3GPP TSG RAN WG1 #111 R1-22XXXXX

Toulouse, France, November 14th – 18th, 2022

Source: Moderator (OPPO)

Title: Summary#1 for other aspects on AI/ML for beam management

Agenda Item: 9.2.3.2

Document for: Discussion and Decision

# Introduction

The Rel-18 WID of AI/ML for NR Air Interface focuses on a subset of three typical use cases:

1. CSI feedback enhancement
2. Beam management
3. Positioning accuracy improvement.

This document focuses on the other aspects of AI/ML for beam managements, including representative sub use cases and potential specification impact.

Regarding the file names, companies are encouraged to follow the guidance of R1-2203012 (Page 16) as below:

|  |
| --- |
| * + - To avoid ending-up with too long file names and downloading/opening issues, the following naming convention is recommended:
			* Keep the previous company’s name (only the most recent one) in the filename, e.g.
				+ 5/Summary-1-v000-Moderator (HW)
				+ 5/Summary-1-v001-LG
				+ 5/Summary-1-v002-LG-CATT
				+ 5/Summary-1-v003-CATT-vivo
				+ 5/Summary-1-v004-Moderator(HW)
			* It helps identifying on which previous version your input is based on and solve any crossing emails issue. Note the use of 3digit version numbers in the file names.
 |

In the following sections, the company proposals are summarized, and offline proposals drafted based on company contributions for discussion/input.

**Additional information**

As RAN1#111 is a F2F meeting,

* + Companies can provide inputs in the draft folder at will.
	+ Companies are encouraged to share comments/suggestions (if any) as early as possible by various ways (draft folder, F2F talk, email, …) so that moderator can have enough time to coordinate companies and try to make stable proposals.

# Training and deployment of AI/ML model

## Training/inference at UE/NW side

In previous RAN1 meeting(s), the following agreements were made:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RAN1#109-eAgreement For the sub use case BM-Case1, consider both Alt.1 and Alt.2 for further study:* Alt.1: AI/ML inference at NW side
* Alt.2: AI/ML inference at UE side

Agreement For the sub use case BM-Case2, consider both Alt.1 and Alt.2 for further study:* Alt.1: AI/ML inference at NW side
* Alt.2: AI/ML inference at UE side

RAN1#110Agreement At least for the sub use case BM-Case1 and BM-Case2, support both Alt.1 and Alt.2 for the study of AI/ML model training:* Alt.1: AI/ML model training at NW side;
* Alt.2: AI/ML model training at UE side.

Note: Whether it is online or offline training is a separate discussion.Working AssumptionInclude the following into a working list of terminologies to be used for RAN1 AI/ML air interface SI discussion.

|  |  |
| --- | --- |
| Terminology | Description |
| AI/ML model delivery | A generic term referring to delivery of an AI/ML model from one entity to another entity in any manner.Note: An entity could mean a network node/function (e.g., gNB, LMF, etc.), UE, proprietary server, etc. |

 |

The related proposals/observations from the contributions are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 19: For the sub use case BM-Case1 and BM-Case2, at least support Alt.1 and Alt.2 for AI/ML model training and inference for further study** *Alt.1. AI/ML model training and inference at NW-side*
* *Alt.2. AI/ML model training and inference at UE-side*
* *Note: Alt.3 (AI/ML model training at NW-side, AI/ML model inference at UE-side) can be discussed to be included in the study once its enabling mechanism of model transfer has been agreed in 9.2.1*
 |
| Vivo[3] | *Observation 1: The memory storage requirement in NW side for UE specific models seems unaccepted for Alt.4.**Observation 2: Report overhead may increase dramatically but with less specification impacts for Alt. 1 with enhanced beam pair prediction solution and DL Tx beam prediction solution.**Observation 3: Report overhead can be reduced to top-k L1-RSRP and its related Rx beam information, but assistance information including NW-side information, such as antenna configuration, Tx beam angle, etc., should be signaled to UE for Alt.2.**Observation 4: Due to UE side model training, if mismatch NW-side beam information is signaled to UE, significant performance deterioration can be observed for AI based beam prediction scheme in Alt.2.**Observation 5: Report overhead and UE energy/complexity is limited for Alt.3, but model transfer is needed.**Observation 6: For Alt.3, a cell specific AI solution can be achieved with generalization consideration and infra vendor may not need to disclose NW-side information such as antenna configuration, Tx beam angle, etc.**Proposal 2: For the sub use case BM-Case1 and BM-Case2, at least support to further study Alt.1, Alt.2 and Alt.3 for AI/ML model training and inference:** *Alt.1. AI/ML model training and inference at NW side*
* *Alt.2. AI/ML model training and inference at UE side*
* *Alt.3. AI/ML model training at NW side, AI/ML model inference at UE side*
 |
| CATT[8] | *Proposal 1: For the sub use case BM-Case1 and BM-Case2, support the following alternatives for further study:** *Alt.1: AI/ML training and inference at NW side;*
* *Alt.2: AI/ML training and inference at UE side;*
* *Alt.3: AI/ML training at NW side and inference at UE side.*
 |
| Ericsson[10] | *Proposal 5 RAN1 can provide RAN2 with requirements of the beam prediction AI/ML model to be transferred. It is up to RAN2 to address the feasibility of the transfer.* |
| IDC[11] | *Observation 3: AI/ML inference/training at NW side (Alt.1) could be a good implementation option as UE implementation is generally limited due to computational power and battery consumption than gNB implementation. However, AI/ML inference/training generally requires more detailed explicit information which leads significant reporting overhead.**Observation 4: AI/ML inference/training at UE side (Alt.2) can be limited due to limited computational power and battery consumption at UE implementation, however, UE can easily utilize more information that the UE acquired by measuring SSB/CSI-RS without consuming any reporting overhead.**Proposal 2: Support both AI/ML inference/training at NW side (Alt.1) and UE side (Alt.2) for both BM-Case1 and BM-Case2.* |
| OPPO[14] | *Observation 1: For BM-Case1, deploying AI/ML inference at UE side can avoid beam reporting on Set B, therefore resulting in minimum standard impact.**Observation 2: For BM-Case2, deploying AI/ML inference at UE side seems more reasonable, otherwise (inference at NW side) there could be overwhelming beam reporting on Set B when Set B is the same as Set A.**Proposal 1: For BM-Case1 and BM-Case2, study the following two options**• AI/ML model training and inference at NW side* *• AI/ML model training and inference at UE side.**Observation 3: Whether to study the case that AI/ML model for BM-Case1 and BM-Case2 trained at NW side and delivered to UE side, highly depends on the outcome of model transfer issue discussed in agenda item 9.2.1.* |
| CMCC[20] | *Proposal 1: For BM-Case1 and BM-Case2, at least support Alt.1 and Alt.2 for AI/ML model training and inference for further study:**• Alt.1. AI/ML model training and inference at NW side**• Alt.2. AI/ML model training and inference at UE side**• Regarding whether to support Alt.3 for BM-Case1 and BM-Case2, wait for the conclusion/agreement of Agenda item 9.2.1 on whether to support mode transfer or not** *Alt.3. AI/ML model training at NW side, AI/ML model inference at UE side*
 |
| NVIDIA[21] | *Proposal 3: For the sub use case BM-Case1 and BM-Case2, support both Alt.1 and Alt.2 for the study of AI/ML model training and inference:**• Alt.1: AI/ML model training and inference at network side* *• Alt.2: AI/ML model training and inference at UE side* |

In latest meetings, there were intensive discussions on whether the AI/model training and inference are at the same node or different nodes (e.g., the AI/ML model can be trained by one node and be used for inference by a node of the opposite side), and the alternatives under discussion were as below:

* Alt.1. AI/ML model training and inference at NW side
* Alt.2. AI/ML model training and inference at UE side
* Alt.3. AI/ML model training at NW side, AI/ML model inference at UE side
* Alt.4. AI/ML model training at UE side, AI/ML model inference at gNB side

After several rounds of discussions and revisions, all companies can agree to not support Alt.4. The following proposal was provided, but not agreed

|  |
| --- |
| (RAN1#110bis-e) Proposal 2.1b: For the sub use case BM-Case1 and BM-Case2, at least support Alt.1 and Alt.2 for AI/ML model training and inference for further study:* Alt.1. AI/ML model training and inference at NW side
* Alt.2. AI/ML model training and inference at UE side
* Defer the discussion on Alt.3 for BM-Case1 and BM-Case2 to wait for the conclusion/agreement of Agenda item 9.2.1 on whether to support model transfer for UE-side AI/ML model or not
	+ Alt.3. AI/ML model training at NW side, AI/ML model inference at UE side
 |

Based on tdocs submitted to RAN1#111 meetings, the positions of companies have not changed. In summary, Alt.1 and Alt.2 are supported by all companies and the views are quite controversial for Alt.3. Some companies have strong motivation to support Alt3., but 16 companies don’t support Alt.3.

Proposal 2.1b of RAN1#110bis-e was supported by all companies except vivo and Apple. Meanwhile, we should have some focused scope for the further study based on whatever we have achieved so far. A new version of the proposal is proposed for discussion:

* The 2nd sub-bullet of the 3rd bullet is added based on the comment and request of vivo in RAN1#110bis-e
* Since many companies have strong concern for Alt.3, it would be good to provide the specific requirements to check whether model transfer is feasible and beneficial for BM cases. Thus, the 3rd sub-bullet of the 3rd bullet is added based on Ericsson’s suggestion.

###### Proposal 2.1

***Proposal 2.1: For the sub use case BM-Case1 and BM-Case2, at least support Alt.1 and Alt.2 for AI/ML model training and inference for further study:***

* ***Alt.1. AI/ML model training and inference at NW side***
* ***Alt.2. AI/ML model training and inference at UE side***
* ***Defer the discussion on whether Alt.3 for BM-Case1 and BM-Case2 is supported or not to wait for the conclusion/agreement of Agenda item 9.2.1 of RAN1 and/or RAN2 on whether to support model transfer for UE-side AI/ML model or not***
	+ ***Alt.3. AI/ML model training at NW side, AI/ML model inference at UE side***
	+ ***It is up to companies whether or not to submit evaluation results to agenda item 9.2.3.1***
	+ ***RAN1 can provide RAN2 with the specific requirements of the mode transfer for BM-Case1 and BM-Case2 if any consensus can be achieved further.***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Online/offline training

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **RAN1#110**Working Assumption

|  |  |
| --- | --- |
| Terminology | Description |
| Online training | An AI/ML training process where the model being used for inference) is (typically continuously) trained in (near) real-time with the arrival of new training samples. Note: the notion of (near) real-time vs. non real-time is context-dependent and is relative to the inference time-scale.Note: This definition only serves as a guidance. There may be cases that may not exactly conform to this definition but could still be categorized as online training by commonly accepted conventions.Note: Fine-tuning/re-training may be done via online or offline training. (This note could be removed when we define the term fine-tuning.) |
| Offline training | An AI/ML training process where the model is trained based on collected dataset, and where the trained model is later used or delivered for inference.Note: This definition only serves as a guidance. There may be cases that may not exactly conform to this definition but could still be categorized as offline training by commonly accepted conventions. |

Note: It is encouraged for the 3gpp discussion to proceed without waiting for online/offline training terminologies. |

The related proposals/observations from the contributions are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 1: For the sub use case BM-Case1 and BM-Case2, study** *whether there is spec impact for offline training and online training.*
* *the associated spec impact of each training type if exists*

*FFS: whether online training is supported (being discussed under 9.2.1)* |
| Huawei[2] | *Observation 8: For NW-side model, model training under online/offline manner is up to implementation.**Proposal 20: If an online/offline discussion shall be conducted for the UE-side model, this discussion should be kept separated from the issue whether data set collection is via air-interface or non-air-interface.* |
| OPPO[14] | *Proposal 2: For AI/ML beam management, focus on offline model training at least at current stage.* |
| NVIDIA[21] | *Observation 1: Offline training may be more feasible for the near future. But in the long run, it is vital that the AI/ML models can learn continuously to adapt to varying environments, site-specific conditions, and heterogenous configurations.**Proposal 2: For the sub use case BM-Case1 and BM-Case2, support both Alt.1 and Alt.2 for the study of AI/ML model training:**• Alt.1: offline training**• Alt.2: online training* |
| QC[28] | *Proposal 1: For the sub use case BM-Case1 and BM-Case2 and for UE-side AI/ML models, Agenda item 9.2.3.2 should focus on offline training scenario, in which the development and training of the AI/ML model happens offline without the need to involve 3gpp signaling.* |
| NEC[34] | *Proposal 5: For the trained AI/ML model in offline, study the mechanism of model update (e.g., fine-tuning) based on the online data.* |
|  |  |

Based on previous discussions and the tdocs submitted to this meeting, all companies support offline training. The controversial point is whether to support online training or not**.** Meanwhile, some companies think whether online and offline training are up to implementation and can be transparent from the perspective of 3GPP specification. By considering the afore-mentioned information, moderator feels that it would be more productive to identify the spec impact of online training different from offline training for BM cases as the first step. Thus, the following is proposed for discussion

###### Proposal 2.2

***Proposal 2.2: For the sub use case BM-Case1 and BM-Case2,***

* ***study the potential spec impact of AI/ML offline training***
* ***study and identify (if any) the spec impact of AI/ML online training different from offline training***
	+ ***Note: Support/de-prioritization of online training is discussed in Agenda item 9.2.1***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Sub use cases of BM-Case1 and BM-Case2

## General views

In previous meetings, the following agreements/conclusion were made as below:

|  |
| --- |
| **RAN1#109-e**AgreementFor AI/ML-based beam management, support BM-Case1 and BM-Case2 for characterization and baseline performance evaluations* BM-Case1: Spatial-domain DL beam prediction for Set A of beams based on measurement results of Set B of beams
* BM-Case2: Temporal DL beam prediction for Set A of beams based on the historic measurement results of Set B of beams
* FFS: details of BM-Case1 and BM-Case2
* FFS: other sub use cases

Note: For BM-Case1 and BM-Case2, Beams in Set A and Set B can be in the same Frequency Range |

The related proposals/observations from the contributions are copied as below:

|  |  |
| --- | --- |
| IDC[11] | *Observation 1: The agreements made in RAN1#109 and RAN1#110bis-e do not preclude the case that beams in Set A and Set B in different frequency ranges.**Observation 2: Supporting Set A and Set B in different frequency ranges is beneficial considering different beamwidths especially when multiple cells in different Frequency Ranges are implemented toward to an identical direction.**Proposal 1: BM-Case 1 and BM-Case 2 with Set A and Set B in different frequency ranges are supported as well as in a same frequency range.* |
| Xiaomi[12] | *Proposal 1: For BM-Case2, support the periodicity of future time instance is same or 1/N of measurement/report instance* |
| Sony[19] | *Proposal 4 : the time window size of N and F in BM-Case2 can be determined by properties of time domain channel, e.g. coherence time, etc.* |
| NVIDIA[21] | *Proposal 1: Beam prediction in spatial domain and beam prediction in time domain should be the focal point for studying AI/ML based algorithms for beam management.* |
| DCM[26] | *Proposal 10: Study two-stage beam measurements with top-N predicted beams, since it reduces RS measurement overhead and increases the reliability of beam selection compared to top-1 beam prediction.* |
| Rakuten[32] | *Proposal 1: Consider a two-step beam management procedure where existing beam management mechanism is used to choose the best beam from a set of beam recommendations from the AI/ML model* |
| Nokia[33] | *Proposal 20: For UE side DL Tx beam prediction with inter-cell beam measurements and reporting, RAN1 shall further study the feasibility of applying beam predictions (BM-Case1 and BM-Case2) across different PCIs or within one PCI.* *Proposal 21: For UE side DL Tx beam prediction, Ran1 shall further study group-based beam reporting supported for mTRP operation, including whether Set B measurements can be from two TRPs and UE can report beam pairs from Set A.*  |
|  |  |

**Mod recommendation**: TBD

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |

## Type of beam prediction

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110**Agreement For the sub use case BM-Case1 and BM-Case2, further study the following alternatives for the predicted beams:* Alt.1: DL Tx beam prediction
* Alt.2: DL Rx beam prediction
* Alt.3: Beam pair prediction (a beam pair consists of a DL Tx beam and a corresponding DL Rx beam)
* Note1: DL Rx beam prediction may or may not have spec impact
 |

The related proposals/observations from the contributions are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 2: For the sub use case BM-Case1 and BM-Case2, study both Alt.1 and Alt.3 for the predicted beams without down-selection.** *Note: Alt.1 and Alt.3 were agreed in RAN1#110 meeting as below*
	+ - *Alt.1: DL Tx beam prediction*
		- *Alt.3: Beam pair prediction (a beam pair consists of a DL Tx beam and a corresponding DL Rx beam)*
 |
| Huawei[2] | *Observation 6: For the beam prediction mechanisms for BM-Case1 and BM-Case2, the performance gain of Alt.2 (DL Rx beam prediction) may be limited due to a relatively small number of wide Rx beams at UE.**Observation 7: For the beam prediction mechanisms for BM-Case1 and BM-Case2, Alt.3 (beam pair prediction) needs the NW to be aware of the numbers/patterns of Rx beams and needs the UE to be aware of the numbers/patterns of Tx beams.**Proposal 17: For the beam prediction mechanisms for BM-Case1 and BM-Case2, consider Alt.1 (DL Tx beam prediction) as a starting point due to its simplicity and flexibility.**Proposal 18: For the beam prediction mechanisms for BM-Case1 and BM-Case2, if Alt.3 (beam pair prediction) is to be further studied, it should be studied for both NW-side AI/ML model and UE-side AI/ML model with the same priority.* |
| Vivo[3] | *Proposal 1: Study the two AI-based beam prediction solutions for both BM-Case1 and BM-Case2, i.e. enhanced beam pair prediction scheme and DL Tx beam prediction scheme, and considering specification impacts with generalization aspects, such as Set B construction, supported number of Tx/Rx beams, various number of antenna configurations, etc.* |
| ZTE[5] | *Proposal 1: Both DL Tx beam prediction and beam pair prediction can be supported without any further down-selection.**Observation 1: The Tx beam prediction can be achieved by initiating P2 beam sweeping procedure for beam measurement and data collection.**Observation 2: Since the UE Rx beam is up to implementation in current specification, the P1 beam sweeping procedure is still conceptual and there is no explicit signaling/configuration for P1 in current specification.**Proposal 2: To facilitate the beam pair prediction, P1 may need to be specified clearly in the spec with potentially enhanced RS resource set configuration and reporting mechanism.**Proposal 3: For the NW-side beam pair prediction, it may be desirable to implicitly indicate the Rx beam ID to facilitate data collection at the NW side and avoid UE proprietary information disclosure issue.**Observation 3: For the UE-side beam pair prediction, it brings less standardization work and does not involve sensitive proprietary information disclosure issues.* |
| CATT[8] | *Proposal 5: For the sub use case BM-Case1 and BM-Case2, focus on Alt.1 and Alt.3 for further study.** *Alt.1: DL Tx beam prediction;*
* *Alt.3: Beam pair prediction.*
 |
| Spreadtrum[9] | *Proposal 4: For sub use cases BM-Case1 and BM-Case2, support Alt3 Beam pair prediction as baseline.* |
| IDC[11] | *Observation 16: For Rel-15 beam management, actual mapping between DL Tx beam and UE Rx beam is totally based on UE implementation.**Observation 17: The implementation-based UE Rx beam selection works for Rel-15, however, UE Rx beam information is crucial to accurately predict beam qualities for AI/ML based beam prediction.* *Proposal 17: Study benefits of specification enhancements on acquiring UE Rx beam information for DL Tx beam prediction (Alt. 1) and beam pair prediction (Alt. 3).**Proposal 18: DL Rx beam prediction (Alt. 2) should be a part of UE implementation.* |
| Intel[13] | *Proposal 1: Beam Pair prediction (Alt-3) should be supported, at least for BM-Case 1 since it can provide large latency and measurement gains for joint P2/P3 procedure* |
| OPPO[14] | *Proposal 3: For BM-Case1 and BM-Case2, at least support beam pair prediction (Alt.3) as the key feature of representative sub use cases.* |
| CT[16] | *Proposal 1: We slightly prefer Alt.1, and Alt.3 can also be considered.**Proposal 2: Design procedure and signalling for joint prediction in Alt.3, and study the potential spec. impact.* |
| CIACT[25] | *Proposal 1: For BM-Case1 and BM-Case2, DL tx beam prediction and Beam pair predictions should be supported.* |
| SS[27] | *Proposal 10: For predicted beams, Alt 1 (DL Tx beam prediction) is preferred.* |
| Nokia[33] | *Proposal 3: For BM-Case1, considering beam types of Set A/B, prioritize Alt.1: DL Tx beam prediction for further study.**• RAN1 may consider Alt.2: DL Rx beam prediction and Alt.3: Beam pair prediction as an additional scenario if the benefits are identified in 9.2.3.1.**Proposal 6: For BM-Case2 construction of Set A/B, prioritize Alt.1: DL Tx beam prediction for further study.**• RAN1 may consider Alt.2: DL Rx beam prediction and Alt.3: Beam pair prediction as an additional scenario if the benefits are identified in 9.3.2.1.*  |
|  |  |

In the last meeting, most companies can live with the proposal to focus on Alt.1 and Alt.3. One company requested to add “potential down-selection” in the proposal whereas 7 companies suggested to remove it. As it is in the 1st half of the study item, we can keep open for more study if most companies want to. We may have more better understanding whether any further down-selection is needed or not when we have more progress. Thus, the potential down-selection can be discussed later. Accordingly, the following proposal is suggested for discussion.

