systems (GNSS), measuring the time difference between their local UTC(k) and GNSS time, and regularly transmitting these data to the Bureau International des Poids et Mesures (BIPM). The submitted clock and time-transfer data form the basis for the computation of International Atomic Time (TAI) and UTC, which are published monthly in BIPM Circular T [1, 2].

The Consultative Committee for Time and Frequency (CCTF) has established the CCTF-K001-UTC key comparison as the formal framework ensuring traceability of national time scales to UTC. Within this framework, the BIPM acts as the pilot laboratory, responsible for the coordination, analysis, and publication of UTC comparisons. Laboratories seeking traceability to UTC must therefore demonstrate the performance of their UTC(k) through participation in recognized comparisons and through submission of validated GNSS time-transfer data.

Before participating in global UTC contributions or key comparisons, newly established or developing laboratories typically engage in bilateral or multilateral time scale comparisons within their Regional Metrology Organization (RMO). Such comparisons are essential for validating local time scale generation, assessing relative frequency accuracy, and gaining

experience in time-transfer data analysis. In particular, comparisons with laboratories already contributing to UTC are highly beneficial for ensuring metrological consistency and readiness for future participation in global coordination continuous bilateral and multilateral time scale comparisons using the Global Navigation Satellite Systems Common-View (GNSS CV) method provide laboratories with sustained benefits through systematic data analysis, long-term performance monitoring, and a deeper understanding of their local time scale behavior, including stability, drift, and noise characteristics.

In this context, the present study reports a supplementary regional time scale comparison conducted under the GULFMET framework, involving six National Metrology Institutes (NMIs). The participating laboratories are Saudi Arabia's SASO-NMCC (pilot laboratory), TÜBİTAK UME (Türkiye), NIS (Egypt), EMI (United Arab Emirates), UzNIM (Uzbekistan), and AzMI (Azerbaijan). These institutes span West and Central Asia and North Africa, forming inter-laboratory baselines of up to several thousand kilometers.

The GNSS Common-View (CV) method was selected as the comparison technique due to its proven precision [3], robustness, and long-standing use by the BIPM since the 1980s [4-6]. In the CV approach, each laboratory measures the time difference between its local UTC(k) and GNSS time using a co-located receiver. Simultaneous observations of the same GNSS satellites at two remote sites are differenced, effectively canceling common satellite and propagation errors. This yields the relative time offset between two UTC(k) realizations with sub-nanosecond-level uncertainty, provided that identical signals and observation schedules are used [7].

The primary objective of this comparison is to evaluate the mutual consistency of the participating UTC(k) time scales, quantify relative offsets, frequency drifts, and stability characteristics, and identify any systematic effects requiring corrective action. Beyond its immediate technical results, this comparison serves as a foundational step toward enhanced regional coordination in time and frequency metrology within GULFMET. Prior to this work, no formal regional time scale comparison had been conducted under the GULFMET framework; thus, the present study represents a first assessment of inter-laboratory synchronization in the region [8-9].

Furthermore, the outcomes of this comparison provide critical input for future calibration activities, including on-site atomic clock calibrations using GNSS-based time transfer, and demonstrate the region's readiness to participate more fully in global UTC coordination. The comparison results are presented in a format suitable for international scrutiny, consistent with RMO-level studies supporting eventual participation in CCTF-K001-UTC and contribution to UTC.

2. METHODOLGY

2.1 GNSS Common-View Time Transfer and Data Processing

Time scale comparisons between participating laboratories were performed using the GNSS Common-View (CV) time transfer method [10], as schematically illustrated in Figure 1. This technique enables high-precision comparison of remote time scales by exploiting simultaneous observations of signals transmitted from the same GNSS satellite at geographically separated sites.

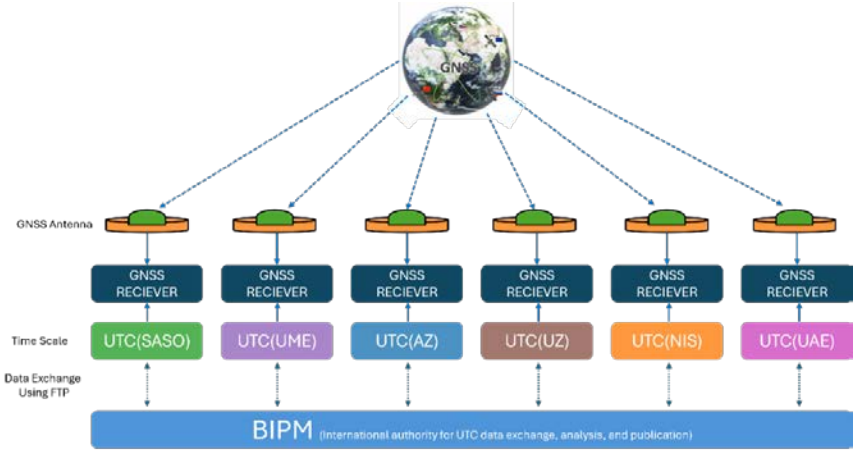

Figure 1. GNSS Common-View time transfer scheme for intercomparison of UTC(k) time scales. Participating laboratories observe the same GNSS satellites simultaneously, derive relative time offsets, and exchange validated data with the BIPM for UTC contribution and traceability.

Global Navigation Satellite Systems (GNSS), including GPS, GLONASS, Galileo, and BeiDou, consist of constellations of medium-Earth-orbit satellites equipped with stable onboard atomic clocks [11]. These satellites continuously broadcast timing signals referenced to their respective system time scales. Each participating laboratory receives these signals using a calibrated GNSS receiver connected to its local realization of Coordinated Universal Time, denoted UTC(k).

In the Common-View approach, two laboratories (A and B) track the same GNSS satellite at the same epoch. Each laboratory measures the time difference between its local clock and the received satellite signal. Differencing these measurements cancels common satellite-related errors such as satellite clock offsets and orbit errors yielding a precise estimate of the relative offset between the two UTC(k) realizations [10].

Mathematically, the Common-View time difference between Laboratory A and Laboratory B is expressed as:

$$[A - (R + d_A)] - [B - (R + d_B)] = A - B - (d_A - d_B) \quad (1)$$

Where:

- A and B denote the local time scales UTC(A) and UTC(B)
- R represents the reference signal of GNSS satellite's that is observed in a common view
- d_A and d_B are the total signal propagation delays from the satellite to the receivers at laboratories A and B, respectively.

The propagation delays d_A and d_B include geometric path delay, ionospheric and tropospheric delays, and receiver-specific hardware delays. These delays are corrected using standard GNSS models and laboratory-specific calibration values.

2.2 Dual-Frequency Combination and L3P Observable

In this comparison to mitigate first-order ionospheric delay effects, dual-frequency GNSS observations were processed using the L3P combination from GPS, which is the standard ionosphere-free pseudorange observable used in time transfer applications [12]. The L3P observable is formed as a linear combination of measurements from two carrier frequencies $f_1$and f_2, expressed as:

$$P_{L3} = \frac{f_1^2 P_1 - f_2^2 P_2}{f_1^2 - f_2^2} \quad (2)$$

where P_1 and P_2 are the pseudorange measurements on frequencies f_1 and f_2 , respectively [1]. This combination effectively removes the first-order ionospheric delay, significantly improving time transfer accuracy and long-term stability.

2.3 Data Screening and Quality Control

To ensure data quality and consistency across all laboratories, a minimum satellite elevation mask of 30° was applied [13]. Observations below this elevation were excluded to reduce multipath effects and tropospheric modeling errors.

Outlier detection was performed using a 1.5σ (where σ is standard deviation) statistical criterion, applied to the Common-View time difference series. Measurements deviating more than 1.5σ from the local mean were rejected. This procedure suppresses sporadic disturbances caused by receiver glitches, tracking anomalies, or unmodeled propagation effects, while preserving the underlying time scale behavior.

