3GPP TSG-RAN WG4 Meeting #116 R4-2509098

**Bengaluru, India, 25th – 29th August, 2025**

**Agenda item:** 7.17.1

**Source:** Moderator (Qualcomm)

**Title:** Topic summary for [116][131] NR\_AIML\_air\_part1

**Document for:** Information

# Introduction

This is the summary thread for issues related to the NR AI/ML WI. The topics handled in this thread are the following:

* CSI reporting requirement and testing framework for CSI prediction
* RRM core requirement and testing framework for beam management
* RRM core requirement and testing framework for Positioning accuracy enhancement

Issues related to the general part (agenda item 17.17.2) are treated in the AI/ML part 2 thread ([116][132])

A WF summarizing the agreements from RAN4#115 was agreed in RP-2508080. The discussion will take into account these previous agreements.

# Topic #1: CSI reporting requirement and testing framework for CSI prediction

## Companies’ contributions summary

|  |  |  |
| --- | --- | --- |
| **T-doc number** | **Company** | **Proposals / Observations** |
| [**R4-2509133**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509133.zip) | Anritsu Corporation | **Observation 1: It is possible to schedule PDSCH at slot#(n+2) based on the received PMI report from TE point of view.**  **Observation 2: It is possible to reduce time separation between PMI report and scheduled PDSCH and change from “before slot#(n+4)” to “before slot#(n+3)” from test equipment viewpoint.** |
| [**R4-2509411**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509411.zip) | Samsung | **Observation 1: For performance monitoring metric report, SGCS will be used and is quantized with 4-bit. For each report, two SGCS are included to reflect the Quantity reported.**  **Proposal 1: RAN4 should discuss the testability of performance monitoring accuracy with SGCS based on RAN1 definition for CSI prediction**  **Proposal 2: Activation delay requirements created for semi-persistent and aperiodic CSI reporting for CSI prediction**  **Observation 3: NW can use the ID of interference report configured in the configuration for monitoring to differentiate the predicated CSI reporting and the measured CSI reporting.**  **Observation 4: The existing performance monitoring report can be used to differentiate the reports before and after the deactivation**  **Proposal 3: RAN4 should clarify how the generalization testing will be conducted, whether it will be tested in one case with multiple configurations combination, or it will be tested in multiple cases with individual configurations**  **Proposal 4: If one test case is conducted with multiple test configurations, one case per UE capability should be enough. If one test case is conducted with individual test configuration, multiple cases with different configurations can be considered.**  **Proposal 5: If non-static scenario/configuration introduced, RAN4 should discuss how to perform the test with different parameters during one test setup.**  **Proposal 6: There is no need the generalization testing related with indoor/outdoor scenario.**  **Proposal 7: There is no need the generalization testing related with channel bandwidth**  **Observation 5: Rel-19 SCM channel model study is expected to only focus on limited scenarios. The testability and test repeatability for CDL channel model on demodulation and testing should be further studied**  **Proposal 8: Prioritize to use TDL channel model for AI/ML based CSI prediction for generalization testing in Rel-19. Further study the performance impact on AI/ML based CSI prediction with spatial channel model (e.g., CDL based channel model, or TDL based extended channel model) in Rel-20 pending on the progress of SCM channel model study in Rel-19 and the specific channel model in Rel-20.**  **Proposal 9: RAN4 should generate the training data with the same gNB array parameters (NW-side) specified for test. For different gNB array parameters, RAN4 may need to define the multiple cases.**  **Observation 6: AI/ML based CSI prediction shows better performance compared with the nearest historical CSI scheme, while generalization problem related different doppler needs further investigated**  **Observation 7: With including the data from different doppler values for AI/ML model training, good generalization can be achieved for different doppler scenario, although there is a performance degradation**  **Proposal 10: RAN4 can define requirements with considering different scenarios to guarantee the generalization of AL/ML based CSI prediction.**  **Proposal 11: RAN4 can consider to define different MCS requirement targeting different SNR range. The specific SNR range associated with MCS for generalization test should be further discussed. Another alternative solution is to check the SNR value associated with X% of maximum throughput, beside the 90%, for example, X can be considered 30% or 70%.** |
| [**R4-2509424**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509424.zip) | Apple | **Proposal 1: RAN4 is invited to consider curriculum training for AI/ML-based CSI prediction to improve simulation alignment across SNR ranges. Instead of training on a wide ±X dB range from the start, the model is first exposed to high-SNR samples where CSI structure is clean, and then progressively trained on lower-SNR data. This staged approach enhances model stability, improves learning of temporal patterns, and better reflects the signal conditions encountered under practical channel estimation.**  **Proposal 2: RAN4 is invited to study the generalization behavior of AI/ML-based CSI prediction models under practical deployment variations, including (i) differences in gNB-specific CSI port virtualization, (ii) Doppler spread due to UE mobility, and (iii) prediction window length (Δ). These factors affect the structure and dynamics of the CSI input and may degrade model performance when training and deployment conditions differ. The degree of generalization achievable across these axes will directly influence the number of test cases needed for performance verification, with better generalization allowing a more streamlined test framework.** |
| [**R4-2510161**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510161.zip) | CMCC | ***Proposal 1: for CSI prediction, it is proposed to follow RAN1 to use SGCS as metric for performance monitoring.***  ***Proposal 2: it is proposed to define report mapping for SGCS based on RAN1 agreements.***  ***Proposal 3: it is proposed to define monitoring delay for CSI prediction performance monitoring.*** |
| [**R4-2510336**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510336.zip) | vivo | Proposal 1: RAN4 to define delay requirement and accuracy requirement type 3 performance monitoring for CSI prediction. |
| [**R4-2510671**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510671.zip) | ROHDE & SCHWARZ | **Proposal: RAN4 considers “This reported PMI cannot be applied at the gNB downlink before slot#(n+3)” as feasible. Further reduction in PMI delay shall not be considered in Rel-19.** |
| [**R4-2510805**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510805.zip) | OPPO | **Proposal 1: For AI-based CSI prediction performance:**  **- When compared to R16 eTypeII using 'sample and hold' reporting, it should not be worse than R16 eTypeII.**  **- When compared to R18 eTypeII codebook for predicted PMI, it should not be worse than R18 eTypeII codebook.**  **Proposal 2: Reuse the legacy PMI requirement (γ=1.8 for Enhanced Type II codebook for predicted PMI @ 20Hz Doppler spread), FFS γ value for 50Hz Doppler spread.**  **Minimum requirement for AI/ML based CSI prediction**   |  |  |  | | --- | --- | --- | | **Parameter** | **Test 1 (20Hz Doppler spread)** | **Test 2 (50Hz Doppler spread)** | |  | **1.8** | **TBD** |   **Proposal 3: could be obtained within a CSI prediction window. If a CSI prediction window contains more than one CSI prediction instance,**  **Option1: is calculated for each predicted CSI and then do the average within the CSI prediction window**  **Option2: and are calculated within the CSI prediction window first and then get the averaged value.**  **Proposal 4: For CSI prediction, following test steps could be considered:**   * **Step1: TE sends CSI-RS to DUT, within a CSI observation window (signal generating and transmitting)** * **Step2: DUT conducts the CSI measurement within the CSI observation window (CSI-RS receiving)** * **Step3: DUT conducts the CSI prediction within a CSI prediction window (Inference)** * **Step4: DUT reports the predicted CSI (CSI reporting)** * **Step5: TE calculates the for performance [and calculates the measurement/inference latency for LCM]** |
| [**R4-2510835**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510835.zip) | Ericsson | Observations:  [Observation 1: CSI-PAI SGCS1 and SGCS2 are close for TDLA30 with 20Hz scenario. Ratio of SGCS1 and SGCS2 > 1.06 for 90% of test time.](#_Toc206186677)  [Observation 2: CSI-PAI SGCS1 is higher compared with SGCS2 for TDLA30 with 50Hz scenario. Ratio of SGCS1 and SGCS2 > 1.52 for 90% of test time.](#_Toc206186678)  Proposals:  [Proposal 1: During the AI/ML-based predicted PMI reporting test, TE also collects SGCS1 and SGCS2 using CSI-PAI reporting.](#_Toc206186679)  [Proposal 2: Study the feasibility to set the CSI-PAI reporting requirements with that the SGCS1 should be more than X in [90]% of the test time, e.g., X=0.85.](#_Toc206186680)  [Proposal 3: Study the feasibility to set the CSI-PAI reporting requirements with that the radio SGCS1/SGCS2 should be more than Y in [90]% of the test time, e.g., Y=1.0.](#_Toc206186681) |
| [**R4-2510872**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510872.zip) | Huawei,HiSilicon | **Proposal 1**: CDL channel model is not used in RAN4 for AI CSI prediction.  **Proposal 2:** For generalization testing, there is no need to use non-static condition.  **Proposal 3:** RAN4 will start to discuss whether/how to define requirements for UE-side monitoring if other WGs achieve sufficient progress.  **Proposal 4:** RAN4 will focus on NW-side monitoring for AI CSI prediction and legacy measurements reporting requirements are reused. |
| [**R4-2511221**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511221.zip) | Nokia | 1. Reporting delay of AIML CSI prediction should be defined as the time delay from the RRC (re)configuration complete message, including AIML functionality activation and monitoring resource configuration, to the transmission of the first monitoring report. 2. RAN4 should define the reporting delay requirements for the first monitoring report based on the agreed framework from beam management use cases. 3. The accuracy requirements for CSI prediction performance monitoring can be defined specifically for UE-reported SGCS1 values. 4. RAN4 to define accuracy requirements for CSI prediction performance monitoring, specifically for UE-reported SGCS1 values.   The complementary use of SGCS1 and SGCS2 enables the testing framework to assess AI/ML-based CSI prediction performance under both static and dynamic channel conditions, even without visibility into the DUT’s internal model.   1. RAN4 should study the feasibility of defining a standardized testing framework for performance monitoring of AI/ML-based CSI prediction using SGCS1 and SGCS2 values. 2. Delay requirements for CSI prediction performance monitoring should use the approach established in the BM use case as a baseline. |
| [**R4-2511568**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511568.zip) | Qualcomm Incorporated | **Observation 1: RAN4 is currently considering 4ms time separation between PUSCH containing predicted PMI and the last CSI-RS occasion of the observation window in tests.**  **Observation 2: The 4ms time separation is aligned with RAN1 spec defined minimum time separation between CSI-RS reference resource and SP reporting of non-AI-ML based CSI report.**  **Observation 3: 4ms time separation may not be sufficient to process CSI-RS, run inference and generate reports containing AI-ML based predicted PMI.**  **Observation 4: RAN1 will discuss potential extension of 4ms time separation in RAN1 122, i.e., during August 2025.**  **Proposal 1: RAN4 extends the time separation between the last CSI-RS occasion of the observation window and the PUSCH slot containing SP configured predicted PMI report from 4ms to 4+X ms.**   * **Note: (4+X) ms will be aligned with the RAN1 spec defined minimum time separation between CSI-RS reference and corresponding AI-ML based predicted PMI in SP reporting.** |
| [**R4-2509412**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509412.zip) | Samsung | **Proposal 1: Existing test configuration with aperiodic CSI-RS configuration can be considered as starting point for AI/ML based CSI prediction.**  **Proposal 2: Focus on periodic CSI-RS configuration for AI/ML based CSI prediction firstly, after the alignment work of periodic CSI-RS Configuration is stable, RAN4 can proceed the related alignment and requirement discussion for Aperiodic CSI-RS configuration**  **Proposal 3: RAN4 can discuss whether additional SGCS in step 3 is need to be output, where the SGCS is calculated by comparing the SVD of model output (before feedback) with the SVD of the measured CSI.**  **Proposal 4: Beside MCS 13, RAN4 can consider MCS 4 as QPSK, MCS 17 or MCS 19 for 64QAM can be considered as starting point to check the feasibility of different MCS with AI/ML based CSI prediction**  **Observation 1: With mix dataset training, the performance has minor performance degradation. Good generalization performance can be achieved with mix dataset training**  **Observation 2: Compared with Rel-15 Type I random PMI, both AI and non-AI schemes can achieve large performance gain for different doppler scenario. Compared with non-AI scheme Rel-18 eType II Doppler with K=1, the gain of AI scheme is minor, especially under the scenario with doppler as 50Hz**  **Observation 3: With doppler increasing, all the schemes show the performance degradation**  **Proposal 3: RAN4 can consider the intermediate SGCS under the practical channel estimation for AI/ML based CSI prediction.**  **Observation 4: With the AI/ML model training under high SNR condition, better performance can be achieved compared with the AI/ML model training with the test SNR**  **Proposal 4: RAN4 should further check the generalization related with SNR for AI/ML based CSI prediction**  **Proposal 5: RAN4 should clarify the SNR range of dataset for AI/ML model training if the final results among companies have large gap.** |
| [**R4-2509428**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509428.zip) | Apple | 1. RAN4 is invited to focus on **static test conditions** for CSI prediction performance evaluation in Rel-19, while explicitly defining **multiple test cases** to reflect the dependency of AI/ML model performance on deployment scenarios. These test cases should vary in terms of key parameters such as **antenna configurations**, **UE speed**, **observation window length**, and **prediction window length**. 2. RAN4 is invited to define a Type 3 performance monitoring framework for AI/ML-based CSI prediction, using SGCS or a similar metric to assess model prediction accuracy at runtime. The framework should specify how the metric is computed, the format of the reported results (e.g., CDF-based thresholds or per-instance values), the reporting granularity (such as wideband or subband level), and the conditions under which monitoring results should trigger model fallback or AI/ML model reconfiguration. |
| [**R4-2509655**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509655.zip) | MediaTek inc. | **Observation #1: FDD 16TX and FDD 32TX results with the same Doppler perform quite similarly.**  **Observation #2: In lower Doppler accuracy is better than in higher Doppler.**  **Observation #3: Rel-18 codebook quantization performance is close to raw channel prediction performance.**  **Observation #4: In simulation assumptions there exists sufficient test setup for FDD with Periodic CSI-RS configuration.**  **Observation #4: Prediction gain measured in SNR varies between 1.6 to 2.0dB in Doppler 20Hz.**  **Observation #5: Prediction gain measured in γ varies between 0.6 to 1.3 (27% to 48%) in Doppler 20Hz.**  **Observation #6: There is no prediction gain in Doppler 50Hz.**  **Observation #7: Operation point of MCS13 in Doppler 20Hz is close to SNR 0dB.**  **Proposal #1: Exclude Doppler 50Hz in further studies.**  **Proposal #2: Focus on Doppler 20Hz in further studies.**  **Proposal #3: Focus on MCS20 or higher for robust testing SNR.**  **Proposal #4: Focus on FDD to find feasible test configuration before using effort on more complex TDD configurations.** |
| [**R4-2510162**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510162.zip) | CMCC | ***Proposal 1: for CSI prediction test with static condition, it is proposed to define test cases with different configuration/parameters, e.g. antenna configuration, observation window length, prediction window length, and UE speed .***  ***Proposal 2: For channel model for CSI prediction tests, it is proposed to focus on TDL channel models in Rel-19. Using CDL channel model for CSI prediction test can be further dicsussed in later release based on the progress on SCM.*** |
| [**R4-2510340**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510340.zip) | vivo | **Observation 1: Initial evaluation results for CSI prediction AI/ML model is shown in Table 2-1, for FDD and 32gNB Tx.**  Table 2-1. Initial evaluation results for CSI prediction AI/ML model, for FDD and 32gNB Tx.   |  |  |  |  | | --- | --- | --- | --- | | FDD, 32 gNB Tx | Test on 20Hz Doppler spread | Test on 50Hz Doppler spread | Test on 100Hz Doppler spread | | Model trained by 20Hz Doppler spread | 0.993 | / | / | | Model trained by 50Hz Doppler spread | / | 0.902 | / | | Model trained by 100Hz Doppler spread | / | / | 0.298 | | Model trained by 20Hz, 50Hz and 100Hz Doppler spread | 0.994 | 0.907 | 0.296 |   **Observation 1: Initial evaluation results for CSI prediction AI/ML model is shown in Table 2-1, for FDD and 32gNB Tx.** |
| [**R4-2510809**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510809.zip) | OPPO | **Proposal 1: Regarding the RAN4 Simulation for CSI prediction and the Reference model for CSI prediction, RAN4 should first assess the CSI prediction evaluation results across different companies before determining whether a reference model is necessary.**  **Proposal 2: Regarding the RAN4 Simulation for CSI prediction and the Reference model for CSI prediction,**   * + - **If RAN4 achieves a relatively aligned CSI prediction evaluation results in step-2, further evaluation (step-3 comparisons) can proceed (during/after RAN4#116).**     - **If alignment in step-2 proves difficult, RAN4 should identify the key factors causing discrepancies and may need to initiate following evaluations to reassess step-2 alignment (during/after RAN4#116).**     - **If alignment remains unachievable after step-3, RAN4 may then trigger discussions and evaluations on a reference model for CSI prediction (e.g., during/after RAN4#116bis).**     - **If alignment achieves after step-3, RAN4 will trigger evaluations on the CSI reporting requirements for CSI prediction(e.g., during/after RAN4#116bis).**   **Observation 1: Regarding the updated simulation for AI/ML-based CSI prediction, the updated simulation results are shown in Table 2, Table 3 and Table 4.**  **Observation 2: For step-2, the SGCS calculated by comparing the feedbacked SVD of the model output with the SVD of the ground truth raw channel deteriorates significantly. The SGCS calculated by comparing the feedbacked SVD of the model output with the feedbacked SVD of the ground truth raw channel shows much less degradation.**  **Observation 3: For step-3, under the condition of TX=16 and rank=2, the throughput performance obtained by randomly selecting a Type I codebook is relatively poor. The key difference between these schemes (e.g. from different companies) will primarily be reflected in their γ values relative to the baseline benchmark.**  **Proposal 3: Consistency between training data and test data should be ensured**   * + - **FFS: Whether the test emulated channel will be similar to the one used in RAN4 simulations**     - **FFS: Whether anything special should be done during the test** |
| [**R4-2510836**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510836.zip) | Ericsson | Observations:  [Observation 1: Average SGCS does not change whether the AL/ML-based CSI prediction model uses matched training dataset or mixed training dataset, at least for TDLA30-20/50.](#_Toc206187629)  [Observation 2: Degradation of SGCS after the quantization (SGCS-3) is not significant compared with SGCS-1 (before quantization).](#_Toc206187630)  [Observation 3: AI/ML-based PMI prediction shows slightly better performance compared with Rel-18 with the same test setup.](#_Toc206187631)  [Observation 4: The performance gain of AI/ML-based PMI prediction over Rel-18 PMI prediction is larger for 16Tx compared with 32Tx.](#_Toc206187632)  [Observation 5: RAN1 is still discussing the CSI processing time for Rel-19 AI/ML-based PMI prediction.](#_Toc206187633)  Proposals:  [Proposal 1: RAN4 evaluates further PDSCH demodulation performance with AI/ML-based PMI.](#_Toc206187634)  [Proposal 2: RAN4 continues the AI/ML-based PMI prediction performance evaluation based on the CSI processing time assumption with t=0. RAN4 should revisit the scheduling when RAN1 concludes the UE capability of reporting delay t.](#_Toc206187635) |
| [**R4-2510876**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510876.zip) | Huawei,HiSilicon | **Proposal 1**: For step-1, SGCS performance results of AI CSI prediction are provided in Table 1.  Table 1. SGCS performance results of AI CSI prediction (step-1)   |  |  |  | | --- | --- | --- | |  | 20Hz | 50Hz | | Rank 1 | 0.9985413 | 0.8673333 | | Rank 2 | 0.99800384 | 0.8166052 | |
| [**R4-2510994**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510994.zip) | Nokia | [Proposal 1: Concerning the time separation between PMI report and scheduled PDSCH, RAN4 to keep the same value as in the previously agreed simulation assumption (i.e., slot#(n+4)) and not to reduce it (to slot#(n+3)).](#_Toc206177876)  [Proposal 2: Concerning the SNR handling for data collection for AI/ML training, RAN4 to first discuss whether fixing the SNR within only a certain range would not result in overfitting.](#_Toc206177877)  [Proposal 3: If RAN4 agrees to fix the SNR range, the SNR range should cover both the SNRs of 90% and 70% throughput.](#_Toc206177878)  [**Observation 1:** From the simulation results collection in RAN4#115, for CSI prediction step2, SGCS2 and SGCS3 show different results.](#_Toc206177879)  [**Observation 2:** As SGCS3 considers the same codebook for both the model and the ground truth, it is a more appropriate measure of similarity compared to SGCS2 which only considers the codebook only for the model.](#_Toc206177880)  [Proposal 4: RAN4 to prioritize SGCS3 over SGCS2 in the evaluation.](#_Toc206177881) |
| [**R4-2511573**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511573.zip) | Qualcomm Incorporated | **Observation 1: Both SGCS2 and SGC3 of Figure 1 are reported as SGCS after CSI feedback in 2nd step.**  **Observation 2: For 25Hz and 50Hz Doppler frequencies, the SGCS obtained with AI-ML based CSI prediction in both step 1 and step 2 are significantly higher than those obtained with “sample and hold” approach.**  **Observation 3: The SGCS performances of AI-ML model, trained based on the mixed dataset of 20, 50 and 100Hz, are almost identical to those of AI-ML model, trained based on individual datasets.**  **Observation 4: The number of parameters of our AI-ML CSI prediction model, that got used to generate the results of this contribution, is in the order of several millions.** |

## Open issues summary

The agreements from the previous meetings on this topic are listed below for reference:

Agreements in RAN4#114bis(R4-2505105)

**Issue 1-1: Codebook to be used in the tests**

Agreement:

* In Rel-19 only PMI requirements based on AI/ML CSI prediction will be defined
* use Rel-18 eType II Doppler codebook for AI/ML based CSI predictions

**Issue 1-2: Requirement baseline**

Agreement:

* + Requirement to be introduced for Dopler domain basis N4=1
    - * N4>1 can be further discussed in the future if needed

**Issue 1-4: Test metric**

Requirements to be defined based on the ratio(same metric as for the current PMI reporting requirements in 38.101-4)

* + - 
      * tue\_AI/ML is 90 % of the maximum throughput obtained at SNRue\_AI/ML  ~~[a given SNR]~~using the precoders configured according to the UE reports,
      * trnd is the throughput measured at SNRue\_AI/ML  ~~[the same SNR]~~ with random precoding.
        + trnd is obtained with Rel-15 type I single panel codebook
      * Note: SNRue\_AI/ML is calculated in the same way as SNRue and SNRfollow1,follow2 in section 6.3 of 38.101-4

Also, simulation assumptions were agreed in R4-2502970.