###### Proposal 3.2

***Proposal 3.2: For the sub use case BM-Case1 and BM-Case2, focus on Alt.1 and Alt.3 for the predicted beams for further study***

* ***Note: Alt.1 and Alt.3 were agreed in RAN1#110 meeting as below***
	+ ***Alt.1: DL Tx beam prediction***
	+ ***Alt.3: Beam pair prediction (a beam pair consists of a DL Tx beam and a corresponding DL Rx beam)***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Construction of Set A and Set B

In previous RAN1 meeting(s), the following agreements and conclusions were made:

|  |
| --- |
| **RAN1#109-e**ConclusionFor the sub use case BM-Case1, consider the following alternatives for further study:* Alt.1: Set B is a subset of Set A
	+ FFS: the number of beams in Set A and B
	+ FFS: how to determine Set B out of the beams in Set A (e.g., fixed pattern, random pattern, …)
* Alt.2: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)
	+ FFS: the number of beams in Set A and B
	+ FFS: QCL relation between beams in Set A and beams in Set B
	+ ~~FFS: construction of Set B (e.g., regular pre-defined codebook, codebook other than regular pre-defined one)~~
* Note1: Set A is for DL beam prediction and Set B is for DL beam measurement.
* Note2: The narrow and wide beam terminology is for SI discussion only and have no specification impact
* Note3: The codebook constructions of Set A and Set B can be clarified by the companies.

ConclusionFor the sub use case BM-Case2, further study the following alternatives with potential down-selection:* Alt.1: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)
	+ FFS: QCL relation between beams in Set A and beams in Set B
* Alt.2: Set B is a subset of Set A (Set A and Set B are not the same)
	+ FFS: how to determine Set B out of the beams in Set A (e.g., fixed pattern, random pattern, …)
* Alt.3: Set A and Set B are the same
* Note1: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.
* Note2: It is up to companies to provide other alternative(s)
* Note3: The narrow and wide beam terminology is for SI discussion only and have no specification impact

**RAN1#110**Agreement For the sub use case BM-Case1, support the following alternatives for further study:* Alt.1: Set A and Set B are different (Set B is NOT a subset of Set A)
* Alt.2: Set B is a subset of Set A
* Note1: Set A is for DL beam prediction and Set B is for DL beam measurement.
* Note2: The beam patterns of Set A and Set B can be clarified by the companies.

AgreementFor the sub use case BM-Case2, further study the following alternatives:* Alt.1: Set A and Set B are different (Set B is NOT a subset of Set A)
* Alt.2: Set B is a subset of Set A (Set A and Set B are not the same)
* Alt.3: Set A and Set B are the same
* Note1: The beam pattern of Set A and Set B can be clarified by the companies.
 |

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Observation 2: For the alternatives of the Set A and Set B relationship under BM-Case 2, Alt.3 (Set A and Set B are the same)** *Can inflict compatibility issues with non-AI/ML-based UEs*
* *Results into a large beam sweeping overhead during the observation phase*
* *May cause unnecessary high interference to cells from neighbor UEs.*

*Proposal 12: For the study of the alternatives of the Set A and Set B relationship under BM-Case 2,** *Prioritize the study of Alt.1 (Set A and Set B are different) and Alt.2 (Set B is a subset of Set A).*
* *Alt.3 (Set A and Set B are the same) can be used as a benchmark for performance comparison in evaluations.*
 |
| ZTE[5] | *Observation 4: The number of beams for measurement (i.e., set B) and for prediction (i.e., set A) is related to the trade-off between inference performance and RS overhead for beam measurement.* *Proposal 4: For BM-Case1, the sub-sampling based method in Alt.2 can be a starting point for the study of spatial domain beam prediction.**Proposal 5: For BM-Case2, both Alt.2 and Alt.3 can be further evaluated in agenda 9.2.3.1, including beam prediction accuracy, RS overhead reduction, and model generalization performance in various scenarios/configurations.* |
| IDC[11] | *Observation 5: As using same beamwidth for all channels and signals is a general implementation* *within a frequency range, using a subset of Set A as Set B is a reasonable option if Set A and Set B are utilized in a same frequency range.* *Observation 6: It is difficult to use a subset of Set A considering different beamwidths for beam management between different frequency ranges.**Observation 7: Utilization of wide beam information from a low frequency range has great potential as a low frequency range is more reliable and utilization of wide beam requires much less time and frequency resources for beam management.**Proposal 3: Support ‘Set B is a subset of Set A’ when Set A and Set B are utilized in a same frequency range for both BM-Case1 and BM-Case2.* *Proposal 4: Support ‘Set A and Set B are different’ when Set A and Set B are utilized in different frequency ranges for both BM-Case1 and BM-Case2.* *Proposal 5: AI/ML based beam management based on association between different frequency ranges should supported for both between FR1 and FR2-1 and between FR2-1 and FR2-2.* |
| Xiaomi[12] | *Proposal 2: For spatial domain beam prediction, consider set B is a subset of set A with high priority.**Proposal 10: For temporal beam prediction, consider set B is same as set A with high priority.* |
| OPPO[14] | *Proposal 4: For BM-Case1, Set B is a subset of Set A.**Proposal 5: For BM-Case2, Set B and Set A are the same.* |
| CT[16] | *Proposal 3: At least Alt.2 should be the baseline for performance evaluation of BM-Case1.**Mod’s note: Alt.1 seems a typo. Please check whether the modification is correct or not* |
| Sony[19] | *Proposal 1 : Support Set A and Set B are different, Set B with wide beams and Set A with narrow beams.* |
| LG[24] | *Proposal #7: For the relation between Set A and Set B of BM-Case2, start from Alt3 to see the feasibility and performance gain of pure TD prediction as an independent approach as SD prediction. After studying this, joint SD and TD prediction (i.e. Alt1 and Alt2) can be studied as a next step* |
| CIACT[25] | *Proposal 2: Conclusions on Set A and Set B could be deferred till further down-selection on use cases and simulation results from 9.2.3.1* |
| SS[27] | *Proposal 1: For BM-Case1 with a NW AI/ML model, consider to define Set B’ for the beams measured by UE.** Set B consists of the beams reported by UE from Set B’.**Proposal 5: For BM-Case2 with a NW AI/ML model, consider to define Set B’ for the beams measured by UE.** Set B consists of the beams reported by UE from Set B’.* |
| Nokia[33] | *Proposal 1: For BM-Case1, considering the construction of Set A/B, prioritize Alt.2: Set B is a subset of Set A for further studies.**• RAN1 may consider Alt1: Set A and Set B are different (Set B is NOT a subset of Set A) as an additional scenario if the benefits are identified in 9.3.2.1.* *Proposal 4: In BM-Case2, “Set B and Set A are the same” should be the baseline to study the prediction performance.**• FFS relation between K and F with different UE speeds, different channel assumptions, and different measurement periods.**Proposal 5: In BM-Case2, prioritize studying “Alt.3 Set B and Set A are the same” and “Alt.2 Set B is a subset of Set A”**• FFS use cases of Alt.1 Set B and Set A are different.* |

**Mod recommendation**: Potential down-selection (if any) will be discussed when we have more progress (e.g., observations of evaluation results, different spec impacts, …). Let’s focus on the study on the potential spec impact of different alternatives in other sections.

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |

## Set B

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110**Agreement* Study the following options on the selection of Set B of beams (pairs)
	+ Option 1: Set B is fixed across training and inference
		- FFS on the beams of Set B
	+ Option 2: Set B is variable (e.g., different beams (pairs) patterns in each report/measurement during training and/or inference)
		- FFS on fixed or variable number of beams (pairs)
		- FFS on the details
	+ Other options are not precluded.
	+ FFS on the number of beams (pairs) in Set B

**RAN1#110bis-e**Agreement* Study the following options on the selection of Set B of beams (pairs)
	+ Option 1: Set B is fixed across training and inference
	+ Option 2: Set B is variable (e.g., different beams (pairs) patterns in each time instance/report/measurement during training and/or inference), FFS:
		- Opt A: Set B is changed following a set of pre-configured patterns
		- Opt B: Set B is randomly changed among pre-configured patterns
		- Opt C: Set B is randomly changed among Set A beams (pairs)
		- The number of beams(pairs) in Set B can be fixed or variable
		- Note: BM-Case1 and BM-Case2 may be considered for different option.
	+ Other options are not precluded.

Conclusion For the sub use case BM-Case1 and BM-Case2, Set B is a set of beams whose measurements are taken as inputs of the AI/ML model, |

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| Vivo[3] | *Proposal 4: Support to study specification impact on Set B selection with semi-random beam subset selection scheme which can provided comparable gain to fixed scheme but with higher flexibility and better generalization performance.* |
| Spreadtrum[9] | *Proposal 1: For Alt.2 (Set B is a subset of Set A) of sub use cases BM-Case1,* * *If AI/ML inference is at NW side, beams in Set B can be determined by NW implementation.*
* *If AI/ML inference is at UE side, beams in Set B can be determined with a fix pattern.*
 |
| TCL[15] | *Proposal 3: Some patterns can be designed for the input set B of beam prediction in spatial domain.** *A fixed pattern;*
* *A random pattern.*
 |
| Lenovo[22] | *Proposal 1: Selection of beams for Set B should allow for variable beams, i.e., different beams (pairs) patterns during training and/or inference.* |
|  |  |

**Moderator recommendation**: Let’s focus on the study on the potential spec impact of different alternatives in other sections. Any further down-selection (if any) can be discussed once we have more progress (e.g., observations based on evaluation results, spec impacts, …)

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |

## Input of BM-Case1 and BM-Case2

In previous RAN1 meeting(s), the agreements/conclusions were made as below:

|  |
| --- |
| **RAN1#109-e**ConclusionRegarding the sub use case BM-Case1, further study the following alternatives for AI/ML input:* Alt.1: Only L1-RSRP measurement based on Set B
* Alt.2: L1-RSRP measurement based on Set B and assistance information
	+ FFS: Assistance information. The following were mentioned by companions in the discussion:  Tx and/or Rx beam shape information (e.g., Tx and/or Rx beam pattern, Tx and/or Rx beam boresight direction (azimuth and elevation), 3dB beamwidth, etc.), expected Tx and/or Rx beam for the prediction (e.g., expected Tx and/or Rx angle, Tx and/or Rx beam ID for the prediction), UE position information, UE direction information, Tx beam usage information, UE orientation information, etc.
		- Note: The provision of assistance information may be infeasible due to the concern of disclosing proprietary information to the other side.
* Alt.3: CIR based on Set B
* Alt.4: L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID
* Note1: It is up to companies to provide other alternative(s) including the combination of some alternatives
* Note2: All the inputs are “nominal” and only for discussion purpose.

ConclusionRegarding the sub use case BM-Case2, further study the following alternatives of measurement results for AI/ML input (for each past measurement instance):* Alt.1: Only L1-RSRP measurement based on Set B
* Alt 2: L1-RSRP measurement based on Set B and assistance information
	+ FFS: Assistance information. The following were mentioned by companies in the discussion:, Tx and/or Rx beam angle, position information, UE direction information, positioning-related measurement (such as Multi-RTT), expected Tx and/or Rx beam/occasion for the prediction (e.g., expected Tx and/or Rx beam angle for the prediction, expected occasions of the prediction), Tx and/or Rx beam shape information (e.g., Tx and/or Rx beam pattern, Tx and/or Rx beam boresight directions (azimuth and elevation), 3dB beamwidth, etc.) , increase ratio of L1-RSRP for best N beams, UE orientation information
		- Note: The provision of assistance information may be infeasible due to the concern of disclosing proprietary information to the other side.
* Alt.3: L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID
* Note1: It is up to companies to provide other alternative(s) including the combination of some alternatives
* Note2: All the inputs are “nominal” and only for discussion purpose.
 |

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 1: For the BM-Case 1 study of the AI/ML model input, study Alt.1 (Only L1-RSRP for Set B) as a starting point.**Proposal 4: For the BM-Case 1 study of the AI/ML model input, Alt.4 (L1-RSRP for Set B and DL Tx and/or Rx beam ID) can be studied if benefits are justified by evaluation.**Proposal 5: For the study of AI/ML model input for BM-Case 1 and BM-Case 2, consider a fixed beam as a starting point.* *Proposal 9: For the BM-Case 2 study of the AI/ML model input, study Alt.1 (Only L1-RSRP for Set B) as a starting point.* *Proposal 11: For the BM-Case 2 study of the AI/ML model input, Alt.3 (L1-RSRP for Set B and DL Tx and/or Rx beam ID) can be studied if benefits are justified by evaluation.* |
| Vivo[3] | *Proposal 3: Regarding to BM-Case1 and BM-Case 2, at least prioritize following AI input information for further study on specification impact:** L1-RSPR measurement based on Set B** Corresponding DL Tx beam pointing angle/ID** Corresponding DL Rx beam pointing angle/ID** Expected Tx and/or expected Rx beam angle/ID** Further discuss other information, such as Tx and/or Rx beam shape information, 3dB beam-width, etc.* |
| Fujitsu[6] | *Observation 1: For Alt. 1, the beam IDs are implicitly included with the measured L1-RSRPs on Set B constructed by fixed or a set of pre-configured patterns, and may not be included with the measured L1-RSRP on Set B randomly selected from Set A.**Proposal 1: Regarding the sub use case BM-Case1, the measured L1-RSRPs on Set B and corresponding Tx and/or Rx beam IDs are suggested as baseline for the study of AI/ML input.* |
| Google[7] | *Proposal 1: For spatial domain beam prediction, support Alt3 (CIR based on set B).**Proposal 6: For time-domain beam prediction, support to add CIR measurement based on set B as one alternative.* |
| CATT[8] | *Proposal 4: For the sub use case BM-Case1 and BM-Case2, study the following alternatives for AI/ML input:** *Alt.1: Only L1-RSRP measurement based on Set B;*
* *Alt.2: L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID;*
* *Alt.3: L1-RSRP measurement based on Set B and assistance information.*
	+ *Whether and how the assistance information can be used as model inputs should be discussed for UE-sided model and NW-sided model, separately.*
 |
| Sptreadtrum[9] | *Proposal 2: For the AI/ML model input of BM-Case 1,** *Alt 1 and Alt 4 should be studied with priority.*
* *Whether to choose Alt 1 or Alt 4 needs further discussion according to the beam pattern selection.*
	+ *If Set B is fixed, Alt 1 will be selected;*
	+ *If Set B is variable, Alt 4 will be selected.*
* *For the corresponding DL Tx and/or Rx beam ID in Alt 4, it should be input in AI model explicitly.*
* *Alt 2 should be clarified which assistance information can be used as AI model input.*

*Proposal 3: For the AI/ML model input of BM-Case 2,** *Alt 1 and Alt 3 should be studied with priority.*
* *Whether to choose Alt 1 or Alt 3 needs further discussion according to the beam pattern selection.*
	+ *If Set B is fixed, Alt 1 will be selected;*
	+ *If Set B is variable, Alt 3 will be selected.*
* *For the corresponding DL Tx and/or Rx beam ID in Alt 3, it should be input in AI model explicitly.*
* *Alt 2 should be clarified which assistance information can be used as AI model input.*
 |
| IDC[11] | *Observation 8: ‘Only L1-RSRP measurement based on Set B’ is not clear enough as the alternative does not provide any beam related information.** *If ‘Only L1-RSRP measurement based on Set B’ means that L1-RSRP measurements are provided in a fixed order, in our view, the input is not ‘Only L1-RSRP measurement based on Set B’.*
* *Reporting L1-RSRP measurements in a fixed order is indicating L1-RSRP measurement with implicit beam related information.*

