**3GPP TSG RAN WG1 #118 R1-240xxxx**

**Maastricht, Netherland, August 19th – 23th, 2024**

**Agenda Item:** 9.1.3.1

**Source:** Moderator (LG Electronics)

**Title:** Discussion on the TPs for pCR

**Document for:** Discussion and decision

# Introduction

**Purpose**: Just to collect offline comment and to prepare for the CR

**Principle**: Copy and paste RAN1 agreements/conclusions/observation

* Minor wording change highlighted by Yellow to adapt to the context

# Draft TP

## TP 1

|  |
| --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------6.2.2 Performance resultsCSI\_Table 1 through CSI\_Table 8 in attached Spreadsheets for CSI feedback enhancement evaluations present the performance results for:- CSI\_Table 1. Evaluation results for CSI compression of 1-on-1 joint training without model generalization/scalability- CSI\_Table 2. Evaluation results for CSI compression with model generalization- CSI\_Table 3. Evaluation results for CSI compression with model scalability- CSI\_Table 4. Evaluation results for CSI compression of multi-vendor joint training without model generalization/scalability- CSI\_Table 5. Evaluation results for CSI compression of separate training without model generalization/scalability- CSI\_Table 6. Evaluation results for CSI prediction without model generalization/scalability- CSI\_Table 7. Evaluation results for CSI prediction with model generalization- CSI\_Table 6-A. Evaluation results for CSI prediction without model generalization/scalability- CSI\_Table 7-A. Evaluation results for CSI prediction with model generalization- CSI\_Table 8. Evaluation results for CSI prediction with localized model--------------------------------------------------------Text omitted --------------------------------------------------------- |

## TP 2

|  |
| --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------***KPIs and Evaluation metrics*:** - Capability/complexity: Floating point operations (FLOPs), AI/ML memory storage in terms of AI/ML model size and number of AI/ML parameters reported by companies who may select either or both- Reported separately for the CSI generation part and the CSI reconstruction part (for CSI compression sub-use case) - When reporting the computational complexity including the pre-processing and post-processing, the complexity metric of FLOPs may be reported separately for the AI/ML model and the pre/post processing. While reporting the FLOPs of pre-processing and post-processing the following boundaries are considered:- Estimated raw channel matrix per each frequency unit as an input for pre-processing of the CSI generation part.- Precoding vectors per each frequency unit as an output of post-processing of the CSI reconstruction part.- Additionally, reported for non-AI/ML based CSI prediction including additional complexity if applicable, e.g., update of filter, assuming whole bandwidth and one prediction sample (for CSI prediction sub-use case)- CSI compression: Intermediate KPIs: SGCS and/or NMSE to evaluate the accuracy of the AI/ML output CSI- For rank>1 cases, SGCS calculation/extension methods are to be reported:- SGCS separately calculated for each layer (e.g., for K layers, K SGCS values are derived respectively, and comparison is performed per layer). Companies to ensure the correct calculation of SGCS and to avoid disorder issue of the output eigenvectors. Note: Eventual KPI can still be used to compare the performance. - The granularity of the frequency unit for averaging operation is assumed to be: - For 15kHz SCS: For 10MHz bandwidth: 4 RBs; for 20MHz bandwidth: 8 RBs- For 30kHz SCS: For 10MHz bandwidth: 2 RBs; for 20MHz bandwidth: 4 RBs- Other frequency unit granularities not precluded.- CSI compression: Intermediate KPI: model monitoring mechanism is considered as: - Step 1: Generate test dataset including K test samples.- Step 2: For each of the K test samples, a bias factor of monitored intermediate KPI (KPI*Diff*) is calculated as a function of KPI*Diff* = *f* ( KPI*Actual* , KPI*Genie* ), where KPI*Actual* is the actual intermediate KPI, and KPI*Genie* is the genie-aided intermediate KPI. - KPI*Diff* is considered for:- Case 1: NW side monitoring of intermediate KPI, where the monitoring accuracy is evaluated for a given ground-truth CSI format (e.g., quantized ground-truth CSI with 8 bits scalar, R16 eType II-like method, etc.) or SRS measurements, where- KPI*Actual* is calculated with the output CSI at the NW side and the given ground-truth CSI format or SRS measurements.- KPI*Genie* is calculated with output CSI (as for KPI*Actual*) and the ground-truth CSI of Float32- Note: if Float32 is used for KPI*Actual*, the monitoring accuracy is 100% if KPI*Actual* and KPI*Genie* are based on the same CSI sample. - Case 2: UE side monitoring of intermediate KPI with a proxy model, where the monitoring accuracy is evaluated for the output of the proxy model at UE:- Case 2-1: the proxy model is a proxy CSI reconstruction part, and KPI*Actual* is calculated based on the inference output of the proxy CSI reconstruction part at UE and the ground-truth CSI. Note: if the proxy CSI reconstruction model is the same as the actual CSI reconstruction model at the NW, the monitoring accuracy is 100%. - Case 2-2: the proxy model directly outputs intermediate KPI (KPI*Actual*)- KPI*Genie* is calculated with the output CSI at the NW side and the same ground-truth CSI. - KPI*Diff* = *f* ( KPI*Actual* , KPI*Genie* ) can take the following forms: - Option 1 (baseline for calibration): Gap between KPI*Actual* and KPI*Genie*, i.e. KPI*Diff* = (KPI*Actual* - KPI*Genie*); Monitoring accuracy is the percentage of samples for which | KPI*Diff*| < KPI*th 1*, where KPI*th 1* is a threshold of the intermediate KPI gap which can take the following values: 0.02, 0.05 and 0.1.- Option 2 (optional and up to companies to report): Binary state where KPI*Actual* and KPI*Genie*, have different relationships to their threshold(s), i.e., KPI*Diff* = (KPI*Actual* > KPI*th 2*, KPI*Genie* < KPI*th 3*) OR (KPI*Actual* < KPI*th 2*, KPI*Genie* > KPI*th 3*), where KPI*th 2* is considered to be the same as KPI*th 3*. Monitoring accuracy is the percentage of samples for which KPI*Diff* = 0. - Step 3: Calculate the statistical result of the KPI*Diff* over K test samples which represents the monitoring accuracy performance.- Note: is introduced for the evaluation and comparison purpose; it may not be available in the real network.- Note: the complexity, overhead and latency of the monitoring scheme are to be reported. - CSI prediction: Intermediate KPIs: calculated for each predicted instance if AI/ML model outputs multiple predicted instances - If collaboration level x is reported as the benchmark, the EVM to distinguish level x and level y/z based AI/ML CSI prediction is considered from the generalization aspect, e.g., collaboration level y/z based CSI prediction is modelled as the fine-tuning case or generalization Case 1, while collaboration level x based CSI prediction is modelled as generalization Case 2 or Case 3.- Throughput including: average UPT, 5%-ile UE throughput, and CDF of UPT--------------------------------------------------------Text omitted --------------------------------------------------------- |

## TP 3

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------Table 6.2.1-1: Baseline System Level Simulation assumptions for AI/ML based CSI feedback enhancement evaluations