Agreements in RAN4#115(R4-2508080)

**Issue 1-1: Requirement considerations**

Agreement:

The performance of AI based CSI prediction should not be worse than R16 etypeII.

* The exact test procedure and configurations are FFS
* The impact of different timeline between AI and R16 etype II should be also carefully analysed

**Issue 1-3: Doppler handling**

Agreement:

* + De-prioritize 100Hz Doppler in R19
    - Consider 100Hz Doppler in the future release if needed

**Issue 1-5: SNR levels for training data and generalization**

Agreement:

No need to capture anything explicit in RAN4 about training data in the spec,

* For simulation calibration, SNR range alignment for training can be discussed separately

UE performance at more than one MCS should be tested

* Model generalization can be discussed separately
* Each individual test is expected to involve in single MCS

The open issues were grouped in the following sub-topics for further discussion:

* + - 1. Performance monitoring requirements introduction
      2. Requirements for performance monitoring
      3. CSI prediction activation delay
      4. Reporting delay requirements
      5. Scheduling delay in tests
      6. Generalization
      7. Doppler values
      8. Simulation results and next steps

### Sub-topic 1-1

*Performance monitoring*

Several companies proposed to introduce Type 3 performance with SGCS reporting

**Issue 1-1: Performance monitoring**

* Proposals
  + Option 1: Introduce requirements for Type 3 performance monitoring
  + Option 2: other performance monitoring framework/requirement
  + Option 3: do not introduce any monitoring requirements
* Recommended WF

Option 1

### Sub-topic 1-2

*Requirements for monitoring*

A framework for the monitoring requirements should be discussed and agreed

**Issue 1-2: Requirement baseline**

* Proposals
  + Option 1: Introduce reporting delay, accuracy and reporting mapping requirements
    - Reporting delay should follow the same framework as for beam management use case
  + Delay to be defined as the period from the time when UE sends RRCReconfigurationComplete message in response to the configuration of monitoring RS resources via RRCReconfiguration, to the time when UE reports the first performance monitoring metric.
  + Option 2: Others
* Recommended WF
  + Option 1, discuss if the proposed delay definition can be agreed

### Sub-topic 1-3

*CSI Prediction activation delay*

**Issue 1-3: Activation delay for CSI Prediction**

|  |
| --- |
| **Agreement from General Aspect (BM use case) RAN4#115:**  **Activation delay:**   * **For semi-persistent CSI reporting** * **Activation delay starts at the reception of the MAC-CE/DCI** * **For aperiodic CSI reporting** * **Activation delay starts at the reception of the DCI** |

* Proposals
  + Option 1: Introduce activation delay for CSI prediction for both periodic and aperiodic (or semi-persistent) reporting

Activation delay starting time

* + For period CSI reporting, activation delay starts when UE sends RRCReconfigurationComplete message containing applicable functionality report.
  + For aperiodic CSI reporting, activation delay starts at the reception of the DCI.
  + For semi-persistent CSI reporting, activation delay starts at the reception of the MAC-CE/DCI.

Activation delay ending time

* + RAN4 to consider first inference report as the end point to define activation delay requirements for periodic/aperiodic/semi-persistent CSI reporting for CSI prediction use cases.
  + Delay value TBD, to be discussed/agreed in next meeting
  + Option 2: No need for any explicit activation delay, can be handled as any other RRC reconfiguration
  + Option 3: others
* Recommended WF
  + To be discussed

### Sub-topic 1-4

*Reporting delay requirement*

The delay between the CSI-RS and UCI containing the report with the predicted PMI should be agreed

**Issue 1-4: Reporting delay requirement**

* Proposals
  + Option 1: Use 4ms separation between CSI-RS and predicted PMI report
  + Option 2: Use 4+Xms separation between CSI-RS and predicted PMI report
    - X to be defined by RAN1
  + Option 3: wait for RAN1 decision
  + Option 4: other
* Recommended WF
  + Option 3

### Sub-topic 1-5

*Scheduling delay*

The delay from when the UE prediction is send until the time the TE applies should be discussed and agreed

**Issue 1-5: Scheduling delay**

* Proposals
  + Option 1: n+4
  + Option 2: n+3
  + Option 3: n+2
  + Option 4: others
* Recommended WF
  + Option 1

Agreement:

* + The delay from when the UE prediction is send until the time the TE applies is n+4 (slots) for FDD

### Sub-topic 1-6

*Generalization*

Several companies proposed to further study generalization issues and how to ensure that the UE performance does not degrade under different conditions.

**Issue 1-6: Generalization**

* Proposals
  + Option 1: Introduce tests with different MCSs
    - MCS 13
    - MCS4
    - MCS19
    - MCS20
    - >MCS20
  + Option 2: Introduce tests for different throughput metrics:
    - 90%
    - 70%
    - 30%
  + Option 3: different SNR points
  + Option 4: Combinations of the above
  + Option 5: further discuss after more evaluation through simulation
* Recommended WF
  + To be discussed

### Sub-topic 1-7

*Doppler values*

There are proposals to further limit the Doppler values being consider

**Issue 1-7: Doppler values**

* Proposals
  + Option 1: keep 20Hz and 50Hz
  + Option 2: keep only 20Hz
  + Option 3: others
* Recommended WF
  + Option 1

### Sub-topic 1-8

*Simulation results and next steps*

Several companies provided simulation results based on the agreed assumptions. The results and next steps in the simulation evaluation campaign should be discussed

**Issue 1-8: Simulation results**

* Proposals
  + Discuss the simulation results and next steps
    - Simulation results
    - Refinement of parameters:
      * Introduce realistic channel estimation
      * Further discuss CSI-RS configuration, codebook configuration, Dopller, SNR, etc
* Recommended WF
  + To be discussed

Discussion on refinement of simulation parameters to be done mainly offline

# Topic #2: RRM core requirement and testing framework for beam management

This section contains the sub-topics regarding specific issues for beam management.