*Proposal 6: Companies supporting the alternative should provide more details for predicting L1-RSRP values without any beam information.**Observation 9: ‘L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID’ can be a baseline option as AI/ML model can predict RSRP measurements with Tx and Rx beam IDs which are not provided.**Proposal 7: Support ‘L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID’ as a baseline.**Proposal 8: Additional information such as TRP IDs and Panels IDs should be considered.**Proposal 9: ‘CIR based on Set B’ can be considered as an alternative only for beam management based on FR1 information.* |
| Xiaomi[12] | *Proposal 3: Support L1-RSRP and beam (pair) ID as AI/ML model input with high priority for variable set B.* |
| OPPO[14] | *Proposal 6: For BM-Case1 and BM-Case2, at least support Tx-Rx beam pair prediction (Alt.3) as the key feature of representative sub use cases.**Proposal 7: The input of AI/ML model for beam prediction are element-wise sensitive, therefore the L1-RSRPs of Tx and/or Rx beams in Set B should be input in proper order.**Proposal 8: For BM-Case1 and BM-Case2, suggest to adopt L1-RSRP measurement based on Set B as input of AI/ML model.* |
| CT[16] | *Observation 1: How to guarantee the quality of assistance information is a critical step for the conduction of Alt. 2.**Observation 2: The choice of assistance information including number and type should be well balanced so as to obtain an adequate trade-off between prediction performance and overhead.**Proposal 4: For BM-Case1 and BM-Case2, suggest to give priority to only L1-RSRP measurement based on Set B. L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID can be as a supplement for input when random patterns are selected.* |
| Sony[19] | *Proposal 2 : Support Set A and Set B in different frequency bandwidth and channel observation as the input of AI/ML model (e.g., channel information matrix by estimating, CIR, etc).* |
| NVIDIA[21] | *Observation 2: Evaluation results show that by using L1-RSRP measurement based on Set B of beams, the AI/ML-based algorithm can achieve performance comparable to that of exhaustive beam search in Set A of beams.**Proposal 4: For BM-Case 1, at least support L1-RSRP measurement based on Set B of beams as AI/ML model input.**Observation 3: Evaluation results show that by using historical optimal index, the AI/ML-based algorithm can satisfactorily yield optimal beam index prediction for future time instances.**Proposal 6: For BM-Case 2 (temporal DL beam prediction), at least support using historical optimal beam index based on Set B of beams as AI/ML model input.* |
| LG[24] | *Proposal #1: For the UE AI/ML input, Alt2 can be considered. For the assist information for input, output, training, and inference, relation between Set A beams and Set B beams should be provided to UE. In this regard, Set A beams can be represented as a linear combination of Set B beams or Set A/B beams can be represented on a 2D/3D coordinate.**Proposal #2: It should be assumed that measurements of Set A beams are not available or is available but can be outdated at UE side, and UE is not expected to manage a good Rx beam for each of the Set A beams.**Proposal #3: Consider UE assistance/reporting for determining Set A, e.g. UE to report preferred Set A among candidate beams of Set A.* |
| KT[29] | *Proposal 1. For BM-Case1 and BM-Case2, support L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID as a baseline of AI/ML input.* |
| Nokia[33] | *Proposal 7: Regarding the sub-use case BM-Case1, select the following alternatives for AI/ML input:** *Alt.2: L1-RSRP measurement based on Set B and assistance information*
	+ *FFS: Assistance information. The following were mentioned by companions in the discussion:  Tx and/or Rx beam shape information (e.g., Tx and/or Rx beam pattern, Tx and/or Rx beam boresight direction (azimuth and elevation), 3dB beamwidth, etc.), expected Tx and/or Rx beam for the prediction (e.g., expected Tx and/or Rx angle, Tx and/or Rx beam ID for the prediction), UE position information, UE direction information, Tx beam usage information, UE orientation information, gNB panel array parameters (bearing angle, mechanical downtilt, slant angle), etc.*

*Proposal 8: Regarding the sub-use case BM-Case2, select the following alternatives for AI/ML input:** *Alt 2: L1-RSRP measurement based on Set B and assistance information*
	+ *FFS: Assistance information. The following were mentioned by companies in the discussion:, Tx and/or Rx beam angle, position information, UE direction information, positioning-related measurement (such as Multi-RTT), expected Tx and/or Rx beam/occasion for the prediction (e.g., expected Tx and/or Rx beam angle for the prediction, expected occasions of the prediction), Tx and/or Rx  beam shape information (e.g., Tx and/or Rx beam pattern, Tx and/or Rx beam pointing angles beam boresight directions (azimuth and elevation), 3dB beamwidth, etc.) , increase ratio of L1-RSRP for best N beams, UE orientation information*
 |
| NEC[34] | *Proposal 1: Support Alt.2: L1-RSRP measurement based on Set B and assistance information (e.g., Rx-beam angle, UE’s angle related to a reference location).* |
|  |  |

Different companies have different preferences on the alternatives of AI model inputs. We can wait for more progress and then discuss whether any down-selection is needed or not if needed. Let’s focus on the study on the potential spec impact of different alternatives in other sections.

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |

## Output of BM-Case1 and BM-Case2

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) are made as below:

|  |
| --- |
| **RAN1#110**AgreementRegarding the sub use case BM-Case1 and BM-Case2, study the following alternatives for AI/ML output:* Alt.1: Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams
	+ E.g., N predicted beams can be the top-N predicted beams
* Alt.2: Tx and/or Rx Beam ID(s) of the N predicted DL Tx and/or Rx beams and other information
	+ FFS: other information (e.g., probability for the beam to be the best beam, the associated confidence, beam application time/dwelling time, Predicted Beam failure)
	+ E.g., N predicted beams can be the top-N predicted beams
* Alt.3: Tx and/or Rx Beam angle(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams
	+ E.g., N predicted beams can be the top-N predicted beams
	+ FFS: details of Beam angle(s)
* FFS: how to select the N DL Tx and/or Rx beams (e.g., L1-RSRP higher than a threshold, a sum probability of being the best beams higher than a threshold, RSRP corresponding to the expected Tx and/or Rx beam direction(s))
* Note1: It is up to companies to provide other alternative(s)
* Note2: Beam ID is only used for discussion purpose
* Note3: All the outputs are “nominal” and only for discussion purpose
* Note4: Values of N is up to each company.
* Note5: All of the outputs in the above alternatives may vary based on whether the AI/ML model inference is at UE side or gNB side.
* Note 6: The Top-N beam IDs might have been derived via post-processing of the ML-model output
 |

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Observation 2: Model outputs are typically used internally and hence without standards impact. Therefore, unless there are standards impacts involved, model outputs don’t need to be explicitly specified in the standards.* *Proposal 5: Specify model outputs only when standards impact is involved while companies are encouraged to share their model output for AI/ML based beam management.* |
| Huawei[2] | *Observation 4: For the alternatives for AI/ML output for BM-Case 1 and BM-Case 2,* * *Alt.1 is straightforward and can achieve significant performance gain without other output information.*
* *Alt.2 (beam ID and other information) has too many sub-options and for its further study a down-selection within Alt.2 is necessary.*
* *Alt.3 (beam angle and RSRP) can be seen as a further sub-option of Alt.2. Before studying output options of Alt.2, more details on their usage and their potential benefits are necessary.*

*Proposal 16: For BM-Case1 and BM-Case2, consider Alt. 1 as the baseline for the assumption on the AI/ML model output:** *Alt.1: Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams*
	+ *E.g., N predicted beams can be the Top-N predicted beams*
 |
| Vivo[3] | *Proposal 6: Support to prioritize following AI output for further study on specification impact:** Tx and/or Rx Beam ID(s)/angle(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams.** The N predicted Tx/Rx beams can be produced according to the expected beam information input to the AI model** FFS: study global beam ID or local beam ID** FFS: study global beam information, e.g. global beam ID or beam angle, with minimum exposures of implementation details**Proposal 7: Suggest to deprioritize Alt.2, i.e. Tx and/or Rx Beam ID(s) of the N predicted DL Tx and/or Rx beams and other information, for further study specification impact.* |
| ZTE[5] | *Proposal 8: For BM-Case1 and BM-Case2, focusing the AI input and output on measured RSRP and/or beam ID is a good starting point, in which case the standardization workload and AI model complexity would be relatively low.* |
| Fujitsu[6] | *Observation 3: The NW-side model does not need to study the output of AI/ML model since there are no specification impacts.* *Proposal 4: Regarding the sub use case BM-Case1 and BM-Case2, the following outputs of UE-side model are suggested as baseline:** *Tx and/or Rx Beam ID(s)*
* *The predicted L1-RSRP or associated confidence of the N predicted DL Tx and/or Rx beams*
 |
| Google[7] | *Proposal 2: For spatial domain beam prediction, support the best beam possibility for each beam in Set A as the output.* *Proposal 3: When AI/ML model is implemented in the NW side, the output for the AI/ML for spatial domain beam prediction with spec impact should be the reference angle for DL Rx beam refinement (Alt3).**Proposal 4: When AI/ML model is implemented in the UE side, the output for the AI/ML model for spatial domain beam prediction with spec impact should be the reference angle for DL Tx beam refinement (Alt3).**Proposal 7: For time-domain beam prediction, support the best beam possibility for each beam in Set A as the output.* *Proposal 8: When AI/ML model is implemented in the NW side, the output for the AI/ML for time domain beam prediction with spec impact should be the reference angle for DL Rx beam refinement (Alt3).**Proposal 9: When AI/ML model is implemented in the UE side, the output for the AI/ML model for time domain beam prediction with spec impact should be the reference angle for DL Tx beam refinement (Alt3).* |
| Ericsson[10]  | *Proposal 4 Further define the FFS on AI/ML output after sufficient progress is made on studying the specification impact for AI/ML model inference aspects* |
| IDC[11] | *Proposal 10: Support ‘Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams’ as a baseline.**Proposal 11: ‘Tx and/or Rx Beam ID(s) of the N predicted DL Tx and/or Rx beams and other information’ can be considered with LOS probability.**Proposal 12: Benefits from utilization of TX/Rx beam angles should be clarified.* |
| Xiaomi[12] | *Proposal 8: Support Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams as the AI/ML model output with high priority.* |
| Intel[13] | *Proposal 2: For BM-Case1 and 2, Alt-1 (Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams) should be considered as the baseline use case, with potential specification impact on how beam IDs are mapped in the spatial domain.* |
| OPPO[14] | *Proposal 10: For the output of AI/ML model for BM-Case1 and BM-Case2, suggest to include at least* *• Tx and/or Rx Beam ID(s)**• The predicted L1-RSRP of the predicted Top-K DL Tx and/or Rx beams**• Note: the above output should be extended for F time instances for BM-Case2* |
| CT[16] | *Proposal 5: For the output of the AI model, Alt.1 can serve as a benchmark for performance comparison. And other options need further discussion.* |
| Sony[19] | *Observation 1 : The complexity of beam measurement can be reduced if a subset of candidate beams can be predicted by AI/ML model.**Proposal 3 : In output of AI/ML, should clearly indicate the criterion associated with the predicted beam ID in BM-case1 and BM-case2, for example, sum probabilities of being the best beams higher than a threshold, maximum dwelling time, maximum RSRP, etc.*  |
| Lenovo[22] | *Proposal 3: Tx and/or Rx Beam ID(s) and the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams should be taken as the AI/ML model output at least for UE-centric AI inference**Proposal 4: When specifying the AI/ML model output, we should consider that it may be used for model monitoring* |
| SS[27] | *Proposal 9: For AI/ML output for beam prediction, Alt 1 (e.g., Rx beam ID) is preferred.* |
| MTK[30] | *Proposal 6: RAN1 will discuss and define the method(s) to select the top N DL Tx and/or Rx beams for AI/ML model output.* |
| NEC[34] | *Proposal 2: Support selecting Top-N1 DL Tx and/or Rx beams according to some pre-defined rules, e.g., a sum probability of being the best beam higher than a threshold, L1-RSRP higher than a threshold.* |
|  |  |

Based on the discussions of RAN1#110bis-e meeting, down-selection seems not achievable at the current stage. Thus, we will focus on the spec impact (if any) of AI model output in other section(s).

For the proposed types of “other information”, in the last RAN1 meeting many companies thought “too little evaluation and description for these parameters” so that it was hard to make any discussion/decision. By reviewing all the submitted tdocs, moderator feels that there seems no much additional information for the proposed “other information”. Thus, proponents of a given type of “other information” are encouraged to provide more inputs to convince other companies. Otherwise, there cannot be any progress.

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Assistance information

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Observation 1: Assistance information may come with additional cost like signalling overhead, extra UE measurement overhead (including complexity, power consumption, etc.) There is usually a trade-off between performance gain and the associated overhead.* *Proposal 4: When assistance information is used as input, study its performance gain vs. the standards impacts and overhead.* |
| Huawei[2] | *Proposal 2: Regarding the assistance information for the input of NW-side AI/ML model (if supported), study of the following information is deprioritized:** *UE location*
* *UE moving direction*
* *NW-side beam shape information*
	+ *E.g., 3dB beamwidth, beam boresight directions, beam shape, Tx beam angle, etc.*

*Proposal 3: For the BM-Case 1 study of the AI/ML model input, if Alt.2 is included, for the determination/selection of assistance information:* * *Information that currently is proprietary/privacy information should not be disclosed*
* *The performance benefits of non-proprietary/non-privacy assistance information should be evaluated firstly to justify a study of their specification impact.*
	+ *Note: Generalization is included in performance*

*Proposal 10: For the BM-Case 2 study of the AI/ML model input, if Alt.2 is studied, for the determination/selection of assistance information:* * *Information that currently is proprietary/privacy information should not be disclosed*
* *The performance benefits of remaining assistance information should be evaluated firstly to justify a study of their specification impact.*
	+ *Note: Generalization is included in performance*
 |
| Vivo[3] | *Proposal 5: For the determination/selection of assistance information,* *• The performance, model generalization and potential specification impacts should be considered.**• Study how to protect sensitive proprietary/privacy information and disclose beam specific related assistance information.* |
| Fujitsu[6] | *Proposal 2: Regarding the sub use case BM-Case1, if one assistance information other than beam IDs is supported as input of AI/ML model, it’s suggested to further investigate* * *Additional performance gains*
* *The proprietary issue*

*Proposal 3: Regarding the sub use case BM-Case1, it is suggested that the assistance information other than beam ID is optional input of AI/ML model.* |
| CATT[8] | *Proposal 2: For the assistance information as model inputs, the proprietary/privacy information should not be disclosed.**Proposal 3: Whether the assistance information is proprietary/privacy information or not should be discussed separately with UE-sided model and NW-sided model.* |
| Ericsson[10] | *Observation 1 The feasibility of defining a meaningful TX/RX beam shape information for beam prediction is questionable.**Observation 2 It is possible to deduce AoD while keeping the beam shape information proprietary**Observation 3 There is no precedent for disclosing beam radiation patterns from the Rel-17 work on positioning**Proposal 1 Assistance information related to “beams” should focus on information related to NW antenna/beam configuration ID or UE antenna/beam configuration ID**Proposal 2 Prioritize assistance information that can be obtained with low standardization effort, such as UE position information**Proposal 3 Study assistance information that captures dynamic UE movement (e.g. using sensors)* |
| OPPO[14] | *Proposal 9: For the assistance information of BM-Case1 and BM-Case2, suggest to**• Justify the performance benefits if assistance information is input to model**• Identify whether assistance information would expose proprietary and/or privacy information of NW-side or UE-side.* |
| TCL[15] | *Proposal 1: The UE position information is not necessary for predictive beam switching.**Proposal 2: Support the UE moving speed as a kind of assistant information for beam prediction in time domain.* |
| ETRI[17] | *Proposal 1. For BM-Case1 and BM-Case2 with a network-side AI/ML model, positioning-related information may be included as assistance information.* |
| Panasonic[18] | *Observation 1: Beam pattern information can be defined as model input to make the model more general. Otherwise, multiple models need to be trained with each corresponding to one specific assumption of Tx beam pattern.**Observation 2: For UE-side inference, Tx beam pattern information needs to be made available at UE side.**Proposal 1: Relative angle/direction among beams are signaled to UE.* |
| NVIDIA[21] | *Proposal 5: Comprehensive evaluation results showing convincing performance gains is needed to nail down the essential assistance information needed for the spatial-domain DL beam prediction.**Proposal 7: Comprehensive evaluation results showing convincing performance gains is needed to nail down the essential assistance information needed for the temporal DL beam prediction.* |
| Lenovo[22] | *Proposal 2: Assistance information for AI/ML input should be carefully studied considering the availability of different kinds of assistance information for UE-centric or NW-centric AI/ML inference.* |
| Apple[23] | *Observation 2: the Tx analog beam information is already embedded in the training data. Whether additional information about Tx beams such as Tx beam shape and Tx beam angle can be useful, or concepts such as Tx beam shape and/or Tx beam orientation can be used in practice need further study.* *Observation 3: conventionally Rx beam design is transparent to network operation, AI/ML aided/enabled beam management does not need to depart from that. Whether additional information about Rx beams such as Rx beam shape and Rx beam angle can be useful, or concepts such as Rx beam shape and/or Rx beam orientation can be used in practice need further study.**Observation 4: If explicit Tx beam shape information for different datasets is not available to model trainer, it may be difficult to design AI model to generalize well over different scenarios/configurations. However, acquiring explicit Tx beam shape information at the UE side may be difficult due to concerns on disclosing proprietary information.**Proposal 2: If UE position information is used AI/ML aided beam management, user privacy needs to be considered in data collection for model training and input for inference with UE position information.* |
| DCM[26] | *Proposal 1: Study the performance/mechanism of assistance information only if at least one company from each UE vendor, gNB vendor and operator agrees with the feasibility of disclosing the information to the other side from the proprietary perspective.**Observation 6: Boresight direction and/or (relative) power per angle for each reference signal can be potential assistance information of Tx beam in DL beam prediction.* |
| QC[28] | *Proposal 2: Study the signalling aspects related to gNB sending assistance information to help UE-side AI/ML models**• Examples of such assistance information: information about gNB beam shape, beam boresight directions, 3dB beamwidth, etc.**• Study means to provide beam-shape related assistance information while preserving sensitive proprietary information**o Consider Rel-17 positioning agreement as a starting point* |
| MTK[30] | *Proposal 5: RAN1 will discuss and agree on prioritizing and down-scoping alternatives of UE assistance information.* |
| Nokia[33] | *Proposal 2: For BM-Case1 with Set A/B considering Tx-Rx pairs, further discussion may be needed on NW side DL Tx-AoA prediction, UE position information as assistant info to the input of ML model.* |
|  |  |

In the last RAN1 meeting, most companies support to not disclose proprietary/privacy information, which is widely accepted as a best practice in 3GPP so far. Meanwhile, some companies thought there is some mechanism(s) not disclosing proprietary/privacy information to get some given assistance information. Thus, it may be a more constructive way to study whether there is some mechanism not disclosing proprietary/privacy information for a given assistance information before the group decide whether to support it or not.

Among all the proposed assistance information, the Tx beam shape related and UE position related information get more supporters than others. Thus, the following proposals are suggested for discussion.