|  |  |
| --- | --- |
| Parameter | Value |
| Duplex, Waveform | FDD (TDD is not precluded), OFDM |
| Multiple access | OFDMA |
| Scenario | Dense Urban (Macro only) is a baseline.Other scenarios (e.g., UMi@4GHz 2GHz, Urban Macro) are not precluded. |
| Frequency Range | FR1 only, 2GHz as baseline, optional for 4GHz (if R16 as baseline)FR1 only, 2GHz with duplexing gap of 200MHz between DL and UL, optional for 4GHz (if R17 as baseline) |
| Inter-BS distance | 200m |
| Channel model         | According to TR 38.901 |
| Antenna setup and port layouts at gNB | Companies need to report which option(s) are used between- 32 ports: (8,8,2,1,1,2,8), (dH,dV) = (0.5, 0.8)λ- 16 ports: (8,4,2,1,1,2,4), (dH,dV) = (0.5, 0.8)λOther configurations are not precluded. |
| Antenna setup and port layouts at UE | 4RX: (1,2,2,1,1,1,2), (dH,dV) = (0.5, 0.5)λ for (rank 1-4)2RX: (1,1,2,1,1,1,1), (dH,dV) = (0.5, 0.5)λ for (rank 1,2)Other configuration is not precluded. |
| BS Tx power | 41 dBm for 10MHz, 44dBm for 20MHz, 47dBm for 40MHz |
| BS antenna height | 25m |
| UE antenna height & gain | Follow TR36.873 |
| UE receiver noise figure | 9dB |
| Modulation | Up to 256QAM |
| Coding on PDSCH | LDPCMax code-block size=8448bit |
| Numerology | Slot/non-slot | 14 OFDM symbol slot |
| SCS | 15kHz for 2GHz, 30kHz for 4GHz |
| Simulation bandwidth | 10 MHz for 15kHz as a baseline, and configurations which emulate larger BW, e.g., same sub-band size as 40/100 MHz with 30kHz, may be optionally considered. Above 15kHz is replaced with 30kHz SCS for 4GHz (if R16 as baseline)20 MHz for 15kHz as a baseline (optional for 10 MHz with 15KHz), and configurations which emulate larger BW, e.g., same sub-band size as 40/100 MHz with 30kHz, may be optionally considered. Above 15kHz is replaced with 30kHz SCS for 4GHz (if R17 as baseline) |
| Frame structure | Slot Format 0 (all downlink) for all slots |
| MIMO scheme | SU/MU-MIMO with rank adaptation. Companies are encouraged to report the SU/MU-MIMO with RU.  |
| MIMO layers | For all evaluation, companies to provide the assumption on the maximum MU layers (e.g., 8 or 12) |
| CSI feedback | Feedback assumption at least for baseline scheme- CSI feedback periodicity (full CSI feedback): 5 ms (baseline)- Scheduling delay (from CSI feedback to time to apply in scheduling): 4 ms |
| Overhead | Companies shall provide the downlink overhead assumption (i.e., whether the CSI-RS transmission is UE-specific or not and take that into account for overhead computation) |
| Traffic model | At least, FTP model 1 with packet size 0.5 Mbytes is assumed.Other options are not precluded |
| Traffic load (Resource utilization) | 20/50/70%. Companies are encouraged to report the MU-MIMO utilization.  |
| UE distribution | CSI compression: 80% indoor (3 km/h), 20% outdoor (30 km/h)CSI prediction: 100% outdoor (10, 20, 30, 60, 120 km/h) including outdoor-to-indoor car penetration loss per TR 38.901 if the simulation assumes UEs inside vehicles. No explicit trajectory modeling considered for evaluations.  |
| UE receiver | MMSE-IRC as the baseline receiver |
| Feedback assumption | Realistic |
| Channel estimation          | Realistic as a baseline. Up to companies to choose the error modelling method for realistic channel estimation. Error modelling in TR36.897 Table A.1-2 as a baseline if channel estimation error is modelled.Ideal DL channel estimation is optionally taken into the baseline of evaluation methodology for the purpose of calibration and/or comparing intermediate results (e.g., accuracy of AI/ML output CSI, etc.). Up to companies to report whether/how ideal channel is used in the dataset construction and performance evaluation/inference.Note: Eventual performance comparison with the benchmark release and drawing SI conclusions should be based on realistic DL channel estimation. |
| Phase discontinuity | A uniform distribution between within a time window of , where =40 degrees and =20ms can be a baseline. * A fixed phase for all CSI-RS observations within the time window, and another fixed phase for the next time window. The phases are according to uniform distribution.
 |
| CSI-RS configuration | Periodic: 5 ms periodicity (baseline), 20 ms periodicity (encouraged) Aperiodic: Optional, CSI-RS burst with K resources and time interval m slots (based on R18 MIMO eType-II) |
| Evaluation Metric | Throughput and CSI feedback overhead as baseline metrics.The CSI feedback overhead is calculated as the weighted average of CSI payload per rank and the distribution of ranks reported by the UE. - For AI/ML based solutions: The above-mentioned "CSI feedback overhead" is calculated as max allowed bits at the given rank. - For legacy Type II CB: Option 2b is mandatorily reported by companies, while Option 2a can be optionally reported up to companies if partial NZC report is assumed for the legacy Type II CB- Option 2a: The above-mentioned "CSI feedback overhead" is calculated as each CSI reported payload with a given rank- Option 2b: The above-mentioned "CSI feedback overhead" is calculated as max allowed bits at the given rankAdditional metrics, e.g., ratio between throughput and CSI feedback overhead, can be used.Maximum overhead (payload size for CSI feedback)for each rank at one feedback instance is the baseline metric for CSI feedback overhead, and companies can provide other metrics. |
| Baseline for performance evaluation | For CSI compression:Companies need to report which option is used between:- Rel-16 TypeII Codebook as the baseline for performance and overhead evaluation.- Rel-17 TypeII Codebook as the baseline for performance and overhead evaluation.Additional assumptions from R17 TypeII EVM: Same consideration with respect to utilizing angle-delay reciprocity should be considered for the AI/ML based CSI feedback and the baseline scheme if R17 TypeII codebook is selected as baseline.Optionally, Type I Codebook (if it outperforms Type II Codebook) can be considered for comparing AI/ML schemes.For CSI-prediction: Both of the following are taken as baseline: - The nearest historical CSI without prediction- Non-AI/ML or AI/ML with collaboration Level x based CSI prediction for which corresponding details would need to be reportedNote: the specific non-AI/ML based CSI prediction is compatible with R18 MIMO; collaboration level x AI/ML based CSI prediction could be implementation based AI/ML compatible with R18 MIMO as an example.Note: R18 eType II doppler codebook is assumed for CSI report for both AI/ML and Non AI/ML prediction. Companies can report the assumption for N4: 1,4 (baseline), 2,8 (optional), and the assumption for paramCombination-Doppler-r18: 6,7 or paramCombination -r16 = 5,6 (for Benchmark 1).sFor the evaluation of CSI enhancements, companies can optionally provide the additional throughput baseline based on CSI without compression (e.g., eigenvector from measured channel), which is taken as an upper bound for performance comparison. |
| Note: the baseline EVM is used to compare the performance with the benchmark release, while the AI/ML related parameters (e.g., dataset construction, generalization verification, and AI/ML related metrics) can be of additional/different assumptions. The conclusions for the use cases in the SI should be drawn based on generalization verification over potentially multiple scenarios/configurations. |

--------------------------------------------------------Text omitted --------------------------------------------------------- |

## TP 4

|  |
| --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------6.2 CSI feedback enhancement6.2.1 Evaluation assumptions, methodology and KPIs--------------------------------------------------------Text omitted ---------------------------------------------------------***Model Fine-tuning*:** For the evaluation of the potential performance benefits of model fine-tuning of CSI feedback enhancement, which is optionally assessed, the following case is considered:- The AI/ML model is trained based on training dataset from a different dataset than Scenario#B/Configuration#B, e.g., Scenario#A/Configuration#A, Scenario#A/Configuration#B, Scenario#B/Configuration#A, and then the AI/ML model is updated based on a fine-tuning dataset Scenario#B/Configuration#B. After that, the AI/ML model is tested on Scenario#B/Configuration#B. - In this case, the fine-tuning dataset setting (e.g., size of dataset) is to be reported along with the improvement of performance.***Localized model:***For the evaluation of AI/ML-based CSI feedback enhancement using localized models, consider the following options to model the spatial correlation in the dataset for a local region:* Option 1: The dataset is derived from UEs dropped within the local region, with spatial consistency modelling as per TR 38.901.
	+ - E.g., Dropped in a specific cell or within a specific boundary.
* Option 2: By using a scenario/configuration specific to the local region.
	+ - E.g., Indoor-outdoor ratio, LOS-NLOS ratio, TXRU mapping, etc.

Note: While modelling the spatial correlation, strive to ensure that the dataset distribution also correctly captures the decorrelation due to temporal variations in the channel. To report methods to generate training and testing dataset.For the evaluation of AI/ML-based CSI feedback enhancement using localized models, regarding training,* The k-th local model is trained on region #B\_k (the k-th local region), 1<=k<=N.
* The generalized model is trained on Region #A that may be constructed via any of the following methods that is appropriate for the given generalized/local region modeling approach.
	+ Region #A is the same as the union of regions #B\_1, …, #B\_N.
	+ Region #A is a proper superset of the union of regions #B\_1, …, #B\_N.
	+ Region #A is generated separately from regions #B\_1, …, #B\_N.
	+ Note: companies to report which method was used.

For the evaluation of AI/ML-based CSI feedback enhancement using localized models, regarding testing,* The trained generalized model, local model, and the non-AI/ML benchmark are tested on the regions #B\_1, …, #B\_N.
* In case N>1, when reporting the results, companies may report the performance of the generalized model, the local models, and the non-AI/ML benchmark, by averaging the performance over the regions #B\_1,…,B\_N. Companies to report the value of N.

--------------------------------------------------------Text omitted --------------------------------------------------------- |

## TP 5

|  |
| --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------7.1.2 CSI feedback enhancement ***Items considered for studying the necessity, feasibility, potential specification impact:*** --------------------------------------------------------Text omitted ---------------------------------------------------------**In CSI prediction using UE-sided model use case:***Data collection:*In CSI prediction using UE sided model use case, at least the following aspects have been proposed by companies on data collection, including:- Signalling and procedures for the data collection - Data collection indicated by NW - Requested from UE for data collection - CSI-RS configuration - Assistance information for categorizing the data, if needed- The provision of assistance information needs to consider feasibility of disclosing proprietary information to the other side.*Performance monitoring:* For CSI prediction using UE side model use case, at least the following aspects have been proposed by companies on performance monitoring for functionality-based LCM:- Type 1:- UE calculates the performance metric(s)- UE reports performance monitoring output that facilitates functionality fallback decision at the network- Performance monitoring output details can be further defined - NW may configure threshold criterion to facilitate UE side performance monitoring (if needed). - NW makes decision(s) of functionality fallback operation (fallback mechanism to legacy CSI reporting).- Type 2: - UE reports *predicted CSI* and/or the corresponding *ground-truth* - NW calculates the *performance metrics*. - NW makes decision(s) of functionality fallback operation (fallback mechanism to legacy CSI reporting).- Type 3: - UE calculates the *performance metric(s)* - UE reports *performance metric(s)* to the NW- NW makes decision(s) of functionality fallback operation (fallback mechanism to legacy CSI reporting). - Functionality selection/activation/deactivation/switching as defined for other UE side use cases can be reused, if applicable. - Configuration and procedure for performance monitoring - CSI-RS configuration for performance monitoring- Performance metric including at least intermediate KPI (e.g., NMSE or SGCS)- UE report, including periodic/semi-persistent/aperiodic reporting, and event driven report- Note: down selection is not precluded.- Note: UE may make decision within the same functionality on model selection, activation, deactivation, switching operation transparent to the NW.For the boundary between Type 3 and Type 1 performance monitoring, the difference is whether UE reports performance metric or performance monitoring output to NW, respectively. * The monitoring output is determined based on performance metric, and additionally, baseline and/or threshold criterion if configured.