2.4 Weighted REFSYS Computation

For each laboratory, a weighted REFSYS (reference system) time series was computed from the validated Common-View observations. The REFSYS represents a weighted average of all available GNSS satellite links at each epoch, derived from

UTC(k) − GPS time measurements, and is used as a stable reference for assessing the behavior of the local UTC(k). The analysis was performed over the interval MJD 61001–61036 [14].

The weighted REFSYS time difference is calculated as:

$$REFSYS(t) = \frac{\sum_i w_i . \Delta t_i(t)}{\sum_i w_i} \quad (3)$$

Where:

- $\Delta t_i(t)$is the Common-View time difference associated with satellite i at epoch t,
- $w_i = 1/\sigma_i^2$ is the statistical weight assigned to each observation, derived from its estimated uncertainty σ_i.

This weighting scheme emphasizes higher-quality satellite links and reduces the influence of noisy observations. The resulting REFSYS time series provides a robust basis for evaluating time offsets, relative frequency drift, and stability metrics. Validated Common-View data and derived time differences were exchanged daily with the BIPM, which acts as the central authority for UTC data collection [1, 9], validation, and analysis. This methodology is fully consistent with international practices used for UTC generation, RMO-level comparisons, and preparation for participation in CCTF-K001-UTC.

3. RESULTS AND DISCUSSION

3.1 UTCr Profiling

To support the profiling and performance assessment of the participating laboratories, UTCr values published by the Bureau International des Poids et Mesures (BIPM) were used as an additional validation. Specifically, UTCr Circulars 2547–2552 were employed to analyze the behavior of the contributing NMIs over the comparison interval [14]. For each laboratory k, the time differences UTCr − UTC(k) were computed and evaluated alongside the inter-laboratory GNSS Common-View links.

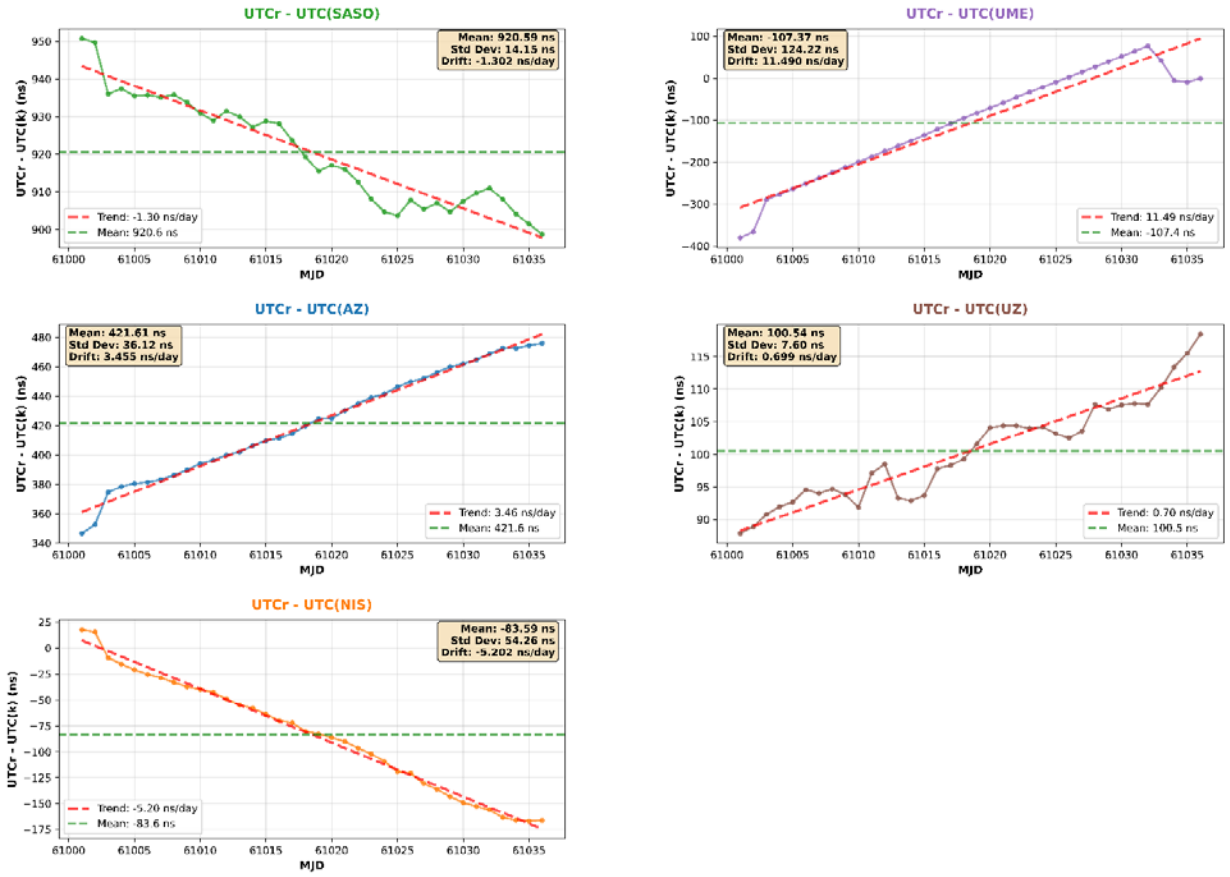

Figure 2. Time series comparison of five NMI time scales against UTCr reference (MJD 61001-61036). Each panel shows [UTCr - UTC(k)] deviations, linear regression (red dashed line), and mean offset (green dashed line). Statistical metrics (mean, standard deviation, drift rate) are displayed for each NMI: SASO, UME, AzMI, UzNIM, and NIS (UAE data excluded as no rapid UTC available).

This UTCr-referenced analysis provides an independent characterization of time offset, linear drift, and stability, relative to the global UTC ensemble. UTCr data were not available for UTC(UAE) during the considered period; therefore, UAE was excluded from the UTCr-based profiling while remaining included in the GNSS Common-View intercomparison results. The combined use of UTCr − UTC(k) time series and bilateral GNSS Common-View results enables a coherent evaluation of both global alignment to UTC and regional synchronization performance among the participating NMIs.

UTC(SASO) against UTCr exhibits a mean offset of approximately +920 ns with a small standard deviation (~14 ns), indicating very low short-term noise. A modest negative drift (−1.3 ns/day) is observed, reflecting a small relative frequency offset with respect to UTCr. The time series remains smooth, with no abrupt jumps, demonstrating good internal stability.

UTC(UME) against UTCr shows a mean offset of −107 ns and a moderate standard deviation (~124 ns). The dominant feature is a clear positive linear drift (~+11.5 ns/day), indicating a systematic frequency difference relative to UTCr throughout the comparison period. The largely linear behavior suggests deterministic frequency offset rather than stochastic instability.

UTC(AZ) against UTCr presents a larger mean offset (~421 ns) and standard deviation (~36 ns). The time series is highly linear with a positive drift (~+3.5 ns/day), indicating stable operation but with a persistent relative frequency offset. Short-term fluctuations are limited compared to the overall trend.

UTC(UZ) against UTCr shows a mean offset of ~100 ns and a standard deviation of ~7 ns. The drift is the smallest among the participating laboratories (~+0.7 ns/day), indicating good long-term frequency agreement with UTCr, despite some visible short-term variations.

UTC(NIS) against UTCr exhibits a mean offset of (~−84 ns) with a relatively larger standard deviation (~54 ns). A pronounced negative drift (−5.2 ns/day) is observed, indicating a systematic frequency deviation relative to UTCr. The near-linear trend suggests that the dominant contribution is frequency offset rather than random noise.

To complement the qualitative assessment from the time-series plots, Table 1 summarizes the key statistical indicators derived using UTCr as a common reference, including the mean offset, standard deviation, linear drift rate, and the Allan deviation evaluated at an averaging time of $\tau = 5$ days. While the mean offset and drift characterize the long-term alignment and relative frequency behavior of each time scale, the Allan deviation provides an independent measure of mid-term frequency stability, largely insensitive to constant offsets and linear trends.