###### Proposal 4.1

***Proposal 4.1:*** ***For assistance information related to gNB Tx beams (e.g., Tx beam shape, …)***

* ***Study the feasibility and how (if feasible) a mechanism not disclosing the proprietary information***
* ***Whether to support this assistance information or not is a separate discussion***
	+ ***At least the performance and specification impacts should be considered***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Spec impact of Data collection

###### Proposal 4.2

***Proposal 4.2:*** ***For assistance information related to UE position/moving speed***

* ***Study the feasibility and how (if feasible) a mechanism not disclosing the privacy information***
* ***Whether to support this assistance information or not is a separate discussion***
	+ ***At least the performance and specification impacts should be considered***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## General/common aspects

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110**AgreementFor the data collection for AI/ML model training (if supported), study the following aspects as a starting point for potential necessary specification impact:* Signaling/configuration/measurement/report for data collection, e.g., signaling aspects related to assistance information (if supported), Reference signals
* Content/type of the collected data
* Other aspect(s) is not precluded
 |

The related proposals/ observations related are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 22: RAN1 to further study the potential spec impact of data collection from a realistic network for training from the following aspects:** *For reference signal, enhanced RS design can be considered, e.g., RS design for AI/ML specific RSRP measurement and enhancement of RS for improving data sample accuracy*
* *For UE measurement/report, new RSRP and/or CRI/SSBRI report behavior can be considered*
* *For the signaling/configuration, signaling to trigger/configure/request data collection window can be considered*
 |
| Ericsson[10] | *Proposal 6 Conclude that the time scale on which data collection for NW for model training for beam management use cases is fulfilled by an RRC-message based approach.**Proposal 7 Study details of an RRC-message based data collection framework for supporting UE performing data logging/collection and reporting the collected data to NW for model training for beam management use cases.**Proposal 8 For data collection for model training, study a measurement occasion which can consist of a single RS resource set (e.g., for spatial beam prediction) or a burst of multiple RS resource sets at different time instances (e.g., for temporal beam prediction).**Proposal 9 For data collection for model training for the beam management use cases, study and define the RS resource set(s) configuration parameters per measurement occasion.**Proposal 12 For data collection for model training for the beam management use cases, study and define the candidate values of measurement occasion interval (data logging interval).**Proposal 13 For data collection for model training for the beam management use cases, study and define the candidate values of data collection duration (logging duration).**Proposal 14 For data collection for AI/ML model training for the beam management use cases, study and define event types that may be used to trigger a UE to perform measurements and data logging.* |
| Intel[13] | *Observation 1: The impact of 3GPP specification related procedures for data collection for training as well as inference depends on where the model resides and if training and inferencing is being performed at the same node.**Observation 2: Training dataset construction using 3GPP specified measurement and reporting framework may be advantageous for harmonizing deployment of proprietary AI/ML models.* |
| OPPO[14] | *Proposal 11: Study data collection for AI/ML model offline training and fine-tuning with legacy beam measurement as a starting point.* |
| CT[16] | *Proposal 6: Study the effect of different data types on prediction accuracy or performance monitoring.**Proposal 7: Further study the effects of the amount of data on model training, signaling overhead, storage consumption, model inference and performance monitoring.* |
| CMCC[20] | *Proposal 2: For data collection of AI/ML based TX-RX beam pair prediction, study how to define and map the beam pair ID to align the understanding between the network and the UE.* |
| NVIDIA[21] | *Proposal 8: For AI/ML model training for beam prediction in spatial/time domain, study potential specification impact related to training data type/size, training data source determination, and assistance signalling and procedure for training data collection.* |
| NEC[34] | *Proposal 3: Study explicit configuration of AI/ML model specific RSs for data collection in model training, model update and model monitoring.**Observation 1: For data collection in model training, model monitoring or model update, the beam information corresponding to input and output (i.e., partial beams in Set A) are needed, rather than the beams information of all beams in Set A.**Proposal 4: Study enhanced beam reporting for data collection in model training, model monitoring and model update.**Proposal 5: For the trained AI/ML model in offline, study the mechanism of model update (e.g., fine-tuning) based on the online data.**Proposal 6: Study the mechanism of online data processing.* |
|  |  |

Based on the agreement of RAN1#110, it would be beneficial to study more details of data collection. Thus, a proposal is proposed as below based on the submitted tdocs.

###### Proposal 5.1

***Proposal 5.1: For BM-Case1 and BM-Case2, study beam management specific requirement(s)/potential specification impacts of data collection for AI/ML model training from the following aspects***

* ***Requirements of the data set for BM-related AI model training (e.g., size of training data)***
* ***Mechanism to configure/trigger the corresponding measurement (e.g., measurement occasion, enhanced RS, …)***
* ***Mechanism to configure/trigger data collection/logging (e.g., data collection/logging window)***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## AI model training at NW side

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 23: At least for data collection for AI/ML model training at NW-side, support the following to options to obtain AI/ML input and labels:** *Opt.1: UE reports RSRPs for beams in Set A, and Set B if applicable*
* *Opt.2: UE reports RSRPs for beams in Set B and the best genie-aided beam ID in Set A*

*Proposal 24: Regarding the data collection for AI/ML model training at NW side, study the following information for UE reporting as a starting point:** *M L1-RSRPs and optionally with the corresponding RS indicator and/or best genie aided beam (pair) indicator, where M can be larger than 4*
	+ *FFS: the range of M*
* *Other information may be added based on further discussion*
 |
| Vivo[3] | *Proposal 10: For AI/ML model training and inference at NW side with enhanced beam pair prediction, at least study the following aspects for potential specification impact in data collection procedure:**• Enhanced resource configuration**- Specific beam pair resource configuration**• Enhanced beam report**- All measured L1-RSRP should be reported to gNB**- Relative Rx beam information (including measured Rx beam information and/or expected Rx beam information) as assistance information may be reported**- Considering report overhead reduction, such as enhanced L1-RSRP quantization method, reducing unnecessary L1-RSRPs in a beam report where the omitted L1-RSRPs may be lower than a pre-defined threshold**Proposal 11: For AI/ML model training and inference at NW side with DL Tx beam prediction, at least study the following aspects for potential specification impact in data collection procedure:**• Enhanced resource configuration**- P3+P2 resource configuration that Rx beam assumption of P2 resource measurement is the best Rx beam searched from P3 procedure**• Enhanced beam report**- Only P2 measured L1-RSRP should be reported to gNB**- Best Rx beam information as assistance information may be reported**- Considering report overhead reduction, such as enhanced L1-RSRP quantization method, reducing unnecessary L1-RSRP reporting where the omitted L1-RSRPs may be lower than a pre-defined threshold* |
| H3C[4] | *Proposal 1: Regarding the data collection for AI/ML model training at NW side (if the data collection is optionally supported from the perspective of 3GPP specification), study the following information for UE reporting as a starting point.** + - *M L1-RSRPs with the corresponding RS indicator*
* *FFS: the range of M, same as the beam number in SetB*
	+ - *Note: Label the collected data should be done by gNB*
 |
| ZTE[5] | *Proposal 9: For data collection at NW side, support to study explicit or implicit Rx beam ID reporting method, especially for the sub use case of beam pair prediction.**Observation 7: For data collection at NW side, reporting of model label from UE can be specified as the genie-aided best beam ID from set A, all measurement results of set A or other post-processing of measurement results of set A, which depends on the model training strategy.**Proposal 10: If all measurement results of set A/B need to be reported to gNB, suggest to further study reporting overhead reduction, e.g., beam ID can be obtained implicitly from the reporting order of all measured RSRPs.**Proposal 11: Support to study resource configuration aspects of data collection at NW side and associated spec impact with potentially enhanced signaling mechanisms and auxiliary information transmission.* |
| Fujitsu[6] | *Observation 4: UCI reporting overhead is increased a lot on data collection for AI/ML model training at NW side.**Proposal 6: Regarding the data collection for AI/ML model training at NW side, the mechanism of UCI reporting overhead reduction is suggested to be studied.* |
| CATT[8] | *Proposal 6: Regarding the data collection for AI/ML model training in BM-Case1 and BM-Case2, at least study the following aspects on RS transmission.** *Alt1: gNB transmits RS in both Set A and Set B to UE;*
* *Alt2: gNB transmits RS in Set A and informs the beam pattern of Set B to UE.*

*Proposal 7: Regarding the data collection for AI/ML model training at NW side, study the following information for UE reporting as a starting point:** *For model inputs, report the measurement results (e.g., L1-RSRP) of Set B;*
* *For the label of model outputs, study the following alternatives depend on AI algorithm and model functionality:*
	+ *Alt1: Best genie-aided beam ID from Set A;*
	+ *Alt2: Measurement results (e.g., L1-RSRP) of Set A;*
	+ *Other alternatives are not precluded.*
 |
| Ericsson[10] | *Proposal 10 For data collection for model training for the beam management use cases, a UE should log and report all measurements performed on CSI-RS/SSB resources associated with Set A and Set B of beams at each measurement occasion.**Proposal 11 For data collection for model training for the beam management use cases, study and define signaling and report format to support a UE reporting logged data to the NW.**Proposal 15 For data collection for AI/ML model training for the beam management use cases, study and define event types that may be used to trigger a UE to report the logged data to the NW.* |
| Lenovo[22] | *Proposal 6: Study data collection procedure to support both UE-side and NW-side AI/ML model training and model update** For UE-centric model training, study procedure to support UE triggered data collection for model update** For NW-centric model training, support to report larger number of beams in one beam report.* |
| CIACT[25] | *Proposal 4: For AI model training and inference at NW side, UE should report M L1-RSRPs with the corresponding RS indicator and the best beam ID for Set A to NW for AI model training.* |
| SS[27] | *Proposal 11. For the data collection for AI/ML model training, in the case that AI/ML model is at gNB-side, the following aspects can be further study:** Potential enhancement for the measurement and report for data collection** The handling/buffering of the collected data* |
| Nokia[33] | *Proposal 11: For data collection purpose at the NW side, study the CSI reporting enhancement (e.g., reporting more than 4 beams and associated L1-RSRP) such that NW may update the data set for model training/update/fine-tuning.**Proposal 12: For model inference at the NW side, study the CSI reporting enhancement on how to configure measurements of a fixed or variable set for Set B measurements.*  |

In the last meeting, there were some intensive discussions on whether the best genie-aided beam ID from Set A should be reported or not, and no consensus was made. In the submitted tdocs, there are different assumptions/alternatives for the data collection mechanism. It would be good to have better understanding of these different alternatives before we do the potential down-selection (if any). Thus, the following proposal is suggested to encourage the study on pros/cons of each alternatives

###### Proposal 5.2.1

***Proposal 5.2.1: Regarding the data collection mechanism for AI/ML model training at NW side, study the following alternatives as a starting point.***

* ***Alt.1: UE measures the beams of Set A and report M1 L1-RSRPs with the corresponding RS indicators, where M1 can be larger than 4***
	+ ***FFS: the range of M1***
* ***Alt.2: UE measures the beams of Set B and report M2 L1-RSRPs with the corresponding RS indicators, where M2 can be larger than 4, measures the beams of Set A and report M3 L1-RSRPs with the corresponding RS indicator, where M3 can be larger than 4,***
	+ ***FFS: the range of M2, M3***
	+ ***Note1: the measurement and reporting related to Set A may be separate from/transparent to the operations related to Set B***
* ***Alt.3: UE measures the beams of Set B and report M4 L1-RSRPs with the corresponding RS indicator, where M4 can be larger than 4, measures the beams of Set A and report M5 RS indicator s corresponding to the best beam(s)***
	+ ***FFS: the range of M4, M5***
	+ ***Note2: the measurement and reporting related to Set A may be separate from/transparent to the operations related to Set B***
* ***Other alternative(s) is not precluded***
* ***Note3: Data collection for model training may be implemented by gNB in a transparent way***
* ***Note4: Potential down-selection/prioritization will be discussed later***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## AI model training at UE side

|  |  |
| --- | --- |
| Vivo[3] | *Proposal 12: For AI/ML model training and inference at UE side with enhanced beam pair prediction, at least study the following aspects for potential necessary specification impact in data collection procedure:**• Enhanced request signaling**- Beam pair resources request**- Minimum number of beams request in advance or with resource request* *- Tx beam information (including measured Tx beam and/or expected Tx beam) request**• Enhanced resource configuration**- Specific beam pair resource configuration**- Tx beam information (such as 3dB beam width, ID or angle for measured Tx beam and/or expected Tx beam in model input) may be signaled with resource configuration**Proposal 13: For AI/ML model training and inference at UE side with enhanced DL Tx beam prediction, at least study the following aspects for potential necessary specification impact in data collection procedure:**• Enhanced request signaling**- P3+P2 beam sweeping resources request**- Minimum number of beams request in advance or with resource request* *- Tx beam information request**• Enhanced resource configuration**- P3+P2 resource configuration that Rx beam assumption of P2 resource measurement is the best Rx beam searched from P3 procedure**- Tx beam information (such as 3dB beam width, ID or angle for measured Tx beam and/or expected Tx beam in model input) may be signaled with resource configuration**Proposal 14: For AI/ML model training and inference at UE side with DL Rx beam prediction, at least study the following aspects for potential necessary specification impact in data collection procedure:**• Enhanced request signaling**- P3 beam sweeping resources request**- Minimum number of repetition request in advance or with resource request* *- Tx beam information request**Proposal 15: For Alt.3. which is AI/ML model training at NW side and inference at UE side, it has similar data collection procedure and potential specification impacts as Alt.1, i.e. both model training and model inference at NW side, for enhanced beam pair prediction and DL Tx beam prediction scheme.* |
| H3C[4] | *Proposal 2: Regarding the data collection for AI/ML model training at UE side, study the following information as a starting point.** + - *UE request the gNB to send beams of SetA and/or SetB with L1 signaling*
		- *Note: Label the measurement data should be done by UE*
 |
| Lenovo[22] | *Proposal 6: Study data collection procedure to support both UE-side and NW-side AI/ML model training and model update** For UE-centric model training, study procedure to support UE triggered data collection for model update** For NW-centric model training, support to report larger number of beams in one beam report.* |
| CIACT[25] | *Proposal 3: For AI model training and inference at UE side, NW could provide some assistant information for data collection and /or AI model selection at UE side.* |
| SS[27] | *Proposal 12. For the data collection for AI/ML model training, in the case that AI/ML model is at UE-side, the following aspects can be further study:** UE report for the preference of data collection, e.g., intended/preferred RS transmission for UE measurement, intended/preferred time domain pattern of the RS transmission** RS measurement configuration for data collection* |
|  |  |

###### Proposal 5.3.1

***Proposal 5.3.1: Regarding the data collection for AI/ML model training at UE side, study the potential specification impact from the following aspects as a starting point.***

* ***Request message for data collection from UE to NW***
	+ ***FFS: contents of the request message***
* ***Configuration of Set A and Set B from NW to UE (e.g., association/mapping of Set A and Set B)***
* ***Other aspect(s) is not precluded***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Spec impact of AI/ML inference for BM-Case1 & BM-Case2

## General/common aspects

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110**AgreementIn order to facilitate the AI/ML model inference, study the following aspects as a starting point:* Enhanced or new configurations/UE reporting/UE measurement, e.g., Enhanced or new beam measurement and/or beam reporting
* Enhanced or new signaling for measurement configuration/triggering
* Signaling of assistance information (if applicable)
* Other aspect(s) is not precluded
 |

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| Vivo[3] | *Proposal 22: In model inference procedure, Alt.3, i.e. model training at NW side and model inference at UE side, with enhanced beam pair prediction and DL Tx beam prediction scheme has similar specification impacts as an AI model trained and inferenced at UE side.**Proposal 23: Study signaling aspects enhancement related to the procedure of model transfer, model registration and model activation, for the case with AI/ML model training at NW side and AI/ML model inference at UE side.* |
| ZTE[5] | *Observation 5: For Alt.3 in BM-Case2, a feasible working mode is to configure/transmit the RS resource set (i.e., set A/B) only in the measurement window to minimize the RS overhead for beam measurement.**Proposal 6: For Alt.3 in BM-Case2, it is necessary to study flexible RS resource set and report configuration within the measurement window and prediction window, regardless of NW-side model or UE-side model.**Observation 6: For Alt.2 in BM-Case2, a feasible working mode is to configure/transmit the RS resource set (i.e., set B) in both the measurement window and the prediction window.**Proposal 7: For Alt.2 in BM-Case2, if set B can be randomly changed, support to study enhanced resource configuration and activation method to flexibly activate/deactivate arbitrary beams or beam subset among pre-configured patterns in set A beams (pairs).**Proposal 12: In order to facilitate AI/ML operations for BM-Case1 and BM-Case2, study the following additional aspects:** *Beam indication of the unmeasured Tx beam from network to UE*
* *Beam indication of the predicted DL beam pair from network to UE*
* *Beam indication of multiple future time instances*
 |
| Spreadtrum[9] | *Observation 2: For beam indication, the Rel15/16/17 TCI framework can be considered as starting point.** *If AI/ML inference is at NW side, how to determine/indicate the best Rx beam needs further study*
* *If AI/ML inference is at UE side, no specification impact is identified*
 |
| IDC[11] | *Observation 14: The current NR specification does not consider association between beams with different beam widths.**Observation 15: Utilizing association between beams with different beam widths can provide benefits for prediction accuracy e.g., robust estimation/identification of whole spatial characteristics with wide beams and accurate beam identification with narrow beams.**Proposal 16: Study benefits of specification enhancements on association between beams with different beam widths.* |
| Xiaomi[12] | *Proposal 6: For the case of Tx beam or TxRx beam pair inference with specific Rx, support to indicate Rx beam information to UE for obtaining L1-RSRP input to AI/ML model.**Proposal 7: Support a common AI model for UE with different number of Rx beam.* |
| TCL[15] | *Proposal 4: How to merge the beam prediction into beam training steps is needed to be considered.* |
| NVIDIA[21] | *Proposal 11: For AI/ML based beam prediction in spatial/time domain, study potential specification impact related to report/feedback of model input for inference, type of model input, and model input acquisition and pre-processing.**Proposal 12: For AI/ML based beam prediction in spatial/time domain, study potential specification impact related to report/feedback of model inference output and post-processing.* |
| Lenovo[22] | *Proposal 11: Study on how to obtain the assisting information for AI/ML model input.* |
| NEC[34] | *Proposal 9: Study the mechanism of reducing the overhead of beam measurement and reporting in model inference.**Proposal 10: Study the mechanism of discontinuous reporting in periodic or semi-persistent beam reporting. (for BM-Case2)**Proposal 11: Study the method of indicating the predicted beams and corresponding beam application/dwelling times. (for BM-Case2)* |
|  |  |

In the last meeting, there were intensive discussion on the indication of predicted beams. There were some comments raised during the discussion:

* Clarification is needed for the details of “indication of predicted beams”.
	+ Two bullets are used to list all the detailed indications.
* Legacy mechanism can be reused and no spec enhancement is needed.
	+ A note is added for the 1st bullet to emphasize the above information

###### Proposal 6.1

***Proposal 6.1: In order to facilitate AI/ML operations for BM-Case1 and BM-Case2, study the following additional aspects:***

* ***Beam indication of the unmeasured Tx beam from network to UE***
	+ ***Note: This may or may not have specification impact (e.g., legacy mechanism may be reused).***
* ***Beam indication of multiple future time instances for BM-Case2***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## AL/ML inference at gNB side