For CSI prediction using UE-sided model, for performance monitoring, at least following specification impacts are additionally identified compared to that has been captured in TR38.843, * Type 1
	+ Definition/configuration of performance metric
	+ Definition of threshold criterion, if configured
	+ Definition/configuration and report of monitoring output, and corresponding report mechanism
* Type 2
	+ Definition/configuration and report of ground truth CSI, and corresponding report mechanism.
* Type 3
	+ Definition/configuration and report of performance metric, and corresponding report mechanism.
* For all types of performance monitoring, NW indication to the UE of the decision regarding the monitoring action

--------------------------------------------------------Text omitted --------------------------------------------------------- |

## TP 6

|  |
| --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------6.2.2.6-A Basic performance for CSI predictionThe complexity values in terms of FLOPs and number of parameters of AI/ML models adopted in the evaluations of CSI prediction are summarized in Figure 6.2.2.6A-1.* Results refer to Table 2 of clause 7.3, R1-2407341.

Figure 6.2.2.6A-1: Complexity of AI/ML models from evaluation results in terms of FLOPs and number of parameters for CSI prediction. The complexity values in terms of FLOPs for AI/ML model and benchmark 2 of an auto-regression/Kalman filter adopted in the evaluations of CSI prediction are summarized in Figure 6.2.2.6A-2 and Figure 6.2.2.6A-3* Results refer to Table 2-9, and Table 2-10 in R1-2407341

Figure 6.2.2.6A-2: Complexity of AI/ML models and benchmark 2 of an auto-regression/Kalman filter from evaluation results in terms of FLOPs for CSI prediction. From a perspective of AI/ML complexity, 19 sources adopt the model subject to the computational complexity in units of FLOPs from 0.05M to 3000M. The actual model complexity may differ from the model complexity in the evaluation with respect to platform-dependent optimization on model implementations. From a perspective of complexity of non-AI/ML benchmark, 16 sources adopt the algorithm (e.g., Kalman filter, Auto-regression, Wiener filter) subject to the computational complexity in units of FLOPs from 0.14M to 107M. For non-AI/ML benchmark, main computation complexity is dominated by filter updates, which may not need to be updated per inference at the expense of performance loss. For example, 7 sources adopt the algorithm subject to the computational complexity of filter updates and inference in units of FLOPs from 0.47M to 106M and 0.067M to 3M, respectively. ***SGCS performance over benchmark 1 of the nearest historical CSI***For the CSI prediction using UE-sided model, compared to the Benchmark 1 of the nearest historical CSI, in terms of SGCS, from UE speed perspective, * + If spatial consistency is not adopted, and if N4=1
		- For 10km/h UE speed, 1 source observes 6% gain
		- For 30km/h UE speed, 9 sources observe 5.9%~20.6% gain and 4 sources observe 23.2%~35.4% gain. 2 sources observe 54%~106% gain
		- For 60km/h UE speed, 4 sources observe 10.5%~26.53% gain
	+ If spatial consistency is adopted, and if N4=1
		- For 30km/h UE speed, 2 sources observe 22.93%~23% gain, 1 source observe 68.5% gain
		- For 60km/h UE speed, 3 sources observe 16.2%~24.3% gain
	+ If spatial consistency is not adopted, and if N4=4
		- * For 10km/h UE speed, 1 source observes 9.1%~29.7% gain
			* For 30km/h UE speed, 1 source observes 4.3%~12.3% gain, and 2 sources observe 17.6%~35.4%gain
	+ If spatial consistency is adopted, and if N4=4
		- * For 10km/h UE speed, 1 source observes -1.61%~62.9% gain
			* For 30km/h UE speed, 1 source observes 23%~34% gain, 1 source observe 20.9%~76.4% gain
			* For 60km/h UE speed, 2 sources observe 5.96%~-22% gain,
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 20ms~50ms.
	+ A future 4ms or 5ms instance from the prediction output is considered for calculating the metric.
	+ 7 sources consider realistic channel estimation, and other sources consider ideal channel estimation.
	+ 1 source considers eigenvector as model input, and other sources considers Raw channel matrix as model input.
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, 1 source considers per layer raw channel matrix after pre-processing, and other sources do not consider pre/post processing.
	+ The performance metric is SGCS in linear value for layer 1.
* Note: N4 refers to the number of predicted CSI instances
* Note: Results refer to Table 2-1 of R1-2407338

***SGCS performance over benchmark 2 of the nearest historical CSI***For the CSI prediction using UE-sided model, compared to the Benchmark 2 of non-AI based CSI prediction, in terms of SGCS, from channel estimation perspective* If ideal channel estimation is adopted
	+ For N4=1
		- 9 sources observe -1.2%~6.3% gain
		- 1 source observes -2% ~17% gain depending on filter complexity and filter update
	+ For N4=4
		- 2 sources observe -1.6%~5.56%
		- 1 source observes -4.5%~6.93%
		- 1 source observes 3.1%~40.5%
* If realistic channel estimation is adopted,
	+ For N4=1
		- 5 sources observe 0.43%~7.26% gain
		- 2 sources observe 15.2%~19.5 gain
		- 1 source observes 68% gain
		- 1 source observes 5.8%~16.4% gain depending on traffic load
	+ For N4=4
		- 2 sources observe 0.1%~1.4% gain
		- 1 source observes 5%~29% gain
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms or 5ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix as model input and UE speed of 30km/h is assumed.
	+ 3 sources consider spatial consistency, and other sources do not consider spatial consistency.
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources do not consider pre/post processing
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
	+ The performance metric is SGCS in linear value for layer 1.
* Note: Results refer to Table 2-3 of R1-2407338
* Note: N4 refers to the number of predicted CSI instances

***Mean UPT performance over benchmark 1 of the nearest historical CSI***For the CSI prediction using UE-sided model, in terms of mean UPT, gains are observed compared to Benchmark 1 of the nearest historical CSI:* For FTP traffic with low RU (RU<=39%)
	+ For 30km/h UE speed and N4=1
		- 2 sources observe -0.2%~1% gain
		- 4 sources observe 4.4%~9% gain.
	+ For 60km/h UE speed, and N4=1
		- 2 sources observe 1.2%~5% gain;
	+ For 30km/h UE speed and N4=4
		- 1 source observes 0.3% gain.
		- 1 source observes 7% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 5% gain.
* For FTP traffic with mid RU (40<=RU<=69%)
	+ For 30km/h UE speed and N4=1
		- 2 sources observe 0.9%~2.3% gain
		- 1 source observes 8.5% gain
		- 2 sources observe 22.3%~37% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 1.8%~3.5% gain;
		- 1 source observes 21% gain.
	+ For 30km/h UE speed and N4=4
		- 1 source observes 2.3% gain.
		- 1 source observes 29% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 21% gain.
* For FTP traffic with high RU (RU>=70%)
	+ For 30km/h UE speed and N4=1
		- 4 sources observe 2%~3.6% gain
		- 2 sources observe 8.5%~31.7% gain.
	+ For 30km/h UE speed and N4=4
		- 1 source observes 7.2% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 4.2% gain;
* For full buffer traffic:
	+ For 30km/h UE speed and N4=1
		- 3 sources observe 27%~51% gain.
		- 2 sources observe 1.2%~8.7% gain.
	+ For 60km/h UE speed and N4=1
		- 1 source observes 12.1% gain.
	+ For 30km/h UE speed and N4=4
		- 1 source observes 11.6% gain.
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms ~ 20ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix is considered as model input
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources do not consider pre/post processing
	+ 3 sources consider spatial consistency, and other sources do not consider spatial consistency.
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
	+ 5 sources consider realistic channel estimation, and other sources consider ideal channel estimation.
	+ 1 source are modelled phase discontinuity, and other sources do not consider phase discontinuity
* Note: N4 refers to the number of predicted CSI instances
* Note: Results refer to Table 2-5/2-7 of R1-2407338

***5% UE UPT performance over benchmark 1 of the nearest historical CSI***For the CSI prediction using UE-sided model, in terms of 5% UE UPT, gains are observed compared to Benchmark 1 of the nearest historical CSI:* For FTP traffic with low RU (RU<=39%)
	+ For 30km/h UE speed and N4=1
		- 1 source observes 1.5% gain.
		- 4 sources observe 6.7%~15.7% gain.
		- 1 source observes 27% gain.
	+ For 60km/h UE speed, and N4=1
		- 2 sources observe 9%~18.3% gain;
	+ For 30km/h UE speed and N4=4
		- 1 source observes 3.6% gain.
		- 1 source observes 14% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 4% gain.
* For FTP traffic with mid RU (40<=RU<=69%)
	+ For 30km/h UE speed and N4=1
		- 1 source observes 4.1% gain.
		- 2 sources observe 14.3%~38.9% gain.
		- 1 source observes 100% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 11.3% gain;
		- 1 source observes 45% gain.
	+ For 30km/h UE speed and N4=4
		- 1 source observes 10% gain.
		- 1 source observes 77% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 26% gain.
* For FTP traffic with high RU (RU>=70%)
	+ For 30km/h UE speed, and N4=1
		- 1 source observes 51.7% gain.
		- 3 sources observe 7.8%~15.8% gain
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 17.8% gain;
	+ For 30km/h UE speed, and N4=4
		- 1 source observes 6.2% gain
* For full buffer traffic:
	+ For 30km/h UE speed and N4=1
		- 4 sources observe 7.7%~10.5% gain.
		- 1 source observes 39.7% gain.
	+ For 60km/h UE speed and N4=1
		- 1 source observes 20.6% gain.
	+ For 30km/h UE speed and N4=4
		- 1 source observes 2.1% gain.
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms ~ 20ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix is considered as model input
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources do not consider pre/post processing
	+ 3 sources consider spatial consistency, and other sources do not consider spatial consistency.
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
	+ 5 sources consider realistic channel estimation, and other sources consider ideal channel estimation.
	+ 1 source are modelled phase discontinuity, and other sources do not consider phase discontinuity
* Note: N4 refers to the number of predicted CSI instances
* Note: Results refer to Table 2-5/2-7 of R1-2407338