## Companies’ contributions summary

|  |  |  |
| --- | --- | --- |
| **T-doc number** | **Company** | **Proposals / Observations** |
| [**R4-2509298**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509298.zip) | CATT | **Proposal 1: RAN4 to verify at least half of predicted beams, i.e., N=.**  **Proposal 2: The same x value applies for each comparison between the same ranked predicted beam and genie-aided beam. Value of x is further discussed based on the simulation results.**  **Proposal 3: RAN4 not to apply absolute RSRP accuracy requirement to other beams, i.e., predicted beams with lower RSRPs.**  **Proposal 4: RAN4 to remove the square bracket in the definition of relative RSRP accuracy, i.e.,**  **Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - reported L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), ~~[~~where the beam index n owns the largest reported value~~]~~.**  **Proposal 5: The** **reported L1-RSRP in the relative RSRP accuracy definition can be measured RSRP.**  **Proposal 6: The relative RSRP accuracy definition applies for BM-Case 2.**  **Observation 1: For BM-Case 1, UE measurement behaviour on SetB beams is the same as L1-RSRP measurement. The prediction for SetA beams in spatial domain is a short UE internal procedure. The legacy L1-RSRP measurement requirement can be reused for BM-Case 1.**  **Observation 2: For BM-Case 2, UE measurement behaviour on SetB beams is mostly the same as L1-RSRP measurement. The prediction for SetA beams over multiple future time instances is a one-shot UE internal procedure. The legacy L1-RSRP measurement requirement can also be reused as starting point for BM-Case 2.**  **Proposal 7: RAN4 to reuse legacy L1-RSRP measurement requirement for both BM-Case 1 and BM-Case 2 and discuss value of M for BM-Case 2 if necessary.** |
| [**R4-2509425**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509425.zip) | Apple | 1. **If we could truly guarantee that the predicted RSRP accuracy in AI/ML is the same as the measured RSRP accuracy in legacy, for every predicted beam the UE reports, then in principle, a separate beam prediction accuracy metric would not be strictly necessary.**   **The testing environment for Beam Management (BM) case 1 and BM case 2 should replicate conditions that sufficiently capture the correlations of the transmit (Tx) and receive (Rx) beam patterns across the entire spectrum of propagation conditions in both spatial (angles of arrival and departure - AoAs and AoDs) and temporal domains.**  **The following conditions will introduce randomness and variations in propagation conditions across both time and spatial domains for the computation of L1-RSRP.**   1. **Different AoDs with respect to the Tx antenna array** 2. **Different AoAs with respect to the Rx antenna array** 3. **Different superpositions of {AoA,AoD} pairs** 4. **Fading/Variation in time domain ( different {AoA,AoD} pairs per resolvable delay bin path)** 5. **UE movement (including rotation)**   **During real-world deployment, UE will encounter random radio propagations characterized by variations in both spatial and temporal domains (fading). Testing UE under similar conditions is important to reflect the realities of deployment accurately**  **The following questions need to be answered to evaluate the feasibility of the FR2 OTA-based test procedure:**   1. **How can we generate multiple beams from the Set-B and Set-A Tx beams given the limitation of two AoAs?** 2. **What assumption is made regarding Rx beam sweeping? Does the UE utilize a fixed Rx beam, or does it sweep to find the optimal Rx beam?** 3. **How can we simultaneously emulate different AoAs (Rx beam) and AoD (Tx beams)?** 4. **How can we achieve dynamic variation in the AoD domain (Tx beam sweeping) for BM case 2 prediction?**   **For RAN4 testing we should aim at satisfying the following goals: (1) Provide confidence that if a DUT passes the test it will also perform well in the field (2) Ensure reliability of using synthetic channels for test data in evaluating models trained on real data**  **Rather than simulating CDL channels and [64-128] Tx beams using multiple angles of arrival (AoAs) and departure (AoDs), the spatial-selective and time-varying propagation conditions can be effectively represented by adjusting the Tx power at the test equipment (TE) for each beam at different times**   1. (BM-Case 1, spatial prediction) RAN4 should not alter the existing TS 38.133 measurement‐delay formula for BM‑Case 1, instead, networks should be encouraged to **optimize the spatial‐sweep factor N** by sizing their CSI‑RS resource‐sets to match exactly the number of Rx‑beam measurements required for AI/ML inference. By keeping **M=1** and **P=1** and simply adjusting the resource‑set count, the gNB can directly trade CSI‑RS overhead against measurement latency. This approach preserves full compatibility with legacy and gives operators freedom to hit a desired delay‑overhead point for single‑snapshot inference 2. (BM-Case 2, temporal/spatial prediction) RAN4 should make the temporal depth **M** a configurable RRC parameter with a defined range, retain the existing spatial-sweep rule **N = ⌈maxNumberRxBeam / N\_res\_per\_set⌉** while allowing limited-size resource sets to keep latency bounded as **M** grows, and mandate **P = 1** by default via CSI-RS scheduling that avoids SMTC/DRX collisions; when perfect alignment isn’t possible, require the UE to add a small per-report “gap” flag indicating which of the expected measurement instants were actually captured/used, so the network can judge prediction quality. 3. RAN4 should first decide whether truly scenario-agnostic performance requirements are achievable; if not, we propose agreeing a scenario list to anchor requirements and conformance tests, starting with BM-Case 1. Concretely, define: (i) Scenario 1 (Set A ≠ Set B; wide→narrow), e.g., use SSB beams to predict CSI-RS beams, with two input options: 1a “best-Rx-beam RSRP per Tx in Set B” (repetition ON) and 1b “specific-Rx-beam RSRP per Tx in Set B”; and (ii) Scenario 2 (Set B ⊂ Set A; narrow→narrow or wide→wide), comprising 2-1 (SSB), predict all SSB beams from a subset, with 2-1a best-Rx input (repetition ON) and 2-1b specific-Rx input, and 2-2 (CSI-RS), predict additional CSI-RS beams from a subset with repetition OFF. The use of beam-pair prediction remains a UE implementation choice and is out of scope for this WI; we propose no separate requirements or test cases for beam-pair prediction in this release and instead record it as an item for future study (e.g., Rx-beam prediction for RRM measurement reduction) in a next 3GPP/6G release 4. RAN4 is invited to study under realistic Set A/Set B CSI‑RS patterns the impact of elongated  ​ on beam‑failure rate, and to evaluate mitigation options such as (i) defining a separate, short‑period “health‑check” for candidate beams, (ii) capping  ​ to a fixed upper limit (e.g. 40 ms or one DRX cycle), or (iii) scaling the freshness factor ​ downward when the serving beam is AI/ML‑predicted 5. For UE side model known TCI state is determined based on Set B transmission and UE report of top-K. 6. For NW side model: (1) TCI state is unknown if RS of target beam is not transmitted to UE prior to TCI state switch (2) If RS of target beam is transmitted to the UE and reported, legacy definition of known TCI state is applicable 7. NW side data collection, QCL type D is signalled between set A and set B RS. 8. For NW side model, TCI state is known if QCL source is signalled in both data collection and inference (to ensure UE Rx consistency), it is known. Otherwise, it is unknown 9. (for Alt 1): Define UE capability for Alt 1 of UE implementation. When UE does not support this capability, the TCI state is unknown. Unknown TCI state switching delay is applied 10. (for Alt 2): When gNB signal the QCL relationship between set A and set B, in training RS configuration, inference RS configuration and performance monitoring RS configuration (example of Alt 2 NW beam implementation of set A and set B),   **The TCI state is known when the following is met:**  **Target TCI state is based on RS from set A beams with QCL relationship configured to a set B beam**  **The time between TCI state switch command and transmission of the QCLed RS from set B shall not exceed X ms**  **X = 1280 ms (legacy)**  **UE shall send measurement report for the QCLed RS from Set B between transmission and TCI state switch**  **TCI state and SSB of TCI state remain detectable during TCI state switching period**  **Otherwise, it is unknown.**   1. The ability to identify the Rx beam associated with an unmeasured predicted Tx beam may rely on the UE’s supported features. 2. Because AI-driven beam prediction introduces timing uncertainty, adopt a lightweight and flexible LCM framework that (i) exposes the UE’s worst-case preparation time, (ii) signals the required measurement-history depth, and (iii) flags whether activation uses a fast toggle or a full reconfiguration. With these inputs, the scheduler can predict the first valid report, handle rapid on/off cycles cleanly, and place uplink grants accurately for both quick toggles and full reconfigs. 3. RAN4 should address the lack of end-to-end delay visibility for UE-side AI/ML performance monitoring (window → compute → uplink) by adopting: (1) a single “performance-monitoring latency” budget that bounds the whole chain; (2) a two-step “progress + result” scheme where the UE first flags the window is ready and later sends the full KPI/event; and (3) a “fast event” path allowing the UE to immediately signal threshold breaches via a lightweight aperiodic report so the gNB can fall back. 4. RAN4 should set one end‑to‑end timer, , so that from the instant the UE flags inapplicability to the first reliable fallback report, everything (UL flag, NW reaction, signalling, new report) finishes within **K × T\_CSI‑RS\_max**, the slower of the relevant CSI‑RS periods. This single cap prevents drawn‑out beam recovery yet relies only on existing UAI signalling and CSI‑report reconfiguration. 5. Multi-AoA based Testing Setup for BM: Evaluate the testability requirements for simulating time-varying input power to a sparse probe layout based on a simplified channel CDL model. The key aspects to be considered are: 6. **How to determine the minimum number of clusters to be emulated without introducing bias in the results. Investigate a quantitative “goodness” criterion to selecting the number of probes/clusters.** 7. **Consider incorporating UE rotation for the test** 8. **How many different CDL channels we need to support (considering generalization purposes as well)** 9. To establish a criterion for assessing the feasibility of a simplified sparse probe layout for beam management (BM) testing, the following procedure can be considered: 10. **Train a model using the reference CDL channel and test it with the simplified sparse layout-based CDL channel.** 11. **Train a model using the simplified CDL channel and test it with the reference CDL channel model.** 12. **Train and test a model using the reference CDL channel.**   **A similarity metric should be defined to compare the RSRP accuracies across these test cases. This metric would indicate the effectiveness of testing with the simplified CDL model. The similarity metric can be derived through simulations.**   1. Train a model using the re For BM-Case 1 spatial prediction and for verifying RSRP accuracy we propose to use the testing setup and channel emulator functionality as described in Fig 6. (multiple static configurations 2. A screenshot of a computer     AI-generated content may be incorrect.For testing the beam prediction/RSRP accuracy for BM-Case 2 temporal prediction we propose the framework described in this section and visualized in Fig 7, through a non static configuration of the channel emulator   **Fig. 9: Proposal for single AoA BM testing setup (flowchart)**   1. RAN4 to consider the set of procedures/steps shown in the flowchart for adopting a single AoA BM testing setup 2. RAN4 to evaluate the tradeoff between testing complexity (number of different channel realizations for power measurements, different number of parameters that need to be emulated in the fading simulator) and establishing accuracy on the test metric (top-K) beam prediction accuracy 3. RAN4 to evaluate if a power measurement should be performed with a single Rx beam or the optimum Rx beam should be determined by sweeping Rx beams from an Rx codebook to measure the power of set B Tx beams 4. RAN4 to discuss Option2 (a Reference Complex UE Rx Beam/Antenna pattern) for emulation of Rx patterns in the fading simulator 5. Distinguish between the descriptions of power based CDL channel model and the original complex domain CDL channel fading simulator 6. RAN4 to evaluate the prediction performance difference between a reference complex UE Rx beam/antenna pattern and its own complex pattern to ensure that testing will provide confidence that UE will perform well in the field if it passes the test |
| [**R4-2509761**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509761.zip) | Xiaomi | **Proposal 1: For BM-case 2, introduce a scaling factor Q to adjust the measurement time for RSRP predictions, where Q is determined by UE capability, to adopt to different UE implementations.**  **Proposal 2: For BM-case 2, RAN4 to define prediction window:**   * **Stating point of prediction window is:** * **The most recent occasion of the CSI-RS/SSB resource in Set B for measurement + Reference time of the earliest time instance for the predicted results** * **Ending point of prediction window is:** * **Starting point + Ttime\_gap \*N, where Ttime\_gap is the time gap between two consecutive future time instances configured by RRC, and N is the number of future time instance(s) configured by RRC.**   **Proposal 3: For BM-case 2, the scheduling and measurement restriction can be relaxed in the prediction window.**  **Proposal 4: Define a new UE capability with the following table:**    **Proposal 5: For BM case-1, the TCI state switch command is received within Tprediction\_valid ms after latest predicted L1 RSRP report of the target TCI state. Where Tprediction\_valid ms is prediction valid time which is related to channel coherence time. FFS Tprediction\_valid ms.**  **Proposal 6: For BM case-1, when target TCI state is in set A but is not QCL-D to any measured RS before, there is no detectable RS and SNR condition.**  **Proposal 7: For BM case-1, if the predicted Tx beam in Set A is QCL Type-D to a known measured Tx beam, where TX beam can be both inside or outside set B, the corresponding Rx beam is known.**  **Proposal 8: For BM case-2:**   * **The target TCI state is known when the following is met: The TCI state switch command is received within Tprediction\_time\_period + Tprediction\_valid ms upon the last transmission of the RS resource for beam reporting or measurement and the UE has sent at least one L1 RSRP report of the target TCI state** * **there is no detectable RS and SNR condition.**   **Proposal 9: For OTA test, prefer to split the test into two parts.**   * **Part 1: Verify AI model prediction algorithm, using single AOA model.** * **Part 2: Verify input data accuracy of model AI based on RX beam sweeping, using multi-AOA model.**   **Observation 2: If Emulated signal considers both TX and RX beamforming gain, RX beam pattern is UE specific. In order to reduce test complex, prefer to design a common pre-defined RX beam pattern for both TE and UE.**  **Observation 3: if a fixed RX beamforming gain is used for multi-path TX signals, there will be some mismatch between emulated signal for test and original signals for training. The mismatch impact needs to be evaluated.**  **Observation 4: For simplified channel model, L1-RSRP delta between different TX beams will increase if cluser number and AOA number is further reduced. It’s more easiler for AI to distingwish the best Top-K TX beam. The AI prediction requirement will be relaxed.**  **Proposal 10: For single AOA, RAN4 to discuss whether it’s possible to design common pre-defined RX beam pattern to generate emulated single AOA signals.**  **Proposal 11: For simplified multi-AOA CDL model, cluster number can be reduced from spatial angle offset and power aspect:**   * **Combine clusters if AOA offset between these clusters is smaller than a threshold** * **Remove cluster with low power.**   **Proposal 12: If the test metric is beam index prediction accuracy, UE may need to report ground truth of N best beam index for set A.**  **Observation 5: For deriving ground truth of set A, SNR level should satisfy that for at least Top-M beams, the measurement error is small. SNR didn’t need to guarantee that all beams in set A can be measured accurately.**  **Proposal 13: For deriving ground truth of set A, SNR level should satisfy that for Top-M beams, the measurement error is small. RAN4 to discuss how to decide M and how to set SNR.**  **Proposal 14: For BM case-1, channel doppler can set to 0 or a small value to guarantee that there is neglectable L1-RSRP variation.**  **Proposal 15: for BM case-2, channel doppler will depend on UE speed and UE trajectory. RSRP variation impact can be considered in RSRP accuracy requirement.**  **Observation 6: UE will report best predicted L1-RSRP/beam index at T1 and measured RSRP/beam index of best beam at T2 to TE, it’s easy for UE to pass the test by cheating.**  **Proposal 16: RAN4 to discuss how to solve the UE cheating issue if UE report both predicted result and ground truth.** |
| [**R4-2509921**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509921.zip) | Nokia | **Proposal 1: RAN4 should consider the legacy L1-RSRP measurements with the same values of M, N, and P for measurement delay requirements for BM-Case1.**  **Proposal 2: RAN4 should consider the enhancement of measurement period for L1-RSRP measurement for BM-Case2, where the values of M higher than 3 should be considered depending upon the observation window size.**  **Proposal 3: RAN4 should consider the impact in RAN4 requirements for the Evaluation period due to the presence of predicted beams.**  **Proposal 4: RAN4 should consider the Table below as a baseline for further discussion on LCM operations related RRM core requirements.**  Table : Delay components related LCM operations for different CSI-report configurations   |  |  |  | | --- | --- | --- | | CSI-Report Configuration | Activation | Deactivation | | Periodic CSI-reporting | Delay from RRC Reconfiguration  Delay from UE processing related to capability report  Delay due to CSI-RS measurement period | Delay from re-configuration (RRC Reconfiguration)  Delay from UE processing related to capability report  (Option 1) Delay due to CSI-RS measurement period  (Option 2) Delay due to UE’s acknowledgement upon reconfiguration (RRCReconfigComplete) | | Semi-persistent CSI-reporting | Delay for activating semi-persistent CSI resource set  Delay for selecting of semi-persistent CSI resource set  Delay from UE processing related to capability report  Delay due to CSI-RS measurement period | Delay for activating semi-persistent CSI resource set  Delay from re-configuration (RRC Reconfiguration) | | Aperiodic CSI reporting | Delay related activating aperiodic CSI resource set,  Delay from UE processing related to capability report  Delay due to CSI-RS measurement period | Delay related activating aperiodic CSI resource set  Delay from UE processing related to capability report  Delay due to CSI-RS measurement period |   **Proposal 5: RAN4 to consider RRCReconfigurationComplete message containing applicable functionality report as the starting point to define activation delay requirement for periodic CSI reporting for AI/ML BM use case.**  **Proposal 6: RAN4 to consider first inference report as the end point to define activation delay requirements for periodic/aperiodic/semi-persistent CSI reporting for AI/ML BM use cases.**  **Proposal 7: RAN4 to define accuracy requirements for UE-assisted monitoring where the UE reports the monitoring metric (i.e., monitoring Type 1 Option 2) in BM-Case1 and BM-Case2. The accuracy of the performance monitoring metric should be tested at least in static radio conditions.**  **Proposal 8: RAN4 to define delay requirements for UE-assisted monitoring where the UE reports the performance monitoring metric, RS-PAI, in BM-Case1 and BM-Case2. The starting point of this delay can be when UE sends RRCReconfigurationComplete message in response to the configuration of monitoring RS resources via RRCReconfiguration, and the ending point can be when UE reports the first performance monitoring metric, RS-PAI.**  **Proposal 9: RAN4 to further discuss the delay requirement for AI/ML BM functionality pre-activation when performance monitoring is considered when verifying AI/ML BM functionality to be activated**  **Proposal 10: RAN4 to further discuss the delay requirement for AI/ML BM functionality failure recovery success when AI/ML BM functionality failure happens and UE is enabled for recovery mode.**  **Proposal 11: RAN4 should consider relaxing known conditions for TCI states (Clause 8.10.2 of TS 38.133) for both BM-Case1 and BM-Case2. The TCI state switch command is received within a delay X\*1280 ms or the prediction should be added in the TCI state condition.** |
| [**R4-2510105**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510105.zip) | MediaTek Inc. | **Proposal 1**: If the predicted Tx beam in Set A is QCL Type-D to a known Tx beam, then legacy known TCI state conditions are applicable. If the predicted Tx beam in Set A is not QCL Type-D to a known Tx beam, known TCI state conditions shall be updated as  - UE reports [TCI state known] in Capability X, and  - TCI state switch command is received within 1280 ms upon transmitting the predicted L1-RSRP report for the target TCI state  - The TCI state remains detectable during the TCI state switching period, and  - The SSB associated with the TCI state remain detectable during the TCI switching period  - SNR of the TCI state ≥ -3 dB  **Observation 1**: In terms of Top K/1 prediction accuracy, the performance of AI/ML BM case 1 models for different scenarios may have much difference.  **Proposal 2:** Further study and discuss whether it is possible to define scenario-agnostic performance requirements for AI/ML BM. If not, maybe we can seek for similar solutions as TDCP, i.e., define the test cases for limited scenarios instead of defining scenario-agnostic accuracy requirements.  **Proposal 3**: Both BB error and RF error should be considered when define the requirements and test case.  **Observation 2**: There is little impact on either Top K prediction accuracy or absolute RSRP accuracy training with measurement error or not.  **Observation 3**: Comparing to the genie-aided ground-truth or ground-truth with measurement error, Top K prediction accuracy is a bit different and absolute RSRP accuracy of the 1st and 3rd predicted beams are similar.  **Observation 4**: Using different dataset, for some metric, the accuracy performance are similar, while for some other metric, the accuracy performance may have some divergence.  **Proposal 4**: Further discuss how to handle the performance divergence between different datasets.  **Proposal 5**: Prefer that the metric to be used in RAN4 is also used for monitoring, i.e., x shall be 0.  **Proposal 6**: From the point of necessity, the absolute RSRP accuracy requirements should be defined for none Top-1 beams too.  **Proposal 7**: RSRP accuracy requirements of Top-1 of predicted beams can be defined by filtering out the data with SNR of the Top-1 beam in set A no less than -3dB. Further discuss how to deal with the issue that the distribution of SNR of the Top-1 beam in set A vary among different companies’ dataset as well as the test dataset generated by TE vendors.  **Observation 5**: For different beams, e.g., predicted 1st strongest and 5th strongest beam, the predicted absolute RSRP accuracy varies. It may be the SINR difference that results in the difference between the predicted absolute RSRP accuracy of top-1 beam and other beams.  **Proposal 8**: Further discuss whether it is workable to define a single absolute accuracy requirement which is applicable to all/some of the predicted beams with some side conditions.  **Proposal 9**: For AI/ML BM, further discuss whether the performance under synthetic channels can reflect the performance of AI/ML models in the real field.  **Proposal 10**: For AI/ML BM, discuss what the testing goal is, to verify the model can work well in real field or just to test UE can train a model that works well under a specific scenario.  **Proposal 11:** If the testing goal is to verify the model can work well in real field, and the performance under a specific channel/scenario cannot reflect the performance in the real field, it is necessary to design a test system that is possible to use field data in the test.  **Proposal 12:** If the testing goal is to test UE can train a model that works well under a specific scenario, prefer to choose a test system that is easy to implement and evaluate.  **Proposal 13**: In AI/ML BM test, emulate the spatial-selective and time-varying in propagation conditions through adjusting Tx power at TE on each beam at different time.  **Observation 6**: Single AoA test system can be applicable to all kinds of synthetic channels as well as using real field data. It can be easily extended to different scenarios and settings.  **Observation 7**: With single AoA, TE can know the expected RSRP of Tx beams at UE baseband.  **Proposal 14**: Use existing IFF or enhance IFF test systems with single AoA to test R19 DL Tx beam prediction models.  **Proposal 15**: Standardized UE radiation pattern in 38.901 can be assumed in R19 AI/ML BM test with single AoA test system instead of using test UE’s radiation pattern.  **Proposal 16**: The procedures of AI/ML BM test with single AoA:  Beam lock at test UE during the test.   * Step 1: Overall coupling loss calibration   + TE transmit the signal with a fixed power, UE receive signal and report RSRP to TE. TE can get the overall coupling loss through the relationship between TE’s transmit power and the reported RSRP (i.e., including the antenna gain of the probe, OTA pathloss and UE’s Rx beamforming gain). TE can ask UE to report multiple times to reduce the impact of noise in the coupling loss calculation. * Step 2: Main test   + TE gets the test dataset. The dataset can be early prepared through computer-like simulation using standardized UE radiation pattern in 38.901 and is the same for all the test UEs.   + TE transmit the calibrated power on each Tx beam in set B from the test dataset by the overall coupling loss. UE measures set B and predicts Top K or RSRP of set A. In such a way, the uncertainty of beamforming gain can be removed when deciding pass/fail for each predicted RSRP.   **Observation 8**: Simulation workload will increase linearly with the number of candidate methods to simply the channels.  **Proposal 17**: Make a down-selection on the candidate methods of how to simply CDL channels.  **Proposal 18**: Discuss how to emulate UE movement in multiple AoA test systems.  **Proposal 19**: How to avoid UE cheating in the test, such as reports fake values according to the previous predicted results, shall be further discussed to use the reported RSRP measurement as ground truth.  **Proposal 20**: Discuss how to get training dataset if to use multiple AoA test systems.  **Observation 9**: The upper bound of SNR with rough beam are much lower than that with fine beam.  **Proposal 21**: Fine beam should be used in AI/ML test.  **Observation 10**: Among all the test methods, only using peak fine beam direction can make sure the lowest SNR above -6dB with 30dB dynamic power range.  **Observation 11**: Multiple AoA test systems cannot use peak fine beam direction only.  **Proposal 22**: Using peak fine beam direction in AI/ML BM test with single AoA test system.  **Observation 12**: SNR dynamic range of multiple AoA test system would be at least 12dB narrower than single AoA test system. |