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110bis-e**Working AssumptionFor BM-Case1 and BM-Case2 with a network-side AI/ML model, study the following L1 beam reporting enhancement for AI/ML model inference* UE to report the measurement results of more than 4 beams in one reporting instance
* Other L1 reporting enhancements can be considered
 |

The related proposals/observations for both BM-Case1 and BM-Case2 are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 3: For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the L1 beam reporting enhancement for AI/ML model inference, including the mechanism that enables gNB to map the received L1-RSRP measurements to the corresponding Tx-Rx beam pairs.* |
| Huawei[2] | *Observation 1: For NW-side model and BM-Case 1, the relationship between Set B and Set A is transparent to UE for DL Tx beam prediction.**Proposal 8: For NW-side model and BM-Case 1, the relationship between Set B and Set A can be studied for DL Tx-Rx beam pair prediction, where the Rx beam IDs for Set B and the Rx beam IDs for Set A may need to be known by the gNB.**Observation 3: For NW-side model and BM-Case 2, the relationship between Set B and Set A is transparent to UE for DL Tx beam prediction.**Proposal 15: For NW-side model and BM-Case 2, the relationship between Set B and Set A may be studied for DL Tx-Rx beam pair prediction, where the Rx beam IDs for Set B and the Rx beam IDs for Set A are informed to gNB.**Proposal 29: Confirm the Working Assumption from RAN1#110bis-e: For BM-Case1 and BM-Case2 with a NW-side AI/ML model, study the following L1 beam reporting enhancement for AI/ML model inference** *UE to report the measurement results of more than 4 beams in one reporting instance*
* *Other L1 reporting enhancements can be considered*

*Observation 10: For AI/ML model inference at the NW-side, at least for BM-Case 1, there is no need to signal the predicted beams from network to UE.*  |
| Vivo[3] | *Proposal 16: For AI/ML model training and inference at NW side with enhanced beam pair prediction, at least study the following aspects for potential necessary specification impact in model inference procedure:**• New signaling**- Expected Rx beam information should be reported, such as supported DL Rx beam angle or scope at UE side**• Enhanced resource configuration**- Specific beam pair resource configuration**• Enhanced beam report**- All measured L1-RSRP should be reported to gNB**- Relative Rx beam information as assistance information may be reported**- Considering report overhead reduction, such as enhanced L1-RSRP quantization method, reducing unnecessary L1-RSRP in beam report where the omitted L1-RSRPs may be lower than a pre-defined threshold**Proposal 17: For AI/ML model training and inference at NW side with DL Tx beam prediction, at least study the following aspects for potential necessary specification impact in model inference procedure:**• Enhanced resource configuration**- P3+P2 resource configuration that Rx beam assumption of P2 resource measurement is the best Rx beam searched from P3 procedure**• Enhanced beam report**- Only P2 measured L1-RSRP should be reported to gNB**- Best Rx beam information as assistance information may be reported**- Considering report overhead reduction, such as enhanced L1-RSRP quantization method, reducing unnecessary L1-RSRP reporting where the omitted L1-RSRPs may be lower than a pre-defined threshold**Proposal 18: Confirm above working assumption that a UE can report measurement results of more than 4 beams in one reporting instance for BM-Case1 and BM-Case2 with a network-side AI/ML model.* |
| ZTE[5] | *Proposal 3: For the NW-side beam pair prediction, it may be desirable to implicitly indicate the Rx beam ID to facilitate data collection at the NW side and avoid UE proprietary information disclosure issue.**Proposal 13: For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the following L1 beam reporting enhancements for AI/ML model inference:** *Reporting resolution enhancement*
* *Reporting overhead reduction*
* *Assistance information reporting*
* *Other reporting enhancements to handle inference latency*
 |
| Fujitsu[6] | *Proposal 5: For NW-side model, study the potential specification impacts on UE behavior of beam reporting.**Proposal 7: For BM-Case1 and BM-Case2 of DL beam pair prediction with a network-side AI/ML model, study the potential specification impacts on Rx beam information included in report instance for AI/ML model inference.** *FFS: the definition of the Rx beam information (e.g., beam IDs/angels)*
 |
| Google[7] | *Proposal 14: For AI/ML in gNB side, study the following potential enhancement to reduce the L1-RSRP measurement and quantization error.** *CSI-RS coverage enhancement*
* *More advanced receiver to reduce measurement error*
* *High-resolution quantization scheme to reduce quantization error*
 |
| CATT[8] | *Proposal 8: Regarding the data collection for AI/ML model inference at NW side, the UE needs to report the measurement results (e.g., L1-RSRP) based on the RS received in Set B to gNB as model inputs.**Proposal 11: Confirm the working assumption for BM-Case1, i.e., For BM-Case1 with a network-side AI/ML model, study the following L1 beam reporting enhancement for AI/ML model inference:** *UE to report the measurement results of more than 4 beams in one reporting instance;*
* *Other L1 reporting enhancements can be considered.*

*Proposal 12: For BM-Case2 with a network-side AI/ML model, study the following L1 beam reporting enhancement for AI/ML model inference:** *UE to report the measurement results of more than 4 beams in one reporting instance for each measurement instance among K (K>=1) latest measurement instances;*
* *Other L1 reporting enhancements can be considered.*

*Proposal 13: For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the following TCI states enhancement for AI/ML model inference:** *How to indicate the predicted best beam in TCI states.*
 |
| Spreadtrum[9] | *Proposal 5: For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the enhancement for beam reporting to report one DL Tx beam received by multiple Rx beams.* |
| Ericsson[10] | *Proposal 16 Further study specification impact of enhanced UE reporting of measurement accuracies and increased resolution**Proposal 17 Consider mechanism to signal UE assistance data (e.g. UE direction, UE orientation information, UE probability of being subject to dynamic blocker, UE speed) associated with beam measurement report for NW-sided model inference**Proposal 18 Consider enhanced UE configurations for NW-sided AI/ML model inference, for example NW indicates potential measurement pre-processing for reducing the UE uplink reporting overhead* |
| IDC[11] | *Observation 13: The current NR specification supporting UE reporting with up to 4 best CRIs/SSBRIs with L1-RSRP or L1-SINR can be very limited for a network-side AI/ML model.* *Proposal 15: Consider increasing number of CRIs/SSBRIs (e.g., 8 CRIs/SSBRIs).* |
| Xiaomi[12] | *Proposal 5: For spatial domain beam prediction, study to report Rx beam information, including Rx beam ID/Rx beam shape information of UE to gNB for gNB side inference.* |
| OPPO[14] | *Proposal 15: For BM-Case1 and BM-Case2 when inference at NW side, study L1 beam reporting mechanism to convey the measurements of Set B to NW.**Proposal 16: For BM-Case1 and BM-Case2 when inference at NW side, study the beam indication mechanism for Tx beam only prediction and Tx-Rx beam pair prediction.* |
| Panasonic[18] | *Observation 3: Existing TCI framework can be reused for the beam indication of the predicted DL Tx beam(s) from NW to UE.* |
| CMCC[20] | *Proposal 3: For BM-Case1 with a network-side AI/ML model, study the following L1 beam reporting enhancement for AI/ML model inference**• UE to report the measurement results of more than 4 beams (pairs) in one reporting instance**Proposal 4: For BM-Case1 with a network-side AI/ML model, study how to indicate the beam pair to UE.* |
| Lenovo[22] | *Proposal 9: Rel-17 CSI reporting framework can be reused for NW-side beam prediction by increasing the number of beams in a beam report.* |
| Apple[23] | *Proposal 1:** *For Model training at the NW side & inference at the NW side, study efficient signalling of set B selection or beam selection and RSRP representation.*
* *For Model training at the NW side & inference at the UE side, study model generalization performance, study model transfer/model delivery for cell-specific AI models and non cell-specific AI models.*
* *For Model training at the UE side, and inference at the UE side, study cell-specific signals to facilitate data collection.*
 |
| LG[24] | *Proposal #6: For NW-sided models, in addition to beam reporting enhancements, beam indication enhancement should be considered that TCI/QCL RS should be represented based on Set B beams of which UE can measure and maintain its Rx beam.**Proposal #12: For NW-sided AI/ML in BM-Case2, consider enhancements on UE reporting and beam indication.* |
| DCM[26] | *Proposal 4: Confirm the working assumption to study L1 beam reporting enhancements for model inference, such as more than 4 beams in one reporting instance.**Proposal 5: In DL beam prediction with NW side model, some mechanisms to report Rx beam ID used for beam measurements can be considered as potential specification impacts.**Observation 3: Enhancements on beam selection policy in CSI reports might be potential specification impacts for spatial domain beam estimation.**Proposal 6: CSI report should be enhanced to facilitate the model inference of temporal beam prediction with the overhead reduction.* |
| SS[27] | *Proposal 2: For BM-Case1 with a NW AI/ML model, further study the specification impacts for AI/ML inference considering the following aspects:** Enhancement on L1 beam report mechanism (e.g., report more than 4 beams in one reporting instance, enhanced granularity of L1-RSRP)** Assistance information for beam prediction**Proposal 6: For BM-Case2 with a NW AI/ML model, further study the specification impacts for AI/ML inference considering the following aspects:** Enhancement on L1 beam report mechanism (e.g., report more than 4 beams in one reporting instance, enhanced granularity of L1-RSRP)* |
| QC[28] | *Proposal 4**For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the following L1 beam reporting enhancement for AI/ML model inference**• Resolution enhancement, report of temporal variance of L1-RSRP/L1-SINR measurements, measurement imperfection indication* |
| KT[29] | *Proposal 2. For BM-Case1 and BM-Case2, study the potential specification impact of L1 signaling to collect the following inputs:**• L1-RSRP measurement based on Set B**• The corresponding DL Tx and/or Rx beam ID* |
| MTK[30] | *Proposal 4: RAN1 will study on the details and advancement of UE’s beam-related L1-RSRP report. UE can report more than 4 beams in one reporting instance.* |
| Nokia[33] | *Proposal 13: Confirm the following working assumption,* *Working Assumption**For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the following L1 beam reporting enhancement for AI/ML model inference** *UE to report the measurement results of more than 4 beams in one reporting instance*
* *Other L1 reporting enhancements can be considered*
 |
|  |  |

Several companies suggested to confirm the working assumption of RAN1#110bis-e. A minor modification is suggested in order to address SS’s concern raised in the last meeting.

###### Proposal 6.2.1

***Proposal 6.2.1: Confirm the working assumption of RAN1#110bis-e with the following modifications:***

***For BM-Case1 and BM-Case2 with a network-side AI/ML model, study potential specification impact on the following L1 beam reporting enhancement for AI/ML model inference***

* ***UE to report the measurement results of more than 4 beams in one reporting instance***
* ***Other L1 reporting enhancements can be considered***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Several companies suggest more bits for L1-RSRP reporting for AI model inference and/or training. Thus, the following proposal is suggested for discussion

###### Proposal 6.2.2

***Proposal 6.2.2: For BM-Case1 and BM-Case2 with a network-side AI/ML model, study potential specification impact from the additional aspects for UE reporting***

* ***Finer resolution for the UE reporting of the measurement results (e.g., more bits for L1-RSRP reporting)***
* ***Note: it is applicable to both AI model inference and training***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

For network-side AI/ML model, there are many tdocs to support and discuss the potential spec impacts of beam pair prediction. Meanwhile, there are also some companies not supporting beam pair prediction. It would be a natural way to discuss the feasibility, spec impacts and performance before the group decide the potential down-selection/prioritization (if any). One import issue is how to align the common understanding on the mapping of beam pairs and UE’s Rx beams. For example, let’s consider two UEs with the same number of Rx beams and the same Rx beam patterns (we denote the physical Rx beams by B1 and B2, respectively):

* The mapping of the physical Rx beams to some indicators (e.g., beam ID) is usually up to UE implementation.
* The same physical Rx beam (e.g., B1) of these two UEs may be associated with different beam IDs or different DL beam pairs.

Thus, the following proposal is suggested for discussion.

###### Proposal 6.2.3

***Proposal 6.2.3: For BM-Case1 and BM-Case2 with a network-side AI/ML model, study DL beam pair prediction from the following aspects as a starting point***

* ***How to align the common understanding between NW and UE on the mapping between beam pairs and UE’s associated Rx beams***
* ***How to indicate a beam pair from NW to UE***
* ***whether/how Rx beam related information reported from UE to NW***
* ***Note: The potential down-selection/prioritization (if any) on the types of beam prediction is a separate discussion***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## AL/ML inference at UE side

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110bis-e**AgreementFor BM-Case1 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to report the following information of AI/ML model inference to NW * The beam(s) that is based on the output of AI/ML model inference
* FFS: Predicted L1-RSRP corresponding to the beam(s)
* FFS: other information

AgreementFor BM-Case2 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to report the following information of AI/ML model inference to NW* The beam(s) of N future time instance(s) that is based on the output of AI/ML model inference
	+ FFS: value of N
* FFS: Predicted L1-RSRP corresponding to the beam(s)
* Information about the timestamp corresponding the reported beam(s)
	+ FFS: explicit or implicit
* FFS: other information
 |

The related proposals/observations for BM-Case1 and BM-Case2 are copied as below:

|  |  |
| --- | --- |
| Futurewei[1] | *Observation 3: For BM-Case1, predicted L1-RSRP is used together with other information available at the inference side as proxy to determine the Top-1/K beam(s) or beam pair(s). Using the predicted L1-RSRP values or the accuracy of predicted L1-RSRP alone cannot indicate the performance of Top-1/K prediction.* *Proposal 8: For BM-Case1 with a UE-side AI/ML model, if the predicted L1-RSRP is part of the output of AI/ML model at the inference side, do not include the Predicted L1-RSRP of the Beam(s) as information to be reported or shared to the other side.* |
| Huawei[2] | *Proposal 6: For UE-side model and BM-Case 1, for a specific Set B which the UE measures, it may be studied how to indicate the UE with the associated Set A which it may or may not have measured.* *Proposal 7: For UE-side model and BM-Case 1, the mapping between Set B and Set A (i.e., which beams from Set A are used to construct/associate with Set B) may be studied at least for following cases:** *When Set B is a subset of Set A, and when variable beams are used in Set B (if applicable)*
* *FFS when Set B is a set of wide beams different from Set*

*Proposal 13: For UE-side model and BM-Case 2, for a specific Set B which the UE measures, it can be studied to indicate UE with its associated Set A which it may or may not measure.**Proposal 14: For UE-side model and BM-Case 2, the mapping between Set B and Set A (i.e., which beams from Set A are used to construct Set B) may be studied at least for following cases:** *When Set B is a subset of Set A, and when variable beams are used in Set B (if applicable), including DL Tx prediction and DL Tx-Rx beam pair prediction*
* *FFS when Set B is a set of wide beams different from Set A*

*Proposal 28: For AI/ML model inference at the UE-side under BM-Case 1 and BM-Case 2, study the potential spec impact of L1 signaling to report the predicted beam IDs of more than 4 beams in one reporting instance.* |
| Vivo[3] | *Proposal 19: For AI/ML model training and inference at UE side with enhanced beam pair prediction, at least study the following aspects for potential specification impact in model inference procedure:**• New Signaling**- Expected DL Tx beam information indication, such as DL Tx beam angles or angle scopes.**- Renew beam pattern request from UE**• Enhanced resource configuration**- Beam pair resources configuration**- Tx beam information indication* *• Enhanced beam report**- Predicted L1-RSRP and relative Tx/Rx beam information report**- A fallback beam report including only measured results, rather than predicted results, as well as non-predicted L1-RSRP indication if all measured results may not suitable use for AI based beam prediction**Proposal 20: For AI/ML model training and inference at UE side with enhanced DL Tx beam prediction, at least study the following aspects for potential specification impact in model inference procedure:**• New Signaling**- Expected DL Tx beam information indication, such as DL Tx beam angles or angle scopes.**- Renew beam pattern request from UE**• Enhanced resource configuration**- P3 + P2 resource configuration that Rx beam assumption of P2 resource measurement is the best Rx beam searched from P3 procedure* *- Tx beam information indication* *• Enhanced beam report**- Predicted L1-RSRP and relative Tx/Rx beam information report**- A fallback beam report including only measured results, rather than predicted results, as well as non-predicted L1-RSRP indication if all measured results may not suitable use for AI based beam prediction**Proposal 21: For AI/ML model training and inference at UE side with DL Rx beam prediction, at least study the following aspects for potential specification impact in model inference procedure:**• Enhanced resource configuration**- UE may report minimum number of repetitions for model inference before P3 resource configuration**- Tx beam information may be signaled to UE* |
| H3C[4] | *Observation1: For BM-Case1 and BM-Case2 with a UE-side AI/ML model reference, report the Predicted L1-RSRP of the top-K beam(s) to NW can’t provide obvious benefits.**Proposal 3: For BM-Case1 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to report the following information of AI/ML model inference to NW* * *The beam(s) that is based on the output of AI/ML model inference*
* *FFS: other information*
 |
| Google[7] | *Proposal 15: For AI/ML in UE side, study the potential enhancement to maintain the same understanding between the gNB and UE with regard to the reported beam information based on a beam-codebook similar to CSI feedback based on a codebook** *The UE can report a beam matrix indicator (BMI) based on the beam-codebook*
 |
| CATT[8] | *Proposal 14: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to report the following information of AI/ML model inference to NW:** *Whether to report predicted L1-RSRP depends on the model outputs.*
 |
| Spreadtrum[9] | *Proposal 5: For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the enhancement for beam reporting to report one DL Tx beam received by multiple Rx beams.**Proposal 6: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the enhancement for beam report without RSRP.* |
| Ericsson[10] | *Proposal 19: Update agreement with changes marked in red: “For BM-Case1 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to report the following information of AI/ML model inference to NW* * *The beam(s) that is based on the output of AI/ML model inference*
* *~~FFS:~~ Predicted L1-RSRP corresponding to the beam(s) (if applicable)*
* *FFS: other information*
* *Confidence/uncertainty information related to the output of AI/ML model inference*