***Mean UPT performance over benchmark 2 of non-AI based CSI prediction*** For the CSI prediction using UE-sided model, in terms of mean UPT, gains are observed compared to Benchmark 2 of a non-AI/ML based CSI prediction approach:* For FTP traffic, with low RU (RU<=39%)
	+ For 30km/h UE speed and N4=1,
		- 2 sources observe 7.6%~9% gain.
		- 4 sources observe -2.41%~1.1% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes -3.4% gain
		- 2 sources observe 0.9%~1.2% gain;
		- 1 source observes 11% gain.
	+ For 30km/h UE speed and N4=4,
		- 1 source observes 13% gain.
		- 2 sources observe 0%~0.3% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 13%
		- 1 source observes 0.14% gain
* For FTP traffic, with mid RU (40<=RU<=69%)
	+ For 30km/h UE speed and N4=1,
		- 1 source observes 24% gain.
		- 3 sources observe 0.2%~5.1% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes -29.4% gain;
		- 1 source observes 3.1% gain;
		- 1 source observes 31% gain.
	+ For 30km/h UE speed and N4=4,
		- 1 source observes 35% gain.
		- 2 sources observe -0.25%~2% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes -16.2% gain.
		- 1 source observes 0.25% gain.
		- 1 source observes 32% gain.
* For FTP traffic, with high RU (RU>=70%)
	+ For 30km/h UE speed, and N4=1
		- 4 sources observe -0.2%~9.2% gain;
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 2.5% gain;
	+ For 30km/h UE speed and N4=4,
		- 2 sources observe -0.8%~0.11% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes -19.1% gain.
		- 1 source observes 0.92% gain.
* For full buffer traffic:
	+ For 30km/h UE speed and N4=1
		- 2 sources observe 10.6%~24% gain.
		- 2 sources observe 7.8%~9.7% gain.
		- 3 sources observe -0.6%~1.2% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 0.2% gain
		- 1 source observes 8.4% gain
	+ For 30km/h UE speed and N4= 4
		- 1 source observes 7% gain.
		- 1 source observes 6.8% gain
	+ For 60km/h UE speed and N4=4
		- 1 source observes 11.6% gain
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms ~ 20ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix is considered as model input
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources considers no pre/post processing.
	+ 3 sources consider spatial consistency, and other sources do not consider spatial consistency.
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
	+ 5 sources consider realistic channel estimation, and other sources consider ideal channel estimation.
	+ 1 source are modelled phase discontinuity, and other sources do not consider phase discontinuity
* Note: N4 refers to the number of predicted CSI instances
* Note: Results refer to Table 2-6/2-8 of R1-247339
* Note: The following boxchart shows the median, 0.75 quantile, 0.25 quantile, outliers, and min/max values excluding outliers, for low RU, Mid RU, High RU and full buffer with UE speed of 30km/h.

C:\Users\admin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\meanBM2.emf***5% UE UPT performance over benchmark 2 of non-AI based CSI prediction***For the CSI prediction using UE-sided model, in terms of 5% UE UPT, gains are observed compared to Benchmark 2 of a non-AI/ML based CSI prediction approach:* For FTP traffic, with low RU (RU<=39%)
	+ For 30km/h UE speed and N4=1,
		- 2 sources observe 18%~26% gain.
		- 4 sources observe -1.1%~7.6% gain.
	+ For 60km/h UE speed, and N4=1
		- 3 sources observe 1.9%~5.2% gain;
		- 1 source observes 17% gain.
	+ For 30km/h UE speed and N4=4,
		- 1 source observes 23% gain.
		- 1 source observes -3.7% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 19% gain.
		- 1 source observes 4% gain.
* For FTP traffic, with mid RU (40<=RU<=69%)
	+ For 30km/h UE speed and N4=1,
		- 1 source observes 46% gain.
		- 1 source observes 18.7% gain.
		- 2 sources observe 4%~6.6% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 8.6% gain;
		- 2 sources observe 23%~66% gain.
	+ For 30km/h UE speed and N4=4,
		- 1 source observes 73% gain.
		- 1 source observes -9% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 56% gain.
		- 1 source observes 11.5% gain.
* For FTP traffic, with high RU (RU>=70%)
	+ For 30km/h UE speed, and N4=1
		- 1 source observes 1.9% gain;
		- 2 sources observe 20.7%~26.3% gain;
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 14.8%~16% gain;
	+ For 30km/h UE speed and N4=4,
		- 1 source observes 0.9% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 22.8% gain.
* For full buffer traffic:
	+ For 30km/h UE speed and N4=1
		- 2 sources observe 15.7%~16.1% gain.
		- 4 sources observe 0.2%~6% gain.
		- 1 source observes -2% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 0.4% gain
		- 1 source observes 11.6% gain
	+ For 30km/h UE speed and N4= 4
		- 1 source observes 6.3% gain.
		- 1 source observes 21% gain
	+ For 60km/h UE speed and N4=4
	+ 1 source observes 26.7% gain
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms ~ 20ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix is considered as model input
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources considers no pre/post processing.
	+ 3 sources consider spatial consistency, and other sources do not consider spatial consistency.
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
	+ 5 sources consider realistic channel estimation, and other sources consider ideal channel estimation.
	+ 1 source are modelled phase discontinuity, and other sources do not consider phase discontinuity
* Note: N4 refers to the number of predicted CSI instances
* Note: Results refer to Table 2-6/2-8 of R1-2407339
* Note: The following boxchart shows the median, 0.75 quantile, 0.25 quantile, outliers, and min/max values excluding outliers, for low RU, Mid RU, High RU and full buffer with UE speed of 30km/h.

C:\Users\admin\AppData\Local\Microsoft\Windows\INetCache\Content.Word\5%BM2.emf***SGCS performance over benchmark 2 of non-AI based CSI prediction, impact of channel estimation error***For the CSI prediction using UE-sided model, compared to the Benchmark 2 of non-AI based CSI prediction, in terms of SGCS, from channel estimation perspective* If ideal channel estimation is adopted
	+ For N4=1
		- 9 sources observe -1.2%~6.3% gain
		- 1 source observes -2% ~17% gain depending on filter complexity and filter update
	+ For N4=4
		- 2 sources observe -1.6%~5.56%
		- 1 source observes -4.5%~6.93%
		- 1 source observes 3.1%~40.5%
* If realistic channel estimation is adopted,
	+ For N4=1
		- 5 sources observe 0.43%~7.26% gain
		- 2 sources observe 15.2%~19.5 gain
		- 1 source observes 68% gain
		- 1 source observes 5.8%~16.4% gain depending on traffic load
	+ For N4=4
		- 2 sources observe 0.1%~1.4% gain
		- 1 source observes 5%~29% gain
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms or 5ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix as model input and UE speed of 30km/h is assumed.
	+ 3 sources consider spatial consistency, and other sources do not consider spatial consistency.
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources do not consider pre/post processing
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
	+ The performance metric is SGCS in linear value for layer 1.
* Note: Results refer to Table 2-3 of R1-2407338
* Note: N4 refers to the number of predicted CSI instances

***Mean UPT performance over benchmark 2 of non-AI based CSI prediction, impact of channel estimation error***For the CSI prediction using UE-sided model, till the RAN1#118 meeting, in terms of mean UPT, gains are observed compared to Benchmark#2 of a non-AI/ML based CSI prediction, from channel estimation perspective:* For FTP traffic, with low RU (RU<=39%)
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 2 sources observe -2.41%~1.8% gain.
		- For 60km/h UE speed, and N4=1
			* 2 sources observe -3.4%~0.9% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes 0.3% gain.
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1,
			* 2 sources observe 7.6%~9% gain.
			* 3 sources observe 0%~1.1% gain.
		- For 60km/h UE speed, and N4=1
			* 2 sources observe -3.4%~1.2% gain
			* 1 source observes 11% gain.
		- For 30km/h UE speed and N4=4,
			* 1 source observes 13% gain.
			* 1 source observes 0% gain.
		- For 60km/h UE speed, and N4=4
			* 1 source observes 13%
			* 1 source observes 0.14% gain
* For FTP traffic, with mid RU (40<=RU<=69%)
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes -4.5% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes -7.1% gain
			* 1 source observes 3.1% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes -2% gain.
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1,
			* 2 sources observe 24% gain.
			* 3 sources observe 0.2%~5.1% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes 31% gain
			* 1 source observes -29.4% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes 35% gain.
			* 2 sources observe -0.25%~1.1% gain
		- For 60km/h UE speed, and N4=4
			* 1 source observes 32%
			* 1 source observes 0.25% gain
* For FTP traffic, with high RU (RU>=70%)
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes -4.8% gain
		- For 60km/h UE speed, and N4=1
			* 1 source observes -9% gain
			* 1 source observes 2.5% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes -0.8% gain.
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1,
			* 3 sources observe 0%~0.8%
			* 1 source observes 9.2% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes -9% gain
		- For 30km/h UE speed and N4=4,
			* 2 sources observe 0%~0.1% gain
		- For 60km/h UE speed, and N4=4
			* 1 source observes 0.92% gain
* For full buffer model,
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes 24%
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1
			* 3 sources observe 7.8%~10.6% gain.
			* 3 sources observe -0.6%~1.2% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes 0.2% gain
			* 1 source observes 8.4% gain
		- For 30km/h UE speed and N4= 4
			* 1 source observes 7% gain.
			* 1 source observes 6.8% gain
		- For 60km/h UE speed and N4=4
			* 1 source observes 11.6% gain
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms ~ 20ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix is considered as model input
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources considers no pre/post processing.
	+ 3 sources consider spatial consistency, and other sources do not consider spatial consistency.
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
* Note: N4 refers to the number of predicted CSI instances
* Note: Results refer to Table 2-6/2-8 of R1-2407340