NMI	Mean (ns)	Std Dev (ns)	Drift Rate (ns/day)	ADEV (5 days) (ns)
SASO	+920.59	14.15	-1.3022	5.03
UME	-107.37	124.22	+11.4901	43.47
AzMI	+421.61	36.12	+3.4552	11.95
UzNIM	+100.54	7.60	+0.6989	2.74
NIS	-83.59	54.26	-5.2015	18.45

Table 1. Summary statistics of NMI time scales versus UTCr (MJD 61001-61036) shown in Figure 2. Mean offset, standard deviation, drift rate, and Allan deviation (τ=5 days) for five participating NMIs: SASO, UME, AzMI, UzNIM, and NIS.

The ADEV results are consistent with the trends observed in the time series. UzNIM, which exhibits the smallest drift, also shows the lowest ADEV (2.74 ns at 5 days), confirming its good frequency stability over multi-day averaging times. SASO similarly demonstrates low ADEV (5.03 ns), reflecting its smooth time evolution and limited short-term noise. In contrast, UME and NIS, which display larger drift magnitudes and increased scatter, show higher ADEV values (43.47 ns and 18.45 ns, respectively), indicating reduced stability at the 5-day averaging time. AzMI occupies an intermediate position, with moderate drift and corresponding ADEV. In this context, ADEV($\tau = 5$ days) in ns quantifies the multi-day wander of the UTCr–UTC(k) offset (i.e., the typical change between adjacent τ-day averaged offsets), and therefore represents the relative stability to the UTCr ensemble, which may include correlation effects because UTCr is not an independent reference.

To further illustrate these findings, Figure 3 provides a complementary UTCr-based performance visualization of the participating NMIs.

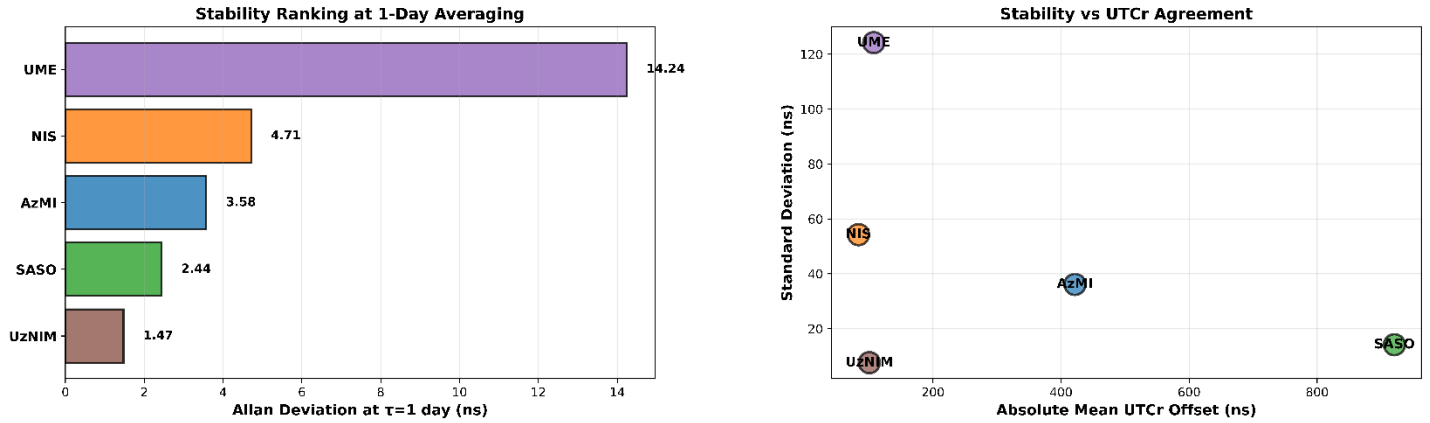

Figure 3. UTCr-based performance assessment for participating NMIs. **(Left)** Allan deviation ranking at τ=1 day for five NMIs, color-coded by laboratory. Lower values indicate better frequency stability. **(Right)** Scatter plot of standard deviation versus absolute mean UTCr offset. Ideal performance is in the bottom-left region (low offset, low scatter). Data period: MJD 61001-61036.

The left panel presents a stability ranking based on the Allan deviation at $\tau = 1$ day, highlighting relative short-term offset wander with respect to UTCr. This ranking is consistent with the multi-day ADEV results in Table 1, with UzNIM and SASO exhibiting the lowest values, while UME shows the largest short-term variability. The right panel places stability in the context of agreement with the reference by plotting the standard deviation versus the absolute mean UTCr offset. This representation enables simultaneous assessment of offset magnitude and dispersion, where optimal performance corresponds to the lower-left region of the plot. Together, these visualizations reinforce the separation between long-term alignment (mean offset), relative frequency behavior (drift), and offset stability (ADEV), providing an integrated view of UTC(k) performance relative to the UTCr ensemble.

3.2 Bilateral Comparison with UTC(SASO)

Following the UTCr-based profiling presented in Section 3.1, the analysis is extended to bilateral time-scale comparisons using UTC(SASO) as a reference**.** This approach allows direct pairwise assessment of time-scale behavior and enables inclusion of laboratories for which a direct UTCr comparison is not available.

In particular, the use of UTC(SASO) as a common reference permits the evaluation of UTC(UAE), which could not be included in the UTCr-based analysis due to the absence of rapid UTC data. By adopting a single-laboratory reference, consistent time-difference series can be formed for all participating NMIs, ensuring continuity of the comparison framework.

The bilateral comparison focuses on the analysis of UTC(SASO) − UTC(k) behavior, examining mean offset, linear drift, and short-term dispersion derived from GNSS Common-View observations. Compared to the ensemble-based UTCr analysis, this approach emphasizes relative alignment and frequency behavior between individual laboratories, providing additional insight into link performance and inter-laboratory consistency.

To ensure the validity of the evaluation, data affected by known clock adjustments were excluded. For UTC(UAE), measurements from MJD 61001 to 61015 were removed due to adjustment activities performed prior to stable operation. Similarly, for UTC(UME), data after MJD 61032 were excluded, as adjustments carried out beyond this point would affect the interpretation of offsets, drift, and stability metrics.

Figure 4 presents the corresponding time series derived from UTC(k) - GPS time at each laboratory. The results are discussed in terms of relative offsets, drift characteristics, and standard deviation.