*Proposal 20 Study enhanced CSI report configuration to facilitate temporal and spatial beam predictions**Proposal 20 Study enhanced CSI report configuration to facilitate temporal and spatial beam predictions**Proposal 21 The investigation of assistance information signalling should prioritize efficient mechanisms for NW to indicate beam IDs to the UE* |
| IDC[11] | *Observation 11: For BM-Case 1 with a UE-side AI/ML model, the current beam reporting with CRIs/SSBRIs and corresponding L1-RSRP values is enough to indicate the best beam(s) from AI/ML model inference.* *Observation 12: For BM-Case 2 with a UE-side AI/ML model, information about the time stamp for the reported CRIs/SSBRIs can be further considered.* *Proposal 14: For a UE-side AI/ML model, consider information about the time stamp for potential specification impact.* |
| Xiaomi[12] | *Proposal 4: For spatial domain beam prediction, study how to indicate the Tx beam information, including Tx beam ID/Tx beam shape information of gNB to UE for UE side inference.**Proposal 9: Support to report predicted L1-RSRP in the L1-beam report with an indication to let gNB know which L1-RSRP is a predicted L1-RSRP.**Proposal 11: Consider one absolute L1-RSRP for each time instance or one absolute L1-RSRP for all time instance in one beam report including beam reports of more than one time instance for BM-case 2.**Proposal 12: Consider UE to report the number/ periodicity of the time instance in beam report for BM-case 2.* |
| Intel[13] | *Proposal 3: If the predicted RSRP is directly obtained from the output of the AI/ML model, such metrics should not be reported back to the network. If the predicted RSRP is defined as the actual measured RSRP on the best beam predicted by the model, such metric can be reported back to the network assuming that the overhead due to the additional measurement, if needed, is accounted for in the KPI calculation.* |
| OPPO[14] | *Proposal 12: For BM-Case1, if Rx beam(s) predicted at UE side, there seems no strong specification impact.**Proposal 13: For BM-Case2 beam predicted at UE side, one beam reporting instance should include beam prediction of N future time instance(s).**Proposal 14: For BM-Case2 beam predicted at UE side, the timestamp of N time instance(s) can be implicitly reported to NW.* |
| ETRI[17] | *Proposal 2. For BM-Case2 with a UE-side AI/ML model, N for future time instances can be static/semi-static/dynamic indication.* |
| Panasonic[18] | *Proposal 2: CSI reporting framework can be considered as starting point for UE to report beam prediction to NW in case of UE-side inference.**Proposal 3: Mechanism to distinguish between measured result and predicted result when UE reports CSI for beam management is necessary.* |
| CMCC[20] | *Proposal 5: For BM-Case1 with a UE-side AI/ML model, study whether to support UE to report the measurement results of more than 4 beams (pairs) in one reporting instance**Proposal 6: For BM-Case1 with a UE-side AI/ML model, whether the predicted L1-RSRP need to be reported can be configured by the gNB.* |
| Lenovo[22] | *Proposal 10: Rel-17 CSI reporting framework can be reused for UE-side beam prediction by configuring measurement beam Set B as the channel measurement resource but the reported beam is selected from another prediction beam Set A.* |
| Apple[23] | *Proposal 1:** *For Model training at the NW side & inference at the NW side, study efficient signalling of set B selection or beam selection and RSRP representation.*
* *For Model training at the NW side & inference at the UE side, study model generalization performance, study model transfer/model delivery for cell-specific AI models and non cell-specific AI models.*
* *For Model training at the UE side, and inference at the UE side, study cell-specific signals to facilitate data collection.*
 |
| LG[24] | *Proposal #4: For UE-sided models, support predicted L1-RSRP reporting together with beam information for NW to compare beam quality of multiple UEs and/or multiple beams of the same UE.**Proposal #5: For UE-sided models, the reported beam(s) that is based on the output of AI/ML model inference should be represented based on Set B beams of which UE can measure and maintain its Rx beam.**Proposal #8: For the beam reporting of N future time instance(s), consider UE to report beam(s) for each time instance or beam(s) for the time duration, i.e. from the first time instance to the last time instance.**Proposal #9: For fallback operation, it should be considered that UE to report the beam(s) of current time instance as well as the N future time instance(s).**Proposal #10: Support predicted L1-RSRP report together with beam(s). In case of N>1, information on time-variation of L1-RSRP can also be included in the report for helping intra-/extra-polation at NW side.**Proposal #11: Consider NW to indicate timestamp(s) for the predicted beam(s).* |
| DCM[26] | *Observation 4: Active models determine what information can be reported based on the model inference.**Proposal 9: Study the L1 signalling to report the top-1 or/and top-N/1 probability of certain beam(s) from each model inference as a potential specification impact.**Observation 5: Mechanisms to provide DL Tx beam information from NW to UE could be potential specification impacts in DL beam prediction**Proposal 10: Study two-stage beam measurements with top-N predicted beams, since it reduces RS measurement overhead and increases the reliability of beam selection compared to top-1 beam prediction.**Observation 7: Time instances corresponding to the reported beam(s) does not need to be reported when CSI reference resource is aligned between UE and NW, since it can be implicitly determined based on the measured RS occasion and the time offset.**Observation 8: It is beneficial to report the explicit predicted time instances in the reporting, if CSI reference resource is not always aligned between UE and NW.* |
| SS[27] | *Proposal 3: For BM-Case1 with a UE-side AI/ML model, further study the specification impacts for AI/ML inference considering the following aspects:** Assistance information for the determination of Set A** Enhancement on L1 beam report mechanism**Proposal 4: For BM-Case1 with a UE-side AI/ML model, the report of predicted L1-RSRP corresponding to a predicted beam is not necessary.**Proposal 7: For BM-Case2 with a UE-side AI/ML model, consider the following for L1 reporting of future beams:** For the beam(s) of N future time instance(s), N = 1 is baseline** Implicit timestamp corresponding the reported beam(s) is baseline**Proposal 8: For BM-Case2 with a UE-side AI/ML model, the report of predicted L1-RSRP corresponding to a predicted beam is not necessary.* |
| QC[28] | *Proposal 3: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the potential specification impact of reporting the following information of AI/ML model inference to NW:**• A measure of confidence for the predicted beam(s) and predicted L1-RSRP(s)**• Probability for the predicted beams to be the best beam**• FFS: frequency of such a report**Proposal 5: For BM-Case1 with a UE-side AI/ML model, study the potential specification impact of L1 signalling to report the following information of AI/ML model inference to NW:**• Information about NW DL TX beam angles from Set A**o UE may predict best beam angles from Set A by measuring Set B of DL TX beams**o FFS: details of beam angle, e.g., beam boresight direction**Proposal 6: For BM-Case2 with a UE-side AI/ML model, study the potential specification impact of L1 signalling to report the following information of AI/ML model inference to NW:**• Predicted beam blockage/failure* |
| KT[29] | *Proposal 3. For BM-Case1 and BM-Case2, study the potential specification impact of L1 signaling to report the L1-RSRP of AI/ML model inference to NW.* |
| Nokia[33] | *Proposal 9: For BM-Case1 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to optionally report the following information of AI/ML model inference to NW,* *• Predicted L1-RSRP corresponding to the beam(s)**• Other information (FFS: details)* *Proposal 10: For BM-Case2 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to optionally report the following information of AI/ML model inference to NW,* *• Other information (FFS: details)**Proposal 14: For UE side DL Tx beam or DL Tx-Rx beam pair prediction with collaboration level-y, RAN1 shall investigate further details about UE side model generalization and the corresponding NW-UE model alignment scheme.**Proposal 15: For UE side DL Tx beam or Tx-Rx beam pair prediction, further study configuring different RS resource sets for beam prediction and beam measurements.* |
|  |  |

Based on the tdocs, Futurewei/H3C/Intel/Spreatrum/SS (4 companies) thought reporting of the predicted L1-RSRP is not useful. In contrast, vivo/CATT/Ericsson/Xiaomi/Panasonic/CMCC/LG/QC/KT/Nokia (10 companies) supported to study it. Thus, the first bullet is suggested for discussion. It is only for further study, not decide to support it in this proposal. There are also several tdocs suggesting the reporting of confidence information of the top-K predicted beams). According, the 2nd bullet is suggested.

###### Proposal 6.3.1

***Proposal 6.3.1: For BM-Case1 with a UE-side AI/ML model, study the potential specification impact of additional information reported optionally from UE:***

* ***Predicted L1-RSRP corresponding to the beam(s)***
	+ ***Whether/how to differentiate predicted L1-RSRP and measured L1-RSRP***
* ***Confidence information related to the output of AI/ML model inference (e.g., predicted beams)***
	+ ***FFS: definition/content of confidence information***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

###### Proposal 6.3.2

***Proposal 6.3.2: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study potential specification impact from the following aspect(s):***

* ***Indication of the associated Set A from NW to UE***
* ***Indication of the beams of Set C for UE measurement***
	+ ***Note: Set C is for DL beam measurement. Set B may be Set C or a subset of Set C***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Spec impact of Model monitoring

## General aspects

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110**AgreementRegarding the model monitoring for BM-Case1 and BM-Case2, to investigate specification impacts from the following aspects* Performance metric(s)
* Benchmark/reference for the performance comparison
* Signaling/configuration/measurement/report for model monitoring, e.g., signaling aspects related to assistance information (if supported), Reference signals
* Other aspect(s) is not precluded

**RAN1#110bis-e**AgreementStudy AI/ML model monitoring for at least the following purposes: model activation, deactivation, selection, switching, fallback, and update (including re-training).FFS: Model selection refers to the selection of an AI/ML model among models for the same functionality. (Exact terminology to be discussed/defined)AgreementStudy at least the following metrics/methods for AI/ML model monitoring in lifecycle management per use case:* Monitoring based on inference accuracy, including metrics related to intermediate KPIs
* Monitoring based on system performance, including metrics related to system peformance KPIs
* Other monitoring solutions, at least following 2 options.
	+ Monitoring based on data distribution
		- Input-based: e.g., Monitoring the validity of the AI/ML input, e.g., out-of-distribution detection, drift detection of input data, or ~~something simple like checking~~ SNR, delay spread, etc.
		- Output-based: e.g., drift detection of output data
	+ Monitoring based on applicable condition

Note: Model monitoring metric calculation may be done at NW or UEAgreementStudy performance monitoring approaches, considering the following model monitoring KPIs as general guidance* Accuracy and relevance (i.e., how well does the given monitoring metric/methods reflect the model and system performance)
* Overhead (e.g., signaling overhead associated with model monitoring)
* Complexity (e.g., computation and memory cost for model monitoring)
* Latency (i.e., timeliness of monitoring result, from model failure to action, given the purpose of model monitoring)
* FFS: Power consumption
* Other KPIs are not precluded.

Note: Relevant KPIs may vary across different model monitoring approaches.FFS: Discussion of KPIs for other LCM procedures |

The related proposals/ observations are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 25: For BM-Case1 and BM-Case2, study at least the following performance metrics for AI/ML model monitoring:** *Final KPIs, e.g. throughput, RSRP, SINR, BLER*
* *Intermediate KPIs, e.g. accuracy of predicted beam ID and/or predicted RSRP*
 |
| ZTE[5] | *Observation 8: Model monitoring can be performed by comparing the predicted optimal beam with the realistic optimal beam, which is obtained by measuring one or more resource sets consisting of CSI-RS or selected SS blocks that correspond to different downlink beams at the whole beam space.**Proposal 14: To align understanding across companies, it is necessary to further clarify the model monitoring procedure and the meaning of ‘monitoring performance metric’ and ‘making decision’.**Proposal 15: For the study of model monitoring, at least the following aspects should be considered:** *Which entity is responsible for metric calculation*
* *What is the UE reporting for NW-side model and UE-side model, e.g, performance metric, intermediate results, validation of the model.*
* *Which entity can make a final decision of model selection/activation/deactivation/switching/fallback operation*
 |
| Google[7] | *Proposal 5: For spatial domain beam prediction, the beam quality for current beam from an indicated TCI can be used for performance validation, and if none of the predicted beam(s) can provide better beam quality than current beam, the predicted beam(s) are assumed to fall to pass the performance validation.**Proposal 10: For time-domain beam prediction, the beam quality for current beam from an indicated TCI can be used for performance validation, and if none of the predicted beam(s) can provide better beam quality than current beam, the predicted beam(s) are assumed to fall to pass the performance validation.**Proposal 11: Study UE feedback before the beam action time for performance validation for predicted beam in addition to the ACK/NACK for the TCI update signaling.* |
| CATT[8] | *Proposal 15: Regarding the model monitoring for BM-Case1 and BM-Case2, further study at least the following options for performance monitoring metrics/methods:** *Intermediate KPIs, e.g., the beam prediction accuracy KPIs;*
* *Eventual KPIs, e.g., throughput.*

*Proposal 16: Regarding the model monitoring for BM-Case1 and BM-Case2, the benchmark/reference for the performance comparison can be the best beam based on the baseline schemes.**Proposal 17: Regarding the model monitoring for BM-Case1 and BM-Case2, the spec impacts of following procedures based on model monitoring results should be studied, e.g., model update/switching/fallback.* |
| Sptreadtrum[9] | *Observation 3: Considering the reference for the performance comparison,** *If set A is used as the reference, UE reporting overhead may be significant.*

*- If set B is used as the reference, only part of the output results will be compared.**Proposal 7: Actual RSRP of Set A/ Set B used as the reference for the performance comparison needs to be further studied.**Proposal 8: The RSRP difference evaluated by comparing actual RSRP and predicted RSRP can be used as a performance metric.* |
| IDC[11] | *Observation 10: Definition of ‘monitoring’ in the agreement for model monitoring is not clear enough.**Proposal 13: Clarify the details of ‘monitoring for each alternative.* |
| Xiaomi[11] | *Proposal 13: gNB to transmit all beams in set A periodically/semi-persistently/ a-periodically for performance monitoring.* |
| CMCC[20] | *Proposal 7: For model inference of BM-Case1, beam prediction accuracy related KPI can be used as the metric of model performance monitoring.* |
| NVIDIA[21] | *Proposal 10: For AI/ML based beam prediction in spatial/time domain, study potential specification impact related to assistance signalling and procedure for model performance monitoring and model update/tuning.* |
| DCM[26] | *Observation 1: The following values are necessary for evaluating model accuracy.* *・Model inference results: predicted beam quality of Set A (e.g. estimated L1-RSRP of Set A)* *・Ground truth data: actual beam quality of Set A (e.g. L1-RSRP of Set A)**Proposal 2: Consider RS configuration to enable both Set A and Set B beam measurement with the following condition.**・Spatial domain beam prediction: SetA and SetB beam measurements at close time**・Temporal beam prediction: SetA and SetB beam measurements with certain prediction time offset**Observation 2: System performance can be obtained by the empirical observation and the calculation with CSI accuracy and the channel measurements.**Proposal 3: Discuss the feasibility of the model monitoring based on the input/output data distribution in the beam prediction, before the specification impact discussion related to it.* |
| SS[27] | *Proposal 13. For the performance metric(s) of AI/ML model monitoring, beam prediction accuracy related KPIs agreed in 9.2.3.1 can be considered.**Proposal 14. For benchmark/reference for the performance comparison of the AI/ML model monitoring, baseline performance options for spatial-domain beam prediction and temporal beam prediction agreed in 9.2.3.1 can be considered.* |
| MTK[30] | *Proposal 3: As model monitoring details (e.g., UE-side, NW-side and Hybrid) are generic and not particular to only Beam Management, RAN1 should not discuss such model monitoring details in AI/ML BM. Instead, these details should be discussed in Architecture/Framework Section.* |
| Rakuten[32] | *Proposal 2: Support periodic and aperiodic AI/ML model monitoring.**Proposal 3: Support UE calculating KPI based on transmissions using CSI derived from legacy method and AI/ML model.* |
| NEC[34] | *Proposal 8: Study direct and indirect* *performance metrics and methods of model monitoring, e.g., direct performance metric and method based on intermediate KPIs, indirect performance metric and method based on input data.* |
|  |  |
|  |  |

Based on the tdocs and the discussion of the last meeting, several companies suggest to clarify what “monitor the performance metric” means. Thus, the following proposals is suggested for discussion.

###### Conclusion 7.1.1

***Conclusion 7.1.1: Regarding the previous agreements for the AI/ML model monitoring for BM-Case1 and BM-Case2, “monitor the performance metric(s)” refers to “calculate/measure the performance metric(s)”***

* ***The calculation of the performance metric(s) may be based on the information reported by the other side***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |

There are some detailed proposals for the performance metrics and the benchmark/reference for performance comparison in the submitted tdocs. Thus, the following two proposals are suggested for discussion.

###### Proposal 7.1.2

***Proposal 7.1.2: Regarding the performance metric(s) of AI/ML model monitoring for BM-Case1 and BM-Case2, study the following alternatives as a starting point:***

* ***Alt.1: Beam prediction accuracy related KPIs, e.g., Top-K/1 beam prediction accuracy***
* ***Alt.2: Link quality related KPIs, e.g., throughput, RSRP, SINR***
* ***Alt.3: Performance metric based on input/output data distribution of AI/ML***
* ***Alt.4: The RSRP difference evaluated by comparing actual RSRP and predicted RSRP***
* ***Other alternatives are not precluded***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |

###### Proposal 7.1.3

***Proposal 7.1.3: For AI/ML model monitoring for BM-Case1 and BM-Case2, study the following alternatives for the benchmark/reference for performance comparison as a starting point:***

* ***Alt.1: The best beam(s) obtained by measuring beams of a set indicated by gNB (e.g., Beams of Set A)***
* ***Alt.2: The best beam(s) among those used for AI/ML model inputs (e.g., Beams of Set B)***
* ***Alt.3: The beam corresponding to some indicated TCI state(s)***
* ***Other alternatives are not precluded***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |

## NW-side model

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110bis-e**AgreementFor BM-Case1 and BM-Case2 with a network-side AI/ML model, study the NW-side model monitoring:* NW monitors the performance metric(s) and makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation

AgreementRegarding NW-side model monitoring for a network-side AI/ML model of BM-Case1 and BM-Case2, study the potential specification impacts from the following aspects* Beam measurement and report for model monitoring
* Note: This may or may not have specification impact.
 |

|  |  |
| --- | --- |
| Vivo[3] | *Proposal 24: For a network-side AI model with enhanced beam pair prediction, at least study the following aspects for potential specification impact in monitoring procedure:**• Enhanced resource configuration**- Specific beam pair resource configuration**• Enhanced beam report**- All measured L1-RSRP should be reported to gNB as labels**- Relative Rx beam information as assistance information may be reported**- Considering report overhead reduction, such as enhanced L1-RSRP quantization method, reducing unnecessary L1-RSRPs in beam report where the omitted L1-RSRP may be lower than a pre-defined threshold**Proposal 25: For a network-side AI model with DL Tx beam prediction, at least study the following aspects for potential specification impact in model monitoring procedure:**• Enhanced resource configuration**- P3+P2 resource configuration that Rx beam assumption of P2 resource measurement is the best Rx beam searched from P3 procedure**• Enhanced beam report**- Only P2 measured L1-RSRP should be reported to gNB as labels**- Best Rx beam information as assistance information may be reported**- Considering report overhead reduction, such as enhanced L1-RSRP quantization method, reducing unnecessary L1-RSRP reporting where the omitted L1-RSRPs may be lower than a pre-defined threshold* |
| ZTE[5] | *Proposal 16: For BM-Case1 and BM-Case2 with a network-side AI/ML model, NW CALCULATES the performance metric and makes decision of model selection/activation/ deactivation/switching/ fallback operation.* |
| Ericsson[10] | *Proposal 22 Study both an RRC-message based and L1 fast CSI reporting-based data collection methods for model monitoring for NW-sided beam prediction use cases.* |
| xiaomi [12] | *Proposal 19: For NW-side model monitoring for network-side AI/ML model, support to report both set B and set C, where set B will be used as network-side AI/ML model input, and set C consists of Top-K beams by UE measurement of set A.* *Proposal 20: Study potential signaling for NW-side initiated performance monitoring for NW-side AI/ML model.* |
| OPPO[14] | *Observation 4: For BM-Case1 and BM-Case2 with a network-side AI/ML model, there may be no additional specification impact on LCM and increased overhead for DL measurement and reporting.* |
| CMCC[20] | *Proposal 8: For BM-Case1 with a NW-side AI/ML model, study the following mechanism for model monitoring:** *Atl1. NW-side Model monitoring*
	+ *NW monitors the performance metric(s)*
	+ *NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation*
 |
| Lenovo[22] | *Proposal 7: NW-side model monitoring is preferred for NW-side AI/ML inference, and the Rel-15 beam report procedure can be reused with necessary enhancements.* |
| LG[24] | *Proposal #14: For NW-sided model monitoring, UE reporting based on Set A can be considered but with an assumption that measurement and reporting on Set A or potential beams of Set A shall happen rarely.* |
| CIACT[25] | *Proposal 7: For NW-side AI/ML model monitoring, system performance related KPI should be considered.* |
| DCM[26] | *Proposal 7: Study the overhead reduction of L1 signalling to report SetA beam measurements for NW-based model monitoring.* |
| SS[27] | *Proposal 15. For the AI/ML model monitoring, in the case that AI/ML model is at gNB-side, the following aspect can be further study on beam measurement and report for model monitoring:** UE to report the measurement results of more than 4 beams in one reporting instance* |
| Nokia[33] | *Proposal 17: For the NW-sided beam prediction, further study the model monitoring considering,* * *NW-side model monitoring as the baseline*
* *NW monitors the performance metric(s) based on frequent measurement/reporting of Set A beams*
 |
|  |  |
|  |  |

Several tdocs proposed some discussion points for the NW-side model monitoring for a network-side AI/ML model. The following proposal is tried to list the afore-mentioned discussion points.