| [**R4-2510159**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510159.zip) | CMCC | ***Proposal 1: it is proposed that the conditions for known TCI state is updated as following to cover AI/ML based beam prediction. And the updated parts are applied to both case 1 and case 2.***   |  | | --- | | ***The TCI state is known if the following conditions are met:***  ***-*** ***During the period from the last transmission of the RS resource used for the L1-RSRP measurement reporting for the target TCI state to the completion of active TCI state switch, where the RS resource for L1-RSRP measurement is the RS in target TCI state or QCLed to the target TCI state, or***  ***- TCI state switch command is received within 1280 ms upon the last transmission of the RS resource for beam reporting or measurement***  ***- The UE has sent at least 1 L1-RSRP report for the target TCI state before the TCI state switch command***  ***- The TCI state remains detectable during the TCI state switching period***  ***- The SSB associated with the TCI state remain detectable during the TCI switching period***  ***- SNR of the TCI state ≥ -3 dB***  ***- TCI state switch command is received within T ms upon the report of predicted results.***  ***- If the predicted Tx beam in Set A is QCL Type-D to a known Tx beam***  ***- If the predicted Tx beam in Set A is not QCL Type-D to a known Tx beam, and UE is capable of knowing RX beam after beam prediction***  ***Note: the value of T is FFS. The value of T for case 1 and case 2 can be different.*** |   ***Proposal 2: for TCI state switch to a predicted beam, no need to consider the delay for RX beam refinement in the TCI state switch delay with following conditions:***   * ***if the predicted Tx beam in Set A is QCL Type-D to a known Tx beam, the corresponding Rx beam is known, or*** * ***if the predicted Tx beam in Set A is not QCL Type-D to a known Tx beam, and UE is capable of knowing RX beam after beam prediction***   ***Proposal 3: for Metrics/KPIs for beam ID prediction, for the FFS on top-N (N</=K) predicted beams have to be considered in the success rate evaluation, it is proposed that whether N value can be larger than 1 is related with x value.***   * ***Either x =0 and N>1, or x of non-zero value and N=1 are OK;*** * ***It is not necessary that N>1 and x of non-zero value are considered together***   ***Proposal 4: for the case that both RSRP and beam ID are reported, it is proposed that both RSRP accuracy requirements and successful rate for the correct prediction are applied.***  ***Proposal 5: it is proposed that absolute RSRP accuracy requirement only applies to Top-1 of predicted beams. And relative accuracy requirements applies to other reported predicted beams.***  ***Proposal 6: for relative RSRP accuracy, it is proposed that the reported L1-RSRP cannot be measured RSRP, since RAN1 agreed that the RSRP of predicted beam(s)in the report of inference results is the predicted RSRP which is based on AI/ML output.***  ***Proposal 7: for relative RSRP accuracy, it is proposed that beam index n owns the largest reported value.***  ***Proposal 8: for BM case 1, it is proposed that***   * ***The relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), ~~[~~where the beam index n owns the largest reported value~~]~~***   ***Proposal 9: for BM case 2, it is proposed that***   * ***The relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i for time instance m - predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i for time instance m - ground truth of L1-RSRP of beam index n), where the beam index n owns the largest reported value among all the predicted beams. 1<=m<=M where M is the number of time instance*** |
| [**R4-2510234**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510234.zip) | Qualcomm Incorporated | **Observation 1: The key requirements for AI/ML BM test setup are listed following:**   * **Requirement#1: Channel model in the test**   + **The test should reflect performance in the field. Hence, CDL or simplified CDL to properly verify beam management performance.** * **Requirement#2: Avoid the cheating**   + **Progress A: Model training for Set A’s is generated with channel parameters #1. Set B in UE measurement should have enough randomness compared to Set A’.**   + **Progress B: Set A is generated to make sure to use the same parameters (initial phase, angular spread, delay spread) as the Set B in the Progress A measurement step. NW antenna/beam configurations for Set A and Set B should be deterministic. For some prediction metrics (e.g., option 2 and 3), UE can pass the test without actual measurement if TE does not randomize the transmission of its beams.** * **Requirement#3: Support the beam number of Set A**   + **TE needs to support Set A for ground truth verification.**   **Observation 2: Verifying the AI/ML beam management performance in a TDL channel with single AoA without spatial characteristics makes the test too simple and inconsistent with the actual field environment. Thus it is not feasible to general proper Set A and Set B with single AoA setup.**  **Observation 3: For single AoA test setup with two-step approach, there are following limitations:**   * **UE has to report the UE antenna pattern with X\*Y test points that is very time consuming.** * **UE needs to support additional test function for RSRP reporting** * **If Rx beam ID is one of the input for AI/ML model, additional test mode/signalling is needed** * **UE’s beam management performance can’t be verified properly since for each of UE position there is enough dewll time and UE beam is locked.**   **Observation 4: There will be plenty of additional probes if BS Tx beams point to other clusters even though some of the probes could be reused.**  **Observation 5: The UE’s antenna gain gap between peak and 50%-tile spherical coverage is over 10dB for PC3. The very weak clusters would not influence the UE beam management even if they reach the UE within the UE’s spherical coverage.**  **Observation 6: The AoDs/ZoDs of very weak clusters are unlikely to influence the selection of the strongest beam in any reasonably designed codebook of Set A and Set B.**  **Observation 7: The weak clusters that have a limited impact on the beam management could be removed to reduce the number of clusters in the CDL channel model.**  **Observation 8: The clusters that have the same/similar AoAs could be merged. And the intra-cluster angle spread, e.g., CASD, CASA, CZSD, CZSA can be reduced to simplify the test setup. In this case, a new simplified channel model is emulated by test setup.**  **Observation 9: To generate Set A and Set B, multiple AoAs/probes test setup is necessary.**  **Observation 10: Using Multiple AoA test setup with FR2 RRM test can reuse the FR2 existing test setup to save the test cost and a simplified CDL channel model needs to be specified.**  **Observation 11: Pros and Cons for single AoA and Multiple AoA test setup presented in Table 2.**  **Table 2: Analysis for single AoA and Multiple AoA test setup**   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | **Single AoA test setup** | **Single AoA test setup with two steps approach** | **Multiple AoA test setup with 3D MPAC** | **Multiple AoA test setup with FR2 RRM test setup (eIFF)** | | **Chanel model** | **TDL** | **CDL** | **Simplified CDL** | **Simplified CDL** | | **Number of probes** | **1** | **1** | **6-8** | **4** | | **Complexity of testing and TE implementation** | **Low** | **High**  **Note: Very time-consuming for Step 1. And additional test mode or signaling is needed.** | **Medium**  **Note: Additional probes are needed on top of 3D-MPC** | **Low**  **Note: Reuse existing FR2 RRM test setup** | | **Realistic environment** | **Low** | **Low**  **Note: UE beam management performance can not be verified properly.** | **High/Medium**  **Note: It depends on what does simplified CDL look like** | **High/Medium**  **Note: It depends on what does simplified CDL look like** | | **Whether the test system can support spatial information like UE Rx beam, UE beam pattern, etc is used** | **No** | **No** | **Yes** | **Yes** | | **Support of randomness for transmission beams** | **No** | **Yes** | **Yes** | **Yes** |   **Proposal 1: Single AoA test setup incluidng the use or non-use of the two-step method should not be considerd for AI/ML BM testing.**  **Proposal 2: RAN4 should consider the follwoing two multiple AoAs/probes options as the test seup for AI/ML beam management testing**   * **Option 1: Multiple AoA test setup with 3D MPAC**    + **Consider reducing the cluster number with weak power level, merging the cluster with the same/similar AoA, and decreasing intra-cluster angle spread in CDL channel model to generte a new simplified CDL channel model**   + **Define the paramters for the new simplified CDL channel model based on that 3D-MPAC test setup is used**   + **Take CDL-C UMi as the starting point** * **Option 2: Multiple AoA test setup with FR2 RRM**    + **Conider multiple AoA test setup with FR2 RRM test setup to emulate compareble Tx beam with simplifed CDL, e.g., generating enough Tx beam number the solutions such as combinaiton of probes transmission and power offset for each of probe can be considered.**   + **Define the paramters for the new simplified CDL channel model based on that FR2 RRM eIFF test setup is used**   + **Take CDL-C UMi as the starting point**   **Propsoal 3: Companeis to evlaute whether the new simiplied CDL channel model generated from candidate test sysetms can be used to properly verify AI/ML BM perforance. Validate the proposed simplified channels from Option 1 and Option 2 through AIML BM system level simulations.**  **Proposal 4: RAN4 to investigate whether measurement uncertainty should be considered when defining AI/ML BM core requirements.** |
| [**R4-2510334**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510334.zip) | vivo | Draft CR |
| [**R4-2510337**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510337.zip) | vivo | **Proposal 1: For spatial prediction, the measurement of set B relies on one-shot measurement**  **Proposal 2: For L1-RSRP measurement of set B, the definition and value of the beam sweeping factor *N* follows existing requirements.**  **Proposal 3: For requirements on AI/ML-based L1-RSRP measurement, when referencing values from an existing clause in the same spec version where the measurement period is defined, RAN4 will discuss the coexistence supportability of the prediction with other previously enhanced features. These enhanced features primarily include:**  ***The UE capability specified in clause 9.5.4 L1-RSRP measurement requirements in Rel-19 spec：***  ***­ 1. bwpOperationMeasWithoutInterrupt-r18***  ***­ 2. ncd-SSB-BWP-Wor-r18***  ***­ 3. groupBasedBeamReporting-r17***  ***­ 4. multiDCI-IntraCellMultiTRP-TwoTA-r18***  ***­ 5. multiDCI-InterCellMultiTRP-TwoTA-r18 and rxTimingDiff-r18,***  ***­ 6. rxTimingDiff-r18***  ***­ 7. highSpeedMeasFlagFR2-r17***  ***­ 8. concurrentMeasGap-r17***  ***­ 9. musim-GapPreference-r17***  ***­ 10. multiCellL1-measRTD-greaterThan-CP-r18***  ***­ 11.*** ***fastBeamSweepingMultiRx-r18***  **Proposal 4: For the above enhanced features, RAN4 not to consider combining the AI/ML-based BM with the features related to MultiTRP, HST, group-based beam reporting, LTM, Multi-Rx when defining requirements for AI/ML-based L1-RSRP measurement. The specific UE capabilities include:**  ***­ groupBasedBeamReporting-r17***  ***­ multiDCI-IntraCellMultiTRP-TwoTA-r18***  ***­ multiDCI-InterCellMultiTRP-TwoTA-r18 and rxTimingDiff-r18,***  ***­ rxTimingDiff-r18***  ***­ highSpeedMeasFlagFR2-r17***  ***­ multiCellL1-measRTD-greaterThan-CP-r18***  ***­ fastBeamSweepingMultiRx-r18***  **Observation 1: The inference latency is reflected in the RAN1 specifications. As for the prediction delay, this inference latency is already included within TReport.**  **Proposal 5: The prediction processing delay component does not need to be explicitly defined within the prediction delay requirements TL1-RSRP\_Prediction\_Period\_CSI-RS.**  **Proposal 6:**   * **For CSI-RS based prediction , the prediction period requirements TL1-RSRP\_Prediction\_Period\_CSI-RS for BM can be defined as**   **Table : Prediction period TL1-RSRP\_Prediction\_Period\_CSI-RS for FR2**   |  |  | | --- | --- | | Configuration | TL1-RSRP\_Prediction\_Period\_CSI-RS (ms) | | non-DRX | max(TReport, ceil(M\*P\*N)\*TCSI-RS) | | DRX cycle ≤ 320 ms | max(TReport, ceil(1.5\*M\*P\*N)\*max(TDRX,TCSI-RS)) | | DRX cycle > 320 ms | ceil(M\*P\*N)\*TDRX | | NOTE 1: TCSI-RS is the periodicity of CSI-RS configured for L1-RSRP measurement. TDRX is the DRX cycle length. TReport is configured periodicity for reporting.  NOTE 2: the requirements are applicable provided that the CSI-RS resource configured for L1-RSRP measurement is transmitted with Density = 3. | |  * **For SSB-based prediction , the prediction period requirements TL1-RSRP\_Prediction\_Period\_SSB for BM can be defined as**   Table : Prediction period TL1-RSRP\_Prediction\_Period\_SSB for FR2   |  |  | | --- | --- | | Configuration | TL1-RSRP\_Prediction\_Period\_SSB (ms) | | non-DRX | Max (TReport, ceil (M\*P\*N) \*TSSB) | | DRX cycle ≤ 320 ms | Max (TReport, ceil (1.5\*M\*P\*N) \*max (TDRX, TSSB)) | | DRX cycle > 320 ms | Ceil (1.5\*M\*P\*N) \*TDRX | | Note: TSSB = *ssb-periodicityServingCell* is the periodicity of the SSB-Index configured for L1-RSRP measurement. TDRX is the DRX cycle length. TReport is configured periodicity for reporting. | |   **Proposal 7：For prediction delay requirement for BM-Case 2 (if defined), RAN4 not to consider aperiodic CSI-RS reporting scenario.**  **Proposal 8: For BM-Case 1 and BM-Case 2, the target TCI state is known when the following is met:**   * ***The TCI state switch command is received within 1280 ms upon the last transmission of the RS resource for beam reporting or measurement and the UE has sent at least one L-1 RSRP measurement or predicted RS quantities report of the target TCI state (clause 8.10.2 of 38.133)*** * **Prediction would not be used as a condition for known TCI state (i.e., not to be reflected as the time delay restriction between prediction results reporting and TCI switch command)**   **Proposal 9: RAN4 to update the known TCI state conditions as**  **Proposal 10: For BM-Case 1, when the UE has the capability to know the corresponding Rx beam, RAN4 needs to further discuss:**   * **If the UE supports this capability, for MAC-CE-based TCI state switch delay for known TCI state, whether the UE still needs to perform fine time tracking—i.e., whether TOk can be 0 and under what conditions.**   **Proposal 11: For BM-Case 2, regarding the MAC-CE-based TCI state switch delay, when the UE has the capability to know the corresponding Rx beam, RAN4 needs to further discuss:**   * **If the UE supports this capability, should the L1-RSRP measurement and fine time tracking delay always be defined according to the unknown TCI state case?** * **Potential side condition: The L1-RSRP measurement must wait until the corresponding TCI state takes effect before it can be performed (i.e., the L1-RSRP measurement cannot be executed earlier than the earliest predicted time instance of the target TCI state).**   **Proposal 12: RAN4 to define absolute RSRP accuracy as one of metrics for beam management requirements/tests, the corresponding definition is:**   * ***For BM-Case 1, the absolute RSRP accuracy = predicted L1-RSRP of beam index i – ground-truth of L1-RSRP of beam index i. The index i corresponds to the index of Top-1 beam of predicted beams***   **Proposal 13: RAN4 to use the follow definition to define the relative RSRP accuracy requirement:**   * + **For BM-Case 1, Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - Reported L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), where the beam index n owns the largest reported value** * **Where, the reported L1-RSRP refers to predicted L1-RSRP**   **Proposal 14: Before discussing KPI definitions, we propose that RAN4 first reaches an agreement on whether performance requirements should be defined for this case.**  **Proposal 15: For Metrics/KPIs for beam ID prediction agreed in RAN4#114bis, RAN4 not to further consider the verification of top-N (N</=K) predicted beams in the success rate evaluation**  **Proposal 16: For Metrics/KPIs for beam ID prediction agreed in RAN4#114bis, it will apply to the case only beam ID is reported and the case both beam ID and L1-RSRP are reported.**  **Proposal 17: For the X value of Metrics/KPIs for beam ID prediction agreed in RAN4#114bis, it can be determined based on the simulation results**  **Observation 2: The simplified CDL channel with small number clusters will degrade the test dependability.**  **Observation 3: The number of probes can be greatly reduced by grouping different Tx beams together. Only one probe or one set of probes can emulate different Tx beams in one group.**  **Observation 4: Different test system with reduced cost and complexity can be introduced by using the Tx beam grouping method.**  **Proposal 18: RAN4 to consider methods to group Tx beams together to reduce the number of probes in the test system.**  **Proposal 19: RAN4 to consider the following procedures to simplify the test system:**   * **Step 1. Generate the BS beamforming power pattern of every Tx beam** * **Step 2. Add the BS beamforming power pattern into the power of different clusters in specific ZoD and AoD, and derive the new power of every cluster** * **Step 3. Use the new power of clusters to generate the power angular spectrum (PAS) of ZoA and AoA** * **Step 4. Some Tx beams with the corresponding ZoA/AoA are to be grouped together based on a certain rule (e.g., the ZoA/AoA of the strongest power derived from the PAS in step 3 are the same angle or in a pre-defined range)** * **Step 5. Using one or small number of probes to emulate one beam group.** * **Step 6. Delete the unnecessary probes and merge some adjacent probes.**   **Proposal 20: RAN4 to consider the following two approaches for design the test system**    **approach a**    **approach b**  Proposal 21: Companies to provide candidate probe locations for CDL-A/B/C for merging and rotating. E.g.,  Proposal 22: RAN4 to consider the UE Rx beamforming resolution in angular domain to perform OTA probe allocations. |
| [**R4-2510374**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510374.zip) | Ericsson | **Observation 1: For the case that only beam ID is reported, the order of top-K beams indicates the order of RSRP or the probability of each beam in Set A to be the Top-1 beam of implicitly.**  **Observation 2: Only considering the baseline:**   * **Top-K/1 (%), and** * **The correct prediction is considered as maximum ground-truth RSRP among top-K predicted beams larger than or equal to the ground-truth RSRP of the strongest genie-aided beam(s) – x dB**   **may result in a wrong, e.g., bad quality, beam being adopted by NW, and results in degraded performance than the legacy measurement solution.**  **Observation 3: From the network perspective, obtaining a plurality of predicted beams, each subject to specific prediction accuracy requirements, is advantageous for enhancing reliability and robustness in beam management.**  **Observation 4: For AI/ML model supporting RSRP and beam ID reporting, if only RSRP accuracy is required, an underlining problem is that the Top-1/top-K predicted beam is far away from the Top-1/top-K genie-aided beam since no comparison to genie-aided beam** **is performed.**  **Observation 5: For the** **case of the AI/ML model with label (beam ID) + RSRP output type, fulfilling beam prediction accuracy and RSRP accuracy may result in more stringent requirements compared with the case of the AI/ML model only with label (beam ID) output type.**  **Observation 6: RAN1 has agreed that BM-Case 1, RSRP of predicted beam(s)in the report of inference results, is the predicted RSRP. It shall be also be valid for BM-Case 2.**  **Observation 7: RAN1 has defined inference delay, i.e., time gap from completion of Set B measurements to the time at which Set A prediction is valid.**  **Observation 8: Even no QCL to a known Tx beam, the UE is able to know the Rx beam of TCI state if the prediction beam is the RS in target TCI state or QCLed to the target TCI state.**  **Observation 9: The known conditions based on prediction are different for BM-case 1 and BM-case 2.**  **Observation 10: In prediction, the TCI state or The SSB associated with the TCI state may not keep transmitting during the TCI state switching period.**  **Proposal 1: Enhance the baseline metric for beam prediction by testing both cases: K = 1 and K = [2], with the beam prediction metrics baseline as follow:**   * **‘Top-K/1 (%)’, and** * **‘The correct prediction is considered as maximum ground-truth RSRP among top-K predicted beams larger than or equal to the ground-truth RSRP of the strongest genie-aided beam(s) – x dB’**   **Proposal 2: The parameter x in beam prediction accuracy metrics should be set to a value that reflects the expected magnitude of measurement error and other errors, [1] dB may be taken as a baseline.**  **Proposal 3: RAN4 to investigate the SNR side condition for prediction. Potential options are under consideration in the context of beam ID reporting:**   * **Option 1: If Set B beams meet the SNR side condition, then the top‑K beams shall meet the measurement accuracy requirements.** * **Option 2: If Set B beams meet the SNR side condition, then the top‑K beams—which also meet the SNR side condition—shall meet the measurement accuracy requirements.** * **Option 3: Both Set A and Set B beams shall meet the SNR side condition, after which the top‑K beams shall meet the measurement accuracy requirements.**   **Among the options, Option 3 is preferred.**  **Proposal 4: Metrics/KPIs for beam management requirements/tests are outlined by the below options:**   * **In case of the AI/ML model with label (beam ID) output type**   + **Beam prediction accuracy is adopted.** * **In case of the AI/ML model with label (beam ID) + RSRP output type**    + **Beam prediction accuracy and RSRP accuracy, are adopted simultaneously.**   **Proposal5: The absolute RSRP accuracy requirement applies to each of Top-K beams, not only the 1st beam in top-K predicted beam.**  **Proposal 6: For the absolute RSRP, the index i shall be any beam index in all beams based on predicted L1-RSRP as per UE reporting, i.e.,**   * **The absolute RSRP accuracy = predicted L1-RSRP of beam index i – groundtruth of L1-RSRP of beam index i. The index i may be any beam index in top-K beams based on prediction report.**   **Proposal 7: Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), [where the beam index n owns the largest reported value].**  **Proposal 8: Relative RSRP accuracy requirement is applicable for Case 1 and Case 2.**  **Proposal 9: Regarding measurement period for prediction, for both BM-case1 and BM-case2, the legacy L1-RSRP measurement delay shall be applied (i.e., reusing M, N, P), with the assumption that the measurement accuracy remains consistent.**  **Proposal 10: For BM-case 2, RAN4 not to design a new measurement period for tracking changes of L1-RSRP by obtaining more than one L1-RSRP measurement result.**  **Proposal 11: For BM-case 2, RAN4 not to design inference delay, since time gap/delay from completion of Set B measurements until the ready of Set A prediction is configurable according to RAN1 agreements.**  **Proposal 12: RAN4 to define the known conditions based on prediction, i.e., the known condition with respect to Set A. the known conditions for Set B, since it is measured, follows the legacy way.**  **Proposal 13: The known conditions for BM-case 1 based on prediction are proposed as follow.**   * **During the period from the last time instance of the RS resource for beam prediction for the target TCI state to the completion of active TCI state switch, where the RS resource for beam prediction is the RS in target TCI state or QCLed to the target TCI state**   **Proposal 14: The known conditions for BM-case 2 based on prediction are proposed as follow.**   * **During the period from the last time instance of the RS resource for beam prediction for the target TCI state to the completion of active TCI state switch, where the RS resource for beam prediction is the RS in target TCI state or QCLed to the target TCI state**   + **The TCI state switch command is received within the valid prediction window of the RS resource for beam prediction**   + **The UE has sent at least 1 inference report for the target TCI state before the TCI state switch command** * **TCI state and SSB of TCI state remain detectable during TCI state switching period**   + **SNR of the TCI state is ≥ -3 dB**   **Proposal 15: The term “detectable” in the condition “The TCI state and the SSB of the TCI state remain detectable during the TCI state switching period” should be updated to reflect prediction procedures, in which the predicted RS may not be transmitted.**   * **Example: Replace “detectable” with “predictable.”**   **Proposal 16: RAN4 to study activation delay requirements for LCM operations, with respect to to different CSI-report configurations (i) periodic CSI report (ii) semi-persistent CSI report and (iii) aperiodic CSI report for both BM-Case1 and BM-Case2.**  **Proposal 17: Activation delay requirements for periodic CSI report shall comprise the time delays:**   * **Processing time, from sending RRCReconfigurationComplete ( if the inference configuration is provided in Step 3), including the time for UE readying AI/ML mode and signal processing time e.g., TRRC\_processing.** * **prediction delay, covering the time needed for measurement acquisition and preparation of the AI/ML inference.**   **Proposal 18: Activation delay requirements for semi-persistent and aperiodic CSI report shall comprise the time delays:**   * **Processing time, from MAC-CE/DCI, including the time for UE readying AI/ML mode and signal processing time, e.g., THARQ  or time required by the UE to perform PDCCH reception.** * **prediction delay, covering the time needed for measurement acquisition and preparation of the AI/ML inference.**   **Proposal 19: RAN4 to define prediction monitoring requirements for performance monitoring Type 1 Option 2 (UE-assisted performance monitoring)**  **Proposal 20: RAN4 shall take the below RAN1 agreements into performance monitoring requirements:**   * **Evaluation period for performance monitoring:**   + **N transmission occasion of the CSI-RS/SSB resources for monitoring, at least for BM-case 1, Where N = 1, 3, 7, 15 is configured in CSI-ReportConfig.** * **Report quantity RS-PAI,**    + **is defined as the total count of accurate reference signal prediction instance(s) that meets the condition, among N latest transmission occasion(s).**   **Proposal 21: RAN4 not to study BFD, RLM, or CBD until clear solutions have been defined and agreed upon in RAN1/2.**  **Proposal 22: RAN4 to study the two questions regarding multiple/single AoA:**   * **What is the (maximal) difference between really tested Tx/Rx beam and the artificial Tx/Rx beam?** * **If the difference can impact AI/ML model and the prediction accuracy?**   **Proposal 23: RAN4 to check if the 32 beams after simplified CDL channel can be produced by [8] (if it is feasible to be realized in test bed) probes which may achieve the balance between performance and implementation complexity.**  **Proposal 24: Measuring Tx/Rx beams using a TDM approach may be a valid method; however, the control mechanism for Rx switching must be explicitly addressed.**  **Proposal 25: The simplified CDL channel only is considered to be applied in inference, no need to be applied in training.**  **Proposal 26: The metric to estimate the “goodness” of a simplified CDL channel model may be represented by the below steps:**   * **Step 1: Test (inference) with the reference CDL channel.** * **Step 2: Test (inference) with the simplified CDL channel.** * **Step 3: The prediction difference between the results of the two tests in the above steps is less than a threshold with [95] percentile.**   **Proposal 27: Test procedure shall at least comprise the below three steps:**   * **Acquire the predicted RSRP under the required corresponding SNR condition** * **Acquire the ideal measurement of RSRP under high SNR condition** * **Compare the two RSRPs acquired in the above steps**   **Note 1: Test configurations are kept same in the steps except for different SNR conditions.**  **Note 2: Arranging the details on the sequence and test configuration may need further investigation in performance part.**  **Proposal 28: For BM-case1, same channel is applied for measurement and inference.**  **Proposal 29: For BM-case2, plural channel instants with a series of changes of the channel parameters are necessary to reflect temporal correlation between channels. At least, it includes:**   * **Generating a set of channel instants, from channel #a+1 to channel #a+m, for measurement, and** * **Generating a set of channel instants, from channel #b+1 to channel #b+n for inference**   **Where,**   * **The channels instants comprise sequent and grade updates of a set of parameters.** * **The numbers of m and n depend on the measurement window and prediction window of BM-case2.** * **The channel parameters may comprise one or more than one of**    + **Departure angles (AOD, ZOD)**   + **Arrival angles (AOA, ZOA)**   + **UE position and/or rotation**   + **Channel coefficient** |