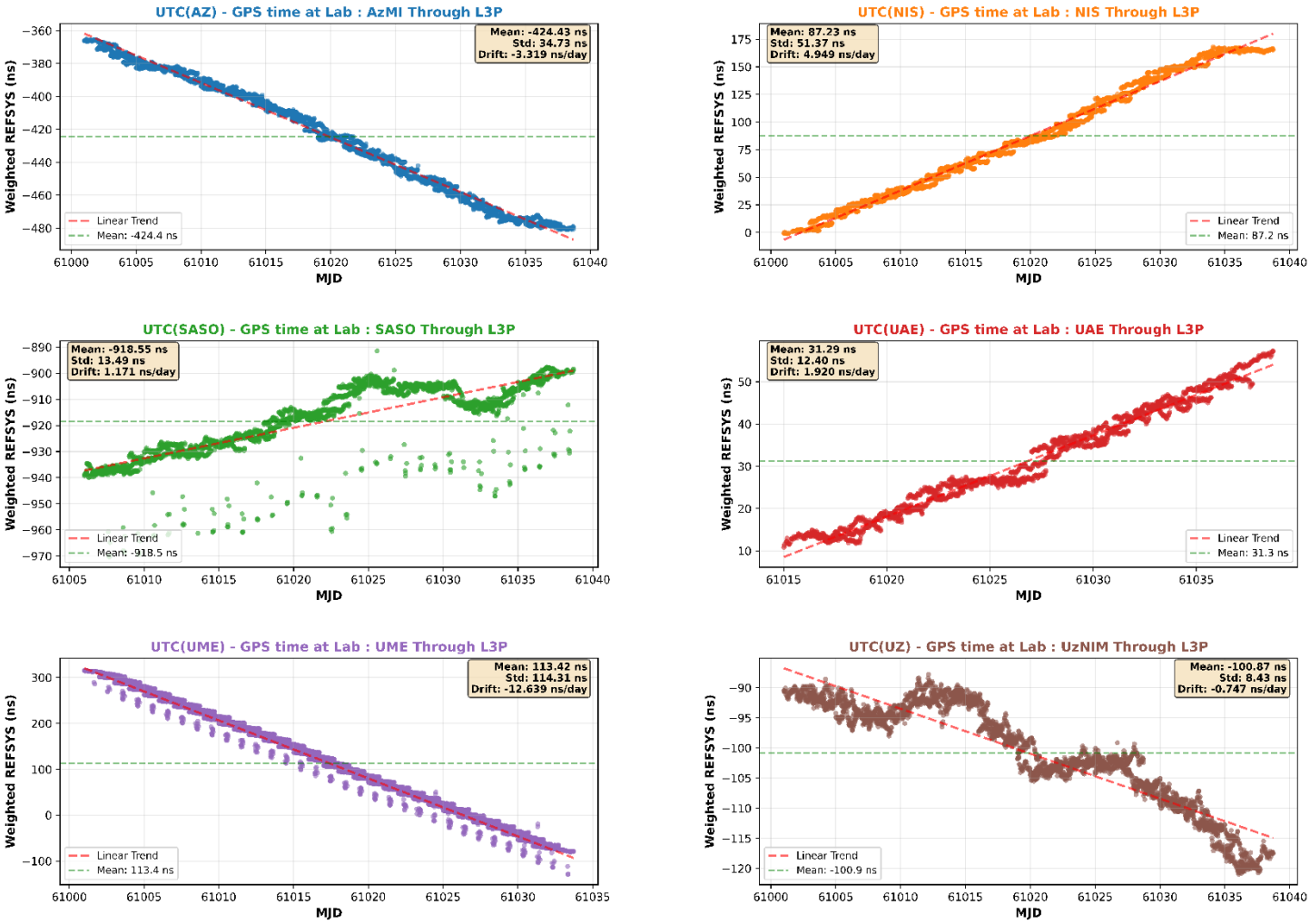

Figure 4. Comparison of NMI time scales with GPS time. Six panels display the offset [UTC(k) − GPS time] measured at each participating laboratory as a function of Modified Julian Date. (a) UTC(AZ) − GPS time at Azerbaijan Metrology Institute. (b) UTC(NIS) − GPS time at National Institute of Standards, Egypt. (c) UTC(SASO) − GPS time at Saudi Standards, Metrology and Quality Organization; the low-lying branch of scattered points corresponds to low-elevation satellite tracks with larger residual tropospheric (wet-delay) and multipath errors. (d) UTC(UAE) − GPS time at Emirates Metrology Institute. (e) UTC(UME) − GPS time at TÜBİTAK National Metrology Institute, Turkey; the banded structure reflects per-track quantization of the REFSYS data rather than physical clock variations. (f) UTC(UZ) − GPS time at Uzbekistan National Institute of Metrology. Each panel shows individual GPS measurement points (colored markers), the linear least-squares fit (red dashed line), and the mean offset (green dashed line). Statistical information including mean offset (ns), standard deviation (ns), and drift rate (ns/day) is displayed in each panel. REFSYS values represent the weighted time difference measurements.

Building on the laboratory-level UTC(k)–GPS time series shown in Figure 4, the analysis is further consolidated by forming direct bilateral time-scale differences with respect to UTC(SASO). This representation removes the GPS reference and places all participating laboratories on a common time axis, enabling direct comparison of relative offsets, drift behavior, and stability across NMIs.

Figure 5 presents the resulting time series of UTC(SASO) − UTC(k) for all participating laboratories over the observation period. By plotting all bilateral links on a single axis, this figure provides a compact overview of inter-laboratory time-scale relationships and highlights differences in relative alignment and temporal evolution. The slopes of the smoothed trends reflect relative frequency offsets with respect to UTC(SASO), while the scatter around each trend indicates short-term dispersion of the bilateral links.

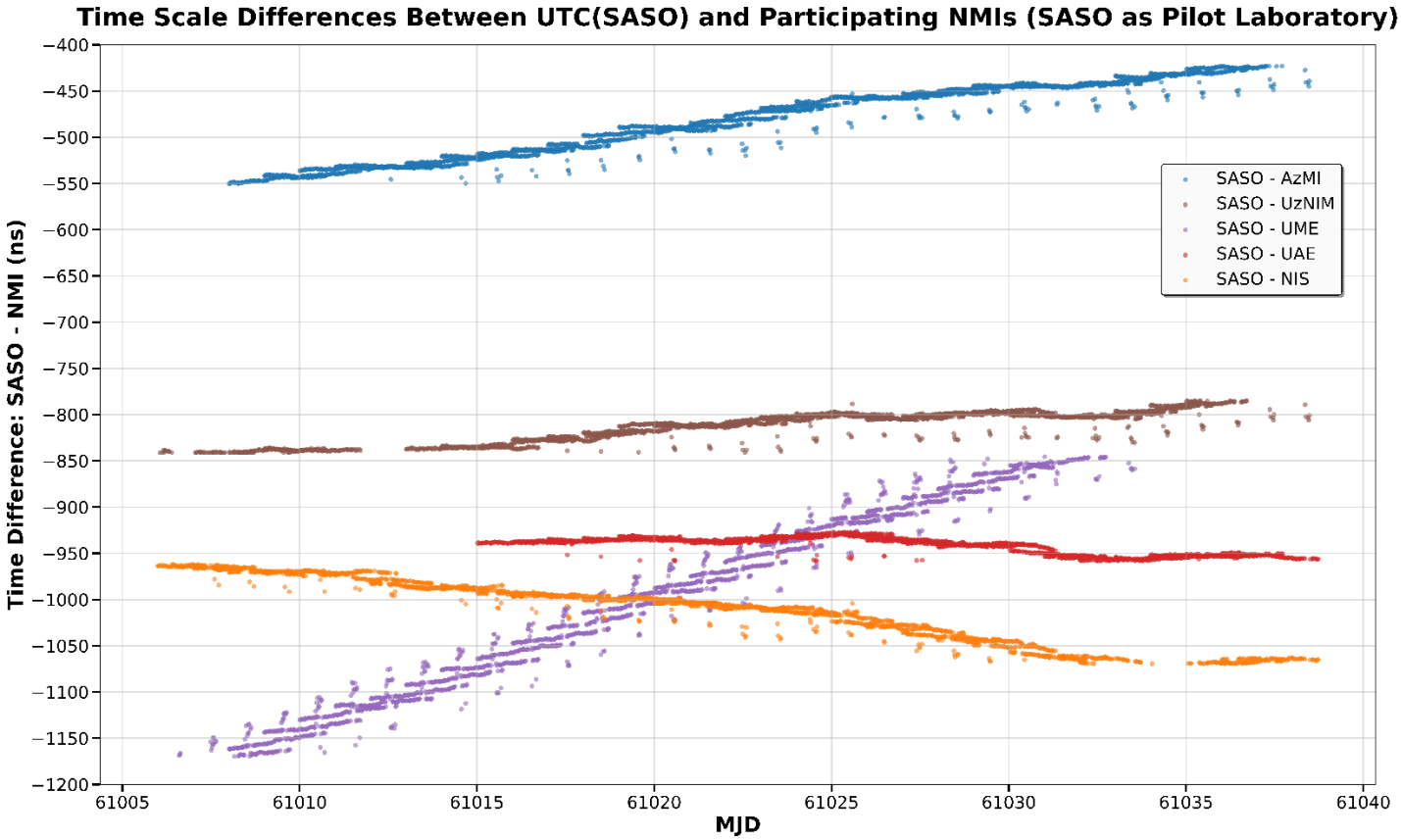

Figure 5. Time scale differences between UTC(SASO) and participating NMIs. The plot displays the time difference [UTC(SASO) - UTC(k)] as a function of Modified Julian Date for five NMI pairs, with SASO serving as the pilot laboratory. Each colored line represents the bilateral comparison between UTC(SASO) and another participating NMI: AzMI (blue), UzNIM (brown), UME (purple), UAE (red), and NIS (orange). Data points show individual measurement. y-axis shows the time difference in nanoseconds, with negative values indicating that the remote NMI time scale is ahead of UTC(SASO). The temporal evolution reveals the relative drift rates and stability characteristics between SASO and each participating laboratory over the observation period.