***5% UE UPT performance over benchmark 2 of non-AI based CSI prediction, impact of channel estimation error***For the CSI prediction using UE-sided model, in terms of 5% UE UPT, gains are observed compared to Benchmark 2 of a non-AI/ML based CSI prediction, from channel estimation perspective:* For FTP traffic, with low RU (RU<=39%)
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes -5.5% gain.
		- For 60km/h UE speed, and N4=1
			* 2 sources observe 4%~4.3% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes -3.7% gain.
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1,
			* 2 sources observe 17% gain.
			* 2 sources observe 0%~4% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes 1.9% gain
			* 1 source observes 17% gain.
		- For 30km/h UE speed and N4=4,
			* 1 source observes 23% gain.
		- For 60km/h UE speed, and N4=4
			* 1 source observes 19%
* For FTP traffic, with mid RU (40<=RU<=69%)
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes -12.9% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes 2.6% gain
			* 1 source observes 8.6% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes -9% gain.
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1,
			* 2 sources observe 4%~6.6% gain.
			* 1 source observes 18.7% gain.
			* 1 source observes 46% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes 66% gain
			* 1 source observes 2.6% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes 73% gain.
			* 1 source observes 18.7% gain
		- For 60km/h UE speed, and N4=4
			* 1 source observes 56%
* For FTP traffic, with high RU (RU>=70%)
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes 3.6% gain
		- For 60km/h UE speed, and N4=1
			* 1 source observes -10.7% gain
			* 1 source observes 14.8% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes 0.9% gain.
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1,
			* 2 sources observe 20.7%~26.3%
			* 1 source observes 1.9% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes 3.6% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes 0.9% gain
* For full buffer model,
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes 0.2%
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1
			* 2 sources observe 15.7%~16.1% gain.
			* 3 sources observe 2.6%~7.7% gain.
			* 1 source observes -2% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes 0.4% gain
			* 1 source observes 11.6% gain
		- For 30km/h UE speed and N4= 4
			* 1 source observes 6.3% gain.
			* 1 source observes 21% gain
		- For 60km/h UE speed and N4=4
			* 1 source observes 26.7% gain
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms ~ 20ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix is considered as model input
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources considers no pre/post processing.
	+ 3 sources consider spatial consistency, and other sources do not consider spatial consistency.
	+ 1 source considers 100% in car UE distribution and other sources consider 100% outdoor UE distribution.
* Note: N4 refers to the number of predicted CSI instances
* Note: Results refer to Table 2-6/2-8 of R1-2407340

***SGCS performance over benchmark 2 of non-AI based CSI prediction, impact of phase discontinuity***For the CSI prediction using UE-sided model, compared to the Benchmark 2 of non-AI based CSI prediction, in terms of SGCS, from perspective of phase discontinuity modelling* If phase discontinuity modelling is not adopted
	+ For 30km/h UE speed,
		- 2 sources observe 4.24%~4.3% gain with N4=1
		- 1 source observes 3.1% ~ 40.5% with N4=4
		- 1 source observes -1.2%~6.3% with N4=1 depending on filter updates between 10 and 70 learning window size
	+ For 60km/h UE speed,
		- 1 source observes 13.5% gain with N4=1, and 1 source observes 8.2%~34.6% gain with N4=4
		- 1 source observes 0%~8% with N4=1 depending on filter updates between 10 and 70 learning window size
* If phase discontinuity modelling is adopted,
	+ For 30km/h UE speed,
		- Without phase adjustment,
			* 2 sources observe 7.26%~7.8% gain with N4=1
			* 1 source observes 6.5%~23.5% gain with N4=4
			* 1 source observes -0.8%~8% with N4=1 depending on filter updates between 10 and 70 learning window size
		- With phase adjustment,
			* 1 source observes 1.2%~7.7% gain with N4=1 depending on filter updates between 10 and 70 learning window size
			* 1 source observes 25.6% gain
	+ For 60km/h UE speed,
		- Without phase adjustment,
			* 1 source observes 6.6% gain with N4=1 and 6.4%~55% with N4=4
			* 1 source observes 5.7%~17.5% gain with N4=1 depending on filter updates between 10 and 70 learning window size
		- With phase adjustment, 1 source observes 4.2%~11.9% gain with N4=1 depending on filter updates between 10 and 70 learning window size
* Note: the above results are based on the following assumptions
	+ The observation window considers to start as early as 15ms~50ms.
	+ A future 4ms or 5ms instance from the prediction output is considered for calculating the metric.
	+ Raw channel matrix as model input is assumed.
	+ The performance metric is SGCS in linear value for layer 1.
* Note: Results refer to Table 2-12 of R1-2407338
* Note: N4 refers to the number of predicted CSI instances

***UPT performance with 20ms CSI-RS periodicity*** For the CSI prediction using CSI-RS with 20ms periodicity, till the RAN1#118 meeting, in terms of mean and 5% UE UPT, gains are observed compared to Benchmark #1 of the nearest historical CSI and Benchmark #2 of non-AI/ML based CSI prediction, * Compared to Benchmark 1
	+ For N4=1 and UE speed of 3km/h, 1 source observes -18%~-7% and -26%~-8% degradation in terms of mean and 5% UE UPT, respectively.
	+ For N4=4 and UE speed of 3km/h, 1 source observes 1%~6% and 6%~17% gain in terms of mean and 5% UE UPT, respectively.
	+ For N4=4 and UE speed of 30km/h, 1 source observes 2%~17% and 13%~19% gain in terms of mean and 5% UE UPT, respectively, by using one 20ms periodic CSI-RS and three aperiodic CSI-RS.
* Compared to Benchmark 2
	+ For N4=1 and UE speed of 3km/h,
		- 1 source observes 10%~19% and 23%~31% gain in terms of mean and 5% UE UPT, respectively.
		- 1 source observes -2.7%~-1% and -10.4%~-4.9% degradation in terms of mean and 5% UE UPT, respectively.
	+ For N4=1 and UE speed of 10km/h, 1 source observes 0.6%~1% gain and -14.2%~-8% gain in terms of mean and 5% UE UPT, respectively.
	+ For N4=4 and UE speed of 3km/h, 1 source observes 14%~30% and 19%~56% gain in terms of mean and 5% UE UPT, respectively.
	+ For N4=4 and UE speed of 30km/h,
		- 1 source observes 17.3% and 16.1 gain in terms of mean and 5% UE UPT, receptively.
		- 1 source observes 5%~19% and 13%~40% gain in terms of mean and 5% UE UPT, respectively, by using one 20ms periodic CSI-RS and three aperiodic CSI-RS.
* Note: the above results are based on the following assumptions besides the assumptions of the agreed EVM table
	+ 1 source considers full buffer model, other sources consider FTP traffic model
	+ Raw channel matrix is used as the model input.
	+ 2 sources consider realistic channel estimation, and other sources consider ideal channel estimation.
	+ The performance metric is SGCS in linear value for layer 1/2/3/4 and mean UPT, 5% UPT.
	+ 1 sources consider spatial consistency. Other sources do not consider spatial consistency.
	+ Note: Results refer to Table 2-13, 2-14, 2-15, 2-16 of R1-2407339

--------------------------------------------------------Text omitted --------------------------------------------------------- |
|  |