| [**R4-2510806**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510806.zip) | OPPO | **Proposal 1: For BM testability, RAN4 needs to:**   1. **Clarify the limitations regarding the FR2 beams or probes that TE vendors could support** 2. **To determine the test setup for BM, a potential approach could be:**  * **Assuming that TE supports X probes in a chamber** * **By utilizing X probes, X different angles of arrival (AOA) can be achieved and can be mapped to a TX beam transmission with X paths(clusters)** * **Different TX beams in BM set A or set B can be represented by setting different pathloss for the different AOA paths, to mimic the beam sweeping procedure** * **FFS the number of probes that could be supported by TE vendors** * **FFS how to map the X AoAs and corresponding X pathloss to different beams in BM set A/B**   **Proposal 2: Further consider how to ensure that the BM model constructed on the DUT side can match(or approximate match) and be utilized in the testing environment on the TE side.**  **Proposal 3: For both BM-Case1 and BM-Case2,** **the target TCI state is known when the following is met: The TCI state switch command is received within 1280ms upon the last transmission of the RS resource for beam reporting or measurement and the UE has sent at least one L1 RSRP report of the target TCI state.** |
| [**R4-2510873**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510873.zip) | Huawei,HiSilicon | **Proposal 1:**RAN4 to define the requirement of prediction delay in AI BM, where the measurement delay and inference delay are not treated as separate requirements.  **Proposal 2*:*** For data collection in NW-sided model, take the existing core requirement for beam related information reporting as the starting point.  **Proposal 3:** RAN4 will focus on NW-side performance monitoring Type 1 option 1 and reuse legacy requirements for measurement reporting. |
| [**R4-2511018**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511018.zip) | ZTECorporation,Sanechips | **Observation 1: The top-K beams are reported or predicted beams or top-K beams are groundtruth beams.**  **Observation 2: For the top-K beams are groundtruth beams but not the predicted beams, may be some UE reported beams could not be verified or why should we require the RSRP accuracy of the beam that the UE doesn’t even report.**  **Observation 3: What is the standard to verify the performance of AI system based on the absolute RSRP accuracy shall be considered or how to quantify the the small or large L1-RSRP difference shall be studied.**  **Observation 4: One question is that the proposed margin shall be maintained or smaller than legacy value. There are two options to consider the change of proposed margin:**   * **Option 1: Maintain the legacy measurement accuracy. In legacy, UE shall do the full beam sweeping in order to confirm the best Rx beam to receive the downlink RS and get the better communication with network. The AI system has already improve the overhead and reduce the complexity than legacy. Thus, maintain the legacy measurement accuracy could be understood.** * **Option 2: Smaller than legacy measurement accuracy. This is a more direct understanding since the AI performance shall have the better performance than legacy and the measurement accuracy shall be tightened. In this way, the performance of AI method could be displayed.**   **Proposal 1: For the absolute RSRP accuracy, the index i shall be the top-K beams based on the predicted beams or UE reported beams instead of the groundtruth beams.**  **Proposal 2: One margin for the absolute RSRP accuracy shall be studied in order to verify the performance of AI system. The proposed margin could be equal to and smaller than the legacy measurement accuracy requirements.**  **Observation 5: In current RAN4 spec, we do not observe the limitation on absolute RSRP accuracy.**  **Proposal 3: RAN4 shall consider absolute RSRP accuracy requirement applying to Top-1 of predicted beams and other beams.**  **Observation 6: The whole principle for defining relative RSRP accuracy is that the predicted L1-RSRP minus ground truth beams which is similar to legacy definition.**  **Observation 7: If the reported L1-RSRP is reported measured RSRP, this components will disappear based on current formulation.**  **Proposal 4: The reported L1-RSRP shall be reported predicted L1-RSRP not the measured RSRP.**  **Observation 8: Top-K/1 is the relaxation of Top-1/1, UE vendors has extended margin on this performance metric and the prediction accuracy can be up to 97% even the measurement error has been added.**  **Observation 9: If the margin is set too large, the threshold for performance validation becomes low enough that any UE can pass the test, leading to the realization that AI’s predictive capabilities are actually not as good as legacy methods.**  **Proposal 5: At least x=0 shall be considered in performance metrics.**  **Proposal 6: The concrete value of x when x>0 shall be defined based on simulation results.**  **Proposal 7: The value of x shall not be set too large in order to guarantee the AI performance test and prevent UE easily passes the test.**  **Observation 10: the maximum ground truth beam among Top-K beams has small difference or equal to Top-1 beam if only Top-1 beam is considered. The test will definitely pass. Nonetheless, if the rest of beams in Top-K predicted beams are worse beams, the whole AI/ML performance will be wrongly judged as successful.**  **Proposal 8: RAN4 shall support to introduce Top-N beams for beam ID prediction to ensure high reliability of AI/ML.**  **Proposal 9: Different interpretation of Top-K/N% could be seen as below:**   1. **One of Top-N beams contains in Top-K predicted beams; (like RAN1’s agreement on performance monitoring)** 2. **All of Top-N beams contains in Top-K predicted beams with corresponding order.** 3. **All of Top-N beams contains in Top-K predicted beams without corresponding order.**   **Observation 11: For interpretation A, the test will be passed easily if only one of Top-N in Top-K predicted beams.**  **Observation 12: For interpretation B, compared to interpretation A, this metric will not cause test passes easily, but this metric has large probability that the test will not be passed since the conditions are quite strict.**  **Observation 13: For interpretation C, only guarantee all Top-N beams contain in Top-K beams, Top-N beams are measured by UE which can be assumed as groundtruth beams. That is, if all groundtruth beams contain in AI/ML predicted Top-K beams, the test will be passed.**  **Proposal 10: The definition of Top-K/N% shall be all of Top-N beams contains in Top-K predicted beams without corresponding order.**  **Proposal 11: The potential components in prediction delay contains three parts:**   * **Measurement delay: The time for measurement of SetB which is the input for inference.** * **Inference delay: The time for inference from starting performing inference to generate the prediction results.** * **Reporting delay: The time between the report triggering and the point when the UE starts to transmit the outputs over the air.**   **Proposal 12: For measurement delay, the legacy L1-RSRP measurement period shall be reused.**  **Observation 14: The inference delay maybe different due to model complexity, such as the number of hidden layers, hardware architecture (CPU/GPU) or other parameters of AI/ML models. Thus, the reporting delay including inference delay vary from UE to UE.**  **Proposal 13: For reporting delay, RAN4 shall not define one fixed value and wait for RAN1’s agreement.**  **Proposal 14: The test procedure shall be clarified and unified firstly:**   |  |  | | --- | --- | | **Training phase** | **1.TE configure resource sets for measurement of Set B and beam report of Set A** | | **2.UE measures Set B as the model input.** | | **Inference phase** | **3.Generates the test setup environment for Set B.** | | **4.UE predicts best Tx beam/ L1-RSRP in Set A.** | | **5.UE reports predicted best Tx beam/ L1-RSRP to TE in Set A.** | | **Groundtruth phase (measurement of RSRP with high SNR)** | **6.TE configures resource sets for measurement of Set A.** | | **7.Generates the test setup environment for Set A (including high SNR).** | | **8.UE measures Set A under high SNR condition as ground truth.** | | **9.UE reports groundtruth (measurement on Set A) to TE.** | | **TE evaluation phase** | **10.Evaluates the predicted results and groundtruth and judge the performance degradation** |   **Observation 15: The smaller SNR conditions, the larger measurement error.**  **Observation 16: The performance degradation will be occurred due to added measurement error.**  **Observation 17: For Top-1, the prediction accuracy is below 86% even if there is no error on test data.**  **Observation 18: The prediction is getting larger and larger when K is larger and larger whatever the measurement error is added or not.**  **Proposal 15: RAN4 shall further study the measurement error impact on different performance metrics if new metrics are introduced.**  **Proposal 16: Except measurement error impact, RF error influence shall also be considered in RAN4.** |
| [**R4-2511249**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511249.zip) | Samsung | *Performance Evaluation for AI Beam Management Model*  **Observation-1:** The reference model agreed shows limited performance, and especially showing poor generalization performance with measurement errors added.  **Observation-2:** With measurement errors considered in reference, the performance for Top-1 prediction is lower than reasonable performance even with X = 3dB margin considered (i.e., 42.75% for Case 2a and 38.82% for Case 2b).  **Observation-3:** Even without measurement errors considered in reference, RSRP accuracy is worse than 7dB.  *RRM core requirement for model Inference*  **Proposal 1:** For absolute RSRP accuracy requirement, the requirement shall be defined as:   * For core requirement and conformance testing, the absolute RSRP accuracy requirement shall only be applied to the strongest beam in the configured top-K beam RSRP report. * The existing measurement accuracy requirement is used as the baseline for prediction accuracy.   **Proposal 2:** For relative RSRP accuracy requirement, if defined, the requirement shall be defined as:   * Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - reported L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), where the beam index n owns the largest reported value given by absolute RSRP reporting.   + Reported L1-RSRP can be predicted,   + Reported L1-RSRP shall be predicted RSRP * Relative RSRP accuracy requirements apply for Case 1, and also for Case 2   **Proposal 3:** For the beam ID predication accuracy in Rel-19, only top-1 (i.e., Top-N, N=1) predicted beam is considered in the successful rate evaluation/requirement.  **Proposal 4:** For AI-BM Case-1, the legacy L1 measurement period (the same M, N, P) is used for Set-B measurement to evaluate the achievable prediction accuracy. |
| [**R4-2511570**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511570.zip) | Qualcomm Incorporated | **Observation 1: If K > 1, network has to transmit the top-K predicted beams in a 2nd round of beam sweep and collect UE’s report regarding the strongest beam among top-K predicted beams before changing active TCI state.**  **Observation 2: Currently agreed procedures do not allow network to dynamically configure (e.g., via DCI or MAC-CE) the 2nd round of beam sweep based on UE’s inference results regarding top-K predicted beams.**  **Observation 3: The issue of 2nd round of beam sweep can be avoided by making K = 1.**  **Observation 4: In legacy BM,**   * **The core requirement of L1-RSRP absolute accuracy applies to all reported beams.** * **Performance test of absolute L1-RSRP accuracy only uses reporting of one beam.**   **Observation 5: In legacy BM,**   * **The core requirement of L1-RSRP relative accuracy applies to all reported beams.** * **Performance test of absolute L1-RSRP accuracy only uses reporting of two beams.**   **Observation 6: RAN1 has agreed to increase the minimum time separation between reference signal and aperiodic L1-RSRP report by d’ in AI-ML beam prediction.**   * **Exact value of d’ is TBD in RAN1.**   **Observation 7: RAN4 has agreed to include inference delay in prediction delay of AI-ML BM. This is equivalent to the parameter d’ of RAN1 spec.**  **Proposal 1: Update the previous agreement with the following changes:**   * **“The successful rate for the correct prediction,**    + **The correct prediction is considered as ~~maximum~~ ground-truth RSRP ~~among~~ of top-~~K~~1 predicted beam~~s~~ being larger than or equal to the ground-truth RSRP of the strongest genie-aided beam(s) – x dB,**   + **~~FFS on top-N (N</=K) predicted beams have to be considered in the success rate evaluation~~**     - **~~FFS on N values~~**   + **FFS on x values** * **Note: if x=0 ~~and N=1~~, it means Top-~~K~~1/1 (%) : the percentage of "the Top-1 strongest beam is ~~one of~~ the Top-~~K~~1 predicted beam~~s~~"**   **Proposal 2: RAN4 ensures the following in AI-ML BM case 1,**   * **The core requirement of absolute accuracy of predicted L1-RSRP applies to all top-K predicted beams.** * **Performance test of absolute accuracy of predicted L1-RSRP uses prediction and reporting of one beam, i.e., K = 1.**   **Proposal 3: RAN4 defines relative RSRP accuracy requirement for Case 1 and uses the agreed baseline with the following changes:**   * + **Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i – ~~reported~~ predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), [where the beam index n owns the largest reported value].**     - **~~Reported L1-RSRP can be predicted,~~**     - **~~FFS whether reported L1-RSRP can be measured RSRP~~**     **Proposal 4: RAN4 ensures the following in AI-ML BM case 1,**   * **The core requirement of relative accuracy of predicted L1-RSRP applies to all top-K predicted beams.** * **Performance test of relative accuracy of predicted L1-RSRP uses prediction and reporting of two beams, i.e., K = 2.**   **Proposal 5:**  **Known conditions for TCI state in AI-ML beam prediction:**  **For BM-case 1, the TCI state is known if the following conditions are met:**  **- During the period from the last transmission of the RS configured via resourcesForChannelMeasurement ~~RS resource~~ that gets used to generate ~~for~~ the predicted-L1-RSRP ~~measurement~~ reporting for the target TCI state to the completion of active TCI state switch, where the RS resource ~~for~~ of predicted-L1-RSRP ~~measurement~~ is the RS in target TCI state or QCLed to the target TCI state**  **- TCI state switch command is received within 1280 ms upon the last transmission of the RS resource configured via resourcesForChannelMeasurement ~~for beam reporting or measurement~~**  **~~-~~ UE supports [RxBeamKnownDuringTxBeamBeamPrediction] feature or the RS in target TCI state is known according to clause 8.10.2 or an RS that is QCL-ed Type-D with the RS in target TCI state is known according to clause 8.10.2**  **- The UE has sent at least 1 predicted L1-RSRP report for the target TCI state before the TCI state switch command**  **- ~~The TCI state remains detectable during the TCI state switching period~~**  **- The SSB associated with the TCI state remain detectable during the TCI switching period**  **- predicted SNR of the TCI state ≥ -3dB**  **Otherwise, the TCI state is unknown.**  **Proposal 6: The inference delay mentioned in RAN4 spec should be aligned with d’ in RAN1 spec.**   * **Note 1: d’ is the extension in time separation between RS and AP L1-RSRP report in AI-ML based beam prediction.** * **Note 2: d’ can vary between spatial only and spatio-temporal beam prediction. Exact values of d’ will be discussed on August 2025 meeting in RAN1.** |
| [**R4-2511597**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511597.zip) | Rohde & Schwarz | Proposal 1: (Real Data) Support of real data training and testing is recommended only in the form of 1AoA testing and for R&D purposes, as a complement to the conformance requirement defined with synthetic channels.  Proposal 2: (Real Data) Support of real data training and testing is recommended only in the form of 1AoA testing and for R&D purposes, as a complement to the conformance requirement defined with synthetic channels.  Observation 1: (Single/Multi-AoA) Several aspects related to the UE Rx Antenna assumptions, the input to the AI/ML model and the real data usage for training and testing, have a major impact on the discussion about single- or multi-AoA testing.  Proposal 3: (Single/Multi-AoA) RAN4 to discuss and conclude on the aspects in Table 1 as part of the single-/multi-AoA discussion.  Table 1 – Single vs. multi AoA testing for AI/ML BM with CDL based channel   |  |  |  |  | | --- | --- | --- | --- | | Parameter | Single AoA | Multi AoA | Notes | | CDL Channel Support | Only if UE Rx antenna  performance is emulated as part of the channel (WCM\*) | Direct,  clusters are emulated  on different probes / AoAs | \* Wireless Cable Method | | CDL Channel Complexity | Complex Channel  (with WCM) | Simplified Channel |  | | UE Rx Antenna  Pattern Knowledge | Required for WCM  (reference / declared /  measured) | Not required |  | | UE Rx Antenna  Performance | Static pattern  during testing simulating the dynamic pattern as part of the WCM | Dynamic pattern |  | | UE Rx Beam Sweeping  Timing | Not considered | Considered |  | | UE Rx Beam Index  as input for AI/ML Model | Supported with test mode\* | Supported | \* Depends on the complexity of the UE Rx Beam/Antenna pattern (Codebook) considered for single AoA testing | | Training / testing  with real data | Supported with WCM | Unfeasible  (fixed probe layout) |  | | TE Complexity | Low HW Complexity  Longer test time  WCM & Test mode required | High HW Complexity  Shorter test time |  |   Proposal 4: (Multi-AoA Candidate Setup) Define the implementation of eIFF test setup with 4x active 2x2 MIMO probes with probes (CATRs) placed in a single plane and separated by 30º, 60º and 60º respectively, as feIFF (further enhanced IFF) test setup. The concrete probe layout and the alignment with the coordinate system to be used for testing is FFS.  A screen shot of a video game  Description automatically generatedA screen shot of a cell phone  Description automatically generated  Figure 1 – Example of a commercial implementation of eIFF system with 4 IFF probes / CATRs  Observation 2: (Multi-AoA Candidate Setup) The feIFF test setup supports 4 Active 2x2 MIMO probes with large distinctive angular separation of up to 150°, overall big quiet zone of at least 30 cm, lowest MUs and ensures preservation of (RF-RRM) side conditions.  Observation 3: (Multi-AoA Candidate Setup) The feIFF test setup is a feasible upgrade (superset) of the existing legacy eIFF setup. As such, if feIFF is concluded as capable for AI/ML BM testing, a single system covers the the whole conformance testing, ensuring continuity in the cost sensitive FR2 testing ecosystem.  Proposal 5: (Multi-AoA Candidate Setup) Consider the feIFF test setup as candidate for the AI/ML BM baseline system and analyze its suitability.  Observation 4: (Multi-AoA Simplified Channel) The required total AoA spread of 150° of the simplified channel in Table 2 can be emulated with the feIFF probe layout.  Observation 5: (Multi-AoA Simplified Channel) The simplified channel model parameters for UMi CDL-C can be implemented in the candidate test system feIFF, with 3 possible options (Figure 3): Option A: with flattened elevation (Table 3) and strongest cluster aligned to one probe; Option B: with flattened elevation (Table 3) and mid-point of the total angular spread of the clusters is aligned to the mid-point of the spread of the probes; Option C: with flattened elevation, AoAs aligned to probe layout and no intra‑cluster angle spread (Table 4).  Proposal 6: (Multi-AoA Simplified Channel) Consider the channel models described Table 3 (Option A/B) and Table 4 (Option C) as candidates for CDL-based simplified channel models for multi-AoA testing of AI/ML BM.  Table 3 – Channel model parameters for UMi CDL-C at – Channel model parameters for UMi CDL-C at 28 GHz with flat ZoA   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | Cluster # | Absolute Delay [ns] | Power in [dB] | AOD in [°] | AOA in [°] | ZOD in [°] | ZOA in [°] | | 1 | 0 | -7.4318 | -30.4353 | -134.4434 | 98.9242 | 74.51134 | | 2 | 12.594 | -1.2500 | -20.9269 | 129.1633 | 99.1915 | 74.51134 | | 5 | 13.056 | -5.5318 | -28.0782 | -152.8206 | 99.5732 | 74.51134 | | 6 | 38.196 | 0.0000 | -11.6982 | 164.1145 | 99.306 | 74.51134 | | 13 | 73.71 | -8.1318 | -33.911 | 93.1719 | 100.165 | 74.51134 | | 14 | 78.498 | -9.8318 | -37.5066 | -112.0441 | 100.2604 | 74.51134 | | Per-Cluster Parameters | | | | | | | | Parameter | CASD in [°] | CASA in [°] | CZSD in [°] | CZSA in [°] | XPR in [dB] |  | | Value | 0.799 | 10.4021 | 0.5726 | 0 | 7 |  |   Table 4 – Channel model parameters for UMi CDL-C at 28 GHz  with AoA aligned to probe layout and no intra-cluster angle spread   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | Cluster # | Absolute Delay [ns] | Power in [dB] | AOD in [°] | AOA in [°] | ZOD in [°] | ZOA in [°] | | 1 | 0 | -7.4318 | -30.4353 | **-114.436** | 98.9242 | 74.51134 | | 2 | 12.594 | -1.2500 | -20.9269 | **125.5639** | 99.1915 | 74.51134 | | 5 | 13.056 | -5.5318 | -28.0782 | **-174.436** | 99.5732 | 74.51134 | | 6 | 38.196 | 0.0000 | -11.6982 | **-174.436** | 99.306 | 74.51134 | | 13 | 73.71 | -8.1318 | -33.911 | **95.5639** | 100.165 | 74.51134 | | 14 | 78.498 | -9.8318 | -37.5066 | **-114.436** | 100.2604 | 74.51134 | | Per-Cluster Parameters | | | | | | | | | Parameter | CASD in [°] | CASA in [°] | CZSD in [°] | CZSA in [°] | XPR in [dB] |  | | Value | 0.799 | 0 | 0.5726 | 0 | 7 |  |   A circular graph with different colored circles  Description automatically generated  Figure 3 – Simplified UMi CDL-C at 28 GHz alignment to the candidate test system feIFF  Observation 6: (Multi-AoA Simplified Channel) The evaluation of Options A and B (Table 3) becomes more complex, if the effect of the test setup / probe layout has to be included.  Observation 7: (Multi-AoA Simplified Channel) The emulation of the relative ray phases with the feIFF is highly accurate and should satisfy the requirement for the accuracy for the cluster power fluctuations over time.  Proposal 7: (Multi-AoA Simplified Channel) Further discuss the metric(s) to evaluate channel model implementation with the feIFF candidate test setup.  Observation 8: (Multi-AoA Simplified Channel) For Option C (Table 4), there is no need to optimize the probe weights, and channel model parameters in Table 4 contain all the required information. This option is thus preferred from a TE point of view.  Observation 9: (Multi-AoA Channel Validation) Based on preliminary simulations with 8 Set B & 32 Set A beams, there is only a negligible effect (deviation of max ~1%) on the prediction accuracy for Top 3 & Top 1 beams between the cases with full CDL channel model (Reference - 24 clusters) and simplified CDL channel model suitable for feIFF (Options A/B/C - 6 clusters).  Proposal 8: (Multi-AoA Channel Validation) Consider preliminarily the CDL-C simplified channel models for feIFF with 6 Clusters in Table 3 and Table 4 as suitable for AI/ML BM testing.  Proposal 9: (Multi-AoA Channel Validation) Start the validation of CDL-C simplified channel models for feIFF with 6 Clusters in Table 3 and Table 4 through AI/ML BM system level simulations.  Proposal 10: (Multi-AoA Channel Validation) For the scope of validation, compare the prediction accuracy between the full channel and the simplified channel.  Proposal 11: (Multi-AoA Channel Validation) For the scope of validation, compare the prediction accuracy, when the model is trained with the full channel before the test and with the simplified channel during the test. |
| [**R4-2509134**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509134.zip) | Anritsu Corporation | **Observation 1: TR 38.843 and the WID for AI/ML air interface is assuming both BM-Case 1 and BM-Case 2.**  **Proposal 1: Clarify whether the test setup for the beam management feature shall support both BM-Case 1 and Case 2 tests.**  **Observation 2: Considering the UE behavior in the actual field, there might be a need to conduct tests with mixed conditions of BM-Case 1 and Case 2 and run the tests in series. For example, we conduct a test with BM-Case 1 condition as a first step and then another test with BM-Case 2 condition follows later based on the predicted results from the first BM-Case 1 test. Additional tests with BM-Case 1 condition might also be run at some points during the BM-Case 2 test.**  **Proposal 2: Companies are encouraged to bring views on the necessity of such mixed conditional tests. (i.e. BM-Case 1 and/or BM-Case 2)**  **Observation 3: Since the group has not completed analyzing which type of channel model can ensure the field performance of the AI model, it is not clear at this moment whether the legacy CDL channel model is necessary or the simplified CDL model can suffice.**  **Observation 4: Depending on the results of the required channel model analysis, the test system configuration and test procedures may be significantly affected.**  **Observation 5: In a real field propagation path, there can be variations in the direction of arrival of each path (direction of clusters), etc. Therefore, a test result may not fully reflect its performance in the real field when an AI/ML BM-enabled UE is tested using only one channel model.**  **Proposal 3: Study the necessary channel models to ensure the prediction performance of the AI model in the field. (Legacy CDL or Simplified CDL, if we need to assume a situation that UE is moving, etc.)**  **Proposal 4: Define multiple channel models like the existing CDL models (e.g. CDL-A/B/C) to ensure the field performance of the AI model for beam management. FFS for the necessary number of channel models.**  **Observation 6: According to the previous RAN1 agreement, it is still up to UE implementation if the UE inputs its Rx beam information into its AI model for inference of the Tx beam ID.**  **Proposal 5: Confirm the latest assumptions among companies in RAN4 regarding valuations of Rx beam information to be input to the AI model per measurement of one Tx beam ID.**  **Observation 7: If multiple channel models are necessary to evaluate by multi-AoA test setup, or in a case the group identifies the performance issue with the simplified CDL model, there will be a testability issue.**  **Observation 8: Although there is a benefit that 1AoA test setup has almost no restrictions on the channel model (or number of clusters) to simulate by fading simulator, it may not be able to handle a case where UE sweeps its Rx beam during the tests.**  **Observation 9: Link level simulation in RAN1 appears to apply the assumption of the channel model described in TR 38.901 clause 7.6.3.2 [4], i.e. Spatially-consistent UT/BS mobility modelling.**  **Observation 10: We need to study how we can create the channel environment by the candidate test setups in a case that the UE moves from one point to another.**  **Observation 11: For BM-Case 2, it is expected that there are testability issues with both the 1AoA setup and multi-AoA setup.**  **1AoA setup**   * **The 1AoA test setup may not be able to treat a situation where UE sweeps its Rx beam.**   **Multi-AoA setup**   * **Channel models described in TR 38.901 clause 7.6.3.2 (i.e. Spatially-consistent UT/BS mobility modelling.) might be difficult to reproduce.** * **Positions and directions of clusters cannot be changed since measurement probes are fixed in the OTA chamber.**   **Observation 12: Depending on the outcome of the study, there may still be cases where there is no comprehensive test solution to evaluate BM-Case 1 and/or BM-Case 2 at this moment.**  **Proposal 6: Leave an option that we conduct the evaluation of these beam management AI performances by two steps, first we evaluate only the prediction performance of AI model without the Rx beam sweeping, and second, we evaluate the UE behavior of the Rx beam sweeping.** |