Compared to the UTCr-based profiling in Section 3.1, this SASO-referenced view emphasizes pairwise consistency and relative behavior between laboratories, including those not represented in the UTCr analysis. The combined use of Figures 4 and 5 therefore enables a comprehensive interpretation of both absolute laboratory behavior with respect to GPS time and direct inter-laboratory time-scale differences within the comparison.

3.3 Multi-Link Analysis of Bilateral Time Scale Comparisons

Following the UTCr-based profiling in Section 3.1 and the SASO-referenced bilateral analysis in Section 3.2, the evaluation is extended to a multi-link analysis of all bilateral time scale comparisons formed between the participating NMIs. This approach treats each bilateral link as an independent realization of relative time-scale behavior and allows systematic patterns in stability, offset, and drift to be examined without reliance on a single reference time scale. In a GNSS Common-View comparison, the observed time differences reflect the combined behavior of two local time scales and the associated measurement link. Analyzing all links collectively therefore provides insight into whether observed performance characteristics are consistent across multiple pairings or are specific to individual links. To enable direct comparison of temporal behavior, time-difference series are examined both in their absolute form and after normalization by removal of the mean offset, which emphasizes dispersion and drift structure rather than alignment.

The multi-link analysis focuses on short-term dispersion, mean offset, and linear drift as complementary indicators of link performance. Stability is further characterized using the Allan deviation evaluated at $\tau = 1$ day, which, when applied to time-difference data, quantifies the day-to-day wander of the relative offset and provides a measure of short-term link stability that is largely insensitive to constant offsets. The results of this analysis are presented in the following figures, beginning with a time-domain overview of all links grouped by performance tier.

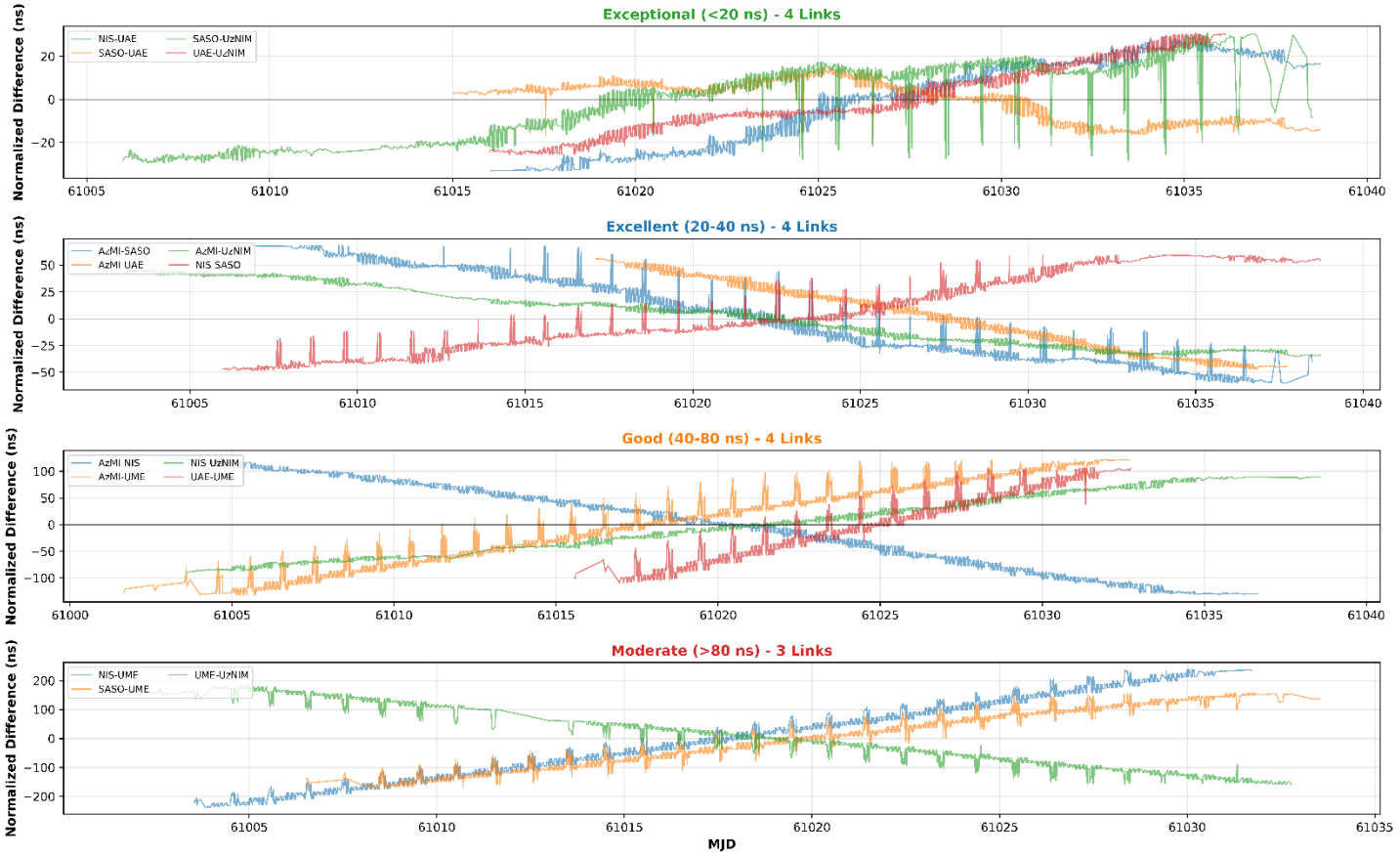

Figure 6. Time series overview of all NMI comparison links grouped by performance tier. Four panels display normalized time differences as a function of Modified Julian Date, with data grouped according to Normalized Difference (N_D): (a) Exceptional (N_D < 20 ns) - 4 links showing the highest stability performance (NIS-UAE, SASO-AzMI, SASO-UAE, UAE-UzNIM), (b) Excellent (20-40 ns) - 4 links with very good stability (AzMI-SASO, AzMI-UzNIM, UME-SASO, NIS-SASO), (c) Good (40-80 ns) - 4 links with good performance (AzMI-NIS, AzMI-UME, NIS-UME, UzNIM-UME), and (d) Moderate (N_D > 80 ns) - 3 links showing moderate stability (NIS-UAE, UAE-AzMI, SASO-UME). All time series are normalized to their respective mean values to facilitate visual comparison of temporal behavior and drift patterns. Each colored line represents one bilateral comparison link between participating NMIs. The y-axis shows the normalized difference in nanoseconds, while the x-axis displays Modified Julian Date.

Figure 6 presents a consolidated time-domain overview of all bilateral NMI time-scale comparison links, grouped according to their short-term stability performance. Each panel displays normalized Common-View time differences as a function of Modified Julian Date, with links classified into four performance tiers based on the normalized difference of the time-difference series. For each bilateral link, the Common-View time difference sequence Δt(t) was first computed from validated GNSS observations. To allow direct visual comparison of temporal behavior across links with different absolute offsets, each series was normalized by subtracting its own mean value over the comparison interval. This normalization removes constant offsets and emphasizes relative temporal behavior, such as drift, noise level, and transient features, while preserving the intrinsic time-scale dynamics of each link.

The classification into performance tiers is based on the normalized difference (N_D) of time-difference series. Links with N_D < 20 ns are labeled *Exceptional*, 20–40 ns *Excellent*, 40–80 ns *Good*, and N_D > 80 ns *Moderate*. These thresholds provide a practical stability-based segmentation that reflects daily drift coherence between the compared time scales.

The *Exceptional* and *Excellent* groups exhibit low drift, indicating high-quality GNSS links and well-behaved local clock ensembles. In contrast, links in the *Good* and *Moderate* categories show higher drifts.