## TP 7

|  |
| --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------6.2.2.7-A Generalization evaluations for CSI prediction***Generalization over UE speeds***For the generalization verification of CSI prediction using UE sided model over various UE speeds, compared to the generalization Case 1 where the AI/ML model is trained with dataset subject to a certain UE speed#B and applied for inference with a same UE speed#B,* For generalization Case 2, generalized performance may be achieved for some certain combinations of UE speed#A and UE speed#B but not for others:
	+ If UE speed#B is 10 km/h and
		- UE speed#A is 30 km/h or 120km/h, 2 sources observe a generalized performance of less than -5.5% degradation, and 1 source observes -16.2% degradation.
		- UE speed#A is 60 km/h, 2 sources observe -26.79%~-13.3% degradation.
	+ If UE speed#B is 30 km/h and
		- UE speed#A is 60km/h, 4 sources observe a generalized performance of -11.4%~-2.7% degradation and 4 sources observe a generalized performance of -33.6%~-19% degradation.
		- UE speed#A is 10km/h or 120km/h, 3 sources observe a generalized performance of -72.37%~-51.5% degradation.
	+ If UE speed#B is 60 km/h and
		- UE speed#A is 10km/h, 1 source observes a generalized performance of -16.4 degradation
		- UE speed#A is 30km/h, 1 source observes a generalized performance of -5.4% degradation and 6 sources observe a generalized performance of -56.3%~-13.8% degradation
	+ If UE speed#B is 120 km/h and
		- UE speed#A is 30km/h, 1 source observes a generalized performance of -56.3% degradation
* For generalization Case 3, generalized performance of the AI/ML model can be achieved in general (0%~-3.8% loss) for UE speed#B subject to any of 10 km/h, 30 km/h, 60 km/h and 120 km/h, if the training dataset is constructed with data samples subject to multiple UE speeds including UE speed#B, as observed by 8 sources
	+ If UE speed#B is 10 km/h
		- 2 sources observe a generalized performance of less than -0.2% degradation, 1 source observes -5.1% degradation, and 1 source observes -16.87% degradation.
	+ If UE speed#B is 30 km/h
		- 9 sources observe a generalized performance of less than -3.8% degradation, 1 source observes a generalized performance of -15.44% degradation.
	+ If UE speed#B is 60 km/h and
		- 8 sources observe a generalized performance of less than -3.63% degradation, 2 sources observe a generalized performance of -13.5%~-6.77% degradation.
	+ If UE speed#B is 120 km/h and
		- 1 source observes a generalized performance of -43.6% degradation
	+ Note: For generalization Case 3, 2 sources observe performance degradations (-43.6%~-13.5% loss) for UE speed#B subject to 10 km/h, 30 km/h, 60 km/h, 120 km/h, but compared with generalization Case 2, in general the performance are still improved.
* Note: the above results are based on the following assumptions besides the assumptions of the agreed EVM table
	+ Raw channel matrix is used as the model input.
	+ The performance metric is SGCS in linear value for layer 1/2/3/4.
	+ 4 sources consider spatial consistency. Other sources do not consider spatial consistency.
	+ 2 sources consider beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, 1 source considered antenna(port)-delay domain transformation/ antenna(port)-frequency domain transformation as pre/post processing, and other sources do not consider pre/post processing
	+ Note: Results refer to Table 3-1 of R1-2407338

***Generalization over deployment scenarios***For the generalization verification of CSI prediction using UE sided model over various deployment scenarios, compared to the generalization Case 1 where the AI/ML model is trained with dataset subject to a certain deployment scenario#B applied for inference with a same deployment scenario#B * For generalization Case 2, generalized performance may be achieved for some certain combinations of deployment scenario#A and deployment scenario#B but not for others:
	+ For deployment scenario#B is UMa
		- 2 sources observe -1.88%~0% degradation
		- 1 source observes -6.8% degradation
	+ For deployment scenario#B is UMi
		- 1 source observes 0% degradation
		- 3 sources observe -4.85%~-3.03% degradation
* For generalization Case 3, generalized performance of the AI/ML model can be achieved (-1.95%~0% loss) for deployment scenario#B subject to any of UMa and UMi, if the training dataset is constructed with data samples subject to multiple deployment scenarios including deployment scenario#B as observed by 3 sources.
	+ Minor loss (0%~-1.95%) are observed by 3 sources
	+ Note: Moderate degradations of -5.2% are observed by 1 source for deployment scenario#B subject to Uma
* Note: the above results are based on the following assumptions besides the assumptions of the agreed EVM table
	+ Raw channel matrix is used as the model input.
	+ The performance metric is SGCS in linear value for layer 1/2/3/4.
	+ 3 sources consider spatial consistency. Other sources do not consider spatial consistency.
	+ 1 source considers beam-delay domain transformation/antenna-frequency domain transformation as pre/post processing, and other sources considers no pre/post processing.
* Note: Results refer to Table 3-2 of R1-2407339

***Generalization over carrier frequency***For the generalization verification of CSI prediction using UE sided model over various carrier frequency, compared to the generalization Case 1 where the AI/ML model is trained with dataset subject to a certain carrier frequency#B applied for inference with a same carrier frequency#B * For generalization Case 2, significant degradations are suffered in general from the perspective of the layouts of antenna ports, as observed by 3 sources:
	+ For carrier frequency#B is 2GHz
		- 1 source observes -11.4% degradation
		- 1 source observes -80.53% degradation
	+ For carrier frequency#B is 3GHz or 4GHz
		- 2 sources observe -34.23%~-80.53% degradation
		- 1 source observes -14.21% degradation
* For generalization Case 3, generalized performance may be achieved for some certain combinations of carrier frequency#A and carrier frequency#B but not for others, if the training dataset is constructed with data samples subject to multiple carrier frequencies including carrier frequency#B
	+ For carrier frequency#B is 2GHz
		- 1 source observes -0.5% degradation
		- 1 source observes -9.27% degradation
	+ For carrier frequency#B is 3GHz or 4GHz
		- 2 sources observe -1.93%~-5.1% degradation
		- 1 source observes -14.94% degradation
* Note: the above results are based on the following assumptions besides the assumptions of the agreed EVM table
	+ Raw channel matrix is used as the model input.
	+ The performance metric is SGCS in linear value for layer 1/2/3/4.
	+ 2 sources consider spatial consistency. Other sources do not consider spatial consistency.
* Note: Results refer to Table 3-3 of R1-2407339

***Generalization over multiple aspects***For the generalization verification of CSI prediction using UE sided model over multiple aspects, compared to the generalization Case 1 where the AI/ML model is trained with dataset subject to certain aspects #B applied for inference with the same aspects #B,* For generalization Case 2,
	+ - 1 source observes -9.8% ~ -1.5% degradation when the aspects #A is (2 GHz carrier frequency, 100% outdoor UE) and the aspects #B is (4GHz carrier frequency, 20% outdoor UE+80% indoor UE)
		- 1 source observes -10.7%~-1.8% when the aspects#A is (UMa, 2 GHz carrier frequency, 100% outdoor UE distribution) and the aspects#B is (UMi, 4 GHz carrier frequency, 100 outdoor UE)
		- 1 source observes -21.2%~-2.4% when the aspects#A is (UMa, 2 GHz carrier frequency, 100% outdoor UE distribution) and the aspects#B is (UMi, 4 GHz carrier frequency, 20% outdoor UE+80% indoor UE distribution)
		- 1 source observes -3% degradation when aspects#A is (30km/h UE speed, 100% UE in a car, 2GHz carrier frequency) and the aspects#B is (3km/h UE speed, 20% outdoor UE, 4GHz carrier frequency)
* For generalization Case 3,
	+ - 1 source observes -1.1%~0% when the aspects#A is (2 GHz carrier frequency, 100% outdoor UE distribution) and the aspects #B is (4GHz carrier frequency, 20% outdoor UE+80% indoor UE distribution)
		- 1 source observes -2.4%~-0.8% when the aspects#A is (UMa, 2 GHz carrier frequency, 100% outdoor UE distribution) and the aspects#B is (UMi, 4 GHz carrier frequency, 100 outdoor UE)
		- 1 source observes -3.4%~-0.3% when the aspects#A is (UMa, 2 GHz carrier frequency, 100% outdoor UE distribution) and the aspects#B is (UMi, 4 GHz carrier frequency, 20% outdoor UE+80% indoor UE distribution)
* Note: the above results are based on the following assumptions besides the assumptions of the agreed EVM table
	+ The performance metric is SGCS in linear value for layer 1/2/3/4.
	+ 1 source considers eigenvector as model input, and 1 source considers Raw channel matrix as model input.
* Note: Results refer to Table 3-4 of R1-2407339

--------------------------------------------------------Text omitted --------------------------------------------------------- |

## TP 8

|  |
| --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------6.2.2.8-A Summary of Performance Results for CSI feedback enhancement The following aspects have been studied for CSI prediction using UE-sided model in Rel-19:* From the perspective of basic performance gain over non-AI/ML benchmark (without considering generalization),
	+ It has been studied with corresponding observations on:
		- the metrics of SGCS, mean UPT, 5% UPT;
		- the benchmarks of nearest historical CSI and non-AI/ML based CSI prediction.
		- The impact of channel estimation error, phase discontinuity, spatial consistency, UE Speed, observation window, prediction window, CSI-RS periodicity.
* It has been studied with corresponding observations on complexity for both AI/ML based CSI prediction and non-AI/ML based CSI prediction.
* It has been studied on localized model including evaluation methodology, but is lack of observations.
* From the perspective of generalization over various scenarios,
	+ It has been studied with corresponding observations on (with the metric of SGCS):
		- the scenario including various UE speeds, deployment scenarios, carrier frequency
* From the perspective of performance monitoring, it has been studied on boundary between Type 1 and 3 performance monitoring, and potential specification impact for each performance monitoring type 1,2 and 3.