| [**R4-2509242**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509242.zip) | NTU | Based on the discussion in the previous meeting, network vendors want to check N largest RSRP (according to ground truth) beams on top of the (one) largest RSRP (according to ground truth) beam. We propose an enhanced version below:  **Proposal 1: consider the following performance metrics if RAN4 decides K>1:**  **For all *n<=N***  **Where ’s are the *N* largest ground truth RSRPs among all the beams, ’s are the predicted RSRPs, is the set of the index of the predicted largest *K* beams. We need to have *K*>=N, and preferably the difference is larger than 2 or 3.**  This proposal is similar to the ones discussed in the previous meeting, but with better formulation to avoid misunderstanding. With the above proposal, we can check whether among all the reported best beams, there are at least one beam that has RSRP close to the ground truth top-N beams. |
| [**R4-2509300**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509300.zip) | CATT | **Observation 1: Different performance requirements need to be separately defined for beam ID prediction and RSRP prediction.**  **Proposal 1: For AI/ML based beam management with both predicted RSRP and beam ID reported, two options can be considered for defining performance requirements:**   * **Option 1: The legacy L1-RSRP performance requirements are reused, regardless of overhead reduction ratio.** * **Option 2: The legacy L1-RSRP performance requirements are relaxed with margins, and larger margin values are used for higher overhead reduction ratios.**   **Proposal 2: For AI/ML based beam management with predicted beam ID reported only, RAN4 can define a lower bound for the successful rate as performance requirement regardless of the values of K, N, x and SINR.** |
| [**R4-2509427**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509427.zip) | Apple | 1. **When we pick the Top-1 predicted beam, we are not just selecting the strongest candidate, we are also favoring the one that happened to be overestimated due to random noise in the predictions. This means the “winner” often looks slightly better in the test results than it is really in reality.** 2. **If we could truly guarantee that the predicted RSRP accuracy in AI/ML is the same as the measured RSRP accuracy in legacy, for every predicted beam the UE reports**, **then in principle, a separate beam prediction accuracy metric would not be strictly necessary.** 3. RAN4 is invited to adopt and focus on a modified Case 3a simulation setup using error-free ground truth labels to isolate model performance from training noise. Additionally, we propose plotting full error PDFs (e.g., RSRP\_pred − RSRP\_true) to identify bias, skewness, or non-Gaussian behavior, rather than relying solely on percentile metrics. This supports robust AI/ML evaluation and simulation alignment. 4. In order for RAN4 to meet the P90(|error|) ≤ 6 dB requirement (FR2), it is proposed to (i) reduce SNR-sensitive errors by ensuring higher-SNR side conditions for the reference signals used in AI/ML inputs, (ii) further improve SNR through measures such as RS power boosting or 1–2 non-coherent repetitions, (iii) revisit RF assumptions by defining deployment RF profiles with tighter variance where feasible, and (iv) apply noise-aware training with augmentation matched to the declared RF/LLS error models to improve model robustness and reduce the baseline prediction error. 5. For UEs reporting only beam IDs, we propose an evaluation framework for beam index prediction accuracy in which the UE predicts a Top-K set of beam candidates via AI/ML, and the final ordering is determined offline using genie RSRP values. The metric checks whether the UE’s predicted Top-K set includes the correct beams, in the offline-determined order, for multiple beam positions, within an allowed RSRP margin (X) from the corresponding strongest genie beams. 6. RAN4 to adopt a beam prediction accuracy metric based on a RSRP threshold model x = x₀ + Δ, where x₀ is the acceptable performance gap between predicted and genie beams, and Δ accounts for RF-induced uncertainty in RSRP measurements and it should be derived statistically from simulations modeling realistic RF error distributions and correlation across Tx beams. 7. We propose setting **N > 1** when evaluating Top-K beam prediction accuracy to reflect realistic gNB usage of multiple fallback beams. Considering more than one strong beam improves robustness, and avoids overfitting to a single beam index. This approach aligns with diverse deployment scenarios and enables more resilient AI/ML evaluation. 8. **RSRP prediction accuracy should not be a test requirement in Case 2**.   A diagram of a test procedure  AI-generated content may be incorrect.   1. RAN4 is invited to adopt a unified AI/ML beam management testing framework that categorizes UE performance outcomes according to the reliability of both RSRP prediction accuracy and beam ID prediction accuracy. If a UE reports only beam IDs, Case 2 is applied, where success is evaluated using the weighted Option 3 formula. If a UE reports RSRP values and its RSRP prediction accuracy meets or exceeds legacy measurement accuracy, only RSRP prediction accuracy is tested. If RSRP accuracy cannot be guaranteed, testing falls back to Case 2 to ensure robustness in beam selection 2. RAN4 to adopt **beam ID prediction accuracy** (Case 2) as the sole conformance metric when the UE reports only beam indices. The test shall be performed offline using the genie RSRP values for all beams in Set A, with no re-measurement required. For each snapshot, the genie Top-N beams are identified, and the UE’s Top-K predicted beams are compared against them. A predicted beam is considered a match to a genie beam if its genie-RSRP is within x = x₀ + Δ dB of the genie beam’s RSRP. 3. For **Case 2** (Probe-Ranked Top-K), we recommend **K > N**, allowing the UE to provide a broader beam pool for post-measurement ranking. A slightly larger K (e.g., K = N+1 or N+2) improves robustness and fallback without adding significant overhead during test 4. RAN4 is invited to study the model robustness across scenarios of AI/ML models for beam management (BM) by evaluating three training approaches: (1) a universal model trained on a combined dataset of multiple deployment scenarios, (2) stand-alone models trained per scenario, and (3) a shared backbone model with lightweight scenario-specific experts. The study will assess performance on both training scenarios (Seen) and scenarios not used in training (Unseen) through systematic testing across diverse deployments. The outcome will clarify whether BM requires many specialized models or if a shared architecture with minimal specialization can handle realistic deployment variations, guiding model deployment strategies and standardization in RAN4. |
| [**R4-2509762**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509762.zip) | Xiaomi | **Observation 1: The similar metric Top-K/M is agreed in RAN1 performance monitoring, where at least one of the Top M beam(s) of the resource set(s) for monitoring is among Top-K predicted beam(s).**  **Proposal 1: In order to align with RAN1, Beam index prediction definition should at least include original option 2:**   * **For N value:**    + **When X = 1 dB, N = 1.**   + **When X = 0dB, N = 1 or 2 depending on K value.**   **Observation 2: In legacy, absolute RSRP is only defined for the largest RSRP and relative accuracy is defined for other differential RSRP in the report.**  **Proposal 2: Only define Absolute RSRP accuracy requirement for Top-1 of predicted beam.**  **Proposal 3: For relative predicted RSRP accuracy, define unified definition for both BM-case 1 and BM-case 2:**  **Where is the largest RSRP across all time instance and beams in the report。 is RSRP of beam index i at time t。 is the true RSRP corresponding to the beam/time of 。Where is predicted differential RSRP in the report.**  **Observation 3: Significant Performance Degradation in Low SNR Groups. The performance degradation is more significant in UE groups with lower maximum SNR.**  **Observation 4: Higher Stability in High SNR Groups. UE groups with higher maximum SNR (e.g., UE group 1 and UE group 2) show better performance stability under different measurement errors.**  **Observation 5: The SINR distribution for all UE groups spans a large range, and this distribution can significantly impact the final accuracy performance.**  **Proposal 4: RAN4 to align simulation results without measurement error first.**  **Proposal 5: To ensure consistent and reliable simulation results, we recommend that RAN4 align the following factors:**   * **1. Method of Adding Measurement Error for Different TX Beams：**   + - **Generate error distributions separately for each individual Tx-Rx beam pair based on different SNRs and apply different Gaussian standard deviations to the corresponding Tx-Rx beam pair measurements in the SLS.** * **2. Gaussian Standard Deviation for Different SNRs**   + - **Establish reference standard deviation values for critical SNR operating points.** * **3. Largest SNR Distribution Among All Selected UEs**   + - **Divide the largest SNR range into several groups and define reference percentages for each group.** * **4. SNR Distribution for TX Beams in Set B**   + - **Align the SNR distribution for all TX beams in set B to ensure consistency across different companies.** |
| [**R4-2509932**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509932.zip) | Nokia | 1. Based on simulation results, RAN4 to define values of and the success rate of beam id prediction for performance requirement of BM-case1 and BM-case2 2. RAN4 to discuss the adaptation of Option 3 with the following changes: 3. At least strongest beams among the top- predicted beams must be considered in the success rate evaluation, where    * RAN4 should consider N=2 if there is no consensus in RAN4 for N>2. 4. RSRP of each of the strongest predicted beams must be compared with the RSRP of its equal ranked genie-aided beam – X 5. Allow the evaluation of Beam Id prediction accuracy as a weighted success rate using the following formulation and weighing factors   **with RSRPs in linear scale**   1. RAN4 should consider defining a test case for the measurement period for L1-RSRP measurement for BM-Case2, where the parameter M needs to be redefined. 2. RAN4 to consider having a Test case for verifying the core requirements for LCM operation related to activation/switching delay requirements. 3. RAN4 to consider having Test case for verifying the core requirements for LCM operation related to deactivation delay requirements. 4. RAN4 to consider having Test case for verifying the core requirements for Conditions for known TCI state and TCI state switch delay requirements. 5. RAN4 to agree on using multi-AoA based test systems for the conformance testing of AI/ML based BM use case. 6. RAN4 to discuss and explore the feasibility of generating Set A and Set B beams during the conformance testing of the AI/ML-BM use case by selecting unique combinations of available test probes and decide the transmit power of each selected probe. Tx power of each selected probe is FFS. 7. RAN4 to discuss and decide the contents of pre-alignment check which needs to be performed between the TE and DUT before the exact start of the conformance test. 8. RAN4 to discuss the feasibility of performing beam sweeping using UE rotation in multiple iterations during the conformance testing of AI/ML based BM use case. 9. RAN4 to discuss the possibility of testing AI/ML BM use-case considering (1) No transmission of Set A beams during the test, (2) exploit and reuse the characteristics of the testing environment between the test probes and UE baseband during the test. 10. We need to check what is the directional channel based on beam patterns, and how many clusters those directional channels interact with in general. The clusters numbers will be greatly reduced compared with the number of clusters in current CDL models. 11. We can use the spatial filter corresponding to the beam width of SSB to filter out the clusters of omni-directional CDL channel models (current CDL models). The spatial filter is defined in 38.901 Chapter 7.7.4 Spatial filter for generating TDL channel model to generate TDL model, so, we need to modify the procedure to keep the angular information of each cluster. 12. Use spatial filter to find out how many clusters in the directional model for SSB before doing simplification of CDL model. 13. Bear in mind that removing the clusters of a CDL model for simplification will change the channel properties. Or we can use the alternatives that use ray tracing or other simulation to simulate directional channels with 4 clusters directly. 14. If the number of clusters interact with each SSB/CSI-RS is less than 4 (or the number of clusters the probes of TE can simulate at the same time), we may only need to divide 24 clusters into “p” sub-groups and each sub-group match the clusters one SSB/CSI-RS interacts with. 15. RAN4 to consider designing a new metric to verify the generalization performance of the AI/ML BM functionality in different (static) scenarios. 16. For the verification/testing of generalization related aspects in RAN4 for AI/ML BM, RAN4 should define the generalization test which consists of multiple subtests each associated with the testing of different scenarios based on parameters listed in Table 1. |
| [**R4-2510160**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510160.zip) | CMCC | ***Proposal 1: For RSRP accuracy for AI/ML based beam management, it is proposed that AI/ML based performance requirement should not be worse than the legacy measurement accuracy requirements.***  ***Proposal 2: According to RAN1 agreements, it is proposed that legacy L1 report mapping, i.e. L1 part of Table 10.1.6.1-1 and Table 10.1.6.1-2 in TS38.133 can be reused for the report mapping of inferenced L1-RSRP.***   * ***The absolute L1-RSRP is is quantized to a 7-bit value in the range [-140, -44] dBm with 1dB step size*** * ***The differential L1-RSRP is quantized to a 4-bit value with 2 dB step size*** |
| [**R4-2510339**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510339.zip) | vivo | **Proposal 1: For RRM performance requirements for beam management, RAN4 to consider the following requirements；**   * **Beam prediction accuracy** * **Absolute Predicted RSRP accuracy** * **Relative Predicted RSRP accuracy** * **Accuracy requirements for performance monitoring(to verify whether the UE correctly reports the number of accurate predictions out of N inference attempts.)**   **Proposal 2: For beam prediction accuracy and RSRP accuracy, when defining specific performance requirements, it is necessary to determine how many different sets of metric requirements should be defined to accommodate various scenarios. RAN4 to first identify the factors that influence prediction performance, potentially including:**   * **The number of set B beams, the number of set A beams**   + **candidate value:**      - **With measurements of Set B of beams that of 1/4 of Set A of beams (e.g., 8 for set B, 32 for set A)**     - **With measurements of Set B of beams that of 1/8 of Set A of beams (e.g., 8 for set B, 64 for set A)** * **Beam pattern**    + **Consider the beam pattern defined in Table 6 of the existing simulation assumptions [1]** * **Prediction scenario:**    + **Wide beam to narrow beam (SSB to CSI-RS)**   + **narrow to narrow (CSI-RS to CSI-RS)** * **The value of N**   + **candidate value: 1,2,3,4** * **The value of X**   + **0,1,2,3**   + **Specific values should be determined based on N and simulation results**   **Proposal 3: For test cases for AI/ML based beam management, RAN4 to define RRM test cases for the following**  **requirements:**   * **Prediction delay requirement** * **Performance monitoring reporting delay requirements** * **Functionality activation delay requirement** |
| [**R4-2510808**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510808.zip) | OPPO | **Proposal 1: The definition of top-N needs further clarification to determine whether additional top-N evaluation is required.**  **Proposal 2: If both RSRP and beam ID are reported, but not for all beams in BM-set A, both Beam ID prediction accuracy and RSRP accuracy of the reported beams should be considered in RAN4 tests.**  **Proposal 3: If both RSRP and beam ID are reported, for all beams in BM-set A, RAN4 could only evaluate the RSRP accuracy of all beams.**  **Proposal 4: Absolute RSRP accuracy requirement applies to Top-1 of predicted beams, relative RSRP accuracy requirement could be applied to other predicted beams with differential RSRP reporting in RAN4 tests.**  **Observation 1: Regarding the impact of measurement error on BM performance evaluation, the following aspects should be further considered**   * **Whether the impact of measurement error obtained from the SLS UMA channel (measuring error added from a CDL channel evaluation) can be directly applied to the OTA testing environment** * **Differences in implementation (e.g., BS antenna configurations, BS antenna element radiation pattern, Beamforming characteristic of the BS pattern) among companies may result in inconsistent evaluation results** |
| [**R4-2510875**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510875.zip) | Huawei,HiSilicon | **Observation 1:** With Gaussian distribution assumption of baseband measurement error, under different SNRs, the mean values are all close to 0, while the variance increases as the SNR drops.  **Proposal 1:**The simulation results for different prediction accuracy regarding to KPI 1 is shown in Table 1.  **Table 1. Measurement error impact on prediction accuracy regarding to KPI 1**   |  |  |  |  | | --- | --- | --- | --- | | KPI-1 | | Model input w/o measurement error | Model input w measurement error | | maximum RSRP among top-K predicted beams is larger than the RSRP of the strongest beam – x dB | x = 1dB | 83.02% | 60.75% | | x = 2dB | 92.49% | 72.87% | | x = 3dB | 96.29% | 81.84% |   **Observation 2:**The more stringent the KPIs are, the greater the impact of the measurement error on the prediction accuracy will be.  **Proposal 2:** The simulation results for different prediction accuracy regarding to KPI 2 is shown in Table 2.  **Table 2. Measurement error impact on prediction accuracy regarding to KPI 2**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | KPI-2 | | | Model input w/o measurement error  (case 1) | Model input w measurement error  (case 2b) | | 90%-tile L1-RSRP difference between the maximum RSRP of the Top-1/ Top-3/ Top-5 predicted beam(s) and the ground truth L1- RSRP of the genie aided strongest beam | Avg. | CDF\_strong\_1 | 2.547dB | 5.871dB | | CDF\_strong\_3 | 2.345dB | 5.392dB | | CDF\_strong\_5 | 2.209dB | 4.893dB | | w/o avg. | CDF\_strong\_Top1 | 2.547dB | 5.871dB | | CDF\_strong\_Top2 | 2.380dB | 5.492dB | | CDF\_strong\_Top3 | 2.139dB | 4.781dB | | CDF\_strong\_Top4 | 2.049dB | 4.445dB | | CDF\_strong\_Top5 | 1.967dB | 4.044dB | | 90%-tile L1-RSRP difference between the predicted L1-RSRP of the Top-1/ Top-3/ Top-5 predicted beam(s) and the ground truth L1-RSRP of the same beams | Avg. | CDF\_same\_Top1 | 2.151dB | 7.250dB | | CDF\_same\_Top1&2&3 | 2.236dB | 7.406dB | | CDF\_same\_Top1&2&3&4&5 | 2.419dB | 7.444dB | | w/o avg. | CDF\_same\_Top1 | 2.151dB | 7.250dB | | CDF\_same\_Top2 | 2.200dB | 7.314dB | | CDF\_same\_Top3 | 2.435dB | 7.702dB | | CDF\_same\_Top4 | 2.674dB | 7.730dB | | CDF\_same\_Top5 | 2.689dB | 7.287dB |   **Observation 3:** Regarding RSRP prediction accuracy, measurement errors have a larger impact on the Top 1 predicted beam than on the Top 5th predicted beam.  **Proposal 3:** RAN4 define requirements with measurement error as model input.  **Proposal 4:** RAN4 will not select KPI under which acceptable performance cannot be achieved, including   * the KPI under which the prediction accuracy is less than 80%, e.g, Top K/1 with K<4 without RSRP margin. * the KPI under which the 90%-tile L1-RSRP difference is larger than 5dB, e.g, maximum RSRP among top-K predicted beams is larger than the RSRP of the strongest beam – x dB with x<3dB. |
| [**R4-2511574**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511574.zip) | Qualcomm Incorporated | **Observation 1: RAN4 is currently focusing on system level channel, e.g., dense urban macro, to simulate and investigate the performance of spatial only beam prediction.**  **Observation 2: AI-ML beam prediction requirement defined based on system level channel is more realistic for field deployment.**  **Observation 3: RAN4 has not calibrated system level channel for FR2 OTA testing. RAN4 has calibrated UMi CDL-C channel for FR2 MIMO OTA testing and captured it in Table D.1-1 of 38.151.**  **Observation 4: RAN4 has already agreed to focus on CDL based channel model as a starting point for testing.**  **Observation 5: Baseband error will reduce at high SNR.**  **Observation 6: RF error may not decrease at high SNR because UE’s Rx AGC settings might be calibrated for a particular SNR value.**  **Observation 7: Ground truth will be based on UE’s report and UE’s reported RSRP will include RF error that may not decrease with SNR.**  **Observation 8: Baseband error at different SNRs can be modelled with Gaussian distribution. For example, at -3 dB SNR, the distribution of absolute baseband measurement error can be fitted with following Gaussian distributions:**   * **AWGN: Gaussian with zero mean and sigma = 0.6 dB** * **TDL-C: Gaussian with mean = -0.12 dB and sigma = 0.85 dB**   **Table 2: Definitions of Metrics for AI-ML based Beam Prediction**   |  |  | | --- | --- | | Metric | Definition | | **L1-RSRP absolute accuracy** | 90%-tile L1-RSRP difference between the predicted L1-RSRP of the Top-1 predicted beam and the ground truth L1-RSRP of the same beam. | | **Worst case L1-RSRP absolute accuracy among top-K predicted beams** | 90%-ile of the worst case L1-RSRP difference between the predicted L1-RSRP of the top-K predicted beams and the ground truth L1-RSRP of the same beams.  This gets calculated in following steps:   * In one sample of inference reporting, assume I = arg max\_{k \in K} (predicted RSRP of beam k – ground truth of beam k) * Store (predicted RSRP of beam I – ground truth of beam I) for this sample of inference * Report 90%-ile distribution of above metric | | **Relative L1-RSRP of beam owning the 2nd largest predicted reported value** | 90%-ile of following metric:  (predicted L1-RSRP of beam index i - predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), [where the beam index n owns the largest reported value]. | | **Top K/1 without margin** | The percentage of the time when the Top-1 strongest beam (ground truth based) is one of the Top-K predicted beams | | **Top 1/K without margin** | The percentage of the time when the Top-1 predicted beam is one of the Top-K strongest beams (ground truth based) | | **Top K/N (m) without margin** | The percentage of the time when the Top-K predicted beams contain m of the top-N strongest beams (ground truth based) | | **Top 1/1 with margin of X dB** | The successful rate for the correct prediction which is considered as maximum ideal RSRP among top-1 predicted beams is larger than the ideal RSRP of the strongest genie-aided beam – x dB |   **Table 3: Results of AI-ML based “spatial only” beam prediction (placeholder to be filled by the meeting week)**   |  |  |  |  | | --- | --- | --- | --- | | Metrics | | Training and testing without measurement error | Training and testing dataset, along with ground truth for training and testing, with measurement error | | L1-RSRP absolute accuracy (dB) | |  |  | | Worst case L1-RSRP absolute accuracy among top-K predicted beams (dB) | K = 2 |  |  | | K = 3 |  |  | | L1-RSRP relative accuracy of beam containing the 2nd largest predicted value | |  |  | | Top K/N (m = 2) without margin (%) | 2/2 |  |  | | 2/4 |  |  | | 3/2 |  |  | | 3/4 |  |  | | 4/4 |  |  | | 5/4 |  |  | | Top K/1 without margin (%) | K = 1 |  |  | | K = 3 |  |  | | K = 5 |  |  | | Top 1/1 with margin of X dB (%) | X = 1 |  |  | | X = 3 |  |  | | X = 5 |  |  |   **Proposal 1: RAN4 agrees CDL-based channel model to test AI-ML BM performance.**  **Proposal 2: Accuracy of AI-ML BM-case 1 is defined based on the worst-case performance in following two scenarios:**   * **Dense urban macro**    + **Note: This is currently under consideration and captured in the recently agreed simulation assumption of** [**R4-2508081**](http://10.10.10.10/ftp/RAN/RAN4/Inbox/R4-2508081.zip) * **CDL-based channel model (details are TBD)**   + **Note: The UMi CDL-C channel of Table D.1-1 of 38.151 can be used as a starting point (shown in Table 1)**   **Table 1: Channel model parameters for UMi CDL-C at 28 GHz**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Cluster #** | **Absolute Delay [ns]** | **Power in [dB]** | **AOD in [°]** | **AOA in [°]** | **ZOD in [°]** | **ZOA in [°]** | | **1** | **0** | **-4.4215** | **-30.4353** | **-134.4434** | **98.9242** | **83.3318** | | **2** | **12.594** | **-1.25** | **-20.9269** | **129.1633** | **99.1915** | **72.5229** | | **3** | **13.314** | **-3.4684** | **-20.9269** | **129.1633** | **99.1915** | **72.5229** | | **4** | **13.974** | **-5.2294** | **-20.9269** | **129.1633** | **99.1915** | **72.5229** | | **5** | **13.056** | **-2.5215** | **-28.0782** | **-152.8206** | **99.5732** | **71.1282** | | **6** | **38.196** | **0** | **-11.6982** | **164.1145** | **99.306** | **74.7544** | | **7** | **38.688** | **-2.2185** | **-11.6982** | **164.1145** | **99.306** | **74.7544** | | **8** | **39.36** | **-3.9794** | **-11.6982** | **164.1145** | **99.306** | **74.7544** | | **9** | **39.504** | **-7.4215** | **17.3861** | **84.3647** | **100.4513** | **69.2454** | | **10** | **47.61** | **-7.1215** | **-37.5865** | **92.0623** | **98.5616** | **66.7349** | | **11** | **49.278** | **-10.7215** | **20.2226** | **-97.7585** | **100.6231** | **72.0348** | | **12** | **56.016** | **-11.1215** | **-50.6106** | **78.4702** | **98.218** | **64.4337** | | **13** | **73.71** | **-5.1215** | **-33.911** | **93.1719** | **100.165** | **85.4238** | | **14** | **78.498** | **-6.8215** | **-37.5066** | **-112.0441** | **100.2604** | **64.1548** | | **15** | **130.224** | **-8.7215** | **-43.1797** | **102.4645** | **98.1225** | **64.7824** | | **16** | **162.63** | **-13.2215** | **29.2116** | **67.2359** | **100.2604** | **92.467** | | **17** | **255.534** | **-13.9215** | **27.8133** | **34.5731** | **98.4852** | **65.6889** | | **18** | **276.018** | **-13.9215** | **23.6584** | **48.5813** | **98.1416** | **68.7572** | | **19** | **329.412** | **-15.8215** | **-52.5282** | **36.4455** | **97.9698** | **59.1339** | | **20** | **336.462** | **-17.1215** | **25.0168** | **52.6729** | **100.7376** | **65.3402** | | **21** | **378.39** | **-16.0215** | **25.4562** | **49.8296** | **98.1225** | **58.4365** | | **22** | **398.244** | **-15.7215** | **30.7697** | **46.4316** | **98.1034** | **65.2705** | | **23** | **422.562** | **-21.6215** | **35.9234** | **30.759** | **100.4513** | **62.6903** | | **24** | **519.138** | **-22.8215** | **-61.2775** | **69.2469** | **100.9476** | **61.993** | | **Per-Cluster Parameters** | | | | | | | | **Parameter** | **CASD in [°]** | **CASA in [°]** | **CZSD in [°]** | **CZSA in [°]** | **XPR in [dB]** |  | | **Value** | **0.799** | **10.4021** | **0.5726** | **4.8814** | **7** |  |   **Proposal 3: RAN4 allows each company the flexibility to use individual RX beamforming codebook while reporting simulation results.**  **Proposal 4:**   * **RAN4 defines performance based on case 3 of last meeting where measurement error will be considered in training dataset, model input during inference and ground-truth during training and testing.** |