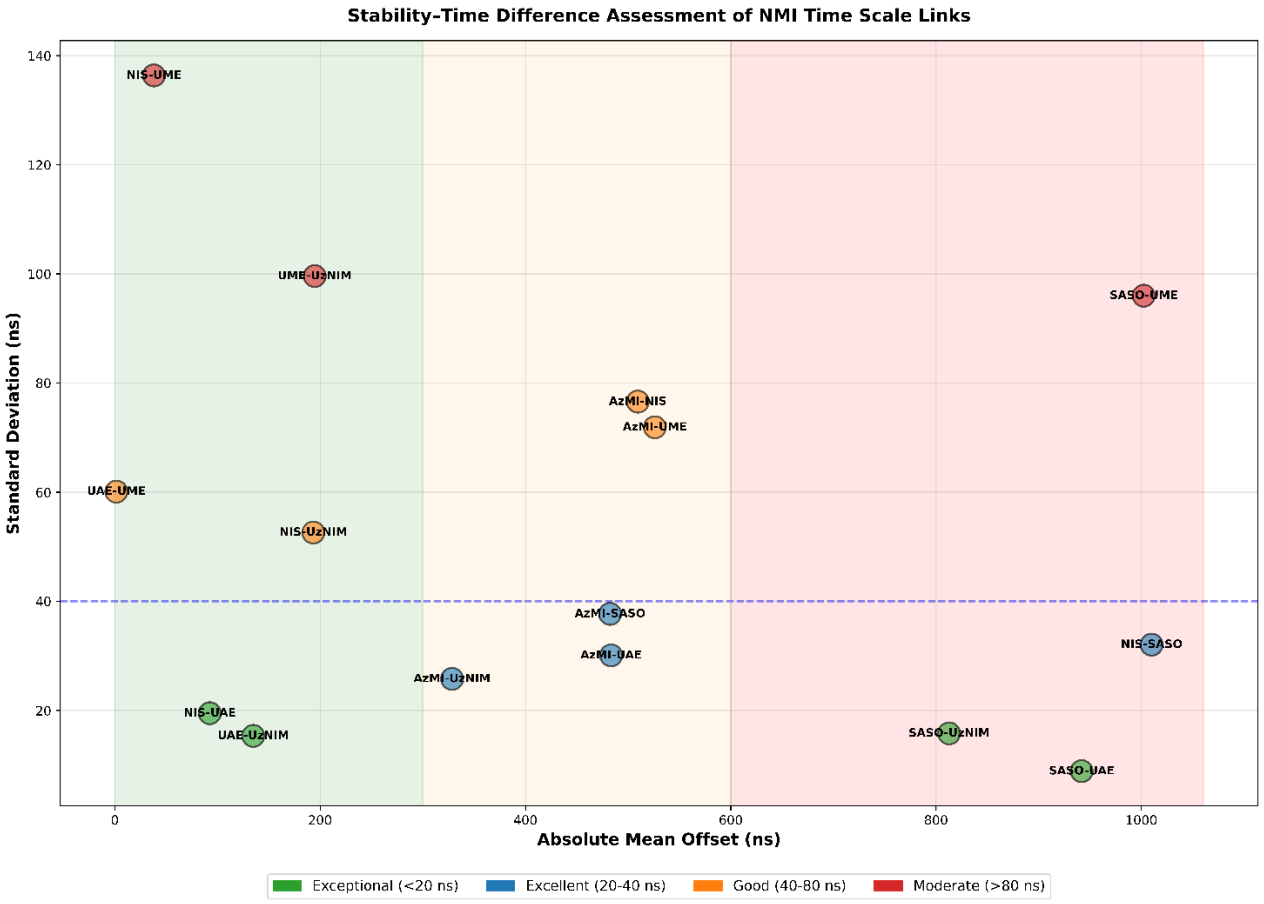

Figure 7. Stability-time difference assessment of NMI time scale links. Scatter plot displaying the relationship between absolute mean offset (x-axis, in nanoseconds) and standard deviation (y-axis, in nanoseconds) for all bilateral comparison links between participating NMIs. Each data point represents one link, labeled with the NMI pair designation. Background shading indicates performance regions: green (Exceptional, $\sigma < 20$ ns), blue (Excellent, 20-40 ns), orange (Good, 40-80 ns), and red (Moderate, $\sigma > 80$ ns). The horizontal dashed line at 40 ns marks the boundary between excellent and good performance tiers. Links positioned in the lower-left region demonstrate both high stability (low standard deviation) and good agreement (low mean offset), while those in the upper-right show higher variability and larger offsets. The plot enables simultaneous assessment of both short-term stability and systematic offset characteristics for each bilateral comparison.

Figure 7 summarizes the performance of all bilateral NMI time-scale links by jointly examining short-term dispersion and systematic alignment. Each point in the plot corresponds to one bilateral link and is positioned according to its standard deviation (vertical axis) and absolute mean time offset (horizontal axis), both derived from the full Common-View time-difference series.

The standard deviation represents the short-term variability of the link, reflecting the combined effects of clock noise and measurement uncertainty, while the mean offset quantifies the average alignment between the two time scales over the comparison interval. By displaying these two metrics simultaneously, the figure enables direct assessment of whether a given link combines good stability with good agreement.

Links located in the lower-left region of the plot exhibit both low dispersion and small mean offset, indicating favorable performance in terms of relative stability and alignment. Conversely, links in the upper regions show increased variability, while those positioned toward the right exhibit larger systematic offsets. The horizontal dashed line at 40 ns marks the boundary between the *Excellent* and *Good* stability categories and provides a visual reference for stability classification.

The distribution of points demonstrates that stability and agreement are not strictly correlated. Several links achieve low standard deviation despite relatively large mean offsets, indicating stable but systematically shifted time scales, while others exhibit small offsets but higher dispersion. This highlights the importance of considering both metrics together when evaluating bilateral comparison performance and motivates the subsequent analysis of drift behavior across all links.