Based on the evaluation for CSI prediction in Rel-19, the following high-level observations are provided:* From the perspective of basic performance gain over benchmark of non-AI/ML based CSI prediction, under the same UE speed for training and inference
	+ for AI/ML based CSI prediction over non-AI/ML based CSI prediction, 0%~7.8% gain depending on traffic model, in terms of mean UPT, is observed by 7 sources
	+ for AI/ML based CSI prediction over non-AI/ML based CSI prediction, 3.8%~20.7% gain depending on traffic model, in terms of 5% UE UPT, is observed by 7 sources.
* From a perspective of AI/ML complexity, 19 sources adopt the model subject to the computational complexity in units of FLOPs from 0.05M to 3000M. For complexity of non-AI/ML benchmark, 16 sources adopt the algorithm (e.g., Kalman filter, Auto-regression, Wiener filter) subject to the computational complexity in units of FLOPs from 0.14M to 107M. The ratio of FLOPs (AI/ML over benchmark 2) ranges from 1 to 35, which is used by majority sources. For non-AI/ML benchmark, main computation complexity is dominated by filter updates, which may not be need to be updated per inference at the expense of performance loss.
* From the perspective of performance impact on channel estimation error and phase discontinuity, compared to non-AI/ML CSI prediction, higher gain is observed by 10 sources for AI/ML based CSI prediction in the presence of channel estimation error or phase discontinuity.
* From the perspective of generalization over various scenarios/configurations (e.g., various UE speed, deployment scenario, carrier frequency) that have been evaluated, compared to generalization Case 1 where the AI/ML model is trained with dataset subject to a certain scenario#B/configuration#B and applied for inference with a same scenario#B/configuration#B
	+ For generalization Case 2 where the AI/ML model is trained with dataset from a different scenario#A/configuration#A, generalized performance may be achieved for some certain combinations of scenario#A/configuration#A and scenario#B/configuration#B but not for others.
	+ For generalization Case 3 where the training dataset is constructed with data samples subject to multiple scenarios/configurations including scenario#B/configuration#B, generalized performance of the AI/ML model can be achieved.

--------------------------------------------------------Text omitted --------------------------------------------------------- |

## TP 9

|  |
| --- |
| --------------------------------------------------------Text omitted ---------------------------------------------------------8 Conclusions**CSI prediction sub use case in Rel-19:** From RAN1 perspective, study of CSI prediction has been completed and performance improvement is observed with increased complexity.--------------------------------------------------------Text omitted --------------------------------------------------------- |

## Comments/input

Please provide your input, if any.

|  |  |
| --- | --- |
| **Company** | **Views** |
| New H3C | For TP#9, from our perspective, 2nd bullet from RAN1#118 agreement is fine. I wonder whether there is agreement to support/correspond to 1st bullet as follows:The performance gain over Rel-18 non-AI/ML based approach and associated complexity were studied for AI/ML based UE side CSI prediction sub use case in Rel-19. The intention of this sentence isn’t clear to us and plz elaborate the reason on why to add this sentence in TR in detail.[Mod] Thanks for the comment. I don’t have any intention on first and second bullet. And as you indicated, there is no such agreement on that. It is just coming from “SID context” as below.“For CSI prediction (UE-sided model), further study performance gain over Rel-18 non-AI/ML based approach and associated complexity, while addressing other aspects requiring further study/conclusion as captured in the conclusions section of the TR 38.843 (e.g., cell/site specific model could be considered to improve performance gain).”If companies are not comfortable with first and second bullet, I’m ok to delete it. Please check current version. |
| Samsung | Please check the comments above.* [Mod] Thanks for the comment, I only can see one comment of T\_window which is now fixed. If you have other comments that I cannot find, please let me know.
* Please check above on your comments.
 |
| CATT | For TP#6 *UPT performance with 20ms CSI-RS periodicity* , we noticed that in RAN1#118 Chair’s Note, the number of source companies is not correct in some bullets in the Note:ObservationFor the CSI prediction using CSI-RS with 20ms periodicity, till the RAN1#118 meeting, in terms of mean and 5% UE UPT, gains are observed compared to Benchmark #1 of the nearest historical CSI and Benchmark #2 of non-AI/ML based CSI prediction, …(omitted)* Note: the above results are based on the following assumptions besides the assumptions of the agreed EVM table
	+ 1 source [ZTE] considers full buffer model, other sources consider FTP traffic model
	+ Raw channel matrix is used as the model input.
	+ ~~5~~ 2 sources [Ericsson, ZTE] consider realistic channel estimation, and other sources consider ideal channel estimation.
	+ The performance metric is SGCS in linear value for layer 1/2/3/4 and mean UPT, 5% UPT.
	+ ~~2~~1 sources [Ericsson] consider spatial consistency. Other sources do not consider spatial consistency.
	+ Note: Results refer to Table 2-13, 2-14, 2-15, 2-16 of R1-2407339

We suggest to capture the corrected number in the TP and to fix the number in RAN1#118 Chair’s Note. [Mod] Thanks for pointing this typo out. I fixed it.  |
| ZTE | For “Generalization over carrier frequency” inTP7, there is a wrong result of our company captured in the observation agreed in RAN1#118 meeting, as yellow highlighted in the following text. The correct number should be **-14.21%** from the evaluation Table 7-A, which is not -4.21%. So, please FL double check and further modify it to the correct number. If correct, TP7 should be further modified as well, thanks. **-----------------**Text omitted **---------------*** For generalization Case 2, significant degradations are suffered in general from the perspective of the layouts of antenna ports, as observed by 3 sources:
	+ For carrier frequency#B is 2GHz
		- 1 source [MediaTek] observe -11.4% degradation
		- 1 source [vivo] observe -80.53% degradation
	+ For carrier frequency#B is 3GHz or 4GHz
		- 2 sources [MediaTek, vivo] observe -34.23%~-80.53% degradation
		- 1 source [ZTE] observe ~~-4.21%~~14.21% degradation

**-----------------**Text omitted **---------------**[Mod] Thanks for pointing this typo out, and sorry for incorrectly capturing your results. I fixed it. |
| Mod | Please provide input if you have concern on TP or correction.  |
| Panasonic | The following potential typos could be corrected in the TP.**[TP6]**From a perspective of complexity of non-AI/ML benchmark, 16 sources adopt the algorithm (e.g., Kalman filter, Auto-regression, Wiener filter) subject to the computational complexity in units of FLOPs from 0.14M to 107M. For non-AI/ML benchmark, main computation complexity is dominated by filter updates, which may not ~~be~~ need to be updated per inference at the expense of performance loss. For example, 7 sources adopt the algorithm subject to the computational complexity of filter updates and inference in units of FLOPs from 0.47M to 106M and 0.067M to 3M, respectively.***SGCS performance over benchmark 1 of the nearest historical CSI***<…>* + If spatial consistency is adopted, and if N4=1
		- For 30km/h UE speed, 2 sources observe 22.93%~23% gain, 1 source observe 68.5% gain
		- For 60km/h UE speed, 3 sources observe 16.2%~24.3% gain

***Mean UPT performance over benchmark 1 of the nearest historical CSI***<…>* For FTP traffic with mid RU (40<=RU<=69%)
	+ For 30km/h UE speed and N4=1
		- 2 sources observe 0.9%~2.3% gain
		- 1 source observes 8.5% gain
		- 2 sources observe 22.3%~37% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 1.8%~3.5% gain;
		- 1 source observes 21% gain.
	+ For 30km/h UE speed and N4=4
		- 1 source observes 2.3% gain.
		- 1 source observes 29% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 21% gain.

<…>* For full buffer traffic:
	+ For 30km/h UE speed and N4=1
		- 3 sources observe 27%~51% gain.
		- 2 sources observe 1.2%~8.7% gain.
	+ For 60km/h UE speed and N4=1
		- 1 source observes 12.1% gain.
	+ For 30km/h UE speed and N4=4
		- 1 source ~~[Fujitsu]~~ observes 11.6% gain.

***5% UE UPT performance over benchmark 1 of the nearest historical CSI***<…>* For FTP traffic with high RU (RU>=70%)
	+ For 30km/h UE speed, and N4=1
		- 1 source observes 51.7% gain.
		- 3 sources observe 7.8%~15.8% gain
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 17.8% gain;
	+ For 30km/h UE speed, and N4=4
		- 1 source observes 6.2% gain

***Mean UPT performance over benchmark 2 of non-AI based CSI prediction*** * For FTP traffic, with low RU (RU<=39%)
	+ For 30km/h UE speed and N4=1,
		- 2 sources observe 7.6%~9% gain.
		- 4 sources observe -2.41%~1.1% gain.
	+ For 60km/h UE speed, and N4=1
		- 1 source observes -3.4% gain
		- 2 sources observe~~s~~ 0.9%~1.2% gain;
		- 1 source observes 11% gain.
	+ For 30km/h UE speed and N4=4,
		- 1 source observes 13% gain.
		- 2 sources observe 0%~0.3% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 13%
		- 1 source observes 0.14% gain

<…>* For FTP traffic, with high RU (RU>=70%)
	+ For 30km/h UE speed, and N4=1
		- 4 sources observe -0.2%~9.2% gain;
	+ For 60km/h UE speed, and N4=1
		- 1 source observes 2.5% gain;
	+ For 30km/h UE speed and N4=4,
		- 2 sources observe~~s~~ -0.8%~0.11% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes -19.1% gain.
		- 1 source observes 0.92% gain.

***5% UE UPT performance over benchmark 2 of non-AI based CSI prediction**** For FTP traffic, with low RU (RU<=39%)
	+ For 30km/h UE speed and N4=1,
		- 2 sources observe 18%~26% gain.
		- 4 sources observe -1.1%~7.6% gain.
	+ For 60km/h UE speed, and N4=1
		- 3 sources observe 1.9%~5.2% gain;
		- 1 source observes 17% gain.
	+ For 30km/h UE speed and N4=4,
		- 1 source observes 23% gain.
		- 1 source observes -3.7% gain.
	+ For 60km/h UE speed, and N4=4
		- 1 source observes 19% gain.
		- 1 source observes 4% gain.

***5% UE UPT performance over benchmark 2 of non-AI based CSI prediction, impact of channel estimation error**** For FTP traffic, with low RU (RU<=39%)
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes -5.5% gain.
		- For 60km/h UE speed, and N4=1
			* 2 sources observe 4%~4.3% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes -3.7% gain.