## Open issues summary

Previous agreements on this topics are listed below for reference (R4-2505105):

**Issue 2-1: Metrics/KPIs for beam ID prediction**

**Agreement:**

* Metrics/KPIs for beam ID prediction, at least for the case if only beam ID is reported:

The successful rate for the correct prediction,

* + The correct prediction is considered as maximum ground-truth RSRP among top-K predicted beams larger than or equal to the ground-truth RSRP of the strongest genie-aided beam(s) – x dB,
  + FFS on top-N (N</=K) predicted beams have to be considered in the success rate evaluation
    - FFS on N values
  + FFS on x values

Note: if x=0 and N=1, it means Top-K/1 (%) : the percentage of "the Top-1 strongest beam is one of the Top-K predicted beams"

**Issue 2-2: Absolute RSRP accuracy**

Agreement:

* Absolute RSRP accuracy requirement applies to Top-1 of predicted beams at least
  + FFS whether to apply to other beams

**Issue 2-3: Relative RSRP accuracy**

**Agreement:**

* RAN4 to use the follow as baseline for further discussions, if the relative RSRP accuracy requirement is defined
  + Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - reported L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), [where the beam index n owns the largest reported value].
    - Reported L1-RSRP can be predicted,
    - FFS whether reported L1-RSRP can be measured RSRP
  + Relative RSRP accuracy requirements apply for Case 1, and FFS on whether it can apply for Case 2

**Issue 2-4: Measurement period for inference**

Agreement:

* For measurement delay component of the prediction delay:
  + Use “legacy” L1-RSRP measurement as the baseline for Case 1
    - FFS whether same values for M,N,P should be reused
  + FFS for case 2

**Issue 2-6: Test system setup**

Companies are invited to bring further analysis on the test system setup.

At least the following aspects to be considered in the analysis:

* + - Whether the test can/should reflect performance in the field and whether field data should be usable in the test
    - What channel model should be used
    - UE prediction model input (e.g. whether any spatial information like UE Rx beam, UE beam pattern, etc is used)

Pros and cons of proposed single AoA and multi AoA setups

The agreements from RAN4#115 are listed below (R4-2508080)

**Issue 2-2: Measurement period for inference**

Agreement:

Case 1: The starting point is to use M, N, P (reference the values from an existing clause in the same version of the spec where the measurement period is defined)

Case 2: postpone discussion until RAN1 concludes

**Issue 2-3: Rx beam knowledge and TCI handling**

Agreement:

* + if the predicted Tx beam in Set A is QCL Type-D to a known Tx beam, the corresponding Rx beam is known.
  + If the predicted Tx beam in Set A is not QCL Type-D to a known Tx beam, depending on UE capability, the corresponding Rx beam may be known or unknown.
  + Known Rx beam means no additional Rx beam sweeping is needed
  + Unknown Rx beam means additional Rx beam sweeping is needed.
  + Rx beam is known mean TCI state is known if conditions below are also met

The following conditions related to whether the TCI state is known or unknown is defined as the starting point. Companies are invited to bring analysis.

* The target TCI state is known when the following is met:
  + The TCI state switch command is received within 1280 ms upon the last transmission of the RS resource for beam reporting or measurement and the UE has sent at least one L-1 RSRP report of the target TCI state (clause 8.10.2 of 38.133)
  + Applies for case 1
  + FFS for case 2
  + Whether RS resource is only from set B should be clarified
    - FFS on whether prediction also impacts if a TCI state is known or unknown
      * Companies are invited to bring analysis on how prediction would impact whether a TCI state is known or unknown
      * FFS on the time delay if prediction is used as a condition for a known TCI state

The open issues were grouped in the following sub-topics for further discussion:

1. Measurement period for inference
2. Prediction report delay
3. TCI state handling
4. Activation delay
5. Simulation results and next steps
6. Test system basic setup
7. Test system channel model
8. Draft CR for 38.133
9. Prediction KPIs
10. Relative RSRP definition

### Sub-topic 2-1

*Measurement period for inference*

Some agreements were made for case 1 but case 2 was still TBD. As this requirement will impact the core, it is important that progress is made

**Issue 2-1: Measurement period for inference**

Proposals

* + Option 1: reuse existing measurement period (existing M, N, P) for both case 1 and case 2
    - Agree the baseline requirements, discuss if this should be combined with any other features in future meetings
  + Option 2: reusing existing measurement period (existing M,N,P) for case 1, introduce a scaling factor T to increase the delay for case 2
    - T defined as a capability or dependent on M, N, P
  + Option 3: other options
* Recommended WF
  + Option 1

From a network operation point of view, it seems best to have consistent delays irrespective of which case is used.

If option 2 is preferred, how to define the additional scaling factor should be further discussed

Moderator recommends we do not combine the BM prediction feature with any other feature because there would be no time to finalize the requirements and specification would become unnecessarily complex.

### Sub-topic 2-2

*Prediction report delay*

For the final requirement, a timeline between measurement and the final report should be established

**Issue 2-2: Prediction report delay**

* Proposals
  + Option 1: Overall timeline to be measurement delay + inference delay + report
  + Option 2: others
* Recommended WF
  + Option 1

Inference delay should be further discussed. For now can be introduced a parameter but some numbers will be needed to finalize the requirements

ZTE: based on RAN1 discussion, inference delay can be part of reporting delay.

Xiaomi:it should be the same activation delay.

Apple:measurement delay covers all the measurement for inference. Reporting delay incude inference delay and time to make report.

Qualcomm:describe RAN1 design

Agreement:

* + For aperiodic report, the overall prediction reporting delay at least includes measurement delay +reporting delay
    - The reporting delay includes the inference delay
  + For semi-persistent and periodic report, the overall prediction reporting delay at least includes measurement delay +inference delay+time for the first available reporting occasion

### Sub-topic 2-3

*TCI State Handling*

In the previous meeting several agreements were reached regarding the handling of TCI state switching relative to whether it is known/unknown and the UE Rx beam knowledge. The agreements are listed above.

**Issue 2-3: TCI State Handling**

* Agreement:
  + On Detectability and SNR conditions:
    - The UE has sent at least 1 L1-RSRP inference report for the target TCI state before the TCI state switch command
    - The TCI state remains detectable during the TCI state switching period
    - The SSB associated with the TCI state remain detectable during the TCI switching period
      * SNR of the TCI state ≥ -3 dB
* Proposals for known TCI state
  + Detectability and SNR conditions:
    - Option 1:
      * The UE has sent at least 1 L1-RSRP inference report for the target TCI state before the TCI state switch command
      * The TCI state remains detectable during the TCI state switching period
      * The SSB associated with the TCI state remain detectable during the TCI switching period
        + SNR of the TCI state ≥ -3 dB
      * For BM case-1 and case-2, when the target TCI state is predicted (in Set A) and is not QCL-D to any previously measured RS, there is no physical RS transmission for the target. Therefore, the conditions of "detectable RS" and "SNR ≥ -3dB" cannot be applied and should be exempted.​
    - Option 2: others
  + Time conditions:
  + Agreement:
    - TCI state switch command is received within 1280 ms upon the last transmission of the RS resource for beam predict reporting or measurement
    - Option 2: TCI state switch command is received within X\*1280 ms upon the last transmission of the RS resources for beam reporting or measurement
      * X can depend on case 1/2 and can be 1
    - Option 3:
      * For BM case 1, TCI state switch command is received within T\_prediction\_valid upon the last transmission of the RS resources for reporting. T\_prediction\_valid is a prediction-validity time related to channel coherence; FFS value/range.
      * For BM case 2, TCI state switch command is received within Tprediction\_time\_period + Tprediction\_valid upon the last transmission of the RS resources for beam reporting .
    - Option 4:
      * option 1 when target TCI state is based on RS from set A that is QCL type-D to a known measured TX beam.
      * Option 2 when target TCI state is based on RS from set A that is not QCL type-D to a known Tx beam and UE reports [TCI state known] capability.
    - Option 5: others
  + QCL relationship:

Qualcomm: option 1 and 3

Ericsson: option 1 and 3

Apple: option 1 and 3

* + - Option 1: Target TCI state is based on RS from set A beams with QCL relationship configured to a set B beam
    - Agreement on QCL relationship
    - If the predicted Tx beam in Set A is QCL Type-D to a known measured Tx beam, where TX beam can be both inside or outside set B, the corresponding Rx beam is known.
    - If the predicted Tx beam in Set A is not QCL Type-D to a known Tx beam, known TCI state conditions shall be based on UE capability.
    - Option 4:
      * For BM-Case 1, when the UE has the capability to know the corresponding Rx beam, RAN4 needs to further discuss:
        + If the UE supports this capability, for MAC-CE-based TCI state switch delay for known TCI state, whether the UE still needs to perform fine time tracking—i.e., whether TOk can be 0 and under what conditions.
      * For BM-Case 2, regarding the MAC-CE-based TCI state switch delay, when the UE has the capability to know the corresponding Rx beam, RAN4 needs to further discuss:
        + If the UE supports this capability, should the L1-RSRP measurement and fine time tracking delay always be defined according to the unknown TCI state case?
        + Potential side condition: The L1-RSRP measurement must wait until the corresponding TCI state takes effect before it can be performed (i.e., the L1-RSRP measurement cannot be executed earlier than the earliest predicted time instance of the target TCI state).
    - Option 5: Known if predicted beam is the RS in target TCI or QCL-ed to the target TCI state
    - Option 6: RS resource for of predicted-L1-RSRP measurement is the RS in target TCI state or QCLed to the target TCI state
* Recommended WF
  + Detectability and SNR conditions:
    - Option 1
  + Time conditions:
    - Option 1
  + QCL relationship:
    - To be discussed

### Sub-topic 2-4

*Activation delay*

Some companies are proposing to discuss/introduce an activation delay when the UE is configured to report predictions

**Issue 2-4: Activation delay**

* Proposals
  + Option 1: No need for any additional delay definition, delay of legacy procedures (RRC reconfiguration, etc) can be reused
  + Option 2: Introduce a separate delay from RRC reconfiguration until UE starts sending prediction reports
    - Requirement definition is FFS, should be agreed in the next meeting
  + Option 3: others
* Recommended WF
  + Option 1

If option 2 is to be agreed, concrete proposals on how to define the requirement and what should be studied/considered should be presented

### Sub-topic 2-5

*Simulation results*

**Issue 2-5: How to proceed with defining metrics based on Simulation results**

Discuss the following Options based on summary of simulation results–prepared by vivo

* Proposals
  + Option 1: Continue aligning on the following aspects
    - Prediction model
      * Companies clarify the model they used (or whether they applied the reference model provided in the simulation assumptions) when submitting results in the next meeting.
    - Dataset
      * Companies should perform simulations based on the reference dataset and submit results in the next meeting.
      * Alternatively, agree on dataset-related parameter settings (e.g., dataset size) in this meeting, and companies should submit results based on the aligned dataset parameters in the next meeting.
  + Option 2: In the next meeting, decide the value of metric based on the simulation results submitted by companies.
    - The source for the metric value can be determined by directly taking the average of the results from companies and applying certain criteria to exclude results with excessively large deviations
  + Option 3: other parameters or assumptions to be clarified/modified
* Recommended WF
  + To be discussed

### Sub-topic 2-6

*Test system basic setup*

The basic test setup has been discussed for several meetings without much progress.

**Issue 2-6: Test system setup**

* Proposals
  + Option 1: Single AoA with UE antenna pattern knowledge
  + Option 2: Single AoA without any UE antenna pattern knowledge
  + Option 3: Multi AoA based on MPAC
  + Option 4: Multi AoA based on RRM testing setup
* Recommended WF
  + Focus the discussion on the multi AoA based on RRM testing setup

Multi AoA based RRM testing setup can also be used to run some tests with single AoA while also having the potential to provide a more realistic environment by using multiple AoAs

### Sub-topic 2-7

*Test system channel model*

In the previous meetings it was agreed that a CDL-based channel model will be taken as starting point. Further details need to be discussed

**Issue 2-7: Test system setup**

* Proposals
  + Option 1: “full” CDL model to be emulated in the test
  + Option 2: simplified CDL should be used
  + Option 3: others
* Recommended WF
  + To be discussed

If Option 2 is chosen, simplification methods should be discussed.

### Sub-topic 2-8

*Draft CR*

A draft CR for the introduction of the core requirements for beam prediction was submitted in R4-2510334

**Issue 2-8: draft CR**

* Proposals
  + Option 1: Discuss the draft CR
* Recommended WF
  + Discuss the structure of the CR and next steps until formal approval

### Sub-topic 2-9

*Prediction KPIs*

Several agreements regarding the prediction KPIs were reached in the previous meetings, however, there are still some open issues regarding whether the accuracy(in absolute or relative terms) should apply to more or all beams predicted and how to make sure that as many of the predicted beams have a certain level of accuracy such that the network can have confidence in the report.

**Issue 2-9: Metrics/KPIs for beam ID prediction**

Proposals

* + Option 1: N=1 (KPI applies only to top predicted beam
  + Option 2: absolute RSRP applies to all predicted beams
    - Predicted beams should meet an SNR side condition like >0 dB
  + Option 3: **consider the following performance metrics if RAN4 decides K>1:**
    - * **For all *n<=N***
      * **Where ’s are the *N* largest ground truth RSRPs among all the beams, ’s are the predicted RSRPs, is the set of the index of the predicted largest *K* beams. We need to have *K*>=N, and preferably the difference is larger than 2 or 3.**
  + Option 4: N=3
  + Option 5: others
  + K is the predicted beams
  + N is the ground truth beams
  + Agreement:
    - If RSRP prediction is reported, absolute/relative RSRP requirements applies to
      * reported beams where SNR side condition is met
      * The exact RSRP reporting including both absolute and relative will be defined in RAN1
      * The related test can only apply to the top-K, where the value of K is FFS, beams.
        + The legacy beam management test and parameters can be taken as the reference to decide the exact value of K
        + The feasibility of the test will be further discussed and decided.

### Sub-topic 2-10

*Relative RSRP accuracy*

Relative RSRP accuracy has been discussed in previous meetings and some agreements were reached, there are proposals to make further changes/refinements to the definition.

**Issue 2-10: Relative RSRP accuracy**

* Proposals
  + Option 1: **RAN4 to remove the square bracket in the definition of relative RSRP accuracy, i.e.,**
    - **Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - reported L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), ~~[~~where the beam index n owns the largest reported value~~]~~.**
  + Option 2:  ***for BM case 1, it is proposed that*** 
    - * + ***The relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), [where the beam index n owns the largest reported value]***
      * ***for BM case 2, it is proposed that*** 
        + ***The relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i for time instance m - predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i for time instance m - ground truth of L1-RSRP of beam index n), where the beam index n owns the largest reported value among all the predicted beams. 1<=m<=M where M is the number of time instance***
  + Option 3: **For BM-Case 1, Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - Reported L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), where the beam index n owns the largest reported value**
  + **Where, the reported L1-RSRP refers to predicted L1-RSRP**
  + Option 4: **Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), [where the beam index n owns the largest reported value].**
  + Option 5: Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i - reported L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), where the beam index n owns the largest reported value given by absolute RSRP reporting.
    - * Reported L1-RSRP can be predicted,
      * Reported L1-RSRP shall be predicted RSRP
    - Relative RSRP accuracy requirements apply for Case 1, and also for Case 2
  + Option 6: **RAN4 defines relative RSRP accuracy requirement for Case 1 and uses the agreed baseline with the following changes:**
    - **Relative RSRP accuracy for reported beams during inference reporting = (predicted L1-RSRP of beam index i – ~~reported~~ predicted L1-RSRP of beam index n) - (ground truth of L1-RSRP of beam index i - ground truth of L1-RSRP of beam index n), [where the beam index n owns the largest reported value].**
      * + **~~Reported L1-RSRP can be predicted,~~**
        + **~~FFS whether reported L1-RSRP can be measured RSRP~~**

**RAN4 ensures the following in AI-ML BM case 1,**

* + - * **The core requirement of relative accuracy of predicted L1-RSRP applies to all top-K predicted beams.**
      * **Performance test of relative accuracy of predicted L1-RSRP uses prediction and reporting of two beams, i.e., K = 2.**
  + Option 7: **For relative predicted RSRP accuracy, define unified definition for both BM-case 1 and BM-case 2:**
    - * **Where is the largest RSRP across all time instance and beams in the report。 is RSRP of beam index i at time t。 is the true RSRP corresponding to the beam/time of 。Where is predicted differential RSRP in the report.**
  + Option 8:others
* Recommended WF
  + Option 1

### Sub-topic 2-11

The session has been looking into dense urban macro to investigate the performance of AI-ML BM-case 1 (i.e., spatial only beam prediction). On the other hand, RAN4 agreed to consider CDL as a baseline to test AI-ML BM performance. This issue checks if and how system level and link levels will be considered to define performance requirement.