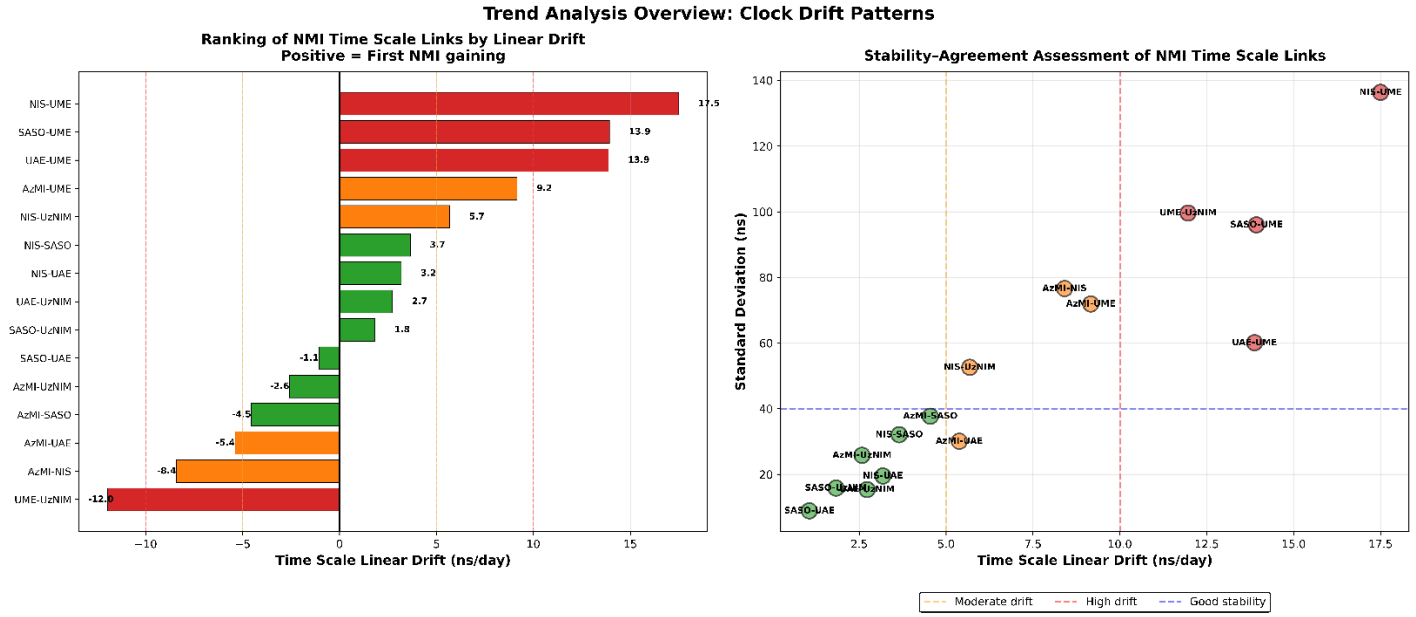

Figure 8. Trend analysis overview of clock drift patterns for all NMI time scale links. Two-panel display showing: **(Left panel)** Horizontal bar chart ranking all bilateral comparison links by linear daily drift rate (ns/day). Bars are colored according to drift magnitude: green (good stability, |drift| < 5 ns/day), orange (moderate drift, 5-10 ns/day), and red (high drift, |drift| > 10 ns/day). Positive values indicate the first NMI is gaining time relative to the second, while negative values indicate time loss. A vertical dashed line at zero marks no drift. **(Right panel)** Scatter plot showing the relationship between daily drift or absolute trend magnitude (x-axis, ns/day) and standard deviation (y-axis, ns). Each point represents one bilateral link, labeled with the NMI pair. Vertical and horizontal dashed lines at 10 ns/day and 40 ns respectively divide the plot into quadrants representing different combinations of drift and stability characteristics. Points are color-coded by drift category matching the left panel. The combined visualization enables assessment of both systematic drift trends and their relationship to measurement stability.

Figure 8 provides a consolidated view of linear drift behavior across all bilateral links. The left panel ranks links by daily drift magnitude (ns/day), while the right panel relates drift magnitude to standard deviation. This combined view enables identification of links dominated by systematic frequency differences versus those primarily affected by random noise.

Links exhibiting both large drift and large standard deviation correspond to reduced overall performance, while links with small drift and low scatter indicate good relative frequency agreement. The visualization confirms that drift and stability are related but independent characteristics, both of which must be evaluated to fully characterize time-scale link behavior.

Figure 9 ranks all bilateral NMI time-scale links according to the Allan deviation evaluated at averaging time $\tau = 1$ day, with lower values indicating better short-term stability. In this context, the Allan deviation quantifies the day-to-day wander of the relative time offset between two time scales and provides a stability metric that is largely insensitive to constant offsets. The ranking is consistent with the dispersion-based classification presented earlier, with links exhibiting low standard deviation also showing low ADEV values. Links with higher ADEV correspond to increased short-term variability, confirming that Allan deviation provides a complementary and robust measure of bilateral link stability.

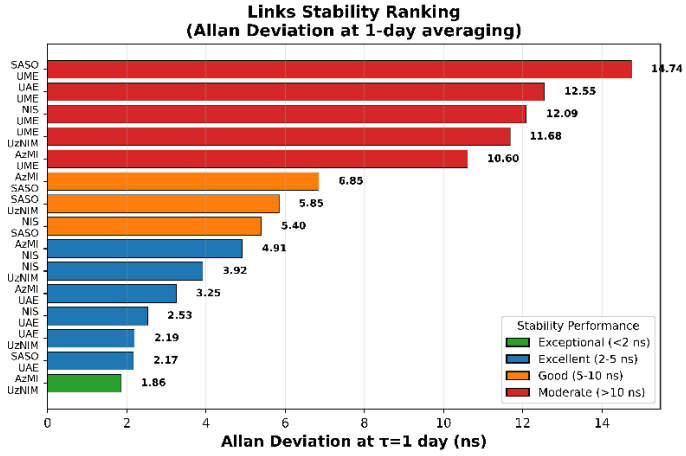

Figure 9. Allan deviation ranking of bilateral NMI links at τ=1 day averaging time. All 15 comparison links ranked by stability (lower is better).

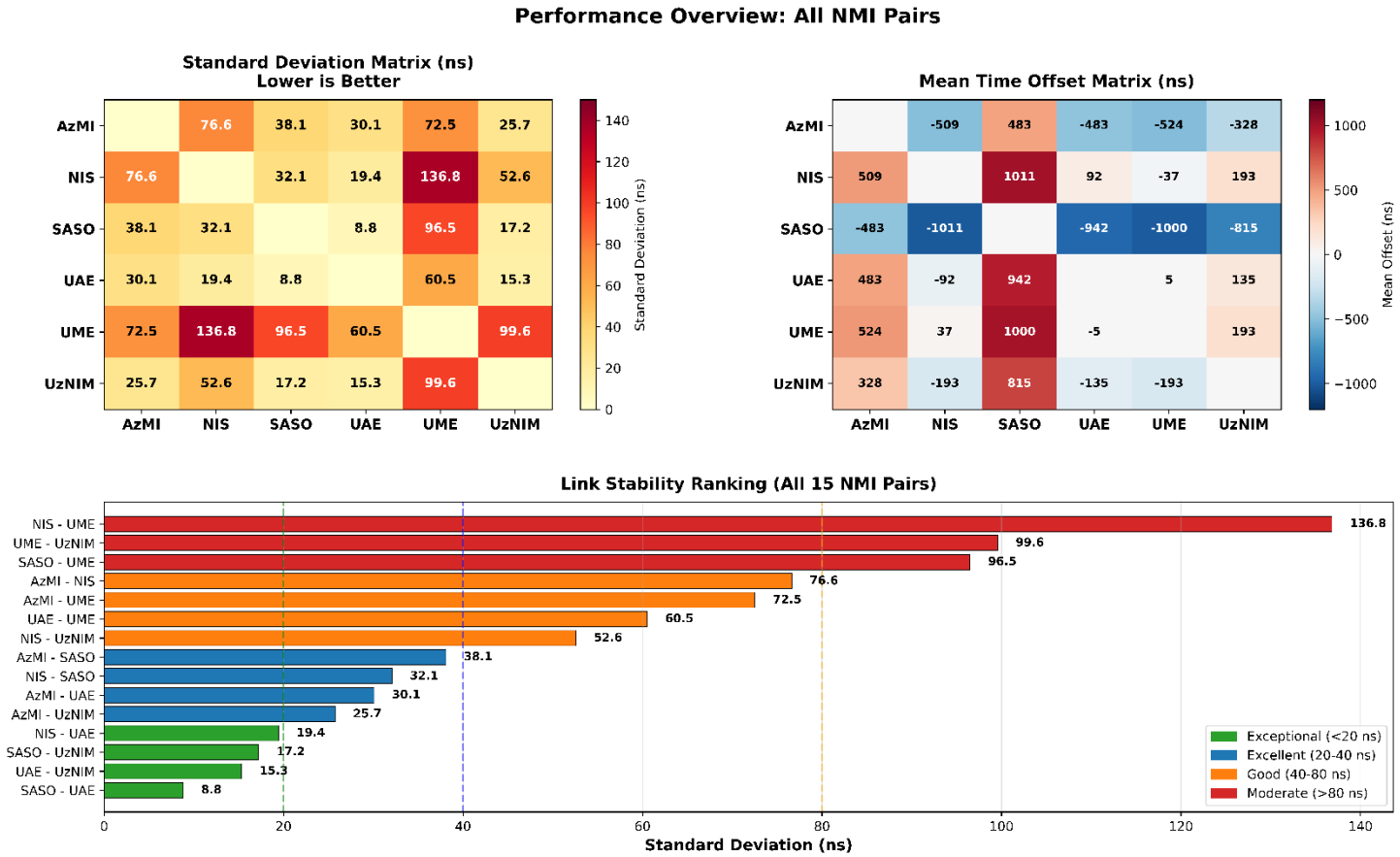

Figure 10. Performance overview of all NMI bilateral comparison pairs**.** Three-panel display showing: **(Top left)** Standard deviation matrix displaying the measurement variability (ns) for all 15 bilateral links. Matrix cells are color-coded from yellow (low standard deviation, better stability) to dark red (high standard deviation, poorer stability). The matrix is symmetric with values representing the short-term stability of each NMI pair comparison. **(Top right)** Mean time offset matrix showing the average time difference (ns) between each NMI pair. Color coding ranges from dark blue (large negative offset) through white (near-zero offset) to dark red (large positive offset), with values closer to zero indicating better agreement between time scales. **(Bottom)** Horizontal bar chart ranking all 15 NMI pairs by standard deviation, from lowest (best) to highest (worst). Bars are color-coded by performance tier: green (Exceptional, $\sigma < 20$ ns), blue (Excellent, 20-40 ns), orange (Good, 40-80 ns), and red (Moderate, $\sigma > 80$ ns). A vertical dashed line at 40 ns marks the boundary between excellent and good performance. The combined visualization enables comprehensive assessment of both stability and systematic offsets across all bilateral comparisons.