###### Proposal 7.2

***Proposal 7.2: Regarding NW-side model monitoring for a network-side AI/ML model of BM-Case1 and BM-Case2, study the potential specification impacts from the following aspects:***

* ***UE reporting of measurement based on Set A***
* ***UE to report the measurement results of more than 4 beams in one reporting***
* ***Signaling, e.g., RRC-based, L1-based***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |

## UE-side model

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110bis-e**AgreementFor BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the following alternatives for model monitoring with potential down-selection: * Atl1. UE-side Model monitoring
	+ UE monitors the performance metric(s)
	+ UE makes decision(s) of model selection/activation/ deactivation/switching/fallback operation
* Atl2. NW-side Model monitoring
	+ NW monitors the performance metric(s)
	+ NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation
* Alt3. Hybrid model monitoring
	+ UE monitors the performance metric(s)
	+ NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation
 |

|  |  |
| --- | --- |
| Huawei[2] | *Observation 9: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, for Alt.1 UE-side Model monitoring,* * *The UE may not be aware of all aspects impacting the AI/ML model operation.*
* *NW may suffer unknown performance fluctuation if the UE autonomously performs the model adaption without notifying the NW.*

*Proposal 27: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, for Alt.1 UE-side Model monitoring, the UE should report the decision to the NW, and then the NW could indicate the UE a corresponding execution of the decision.* |
| Vivo[3] | *Proposal 26: For a UE-side AI model with enhanced beam pair prediction, at least study the following aspects for potential specification impact in data collection procedure:**• Enhanced request signaling* *- UE may request beam pair sweeping resources for model monitoring purpose including the number of requested labels, and potentially some associated triggering events to be defined**- UE may request the minimum number of beams in advance or with resource request* *- UE may request Tx beam information that should be signaled from gNB**• Enhanced resource configuration**- Specific beam pair resource configuration**- Tx beam information (such as ID or angle) may be signaled with resource configuration**• New signaling**- Monitoring result report**- Monitoring indication from network for further model management decision**Proposal 27: For a UE-side AI model with DL Tx beam prediction, at least study the following aspects for potential specification impact in data collection procedure:**• Enhanced request signaling* *- UE may request P3+P2 beam sweeping resources for model monitoring purpose including the number of requested labels, and potentially some associated triggering events to be defined**- UE may request the minimum number of beams in advance or with P2+P3 resource request**- UE may request Tx beam information that should be signaled from gNB**• Enhanced resource configuration**- P3 + P2 resource configuration that Rx beam assumption of P2 resource measurement is the best Rx beam searched from P3 procedure.**- Tx beam information (such as ID or angle) may be signaled with resource configuration**• New signaling**- Monitoring result report**- Monitoring indication from network for further model management decision* |
| H3C[4] | *Proposal 4:* *For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the following alternatives for model monitoring with potential down-selection:* * *Atl1. UE-side Model monitoring*
	+ *UE monitors the performance metric(s)*
	+ *UE makes decision(s) of model selection/activation/ deactivation/switching/fallback operation*
* *Atl2. NW-side Model monitoring*
	+ *NW monitors the performance metric(s)*
	+ *NW makes decision(s) of model activation/ deactivation operation*
* *Alt3. Hybrid model monitoring*
	+ *UE monitors the performance metric(s)*
	+ *NW makes decision(s) of model activation/ deactivation/ operation*

*Proposal 5: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the potential specification impact of L1 signalling to trigger SetA beam test for performance monitoring.* |
| ZTE[5] | *Proposal 17: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, support to study both the NW-side model monitoring in Alt.2 and hybrid model monitoring in Alt.3.* |
| Fujitsu[6] | *Proposal 8: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, the Alt.2 and Alt.3 is suggested to have high priority for further study of model monitoring.* |
| Ericsson[10] | *Proposal 23 Study performance metrics comprising both per-sample prediction error and statistical metrics**Proposal 24 For BM-Case1 and BM-Case2 with a UE-side AI/ML model, prioritize Alt 3 for model monitoring (hybrid model monitoring)* |
| Xiaomi[12] | *Proposal 14: For UE-side AI/ML model with UE-side model monitoring, support UE to indicate the decision to NW.**Proposal 15: For UE-side AI/ML model with NW-side model monitoring, support UE to calculate the performance metric and report to NW.**Proposal 16: For UE-side AI/ML model with hybrid model monitoring, support an event-triggered report of performance metric from UE based on a threshold configured by gNB.**Proposal 17: Threshold of beam prediction accuracy related KPIs can be used for performance monitoring.**Proposal 18: Study potential signaling for both NW-side initiated and UE-side initiated performance monitoring for UE-side AI/ML model.* |
| Intel[13] | *Proposal 4: For UE-side AI/ML model, support UE-side and Hybrid model monitoring. NW-side monitoring can be further studied and used in specific cases if the model is transferred from the NW to the UE.* |
| OPPO[14] | *Proposal 17: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the (Alt1) UE-side model monitoring as a starting point.* |
| ETRI[17] | *Proposal 3. For BM case-1 and BM-case 2 with a UE-side AI/ML model, UE can monitor the performance metrics. Since UE decision or network decision may have different needs depending on operating conditions, UE decision or network decision can be required depending on operating methods.* |
| Lenovo[22] | *Proposal 8: For UE-side AI/ML inference, support NW-side Model monitoring as well as Hybrid model monitoring as follows:** Alt2. NW-side Model monitoring**○ NW monitors the performance metric(s)* *○ NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation** Alt3. Hybrid model monitoring**○ UE monitors the performance metric(s)* *○ NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation* |
| LG[24] | *Proposal #13: For UE-sided AI/ML model, Alt1(UE-side model monitoring) should be supported.* |
| CIACT[25] | *Proposal 6: For UE-side AI/ML model monitoring, UE side directly monitoring (Alt.1) should be baseline.* |
| DCM[26] | *Proposal 8: NW should control which UE side model to be activated or deactivated based the NW operation in beam prediction.**Proposal 11: Hybrid model monitoring should be prioritized than NW side model monitoring in terms of signalling overhead, when model monitoring is based on model accuracy.* |
| SS[27] | *Proposal 16. For BM-Case1 and BM-Case2 with a UE-side AI/ML model, Alt2 (i.e., NW-side model monitoring) and Alt3 (Hybrid model monitoring) are preferred.* |
| QC[28] | *Proposal 7: For BM-Case1 and BM-Case2, study the signalling aspects related to gNB sending assistance signalling to help UE in comparing predicted measurements with actual measurements.**• This assistance signalling can be in the form of auxiliary reference signals.* |
| KT[29] | *Proposal 4. For model monitoring with a UE-side AI/ML model, Alt 1. UE-side model monitoring and Alt 3. Hybrid model monitoring are preferred.**Proposal 5. Regarding model monitoring for a UE-side AI/ML model of BM-Case1 and BM-Case2, study the potential specification impacts from the following aspects**• Monitoring beam and corresponding resource configuration**• Trigger conditions for monitoring and/or reporting**• Reporting parameter(s)**Proposal 5. Regarding model monitoring for a UE-side AI/ML model of BM-Case1 and BM-Case2, study the potential specification impacts from the following aspects**• Monitoring beam and corresponding resource configuration**• Trigger conditions for monitoring and/or reporting**• Reporting parameter(s)* |
| MTK[30] | *Proposal 1: Regarding specification impact of L1 signalling in BM-Case1 with a UE-side AI/ML model, RAN1 will study and decide on the predicted L1-RSRP corresponding to the beams, reported by the UE.**Proposal 2: Regarding specification impact of L1 signalling in BM-Case2 with a UE-side AI/ML model, RAN1 will study and find out a process to estimate the value of N, corresponding to the beam(s) of N future time instance(s), reported by the UE, based on its output of AI/ML model inference.* |
| Nokia[33] | *Proposal 16: For BM-Case1 and BM-Case2 with a UE-side AI/ML model, select the following alternatives for model monitoring:* * *Atl2. NW-side Model monitoring*
	+ *NW monitors the performance metric(s)*
	+ *NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation*
* *Alt3. Hybrid model monitoring*
	+ *UE monitors the performance metric(s)*
	+ *NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation*

*Proposal 18: For the UE-sided beam prediction, further study the model monitoring considering,* * *Hybrid model monitoring as the baseline*
* *UE monitors the performance metric(s) based on frequent measurement/reporting of Set A beams and reports the metrics to the NW*
 |
|  |  |
|  |  |

The preferences on the three alternatives are quite diverging among companies. One possible way is to investigate more details and pros/cons of these alternative so that we can have better understanding and then decide the potential down-selection/prioritization (if any).

###### Proposal 7.3.1

***Proposal 7.3.1:******For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the following aspects for Atl1. UE-side model monitoring as a starting point:***

* ***Signaling from gNB to facilitate UE for performance monitoring (e.g., dedicated RS configuration for measurement)***
* ***Whether/how UE will report the decision to NW***
* ***UE performance and/or system performance***
* ***Other aspect(s) is not precluded***
* ***Note: The potential down-selection/ prioritization (if any) of these three alternatives is a separate discussion.***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |

###### Proposal 7.3.2

***Proposal 7.3.2:******For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the following aspects for Atl2. NW-side model monitoring as a starting point:***

* ***Signaling from gNB to facilitate UE for the corresponding measurement (if applicable)***
* ***What UE reporting needed to facilitate NW to calculate the performance metric***
* ***UE performance and/or system performance***
* ***Other aspect(s) is not precluded***
* ***Note: The potential down-selection/ prioritization (if any) of these three alternatives is a separate discussion.***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |

###### Proposal 7.3.3

***Proposal 7.3.3:******For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the following aspects for Atl3. Hybrid model monitoring as a starting point:***

* ***Signaling from gNB to facilitate UE for performance monitoring (e.g., dedicated RS configuration for measurement)***
* ***UE reporting mechanism to NW***
* ***UE performance and/or system performance***
* ***Other aspect(s) is not precluded***
* ***Note: The potential down-selection/ prioritization (if any) of these three alternatives is a separate discussion.***

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |

# Spec impact of Model selection, activation, deactivation, switching, and fallback operation

In previous RAN1 meeting(s), the following agreements were made:

|  |
| --- |
| **RAN1#110bis-e**AgreementFor model selection, activation, deactivation, switching, and fallback at least for UE sided models and two-sided models, study the following mechanisms:* Decision by the network
	+ Network-initiated
	+ UE-initiated, requested to the network
* Decision by the UE
	+ Event-triggered as configured by the network, UE’s decision is reported to network
	+ UE-autonomous, UE’s decision is reported to the network
	+ UE-autonomous, UE’s decision is not reported to the network

FFS: for network sided modelsFFS: other mechanismsAgreementStudy the specification impact to support multiple AI models for the same functionality, at least including the following aspects:* Procedure and assistance signaling for the AI model switching and/or selection

FFS: Model selection refers to the selection of an AI/ML model among models for the same functionality. (Exact terminology to be discussed/defined) |

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 26: For the spec impact of model monitoring, RAN1 studies the co-existence and fallback mechanisms between AI/ML-based beam prediction approach and legacy non-AI/ML-based beam management approach.* |
| Fujitsu[6] | *Proposal 9: Study the potential specification impacts on the ping-pong decision for the model selection/activation/deactivation/switching/fallback operation.* |
| NVIDIA[21] | *Proposal 9: For AI/ML based beam prediction in spatial/time domain, study potential specification impact related to assistance signalling and procedure for model configuration, model activation/deactivation, model recovery/termination, and model selection.* |
| Lenovo[22] | *Proposal 12: Dynamic switching between AI/ML based beam prediction and non-AI/ML based beam report schemes as well as dynamic switching between different AI/ML models should be supported.* |
| Nokia[33] | *Proposal 19: RAN1 to study NW input to support ML model switching/(de)activation at UE for DL Tx beam or DL Tx-Rx beam pair prediction.* |
| NEC[34] | *Observation 2: For a sub use case (or the same functionality), multiple AI/ML models corresponding to different scenarios (e.g., LOS, NLOS) or different inputs/outputs may be arranged.**Proposal 7: Study the mechanism of model selection, e.g.,* *the following information should be exchanged between gNB and UE.** *Information related to multiple AI/ML models.*
* *Information indicating the selected AI/ML model.*
 |
|  |  |

**Moderator Note**: The above proposals seem quite general and applicable to all sub use cases (e.g., CSI compression, positioning accuracy enhancement). Not sure what spec impact is specific to BM use cases based on the above proposals.

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |

# UE/NW Capability

The related proposals/ observations are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 6: Regarding AI/ML-based beam management, study the standards impact, including AI/ML related UE configuration/capability reporting, related to AI/ML model selection/configuration (like activation/deactivation) in case multiple trained AI/ML models are deployed.* |
| Huawei[2] | *Proposal 30: Study the potential specification impact for UE capability, including the following aspects as a starting point:* * *Data collection, model training, inference latency, monitoring, models switching, model updating.*
* *Details can be discussed until further progress has been made for schemes themselves and their related spec impact.*
 |
| OPPO[14] | *Proposal 18: For BM-Case1 and BM-Case2, consider the UE capability on AI/ML beam prediction when stable.* |
| NVIDIA[21] | *Proposal 13: For AI/ML based beam prediction in spatial/time domain, study potential specification impact related to UE capability for AI/ML based beam prediction including model training, model inference and model monitoring.* |
| Lenovo[22] | *Proposal 5: Study UE/NW capability related signaling corresponding to AI/ML-based beam management under different network-UE collaboration levels.* |
|  |  |

**Mod recommendation**: TBD

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Other aspects of LCM

In previous RAN1 meeting(s), the agreement(s)/conclusion(s) were made as below:

|  |
| --- |
| **RAN1#110**Agreement Study the following aspects, including the definition of components (if needed) and necessity, in Life Cycle Management* Data collection
	+ Note: This also includes associated assistance information, if applicable.
* Model training
* [Model registration]
* Model deployment
	+ Note: Terminology is to be defined. ~~This includes process of compiling a trained AI/ML model and packaging it into an executable format and delivering to a target device.~~
* [Model configuration]
* Model inference operation
* Model selection, activation, deactivation, switching, and fallback operation
	+ ~~Note: some of them to be refined~~
* Model monitoring
* Model update
	+ Note: Terminology is to be defined. This includes model finetuning, retraining, and re-development via online/offline training.
* Model transfer
* UE capability

Note: Some aspects in the list may not have specification impact.Note: Aspects with square brackets are tentative ~~and pending terminology definition~~.Note: More aspects may be added as study progresses. **RAN1#110bis-e**AgreementStudy LCM procedure on the basis that an AI/ML model has a model ID with associated information and/or model functionality at least for some AI/ML operations ~~when network needs to be aware of UE AI/ML models~~* FFS: Detailed discussion of model ID with associated information and/or model functionality.
* FFS: usage of model ID with associated information and/or model functionality based LCM procedure
* FFS: whether support of model ID
* FFS: the detailed applicable AI/ML operations

AgreementStudy various approaches for achieving good performance across different scenarios/configurations/sites, including* Model generalization, i.e., using one model that is generalizable to different scenarios/configurations/sites
* Model switching, i.e., switching among a group of models where each model is for a particular scenario/configuration/site
	+ [Models in a group of models may have varying model structures, share a common model structure, or partially share a common sub-structure. Models in a group of models may have different input/output format and/or different pre-/post-processing.]
* Model update, i.e., using one model whose parameters are flexibly updated as the scenario/configuration/site that the device experiences changes over time. Fine-tuning is one example.
 |

The related proposals/ observations are copied as below:

|  |  |
| --- | --- |
| Google[7] | *Proposal 12: For AI/ML based BM, the study should be based on both Rel-17 unified TCI framework and Rel-15/Rel-16 BM framework.**Proposal 13: The study of AI/ML based BM should consider both FR1 and FR2.* |
|  |  |

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Model registration

|  |  |
| --- | --- |
| CATT[8] | *Proposal 9: Regarding BM-Case1 and BM-Case2, for model registration, study the following aspects as a starting point for model description information which UE should provide to gNB:** *Model functionality, e.g., BM-Case1/BM-Case2 or DL beam pair/Tx beam prediction;*
* *Information of model inputs, e.g., the number of DL Tx beams or beam pairs in Set B;*
* *Information of model outputs, e.g., the number of predicted beam and/or L1-RSRP;*
* *Information on assistance information for inference,*
* *Information on model performance,*
* *Information on co-existence of other AI/ML models and/or non-AI/ML features,*
* *Information on applicable scenarios/configurations/sites.*

*Proposal 10: For model registration, further study the following aspects on the model ID of a registered AI/ML model:** *Whether the model ID is reported by UE or assigned by network;*
* *Whether the model ID is explicit or implicit.*
 |
|  |  |

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |

## Training

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 21: For UE-side model, study the UE report of training-required information, including the size of the needed training data samples, the supported configurations of Set A and/or Set B, the supported values of Top-K, etc.* |
|  |  |

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |
|  |  |

## Model transfer

|  |  |
| --- | --- |
| Vivo[3] | *Proposal 8: For case 1 and case 2 of beam management, both collaboration level level-y, and collaboration level-z can be considered.**Proposal 9: Take the following supportable model update choices as one aspect for defining model update levels of beam management.**- Choice 0: No model update during lifecycle management**- Choice 1: Updating model parameter or structure w/o model transfer**- Choice 2: Updating model parameter or structure with model transfer**- Study the lifecycle management signaling and procedures for each of the collaboration levels and model updating choices.**Proposal 22: In model inference procedure, Alt.3, i.e. model training at NW side and model inference at UE side, with enhanced beam pair prediction and DL Tx beam prediction scheme has similar specification impacts as an AI model trained and inferenced at UE side.**Proposal 23: Study signaling aspects enhancement related to the procedure of model transfer, model registration and model activation, for the case with AI/ML model training at NW side and AI/ML model inference at UE side.* |
| CATT[8] | *Proposal 18: Regarding the model transfer, the following aspects can be further studied in RAN1:** *Full or partial model transfer;*
* *Data size of model transfer;*
* *Model transfer frequency for model deployment/update;*
* *Latency and reliability requirements for model transfer;*
* *Model delivery format for model transfer, e.g., ONNX or 3GPP-standadized model representation format.*

*Proposal 19: Regarding the model transfer, the signaling and model representation format can be further studied in RAN2 based on RAN1 progress.* |
| Spreadtrum[9] | *Observation 1: If AI/ML training is at NW side while AI/ML inference is at UE side, signaling related to AI/ML transfer should be defined.* |
| CIACT[25] | *Proposal 5: If AI model transfer from NW side to UE side is supported, AI model transfer over air interface should be specified.* |
|  |  |

**Moderator recommendation**: Please refer to Section 2.1. The discussion on spec impacts of model transfer is deferred at the current stage until RAN1/RAN2 has more progress.