Issue 2-11: Simulation Setting to Define Performance Requirement

* Proposals
  + Option 1:

Accuracy of AI-ML BM-case 1 is defined based on the worst-case performance of two test dataset where datasets come from following two scenarios:

* + - Dense urban macro
      * Note: This is currently under consideration and captured in the recently agreed simulation assumption of [R4-2508081](http://10.10.10.10/ftp/RAN/RAN4/Inbox/R4-2508081.zip)
    - CDL-based channel model (details are TBD)
      * Note: The UMi CDL-C channel of Table D.1-1 of 38.151 can be used as a starting point
  + Option 2: other options
* Recommended WF
  + Discuss

# Topic #3: RRM core requirement and testing framework for Positioning accuracy enhancement

This section contains the sub-topics regarding specific issues for positioning

## Companies’ contributions summary

|  |  |  |
| --- | --- | --- |
| **T-doc number** | **Company** | **Proposals / Observations** |
| [**R4-2509299**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509299.zip) | CATT | **Proposal 1: RAN4 to define the following report mapping requirements for additional sample report:**  Table 1: Measurement report mapping for k=0   |  |  |  | | --- | --- | --- | | Reported Quantity Value,  sample\_i | Measured Quantity Value,  Δsample | Unit | | | sample\_000 | Δsample = 1 | Tc | | sample\_001 | Δsample = 2 | Tc | | sample\_002 | Δsample = 3 | Tc | | … | … | … | | sample\_063 | Δsample = 64 | Tc | | sample\_064 | Δsample = 65 | Tc | | … | … | … | | sample\_124 | Δsample = 125 | Tc | | sample\_125 | Δsample = 126 | Tc | | sample\_126 | Δsample = 127 | Tc |   Table 2: Measurement report mapping for k=1   |  |  |  | | --- | --- | --- | | Reported Quantity Value,  sample\_i | Measured Quantity Value,  Δsample | Unit | | | sample\_000 | Δsample = 2 | Tc | | sample\_001 | Δsample = 4 | Tc | | sample\_002 | Δsample = 6 | Tc | | … | … | … | | sample\_063 | Δsample = 128 | Tc | | sample\_064 | Δsample = 130 | Tc | | … | … | … | | sample\_124 | Δsample = 250 | Tc | | sample\_125 | Δsample = 252 | Tc | | sample\_126 | Δsample = 254 | Tc |   Table 3: Measurement report mapping for k=2   |  |  |  | | --- | --- | --- | | Reported Quantity Value,  sample\_i | Measured Quantity Value,  Δsample | Unit | | | sample\_000 | Δsample = 4 | Tc | | sample\_001 | Δsample = 8 | Tc | | sample\_002 | Δsample = 12 | Tc | | … | … | … | | sample\_063 | Δsample = 256 | Tc | | sample\_064 | Δsample = 260 | Tc | | … | … | … | | sample\_124 | Δsample = 500 | Tc | | sample\_125 | Δsample = 504 | Tc | | sample\_126 | Δsample = 508 | Tc |   Table 4: Measurement report mapping for k=3   |  |  |  | | --- | --- | --- | | Reported Quantity Value,  sample\_i | Measured Quantity Value,  Δsample | Unit | | | sample\_000 | Δsample = 8 | Tc | | sample\_001 | Δsample = 16 | Tc | | sample\_002 | Δsample = 24 | Tc | | … | … | … | | sample\_063 | Δsample = 520 | Tc | | sample\_064 | Δsample = 528 | Tc | | … | … | … | | sample\_124 | Δsample = 1000 | Tc | | sample\_125 | Δsample = 1008 | Tc | | sample\_126 | Δsample = 1016 | Tc |   Table 5: Measurement report mapping for k=4   |  |  |  | | --- | --- | --- | | Reported Quantity Value,  sample\_i | Measured Quantity Value,  Δsample | Unit | | | sample\_000 | Δsample = 16 | Tc | | sample\_001 | Δsample = 32 | Tc | | sample\_002 | Δsample = 48 | Tc | | … | … | … | | sample\_063 | Δsample = 1024 | Tc | | sample\_064 | Δsample = 1040 | Tc | | … | … | … | | sample\_124 | Δsample = 2000 | Tc | | sample\_125 | Δsample = 2016 | Tc | | sample\_126 | Δsample = 2032 | Tc |   Table 6: Measurement report mapping for k=5   |  |  |  | | --- | --- | --- | | Reported Quantity Value,  sample\_i | Measured Quantity Value,  Δsample | Unit | | | sample\_000 | Δsample = 32 | Tc | | sample\_001 | Δsample = 64 | Tc | | sample\_002 | Δsample = 96 | Tc | | … | … | … | | sample\_063 | Δsample = 2048 | Tc | | sample\_064 | Δsample = 2080 | Tc | | … | … | … | | sample\_124 | Δsample = 4000 | Tc | | sample\_125 | Δsample = 4032 | Tc | | sample\_126 | Δsample = 4064 | Tc | |
| [**R4-2509426**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509426.zip) | Apple | **Proposal 1:** **RAN4 is invited to define the mapping, granularity, quantization, and reporting range for additional samples, and to decide whether phase information should be included and, if so, its representation. Completing these specifications will prevent vendor-specific behaviors, ensure interoperability, and support robust LMF-side model performance** |
| [**R4-2510164**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510164.zip) | CMCC | ***Proposal 1: for case 1, it is proposed to define delay related requirements, e.g. prediction delay, LCM related delay.***  ***Proposal 2: It is proposed to reuse legacy UL SRS RSRP measurement accuracy requirements for case 3a.***  ***Proposal 3: same as case 3a, it is proposed that legacy report mapping for UL SRS-RSRP, legacy report mapping for gNB Rx-Tx time difference, and legacy UL SRS RSRP measurement accuracy requirements are reused for case 3b.*** |
| [**R4-2510338**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510338.zip) | vivo | **Proposal 1: RAN4 to define delay requirement for case 1 at least considering the following procedures: PRS measuring, AI processing. Legacy requirement (if exists) can be reused.**  **Proposal 2: RAN4 not to define new requirements for positioning case 3. Legacy requirement (if exists) can be reused.**  **Proposal 3: RAN4 to consider define delay requirements for monitoring Option A-1, A-2 and A-3 for positioning case 1.** |
| [**R4-2510758**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510758.zip) | Ericsson | **Observation 1**: To ensure that the inferred location information is reported by the UE to the network in a timely manner, reporting delay requirement for positioning use case #1 needs to be specified.  **Proposal 1**: Measurement reporting delay for AI/ML positioning use case #1 should be defined as the time between the moment the position inference reporting request from the network is received by the UE and the moment when the UE starts to transmit the inferred UE location report over the air interface to be received by the network.  **Observation 2**: Measurement delay for model input generation depends on the number of measured PFLs, scaling factor associated to the measured PFL, Rx beam sweeping factor, number of samples for the PRS measurement, and the PRS processing capability of the UE.  **Proposal 2**: Measurement period requirement for legacy positioning measurements is used as a baseline to define the measurement period for the AI/ML model input generation for the UE capable of supporting AI/ML positioning use case #1.  **Proposal 3**: Impact of the number of Rx TEGs on the measurement period requirement for measurements performed by UE in use case #1 for positioning inference is not considered. |
| [**R4-2510807**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510807.zip) | OPPO | **Observation 1: RAN4 does not define positioning accuracy requirements for AI/ML-based positioning accuracy enhancement.**  **Proposal 1: For case1, the impact of measurement(PRS measurement) and AI/ML processing (may involve model activation delay and inference delay) should be taken into account when defining the delay requirement.**  **Proposal 2: For case3, RAN4 does not need to impose additional measurement delay requirements (reporting delay) specifically for AI/ML based Positioning enhancement.** |
| [**R4-2510874**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510874.zip) | Huawei,HiSilicon | **Proposal 1**: For Case 3b, reuse the legacy requirement for reporting of timing information or timing and power information from gNB to LMF.  **Proposal 2**: For Case 3a, RAN4 to discuss whether to reuse the legacy requirement for reporting of timing information from gNB to LMF, if there is no explicit indicator introduced to distinguish whether the timing information is obtained by legacy method or by Rel-19 AI/ML.  **Proposal 3**: RAN4 to discuss the requirement for Rel-19 AI/ML based timing information reporting if introduced by other WGs.  **Proposal 4*:*** RAN4 to discuss requirements for LCM procedure especially performance monitoring for AI/ML positioning, based on RAN1 conclusion on performance monitoring schemes and also RAN4 conclusion on the requirements for inference. |
| [**R4-2511019**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511019.zip) | ZTECorporation,Sanechips | **Observation 1: In legacy, there are two kinds of location request: UE-triggered location request and NW-triggered location request. The latest agreement contains the implicit location request which is the NW triggers the location request, otherwise no reporting scheme is needed.**  **Observation 2: The legacy measurement delay is defined as the time between the moment when the periodic measurement report is triggered and the moment when the UE starts to transmit the measurement report over the air interface.**  **Proposal 1: For AI positioning case 1, the legacy positioning reporting delay could be as baseline.** |
| [**R4-2511222**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2511222.zip) | Nokia | **Case 1**   1. A comprehensive reporting delay can be defined by encompassing measurement, AI/ML inference, and reporting time intervals. 2. It would be needed to categorize AI/ML inference delays based on model complexity and UE processing capabilities, to support flexible but bounded delay requirements. 3. RAN4 could be required to investigate the feasibility of defining a model verification framework to ensure that AI/ML-based positioning outputs meet minimum reliability and performance standards. 4. RAN4 should consider whether a reporting delay requirement is appropriate for UE-based positioning in Case 1, given the variability introduced by AI/ML inference on the UE and the current lack of standardized delay, accuracy, and verification frameworks. 5. RAN4 should not define explicit performance monitoring requirements for UE-based positioning in Case 1. 6. Postpone the discussion on the testing framework for Case 1 until testable core requirements are specified.   Case 3a/3b   1. Do not specify a testing framework for Case 3a. 2. Do not specify a testing framework for path-based or sample-based measurements for Case 3b. |
| [**R4-2509932**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2509932.zip) | Nokia | 1. Reuse report mapping tables in TS 38.133 clause 13.1.1 for UL RTOA first path reporting and report mapping tables in TS 38.133 clause 13.1.1A for UL RTOA additional path reporting in Case 3a. 2. Reuse report mapping table in TS 38.133 clause 13.3.1 for UL SRS RSRP reporting and report mapping table in TS 38.133 clause 13.6.1 for UL SRS RSRPP reporting in Case 3a. 3. No RAN4 requirements for inference need to be defined for the LOS/NLOS indicator agreed by RAN1 for Case 3a. 4. Measurement accuracy requirements for gNB Rx-Tx time difference apply for UL SRS-TDCT first path under the condition that the reported measurement is based on N=4 samples. 5. Measurement accuracy requirements for UL SRS RSRP apply for UL SRS-TDCP under the condition that the reported measurement is based on N=4 samples. 6. Add new report mapping tables for additional samples for UL RTOA by reusing single-sided range of existing additional path reporting for UL RTOA and keeping the reporting range agnostic to the number of configured by LMF. 7. Add following tables to the new clause 13.1.1X of TS 38.133 for additional sample reporting for sample-based reporting:   **Table 13.1.1X-1: Measurement report mapping for k=0**   |  |  |  | | --- | --- | --- | | **Reported Quantity Value,**  **sample\_i** | **Measured Quantity Value,**  **sample** | **Unit** | | | sample\_0000 | 0  sample < 1 | Tc | | sample\_0001 | 1  sample < 2 | Tc | | sample\_0002 | 2  sample < 3 | Tc | | … | … | … | | sample\_8173 | 8173  sample < 8174 | Tc | | sample\_8174 | 8174  sample < 8175 | Tc | | sample\_8175 | 8175  sample | Tc |   **Table 13.1.1X-2: Measurement report mapping for k=1**   |  |  |  | | --- | --- | --- | | **Reported Quantity Value,**  **sample\_i** | **Measured Quantity Value,**  **sample** | **Unit** | | | sample\_0000 | 0  sample < 2 | Tc | | sample\_0001 | 2  sample < 4 | Tc | | sample\_0002 | 4  sample < 6 | Tc | | … | … | … | | sample\_4086 | 8172  sample < 8174 | Tc | | sample\_4087 | 8174  sample | Tc |   **Table 13.1.1X-3: Measurement report mapping for k=2**   |  |  |  | | --- | --- | --- | | **Reported Quantity Value,**  **sample\_i** | **Measured Quantity Value,**  **sample** | **Unit** | | | sample\_0000 | 0  sample < 4 | Tc | | sample\_0001 | 4  sample < 8 | Tc | | sample\_0002 | 8  sample < 12 | Tc | | … | … | … | | sample\_2042 | 8172  sample < 8174 | Tc | | sample\_2043 | 8174  sample | Tc |   **Table 13.1.1X-4: Measurement report mapping for k=3**   |  |  |  | | --- | --- | --- | | **Reported Quantity Value,**  **sample\_i** | **Measured Quantity Value,**  **sample** | **Unit** | | | sample\_0000 | 0  sample < 8 | Tc | | sample\_0001 | 8  sample < 16 | Tc | | sample\_0002 | 16  sample < 24 | Tc | | … | … | … | | sample\_1019 | 8152  sample < 8160 | Tc | | sample\_1020 | 8160  sample < 8168 | Tc | | sample\_1021 | 8168  sample | Tc |   **Table 13.1.1X-5: Measurement report mapping for k=4**   |  |  |  | | --- | --- | --- | | **Reported Quantity Value,**  **sample\_i** | **Measured Quantity Value,**  **sample** | **Unit** | | | sample\_000 | 0  sample < 16 | Tc | | sample\_001 | 16  sample < 32 | Tc | | sample\_002 | 32  sample < 48 | Tc | | … | … | … | | sample\_508 | 8128  sample < 8144 | Tc | | sample\_509 | 8144  sample < 8160 | Tc | | sample\_510 | 8160  sample | Tc |   **Table 13.1.1X-6: Measurement report mapping for k=5**   |  |  |  | | --- | --- | --- | | **Reported Quantity Value,**  **sample\_i** | **Measured Quantity Value,**  **sample** | **Unit** | | | sample\_000 | 0  sample < 32 | Tc | | sample\_001 | 32  sample < 64 | Tc | | sample\_002 | 64  sample < 96 | Tc | | … | … | … | | sample\_253 | 8098  sample < 8128 | Tc | | sample\_254 | 8128  sample < 8160 | Tc | | sample\_255 | 8160  sample | Tc |  1. Reuse the UL SRS RSRPP report mapping table in TS 38.133 clause 13.3.1 for the reported samples for UL SRS TDCP for sample-based reporting. 2. RAN4 to not specify a requirement on training data collection of Part B in Case 3a/3b. |
| [**R4-2510759**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510759.zip) | Ericsson | **Observation 1**: Depending on the request from LMF, gNB can report path-based measurements to the LMF. For such measurement reporting, no enhancement to Rel. 18 reporting has been introduced.   **Observation 2**: Depending on the request from LMF, gNB can report sample-based measurements to the LMF. For such measurement reporting, report mapping table needs to be specified in RAN4 specification.  **Observation 3**: For sample-based reporting, gNB may either report timing measurement or timing and power measurements associated to Nt´ samples of the estimated channel response in time domain.  **Observation 4**: The minimum number of samples for which the timing or timing and power information are reported is 8 and the maximum number of samples for which the timing or timing and power information are reported is 24.  **Proposal 1**: Reporting range for UL SRS-TDCP is defined from -156 dBm to -31 dBm with resolution 1 dB.  **Proposal 2**: Table 1 is used to define measurement report mapping for UL SRS-TDCP measurement.  Table 1. UL SRS-TDCP report mapping   |  |  |  | | --- | --- | --- | | Reported value | Measured quantity value | Unit | | SRS\_TDCP\_000 | SRS-TDCP < -156 | dBm | | SRS\_TDCP\_001 | -156 ≤ SRS-TDCP < -155 | dBm | | SRS\_TDCP\_002 | -155 ≤ SRS-TDCP < -154 | dBm | | SRS\_TDCP\_003 | -154 ≤ SRS-TDCP < -153 | dBm | | SRS\_TDCP\_004 | -153 ≤ SRS-TDCP < -152 | dBm | | SRS\_TDCP\_005 | -152 ≤ SRS-TDCP < -151 | dBm | | SRS\_TDCP\_006 | -151 ≤ SRS-TDCP < -150 | dBm | | SRS\_TDCP\_007 | -150 ≤ SRS-TDCP < -149 | dBm | | SRS\_TDCP\_008 | -149 ≤ SRS-TDCP < -148 | dBm | | SRS\_TDCP\_009 | -148 ≤ SRS-TDCP < -147 | dBm | | SRS\_TDCP\_010 | -147 ≤ SRS-TDCP < -146 | dBm | | SRS\_TDCP\_011 | -146 ≤ SRS-TDCP < -145 | dBm | | SRS\_TDCP\_012 | -145 ≤ SRS-TDCP < -144 | dBm | | SRS\_TDCP\_013 | -144 ≤ SRS-TDCP < -143 | dBm | | SRS\_TDCP\_014 | -143 ≤ SRS-TDCP < -142 | dBm | | SRS\_TDCP\_015 | -142 ≤ SRS-TDCP < -141 | dBm | | SRS\_TDCP\_016 | -141 ≤ SRS-TDCP < -140 | dBm | | SRS\_TDCP\_017 | -140 ≤ SRS-TDCP < -139 | dBm | | SRS\_TDCP\_018 | -139 ≤ SRS-TDCP < -138 | dBm | | … | … | … | | SRS\_TDCP\_111 | -46 ≤ SRS-TDCP < -45 | dBm | | SRS\_TDCP\_112 | -45 ≤ SRS-TDCP < -44 | dBm | | SRS\_TDCP\_113 | -44 ≤ SRS-TDCP < -43 | dBm | | SRS\_TDCP\_114 | -43 ≤ SRS-TDCP < -42 | dBm | | SRS\_TDCP\_115 | -42 ≤ SRS-TDCP < -41 | dBm | | SRS\_TDCP\_116 | -41 ≤ SRS-TDCP < -40 | dBm | | SRS\_TDCP\_117 | -40 ≤ SRS-TDCP < -39 | dBm | | SRS\_TDCP\_118 | -39 ≤ SRS-TDCP < -38 | dBm | | SRS\_TDCP\_119 | -38 ≤ SRS-TDCP < -37 | dBm | | SRS\_TDCP\_120 | -37 ≤ SRS-TDCP < -36 | dBm | | SRS\_TDCP\_121 | -36 ≤ SRS-TDCP < -35 | dBm | | SRS\_TDCP\_122 | -35 ≤ SRS-TDCP < -34 | dBm | | SRS\_TDCP\_123 | -34 ≤ SRS-TDCP < -33 | dBm | | SRS\_TDCP\_124 | -33 ≤ SRS-TDCP < -32 | dBm | | SRS\_TDCP\_125 | -32 ≤ SRS-TDCP < -31 | dBm | | SRS\_TDCP\_126 | -31 ≤ SRS-TDCP | dBm | |
| [**R4-2510760**](https://www.3gpp.org/ftp/tsg_ran/WG4_Radio/TSGR4_116/Docs/R4-2510760.zip) | Ericsson | draftCR to 38.133 on report mapping for UL SRS-TDCT and UL SRS-TDCP measurements (Rel.19)   |  |  | | --- | --- | | ***Reason for change:*** | * To update list of acronyms in Chapter 3.3 with the expanded forms of UL SRS-TDCT and UL SRS-TDCP measurements. * To implement new clauses for report mapping for UL SRS-TDCT and UL SRS-TDCP measurements. | |  |  | | ***Summary of change:*** | * List of acronyms in Chapter 3.3 is updated with the expanded forms of UL SRS-TDCT and UL SRS-TDCP measurements. * New clauses for report mapping for UL SRS-TDCT and UL SRS-TDCP measurements are introduced. | |  |  | | ***Consequences if not approved:*** | * List of acronyms in chapter 3.3 is incomplete. * Report mappings for UL SRS-TDCT and UL SRS-TDCP measurements remain undefined. | |

## Open issues summary

The agreements from RAN4#114-bis are listed below for convenience:

**Issue 3-2: LOS/NLOS indicator**

Agreement:

the LOS/NLOS indicator is out of scope of WID. Stop discussions on it.

**Issue 3-3: Rx-Tx time difference for case 3a**

Agreement:

* Introduce gNB Rx-Tx time difference reporting accuracy requirements
  + - Reuse legacy accuracy requirements

**Issue 3-4: Report mapping for gNB Rx-Tx time difference**

Agreement:

* Reuse the existing report mapping for gNB Rx-Tx time difference measurements

The agreements from RAN4#115 are listed below for convenience(R5-2508080):

**Issue 3-1: Requirements for case 1**

Agreement:

* Discuss how to define a measurement delay requirement(reporting delay)
* Do not define any positioning accuracy requirement

**Issue 3-2: Report mapping for sample based timing measurement reporting**

Agreement:

* + Use the value of the legacy UL RTOA mapping for 1st sample
  + FFS for the additional sample reporting

**Issue 3-3: Requirements for case 3b**

Agreement:

* Reuse legacy path based Rx-Tx accuracy requirements for case 3b
* Requirement applies only for the 1st path

The open issues were grouped in the following sub-topics for further discussion:

Case 1 requirements

Sample based report mapping

Report mapping for UL SRS-TDCP measurement

Draft CR for UL SRS-TDCT and UL SRS-TDCP measurements

### Sub-topic 3-1

*Requirements for case 1*

Several companies proposed to introduce a reporting delay requirements for case 1

**Issue 3-1: Reporting Delay Requirements for case 1**

Proposals

* + Option 1: Introduce a reporting delay requirement defined as the delay from the NW triggered location request until the UE sends the location report
    - Take the existing requirement as baseline
    - Further discuss if this delay should be extended due additional time needed for processing
  + Option 2: others
* Recommended WF
  + Option 1

CMCC: support option 1

Ericsson: support option 1

Nokia: since we don’t have accuracy requirement, what’s the point to define the delay requirement.

ZTE: what’s the existing requirement in option 1?

### Sub-topic 3-2

*Sample Based Report Mapping*

There are two proposals for the sample based mapping which are very similar. It should be decided which one is used

**Issue 3-2: Sample Based Report Mapping**

Proposals

* + Option 1: Adopt the report mapping in R4-2509299 (CATT)
  + Option 2: Adopt the report mapping in R4-2509932 (Nokia)
  + Option 3: Others
* Recommended WF
  + Option 2

Option 2 seems more general and uses inequalities

### Sub-topic 3-3

*Report mapping for UL SRS-TDCP measurement*

**Issue 3-3: Report mapping for UL SRS-TDCP measurement**

* Proposals
  + Option 1: Agree the mapping for UL SRS-TDCP as proposed in R4-2510759
  + Option 2: others
* Recommended WF
  + Option 1

### Sub-topic 3-4

*Draft CR for UL SRS-TDCT and UL SRS-TDCP*

**Issue 3-4: Report mapping for UL SRS-TDCT and UL SRS-TDCP measurements**

Proposals

* + Option 1: Endorse CR in R4-2510760, to be later merge with the rest of positioning requirements
  + Option 2: Further modify the draft CR
* Recommended WF
  + To be discussed

The draft could be technically endorsed, would be later merged with other drafts for other positioning requirements