<…>* For FTP traffic, with high RU (RU>=70%)
	+ With ideal channel estimation
		- For 30km/h UE speed and N4=1,
			* 1 source observes 3.6% gain
		- For 60km/h UE speed, and N4=1
			* 1 source observes -10.7% gain
			* 1 source observes 14.8% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes 0.9% gain.
	+ With realistic channel estimation
		- For 30km/h UE speed and N4=1,
			* 2 sources observe~~s~~ 20.7%~26.3%
			* 1 source observes 1.9% gain.
		- For 60km/h UE speed, and N4=1
			* 1 source observes 3.6% gain
		- For 30km/h UE speed and N4=4,
			* 1 source observes 0.9% gain

***SGCS performance over benchmark 2 of non-AI based CSI prediction, impact of phase discontinuity**** If phase discontinuity modelling is not adopted
	+ For 30km/h UE speed,
		- 2 sources observe 4.24%~4.3% gain with N4=1
		- 1 source observes 3.1% ~ 40.5% with N4=4
		- 1 source observes -1.2%~6.3% with N4=1 depending on filter updates between 10 and 70 learning window size
	+ For 60km/h UE speed,
		- 1 source observes 13.5% gain with N4=1, and 1 source ~~[vivo]~~ observes 8.2%~34.6% gain with N4=4
		- 1 source observes 0%~8% with N4=1 depending on filter updates between 10 and 70 learning window size
* If phase discontinuity modelling is adopted,

<…>* + For 60km/h UE speed,
		- Without phase adjustment,
			* 1 source observes 6.6% gain with N4=1 and 6.4%~55% with N4=4
			* 1 source observes 5.7%~17.5% gain with N4=1 depending on filter updates between 10 and 70 learning window size
		- With phase adjustment, 1 source observes 4.2%~11.9% gain with N4=1 depending on filter updates between 10 and 70 learning window size

**[TP7]*****Generalization over UE speeds**** For generalization Case 2, generalized performance may be achieved for some certain combinations of UE speed#A and UE speed#B but not for others:

<…>* + If UE speed#B is 60 km/h and
		- UE speed#A is 10km/h, 1 source observes a generalized performance of -16.4 degradation
		- UE speed#A is 30km/h, 1 source observes a generalized performance of -5.4% degradation and 6 sources observe a generalized performance of -56.3%~-13.8% degradation
	+ If UE speed#B is 120 km/h and
		- UE speed#A is 30km/h, 1 source observes a generalized performance of -56.3% degradation
* For generalization Case 3, generalized performance of the AI/ML model can be achieved in general (0%~-3.8% loss) for UE speed#B subject to any of 10 km/h, 30 km/h, 60 km/h and 120 km/h, if the training dataset is constructed with data samples subject to multiple UE speeds including UE speed#B, as observed by 8 sources
	+ If UE speed#B is 10 km/h
		- 2 sources observe a generalized performance of less than -0.2% degradation, 1 source observes -5.1% degradation, and 1 source observes -16.87% degradation.

***Generalization over deployment scenarios**** For generalization Case 2, generalized performance may be achieved for some certain combinations of deployment scenario#A and deployment scenario#B but not for others:
	+ For deployment scenario#B is UMa
		- 2 sources observe -1.88%~0% degradation
		- 1 source observes -6.8% degradation
	+ For deployment scenario#B is UMi
		- 1 source observes 0% degradation
		- 3 sources observe -4.85%~-3.03% degradation

***Generalization over carrier frequency**** For generalization Case 2, significant degradations are suffered in general from the perspective of the layouts of antenna ports, as observed by 3 sources:
	+ For carrier frequency#B is 2GHz
		- 1 source observes -11.4% degradation
		- 1 source observes -80.53% degradation
	+ For carrier frequency#B is 3GHz or 4GHz
		- 2 sources observe -34.23%~-80.53% degradation
		- 1 source observes -14.21% degradation
* For generalization Case 3, generalized performance may be achieved for some certain combinations of carrier frequency#A and carrier frequency#B but not for others, if the training dataset is constructed with data samples subject to multiple carrier frequencies including carrier frequency#B
	+ For carrier frequency#B is 2GHz
		- 1 source observes -0.5% degradation
		- 1 source observes -9.27% degradation
	+ For carrier frequency#B is 3GHz or 4GHz
		- 2 sources observe -1.93%~-5.1% degradation
		- 1 source observes -14.94% degradation

***Generalization over multiple aspects***For the generalization verification of CSI prediction using UE sided model over multiple aspects, compared to the generalization Case 1 where the AI/ML model is trained with dataset subject to certain aspects #B applied for inference with the same aspects #B,* For generalization Case 2,
	+ - 1 source observes -9.8% ~ -1.5% degradation when the aspects #A is (2 GHz carrier frequency, 100% outdoor UE) and the aspects #B is (4GHz carrier frequency, 20% outdoor UE+80% indoor UE)
		- 1 source observes -10.7%~-1.8% when the aspects#A is (UMa, 2 GHz carrier frequency, 100% outdoor UE distribution) and the aspects#B is (UMi, 4 GHz carrier frequency, 100 outdoor UE)
		- 1 source observes -21.2%~-2.4% when the aspects#A is (UMa, 2 GHz carrier frequency, 100% outdoor UE distribution) and the aspects#B is (UMi, 4 GHz carrier frequency, 20% outdoor UE+80% indoor UE distribution)
		- 1 source observes -3% degradation when aspects#A is (30km/h UE speed, 100% UE in a car, 2GHz carrier frequency) and the aspects#B is (3km/h UE speed, 20% outdoor UE, 4GHz carrier frequency)

[Mod] Thanks you very much for your careful reading. According to your comment, I fixed typos.  |
| Huawei, HiSilicon | **[TP3]**Error modelling and phase discontinuity modeling are only applicable for CSI prediction – should be pointed out in the common EVM table.

|  |  |
| --- | --- |
| Channel estimation          | Realistic as a baseline. Up to companies to choose the error modelling method for realistic channel estimation. Error modelling in TR36.897 Table A.1-2 as a baseline if channel estimation error is modelled for CSI prediction.Ideal DL channel estimation is optionally taken into the baseline of evaluation methodology for the purpose of calibration and/or comparing intermediate results (e.g., accuracy of AI/ML output CSI, etc.). Up to companies to report whether/how ideal channel is used in the dataset construction and performance evaluation/inference.Note: Eventual performance comparison with the benchmark release and drawing SI conclusions should be based on realistic DL channel estimation. |
| Phase discontinuity | For CSI prediction: A uniform distribution between within a time window of , where =40 degrees and =20ms can be a baseline. * A fixed phase for all CSI-RS observations within the time window, and another fixed phase for the next time window. The phases are according to uniform distribution.
 |

BTW, there seems a **typo** at TP3 “Baseline for performance evaluation” entry, which is marked with a side note.**[TP6], [TP7], [TP8], and [TP1]**For the titles of 6.2.2.6-A, 6.2.2.7-A, and 6.2.2.8-A, they should be distinguished from the 6.2.2.6, 6.2.2.7, and 6.2.2.8 (R18 study). Otherwise it is not clear the relationship between the sections without “-A” and with “-A”.6.2.2.6-A Basic performance for CSI prediction for the continued studyThe complexity values in terms of FLOPs and number of parameters of AI/ML models adopted in the evaluations of CSI prediction are summarized in Figure 6.2.2.6A-1.…6.2.2.7-A Generalization evaluations for CSI prediction for the continued study***Generalization over UE speeds***…6.2.2.8-A Summary of Performance Results for CSI feedback enhancement for the continued study……Similar, for TP1, the caption of table6-A and 7-A should be corrected.6.2.2 Performance resultsCSI\_Table 1 through CSI\_Table 8 in attached Spreadsheets for CSI feedback enhancement evaluations present the performance results for:……- CSI\_Table 6. Evaluation results for CSI prediction without model generalization/scalability- CSI\_Table 7. Evaluation results for CSI prediction with model generalization- CSI\_Table 6-A. Evaluation results for CSI prediction without model generalization/scalability for the continued study- CSI\_Table 7-A. Evaluation results for CSI prediction with model generalization for the continued study- CSI\_Table 8. Evaluation results for CSI prediction with localized model**[TP9]**Where is this paragraph placed? Should be like the following?8 Conclusions The following aspects have been studied for the general framework of AI/ML over air interface for one-sided models and two-sided models:……**CSI prediction sub use case:** The performance and potential specification impact were studied for AI/ML based UE side CSI prediction sub use case. Evaluations have been performed to assess AI/ML based CSI prediction from various aspects, including performance compared to baseline, model input/output type, generalization over UE speed, etc. Some aspects are studied but lack observations, including scalability over various configurations and generalization over other scenarios and approach of fine tuning. Performance monitoring accuracy has not been evaluated.Performance compared with baseline is summarized in clause 6.2.2.8.Potential specification impact on data collection and performance monitoring are discussed in clause 7.2.2. Limited specification aspects were considered.From RAN1 perspective, there is no consensus on the recommendation of CSI prediction for normative work.The reason for the lack of RAN1 consensus on the recommendation of CSI prediction for normative work:* Lack of results on the performance gain over non-AI/ML based approach and associated complexity.

Other aspects that require further study/conclusion are captured in the summary above. After the continued study of Rel-19, from RAN1 perspective, study of CSI prediction has been completed and performance improvement is observed with increased complexity. |