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |

## Generalization

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 7: Study Standards impact related to supporting model generalization across scenarios and/or configurations.* |
| OPPO[14] | *Proposal 19: For BM-Case1 and BM-Case2, study enhancement on generalization of AI/ML model (if necessary) under heterogeneous scenarios and different Tx and/or Rx beam configurations.* |
|  |  |

|  |  |
| --- | --- |
| Company | Comments |
|  |  |
|  |  |

# Summary of Discussion

## Proposal for 1st GTW

# Reference

1. R1-2210844 Continued discussion on other aspects of AI/ML for beam management FUTUREWEI
2. R1-2210888 Discussion on AI/ML for beam management Huawei, HiSilicon
3. R1-2211001 Other aspects on AI/ML for beam management vivo
4. R1-2211038 Discussion on other aspects of AI/ML beam management New H3C Technologies Co., Ltd.
5. R1-2211060 Discussion on other aspects for AI beam management ZTE
6. R1-2211076 Sub use cases and specification impact on AI/ML for beam management Fujitsu
7. R1-2211127 On Enhancement of AI/ML based Beam Management Google
8. R1-2211192 Other aspects on AI/ML for beam management CATT
9. R1-2211230 Discussion on other aspects on AIML for beam management Spreadtrum Communications
10. R1-2211289 Discussion on AI/ML for beam management Ericsson
11. R1-2211316 Discussion for other aspects on AI/ML for beam management InterDigital, Inc.
12. R1-2211358 Potential specification impact on AI/ML for beam management xiaomi
13. R1-2211396 Use-cases and Specification Impact for AI/ML beam management Intel Corporation
14. R1-2211481 Other aspects of AI/ML for beam management OPPO
15. R1-2211510 Discussions on Sub-Use Cases in AI/ML for Beam Management TCL Communication
16. R1-2211528 Other aspects on AI/ML for beam management China Telecom
17. R1-2211558 Discussion on other aspects on AI/ML for beam management ETRI
18. R1-2211590 Discussion on sub use cases of AI/ML beam management Panasonic
19. R1-2211608 Consideration on AI/ML for beam management Sony
20. R1-2211675 Discussion on other aspects on AI/ML for beam management CMCC
21. R1-2211721 AI and ML for beam management NVIDIA
22. R1-2211776 Further aspects of AI/ML for beam management Lenovo
23. R1-2211808 Discussion on other aspects of AI/ML for beam management Apple
24. R1-2211870 Other aspects on AI/ML for beam management LG Electronics
25. R1-2211914 Discussions on AI-ML for Beam management CAICT
26. R1-2211980 Discussion on AI/ML for beam management NTT DOCOMO, INC.
27. R1-2212039 Representative sub use cases for beam management Samsung
28. R1-2212111 Other aspects on AI/ML for beam management Qualcomm Incorporated
29. R1-2212150 Discussion on other aspects on AI/ML for beam management KT Corp.
30. R1-2212229 Other aspects on AI/ML for beam management MediaTek Inc.
31. R1-2212292 Continued Discussion on Performance Related Aspects of Codebook Enhancement with AI/ML Charter Communications, Inc
32. R1-2212320 Other aspects on AI/ML for beam management Rakuten Symphony
33. R1-2212330 Other aspects on ML for beam management Nokia, Nokia Shanghai Bell
34. R1-2212372 Discussion on AI/ML for beam management NEC

# Appendix A: Contact Information

The following information was collected in the last meeting(s). Please feel free to update/correct the contact information if needed.

|  |  |  |
| --- | --- | --- |
| Company | Name | Email |
| Moderator | Zhihua SHI | szh@oppo.com |
| Apple | Weidong Yang | Wyang23@apple.com |
| vivo | Peng SUN | sunpeng@vivo.com |
| AT&T | Thomas Novlan | thomas\_novlan@labs.att.com |
| Futurewei | Chunhui Zhu | czhu@futurewei.com |
| Xiaomi | Mingju Li | limingju@xiaomi.com |
| Lenovo | Bingchao LIU | liubc2@lenovo.com  |
| Sony | Chen SUN | Chen.sun@sony.com |
| Huawei, HiSilicon | Thorsten Schier | thorsten.schier@huawei.com |
| NEC | Zhen He | he\_zhen@nec.cn |
| LG Electronics | Jiwon KangHaewook Park | jw.kang@lge.comhaewook.park@lge.com |
| Panasonic | Quan Kuang | quan.kuang@eu.panasonic.com |
| Ericsson | Henrik Ryden | Henrik.a.ryden@ericsson.com |
| Nokia, NSB | Keeth JayasingheMihai Enescu | keeth.jayasinghe@nokia.com, mihai.enescu@nokia.com |
| CATT | Yongqiang FEI | feiyongqiang@catt.cn |
| Fujitsu | Wang Xin | wangxin@fujitsu.com |
| Samsung | Zhe Chen | tom.chenzhe@samsung.com |
| CMCC | Jun ZUO | zuojun@chinamobile.com |
| NVIDIA | Xingqin Lin | xingqinl@nvidia.com |
| CAICT | Xiaofeng Liu | Liuxiaofeng1@caict.ac.cn |
| OPPO | Jianfei Cao | caojianfei@oppo.com |
| MediaTek | Gyu Bum KyungYu-Jen Ku | gyubum.kyung@mediatek.comyu-jen.ku@mediatek.com |
| Intel | Avik Sengupta | avik.sengupta@intel.com |
| NTT DOCOMO | Haruhi Echigo | haruhi.echigo.fw@nttdocomo.com |
| Beijing Jiaotong University (BJTU) | Wei Chen | weich@bjtu.edu.cn |
| ZTE, Sanechips | Wenfeng Liu | liu.wenfeng@zte.com.cn |
| InterDigital | Youngwoo Kwak | youngwoo.kwak@interdigital.com |
| Qualcomm | Hamed Pezeshki | hamedp@qti.qualcomm.com |
| Spreadtrum | Dawei Ma | dawei.ma@unisoc.com |
| Charter Communications | Dumitru M. IonescuSamer Henry | dumitru.ionescu@charter.comC-Samer.Henry@charter.com |
|  |  |  |

# Appendix B: Agreements

## RAN1#111

## RAN1#110bis-e

Conclusion

For AI/ML based beam management, RAN1 has no consensus to support on studying any other sub use case in addition to BM-Case1 and BM-Case2

Note: this conclusion is independent of the discussion on the alternatives of AI/ML model inputs for BM-Case1 and BM-Case2

Conclusion

For the sub use case BM-Case1 and BM-Case2, Set B is a set of beams whose measurements are taken as inputs of the AI/ML model

***Agreement***

***For BM-Case1 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to report the following information of AI/ML model inference to NW***

* ***The beam(s) that is based on the output of AI/ML model inference***
* ***FFS: Predicted L1-RSRP corresponding to the beam(s)***
* ***FFS: other information***

***Agreement***

***For BM-Case2 with a UE-side AI/ML model, study the potential specification impact of L1 signaling to report the following information of AI/ML model inference to NW***

* ***The beam(s)*** ***of N future time instance(s) that is based on the output of AI/ML model inference***
	+ ***FFS: value of N***
* ***FFS: Predicted L1-RSRP corresponding to the beam(s)***
* ***Information about the timestamp corresponding the reported beam(s)***
	+ ***FFS: explicit or implicit***
* ***FFS: other information***

***Working Assumption***

***For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the following L1 beam reporting enhancement for AI/ML model inference***

* ***UE to report the measurement results of more than 4 beams in one reporting instance***
* ***Other L1 reporting enhancements can be considered***

***Agreement***

***For BM-Case1 and BM-Case2 with a network-side AI/ML model, study the NW-side model monitoring:***

* ***NW monitors the performance metric(s) and makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation***

***Agreement***

***Regarding NW-side model monitoring for a network-side AI/ML model of BM-Case1 and BM-Case2, study the potential specification impacts from the following aspects***

* ***Beam measurement and report for model monitoring***
* ***Note: This may or may not have specification impact.***

***Agreement***

***For BM-Case1 and BM-Case2 with a UE-side AI/ML model, study the following alternatives for model monitoring with potential down-selection:***

* ***Atl1. UE-side Model monitoring***
	+ ***UE monitors the performance metric(s)***
	+ ***UE makes decision(s) of model selection/activation/ deactivation/switching/fallback operation***
* ***Atl2. NW-side Model monitoring***
	+ ***NW monitors the performance metric(s)***
	+ ***NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation***
* ***Alt3. Hybrid model monitoring***
	+ ***UE monitors the performance metric(s)***
	+ ***NW makes decision(s) of model selection/activation/ deactivation/switching/ fallback operation***

## RAN1#110

Agreement

For the sub use case BM-Case1, support the following alternatives for further study:

* Alt.1: Set A and Set B are different (Set B is NOT a subset of Set A)
* Alt.2: Set B is a subset of Set A
* Note1: Set A is for DL beam prediction and Set B is for DL beam measurement.
* Note2: The beam patterns of Set A and Set B can be clarified by the companies.

Agreement

For the data collection for AI/ML model training (if supported), study the following aspects as a starting point for potential necessary specification impact:

* Signaling/configuration/measurement/report for data collection, e.g., signaling aspects related to assistance information (if supported), Reference signals
* Content/type of the collected data
* Other aspect(s) is not precluded

Agreement

At least for the sub use case BM-Case1 and BM-Case2, support both Alt.1 and Alt.2 for the study of AI/ML model training:

* Alt.1: AI/ML model training at NW side;
* Alt.2: AI/ML model training at UE side.

Note: Whether it is online or offline training is a separate discussion.

Agreement

For the sub use case BM-Case1 and BM-Case2, further study the following alternatives for the predicted beams:

* Alt.1: DL Tx beam prediction
* Alt.2: DL Rx beam prediction
* Alt.3: Beam pair prediction (a beam pair consists of a DL Tx beam and a corresponding DL Rx beam)
* Note1: DL Rx beam prediction may or may not have spec impact

Agreement

For the sub use case BM-Case2, further study the following alternatives:

* Alt.1: Set A and Set B are different (Set B is NOT a subset of Set A)
* Alt.2: Set B is a subset of Set A (Set A and Set B are not the same)
* Alt.3: Set A and Set B are the same
* Note1: The beam pattern of Set A and Set B can be clarified by the companies.

Agreement

Regarding the model monitoring for BM-Case1 and BM-Case2, to investigate specification impacts from the following aspects

* Performance metric(s)
* Benchmark/reference for the performance comparison
* Signaling/configuration/measurement/report for model monitoring, e.g., signaling aspects related to assistance information (if supported), Reference signals
* Other aspect(s) is not precluded

Agreement

In order to facilitate the AI/ML model inference, study the following aspects as a starting point:

* Enhanced or new configurations/UE reporting/UE measurement, e.g., Enhanced or new beam measurement and/or beam reporting
* Enhanced or new signaling for measurement configuration/triggering
* Signaling of assistance information (if applicable)
* Other aspect(s) is not precluded

Agreement

Regarding the sub use case BM-Case1 and BM-Case2, study the following alternatives for AI/ML output:

* Alt.1: Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams
	+ E.g., N predicted beams can be the top-N predicted beams
* Alt.2: Tx and/or Rx Beam ID(s) of the N predicted DL Tx and/or Rx beams and other information
	+ FFS: other information (e.g., probability for the beam to be the best beam, the associated confidence, beam application time/dwelling time, Predicted Beam failure)
	+ E.g., N predicted beams can be the top-N predicted beams
* Alt.3: Tx and/or Rx Beam angle(s) and/or the predicted L1-RSRP of the N predicted DL Tx and/or Rx beams
	+ E.g., N predicted beams can be the top-N predicted beams
	+ FFS: details of Beam angle(s)
* FFS: how to select the N DL Tx and/or Rx beams (e.g., L1-RSRP higher than a threshold, a sum probability of being the best beams higher than a threshold, RSRP corresponding to the expected Tx and/or Rx beam direction(s))
* Note1: It is up to companies to provide other alternative(s)
* Note2: Beam ID is only used for discussion purpose
* Note3: All the outputs are “nominal” and only for discussion purpose
* Note4: Values of N is up to each company.
* Note5: All of the outputs in the above alternatives may vary based on whether the AI/ML model inference is at UE side or gNB side.
* Note 6: The Top-N beam IDs might have been derived via post-processing of the ML-model output

## RAN1#109-e

Agreement

For AI/ML-based beam management, support BM-Case1 and BM-Case2 for characterization and baseline performance evaluations

* BM-Case1: Spatial-domain DL beam prediction for Set A of beams based on measurement results of Set B of beams
* BM-Case2: Temporal DL beam prediction for Set A of beams based on the historic measurement results of Set B of beams
* FFS: details of BM-Case1 and BM-Case2
* FFS: other sub use cases

Note: For BM-Case1 and BM-Case2, Beams in Set A and Set B can be in the same Frequency Range

Agreement

Regarding the sub use case BM-Case2, the measurement results of K (K>=1) latest measurement instances are used for AI/ML model input:

* The value of K is up to companies

Agreement

Regarding the sub use case BM-Case2, AI/ML model output should be F predictions for F future time instances, where each prediction is for each time instance.

* At least F = 1
* The other value(s) of F is up to companies

Agreement

For the sub use case BM-Case1, consider both Alt.1 and Alt.2 for further study:

* Alt.1: AI/ML inference at NW side
* Alt.2: AI/ML inference at UE side

Agreement

For the sub use case BM-Case2, consider both Alt.1 and Alt.2 for further study:

* Alt.1: AI/ML inference at NW side
* Alt.2: AI/ML inference at UE side

Conclusion

For the sub use case BM-Case1, consider the following alternatives for further study:

* Alt.1: Set B is a subset of Set A
	+ FFS: the number of beams in Set A and B
	+ FFS: how to determine Set B out of the beams in Set A (e.g., fixed pattern, random pattern, …)
* Alt.2: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)
	+ FFS: the number of beams in Set A and B
	+ FFS: QCL relation between beams in Set A and beams in Set B
	+ ~~FFS: construction of Set B (e.g., regular pre-defined codebook, codebook other than regular pre-defined one)~~
* Note1: Set A is for DL beam prediction and Set B is for DL beam measurement.
* Note2: The narrow and wide beam terminology is for SI discussion only and have no specification impact
* Note3: The codebook constructions of Set A and Set B can be clarified by the companies.

Conclusion

Regarding the sub use case BM-Case1, further study the following alternatives for AI/ML input:

* Alt.1: Only L1-RSRP measurement based on Set B
* Alt.2: L1-RSRP measurement based on Set B and assistance information
	+ FFS: Assistance information. The following were mentioned by companions in the discussion:  Tx and/or Rx beam shape information (e.g., Tx and/or Rx beam pattern, Tx and/or Rx beam boresight direction (azimuth and elevation), 3dB beamwidth, etc.), expected Tx and/or Rx beam for the prediction (e.g., expected Tx and/or Rx angle, Tx and/or Rx beam ID for the prediction), UE position information, UE direction information, Tx beam usage information, UE orientation information, etc.
		- Note: The provision of assistance information may be infeasible due to the concern of disclosing proprietary information to the other side.
* Alt.3: CIR based on Set B
* Alt.4: L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID
* Note1: It is up to companies to provide other alternative(s) including the combination of some alternatives
* Note2: All the inputs are “nominal” and only for discussion purpose.

Conclusion

For the sub use case BM-Case2, further study the following alternatives with potential down-selection:

* Alt.1: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)
	+ FFS: QCL relation between beams in Set A and beams in Set B
* Alt.2: Set B is a subset of Set A (Set A and Set B are not the same)
	+ FFS: how to determine Set B out of the beams in Set A (e.g., fixed pattern, random pattern, …)
* Alt.3: Set A and Set B are the same
* Note1: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.
* Note2: It is up to companies to provide other alternative(s)
* Note3: The narrow and wide beam terminology is for SI discussion only and have no specification impact

Conclusion

Regarding the sub use case BM-Case2, further study the following alternatives of measurement results for AI/ML input (for each past measurement instance):

* Alt.1: Only L1-RSRP measurement based on Set B
* Alt 2: L1-RSRP measurement based on Set B and assistance information
	+ FFS: Assistance information. The following were mentioned by companies in the discussion:, Tx and/or Rx beam angle, position information, UE direction information, positioning-related measurement (such as Multi-RTT), expected Tx and/or Rx beam/occasion for the prediction (e.g., expected Tx and/or Rx beam angle for the prediction, expected occasions of the prediction), Tx and/or Rx beam shape information (e.g., Tx and/or Rx beam pattern, Tx and/or Rx beam boresight directions (azimuth and elevation), 3dB beamwidth, etc.) , increase ratio of L1-RSRP for best N beams, UE orientation information
		- Note: The provision of assistance information may be infeasible due to the concern of disclosing proprietary information to the other side.
* Alt.3: L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID
* Note1: It is up to companies to provide other alternative(s) including the combination of some alternatives
* Note2: All the inputs are “nominal” and only for discussion purpose.