Figure 10 provides a consolidated overview of the performance of all bilateral NMI time-scale links by combining stability and alignment information into a single set of complementary visualizations. The standard deviation matrix summarizes short-term variability across all link pairs, while the mean offset matrix highlights systematic time-scale alignment. The accompanying ranking further orders all links according to their dispersion, enabling direct comparison of relative performance.

Together, these representations confirm that link quality varies significantly across the comparison network and that stability and offset are largely independent characteristics. Several links exhibit good stability despite relatively large systematic offsets, while others show increased variability even when mean offsets are small. By presenting all bilateral results in a unified framework, Figure 10 enables rapid identification of both well-performing and degraded links and provides a compact summary of the overall comparison quality.

4. CONCLUSION

This comparison demonstrates the effectiveness of a structured, multi-layered methodology for evaluating national time scale performance using GNSS Common-View techniques within a regional metrology framework. By combining UTCr-based profiling, single-reference bilateral analysis using UTC(SASO), and a fully reference-independent multi-link assessment, the study provides a comprehensive and internally consistent characterization of inter-laboratory time-scale behavior.

A key strength of the adopted approach lies in its progressive analysis hierarchy. The use of UTCr enables direct benchmarking against the global UTC ensemble, separating absolute alignment and frequency behavior from local effects. The subsequent UTC(SASO)-referenced bilateral analysis ensures continuity across all participating NMIs, including laboratories without rapid UTC availability, while preserving direct pairwise interpretability. Finally, the multi-link analysis removes dependence on any single reference and treats each bilateral link as an independent realization, allowing systematic patterns in stability, drift, and offset to be identified across the entire comparison network.

Methodologically, the analysis is grounded on standardized CGGTTS data processing using the BIPM-provided CGGTTS Analyser, ensuring consistency with established UTC comparison practices. This was complemented by custom Python-based analysis workflows, leveraging open-source scientific libraries for data screening, weighted REFSYS construction, statistical evaluation, and advanced visualization. The combined use of standardized tools and flexible open-source processing ensures transparency, reproducibility, and traceability of all analysis steps. The application of weighted REFSYS construction, strict data screening, and normalization of time-difference series emphasizes intrinsic clock and link behavior rather than artifacts of offset magnitude. The classification of links by standard deviation, combined with complementary metrics such as linear drift and Allan deviation at $\tau = 1$ day, enables clear separation between short-term dispersion, long-term frequency offset, and systematic alignment. Importantly, the results show that stability, drift, and mean offset are largely independent characteristics, underscoring the necessity of multi-metric evaluation rather than reliance on a single performance indicator.

From a practical perspective, the applied framework goes beyond a conventional pass–fail comparison and instead provides diagnostic insight into link quality, clock behavior, and measurement performance. This is particularly valuable for developing laboratories, as it supports targeted corrective actions, informed system improvements, and readiness assessment for future participation in formal UTC comparisons such as CCTF-K001-UTC.

Overall, this work demonstrates that a carefully designed multi-link GNSS Common-View analysis, supported by robust statistical profiling, offers a transparent, scalable, and reproducible methodology for regional time-scale comparisons. The approach is well suited not only for GULFMET activities but also as a general model for RMO-level studies, contributing to stronger regional integration and enhanced global coherence of UTC [14].

REFERENCES

[1] Allan, D.W. and Weiss, M.A. (1980) ‘Accurate Time and Frequency Transfer During Common-View of a GPS Satellite’, in *Proceedings of the 34th Annual Symposium on Frequency Control*. Philadelphia, PA, USA, 28–30 May 1980. Piscataway, NJ: IEEE, pp. 334–346. doi: 10.1109/FREQ.1980.200424.

[2] Gamidov (Hamid), R. and Çetintaş, M. (1999) 'Time Comparison of atomic clocks using counter and GPS system', Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and the IEEE International Frequency Control Symposium, 1, pp. 228-270.

[3] Alassaf, A., Al Harbi, W.M., Al Dawood, K.S. and Hamid, R. (2026) 'Monte Carlo Analysis of Gate-Time-Resolved Uncertainty and Oscillator Noise in Frequency Measurements'. (To be published).

[4] Galliana, F., Capra, P.P., Cerri, R., D’Elia, V., Gasparotto, E., Roncaglione Tet, L., Del Moro, M. and Cozzani, A. (2017) ‘Inter-laboratory measurement comparison between INRIM and ESA on electrical quantities’, in *18th International Congress of Metrology*. EDP Sciences, 07005. doi: 10.1051/metrology/201707005.

[5] Chen, R., Fan, D., Zhao, Z., Meng, L. and Liu, Y. (2024) ‘Research on Common-View Comparison Method Based on GNSS Satellite’, in *2024 IEEE 7th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. Chongqing, China, 20–22 September 2024. Piscataway, NJ: IEEE, pp. 746–750. doi: 10.1109/ITNEC60942.2024.10732942.

[6] National Institute of Standards and Technology (2024) *Common View GNSS Time Transfer*. Available at: https://www.nist.gov/pml/time-and-frequency-division/time-services/common-view-gnss-time-transfer (Accessed: 6 January 2026).

[7] Enzer, D.G., Murphy, D.W. and Burt, E.A. (2021) ‘Allan Deviation of Atomic Clock Frequency Corrections: A New Diagnostic Tool for Characterizing Clock Disturbances’, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 68(7), pp. 2590–2604 . doi: 10.1109/TUFFC.2021.3061005.

[8] Allan, D.W., Ashby, N. and Hodge, C.C. (1997) *The Science of Timekeeping*. Application Note 1289. Santa Clara, CA: Agilent Technologies

[9] Abbasov, S., Gedik, A., Aldawood, K. H., & Hamid, R. (2025, November 11–13). Multi-region comparisons of national time scales using the common view method and NTP server. XI International Competition “Best Young Metrologist of COOMET 2025”, Uzbek National Institute of Metrology, Tashkent, Uzbekistan.

[10] National Institute of Standards and Technology (2024) *Common View GNSS Time Transfer*, NIST Time and Frequency Division. Available at: https://www.nist.gov/pml/time-and-frequency-division/time-services/common-view-gnss-time-transfer (Accessed: 21 January 2026).

[11] Arias, F., Jiang, Z., Lewandowski, W. & Petit, G. (2005) *BIPM comparison of time transfer techniques*. Proceedings of the 2005 IEEE International Frequency Control Symposium. IEEE, pp. 4. https://doi.org/10.1109/FREQ.2005.1573951

[12] Senior, K., Koppang, P. & Matsakis, D. (2016) ‘GNSS-based time transfer’, in *Springer Handbook of Global Navigation Satellite Systems*.

[13] G. G. Hamza, “Mitigating the effect of multipath on the stability of time transfer using GNSS,” preprint, 2022.

[14] BIPM (2025) *UTCr: rapid realization of Coordinated Universal Time (UTC)*. Bureau International des Poids et Mesures. Available at: https://www.bipm.org/en/time-ftp/utcr (Accessed: 21 January 2026).